



C I E N C I A S
UNIVERSIDAD DE LOS ANDES
M É R I D A V E N E Z U E L A

República Bolivariana de Venezuela

Universidad de los Andes

Departamento de Física

SUMA

Datos Financieros y Cómo minarlos

Trabajo de grado presentado ante la ilustre

Universidad de Los Andes

como requisito parcial para optar al Título de:

Licenciado en Física

Presentado por:

Br. Victor T. Marchetta M.

Tutor:

MSc. Mayerlin Y. Uzcátegui S.

www.bdigital.ula.ve

Resumen

Dada la relevancia de las criptomonedas, hemos puesto la mirada en conceptos de econofísica y sociofísica con el fin de aplicarlos al análisis en los mercados de intercambios de criptomonedas. Es así como presentamos un método diseñado para la minería de datos basado en la extracción de datos financieros provenientes de múltiples fuentes con el fin de proporcionar una herramienta que facilite el análisis de cotizaciones en estos mercados y la toma de decisiones al momento de realizar intercambios. Para llevar a cabo este diseño fue necesario documentarnos sobre el mundo de las finanzas el tipo de dato con el que contamos en este caso son cotizaciones de pares de criptomoneda y dinero fiat a los cuales les aplicamos las técnicas de minería de datos en tres fases: extracción, procesamiento y reportes. En este documento se expone como se realizó la extracción de los datos de sitios web, un breve resumen de los estudios que se consideraron y la manera en que se presentan los resultados.

Basado en este método presentamos un prototipo de software funcional que nos permite validar el modelo, mostrar detalles implementación y ratificar que con la ayuda de una interfaz gráfica de usuario (GUI) amigable es posible presentar las múltiples alternativas de estudio de una manera cómoda.

Palabras clave: *minería de datos; criptomoneda; GUI.*

Índice general

Resumen	I
Introducción	1
1. Datos financieros.	4
1.1. ¿Qué es el dinero?	4
1.1.1. Tipos de dinero.	5
1.2. Dinero en los mercados.	5
1.2.1. Criptomonedas.	5
1.2.2. Dinero fíat y bancario.	13
1.2.3. Metales Preciosos.	23
1.2.4. Conceptos de mercados de intercambio	29
1.3. Dato e información.	37
1.4. Fuente de datos.	38
1.4.1. Almacenes de datos.	39
1.5. Sitios web de cotizaciones.	42
1.5.1. Datos de nuestro interés.	45
1.5.2. Opciones de interactividad.	46
2. Minería de datos.	50
2.1. Extracción	50
2.1.1. Interfaz de Programación de Aplicaciones.	51
2.1.2. Automatización de descargas.	55
2.2. Procesamiento.	57
2.2.1. Muestreo.	60
2.2.2. Estadística.	64
2.2.3. Medidas Estadísticas.	65

2.2.4.	Probabilidad.	86
2.2.5.	Inferencia.	97
2.2.6.	Modelado y Simulación.	103
2.3.	Reportes.	116
2.3.1.	Entradas.	116
2.3.2.	Archivo de datos.	118
2.3.3.	Tipos de Reportes.	119
2.3.4.	Salidas.	121
3.	Método Propuesto para la Minería de Datos.	122
3.1.	Diseño y arquitectura del método propuesto.	122
3.1.1.	Extracción	124
3.1.2.	Procesamiento.	131
3.1.3.	Reportes.	133
3.1.4.	Interfaz Gráfica de Usuario.	134
3.2.	Prototipo Funcional	135
3.2.1.	Especificaciones del prototipo.	135
3.2.2.	Inicio de sesión.	139
3.2.3.	Descarga de datos.	141
3.2.4.	Sesión de estudio.	144
3.2.5.	Selección de estudio.	145
3.2.6.	Comparación	146
3.2.7.	Estadística	149
3.2.8.	Inferencia	156
3.2.9.	Simulación	157
3.2.10.	Calcular y cancelar.	163
3.2.11.	Generación de Reportes.	164
3.2.12.	Archivos del Prototipo.	170
3.2.13.	Instalación y ejecución.	170
	Conclusiones	172

Índice de figuras

1. Áreas que integran la Física Computacional.	2
1.1. Las principales casas de intercambio de criptomonedas segun CoinMarketCap. . . .	13
1.2. Países incluidos en la Unión Europea.	16
1.3. Ejemplo de Área de Soporte	31
1.4. Ejemplo de Área de Resistencia	32
1.5. Ejemplo de soporte y resistencia psicológico	33
1.6. Ejemplo de soporte y resistencia de la linea de tendencia.	33
1.7. Ejemplo de soporte y resistencia debido media móvil	34
1.8. Instantes de tiempo.	41
1.9. Opciones disponibles en las Fuentes de datos pre-seleccionadas.	48
2.1. Métodos del Protocolo de Transferencia de Hipertexto.	55
2.2. Ejemplo de muestreo sistemático	63
2.3. Enumeracion de muestras en para el muestreo por grupos.	64
2.4. Gráfico de promedio y dos medias móviles.	72
2.5. Bandas de Bollinger.	74
2.6. Retrocesos de Fibonacci.	75
2.7. Media móvil de Convergencia Divergencia.	77
2.8. Índice de fuerza relativa.	80
2.9. Índice direccional promedio.	84
2.10. Nube Ichimoku.	86
2.11. Distribución Uniforme	91
2.12. Distribución Normal	92
2.13. Distribución Exponencial	93
2.14. Distribución Triangular	93
2.15. Distribución Beta.	95

2.16. Distribución Semicircular	95
2.17. Distribución Semicircular de radio 1.	96
2.18. Distribución Gamma	96
2.19. Elementos básicos del modelado y simulación de sistemas.	103
2.20. Concepto básico de Sistema.	106
2.21. Especificaciones del modelo basado en Ecuaciones Diferenciales.	107
2.22. Especificaciones del modelo basado en Ecuaciones en Diferencia.	108
2.23. Especificaciones del modelo basado en Eventos Discretos.	108
2.24. Jerarquía de Especificación de Sistemas.	112
2.25. Algoritmo básico de un simulador atómico.	113
3.1. Fases del método propuesto.	123
3.2. Pasos de la minería de datos	124
3.3. Arquitectura detallada del método propuesto.	125
3.4. Fase de Extracción.	126
3.5. Formato propuesto para los archivos de datos.	129
3.6. Fase de Procesamiento.	131
3.7. Fase de Reportes.	133
3.8. Software requerido por el prototipo.	136
3.9. Librería de python3 requeridas	136
3.10. Implementación del método propuesto en el prototipo	137
3.11. Pantalla de inicio de prototipo MAillardFinance	139
3.12. Ejecución de GesDescargas en el prototipo.	140
3.13. Tareas del script GesDescargas	141
3.14. Tareas de un script Extractor_SXX	142
3.15. Datos procesados del sitio BitcoinAverage.	143
3.16. Datos procesados del sitio CoinMarketCap	144
3.17. Extracción completada por el prototipo.	144
3.18. Archivos resultantes de una extracción para ambas fuentes.	145
3.19. Archivos diarios de la fuente BitcoinAverage para 8 monedas fiat	145
3.20. Archivos diarios de la fuente CoinMarketCap para 8 monedas fiat	146
3.21. Estudio Moneda vs Moneda del prototipo.	147
3.22. Formato de resultado para Moneda vs Moneda.	147
3.23. Estudio Varias Monedas del prototipo.	148

3.24. Formato de resultado para Varias monedas.	148
3.25. Formato de resultado para Varias monedas normalizadas.	149
3.26. Estudio Moneda + Trading del prototipo.	149
3.27. Formato de resultado para Media Móvil Simple.	150
3.28. Formato de resultado para Media Móvil Exponencial.	151
3.29. Formato de resultado para Retrocesos de Fibonacci.	151
3.30. Formato de resultado para Bandas de Bollinger.	152
3.31. Formato de resultado para Media Móvil de Convergencia Divergencia	152
3.32. Formato de resultado para Índice de Fuerza Relativa.	153
3.33. Formato de resultado para Oscilador Estocástico.	153
3.34. Formato de resultado para Índice Direccional Promedio	154
3.35. Formato de resultado para Nube Ichimoku.	155
3.36. Estudio Tendencia Central.	156
3.37. Formato de resultado para Tendencia Central	157
3.38. Estudio Inferencia Basada en Medias Móviles.	158
3.39. Estudio Simulación mediante el Método Distri	159
3.40. Estudio Simulación mediante el Método MAillard.	161
3.41. Formato de resultado para Simulación Método Distri y MAillard.	163
3.42. Cálculo de Simulación Método MAillard.	163
3.43. Barra de herramientas del gráfico.	164
3.44. Archivo de Datos de una simulación según el Método MAillard.	165
3.45. Gráfico de una simulación según el Método MAillard.	166
3.46. Formato de nombre para reportes del área Comparación.	167
3.47. Formato de nombre para reportes del área Estadística	168
3.48. Formato de nombre para reportes del área Inferencia	169
3.49. Formato de nombre para reportes del área Simulación	169
3.50. Archivos necesarios para la ejecución del prototipo.	170
3.51. Tareas desempeñadas por el script setup	171

“Somos lo que hacemos repetidamente. La excelencia, entonces, no es un acto, es un hábito.”

Will Durant

Introducción

Las criptomonedas hoy día ya se han instaurado en la sociedad a un nivel global, esto ha llevado a que las personas las usen no sólo como medio de pago sino también como objeto de resguardo de valor y negocio. Un ejemplo es su uso como activo para generar ingresos de la compra-venta, esto de alguna manera incentiva el desarrollo de estudios con respecto a su comportamiento.

En nuestro caso proponemos el diseño de un método que usa minería de datos con el objetivo de proporcionar herramientas que faciliten el análisis de datos y permitan al usuario estimar el rumbo del mercado mediante comparaciones y estudios de estadística, inferencia y simulación.

La física computacional se centra en la elaboración de modelos computacionales para resolver problemas que generalmente abarcan múltiples variables para los que incluso puede existir algún método de solución cuantitativo. Esta integran aplicaciones en física, técnicas matemáticas y soluciones del campo de la ciencia computacional como se ve en la figura 1. En nuestro caso el problema es el diseño de un método de minería de datos que permita generar estimaciones para el mercado de criptomonedas, este problema integra tanto fenómenos sociales como económicos de aquí que se trate como econofísica y sociofísica. Las matemáticas nos permiten proporcionar herramientas útiles basadas en comparación, estadística, probabilidades, inferencia y simulación que facilitan al usuario las tareas vinculadas al proceso intercambio (*trading* o especulación sobre instrumentos financieros) para que junto a su experiencia previa facilite el análisis de estos mercados y permita estimar su comportamiento.

Proponemos un método que implementa la minería de datos en tres fases y una interfaz gráfica de usuario como apoyo, adicionalmente para validar este método se presentan los detalles de implementación de un prototipo funcional.

En este documento, se presenta la información que consideramos necesaria para el entendimiento del proyecto, su manipulación y futuros trabajos.

El capítulo 1 establece las bases de conocimiento necesario. Comenzamos centrándonos en

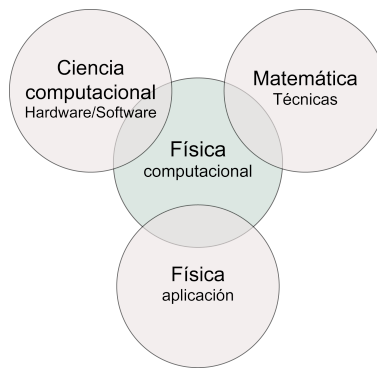


Figura 1: Áreas que integran la Física Computacional.

nuestro objeto de estudio (cotizaciones) y luego presentamos las fuentes de datos que usaremos. Como bien el título lo dice, hablaremos de datos financieros y como el término finanza engloba a todas las actividades que tienen que ver con dinero, definiremos ¿qué es el dinero?, presentaremos sus tipos luego pasaremos a clasificarlo respecto a su uso en las cotizaciones del mercado de criptomonedas y mas específicamente en mercados de intercambio de divisas. A lo largo del capítulo dejaremos claro ¿qué es esto de las criptomonedas?, ¿a qué se debe el revuelo que las criptomonedas han causado?, ¿por qué se incluyen algunos metales preciosos en las listas de monedas fíat? y otras dudas comunes. Finalizaremos el capítulo presentando algunos de los sitios web más comunes dedicados a la divulgación de datos sobre mercados de criptomonedas y detallaremos lo que nos ofrecen.

En el capítulo 2 explicamos de manera detallada lo que implica la minería de datos, específicamente de la manera cómo se propone en el método donde integramos datos provenientes de múltiples sitios web que reflejan el comportamiento histórico del mercado de criptomonedas desde su inicio hasta el momento de la consulta y un conjunto de utilidades que pretenden facilitar la toma de decisiones al momento de realizar intercambios. Finalmente expondremos las herramientas y técnicas que se aplicarán en las tres fases del método:

1. Extracción. Presentamos los métodos de extracción y las herramientas para automatizar descarga de datos de los múltiples servidores.
2. Procesamiento. Se unifica el formato de los datos a la medida de nuestras necesidades facilitando la utilización en los métodos matemáticos seleccionados que son ampliamente usados para el análisis de datos.
3. Reportes. Mostramos los posibles reportes que se generan a partir de los datos descargados y procesados, los cuales se espera brinden información útil y personalizada de acuerdo a la estrategia que decida utilizar el usuario.

En el capítulo 3 explicamos detalles del diseño del método. Partimos de una explicación desde una perspectiva general hasta una explicación detallada sobre que sucede específicamente en cada una de las fases y sus ventajas, haciendo uso de diagramas de flujo y gráficos representativos. Adicionalmente se presentan los detalles del prototipo funcional que implementa dicho método y de la forma en que abordamos cada proceso dentro del mismo.

www.bdigital.ula.ve

“La naturaleza nos ha dado las semillas del conocimiento, no el conocimiento mismo”

Séneca, Lucio Anneo

Capítulo 1

Datos financieros.

Este capítulo va destinado a sentar las bases del conocimiento necesarias para comprender este proyecto, primero centrándonos en nuestro objeto de estudio y luego exponiendo la fuentes de datos que usaremos para realizar nuestra extracción de datos. Como bien el título lo dice hablaremos de datos financieros y como el termino finanza engloba a todas las actividades que tienen que ver con dinero, definiremos qué es el dinero y sus tipos de manera amplia, luego pasaremos a clasificarlo de una manera más específica con respecto a como luce en los mercados de nuestro interés y más específicamente en mercados de intercambio de divisas.

Con el tema del dinero aclarado continuaremos con el tema de datos, ya que nuestro fin es hacer minería de datos, primero debemos definirlos y diferenciar dato de información, definir tipos de fuentes de datos y finalizar con donde y como se almacenan estos datos, en este último punto comentaremos algunos factores a tomar en cuenta a la hora de elegir y estudiar almacenes de datos en línea. Por último cerraremos el capítulo hablando de la fuente de datos que usaremos y discutiremos los rasgos más relevantes que contemplan estas fuentes de datos en específico al proporcionarnos datos e información, lo que llamaremos opciones de visualización.

1.1. ¿Qué es el dinero?

Hablar de dinero puede llegar a ser un poco ambiguo ya que según cada fuente que consultemos puede tener significados y clasificaciones distintas, pero lo más básico es definirlo como un medio de pago, separando así compra-venta de trueque.

1.1.1. Tipos de dinero.

De curso legal o fiat: el cual es emitido por un Banco Central, su característica resaltante es que es de aceptación forzosa, declarada por el estado como en nuestro caso ha sido el Bolívar, luego el Bolívar Fuerte, después el Bolívar Soberano y por último el Bolívar Digital.

Bancario: este es creado por los bancos a través de las operaciones de crédito y se expresan en forma de depósitos a la vista. También denominado dinero secundario o depósito.

Mercancía: el cual es intrínsecamente útil y se valoraría aun cuando no sirviese como dinero. Antigua forma de pagos cuyo valor estaba asociado al metal o material del dinero mismo.

Fiduciario: este no posee valor ni intrínseco ni por ley, se basa en la aceptación y confianza de sus usuarios.

Comúnmente el término moneda en el sentido más específico es dinero en cualquiera de sus presentaciones, cuando está en uso o en circulación como medio de cambio, especialmente billetes y monedas circulantes. Ahora bien, con respecto a los mercados de divisas, fondos y acciones podríamos clasificarlos de manera distinta, una más práctica y más tangible para nuestro proyecto.

1.2. Dinero en los mercados.

Si bien ya hablamos sobre los tipos de dinero, ahora explicaremos de manera más detallada los tres subgrupos que forman parte de nuestro estudio. Tomaremos algunas criptomonedas como referencia y tendremos acceso a su cotización con respecto a una gran cantidad de monedas fíat y cuatro metales preciosos. Con esto en mente, las criptomonedas nos servirán como base para la comparación de las monedas fíat, claro que si bien las monedas fíat se nos hacen familiar, las criptomonedas a pesar de llevar más de diez años entre nosotros aun están rodeadas de gran cantidad de desinformación que trataremos de aclarar.

1.2.1. Criptomonedas.

Una criptomoneda es un medio de intercambio digital, cifrado y descentralizado. Estas utilizan técnicas de cifrado para regular la generación de unidades monetarias y verificar la transferencia de fondos. A diferencia del dólar estadounidense (USD) o el euro (EUR) por ejemplo, no existe una autoridad central que administre, rija su emisión y mantenga el valor de una criptomoneda o al menos esto era su finalidad inicial, ya que como veremos más adelante algunas se han alejado

un poco de este propósito de ser netamente descentralizadas pero siguen estando dentro de lo que llamamos criptomoneda.

Hoy en día se pueden usar criptomonedas para comprar bienes y pagar servicios regulares. Adoptadas como un medio de intercambio serio, descentralizado y a nivel mundial, donde su valor recae en la solidez y apoyo que posea el proyecto de la criptomoneda en sí, algunos proyectos sólidos han llegado a ser tomados como inversiones para resguardar valor como el caso de la criptomoneda Bitcoin.

Las criptomonedas se sustentan en una innovación llamada cadena de bloques, de ahora en adelante blockchain del inglés, que como su nombre lo indica, es una cadena de bloques, cada uno de estos bloques que es agregado a la cadena funciona como una página nueva de un libro mayor público, el cual está distribuido en todos los ordenadores que forman parte de la red. Estos ordenadores tienen su propia copia del libro para crear un registro de transacciones unificado. El software registra cada nueva transacción a medida que ocurre y cada copia de la cadena de bloques se actualiza simultáneamente con la nueva información, manteniendo todos los registros idénticos y precisos.

Para evitar el fraude, cada transacción de una criptomoneda en específico es agregada al bloque que se está minando y este se verifica utilizando una de las dos técnicas principales de validación: prueba de trabajo (Proof-of-Work, PoW por sus siglas en inglés) [1] o prueba de participación (Proof-of-Stake, PoS por sus siglas en inglés) [2], una vez que se alcanza la cantidad máxima de transacciones posibles dentro del bloque en cuestión y es realizada la prueba, el bloque es agregado a la cadena y el autor o los participantes de esta acción reciben recompensas en la misma criptomoneda en cuestión, esto es a lo que llamamos minado, y es como se generan nuevas criptomonedas, cabe destacar que algunas criptomonedas poseen otros métodos de minado, pero éste es el más común.

Hoy día hay cierto debate entre estas dos modalidades de validación ya que la PoW conlleva un consumo mayor de energía debido a que el computo necesario es mayor, por esto algunas criptomonedas han tomado la iniciativa de migrar a PoS o implementar energías limpias para efectuar el minado. Las críticas más fuertes o relevantes en contra de las criptomonedas, coinciden en que mantener la red acarrea un costo energético alto, claro que esto es relativo porque ¿cuánto cuesta mantener toda la infraestructura humana y de activos de un banco o de una moneda de algún país en su totalidad?

Ahora bien, entendiendo de manera básica como funcionan las criptomonedas, ¿cómo las manejo?. La manera de manipularlas es mediante una billetera/cartera electrónica [3], aunque el

término billetera es un poco inapropiado, ya que una billetera de criptomonedas en realidad no almacena fondos, en cambio, cada billetera almacena una o más claves criptográficas privadas que se requieren para acceder a los fondos dentro de la red de la criptomoneda en cuestión, es por esto que perder su o sus claves a menudo puede significar perder sus criptomonedas, ya que al ser un sistema descentralizado no hay un ente donde reclamar o que posea un acceso indirecto a éstos para otorgar de nuevo el acceso.

Hay tres grupos principales de billeteras de criptomonedas: software, hardware y billeteras de papel [4]. Pero también pueden definirse como monederos calientes o monederos fríos según la forma en que funcionan. Las billeteras calientes son las que de alguna manera están conectadas a internet y por lo tanto son más susceptibles a los ataques de piratería. Las billeteras frías son las que generan claves sin conexión a internet, lo que las hace altamente resistentes a los ataques cibernéticos.

Carteras de software: Los tipos más comunes de carteras de software incluyen carteras web, carteras de escritorio y carteras móviles.

- Monedero web: consiste en una interfaz de navegador que no requiere ninguna descarga o instalación. Más cómodo pero también más peligroso, ya que las claves privadas suelen ser gestionadas por terceros.
- Cartera de escritorio: software que se puede descargar y ejecutar localmente. Menos conveniente que las billeteras web pero más seguras porque las claves privadas se almacenan localmente y las administran los usuarios.
- Billetera móvil: son similares a las billeteras de escritorio pero diseñadas para teléfonos inteligentes. El uso de códigos QR (Quick Response code, código de respuesta rápida o QR por sus siglas en inglés) los convierte en una alternativa conveniente para enviar y recibir criptomonedas.

Carteras de hardware: Las billeteras de hardware consisten en dispositivos físicos que generan y almacenan claves sin ninguna conexión a internet y como tales entran en la categoría de billeteras frías. Por lo general, las claves se crean en base a algoritmos de generación de números aleatorios (Random Number Generator, RNG por sus siglas en inglés) y se almacenan en el propio dispositivo (en ningún otro lugar). A pesar de ser menos convenientes debido a la accesibilidad limitada, las billeteras de hardware se consideran una de las alternativas más seguras para “almacenar” y administrar criptomonedas.

Carteras de papel: Una billetera de papel consiste en una hoja de papel con una dirección de cadena de bloques y su clave privada correspondiente. Las claves generalmente se imprimen como largas cadenas de números y letras junto con códigos QR, que se pueden escanear para ejecutar transacciones de criptomonedas. Si las billeteras de papel se utilizan para generar claves fuera de línea, también pueden considerarse billeteras frías. Sin embargo, se desaconseja su uso porque presentan numerosas fallas y riesgos potenciales para los usuarios que carecen de conocimientos técnicos.

Si bien las criptomonedas se pueden generar a través de la minería al ser parte de la red, también se pueden crear de manera unilateral por los o el creador del proyecto, de manera totalmente opuesta a la creación unilateral de dichas monedas algunos proyectos también implementan lo que se denomina métodos de quemado, la quema de criptomonedas es un mecanismo utilizado para retirar monedas de circulación de manera permanente, reduciendo así su suministro total de forma permanente. De esta manera es común ver proyectos de criptomonedas que realizan quemas de monedas periódicamente o mediante cobro de comisiones al realizar algún tipo de transacción con el fin crear un efecto deflacionario sobre la criptomoneda, y por otra parte, existen criptomonedas las cuales han sido quemadas mediante la destrucción de billeteras frías o pérdidas de claves privadas ya que sin las claves es imposible acceder a estos fondos.

Mediante este proceso se suelen enviar dichas criptomonedas a una cartera, llamada comúnmente un “agujero negro”, a la que nadie tiene acceso, sacándolas así de circulación y por ende quemándolas.

Pero hay otros usos de la quema de monedas y otras razones por la que se puede querer quemar una criptomoneda. En comparación con los mecanismos PoW y PoS que utilizan grandes cantidades de energía para realizar los cálculos necesarios, el mecanismo prueba de quema (Proof-of-Burn, PoB por sus siglas en inglés) es diferente [5], en este caso los mineros deben quemar criptomonedas para obtener acceso a minar y escribir bloques a un coste mucho más barato. El sistema de prueba de quema también utiliza muy poca energía, ya que se quema la criptomoneda, en lugar de crearla o minarla.

Ahora bien hablemos de los proyectos de criptomonedas y sus criptomonedas más usadas hoy en día, acompañaremos el nombre del proyecto seguido del código de su moneda nativa, este es el usado para hacer referencia a su cotización en los mercados que es como nos referiremos a ellos de aquí en adelante.

Bitcoin, BTC [6,7]: fue la primera criptomoneda, descrita por primera vez en principio por Satoshi Nakamoto en un artículo de 2008 titulado “Bitcoin: A Peer-to-Peer Electronic Cash

System” (Bitcoin: un sistema de dinero electrónico de igual a igual). Nakamoto describió el proyecto como un sistema de pago electrónico basado en pruebas criptográficas en lugar de confianza.

Bitcoin fue la primera red de pago entre pares completamente descentralizada e impulsada por sus mismos usuarios, en este caso la criptomoneda dentro de la red posee el mismo nombre, Bitcoin. Representa una forma de pago global sin fronteras, transacciones entre monederos los cuales son creados de manera anónima. A diferencia de algunas criptomonedas, Bitcoin sigue siendo la que representa en todo su potencial la filosofía de descentralización, ya que al ser su inventor anónimo hasta ahora, no hay nadie dando la cara o tomando decisiones sobre el rumbo del proyecto, su valor y desarrollo recae en la comunidad que la usa y la mantiene en funcionamiento. Sigue siendo hoy día la criptomoneda líder siendo catalogada como el oro de las criptomonedas e incluso a ser usada con un propósito en común, el de guardar valor a través de los años.

El BTC es una de las criptomonedas que posee límite de emisión o minado mejor dicho, estipulado en 21 millones de bitcoins y este no posee mecanismos de quema de criptomonedas por los momentos.

Ethereum, ETH [8–10]: es un sistema blockchain de código abierto descentralizado que cuenta con su criptomoneda nativa de la red, esta lleva el nombre de Ether, que es lo que verdaderamente se intercambia, a diferencia de la red BTC aquí la red y la moneda nativa no poseen el mismo nombre pero es usual el uso de las siglas ETH para hacer referencia a ambas, distinguiéndolas solamente por el contexto donde se usen. ETH, la red, funciona como plataforma para muchas otras criptomonedas, así como para la ejecución de contratos inteligentes descentralizados.

Ethereum se describió por primera vez en un documento técnico de 2013 de Vitalik Buterin. Buterin, junto con otros cofundadores, aseguró la financiación del proyecto en una venta pública en línea en el verano de 2014. El equipo del proyecto logró recaudar \$18.3 millones en Bitcoin y el precio de ETH en la oferta inicial de monedas (Initial Coin Offering o ICO por sus siglas en inglés) fue de \$0.311, con más de 60 millones de ETH vendidos. Tomando el precio de ETH ahora, esto pone el retorno de la inversión a una tasa anualizada de más del 270 %, esencialmente casi cuadruplicando su inversión cada año desde el verano de 2014.

El supuesto objetivo de Ethereum es convertirse en una plataforma global para aplicaciones descentralizadas que permita a los usuarios de todo el mundo escribir y ejecutar software

resistente a la censura, al tiempo de inactividad y al fraude. Su principal innovación fue diseñar una plataforma que le permitiera ejecutar contratos inteligentes usando blockchain, lo que refuerza aún más los beneficios ya existentes de la tecnología de contratos inteligentes. La cadena de bloques de Ethereum fue diseñada, según el cofundador Gavin Wood, como una especie de “computadora para todo el planeta”, teóricamente capaz de hacer que cualquier programa sea más robusto, resistente a la censura y menos propenso al fraude al ejecutarlo en una red distribuida globalmente de nodos públicos.

Además de los contratos inteligentes, la cadena de bloques de Ethereum puede alojar otras criptomonedas llamadas “tokens” mediante el uso de su estándar de compatibilidad ERC-20, los tokens, en términos generales son unidades de valor digitales que no se pueden minar y que existen como entradas de registro en las cadenas de bloques. Los tokens generalmente son emitidos por empresas que utilizan cadenas de bloques de terceros existente. Estrictamente hablando, los tokens no son criptomonedas como Bitcoin o Ether, pero si unidades de valor transferibles emitidas sobre una cadena de bloques [11], aunque usualmente son vistas como iguales. De hecho, este ha sido el uso más común de la plataforma ETH hasta el momento, se han lanzado más de 280,000 tokens compatibles con ERC-20. Más de 40 de estas se encuentran entre las 100 principales criptomonedas por capitalización de mercado.

En sus últimas actualizaciones se añade un mecanismo de quema de tarifas. Una parte de cada tarifa de transacción (la tarifa base) se quema y se retira de circulación. Esto tiene como objetivo reducir el suministro circulante de Ether y potencialmente aumentar el valor del token con el tiempo.

Tether, USDT [12, 13]: es parte de un grupo de criptomonedas que se denominan criptomonedas estables o solo monedas estables de ahora en adelante “stablecoin(s)” del inglés, estas usualmente son tokens que están asociados al valor de una moneda fiat (como el USD o el EUR), a bienes materiales como el oro, bienes inmuebles o a otra criptomoneda. También hay stablecoins que no están asociadas a ninguna otra moneda sino que están controladas mediante algoritmos para mantener un precio estable. La principal motivación para crear una stablecoin es tratar de dar refugio a los inversores en momentos de alta volatilidad.

Se espera que el precio del USDT sea igual a un dólar estadounidense, esta es emitida por Tether Operations Limited, una empresa con sede en Hong Kong. La vinculación del token al USD se logra mediante el mantenimiento de una suma de papel comercial, depósitos fiduciarios, efectivo, notas de recompra de reserva y letras del Tesoro en reservas que es igual

en valor en dólares al número de USDT en circulación.

Lanzado originalmente en julio de 2014 como Realcoin, un token de criptomoneda construido sobre la cadena de bloques de Bitcoin, luego se le cambió el nombre a USTether y finalmente a USDT. Tether se actualizó más tarde para trabajar en las cadenas de bloques Ethereum, EOS [14], Tron [15], Algorand [16] y OMG [17].

El propósito declarado de USDT es combinar la naturaleza sin restricciones de las criptomonedas que pueden enviarse entre usuarios sin un tercero intermediario de confianza con el valor estable del dólar estadounidense. La empresa Tether garantiza que su valor permanecerá vinculado al USD y asegurando que cada vez que emite nuevos tokens USDT, asigna la misma cantidad de USD a sus reservas, lo que garantiza que USDT esté completamente respaldado por efectivo y equivalentes de efectivo con el objetivo de proteger al USDT de las altas fluctuaciones que hasta ahora presentan las criptomonedas.

Convirtiéndola así en un tipo de refugio seguro durante los períodos de alta volatilidad, transfiriendo sus fondos a USDT sin tener que retirar dinero por completo en USD o alguna otra moneda fiat. Además USDT ya cuenta con un uso común para realizar transacciones equivalentes en dólares estadounidenses entre regiones, países e incluso continentes a través de cadenas de bloques, sin tener que depender de un intermediario como un banco o un proveedor de servicios financieros.

Sin embargo, a lo largo de los años han habido una serie de controversias con respecto a la validez de las afirmaciones de Tether sobre sus reservas de USD. Muchos han expresado su preocupación por el hecho de que las reservas de Tether nunca han sido auditadas por completo por un tercero independiente. A pesar de esto En 2019, Tether superó a Bitcoin en volumen de operaciones con el volumen de operaciones diario y mensual más alto de todas las criptomonedas del mercado.

Se dice que las monedas estables canjeables por divisas, mercancías, o dinero fiat están respaldadas; mientras que aquellas ligadas a un algoritmo se denominan de tipo señoreaje (no respaldadas).

Binance Coin, BNB [18–20]: Lanzado en julio de 2017, Binance es el mayor intercambio de criptomonedas a nivel mundial en función del volumen de negociación diario. Binance tiene como objetivo llevar los intercambios de criptomonedas a la vanguardia de la actividad financiera a nivel mundial. La idea detrás del nombre de Binance es mostrar este nuevo paradigma en las finanzas globales: Finanzas binarias (Binary Finance o Binance, en inglés).

Además de ser el intercambio de criptomonedas más grande del mundo, Binance ha lanzado todo un ecosistema de funcionalidades para sus usuarios. La red de Binance incluye los proyectos Binance Chain [21], Binance Smart Chain [22], Binance Academy [23], Trust Wallet [24] y Research [25], que emplean los poderes de la tecnología blockchain para llevar las finanzas de la nueva era al mundo. BNB es una parte integral del funcionamiento exitoso de muchos de los subproyectos de Binance.

La criptomoneda cuenta con el símbolo BNB y se basaba en la cadena de bloques de Ethereum, desde el lanzamiento de su propia cadena, la “Binance Chain” se introdujo el tipo de token BEP-2, de esta manera los tokens BNB de tipo ERC-20 fueron intercambiados gradualmente por tokens BNB de tipo BEP-2 en una proporción 1:1, convirtiéndose en la moneda nativa de la Binance Chain.

La moneda nació básicamente para soportar transacciones dentro de la propia plataforma de Binance por lo que no se pueden adquirir en otras plataformas. A diferencia de otras criptomonedas como el BTC, no se puede minar BNB. Todos los tokens fueron minados previamente por Binance en su comienzo. Su estructura tiene un soporte de datos con diversas unidades de asignación que incrementan la seguridad del sistema por un lado y aceleran la búsqueda de datos, gracias a la fragmentación.

Hay un suministro máximo inicial de 200,000,000 BNB, de los cuales 168,137,036 están actualmente en circulación a septiembre de 2021.

Binance realiza una quema trimestral del suministro de BNB, el monto a quemar representa el 20 % de sus beneficios en BNB cada trimestre, aunque en sus inicios el cálculo de la cantidad de BNB a quemar era más complicado y/o desconocido, esta quema tiene como objetivo reducir el suministro total a 100,000,000 BNB. Según el documento técnico de Binance, exactamente la mitad del suministro máximo de tokens BNB se asignó a la ICO y la venta pública del token. La ICO se llevó a cabo en julio de 2017 poco después del lanzamiento del intercambio y Binance recaudó \$15 millones en fondos, con un precio de BNB igual a \$0.10 durante la ICO. Tomando el precio de BNB hoy, representa un retorno de la inversión de casi 4200 veces. Otro 40 % del suministro total, o alrededor de 80,000,000 de monedas BNB, se distribuyó entre los miembros fundadores y el equipo. Finalmente, el 10 % restante de la oferta máxima se repartió entre ángeles inversores [26].

Binance como acabamos de mencionar es una casa de intercambio de criptomonedas de ahora en adelante **exchange** del inglés, es un negocio que permite a los clientes intercambiar criptomonedas

o monedas digitales por otros activos, como dinero fíat convencional u otras monedas digitales. Los intercambios suelen aceptar pagos con tarjeta de crédito, transferencias electrónicas u otras formas de pago a cambio de monedas digitales o criptomonedas. Un intercambio de criptomonedas puede ser un creador de mercado [27] que generalmente toma los diferenciales de oferta y demanda como una comisión de transacción por su servicio o, como una plataforma de coincidencia, simplemente cobra tarifas o comisiones por cada transacción.

Algunas casas de bolsa que se enfocan en otros activos como acciones tales como Robinhood [28] y eToro [29], permiten a los usuarios comprar y canjear dentro criptomonedas por otras pero no permite su retiro en la criptomoneda misma. Sin embargo, los exchange dedicados, como Binance y Coinbase [30] permiten depósitos y retiros de criptomonedas. Según el sitio web CoinMarketCap las principales casas de intercambio de criptomonedas en función del tráfico, la liquidez y los volúmenes de comercio son Binance, Coinbase, FTX, Kraken y Cucoin [31], estos datos se basan en la confianza y legitimidad de los volúmenes de comercio reportados, más detalles en la figura 1.1.

Casa de Intercambio	Volumen(\$/24h)	Criptomonedas	Monedas Fíat
Binance [32]	>15 mil millones	>400	46
Coinbase [30]	3 mil millones	146	{USD,EUR,GBP*}
FTX [33]	>1 mil millones	303	10
Kraken [34]	1 mil millones	105	7
KuCoin [35]	>2 mil millones	593	48

Figura 1.1: Las principales casas de intercambio de criptomonedas segun CoinMarketCap.

Ya habiendo explicado de una manera suficiente sobre las criptomonedas, veamos a detalle ahora las principales monedas fíat del mundo y su respectiva parte como dinero bancario.

1.2.2. Dinero fíat y bancario.

Primero hablemos de como terminamos con el dinero fíat que conocemos hoy por hoy, en sus inicios en su forma de papel moneda o notas bancarias se desarrolló por primera vez en la dinastía Tang de China durante el siglo VII, aunque el verdadero papel moneda no apareció hasta el siglo XI, durante la dinastía Song. Exploradores europeos como Marco Polo introdujeron el concepto en Europa durante el siglo XIII.

En Europa, los primeros billetes de los que hay constancia aparecen en Suecia en el año 1661 (siglo XVII) de la mano del cambista Johan Palmstruch, quien los entregaba como recibo para quien depositaba oro u otro metal precioso en el Banco de Estocolmo, fundado por él mismo. A

España llegaron en 1780, durante el reinado de Carlos III, y su uso se popularizó rápidamente por ser mucho más cómodos de llevar. Así, no hacía falta cargar con la famosa bolsa llena de monedas, mucho más llamativa y pesada [36].

La percepción de los billetes como dinero ha evolucionado con el tiempo. Originalmente, el dinero se basaba en metales preciosos. Algunos veían los billetes como un pagaré, una promesa de pagar a alguien en metal precioso. Fueron fácilmente aceptados, por comodidad y seguridad, en Londres, por ejemplo, desde finales del siglo XVII en adelante. Después de la Primera Guerra Mundial, los gobiernos y los bancos en general aún prometían convertir billetes y monedas en su mercancía nominal (reembolso en especie, típicamente oro) a pedido. Sin embargo los costos de la guerra y las reparaciones necesarias y el crecimiento económico basado en los préstamos del gobierno hicieron que los gobiernos suspendieran la redención en especie.

Algunos gobiernos tuvieron cuidado de evitar el incumplimiento soberano, pero no desconfiaron de las consecuencias de pagar deudas al consignar efectivo recién impreso no asociado con un patrón de metal a sus acreedores, lo que resultó en hiperinflación.

De 1944 a 1971, el acuerdo de Bretton Woods fijó el valor de 35 USD a una onza troy de oro [37]. Otras monedas se calibraron con el USD a tasas fijas. Estados Unidos prometió canjear USD por oro transferido a otros bancos nacionales. Los desequilibrios comerciales se corrigieron mediante canjes de reservas de oro o mediante préstamos del Fondo Monetario Internacional (FMI).

El sistema de Bretton Woods terminó con lo que se conoció como el “shock de Nixon” [38]. Esta fue una serie de cambios económicos efectuados por presidente de los Estados Unidos, Richard Nixon, en 1971, estos incluían la cancelación unilateral de la convertibilidad directa del dólar de los Estados Unidos en oro. Desde entonces, se ha utilizado a nivel mundial un sistema de dinero fiduciario nacional, con tipos de cambio variables entre las principales monedas. Con la eliminación de los metales preciosos del sistema monetario, los billetes se convirtieron en dinero fiduciario puro, también llamado dinero por decreto, ya que por ley se debe aceptar como forma de pago, claro que respectivamente en la nación que lo posea como dinero de curso legal [36, 39].

En cuanto a la creación de este dinero luego de la finalización del patrón oro, podríamos decir que no hay límite para su impresión, más allá de todas las implicaciones que esto acarrea, es por esto que se toman ciertas acciones previas a imprimir más dinero, como respaldarlo previamente con algo. A este proceso previo a la impresión es a lo que denominamos creación de base monetaria y consiste en aumentar los activos del Banco Central, para esto existen diversas formas de hacerlo [40]:

- Aumentando la cantidad de oro y divisas que posee el Banco Central.
- Aumentando el número de créditos que el Banco Central otorga a los bancos comerciales que

necesitan liquidez.

- Comprando deuda pública mediante operaciones de mercado abierto.
- Aumentando los créditos que el Banco Central realiza al sector público.

Frente a la forma oficial de crear el dinero, existe otra manera que podríamos denominar “virtual”. Esta es el procedimiento por el cual los bancos comerciales también pueden crear dinero. La actividad principal del sector bancario es “vender dinero” y el precio de éste está marcado por el tipo de interés.

Los bancos centrales fijan el coeficiente de reservas, que no es sino el porcentaje de dinero que basado en los depósitos, el banco ha de tener en efectivo para satisfacer las necesidades de sus clientes. Cuanto menor sea este coeficiente, más dinero podrá prestar el banco. El fenómeno de creación de dinero bancario se conoce como “expansión múltiple de los depósitos bancarios”. Lo que hace el proceso de expansión es crear dinero virtual mediante la concesión de préstamos.

Luego, los bancos comerciales redistribuyen o reutilizan este dinero base mediante la creación de crédito a través de la banca de reserva fraccionaria, lo que expande la oferta total de “dinero en sentido amplio” (efectivo más depósitos a la vista) es por esto que hemos unido dinero fiat y dinero bancario como uno mismo ya que si bien son distintos, cuando hablamos de dinero fiat comúnmente nos referimos a toda la moneda como tal.

En las economías modernas, relativamente poco de la oferta de dinero en sentido amplio es moneda física. La fabricación de nuevo dinero físico suele ser responsabilidad del banco nacional o a veces de la Tesorería del gobierno.

A continuación las monedas más usadas en intercambios a nivel mundial, su nombre seguido de su símbolo y su código según el estándar ISO 4217 [41].

El Dólar estadounidense, \$, USD [37, 39, 42]: El dólar es la moneda de curso legal de Estados Unidos, sus dependencias y otros países. El dólar es la moneda de reserva internacional más importante del mundo y también la más utilizada en las transacciones internacionales. Su condición de divisa de referencia mundial la ha convertido en la moneda oficial de varios países y la moneda de facto en muchos otros que utilizan su papel moneda para las transacciones corrientes o dan un tipo de cambio fijo a sus divisas nacionales respecto al dólar. La política monetaria de los Estados Unidos está a cargo del Sistema de la Reserva Federal, que actúa como el Banco Central de la nación y que también se encarga de la emisión de los dólares.

La Ley de acuñación de monedas de 1792 introdujo el dólar estadounidense a la par del dólar de plata español, lo dividió en 100 centavos y autorizó la acuñación de monedas denominadas en dólares y centavos. Los billetes de banco de Estados Unidos, se emiten en forma de billetes de la Reserva Federal, popularmente llamados billetes verdes debido a su color históricamente predominantemente verde.

En 1900 se rompió con el bimetalismo al vincular al dólar únicamente con el oro. El dólar se convirtió en una moneda de reserva importante desde el final de la Primera Guerra Mundial y desplazó a la libra esterlina como principal moneda de reserva en el mundo tras el acuerdo de Bretton Woods, al final de la Segunda Guerra Mundial. En 1934, se revisó su equivalencia con el oro hasta los 35 dólares por onza-troy, valor que mantendría a lo largo de varias décadas, hasta que en 1971 quedó definitivamente desvinculada del patrón oro, por lo que la divisa se convirtió de facto en una moneda fiat.

El Euro, €, EUR [43–45]: Es la moneda usada por las instituciones de la Unión Europea (UE), así como la moneda oficial de la eurozona, que es la unión monetaria del conjunto de Estados miembros de la Unión Europea que han adoptado el euro como moneda oficial. En total son 19 países y sus fechas de adopción del Euro lo podemos ver en la figura 1.2.

Fecha de adopción	País
1999-2004	Alemania, Austria, Bélgica, España, Finlandia, Francia, Irlanda, Italia, Luxemburgo, Países Bajos, Portugal y Grecia
2005-2009	Eslovenia, Malta, Chipre y Eslovaquia
2009-2014	Estonia y Letonia
2014-2019	Lituania

Figura 1.2: Países incluidos en la Unión Europea.

El nombre de “Euro” fue adoptado oficialmente el 16 de diciembre de 1995 en Madrid. Por su parte, las monedas y billetes entraron en circulación el 1 de enero de 2002. Entre tanto, de los veintisiete Estados de la UE, solo ocho no han adoptado el euro, tampoco lo hizo el Reino Unido durante sus años de pertenencia a la UE, si bien todos ellos, excepto Dinamarca, que tiene una cláusula de exclusión, están obligados a unirse a la eurozona.

La ampliación de la eurozona es un proceso por el que los estados miembros de la Unión Europea abandonan sus monedas nacionales y adoptan el euro, también llamado moneda única. El 1 de enero de 1999 se introdujo la nueva moneda de manera oficial.

El euro es gestionado y administrado por el Banco Central Europeo [46] (BCE) con sede en Frankfurt (Alemania) y el Eurosistema (compuesto por los bancos centrales de los países

de la eurozona). Como Banco Central independiente, el BCE tiene la autoridad exclusiva para establecer la política monetaria. El Eurosistema participa en la impresión, acuñación y distribución de billetes y monedas en todos los estados miembros, y en el funcionamiento de los sistemas de pago de la eurozona.

El Tratado de Maastricht de 1992 [47] obliga a la mayoría de los estados miembros de la UE a adoptar el euro si cumplen ciertos criterios de convergencia monetaria y presupuestaria, aunque no todos los estados lo han hecho. Dinamarca ha negociado exenciones, mientras que Suecia (que se unió a la UE en 1995, después de la firma del Tratado de Maastricht) rechazó el euro en un referéndum no vinculante en 2003 y eludió la obligación de adoptar el euro al no cumplir las necesidades monetarias y presupuestarias. Todas las naciones que se han unido a la UE desde 1993 se han comprometido a adoptar el euro a su debido tiempo.

El euro es la segunda moneda de reserva así como la segunda moneda más negociada en el mundo después del dólar estadounidense, millones de personas alrededor del mundo usan monedas fijadas al euro.

El Yen, ¥, JPY [48]: Es la moneda oficial de Japón. Es la tercera moneda más negociada en el mercado de divisas, después del dólar estadounidense y el euro. También es ampliamente utilizado como tercera moneda de reserva después del dólar estadounidense y el euro.

El concepto del yen fue un componente del programa de modernización de la economía de Japón del gobierno Meiji de finales del siglo XIX, que postulaba la búsqueda de una moneda uniforme en todo el país, siguiendo el modelo del sistema monetario decimal europeo. Antes de la Restauración Meiji, todos los feudos de Japón emitían su propia moneda. La Ley de la Nueva Moneda de 1871 los eliminó y estableció el yen, que se definió como 1.5g de oro o 24.26g de plata, como la nueva moneda decimal. Los antiguos *Han* (nombre dado a los feudos en Japón) se convirtieron en prefecturas y sus casas de moneda en bancos autorizados privados que inicialmente conservaron el derecho a imprimir dinero. Para poner fin a esta situación, se fundó el Banco de Japón en 1882 y se le otorgó el monopolio del control de la oferta monetaria.

Después de la Segunda Guerra Mundial, el yen perdió gran parte de su valor anterior a la guerra. Para estabilizar la economía japonesa, el tipo de cambio del yen se fijó en 360 yenes por dólar estadounidense como parte del sistema de Bretton Woods. Cuando se abandonó ese sistema en 1971, el yen quedó infravalorado y se le permitió flotar. El yen se había apreciado hasta un máximo de 271 yenes por dólar estadounidense en 1973, luego pasó por períodos

de depreciación y apreciación debido a la crisis del petróleo de 1973, llegando a un valor de 227 yenes por dólar estadounidense en 1980.

Desde 1973, el gobierno japonés ha mantenido una política de intervención monetaria y, por lo tanto, el yen se encuentra bajo un régimen de flotación administrada o flotación sucia (dirty float, en inglés). El gobierno japonés se centró en un mercado de exportación competitivo y trató de garantizar un tipo de cambio bajo para el yen a través de un superávit comercial. El Acuerdo Plaza de 1985 [49] cambió temporalmente esta situación: el tipo de cambio cayó de su promedio de 239 yenes por dólar estadounidense en 1985 a 128 yenes en 1988 y llevó a un tipo máximo de 80 yenes frente al dólar estadounidense en 1995, aumentando efectivamente el valor del producto interno bruto (PIB) de Japón en términos de USD a casi el de los Estados Unidos. Desde entonces, sin embargo, el precio mundial del yen ha disminuido considerablemente. El Banco de Japón mantiene una política de tasas de interés de cero a casi cero y el gobierno japonés ha tenido previamente una estricta política anti-inflacionaria.

La Libra Esterlina, £, GBP [50]: Es la moneda del Reino Unido así como de las Dependencias de la Corona y de algunos territorios británicos de ultramar. En sus demás territorios coloniales se usan diferentes divisas pero fijadas a la libra esterlina: la libra gibraltareña, libra malvinense y libra de Santa Elena. Su símbolo monetario es £ y proviene del latín libra, que se refería a la unidad de masa. Una libra se divide en cien peñiques.

Las Dependencias de la Corona usan sus propios códigos (no ISO) cuando desean reflejar su distinción. Los recursos generalmente son intercambiados por peñiques, así que los comerciantes pueden referirse al peñique esterlino, a veces abreviado GBp cuando enlistan sus precios.

Después de la adopción del euro por varios países, la libra se convirtió en la divisa más antigua del mundo aún en uso, y actualmente tiene la tercera porción más grande de monedas de reserva a nivel mundial, después del dólar estadounidense y el euro. La libra esterlina es la cuarta divisa más intercambiada en el mercado internacional de intercambio detrás del dólar, el euro y el yen japonés.

El Renminbi, ¥, CNY [51]: Es la moneda oficial de la República Popular China y una de las monedas de reserva del mundo, clasificándose como la octava moneda más negociada del mundo a partir de abril de 2019. Si bien su código es CNY, se suele abreviar su nombre Renminbi como RMB.

El yuan es la unidad básica del renminbi, pero la palabra también se usa para referirse a la

moneda china en general, especialmente en contextos internacionales. Un yuan se divide en diez jiao y un jiao a su vez se divide en diez fen. El renminbi es emitido por el Banco Popular de China, la autoridad monetaria de China.

Hasta 2005, el valor del renminbi estaba vinculado al dólar estadounidense. A medida que China avanzaba en su transición de la planificación central a una economía de mercado y aumentaba su participación en el comercio exterior, el renminbi se devaluó para aumentar la competitividad de la industria china. Anteriormente se afirmó que el tipo de cambio oficial del renminbi estaba subvaluado hasta en un 37.5 % frente a su paridad de poder adquisitivo. Sin embargo, más recientemente, las acciones de apreciación del gobierno chino, así como las medidas de flexibilización cuantitativa adoptadas por la Reserva Federal estadounidense y otros bancos centrales importantes, han provocado que el renminbi se sitúe en tan solo un 8 % de su valor de equilibrio en el segundo semestre de 2012. Desde 2006, se ha permitido que el tipo de cambio del renminbi flote en un margen estrecho en torno a un tipo base fijo determinado con referencia a una cesta de monedas mundiales. El gobierno chino ha anunciado que aumentará gradualmente la flexibilidad del tipo de cambio. Como resultado de la rápida internacionalización del renminbi.

El 1 de octubre de 2016, el renminbi se convirtió en la primera moneda de mercado emergente incluida en la canasta de derechos especiales de giro del FMI, la canasta de monedas que utiliza el FMI como moneda de reserva.

Al igual que con las criptomonedas, existe un mercado exclusivo de intercambio de dinero fiat. El mercado de divisas o mercado cambiario es un mercado que se caracteriza por el libre cambio de divisas, es decir, su objetivo principal es el de facilitar el comercio internacional y la inversión. También se conoce como intercambio de monedas extranjeras o Forex (**Foreign Exchange**, en inglés) [52].

