

De lo tradicional a lo computacional: enseñanza y aprendizaje de la estadística con Jupyter Notebook

From Traditional to Computational: Teaching and Learning Statistics with Jupyter Notebook



Fabricio Vladimir Vines-Vinces¹

Universidad Nacional de Loja (UNL) | Loja - Ecuador | CP 110101 |

<https://orcid.org/0000-0002-5912-6367>

<https://ror.org/03a5x6z77>

Wilmer Ríos-Cuesta²

Universidad de Antioquia (UdeA) | Medellín - Colombia | CP 050001 |

<https://orcid.org/0000-0001-8129-2137>

<https://ror.org/03bp5hc83>

Correspondencia: fabricio.vinces@unl.edu.ec

<https://doi.org/10.26423/xw1a4c13>

Páginas: 55-68

RESUMEN

El aprendizaje de la estadística resulta desafiante cuando se aborda de forma abstracta y desvinculada de contextos reales. Esta investigación presenta una experiencia didáctica estadística descriptiva en educación superior mediante Jupyter Notebook con librerías de Python y *dataset* de contexto real. Se adoptó un estudio de caso con 28 estudiantes de la carrera de Pedagogía de las Ciencias Experimentales, Matemáticas y Física de una universidad pública ecuatoriana. La intervención se estructuró en tres fases cíclicas (inicio, desarrollo y cierre) implementado en 10 horas presenciales durante dos semanas. El análisis cualitativo de testimonios estudiantiles evidenció que la experiencia fue percibida como innovadora y satisfactoria, favoreciendo la comprensión de conceptos estadísticos fundamentales, el manejo de *datasets* y la aplicación de técnicas de visualización. Los resultados sugieren que la herramienta Jupyter Notebook puede constituirse en una herramienta pedagógica pertinente para favorecer experiencias activas y contextualizadas de aprendizaje estadístico, siempre que su uso esté acompañado de una adecuada mediación docente con competencias básicas en Python y en diseño de entornos mediados por tecnología.

Palabras clave: estadística; experiencia de aprendizaje; innovación pedagógica; método de enseñanza.

ABSTRACT

Learning statistics is challenging when approached in an abstract manner and disconnected from real-world contexts. This research presents a descriptive statistic teaching experience in higher education using Jupyter Notebook with Python libraries and real-world *datasets*. A case study was conducted with 28 students in the Experimental Sciences, Mathematics, and Physics Education program at a public university in Ecuador. The intervention was structured into three cyclical phases (initiation, development, and closure) and implemented over 10 hours of in-person instruction spread across two weeks. Qualitative analysis of student testimonials revealed that the experience was perceived as innovative and satisfying, promoting understanding of fundamental statistical concepts, the handling of *datasets*, and the application of visualization techniques. The results suggest that Jupyter Notebook can serve as a relevant pedagogical tool for fostering active and contextualized statistical learning experiences, provided its use is accompanied by appropriate faculty guidance from instructors with basic proficiency in Python and in the design of technology-mediated learning environments.

Keywords: statistics, learning experience, pedagogical innovation, teaching method.

Recepción: 9 abril 2026 | Aprobación: 23 junio 2026 | Publicación: 30 junio 2026

¹ Magíster en Estadística Aplicada, por la Universidad Politécnica Estatal del Carchi (UPEC) – Ecuador; Magíster en Modelación y Ciencia Computacional, por la Universidad de Medellín – Colombia.

² Doctor en Educación, por la Universidad del Valle – Colombia.

INTRODUCCIÓN

La estadística ocupa un lugar central en la formación de profesores de ciencias y matemáticas, ya que constituye una herramienta fundamental tanto para la investigación educativa como para el análisis crítico de información en contextos sociales y científicos. Sin embargo, su enseñanza en educación superior enfrenta desafíos persistentes y bien documentados, entre ellos, que los estudiantes tienden a memorizar procedimientos sin comprender su significado, muestran dificultades para interpretar resultados en contextos reales y desarrollan actitudes negativas hacia la materia cuando la instrucción se limita al cálculo algorítmico (e.g. Garfield y Ben-Zvi, 2007; Biehler *et al.*, 2018). Esta situación se agudiza en programas de formación docente, donde los futuros profesores necesitan dominar los conceptos estadísticos como parte del saber matemático del profesor, así como desarrollar competencias para enseñarlos de manera significativa a sus propios estudiantes (Vásquez y Alsina, 2021).

En América Latina, tiene particularidades que merecen especial atención. La enseñanza de la estadística en educación superior suele estar desconectada de datos reales y de herramientas computacionales vigentes, esto limita el desarrollo del pensamiento estadístico y la literacidad de datos en los estudiantes (Larreamendy-Joerns *et al.*, 2005; Ríos-Cuesta, 2021). A esto se suma que los reportes de investigación sobre el uso de entornos computacionales para la enseñanza de estadística en idioma español son escasos, lo que representa una brecha importante en la literatura regional (Cardoso *et al.*, 2019). En este aspecto, Ecuador no es la excepción, pues la formación estadística en las carreras de Pedagogía de Ciencias Exactas demanda estrategias innovadoras que respondan a las exigencias curriculares y al perfil tecnológico que la sociedad actual requiere de sus egresados.

Frente a este panorama, la integración de herramientas computacionales interactivas emerge como una alternativa pedagógica con fundamento empírico creciente. En particular, los entornos de cuadernos computacionales, como Jupyter Notebook, permiten combinar código ejecutable, texto explicativo, visualizaciones y datos en un mismo documento, lo que favorece un aprendizaje activo, situado y reproducible (Rule *et al.*, 2018; Çetinkaya-Rundel y Ellison, 2021), a diferencia de los enfoques tradicionales centrados en el cálculo manual. Estos entornos interactivos le permiten que los estudiantes trabajen con conjuntos de datos (*dataset*), generen gráficas estadísticas y complejas e interpreten resultados que contribuyan al desarrollo del pensamiento estadístico y computacional (Nwulu *et al.*, 2021; Kaganovskiy, 2025). A pesar de esto, la eficacia de estas herramientas no es automática, puesto que depende de manera crítica del diseño didáctico que las acompaña y de la mediación docente que guía su uso

(Amoudi y Tbaishat, 2023; Engelbrecht y Borba, 2024).

Si bien existe las evidencias sobre el potencial pedagógico de Jupyter Notebook, reportadas en inglés principalmente y en contextos universitarios en Norte América, son escasos los estudios que documentan experiencias de aula, concretamente de la enseñanza estadística en el contexto latinoamericano en la formación académica de docente. Esta brecha limita la posibilidad de que otros docentes de la región, adapten e implementen estrategias similares con respaldo empírico. Por tanto, documentar, analizar y compartir estas experiencias constituye una necesidad tanto para la investigación en educación matemática como para el fortalecimiento de la práctica docente.

Pensamiento estadístico aplicado a la enseñanza de la estadística descriptiva

La educación estadística ha desarrollado un corpus teórico propio que distingue entre el conocimiento procedimental, el cálculo de medidas, y el pensamiento estadístico, entendido como la capacidad de razonar con y sobre los datos para construir argumentos y tomar decisiones en contextos de incertidumbre (Wild y Pfannkuch, 1999). Esta distinción es fundamental, mientras el primero puede lograrse mediante la memorización de algoritmos, el segundo requiere que el estudiante comprenda por qué se calcula cada medida, qué significa en su contexto y cómo se relaciona con la variación inherente a los datos reales (Garfield y Ben-Zvi, 2007).

Por otro lado, la investigación en educación estadística ha documentado forma consistente que los enfoques tradicionales basados en el cálculo manual con datos pequeños y artificiales limitan el desarrollo del pensamiento estadístico (Biehler *et al.*, 2018; Larreamendy-Joerns *et al.*, 2005). Los estudiantes aprenden a ejecutar fórmulas, pero muestran dificultades para interpretar los resultados, identificar patrones en distribuciones o vincular los estadísticos con preguntas relevantes sobre el mundo real (Garfield y Ben-Zvi, 2007; Ríos-Cuesta, 2021). En el caso específico de la formación de profesores de matemáticas, es una limitación con implicaciones adicionales, puesto que un docente que no ha desarrollado pensamiento estadístico difícilmente podrá extenderlo a sus propios estudiantes (Vásquez y Alsina, 2021).

