

Relación entre la formación académica de los estudiantes universitarios y el éxito de sus proyectos en el Nasa Space Apps Challenge, Guayaquil (2019–2023)

Relationship between the academic training of university students and the success of their projects in the NASA Space Apps Challenge, Guayaquil (2019–2023)

<https://doi.org/10.47606/ACVEN/PH0371>

Roberto Augusto Poveda Páez^{1*}

<https://orcid.org/0000-0001-9503-4534>
rpovedap@unemi.edu.ec

John Kevin Tigre Méndez¹

<https://orcid.org/0000-0002-0293-7484>
jtigrem@unemi.edu.ec

Josué Jandry Aranea Álava²

<https://orcid.org/0009-0004-9993-0589>
jaranea@unemi.edu.ec

Recibido: 07/07/2025

Aceptado: 01/09/2025

RESUMEN

Los hackatones universitarios promueven colaboración e innovación, pero la evidencia sobre cómo la formación académica de los equipos se relaciona con el éxito aún es limitada. **Objetivo.** Analizar la relación entre la formación académica (carrera/área, STEM vs. no-STEM, composición mono vs. interdisciplinaria) y el éxito de proyectos en el NASA Space Apps Challenge—Guayaquil (2019–2023). **Diseño.** Estudio cuantitativo, no experimental, transversal y correlacional con unidad de análisis equipo—año (censo 2019–2023). **Variable dependiente:** éxito (1 = Global Winner; 0 = Global Nominee). **Predictores:** %STEM, índice de Blau (diversidad disciplinaria), interdisciplinariedad (≥ 2 macroáreas), dominancia de Computación/Ingenierías (≥ 50 %). **Covariables de proceso:** % mentoría, % experiencia previa, % mujeres e índice de Blau de género. El desempeño local se midió con SCL-z (puntaje de calidad local estandarizado por año), derivado de cinco criterios de rúbrica ($\alpha = 0,896$). **Análisis:** descriptivos; correlaciones de Pearson/Spearman con SCL-z; pruebas exactas de Fisher para asociaciones 2x2; análisis de sensibilidad con umbrales alternativos del 40 %/60 %. **Resultados.** No se hallaron asociaciones lineales fuertes entre %STEM o diversidad disciplinaria y SCL-z. La experiencia previa mostró asociación positiva moderada con SCL-z ($p < .05$); la mentoría y la composición de género no evidenciaron relaciones robustas. **Conclusiones.** El éxito no parece depender de un perfil académico único; prácticas de equipo y mentoría estructurada probablemente sean más determinantes. Se recomienda ampliar la muestra y emplear métodos para eventos raros, además de usar desenlaces continuos/ordinales basados en la rúbrica.

Palabras Clave: Aprendizaje basado en proyectos; NASA Space Apps Challenge; interdisciplinariedad; composición de equipos; rendimiento académico

1. Universidad Estatal De Milagro (UNEMI) / Instituto Superior Tecnológico Vicente Rocafuerte

2. Universidad Estatal De Milagro (UNEMI)

* Autor de correspondencia: rpovedap@unemi.edu.ec

ABSTRACT

University hackathons foster collaboration and innovation, yet evidence on how teams' academic background relates to success remains limited. Objective. To examine the relationship between academic formation (major/area, STEM vs. non-STEM, mono- vs. interdisciplinary composition) and project success in the NASA Space Apps Challenge—Guayaquil (2019–2023). Design. Quantitative, non-experimental, cross-sectional, correlational study with team–year as the unit of analysis (2019–2023 census). Outcome: success (1 = Global Winner; 0 = Global Nominee). Predictors: %STEM, Blau's index (disciplinary diversity), interdisciplinarity (≥ 2 macro-areas), and Computing/Engineering dominance (≥ 50 %). Process covariates: % mentoring, % prior experience, % women, and gender Blau index. Local performance was measured with SCL-z (year-standardized local quality score) derived from five rubric criteria ($\alpha = .896$). Analyses: descriptives; Pearson/Spearman correlations with SCL-z; Fisher's exact tests for 2×2 associations; sensitivity analyses using 40 %/60 % thresholds. Results. No strong linear associations were found between %STEM or disciplinary diversity and SCL-z. Prior experience showed a positive, moderate association with SCL-z ($p < .05$), whereas mentoring and gender composition showed no robust relationships. Conclusions. Success does not appear to depend on a single academic profile; team practices and structured mentoring likely matter more. We recommend enlarging the sample, applying rare-event methods, and using continuous/ordinal rubric-based outcomes.

Keywords: project-based learning; NASA Space Apps Challenge; interdisciplinarity; team composition; academic achievement; Ecuador.

INTRODUCCIÓN

Los hackatones se han consolidado en la educación superior como entornos de aprendizaje intensivo. En ellos, equipos multidisciplinarios resuelven problemas reales mediante propuestas creativas en tiempos breves. Este formato potencia la creatividad, la colaboración y la innovación. No obstante, la evidencia académica sigue siendo dispersa y heterogénea. Una revisión sistemática sobre hackatones como oportunidades de aprendizaje indica la falta de estandarización para evaluar aprendizajes y resultados, así como la prevalencia de medidas autorreportadas y escasez de indicadores objetivos (Schulten & Chounta, 2024).

En la misma línea, un mapeo bibliométrico con alcance en la educación, identifica un grupo de 249 publicaciones, entre los años 2014 y 2022, y concluye que el área está en fase de consolidación y persisten brechas respecto de resultados verificables y efectos a largo plazo (García, 2023). Estas conclusiones justifican estudios que utilicen métricas de desempeño comparables entre eventos y localidades.