Este mercado determina los tipos de cambio de cada divisa. Incluye todos los aspectos de compra, venta e intercambio de divisas a precios actuales o determinados. En términos de volumen de negociación, es, con mucho, el mercado más grande del mundo, seguido por el mercado de crédito.

Los principales participantes en este mercado son los bancos internacionales más grandes. Los centros financieros de todo el mundo funcionan como anclas de negociación entre una amplia gama de múltiples tipos de compradores y vendedores durante todo el día, con la excepción de los fines de semana. Dado que las divisas siempre se negocian en pares, el mercado de divisas no establece el valor absoluto de una divisa, sino que determina su valor relativo al establecer el precio de mercado

de una divisa si se paga con otra.

Forex existe para que se puedan cambiar grandes cantidades de una moneda por el valor equivalente en otra moneda al tipo de cambio actual del mercado.

Algunas de estas transacciones ocurren porque las instituciones financieras, las empresas o las personas tienen la necesidad comercial de cambiar una moneda por otra. Por ejemplo, una empresa estadounidense puede cambiar dólares estadounidenses por yenes japoneses para pagar la mercancía que se ha pedido a Japón y que se paga en yenes.

En el pasado, el comercio de divisas se limitaba en gran medida a los gobiernos, las grandes empresas y los fondos de cobertura. Ahora, cualquiera puede operar en forex. Muchas empresas de inversión, bancos y corredores minoristas permiten a las personas abrir cuentas y operar con divisas.

Al operar en el mercado de divisas, está comprando o vendiendo la moneda de un país en particular, en relación con otra moneda. Pero no hay intercambio físico de dinero de una parte a otra como en un quiosco de cambio de divisas.

En el mundo de los mercados electrónicos, los comerciantes generalmente toman una posición en una moneda específica con la esperanza de que habrá algún movimiento alcista y fortaleza en la moneda que están comprando (o debilidad si están vendiendo) para que puedan hacer una ganancia.

Una divisa siempre se negocia en relación con otra divisa. Si vendes una moneda, estás comprando otra, y si compras una moneda, estás vendiendo otra. El beneficio se obtiene de la diferencia entre los precios de transacción [53].

Cómo se diferencia Forex de otros mercados:

Menos reglas: Esto significa que los inversores no están sujetos a normas o reglamentos tan estrictos como los de los mercados de acciones, futuros u opciones. No hay cámaras de compensación ni organismos centrales que supervisen todo el mercado de divisas.

Tarifas y Comisiones: Dado que el mercado no está regulado, las tarifas y comisiones varían ampliamente entre los corredores. La mayoría de los corredores de divisas ganan dinero marcando el diferencial en los pares de divisas. Otros ganan dinero cobrando una comisión, que fluctúa según la cantidad de moneda negociada. Algunos corredores usan ambos.

Acceso completo: No hay límite en cuanto a cuándo puede y no puede operar. Debido a que el mercado está abierto las 24 horas del día, puede operar en cualquier

momento del día. La excepción son los fines de semana o cuando ningún centro financiero global está abierto debido a un feriado.

Apalancamiento: El mercado de divisas permite un apalancamiento de hasta 50:1 en los Estados Unidos e incluso más alto en algunas partes del mundo. Eso significa que un operador puede abrir una cuenta por \$1,000 y comprar o vender hasta \$50,000 en divisas. El apalancamiento es una espada de doble filo; magnifica tanto las ganancias como las pérdidas.

Claro que si hablamos de Forex, resalta el tema que estás siempre cambiando una moneda por otra dentro de un mismo ecosistema, fuera de este mercado, la manera de transferir dinero fíat entre bancos de distintos países y/o continentes con forma de una simple transferencia de un cierto monto es un poco más complicado y para estos casos es común que se recurra a otro tipo de métodos.

Detrás de la mayoría de las transferencias internacionales de dinero y valores se encuentra el sistema de la Sociedad para las Telecomunicaciones Financieras Interbancarias Mundiales (Society for Worldwide Interbank Financial Telecommunications o **SWIFT** por sus siglas en inglés). El SWIFT es una amplia red de mensajería que los bancos y otras instituciones financieras utilizan para enviar y recibir información de forma rápida, precisa y segura, como instrucciones de transferencia de dinero.

Más de 11,000 instituciones miembros globales de SWIFT enviaron un promedio de 42 millones de mensajes por día a través de la red en 2021, lo que representa un aumento del 11.4% con respecto a 2020.

SWIFT es una red de mensajería que utilizan las instituciones financieras para transmitir información e instrucciones de forma segura a través de un sistema estandarizado de códigos. Aunque SWIFT se ha convertido en una parte crucial de la infraestructura financiera mundial, no es una institución financiera en sí misma: SWIFT no posee ni transfiere activos. Más bien, su utilidad radica en su poder para facilitar una comunicación segura y eficiente entre las instituciones miembros.

El dominio global de SWIFT lo ha llevado a convertirse en una consideración importante en la geopolítica. En 2012, la Unión Europea sancionó a los bancos en Irán, desconectándolos del sistema SWIFT. En febrero de 2022, los líderes de Estados Unidos y la UE acordaron eliminar algunos bancos de Rusia de SWIFT [54].

A mediados de febrero de 2022 Estados Unidos, el Reino Unido y la Unión Europea han anunciado oleadas de sanciones contra Rusia para presionar al presidente Vladimir Putin a que

retroceda en su ataque contra Ucrania. Eso incluye sacar al país del sistema de comunicación de pagos global ampliamente utilizado, SWIFT [55].

Binance, el intercambio de criptomonedas más grande del mundo, en una entrevista para Fortune [56] comenta que no tiene planes para un congelamiento unilateral, aunque no negó que bloquearía las cuentas de personas específicas que se vean afectadas por las sanciones.

“No vamos a congelar unilateralmente las cuentas de millones de usuarios inocentes”, dijo la compañía a Fortune en un comunicado. “Crypto está destinado a proporcionar una mayor libertad financiera a las personas de todo el mundo. Decidir unilateralmente prohibir el acceso de las personas a sus activos en criptomonedas iría en contra de la razón por la que existe la criptomonedas”.

Los competidores de Kraken, Binance, Coinbase y FTX, también siguen operando en Rusia y, según la firma de datos Kakio, han visto más actividad en Europa del Este desde que comenzó el conflicto [55].

Lo cual revela un poco lo centralizado que están los intercambios de divisas y que suelen poseer cierto poder sobre los que lo usan ya que de no querer usarlos, te estarías enfrentando a un sin fin de consecuencias, es por esta razón que se han creado sistemas análogos al SWIFT pero entre otro grupo de empresas o países como lo son:

INSTEX [57]: El instrumento de apoyo a los intercambios comerciales (Instrument in Support of Trade Exchanges, en inglés) tiene el objetivo de facilitar el comercio legítimo con Irán a la luz de los desafíos que enfrentan los participantes del mercado europeo que desean continuar comerciando con Irán sin obstáculos, fue fundada en 2019 como una empresa con sede en París regida por la ley francesa.

CIPS [58]: Sistema de pago interbancario transfronterizo (Cross-border Interbank Payment System, en inglés) es un sistema de pago mayorista autorizado por el Banco Popular de China (PBC). Al especializarse en la compensación de pagos transfronterizos en RMB, CIPS se esfuerza por brindar servicios de compensación y liquidación confiables, eficientes, convenientes y de bajo costo. Como una de las importantes infraestructuras del mercado financiero (FMIs) en China, CIPS ha desempeñado un papel positivo en la contribución al desarrollo del centro financiero internacional de Shanghái, promoviendo la apertura bidireccional del sector financiero y fortaleciendo el apoyo financiero a la economía del sector real de China, sirviendo a la Iniciativa Belt and Road (BRI) financiando y facilitando el uso global de RMB [58].

SPFS [59]: El sistema de mensajería financiera del Banco de Rusia (SPFS por sus

siglas de la pronunciación en ruso y escrito en nuestro alfabeto “*Sistema Peredachi Finansovyykh Soobshcheniy*” forma un canal alternativo de mensajería financiera electrónica que a menudo es visto como análogo del SWIFT. SPFS garantiza la transmisión ininterrumpida de mensajes financieros tanto dentro del país como en el exterior [59].

Las instituciones de crédito y sus clientes corporativos se conectan al SPFS cuando están técnicamente preparados y establecen acuerdos contractuales con el Banco de Rusia. El procedimiento está definido por una regulación especial del Banco de Rusia.

Si bien hemos hablado de criptomonedas y dinero fíat, hay un tercer grupo que puede parecer muy distinto, pero es el que le dio origen a los ya mencionados y son los metales preciosos. Claro que si bien una amplia gama de metales se comercian día a día, solo un grupo selecto conforman los metales preciosos y algunos de estos están incluidos en la lista del estándar ISO 4217 dedicado a códigos de monedas, ya que son vistos tanto como una buena opción para resguardar valor como para efectuar pagos y cobros de ser requerido.

1.2.3. Metales Preciosos.

Los metales preciosos son elementos químicos metálicos raros y naturales de alto valor económico. Químicamente, los metales preciosos forman parte de un grupo de metales que tienden a ser menos reactivos en comparación a otros metales, a este grupo se les denomina, metales nobles. Por lo general, son dúctiles y tienen un alto brillo. Históricamente, los metales preciosos fueron importantes como moneda, pero ahora se consideran principalmente como bienes de inversión e industriales. El oro, la plata, el platino y el paladio tienen cada uno un código de moneda según el estándar ISO 4217 y pese a esto no es común conseguir datos de cotizaciones de estas consideradas monedas con respecto a otras monedas fíat, lo más común es conseguir su valor expresado en USD, así como es el caso de muchos productos intercambiados a nivel mundial como el petróleo, para el cual la Organización de Países Exportadores de Petróleo (OPEP o más conocida por sus siglas en inglés Organization of the Petroleum Exporting Countries, OPEC) presenta el precio de su cesta diaria en USD [60].

Los metales preciosos más conocidos son los metales de acuñación, que son el oro y la plata. Aunque ambos tienen usos industriales, son más conocidos por sus usos en el arte, la joyería y la acuñación de monedas. Otros metales preciosos incluyen los metales del grupo del platino: rutenio, rodio, paladio, osmio, iridio y platino, de los cuales el platino es el más comercializado. La demanda

de metales preciosos está impulsada no sólo por su uso práctico, sino también por su papel como inversiones y depósito de valor.

Usualmente el descubrimiento de nuevas fuentes de mineral o mejoras en los procesos de extracción o refinación puede hacer que disminuya el valor de un metal precioso. El estado de un metal precioso también puede estar determinado por la alta demanda o el valor de mercado. Los metales preciosos a granel se conocen como lingotes y se comercializan en los mercados de productos básicos. Los metales en lingotes pueden fundirse en lingotes o acuñarse en monedas. El atributo definitorio de los lingotes es que se valora por su masa y pureza más que por su valor nominal como dinero. Una de las características fundamentales cuando hablamos de su presencia en joyería es que a veces el metal precioso en cuestión no presenta todas las cualidades mecánicas que serían deseables para su uso en orfebrería, joyería o bisutería, por eso en este campo es usual el uso de aleaciones [61]

Muchas naciones acuñan monedas en lingotes. Aunque nominalmente emitidas como moneda de curso legal, el valor nominal de estas monedas como moneda es muy inferior al de su valor como lingotes. Por ejemplo, Canadá acuña varias monedas de oro en lingotes de distintos pesos y distintos diseños, entre ellos esta el “W” *Mint Mark: Gold Maple Leaf (Winnipeg, 2021)* [62] con un valor nominal de \$50 que contiene 1oz (31.16 g) de oro; para marzo de 2022 representa alrededor de \$1800 por su precio en oro y \$3414.77 como lingote. La acuñación de monedas en lingotes por parte de los gobiernos nacionales les otorga cierto valor numismático además de su valor en lingotes, además de certificar su pureza.

Algo importante a tomar en cuenta es que si se desea invertir en metales preciosos hay dos posibles caminos, pero uno sólo de ellos es de manera directa:

Como primera opción tenemos comprar metales físicamente, tanto en su forma de lingotes o monedas usualmente de puridades mayores a 90 % o en forma de joyería, en este último caso la pureza podría llegar a bajar hasta 58 %, ambos acarrearán consecuencias de logística como almacenamiento y transporte del mismo.

Como segunda opción tenemos invertir en acciones de metales preciosos y fondos cotizados en bolsa (Exchange Traded Funds, ETF por sus siglas en inglés). Las principales ventajas de las acciones y los ETF son que no tiene que almacenar el metal y existe la posibilidad de obtener dividendos. Además de las acciones mineras individuales, también hay fondos mutuos disponibles que invierten parcial o exclusivamente en empresas mineras. Estos pueden proporcionar diversificación de metales preciosos como oro, platino, paladio y plata. Las acciones de oro no necesariamente se mueven de acuerdo con los precios de los lingotes, porque las empresas mineras tienen éxito o

fracasan en función de su desempeño operativo individual. [63]

Ahora bien veamos más a detalle los metales que poseen código ISO y suelen poseer datos de cotizaciones en dólares en los mercados de intercambio. A continuación su nombre y su código respectivo.

Oro, XAU [64–66]: El principal problema del oro y los demás metales preciosos es que a diferencia de otras materias primas como el petróleo o el trigo, no se agota ni se consume. Una vez que se extrae el oro, permanece en el mundo. Independientemente del destino final del oro, su composición química es tal que el metal precioso no puede agotarse, es permanente. Debido a esto, el argumento de oferta/demanda que se puede hacer para materias primas como el petróleo, no se sostiene tan bien como para el oro. En otras palabras, la oferta solo aumentará con el tiempo, incluso si la demanda del metal se agota.

Desde una perspectiva fundamental, el oro generalmente se considera una cobertura favorable contra la inflación. El oro funciona como una buena reserva de valor frente a una moneda en declive y aunque se piense que ya no, aun hay lugares donde se usa como moneda de intercambio día a día.

Nuestros antepasados se enfrentaron a encontrar un método de intercambio que fuera más fácil de implementar que un sistema de trueque. Una moneda es uno de esos medios de intercambio. De todos los metales en la tabla periódica de elementos, el oro es la elección lógica. Podemos descartar elementos distintos a los metales porque una moneda gaseosa o líquida no es muy práctica desde el punto de vista de la portabilidad personal. Esto deja metales como hierro, cobre, plomo, plata, oro, paladio, platino y aluminio.

Hierro, plomo, cobre y aluminio son propensos a corroerse con el tiempo, por lo que no serían un buen valor en términos de almacenamiento, el aluminio se siente muy liviano e insustancial, no es ideal para una moneda de metal que podría invocar sentimientos de seguridad y valor. El platino o el paladio son opciones razonables porque en su mayoría no reaccionan con otros elementos, es decir, producen poca corrosión, pero son demasiado raros para generar suficientes monedas para circular. Para asignar valor a un metal, debe ser algo raro, de modo que no todos produzcan monedas, pero lo suficientemente disponible como para que se pueda crear una cantidad razonable de monedas para el comercio. Quedándonos así el oro y la plata. El oro no se corroe y se puede derretir sobre una llama, lo que facilita trabajar con él y estamparlo como una moneda. La plata y el oro son metales hermosos que son fáciles de convertir en joyas, y ambos metales preciosos tienen sus propios devotos en los

círculos de joyería fina.

Muchas personas no se dan cuenta de que el oro es una moneda en el sistema actual. A menudo se ha pensado en el oro en relación con el dólar estadounidense, principalmente porque su precio suele ser en dólares estadounidenses. Existe una correlación negativa a largo plazo entre el dólar y los precios del oro. Estos factores deben ser considerados cuando vemos que el precio del oro es simplemente un tipo de cambio. Así como uno puede cambiar dólares estadounidenses por yenes japoneses, una moneda de papel puede cambiarse por oro. El oro también desempeñó un papel esencial en el origen del dinero.

Para junio de 2021, los cinco países líderes en minería de oro son: China 370t, Australia 330t, Rusia 300t, Estados Unidos 180t y Canadá con 170t [67].

Plata, XAG [68–71]: Las monedas de plata son posiblemente la forma de acuñación más antigua producida en masa. La plata se ha utilizado como metal de acuñación desde la época de los griegos; sus dracmas de plata eran monedas comerciales populares. Los antiguos persas usaron monedas de plata entre el 612 aC y el 330 aC. Antes de 1797, los centavos británicos estaban hechos de plata.

Al igual que con todas las monedas coleccionables, muchos factores determinan el valor de una moneda de plata, como su rareza, demanda, condición y el número acuñado originalmente. Las monedas de plata antiguas codiciadas por los coleccionistas incluyen el denario y el miliarensis, mientras que las monedas de plata coleccionables más recientes incluyen el dólar Morgan. Además de las monedas de plata de colección, las monedas de lingotes de plata son populares entre las personas que desean una “cobertura” contra la inflación de la moneda o una reserva de valor.

Los inversores y comerciantes compran plata a través de los mercados de materias primas. Existen mercados comunes de materias primas para metales preciosos en Japón, Londres, Europa continental y los Estados Unidos. Las personas pueden comprar plata en barras, monedas y lingotes. Los inversores también pueden comprar activos respaldados por el metal precioso sin tener que poseer el producto básico real, ETF, acciones en compañías de plata y fondos mutuos.

A diferencia del oro que se considera mayormente como una reserva de valor, la plata también se beneficia de su amplio uso en muchas aplicaciones industriales. El metal ha establecido usos en el sector automotriz, en varios productos electrónicos, en paneles solares y en fotografía. Las nuevas tecnologías, como las baterías de óxido de plata, las tintas conductoras de plata

y varias nanotecnologías basadas en plata con aplicaciones médicas se están convirtiendo rápidamente en estándares industriales.

Esta demanda industrial hace que los precios de la plata sean más volátiles que los del oro y en general reactivos a varias medidas de datos de fabricación. Dado este hecho, los ETF que rastrean los precios o los futuros de la plata podrían ser una mejor apuesta que los lingotes físicos, ya que se pueden vender con bastante facilidad si los inversores piensan que los precios son demasiado volátiles.

El precio de la plata está determinado por muchos de los mismos factores que el oro: la tasa de inflación y las expectativas de inflación, el crecimiento del PIB global y las tasas de interés, así como las políticas monetarias de algunos de los bancos centrales más grandes del mundo: la Reserva Federal de Estados Unidos, el Banco de Inglaterra, el Banco de Japón y el Banco Popular de China. Hasta cierto punto, los factores de oferta/demanda tienden a desempeñar un papel más importante en las fluctuaciones de precios de la plata que en las del oro debido a su relativa falta de liquidez en el mercado. Por último, el sentimiento del mercado juega un papel importante en el precio del “oro de los pobres”, por lo que los comerciantes de plata querrán estar atentos a los datos de confianza del consumidor, así como a los movimientos en muchos de los mercados mundiales de bonos y acciones.

Los siete países con mayor producción mundial en toneladas métricas son: México con 5,600t, China y Perú con 3,400t y 3,000t respectivamente, Chile 1,600t y Australia, Polonia y Rusia 1,300t cada una [67].

Luego del oro y la plata, existe un subgrupo de metales integrado por: paladio, platino, rodio, rutenio, iridio y osmio, estos forman un grupo de elementos denominados metales del grupo del platino (Platinum-Group Metals, PGM por sus siglas en inglés) [72]. Estos también catalogados como metales preciosos, utilizados en catalizadores para la producción de productos químicos a granel y refinación de petróleo, dispositivos dentales y médicos, aparatos electrónicos como en discos duros de computadora, circuitos integrados híbridos y condensadores cerámicos multicapa; fabricación de vidrio, inversión, joyas, y equipo de laboratorio.

Los precios medios anuales estimados del paladio, el platino y el rutenio aumentaron un 18 %, 35 % y 88 % respectivamente, en comparación con los de 2020, los precios estimados del rodio se duplicaron y los del iridio se triplicaron. Por otra parte se espera que entre 2021 y 2022 la producción de automóviles restringida debido a la escasez de chips semiconductores y una disminución en la producción de vehículos de pasajeros diésel resulten en una disminución de la demanda de paladio,

platino y rodio utilizados en convertidores catalíticos, lo que podría influir en una bajada del precio de los mismos. Veamos un poco más sobre el paladio y el platino.

Paladio XPD [73–75]: El paladio es un componente importante en la electrónica y se utiliza en muchas tecnologías nuevas, como las pilas de combustible. Como materia prima, ha llamado la atención de los inversores porque no es fácil de sustituir por otros metales. Por ejemplo, el elemento es un componente importante de los convertidores catalíticos.

El paladio es treinta veces más raro que el oro. Esta rareza afecta su precio en los mercados de materias primas y el metal alcanzó máximos históricos de más de \$2981 por onza en febrero de 2022. El mínimo histórico de \$42.20 se estableció en agosto de 1977.

Los inversores pueden ganar exposición al paladio de varias formas. Una de las formas más sencillas de hacerlo es invertir en ETF centrados en el paladio que rastrean los lingotes de paladio, como *Sprott Physical Platinum and Palladium Trust (SPPP.U)* [76] o *Aberdeen Physical Palladium Shares (PALL)* [77]. Estos ETF contienen paladio físico, lo que brinda el beneficio de no tener que almacenar el metal por sí mismo, pero aun así obtener exposición a su valor. La forma opuesta de invertir en paladio en forma de ETF es comprar lingotes físicos pero esto genera el costo de almacenamiento del metal en físico.

Otra opción para invertir en paladio es comprar acciones de empresas que están involucradas en el negocio del paladio, generalmente su extracción y venta. En este caso, puede ser difícil concentrarse solo en el paladio, ya que las empresas mineras suelen extraer una multitud de metales en lugar de solo uno, por lo que estará expuesto a más metales que solo paladio.

El paladio ha sido sustituido por platino en la mayoría de los convertidores catalíticos de motores de gasolina debido al precio llegando a costar el paladio el doble que el platino. Aproximadamente el 25 % del paladio se puede sustituir rutinariamente por platino en los convertidores catalíticos diésel; la proporción puede llegar al 50 % en algunas aplicaciones. Para algunos usos finales industriales, un PGM puede sustituir a otro, pero con pérdidas de eficiencia.

Los cinco países con mayor producción de paladio son: Sur África con 80,000kg, seguido de Rusia con 74,000kg, Canadá 17,000kg, Estados Unidos 14,000kg y por último Zimbabue con 13,000kg [67].

Platino, XPT [78, 79]: Por su parte el platino también se negocia las 24 horas del día en los mercados mundiales de materias primas. El platino se considera un metal industrial. La mayor

demanda de platino proviene de los catalizadores para automóviles al igual que el paladio, se utiliza para reducir la nocividad de las emisiones. Después de esto, las joyas representan la mayor parte de la demanda. Los catalizadores de refinación de petróleo y químicos y la industria informática consumen el resto.

Debido a la fuerte dependencia de la industria automotriz en el metal, los precios del platino están determinados en gran parte por las ventas de automóviles y las cifras de producción. La legislación de ‘aire limpio’ [80] podría requerir que los fabricantes de automóviles instalen más convertidores catalíticos, aumentando la demanda.

El platino es uno de los elementos más valiosos del mundo y se considera uno de los metales preciosos más costosos. Sin embargo, mientras que el platino se negoció con una prima significativa frente al oro durante décadas, no lo ha hecho desde 2008. Desde la crisis financiera de 2007-2008, el platino generalmente se ha comportado peor que otros metales como el oro, la plata y el paladio.

Los observadores del mercado creen que una caída en los mercados del platino en 2008 ahuyentó a la clase inversora del metal, dejando a las industrias automotriz y de joyería como la única fuente de demanda de platino. Además, las minas sudafricanas han aumentado sustancialmente la producción de platino desde 2014 y se han sumado a los suministros globales.

Los cinco países con mayor producción de Platino en kilogramos son: Sur África con 130,000kg, Rusia 19,000kg, Zimbabue 15,000kg, Canadá 6,000kg y Estados Unidos con 4,200kg [67].

Si bien estos dos últimos metales pueden no sonar tan común como los dos anteriores, debido a su alto precio incluso más que el oro en el caso del paladio y un poco más de la mitad que la del oro para el platino y acompañado de una cantidad de usos inmensa han ganado relevancia, como el hecho de ser incluidos en la tabla de estándar de nomenclaturas de monedas, que se usen en joyería fina y como resguardo de valor en forma de monedas acuñadas o lingotes.

Ahora bien veamos algunos términos y definiciones que se aplican usualmente en el campo de la compra-venta de dinero, acciones y bonos, pero mas que todo en el campo de las criptomonedas y dinero fiat.

1.2.4. Conceptos de mercados de intercambio

Intercambio: Trading en inglés. Se denomina así a la acción de comprar y vender bienes, acciones o dinero en todas sus formas y tipos posibles. Si bien esta palabra viene del inglés

y la podríamos traducir simplemente como comercio, tomamos la decisión de seguir usando la palabra **trading**, al igual que otros términos, ya que estos se utilizan de manera común en el lenguaje español, en este caso para definir la acción de comprar y vender en su mayoría acciones, divisas en Forex o cualquier par de criptomoneda contra dinero fíat u otra criptomoneda.

De esta manera se le denomina “trader(s)” a la persona que realiza tal acción de hacer trading.

Análisis Fundamental (AF): El análisis fundamental es un método empleado por inversores y traders para tratar de determinar el valor intrínseco de activos o negocios. Para valorarlos de forma precisa, estudiarán de una forma rigurosa tanto factores internos como externos, con el objetivo de determinar si el activo o negocio en cuestión se haya sobrevalorado o infravalorado. Estas conclusiones podrán ayudar a formular de manera más adecuada una estrategia que tenga más posibilidades de generar buenos retornos [81].

Los analistas suelen estudiar en orden el estado general de la economía y luego la fortaleza de la industria específica antes de concentrarse en el desempeño de la empresa individual para llegar a un valor de mercado justo para las acciones.

El análisis fundamental utiliza datos públicos para evaluar el valor de una acción o cualquier otro tipo de valor. Por ejemplo, un inversionista puede realizar un análisis fundamental sobre el valor de un bono observando factores económicos como las tasas de interés y el estado general de la economía, y luego estudiando información sobre el emisor del bono, como cambios potenciales en su calificación crediticia.

Para las acciones, el análisis fundamental utiliza los ingresos, las ganancias, el crecimiento futuro, el rendimiento del capital, los márgenes de beneficio y otros datos para determinar el valor subyacente y el potencial de crecimiento futuro de una empresa. Todos estos datos están disponibles en los estados financieros de una empresa (más sobre eso a continuación).

Análisis Técnico (AT): a menudo denominado “chartismo”, es un tipo de análisis enfocado a predecir el comportamiento futuro del mercado en base a las acciones del precio y datos sobre volúmenes previos. El enfoque AT se aplica de forma extensiva a valores y otros activos de los mercados financieros tradicionales, pero también es un componente integral del trading de monedas digitales en el mercado de criptomonedas.

Las herramientas de análisis técnico se utilizan para examinar las formas en que la oferta y la demanda de un valor afectarán los cambios en el precio, el volumen y la volatilidad implícita. Opera a partir de la suposición de que la actividad comercial anterior y los cambios de precio

de un valor pueden ser indicadores valiosos de los movimientos futuros del precio del valor cuando se combinan con las reglas de inversión o negociación adecuadas.

A menudo se usa para generar señales comerciales a corto plazo a partir de varias herramientas de gráficos, pero también puede ayudar a mejorar la evaluación de la fortaleza o debilidad de un valor en relación con el mercado más amplio o uno de sus sectores. Esta información ayuda a los analistas a mejorar su estimación de valoración general [82].

Cabe señalar que el AT se considera más fiable y efectivo en mercados que operan bajo condiciones normales es decir, con un elevado volumen y liquidez. Los mercados de gran volumen están menos expuestos a la manipulación de precios y a influencias externas anormales que pudieran crear falsas señales lo que haría del AT una herramienta inútil [83].

Áreas de soporte y resistencia: Desde un punto de vista general, el soporte y la resistencia son conceptos simples. El precio encuentra un nivel que no es capaz de atravesar, lo que hace que el mismo actúe como una especie de barrera. En el caso del soporte, el precio encuentra un “suelo”, mientras que en el caso de la resistencia, encuentra un “techo”. Básicamente, puedes imaginarte el soporte como una zona de demanda, y la resistencia como una zona de oferta. Aunque de forma más tradicional soporte y resistencia se indique como líneas, los ejemplos del mundo real no son tan precisos. Hay que tener en cuenta que los mercados no se ven dirigidos por una ley física que impide que atraviesen cierto nivel. Es por ello que puede resultar más beneficioso imaginarte el soporte y la resistencia como áreas. Se pueden considerar dichas áreas como rangos en el gráfico de un precio, que probablemente provocarán en los traders una mayor actividad.

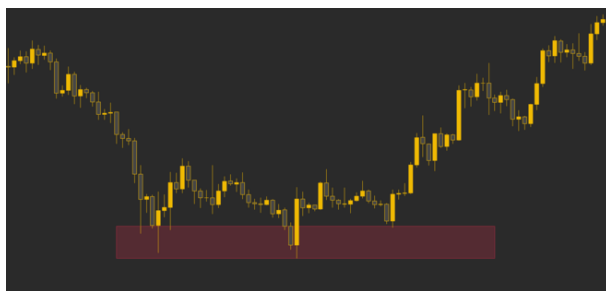


Figura 1.3: Ejemplo de Área de Soporte

En la figura 1.3 podemos observar un ejemplo de un nivel de soporte [84], el precio continuamente entra en un área en la que el activo es comprado. Se ha formado un rango de soporte a medida que el área ha sido testeada múltiples veces. Y dado que los vendedores no han

sido capaces de llevar el precio más abajo, al final ha rebotado lo que potencialmente puede suponer el inicio de una nueva tendencia al alza.

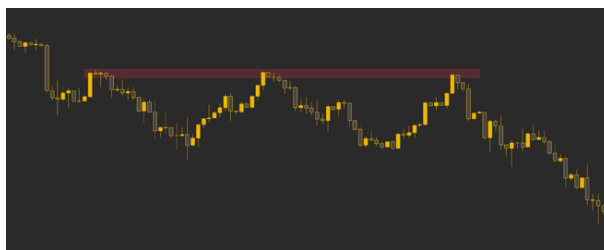


Figura 1.4: Ejemplo de Área de Resistencia

Ahora fijémonos en el nivel de resistencia mostrado en la figura 1.4 [84], allí se puede ver que el precio se hallaba en una tendencia a la baja pero después de cada rebote, no era capaz de atravesar una misma área múltiples veces. El nivel de resistencia se forma porque los compradores no han sido capaces de tomar el control del mercado y conducir el precio hacia el alta lo que provoca que la tendencia a la baja prosiga.

Por otro lado, dentro de los soportes y resistencia, existe un subgrupo de tipos, dependiendo el origen del fenómeno.

Soportes y resistencias psicológicos [84]: El primer tipo que analizaremos es el de los denominados soportes y resistencias psicológicos. Dichas áreas no presentan necesariamente una correlación con cualquier patrón técnico, pero existen debido a cómo la mente humana intenta darle sentido al mundo. Debido a la complejidad del mundo en el que vivimos, diariamente estamos en busca de simplificar cada detalle o cada cosa que este a nuestro alcance simplificar, y así intentar darle un mayor sentido; esto incluye redondear los números hacia arriba.

Un efecto similar está en juego en los mercados financieros. Es especialmente cierto para el trading de criptomonedas, que implica unidades digitales fácilmente divisibles. Comprar un activo a \$8.0674 y venderlo a \$9.9765 simplemente no se procesa de la misma manera que comprarlo a \$8 y venderlo a \$10. Es por eso que los números redondos también pueden actuar como soporte o resistencia en un gráfico de precios.

Como vemos en la figura 1.5, a medida que el *US Dollar Index* (DXY) se acerca a 100, algunos traders colocan órdenes de venta justo debajo de ese nivel para asegurarse de que esas órdenes se llenen, de esta forma el mercado nunca lo alcanza y se revierte justo antes [84]. Este fenómeno se ha hecho conocido a lo largo de los años. Como tal, algunos traders pueden tratar de “adelantar” el apoyo psicológico obvio o las áreas de resistencia.



Figura 1.5: Ejemplo de soporte y resistencia psicológico

Inversión ventajista, en este caso, significa hacer pedidos justo por encima o por debajo de un área de soporte o resistencia anticipada.

Soporte y resistencia de la línea de tendencia. Es común ver que patrones de gráficos también actúan como barreras para el precio. En la figura 1.6 se puede observar un triángulo ascendente mantiene el precio del índice *Standard & Poor's 500* (S&P 500) contenido hasta que el patrón se rompe al alza [84].



Figura 1.6: Ejemplo de soporte y resistencia de la línea de tendencia.

Se pueden utilizar estos patrones como ventaja para identificar áreas de soporte y resistencia que coincidan con las líneas de tendencia. Pueden ser especialmente útiles si se logran detectar temprano, antes de que el patrón se desarrolle completamente.

Soporte y resistencia debido a indicadores: Muchos indicadores también pueden proporcionar soporte o resistencia cuando interactúan con el precio. Uno de los ejemplos más directos de esto son las medias móviles. Como una media móvil actúa como soporte o resistencia para el precio, muchos traders la usan como un instrumento para medir la salud general del mercado. Las medias móviles también pueden ser útiles al tratar de detectar reversiones de tendencia o puntos de pivote. En la figura 1.2.4 se puede ver como una media móvil de 200 semanas actúa como área de soporte para el precio de Bitcoin [84].

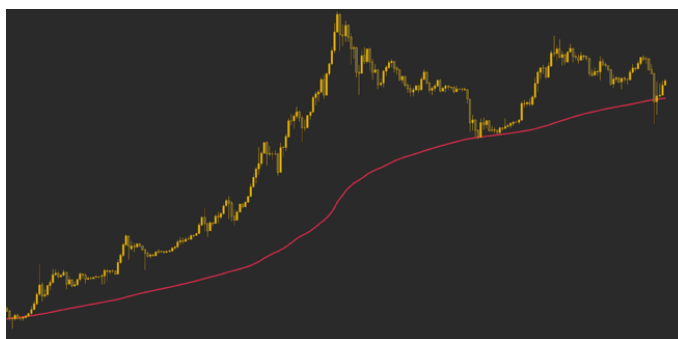


Figura 1.7: Ejemplo de soporte y resistencia debido media móvil

Confluencia en el análisis técnico: Hasta ahora, hemos discutido qué son el soporte y la resistencia, y algunos de sus diferentes tipos. Pero, ¿cuál es la forma más efectiva de desarrollar estrategias de trading a su alrededor?. Una cosa clave para entender es un concepto llamado confluencia. La confluencia es cuando una combinación de múltiples estrategias se usan juntas para crear una estrategia. Los niveles de soporte y resistencia tienden a ser más fuertes cuando caen en varias de estas categorías que hemos discutido [84].

Sobrecompra: es un término que se utiliza cuando se cree que un valor se cotiza a un nivel superior a su valor intrínseco o justo. La sobrecompra generalmente describe un movimiento reciente o de corto plazo en el precio del valor y refleja una expectativa de que el mercado corregirá el precio en un futuro cercano. La tendencia alcista puede deberse a noticias positivas sobre la empresa, la industria o el mercado subyacente en general. La presión de compra puede alimentarse a sí misma y conducir a un optimismo continuo más allá de lo que muchos comerciantes consideran razonable. Cuando este es el caso, los comerciantes se refieren al activo como sobrecomprado y muchos apostarán por una reversión en el precio [85].

Sobreventa: se refiere a una condición en la que un activo se ha negociado a un precio más bajo y tiene el potencial de un rebote de precios. Una condición de sobreventa puede durar mucho tiempo y, por lo tanto, estar sobrevendido no significa que un aumento de precios llegará pronto, o en absoluto. Muchos indicadores técnicos identifican los niveles de sobreventa y sobrecompra. Estos indicadores basan su evaluación en dónde se negocia actualmente el precio en relación con los precios anteriores. Los fundamentos también se pueden usar para evaluar si un activo está potencialmente sobrevendido y se ha desviado de sus métricas de valor típicas.

Esto podría ser el resultado de malas noticias con respecto a la empresa en cuestión, una mala perspectiva para la empresa en el futuro, una industria desfavorecida o un mercado

general en caída.

Los comerciantes también pueden usar indicadores técnicos para establecer niveles de sobreventa. Un indicador técnico solo mira el precio actual en relación con los precios anteriores. No tiene en cuenta los datos fundamentales [86].

Impulso: es la tasa de aceleración del precio de un valor, es decir, la velocidad a la que cambia el precio. El impulso del mercado se refiere a la tasa agregada de aceleración para el mercado más amplio en su conjunto.

El impulso del mercado se puede utilizar como una medida del sentimiento general del mercado que puede respaldar la compra y venta con y contra las tendencias del mercado. Es uno de varios indicadores que pueden ayudar a un inversor a seguir el impulso del mercado.

En general, el impulso del mercado se puede definir a partir de la siguiente ecuación:

$$M = V - V_x \quad (1.1)$$

Donde:

V = El último precio

V_x = El precio de cierre x número de días atrás.

Esta ecuación puede conducir al dibujo de una línea de tendencia con diferentes períodos utilizados en el cálculo.

El impulso positivo puede indicar una tendencia alcista potencial, mientras que el impulso negativo puede indicar una tendencia bajista. En términos generales, el impulso se puede medir tanto en clases de activos como en valores individuales, y el impulso del mercado, en particular, se refiere al mercado en general.

El comercio de impulso es una estrategia que busca capitalizar el impulso para entrar en una tendencia a medida que se acelera. Los inversores pueden utilizar el impulso como una técnica comercial que busca sacar provecho del comportamiento de pastoreo de la psicología del mercado. En lugar de “comprar barato, vender caro”, el comercio de impulso sigue una estrategia de “compra caro, vende caro”. Una vez que un comerciante de impulso ve una aceleración en el precio, las ganancias o los ingresos de una acción, el comerciante a menudo tomará una posición larga o corta en la acción con la esperanza de que su impulso continúe

en una dirección ascendente o descendente. Esta estrategia se basa en movimientos a corto plazo en el precio de una acción en lugar del valor fundamental.

Indicadores de impulso: son herramientas de análisis técnico que se utilizan para determinar la fortaleza o debilidad del precio de una acción. El impulso mide la tasa de subida o bajada de los precios de las acciones. Los indicadores de impulso comunes incluyen el índice de fuerza relativa (RSI) y la divergencia de convergencia del promedio móvil (MACD).

El impulso mide la tasa de subida o bajada de los precios de las acciones. Desde el punto de vista de la tendencia, el impulso es un indicador muy útil de la fortaleza o debilidad del precio de la emisión. La historia nos ha demostrado que el impulso es mucho más útil durante los mercados alcistas que durante los mercados bajistas; el hecho de que los mercados suban con más frecuencia de lo que caen es la razón de esto. En otras palabras, los mercados alcistas tienden a durar más que los mercados bajistas.

La fuerza relativa: es una estrategia utilizada en la inversión de impulso y en la identificación de acciones de valor. Se centra en invertir en acciones u otras inversiones que han tenido un buen desempeño en relación con el mercado en su conjunto o con un punto de referencia relevante. Por ejemplo, un inversor de fuerza relativa podría seleccionar empresas de tecnología que hayan superado al índice compuesto **Nasdaq** o acciones que estén superando al índice **S&P 500**.

Los analistas técnicos utilizan un indicador conocido como índice de fuerza relativa (RSI) para generar señales de sobrecompra o sobreventa.

Mientras que el objetivo de la inversión en valor es comprar barato y vender caro, el objetivo de la inversión de fuerza relativa es comprar caro y vender aún más caro. Como tal, los inversores de fuerza relativa asumen que las tendencias que actualmente muestra el mercado continuarán durante el tiempo suficiente para permitirles obtener un rendimiento positivo. Cualquier reversión repentina de esa tendencia dará lugar a resultados negativos.

Y así culminamos los tipos de dinero en los cuales se basará nuestra minería de datos, para esto necesitamos datos de cotizaciones de criptomonedas/fiat tanto históricas como actuales. Si bien lo que queremos es hacer minería de datos y ya aclaramos que estas cotizaciones serán nuestros datos.

Ahora veamos con detalle a qué llamamos datos, donde y cómo podemos conseguir los que necesitamos en particular.

1.3. Dato e información.

Aunque parezca clara la distinción entre ambos, se manejan como sinónimos en la vida cotidiana, pero no lo son. Un dato es un registro de un fenómeno en específico; éste comúnmente es un número, letra, representación cuantitativa o cualitativa, el cual de manera aislada difícilmente puede tener algún tipo de valor intrínseco; por el contrario la información es un conjunto organizado de datos relacionados que procesados e interpretados constituyen una visión general del hecho en cuestión.

Nuestra vida diaria está rodeada de todo tipo de instrumentos o programas los cuales registran un sin fin de datos, por ejemplo cada uno de nosotros generamos en nuestra vida cotidiana una amplia gama de datos de todo tipo, como las series que vemos en algún portal audiovisual o que tipo de crema dental usamos, claro que su recolección y procesamiento dependerá del valor que se espere extraer de esta información. En este punto es bueno resaltar que en los últimos años se ha visto cómo el acertado análisis y procesamiento de datos puede separar grandes estrategias económicas de fallidos negocios, lo que ha llevado a que se nombre a los datos como el petróleo del siglo XXI, ya que los buenos datos superan la opinión [87].

Es de esta manera como con el uso adecuado de datos y de modelos matemáticos, logramos llegar al entendimiento de los fenómenos y su posible predicción a corto, mediano o largo plazo. Hay consideraciones a tomar en cuenta, por un lado con respecto a los datos basta con tener en cuenta que sean provenientes de una fuente fiable, ya que si es cierto que no siempre obtendremos buenas conclusiones de datos correctos, datos incorrectos siempre nos llevarán a conclusiones erróneas. Con respecto a los modelos matemáticos a aplicar; no todo está escrito, claro que hay modelos de análisis de datos hoy día, distintos indicadores y técnicas de procesamiento las cuales tienen cierta efectividad, incluso algunas dependientes del contexto, es en este campo donde siempre podremos tratar de hacer algo nuevo con los datos esperando extraer más información de ellos.

En el caso de hablar sobre registros de un fenómeno en particular como por ejemplo el precio del Dólar/Euro, de poseer dos registros consecutivos, podemos registrar su cambio en el tiempo, de tener varios registros consecutivos podríamos representar su comportamiento a lo largo del tiempo y de contar con una cantidad representativa de registros incluso ir en búsqueda de algún tipo de patrón de la serie temporal, en este caso más siempre será más, mientras más datos poseamos de un fenómeno en específico más acertado será nuestro análisis.

Claro que en el caso de buscar patrones e intentar predecir su comportamiento, puede ser un camino muy escabroso ya que estos sistemas pueden tener agentes tanto directos como indirectos, tan indirectos como el caso de algunas criptomonedas que luego de ser nombradas por algún famoso

reporten máximos en sus precios.

Recordemos que el propósito de los datos y su procesamiento en el fondo, busca estudiar y explicar los sistemas sin importar de la índole que sean. Luego el aprovechamiento o no de ésta información dependerá de nosotros. En conclusión la información es generada por nosotros a partir de los datos, es nuestra interpretación de los mismos.

Nuestra aplicación busca ser una herramienta capaz de recolectar datos de distintas fuentes y facilitar su procesamiento mediante diversas técnicas con la intención de obtener información valiosa que nos ayude en la toma de decisiones. Con esto en mente, veremos a continuación todo lo necesario con respecto a fuentes de datos y nuestras primeras opciones.

1.4. Fuente de datos.

Como habíamos mencionado antes, nuestras conclusiones están basadas en los datos, si éstos presentan algún tipo de fallas ya sean datos erróneos, faltantes o algún otro problema, éstas resultarían comprometidas. De esta manera la veracidad y calidad de nuestra fuente es algo crucial para nuestra minería de datos. Las fuentes podemos dividirlos en tres grupos, según su cercanía al origen de los datos/hechos [88]:

- **Fuentes primarias:** presentan información de primera mano, lo más cercanas al origen de la información o idea objeto de estudio. Usualmente sin filtro, interpretación o evaluación. Aportan material en los que posiblemente se basaran investigaciones o reseñas posteriores. Suelen ser la primer aparición formal de los resultados de un hecho o toma de datos en formato físico, impreso o electrónico.
- **Fuentes secundarias:** generalmente son recopilaciones de fuentes primarias escritas después del hecho con el beneficio de la retrospectiva o sencillamente recopilaciones con el objeto de difundirlas, compararlas y/o contrastarlas. Las fuentes secundarias no son evidencia, sino comentarios y discusión sobre la evidencia. En algunos casos se pueden definir como fuente secundaria lo que otros definen como fuente terciaria. El contexto lo es todo.
- **Fuentes terciarias:** consisten en información que es una destilación y recopilación de fuentes primarias y secundarias.

Claro que también las podemos clasificar según su origen físico, como escuchar un relato, leer un libro, periódico o lo más moderno, visualizarlo en la internet. Este último claramente hoy día

puede llegar a englobar a cada uno de los otros, pasando a ser un medio por el cual accedemos a las anteriores mencionadas más que una fuente como tal.

El medio más usado hoy por hoy para acceder a información y descargar datos de todo tipo es la internet, mediante la cual podemos conocer en tiempo real datos e información de toda índole a nivel global, pero además también podemos dar a conocer datos e información de primera o segunda mano a cualquier rincón del mundo. Al usarla como medio de acceso a datos algo de suma importancia es tomar en cuenta la calidad de los datos o información a la que tendremos acceso.

Nuestras fuentes seleccionadas las cuales se explicarán en breve, son fuentes secundarias, fuentes que recopilan datos a nivel global para proporcionar una perspectiva general de los datos en cuestión, esto mediante el uso de la internet, con la posibilidad de estar interconectado a múltiples fuentes de datos en simultaneo. Estas fuentes de datos que usaremos, funcionan también como grandes almacenes de datos, proporcionando a los usuarios visualizaciones de datos actuales e históricos en forma de gráficos y tablas.

Ahora bien, si ya decidimos que usaremos fuentes secundarias consultadas mediante la internet, debemos tomar varios aspectos en cuenta, el primero que abordaremos es el hecho de como están almacenados los datos a los que queremos acceder, para luego extraerlos y procesarlos y que estén acorde a nuestras necesidades. En este capítulo abordaremos el primer aspecto que es almacenes de datos y ya en el segundo capítulo explicaremos más sobre su extracción y procesamiento.

1.4.1. Almacenes de datos.

En nuestro caso, almacenamiento de datos es la retención de información mediante el uso de tecnología desarrollada especialmente para guardar esos datos y mantenerlos lo más accesibles posible. Hoy día se refiere al uso de medios de grabación para conservar los datos utilizando PC y otros dispositivos. Las formas más frecuentes de almacenamiento de datos son el almacenamiento de archivos, el almacenamiento en bloque y el almacenamiento de objetos, cada uno de los cuales resulta adecuado para un fin diferente [89].

Almacenamiento de archivos: Económico y de infraestructura sencilla, los datos se almacenan en archivos y carpetas. Se encuentra frecuentemente en discos duros, lo que supone que los archivos tienen exactamente el mismo aspecto para el disco duro y para el usuario.

Almacenamiento en bloque: Los datos se almacenan en bloques de tamaño uniforme. Si bien es más caro y complejo, y menos escalable, el almacenamiento en bloque resulta ideal para los datos de acceso y edición frecuentes.

Almacenamiento de objetos: Los datos se almacenan en forma de objeto con metadatos e identificadores únicos. Aunque generalmente es más económico almacenar datos de esta forma, el almacenamiento de objetos es solamente ideal para almacenar datos que no necesitan ser editados.

Dirijamos nuestro enfoque a los de tipo almacenamiento de objetos; ya que son los de nuestro interés en específico.

Esta forma de almacenamiento es ideal para la conservación de cantidades masivas de datos a los que se debe acceder pero sin posibilidad de editarlos, el almacenamiento de objetos alude a una forma de almacenamiento de datos en la que los archivos se almacenan como objetos con metadatos asociados e identificadores únicos que permiten acceder a ellos en una red. El almacenamiento de objetos se emplea cada vez más para aplicaciones nativas de la nube, como el procesamiento de análisis, la inteligencia artificial y el aprendizaje automático.

En la práctica resulta casi infinitamente escalable, y esto, en un mundo en el que la tasa de datos que se generan aumenta exponencialmente, es una buena noticia. Las aplicaciones nativas de la nube modernas que emplean micro servicios utilizan ahora también el almacenamiento de objetos como su capa de datos persistente [90].

De esta manera, si bien ya sabemos que accederemos a almacenamiento de objetos y esto será mediante internet, pasemos ahora a describir factores generales a tener en cuenta con respecto a los almacenes de datos a la hora de seleccionar nuestra fuente.

Almacenamiento: en cuanto al almacenamiento de los datos, hemos destacado tres factores a tomar en cuenta.

1. Formato del archivo: si logramos acceder en línea o descargarlo, debemos contar con los programas o herramientas para poder leerlo y/o editarlo, claro que esto se podría complicar si tocásemos el mundo de multimedia o formatos complejos, pero en casos como el nuestro es común encontrarnos con archivos del tipo texto plano, compuestos de bytes que representan caracteres ordinarios como letras, números y signos de puntuación (incluyendo espacios en blanco) o que puedan ser abiertos como tal, eso si, ahora tendríamos que tener en cuenta si tienen algún tipo de formato de codificación distinto al usual UTF-8 (8-bit Unicode Transformation Format, en inglés).
2. Marca de tiempo: al recabar series temporales, es necesario agregarles su marca de tiempo a cada dato respectivamente, que por redundante que suene le

otorga al registro un identificador único de la generación o almacenamiento de dicho dato, junto a esto se han desarrollado distintos estándares para estar en sintonía a nivel mundial, el más usado y aceptado es el tiempo universal coordinado (UTC) bajo la norma ISO 8601 [91], la cual especifica la notación estándar utilizada para representar instantes, intervalos e intervalos recurrentes de tiempo evitando ambigüedades.

Claro que si bien el formato UTC es el más usado, la mayoría de las veces el navegador o dispositivo de donde estemos haciendo la visualización o consulta se encarga de hacer la equivalencia a nuestra zona horaria con datos de localización suministrados por nuestro navegador. En la tabla de la figura 1.8 podemos ver en varios formatos las dos de la tarde con once minutos y trece segundos del 31 de enero de 1995.

Nombre del formato	Formato	Ejemplo
Formato básico	YYYYMMDDTHHMMSS	19950131T141113
UTC	YYYYMMDDTHHMMSSZ	19950131T141113Z
Extendido	YYYY-MM-DDTHH:MM:SSZ	1995-01-31T14:11:13Z

Figura 1.8: Instantes de tiempo.

3. Frecuencia de los datos: Al realizar un estudio el cual esta relacionado con su estado y evolución temporal; es un factor fundamental, ya que a menor espaciamiento temporal entre los datos/registros, la reconstrucción o representación de lo sucedido tendrá mayor apreciación al cambio. Reflejado en una representación más fiel a la realidad. Podríamos verificar si ésta es estable a lo largo de los datos o es cambiante, para evitar posibles errores en el análisis de los datos.

Alcance: En cuanto al alcance de los datos, pudiese ser limitado su acceso según el tipo de usuario el cual realiza la consulta o según la cantidad de datos o peticiones al servidor en un tiempo determinado.