Frente a este diagnóstico, diversas propuestas curriculares han enfatizado la necesidad de incorporar datos reales, herramientas computacionales y tareas que requieren interpretación contextualizada para el desarrollo del pensamiento estadístico (Fielding *et al.*, 2025; Howley y Roberts, 2020). Desde esta perspectiva, los problemas estadísticos auténticos, aquellos que involucran conjuntos de datos reales, preguntas abiertas y decisiones analíticas, funcionan como situaciones que

requieren exploración y no solo aplicación mecánica de procedimientos (Schoenfeld, 2016). Esta idea es fundamental en este estudio, el uso de *datasets* abiertos de gran volumen como detonadores del aprendizaje estadístico que va más allá del cálculo y se orienta en la resolución de problemas que demandan comprensión, decisión e interpretación.

Aprendizaje mediado por tecnologías digitales y modelo TPACK

Las tecnologías digitales tienen el potencial de enriquecer el aprendizaje matemático en la medida en que amplían los registros de representación disponibles para el estudiante, favorecen la modelación y visualización de relaciones, y permiten someter a prueba conjeturas de manera dinámica (Campo-Meneses y García-García, 2020). En el contexto específico de la estadística, las herramientas computacionales permiten además trabajar con volúmenes de datos imposibles de manejar manualmente, lo que transforma cualitativamente la naturaleza de las tareas disponibles para el aula (Engelbrecht y Borba, 2024).

Sin embargo, diferentes estudios han mostrado con claridad que la incorporación de tecnología por sí misma, no produce mejoras automáticas en el aprendizaje, dado que su eficacia depende de la calidad del diseño didáctico que la acompaña (Darmanova *et al.*, 2025; Ye *et al.*, 2024). Para dar cuenta de esta complejidad, el modelo TPACK (Technological Pedagogical Content Knowledge) de Mishra y Koehler (2006) ofrece un marco analítico pertinente: la integración efectiva de tecnología requiere que el docente articule de manera coherente su conocimiento tecnológico, su conocimiento pedagógico y su conocimiento disciplinar. En ausencia de alguno de estos componentes, la tecnología tiende a incorporarse de manera instrumental, sin transformar las prácticas de enseñanza ni los procesos de aprendizaje (Engelbrecht y Borba, 2024).

Esta consideración es especialmente relevante en contextos de formación inicial docente, donde los estudiantes son, simultáneamente, aprendices de la estadística y futuros profesores que observan e internalizan modelos de enseñanza. El diseño de la experiencia didáctica que se reporta en este artículo tomó en consideración estos dos aspectos: las decisiones didácticas tales como la selección de datos, secuenciación de tareas, uso de analogías y momentos de reflexión, fueron tan deliberadas como la elección de la herramienta computacional.

Pensamiento computacional y Jupyter Notebook como entorno de aprendizaje

El pensamiento computacional se entiende como un conjunto de procesos cognitivos relacionados con la resolución de problemas utilizando algunas

subhabilidades, tales como, la descomposición del problema, reconocimiento de patrones, abstracción, diseño de algoritmos entre otras, que permiten abordar problemas complejos de manera sistemática (Wing, 2006). En el campo de la educación estadística, este tipo de pensamiento cobra especial relevancia porque el análisis de datos reales exige precisamente estas capacidades: descomponer un conjunto de datos complejo en preguntas manejables, identificar patrones en distribuciones, abstraer variables relevantes y diseñar secuencias de código que operacionalicen las decisiones analíticas (Çetinkaya-Rundel y Ellison, 2021; Pruim *et al.*, 2023).

En el caso de Jupyter Notebook constituye un entorno especialmente adecuado para integrar pensamiento computacional y aprendizaje estadístico. A diferencia de los entornos tradicionales de programación, este cuaderno interactivo combina en un único documento celdas de código ejecutable, texto explicativo en formato enriquecido, ecuaciones matemáticas y visualizaciones, lo que permite que el estudiante construya una narrativa computacional coherente que integra el qué, el cómo, y el porqué de cada decisión analítica (Rule *et al.*, 2018). Esta característica lo convierte en una herramienta particularmente valiosa para la educación estadística, ya que favorece que el código emerja como respuesta a una necesidad analítica concreta, y no como un fin en sí mismo (Kaganovskiy, 2025; Krüger, 2022).

Estudios empíricos recientes respaldan la eficacia pedagógica de esta herramienta en distintos contextos universitarios. Amoudi y Tbaishat (2023) reportaron que Jupyter Notebook mejora la comprensión conceptual cuando se acompaña de preguntas de reflexión estructuradas. Nwulu *et al.* (2021) documentaron actitudes favorables hacia su uso en estudiantes de ingeniería. Por otro lado, Carrano *et al.* (2020) destacan tres atributos pedagógicos centrales: permiten aplicaciones de la vida real, cultivan la literacidad computacional y están estructurados para fomentar la autonomía. En conjunto, estos hallazgos sugieren que el valor pedagógico de Jupyter Notebook no reside en sus características técnicas *per se*, la fortaleza está en la forma en que el diseño didáctico aprovecha dichas características para promover aprendizajes profundos y situados.

En respuesta a esta necesidad, el presente estudio se orienta por la siguiente pregunta de investigación: ¿qué aprendizajes consolidan los estudiantes de formación inicial docente al trabajar estadística descriptiva mediante Jupyter Notebook con datos de contextos reales, y cómo perciben esta experiencia?, con el fin de identificar: (a) los aprendizajes que los estudiantes consolidan, (b) las percepciones sobre la experiencia y (c) las condiciones pedagógicas que favorecen su implementación.

MATERIALES Y MÉTODOS

Diseño de investigación y pregunta orientadora

Para responder a la pregunta de investigación planteada al cierre de la Introducción, se adoptó un enfoque metodológico cualitativo bajo el diseño de estudio de caso intrínseco (Stake, 2010). Este diseño resulta pertinente porque el interés del estudio no es generalizar resultados a otros contextos, sino comprender en profundidad la singularidad de una experiencia didáctica específica: sus condiciones de implementación, los aprendizajes que produjo y las percepciones de quienes participaron en ella. El caso está constituido por una intervención didáctica de estadística descriptiva mediada por la herramienta Jupyter Notebook, para un grupo concreto de estudiantes en una universidad pública ecuatoriana.

Participantes y contexto

Participaron 28 estudiantes de cuarto ciclo de la carrera de Pedagogía de las Ciencias Experimentales: Matemáticas y Física de una universidad pública ecuatoriana de Loja, durante el primer semestre de 2022. Sus edades estaban comprendidas entre los 19 y los 24

años, corresponden al perfil estudiantil de formación inicial para docente en carreras de grado. Los participantes cursaban la asignatura de estadística como parte de su formación disciplinar y no contaban con experiencia previa en programación ni en el uso de herramientas computacionales para el análisis de datos. Anteriormente, la asignatura se desarrollaba con métodos tradicionales, cálculo manual con calculadora y conjuntos de datos reducidos, lo que hacía del grupo un caso relevante para explorar la transición hacia entornos computacionales interactivos. Los participantes fueron informados sobre los propósitos del estudio y otorgaron su consentimiento informado, para que sus producciones y testimonios fueran utilizados con fines académicos, preservando en todo momento el anonimato.

Diseño de la intervención didáctica

La intervención se organizó en tres fases didácticas progresivas, inicio, desarrollo y cierre, inspiradas en el ciclo de investigación estadística (Wild y Pfannkuch, 1999), con una duración total de diez horas presenciales distribuidas en dos semanas. La Tabla 1 sintetiza su estructura.