El NASA International Space Apps Challenge constituye un laboratorio global que fomenta la creatividad, colaboración e innovación, para ello emplea datos abiertos, reúne a diversos perfiles (estudiantes y profesionales) y aplica criterios de evaluación comparables entre ediciones anuales y las localidades, lo que facilita analizar el desempeño de equipos bajo condiciones similares (NASA, 2024). En este marco, resulta pertinente examinar si características de la formación académica de los participantes se asocian con el éxito objetivo alcanzado en el evento.

La literatura pedagógica sugiere que enfoques como el aprendizaje basado en proyectos (ABP/PBL) y el design thinking fortalecen la creatividad, la resolución de problemas y otras competencias transversales, justamente las que exige un hackatón. Un meta-análisis con 66 estudios muestra que el ABP mejora de forma significativa el rendimiento académico y las habilidades de orden superior frente a métodos tradicionales (Zhang & Ma, 2023). A su vez, la hoja de ruta de UNESCO para la educación para el desarrollo sostenible enfatiza la necesidad de formar capacidades para abordar problemas complejos de manera colaborativa, lo que se alinea plenamente con las experiencias en los hackatones (UNESCO, 2020).

A pesar de ello, persiste una brecha metodológica, muchos estudios sobre hackatones reportan aprendizajes percibidos y no emplean indicadores verificables de desempeño, lo que limita la comparabilidad entre investigaciones (Schulten & Chounta, 2024; García, 2023). En este trabajo de investigación nos centraremos en el NASA Space Apps Challenge, realizado en la ciudad de Guayaquil, Ecuador, específicamente sobre los ganadores locales a los que se denomina como “Global Nominee”, esto a lo largo de las ediciones generales anuales de los años 2019, 2020, 2021, 2022 y 2023. Para aportar evidencia más robusta, este estudio utilizará como variable dependiente objetiva el resultado del equipo en Space Apps (Global Winner vs. Global Nominee) y relacionará con rasgos de la formación académica de los participantes.

Se plantea como pregunta de investigación: ¿En qué medida la formación académica de los participantes y área de estudio (STEM vs. no-STEM) se asocia con el éxito del proyecto en el NASA Space Apps Challenge, Guayaquil (2019–2023)?

El objetivo general consiste en analizar la relación entre la formación académica de los participantes y el éxito del proyecto en el NASA Space Apps Challenge, Guayaquil (2019–2023). Los objetivos específicos planteados son: a) Describir el perfil de formación (profesión y área), b) Determinar la asociación entre la profesión y éxito y c) Determinar la asociación entre área de estudio (STEM) y éxito.

Se establecen como hipótesis: H1. Existe asociación entre profesión y éxito del proyecto, H2. A mayor %STEM, mayor probabilidad de Winner, y H0. No existe asociación entre formación académica y éxito.

Este trabajo realiza cuatro aportes principales: (1) define un desenlace objetivo de éxito (Global Winner vs. Global Nominee); (2) propone un puntaje de calidad local estandarizado por año (SCL-z) para comparar ediciones; (3) documenta la confiabilidad de la rúbrica ($\alpha = 0,896$) con estimación de

incertidumbre; y (4) caracteriza empíricamente el vínculo entre composición académica, experiencia previa y desempeño.

Con base en este vacío y en nuestros objetivos, procedemos a describir el diseño, las variables y los procedimientos analíticos.

METODOLOGÍA

Este estudio adopta un enfoque cuantitativo con diseño no experimental, transversal y correlacional, pertinente para la búsqueda de asociaciones entre variables sin manipularlas y la medición ocurre en único corte temporal (Hernández-Sampieri & Mendoza, 2020).

El contexto es el NASA International Space Apps Challenge en Guayaquil (2019–2023). Dado que el reconocimiento se otorga al equipo, la unidad de análisis primaria es el equipo quienes generan un proyecto en cada edición anual (2019–2023). Se trabajó con un censo de 11 equipos con reconocimiento local (Global Nominee), incluyendo al Ganador Global (Global Winner) del período 2021, los resultados oficiales se verificaron en el portal del evento, las páginas de los organizadores locales, la base de datos de pre-registros del evento y en páginas de ganadores globales de los años de referencia.

Se modelaron las siguientes variables: Éxito (1=Global Winner; 0=Global Nominee); %STEM (0–100) y composición disciplinaria (Índice de Blau; Interdisciplinaria ≥ 2 macroáreas). Para profesión, se colapsa en Computación/Ingenierías-dominante (≥ 50 %). Puntaje de calidad local = promedio z de cinco criterios por año. En cuanto a la confiabilidad los cinco criterios que se evalúan en escala de Likert (1-5), presentan un α de Cronbach = 0,896.

Es importante mencionar que tanto los proyectos ganadores locales (Global Nominee) como los ganadores mundiales (Global Winners) se encuentran disponibles de manera pública en la página oficial del NASA Space Apps Challenge.

Con fines descriptivos, también se caracterizaron los perfiles individuales de 62 integrantes, para resumir su nivel y área de formación, sin emplear esos datos individuales en las pruebas inferenciales para evitar dependencia intraclase, como se muestra en las Tabla 1 y 2.