Acceso: Nos referimos a la vía para acceder a los datos (descargarlo), entre las más comunes tenemos: mediante visualización, enlaces de descarga, interfaz de programación de aplicaciones (Application Programming Interface, API en inglés), protocolos de intercambio de archivos entre iguales (Peer-to-Peer, P2P en inglés) o solicitudes mediante protocolos de transferencia de hipertexto (HyperText Transfer Protocol, HTTP en inglés). Algunos de estos métodos puede que tengan ciertos

filtros de verificación no computable o necesitar alguna información extra en la solicitud de los datos como podría ser el caso del origen de la petición o una clave pública.

Con esto en mente, veamos ahora las páginas más comunes de consulta de datos sobre cotizaciones de criptomonedas/fiat y a que datos solemos tener acceso.

1.5. Sitios web de cotizaciones.

Dado el auge emergente de las criptomonedas como ya hemos mencionando anteriormente, ha proliferado el surgimiento tanto de casas de cambio de criptomonedas por dinero fiat o solo criptomonedas por otras y a su vez sitios web destinados únicamente a propagar información y datos sobre dichos mercados generalizados o específicos, ya que algunos ofrecen información de primera mano sobre intercambios en específico y otros suelen mostrar información generalizada de varios intercambios combinada y agrupada respecto a cada criptomoneda en específico.

Hemos realizado una pre-selección de cinco de los sitios más frecuentados con el fin de obtener información y datos sobre mercados de criptomonedas, a continuación mientras los vayamos nombrando introduciremos también el acrónimo con el cual los nombraremos por el resto del documento.

Kraken [34], KRK: Con su lema *“Nuestra misión es acelerar la adopción de las criptomonedas para que usted y el resto del mundo puedan lograr la libertad financiera y la inclusión” [92]*. Es un intercambio de criptomonedas y banco, el cual también posee una página web repleta de información de todo lo que tiene que ver con el mundo de las criptomonedas, enlaces a blogs, podcasts [93] y por supuesto ofrece acceso a datos y visualizaciones sobre los intercambios dentro de su plataforma en tiempo real e histórico.

Con respecto a la visualización, cuenta con dos modalidades, avanzada y normal. En cuanto a la modalidad avanzada esta cuenta con una cantidad muy extensa de opciones a aplicar sobre el gráfico, además que en la misma ventana se cuenta con gráficos de cotizaciones, ordenes y volumen de compra-venta, es esta información y este portal el que usaríamos para comprar y vender activos si tuviésemos cuenta en KRK, por otro lado en la modalidad normal, contamos con un gráfico básico con opciones llamémoslas estándar, donde se muestra solo el precio y volumen de datos históricos en diversas escalas de tiempo.

CryptoCompare [94], CCO: se hace llamar una autoridad central para obtener información clara y concisa, que ofrece una amplitud, alcance y una profundidad de datos inigualables, cerrando la brecha entre los criptoactivos y los mercados financieros tradicionales. Es un proveedor global de datos de mercado de criptomonedas que brinda acceso a datos de mercado y precios en tiempo real. Agrega y procesa datos de intercambios reconocidos a nivel mundial. En su página principal cuenta con gráficos de diversas casas de cambio y distintas criptomonedas con datos en tiempo real. Luego tenemos tres opciones, una lista de criptomonedas y sus datos a tiempo real en USD y posibilidad de acceso a más información particular a cada criptomoneda, una sección de foro y una sección de noticias. Al seleccionar una criptomoneda, esto nos dirige a una página donde podremos observar plena información de la criptomoneda en cuestión como cotizaciones, noticias, libro de órdenes, visualización de transacciones en tiempo real y gráficos, algo a resaltar es que algunas opciones están bloqueadas, ya que para su acceso hace falta ingresar a una cuenta de CCO.

Bitcoincharts [95], BCH: Proporciona datos financieros y técnicos relacionados con la red Bitcoin los cuales también puede ser accedidos a través de su JSON-API lo que es una API en formato JSON (JavaScript Object Notation, por sus siglas en inglés). Primero que todo cuenta con noticias sobre la red Bitcoin en su inicio y precios actualizados de Bitcoin con respecto a varias monedas y de varios mercados de intercambio. Posee una pestaña para visualizar precios de Bitcoin de una manera mucho más detallada y de diversos intercambios con posibilidad de obtener más información de cada uno en específico, por otro lado cuenta con una pestaña de gráficos donde contamos con varias entradas con opciones las cuales podemos escoger según lo que necesitemos, entre las más resaltantes esta que podemos de alguna manera jugar con combinaciones de períodos a visualizar y variar la frecuencia de los datos que vamos visualizando, en cuanto a indicadores sobre el gráfico, tenemos una posibilidad finita de seis como máximo. En cuanto a la descarga de la visualización, no es posible descargar la imagen resultante pero se cuenta con un botón con el cual podemos visualizar los datos esenciales plasmados en el gráfico, claro que de quererlos, tendríamos que hacer una copia manual.

BitcoinAverage [96], BAV: fuente de datos históricos de precios de BTC. En 2013, lanzaron lo que ellos llaman el primer índice de precios de Bitcoin del mundo, que se ha convertido en uno de los índices de precios de Bitcoin más respetados y citados.

Si nos dirigimos a la sección de índices globales, a simple vista podríamos concluir que es una

página relativamente sencilla y básica, con un gráfico con opciones estándar, lo más relevante en este caso son los datos que muestran, ellos generan su índice de BTC recopilando datos a nivel global y en todo tipo de monedas. En primer lugar, cada precio promedio del mercado de divisas se convierte a USD utilizando los tipos de cambio globales. Luego usan el volumen diario en cada mercado para proporcionar un factor de ponderación a su precio. Ya con el índice de BTC representado en USD, es posible hacer una conversión cruzada del precio a más de 160 monedas mundiales usando tipos de cambio globales.

El gráfico que se puede visualizar cuenta con precio y volumen, donde la única opción es cambiar el período que estamos visualizando, no hay indicadores o cálculos disponibles.

CoinMarketCap [97], CMC: Ha crecido rápidamente hasta convertirse en lo que ellos llaman la fuente más confiable de usuarios, instituciones y medios para comparar miles de criptoactivos y es citado comúnmente por [CNBC \[98\]](#), [Bloomberg \[99\]](#) y otros importantes medios de comunicación.

En abril de 2020, CoinMarketCap fue adquirida por Binance. Binance es la compañía global de blockchain detrás del intercambio de activos digitales más grande del mundo por volumen de negociación y usuarios, que comparte la visión común de hacer que las criptomonedas sean accesibles y sistemáticamente importantes para las personas de todo el mundo [100].

A través de la adquisición, CoinMarketCap planea aprovechar los amplios recursos de su empresa matriz para adquirir datos más detallados en proyectos, reducir la asimetría de información en la industria de la criptografía y brindar datos de mayor calidad a sus cientos de millones de usuarios.

Cuenta al inicio de su página con noticias relevantes del mundo cripto, como criptomonedas en auge, las mayores perdedoras y ganadoras de las últimas 24h, para luego pasar a una lista de todas las criptomonedas disponibles con información básica, si seleccionamos una en específico, accedemos a una página dedicada la cual muestra un gráfico de cotizaciones en la moneda fiat o criptomoneda que deseemos, con pocas opciones de entrada pero con la opción a abrir un gráfico de TradingView [101] con muchas más posibilidades, luego también tenemos acceso a una amplia selección de información como: titulares, datos históricos, información del proyecto, monederos, redes sociales, análisis de calificaciones, estimaciones de precios y hasta una chat en vivo.

1.5.1. Datos de nuestro interés.

Si bien estos sitios web contienen una infinidad de información, centrémonos en un grupo de datos en específico, que serán con los que trabajaremos, cabe resaltar que usaremos sus nombres en inglés de aquí en adelante ya que es como se suelen usar a nivel mundial:

Apertura (Open, en inglés): si bien, a diferencia de los mercados tradicionales el mundo de las criptomonedas no posee una hora de apertura como tal, si no que es visto como un continuo, el dato de apertura sera entonces el primer dato referente a la escala de visualización que tengas, en este caso para datos diarios, se tomaría el primer registro del día a las 0 horas con 0 minutos y 0 segundos

Cierre (Close, en inglés): de manera análoga al datos de apertura, el dato de cierre para intervalos diarios seria el ultimo durante el intervalo en específico.

Máximo (High, en inglés): representa el valor máximo de la cotización por el cual se realizó una transacción en un período de tiempo.

Mínimo (Low, en inglés): representa el valor mínimo de la cotización por el cual se realizo una transacción en un período de tiempo.

Global Bitcoin price index (GBX): usualmente este es calculado por los autores con el propósito de representar un precio global de una criptomoneda en específico, usualmente se calcula en USD y luego mediante conversiones se representa en otras monedas fiat [102].

OHLC: se le suele llamar así al grupo de datos que agrupa el Open, High, Low y Close.

Volumen: es el número de unidades negociadas en un mercado durante un tiempo determinado. Es una medida del número de unidades individuales de un activo que cambió de manos durante ese período. Cada transacción involucra un comprador y un vendedor. Cuando llegan a un acuerdo a un precio específico, la transacción es registrada por el intercambio facilitador, de esta manera podemos mostrar tanto el volumen como la cantidad de unidades negociadas como la cantidad total de dinero pagado por ellas [103].

Capitalización del mercado (market capitalization, en inglés): a menudo denominada “marketcap”, la capitalización de mercado es el valor de mercado actual de la red de una criptomoneda. Se calcula multiplicando la oferta circulante de un crypto-activo por el precio de una unidad individual [104].

1.5.2. Opciones de interactividad.

Si bien hablamos a grandes rasgos de las fuentes de datos, pasemos ahora a centrarnos en lo que llamaremos opciones de interactividad o simplemente opciones de los gráficos(reportes). Las cuales hacen referencia primero a esas opciones y posibilidades que hemos seleccionado de las páginas web consultadas para generar reportes o visualizaciones a nuestro gusto, a las cuales luego les añadiremos algunas extra creadas por nosotros, las cuales expondremos en el segundo capítulo.

Selección de dato a graficar: en cuanto a cotizaciones, es común que solo haya dos opciones, gráfico de velas japonesas [105] o gráfico de línea, como consecuencia de esto lo más común es que no haya elección más que la posibilidad de graficar la cotización de acuerdo al tipo de intervalo que escojamos o el dato de cotización promedio.

Frecuencia de datos variable: Kraken y CoinmarketCap cuenta con posibilidad de escoger la frecuencia de los datos pero con consecuencia directas sobre el rango de visualización que estemos visualizando, Criptocompare por su parte no cuenta con posibilidad a escoger la frecuencia, esta se adapta al tamaño del período que estemos visualizando, y por último BitcoinAverage, cuenta con un tipo de frecuencia invariante, ya que muestran una sola serie temporal de cotizaciones con tres frecuencias distintas que cambian de un dato por minuto, hora y por día, de manera predeterminada.

Período a visualizar: la selección del período a visualizar va de la mano de la frecuencia de los datos que estemos visualizando, si bien podemos escoger el período manual o escoger de la opciones predeterminadas, la frecuencia podría limitarnos, ya que un patrón a lo largo de estas páginas de cotizaciones es que si bien quieres un período muy amplio, tendrás que sacrificar en cuanto a la frecuencia de los datos y viceversa. De las cinco que inspeccionamos, BitcoinAverage es la única que se diferencia del resto ya que si bien podemos escoger cualquier tipo de período, la frecuencia de los datos está preestablecida en tres formatos, diario, horario y cada minuto, con respecto al momento de la consulta, donde la mayor apreciación está en las últimas 24 horas, y así sucesivamente.

Selección de la fuente: si bien páginas como BitcoinAverage, CoinMarketCap y Criptocompare buscan ofrecer un indicador promedio juntando muchos datos a nivel global, Criptocompare cuenta con la posibilidad de graficar otras cotizaciones además de su índice, pero una a la vez y por su lado CoinMarketCap en su visualización avanzada cuenta con la posibilidad de agregar trazos de otras fuentes sobre el gráfico original. Por último Kraken por su parte

solo cuenta con sus datos ya que éste es el único de los cinco que es un intercambio de criptomonedas como tal.

Información contextual: si bien todas cuentan con esta cualidad, no todas funcionan de igual manera, Kraken y CoinMarketCap que cuentan con una modalidad de visualización avanzada, la muestran de una manera más profesional si se puede decir, ya que la información se visualiza en los bordes del gráfico mientras navegamos con el puntero, más allá de esto, las cinco páginas muestran un recuadro que sigue al puntero mostrándonos información específica de donde vamos señalando sobre el gráfico.

Indicadores (social,estadístico,blockchain): BitcoinAverage no posee opciones de este tipo, Criptocompare si bien posee de los tres tipos de indicadores, sin registrarnos en la página, solo tenemos acceso a los indicadores de tipo estadístico sobre los datos del gráfico; Kraken por su parte al igual que CoinMarketCap en su modalidad de visualización avanzada posee una amplia gama de indicadores de tipo estadístico que podemos aplicar sobre los datos que estamos visualizando, tanto sobre los mismos datos como los de tipo oscilador, los cuales explicaremos en el segundo capítulo.

Gráfico complementario: empezando por BitcoinAverage que no posee ningún tipo de elección, posee un gráfico complementario con el volumen el cual no es opcional, luego tenemos a Criptocompare el cual posee once opciones para el gráfico complementario pero sin registrarse en la página tenemos acceso a siete solamente, los cuales no son manipulables en cuanto a las opciones del gráfico y luego tenemos a Kraken y CoinmarketCap los cuales cuentan con una gran cantidad de opciones para el gráfico complementario, con la posibilidad de opciones sobre los mismos, en estos casos podemos agregar tantos como nos sea posibles y manejables, ya que agregar uno extra, compromete el tamaño de los anteriormente agregados.

Posibilidad de descarga: en este punto podemos hacer la distinción de dos tipos de descargas, entre las cuales están descargas de archivos de datos en formatos de texto como: csv (del inglés Comma-Separated Values), JSON o xlsx (extensión de archivo en Microsoft Excel) con el que cuenta solamente Criptocompare, aquí debemos resaltar que si se añadió un gráfico complementario, el archivo de datos contará con estos datos extra; y descargas de tipo imagen, de los cuales Kraken y CoinMarketcap solo poseen posibilidad de descarga de imágenes en formato PNG (del inglés Portable Network Graphics), BitcoinAverage también posee PNG pero tiene agregado también formatos como jpeg (del inglés Joint Photographic Experts Group), svg (del inglés Scalable Vector Graphics) y pdf (del inglés Portable Docu-

ment Format), por otro lado Criptocompare también posee estos formatos disponibles junto a los antes mencionados.

En la siguiente tabla de la figura 1.9 se puede ver de manera resumida la distribución de las cualidades en nuestras cinco fuentes pre-seleccionadas de manera que ✓ se refiere a que cuenta con la cualidad y x a que no.

	Opciones		Servidores de datos financieros				
			BCH	KRK	CCO	BAV	CMC
1	Selección de dato a graficar		✓	x	x	x	x
2	Frecuencia de datos variable		✓	✓	x	x	✓
3	Período a visualizar		✓	✓	✓	✓	x
4	Selección de la fuente		✓	x	✓	x	✓
5	Información contextual		x	✓	✓	✓	✓
6	Indicadores	Social	x	x	✓	x	x
		Estadístico	✓	✓	✓	x	✓
		Blockchain	x	x	✓	x	x
7	Gráfico complementario		x	✓	x	x	✓
8	Posibilidad de descarga	Tipo imagen	x	✓	✓	✓	✓
		Vectorizado	x	x	✓	x	x
		Archivo de datos	x	x	✓	x	x

Figura 1.9: Opciones disponibles en las Fuentes de datos pre-seleccionadas.

Veamos una por una las entradas de el cuadro 1.9: en cuanto al punto 1 solo se marco con ✓ las páginas que ofrecían una selección individual de cada tipo de dato, ya que la mayoría ofrece gráfico de líneas o velas japonesas, que si bien en las velas japonesas tenemos datos OHLC de esos intervalos, no podemos escoger un gráfico de solo Close por ejemplo.

En cuanto al punto 2 las dos páginas que no cumplen este aspecto tienen modalidades distintas por un lado BAV posee una frecuencia estática en los datos a mostrar y CCO solo es reactivo a la selección de períodos que estemos viendo.

Con respecto al punto 3 la única página que no lo cumple es CMC que si bien en su gráfico sencillo lo cumple, en su gráfico avanzado de TradingView no es posible seleccionar un período en específico, este es reactivo a la frecuencia de datos que estemos visualizando.

El punto 4 como bien lo dice hace referencia a si podemos o no escoger de entre varias fuentes de datos, en este caso diversos mercados de intercambio en particular.

Si bien el punto 5 lo cumplen todas menos BCH, las demás lo cumplen de dos maneras, con una ventana emergente que nos muestra la información detallada de donde tenemos el cursor, o bien con información a los lados del gráfico que siguen al curso por los bordes del mismo.

En cuanto al punto 6 CCO es el único que posee indicadores tanto sociales como basados en la red Blockchain, de resto en cuanto a los indicadores matemáticos, hacemos referencia a esos indicadores que son el fruto de cálculos estadísticos de los datos principales del gráfico.

Con respecto al punto 7, se marcaron con ✓ las páginas que ofrecen opciones y cambios sobre este gráfico ya que las demás solo muestran el volumen, sin opciones sobre él.

Por último con respecto al punto 8, quisimos separar la cualidad de imagen vectorizada ya que esta podría servirnos para otra finalidad de manera más adecuada que una simple imagen.

Y es así como finaliza este primer capítulo del proyecto, manejaremos datos provenientes de páginas web de cotizaciones de intercambio de monedas fiat y cuatro metales contra algunas criptomonedas, si bien haremos uso de estos datos, era necesario observar y analizar lo que ofrecen comúnmente éstas páginas de consulta de datos para tomar lo mejor de ellas y posiblemente agregar aspectos que nos resulten interesantes. En el segundo capítulo expondremos más a fondo la parte que tiene que ver con la extracción en específico de los datos vía API, su procesamiento y análisis mediante algunos métodos estadísticos y probabilísticos.

www.bdigital.ula.ve

“Datos, datos, datos. No puedo hacer ladrillos sin arcilla”

Sherlock Holmes

Capítulo 2

Minería de datos.

El primer capítulo nos deja las bases teóricas sobre los datos financieros, sitios web que reflejan datos en tiempo real e histórico y las utilidades que se suelen ofrecer a la hora visualizar y generar reportes, ahora hablaremos en rasgos generales de las herramientas y técnicas que aplicaremos en nuestro estudio, este estará dividido en tres partes:

Extracción, donde hablaremos de los métodos de extracción y la implementación de herramientas para automatizar descargas a partir de un URL. Una vez que logramos exitosamente la descarga de los datos que deseamos, pasamos al procesamiento de los mismos.

Procesamiento, donde pasaremos de los datos extraídos con formato predeterminado de la fuente a un formato hecho a medida de nuestras necesidades para posteriormente trabajar con estos, aplicando métodos estadísticos y probabilísticos como también usar estos datos para generar una simulación basada en datos históricos.

Reportes, esta sección muestra los posibles reportes que generaran a partir de los datos descargados y procesados, los cuales brindaran información útil y personalizada de acuerdo a la estrategia que este implementando el usuario.

2.1. Extracción

Luego de la localización de los datos que deseamos, tenemos que encontrar la manera de conseguir extraerlos de nuestra fuente, este es el tema que trataremos en esta sección, claro que para llegar a este punto se necesitan pasos previos que mencionamos en el capítulo anterior, es por eso que de los cinco sitios web que expusimos, nos quedaremos con un grupo reducido, luego de tener

un grupo seleccionado de fuentes no solo podemos quedarnos con la que parezca mejor o la que tengan mas datos, debemos hacer un estudio observando la posibilidad que tenemos de extraer dichos datos, ya que como se puede observó en la figura 1.9 una sola página posee descarga directa de datos, es por esto que debemos buscar otros métodos de descargarlos. El escogido por nosotros es mediante la API de cada sitio web. Veamos a continuación ¿qué son? y ¿cómo funcionan?.

2.1.1. Interfaz de Programación de Aplicaciones.

Una API posee la capacidad de prestar comunicación entre componentes de software. Esta engloba un conjunto de subrutinas que ofrecen cierto acceso a servicios o datos de otro software, como una capa de abstracción. Cada página web que se dedica a la divulgación de datos sobre mercados de criptomonedas, posiblemente emplee una API, pero la inclusión de una API no se limita a sitios web de datos financieros; empresas y páginas y sitios de diversos tipos cuentan con estas para acceder a datos; como la red social Twitter también cuentan con una con la cual se tiene acceso de manera practica a datos y estadísticas, con el fin de analizar las interacciones dentro de esta red [106], con la implementación de estas API se proporciona acceso a datos usualmente en un archivo tipo JSON, posibilidad de implementar la API en otro software y así acceder de manera práctica a los datos, para hacer uso de éstas API es necesario contar con su documentación o descifrarlas al menos.

Explicuemos ahora detenidamente como funcionan estas API de sitios web de cotizaciones. Usualmente la solicitud de los datos mediante la API y la implementación de la misma es como si fuera un catálogo y cada opción de este catálogo tiene sus pasos a seguir para obtener lo que queremos. Viéndolo así, es de suma importancia conocer el catálogo y su funcionamiento, es por esto que cada API cuenta con su documentación.

La documentación de una API engloba una colección de referencias, tutoriales y ejemplos que ayudan a los desarrolladores a usar/implementar una API, con el único propósito de describir de manera detallada que servicios ofrece y como hacer uso de ellos. La documentación una API es el recurso principal para explicar lo que es posible con dicha API y ¿cómo comenzar?, esta debe cubrir todo lo que un cliente o usuario necesitaría saber para hacer uso pleno de la API.

Por lo general, la documentación de dichas API se aloja en una sección especial del sitio web o en su propio portal centrado en la API. El contenido suele ser lo más accesible posible para su audiencia, en los casos donde la API es interna de una empresa o un proyecto, es probable que su documentación también sea interna, sin embargo, debe ser fácilmente detectable [107].

En la documentación de la API como bien ya mencionamos se presentan todas las funciona-

lidades de las mismas como si fuese un catálogo. En este caso a cada ítem de este catálogo se le conoce como “punto final(es)” de ahora en adelante: endpoint(s), en inglés. Los endpoints pueden tener un significado tan complicado como se quiera, pero en la mayoría de los casos de las API con acceso a datos financieros, estos puntos finales representan la ruta a los datos que deseamos obtener y van acompañados de algunos argumentos extras, una forma práctica de verlo es que las API poseen una lista de endpoints, los cuales representan los diversos tipos de datos a los que podemos acceder y obviamente si se necesitan argumentos extra, estos deben ser suministrados tal cual dicta la documentación para no obtener un error como respuesta.

Existe un subgrupo dentro de las API, las cuales son de nuestro interés, estas son las API-web. Dichas API-web brindan el acceso a bases de datos a páginas web del tipo interactivas, de ésta manera, el uso de la API-web se realiza de forma indirecta, los llamados a endpoints están programados para hacerse cada vez que la página web lo necesite, para actualizar un gráfico, mostrar una tabla, una foto, entre otros recursos. Claro que si bien estos llamados son hechos de manera indirecta por el usuario, estos si son realizados de manera directa por el navegador al hacer la petición de los datos, lo que entonces nos podría dejar con la duda: ¿Podemos rastrear y capturar estas peticiones?.

Como explicaremos mas adelante, todas estas peticiones se almacenan en la cache de la computadora lo que denominaremos de ahora en adelante: cacheable, y ya estando los datos allí en la cache, podemos capturar dichos datos al realizar una copia ya que la información es descargada temporalmente en el computador y de igual manera cada una de estas peticiones quedan registradas en las interacciones de red hechas por el navegador. Es así como haciendo un seguimiento a estas interacciones del navegador podemos obtener acceso a estas peticiones que como se explicó anteriormente se hacen mediante la solicitud a endpoints, si nos dedicamos al estudio detallado de una página web, observando todas estas peticiones en particular y como es su estructura de endpoints, terminaríamos con una documentación de la API-web rudimentaria pero funcional para nuestras necesidades [108].

Esta documentación rudimentaria la compondríamos al ir observando una por una las peticiones hechas por el ordenador y así al encontrar una por una las de nuestro interés, observaríamos para cada una su Localizador de Recursos Uniforme (Uniform Resource Locator, URL en inglés), un URL es una dirección dada a un recurso único en la Web. En teoría, cada URL valida apunta a un único recurso. Dichos recursos pueden ser páginas con lenguaje de marcado de hipertexto (HyperText Markup Language, HTML por sus siglas en inglés), documentos de hojas de estilo en cascada (Cascading Style Sheets, CSS por sus siglas en inglés), imágenes, etc. Un URL está

compuesto de diferentes partes, algunas obligatorias y otras opcionales.

Veamos las partes más importantes usando el siguiente URL [109], en el cual la parte en azul es particular para cada URL:

<http://www.example.com:80/path/to/myfile.html?key1=v1&key2=v2#SomewhereInTheDocument>

http:// El protocolo, la primera parte de la URL indica qué protocolo debe usar el navegador. Un protocolo es un método establecido para intercambiar o transferir datos alrededor de una red informática. Por lo general, para sitios web es el protocolo de transferencia de hipertexto o HTTP por sus siglas en inglés (HyperText Transfer Protocol) o su versión segura HTTPS (HTTP secure, en inglés).

www.example.com El nombre de dominio, indica qué servidor web se solicita. Alternativamente, es posible usar directamente un dirección IP, pero debido a que es menos conveniente, no se usa con frecuencia en la Web.

:80 El puerto, indica la “puerta” técnica utilizada para acceder a los recursos en el servidor web. Por lo general, se omite si el servidor web utiliza los puertos estándar del protocolo HTTP (80 para HTTP y 443 para HTTPS) para otorgar acceso a sus recursos. De lo contrario es obligatorio.

/path/to/myfile.html La ruta al recurso en el servidor web, en los primeros días de la Web, una ruta como esta representaba la ubicación de un archivo físico en el servidor web. Hoy en día, en muchos casos es una abstracción manejada por servidores web.

?key1=v1&key2=v2 Parámetros adicionales proporcionados al servidor web, esos parámetros son una lista de pares “clave=valor” separados con el símbolo &. El servidor web puede usar esos parámetros para hacer cosas adicionales antes de devolver el recurso. Cada servidor web tiene sus propias reglas con respecto a los parámetros, y la única forma confiable de saber si un servidor web específico está manejando parámetros es preguntando al propietario del servidor web.

#SomewhereInTheDocument Es un “ancla” para otra parte del recurso en sí. Un ancla representa una especie de marcador dentro del recurso, dando al navegador las instrucciones para mostrar el contenido ubicado en ese lugar marcado. En un documento HTML, por ejemplo, el navegador se desplazará hasta el punto donde se define el ancla; en un vídeo o documento de audio, el navegador intentará ir a

la hora que representa el ancla. Vale la pena señalar que la parte después del “#”, también conocido como el identificador de fragmento, nunca se envía al servidor con la solicitud.

Se puede pensar en una URL como una dirección de correo postal normal: el protocolo representa el servicio postal que desea utilizar, el nombre de dominio es la ciudad o el pueblo y el puerto es como el código postal de la zona; la ruta representa el edificio donde se debe entregar su correo; los parámetros representan información adicional como la sección o el piso donde se encuentra el apartamento; y finalmente, el ancla representa el número de apartamento específico a la que va dirigido el correo.

Como se mencionó anteriormente, el protocolo usado generalmente es HTTP o HTTPS, este protocolo de red permite la transferencia de documentos de hipermedia en la red, al nombrar hipermedia, hace referencia a el conjunto de métodos o procedimientos para escribir, diseñar o componer contenidos que integren soportes tales como: texto, imagen, vídeo, audio, mapas y otros soportes de información emergentes. Generalmente entre un navegador y un servidor, para que los humanos puedan leerlos.

HTTP se considera textual porque toda la comunicación se realiza en texto plano. También es un protocolo sin estado, ya que una comunicación no es consciente de las comunicaciones que se hayan podido producir anteriormente. Estas dos propiedades lo hace ideal para la consulta de documentos en la red (sitios y páginas web). Cuenta con un conjunto de métodos de petición para indicar la acción que se desea realizar para un recurso determinado. Aunque estos también pueden ser sustantivos, estos métodos de solicitud a veces son llamados verbos HTTP (HTTP-verbs, en inglés), cada uno de ellos implementa una semántica diferente como podemos ver en la figura 2.1 [110]. Un grupo de ellos comparte algunas características peculiares, un método de solicitud puede ser seguro, idempotente o cacheable. Veamos cada una de estas características por separado:

- Un método HTTP es seguro si no altera el estado del servidor, en otras palabras, un método es seguro si conduce a una operación de solo lectura. Varios métodos HTTP comunes son seguros: GET, HEAD y OPTIONS. Todos los métodos seguros también son idempotentes, pero no todos los métodos idempotentes son seguros. Por ejemplo, PUT y DELETE son idempotentes pero inseguros [111].
- Un método HTTP es idempotente si se puede realizar una solicitud idéntica una o varias veces seguidas con el mismo efecto y dejando el servidor en el mismo estado. En otras palabras, un método idempotente no debería tener efectos secundarios (excepto para llevar estadísticas).

Método	Acción
GET	Solicita una representación del recurso especificado. Las solicitudes que utilizan GET solo deben recuperar datos.
HEAD	Solicita una respuesta idéntica a una solicitud GET , pero sin el cuerpo de la respuesta.
POST	Envía una entidad al recurso especificado, lo que a menudo provoca un cambio de estado o efectos secundarios en el servidor.
PUT	Reemplaza todas las representaciones actuales del recurso de destino con la carga útil de la solicitud.
DELETE	Elimina el recurso especificado.
CONNECT	Establece un túnel al servidor identificado por el recurso de destino.
OPTIONS	Describe las opciones de comunicación para el recurso de destino.
TRACE	Realiza una prueba de bucle invertido de mensajes a lo largo de la ruta al recurso de destino.
PATCH	Aplica modificaciones parciales a un recurso.

Figura 2.1: Métodos del Protocolo de Transferencia de Hipertexto.

- Una respuesta cacheable es una respuesta HTTP que se puede almacenar en caché, que se almacena para recuperarla y usarla más tarde, guardando una nueva solicitud en el servidor [112]. El rendimiento de los sitios web y las aplicaciones se puede mejorar significativamente reutilizando los recursos obtenidos anteriormente.

El método que nos interesa específicamente es el método **GET**, ya que este es el método usado por las páginas web al solicitar datos mediante las API-web y son estas solicitudes las que esperamos reconstruir para realizar nosotros mismos las solicitudes. Este en particular forma parte del grupo que cuenta con las tres cualidades previamente descritas.

Ser idempotente y cacheable es lo mas relevante para nosotros, ya que si tenemos el URL de la solicitud que queremos, el siguiente paso es direccionar esta respuesta a un fichero distinto que no sea la memoria cache al realizar su descarga y además al ser idempotente podemos realizarla cuantas veces queramos sin ninguna restricción. Aquí es donde entra la herramienta que discutiremos a continuación, una herramienta para automatizar las descargas de los datos a partir de los URL.

2.1.2. Automatización de descargas.

Ahora bien si ya conseguimos los URL usados por la API-web y tenemos una idea de como funcionan, procedemos a conseguir como descargar estos datos de una manera automatizada, si bien con el URL podemos visualizar/recuperar los datos mediante el navegador, lo que buscamos

es poder descargar los datos y almacenarlos en forma de archivo y junto a esto que la llamada sea dinámica, dinámica ya que al saber como funciona el endpoint que usaremos, podemos cambiar los parámetros para cada descarga. En nuestro caso la construcción del URL al endpoint que queremos tiene cuatro variables (entradas) que son:

Fecha inicial.	Criptomoneda.
Fecha final.	Moneda fiat.

La herramienta ideal para la automatización de las descargas mediante URL es `curl`, perteneciente al proyecto **cURL** [113], el nombre del proyecto se originó en sus inicios cuando la herramienta consistía en cargar y descargar datos especificados con una URL. Era un programa del lado del cliente, un cliente de URL y mostraba los datos (por omisión). Así que “c” para cliente y **URL**: **cURL**.

Esta es una herramienta que se usa mediante la línea de comandos o con una secuencia de comandos de ahora en adelante script(s), para transferir datos. `curl` también se usa en automóviles, televisores, enrutadores, impresoras, equipos de audio, teléfonos móviles, tabletas, decodificadores, reproductores multimedia y es el motor de transferencia de Internet para miles de aplicaciones de software en más de diez mil millones de instalaciones. `curl` es utilizado diariamente de manera indirecta por prácticamente todos los usuarios de Internet en el mundo.

Hacer uso de `curl` mediante la línea de comandos no es algo poco común, su creador nunca consideró otra modalidad, de esta manera se generan los datos en la salida estándar (**standard out**, `stdout`, en inglés), en la terminal de forma predeterminada.

El mantra de “todo es una tubería” de la filosofía estándar de Unix (familia de sistemas operativos de computadora multitarea y multiusuario) era algo en lo que el creador de `curl` apoyaba, `curl` envía datos a `stdout` para que sea fácil de encadenar junto con otras herramientas para hacer lo que quieras. Esta es también la razón por la que prácticamente todas las opciones de `curl` que permiten leer desde un archivo o escribir en un archivo, también tienen la capacidad de seleccionar hacerlo en `stdout` o desde entrada estándar (**standard in**, `stdin`, en inglés). De aquí que podamos redireccionar la respuesta de la petición a un archivo y hacernos con los datos.

La herramienta de línea de comandos está diseñada para funcionar perfectamente desde scripts u otros medios automáticos. No presenta ninguna otra interfaz gráfica de usuario (GUI, del inglés Graphical User Interface) o interfaz de usuario (UI, del inglés User Interface) que no sea el texto de entrada y la salida de texto.

Específicamente los URL que ocuparemos usan el protocolo HTTP, la creciente cantidad de aplicaciones que se trasladan a la web ha hecho que la creación de secuencias de comandos HTTP

(HTTP scripting, en inglés) sea más solicitado y buscado. Poder extraer automáticamente información de la web, publicar o cargar datos en servidores web son tareas importantes hoy en día.

`curl` es una herramienta capaz de realizar todo tipo de manipulaciones y transferencias de URL, pero no está escrito para hacer todo por el usuario. Hace las solicitudes, obtiene los datos, envía datos y recupera la información. Es necesario que unamos todo utilizando algún tipo de lenguaje de script o invocaciones manuales repetidas.

Con lo que uniremos todo hasta aquí será mediante el software libre *bash* [114], este es un intérprete de órdenes (shell) del Proyecto GNU [115]. Este es un shell que nos permite escribir ejecutables con extensión `.sh` que incorpora características útiles del shell Korn [116] y el shell C [117]. Está diseñado para cumplir con el estándar *IEEE POSIX P1003.2/ISO 9945.2 Shell and Tools* [118].

De manera que haremos un script en *bash* para hacer uso de `curl`, el cual hará los llamados a nuestras direcciones mediante el protocolo HTTP, dichas URL serán construidas de manera automática por dicho script basado en las entradas que el usuario suministre, hasta este punto que es la extracción de datos, las entradas están limitadas a una fecha inicial, una fecha final, una criptomoneda y una moneda fiat, claro que en nuestro caso esto podría estar abierto a estar dentro de un bucle ya que podríamos contar con peticiones a varias páginas cada una con URL distintos, o en bucles con peticiones de una misma criptomoneda con cotizaciones en varias monedas fiat.

Un incentivo para realizar la extracción de ésta manera es que existe una discrepancia de la accesibilidad en cuanto a los datos entre las API-web y la API de una página en específico, si bien la API posee documentación detallada y acceso a una amplia gama de datos, el acceso a estos datos siempre está de alguna manera restringido por uno u otro factor, a diferencia que mediante estos llamados de la API-web tenemos acceso a todos los datos mostrados en la página web como tal, y todos los que esta solicite en el transcurso de su uso si bien los logramos rastrear, claro y recordemos que no tenemos límite de peticiones, ni en cantidad por segundo ni en tamaño. De esta manera si es posible tener acceso durante una sesión normal dentro de la página web, lo podremos obtener mediante la API-web.

2.2. Procesamiento.

En cuanto a la parte de procesamiento primero veremos como podemos generar un archivo a medida de ahora en adelante: `archivo.dat`, mediante los editores de texto, línea y flujo, luego partiendo de que ya contamos con nuestro `archivo.dat` de los datos solicitados, pasaríamos a cen-

trarnos en que se puede hacer con estos datos, los lenguajes que usaremos en el procesamiento y las herramientas para integrar esta sección.

Si bien ya tenemos descargados los datos, no significa que estén ya aptos para nuestro análisis, es por eso que en el mismo script con el que realizaremos la descarga, haremos uso de editores de texto, línea y flujo para obtener un archivo de datos el cual tenga todas las características definidas y uniformes para cada descarga y cada fuente. A continuación veremos las tres herramientas usadas para el procesamiento de los datos descargados, estas son usadas mediante la línea de comandos e incluidos en el script de descarga.

sed: Editor de flujo (**s**tream **e**ditor), es un editor de texto de línea de comandos no interactivo. **sed** se usa comúnmente para filtrar texto, es decir, toma la entrada de texto, realiza alguna operación (o conjunto de operaciones) en él y genera el texto modificado. **sed** se usa normalmente para extraer parte de un archivo mediante la coincidencia de patrones o la sustitución de varias apariciones de una cadena dentro de un archivo [119]. **sed** fue una de las primeras herramientas para admitir expresiones regulares y sigue en uso para el procesamiento de texto, sobre todo con el comando de sustitución. Las herramientas alternativas populares para la manipulación de cadenas de texto sin formato y la “edición de secuencias” incluyen y Perl [120] [121].

El programa **sed** por defecto copia la `stdin` a la `stdout`, quizás ejecutando uno o más comandos de edición en cada línea antes de escribirla en la salida. Este comportamiento puede ser modificado por argumentos en la línea de comando. Es por esto que el tamaño efectivo de un archivo que se puede editar está limitado únicamente por el requisito de que los archivos de entrada y salida quepan simultáneamente en el almacenamiento secundario disponible. Esto se debe a que solo unas pocas líneas del archivo de entrada residen en la memoria a la vez y no se reutilizan archivos temporales.

Los scripts de edición complicados pueden crearse por separado y entregarse al programa **sed** como un archivo de comando. Para ediciones complejas, esto ahorra una cantidad considerable de errores de tipeo y asistentes. El programa **sed** que se ejecuta desde un archivo de comando es más eficiente que un editor interactivo, incluso si ese editor puede ser controlado por un script prescrito [122].

awk: Es un lenguaje de dominio específico, diseñado para el procesamiento de archivos de texto, con el se logran hacer muchas tareas comunes de manipulación de

información y recuperación de texto fácil de enunciar y realizar. Este se ejecuta en un conjunto de archivos de entrada. La operación básica de **awk** es escanear un conjunto de líneas de entrada, en orden, una a la vez. En cada línea, **awk** busca el patrón descrito en el programa **awk**, luego, si ese patrón se encuentra en la línea de entrada, se realiza la acción correspondiente. De esta forma, cada declaración del programa **awk** se ejecuta para una línea de entrada determinada. Cuando se prueban todos los patrones, se continua con la siguiente línea como entrada; y el programa **awk** se vuelve a ejecutar desde el principio.

En el comando **awk**, se omite el patrón o la acción, pero no ambos. Si no hay acción para un patrón, simplemente se imprime la línea coincidente. Si no hay un patrón para una acción, la acción se realiza para cada línea de entrada. El programa nulo **awk** no hace nada. Dado que tanto los patrones como las acciones son opcionales, las acciones se encierran entre llaves para distinguirlas de los patrones [123].

Varios tipos de tareas ocurren repetidamente cuando se trabaja con archivos de texto. Es posible que desee extraer ciertas líneas y descartar el resto. O puede que necesite hacer cambios dondequiera que aparezcan ciertos patrones, pero deje el resto del archivo solo. Estos trabajos suelen ser fáciles con **awk**. La utilidad **awk** interpreta un lenguaje de programación de propósito especial que facilita el manejo de tareas simples de reformato de datos [124].

grep: Es una utilidad de línea de comandos para buscar conjuntos de datos de texto sin formato para líneas que coincidan con una expresión regular [125]. Dado uno o más patrones, **grep** busca en los archivos de entrada coincidencias con los patrones. Cuando encuentra una coincidencia en una línea, copia la línea en la stdout (de forma predeterminada) o produce cualquier otro tipo de salida que haya solicitado con opciones.

Aunque **grep** espera hacer la coincidencia en el texto, no tiene límites en la longitud de la línea de entrada aparte de la memoria disponible, y puede hacer coincidir caracteres arbitrarios dentro de una línea. Si el último byte de un archivo de entrada no es una nueva línea, **grep** silenciosamente proporciona uno. Dado que la nueva línea también es un separador de la lista de patrones, no hay forma de hacer coincidir los caracteres de nueva línea en un texto [126].

Luego del uso de éstas tres herramientas, tendríamos nuestro tan comentado archivo.dat con los datos específicos de una consulta y en nuestro formato, ahora si estaríamos listos para empezar

a procesar y estudiar nuestros datos. Nuestro procesamiento abarca cuatro áreas que se verán marcadas a lo largo del proyecto: estadística, probabilidad, inferencia y simulación. Antes de ver cada una por separado, veamos a detalle una serie de conceptos importante para entender dichas áreas y su funcionamiento.

Población: Conjunto de “objetos” agrupados según alguna característica o propiedad.

Muestra: Subconjunto de “objetos” pertenecientes a una población.

Muestra aleatoria: Subconjunto de “objetos” seleccionados “aleatoriamente” de una población.

Muestreo: Procedimiento y/o técnica que se utiliza para obtener muestras de una población.

Este último lo veremos con mas detalle en la sección siguiente, acompañado de las técnicas de muestreo mas comunes y de las que haremos uso.

2.2.1. Muestreo.

Nuestro conocimiento, se basa en gran medida en muestras. Esto es igualmente cierto en la vida cotidiana y en la investigación científica. La opinión de una persona sobre una institución que realiza miles de transacciones todos los días a menudo está determinada por uno o dos encuentros que ha tenido con la institución en el transcurso de varios años. Tanto en la ciencia como en los asuntos humanos carecemos de los recursos para estudiar más que un fragmento de los fenómenos que podrían hacer avanzar nuestro conocimiento. [127]

El propósito de la teoría del muestreo es hacer que el muestreo sea más eficiente. Intenta desarrollar métodos de selección de muestras y de estimación que proporcionen, al menor costo posible en cuanto a tiempo, gasto de recursos entre otros, estimaciones lo suficientemente precisas para nuestro propósito.

Para aplicar este principio, debemos ser capaces de predecir, para cualquier procedimiento de muestreo que se esté considerando, la precisión y el costo esperados. En lo que se refiere a la precisión, no podemos predecir exactamente qué tan grande será el error en una estimación en cualquier situación específica, porque esto requeriría un conocimiento del valor real para la población. En cambio, la precisión de un procedimiento de muestreo se juzga examinando la distribución de frecuencia generada para la estimación si el procedimiento se aplica una y otra vez a la misma población. Esta es, por supuesto, la técnica estándar por la cual se juzga la precisión en la teoría estadística.

A continuación algunos métodos de muestreo:

Muestreo Aleatorio Simple: El muestreo aleatorio simple es un método para seleccionar n unidades de N , de modo que cada una de las muestras distintas de C_n^N tenga la misma oportunidad de ser extraída. Donde:

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (2.1)$$

En la práctica, se extrae una muestra aleatoria simple unidad por unidad. Las unidades de la población se numeran del 1 al N . Luego se genera una serie de números aleatorios entre 1 y N , ya sea por medio de una tabla de números aleatorios o por medio de un programa de computadora que produce tal tabla. En cualquier sorteo, el proceso utilizado debe dar la misma oportunidad de selección a cualquier número de la población que aún no haya sido sorteado. Las unidades que llevan estos n números constituyen la muestra.

Se verifica fácilmente que todas las muestras distintas de C_n^N tienen la misma posibilidad de ser seleccionadas por este método. Considere una muestra distinta, es decir, un conjunto de n unidades específicas. En el primer sorteo, la probabilidad de que alguna de las n unidades especificadas sea seleccionada es $\frac{n}{N}$. En el segundo sorteo, la probabilidad de que alguna de las restantes $(n-1)$ unidades especificadas sea seleccionada es $\frac{n-1}{N-1}$, y así sucesivamente. Por lo tanto, la probabilidad de que todas las n unidades especificadas sean seleccionadas en n sorteos es:

$$\frac{n}{N} \cdot \frac{n-1}{N-1} \cdot \frac{n-2}{N-2} \cdots \frac{1}{N-n+1} = \frac{n!(N-n)!}{N!} = \frac{1}{C_n^N} \quad (2.2)$$

Dado que un número extraído se elimina de la población para todos los sorteos posteriores, este método también se denomina muestreo aleatorio sin reemplazo. El muestreo aleatorio con reemplazo es completamente factible: en cualquier extracción, todos los N miembros de la población tienen la misma oportunidad de ser extraídos, sin importar la frecuencia con la que ya se hayan extraído. Las fórmulas para las varianzas y las varianzas estimadas de las estimaciones realizadas a partir de la muestra suelen ser más sencillas cuando el muestreo es con reemplazo que cuando es sin reemplazo. Por esta razón, a veces se utiliza el muestreo con reemplazo en los planes de muestreo más complejos, aunque a primera vista parece poco útil tener la misma unidad dos o más veces en la muestra.

Muestreo aleatorio estratificado: En el muestreo estratificado, la población de N unidades se divide primero en subpoblaciones de $N_1, N_2, \dots, N_L$ unidades con L el número de estratos. respectivamente. Estas subpoblaciones no se superponen y juntas comprenden la totalidad de la población, de modo que

$$N_1 + N_2 + \dots + N_L = N \quad (2.3)$$

Las subpoblaciones se denominan estratos. Para obtener el máximo beneficio de la estratificación, se deben conocer los valores de N_h (tamaño de cada estrato). Cuando se han determinado los estratos, se toma una muestra de cada uno, realizándose los sorteos independientemente en diferentes estratos. Los tamaños de muestra dentro de los estratos se denotan por $n_1, n_2, \dots, n_L$, respectivamente. Si se toma una muestra aleatoria simple en cada estrato, todo el procedimiento se describe como muestreo aleatorio estratificado.

La estratificación es una técnica común y en particular como se mencionan tanto los ciclos dentro de las criptomonedas, esto podría ser aplicado conjuntamente como cada ciclo un estrato y así sucesivamente.

Muestreo Sistemático: Este método de muestreo es a primera vista bastante diferente del muestreo aleatorio simple. Suponga que las N unidades de la población están numeradas del 1 al N en algún orden. Para seleccionar una muestra de n unidades, tomamos una unidad al azar de las primeras k unidades y cada k -ésima unidad a partir de entonces. Por ejemplo, si $k = 15$ y si la primera unidad extraída es el número $n = 13$, las unidades siguientes serán los números $n = 28, 43, 58, \dots$. La selección de la primera unidad determina toda la muestra. Este tipo se denomina muestra sistemática cada k -ésima. Las ventajas aparentes de este método sobre el muestreo aleatorio simple son las siguientes.

Intuitivamente, parece probable que el muestreo sistemático sea más preciso que el muestreo aleatorio simple. En efecto, estratifica la población en n estratos, que consisten en las primeras k unidades, las segundas k unidades, y así sucesivamente.

Por lo tanto, podríamos esperar que la muestra sistemática sea tan precisa como la muestra aleatoria estratificada correspondiente con una unidad por estrato. La diferencia es que con la muestra sistemática las unidades se encuentran en la misma posición relativa en el estrato, mientras que con la muestra aleatoria estratificada la posición en el estrato se determina por separado mediante la aleatorización dentro de cada estrato.

La muestra sistemática se distribuye de manera más uniforme entre la población, y este hecho a veces ha hecho que el muestreo sistemático sea considerablemente más preciso que el muestreo aleatorio estratificado. Una variante de la muestra sistemática es elegir cada unidad en o cerca del centro del estrato; es decir, en lugar de comenzar la secuencia con un número aleatorio elegido entre 1 y k , tomamos el número inicial como:

$$\begin{array}{ll} \frac{k+1}{2} & k \text{ impar} \\ \frac{k}{2} \text{ o } \frac{k+2}{2} & k \text{ par} \end{array} \quad (2.4)$$

Note que para el caso de k par, se puede elegir el extremo de la izquierda como el de la derecha. Este procedimiento lleva la idea de muestreo sistemático a su conclusión lógica. Si y_i puede considerarse una función continua de una variable continua i , hay motivos para esperar que esta muestra ubicada centralmente sea más precisa que una ubicada aleatoriamente.

Dado que N no es en general un múltiplo entero de k , diferentes muestras sistemáticas de la misma población finita pueden variar en una unidad de tamaño. Así, con $N = 23$, $k = 5$, los números de las unidades en las cinco muestras sistemáticas se muestran en la figura 2.2. Las tres primeras muestras tienen $n = 5$ y las dos últimas tienen $n = 4$. Este hecho introduce una perturbación en la teoría del muestreo sistemático. La perturbación es probablemente insignificante si n excede de 50 y se ignorará, por simplicidad, en la presentación de la teoría. Es poco probable que sea grande incluso cuando n es pequeño.

I	II	III	IV	V
1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23		

Figura 2.2: Ejemplo de muestreo sistemático

Otro método sugerido proporciona un tamaño de muestra constante y una media de muestra sin sesgo. Considere las N unidades dispuestas alrededor de un círculo y sea ahora k el entero más cercano a $\frac{N}{n}$. Seleccione un número aleatorio entre $[1, N]$ y tome cada k -ésima unidad a partir de entonces, dando vueltas alrededor del círculo hasta que se hayan elegido las n unidades deseadas. Supongamos que queremos $n = 5$ con $N = 23$. Entonces $k = 5$. Si el número aleatorio es 19, tomamos las unidades (19, 1, 6, 11, 16). Se verifica fácilmente que

cada unidad tiene la misma probabilidad de selección con este método. Si se quieren $n = 4$ unidades con $N = 23$ tomamos $k = 6$.

Muestreo por grupos: También denominado muestreo por **clusters**, es otra forma de ver el muestreo sistemático. Con $N = nk$, las k posibles muestras sistemáticas se muestran en las columnas de la 2.3. Es evidente a partir de esta tabla que la población se ha dividido en k grandes unidades de muestreo, cada una de las cuales contiene n de las unidades originales. La operación de elegir una muestra sistemática ubicada al azar es solo la operación de elegir una de estas grandes unidades de muestreo al azar. Así, el muestreo sistemático equivale a la selección de una sola unidad de muestreo compleja que constituye la muestra completa. Una muestra sistemática es una muestra aleatoria simple de un grupo de una población de k grupos.

Número de muestra					
1	2	...	i	...	k
y_1	y_2		y_i		y_k
y_{k+1}	y_{k+2}		y_{k+i}		y_{2k}
...	...		...		...
$y_{(n-1)k+1}$	$y_{(n-1)k+2}$		$y_{(n-1)k+i}$		y_{nk}

Figura 2.3: Enumeración de muestras en para el muestreo por grupos.

Ya con estas definiciones, procedemos a hablar detenidamente de cada una de las técnicas que usaremos a lo largo del proyecto, tomando en cuenta que nuestros datos si bien son muestras para cada página web de datos de criptomonedas, al realizar la descarga y proceder a su estudio, para nosotros estos datos pueden contemplar una muestra o una población dependiendo del estudio en cuestión, a la cual le podríamos aplicar a su vez técnicas de muestreo nuevamente.