Tabla 1. Estructura de la intervención didáctica

Fase	Duración	Actividad principal	Herramienta	Modalidad
Inicio: instrucción teórica	4 horas	Clases expositivas, analogías, resolución de problemas con datos reducidos	Pizarra, calculadora	Presencial guiada
Desarrollo: transición	4 horas	Introducción a Jupyter Notebook; análisis de <i>dataset</i> de calificaciones universitarias	Jupyter Notebook, Python, Pandas, Seaborn	2h guiada + 2h autónoma
Cierre: aplicación autónoma	2 horas	Análisis independiente del <i>dataset</i> SRI_Vehiculos_Nuevos_2022	Jupyter Notebook, Python	Autónoma

Nota: La distribución horaria corresponde al tiempo presencial. Cada fase incluyó momentos de reflexión colectiva sobre los resultados obtenidos.

Fase de inicio: instrucción teórica

Durante esta fase se llevó a cabo la instrucción teórica de los contenidos estadísticos mediante clases expositivas, uso de analogías y resolución de problemas con volúmenes reducidos de datos. El énfasis se puso en la comprensión conceptual de las medidas estadísticas, tendencia central, dispersión, y en la interpretación crítica de representaciones gráficas. Este enfoque responde a la evidencia de que la comprensión conceptual previa es una condición necesaria para que el uso posterior de herramientas computacionales produzca aprendizajes significativos y no se reduzca a la ejecución de código sin comprensión (Amoudi y Tbaishat, 2023).

Fase de desarrollo: transición de lo físico a lo digital

Esta fase tuvo como propósito introducir a los estudiantes en el uso de Jupyter Notebook como entorno de trabajo estadístico. Para facilitar la transición desde los medios físicos habituales, pizarra y cuaderno, hacia el entorno digital, se empleó una estrategia pedagógica basada en analogías conceptuales que evitó deliberadamente el término “programación”, frecuentemente asociado por los estudiantes con abstracción y dificultad técnica, y se lo reemplazó por comparaciones accesibles: el lenguaje de programación fue presentado como una biblioteca; las librerías, como los estantes de esa biblioteca; y las funciones, como los libros individuales. Esta decisión pedagógica buscó reducir la ansiedad tecnológica inicial y construir una disposición favorable hacia la herramienta (Hamer et al., 2024).

Una vez instalado³ el entorno de trabajo Anaconda para la gestión de librerías y herramientas de análisis de datos como Python, Jupyter Notebook, Pandas y NumPy. Luego, los estudiantes analizaron un *dataset* de calificaciones universitarias (33 registros, 8 variables) de cuatro carreras, que contenía variables continuas (notas numéricas) y categóricas (lugar de residencia). El conjunto estaba conformado por ocho variables: cuatro variables continuas correspondientes a las calificaciones numéricas obtenidas en cada carrera (escala de 0 a 20), una variable categórica de identificación de la carrera (Educación Inicial, Comunicación Social, Mecatrónica y Pedagogía) y variables categóricas adicionales como el lugar de residencia del estudiante.

Dado el volumen reducido del conjunto de datos (33 registros), fue revisado previamente por el docente - investigador para verificar la ausencia de valores perdidos, duplicados e inconsistencias antes de su uso en el aula, tal condición favorece la reproducibilidad de la experiencia. Las tareas analíticas que incluyeron: seis tareas analíticas, cada una orientada por una intención didáctica específica vinculada al indicador estadístico trabajado. Cada tarea requirió tanto la implementación del código como la interpretación contextualizada del resultado obtenido.

- 1) Cálculo de promedios por carrera, con la intención de que los estudiantes describieran la tendencia central del rendimiento de cada grupo y la compararan entre carreras.
- 2) Identificación de la carrera con menor dispersión relativa mediante el coeficiente de variación (CV), con el propósito de que los estudiantes compararan la homogeneidad de los resultados entre carreras con medias diferentes, reconociendo cuáles presentaban un rendimiento más estable.
- 3) Cálculo del percentil 78 para una carrera específica, orientado a que los estudiantes interpretaran la posición relativa de una observación dentro de la distribución y comprendieran el significado de los percentiles en términos poblacionales.
- 4) Determinación del tercer cuartil para otra carrera, con la intención de que los estudiantes segmentaran la distribución en cuartos y leyeran el valor que delimita al 75 % de los datos.
- 5) Elaboración de un diagrama de barras según el lugar de residencia, para que los estudiantes representaran la frecuencia de una variable categórica y seleccionaran la técnica gráfica pertinente para datos cualitativos.
- 6) Construcción de diagramas de caja por carrera, con la intención de que los estudiantes visualizaran simultáneamente la mediana, la dispersión y los

valores atípicos, e integraran varias medidas en una sola representación.

Fase de cierre: aplicación autónoma

Como culminación de la experiencia, los estudiantes trabajaron de manera autónoma con la base de datos pública SRI_Vehiculos_Nuevos_2022 (Servicio de Rentas Internas, 2022), un *dataset* de 73 848 registros y 20 variables en formato MS Excel (.xlsx). El archivo original fue descargado desde el portal de datos abiertos del Servicio de Rentas Internas del Ecuador y depurado previamente por el docente-investigador mediante la eliminación de filas duplicadas y la verificación de consistencia en las variables clave antes de su distribución a los estudiantes. Una vez depurado, el conjunto quedó conformado por 73 848 registros y 20 variables, de las cuales se trabajaron principalmente: clase de vehículo (categórica), país de origen (categórica), tipo de servicio (categórica), tipo de combustible (categórica) y año de modelo (discreta). Esta especificación contribuye a la reproducibilidad del estudio. Las tareas propuestas fueron: (1) calcular la frecuencia de vehículos registrados por clase; (2) representar gráficamente mediante diagramas de barras las variables: clase de vehículo, país de origen y tipo de servicio; (3) visualizar la relación entre clase de vehículo y tipo de combustible mediante una técnica gráfica apropiada; y (4) analizar la distribución de vehículos por año de modelo. La autonomía en esta fase fue intencional: buscaba evaluar si los estudiantes habían internalizado el flujo de trabajo estadístico - computacional sin depender de la guía directa del docente.

Recolección y análisis de datos

Los datos del estudio provenían de dos fuentes complementarias. La primera fueron las producciones computacionales de los 28 participantes: los cuadernos de Jupyter Notebook con el código implementado, los resultados obtenidos y las interpretaciones escritas, que constituyen evidencia directa de los aprendizajes consolidados. La segunda fueron dos entrevistas semiestructuradas realizadas al finalizar la intervención a dos estudiantes seleccionados de manera intencional (por conveniencia) por el docente investigador, con el criterio de representar perfiles diferenciados de desempeño durante la experiencia: uno con desempeño alto y otro con desempeño medio. Esta selección responde a las convenciones del estudio de caso intrínseco (Stake, 2010), en el que la profundidad del análisis de casos particulares tiene prioridad sobre la amplitud muestral. En estudios cualitativos exploratorios, dos entrevistas semiestructuradas pueden aportar evidencia ilustrativa sobre percepciones cuando el objetivo no es establecer representatividad estadística, sino documentar experiencias significativas

³ Para detalles, ver tutorial en el minuto 3:39 (Emociones en las Matemáticas, 2022)

desde la perspectiva de los propios participantes. Las entrevistas se orientaron mediante una pregunta central: ¿consideran que lograron aprender estadística descriptiva mediante el uso de Jupyter Notebook?, complementada con preguntas de seguimiento sobre aspectos específicos de la experiencia.