Tabla 1

Frecuencias por profesión

Ítem	Profesión	Frecuencia	%
1	Ingeniería en ciencias computacionales	24	38,7
2	Ingeniería en mecánica	18	29,0
3	Licenciatura en diseño gráfico	2	3,2
4	Ingeniería en software	2	3,2
5	Ingeniería naval	1	1,6
6	Ingeniería eléctrica	1	1,6
7	Producción audiovisual	1	1,6
8	Entrenamiento terapéutico	1	1,6

Ítem	Profesión	Frecuencia	%
9	Ingeniería multimedia	1	1,6
10	Ingeniería empresarial	1	1,6
11	Ingeniería en electrónica	1	1,6
12	Ingeniería en electrónica y telecomunicaciones	1	1,6

Tabla 2

Distribución por macroárea

Ítem	Macroárea	Frecuencia	%
1	STEM	55	88,7
2	Artes/Diseño	4	6,5
3	Otras	3	4,8

El análisis se efectuó a nivel equipo-año (2019–2023), sede Guayaquil, los registros individuales (integrantes) se agregaron al nivel equipo-año para construir indicadores de composición. Previamente, se implementó una depuración: normalización de texto en carrera/profesión, género, mentoría, experiencia previa, proyecto y año, con verificación de unicidad por combinación año-proyecto y resolución de duplicidades conservando la versión más completa. Las carreras se codificaron en macroáreas: STEM, Ciencias Sociales, Salud, Negocios/Economía, Educación, Derecho, Artes/Diseño y Otras, mediante un diccionario controlado y revisión por dos codificadores. La diversidad disciplinaria del equipo se cuantificó con el Índice de Blau (IB), práctica consolidada en estudios recientes de diversidad (Scott, 2024).

En concreto, el índice de Blau se definió como $IB = 1 - \sum p_i^2$, donde p_i es la proporción de integrantes en la categoría i (número integrantes en macroárea i / total de integrantes).

Si el resultado del Índice de Blau es cercano a cero (0) indica mono-disciplina y valores mayores reflejan mayor heterogeneidad (Scott, 2024; Yang, 2023).

A partir de esa codificación se construyeron los indicadores de equipo: %STEM = (integrantes STEM/total)×100, Índice de Blau disciplinario (IB), Interdisciplinaria (INT) = 1 si el equipo incluye 2 o más macroáreas (0 en caso contrario), Dominancia Computación/Ingenierías (DCI) = 1 si ≥50 % de integrantes pertenecen a subáreas de computación/ingenierías (con análisis de sensibilidad a 40 % y 60 %), % experiencia previa (%EXP) y % mentoría (%MEN), % mujeres (%M) y Índice de Blau de género (IB-G). La calidad local se obtuvo como el promedio de cinco criterios (1–5: impacto, creatividad, validez, relevancia y presentación), por otra parte, para asegurar comparabilidad temporal, se estandarizó por año (z-scores) y se definió un puntaje de calidad local (SCL-z) como la media de los cinco z por equipo. La variable dependiente (VD) fue el éxito (1 = Global Winner, 0 = Global Nominee). Las proporciones (% STEM, % experiencia previa, % mentoría y % mujeres) se expresaron en porcentaje (0–100).

La consistencia interna de los cinco criterios se evaluó con alfa de Cronbach (α), en nuestros datos, $\alpha = 0.896$. Dado el interés por estimación e incertidumbre en métricas de confiabilidad, se calculó un IC95 % por bootstrap y se reportaron correlaciones ítem–total corregidas y α si se elimina el ítem (Farkas et al., 2023; Edelsbrunner et al., 2025).

El enfoque fue cuantitativo, seleccionando pruebas coherentes con la naturaleza de las variables y el tamaño muestral.

En cuanto a los descriptivos, las variables continuas (%STEM, IB, SCL-z, %EXP, %MEN, %M, IB-G) se resumen con media (M), desviación estándar (DE), mediana, Q1–Q3 y mín–máx; cuando es pertinente, se incluyen IC95 % de la media. Para proporciones (p. ej., INT, DCI, éxito) se reportan IC95 % exactos (Clopper–Pearson) dada la posible baja incidencia del evento Winner (McGrath & Burke, 2024).

El puntaje de calidad local estandarizado por año (SCL-z) se obtuvo calculando, para cada edición, las puntuaciones z de los cinco criterios (impacto, creatividad, validez, relevancia, presentación) y promediándolas por equipo.

Para evaluar asociación lineal entre SCL-z y predictores continuos (%STEM, IB, %EXP, %MEN, %M, IB-G) se emplea correlación de Pearson (r) con IC95 % vía transformación z de Fisher, como análisis de sensibilidad se utiliza Spearman (ρ) cuando la linealidad/normalidad es dudosa o hay valores atípicos (Bocianowski et al., 2024; Tyagi et al., 2023).

Para las asociaciones 2×2 con el desenlace (Winner), se analizaron cuatro predictores dicotómicos: experiencia previa (≥ 1 integrante), modalidad (virtual vs. presencial), presencia femenina (≥ 1 mujer) y tamaño del equipo (≥ 6). Se empleó la prueba exacta de Fisher; se reportaron odds ratios (OR) —con corrección de Haldane–Anscombe cuando hubo celdas 0—, IC95 % exactos y p de dos colas; este enfoque es adecuado en escenarios donde la χ^2 puede ser inadecuada (Abdul Rahman et al., 2025).

Por otro lado, si hablamos de las comparaciones Winner vs Nominee, para continuas se presentan M (DE) por grupo y, cuando aporta, diferencia de medias con IC95 % como estimación del efecto. Evitamos pruebas paramétricas con N extremadamente pequeño en el grupo Winner (enfoque centrado en estimación y precisión).