2.2.2. Estadística.

Engloba un conjunto de métodos para la recolección, organización, análisis e interpretación de datos. Algunos la consideran una ciencia, otros una rama de las matemáticas.

2.2.2.1. Definiciones básicas.

Dato estadístico: Un valor que se obtiene a través de la observación o medición de algún concepto (atributo).

Valor Cualitativo: No numérico. Ejemplo:

- Género: Masculino, Femenino
- Color: Verde, Amarillo, Rojo

Valor Cuantitativo: Numérico (tiene asociado una unidad de medida). Ejemplo:

- Volumen: 22.5 m^3
- Edad: 12 años

Parámetro: Una característica descriptiva global de una población.

Estadígrafo: característica descriptiva global de una muestra que se usa para estimar o inferir un parámetro.

Estadístico: Es cualquier cantidad cuyo valor se pueda calcular a partir de datos muestrales. Un estadístico es una variable aleatoria y estará denotada por una letra mayúscula; una minúscula se emplea para representar el valor calculado u observado del estadístico.

2.2.2.2. Tipos de Estadística.

Estadística Descriptiva: Describe, de forma cuantitativa, las características principales de un conjunto de datos. Resume dichas características a través del cálculo de un conjunto de métricas.

Estadística Inferencial: Permite deducir (hacer inferencias) características o propiedades de un conjunto de datos (todo) a partir de información tomada para un subconjunto (parte) de éste.

Con un enfoque predominante en estadística descriptiva, nos centraremos en dos grupos de medidas, primero un grupo de medidas comúnmente usadas en ciencia para estudiar a grandes rasgos conjuntos de datos, tanto medidas centrales como medidas de dispersión, luego el segundo grupo lo comprenden medidas usadas comúnmente en el ámbito del trading, tanto de criptomonedas como de valores, acciones y metales.

2.2.3. Medidas Estadísticas.

Ahora, veremos las medidas estadísticas que implementaremos en el estudio de nuestros datos, primero veremos un grupo de medidas de tendencia central, éstas las más comunes a la hora de analizar a grandes rasgos una serie de datos, buscando un resumen promedio de los datos por

esto su nombre, luego veremos algunas medidas de dispersión, las cuales nos indicaran que tan dispersos están nuestros datos los unos de los otros, en nuestro caso estas medidas nos podrán decir que tanta volatilidad posee una cotización en específica, por último tendremos una lista de las medidas o estadísticos más comunes a la hora de hacer análisis de series temporales de precios, estas directamente inspiradas en las más usadas en el ámbito del trading.

2.2.3.1. Medidas de Tendencia Central

Representan un valor numérico que se sitúa hacia el “centro” de la distribución conformada por los datos. Estas medidas constituyen un resumen promedio del comportamiento, pero podrían estar sesgadas si se toman como único parámetro de estudio a un grupo de datos, ya que dos grupos distintos podrían poseer medidas centrales de similar o igual valor pero ser una distribución completamente distinta. A continuación las mas usadas y como se denotan:

Media aritmética, $\bar{x}$ o μ : La suma de cada valor del dato estadístico dividido por el número de muestras (n):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \mu \quad 1 \leq i \leq n \quad (2.5)$$

La media aritmética es la medida más simple y más utilizada de una media o promedio. Simplemente involucra tomar la suma de un grupo de números, luego dividir esa suma por el conteo de los números usados en la serie.

En el mundo de las finanzas, la media aritmética no suele ser un método apropiado para calcular un promedio, especialmente cuando un solo valor atípico puede sesgar la media en gran medida.

La media aritmética tampoco es excelente cuando se calcula el rendimiento de las carteras de inversión, especialmente cuando se trata de capitalización o reinversión de dividendos y ganancias. Por lo general, tampoco se usa para calcular los flujos de efectivo presentes y futuros, que los analistas usan para hacer sus estimaciones. Si lo hace, es casi seguro que conducirá a números engañosos [128].

Promedio Ponderado, $\bar{x}_w$: La media ponderada es una medida de tendencia central, que es apropiada cuando en un conjunto de datos cada uno de ellos tiene una importancia relativa (o peso dato por el arreglo w) respecto de los demás datos. Se obtiene multiplicando cada uno de los datos (x_i) por su ponderación (w_i) para luego sumarlos, obteniendo así una suma

ponderada; después se divide esta entre la suma de los pesos, dando como resultado la media ponderada [129].

$$\bar{x}_w = \frac{\sum_{i=1}^n (w_i \cdot x_i)}{\sum_{i=1}^n w_i} \quad 1 \leq i \leq n \quad (2.6)$$

Media Armónica, H: La media armónica ayuda a encontrar relaciones multiplicativas o divisorias entre fracciones sin preocuparse por los denominadores comunes. Los medios armónicos a menudo se usan para promediar cosas como tasas (por ejemplo, la velocidad de viaje promedio dada la duración de varios viajes).

$$H = n \left[\sum_{i=1}^n \frac{1}{x_i} \right]^{-1} \quad 1 \leq i \leq n \quad (2.7)$$

La media armónica no está definida en el caso de que exista algún valor nulo.

Media Geométrica, GM: es el promedio de un conjunto de productos, cuyo cálculo se utiliza comúnmente para determinar los resultados de rendimiento de una inversión o cartera. Se define técnicamente como “el enésimo producto raíz de n números”. La media geométrica debe usarse cuando se trabaja con porcentajes, que se derivan de valores, mientras que la media aritmética estándar funciona con los valores mismos.

$$GM = \sqrt[n]{\prod_{i=1}^n x_i} \quad 1 \leq i \leq n \quad (2.8)$$

La media geométrica es la tasa de retorno promedio de un conjunto de valores calculada usando los productos de los términos, es más apropiada para las series que exhiben una correlación serial; esto es especialmente cierto para las carteras de inversión.

La mayoría de los rendimientos en finanzas están correlacionados, incluidos los rendimientos de los bonos, los rendimientos de las acciones y las primas de riesgo de mercado.

La media geométrica, a veces denominada tasa de crecimiento anual compuesta o tasa de rendimiento ponderada en el tiempo, es la tasa de rendimiento promedio de un conjunto de valores calculados utilizando los productos de los términos. ¿Qué significa eso? La media geométrica toma varios valores, los multiplica y los eleva a la potencia $1/n$ [130].

Mediana, Med: es el valor central del grupos de datos ordenados. Para un conjunto de datos, puede considerarse como el valor “medio”. La característica básica de la mediana en la descripción de datos en comparación con la media (a menudo descrita simplemente como el “promedio”) es que no está sesgada por una pequeña proporción de valores extremadamente grandes o pequeños y, por lo tanto, proporciona una mejor representación de un “valor típico”. La mediana tiene una importancia central en las estadísticas sólidas, ya que es la estadística más resistente, con un punto de ruptura del 50 %: siempre que no más de la mitad de los datos estén contaminados, la mediana no es un resultado arbitrariamente grande o pequeño [131].

$$\text{Med} = \begin{cases} x_{\frac{n+1}{2}} & n \text{ impar} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & n \text{ par} \end{cases} \quad (2.9)$$

Moda, Mod: Valor que más se repite en la muestra. Un conjunto de datos puede tener una moda o más de una moda.

$$\text{Mod} = x_j \quad \text{Freq}(x_i)_{\max} = \text{Freq}(x_j) \quad 1 \leq i, j \leq n \quad (2.10)$$

Media Cuadrática, $\bar{x}^2$: es una medida estadística de la magnitud de una cantidad variable.

Puede calcularse para una serie de valores discretos o para una función matemática de variable continua. El nombre deriva del hecho de que es la raíz cuadrada de la media aritmética de los cuadrados de los valores.

A veces la variable toma valores positivos y negativos, como ocurre, por ejemplo, en los errores de medida. En tal caso se puede estar interesado en obtener un promedio que no recoja los efectos del signo. Este problema se resuelve, mediante la denominada media cuadrática. Consiste en elevar al cuadrado todas las observaciones (así los signos negativos desaparecen), en obtener después su media aritmética y en extraer, finalmente, la raíz cuadrada de dicha media para volver a la unidad de medida original [132].

$$\bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 \quad 1 \leq i \leq n \quad (2.11)$$

Las medidas de posición central son una ayuda en forma de resumen pero no son categóricas. Como resumen pueden darnos una información de lo que, en promedio, cabría esperar. Pero no siempre son precisas.

Para analizar mejor estas medidas, es recomendable combinar las medidas de tendencia central con medidas de dispersión. Las medidas de dispersión tampoco son infalibles, pero nos ofrecen información sobre la variabilidad de una determinada variable [133].

2.2.3.2. Medidas de Dispersión

La dispersión estadística conocida también como variabilidad o variación estadística, es la variabilidad o extensión en una variable o una distribución de probabilidad. Una medida de dispersión estadística es un número real no negativo, que vale cero si todos los datos son idénticos, y aumenta a medida que los datos están más dispersos. Medidas comunes de dispersión:

Rango, rng: Es la diferencia entre el valor máximo y el mínimo en la muestra.

$$\text{rng} = (x_i)_{\max} - (x_i)_{\min} \quad 1 \leq i \leq n \quad (2.12)$$

Dado que la volatilidad de los precios es equivalente al riesgo, el rango de cotización de un valor es un buen indicador de riesgo. Un inversor conservador prefiere valores con fluctuaciones de precios más pequeñas en comparación con valores que son susceptibles a giros significativos [134].

Varianza, S_x^2 : mide la dispersión de los valores respecto a un valor central (media), específicamente el cuadrado de las desviaciones respecto a dicho valor.

$$S_x^2 = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2 \quad 1 \leq i \leq n \quad (2.13)$$

Desviación Estándar, S_x : también llamada desviación estándar Es la raíz cuadrada (positiva) de la varianza, se utiliza para eliminar el uso de unidades cuadráticas.

$$S_x = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad 1 \leq i \leq n \quad (2.14)$$

Coefficiente de Variación, d : cociente entre la desviación estándar y el promedio. Con esta medida se elimina la influencia de la escala escogida en las mediciones efectuadas.

$$d = \frac{S_x}{\bar{x}} \quad (2.15)$$

Hasta aquí los estadísticos convencionales. Éstos se suelen calcular para muestras de n elementos, mientras que las siguientes se emplean de manera distinta como veremos a continuación.

2.2.3.3. Medidas usadas en Trading

El mundo de la economía por tradición siempre ha querido ir aparte de todas las ciencias con cálculos, ecuaciones y técnicas que mas allá de tener una formula o ecuación para el éxito de la técnica, se basa en parte en astucia y experiencia al interpretar los datos mediante varias técnicas en simultáneo. Muchas de estas técnicas con un fundamento un poco escaso mas allá de que en numerosas ocasiones funciona, todo esto tiene un trasfondo, y es que es un sistema conformado por personas o grupos tratando de sacar el mejor provecho de cada transacción, cada uno por su lado, mientras una parte quiere comprar a precios bajos, otros quieren vender a precios altos y de igual manera muchos quieren vender justo antes de que el precio se desplome y otros quieren comprar antes de que el precio se catapulte a niveles históricos.

A continuación una selección de medias comunes usadas en trading, cabe destacar que la mayoría de estos indicadores se suelen calcular de forma recursiva sobre la serie de datos, ya que estos se usan en parte con el fin de ver la evolución del sistema mas que con el propósito de medirlo en un punto específico, de ésta manera la mayoría se calculan tomando una ventana de m elementos con $m > 1$, la cual se va trasladando sobre los n elementos de una muestra, generando así a su vez $n - m + 1$ medidas distintas de cero, desde los primeros m elementos hasta los últimos m elementos. La notación común es colocar el nombre del estadístico y el numero de la ventana entre paréntesis, ya que varios de estos cuentan con varios cálculos dentro de sí y con ventanas distintas, podemos ver estadísticos de este estilo con varios números como argumento dentro del paréntesis, cada uno refiriéndose a una ventana de un cálculo en específico, es aquí donde debemos tratar de seguir los usos convencionales para no caer en cálculos o interpretaciones erradas.

Media Móvil, $MA(m)$: En finanzas, una media móvil (MA del inglés Moving Average) es un indicador bursátil que se usa comúnmente en el análisis técnico. La razón para calcular el promedio móvil de una acción es ayudar a suavizar los datos de precios mediante la creación de un precio promedio constantemente actualizado.

Al calcular la media móvil, se mitigan los impactos de las fluctuaciones aleatorias a corto plazo en el precio de una acción durante un período de tiempo específico. Debido a la forma como se vaya actualizando esta media podemos encontrar varios tipos de medias móviles, las dos mas comunes son:

MA Simple, $SMA(m)_x$] conocida como media móvil simple (SMA del inglés Simple MA) de la variable x , se calcula tomando la media aritmética de un conjunto dado de m registros recursivamente a lo largo de los n elementos totales. En otras palabras, un conjunto de registros (x_i) o precios (p_i) en el caso de instrumentos financieros, se suman y luego se dividen por el número de registros en el conjunto, luego desechamos el primer registro y añadimos el siguiente, así hasta llegar a añadir el enésimo registro [135]. La ecuación para una $SMA(m)_x$ es la siguiente:

$$SMA_i = \begin{cases} 0 & i < m \\ \frac{1}{m} \sum_{j=1}^m x_j & i = m \\ SMA_{i-1} + \frac{1}{m}(x_{i+m} - x_{i-m-1}) & m < i \leq n \end{cases} \quad (2.16)$$

De esta forma tendríamos la serie:

$$SMA(m)_x = \{0, \dots, 0, SMA_m, \dots, SMA_n\} \quad (2.17)$$

MA Exponencial, $EMA(m)_x$: La media móvil exponencial (EMA del inglés Exponential MA) de la variable x , es un tipo de media móvil que da más peso a los datos recientes en un intento de que responda mejor a la nueva información. Para calcular una $EMA(m)_x$, primero debe calcular la $SMA(m)_x$. A continuación, debe calcular el multiplicador para ponderar la $EMA(m)_x$, denominado “factor de suavizado: α ”, que normalmente sigue la ecuación:

$$\alpha = \frac{2}{m + 1} \quad (2.18)$$

La $EMA(m)_x$ para la serie $SMA(m)_x$ se debe calcular recursivamente:

$$EMA_i = \begin{cases} 0 & i < m \\ SMA_m & i = m \\ \alpha SMA_i + (1 - \alpha) EMA_{i-1} & m < i \leq n \end{cases} \quad (2.19)$$

De esta forma tendríamos la serie:

$$EMA(m)_x = \{0, \dots, 0, SMA_m, EMA_{m+1}, \dots, EMA_n\} \quad (2.20)$$

En finanzas, los analistas técnicos suelen utilizar promedios móviles para realizar un se-

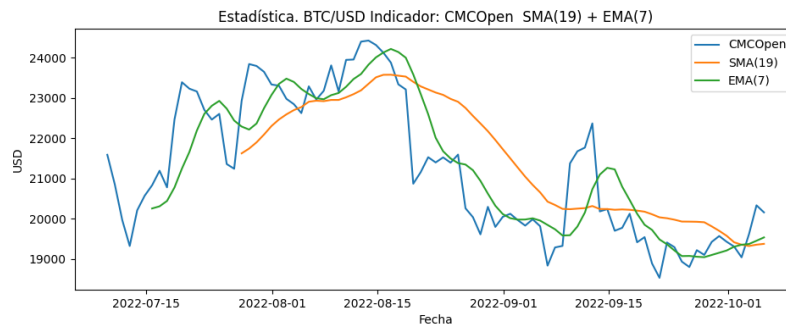


Figura 2.4: Gráfico de promedio y dos medias móviles.

guimiento de las tendencias de precios de valores específicos. Una tendencia alcista en un promedio móvil podría significar un aumento en el precio o el impulso de un valor, mientras que una tendencia a la baja se consideraría una señal de declive. Hoy en día, hay una amplia variedad de promedios móviles para elegir, que van desde medidas simples hasta fórmulas complejas que requieren un programa de computadora para calcular de manera eficiente.

Otras veces, usarán promedios móviles para confirmar sus sospechas de que podría estar ocurriendo un cambio. Por ejemplo, si el precio de las acciones de una empresa sube por encima de su promedio móvil de 200 días, eso podría tomarse como una señal alcista [136]. En la figura 2.4 se puede observar dos medias móviles una SMA con ventana 19 y una EMA con ventana $m=7$ para la cotización BTC/USD, datos diarios Open del sitio CMC.

Bandas de Bollinger, $BB(m, b)_x$: fueron creadas a principios de la década de 1980 por el analista financiero y comerciante John Bollinger. Son ampliamente utilizados como instrumento de análisis técnico sobre la variable x . Básicamente, las Bandas de Bollinger funcionan como un medidor de osciladores. Indica si el mercado tiene alta o baja volatilidad, así como condiciones de sobrecompra o sobreventa. Sus argumentos indican el tamaño de la ventana igual a m , y un valor b que multiplica la S_x de los valores dentro de la ventana.

La idea principal detrás del indicador Bandas de Bollinger es resaltar cómo los precios se dispersan alrededor de un valor promedio. Más específicamente, se compone de una banda superior, una banda inferior y una línea media también conocida como señal que suele ser una SMA de los datos originales. Las dos bandas laterales reaccionan a la acción del precio del mercado, expandiéndose cuando la volatilidad es alta (alejándose de la línea media) y contrayéndose cuando la volatilidad es baja (moviéndose hacia la línea media). La configuración estándar de $BB(m, b)_x$ establece $m = 20$ y $b = 2$.

Veamos ahora como realizar los cálculos, tomando en cuenta que cada uno de estos tres

cálculos, genera una serie de n valores:

1. Línea media, $LM(m)_x$: una serie idéntica a una $SMA(m)_x$ de los datos originales.

$$LM_i = SMA_i \quad 1 \leq i \leq n \quad (2.21)$$

2. Banda superior, $BS(b)$: esta serie es el resultado de multiplicar cada registro de la serie $S(m)_x$ por b y sumarle el valor respectivo de LM_i , de esta forma, realizaremos primero el cálculo de la serie $S(m)_x$:

$$S_{x_i} = \begin{cases} 0 & i < m \\ \sqrt{\frac{1}{(m-1)} \sum_{j=1}^m (x_j - LM_m)^2} & i = m \\ \sqrt{\frac{1}{(m-1)} \sum_{j=i-m-1}^i (x_j - LM_i)^2} & i > m \end{cases} \quad 1 \leq i, j \leq n \quad (2.22)$$

De esta forma tendríamos la serie:

$$S(m)_x = \{0, \dots, 0, S_{x_m}, \dots, S_{x_n}\} \quad (2.23)$$

Y así tendríamos todo para calcular la serie $BS(b)$:

$$BS_i = LM_i + b S_{x_i} \quad 1 \leq i \leq n \quad (2.24)$$

3. Banda inferior, $BI(m)$: línea media menos dos veces la desviación estándar $S(m)_{x_i}$.

$$BI_i = LM_i - b S_{x_i} \quad 1 \leq i \leq n \quad (2.25)$$

El factor b que multiplica la desviación estándar se añade para garantizar que al menos el 85 % de los datos de precios se muevan entre estas dos bandas, pero la configuración puede ajustarse según las diferentes necesidades y estrategias comerciales, en la figura 2.5 se puede apreciar la especificación $BB(21,2)$ para el par BTCGBP y datos Open del sitio CMC para intervalos diarios.

Retrocesos de Fibonacci, Fibo(f_i): Los números de Fibonacci se utilizan para crear indicadores técnicos utilizando una secuencia matemática desarrollada por el matemático italiano,

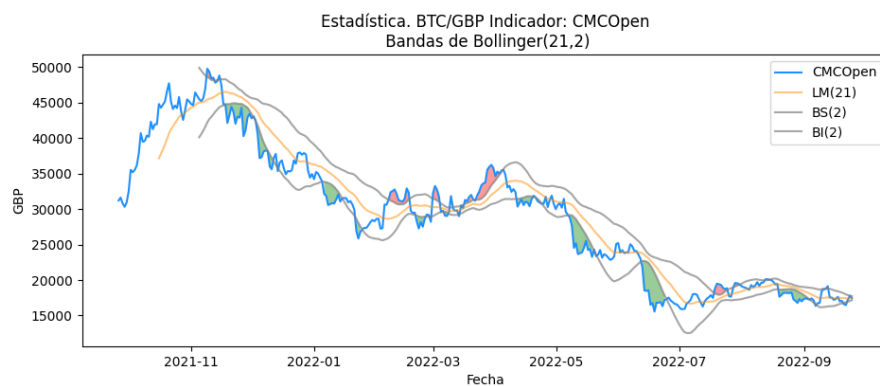


Figura 2.5: Bandas de Bollinger.

comúnmente conocido como “Fibonacci”, en el siglo XIII. La secuencia de números, comenzando con cero y uno, se crea sumando los dos números anteriores.

Esta secuencia se puede desglosar en proporciones que, según algunos, brindan pistas sobre hacia dónde se moverá un mercado financiero determinado.

La sucesión de Fibonacci es significativa para las finanzas debido a la llamada proporción áurea de 1.618, o su inversa 0.618. En la sucesión de Fibonacci, cualquier número dado es aproximadamente 1.618 veces el número anterior, ignorando los primeros 5 números de la serie. Cada número es también 0.618 veces el número siguiente, nuevamente ignorando los primeros números de la secuencia. La proporción áurea es omnipresente en la naturaleza.

La secuencia de Fibonacci se puede usar de diferentes maneras para obtener niveles de retroceso de Fibonacci o niveles de extensión de Fibonacci, que son los que nos interesan aquí, mas allá que la sucesión en si.

Para representar los retrocesos de Fibonacci se requiere que se elijan dos puntos de precio en un gráfico, por lo general, un máximo y un mínimo o dado una período de tiempo, buscar el valor máximo y mínimo que toma la variable. Una vez que se eligen esos dos puntos, los niveles de Fibonacci se representan en forma de líneas horizontales que indican dónde es probable que ocurran el soporte y la resistencia, cada nivel está asociado a un porcentaje.

El porcentaje es cuánto de un movimiento anterior ha retrocedido el precio. Los niveles de retroceso de Fibonacci son 23.6 %, 38.2 %, 61.8 %, 78.6 %, 161.8 %, 261.8 % y 423.6 %. Si bien no es oficialmente una proporción de Fibonacci, también se usa el 0.0 %, 50.0 %, 100.0 % y 200.0 %. Los niveles de retroceso de Fibonacci se grafican como líneas horizontales. De ésta manera los valores porcentuales posibles que puede tomar la variable f_i son:

$$f_i = \{0.0, 23.6, 38.2, 50.0, 61.8, 78.6, 100.0, 161.8, 200.0, 261.8, 423.6\} \quad (2.26)$$

De ésta manera si tenemos los n registros de la variable x durante un período de tiempo, los retrocesos de Fibonacci se calcularían de la siguiente manera:

$$Fibo(f_i) = \begin{cases} \max(x_i) & f_i = 0 \\ \min(x_i) & f_i = 100 \\ \max(x_i) - \left(\frac{\max(x_i) \cdot f_i}{100} \right) & f_i \neq \{0.0, 100.0\} \end{cases} \quad (2.27)$$

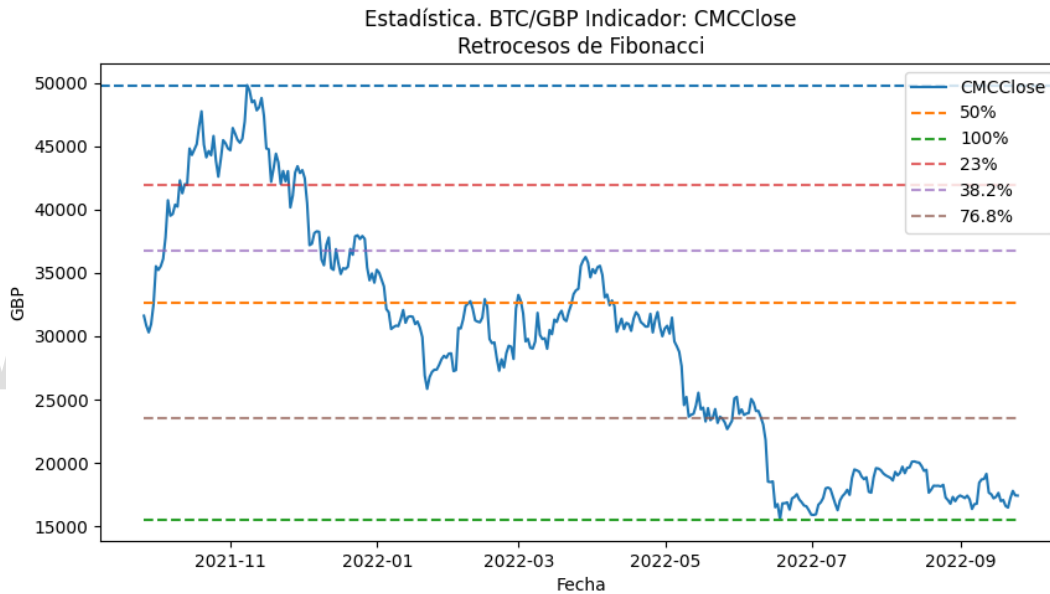


Figura 2.6: Retrocesos de Fibonacci.

El uso de los estudios de Fibonacci es subjetivo, ya que el comerciante debe utilizar los máximos y mínimos de su elección. Los máximos y mínimos que se elijan afectarán los resultados que obtenga un operador, en la figura 2.6 se puede observar este estadístico para el par BTCGBP y datos Close del sitio CMC para intervalos diarios.

Otro argumento en contra de los métodos de comercio de números de Fibonacci es que hay tantos de estos niveles que el mercado está obligado a rebotar o cambiar de dirección cerca de uno de ellos, lo que hace que el indicador parezca significativo en retrospectiva. El problema es que es difícil saber qué número o nivel será importante en tiempo real o en el futuro [137] [138].

Osciladores: es una herramienta de análisis técnico que construye bandas altas y bajas entre

dos valores extremos y luego construye un indicador de tendencia que fluctúa dentro de estos límites. Los operadores utilizan el indicador de tendencia para descubrir condiciones de sobrecompra o sobreventa a corto plazo. Cuando el valor del oscilador se acerca al valor extremo superior, los analistas técnicos interpretan esa información como que el activo está sobrecomprado, y cuando se acerca al extremo inferior, los técnicos consideran que el activo está sobrevendido [139]. Entre los mas comunes tenemos:

MA de Convergencia Divergencia, MACD(m_1, m_2, m_3) : (MACD del inglés MA Convergence Divergence) es un indicador de tipo oscilador ampliamente utilizado por traders para análisis técnico. El MACD(m_1, m_2, m_3) es un instrumento de seguimiento de tendencias que emplea medias móviles para determinar el “momentum” de un valor, criptomoneda u otro activo negociable.

Desarrollado por Gerald Appel a finales de la década de 1970, el indicador de MACD(m_1, m_2, m_3) rastrea acciones del precio que ya han tenido lugar y, por tanto, cae en la categoría de indicadores retrospectivos. El MACD(m_1, m_2, m_3) puede resultar útil para calibrar el momento de mercado y las posibles tendencias del precio, por ello es utilizado por muchos traders para identificar puntos de entrada y salida potenciales [140]. Veamos sus tres partes:

1. La línea MACD: ayuda a determinar el momento alcista o bajista (tendencia de mercado). Se calcula restando dos medias móviles exponenciales de ventanas m_1 y m_2 con $m_1 < m_2$. Primero calculamos las dos EMA:

$$\begin{aligned} \text{EMA1}(m_1)_x &= \text{EMA}(m_1)_x \\ &= \{0, \dots, 0, \text{EMA1}_{m_1}, \dots, \text{EMA1}_n\} \quad 1 \leq i \leq n \end{aligned} \quad (2.28)$$

$$\begin{aligned} \text{EMA2}(m_2)_x &= \text{EMA}(m_2)_x \\ &= \{0, \dots, 0, \text{EMA2}_{m_2}, \dots, \text{EMA2}_n\} \quad 1 \leq i \leq n \end{aligned} \quad (2.29)$$

Y ya con la serie $\text{EMA1}(m_1)_x$ y $\text{EMA2}(m_2)_x$ procedemos a calcular los elementos de la serie línea MACD.

$$\text{MACD}_i = \text{EMA1}_i - \text{EMA2}_i \quad 1 \leq i \leq n \quad (2.30)$$

2. La línea de señal, LS: una EMA de la serie MACD. El análisis combinado de la línea de señal y la línea MACD puede resultar útil para identificar cambios de tendencia o puntos de entrada y salida potenciales.

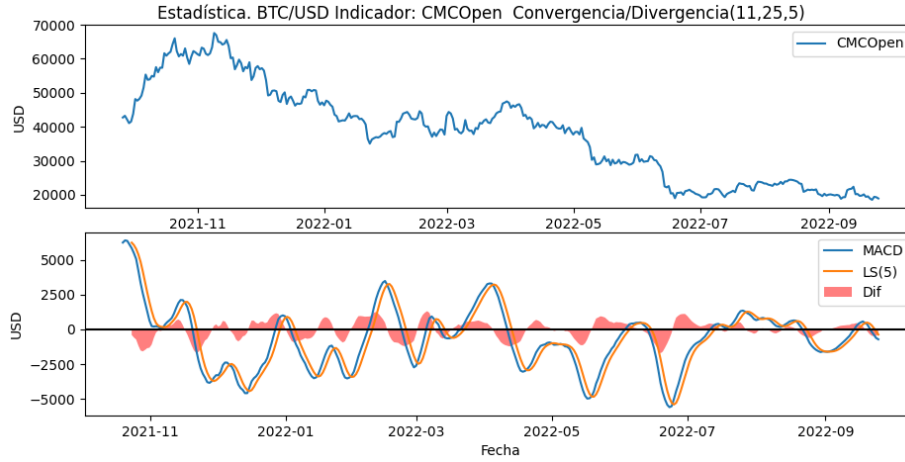


Figura 2.7: Media móvil de Convergencia Divergencia.

$$LS(m_3) = EMA(m_3)_{MACD} \quad (2.31)$$

A pesar de que no siempre son eventos precisos, cuando la líneas MACD y de señal se cruzan, habitualmente se consideran señales de reversión de tendencia, especialmente cuando se producen en los extremos del gráfico $MACD(m_1, m_2, m_3)$, bastante por encima o por debajo de la línea cero.

3. Diferencia, Dif: representación gráfica de la divergencia y convergencia de las líneas MACD y de señal. La diferencia se calcula simplemente restando una de la otra:

$$Dif_i = MACD_i - LS_i \quad 1 \leq i \leq n \quad (2.32)$$

Sin embargo, en lugar de añadir una tercer trazo al gráfico, la diferencia se representa mediante un gráfico de barras, lo que hace que sea visualmente más fácil de leer e interpretar. En la figura 2.7 se puede observar un gráfico de $MACD(11,25,5)$ para el par BTCUSD y datos Open del sitio CMC para intervalos diarios, en este caso la diferencia (Dif) se representa como un área sombreada.

La configuración predeterminada para el $MACD(m_1, m_2, m_3)$ se basa en los valores $m_1 = 12$, $m_2 = 26$ y $m_3 = 9$. Sin embargo, algunos analistas técnicos y chartistas cambian la configuración de períodos con el objetivo de conseguir un indicador más sensible. Por ejemplo, el $MACD(5, 35, 5)$ es uno de los que se usan con frecuencia en los mercados financieros tradicionales junto con intervalos de tiempo más prolongados como por ejemplo gráficos semanales o mensuales.

Vale la pena señalar que debido a la elevada volatilidad de los mercados de cripto-

monedas, el incremento de la sensibilidad del indicador $MACD(m_1, m_2, m_3)$ puede ser arriesgado porque puede conducir con gran probabilidad a un mayor número de falsas señales e informaciones engañosas.

Índice de fuerza relativa, $RSI(m)$: (RSI del inglés Relative Strength Index) es un indicador de impulso utilizado en el análisis técnico que mide la magnitud de los cambios de precios recientes para evaluar las condiciones de sobrecompra o sobreventa en el precio de una acción u otro activo. El RSI se muestra como un oscilador y puede tener una lectura de 0 a 100. El indicador fue desarrollado originalmente por J. Welles Wilder Jr [141].

Este indicador en su forma inicial usa datos de Close (C), pero podría extrapolarse a usar cualquier indicador (x). Primero calculamos las series ganancias G_i (del inglés Gain) y pérdidas L_i (del inglés Loss), ambas series son valores porcentuales:

$$G_i = \begin{cases} 0 & i = 1 \\ 100 \frac{x_i - x_{i-1}}{x_{i-1}} & x_i > x_{i-1} \\ 0 & x_i \leq x_{i-1} \end{cases} \quad 1 \leq i \leq n \quad (2.33)$$

$$L_i = \begin{cases} 0 & i = 1 \\ 100 \frac{x_{i-1} - x_i}{x_{i-1}} & x_i < x_{i-1} \\ 0 & x_i \geq x_{i-1} \end{cases} \quad 1 \leq i \leq n \quad (2.34)$$

Ahora procedemos a calcular la serie de fuerza relativa (RS), para esto necesitaremos tanto el promedio ganancias $AG(m)$ y pérdidas $AL(m)$ (A de average en inglés). De tal manera que están dados por:

$$AG_i = \begin{cases} 0 & i < m \\ \frac{1}{m} \sum_{j=1}^m G_j & i = m \\ \frac{1}{m} (AG_{i-1} \cdot (m-1)) + G_i & i > m \end{cases} \quad 1 \leq i, j \leq n \quad (2.35)$$

$$AL_i = \begin{cases} 0 & i < m \\ \frac{1}{m} \sum_{j=1}^m L_i & i = m \\ \frac{1}{m}(AL_{i-1} \cdot (m-1)) + L_i & i > m \end{cases} \quad (2.36)$$

De ésta manera procedemos al cálculo de la serie RS:

$$RS_i = \begin{cases} 0 & n \leq m \\ \frac{AG_i}{AL_i} & i \geq m \end{cases} \quad (2.37)$$

Pero como se verá a continuación, puede que solo necesitemos un valor de esta serie para proceder al cálculo de toda la serie RSI:

$$RS_m = \frac{AG_m}{AL_m} \quad (2.38)$$

Y de ésta manera:

$$RSI_i = \begin{cases} 0 & i < m \\ 100 - \left[\frac{100}{1 + RS_m} \right] & i = m \\ 100 - \left[\frac{100}{1 + \frac{AL_i}{AG_i}} \right] & i > m \end{cases} \quad (2.39)$$

La interpretación y el uso tradicionales del RSI son que los valores de 70 o más indican que un valor se está sobrecomprando o sobrevaluando y puede estar preparado para una inversión de tendencia o un retroceso correctivo en el precio. Una lectura de RSI de 30 o menos indica una condición de sobreventa o infravaloración [142].

El RSI aumentará a medida que aumente el número y el tamaño de las ganancias, y caerá a medida que aumente el número y el tamaño de las pérdidas. La segunda parte del cálculo suaviza el resultado, por lo que el RSI solo se acercará a 100 o 0 en un mercado con una fuerte tendencia.

Como puede ver en el gráfico de la figura 2.8 un RSI(20) para el par BTCGBP y datos Low del sitio CMC para intervalos diarios, el indicador RSI puede permanecer en la región de sobrecompra durante períodos prolongados mientras las acciones tienen una

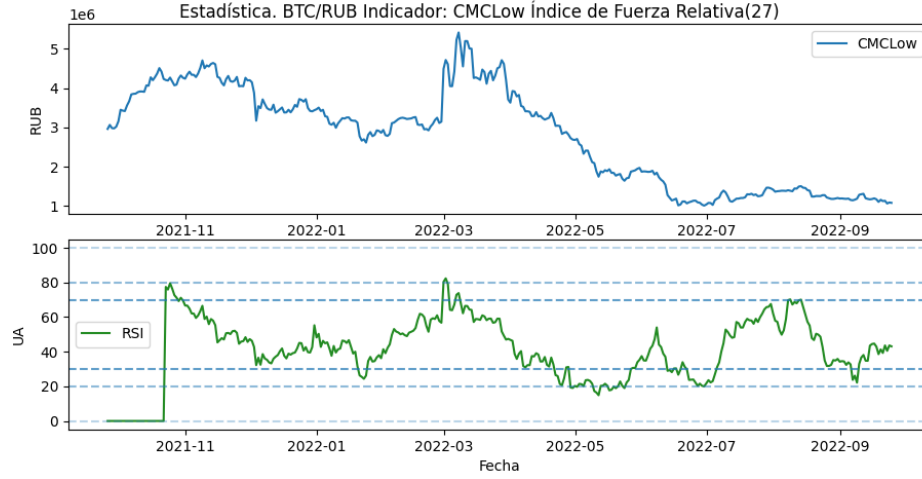


Figura 2.8: Índice de fuerza relativa.

tendencia alcista. El indicador también puede permanecer en territorio de sobreventa durante mucho tiempo cuando las acciones tienen una tendencia bajista.

Oscilador estocástico, $K(m)$: es un indicador de impulso que compara un precio de cierre particular de un valor con un rango de sus precios durante un cierto período de tiempo, de aquí la aparición de dos valores nuevos, el alto del período $PH(m)$ (Period High, en inglés) y el bajo del período $PL(m)$ (Period Low, en inglés). La sensibilidad del oscilador a los movimientos del mercado se puede reducir ajustando ese período de tiempo o tomando un promedio móvil del resultado. Se utiliza para generar señales comerciales de sobrecompra y sobreventa, utilizando un rango de valores limitado de 0 a 100.

Veamos como generar las series $PH(m)$ y $PL(m)$:

$$PH_i = \begin{cases} 0 & i < m \\ \max[x_{i-m+1}, x_{i-m+2}, \dots, x_i] & i \geq m \end{cases} \quad (2.40)$$

$$PL_i = \begin{cases} 0 & i < m \\ \min[x_{i-m+1}, x_{i-m+2}, \dots, x_i] & i \geq m \end{cases} \quad (2.41)$$

Ya con estas series y la serie Close (C_i), podemos proceder a calcular la serie $K(m)$, recordando que este es un valor porcentual.

$$K_i = 100 \left(\frac{C_i - PL_i}{PH_i - PL_i} \right) \quad 1 \leq i \leq n \quad (2.42)$$

En particular, $K(m)$ se conoce a veces como el indicador estocástico rápido. El indicador estocástico “lento” $D(m_2)$, también con un valor porcentual, se consigue al aplicarle una SMA con ventana m_2 a la serie, se suele usar $m_2 = 3$, es común que este valor no se deje a elección del usuario, de tal manera que $K(m)$ tendrá por omisión un valor de $m_2 = 3$.

$$D(m_2)_K = SMA(m_2)_K \quad (2.43)$$

La teoría general que sirve como base para este indicador es que en un mercado con tendencia al alza, los precios cerrarán cerca del máximo, y en un mercado con tendencia a la baja, los precios cerrarán cerca del mínimo.

Tradicionalmente, las lecturas superiores a 80 se consideran en el rango de sobrecompra y las lecturas inferiores a 20 se consideran sobrevendidas. Sin embargo, estos no siempre son indicativos de una reversión inminente; las tendencias muy fuertes pueden mantener condiciones de sobrecompra o sobreventa durante un período prolongado. En su lugar, los operadores deben buscar cambios en el oscilador estocástico en busca de pistas sobre futuros cambios de tendencia.

Los gráficos del oscilador estocástico generalmente consisten en dos líneas: una que refleja el valor real del oscilador para cada sesión y otra que refleja su promedio móvil simple de tres días. Debido a que se cree que el precio sigue el impulso, la intersección de estas dos líneas se considera una señal de que puede estar en marcha una reversión, ya que indica un gran cambio en el impulso de un día a otro.

La divergencia entre el oscilador estocástico y la acción del precio de tendencia también se considera una señal de reversión importante. Por ejemplo, cuando una tendencia bajista alcanza un nuevo mínimo más bajo, pero el oscilador imprime un mínimo más alto, puede ser un indicador de que los bajistas están agotando su impulso y se está gestando una reversión alcista [143].

Índice direccional promedio, $ADX(m)$: (del inglés Average Directional index) es un indicador de análisis técnico utilizado por algunos comerciantes para determinar la fuerza de una tendencia.

La tendencia puede ser alcista o bajista, y esto se muestra mediante dos indicadores adjuntos, el indicador direccional negativo **-DI**, de ahora en adelante NDI y el indicador direccional positivo **+DI**, de ahora en adelante PDI. Por lo tanto, el ADX comúnmente incluye tres líneas separadas. Estas se utilizan para ayudar a evaluar si se debe realizar una operación

larga o corta, o si se debe realizar una operación en absoluto.

Para calcular el ADX primero debemos realizar una serie de cálculos basados en las cotizaciones de H_i , L_i y C_i ; primero generaremos tres series de datos:

La serie movimiento direccional positivo PDM (Positive Directional Movement en inglés):

$$PDM_i = H_i - H_{i-1} \quad PDM_1 = H_1 \quad 1 < i \leq n \quad (2.44)$$

Luego de manera análoga la serie movimiento direccional negativo NDM (Negative Directional Movement en inglés):

$$NDM_i = L_{i-1} - L_i \quad NDM_1 = -L_1 \quad 1 < i \leq n \quad (2.45)$$

Por último en este primer paso, la serie de rango verdadero TR (True Range en inglés):

$$TR_i = \max[(H_i - L_i), |H_i - C_{i-1}|, |L_i - C_{i-1}|] \quad C_0 = 0 \quad 1 \leq i \leq n \quad (2.46)$$

Luego de generar estas series, procedemos a modificar las series PDM y NDM conjuntamente de la siguiente manera, estas serie se suele suavizar, por lo que añadiremos una “S” al nombre, esto se logra mediante la aplicación de una $SMA(m_2)$ donde m_2 suele no ser opcional si no implícito en el cálculo, se suele usar $m_2 = 3$.

$$PDM_i = \begin{cases} PDM_i & PDM_i > NDM_i \\ 0 & PDM_i \leq NDM_i \end{cases} \quad 1 \leq i \leq n \quad (2.47)$$

De manera análoga:

$$NDM_i = \begin{cases} NDM_i & NDM_i > PDM_i \\ 0 & NDM_i \leq PDM_i \end{cases} \quad 1 \leq i \leq n \quad (2.48)$$

El suavizado de ambas sería tal que:

$$PDMS = SMA(m_2)_{PDM} \quad (2.49)$$

$$\text{NDMS} = \text{SMA}(m_2)_{\text{NDM}} \quad (2.50)$$

Luego con la serie de TR pasamos a calcular el TR promedio o $\text{ATR}(m)$ (Average TR en inglés):

$$\text{ATR}_i = \begin{cases} 0 & i < m \\ \frac{1}{m} \sum_{j=1}^m \text{TR}_j & i = m \\ \frac{(i-1)\text{ATR}_{i-1} + \text{TR}_i}{m} & i > m \end{cases} \quad 1 \leq i, j \leq n \quad (2.51)$$

Ya con estas tres series NDMS, PDMS, $\text{ATR}(m)$, pasamos a calcular las series PDI y NDI:

$$\text{PDI}_i = 100 \frac{(\text{PDMS})_i}{\text{ATR}_i} \quad 1 \leq i \leq n \quad (2.52)$$

De manera análoga:

$$\text{NDI}_i = 100 \frac{(\text{NDMS})_i}{\text{ATR}_i} \quad 1 \leq i \leq n \quad (2.53)$$

Y así, procedemos a calcular la serie índice direccional DX:

$$\text{DX}_i = 100 \frac{|\text{PDI}_i - \text{NDI}_i|}{|\text{PDI}_i + \text{NDI}_i|} \quad 1 \leq i \leq n \quad (2.54)$$

Por último calculamos el ADX:

$$\text{ADX}_i = \begin{cases} 0 & i < m \\ \frac{1}{m} \sum_{j=1}^m \text{DX}_j & i = m \\ \frac{(m-1)\text{ADX}_{i-1} + \text{DX}_i}{m} & i > m \end{cases} \quad 1 \leq i, j \leq n \quad (2.55)$$

El $\text{ADX}(m)$, el NDI y el PDI son indicadores de impulso. El $\text{ADX}(m)$ ayuda a los inversores a determinar la fuerza de la tendencia, mientras que NDI y PDI ayudan a determinar la dirección de la tendencia.

El $ADX(m)$ identifica una tendencia fuerte cuando el ADX está por encima de 25 y una tendencia débil cuando el ADX está por debajo de 20. Los cruces de las líneas NDI y PDI se pueden utilizar para generar señales comerciales. En la figura 2.9 podemos observar una configuración $ADX(37)$ para el par BTC/RUB y datos Close del sitio CMC para intervalos diarios.

Como cualquier indicador, el $ADX(m)$ debe combinarse con el análisis de precios y potencialmente con otros indicadores para ayudar a filtrar las señales y controlar el riesgo [144].

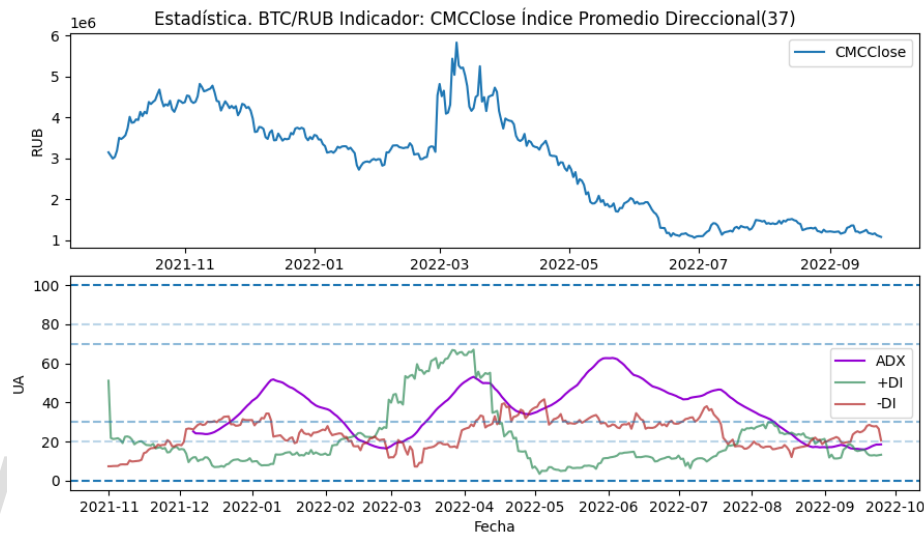


Figura 2.9: Índice direccional promedio.

Nube Ichimoku, $NI(m_1, m_2, m_3)$: es una colección de indicadores técnicos que muestran los niveles de soporte y resistencia, así como el impulso y la dirección de la tendencia. Lo hace tomando múltiples promedios y representándolos en un gráfico. También utiliza estas cifras para calcular una “nube” que intenta pronosticar dónde el precio puede encontrar soporte o resistencia en el futuro.

La Nube Ichimoku fue desarrollada por Goichi Hosoda, un periodista japonés, y publicada a fines de la década de 1960. Proporciona más puntos de datos que el gráfico de velas estándar. Si bien parece complicado a primera vista, aquellos que están familiarizados con la lectura de gráficos a menudo lo encuentran fácil de entender con señales comerciales bien definidas. Las siguientes son las cinco ecuaciones para las líneas que componen el indicador Nube Ichimoku, y sus valores de ventana usuales.

1. Línea de Conversión, $TS(m_1)$: basándonos en las series definidas anteriormente $PH(m_1)$ y $PL(m_1)$, calcularemos su promedio

$$TS(m_1)_i = \frac{PH(m_1)_i - PL(m_1)_i}{2} \quad 1 \leq i \leq n \quad (2.56)$$

2. Línea Base, $KS(m_2)$: se calcula de manera análoga a la serie $TS(m)$ pero con un valor de $m = m_2$, con la condición: $m_2 > m_1$. Tal que con las series $PH(m_2)$ y $PL(m_2)$,

$$KS(m_2)_i = \frac{PH(m_2)_i - PL(m_2)_i}{2} \quad 1 \leq i \leq n \quad (2.57)$$

3. Distancia Adelantada A, $SSA(m_2)$: media móvil del promedio entre la Línea de Conversión y Base proyectada m_2 períodos hacia adelante en el tiempo. Esta proyección se realiza mediante el corrimiento de la serie m_2 posiciones hacia el futuro.

$$SSA(m_2)_i = \frac{TS(m_1)_i + KS(m_2)_i}{2} \quad 1 \leq i \leq n \quad (2.58)$$

4. Distancia Adelantada B, $SSB(m_2, m_3)$: de manera análoga a las series $TS(m_1)$ y $KS(m_2)$ pero con una ventana m_3 proyectada m_2 períodos hacia adelante en el tiempo. El valor de m_3 suele ser el doble de m_2 . La diferencia entre SSA y SSB se colorea para crear la nube.

$$SSB(m_3)_i = \frac{PH(m_3)_i - PL(m_3)_i}{2} \quad 1 \leq i \leq n \quad (2.59)$$

5. Línea de Retraso, CS : precio de cierre del período actual proyectado m_2 períodos hacia atrás en el tiempo.

$$CS_i = C_i \quad 1 \leq i \leq n \quad (2.60)$$

Cuando la línea SSA esté por encima de la línea SSB , se colorea la nube de verde, en caso contrario se colorea la nube de rojo como se puede observar en la figura 2.10 donde tenemos una configuración $NI(37,25,10)$ para el par $BTCUSD$ mediante datos $OHLC$ del sitio CMC e intervalos diarios

La tendencia general es al alza cuando el precio está por encima de la nube, a la baja cuando el precio está por debajo de la nube y sin tendencia o en transición cuando el precio está en la nube.

Cuando la SSA está aumentando y por encima de la SSB , esto ayuda a confirmar la tendencia alcista y el espacio entre las líneas suele ser de color verde. Cuando la SSA está cayendo y

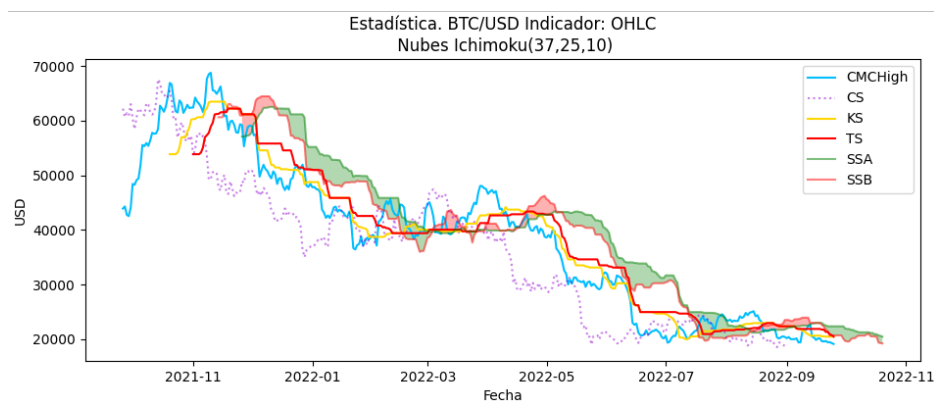


Figura 2.10: Nube Ichimoku.

por debajo de la SSB, esto ayuda a confirmar la tendencia bajista. El espacio entre las líneas suele ser de color rojo en este caso.

Los comerciantes a menudo usarán la Nube Ichimoku como un área de soporte y resistencia dependiendo de la ubicación relativa del precio. La nube proporciona niveles de soporte/resistencia que se pueden proyectar hacia el futuro. Esto distingue a la nube Ichimoku de muchos otros indicadores técnicos que solo brindan niveles de soporte y resistencia para la fecha y hora actual.

Una limitación de la Nube Ichimoku es que se basa en datos históricos. Si bien dos de estos puntos de datos se trazan en el futuro, no hay nada en la fórmula que sea inherentemente predictivo. Los promedios simplemente se trazan en el futuro [145].