El análisis de las producciones computacionales se realizó mediante revisión interpretativa orientada por tres categorías predefinidas, vinculadas a los objetivos de aprendizaje de la intervención: (a) corrección técnica del código, entendida como la capacidad de implementar funciones de librería de Pandas y Seaborn sin errores de sintaxis y con una lógica de flujo adecuada; (b) pertinencia de las visualizaciones

seleccionadas, valorada según la adecuación del tipo de gráfico a la naturaleza de las variables representadas; y (c) idoneidad de las interpretaciones producidas, evaluada como la capacidad del estudiante para articular el resultado numérico o gráfico con su significado, en el contexto del *dataset* analizado. Estas categorías se aplicaron a los cuadernos Jupyter de los 28 participantes y sus resultados se sintetizan en la Tabla 2. El análisis de las entrevistas siguió los procedimientos del análisis de contenido cualitativo (Stake, 2010): transcripción, lectura flotante, identificación de unidades de significado, agrupación en categorías temáticas emergentes relacionadas con la percepción de la experiencia didáctica y los aprendizajes autorreportados.

Tabla 2. Dimensiones y criterios de análisis de las producciones computacionales de los participantes

Dimensión evaluada	Criterio de análisis	Evidencia observada
Corrección técnica del código	Implementación de funciones sin errores de sintaxis; lógica de flujo adecuada	Uso correcto de <i>describe()</i> , <i>groupby()</i> , <i>size()</i> , <i>boxplot()</i>
Pertinencia de las visualizaciones	Adecuación del tipo de gráfico a la naturaleza de las variables (categórica/continua)	Diagramas de barras, boxplot, histogramas, gráficos de conteo
Idoneidad de las interpretaciones	Capacidad de articular el resultado estadístico con su significado en el contexto del <i>dataset</i>	Enunciados que relacionan cuartiles, percentiles o frecuencias con la población analizada
Autonomía en la aplicación	Resolución de tareas sin mediación directa del docente en la fase de cierre	Cuadernos completados de manera independiente con la base SRI_Vehiculos_Nuevos_2022

RESULTADOS

Los hallazgos se organizan en torno a tres dimensiones que emergen de la pregunta de investigación: 1) el desarrollo técnico-estadístico evidenciado en las producciones computacionales durante el ejercicio guiado, 2) el desempeño autónomo en el análisis del *dataset* vehicular y los hallazgos emergentes, y 3) las percepciones de los estudiantes sobre la experiencia, recogidas mediante entrevista semiestructurada.

Producciones computacionales: ejercicios guiados con el *dataset* calificaciones

Las producciones de los estudiantes durante el ejercicio guiado (Figura 1) muestran indicios de dos niveles de logro complementarios. En el plano técnico, los estudiantes utilizaron la librería Pandas para obtener estadísticos descriptivos mediante la función *describe()*. Asimismo, emplearon la librería Seaborn para generar diagramas de caja mediante la función *boxplot()*, sin presentar errores de sintaxis relevantes tras la fase de instrucción. En el plano interpretativo, los estudiantes elaboraron enunciados que evidencian comprensión del significado contextual de los resultados. Un ejemplo representativo corresponde a la interpretación realizada a partir del análisis de cuartiles: el 25 % de los estudiantes de Educación Inicial obtuvo una calificación igual o inferior a 15,35 (escala de 0 a 20, siendo 14 la

nota mínima aprobatoria). Este tipo de enunciado trasciende la lectura del dato numérico, al vincularlo con su significado distribucional y con el contexto académico de las calificaciones.

En la Figura 1 se muestra la integración lograda por los estudiantes entre código, resultado e interpretación en un mismo documento de Jupyter Notebook, lo que permite seguir el razonamiento estadístico de manera transparente y reproducible.

Aplicación autónoma: análisis del *dataset* SRI_Vehiculos_Nuevos_2022

Durante la fase de cierre, los 28 participantes trabajaron de manera independiente con el *dataset* del Servicio de Rentas Internas. Los resultados se organizaron según las cuatro tareas propuestas.

Tarea 1: frecuencia por clase de vehículo.

Los estudiantes aplicaron *groupby()* combinado con *size()* sobre el *dataframe* depurado para generar tablas de frecuencia absoluta por clase de vehículo. El procedimiento demuestra comprensión del flujo de trabajo completo: identificación de registros duplicados, depuración del conjunto de datos mediante la generación de un *dataframe* sin duplicados (*no_duplicates_df*) y obtención de resultados estructurados (Figura 2a)

Importamos librerías (desde Python)

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
```

¿cuáles de las carreras tiene mayor promedio?

```
media_todas_obs = nota["Educacion_inicial"].mean()
print("la media de Educacion_inicial es de:", round(media_todas_obs, 2))

media_todas_obs = nota["ing_mecatronica"].mean()
print("la media de ing_mecatronica es de:", round(media_todas_obs, 2))

media_todas_obs = nota["Comunicación_social_A"].mean()
print("la media de Comunicación_social_A es de:", round(media_todas_obs, 2))

media_todas_obs = nota["Comunicación_social_B"].mean()
print("la media de Comunicación_social_B es de:", round(media_todas_obs, 2))
```

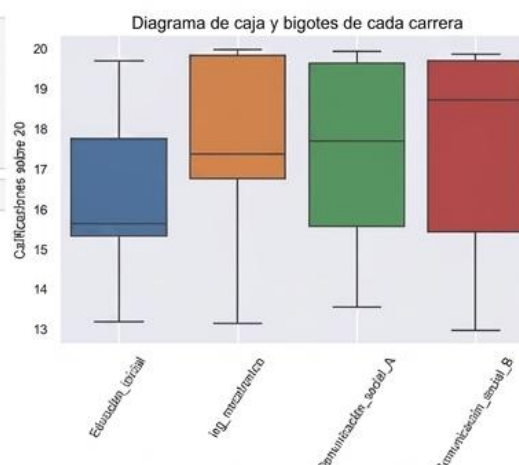
La media de Educacion_inicial es de: 16.3
 La media de ing_mecatronica es de: 17.98
 La media de Comunicación_social_A es de: 17.47
 La media de Comunicación_social_B es de: 17.84

Represente por medio del gráfico de caja y bigotes las calificaciones de cada carrera. Interprete.

```
figura=plt.figure(1,figsize=(7,5))
sns.boxplot(data=df)
plt.xticks(rotation=69)
plt.ylabel("Calificaciones sobre 20")
plt.title("Diagrama de caja y bigotes de cada carrera",size=14)
plt.grid()
plt.show()
```

```
newd(df.describe(),1)
```

	Educacion_inicial	ing_mecatronica	Comunicación_social_A	Comunicación_social_B
count	33.00	33.00	33.00	33.00
mean	16.30	17.98	17.47	17.84
var	1.47	2.40	2.15	2.19
min	13.20	13.15	13.58	12.98
25%	13.33	16.77	15.58	15.44
50%	15.65	17.78	17.72	18.73
75%	17.76	17.85	18.88	18.73
max	19.72	20.00	19.96	19.87



Interpretación

Educación inicial

En la carrera de educación inicial, la nota mínima es 13.19 y la nota máxima es 19.71. El cuartil 1 le corresponde un valor de 15.35, por lo que el 25% cuando mucho llega a 15.33, mientras que el otro 75% son notas mayores a dicho valor. El cuartil 2 le corresponde un valor de 16.85, por lo que el 50% cuando mucho llega a 15.65, mientras que el otro 50% son mayores a dicho valor. El cuartil 3 le corresponde un valor de 17.75, por lo que el 75% cuando mucho llega a 17.75, mientras que el otro 25% son notas mayores a dicho valor.

Ingeniería en mecánica

En la carrera de ingeniería en mecánica, la nota mínima es 13.14 y la nota máxima es 20.00. El cuartil 1 le corresponde un valor de 16.76, por lo que el 25% cuando mucho llega a 16.78, mientras que el otro 75% son notas mayores a dicho valor. El cuartil 2 le corresponde un valor de 17.38, por lo que el 50% cuando mucho llega a 17.38, mientras que el otro 50% son mayores a dicho valor. El cuartil 3 le corresponde un valor de 19.84, por lo que el 75% cuando mucho llega a 19.84, mientras que el otro 25% son notas mayores a dicho valor.

Comunicación Social A

En la carrera de Comunicación Social, la nota mínima es 13.56 y la nota máxima es 19.96. El cuartil 1 le corresponde un valor de 15.58, por lo que el 25% cuando mucho llega a 15.58, mientras que el otro 75% son notas mayores a dicho valor. El cuartil 2 le corresponde un valor de 17.77, por lo que el 55% cuando mucho llega a 17.71, mientras que el otro 50% son mayores a dicho valor. El cuartil 3 le corresponde un valor de 19.65, por lo que el 75% cuando mucho llega a 19.65, mientras que el otro 25% son notas mayores a dicho valor.

Comunicación Social B

En la carrera de Comunicación Social B, la nota mínima es 12.97 y la nota máxima es 19.96. El cuartil 1 le corresponde un valor de 16.44, por lo que el 25% cuando mucho llega a 15.44, mientras que el otro 75% son notas mayores a dicho valor. El cuartil 2 le corresponde un valor de 18.73, por lo que el 55% cuando mucho llega a 18.73, mientras que el otro 50% son mayores a dicho valor. El cuartil 3 le corresponde un valor de 19.71, por lo que el 75% cuando mucho llega a 19.71, mientras que el otro 25% son notas mayores a dicho valor.

Figura 1. Resultados del ejercicio guiado: análisis descriptivo de calificaciones entre carreras.