En las sensibilidades pre-especificadas, tenemos: (i) Umbrales alternativos ≥ 40 % y ≥ 60 % para STEM-dominante y DCI (Fisher), (ii) procedimiento leave-one-year-out (excluyendo 2019–2023 alternadamente) para examinar la estabilidad de r y de los OR. Este énfasis en intervalos de confianza y robustez está alineado con recomendaciones metodológicas recientes que priorizan precisión e inferencia compatible y el uso de bootstrap para IC en métricas de desempeño (McGrath & Burke, 2024; Orawo, 2025).

Para los niveles de decisión y reporte, se adoptó $\alpha = 0.05$ (dos colas), se reportan estimadores puntuales (r, OR), sus IC95 % y valores p cuando aplican, enfatizando la magnitud y precisión frente a la mera significación.

Entre las herramientas y la reproducibilidad, el procesamiento y los análisis se implementaron en Python (pandas, numpy, scipy, matplotlib), generando

programáticamente tablas y figuras desde la base original para asegurar trazabilidad.

Definidas las medidas y justificados los análisis (SCL-z, α , correlaciones y Fisher exacto), presentamos a continuación los resultados, avanzando de los descriptivos generales a las asociaciones y pruebas 2x2.

RESULTADOS

La unidad de análisis fue equipo-año (2019–2023), en la localidad Guayaquil. En total se analizaron los equipos Global Nominee de cada edición y el Global Winner del período. La Tabla 3 resume, por año, el número de proyectos y la ocurrencia del evento Winner, la baja incidencia de este evento justifican el uso posterior de pruebas exactas en las comparaciones dicotómicas.

Tabla 3

Características de la muestra por año

Ítem	Año	Número de proyectos	Número de ganadores globales	%Ganadores globales
1	2019	2	0	0
2	2020	2	0	0
3	2021	3	1	33,3 %
4	2022	2	0	0
5	2023	2	0	0

Nota. % Winner = (N winners/N proyectos)×100.

A nivel equipo-año (2019–2023), los equipos promediaron 5,64 integrantes (Mediana = 6; RIC = 5,5–6). La composición fue predominantemente STEM (M = 85,758 %, Mediana = 100 %; RIC = 73,3–100). La diversidad por macroárea (Índice de Blau) mostró Mediana = 0,00 (Q3 = 0,382), lo que refleja que varios equipos fueron mono-disciplina o de heterogeneidad baja. En procesos de equipo, la experiencia previa alcanzó M = 78,2 % (Mediana = 100 %; RIC = 80–100), la mentoría M = 67,4 % (Mediana = 66,7 %; RIC = 55,0–83,3), y la participación femenina M = 22,9 % (Mediana = 20 %; RIC = 16,7–33,3); el Índice de Blau (género) presentó Mediana = 0,320 (Q3 = 0,444). Véase Tabla 4.

El puntaje compuesto (SCL-z) presentó Mediana = -0,200 (RIC = -0,283–0,276), coherente con la variabilidad entre ediciones.

Tabla 4

Estadísticos descriptivos a nivel equipo-año (2019–2023)

Variable	n	M	DE	Mín	Q1	Mediana	Q3	Máy
% STEM	11	85,758 %	17,262 %	60 %	66,667 %	100 %	100 %	100 %
Índice de Blau	11	0,195	0,231	0	0	0	0,444	0,5
Puntaje compuesto z	11	-0,051	0,655	-1,298	-0,283	-0,2	0,276	0,94

Variable	n	M	DE	Mín	Q1	Mediana	Q3	Máx
(SCL-z)								
% experiencia previa	11	78,182 %	40,452 %	0 %	80 %	100 %	100 %	100 %
% mentoría	11	67,424 %	25,247 %	16,667 %	55 %	66,667 %	83,3 %	100 %
% mujeres	11	22,879 %	15,565 %	0 %	16,667 %	20 %	33,3 %	50 %
Índice de Blau (género)	11	0,309	0,173	0	0,278	0,32	0,44	0,5
N integrantes	11	5,636	0,674	4	5,5	6	6	6

Nota. M = media; DE = desviación estándar; Q1–Q3 = cuartiles. %STEM, % experiencia previa, % mentoría y % mujeres expresan proporciones en porcentaje (0–100). En el índice de Blau, valores cercanos a 0 indican mono-disciplina (o mono-género) y valores mayores, mayor heterogeneidad. SCL-z: puntaje de calidad local estandarizado por año (promedio de los z de los cinco criterios). Unidad analítica: equipo-año.

La rúbrica de cinco criterios (impacto, creatividad, validez, relevancia, presentación; escala 1–5) mostró consistencia interna alta: $\alpha = 0,896$. Para estimar la incertidumbre, se calculó un IC95 % por bootstrap, y se reportan parámetros por ítem: correlaciones ítem–total corregidas y α si se elimina (ver Tabla 5 y Tabla 6). Las Figuras A1–A3 ilustran la contribución de cada criterio y la distribución bootstrap de α .

Tabla 5

Confiabilidad global de la rúbrica.

k (nº ítems)	N (registros crudos)	α de Cronbach	IC95 % LI (bootstrap)	IC95 % LS (bootstrap)
5	62	0,896	0,866	0,917

Nota. IC95 % por bootstrap (B=3000).

Tabla 6

Consistencia por ítem (ítem–total y α si se elimina).

Criterio	M	DE	r(ítem–total) corr.	α si se elimina
Criterio impacto (1-5)	3,71	0,51	0,945	0,824
Criterio creatividad (1-5)	3,96	0,33	0,951	0,84
Criterio validez (1-5)	3,92	0,27	0,7	0,893
Criterio relevancia (1-5)	3,89	0,35	0,675	0,889
Criterio presentación (1-5)	4,22	0,52	0,653	0,909

Nota. r(ítem–total) corregida: correlación del criterio con la suma de los restantes.