La nube también puede volverse irrelevante durante largos períodos de tiempo, ya que el precio se mantiene muy por encima o por debajo de ella. En momentos como estos, la línea de conversión, la línea base y sus cruces se vuelven más importantes, ya que generalmente se mantienen más cerca del precio.

Un indicador no es mejor que otro; simplemente proporcionan información de diferentes maneras y puede que cada uno posea propósitos distintos, queda en manos del usuarios su ajuste, en cuanto a las ventanas de operación, demás opciones, interpretación y su uso en conjunto como mencionamos en el capítulo anterior haciendo referencia a la confluencia de técnicas para obtener un panorama completo lo que esta sucediendo. Adicional a estos estadístico, haremos uso de técnicas basadas en probabilidades, las cuales veremos a continuación.

2.2.4. Probabilidad.

Antes de entrar en materia de probabilidad, tengamos en cuenta las siguientes definiciones:

Experimento: proceso mediante el cual se obtiene alguna observación. Ejemplo: “Lanzar un dado”

Resultado: es un suceso particular proveniente de un experimento.

Evento (S): es un conjunto de uno o más resultados de un experimento.

- Simple: resultado básico que no puede ser dividido. Ejemplo: “Obtener un valor uno al lanzar un dado”
- Compuesto: conjunto de eventos simples. Ejemplo “Obtener un valor impar al lanzar un dado”

Eventos independientes: la ocurrencia de un evento no afecta la ocurrencia del otro. Ejemplo: Obtener uno al lanzar un dado, y dos al lanzar el otro dado.

Algunas operaciones sobre conjuntos: dado dos conjuntos A y B .

- Unión: el evento unión, denotado $A \cup B$, consiste en todos los eventos simples que están contenidos en A o B y en ambos. $A \cup B$ puede ser descrito como que ocurre por lo menos uno de los dos eventos A o B .
- Intersección: El evento intersección, denotado $A \cap B$, consiste de todos los eventos simples comunes de A y B . $A \cap B$ puede ser descrito como que ocurre ambos eventos A y B .
- Complemento: el evento complemento, denotado A' , consiste de todos los eventos simples que no están en A . A' puede ser descrito como que no ocurre A .

Eventos Disjuntos: dos eventos son disjuntos, mutuamente excluyentes, si no pueden ocurrir al mismo tiempo, en otras palabras, ellos no tienen eventos en común. A y B son disjuntos si $A \cap B = \emptyset$.

Espacio muestral (Ω): todos los eventos simples de un experimento. Ejemplo: En el caso de lanzar un dado el espacio muestral es $\Omega = \{1, 2, 3, 4, 5, 6\}$.

De ésta manera, si $S \in \Omega$, entonces $P(S)$: es una medida de la posibilidad relativa de que el evento S ocurra. La probabilidad puede definirse como una función $P : \Omega \rightarrow R$.

2.2.4.1. Tipos de probabilidad.

Clásica: cuando hay n resultados posibles, se asigna la probabilidad a un suceso antes de que éste ocurra.

Empírica: para asignar la probabilidad, el número de veces que ocurre un evento se divide por el número total de observaciones.

Subjetiva: se basa en cualquier información disponible.

2.2.4.2. Reglas de probabilidad.

Básicas: compuestas por tres axiomas.

1. La probabilidad para cualquier evento S estará entre 0 y 1.

$$0 \leq P(S) \leq 1 \quad (2.61)$$

2. La probabilidad para todo el conjunto Ω será igual a 1.

$$P(\Omega) = 1 \quad (2.62)$$

3. La probabilidad del conjunto vacío ($\emptyset$) representa en probabilidad el suceso imposible.

$$P(\emptyset) = 0 \quad (2.63)$$

Otras Reglas:

1. Dado un evento A :

$$P(A') = 1 - P(A) \quad (2.64)$$

2. Si $M = \{\bigcup_i^n A_i\}$, y A_i es disjuncto de A_j , $\forall_i = j$, entonces:

$$P(M) = \sum_i^n P(A_i) \quad (2.65)$$

En general:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (2.66)$$

3. Si A y B son independientes:

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.67)$$

2.2.4.3. Variables Aleatorias.

Es una función que asocia un número real con cada elemento del espacio muestral. Ejemplo: se lanzan tres monedas, con “c” representando que la moneda cayó en cara y “s” en sello; $\Omega = \{ccc, ccs, csc, scc, ssc, sss\}$. Si nos interesa el número de caras que se obtienen, a cada punto de Ω se le asignará 0, 1, 2 o 3. Estos valores son cantidades aleatorias determinadas por el resultado del experimento. Tipos de variables aleatorias:

Variable aleatoria discreta: Cuando está definida sobre un espacio muestral que tiene un conjunto finito o infinito numerable de posibilidades.

Ejemplo: El presentado anteriormente.

Variable aleatoria continua: cuando está definida sobre un espacio muestral que tiene un conjunto infinito no numerable de posibilidades.

Ejemplos: temperaturas, presiones, pesos, cotizaciones (variables correspondientes a datos medidos). En muchos casos, los valores posibles de la variable aleatoria continua son los mismos valores que contiene el espacio muestral continuo.

2.2.4.4. Distribuciones de Probabilidad.

Si es posible representar las probabilidades de una variable aleatoria X mediante alguna fórmula, entonces esa fórmula sería una función de los valores numéricos de x , por ejemplo $f(x)$.

$$f(x) = P(X = x) \quad (2.68)$$

A $(x, f(x))$ se le llama función de probabilidad o distribución de probabilidad de la variable aleatoria X .

Entre los tipos de distribuciones de probabilidad tenemos:

Distribuciones Discretas de Probabilidad: El conjunto de pares ordenados $((x, f(x)))$ es una función de probabilidad de la variable discreta X si, para cada resultado posible x , se cumple:

$$f(x) \geq 0 \quad (2.69)$$

$$\sum_x f(x) = 1 \quad (2.70)$$

$$(X = x) = f(x) \quad (2.71)$$

Distribución uniforme discreta: La variable aleatoria X toma los valores $x_1, x_2, \dots, x_n$ con idénticas probabilidades.

$$f(x; n) = \frac{1}{n} \quad n = 1, 2, \dots, n \quad (2.72)$$

La varianza en este caso mide el tamaño de la dispersión de una distribución.

La varianza de una distribución discreta es representada por σ^2 .

La desviación estándar es: $\sigma = \sqrt{\sigma^2}$

Distribución Binomial (Proceso de Bernoulli): Si se tiene un experimento que tiene como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$ (Experimento de Bernoulli), la distribución de probabilidad, el número de éxitos en n pruebas independientes, de la variable aleatoria X es una Distribución Binomial:

$$f(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad 1 \leq x \leq n \quad (2.73)$$

Distribuciones Continuas de Probabilidad:

- Como X es continua:

$$P(X = x) = 0 \quad (2.74)$$

- Está definido en los casos:

$$P(a \leq X \leq b) \quad \text{y} \quad P(X < b) \quad (2.75)$$

- $f(x)$ en el caso continuo no representa $P(X = x)$. En este caso se llama función de densidad de probabilidad. Esta función de densidad para la variable aleatoria continua X , definida en el conjunto de los números reales, si:

$$f(x) \geq 0 \quad \forall \quad x \in \mathbb{R} \quad (2.76)$$

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (2.77)$$

$$P(a < X < b) = \int_a^b f(x)dx \quad (2.78)$$

- Como X está definida en un espacio continuo, $f(x)$ podría tener discontinuidades; aunque en general se trabaja con funciones continuas.

Estudiaremos detalladamente algunas variables aleatorias continuas importantes y sus características. Como hemos señalado antes, en muchos problemas, se vuelve matemáticamente más simple considerar un espacio de rango “idealizado” para una variable aleatoria X , en el que todos los números reales posibles (en algún intervalo específico o conjunto de intervalos) pueden considerarse como posibles resultados y de ésta manera nos conducen a variables aleatorias continuas.

Ahora bien, detallaremos algunas de las distribuciones de probabilidad que nos serán relevantes mas adelante en el proyecto:

Distribución Uniforme: La función de densidad de la variable uniforme continua en X en el intervalo $[x_1, x_2]$ es:

$$f(x; x_1, x_2) = \begin{cases} 0 & x < x_1 \\ \frac{1}{x_2 - x_1} & x_1 \leq x \leq x_2 \\ 0 & x > x_2 \end{cases} \quad (2.79)$$

En la figura 2.11 podemos observar un ejemplo de distribución uniforme con $x_1 = 2$ y $x_2 = 4$.

$$\mu = \frac{x_1 + x_2}{2} \quad (2.80)$$

$$\sigma^2 = \frac{(x_2 - x_1)^2}{12} \quad (2.81)$$

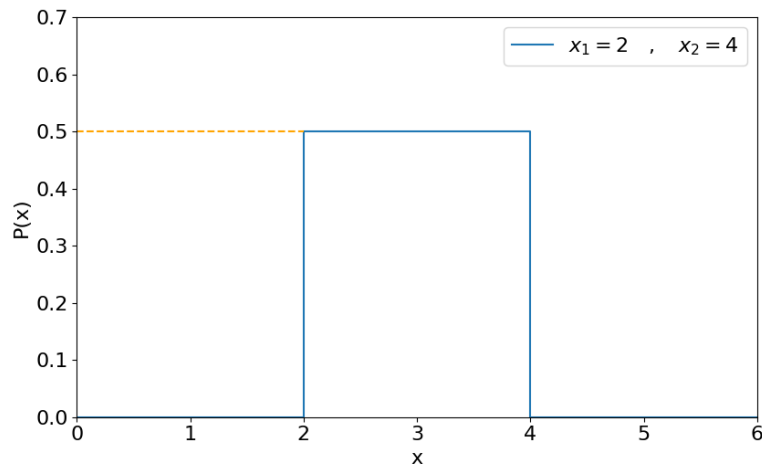


Figura 2.11: Distribución Uniforme

Distribución Normal: La distribución normal es una curva acampanada y presenta sólo un máximo en el centro. La media aritmética, la mediana y la moda de la distribución son iguales y están localizadas en el valor máximo (pico). Es simétrica con respecto a su media decrece en ambas direcciones a partir del valor central (asintótica).

La función de densidad de la variable aleatoria normal X , con media μ y varianza σ^2 , es:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$$

$$-\infty < x < \infty$$

(2.82)

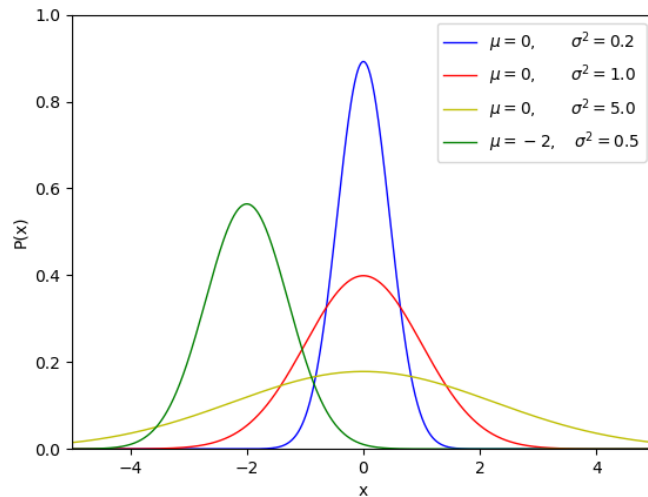


Figura 2.12: Distribución Normal

La distribución de una variable aleatoria normal con $\mu = 0$ y $\sigma^2 = 1$ se llama distribución normal estándar, en la figura 2.12 podemos observar varias distribuciones normales para distintos μ y σ .

Distribución Exponencial: Una variable aleatoria continua X que asume todos los valores no negativos se dice que tiene una distribución exponencial con parámetro $\lambda > 0$:

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (2.83)$$

$$\mu = \frac{1}{\lambda} \quad (2.84)$$

$$\sigma^2 = \frac{1}{\lambda^2} \quad (2.85)$$

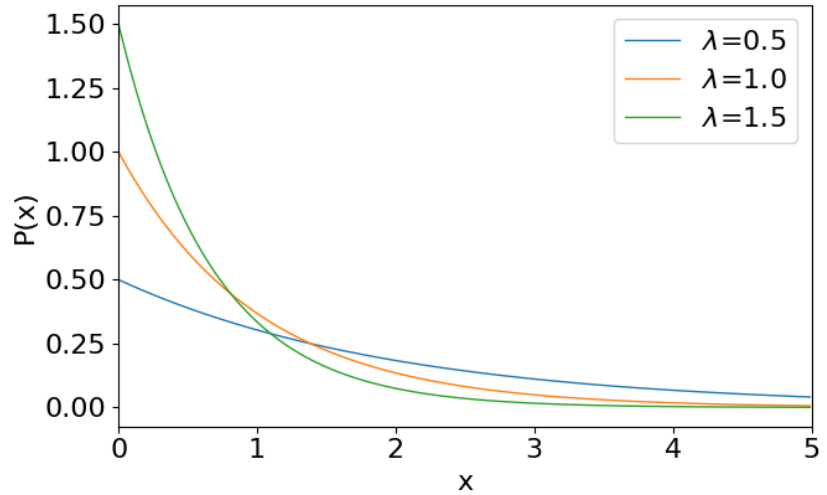


Figura 2.13: Distribución Exponencial

La distribución exponencial juega un papel importante en la descripción de una gran clase de fenómenos, particularmente en el área de la teoría de la confiabilidad. En la figura 2.13 se puede observar una distribución exponencial para $\lambda = 0.5$, $\lambda = 1.0$ y $\lambda = 1.5$.

Distribución Triangular: En probabilidad y estadística, la distribución triangular es una familia de distribuciones continuas de probabilidad que tiene un valor mínimo x_1 , un valor máximo x_2 y una moda x_3 [146]. La distribución triangular es una distribución continua definida en el rango x en $[x_1, x_2]$ con función de densidad de probabilidad:

$$f(x; x_1, x_2, x_3) = \begin{cases} \frac{2(x - x_1)}{(x_2 - x_1)(x_3 - x_1)} & a \leq x \leq c \\ \frac{2(x_2 - x)}{(x_2 - x_1)(x_2 - x_3)} & c < x \leq b \end{cases} \quad (2.86)$$

Donde:

$$x_1: x_1 \in (-\infty, \infty)$$

$$x_2: x_2 > x_1 \quad x_2 \in (x_1, \infty)$$

$$x_3: x_1 \leq x_3 \leq x_2$$

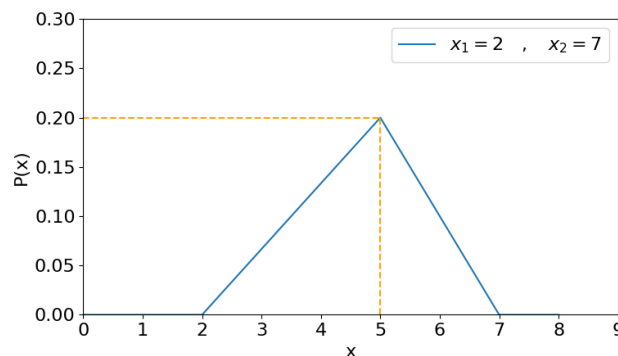


Figura 2.14: Distribución Triangular

$$\mu = \frac{x_1 + x_2 + x_3}{3} \quad (2.87)$$

$$\sigma^2 = \frac{x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3}{18} \quad (2.88)$$

En la figura 2.14 se puede observar una distribución triangular con $x_1 = 2$ y $x_2 = 7$.

Distribución Beta: la distribución beta es una familia de distribuciones de probabilidad continua definida en el intervalo $[0, 1]$ parametrizada por dos parámetros de forma positiva, denotados por alfa (α) y beta (β), que aparecen como exponentes de la variable aleatoria y controlan la forma de la distribución. La generalización a múltiples variables se llama distribución de Dirichlet [147].

La distribución beta se suele aplicar para modelar el comportamiento de variables aleatorias limitadas a intervalos de longitud finita en una amplia variedad de disciplinas.

La función densidad de X ésta dada por:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathbf{B}(\alpha, \beta)} \quad (2.89)$$

Donde:

x : $0 < x < 1$

α : $\alpha > 0$ $\alpha \in \mathbb{R}$, parámetro de forma.

β : $\beta > 0$ $\beta \in \mathbb{R}$, parámetro de forma.

$\mathbf{B}(\alpha, \beta)$:

$$\mathbf{B} = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx \quad (2.90)$$

$$\mathbf{B}(\alpha, \beta) = \mathbf{B}(\beta, \alpha)$$

$$\mathbf{B}(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad \text{donde } \Gamma \text{ es la función Gamma.}$$

$$\mu = \frac{\alpha}{\alpha + \beta} \quad (2.91)$$

$$\sigma = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (2.92)$$

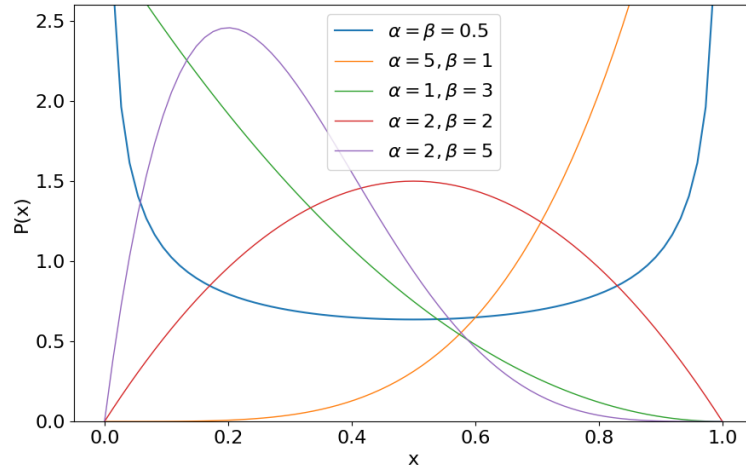


Figura 2.15: Distribución Beta.

En la figura 2.15 se puede observar la distribución Beta para distintos valores de α y β .

Distribución Semicircular: La distribución de semicírculo de Wigner, llamada así por el físico Eugene Wigner, es la distribución de probabilidad en $[-R, R]$ cuya función de densidad de probabilidad f es un semicírculo a escala (es decir, una semi-elipse) centrado en $(0, 0)$ [148]; su función densidad de probabilidad ésta dada por:

$$f(x; R) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2} \quad (2.93)$$

$$-R \leq x \leq R$$

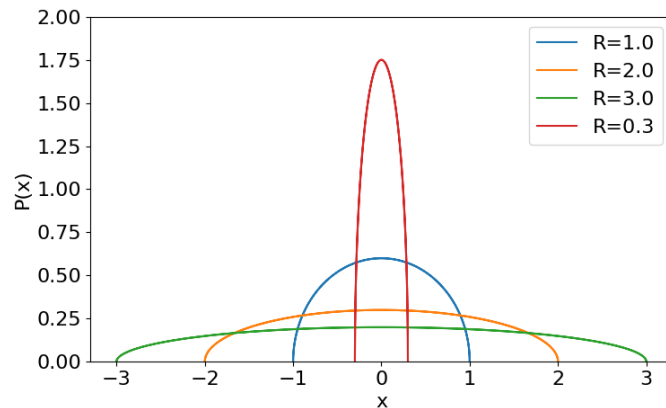


Figura 2.16: Distribución Semicircular

Aunque principalmente se le llama distribución semicircular a las presentes en la figura 2.16 también es usual llamar distribución semicircular a la generalización de radio igual a 1 ($R = 1$) como el caso de la figura 2.17, de ésta manera tendríamos:

$$f(x) = \frac{2}{\pi} \sqrt{1-x^2} \quad (2.94)$$

$$-1 \leq x \leq 1$$

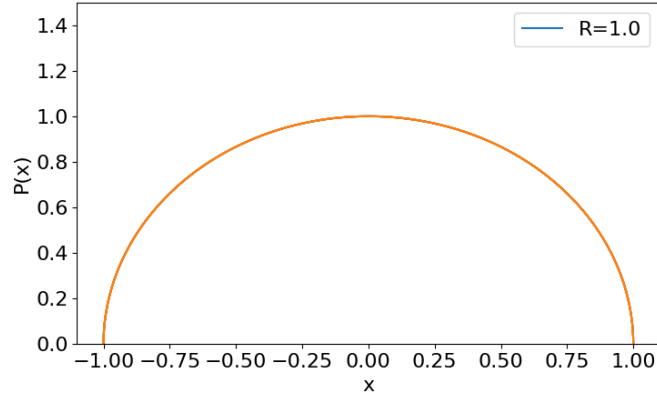


Figura 2.17: Distribución Semicircular de radio 1.

Distribución Gama: Una distribución gamma es un tipo general de distribución estadística que está relacionada con la distribución beta y surge naturalmente en procesos para los cuales los tiempos de espera entre eventos distribuidos son relevantes. Las distribuciones gamma tienen dos parámetros libres, etiquetados α y θ .

Hay dos parametrizaciones diferentes de uso común:

- Con un parámetro de forma k y un parámetro de escala θ .
- Con un parámetro de forma $\alpha = k$ y un parámetro de escala inversa $\beta = 1/\theta$, llamado parámetro de tasa.

En cada una de éstas formas, ambos parámetros son números reales positivos, en la figura 2.18 podemos observar el resultado para distintos valores de k y θ . De ésta manera la función densidad de probabilidad ésta dada por:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad (2.95)$$

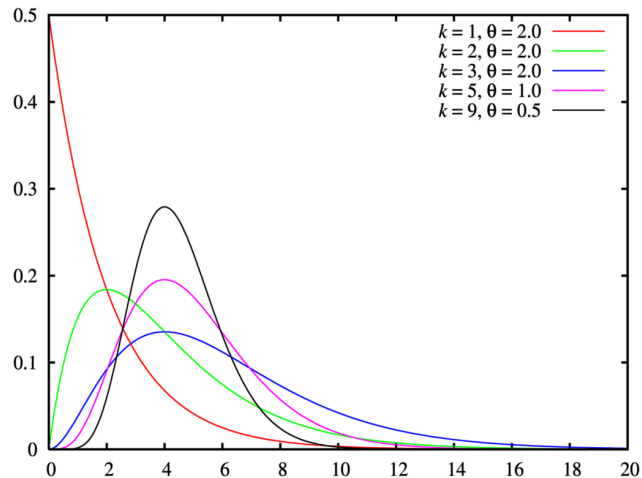


Figura 2.18: Distribución Gamma

La parametrización con k y θ parece ser más común en la econometría y en otros campos aplicados, donde, por ejemplo, la distribución gamma se usa con frecuencia para modelar los

tiempos de espera.

La parametrización con α y β es más común en las estadísticas bayesianas, donde la distribución gamma se usa como una distribución previa conjugada para varios tipos de parámetros de escala inversa (tasa), como el λ de una distribución exponencial o una distribución de Poisson, o para el caso, el β de la distribución gamma en sí. La distribución de gamma inversa estrechamente relacionada se utiliza como un conjugado anterior para los parámetros de escala, como la varianza de una distribución normal.

Estas distribuciones de probabilidad nos serán suficientes para el fin que tenemos en mente, ya que este grupo abarca una gran variedad de formas en cuanto a la distribución de la probabilidad para un intervalo determinado.

La siguiente herramienta de la que haremos uso es de la inferencia, para así en base a datos de eventos que ya sucedieron hacernos una idea de lo que posiblemente ocurrirá en el futuro.

2.2.5. Inferencia.

La estadística usualmente trata de responder dos grupos de preguntas, las del primer grupo constituyen ese grupo de preguntas que buscan obtener una descripción de un grupo de datos los cuales están siendo estudiados, estas respuestas se reducen a ciertos coeficientes como la media y la desviación estándar o a alguna serie como las que vimos en la sección anterior medias móviles entre otros. Usualmente estos estudios llegan hasta aquí, no buscan generar teorías o respuestas a otros posibles escenarios.

En cambio las preguntas del segundo grupo son un poco más amplias, éstas tratan de responder y buscar soluciones a problemas mas generales, tratando así de deducir o sacar una conclusión (inferir), partiendo de respuestas particulares obtenidas de un solo grupo de datos, conclusiones generales y validas para un conjunto mas amplio, del cual el grupo estudiado es solo una fracción. Es a este tipo de prácticas la que llamamos inferencia estadística *causo cita*.

La inferencia estadística comprende un grupo de métodos por lo que se realizan inferencias o generalizaciones acerca de una población, basándonos en una o varias muestras.

En cuanto a los métodos de inferencia tenemos:

Método Clásico: Se basa en la información que se obtiene de una muestra aleatoria.

Método Bayesiano: Además de los datos obtenidos de la muestra, usa conocimiento subjetivo previo sobre la distribución de probabilidad.

Entre las áreas más comunes de la inferencia tenemos:

Estimación: Sobre la base de datos tomados en una muestra, se estima el valor de algún parámetro de la población. (Ejemplo: Encuesta)

Pruebas de Hipótesis: Se plantea una hipótesis sobre algún parámetro de la población y se hace una prueba tomando una o algunas muestras y se acepta o rechaza la hipótesis. (Ejemplo: medición de dos productos)

2.2.5.1. Estimación Puntual.

Una estimación puntual de algún parámetro de la población θ es un solo valor $\hat{\theta}$ de un estadístico $\hat{\Theta}$.

Ejemplo: el valor $\bar{x}$ del estadístico $\bar{X}$, que se calcula con una muestra de tamaño n , es una estimación puntual del parámetro poblacional μ .

Observación: en general, el estimador permite realizar la estimación del parámetro poblacional con un error diferente de cero (0). Se pueden usar para estimar un parámetro poblacional diferentes valores calculados de la muestra. (Ejemplo: Mediana y Media para estimar μ)

2.2.5.2. Estadístico Muestral.

Hemos visto que la inferencia trata de conocer alguna característica de la población, mediante el estudio de una muestra. La información contenida en la muestra puede ser (usualmente conveniente y suficiente) condensada en una magnitud, en una sola variable aleatoria calculada como una función de variables aleatorias que conforman la muestra.

De ésta manera dada una muestra aleatoria: $x_1, x_2 \dots x_n$, de tamaño n de una población con función de densidad $f(x)$, un estadístico es una función $g(x_1, x_2 \dots x_n)$ de las variables aleatorias de la muestra que no contiene ningún parámetro desconocido. Al ser una función de variables aleatorias, un estadístico es también una variable aleatoria.

Siendo esta una aclaratoria importante ya que a lo largo de nuestro proyecto posiblemente tendremos acceso a datos de cotizaciones de pares criptomoneda/dinero fíat los cuales ya han sido manipulados en primera instancia.

Ya que los gráficos o las tablas de datos si bien pudieran proporcionar registros de cada una de las transacciones efectuadas (población total) y nosotros acceder a estas para realizar nuestro análisis, este no es el caso, ya que los datos que se suelen publicar son datos promediados para intervalos de tiempo específico, por ejemplo, gráficos de: un minuto, cinco minutos, dos horas, un

día, tres días, una semana, un mes, entre otros. Con la excepción de los datos OHLC, estos si son datos específicos de transacciones, el Open y el Close a un tiempo en específico (aunque de tomar una fuentes que rastree varios intercambios seguramente obtendremos un promedio) y el High y Low son datos puntuales de la cotización mas alta y baja durante el intervalo, siendo estos dos últimos los que posiblemente tengamos mayor posibilidad de que sean datos originales. De ésta manera nuestros datos podrían ser promedios de promedios o datos originales, en cuanto a los promedios, este será de todas las cotizaciones que se hayan efectuado durante ese intervalo, y nuestro estadístico sería ya, un promedio de promedios. Además, en el caso de hacer uso de los datos del sitio BitcoinAverage, este posee datos de un promedio global, basado ya en promedio de distintos exchange los cuales ya están promediados por intervalos.

Y si bien pudiese parecer que esto nos aleja de la realidad, sólo sería el caso en el que quisiésemos aplicar nuestros resultados a un exchange en particular y operar con márgenes muy finos, si nuestro objetivo es estudiar los mercados con unos márgenes apropiados y con una visión global, esta herramienta nos sera muy útil como herramienta de apoyo para generar, validar o anular conclusiones.

De esta forma, sea $x_1, x_2 \dots x_n$ una muestra aleatoria de tamaño n de una población con función de densidad $f(x)$ con media μ y varianza finita σ^2 , la media muestral representada por $(\bar{x})$, es la media aritmética de los elementos de la muestra.

Al ser $\bar{x}$ una variable aleatoria, tiene un valor esperado y una varianza. El valor esperado de la media muestral es la media de la población, en efecto:

$$\begin{aligned} E(\bar{x}) &= E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{n\mu}{n} \\ E(\bar{x}) &= \mu \end{aligned} \quad (2.96)$$

La varianza de la media muestral $\sigma_{\bar{x}}^2$ es igual a la varianza de la población dividida por el tamaño de la muestra:

$$\begin{aligned} Var(\bar{x}) &= \sigma_{\bar{x}}^2 = Var\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = Var\left(\frac{x_1}{n} + \frac{x_2}{n} + \dots + \frac{x_n}{n}\right) \\ Var(\bar{x}) &= \frac{\sigma^2}{n^2} + \frac{\sigma^2}{n^2} + \dots + \frac{\sigma^2}{n^2} \\ Var(\bar{x}) &= \frac{\sigma^2}{n} \end{aligned} \quad (2.97)$$

La importancia de estos resultados es obvia. Por un lado tenemos que el valor esperado de la media muestral es la media de la población; es decir, que si tomamos todas las muestras posibles de tamaño n de una población con media μ , la media de las medias muestrales es igual a la media de la población. Por otro lado sabemos que la varianza de una variable indica la dispersión de los valores de la variable respecto a su media: en este caso la varianza de la media muestral nos dice cuanto puede esperarse que las medias muestrales varíen de una muestra a otra, y el grado de dispersión de las medias muestrales con respecto de la población. Vemos que esta dispersión es tanto menor cuanto mayor sea el numero de la muestra; en otras palabras la distribución probabilística de la media muestral se concentra alrededor de la media de la población cuanto mayor sea el tamaño de la muestra; es decir, cuando se aumenta el tamaño de la muestra, la media muestral tiende a acercarse a la media de la población.

Ahora bien, sea $x_1, x_2 \dots x_n$ una muestra aleatoria de tamaño n de una población con función de densidad $f(x)$. La varianza muestral representada por (S_x^2) , es el promedio del cuadrado de las desviaciones de los elementos de la muestra respecto a la media muestral, es decir:

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.98)$$

No debe confundirse la varianza muestral (S_x^2) con la varianza de la media muestral $(\sigma_{\bar{x}}^2)$. La primera se refiere a una muestra y mide la dispersión de los datos de esa muestra con respecto de su media. La segunda mide la dispersión de las medias de distintas muestras con respecto a la media de la población; por otro lado (S_x^2) es un estadístico, mientras que $(\sigma_{\bar{x}}^2)$ no lo es pues contiene el parámetro desconocido (σ^2)

2.2.5.3. Estimación por puntos.

Uno de los problemas más importantes que trata de resolver la inferencia estadística es el de la estimación de los parámetros de una población por medio de los datos contenidos en una muestra. Los parámetros de una población tales como la media, la varianza, la proporción, representan características importantes de la población que al investigador le interesa conocer.

Estos parámetros son generalmente desconocidos y es poco práctico o imposible analizar todos y cada uno de los elementos de la población, ni manera de calcular los parámetros en forma exacta. Tenemos que contentarnos con hallar alguna magnitud a través de la información incompleta de una muestra, que nos sirva para aproximar el valor del parámetro desconocido. En esto consiste el problema de estimación en estadística. Esta estimación puede ser de dos clases:

1. Puede usarse una magnitud simple para estimar un parámetro; por ejemplo, usar la media muestral como estimador de la media de la población; en este caso estamos realizando una estimación por puntos o estimación puntual.
2. Se puede calcular un intervalo entre cuyos límites se espera que esté el parámetro con una razonable confianza; en este caso estamos realizando una estimación por intervalos.

Aquí dos ejemplos:

- Si estamos interesados en conocer el salario medio de los obreros de una industria (media de la población: parámetro desconocido), podríamos estimarlo por el salario medio obtenido en una muestra (media muestral: estimador).
- Si queremos estimar la proporción de estudiantes de la universidad que fuman (proporción de la población: parámetro desconocido), se puede hacer por medio de la proporción de fumadores en una muestra aleatoria (proporción - muestral: estimador).

En general, los parámetros desconocidos de la población se estimarán a través de algún estadístico muestral, en consecuencia se puede definir un estimador puntual de la forma siguiente:

Un estimador puntual ($\hat{\theta}$) de un parámetro desconocido (θ) de una población es una función de las variables de una muestra aleatoria $\hat{\theta} = g(x_1 \dots x_n)$, que no contiene ningún parámetro desconocido, y que sirve como una aproximación del parámetro; o sea, es un estadístico que sirve como aproximación del parámetro.

Es preciso hacer una distinción entre “estimador” de un parámetro y “valor estimado” de un parámetro. Un estimador de un parámetro es una función de los valores de una muestra, y en consecuencia es una variable aleatoria que toma distintos valores dependiendo de los valores de las variables aleatorias de la muestra concreta, y calculado el valor del estadístico muestral que usamos como estimador del parámetro, ese valor servirá como “valor estimado” del parámetro. Es decir, el “estimador” del parámetro es el estadístico muestral, el valor concreto del “estimador” obtenido en una muestra, es el “valor estimado” del parámetro.

2.2.5.4. El problema de la elección.

Siendo el estimador del parámetro una función de las variables de una muestra, es evidente que podemos escoger distintos estimadores para un mismo parámetro, puesto que podemos elegir arbitrariamente la función que vamos utilizar. Para cada parámetro desconocido podemos elegir

muchos, prácticamente infinitos, estimadores. Es evidente que se debería tener algún tipo de medida para juzgar si el estimador elegido, es un buen estimador: si uno es mejor que otro es decir, alguna medida que nos indique la eficacia de nuestra elección. En nuestro caso la elección del estimador queda a juicio del usuario al hacer uso de nuestra herramienta.

2.2.5.5. Estimación por intervalos.

La estimación por intervalos busca complementar la estimación por puntos, de forma de tener alguna idea acerca de la precisión de los estimadores puntuales usados.

Como se vio anteriormente, un buen estimador puntual es aquel cuyo valor está cercano al valor del parámetro que se desea estimar. No se debe olvidar que los estimadores son estadísticos muestrales, y en consecuencia son variables aleatorias.

Al tomar una muestra de una población, el valor del estimador obtenido a través de esa muestra, o sea, el valor estimado del parámetro, es un número; pero la muestra concreta que se ha obtenido, es una de las muchas posibles muestras que se podrían haber obtenido. Si se hubiera obtenido otra muestra, el valor estimado hubiere sido posiblemente diferente; en consecuencia, no podemos ser tan optimistas como para pensar que el valor estimado a partir de una determinada muestra, es exactamente igual al valor real del parámetro de la población.

Lo anterior es razón suficiente para considerar que se necesita una cierta medida del posible error, que se comete al usar un estimador puntual del parámetro.

Si queremos estimar, por ejemplo, la media de una población mediante una muestra de tamaño n , podemos obtener muchos diferentes valores de la media muestral, la mayoría de los cuales no son exactamente iguales a la media de la población. Es deseable entonces calcular algún intervalo alrededor del valor del estimador, acompañado de alguna medida que nos diga la confianza que se puede tener de que ese intervalo contenga el verdadero valor del parámetro.

Dada una muestra aleatoria $x_1, x_2, \dots, x_n$ de tamaño n de una población con función de densidad $f(x)$, un intervalo de confianza del $100(1 - \alpha)\%$ para el parámetro desconocido θ , es un intervalo determinado por dos números $\hat{\theta} - \delta_1$, $\hat{\theta} + \delta_2$, calculado con base en los datos de la muestra, tales que:

$$P[\hat{\theta} - \delta_1 \leq \theta \leq \hat{\theta} + \delta_2] = 1 - \alpha \quad (2.99)$$

Se puede observar que los límites del intervalo son variables aleatorias, pues son funciones del estimador y varían de una muestra a otra, de ésta manera cambiará el intervalo según la muestra que usemos, pero el $100(1 - \alpha)\%$ en promedio de los intervalos calculados contendrán el verdadero

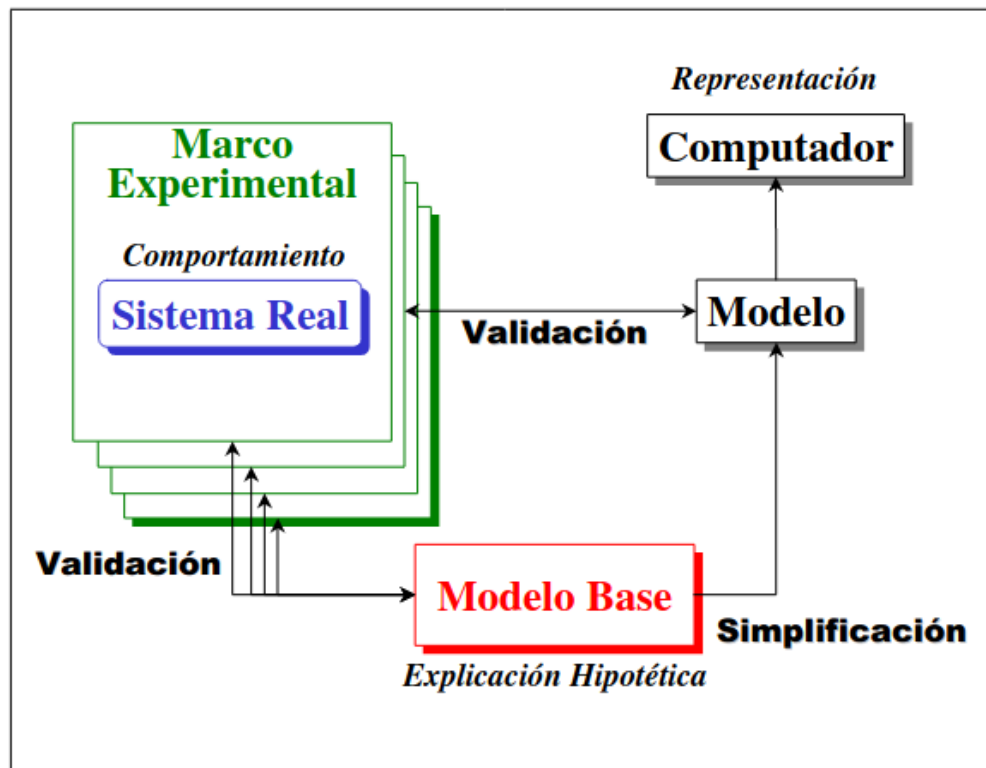


Figura 2.19: Elementos básicos del modelado y simulación de sistemas.

valor del parámetro.

Un tema clave tanto en la inferencia como en la simulación como se verá en la sección siguiente es el muestreo y sus tipos los cuales definimos en la 2.2.1, dada una población o una muestra, la forma como escogemos una muestra o efectuamos un remuestreo puede influir de manera notoria en los resultados. Veamos ahora todo lo pertinente a simulación.

2.2.6. Modelado y Simulación.

La frase “modelado y simulación” designa el conjunto de actividades asociadas con la construcción de modelos de sistemas del mundo real y su simulación haciendo uso de computadores.

La figura 2.19 muestra cinco elementos básicos que están asociados a aspectos conceptualmente diferentes de las actividades de modelado y simulación de sistemas [149].

Cuando se discute el modelado y la simulación se definen estos elementos y se establecen las relaciones entre ellos [150]:

Sistema Real: se refiere a la fuente observable de datos, la cual puede ser natural, artificial o una combinación de ambas. En esta etapa del modelado y simulación, la característica importante es la identificación de una porción de realidad y su distinción con el resto del

sistema, permitiendo realizar observaciones y mediciones sobre dicha porción.

El comportamiento del sistema real constituye todo aquello que podemos conocer directamente del mundo real. Este comportamiento es representado haciendo uso de variables descriptivas, nombres genéricos para atributos del sistema, las cuales son clasificadas como observables y no observables. Las variables observables son aquellas que corresponden directamente a mediciones, mientras que las variables no observables son aquellas que no pueden ser representadas directamente por una medida o juicio de valor.

Las variables observables se pueden clasificar como variables de entrada y de salida. Las variables de entrada permiten describir perturbaciones o influencias que afectan el sistema desde el exterior. Las variables de salida permiten representar los valores resultantes del proceso de simulación. En otras palabras, las variables de entrada son generalmente consideradas las “causas” y las variables de salida los “efectos”.

El avance del tiempo representa uno de los conceptos básicos de simulación. Al agrupar las variables se obtiene el estado del sistema en el transcurso del tiempo. Cada intervalo de tiempo determina los segmentos o trayectorias del modelo y al conjunto de todos los posibles pares de segmentos entrada-salida obtenidos experimentalmente se les denomina el comportamiento del sistema real.

La validación de un modelo es el proceso de evaluar la validez de dicho modelo. Esta validez es relativa a un marco experimental y a los criterios que permitan calibrar las trayectorias. El marco experimental es la definición de un conjunto limitado de circunstancias bajo las cuales se observa o se experimenta con el sistema real.

Modelo base: es un modelo que resume el comportamiento del sistema real. Este modelo es válido en todos los marcos experimentales permitidos pues proporciona una explicación hipotética completa del comportamiento del sistema real.

Modelo: en casos reales, la complejidad del modelo base no permite considerarlo para simulación. En estos casos la simplificación tiene como finalidad la definición de un modelo (agregado) a partir del modelo base. Este modelo debe ser válido con respecto a un marco experimental de interés.

Computador: es el dispositivo de cálculo que ayuda a generar las trayectorias del modelo. Aunque en algunos casos estas trayectorias o sus propiedades pueden ser manipuladas en forma analítica, sin la ayuda del computador, generalmente se desea realizar la representación y la

manipulación de estas trayectorias paso a paso, es decir, de un instante de tiempo al siguiente. Es a este proceso a lo que se le ha llamado simulación.

Las instrucciones para llevar a cabo el proceso de simulación paso a paso están contenidas en el modelo y el computador debe ejecutar correctamente estas instrucciones. El término modelo se refiere al código del programa escrito en el lenguaje utilizado para describir el sistema al simulador (en general, diferente al lenguaje en que se codifica el simulador).

Los motivos o fines que se persiguen con el modelado y la simulación de sistemas pueden verse según dos puntos de vistas diferentes: desde el punto de vista científico, se intenta entender como funciona un sistema, mientras que desde el punto de vista aplicativo lo interesante es notar o explicar que sucede en un sistema cuando se somete a ciertos cambios. Por ende, el modelado y simulación de sistemas constituyen una herramienta básica para la solución de problemas. Como herramienta la simulación ofrece múltiples ventajas:

1. Se apoya en una teoría matemática formal. Esto implica, entre muchas otras cosas, que es posible realizar simplificaciones sistemáticas de las representaciones de un sistema, preservando la estructura, a través de isomorfismos.
2. Es posible representar varios escenarios para el mismo problema, lo cual permite estudiar varias soluciones en forma simultánea.
3. El experimento permanece, no se destruye. Además, la historia generada puede ser extraída completamente.
4. Es posible repetir el experimento, la simulación, tantas veces como sean necesarias con relativa economía.

2.2.6.1. Formalismos para especificación de sistemas.

La teoría de sistemas [149] proporciona una formalización matemática que permite representar en forma rigurosa los sistemas dinámicos. Esta formalización distingue entre la estructura de un sistema, su constitución propia, y su comportamiento o manifestaciones.

Los sistemas dinámicos se representan mediante una estructura de la forma:

$$S = (X, Y, \Omega, Q, \Delta, \Lambda, T) \quad (2.100)$$

Donde:

X : Conjunto de Entradas.

Y : Conjunto de Salidas.

Ω : Conjunto de Segmentos de entrada posibles.

Q : Conjunto de Estados.

Δ : Función de Transición de estados. $Q \times \Omega \rightarrow Q$

Λ : Función de Salida. $Q \times \Omega \rightarrow Y$

T : Base Temporal. Posibles valores para el tiempo.

X , Y y Ω definen las entradas y salidas del sistema, mientras que Q y Δ representan la estructura interna y el comportamiento.

La función de transición de estados Δ debe garantizar que siempre se puede determinar el próximo estado del sistema a partir del estado actual, es decir debe ser un semigrupo.

Por otro lado, visto como una caja negra (figura 2.20 [149]), el comportamiento de un sistema es la relación existente entre las entradas y las salidas. Este comportamiento se representa como una tupla de la forma (entrada, salida) obtenida a partir del sistema real o del modelo. Así mismo, la estructura interna de un sistema incluye el estado del sistema, el mecanismo de transición de estados y el “mapping” hacia los estados de salida.

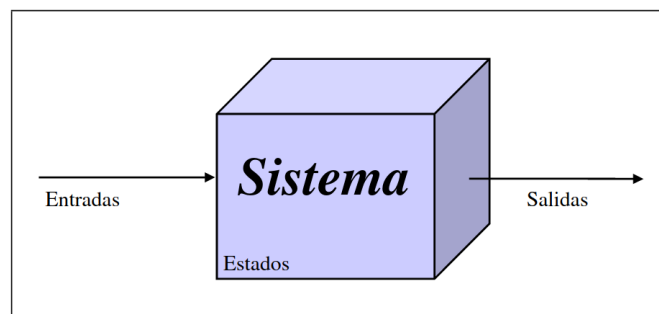


Figura 2.20: Concepto básico de Sistema.

Conocida la estructura de un sistema es posible deducir, analizar o simular su comportamiento. Usualmente en la dirección opuesta (inferir la estructura a partir del comportamiento) no es posible realizar una deducción unívoca. Descifrar una representación válida a partir de un comportamiento observado es uno de los objetivos del modelado y la simulación de sistemas. En el momento en que se definen restricciones sobre las componentes que describen un sistema en el tiempo es posible derivar distintos formalismos del sistema.

Los formalismos más utilizados para la representación de sistemas se describen brevemente a continuación.

Ecuaciones Diferenciales (DESS, Differential Equation Systems Specifications). Se emplean ecuaciones diferenciales ordinarias de primer orden para especificar la tasa de cambio de las variables de estado, las cuales generalmente se resuelven utilizando métodos numéricos. La representación de ecuaciones diferenciales se muestra en la figura 2.21 [149].

$$\text{DESS} = (X, Y, \Omega, Q, f, \lambda, \mathbb{R}) \quad (2.101)$$

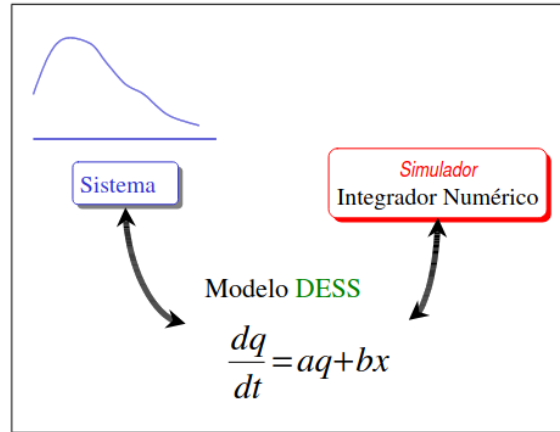


Figura 2.21: Especificaciones del modelo basado en Ecuaciones Diferenciales.

Donde:

X : Espacio vectorial ($\mathbb{R}^m$) de las variables de entrada.

Y : Espacio vectorial ($\mathbb{R}^p$) de las variables de salida.

Ω : Conjunto de todos los segmentos de entrada posibles. Estos deben ser contiguos a trozos.

Q : Espacio vectorial ($\mathbb{R}^n$) de estados accesibles del sistema.

$f : Q \times X \rightarrow Y$.

$\lambda : Q \rightarrow Y \quad \vee \quad Q \times X \rightarrow Y$. Permite obtener la salida en el tiempo t .

$\mathbb{R}$: La base del tiempo son los números reales.

Donde las trayectorias de entrada y salida son trayectorias continuas.

La función de tasa de cambio f debe tener solución única para asegurar que las trayectorias de estado sean únicas.

Ecuaciones en Diferencia (DTSS: Discrete Time Systems Specifications). Se emplea un conjunto de ecuaciones en diferencia para especificar el próximo estado en el tiempo, el cual avanza en forma constante. La representación de este formalismo mostrado en la figura 2.22 [149]:

$$DTSS = (X, Y, \Omega, Q, \delta, \lambda, c) \quad (2.102)$$

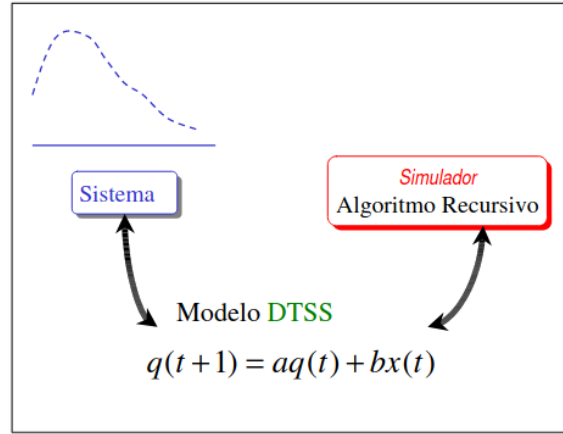


Figura 2.22: Especificaciones del modelo basado en Ecuaciones en Diferencia.

Donde:

X : Conjunto de Entradas.

Y : Conjunto de Salidas.

Ω : Conjunto de las secuencias admisibles sobre X, Y .

Q : Conjunto de Estados.

$\delta : Q \times X \rightarrow Q$. Función de transición de estados.

$\lambda : Q \rightarrow Y \vee Q \times X \rightarrow Y$ Función de Salida.

c : Base temporal. Isomorfa a los enteros.

Eventos Discretos (DEVS: Discrete Event Systems Specifications). El sistema es activado por eventos y obedece a transiciones producidas interna y externamente. La representación de este formalismo mostrado en la figura 2.23.

$$DEVS = (X, Y, S, \delta_{int}, \delta_{ext}, \lambda, ta) \quad (2.103)$$

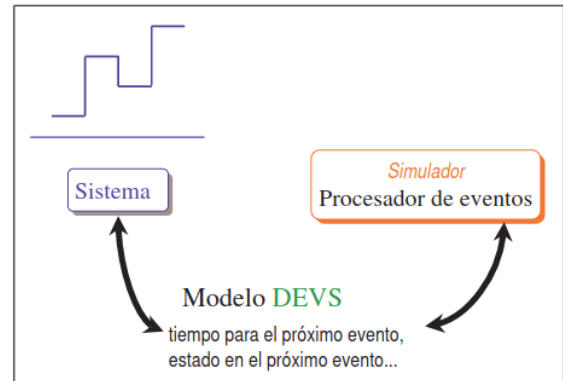


Figura 2.23: Especificaciones del modelo basado en Eventos Discretos.

Donde:

$\mathbf{X}$: Conjunto de eventos externos de entrada.

$\mathbf{Y}$: Conjunto de eventos externos de salida.

$\mathbf{S}$: Conjunto de estados.

$\delta_{int} : S \rightarrow S$. Función de transición interna.

$\delta_{ext} : Q \times X \rightarrow S$. Función de transición externa donde:

$$Q : \{(s, e) | s \in S, 0 \leq e \leq ta(s)\}$$

e : tiempo transcurrido desde la ultima transición

$\lambda : S \rightarrow Y$. Función de salida.

$ta : S \rightarrow \mathbb{R}^+$. Función de avance de tiempo.

A continuación veremos uno de los formalismos que más se asemeja a nuestro sistema en específico.