```
Frecuencia_abs_clase_vehiculo=no_duplicates_df.groupby(["Clase"]).size().reset_index()
Frecuencia_abs_clase_vehiculo.columns=["Clase","frecuencia_abs"]
#pd.options.display.max_rows=None
Frecuencia_abs_clase_vehiculo
```

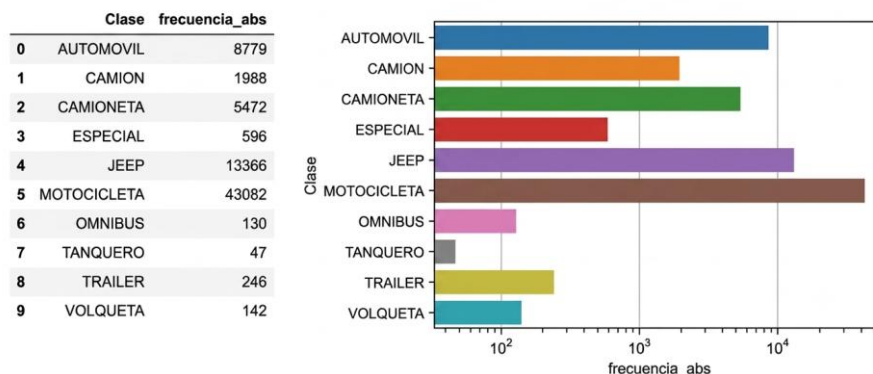


Figura 2a. De los vehículos registrados en el SRI en el 2022, ¿Cuántos hay de cada clase?