Se evaluaron correlaciones de Pearson (r) entre SCL-z y los predictores continuos: %STEM, Índice de Blau (área), % experiencia previa, % mentoría, %

mujeres e Índice de Blau de género. El IC95 % de r se obtuvo con la transformación z de Fisher:

$$z = \text{atanh}(r), \text{ SE} = \frac{1}{\sqrt{n-3}}, \text{ IC95\%}(z) = z \pm 1.96 \cdot \text{SE}, \text{ IC95\%}(r) = \tanh(\text{IC95\%}(z))$$

Como sensibilidad, se reporta Spearman (ρ). La Tabla 7 resume r (IC95 %), p , ρ , y n por predictor.

No se hallaron asociaciones lineales fuertes entre %STEM ni la diversidad disciplinaria (IB) y el desempeño (SCL- z). La experiencia previa del equipo mostró correlación moderada con SCL- z ($r \approx 0,66$). En las comparaciones Winner vs. Nominee se observaron efectos pequeños y alta incertidumbre por baja incidencia del evento Winner. Llama la atención que el equipo Global Winner presentó un SCL- z inferior al promedio de los Nominee en su año ($-1,30$ vs. $0,07$) (“véase Tabla 7 para correlaciones; Tablas 8–9 para contrastes Winner vs. Nominee”).

Tabla 7

Correlaciones (Pearson y Spearman) con SCL- z .

Variable	r de Pearson	IC95 % LI	IC95 % LS	p (Pearson)	ρ de Spearman	p (Spearman)	n
% STEM	0,162	-	0,694	0,635	0,235	0,488	11
Índice de Blau (área)	-0,157	-	0,489	0,644	-0,227	0,502	11
% Experiencia	0,659	0,098	0,902	0,027	0,579	0,062	11
% Mentoría	-0,009	-	0,594	0,979	-0,128	0,707	11
% Mujeres	-0,146	-	0,497	0,669	-0,166	0,626	11
Índice de Blau (género)	-0,19	-	0,463	0,576	-0,166	0,626	11

Nota. IC95 % de r por z de Fisher; ρ = Spearman.

Se realizaron tablas 2×2 y prueba exacta de Fisher para las asociaciones entre éxito (1=Winner, 0=Nominee) además de Interdisciplinaria (≥ 2 macroáreas), STEM-dominante (≥ 50 %) y Computación/Ingenierías-dominante (≥ 50 %). Se reportan OR, IC95 % (cuando todas las celdas > 0) y p exacta (dos colas). La Tabla 8 muestra descriptores de Winner vs Nominee, la Tabla 9 resume los Fisher exactos.

Como contraste entre niveles de logro, describimos medias por grupo; obsérvese que el N de Winner es mínimo, por lo que priorizamos estimación sobre significación.

Tabla 8

Comparación descriptiva Winner vs Nominee.

Variable	Winner (M, DE)	Nominee (M, DE)	n Winner / n Nominee
% STEM	100,00 (—)	88,00 (15,96)	1 / 10
Blau (área)	0,00 (—)	0,18 (0,23)	1 / 10
Puntaje compuesto (z)	-1,30 (—)	0,07 (0,54)	1 / 10

Nota. M (DE) por grupo; N Winner reducido.

Llama la atención la desalineación SCL-z-veredicto global (Tabla 8), lo que motiva explorar asociaciones binarias específicas con Fisher.

Contrastamos ahora cuatro predictores dicotómicos con el desenlace mediante Fisher exacto y OR con corrección de continuidad.

Tabla 9

Asociación entre características del equipo y probabilidad de Global Winner: Fisher exacto, OR e IC95 % (equipo-año)

Predictor (exposición)	Winner & Sí (a)	No Winner & Sí (b)	Winner & No (c)	No Winner & No (d)	OR (Haldane-Anscombe)	IC95 % (bajo)	IC95 % (alto)	p (Fisher, 2-colas)
Experiencia previa (≥ 1)	1	8	0	2	0,882	0,012	20,874	1
Modalidad virtual	1	2	0	8	10,2	0,192	333,987	0,273
≥ 1 mujer en el equipo	1	8	0	2	0,882	0,012	20,874	1
Tamaño del equipo ≥ 6	1	7	0	3	1,4	0,022	33,119	1

Nota. OR = odds ratio; IC95 % = intervalo de confianza del 95 %; p (Fisher) = prueba exacta de Fisher, dos colas. Los OR se estimaron con corrección de Haldane-Anscombe cuando alguna celda fue 0; los IC95 % son intervalos exactos (tipo Baptista-Pike) calculados sobre la tabla original. En la columna "Tabla 0/1 $\times$ 0/1", las celdas siguen el orden a = Winner & Sí, b = No Winner & Sí, c = Winner & No, d = No Winner & No. Unidad analítica: equipo-año (N = 11); incidencia del evento Global Winner: n = 1. Debido al tamaño muestral reducido y evento raro, los IC son anchos; los p-valores se reportan como referencia.

Como era esperable por el evento raro (n=1), los IC95 % son anchos; retenemos dirección y magnitud como evidencia compatible, más que significación puntual.

La Tabla 10 recoge el Top 12 de profesiones/carreras a nivel de integrante, y la Tabla 11 presenta la distribución por macroárea.