2.2.6.2. Formalismo DEVS.

Esta sección introduce el formalismo DEVS para sistemas de eventos discretos que proporcionan un enfoque modular, jerárquico a la construcción de modelos de simulación de eventos discretos [151]. Con este fin se incluyen en el DEVS los conceptos de teoría de sistemas y modelado introducidos al inicio de la sección.

El formalismo DEVS no sólo ha demostrado que es importante para modelo de eventos discretos sino que también proporciona una base computacional para implementar comportamiento de sistemas que pueden ser expresados en diferentes formalismos (DESS, DTSS, ...) y que abarcan los enfoques de simulación tanto secuencial como paralela o distribuida [149]. Por otro lado, el simulador DEVS puede usarse para implementar otras estrategias, por ejemplo, el esquema de tiempo discreto se implementa a través de eventos uniformemente distribuidos en el tiempo, tal y como mostraremos a continuación.

Estructura DEVS

DEVS separa el modelado de la simulación y propicia la reutilización y acoplamiento de componentes. Los modelos DEVS capturan la dinámica del comportamiento de los sistemas, es decir, representan los estados del sistema y el cambio de estos estados a través del tiempo. Estos cambios pueden ser causados por la evolución misma del sistema o en respuesta a eventos externos. Los modelos DEVS se clasifican en:

Atómicos: Representación modular de sistemas. Son modelos de bajo nivel basados en la estructura del sistema. Sus elementos son:

- Entradas.
- Salidas.
- Variables de estado.
- Funciones de transición.
- Función de salida.
- Función de avance de tiempo.

Acoplados: Representación jerárquica de sistemas. Un modelo acoplado esta compuesto por uno o más modelos atómicos o acoplados. Sus elementos son:

- Componentes,
- Reglas internas de acoplamiento,
- Reglas externas de acoplamiento: Entradas y Salidas.

A cada componente de un modelo DEVS acoplado se le puede representar a través de su equivalente escrito como una composición de modelos DEVS atómicos.

Especificación DEVS

La jerarquía de niveles con los que un sistema puede describirse o especificarse se resume en la siguiente descripción. A medida que se incrementa el nivel, se aumenta la especificidad estructural (es decir, nos movemos del comportamiento a la estructura). Cabe mencionar que se incorpora nuevo conocimiento en cada nivel en la jerarquía como se observa en la tabla de la figura 2.24.

Nivel 0: Describe las entradas y salidas de un sistema en el tiempo.

Nivel 1: Se refiere a registros completamente observacionales de la conducta de un sistema a través de un conjunto de segmentos entrada/salida.

Nivel 2: Se estudia el conjunto de funciones E/S que relacionan las partes de un sistema E/S. Cada función es asociada a un estado inicial del sistema.

Nivel 3: El sistema es descrito por conjuntos abstractos y funciones. Se introducen el espacio de estados y las funciones de transición y salida. La función de la transición describe las transiciones de un estado al estado siguiente causadas por los segmentos de entrada; la función de salida describe la transición de un estado a una salida observable.

Nivel 4: El sistema se describe a través del generador de segmentos de entrada y las transiciones del estado ocasionadas por estos segmentos.

Nivel 5: Se presentan los conjuntos abstractos y las funciones de los niveles más bajos como provenientes del producto directo de conjuntos primitivos y funciones. Esta es la forma en que se describen informalmente modelos.

Nivel 6: Un sistema se especifica como un conjunto de componentes. Cada uno de estos componentes con su propio conjunto de estados y su función de transición. Los componentes se acoplan en forma no modular y la influencia de cada uno recae sobre los otros a través de sus conjuntos de estados y las funciones de transición.

Nivel 7: Un sistema se especifica como un acoplamiento modular de componentes en contraste con los acoplamientos no modulares de nivel 6. Los componentes representan características técnicas del sistema y se acoplan conectando sus interfaces de entrada y salida.

Con esta especificación, por un proceso de múltiples pasos podemos expresar la conducta asociada con sistemas a cualquier nivel. El proceso de ir de una especificación del sistema a la especificación de su comportamiento es equivalente a la simulación de un modelo por computadora. La máquina obtiene, paso a paso los estados del sistema y la salida. Cada paso involucra el cálculo de las variables de estado componente a componente. Las secuencias de estados y salidas constituyen trayectorias, y los conjuntos de tales trayectorias constituyen la conducta del modelo. Para los propósitos de la simulación, los niveles más estructurados son muy prácticos. Esta base formal nos servirá ahora para plantear las estrategias de simulación [149].

2.2.6.3. Simuladores DEVS.

Los simuladores son los responsables de ejecutar la dinámica del modelo (en forma de instrucciones). El simulador puede operar de varias maneras:

- Secuencial.
- Paralela.

Nivel	Especificación	Formalismo
0	Marco E/S	(T, X, Y)
1	Relación observación E/S	(T, X, Ω, Y, R)
2	Función observación E/S	(T, X, Ω, Y, F)
3	Sistema E/S	$(T, X, \Omega, Y, Q, \Delta, \Lambda)$
4	Sistema Secuencial	$(T, X, \Omega G, Y, Q, \delta, \lambda)$
5	Sistema Estructurado	$(T, X, \Omega, Y, Q, \Delta, \Lambda)$ con X, Y, Q, Δ, Λ estructurados
6	Sistema no modular acoplado	$(T, X, \Omega, Y, D, M_d d \in D)$ con $M_d = (Q_d, E_d, I_d, \Delta d, \Lambda_d)$
7	Sistema modular acoplado	$N = (T, X_N, Y_N, D, \{M_d d \in D\},$ $\{I_d d \in D \cup \{N\}\}, \{Z_d d \in D \cup \{N\}\})$

Figura 2.24: Jerarquía de Especificación de Sistemas.

- Distribuida (secuencial/paralela).

Existen algoritmos básicos para los simuladores sobre DEVS que se basan en el manejo de una lista de eventos ordenada según el tiempo de activación de los eventos. Los eventos se extraen en orden secuencial de dicha lista y se ejecutan tomando en cuenta las transiciones de estados.

La ejecución de un evento genera nuevos eventos que deben ser incluidos en la lista de eventos. Así mismo una ejecución puede ocasionar cancelación o eliminación de eventos.

Tomando en cuenta estas consideraciones generales, formularemos esquemas para simuladores DEVS siguiendo el modelo básico. En particular, el algoritmo básico para el simulador atómico, mostrado en la figura 2.25, es [149]:

1. Asignar valores iniciales al tiempo actual. $t_A = 0$
2. Actualizar el estado del sistema.
 - a) Obtener el estado del sistema para el tiempo actual
 - b) Generar los mensajes de salida
 - c) Ordenar y distribuir los mensajes de salida
 - d) Si ocurren eventos externos ejecutar la función de transición.
3. Calcular el tiempo del próximo evento a partir de la función $ta()$ $t_A = ta()$
4. Regresar al paso 2

2.2.6.4. Enfoques para Simulación.

Como mencionamos anteriormente, uno de los conceptos clave en simulación es la manera en que se realiza el avance del tiempo. Este concepto permite realizar una distinción básica entre las

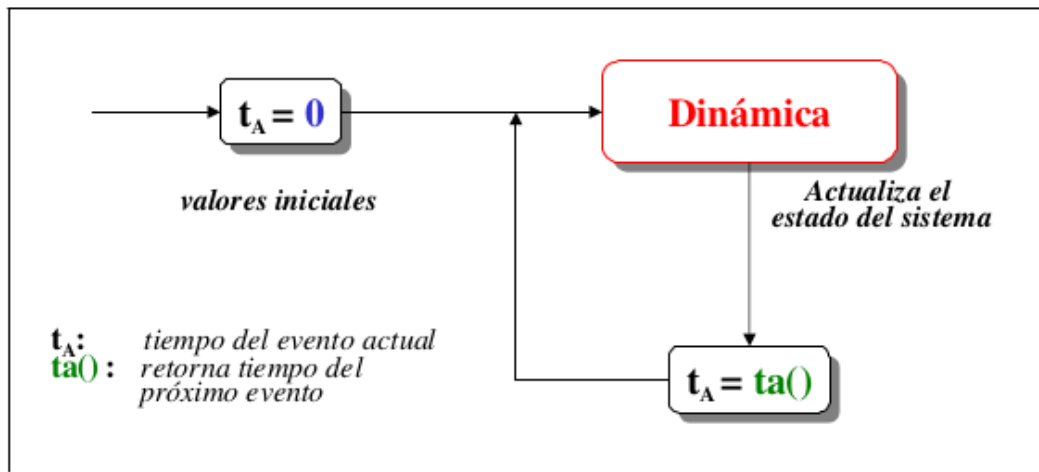


Figura 2.25: Algoritmo básico de un simulador atómico.

simulaciones basadas en eventos discretos y aquellas basadas en modelos continuos.

Las Simulaciones de Eventos Discretos están basadas en modelos que asumen que el sistema no cambia de estado hasta que ocurre un evento. Dicho evento genera un cambio instantáneo en el estado del sistema. En este caso, el avance del tiempo se realiza de evento a evento y los eventos se ordenan cronológicamente. Estas simulaciones se encargan generalmente de representar sistemas humanos o autómatas que deben ejecutarse en forma sincronizada, generando análisis estadísticos.

Por otro lado, los modelos continuos generalmente se encargan de representar sistemas a través de ecuaciones diferenciales que pueden ser ejecutadas simultáneamente, en lotes, generando tablas o gráficos de funciones continuas. Estos modelos asumen que el tiempo avanza en forma continua y los cambios de estado del sistema son también continuos. El avance de tiempo en este caso se representa con una variable que avanza en incrementos constantes lo suficientemente pequeños para simular en forma precisa el comportamiento continuo del tiempo. Para simular el estado continuo las variables que representan el estado del sistema son actualizadas en cada avance de tiempo. Las simulaciones basadas en estos modelos se denominan Simulaciones Continuas.

La Simulación Combinada, propuesta por Fahrland [152], abarca la simulación de sistemas que están constituidos por subsistemas discretos y subsistemas continuos que se ejecutan de manera coordinada y cooperativa. Una simulación combinada es una operación concurrente en el tiempo donde el tiempo para las partes discreta y continua debe avanzar de forma sincronizada. La incorporación de subsistemas discretos hace que el sistema completo pueda ser modelado como discreto. De esta forma los segmentos uniformes de la parte continua son incorporados como eventos, y deben ordenarse cronológicamente con los eventos del subsistema discreto.

La simulación combinada es más que una simple agregación de los componentes discretos y

continuos ya que proporciona los mecanismos para el acceso a las variables de estado, la activación y desactivación de procesos y la interacción estructural mutua entre los componentes discretos y continuos. Además, al establecer interfaces de comunicación entre los componentes discretos y continuo la simulación combinada permite la construcción de módulos y al organizar dichos componentes en distintos niveles de agregación, se permite la construcción de jerarquías.

Una vez aclarada la forma de representar sistemas basados en modelos continuos y en eventos discretos siguiendo el esquema de eventos discretos, no es necesario hacer distinción entre ambos paradigmas de modelado. Por tanto, a partir de este momento nos concentraremos en el estudio de simulación de eventos discretos.

En general, se conocen dos métodos que permiten resolver el problema del avance del tiempo y del manejo y ordenamiento de la lista de eventos.

1. El primero de ellos sigue la perspectiva optimista y supone que cada proceso programa eventos en la lista independientemente del resto de los procesos. En el caso que algún proceso requiera programar un evento en el pasado, es decir, que debe ocurrir antes que el siguiente evento de la lista, este método utiliza los mecanismos de regreso del tiempo para retornar el tiempo y el estado de la simulación al momento en que debe ocurrir el evento que se desea programar.
2. El segundo método sigue la perspectiva conservadora en el cual se utiliza la sincronización de los procesos para detener un avance de tiempo indebido, de ésta manera, no se registra un avance de tiempo hasta que se asegura que no ocurrirán programaciones de eventos en el pasado del componente en cuestión.

Veremos a continuación un enfoque de simulación, en el orden de ideas de las que haremos uso. Primero, las cotizaciones de monedas se actualizan cada vez que se realiza una transacción lo que la define de facto en un sistema de eventos discretos y segundo, como consecuencia de la gran cantidad de transacciones que se manejan, las páginas no ofrecen registros completos (población), por lo tanto en la mayoría de los casos cualquier observador tendrá acceso a una muestra de los mismos, dicha muestra corresponde a datos de promedios o a datos OHLC para intervalos específicos, de esta manera tendremos datos promediados o medidos cada Δ_t , en nuestro caso la medida en cuestión es una variable discreta (precio o volumen).

Es así como tomando la premisa que los eventos en finanzas son cíclicos y específicamente en las criptomonedas, podemos generar distintos modelos compuestos por dos partes en las cuales podemos variar su funcionamiento, las dos partes son:

1. **Selección de la muestra:** dada una población de datos, la toma de la muestra como hemos explicado en varias oportunidades tiene una gran variedad de posibilidades en cuanto a las técnicas de muestreo ya que podríamos usar las ya mencionadas e incluso personalizarlas con la adición de distribuciones de probabilidades al efectuar la selección de muestras para cada estrato y no quedarnos solamente con una muestra aleatoria para cada uno, esto sería similar al muestreo cada k-esima pero un poco más complejo.
2. **Escenario de Simulación:** ya con nuestra muestra de los datos distinguimos dos escenarios para la simulación:

Basado en Réplica: orientados a tomar los valores de esta muestra y mediante la repetición de los mismos datos, construir la simulación. Reordenar o no los datos de nuevo esta abierto a nuestra imaginación, podríamos replicar un solo intervalo de tiempo de la muestra, representarlo en orden inverso, entre otros. La desventaja de este método es que al no estar generando nuevos datos o mediciones, la simulación siempre estará dentro de los parámetros de lo que ya sucedió, en ocasiones esto podría ser una ventaja, pero queda a juicio de cada usuario y cada estudio en específico.

Basado en Predicción: se diferencia del anterior principalmente en que para este escenario estudiaremos los datos que tenemos y procederemos a la generación de datos para representar la simulación. Un enfoque que nos interesa es el que intenta extraer el comportamiento de la muestra analizando cada transición en específico, de esta manera dado un precio x_j en un tiempo $t = t_j$, luego de un intervalo de tiempo específico Δ_t tendríamos x_{j+1} en el tiempo $t = t_j + \Delta_t$, planteamos tres posibilidades para definir la evolución del sistema:

a) Su cambio original:

$$\Omega_{j+1} = x_{j+1} - x_j \quad (2.104)$$

b) Su cambio absoluto:

$$\Omega_{j+1} = |x_{j+1} - x_j| \quad (2.105)$$

c) O su cambio en proporción:

$$\Omega_{j+1} = 100 \frac{x_{j+1}}{x_j} \quad (2.106)$$

Si bien el segundo escenario nos brinda libertad en cuanto a la generación de datos nuevos, no siempre está la necesidad de que la simulación esté o no limitada por los datos del pasado o que específicamente queramos representar una réplica de lo que ya sucedió.

Es por esto que intentaremos abarcar ambos escenarios cada uno con acercamientos distintos y así ofrecer una amplia gama de configuraciones al momento de realizar una simulación. Ya con los formalismos definidos y los enfoques de simulación explicados, pasemos a explicar los tipos de reportes que esperamos obtener y/o generar, estos claramente influenciados por los que vimos en el capítulo 1 los cuales son ofrecidos usualmente por las páginas de consulta de datos y algunos extra.

2.3. Reportes.

Por último en este capítulo, veamos lo que esperamos lograr en cuanto a la generación de reportes y sus respectivas visualizaciones. Algo importante es mencionar que dada una serie de datos, dependiendo del estudio en cuestión si bien los datos descargados representan una muestra, estos serán tomados como una muestra o como una población como es en el caso de realizar una simulación. Estas muestras que descargamos podrían representar muestras de distintos tipos de intervalos entre datos: un minuto, una hora, un día entre otros.

Esperamos que nuestro programa pueda generar tantas visualizaciones y reportes en simultáneo como el equipo lo permita, de ésta manera podremos generar diversos estudios sobre una misma serie de datos y compararlos en simultáneo, claro que la comparación se suele hacer mayormente en forma de gráficos más que con tablas de datos pero esto dependerá de la necesidad del usuario.

A continuación, veamos a detalle la información o especificaciones necesarias (entradas) que se solicitarán para particularizar cada extracción de datos (descarga) y su posterior análisis. De esta manera, cada descarga será en base a los requerimientos del usuario.

2.3.1. Entradas.

No podríamos hablar de los reportes que esperamos ofrecer sin hablar previamente del funcionamiento de la descarga de datos. En ésta sección en específico nos centraremos en los datos que debe suministrar el usuario para la extracción de los datos pertinentes a una consulta y sobre los cuales generaremos reportes.

A continuación las cuatro entradas y sus nombres como nos referiremos a ellas de ahora en adelante:

Sitio Web, sitioW: en principio esta entrada registrará las demás, en cuanto a la dinámica de selección de sitioW esperamos contar con una variedad de páginas, de las cuales podamos escoger cuantas queramos, descargando así un archivo de datos proveniente de cada fuente de datos. Seleccionar varias sitioW irá comprometiendo las entradas siguientes ya que éstas deben ser validadas para cada una de las que estemos seleccionando.

Estas sitioW pueden ser exchange como tal o bien páginas que recopilan datos y los divulgan como vimos en la sección 1.5.

Período de consulta, fi-ff: luego de seleccionar la o las fuentes de datos, se solicitará la introducción de dos instantes de tiempo válidos para las fuentes en formato UTC como vimos en la sección 1.4.1, el cual especifica año, mes, día, hora, minuto y segundo. Esto en teoría lo podríamos realizar de dos maneras:

La primera mediante la selección de uno de los períodos ya establecidos, estos son los de uso mas común como lo son: últimas 24 horas, última semana, último mes, últimos tres meses, último año, últimos tres años o toda la serie posible.

La segunda mediante la elección manualmente de dos instantes de tiempo a los cuales llamaremos:

Fecha Inicial, fi: representa el instante inicial de la serie de datos que descargaremos. Ésta dependerá de las sitioW que hayamos seleccionado anteriormente.

Fecha Final, ff: representa el instante final de la serie de datos que descargaremos, ésta fecha está condicionada a ser estrictamente mayor que la f_i y menor o igual al tiempo de la realización de la descarga.

Al momento de realizar la descarga, ambos instantes se verificarán tanto en cuanto a disponibilidad de datos como en cuanto a que estén en el formato correcto, de no cumplir con alguna de estas condiciones, se le hará saber al usuario qué cuál fué el error y como resolverlo.

Criptomoneda, cript: se podrá seleccionar una criptomoneda de la lista de cript disponibles, dicha lista será el resultado de la intersección de todas las cript ofrecidas por las sitioW que hayamos seleccionado previamente.

Moneda Fíat, fiat: en este caso, es similar a lo anterior ya que se deberá realizar de igual manera una lista con la intersección de todas las monedas fiat disponibles

para las **sitiow** que hayamos seleccionado pero en este caso podremos seleccionar una o mas monedas, y así realizar la descarga de la cotización de un **cript** contra todas las **fiat** que seleccionemos.

Ya con nuestras entradas proporcionadas; procedemos a la descarga de los datos, igualmente que como bien mencionamos con respecto al período, si hubiese algún problema con una entrada, se deberá notificar cual es el error y posteriormente repetir el proceso de suministro de entradas, en el caso de que ésta resulte exitosa se dará inicio a lo que llamaremos “sesión” y se procederá al estudio de los datos de la misma o a la generación de reportes, para este punto ambas representan lo mismo. De esta manera cada sesión contará con sus datos respectivos y los reportes estarán limitados estrictamente a estos.

2.3.2. Archivo de datos.

Luego de especificar nuestras entradas y realizar de manera exitosa la descarga, tendremos como resultado un archivo descargado para cada par **cript/fiat** respectivo a cada **sitiow** que hayamos seleccionado. Los datos con los que esperamos contar de nuestras fuentes y como nos referiremos a ellos son:

Estampa de Tiempo: ésta se refiere a la estampa de tiempo en formato UTC del dato en cuestión, en nuestro caso ésta estampa de tiempo es común para cada grupo de datos, que lo conforman un dato de cada tipo de los que veremos a continuación. Este dato lo usaremos como índice en la tabla de datos que construiremos para cada descarga en específico.

Cotización Promedio, average: representa el promedio de las cotizaciones de todas las transacciones realizadas en un intervalo de tiempo específico. Este según la fuente en cuestión podrá ser un promedio de la población en el caso de consultar a un exchange o puede ser promedio de promedios si consultamos un sitio que recopile datos.

Datos OHLC, OHLC: este grupo de datos puede a o no representar datos únicos, ya que son cuatro datos que representan a cada intervalo de tiempo, pero si bien accedemos a un sitio que rastrea varios exchange o a un exchange en particular, el precio del Open o Close podría estar o no promediado, para el caso del High y Low asumimos que estos son datos únicos independientemente de la fuente. Ya que éste por definición es el máximo o mínimo del intervalo de tiempo en cuestión,

su marca de tiempo es desconocida hasta que sucede. Veamos la definición de éstos cuatro datos:

Apertura, Open: cotización de la primera transacción dentro del intervalo.

Máximo, High: el máximo valor de la cotización durante el intervalo.

Mínimo, Low: el mínimo valor de la cotización durante el intervalo.

Cierre, Close: cotización de la ultima transacción dentro del intervalo.

Volumen Global, Vol(G): cantidad máxima conocida de criptomonedas intercambiadas a nivel global de una criptomoneda en específico.

Volumen Fíat, Vol(Fíat): este representa el volumen de intercambio para un par expresado en la moneda fíat del mismo.

Capitalización del Mercado, MarketCap: calculado mediante la multiplicación de la cotización promedio de una criptomoneda por la cantidad de monedas circulantes (cript_c).

$$\text{MarketCap} = \text{average} \cdot \text{cript}_c \quad (2.107)$$

De esta manera nuestro `archivo.dat` contaría con ocho series para cada par de `cript/fíat` desde una fecha inicial hasta una final y fuente específica. Los futuros reportes serán en base a estos datos, los cuales dependiendo del estudio serán tomados como la muestra que son o como una población para efectuar nuevamente muestreo.

2.3.3. Tipos de Reportes.

Ahora sí, vayamos al punto de esta sección, empezando con los reportes estadísticos sobre los datos. Para realizar éstos reportes se deberá escoger una columna de nuestro `archivo.dat` y el calculo se aplicara a todos los datos. De aquí en adelante los clasificaremos como estadísticos de tipo atómico y de tipo serial.

Atómicos: Estos están integrados por los estadísticos de tendencia central y dispersión vistos en la sección 2.2.3.1, debido a que estos se calculan en base a todos los datos de un tipo específico que tengamos dando como resultado un número o un rango. Esperamos realizar un solo reporte donde se muestren los resultados para todos los estadísticos de este tipo que el usuario seleccione, teniendo como opción rápida el “seleccionar todos”.

Seriales: conformados por los estadísticos usualmente usados en trading de la sección 2.2.3.1, éstos estadísticos como bien su nombre lo dice producen series de datos que siguen el com-

portamiento de los datos originales de tal forma que su representación común es mediante gráficos ya que su análisis es mayormente visual.

Esperamos añadir al archivo.dat todas las series del estadístico en cuestión. Si bien podemos realizar distintos análisis a los mismos datos durante una sesión, es posible que el hecho de ir añadiendo series al archivo.dat cada vez que se pruebe una configuración distinta o algún otro estadístico podría resultar confuso para su lectura si una sesión se extiende a un número grande de estudios, por lo que parece una mejor opción generar un reporte con los datos originales y solo una serie (o series si el estadístico lo requiere) a la vez.

Con respecto a la inferencia, esperamos generar reportes a corto, mediano o largo plazo según los datos extraídos lo permitan, ya que el intervalo de tiempo entre los datos que hayamos extraído estará correlacionado con la inferencia que realizaremos de manera directa, dando así como resultado que si poseemos datos minuto a minuto, la inferencia sera minuto a minuto, si los datos son diarios obtendremos inferencias en intervalos diarios y así para cada tipo de intervalo al que tengamos acceso.

En cuanto a las entradas para un estudio del tipo inferencia esperamos seleccionar una columna de nuestro archivo.dat, utilizar una media móvil como estimador de la serie y así conseguir añadir algunos datos a la misma con el propósito de inferir la dirección del mercado hacia el futuro.

En cuanto a la comparación de datos, esperamos realizar dos tipos de gráficos distintos, los cuales son:

Cotización-Cotización: Un tipo de reporte que no se ofrece hasta ahora en los sitios web de divulgación de datos es el poder comparar cotizaciones entre ellas tomando una criptomoneda como referencia, si bien este estudio podría ser más interesante como estudio hacia las monedas fiat en cuestión, esperamos que nos brinde información útil.

Varias Monedas: esto nos permitirá observar el comportamiento en simultáneo de la cotización de varias monedas fiat con respecto a una criptomoneda, en este caso tendríamos dos escenarios, uno donde debemos que tener cuidado de comparar monedas que estén en rango de precios parecidos a menos que nuestra intención sea observar los diferentes órdenes de magnitud y el otro escenario donde las cotizaciones son normalizadas y de esta manera comparadas, dejándonos así tal vez una comparación mas interesante.

Por último en cuanto al tema de simulación, esperamos presentar dos posibles tipos de reportes de simulaciones que como dijimos anteriormente se espera abarquen los dos tipos de escenarios tanto de réplica como de predicción. Para esto el usuario podrá escoger una nueva fecha inicial y final (en formato UTC) como las datos a tomar en cuenta para las simulaciones, obviamente estas fechas deben estar dentro del período respectivo de la sesión, posteriormente escoger dos fechas más con el mismo formato que las anteriores pero en cuanto a las restricciones éstas dos en teoría sólo cuentan con la restricción de que una debe ser mayor que la otra, las cuales serán el inicio y final del período a simular.

2.3.4. Salidas.

En cuanto a las salidas de nuestro programa, nos referimos a cómo obtendrá el usuario los reportes que va generando a lo largo de la sesión, esperamos contar con dos opciones.

La primera es mediante la visualización de cada reporte al momento de realizarlo, esto sería mediante la generación de una ventana emergente individual para cada uno respectivamente, en principio la cantidad de ventanas estaría limitada solamente por las capacidades del ordenador, teniendo así la posibilidad de comparar una gran cantidad de configuraciones de estadísticos y pares de cotización.

La segunda sería mediante la descarga directa de la tabla de datos, generando así un archivo con la tabla de datos original y la o las series correspondientes a cada estadístico en particular que se esté estudiando, el cual podrá ser utilizado posteriormente por el usuario para lo que requiera.

Con esto damos por terminado el capítulo 2, procedamos ahora a explicar lo que hemos logrado y como ha sido el proceso para ello.

“No es suficiente estar en contra, tienes que estar a favor de algo mejor”

Tony Stark

Capítulo 3

Método Propuesto para la Minería de Datos.

El presente capítulo está destinado a explicar el diseño de un método que facilite las tareas de análisis de cotizaciones en los mercados financieros y la toma de decisiones al momento de realizar intercambios. Con este método proponemos consultar múltiples servidores que contienen datos históricos de cotizaciones para procesarlos, permitir realizar los estudios y emitir reportes que el usuario considere pertinentes al momento de realizar el análisis de cotizaciones o la toma de decisiones según sea el caso.

Este método consta de tres fases. Iniciaremos explicando el método desde una visión general y luego detallaremos cada fase del método y sus ventajas. Adicionalmente mostraremos los detalles de un primer prototipo funcional que implementa el método propuesto así para cada una de estas fases procederemos a explicar cómo hemos llevado el método propuesto al prototipo funcional y de qué forma se ha ido resolviendo cada proceso dentro de cada una de las fases, explicando así en cada proceso las soluciones por las que hemos optado y el por qué de esta decisión.

3.1. Diseño y arquitectura del método propuesto.

El método propuesto nos permite minar datos de cotizaciones de criptomonedas expresadas en monedas fíat, recordando que dentro de lo que llamamos monedas fíat incluimos los cinco metales preciosos mencionados en la sección 1.2.3. El método posee una arquitectura de tres fases: extracción, procesamiento y reportes, los cuales son gestionados por una Interfaz Gráfica de Usuario

(GUI) tal cual muestra la figura 3.1.

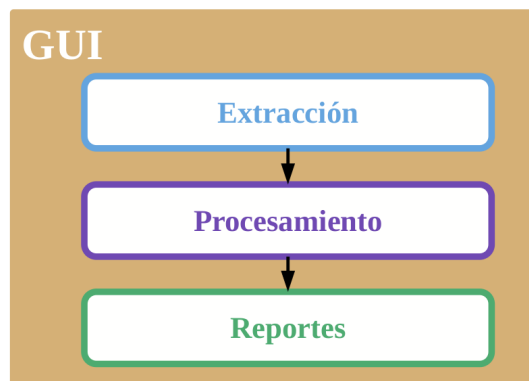


Figura 3.1: Fases del método propuesto.

Este método busca guiar al usuario en la toma de decisiones y el análisis de datos proporcionando una herramienta para realizar minería de datos financieros a partir de la consulta de múltiples servidores web que alojan valores históricos de cotizaciones para criptomonedas en monedas fiat. Se puede observar en la figura 3.2 los pasos detalladamente de lo que se define como minería de datos. Estos pasos pueden ejecutarse mediante la implementación de distintas herramientas dejando libre elección de software para cada uno en específico, pero respetando los lineamientos descritos en el método:

- El método está pensado para brindar acceso a la minería de datos financieros de una forma amigable, es por esto que nuestra propuesta ha ido más allá de la línea de comandos y ejecución de programas por separado. Proponemos una GUI que le permita al usuario de manera amigable, confiable, segura y práctica tener acceso a los estudios en las áreas de comparación, estadística, inferencia y simulación que le facilite las tareas de análisis y estimación de datos financieros
- Con la idea de ser un método escalable, las descargas de datos de los sitios web deben realizarse de forma modular, para ello proponemos un gestor de descargas que se encarga de la ejecución de los scripts específicos para cada una de los servidores, realizando la descarga total de los datos.
- Una vez iniciada la sesión, el usuario escoge de entre una gama de estudios el que desee realizar. De esta manera podrá realizar múltiples estudios, uno a la vez en base a los datos de la sesión. Luego de seleccionar el estudio se efectuaran los cálculos necesarios para obtener el resultado del estudio. Esta sección al igual que la anterior se ha pensado con la idea de dejar todas las áreas de estudio separadas para continuar con el diseño modular.

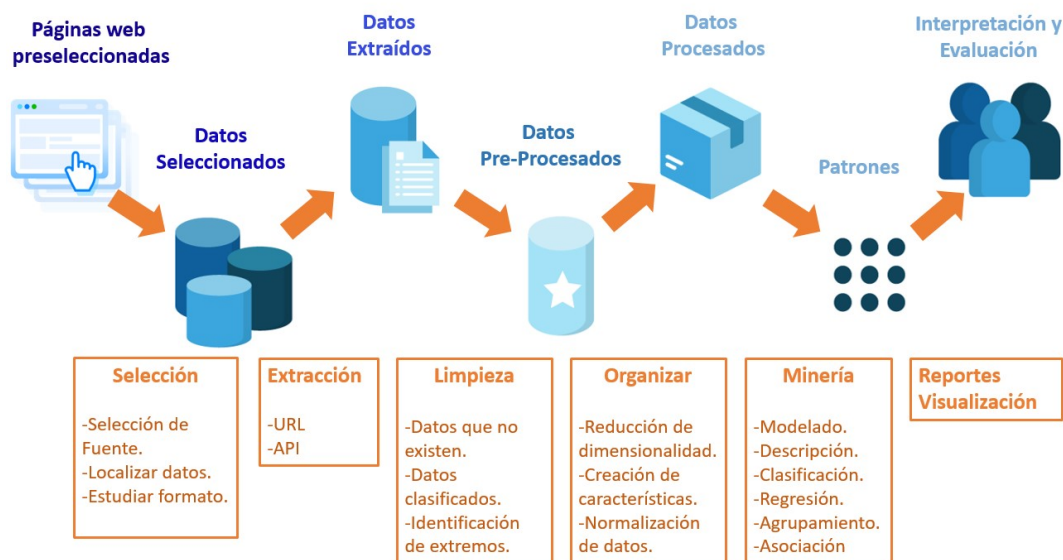


Figura 3.2: Pasos de la minería de datos

- En base a los resultados obtenidos el usuario podrá escoger de entre distintos tipos de reportes, uno a la vez, para visualizar los resultados del estudio realizado.

En resumen el usuario suministra los datos necesarios para el inicio de sesión, a partir de los datos de la sesión el usuario podrá realizar múltiples estudios y conjuntamente sus reportes de manera individual. Todo esto a través de la GUI.

Veamos ahora nuestro modelo en la figura 3.3 donde podremos observar detalladamente la arquitectura de cada fase y los pasos realizados dentro de cada una, en dicha figura podemos observar cuatro tipos de cajas:

- Las cajas rectangulares se refieren a scripts.
- Las cajas de esquinas redondeadas hacen referencia a fases del método.
- Las cajas de laterales circulares se refieren subprogramas.
- Las cajas con fondo rayado se refieren a subprogramas donde el usuario debe intervenir.

Veamos la primera fase del método a detalle:

3.1.1. Extracción

Nuestra fase de extracción engloba todas las tareas necesarias para que dadas las especificaciones del usuario se realice la extracción de datos y se generen los archivos a medida para la fase

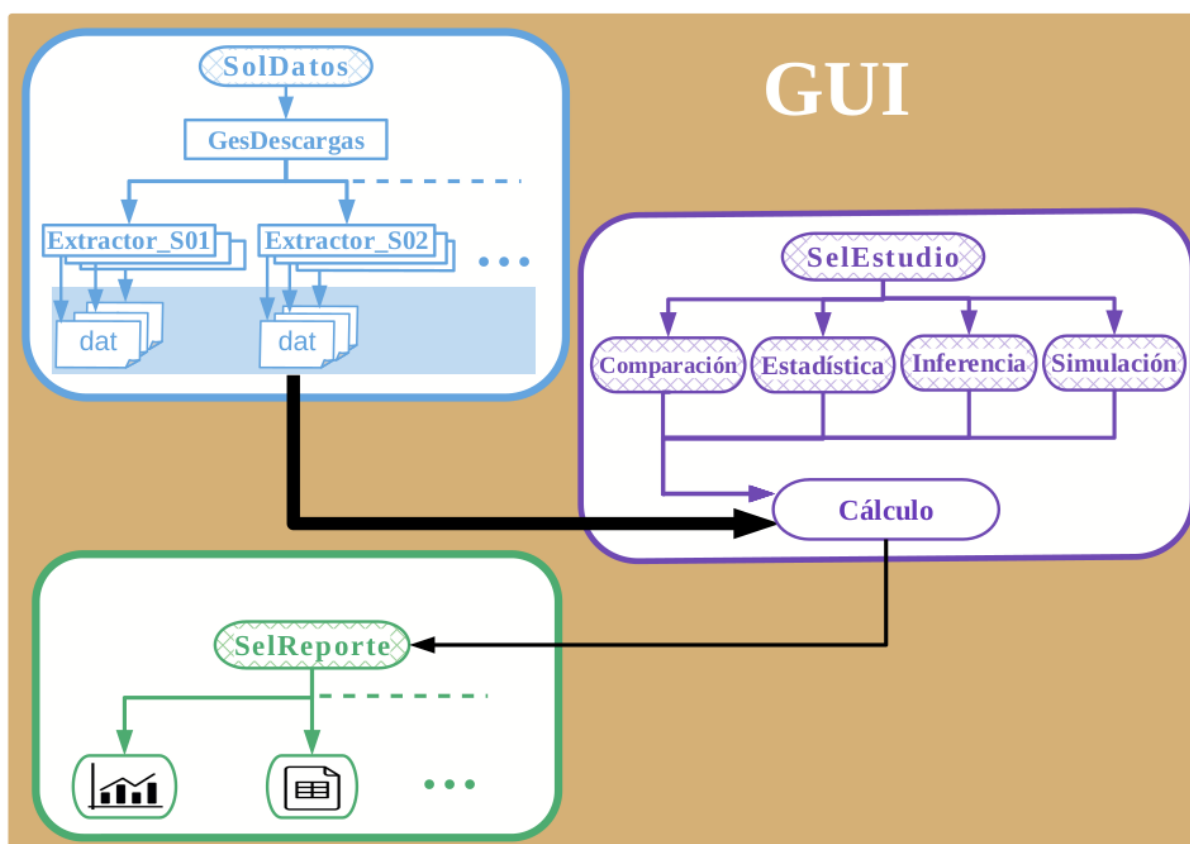


Figura 3.3: Arquitectura detallada del método propuesto.

de procesamiento. Estas tareas las podemos ver representadas en la figura 3.4 donde el primer paso incluye solicitar, recibir y verificar los datos necesarios para obtener los datos de la sesión, este paso se realizará mediante un subprograma que llamamos **SolDatos**. Una vez culminado este proceso, se ejecutará el script que gestiona las descargas y lleva por nombre **GesDescargas**, el cual procederá a crear una carpeta para almacenar los datos descargados y luego realiza invoca la ejecución de los scripts llamados **Extractor_SXX** correspondientes a cada uno de los servidores seleccionados. De resultar la extracción exitosa, **GesDescargas** dará por iniciada una sesión, en el caso de ocurrir un error se aborta el proceso y se elimina todo rastro del inicio de sesión fallida. De esta manera el usuario podrá iniciar una sesión a la vez.

Veamos cada una de éstas tareas por separado:

3.1.1.1. SolDatos.

SolDatos, este subprograma se encarga de la verificación de los datos necesarios para realizar una descarga de datos mediante la ejecución de **GesDescargas**. Este debe obtener los datos para la descarga mediante las herramientas disponibles de la GUI, verificar que sean datos válidos y en el caso de encontrar errores los notifica y ofrecer

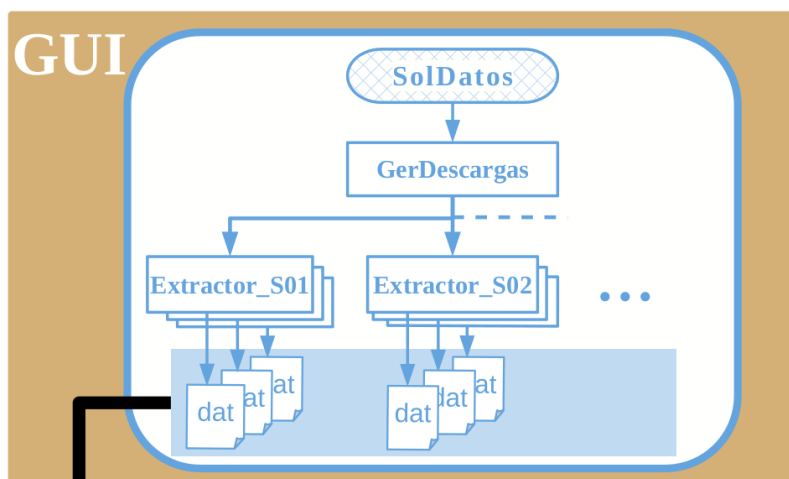


Figura 3.4: Fase de Extracción.

una pista de la solución.

Los datos necesarios para iniciar una sesión:

Fuentes: el usuario selecciona la o las fuentes que desee entre los sitios web de datos financieros ofrecidas en la GUI para la descarga de datos. Estas fuentes deben cumplir con ciertas especificaciones, siendo la más importante que corresponda a un sitio web interactivo, el cual haga uso de llamados `http` del tipo `GET` a endpoints para la descarga de datos en el sitio ya que este será nuestro método de extracción.

Criptomonedas: una vez seleccionada la(s) fuente(s), se crea una lista con las criptomonedas comunes disponibles en todas las fuentes. El usuario seleccionará las que desee para la descarga de datos.

Monedas Fíat: de manera similar a las criptomonedas se genera una lista de monedas fíat comunes presentes en las cotizaciones de las criptomonedas seleccionadas previamente, así el usuario procederá a seleccionar todas las monedas fíat en las que desee la cotización de las criptomonedas.

Período de Estudio: una vez seleccionadas las entradas previas, se procede a seleccionar un período válido para la extracción de los datos, esto se refiere a un período con instante inicial para el cual las fuentes posean datos de la descarga de datos seleccionada y uno final menor o igual al último registrado por las fuentes, con la condición de que el instante final debe ser mayor al inicial y que deben estar separados de tal manera

que al menos se extraigan un mínimo de datos establecido. Además de la selección manual del período de estudio, se cuenta con ocho períodos de tiempo preestablecidos que son:

- Última hora.
- Últimas 12 horas.
- Últimas 24 horas.
- Última semana.
- Último mes.
- Últimos 3 meses.
- Último año.
- Últimos 3 años.

SolDatos en primer lugar limita, en la medida de lo posible, la introducción de entradas no validas, reportando los errores que encuentre. Es aquí donde interviene la GUI y sus artilugios o herramientas (de ahora en adelante widgets). La ejecución de **SolDatos** culmina con el llamado a **GesDescargas**.

3.1.1.2. GesDescargas

Éste script se encarga de realizar las tareas necesarias para el inicio de una sesión.

- Inicia verificando cada uno de las entradas: fuentes, criptomonedas, monedas fiat y periodo de estudio.
- Procede a la creación de un directorio de trabajo para para la sesión
- Para cada sitio web se invoca el **Extractor_SXX** respectivo obteniendo los datos necesarios para cada par criptomoneda moneda fiat. Durante este proceso se verifica el correcto funcionamiento de cada extractor por separado y de resultar exitosa la ejecución de todos los scripts necesarios nos indicará que la descarga de los datos solicitados fue exitosa y se dará inicio a la sesión. En el caso de ocurrir un error se detiene el proceso y se eliminan los rastros de la sesión no valida.

En cuanto a los extractores veamos a continuación los detalles de su funcionamiento.

3.1.1.3. Extractor_SXX.

Planteamos generar un programa de descarga dedicado para cada sitio web el que llamaremos **Extractor_SXX**, donde **SXX** esta pensado para indicar el sitio web mediante tres letra en el caso de que los estemos manipulando de manera individual

y en el caso de utilizarlos mediante el **GesDescargas** previamente se crearán enlaces simbólicos con el mismo formato de nombre pero ahora las **XX** le asignaran a cada extractor existente un número según algún orden en especificado. Como se puede ver en la figura 3.4 que tenemos los **Extractor_SXX** con **XX** igual a cero uno (01) para la primera fuente de la lista de fuentes y cero dos (02) de manera análoga para la segunda.

El **Extractor_SXX** es una de las piezas esenciales del método. El realiza la extracción de los datos mediante las técnicas descritas en la sección 2.1.1, capturando las respuestas de llamados a endpoints diseñados para actualizar páginas de sitios web interactivos.

Sus cinco tareas son:

1. Construir el llamado al endpoint según dicta la API del sitio web y empleando las características específicas de la descarga de datos.
2. Realizar el llamado al endpoint mediante una herramienta que permite capturar la respuesta.
3. Almacenar los datos de la respuesta emitida por del servidor en un archivo.
4. Generar un archivo con el formato preestablecido con los datos descargados.
5. Realizar un filtrado de datos en el caso de que estos posean distintos intervalos de tiempo entre ellos, obteniendo así un archivo por cada tipo de intervalo.
6. Enviar los archivos resultantes a la carpeta de resultados respectiva.

Para la elaboración de un **Extractor_SXX** es necesario el estudio previo del sitio web de datos específico para el que se elabora. Al capturar los llamados a endpoints de los datos que queramos extraer procedemos al análisis de la respuesta que se obtiene con el propósito de especificar cómo están organizados los datos dentro de la respuesta.

Luego de entender tanto el funcionamiento del endpoint como la respuesta al mismo, procedemos a la creación de un script con el fin de automatizar dichas descargas y con esto nos referimos a la construcción del llamado al endpoint de manera adecuada con las entradas suministradas por el usuario. Este llamado debe realizarse mediante un método que pueda redireccionar la respuesta a dicha petición al endpoint hacia un archivo que modificaremos mediante editores de texto de línea y flujo con el único

fin de generar un archivo con un formato específico. En este punto debemos aclarar que invocar un **Extractor_SXX** tiene como resultado la extracción de datos para un único par (*cript,fiat*) y la generación de los archivos respectivos a ese par.

Los archivos generados para un par al invocar un **Extractor_SXX** son:

- Un archivo con la respuesta original del llamado al endpoint.
- Un archivo con los datos del período seleccionado y con un formato preestablecido.
- Archivos con los datos separados por tipo de intervalo en el caso de que la respuesta posea varios intervalos.

El formato es único para todos los datos que utilizaremos de aquí en adelante, independientemente de cómo sean suministrados por cada una de las fuentes.

	A	B	C
1	Date	SW1_C1	
2			...
3	YYYYMMDDhhmmss	C1	
4			
5	...	...	
6			

Figura 3.5: Formato propuesto para los archivos de datos.

El formato mostrado en la figura 3.5, se trata de un archivo del tipo CSV con delimitador punto y coma (;) ya que los datos financieros podrían contener puntos o comas. El primer registro de nuestro archivo posee la estampa de tiempo en formato UTC sin delimitadores y tendrá como etiqueta **Date** y los registros siguientes, a los que les llamaremos **indicadores**, poseen los datos específicos de cada fuente, etiquetados mediante las tres letras que representan el sitio web del que provienen seguido de la especificación del dato que representa, por ejemplo **CMCHigh** haría alusión a la cotización High del sitio CMC.

El formato usado para nombrar nuestros archivos de datos es:

AAA_BBBCCC_YYYYMMDD_YYYYMMDD.dat

Donde:

AAA_

Tres letras representando el sitio web de donde provienen los datos.

BBBCCC_

El par representado por tres letras para la criptomoneda seguido de tres letras para la moneda fíat.

YYYYMMDD_

Fecha inicial representada mediante cuatro caracteres para el año, dos para el mes y dos para el día.

YYYYMMDD

Fecha final representada mediante cuatro caracteres para el año, dos para el mes y dos para el día.

.dat

La extensión particular que le daremos al archivo.

Esta arquitectura de un **GesDescargas** y diversos **Extractor_SXX** simplifica el trabajo a la hora de agregar nuevas fuentes o modificar las presentes, de manera que para agregar una fuente extra o una misma fuente pero un endpoint apuntando a distintos datos tendríamos que:

1. Generar su script **Extractor_SXX** respectivo.
2. Agregar la opción en el **GesDescargas**.
3. Agregar la opción en **SolDatos**.

En el caso de querer bloquear una fuente existente solo hay que limitar su posibilidad de elección en el subprograma **SolDatos** sin tener que eliminar ningún archivo.

Así luego de ejecutar el **GesDescargas** de manera exitosa tendríamos todos los datos de una sesión listos para realizarle estudios dentro de una única carpeta.

Veamos antes de pasar a la sección de Estudios, a que se refieren los datos activos de una sesión.

3.1.1.4. Datos de la sesión.

Todos los datos descargados para una sesión se almacenarán en un directorio generado por el **GesDescargas**.

Los datos activos de una sesión son todos los archivos resultantes de todos los llamados a los **Extractores_SXX** que estén dentro de la carpeta de la sesión activa, representados mediante el área sombreada en la figura 3.4. Es por esto que la flecha de la figura 3.3 de los datos a el subprograma **Cálculo** es más ancha que las demás, ya que representa el uso de cada uno de los archivos.

Si bien al realizar una nueva descarga los datos de la sesión cambian, todos los archivos y reportes de la sesión previa quedan almacenados en su carpeta respectiva. Con los datos de la sesión activa, procederemos a realizar la selección del estudio a realizar.

Ahora teniendo claro a que llamamos datos de la sesión, procedamos con la siguiente fase del método, procesamiento.

3.1.2. Procesamiento.

Esta sección del método está destinada a generar los resultados para una gama de estudios y según unas especificaciones que más adelante se usan para generar los reportes.

El procesamiento consta de tres partes, las cuales podemos observar en la figura 3.6. Primero el usuario debe seleccionar en cuál área desea realizar el estudio mediante el subprograma **SelEstudio**. Mediante este subprograma el usuario selecciona el estudio específico que desea realizar y suministra las especificaciones para llevarlo a cabo. Por último el subprograma **Cálculo** efectúa los cálculos necesarios basado en las especificaciones suministradas por el usuario con el fin de obtener resultados en forma de arreglo de datos, listo para la generación de reportes sobre él.

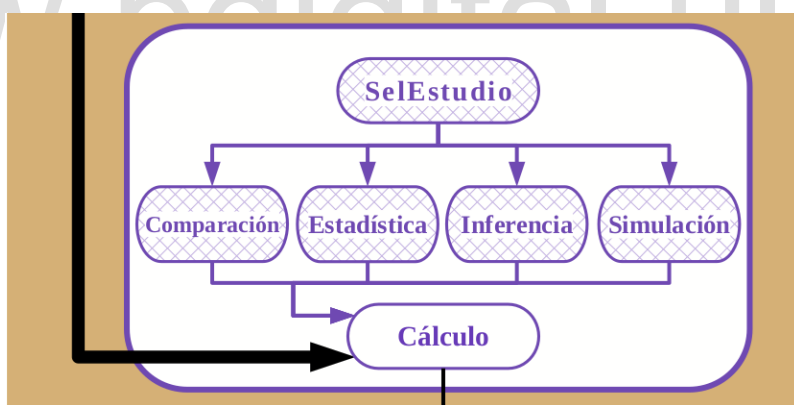


Figura 3.6: Fase de Procesamiento.

3.1.2.1. SelEstudio.

La primer parte es la selección por parte del usuario del área de estudios que se desea realizar mediante **SelEstudio** y luego el suministro de todos los datos necesarios en cuanto a especificaciones para el estudio. Las entradas son verificadas y de ser correctas se procede al cálculo. En cuanto a procesamiento de los datos obtenidos se espera abarcar cuatro tipos de estudio.

Comparación: realiza comparaciones de los datos descargados. Dentro de las comparaciones posibles está, por ejemplo, una serie de datos paralelamente a otra o graficar dos pares de cotizaciones de criptomonedas-dinero fiat una vs. la otra de forma similar a los diagramas de fase vistos en mecánica.

Estadística: estudiar los datos mediante medidas de tendencia central y estadísticos del tipo seriales comúnmente usados en trading, ambos explicados en la sección 2.2.3.

Inferencia: realizar inferencias que se usarán como pronóstico a corto, mediano o largo plazo dependiendo del tipo de intervalo que posean los datos usados, mediante la propuesta de un estimador como se vio en la sección 2.2.5.

Simulación: la simulación es algo no ofrecido por estos sitios web, por esto la inclusión de estudios basados en simulación es hasta cierto punto innovador, por lo que hemos incluido estudios de simulación basados en la teoría expuesta en la sección 2.2.6.

Cada sección de estudio está diseñada de manera modular de forma que, cada área de estudio se encargará del manejo de los estudios que contenga. La adición de nuevos estudios está a la orden del día en el ámbito de la economía y mediante éste diseño agregarlos al método es práctico, ya que lo necesario es la inclusión de la rutina y especificar los parámetros necesarios para realizar este nuevo estudio. Así una vez seleccionado el estudio y habiendo suministrado los datos necesarios para el mismo, procedemos a realizar los cálculos.

3.1.2.2. Cálculo.

Esta sección del método hace referencia al subprograma que mediante unos datos iniciales suministrados por el usuario y todos los datos extraídos por el **GesDescargas** realiza los cálculos necesarios y genera resultados en forma de un arreglo de datos. Dichos arreglos de datos posteriormente, según la decisión del usuario, son la materia prima para la generación de reportes de diversos tipos.