```
sns.catplot(y="Pais", hue="Tipo_servicio",
            row="Clase", kind="count",
            data=no_duplicates_df, height=6, aspect=2)
plt.xscale("log")
```

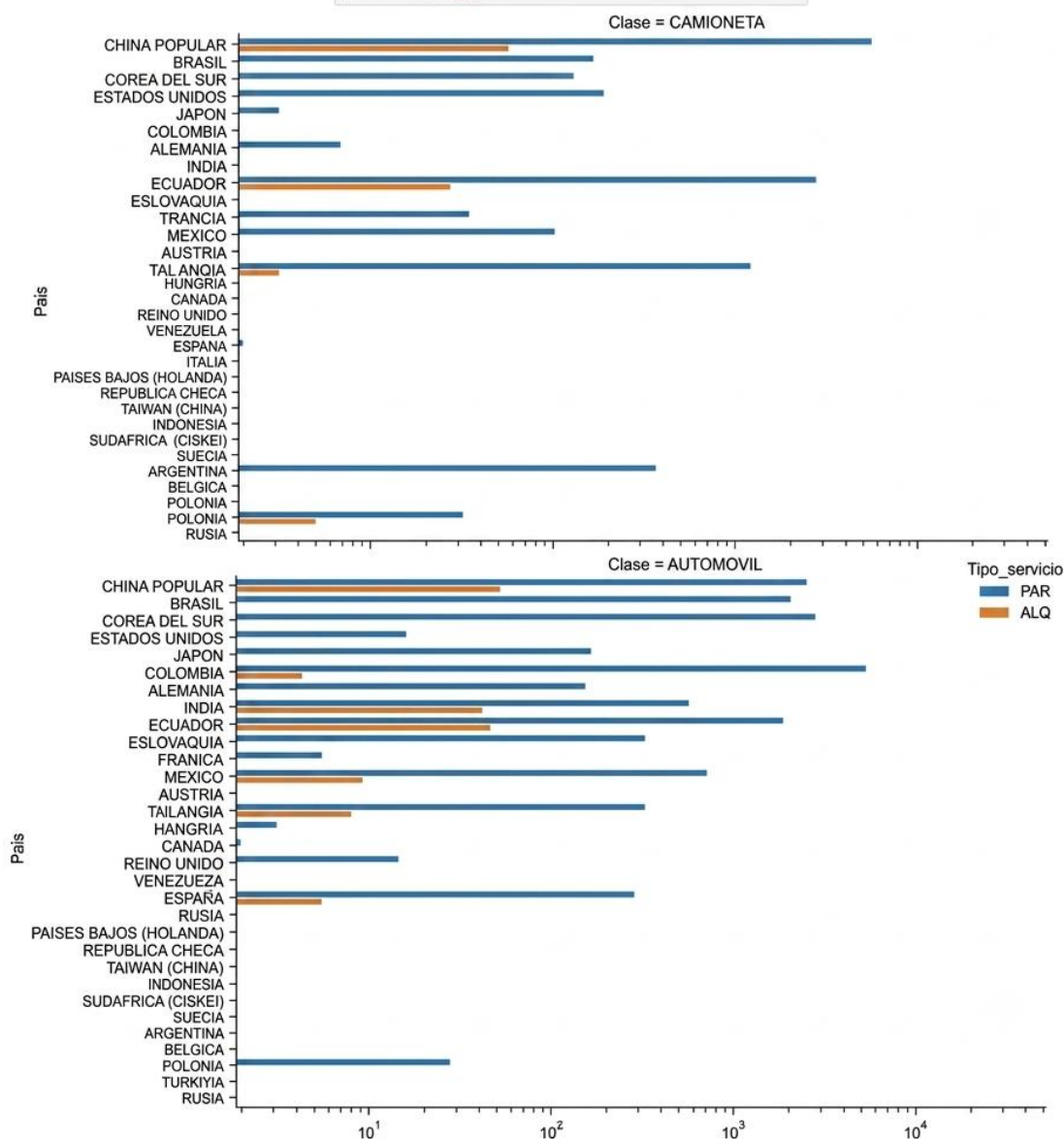


Figura 2b. Gráfico de barras de las variables: clase de vehículo, país y tipo de servicio

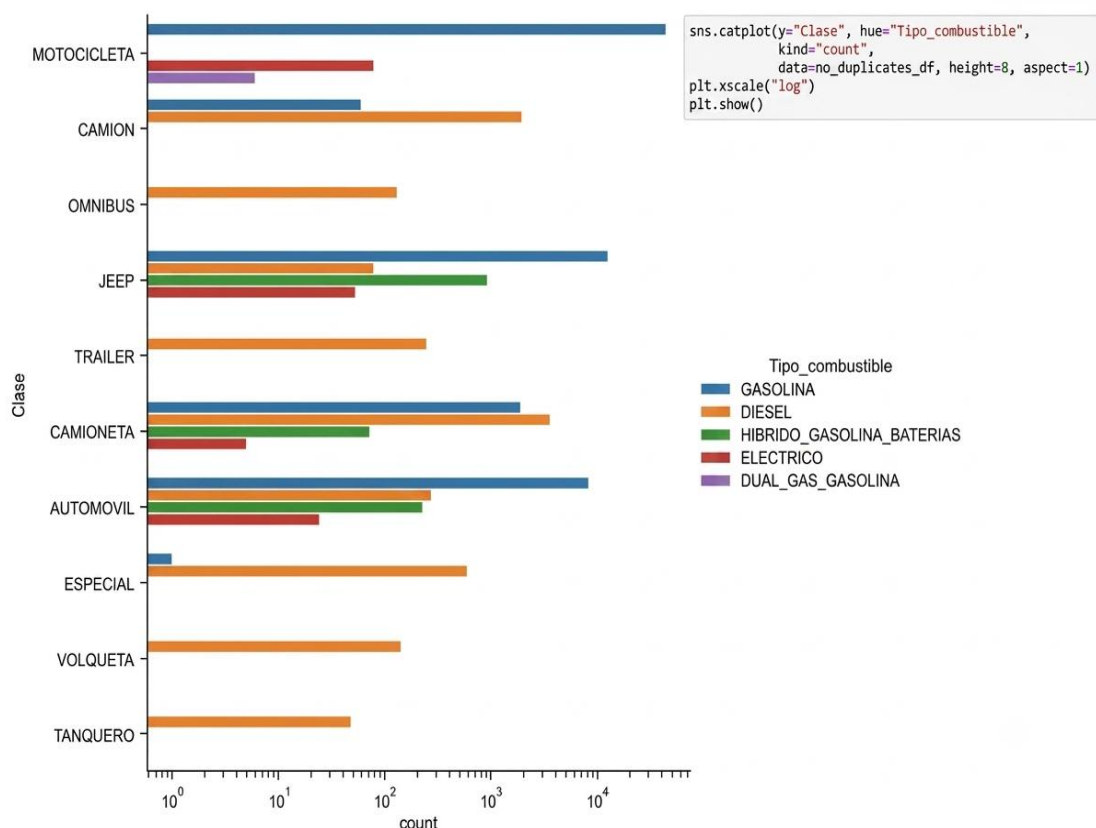


Figura 2c. Representación visual de la clase de vehículo vs tipo de combustible.

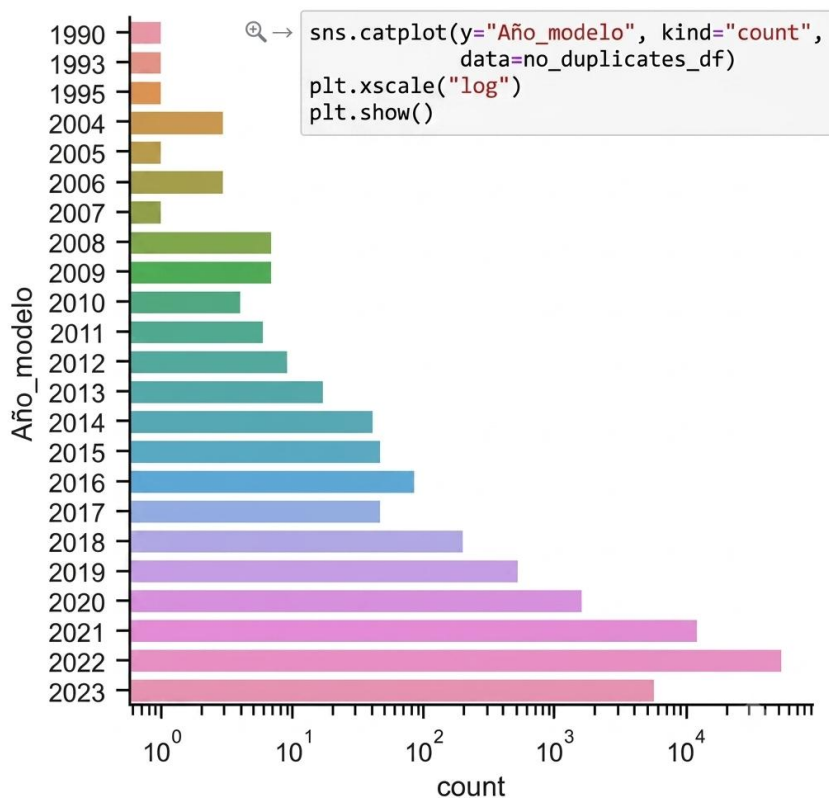


Figura 2d. Distribución de vehículos registrados por año de modelo (escala logarítmica)

Tarea 2: visualización multivariable.

Para representar simultáneamente las variables clase de vehículo, país de origen y tipo de servicio, los estudiantes generaron una cuadrícula de nueve gráficos de barras, uno por clase de vehículo, en escala logarítmica. La elección de la escala logarítmica, no prescrita en la consigna, fue una decisión adoptada de manera independiente por los estudiantes durante la fase de aplicación autónoma, sin mediación directa del docente en ese momento, para manejar la gran disparidad en las frecuencias entre categorías. En la Figura 2b se presentan dos de los nueve gráficos generados a modo ilustrativo.

Tarea 3: clase de vehículo versus tipo de combustible.

La visualización de esta relación reveló un hallazgo no anticipado: la existencia de motocicletas con sistema de combustión dual, gas y gasolina, en el parque automotor ecuatoriano. Este dato era desconocido por todos los participantes y generó una discusión espontánea sobre energía y transporte durante la sesión de cierre (Figura 2c).

Tarea 4: distribución por año de modelo.

Mediante tres líneas de código con la librería Seaborn, los 28 participantes generaron de manera consistente un gráfico de conteo en escala logarítmica para la variable `año_modelo`. El análisis reveló un hallazgo contraintuitivo: la concentración masiva de vehículos de modelo reciente (en este caso 2022), lo que suscitó discusiones críticas sobre las dinámicas de renovación vehicular, políticas públicas y el perfil de consumo en Ecuador. Esta reflexión emergente, basada en evidencia empírica, ejemplifica cómo el uso de herramientas computacionales no solo agiliza el procesamiento de datos, sino que favorece indicios de pensamiento crítico y aproxima la estadística a problemáticas socioeconómicas reales, cumpliendo así con el objetivo pedagógico central de la experiencia (Figura 2d).

Percepciones de los estudiantes sobre la experiencia

Al finalizar la intervención se realizaron dos entrevistas semiestructuradas. Se denominan E1 y E2 a los participantes para preservar su anonimato. Ante la pregunta ¿consideran que lograron aprender estadística descriptiva mediante el uso del cuadernillo Jupyter Notebook?, ambos respondieron afirmativamente, aunque con énfasis distintos.

E1 señaló: “...la experiencia fue relativamente nueva donde se aprendió de manera distinta, donde al final se logró consolidar los conceptos básicos de la estadística descriptiva, a tratar bases de datos numerosas y a utilizar técnicas visuales y particularmente se aprendió a identificar cuando una variable es cuantitativa y

cuando cualitativa...”

E2 indicó: “...al usar este entorno virtual logré desarrollar experiencias nuevas, satisfactorias y divertidas aprendiendo de una manera distinta a lo tradicional, que a la vez me permitieron consolidar los aprendizajes en temas referentes al cálculo de medidas de tendencia central, de dispersión y de forma y su presentación mediante técnicas visuales, como un diagrama de barras e histogramas...”

El análisis de contenido de ambos testimonios permite identificar tres categorías temáticas recurrentes: (a) la percepción de novedad respecto a las prácticas de aprendizaje habituales, presente en ambos participantes mediante el contraste explícito con “lo tradicional”; (b) la consolidación de conceptos estadísticos específicos, reportada con precisión técnica por ambos; y (c) una valoración afectiva positiva de la experiencia, expresada por E2 mediante los calificativos “satisfactoria y divertida”.

DISCUSIÓN

Los aprendizajes evidenciados en los trabajos de los estudiantes, desde la ejecución guiada de código hasta la aplicación autónoma con interpretación contextualizada, no puede atribuirse únicamente a las características técnicas de Jupyter Notebook, dejando de lado el diseño didáctico progresivo que estructuró su uso. La secuencia inicio-desarrollo-cierre, inspirada en el ciclo de investigación estadística (Wild y Pfannkuch, 1999), garantizó que los estudiantes contaran con comprensión conceptual previa antes de enfrentarse al entorno computacional, lo que es consistente con el hallazgo de Amoudi y Tbaishat (2023) de que Jupyter potencia la comprensión conceptual cuando se acompaña del andamiaje estructurado y no cuando se emplea como herramienta autónoma desde el inicio.

Esta observación conecta directamente con el modelo TPACK (Mishra y Koehler, 2006): la eficacia de la experiencia dependió de la articulación coherente entre el conocimiento disciplinar del docente para seleccionar *datasets* con potencial analítico y social, su conocimiento pedagógico para diseñar tareas que demandaran interpretación más que ejecución y su conocimiento tecnológico para anticipar las posibilidades de visualización que el entorno ofrecía. La ausencia de cualquiera de estos componentes habría comprometido la calidad de la experiencia didáctica, independientemente de las capacidades técnicas de la herramienta, tal como advierten Engelbrecht y Borba (2024).

Por otro lado, uno de los hallazgos más relevantes del estudio es que los estudiantes no se limitaron a producir resultados estadísticos correctos, sino que elaboraron interpretaciones contextualmente fundamentadas. Este

tipo de razonamiento es precisamente lo que la investigación en educación estadística ha identificado como difícil de lograr mediante enfoques tradicionales centrados en el cálculo manual (Garfield y Ben-Zvi, 2007; Biehler *et al.*, 2018; Larreamendy-Joerns *et al.*, 2005).

Un aspecto importante por destacar es la decisión analítica de emplear una escala logarítmica, adoptada de manera independiente por los estudiantes durante la fase de aplicación autónoma, sin que estuviera prescrita en la consigna original, ante la disparidad en las frecuencias del *dataset* vehicular. Esta elección revela una comprensión de la naturaleza de los datos que va más allá de la instrucción recibida, y que Howley y Roberts (2020) asocian al desarrollo de la literacidad estadística cuando los estudiantes trabajan con datos reales de variabilidad amplia. El uso de funciones como *groupby()* y *size()* en ese mismo contexto coincide con las prácticas recomendadas en la formación de Ciencia de Datos (Pruim *et al.*, 2023), esto sugiere que la experiencia contribuyó simultáneamente al desarrollo de competencias estadísticas y computacionales.