Tabla 10

Top 12 profesiones/carreras

Posición	Profesión/Carrera	Frecuencia	%
1	Ingeniería en Ciencias Computacionales	24	38,7
2	Ingeniería en Mecatrónica	18	29

3	Licenciatura en Diseño Gráfico	2	3,2
4	Ingeniería en Software	2	3,2
5	Ingeniería Naval	1	1,6
6	Ingeniería Eléctrica	1	1,6
7	Producción Audiovisual	1	1,6
8	Entrenamiento terapéutico	1	1,6
9	Ingeniería Multimedia	1	1,6
10	Ingeniería Empresarial	1	1,6
11	Ingeniería en Electrónica	1	1,6
12	Ingeniería en Electrónica y Telecomunicaciones	1	1,6

Nota. % sobre el total con dato de profesión/carrera.

Tabla 11

Distribución por macroárea (nivel integrante).

Macroárea	Frecuencia	%
STEM	55	88,7
Artes/Diseño	4	6,5
Otras	3	4,8

DISCUSIÓN

Los resultados sugieren que la composición académica por sí sola (mayor %STEM o mayor heterogeneidad disciplinaria) no se traduce automáticamente en mejor desempeño local, mientras que la experiencia previa del equipo sí se asocia con SCL-z (puntaje de calidad local estandarizado por año). La desalineación observada entre SCL-z y el veredicto global indica que el juicio internacional podría ponderar dimensiones adicionales, por ejemplo, narrativa del pitch, alineación al desafío y señal de impacto, no plenamente capturadas por la rúbrica local. Esto abre la puerta a ajustar los pesos/criterios o incorporar indicadores complementarios si se busca mayor validez predictiva del resultado global.

Nuestro análisis cuantitativo, basado en ganadores locales (Global Nominees) y el desenlace objetivo Global Winner, no encontró asociaciones lineales fuertes entre (a) carrera/área (STEM vs. no-STEM), (b) composición disciplinaria (mono vs. interdisciplinaria; índice de Blau por macroáreas) y (c) el nivel de éxito del proyecto, los tamaños de efecto fueron pequeños y los intervalos de confianza amplios, coherentes con la baja incidencia del evento Global Winner y la consiguiente potencia estadística limitada. En contraste, la consistencia interna de la rúbrica local de cinco criterios fue alta ($\alpha = 0,896$), lo que respalda la fiabilidad de las calificaciones y la comparabilidad entre ediciones. En conjunto, los resultados sugieren que no existe un único “perfil académico” determinante; el éxito parece depender de prácticas de equipo y factores situacionales (p. ej., ajuste al reto, mentoreo, coordinación), algo congruente con la naturaleza del NASA International Space Apps Challenge, que evalúa múltiples dimensiones y

distingue entre Global Nominee y Global Winner mediante rondas sucesivas de juzgamiento. (NASA, 2024; Schulten & Chounta, 2024).

En nuestro conjunto, la composición fue mayoritariamente STEM ($M \approx 87,6\%$, Mediana = 100 %) y la heterogeneidad disciplinaria mediana fue baja ($IB = 0,00$; $Q3 = 0,382$). La participación femenina promedió $\approx 22,9\%$ (Mediana = 20 %) con IB de género Mediana = 0,320 ($Q3 = 0,444$). Estos descriptivos contextualizan la ausencia de asociaciones robustas entre “mix” académico y desempeño estandarizado, y sugieren priorizar procesos de equipo (experiencia, mentoría, coordinación).

La evidencia más reciente constata el potencial de los hackatones para fomentar trabajo colaborativo, creatividad y resolución de problemas, pero también subraya la heterogeneidad de diseños y la escasez de métricas objetivas de resultado (predominan medidas autorreportadas). Nuestro uso de un desenlace objetivo (Winner vs. Nominee) y de un puntaje estandarizado basado en rúbrica atiende justamente estas llamadas a estandarizar mediciones (Schulten & Chounta, 2024; Oyetade et al., 2024; Rooholamini et al., 2024).

La investigación organizacional y educativa converge en que los efectos de la diversidad (funcional/disciplinaria, de género) sobre el rendimiento suelen ser pequeños y contingentes a procesos y climas de equipo. La meta-análisis de Triana et al. (2021) muestra efectos modestos y mediados por estados emergentes (p. ej., memoria transaccional, coordinación) y conflicto; estudios recientes siguen empleando el índice de Blau para operacionalizar diversidad categórica en equipos (Vedres, 2023; Yoo et al., 2024). En nuestra muestra, la falta de asociaciones robustas entre %STEM/interdisciplinariedad y desempeño coincide con esa visión matizada: la diversidad por sí sola no garantiza mejores resultados si no se acompaña de prácticas de orquestación (roles claros, integración de perspectivas, gestión del tiempo y del pitch). (Triana et al., 2021; Vedres, 2023; Yoo et al., 2024).

Revisiones recientes en educación con hackatones describen que intervenciones pedagógicas estructuradas (mentoría, andamiajes, guías) suelen asociarse a mejores procesos y experiencias de aprendizaje; en cuanto a género, los hallazgos tienden a efectos pequeños y dependientes del contexto y del tipo de resultado (educativo vs. innovación) (Oyetade et al., 2024; Rooholamini et al., 2024; Vedres, 2023). Nuestros diagramas de dispersión y correlaciones, no concluyentes, para mentoría y equilibrio de género se alinean con esa literatura.