En cuanto al cálculo existen dos caminos posibles. El camino simplificado que se basa en buscar modelos y paquetes de cálculos ya establecidos y hacer uso de ellos pero dejando a un lado la potestad de hacer cambios a bajo nivel y tener el 100 % de certeza de que es

lo que se está haciendo y el segundo camino que contempla realizar subprogramas para realizar cada uno de los cálculos; esto suele ser más complicado pero nos da la certeza de qué estamos específicamente realizando, dándonos la oportunidad de efectuar cambios en todo el estudio; un ejemplo, teniendo en cuenta que algunos de los estadísticos vistos de trading en la sección 2.2.3 poseían ventanas las cuales se suelen dejar por omisión en un número específico en cambio creando nosotros el código podríamos dejar a juicio del usuario la manipulación total, esto podría verse como un arma de doble filo ya que más opciones no siempre se traduce en mejores resultados para el usuario.

Proponemos no tomar un camino en específico sino tomar una posición donde dependiendo de la tarea o el estudio escoger el primer o segundo camino, de esta manera a simple vista no escribiremos absolutamente todo el código ya que cálculos repetitivos y tediosos se podrán realizar mediante el uso de módulos, paquetes y funciones prescritas y la otra parte podría estar escrita desde cero por nosotros. Claro que ésta es nuestra recomendación pero se podría implementar el método solo juntando paquetes ya existentes sin ningún problema o escribiendo funciones con cada uno de los cálculos.

Realizado el cálculo, damos paso a la siguiente fase donde el usuario deberá decidir como querrá obtener los resultados.

Luego de realizar el cálculo según las especificaciones suministradas, el usuario procede a la generación de reportes de los resultados en cuestión. Luego de realizar al menos un cálculo se habilitaran las opciones para generar reportes del área en específica que hayamos realizado un cálculo.

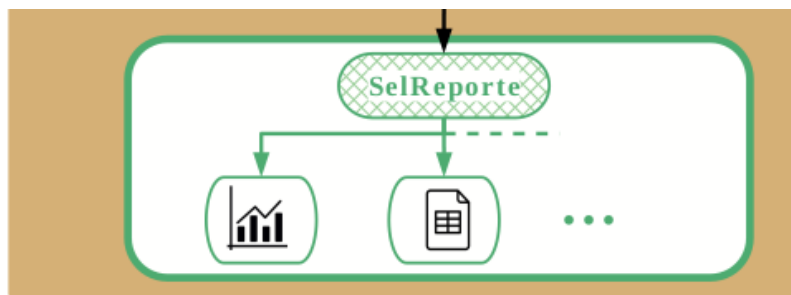


Figura 3.7: Fase de Reportes.

3.1.3. Reportes.

Luego de ser realizado el cálculo respectivo al estudio seleccionado el usuario podrá elegir cómo desea obtener los reportes en base a los resultados, esto mediante el subprograma **SelReporte** el

cual se encargará de obtener los resultados (arreglo de datos) y ofrecer las opciones de reportes que estén a su alcance y luego de la elección, emitirlo como muestra la figura 3.7.

3.1.3.1. SelReporte

Este subprograma se encarga únicamente de ofrecer al usuario una gama de reportes de manera sencilla, el usuario deberá escoger el tipo de reporte y automáticamente se realizarán los procesos respectivos para generar dicho reporte. En el caso de que el reporte necesite algún tipo de dato extra para ser generado, esta petición se efectuará dentro del subprograma respectivo a cada reporte.

Proponemos ofrecer dos tipos de reportes como mínimo:

- Tabla de datos como un archivo **CSV** con los datos originales y los resultados.
- Gráficos.

De nuevo como en las dos secciones anteriores, ya contando con los datos y los resultados del estudio seleccionado el abanico de generación de reportes es amplio, variando el formato de los archivos o incluso añadiendo funcionalidades como enviar los resultados directamente por correo, generar un archivo PDF entre otros. Una vez contando con lo básico podremos ir añadiendo más opciones de manera sencilla al agregarlas o eliminarlas de **SelReporte**.

Este apartado no se debe subestimar ya que podríamos estar en frente de grandes hallazgos mediante el cálculo empleando los datos, pero el objetivo final es la buena interpretación de los datos extraídos y los cálculos realizados, los reportes podrían fácilmente potenciar o limitar la experiencia de la implementación de un software que siga el método descrito.

Para finalizar el diseño hemos decidido implementar todas estas fases mediante una GUI.

3.1.4. Interfaz Gráfica de Usuario.

Ya contando con nuestros programas y subprogramas capaces de realizar minería de datos financieros de una manera robusta cumpliendo el método propuesto y sus tres fases (extracción, procesamiento y reportes), proponemos una GUI, esta tiene el propósito de facilitar y hacer amigable la implementación del método ya que todos estos programas y subprogramas poseen una gran cantidad de opciones y funcionalidades.

Estas opciones y funcionalidades se seleccionarán y apreciarán mediante la GUI y de sus widgets que facilitan tareas de selección y visualización, llevar acabo todas estas tareas mediante el terminal de comandos sería muy ambicioso y complicado para el usuario ya que nuestro método propuesto abarca cuatro características para realizar una extracción de datos y luego cuenta con cuatro áreas de estudio las cuales dentro poseen diversos estudios específicos. Siendo así la GUI un apoyo extra para el usuario y un pilar fundamental para la implementación del método propuesto incluyendo información contextual, ventanas emergentes para notificar errores, limitando así la mala implementación del método.

De manera que una vez bien organizado nuestro diseño procedemos a ponerlo en practica al crear nuestro prototipo completamente funcional del método propuesto.

3.2. Prototipo Funcional

Hemos bautizado nuestro prototipo funcional de software como **MAillardFinance**, el cual integra las iniciales de **Minería** y **Análisis** en la palabra **MAillard**, la cual hace referencia a una serie de reacciones químicas que suceden en los alimentos al aplicarles calor o algún método de conservación, dichas reacciones se suelen traducir en un mejor sabor y un color dentro de un rango de tonalidades de color café en el alimento tratado. Como peculiaridad, dentro de los colores producidos mediante estas reacciones podemos encontrar el color de nuestra Facultad de Ciencias, el cual se ha implementado en nuestra interfaz gráfica como fondo y base para la paleta de colores de la misma. La inclusión de esta palabra hace referencia a tomar los datos y mediante el software que los procesará obtener de ellos información accesible y provechosa mediante un debido tratamiento de los datos. Por otra parte, al agregar la palabra **Finance** se hace referencia a la palabra finanzas en inglés con el propósito de que sea intuitivo y global el área hacia donde esta orientado nuestro software.

3.2.1. Especificaciones del prototipo.

En relación a especificaciones de requerimientos y de implementación, comenzamos por el software requerido para el correcto funcionamiento del prototipo **MAillardFinance**.

3.2.1.1. Requerimientos del prototipo.

En cuanto a la selección de software se ha hecho uso de software libre orientado para el perfecto funcionamiento en dispositivos con sistema operativo tipo UNIX como lo es Linux. Dentro de los

software usados en el diseño de nuestro prototipo tenemos a **bash** como intérprete de comando para el sistema operativo con el cual hemos realizado script los cuales a su vez hacen uso de software como lo son **sed** con el fin de modificar caracteres, **grep** para el filtrado de líneas en particular, **awk** como editor de flujo de texto, **curl** para realizar los llamados a los endpoints, tanto para la GUI como para el procesamiento de los datos hemos hecho uso de **python3** y conjuntamente su instalador de librerías **pip**. Todo este software y sus versiones mínimas se puede visualizar en la figura 3.8. En el caso del sistema operativo *SparkyLinux* se refiere a cualquier versión 5.* con “full system upgrade”.

Software	Tipo	Versión Mínima	Versión Verificada
Ubuntu	Sistema operativo	20.04 LTS	22.04.1 LTS
SparkyLinux	Sistema operativo	5.*	6.3 Po-Tolo
GNU bash	Interprete de sistema operativo	–	5.0.17
curl	Comunicación http	–	7.68.0
GNU awk	Editor de flujo	–	5.2.0
GNU sed	Edición de texto	–	4.7
GNU grep	Filtrado de texto	–	3.4
python3	GUI	3.8	3.10.4
pip3	Instalador de librerías python3	–	22.0.4

Figura 3.8: Software requerido por el prototipo.

En cuanto a librerías de **python3** hemos hecho uso de **pandas** para el manejo de los datos en forma de datos tabulares bidimensionales y de tamaño variable, lo que llamaremos *DataFrame* de ahora en adelante y para exportar los mismos a formato **csv** y **xlsx**, **numpy** para el uso de algunas rutinas de cálculo ya establecidas, **scipy** para la implementación de rutinas preestablecidas en cuanto al uso de distribuciones de probabilidad, **matplotlib** para la generación de gráficos, **random** para la generación de números aleatorios, **os** para funcionalidades dependientes del sistema operativo y **subprocess** para el seguimiento de la ejecución de los mismos, **pillow** para el manejo de imágenes dentro de la interfaz, **Tkinter** para la realización de la interfaz gráfica, esta librería se implementa como un contenedor de **python3** alrededor de un intérprete Tcl (del inglés Tool Command Language) [153] completo integrado en el intérprete de **python3** con la finalidad de que las llamadas de **Tkinter** se traduzcan en comandos Tcl [154], **tkcalendar** como herramienta extra para el suministro de fechas por el usuario. La

Librería	Versión
pandas	1.1.4
numpy	1.19.3
scipy	1.8.1
matplotlib	3.3.2
Pillow	8.0.1
tkcalendar	1.6.1
DateTime	4.3
random	–
os	–
subprocess	–
Tkinter	–

Figura 3.9: Librería de python3 requeridas

lista completa de librerías y sus respectivas versiones se pueden observar en la figura 3.9.

En la figura 3.10 podemos observar detalladamente como hemos implementado el método propuesto en el prototipo y las herramienta que hemos utilizado en cuanto a programa y subprogramas.

Ya expuesto el software requerido por MAillardFinance, veamos ahora las especificaciones de implementación de nuestro prototipo.



Figura 3.10: Implementación del método propuesto en el prototipo

3.2.1.2. Especificaciones de implementación.

En cuanto a las especificaciones de implementación del prototipo MAillardFinance hemos abarcado todos los aspectos mencionados en la sección de diseño, cumpliendo así los requisitos mínimos y apegándonos al diseño modular lo cual se traduce en un software de fácil escalabilidad.

Estas especificaciones de implementación son:

Fuentes de datos: siguiendo las pautas de diseño del apartado Fuentes de la sección 3.1.1.1 hemos seleccionado dos fuentes (estudiadas en la sección 1.5). Por un lado tenemos al sitio web BitcoinAverage el cual es una fuente de datos que monitorea el precio y el volumen de transacción de las criptomonedas mas usadas a nivel global, recopilando así datos de cotización y volumen de la mayor cantidad de fuentes posibles y por el otro lado el sitio web CoinMarketCap el cual es uno de los más respetados en cuanto a divulgación de datos sobre cotizaciones de criptomonedas se refiere, éste cuenta con datos de cotización, volumen y capitalización de mercado.

Criptomonedas: siguiendo las pautas del apartado Criptomonedas de la sección 3.1.1.1 hemos seleccionado dos criptomonedas, las cuales cuentan con la mayor capitalización de mercado, éstas son Bitcoin (BTC) y Ether (ETH).

Monedas fiat: cumple los requisitos expuestos en la sección 3.1.1.1 en el apartado Monedas Fíat.

Período de estudio: cumple con los requisitos mínimos expuestos en la sección 3.1.1.1 en el apartado Período de Estudio.

Estudios: en cuanto a los estudios disponibles, poseemos las cuatro áreas propuestas en la sección 3.1.2.1 y en cada área presentamos los siguientes estudios los cuales serán explicados detenidamente más adelante:

■ Comparación:

1. Moneda vs Moneda.
2. Varias monedas.

■ Estadística:

1. Moneda + Trading.
2. Tendencia Central.

■ Inferencia:

1. Estimador MA.

■ Simulación:

1. Método Distri.
2. Método MAillard.

Período para y de simulación: en cuanto a la simulación, hemos implementado un esquema de tres instantes **a-b-c**, en el cual el usuario seleccionará los instantes **a** y **b** que delimitarán los datos a usar en la generación de la simulación y un instante **c** el cual indicará el final de la simulación. Ya que ésta tendrá lugar de **b** hasta **c**.

Reportes: la selección de reportes disponibles en nuestro prototipo se apega a los propuestos en la sección 3.1.3 con la adición de emitir un reporte directamente en formato **xlsx**. De manera que generaremos reportes mediante dos librerías de **python3**: **pandas** para la generación directa de archivos tipo **csv** y **xlsx** y **matplotlib** para la generación de gráficos, esta librería permite generar un gráfico a la vez y no es hasta que se guarda o se cierra el mismo que se puede proseguir con otras acciones, esto se solucionó mediante un método denominado “Incrustación de Matplotlib en interfaces gráficas de usuario” lo que genera como resultado una ventana emergente de la GUI la cual posee el gráfico y una barra de herramientas.

Ya teniendo claro tanto el software requerido como las especificaciones de nuestro prototipo, veamos como hacer uso de él y como funciona simultáneamente.

3.2.2. Inicio de sesión.

Al iniciar nuestro software nos ubicaremos en la pantalla de inicio a la que hemos denominado **Consulta** y se muestra en la figura 3.11, en esta figura se observa como hemos implementado distintos tipos de widgets [155] como lo son las casillas de verificación (*check box*), botones de opción (*radio button*) y cuadro de lista (*combo box*) para desempeñar las tareas del subprograma de diseño **Soldatos** en nuestro software, el cual cumple todas las especificaciones de la sección 3.1.1.1, en esta figura además se puede apreciar que al iniciar nuestro software la única pestaña habilitada es la pestaña dedicada a realizar consultas (descarga de datos) y por consiguiente el inicio de una sesión de estudio.



Figura 3.11: Pantalla de inicio de prototipo **MAillardFinance**.

Veamos a detalle como funciona esta pestaña **Consulta** y como está organizada:

- La sección **Fuentes** cuenta con el uso de casillas de verificación para realizar la selección de las fuentes deseadas para la descarga de datos. La fuente por omisión es **BitcoinAverage**.
- La sección **Criptomonedas** cuenta con un botón de opción con el cual seleccionaremos la criptomoneda deseada. La criptomoneda por omisión es **BTC**.
- La sección **Monedas** le ofrece al usuario la selección de las monedas fíat para la descarga de datos mediante un cuadro de lista, se realiza la selección de una moneda fíat a la vez y

para añadir o eliminar un cuadro de lista equivalente están disponibles dos botones con el símbolo “+” y “-” respectivamente. El cuadro de lista inicial cuenta con USD como moneda por omisión.

- La sección **Período de Estudio** cuenta con un cuadro de lista para la selección de entre los ocho períodos preestablecidos y la opción manual *seleccionar período*. El período por omisión es *última semana*.

Al lado de este cuadro de lista se muestra el instante inicial y final actualmente seleccionado. La opción *seleccionar período* se realiza mediante una ventana emergente que cuenta con un calendario para la especificación de el instante inicial y final mediante la selección de: año, mes, día, hora y minuto.

Veamos un ejemplo de descarga de datos como se muestra en la figura 3.12 y posteriormente el funcionamiento del script respectivo de esta acción.



Figura 3.12: Ejecución de GesDescargas en el prototipo.

Este ejemplo posee las siguientes especificaciones:

Fuentes: BitcoinAverage y CoinMarketCap.

Criptomoneda: Bitcoin.

Monedas: Dolar Estadounidense, Bolívar Soberano, Euro, Oro, Plata, Rublo ruso, Libra Esterlina y Yen.

Período: período preestablecido último año.

Una vez el usuario haya suministrado las especificaciones mediante el uso de los widgets, este debe proceder a realizar la descarga mediante el uso del botón **Descargar Datos**.

3.2.3. Descarga de datos.

Al hacer uso del botón **Descargar Datos** este es el encargado de proceder a la ejecución del subprograma **GesDescargas** suministrando a su vez las opciones seleccionadas en los widgets de **SolDatos** de nuestro prototipo.

Simultáneamente al llamado a este script como se puede observar en la figura 3.12 debajo del botón **Descargar Datos** se despliega una barra de progreso la cual realizará un movimiento periódico hasta finalizar la descarga. La lista de tareas desempeñadas por el script **GesDescargas** de nuestro software se observan detalladamente en la figura 3.13 descritas paso a paso.

-
1. Crea un directorio temporal que lleva por nombre el identificador de la sesión, donde se almacenarán los archivos necesarios para la sesión.
 2. Crea el directorio de trabajo de la sesión, donde se almacenarán todos los reportes generados por el usuario
 3. Para cada una de las monedas fiat:
 - 3.1 Para cada una de las fuentes:
 - 3.1.1 Realiza el llamado al extractor pertinente.
 - 3.1.2 Filtra los datos correspondientes al período de estudio.
 - 3.1.3 Genera los archivos respectivos para cada intervalo de tiempo.
 - 3.1.4 De ocurrir un error aborta la ejecución, elimina la carpeta de la sesión y genera un archivo donde se indica que la descarga fue fallida.
 4. Genera un archivo donde se indica que la ejecución fue exitosa.
-

Figura 3.13: Tareas del script **GesDescargas**

Para la ejecución del script las especificaciones se le suministran en forma de cinco entradas y específicamente para nuestro ejemplo éstas son:

Identificador de la sesión: tiempo POSIX del momento en que se hizo uso del botón **Descargar Datos** {1660160460}.

Fuentes: {BAV, CMC}

Criptomoneda: {BTC}

Monedas: {USD,VES,EUR,XAU,XAG,RUB,GBP,JPY}

Período: instante inicial y final en formato UTC extendido obviando la Z como se puede ver en la sección 1.8 {2021-08-09T15:41:00,2022-08-10T15:41:00}

Como se puede observar en el punto **3.1.1** de la lista de tareas de la figura 3.13, una de las tareas es realizar los llamados a los scripts respectivos para la descarga de datos según cada fuente especificando el par a consultar, en el caso de nuestro prototipo poseemos dos, uno para el sitio BitcoinAverage y otro para el sitio CoinMarketCap. El funcionamiento de estos scripts en líneas generales se observa en la figura 3.14.

-
1. Construye de manera adecuada el llamado al endpoint
 2. Realiza y captura el llamado mediante `curl`.
 3. Genera un archivo de datos con el formato preestablecido.
-

Figura 3.14: Tareas de un script `Extractor_SXX`

Veamos a detalle los dos extractores respectivos a nuestras fuentes de datos.

3.2.3.1. Extractor BAV.

Luego de un extenso análisis al sitio web BitcoinAverage (BAV) se decidió extraer los datos provenientes del gráfico del índice global para cada criptomoneda, en el sitio web éste gráfico se actualiza mediante un llamado a un endpoint al realizar la selección de una moneda fiat, este llamado se realiza mediante una petición HTTP del tipo GET por lo que es cacheable.

Para esta fuente es necesario la especificación del par criptomoneda y moneda fiat mas la adición de unos encabezados particulares como especificación para realizar una extracción y obtener una respuesta exitosa, estos encabezados son:

- | | | |
|------------------------------|-------------------------------|----------------------------|
| ■ <code>Accept:</code> | ■ <code>Authorization:</code> | ■ <code>Connection:</code> |
| ■ <code>X-session-id:</code> | ■ <code>Origin:</code> | ■ <code>Referer:</code> |

El gráfico en cuestión se crea a partir un archivo de datos el cual posee tres tipos de intervalos de tiempo los cuales se establecen respecto al momento de actualizar el gráfico, los intervalos y los datos que contienen cada uno son:

- Las últimas 28 horas poseen registros en intervalos de 1 minuto del índice global de la criptomoneda seleccionada, a este indicador llamaremos `BAVAverage`.

- Los últimos 35 días poseen registros en intervalos de 1 hora del índice global de la criptomoneda seleccionada **BAVAverage**.
- Para el resto del tiempo poseen registros en intervalos de 1 día del índice global de la criptomoneda seleccionada **BAVAverage** y datos del volumen global de criptomonedas intercambiadas a este indicador lo llamaremos **BAVVolG**.

Es así como la descarga de BAV luego de ser procesada por el gestor resultará en su máxima expresión como un archivo original de descarga y cuatro archivos generados a medida, el primero desde el instante inicial hasta el final sin distinción de intervalos con la extensión **.dat** y otros tres con datos cada día, hora y minuto con las extensiones **_day.dat**, **_hour.dat** y **_min.dat** respectivamente. Estos archivos cuentan con datos del índice global para la criptomoneda seleccionada y datos del volumen global de criptomonedas intercambiadas, en la figura 3.15 se puede apreciar un archivo de datos para el par BTCUSD (intervalos diario). Este extractor sera designado como nuestro extractor **S01**

	A	B	C
1	Date	BAVAverage	BAVVolG
2	20210926000000	42842.85	52657.9757123121
3	20210927000000	43537.46	31143.4309282087
4	20210928000000	41964.17	36789.7815388966
5	20210929000000	41845.97	30085.7109200867
6	20210930000000	43272.16	70804.093379513
7	20211001000000	46057.6	127567.7306274071
8	20211002000000	47827.22	22115.966102441398

Figura 3.15: Datos procesados del sitio BitcoinAverage.

3.2.3.2. Extractor CMC

En cuanto al sitio web CoinMarketCap (CMC) se decidió extraer los datos provenientes de su sección “datos históricos”, la cual cuenta con una tabla de registros diarios de cotizaciones de pares criptomoneda y moneda fiat en forma de datos OHLC (1.5.1) a estos indicadores los llamaremos **CMCOpen**, **CMCHigh**, **CMCLow** y **CMCClose**, volumen de transacción a este indicador le llamaremos **CMCVolF** y capitalización de mercado a cual llamaremos **CMCMarketCap** como se puede ver en la figura 3.16 el archivo de datos para el par BTCUSD. En cuanto a las especificaciones necesarias para realizar un llamado estas son: la especificación del par criptomoneda moneda-fiat, el instante inicial y final.

Al realizar el llamado a este endpoint tendremos como resultado dos archivos, uno con la respuesta original y otro con registros diarios para un par en específico con la extensión **.dat** y

	A	B	C	D	E	F	G
1	Date	CMCOpen	CMCHigh	CMCLow	CMCClose	CMCVoIF	CMCMarketCap
2	20210926000000	42721.62866	43919.3009	40848.4616	43208.539104	306612220	813490665612.76
3	20210927000000	43234.18492	44313.2458	42190.6325	42235.731846	309800290	795216175345.08
4	20210928000000	42200.89661	42775.1461	40931.6624	41034.544665	302149405	772639687708.88
5	20210929000000	41064.98395	42545.2585	40829.6692	41564.363481	306023599	782651228474.32
6	20210930000000	41551.27083	44092.6005	41444.5816	43790.895625	311416819	824619217607.52
7	20211001000000	43816.74064	48436.0112	43320.0229	48116.941218	428506415	906127699407.78

Figura 3.16: Datos procesados del sitio CoinMarketCap

con el formato a medida el cual se puede observar en la figura 3.16. Este extractor sera designado como nuestro extractor **S02**.

De resultar exitosa la descarga de datos, se da por iniciada una nueva sesión de estudio y se habilitan las demás pestañas de la GUI y la barra de progreso se mantiene estática en el 100 % como se puede ver en la figura 3.17. En el caso de ocurrir algún error en el proceso se emitirá el mensaje de error y se debe proceder nuevamente a hacer uso del botón [Descargar Datos](#).



Figura 3.17: Extracción completada por el prototipo.

Veamos a continuación lo que representa una nueva sesión de estudio.

3.2.4. Sesión de estudio.

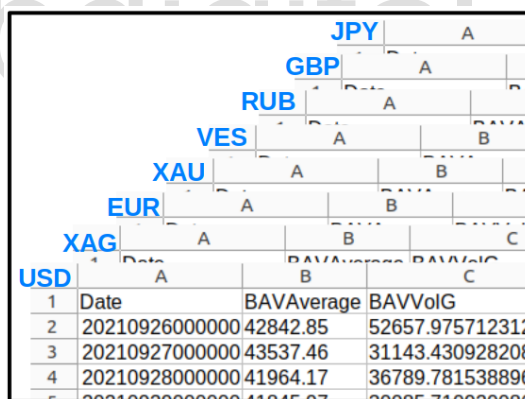
Al realizar la descarga con éxito damos por iniciada una sesión de estudio, con esto nos referimos a que procedemos a realizar estudios en las cuatro áreas disponibles por el software en base al

conglomerado de archivos que resultaron de la extracción. En la figura 3.18 se puede observar el posible conglomerado de archivos resultantes de la extracción de ambas fuentes para un par específico de criptomoneda-dinero fiat (CCCMMM), de manera que el conglomerado total de la sesión sera la suma de los respectivos conglomerados para cada moneda fiat.

Para nuestro ejemplo el conglomerado resultante serían 40 archivos correspondientes a 8 llamados de extractor y obteniendo por cada par CCCMMM 4 archivos del sitio BitcoinAverage y 1 del sitio CoinMarketCap. Los resultados de cada extractor en cuanto a registros diarios para las fuentes BAV y CMC lo podemos visualizar en las figuras 3.19 y 3.20 respectivamente. Con respecto a los otros archivos de intervalos distintos de BAV estos son equivalentes.

Fuente	Archivo	Descripción
BAV	BAV_CCCMMM_20210809_20220810.dat	Datos sin distinción de intervalos
	BAV_CCCMMM_20210809_20220810_day.dat	Datos de intervalos diarios
	BAV_CCCMMM_20210809_20220810_hour.dat	Datos de intervalos de una hora
	BAV_CCCMMM_20210809_20220810_min.dat	Datos de intervalos de un minuto
CMC	CMC_CCCMMM_20210809_20220810.dat	Datos de intervalos diarios

Figura 3.18: Archivos resultantes de una extracción para ambas fuentes.



	Date	BAVAverage	BAVVolG
1	20210926000000	42842.85	52657.975712312
2	20210927000000	43537.46	31143.430928208
3	20210928000000	41964.17	36789.781538896

Figura 3.19: Archivos diarios de la fuente BitcoinAverage para 8 monedas fiat

Estos dos conglomerados de archivos mas los respectivos en cuanto a intervalos de tiempo distintos del sitio BAV serán los datos que tendremos disponibles para la realización de estudios y generación de reportes respectivamente dentro de nuestro ejemplo. Veamos a continuación todo lo respectivo en cuanto a estudios y las especificaciones de los mismos que podremos realizar.

3.2.5. Selección de estudio.

Un vez se haya iniciado una nueva sesión de estudio y se hayan habilitado nuestras cuatro pestañas de áreas de estudio, el usuario debe seleccionar alguna en específico para realizar estudios

1	Date	CMCOpen	CMCHigh	CMCLow	CMCClose	CMCVolF
2	20210926000000	42721.628	43919.30	40848.46	43208.539	30661222
3	20210927000000	43234.184	44313.24	42190.63	42235.731	30980029
4	20210928000000	42200.896	42775.14	40931.66	41034.544	30214940
5	20210929000000	41064.983	42545.25	40829.66	41564.363	30602359
6	20210930000000	41551.270	44092.60	41444.58	43790.895	31141681

Figura 3.20: Archivos diarios de la fuente CoinMarketCap para 8 monedas fiat

respectivos a esa área. El resultado de realizar un estudio es un DataFrame (temporal), es este arreglo el que será usado para la generación de reportes. Algo que debemos recordar antes de continuar con los estudios es la definición que le hemos dado a **indicador**, para nosotros éste constituye la especificación tanto de la fuente como del tipo de dato y de manera práctica el indicador representa la columna específica del DataFrame que haremos uso para el estudio.

Veamos a continuación los estudios disponibles en nuestro software y sus DataFrame resultantes.

3.2.6. Comparación

En la pestaña de comparación de la GUI contaremos con un cuadro de lista en el que podremos escoger de entre dos tipos de estudios disponibles, estos son:

3.2.6.1. Moneda vs Moneda.

Este estudio nos permite comparar dos monedas fiat al enfrentar sus datos en base a un indicador en particular, de manera tal que se pueda apreciar la diferencia entre la subida o bajada del indicador expresado en las monedas fiat y tomando la criptomoneda como base. De esta manera se puede observar la fortaleza o debilidad de una moneda fiat con respecto a otra, tomando en cuenta que la lectura de este estudio puede necesitar de un usuario entrenado que posea conocimientos de ambas monedas fiat y los aspectos externos que puedan introducir sesgos a este estudio.

Como se puede observar en la figura 3.21 se nos permite escoger dos monedas presentes en la



Figura 3.21: Estudio Moneda vs Moneda del prototipo.

sesión y un indicador. El resultado de un estudio de este tipo tiene como resultado el DataFrame presentado en la figura 3.22, este posee la estampa de tiempo de cada registro y los datos de ambas monedas seleccionadas filtradas por el indicador seleccionado. De generar un reporte gráfico la primer moneda sera el eje **x**.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda1_indicador	Datos filtrados según indicador de la moneda principal
Moneda2_indicador	Datos filtrados según indicador de la moneda secundaria

Figura 3.22: Formato de resultado para Moneda vs Moneda.

3.2.6.2. Varias monedas.

Este estudio organiza los datos con respecto a un indicador en específico de todas las monedas seleccionadas de manera que se puedan visualizar todas con respecto a una única línea de tiempo, además se cuenta con la posibilidad de normalizar los datos y así eliminar las diferencias de escala que puedan existir al comparar monedas que estén en rangos distintos.

Como se observa en la figura 3.23 el usuario cuenta con casillas de verificación tanto para

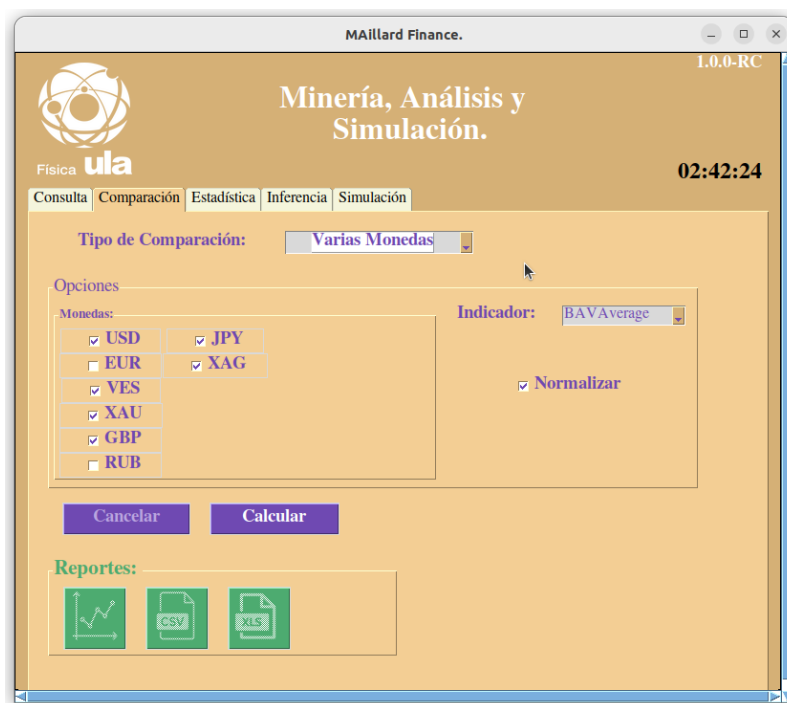


Figura 3.23: Estudio Varias Monedas del prototipo.

seleccionar las monedas fiat para el estudio como para aplicar normalización a los datos y el cuadro de lista para seleccionar el indicador a tomar en cuenta para el estudio.

Mediante este estudio el usuario puede observar paralelamente el comportamiento de las monedas fiat seleccionadas con respecto a una criptomoneda y en el caso de seleccionar datos normalizados apreciar con mas detalle si el comportamiento de al alza o a la baja es particular de una moneda fiat o generalizado para la criptomoneda entre otras cosas. El DataFrame resultante de este estudio se puede observar en la figura 3.24, este posee los datos de todas las monedas seleccionadas filtrados por el indicador seleccionado y su respectiva estampa de tiempo, en el caso de seleccionar la opción **Normalizar** el resultado sería el de la figura 3.25 es cual es análogo en forma pero construido a base de datos normalizados.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda1_indicador	Datos filtrados según indicador de la primera moneda seleccionada
Moneda2_indicador	Datos filtrados según indicador de la segunda moneda seleccionada
...	
MonedaN_indicador	Datos filtrados según indicador de la enésima moneda seleccionada

Figura 3.24: Formato de resultado para Varias monedas.

De manera que ésta sección de comparación tiene el potencial de servirnos tanto para comparar las cotizaciones de una criptomoneda con respecto a varias monedas fiat como también el uso de

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda1_indicador_norm	Datos filtrados y normalizados según indicador de la primera moneda seleccionada
Moneda2_indicador_norm	Datos filtrados y normalizados según indicador de la segunda moneda seleccionada
...	
MonedaN_indicador_norm	Datos filtrados y normalizados según indicador de la enésima moneda seleccionada

Figura 3.25: Formato de resultado para Varias monedas normalizadas.

una criptomoneda como base para la comparación de las monedas fíat entre sí.

3.2.7. Estadística

En cuanto a estudios de tipo estadístico nuestro prototipo cuenta con dos opciones a escoger en la figura de lista, para el cual posteriormente el usuario deberá seleccionar una moneda fíat y un indicador. Los dos tipos de estudios en esta área son:



Figura 3.26: Estudio Moneda + Trading del prototipo.

3.2.7.1. Moneda + Trading

Este estudio permite aplicar los estadísticos clásicos de trading del tipo serial estudiados en la sección 2.2.3.3 a los datos de la sesión, de aquí el nombre del estudio. Se nos permitirá seleccionar una moneda mas un indicador y un estadístico para el cual posteriormente deberemos introducir los valores de las ventanas m_1 , m_2 y m_3 que el estadístico necesite, de manera que dependiendo del estadístico en cuestión, se desplegarán de cero a tres `spinbox` para indicar los tamaños de las ventanas, como se puede observar en la figura 3.26.

El cálculo de los estadísticos mismos no se ha realizado mediante el llamado directo a paquetes ni librerías, pero si la implementación de funciones matemáticas ya existentes para cálculos menores, de esta manera gran parte de la autoría de las funciones que como tal calculan los estadísticos es propia del desarrollo de nuestro prototipo, dejándolos con un buen margen de manipulación posible. A continuación la lista detallada de estadísticos disponibles para este tipo de estudio:

Media Móvil Simple, $SMA(m_1)$: es de los estadísticos mas básicos, se suele usar para reducir ruido en los gráficos o comparar señales lentas y rápidas, para este estudio indicaremos una moneda, un indicador y el tamaño de la ventana para la SMA como se vio en la sección 2.2.3.3 apartado MA Simple. El resultado de este estudio es un DataFrame el cual se puede observar en la figura 3.27 y contiene la estampa de tiempo de los datos, los datos originales filtrados por un indicador y los datos resultantes de aplicar a estos datos una SMA de ventana m_1 .

Columna	Descripción
Date	Estampa de tiempo de los datos..
Moneda_indicador	Datos filtrados según moneda e indicador
$SMA(m_1)$	SMA de ventana m_1 de Moneda_indicador

Figura 3.27: Formato de resultado para Media Móvil Simple.

Media Móvil Exponencial, $EMA(m_1)$: de manera análoga a la SMA, este estadístico tiene la peculiaridad que los datos dentro de la ventana de estudio no poseen el mismo peso, ya que este puede ser modificado por un factor α , en este caso nuestro software posee un valor convencional. Este factor puede otorgar mayor o menor relevancia a los datos mas recientes, lo que se traduce en un estadístico que puede seguir de cerca el comportamiento del objeto de estudio como se explicó en la sección 2.2.3.3 apartado MA Exponencial. El resultado en

forma de Dataframe se puede observar en la figura 3.28 y contiene la estampa de tiempo de los datos, los datos originales filtrados por un indicador y los datos resultantes de aplicar a estos datos una EMA de ventana m_1 .

Columna	Descripción
Date	Estampa de tiempo de los datos..
Moneda_indicador	Datos filtrados según moneda e indicador
EMA(m_1)	EMA de ventana m_1 sobre Moneda_indicador

Figura 3.28: Formato de resultado para Media Móvil Exponencial.

Fibonacci, Fibo(f): el estadístico Retrocesos de Fibonacci es muy común y ampliamente usado en el ámbito del trading, este es sensible con respecto al rango de datos que tomemos en cuenta, ya que se basa en el máximo y mínimo en los datos para generar niveles los cuales pueden ser de resistencia o soporte, más detalles en la sección 2.2.3.3 apartado Retrocesos de Fibonacci.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicador	Datos filtrados según moneda e indicador
R_XXX %	XXX % con respecto al valor máximo (100 %) y mínimo (0 %) de Moneda_indicador
...	
R_XXX %	Análogo al anterior pero con distinto valor de XXX %

Figura 3.29: Formato de resultado para Retrocesos de Fibonacci.

La selección se realiza mediante casillas de verificación, en este caso los niveles disponibles para seleccionar son: {38.2, 50, 61.8, 23.8, 76.8, 132.8, 161.8}. El DataFrame resultante de este estudio se puede observar en la figura 3.29 y contiene la estampa de tiempo de los datos, los datos originales filtrados por un indicador y los valores para los retrocesos seleccionados.

Bandas Bollinger, BB(m_1, m_2): este es un estadístico bastante usado con el fin de percibir cuando el objeto de estudio está sobre vendido o sobre comprado mediante la generación de dos líneas basadas en la desviación estándar de los datos dentro de la ventana de estudio. Para llevar a cabo este estudio debemos indicar el tamaño de la ventana de estudio (m_1) y la

tolerancia (m_2) también nombrada comúnmente **b** como se explicó detalladamente en la sección 2.2.3.3 apartado Bandas de Bollinger. En la figura 3.30 podemos observar el DataFrame resultante de este estudio que contiene la estampa de tiempo de los datos, los datos originales filtrados por moneda e indicador, línea media, desviación estándar simple y multiplicada, banda inferior y banda superior.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicador	Datos filtrados según moneda e indicador
LM	Línea media de Moneda_indicador con ventana $m = 3$
STD	Desviación estándar de Moneda_indicador con ventana m_1
STD(b)	STD multiplicada por b (m_2)
BI	Banda inferior
BS	Banda superior

Figura 3.30: Formato de resultado para Bandas de Bollinger.

Convergencia Divergencia, MACD(m_1, m_2, m_3) : también denominado Media Móvil de Convergencia y Divergencia, es un estadístico que busca revelar cambios en la fuerza, dirección e impulso de una tendencia mediante la comparación de dos EMA de ventanas distintas m_1 y m_2 , usualmente una del doble de valor que la otra y resaltando la diferencia entre la señal de m_3 y el resultado de la resta de las dos EMA, más detalles en la sección 2.2.3.3 apartado MA de Convergencia Divergencia. En la figura 3.31 podemos observar el DataFrame resultante de este estudio el cual posee la estampa de tiempo de los datos filtrados según moneda e indicador y los mismos, una EMA de ventana m_1 y otra de m_2 , las líneas MACD y señal y el Dif.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicador	Datos filtrados según moneda e indicador
EMA(m_1)	EMA de Moneda_indicador con ventana m_1
EMA(m_2)	EMA de Moneda_indicador con ventana m_2
MACD	Diferencia de EMA(m_1) y EMA(m_2)
LS(m_3)	Línea señal de MACD con ventana m_3
Dif	Diferencia entre la MACD y la línea señal

Figura 3.31: Formato de resultado para Media Móvil de Convergencia Divergencia

Índice de Fuerza Relativa, RSI(m_1): es un estadístico del tipo oscilador que busca representar

la fuerza del precio mediante la comparación de los movimientos al alza o baja de los precios de cierre. Para el cálculo del RSI solo debemos indicar el tamaño de la ventana a tomar en cuenta, a pesar de que el cálculo original tomar en cuenta sólo valores de cierre, podríamos aplicarlo a cualquier serie de datos. Para mas información sección 2.2.3.3 apartado Índice de fuerza relativa. El DataFrame resultante de este estudio lo podemos observar en la figura 3.32 y este posee la estampa de tiempo de los datos y los datos filtrados según moneda e indicador, ganancia y perdida promedio, su cociente y el valor porcentual del RSI.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicador	Datos filtrados según moneda e indicador
AG	Ganancia promedio de Moneda_indicador
AL	Perdida promedio de Moneda_indicador
RS	Cociente de AG y AL.
RSI	Valor porcentual del RS

Figura 3.32: Formato de resultado para Índice de Fuerza Relativa.

Oscilador Estocástico, $K(m_1)$: este es un estadístico del tipo oscilador el cual se construye tomando en cuenta el máximo y mínimo valor de una serie de datos en una ventana específica a lo largo de la misma, más detalles en la sección 2.2.3.3 apartado Oscilador estocástico. El DataFrame resultante de este estudio se puede observar en la figura 3.33, este posee la estampa de tiempo de los datos y los datos mismos filtrados según moneda e indicador, el alto y bajo del período, el oscilador rápido y el lento.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicador	Datos filtrados según moneda e indicador
PH	Alto del Período de Moneda_indicador con ventana m_1
PL	Bajo del Período de Moneda_indicador con ventana m_1
K	Oscilador estocástico rápido
D	Oscilador estocástico lento con $m = 3$

Figura 3.33: Formato de resultado para Oscilador Estocástico.

Índice Direccional Promedio , $ADX(m)$: es un estadístico del tipo oscilador creado con el objetivo de medir la fuerza de una tendencia. Se realiza el cálculo mediante el suministro de

el tamaño de la ventana deseado m y datos OHLC, para más detalles 2.2.3.3 apartado Índice direccional promedio. El DataFrame resultante se puede observar en la figura 3.34 y esta conformado por la estampa de tiempo de los datos OHLC filtrados según la moneda especificada y los datos mismos, movimiento direccional positivo y negativo mas sus pares suavizados, rango verdadero y su par suavizado, indicador positivo y negativo, índice direccional y el índice direccional promedio.

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicadorO	Datos filtrados según moneda e indicador (Open)
Moneda_indicadorH	Datos filtrados según moneda e indicador (High)
Moneda_indicadorL	Datos filtrados según moneda e indicador (Low)
Moneda_indicadorC	Datos filtrados según moneda e indicador (Close)
PDM	Movimiento direccional positivo
NDM	Movimiento direccional negativo
TR	Rango verdadero
PDMS	PDM suavizado por una SMA de $m = 3$
NDMS	NDM suavizado por una SMA de $m = 3$
ATR	TR suavizado por una SMA de $m = 3$
PDI	Indicador direccional positivo
NDI	Indicador direccional negativo
DX	Índice direccional
ADX	Índice direccional promedio

Figura 3.34: Formato de resultado para Índice Direccional Promedio .

Nube Ichimoku (m_1, m_2, m_3): este es un estadístico que combina diversos estadísticos dentro de un único estudio. Este estudio busca proporcionar perspectivas sobre posibles zonas de soporte y resistencia de la cotización además de ser utilizado en algunos casos como instrumento de pronóstico con el fin de determinar tendencias futuras (para mas información ver la sección 2.2.3.3, apartado Nube Ichimoku). El DataFrame resultante de este estudio se observa en la figura 3.35 y esta conformado por la estampa de tiempo de los datos OHLC filtrados según la moneda especificada y los datos mismos, el alto y bajo para tres períodos de m_1 , m_2 y m_3 , las líneas de conversión, base y retraso mas las distancias adelantadas A y B.

Algunas de estas medidas de trading tendrán la posibilidad de usarse en simultaneo con otra medida de trading, las medidas con esta opción disponible son:

Columna	Descripción
Date	Estampa de tiempo de los datos.
Moneda_indicadorO	Datos filtrados según moneda e indicador (Open)
Moneda_indicadorH	Datos filtrados según moneda e indicador (High)
Moneda_indicadorL	Datos filtrados según moneda e indicador (Low)
Moneda_indicadorC	Datos filtrados según moneda e indicador (Close)
High(m_1)	Alto del Período de Moneda_indicadorH con ventana m_1
Low(m_1)	Bajo del Período de Moneda_indicadorL con ventana m_1
TS	Línea de conversión
High(m_2)	Alto del Período de Moneda_indicadorH con ventana m_2
Low(m_2)	Bajo del Período de Moneda_indicadorL con ventana m_2
KS	Línea Base
High(m_3)	Alto del Período de Moneda_indicadorH con ventana m_3
Low(m_3)	Bajo del Período de Moneda_indicadorL con ventana m_1
SSA	Distancia adelantada A con m_2
SSB	Distancia adelantada B con m_3
CS	Línea de retraso con m_2

Figura 3.35: Formato de resultado para Nube Ichimoku.

- Media móvil simple
- Media móvil exponencial.
- Bandas de Bolinger.
- Fibonacci.

3.2.7.2. Medidas de tendencia central y dispersión.

Para este tipo de estudio se cuenta con una lista de estadísticos con casillas de verificación para la selección de las medidas de tendencia central que deseemos obtener además de la opción “todos” la cual como su nombre lo dice seleccionará todas las medidas disponibles como se puede observar en la figura 3.36 además se deberá seleccionar una moneda y un indicador. Como resultado de este tipo de estudio se generará una tabla de datos con los nombres de los estadísticos y sus resultados respectivamente. Las medidas disponibles son:

Medidas de tendencia central (sección 2.2.3.1):

- Media aritmética (apartado Media aritmética)
- Media Armónica (apartado Media Armónica)
- Media Geométrica (apartado Media Geométrica)
- Moda (apartado Moda)
- Mediana (apartado Mediana)

Medidas de dispersión (sección 2.2.3.2):

- Rango (apartado Rango)
- Varianza (apartado Varianza)
- Desviación Estándar (apartado Desviación Estándar)
- Coeficiente de variación (apartado Coeficiente de Variación)



Figura 3.36: Estudio Tendencia Central.

En el caso de seleccionar la opción *todos*, el resultante será un DataFrame como el mostrado en la figura 3.37, el cual no cuenta con los datos usados para realizar el estudio sino solo los resultados de las medidas, todas estas se hacen en base a los datos filtrados según la moneda e indicador indicados.

De manera que en el área de estadística poseemos tanto estudios clásicos del trading como medidas de tendencia central y dispersión, las cuales al ser empleadas de manera conjunta proporcionan una visión integral de los datos que se quieran estudiar.

3.2.8. Inferencia

En cuanto a estudios en el campo de la inferencia, nuestro prototipo ofrece la generación de datos mediante el uso de medias móviles sobre los datos, este cálculo se prolongará en el tiempo generando así datos mas allá del instante final indicado en la descarga de datos.

Columna	Descripción
AM	Media aritmética.
HM	Media armónica.
GM	Media geométrica.
Med	Mediana.
Mod	Moda.
Rrng	Rango.
s2	Varianza.
s	Desviación estándar.
d	Coefficiente de variación.

Figura 3.37: Formato de resultado para Tendencia Central

De tal manera que mediante el uso de medias móviles como estimador de la serie de datos, inferir los siguientes movimientos del mercado a corto plazo en relación a los datos empleados para el cálculo, esto haciendo referencia a que la generación de puntos será reducida y acorde al intervalo de tiempo presente en los datos.

Al realizar un estudio de este estilo se hará la elección de la moneda, el indicador que desea tomar en cuenta para el estudio y el tipo de media móvil. En el caso de seleccionar una SMA no será necesario alguna especificación extra pero de seleccionar una EMA se permitirá la selección del valor de α personalizado o el uso de una casilla de verificación la cual indicará el uso del valor estándar para α visto en la sección 2.2.3.3 apartado MA Exponencial. Adicional a esto el usuario podrá seleccionar un estadístico de los usados en trading para añadirlo al estudio como se observa en la figura 3.38.

En cuanto a los resultados, este estudio genera un DataFrame análogo a los vistos en la sección 3.2.7.1 con el extra de que la serie de datos ahora incluye los datos inferidos.

3.2.9. Simulación

Como se puede observar en la figura 3.39 para realizar una simulación primero el usuario debe seleccionar el tipo de simulación que desea realizar, basados en la teoría de la sección 2.2.6.4 hemos creado dos métodos de simulación para los que debemos indicar dos subperiodos, para los que el usuario debe seleccionar tres instantes mediante el despliegue de un calendario para especificar: el año, mes, día, hora y minuto respectivo. Los primeros dos instantes, llamados **a** y **b** delimitan el período de datos a tener en cuenta para la simulación, éstos deben estar dentro del período de estudio activo en la sesión. El tercer instante llamado **c** determina el final de la simulación ya que ésta tendrá lugar desde el instante **b** hasta el instante **c**.



Figura 3.38: Estudio Inferencia Basada en Medias Móviles.

Los dos métodos de simulación disponibles son Distri y Maillard.

3.2.9.1. Método Distri

Este método se basa en filtrar los datos de la sesión en un período de **a** hasta **b** y replicar de estos datos en un período que va desde **b** hasta **c**. Además de moneda e indicador como se observa en la figura 3.39 para realizar la simulación el usuario debe seleccionar:

- Tipo de distribución de probabilidad, seleccionada de un cuadro de lista que posee las distribuciones Normal, Beta, Exponencial, Triangular, Uniforme, Semicircular, Gamma. En el caso de que la distribución necesite parámetros extra, se despliegan los respectivos spinbox para suministrarlos.
- Tipo de muestra, hacen referencia a cómo se tomará muestra de los datos del período $[a, b]$ a replicar. Las opciones disponibles son: Única, Clusters, Clusters aleatorios, Cluster aleatorio.
- Número de simulaciones, que establece la cantidad de realizaciones que se van a simular. En el resultado se muestra cada simulación más el promedio de todas las simulaciones realizadas.



Figura 3.39: Estudio Simulación mediante el Método Distri

Una vez se suministren las especificaciones, veamos a detalle como se realiza esta simulación con el método que hemos denominado **Distri**:

1. Definimos que archivo de datos de la sesión usaremos para la simulación. Como primer filtro tenemos la moneda y el indicador seleccionado y como segundo filtro tenemos el período $[a, b]$ seleccionado. Para la simulación usamos datos que posean intervalos uniformes entre si, de manera si la fuente seleccionada cuenta con distintos intervalos descartaremos los archivos que no abarquen de manera total al subperíodo $[a, b]$. Una vez seleccionado el archivo que contenga $[a, b]$ pasamos al siguiente paso.
2. Para definir cuantos registros vamos a simular, tomamos el período $[b, c]$ y lo dividimos entre el tipo de intervalo que posee el archivo que usaremos.
3. Procedemos a realizar la toma de muestras, el tamaño de estas será según el número de registros a simular y funciona de la siguiente manera:

Única: este tipo de muestra aplica a todos los datos $[a, b]$ el tipo de distribución de probabilidad y procede a la selección de los datos a replicar.

Clusters: como su nombre lo indica procede a dividir el período $[a, b]$ en tantos clusters como se haya indicado para luego aplicar el tipo de distribución de

probabilidad a cada cluster por separado y proceder a la selección de los datos a replicar. Para efectuar la selección de los datos a replicar, el paso por los clusters se hará de manera ordenada, de tal manera que la simulación se realizará por clusters en orden cronológico.

Clusters aleatorios: este muestreo se realiza de manera análoga al anterior con la peculiaridad que luego de dividir el período $[a, b]$ en clusters y aplicarles el tipo de distribución especificado, el orden de la selección del cluster a replicar es aleatorio, añadiendo así una capa mas de aleatoriedad a la simulación.

Cluster aleatorio: de manera análoga a los anteriores se divide el período $[a, b]$ en tantos clusters como se haya indicado y luego se procede a la selección de un único cluster al cual se le aplicará la distribución de probabilidad y se seleccionarán todos los datos a replicar.