En lo relacionado con el uso de *datasets* provenientes de contextos educativos y automotriz, el hallazgo sobre la combustión dual en motocicletas y la predominancia de vehículos de modelo reciente en el parque automotor ecuatoriano, evidencian que el trabajo con datos reales puede detonar discusiones que van más allá del aprendizaje estadístico en sentido estricto. Cuando los datos son reales y suficientemente complejos, generan preguntas que los estudiantes no anticipaban y que los conducen a cuestionar sus propias concepciones sobre el mundo. Este proceso corresponde a lo que Fielding *et al.* (2025) denominan *data-ing*: una práctica en la que los estudiantes no solo analizan datos, sino que también los interrogan, discuten sus implicaciones y construyen argumentos a partir de la evidencia empírica.

Esto es coherente con la perspectiva de Biehler *et al.* (2018), quienes argumentan que el desarrollo del pensamiento crítico en estadística requiere necesariamente el trabajo con datos auténticos que planteen preguntas abiertas. En el contexto de la formación docente, esta dimensión adquiere un valor adicional: los futuros profesores de matemáticas aprenden estadística, a la vez, experimentan de primera mano cómo los datos reales pueden convertirse en punto de partida para discusiones interdisciplinarias sobre energía, transporte y política pública, lo que amplía su repertorio de estrategias didácticas para su futura práctica profesional (Vásquez y Alsina, 2021).

En cuanto a las percepciones de los estudiantes sobre su experiencia reportadas por E1 y E2, quienes comentan sobre la novedad, satisfacción y consolidación de aprendizajes, esto es coherentes con los hallazgos de otros estudios sobre el uso de Jupyter Notebook en educación universitaria, por ejemplo, estudios como los

de Nwulu *et al.* (2021) y Carrano *et al.* (2020) destacan que los estudiantes valoran los entornos de cuadernos computacionales cuando perciben que estos les permiten desarrollar habilidades aplicables más allá del aula. En un plano más amplio, Hamer *et al.* (2024) señalan que los estudiantes valoran especialmente las experiencias de codificación cuando les permiten hacer “cosas reales”, en contraste con ejercicios abstractos y descontextualizados.

La referencia a “lo tradicional” en ambos testimonios sugieren una ruptura con la rutina instruccional, lo que podría tener implicaciones favorables para la motivación intrínseca. No obstante, es necesario reconocer las limitaciones de este hallazgo: las percepciones positivas de dos estudiantes, recogidas inmediatamente después de la intervención, no permiten afirmar que los aprendizajes sean duraderos ni que las competencias computacionales se transferirán a nuevos contextos. Estudios longitudinales, con muestras más amplias e instrumentos de recolección complementarios, como rúbricas de evaluación, pruebas de retención o grupos de comparación, serían necesarios para fortalecer las inferencias sobre el impacto real de este tipo de intervenciones.

CONCLUSIONES

El objetivo de este estudio fue analizar la experiencia sobre la enseñanza estadística descriptiva mediada por la herramienta Jupyter Notebook con estudiantes de formación inicial docente en una universidad pública ecuatoriana, con el propósito de identificar los aprendizajes consolidados, las percepciones sobre la experiencia y las condiciones pedagógicas que favorecen su implementación. Los hallazgos obtenidos permiten ofrecer respuestas parciales a cada uno de estos tres focos, con las reservas que impone el diseño metodológico adoptado.

En cuanto a los aprendizajes, los indicios observados en las producciones computacionales sugieren que los estudiantes transitaron de la ejecución guiada de código hacia formas de razonamiento que articulan el resultado estadístico con su significado en el contexto del *dataset* analizado. La elaboración de interpretaciones contextualmente fundamentadas sobre cuartiles y percentiles, la selección autónoma de escala logarítmica ante la disparidad de frecuencias y la identificación de patrones no anticipados en los datos vehiculares constituyen evidencias que se aproximan a lo que Wild y Pfannkuch (1999) entienden por pensamiento estadístico. Estos indicios son alentadores, aunque dada la ausencia de instrumentos de evaluación pre y post intervención, la presente investigación no permite afirmar con certeza el grado de consolidación alcanzado ni su durabilidad.

Respecto a las percepciones, los dos estudiantes

entrevistados valoraron la experiencia como innovadora y diferenciada de las prácticas habituales, y reportaron la consolidación de conceptos estadísticos específicos con precisión terminológica. Aunque esta evidencia es ilustrativa y no representativa del grupo completo, resulta coherente con los hallazgos de Nwulu *et al.* (2021) y Carrano *et al.* (2020) sobre las actitudes favorables que genera el trabajo con cuadernos computacionales cuando se vincula a aplicaciones reales.

En cuanto a las condiciones pedagógicas, el estudio aporta evidencia de que la secuenciación progresiva inicio-desarrollo-cierre, la instrucción conceptual previa al uso del entorno computacional y la selección de *datasets* con potencial analítico y social son factores que contribuyeron a la calidad de la experiencia. Estas condiciones apuntan a la relevancia del conocimiento tecnológico y pedagógico del docente (TPK) como elemento articulador, en consonancia con el marco TPACK de Mishra y Koehler (2006).

Entre los aportes del estudio se destaca la documentación de una experiencia concreta y reproducible en un contexto latinoamericano de formación inicial docente, con uso de datos abiertos y una progresión didáctica estructurada que otros investigadores y docentes de la región pueden adaptar. Esta contribución responde a una brecha identificada en la literatura, dada la escasez de reportes en idioma español sobre el uso de entornos computacionales interactivos para la enseñanza de estadística en programas de pedagogía.

Las limitaciones del estudio deben leerse con atención. Al tratarse de un estudio de caso intrínseco, los hallazgos son específicos del contexto analizado y no son generalizables. El número reducido de entrevistados limita las inferencias sobre percepciones. La ausencia de seguimiento longitudinal impide pronunciarse sobre la transferencia de las competencias desarrolladas. Futuros estudios con diseños cuasiexperimentales, muestras más amplias e instrumentos de evaluación del razonamiento estadístico permitirán robustecer las inferencias que este trabajo apenas esboza.

INFORMACIÓN COMPLEMENTARIA

Conflicto de intereses: los autores declaran no tener conflicto de intereses.

Financiamiento: los autores expresan que no ha sido necesario financiamiento para realizar esta obra de investigación.