Documentar la fiabilidad de los instrumentos antes de comparar grupos es una buena práctica. Aunque α es ampliamente usado, su interpretación depende de tau-equivalencia y unidimensionalidad; en muchos contextos se recomienda ω como estimador más adecuado cuando las cargas difieren (Hayes & Coutts, 2020; Raykov, 2024). En nuestro caso, $\alpha = 0,896$ indica consistencia interna alta; como trabajo futuro, estimar ω total y análisis factorial confirmatorio reforzaría la validez de la puntuación compuesta. (Hayes & Coutts, 2020; Raykov, 2024).

Implicaciones prácticas

1. Gestión de equipos. En lugar de “maximizar STEM”, conviene definir roles y procesos (datos, ciencia, UX, relato) y practicar coordinación y memoria transaccional —las palancas que la meta-evidencia vincula con rendimiento cuando hay diversidad (Triana et al., 2021).
2. Mentoría y andamiaje. La literatura reciente sugiere mentoría estructurada en momentos críticos (definición del problema, ideación y validación temprana), con rúbricas públicas y alineación con criterios oficiales (Oyetade et al., 2024; NASA, 2023).
3. Métricas comparables. Mantener rúbricas estables, capacitación de jueces y conservar registros (p. ej., calificaciones por criterio) mejora la trazabilidad y la comparabilidad entre sedes y años (Schulten & Chounta, 2024; NASA, 2024).

Fortalezas y limitaciones

Fortalezas. Uso de desenlaces objetivos (Winner/Nominee), alta fiabilidad de la rúbrica ($\alpha = 0,896$) y estandarización anual de los puntajes (z por edición).

Limitaciones. (i) baja frecuencia de Global Winner (evento raro), que limita potencia y estabilidad de estimadores; (ii) posible clasificación heurística de carreras en macroáreas (riesgo de misclasificación leve); (iii) diseño observacional; (iv) variables no observadas (calidad del pitch, dificultad del reto, experiencia previa específica con datos NASA). Dado el evento raro, es recomendable contrastar resultados con regresión logística penalizada de Firth o logística exacta, enfoques sugeridos para muestras pequeñas y separación cuasi-completa (Puhr et al., 2021).

Como trabajo futuro, estimar ω (McDonald) y realizar AFC para reforzar la validez de la puntuación compuesta.

Líneas futuras

1. Ampliar muestra con otras sedes/años para aumentar el número de *Winners* y evaluar moderadores (tipo de desafío, modalidad on-site/virtual). (NASA, 2024).
2. Modelos ordinales/continuos que utilicen la puntuación compuesta (en lugar del binario Winner/Nominee) para ganar potencia.
3. Mecanismos de proceso: medir coordinación y uso de mentores como mediadores entre composición y desempeño (Triana et al., 2021; Oyetade et al., 2024).
4. Intervenciones pedagógicas previas (ABP/PBL y *design thinking*) muestran beneficios en desempeño y habilidades de orden superior; su integración como pre-hackatón (4–8 semanas) podría elevar resultados (Zhang & Ma, 2023; Yu et al., 2024)

CONCLUSIONES

En este estudio cuantitativo sobre el NASA Space Apps Challenge (Guayaquil, 2019–2023) no se hallaron asociaciones lineales fuertes entre la formación académica universitaria, operacionalizada como porcentaje de integrantes STEM, composición mono/interdisciplinaria e índice de Blau por macroáreas, y el éxito del proyecto (Global Winner vs. Global Nominee). La evidencia apunta a tamaños de efecto pequeños y a intervalos de confianza amplios, consistentes con la baja incidencia del evento “Winner”. La rúbrica de evaluación demostró alta consistencia interna ($\alpha = 0,896$), lo que respalda la fiabilidad de los puntajes y la comparabilidad entre ediciones, pero sugiere que el desempeño final depende de factores adicionales al perfil disciplinario de los equipos.

A nivel práctico, los hallazgos desalientan la búsqueda de un “perfil ganador” estrictamente académico y favorecen estrategias centradas en la orquestación del trabajo: definición clara de roles (datos, ingeniería, ciencia, diseño/UX, narrativa), prácticas de coordinación y mentoría estructurada alineada con los criterios oficiales. La diversidad disciplinaria es deseable, pero su valor parece materializarse cuando existen procesos que canalizan esa heterogeneidad hacia la resolución efectiva del reto. De cara a la organización de futuras ediciones, conviene fortalecer la trazabilidad de las decisiones (guías de evaluación y capacitación de jueces) y mantener métricas objetivas y comparables que reflejen la calidad de las soluciones.

El veredicto internacional pondera dimensiones (pitch, alineación, impacto) no capturadas del todo por la rúbrica local, lo que explica la desalineación observada en 2021.

Como agenda futura, se recomienda ampliar la base de datos (más sedes y años) para incrementar la potencia estadística y evaluar moderadores (tipo de desafío, modalidad del evento). Metodológicamente, el uso de modelos adecuados para eventos raros (p. ej., logística penalizada de Firth o enfoques exactos) y de desenlaces continuos/ordinales basados en la puntuación compuesta podría ofrecer estimaciones más estables y matizadas. Asimismo, incorporar medidas breves y estandarizadas de procesos de equipo (coordinación, uso de mentores, decisiones de diseño) permitiría comprender mejor cómo la composición académica interactúa con las prácticas para producir soluciones exitosas.