4. Aplicamos la distribución de probabilidad a los datos, esto implica lo siguiente:

- Dado el número de registros dentro del cluster, generamos una muestra aleatoria con el mismo tamaño que siga la distribución de probabilidad indicada para un intervalo definido.
- Mediante una transformación lineal de intervalos realizamos la transformación de los valores obtenidos en el paso anterior al intervalo que va desde cero hasta n , siendo n el número de registro dentro de una muestra. De esta manera lo que buscamos es aplicar el comportamiento de una distribución de probabilidad en un intervalo al intervalo discreto compuesto por los registros en orden cronológico, en el caso de poseer valores continuos hemos aplicado redondeo.

5. Simultáneamente se van generando los datos de la simulación, los vamos añadiendo al arreglo de datos, de manera que el DataFrame resultante poseerá tantas columnas mas una como simulaciones haya indicado el usuario, ya que se añade el promedio de todas, cada una con registros para el intervalo $[a, c]$.

En la figura 3.41 podemos observar un DataFrame resultante de este estudio, para este caso *Simu* se sustituye por *Distri*.

3.2.9.2. Método Maillard

Este método se basa en filtra los datos de la sesión en un período de **a** hasta **b** y replicar de estos datos de en un período que va desde **b** hasta **c**. Además de moneda e indicador como se observa



Figura 3.40: Estudio Simulación mediante el Método MAillard.

en la figura 3.40 para realizar la simulación el usuario deberá seleccionar:

El segundo método de simulación lo hemos llamado **Método MAillard** de simulación el cual consiste en la replicación del comportamiento de los datos del período $[a, b]$ y no de los datos en sí como el método anterior. Además de moneda e indicador para realizar la simulación el usuario deberá seleccionar:

- Tipo de replicación: nuestro método posee dos forma de abordar la selección de los datos para la generación de la simulación las cuales son:

Cronológicamente inverso: este método consiste tomar los datos y replicar el comportamiento de manera ordena pero cronológicamente inverso.

Aleatorio: de esta manera los datos a replicar se seleccionarán de manera aleatoria a lo largo del período $[a, b]$.

- Diferencial, como bien se ha expuesto este método de simulación se basa en replicar el comportamiento, por lo tanto de los datos del período $[a, b]$ hemos extraído su comportamiento de dos maneras:

Deltas: este comportamiento refleja la diferencia del valor actual con el valor anterior.

Porcentaje: refleja el comportamiento del valor actual con el valor anterior de manera porcentual.

Ambos comportamientos no son tomados como valores absolutos por lo que mantienen su valor negativo o positivo

- Valor de **p**, dado que este método de simulación esta centrado en una distribución de probabilidad de Bernoulli, es necesario que se indique el valor de la probabilidad de éxito, en este caso un éxito se traduce en simular de manera exacta el comportamiento que tuvieron los datos en el pasado y un fracaso se traduce en replicar el comportamiento de manera opuesta.
- Número de simulaciones, además al resultado se añadirá el promedio de todas las simulaciones realizadas.

Veamos a continuación los pasos a seguir para realizar una simulación de este tipo:

1. Definimos que archivo de datos de la sesión usaremos para la simulación de igual manera como se hizo en el caso de la simulación anterior.
2. Para definir cuantos registros vamos a simular, igualmente a la simulación anterior tomamos el período $[b, c]$ y lo dividimos entre el tipo de intervalo que posee el archivo que usaremos.
3. Generamos una muestra aleatoria para distribución Bernoulli con valor de probabilidad de éxito **p** del mismo tamaño que datos a simular.
4. Procedemos a generar la simulación tomando en cuenta el tipo de replicación seleccionado, los datos del diferencial a replicar y si el experimento de Bernoulli para cada registro es un éxito o un fracaso.
5. Simultáneamente se van generando los datos de la simulación, los vamos añadiendo al arreglo de datos, de manera que el DataFrame resultante poseerá tantas columnas mas una como simulaciones haya indicado el usuario, ya que se añade el promedio de todas, cada una con registros para el intervalo $[a, c]$.

En la figura 3.41 podemos observar un DataFrame resultante de este estudio, para este caso *Simu* se sustituye por *Bernu*

Luego de que el usuario haya suministrado las especificaciones de un estudio en particular este debe proceder a realizar el cálculo respectivo. El responsable de esto será el botón **Calcular** que discutiremos a continuación.

Columna	Descripción
Date	Estampa de tiempo de los datos
Simu(1)	Datos originales $[a, b]$ mas simulación $[b, c]$
...	
Simu(n)	Datos originales $[a, b]$ mas simulación $[b, c]$
Simu	Datos originales $[a, b]$ mas promedio de las simulaciones $[b, c]$

Figura 3.41: Formato de resultado para Simulación Método Distri y MAillard.

3.2.10. Calcular y cancelar.

Como se puede observar en las figuras dedicadas a los estudios disponibles para efectuar sobre los datos, la GUI cuenta con un dos botones **calcular** y **cancelar** en la parte inferior de las especificaciones de cada estudio, es a través del botón calcular como su nombre lo indica que se realizan las tareas pertinentes al subprograma **Cálculo** explicado en la sección de diseño 3.1.2.2, esto con respecto al área y estudio en el que estemos posicionados, el resultado de este subprograma es la generación del DataFrame temporal con los resultados del estudio y especificaciones en cuestión. Por otro lado el botón cancelar se implementó con la finalidad de retornar los widgets a la última especificación de estudio para la que se calcularon y obtuvieron resultados, esta acción es independiente para cada pestaña.



Figura 3.42: Cálculo de Simulación Método MAillard.

El siguiente paso es el de generar reportes, estos botones se habilitaran de manera independiente

para cada estudio una vez hayamos realizado como mínimo un cálculo, como se puede observar en la figura 3.42.

3.2.11. Generación de Reportes.

En lo que respecta al **SelReporte** del prototipo, éste está creado según las especificaciones descritas en la sección 3.1.3 y cuenta con tres opciones de reportes. La selección del tipo de reporte se realiza mediante botones dedicados para cada uno en específico, como se puede observar en la figura 3.42. Al hacer uso de ellos, se realiza el llamado al subprograma dedicado al reporte en cuestión y se le suministran los datos del DataFrame temporal provenientes del subprograma **Cálculo** y éste generará el reporte correspondiente.

Los tres reportes disponibles en nuestro prototipo son:

1. **Archivo csv.**

El archivo resultante al generar este reporte contendrá en su totalidad el DataFrame resultante del estudio en cuestión, en otras palabras el archivo resultante del subprograma **calcular** exportado en formato **csv**.

2. **Archivo xlsx.**

Este reporte se realizará de manera análoga al **Archivo csv**, pero ahora generando un **Archivo xlsx** el cual cuenta con metadatos extra, como lo es el tipo de dato en cada columna (fecha o numérico) y el índice de la tabla especificado.

3. **Gráfico.**

La generación de gráficos en base a los datos es la manera mas común de representar datos, en nuestro prototipo estos gráficos se realizarán en ventanas emergentes, de manera que se podrán generar cuantos gráficos se quiera de estudios y combinaciones de especificaciones distintas y así ir comparándolos lado a lado o guardarlos y por ejemplo realizar trazos sobre los mismos mediante otro programa. Además estas ventanas emergentes contarán con una barra de herramientas la cual se puede ver en la figura 3.43, estas herramientas son (explicadas de izquierda a derecha):



Figura 3.43: Barra de herramientas del gráfico.

- a) Volver al gráfico original.
- b) Volver a la vista anterior.
- c) Avanzar a la siguiente vista.
- d) Movimientos:
 - Con el botón principal del ratón (*left-click*) arrastramos el gráfico y la escala de los ejes se mantiene.
 - Con el botón secundario del ratón (*right-click*) al arrastrar el gráfico ajustamos la escala de los ejes de manera independiente.
 - Mediante la tecla **ctrl** al arrastrar el gráfico vamos ajustamos la escala de los ejes pero manteniendo la relación entre ellos (manteniendo el aspecto de forma).
- e) Realiza un zoom a un rectángulo indicado.
- f) Permite manipular la configuración del gráfico dentro de la ventana, margen superior, inferior y laterales.
- g) Guardar el gráfico con la visualización actual, al hacer uso de esta herramienta el directorio y nombre serán por omisión, el directorio de la sesión y el nombre respectivo del reporte.

En la figura 3.44 y 3.45 se puede observar los reportes tanto de archivo de datos que incluyen al **Archivo csv** y el **Archivo.xlsx** como tambien la generacion del **Gráfico** respectivamente, ambos resultantes de las especificaciones de la figura 3.42, una simulación mediante el **Método MAillard** para datos del par **BTCUSD** y filtrados según el indicador **BAVAverage**, valor de **p=0.64** y usando las **deltas** de las cotizaciones para generar una simulación en orden **cronológicamente inverso** y como sub-período $[a, b, c]$ hemos escogido $[15/04/2022, 03/09/2022, 03/10/2022]$. En cuanto al archivo de datos este cumple las especificaciones de la figura 3.41 y en cuanto al gráfico este cuenta con el los datos originales, el sub-período $[a, b, c]$ delimitado con líneas verticales, todas las simulaciones en distintos colores con una opacidad del 30 % y el promedio de todas las simulaciones en color rojo sin opacidad.

	A	B	C	Z	AA	AD	AE
1	Date	Bernu_0	Bernu_1	Bernu_24	Bernu_25	Bernu_28	Bernu_avg
2	03/09/22 00:00:00	19992.72	19992.72	19992.72	19992.72	19992.72	19992.71999
3	04/09/22 00:00:00	19992.0	19992.0	19992.0	19992.0	19992.0	19992.0
4	05/09/22 00:00:00	19909.39	19909.39	20074.61	19909.39	19909.39	19954.96793
5	06/09/22 00:00:00	20562.54	20562.54	20727.76	20562.54	19256.23	19977.49034
6	07/09/22 00:00:00	19179.89	19179.89	19345.11	19179.89	17873.58	20120.52310
7	08/09/22 00:00:00	19581.4	19581.4	19746.62	19581.4	18275.1	20162.05862
8	09/09/22 00:00:00	19821.01	19341.78	19507.00	19821.01	18514.71	20203.37241
9	10/09/22 00:00:00	21376.94	17785.85	21062.93	21376.94	20070.64	19613.19206
10	11/09/22 00:00:00	20911.96	17320.86	20597.94	21841.93	19605.65	19533.02137
11	12/09/22 00:00:00	20946.33	17286.49	20632.31	21807.56	19571.28	19522.35482

Figura 3.44: Archivo de Datos de una simulación según el Método MAillard.

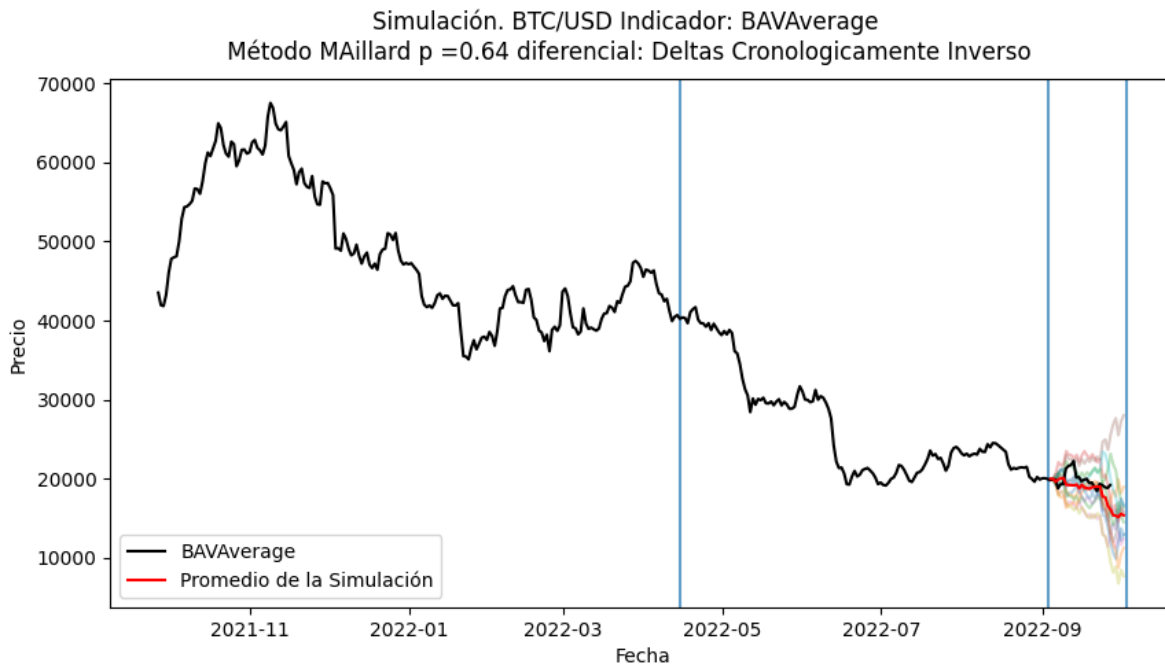


Figura 3.45: Gráfico de una simulación según el Método MAillard.

Además de este ejemplo de reporte gráfico, las figuras de la sección 2.2.3.3 son reportes generados mediante nuestro prototipo como resultado de estudios de Moneda + Trading.

3.2.11.1. Formato para los nombres de los reportes.

Los archivos generados al emitir reportes se almacenarán en la carpeta pertinente a la sesión y llevarán por nombre las especificaciones del mismo y la extensión acorde al tipo de reporte bien sea `csv`, `xlsx` o `png` si se desea almacenar el gráfico. De acuerdo a nuestras cuatro áreas de estudio veamos el formato para nombrar los reportes:

Comparación

- **Moneda vs Moneda**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), criptomoneda (CCC), moneda principal (M01) y moneda Secundaria (M02). Este formato y un ejemplo que representa la comparación del USD con el XAU usando el indicador BAVAverage y la criptomoneda BTC se observa en la figura 3.46.
- **Varias monedas**, poseerán nombres de la siguiente forma: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), criptomoneda (CCC) y en el caso de seleccionar **Normalizar** se añadirá “norm”, primer moneda (M01), segunda

moneda (M02) hasta la enésima moneda (MNN). Este formato y un ejemplo que representa la comparación usando la criptomoneda BTC, el indicador BAVAverage y las monedas fiat USD, VES, JPY, BGP, EUR, XAU, XAG tanto normalizadas como con los datos originales se observa en la figura 3.46.

Moneda vs Moneda:	EEE_Indicador_CCC_M01_M02
	CMP_BAVAverage_BTC_USD_XAU
Varias monedas:	EEE_Indicador_CCC_M01_M02 ... _MNN
	CMP_BAVAverage_BTC_USD_VES_JPY_GBP_EUR_XAU_XAG
normalizadas:	EEE_Indicador_CCC_norm_M01_M02 ... _MNN
	CMP_BAVAverage_BTC_norm_USD_VES_JPY_GBP_EUR_XAU_XAG

Figura 3.46: Formato de nombre para reportes del área Comparación.

Estadística

- **Tendencia central**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), par fiat-criptomoneda (CCCMMM), estadístico central 01 (C01), estadístico central 02 hasta el enésimo estadístico central seleccionado (CNN). Este formato y un ejemplo que representa el estudio estadístico usando el indicador BAVAverage para el par BTCUSD indicando la media geométrica, media aritmética, rango, varianza y desviación estándar se observa en la figura. 3.47.
- **Moneda + Trading**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe(Indicador), par criptomoneda-fiat (CCCMMM), estadístico de trading (Trading), ventanas m_1 , m_2 y m_3 respectivas del estadístico (m1, m2 y m3). Este formato y un ejemplo que representa el estudio estadístico usando el indicador BAVAverage para el par BTCUSD y representando una media móvil de convergencia divergencia con $m_1 = 12$ $m_1 = 9$ $m_1 = 6$ se observa en la figura 3.47.
- **Moneda + 2 Trading**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), par fiat-criptomoneda (CCCMMM), primer estadístico de trading (Trading1), ventanas m_1 , m_2 y m_3 respectivas del estadístico (m11, m12 y m13), segundo estadístico de trading (Trading2), ventanas m_1 , m_2 y m_3 respectivas del estadístico (m21, m22 y

m23). Este formato y un ejemplo que representa el estudio estadístico usando el indicador BAVAverage para el par BTCUSD y representando una media móvil simple de $m_{11} = 12$ y unas bandas de bollinger con $m_{21} = 15$ y $m_{22} = 2$ se observa en la figura 3.47.

Tendencia central:	EEE_Indicador_CCCMMM_C01_C02 ... _CNN EST_BAVAverage_BTCUSD_MG_MA_rng_s_s2
Moneda + Trading:	EEE_Indicador_CCCMMM_Trading_m1_m2_m3 EST_BAVAverage_BTCUSD_MACD_12_9_6
Moneda + 2 Trading:	EEE_Indicador_CCCMMM_Trading1_m11_m12_m13_Trading2_m21_m22_m23 EST_BAVAverage_BTCUSD_SMA_12_BB_15_2

Figura 3.47: Formato de nombre para reportes del área Estadística

Inferencia

- **Inferencia con SMA**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), par fiat-criptomoneda (CCCM), estimador usado (SMA), ventana del estimador usado (m11), estadístico de trading (Trading), ventanas m_1 , m_2 respectivas del estadístico (m21 y m22). Este formato y un ejemplo que representa el estudio de inferencia usando el indicador BAVAverage para el par BTCUSD y usando una media móvil simple de ventana $m_{11} = 12$ como estimador y agregando al estudio unas bandas de bollinger de $m_{21} = 15$ y $m_{22} = 2$ se observa en la figura 3.48.
- **Inferencia con EMA**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), par fiat-criptomoneda (CCCM), estimador usado (EMA), ventana del estimador usado (m11), valor de α usado por el estimador este posee un valor de 0 a 1 pero lo hemos multiplicado por 100 (NNN), estadístico de trading (Trading), ventanas m_1 , m_2 respectivas del estadístico (m21 y m22). Este formato y un ejemplo que representa el estudio de inferencia usando el indicador BAVAverage para el par BTCUSD y usando una media móvil exponencial de ventana $m_{11} = 12$ con $NNN = 078$ lo que se traduce a un valor de $\alpha = 0.78$ como estimador y agregando al estudio unas bandas de bollinger de $m_{21} = 15$ y $m_{22} = 2$ se observa en la figura 3.48.

Inferencia con SMA:	EEE_Indicador_CCCMMM_SMA_m11_Trading_m21_m22 INF_BAVAverage_BTCUSD_SMA_12_BB_15_2
Inferencia con EMA:	EEE_Indicador_CCCMMM_EMA_m11_NNN_Trading_m21_m22 INF_BAVAverage_BTCUSD_EMA_12_078_BB_15_2

Figura 3.48: Formato de nombre para reportes del área Inferencia

Simulación

- **Simulación método Distri**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe(Indicador), par fiat-criptomonedas (CCCM), método de simulación usado (SSS), tipo de distribución usado la selección de datos (dist), valores a y b particulares de la distribución seleccionada (a_b), tipo de muestreo utilizado (muestreo), número de simulaciones (NNN) instante inicial y final de los datos usados para la simulación (Inical_Ifinal). Este formato y un ejemplo donde se representa una simulación para el par BTCUSD usando el indicador BAVAverage, una distribución normal un muestreo por clusters, 15 simulaciones e instante inicial y final 20220101 20220801 respectivamente se observa en la figura 3.49.
- **Simulación método MAillard**, el nombre de estos archivos se conformará de las siguientes partes: especificación del área de estudio (EEE), fuente y columna del Dataframe (Indicador), par fiat-criptomonedas (CCCM), método de simulación usado (SSS), tipo de deltas usados para la simulación (delta), orden para la recreación de eventos (tipo), número de simulaciones (NNN) instante inicial y final de los datos usados para la simulación (Inical_Ifinal). Este formato y un ejemplo donde se representa una simulación para el par BTCUSD usando el indicador BAVAverage, delta de los datos como comportamiento a simular, un muestreo aleatorio, 28 simulaciones e instante inicial y final 20220101 20220801 respectivamente se observa en la figura 3.49.

Método Distri:	EEE_Indicador_CCCMMM_SSS_dist_a_b_muestreo_NNN_Inical_Ifinal SIM_BAVAverage_BTCUSD_Dist_nrm_clusters_15_20220101_20220801
Método MAillard:	EEE_Indicador_CCCMMM_SSS_delta_tipo_NNN_Inical_Ifinal SIM_BAVAverage_BTCUSD_MAi_deltas_random_28_20220101_20220801

Figura 3.49: Formato de nombre para reportes del área Simulación

Ya con los nombres debidamente especificados, solo bastará hacer uso del botón del reporte deseado y este creará el archivo con la extensión indicada. Culminando así lo que respecta a el uso de nuestro software.

Con respecto a la estructura en cuanto a archivos se refiere de nuestro prototipo de software **MAillardFinance**, veamos lo siguiente:

3.2.12. Archivos del Prototipo.

Nuestro prototipo necesita la existencia de dos archivos para su distribución: **MAillardFinance** (comprimido y ejecutable `python3`) y `setup.sh` (script `bash`).

- **MAillardFinance**: archivo comprimido que contiene todos los scripts y recursos necesarios para la ejecución del software, su contenido se puede observar en la figura 3.50.

Nombre	Descripción
<code>__main__.py</code>	Enlace simbólico
<code>MAillardFinance.py</code>	GUI
<code>GesDescargas.sh</code>	Gestor de descargas.
<code>Extractor_CMC.sh</code>	Extractor del sitio CMC
<code>Extractor_BAV.sh</code>	Extractor del sitio BAV
<code>Colector_fiats.sh</code>	Extractor de lista de fiats disponibles
<code>requirements.txt</code>	Lista de librerías usadas por <code>python3</code>
<code>MAillardFinance_icon.png</code>	Icono del software
<code>MAillardFinance_xlsx.png</code>	Imagen para widget
<code>MAillardFinance_csv.png</code>	Imagen para widget
<code>MAillardFinance_graph.png</code>	Imagen para widget
<code>MAillardFinance_logo.png</code>	Logo de ULA Física

Figura 3.50: Archivos necesarios para la ejecución del prototipo.

- `setup.sh` script el cual posee toda la especificación para verificación e instalación del software requerido por **MAillardFinance**.

Si bien se cuenta con los dos archivos necesarios para instalar y ejecutar nuestro software, veamos como es el proceso de instalación y ejecución.

3.2.13. Instalación y ejecución.

El script `setup.sh` es el encargado de realizar todo lo respectivo a la instalación de nuestro prototipo, su funcionamiento lo vemos en la figura 3.51. Es elección del usuario ejecutarlo y dejar las

especificaciones de instalación por omisión o manipular el archivo e introducir unas personalizadas. Dentro de estas especificaciones para realizar la instalación tenemos:

Instalación de software implementado : se cuenta con dos opciones: automatizado, donde se realiza la instalación de los programas necesarios de manera preestablecida y manual si el usuario desea instalarlos por sus propios medios. La opción por omisión es automatizado.

Ruta de instalación : por omisión `/usr/local/bin/MAillardFinance`

Ruta de trabajo : por omisión será `~/home/MAillardFinance/`

-
1. Copia el archivo comprimido en la ruta de instalación
 2. Verifica que exista cada programa necesario y de no existir alguno, procede a instalarlo
 3. Verifica e instala las librerías de `python3` especificadas en `requirements.txt` mediante `pip3`
 4. Genera un ejecutable del software
 5. Pone en marcha el ejecutable.
-

Figura 3.51: Tareas desempeñadas por el script `setup`

www.bdigital.ula.ve

Conclusiones

Se propone el diseño de un método para la minería de datos financieros, este método cuenta con tres fases para la minería y una interfaz gráfica de usuario que tiene como objetivo principal facilitar al usuario la utilización de los programas y subprogramas que poseen diversas alternativas a la hora de extraer y procesar los datos de intercambio de criptomonedas-dinero fiat.

Siguiendo el método propuesto se realizó el prototipo completamente funcional **MAillardFinance**, que implementa la extracción de datos vía web-API mediante `curl`, filtra y da formato a los datos mediante `grep`, `sed` y `awk`; realiza estudios y genera reportes mediante `python3`, todo esto implementado a través de una interfaz gráfica de usuario amigable. Además cada fase se realizó siguiendo el enfoque modular lo que da como resultado un prototipo escalable, es decir, al cual le podemos agregar más fuentes, más estudios o más reportes de manera sencilla.

Se agregaron en **MAillardFinance** los estudios convencionales y algunos estudios no ofrecidos por los sitios web de divulgación, como lo son la comparación tomando las criptomonedas como base, estadísticos de tendencia central, inferencia y simulación.

MAillardFinance es un prototipo dirigido a usuarios que posean algún conocimiento sobre criptomonedas, dinero fiat, metales preciosos y trading.

Bibliografía

- [1] Andrew Tar. Cointelegraph. ¿qué es prueba de trabajo o proof of work (pow)? <https://es.cointelegraph.com/explained/proof-of-work-explained>. Recuperado el 7 de Marzo 2022.
- [2] Paul Wackerow; Corwin Smith; Sam Richards; Victor Luna; Alex Ismodes and Alwin Stockinger. ETH.org. <https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/>. Recuperado el 7 de Marzo 2022.
- [3] Wayne Duggan. U.S.News. Bitcoin wallet definition. <https://money.usnews.com/investing/term/bitcoin-wallet>. Recuperado el 7 de Marzo 2022.
- [4] Dmitri Volhov. Binance Academy. Wallet. <https://academy.binance.com/en/glossary/wallet>. Recuperado el 7 de Marzo 2022.
- [5] Binance Academy. Proof of burn explained. <https://academy.binance.com/en/articles/proof-of-burn-explained>. Recuperado el 7 de Marzo 2022.
- [6] BITCOIN. Preguntas más frecuentes,general. <https://bitcoin.org/es/faq#general>. Recuperado el 29 de Febrero de 2022.
- [7] 2022 CoinMarketCap. Bitcoin price today, btc to usd live, marketcap and chart. <https://coinmarketcap.com/currencies/bitcoin/>. Recuperado el 7 de Marzo 2022.
- [8] <https://ethereum.org/es/>. Recuperado el 7 de Marzo 2022.
- [9] 2022 CoinMarketCap. Ethereum price today, eth to usd live, marketcap and chart. <https://coinmarketcap.com/currencies/ethereum/>. Recuperado el 7 de Marzo 2022.
- [10] <https://ethereum.org/es/what-is-ethereum/>. Recuperado el 7 de Marzo 2022.
- [11] Binance Academy. Token. <https://academy.binance.com/en/glossary/token>. Recuperado el 24 de Marzo 2022.

- [12] Grupo BBVA. ¿qué son las 'stablecoins' y para qué sirven? <https://www.bbva.com/es/que-son-las-stablecoins-y-para-que-sirven/>. Recuperado el 7 de Marzo 2022.
- [13] 2022 CoinMarketCap. Tether price today, usdt to usd live, marketcap and chart. <https://coinmarketcap.com/currencies/tether/>. Recuperado el 7 de Marzo 2022.
- [14] EOS. Eos network. <https://eos.io/>. Recuperado el 7 de Marzo 2022.
- [15] Tron. Tron network. <https://tron.network/>. Recuperado el 7 de Marzo 2022.
- [16] ALGORAND. Algorand. <https://www.algorand.com/>. Recuperado el 7 de Marzo 2022.
- [17] OMG. Omg network. <https://omg.network/>. Recuperado el 7 de Marzo 2022.
- [18] 2022 CoinMarketCap. Bnb price today, bnb to usd live, marketcap and chart. <https://coinmarketcap.com/currencies/bnb/>. Recuperado el 7 de Marzo 2022.
- [19] 2022 Binance Academy. ¿qué es bnb? <https://academy.binance.com/es/articles/what-is-bnb>. Recuperado el 7 de Marzo 2022.
- [20] 2022 Binance Academy. ¿qué es la quema automática de bnb? <https://academy.binance.com/es/articles/what-is-bnb-auto-burn>. Recuperado el 7 de Marzo 2022.
- [21] Binance. Binance chain la blockchain para intercambiar el mundo. <https://www.binance.com/es/blog/all/binance-chain-la-blockchain-para-intercambiar-el-mundo-304219301536473088>. Recuperado el 7 de Marzo 2022.
- [22] Binance. Presentamos bnb chain la evolución de binance smart chain. <https://www.binance.com/es/blog/ecosystem/presentamos-bnb-chain-la-evoluci%C3%B3n-de-binance-smart-chain-421499824684903436>. Recuperado el 7 de Marzo 2022.
- [23] Binance. Binance academy. <https://academy.binance.com/es>. Recuperado el 7 de Marzo 2022.
- [24] Trustwallet. Binance coin wallet. <https://trustwallet.com/es/binance-coin-wallet/>. Recuperado el 7 de Marzo 2022.
- [25] Binance Research. Binance research. <https://research.binance.com/>. Recuperado el 7 de Marzo 2022.

- [26] Akhilesh Ganti. Investopedia. What is an angel investor? <https://www.investopedia.com/terms/a/angelinvestor.asp>. Recuperado el 7 de Marzo 2022.
- [27] Wikipedia. Creador del mercado. https://es.wikipedia.org/wiki/Creador_del_mercado. Recuperado el 7 de Marzo 2022.
- [28] Robinhood. Commission-free stock trading & investing app. <https://robinhood.com/us/en/>. Recuperado el 7 de Marzo 2022.
- [29] eToro. etoro: la red de inversión y trading social líder en el mundo. <https://www.etoro.com/es/>. Recuperado el 7 de Marzo 2022.
- [30] Coinbase. Coinbase. <https://www.coinbase.com/es>. Recuperado el 7 de Marzo 2022.
- [31] 2022 CoinMarketCap. Top cryptocurrency spot exchanges. <https://coinmarketcap.com/rankings/exchanges/>. Recuperado el 7 de Marzo 2022.
- [32] Binance. Binance. <https://www.binance.com/es>. Recuperado el 7 de Marzo 2022.
- [33] FTX. Ftx. <https://ftx.com/es>. Recuperado el 7 de Marzo 2022.
- [34] Kraken. Kraken/prices/btcusd. <https://www.kraken.com/prices/xbt-bitcoin-price-chart/usd-us-dollar?interval=1m>. Recuperado el 19 de Diciembre 2020.
- [35] KuCoin. Kucoin. <https://www.kucoin.com/>. Recuperado el 7 de Marzo 2022.
- [36] Wikipedia. Fiat money. https://en.wikipedia.org/wiki/Fiat_money. Recuperado el 7 de Marzo 2022.
- [37] Daniel Jiménez Bermejo. Acuerdos de bretton woods. economipedia.com. <https://economipedia.com/definiciones/acuerdos-de-bretton-woods.html>. Recuperado el 7 de Marzo 2022.
- [38] Wikipedia. Nixon shock. https://en.wikipedia.org/wiki/Nixon_shock. Recuperado el 7 de Marzo 2022.
- [39] Wikipedia. United states dollar. https://en.wikipedia.org/wiki/United_States_dollar. Recuperado el 7 de Marzo 2022.

- [40] Grupo BBVA. Cómo crean el dinero los bancos centrales. <https://www.bbva.com/es/como-se-crea-el-dinero-2/>. Recuperado el 7 de Marzo 2022.
- [41] ISO. Iso 4217 currency codes. <https://www.iso.org/iso-4217-currency-codes.html>. Recuperado el 7 de Marzo 2022.
- [42] Nick Lioudis. Investopedia. <https://www.investopedia.com/ask/answers/09/gold-standard.asp>. Recuperado el 7 de Marzo 2022.
- [43] Wikipedia. European central bank. https://en.wikipedia.org/wiki/European_Central_Bank. Recuperado el 7 de Marzo 2022.
- [44] Wikipedia. Euro. <https://en.wikipedia.org/wiki/Euro>. Recuperado el 7 de Marzo 2022.
- [45] Sobre el Parlamento. Tratado de la unión europea (tue) / tratado de maasticht. <https://www.europarl.europa.eu/about-parliament/es/in-the-past/the-parliament-and-the-treaties/maastricht-treaty>. Recuperado el 7 de Marzo 2022.
- [46] Banco Central Europeo. Sobre el bce. <https://www.ecb.europa.eu/ecb/html/index.es.html>. Recuperado el 7 de Marzo 2022.
- [47] Banco Central Europeo. Cinco cosas que deberías saber sobre el tratado de maasticht. https://www.ecb.europa.eu/ecb/educational/explainers/tell-me-more/html/25_years_maastricht.es.html. Recuperado el 7 de Marzo 2022.
- [48] Wikipedia. Yen. <https://es.wikipedia.org/wiki/Yen>. Recuperado el 7 de Marzo 2022.
- [49] Wikipedia. Plaza accord. https://en.wikipedia.org/wiki/Plaza_Accord. Recuperado el 16 de Mayo 2022.
- [50] Wikipedia. Libra esterlina. https://es.wikipedia.org/wiki/Libra_esterlina. Recuperado el 7 de Marzo 2022.
- [51] Wikipedia. Renminbi. <https://en.wikipedia.org/wiki/Renminbi>. Recuperado el 7 de Marzo 2022.
- [52] Javier Samper Arias. Mercado de divisas. economipedia.com. <https://economipedia.com/definiciones/mercado-de-divisas-forex.html>. Recuperado el 7 de Marzo 2022.

- [53] Cory Mitchell. Forex (fx) definition, uses, & examples. <https://www.investopedia.com/terms/f/forex.asp>. Recuperado el 7 de Marzo 2022.
- [54] Shobhit Seth. How the swift system works. investopedia. <https://www.investopedia.com/articles/personal-finance/050515/how-swift-system-works.asp>. Recuperado el 7 de Marzo 2022.
- [55] Kate Rooney. CNBC. Kraken crypto exchange won't shut down russian accounts unless legally required, ceo says. <https://www.cnbc.com/2022/03/04/kraken-wont-shut-down-russian-accounts-unless-legally-required-ceo.html>. Recuperado el 7 de Marzo 2022.
- [56] Chris Morris. Fortune. Crypto exchanges choose “financial freedom” and refuse to block russian users. <https://fortune.com/2022/02/28/crypto-exchanges-wont-freeze-russian-accounts-ukraine/>. Recuperado el 7 de Marzo 2022.
- [57] 2022 INSTEX. About us - instex. <https://instex-europe.com/about-us/>. Recuperado el 7 de Marzo 2022.
- [58] Ltd. CIPS Co. Company profile. https://www.cips.com.cn/en/about_us/company_profile/index.html. Recuperado el 7 de Marzo 2022.
- [59] Bank of Russia. Financial messaging system of the bank of russia (spfs). http://www.cbr.ru/PSsystem/fin_msg_transfer_system/. Recuperado el 7 de Marzo 2022.
- [60] Organization of the Petroleum Exporting Countries. Opec: Home. https://www.opec.org/opec_web/en/. Recuperado el 7 de Marzo 2022.
- [61] Wikipedia. Metal precioso. https://es.wikipedia.org/wiki/Metal_precioso. Recuperado el 7 de Marzo 2022.
- [62] Royal Canadian MInt. 1 oz. pure gold coin - “w” mint mark: Gold maple leaf (winnipeg, 2021) - mintage: 500. <https://www.mint.ca/store/coins/1-oz.-pure-gold-coin---%22w%22-mint-mark-gold-maple-leaf-winnipeg-2021---mintage-500-prod3740043>. Recuperado el 7 de Marzo 2022.
- [63] Barclay Palmer. Investopedia. A beginner's guide to precious metals. <https://www.investopedia.com/articles/basics/09/precious-metals-gold-silver-platinum.asp>. Recuperado el 7 de Marzo 2022.

- [64] Jesse Emspak. What moves gold prices? <https://www.investopedia.com/articles/active-trading/031915/what-moves-gold-prices.asp>. Recuperado el 7 de Marzo 2022.
- [65] Cory Mitchell. Gold: The other currency. <https://www.investopedia.com/articles/forex/10/gold-the-other-currency.asp>. Recuperado el 7 de Marzo 2022.
- [66] Geoffrey Michael. Investopedia. 4 ways to buy gold. <https://www.investopedia.com/articles/basics/11/best-bets-buying-gold.asp>. Recuperado el 7 de Marzo 2022.
- [67] U.S. Geological Survey. Mineral commodity summaries 2022. <https://pubs.usgs.gov/periodicals/mcs2022/mcs2022.pdf>. Recuperado el 7 de Marzo 2022.
- [68] Wikipedia. Silve. <https://en.wikipedia.org/wiki/Silver>. Recuperado el 7 de Marzo 2022.
- [69] Lucas Downey. Investopedia. Silver etf. <https://www.investopedia.com/terms/s/silver-etf.asp>. Recuperado el 7 de Marzo 2022.
- [70] James Chen. Investopedia. Silver definition. <https://www.investopedia.com/terms/s/silver.asp>. Recuperado el 7 de Marzo 2022.
- [71] Aaron Levitt. The best way to buy silver. <https://www.investopedia.com/articles/investing/032613/best-way-buy-silver.asp>. Recuperado el 7 de Marzo 2022.
- [72] Wikipedia. Platinum group. https://en.wikipedia.org/wiki/Platinum_group. Recuperado el 7 de Marzo 2022.
- [73] The Investopedia Team. What are the most common etfs that track the metals and mining sector? <https://www.investopedia.com/ask/answers/022015/what-are-most-common-etfs-track-metals-and-mining-sector.asp>. Recuperado el 7 de Marzo 2022.
- [74] David Dierking. Investopedia. 3 funds with exposure to palladium. <https://www.investopedia.com/articles/etfs/071716/top-3-etfs-exposure-palladium-pall-sppp.asp>. Recuperado el 7 de Marzo 2022.
- [75] James Chen. Investopedia. Palladium definition. <https://www.investopedia.com/terms/p/palladium.as>. Recuperado el 7 de Marzo 2022.

- [76] 2022 Sprott Inc. Sprott physical platinum & palladium trust. <https://sprott.com/investment-strategies/physical-bullion-trusts/platinum-and-palladium/#>. Recuperado el 7 de Marzo 2022.
- [77] abrdn. Aberdeen standard physical palladium shares etf. <https://www.abrdn.com/en-us/us/investor/fund-details/aberdeen-standard-physical-palladium-shares-etf/share/us0032621023>. Recuperado el 7 de Marzo 2022.
- [78] Barclay Palmer. Investopedia. A beginner's guide to precious metals. <https://www.investopedia.com/articles/basics/09/precious-metals-gold-silver-platinum.asp>. Recuperado el 7 de Marzo 2022.
- [79] James Chen. Platinum definition. <https://www.investopedia.com/terms/p/platinum.asp>. Recuperado el 7 de Marzo 2022.
- [80] Bureau of Ocean Energy Management. Air quality act (1967). . Recuperado el 7 de Marzo 2022.
- [81] 2022 Binance Academy. ¿qué es el análisis fundamental (af)? <https://academy.binance.com/es/articles/what-is-fundamental-analysis-fa>. Recuperado el 18 de Abril 2022.
- [82] Investopedia Adam Hayes. Technical analysis. <https://www.investopedia.com/terms/t/technicalanalysis.asp>. Recuperado el 18 de Abril 2022.
- [83] 2022 Binance Academy. ¿qué es el análisis técnico? <https://academy.binance.com/es/articles/what-is-technical-analysis>. Recuperado el 18 de Abril 2022.
- [84] 2022 Binance Academy. Guía sobre los principios básicos del soporte y la resistencia. <https://academy.binance.com/es/articles/the-basics-of-support-and-resistance-explained>. Recuperado el 2 de Abril 2022.
- [85] Investopedia James Chen. Overbought. <https://www.investopedia.com/terms/o/overbought.asp>. Recuperado el 18 de Abril 2022.
- [86] Investopedia Cory Mitchell. Oversold. <https://www.investopedia.com/terms/o/oversold.asp>. Recuperado el 18 de Abril 2022.
- [87] JORIS TOONDERS. Data is the new oil of the digital economy. *Wired*, 07 2014. Recuperado el 28 de Febrero de 2020.

- [88] Francisco Coll Morales. Economipedia. Fuente terciaria. <https://economipedia.com/definiciones/fuente-terciaria.html>. Recuperado el 7 de Marzo 2022.
- [89] Hewlett Packard Enterprise Development LP. Almacenamiento de datos. <https://www.hpe.com/es/es/what-is/data-storage.html>. Recuperado el 4 de Marzo 2022.
- [90] Hewlett Packard Enterprise Development LP. Almacenamiento de objetos. <https://www.hpe.com/es/es/what-is/object-storage.html>. Recuperado el 4 de Marzo 2022.
- [91] International Organization for Standardization. Iso 8601, date and time format. <https://www.iso.org/iso-8601-date-and-time-format.html>. Recuperado el 7 de Marzo 2022.
- [92] Kraken. Why kraken. <https://www.kraken.com/en-us/why-kraken/>. Recuperado el 7 de Marzo 2022.
- [93] Wikipedia. Podcast. <https://es.wikipedia.org/wiki/P%C3%B3dcast>. Recuperado el 7 de Marzo 2022.
- [94] 2022 Crypto Coin Comparison Ltd. About cryptocompare. <https://www.cryptocompare.com/about-us/>. Recuperado el 7 de Marzo 2022.
- [95] Bitcoincharts. Bitcoincharts/charts. <https://bitcoincharts.com/charts/bitstampUSD#rg60ztgSzm1g10zm2g25zv>. Recuperado el 19 de Diciembre 2020.
- [96] BitcoinAverage. Bitcoinaverage/globalindices. <https://bitcoinaverage.com/en/bitcoin-price/btc-to-usd>. Recuperado el 19 de Diciembre 2020.
- [97] CoinMarketCap. Bitcoin chart. <https://coinmarketcap.com/currencies/bitcoin/>. Recuperado el 19 de Diciembre 2020.
- [98] CNBC LLC. A Division of NBCUniversal. International business, world news & global stock market analysis. <https://www.cnbc.com/world/?region=world>. Recuperado el 7 de Marzo 2022.
- [99] 2022 Bloomberg Finance L.P. Bloomberg l.p about, careers, products, contacts. <https://www.bloomberg.com/>. Recuperado el 7 de Marzo 2022.
- [100] CoinMarketCap. About coinmarketcap. <https://coinmarketcap.com/about/>. Recuperado el 7 de Marzo 2022.

- [101] 2022 TradingView. Acerca de tradingview. <https://es.tradingview.com/>. Recuperado el 7 de Marzo 2022.
- [102] Blockchain Data LTD. The bitcoinaverage.com bitcoin price index methodology. <https://bitcoinaverage.com/en/methodology>. Recuperado el 7 de Marzo 2022.
- [103] Binance Academy. Volume. <https://academy.binance.com/en/glossary/volume>. Recuperado el 7 de Marzo 2022.
- [104] Binance Academy. Guía sobre la capitalización de mercado cripto. <https://academy.binance.com/es/articles/crypto-market-capitalization-explained>. Recuperado el 8 de Marzo 2022.
- [105] IBM Corporation. Gráficos de velas japonesas - documentación de ibm. <https://www.ibm.com/docs/es/spss-modeler/18.2.0?topic=types-candlestick-charts>. Recuperado el 7 de Marzo 2022.
- [106] Inc 2022 Twitter. Twitter api documentation. <https://developer.twitter.com/en/docs/twitter-api>. Recuperado el 7 de Marzo 2022.
- [107] 2022 Stoplight. Api documentation guide and best practices. <https://stoplight.io/api-documentation-guide/#what-is-api-documentation%3F>. Recuperado el 7 de Marzo 2022.
- [108] Portions of this content are 1998 2022 by individual mozilla.org contributors. Content available under a Creative Commons license. Http caching- http. <https://developer.mozilla.org/en-US/docs/Web/HTTP/Caching>. Recuperado el 7 de Marzo 2022.
- [109] Portions of this content are 1998 2022 by individual mozilla.org contributors. Content available under a Creative Commons license. What is a url - learn web development. https://developer.mozilla.org/es/docs/Learn/Common_questions/What_is_a_URL. Recuperado el 7 de Marzo 2022.
- [110] Portions of this content are 1998 2022 by individual mozilla.org contributors. Content available under a Creative Commons license. Http request methods - http. <https://developer.mozilla.org/en-US/docs/Web/HTTP/Methods>. Recuperado el 7 de Marzo 2022.

- [111] Individual mozilla.org contributors. Safe (http methods) - mdn web docs glossary: Definitions of web-related terms. <https://developer.mozilla.org/en-US/docs/Glossary/Safe/HTTP>. Recuperado el 2 de Abril 2022.
- [112] Individual mozilla.org contributors. Cacheable - mdn web docs glossary: Definitions of web-related terms. <https://developer.mozilla.org/es/docs/Glossary/cacheable>. Recuperado el 2 de Abril 2022.
- [113] Curl. Curl. <https://curl.se/>. Recuperado el 25 de Febrero 2021.
- [114] Inc Free Software Foundation. Bash - gnu project - free software foundation. <https://www.gnu.org/software/bash/>. Recuperado el 2 de Abril 2022.
- [115] Inc. Free Software Foundation. The gnu operating system and the free software movement. <https://www.gnu.org/home.en.html>. Recuperado el 2 de Abril 2022.
- [116] David Korn. Korn shell. <http://www.kornshell.org/>. Recuperado el 23 de Mayo 2022.
- [117] Wikipedia. C shell. https://en.wikipedia.org/wiki/C_shell. Recuperado el 23 de Mayo 2022.
- [118] ISO. Iso/iec/ieee 9945:2009 information technology - portable operating system interface (posix®) base specifications, issue 7. <https://www.iso.org/standard/50516.html>. Recuperado el 2 de Abril 2022.
- [119] Inc Free Software Foundation. Gnu sed - gnu project - free software foundation. <https://www.gnu.org/software/sed/>. Recuperado el 2 de Abril 2022.
- [120] Wikipedia. Perl. <https://en.wikipedia.org/wiki/Perl>. Recuperado el 2 de Abril 2022.
- [121] Wikipedia. sed. <https://en.wikipedia.org/wiki/Sed>. Recuperado el 2 de Abril 2022.
- [122] Por definir. Manual awk y sed. Material de SUMA.
- [123] . Recuperado el 24 de Abril 2022.
- [124] Inc Free Software Foundation. The gnu awk user's guide. <https://www.gnu.org/software/gawk/manual/gawk.html>. Recuperado el 2 de Abril 2022.
- [125] Wikipedia. grep. <https://en.wikipedia.org/wiki/Grep>. Recuperado el 2 de Abril 2022.

- [126] Inc Free Software Foundation. Gnu grep 3.7. <https://www.gnu.org/software/grep/manual/grep.html>. Recuperado el 2 de Abril 2022.
- [127] William G. Cochran. *Sampling Techniques*. JOHN WILEY & SONS. New York, Chichester, Brisbane, Toronto, Singapore, 1977.
- [128] Investopedia James Chen. Arithmetic mean. <https://www.investopedia.com/terms/a/arithmeticmean.asp>. Recuperado el 24 de Abril 2022.
- [129] Wikipedia. Media ponderada. https://es.wikipedia.org/wiki/Media_ponderada. Recuperado el 18 de Abril 2022.
- [130] Investopedia Adam Hayes. Geometric mean definition. <https://www.investopedia.com/terms/g/geometricmean.asp>. Recuperado el 24 de Abril 2022.
- [131] Wikipedia. Median. <https://en.wikipedia.org/wiki/Median>. Recuperado el 24 de Abril 2022.
- [132] Wikipedia. Media cuadratica. https://es.wikipedia.org/wiki/Media_cuadr%C3%A1tica. Recuperado el 24 de Abril 2022.
- [133] Economipedia José Francisco López. Medidas de tendencia central. <https://economipedia.com/definiciones/medidas-de-tendencia-central.html>. Recuperado el 24 de Abril 2022.
- [134] Investopedia Gordon Scott. Range. <https://www.investopedia.com/terms/r/range.asp>. Recuperado el 24 de Abril 2022.
- [135] Wikipedia. Moving average - wikipedia. https://en.wikipedia.org/wiki/Moving_average. Recuperado el 2 de Abril 2022.
- [136] Investopedia Jason Fernando. Moving average (ma) definition. <https://www.investopedia.com/terms/m/movingaverage.asp>. Recuperado el 14 de Abril 2022.
- [137] 2022 Binance Academy. A guide to mastering fibonacci retracement. <https://academy.binance.com/en/articles/a-guide-to-mastering-fibonacci-retracement>. Recuperado el 2 de Abril 2022.
- [138] Investopedia Cory Mitchell. Fibonacci numbers and lines definition and uses. <https://www.investopedia.com/terms/f/fibonaccilines.asp>. Recuperado el 2 de Abril 2022.

- [139] Investopedia James Chen. Oscillator. <https://www.investopedia.com/terms/o/oscillator.asp>. Recuperado el 14 de Abril 2022.
- [140] 2022 Binance Academy. Indicador macd explicado. <https://academy.binance.com/es/articles/macd-indicator-explained>. Recuperado el 18 de Abril 2022.
- [141] J. Welles Wilder JR. *New Concepts in Technical Trading Systems*. Hunter Publishing Company, Winston-Salem, North Carolina, 1978.
- [142] Investopdeia Jason Fernando. Relative strength index (rsi). <https://www.investopedia.com/terms/r/rsi.asp>. Recuperado el 24 de Abril 2022.
- [143] Investopedia Adam Hayes. Stochastic oscillator. <https://www.investopedia.com/terms/s/stochasticoscillator.asp>. Recuperado el 18 de Abril 2022.
- [144] Investopedia Cory Mitchell. Average directional index (adx). <https://www.investopedia.com/terms/a/adx.asp>. Recuperado el 18 de Abril 2022.
- [145] Investopedia Cory Mitchell. Ichimoku cloud definition and uses. <https://www.investopedia.com/terms/i/ichimoku-cloud.asp>. Recuperado el 18 de Abril 2022.
- [146] Wikipedia. Distribucion triangular. https://es.wikipedia.org/wiki/Distribuci%C3%B3n_triangular. Recuperado el 24 de Abril 2022.
- [147] Beta. Beta distribution. https://en.wikipedia.org/wiki/Beta_distribution. Recuperado el 24 de Abril 2022.
- [148] wikipedia. Wigner semicircle distribution. https://en.wikipedia.org/wiki/Wigner_semicircle_distribution. Recuperado el 24 de Mayo 2022.
- [149] Kay A. Tucci K. Mayerlin Y. Uzcátegui S., Jacinto A. Dávila Q. *Simulacion Multi-Agente con Galatea*. Universidad de Los Andes. Mérida, Venezuela, 2004.
- [150] Ernesto Kofman Bernard Zeigler, Alexandre Muzy. *Theory of Modeling and Simulation. Discrete Event and Iterative System Computational Foundations*. Wiley Interscience, 1976 with reissue by Krieger Pub., 1984.
- [151] T. G. Kim B. P. Zeigler, H. Praehofer. *Theory of Modelling and Simulation. 2nd ed.* Academic Press, 2000.
- [152] D. A. Fahrland. Combined discrete event continuous systems simulation, 1970.

- [153] Wikipedia. Tcl. <https://es.wikipedia.org/wiki/Tcl>. Recuperado el 02 de Agosto 2022.
- [154] Wikipedia. Tkinter. <https://en.wikipedia.org/wiki/Tkinter>. Recuperado el 02 de Agosto 2022.
- [155] Wikipedia. Categoría: Widgets. <https://es.wikipedia.org/wiki/Categor%C3%ADa:Widgets>. Recuperado el 09 de Septiembre 2022.

www.bdigital.ula.ve