Declaración sobre uso de inteligencia artificial generativa: los autores utilizaron Claude para hacer una revisión de estilo del manuscrito. El modelo es Claude 1.1062.0 (b81bcd) 2026-04-06T21:20:28.000Z.

Declaración de disponibilidad de datos: los datos que respaldan los resultados del estudio no están disponibles públicamente, pero pueden ser solicitados al autor de correspondencia fabricio.vinces@unl.edu.ec

Contribuciones de los autores: en base a la taxonomía CRediT, las contribuciones fueron: Fabricio Vincés-Vinces (60%) contribuyó en visualización, revisión y edición, redacción del borrador original, supervisión, administración del proyecto, recursos y materiales, software, análisis de datos, conducción de la investigación, curación de datos, metodología, y conceptualización. Wilmer Ríos-Cuesta (40%) participó en visualización, revisión y edición, redacción del borrador original, validación, supervisión, administración del proyecto, análisis de datos, y metodología.

5. REFERENCIAS

- Amoudi, G. y Tbaishat, D. (2023). Interactive notebooks for achieving learning outcomes in a graduate course: a pedagogical approach. *Education and Information Technologies*, 28(12), 16669–16704. <https://doi.org/10.1007/s10639-023-11854-x>
- Biehler, R., Frischemeier, D., Reading, C. y Shaughnessy, J. M. (2018). Reasoning About Data. In D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 139–192). Springer. https://doi.org/10.1007/978-3-319-66195-7_5
- Campo-Meneses, K. G. y García-García, J. (2020). Explorando las conexiones matemáticas asociadas a la función exponencial y logarítmica en estudiantes universitarios colombianos. *Educación Matemática*, 32(3), 209–240. <https://doi.org/10.24844/EM3203.08>
- Cardoso, A., Leitão, J. y Teixeira, C. (2019). Using the Jupyter Notebook as a Tool to Support the Teaching and Learning Processes in Engineering Courses. *arXiv* (Vol. 917, pp. 227–236). https://doi.org/10.1007/978-3-030-11935-5_22
- Carrano, D., Chugunov, I., Lee, J. y Ayazifar, B. (2020). Self-Contained jupyter notebook labs promote scalable signal processing education. *International Conference on Higher Education Advances*, 1409–1416. <https://doi.org/10.4995/HEAd20.2020.11308>
- Çetinkaya-Rundel, M. y Ellison, V. (2021). A Fresh Look at Introductory Data Science. *Journal of Statistics and Data Science Education*,

- 29(sup1), S16–S26.
<https://doi.org/10.1080/10691898.2020.1804497>
- Darmanova, Z., Abylkassymova, A., y Nurmukhamedova, Z. (2025). A systematic review of technology use in middle and high school mathematics education: Insights from contextual, methodological, and evaluation characteristics. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1644284>
- Engelbrecht, J., y Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM – Mathematics Education*, 56(2), 281–292. <https://doi.org/10.1007/s11858-023-01530-2>
- Fielding, J., Makar, K. y Ben-Zvi, D. (2025). Developing students' reasoning with data and data-ing. *ZDM – Mathematics Education*, 57(1), 1–18. <https://doi.org/10.1007/s11858-025-01671-6>
- Garfield, J. y Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372–396. <https://doi.org/10.1111/j.1751-5823.2007.00029.x>
- Hamer, J. M. M., Kemp, P. E. J., Wong, B. y Copsey-Blake, M. (2024). Cracking the code: exploring student attitudes towards coding in secondary education. *Cambridge Journal of Education*, 54(4), 495–516. <https://doi.org/10.1080/0305764X.2024.2387335>
- Howley, P. y Roberts, T. (2020). Engaging school students and educators with the practice of statistics. *Statistics Education Research Journal*, 19(1), 29–38. <https://doi.org/10.52041/serj.v19i1.117>
- Kaganovskiy, L. (2025). *Applied Statistics with Python: Volume I: Introductory Statistics and Regression*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003473114>
- Krüger, R. (2022). Using Jupyter notebooks as didactic instruments in translation technology teaching. *The Interpreter and Translator Trainer*, 16(4), 503–523. <https://doi.org/10.1080/1750399X.2021.2004009>
- Larreamendy-Joerns, J., Leinhardt, G. y Corredor, J. (2005). Six Online Statistics Courses: Examination and Review. *The American Statistician*, 59(3), 240–251. <https://doi.org/10.1198/000313005X54162>
- Mishra, P., y Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Nwulu, N., Damisa, U. y Gbadamosi, S. L. (2021). Students' perception about the use of jupyter notebook in power systems education. *International Journal of Engineering Pedagogy*, 11(1), 78–86. <https://doi.org/10.3991/IJEP.V11I1.14769>
- Pruim, R., Gîrjău, M.-C. y Horton, N. J. (2023). Fostering Better Coding Practices for Data Scientists. *Harvard Data Science Review*, 5(3). <https://doi.org/10.1162/99608f92.97c9f60f>
- Ríos-Cuesta, W. (2021). Alfabetización estadística y probabilística en profesores en activo de primaria (no licenciados en matemáticas) que enseñan matemáticas: un estudio de caso. In L. Tauber & J. P. Sosa (Eds.), *Tendencias y nuevos desafíos de la investigación en Educación Estadística en Latinoamérica* (pp. 48–54). Universidad Nacional del Litoral.
- Rule, A., Tabard, A. y Hollan, J. (2018) Exploration and Explanation in Computational Notebooks. *ACM CHI Conference on Human Factors in Computing Systems*, pp.1-12, 10.1145/3173574.3173606
- Schoenfeld, A. H. (2016). Learning to Think Mathematically: Problem Solving, Metacognition, and Sense Making in Mathematics (Reprint). *Journal of Education*, 196(2), 1–38. <https://doi.org/10.1177/002205741619600202>
- Servicio de Rentas Internas. (2022). *Estadísticas Vehículos 2022*. Datos abiertos Ecuador. <https://www.datosabiertos.gob.ec/dataset/estadisticas-vehiculos-2022>
- Stake, R. E. (2010). *Qualitative Research: Studying How Things Work*. Guilford Press.
- Vásquez, C. y Alsina, Á. (2021). Analysing Probability Teaching Practices in Primary Education: What Tasks Do Teachers Implement? *Mathematics*, 9(19), 2493. <https://doi.org/10.3390/math9192493>

Wild, C. J. y Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67(3), 223–248.
<https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35.
<https://doi.org/10.1145/1118178.1118215>

Ye, L., Ismail, H. H., y Aziz, A. A. (2024). Innovative Strategies for TPACK Development in Pre-Service English Teacher Education in the 21st Century: A Systematic Review. *Forum for Linguistic Studies*, 6(6), 274–294.
<https://doi.org/10.30564/fls.v6i6.7308>

Artículo en normas APA 7ma. Edición.



Artículo de **libre acceso** bajo los términos de la **Licencia Creative Commons Reconocimiento – NoComercial – CompartirIgual 4.0 Internacional**. Se permite que otros remezclen, adapten y construyan a partir de su obra sin fines comerciales, siempre y cuando se otorgue la oportuna autoría y además licencien sus nuevas creaciones bajo los mismos términos.