REFERENCIAS

- Abdul Rahman, H., Noraidi, A. A., Khalid, A. N. H., Mohamad-Adam, A. Z., Zahari, N. H., & Tuming, N. E. (2025). Practical guide to calculate sample size for chi-square test in biomedical research. *BMC Medical Research Methodology*, 25, 144. <https://doi.org/10.1186/s12874-025-02584-4>
- Bocianowski, J., Wrońska-Pilarek, D., Krysztofiak-Kaniewska, A., Matusiak, K., & Wiatrowska, B. (2024). Comparison of Pearson's and Spearman's

- correlation coefficients for selected traits of *Pinus sylvestris* L. *Biometrical Letters*, 61(2), 115–135. <https://doi.org/10.2478/bile-2024-0008>
- Edelsbrunner, P. A., Maier, M., Jauk, E., & Dörfler, T. (2025). The Cronbach's Alpha of domain-specific knowledge tests: A meta-analysis. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-024-09982-y>
- Farkas, B. C., Németh, D., & Janacsek, K. (2023). The complexity of measuring reliability in learning tasks: An illustration using the ASRT task. *Behavior Research Methods*, 55(6), 2925–2944. <https://doi.org/10.3758/s13428-022-02038-5>
- García, M. B. (2023). Fostering an innovation culture in the education sector: A scoping review and bibliometric analysis of hackathon research. *Innovative Higher Education*. <https://doi.org/10.1007/s10755-023-09651-y>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- McGrath, O., & Burke, K. (2024). Binomial confidence intervals for rare events: Importance of defining margin of error relative to magnitude of proportion. *The American Statistician*, 78(4), 437–449. <https://doi.org/10.1080/00031305.2024.2350445>
- Miličević, A., Despotović-Zrakić, M., Stojanović, D., Suvajžić, M., & Labus, A. (2024). Academic performance indicators for the hackathon learning approach—The case of the blockchain hackathon. *Journal of Innovation & Knowledge*, 9(3), 100501. <https://doi.org/10.1016/j.jik.2024.100501>
- NASA Space Apps Challenge. (2023). Global winners. <https://www.spaceappschallenge.org/2023/awards/global-winners/>
- NASA Space Apps Challenge. (2024). Global winners. <https://www.spaceappschallenge.org/nasa-space-apps-2024/awards/global-winners/>
- Orowo, P. (2025). Confidence intervals for the model performance metrics: A tutorial. *Statistical Methods in Medical Research*. <https://doi.org/10.1177/15741699251350368>
- Oyetade, K., Zuva, T., & Harmse, A. (2024). Evaluation of the impact of hackathons in education. *Cogent Education*, 11, 2392420. <https://doi.org/10.1080/2331186X.2024.2392420>
- Puhr, R., Heinze, G., Nold, M., Lusa, L., & Geroldinger, A. (2021). Firth's logistic regression with rare events: Accurate effect estimates and predictions? *arXiv preprint*. <https://doi.org/10.48550/arXiv.2101.07620>
- Raykov, T. (2024). Coefficient alpha and reliability of communication science measures: A note. *Communication Methods and Measures*, 18(1), 1–6. <https://doi.org/10.1080/19312458.2024.2302125>
- Rooholamini, A., Curiel, J., & Reimens, J. (2024). Health profession education hackathons: A scoping review. *BMC Medical Education*, 24, 484. <https://doi.org/10.1186/s12909-024-04901-5>

- Schulten, C., & Chounta, I.-A. (2024). How do we learn in and from hackathons? A systematic literature review. *Education and Information Technologies*, 29, 20103–20134. <https://doi.org/10.1007/s10639-024-12668-1>
- Scott, J. M., & Chang, C. (2024). The imperative to promote diversity post-*Students for Fair Admissions v. Harvard*: Measurement and implications. *The Journal of Higher Education*, 95(6), 907–925. <https://doi.org/10.1080/00221546.2024.2385114>
- Triana, M. del C., Kim, K., Byun, S. Y., Delgado, D. M., & Arthur, W. (2021). The relationship between team deep-level diversity and team performance: A meta-analysis of the main effect, moderators, and mediating mechanisms. *Journal of Management Studies*, 58(8), 2137–2179. <https://doi.org/10.1111/joms.12670>
- Tyagi, A., Salhotra, R., Agrawal, A., & Kumar, R. (2023). Use of Pearson and Spearman correlation testing in Indian anesthesia journals: An audit. *Indian Journal of Anaesthesia*, 66(11), 827–834. https://doi.org/10.4103/joacp.joacp_13_22
- UNESCO. (2020). Education for sustainable development: A roadmap. UNESCO. <https://doi.org/10.54675/YFRE1448>
- Vedres, B., Vasarhelyi, K., Török, J., & Palla, G. (2023). Inclusion unlocks the creative potential of gender diversity in teams. *Nature Communications*, 14, 4810. <https://doi.org/10.1038/s41467-023-40276-8>
- Yang, H., Li, Z., & Zhou, W. (2023). Board diversity and the marginal value of corporate cash holdings. *Pacific-Basin Finance Journal*, 79, 102089. <https://doi.org/10.1016/j.pacfin.2023.102089>
- Yoo, H.-S., Kim, D., Liu, J., & Joo, J. (2024). The interaction of inter-organizational diversity and team composition on innovation performance. *Information & Management*, 61(8), 104115. <https://doi.org/10.1016/j.im.2024.104115>
- Yu, Q., Manalo, E., & Tang, X. (2024). A meta-analysis of the effects of design thinking on student learning. *Humanities and Social Sciences Communications*, 11, 686. <https://doi.org/10.1057/s41599-024-03237-5>
- Zhang, L., & Ma, Y. (2023). A study of the impact of project-based learning on student learning effects: A meta-analysis study. *Frontiers in Psychology*, 14, 1202728. <https://doi.org/10.3389/fpsyg.2023.1202728>