

Artículo de Investigación

Herramienta Stata en el análisis de datos cuantitativos en estudiantes de posgrado en Perú

Stata tool in the analysis of quantitative data in Peruvian graduate students

Lolo Caballero Cifuentes¹: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

lcaballero@une.edu.pe

Richard Quivio Cuno: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

rquivio@une.edu.pe

Sandra Gutiérrez Guadalupe: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

sgutierrezg@une.edu.pe

Mario Jaime Andía: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

mjaime@une.edu.pe

Jesús Alegre Huerta: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

jalegre@une.edu.pe

Juan Carlos Human Hurtado: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

jhuaman@une.edu.pe

Luz Sito Justiniano: Universidad Nacional de Educación Enrique Guzmán y Valle, Perú.

lsito@une.edu.pe

Fecha de Recepción: 20/06/2024

Fecha de Aceptación: 26/07/2024

Fecha de Publicación: 19/09/2024

Cómo citar el artículo:

Caballero, L., Quivio, R., Gutiérrez, S., Jaime, M., Alegre, J., Humanan, J. C. y Sito, L. (2024). Herramienta Stata en el análisis de datos cuantitativos en estudiantes de posgrado en Perú [Stata tool in the analysis of quantitative data in Peruvian graduate students]. *European Public & Social Innovation Review*, 9, 1-19. <https://doi.org/10.31637/epsir-2024-802>

¹ **Autor Correspondiente:** Lolo Caballero Cifuentes. Universidad Nacional de Educación Enrique Guzmán y Valle (Perú).

Resumen:

Introducción: Este artículo analiza el impacto del software Stata en el análisis de datos cuantitativos, una experiencia en estudiantes universitarios de posgrado. **Metodología:** El tipo de investigación fue experimental, aplica un enfoque cuantitativo con diseño cuasi experimental, la muestra fue no probabilístico, conformando 28 estudiantes para grupo experimental y el mismo número para el grupo control, se aplica instrumentos validados y confiables con el método de dos mitades Spearman Brown. **Resultados:** El análisis demuestra la comparación de promedios por grupos en ambos momentos de Preprueba y Posprueba. Los promedios fueron más altos y significativos para grupo experimental (15,86) en comparación al grupo control (12,50), el contraste de la prueba con U. Mann Whitney fue estimado al 95% de nivel de confianza y un 5% de significancia que permitió inferir su impacto positivo del aplicativo Stata. **Discusión:** Los hallazgos encontrados en el estudio fueron muy significativo, mejorando el nivel teórico y práctico en el campo de la investigación científica en los estudiantes al analizar la estadística con el software. **Conclusiones:** Se concluye que el impacto del software estadístico Stata sí influye positivamente en el análisis de datos cuantitativos.

Palabras clave: aprendizaje; análisis; cuantitativo; datos; estadística; impacto; software; universitario.

Abstract:

Introduction: This article analyzes the impact of Stata software on quantitative data analysis, an experience among postgraduate university students. **Methodology:** The type of investigation was experimental, applied a quantitative approach with an experimental design, the sample was non-probabilistic, with 28 students for the experimental group and the same number for the control group, validated and reliable instruments were applied with the dosing method spearman brown myths. **Results:** The analysis demonstrates the comparison of means by groups in both moments of Pre-test and Post-test so that the means were higher and more significant for the experimental group (15,86) in comparison with the control group (12,50), the contrast of The test with U. Mann Whitney was estimated at a 95% confidence level and a 5% significance level that allowed us to infer the positive impact of the Stata application. **Discussions:** The findings found in the study were very significant, improving the theoretical and practical level in the field of scientific research among students when analyzing statistics with the software. **Conclusions:** It is concluded that the impact of the Stata statistical software only positively influences the analysis of quantitative data.

Keywords: learning; analysis; quantitative; data; statistics; impact; software; university.

1. Introducción

En la actualidad todavía en el Perú se observa en algunas universidades docentes que tienen todavía un bajo nivel de manejo de herramientas informáticas, esto se debe a que se rehúsen a utilizar en la enseñanza (Sierra *et al.*, 2022).

Después de la pandemia la sociedad se vio obligada al uso y manejo de las Tecnologías de la Información y la Comunicación, que pasaron de ser consideradas como simples sistemas computacionales a constituirse en herramientas fundamentales que apoyan el desarrollo científico y académico. (Cisneros *et al.*, 2022, p. 8)

En esa misma línea Millones *et al.* (2021) manifiestan acerca de la similitud de dificultades de la enseñanza de la estadística a nivel universitario en Perú con el de Brasil, que requiere fortalecer la necesidad del conocimiento en el análisis de datos cuantitativos.

El aplicativo Stata como una herramienta estadística es importante en el manejo de datos y el análisis estadístico, esta herramienta permite almacenar, crear y la ejecución de operaciones de manejo de datos vinculadas a la investigación (Escobar *et al.*, 2022).

Por otro lado, Hidalgo (2019), en su estudio denominada las técnicas estadísticas en el análisis cuantitativo de datos, señala que un diseño cuantitativo es una manera de aproximación sistemática que hace uso primordialmente del análisis de variables categóricas y numéricas, haciendo uso de diversas técnicas estadísticas que permiten mostrar evidencias para establecer interrelaciones entre variables con fines explicativos y predictivos.

En esa perspectiva, en nuestro trabajo de investigación se enfoca en la aplicación del software estadístico Stata en el análisis de datos cuantitativos para el ámbito educativo en la asignatura de seminario de estadística a nivel de posgrado.

A continuación, doy conocer algunos antecedentes relacionados con la investigación:

Vásquez y García (2020), en su estudio educación estadística, enfatizan la necesidad de generar acciones formativas para el docente a guiar la enseñanza estadística con proyectos de desarrollo sostenible. El estudio presentó análisis de un conjunto de unidades didácticas para la enseñanza y aprendizaje de la estadística, promoviendo el desarrollo de competencias en los estudiantes. En esa línea, nuestro estudio utilizó la herramienta estadística Stata como estrategia didáctica a para el docente, fomentando un aprendizaje más cativo e interactivo para los estudiantes de posgrado.

La investigación de Matzumura *et al.* (2021) tuvo como objetivo determinar el vínculo entre el aplicativo Remind como herramienta de apoyo y el rendimiento académico tipo de investigación correlacional, con una muestra no probabilística de 96 estudiantes los resultados señalan que el 98 % manifiestan que los aplicativos mejoran la gestión y comunicación, concluyendo que la herramienta fue favorable, contribuyendo al aprendizaje en forma positiva.

Vela y Guerrero (2020) en su estudio señala que el software Stata al igual que las herramientas Spss, Eviews, Gretl, Excel, etc. es importante porque permite apoyar los procesos estadísticos matemáticos, econométricos mediante el análisis de datos producto de una recolección, este programa y los mencionados permiten estimaciones y cálculos estadísticos aplicado a diferentes disciplinas de manera rápida y eficaz.

Mercado *et al.* (2012) señala que el software Stata es un paquete estadístico usado en diferentes áreas que trabaja con diferentes formatos almacenados en diferentes archivos que guardan información, que contiene datos y almacena información.

El software Stata es una herramienta de análisis de datos estadísticos creada por STATA CORP en la década de los 80 para Windows, tiene más de 35 años de experiencia en todos los campos. Su uso abarca una gama de tareas incluidos como gestión, cálculo, gráficos, otros. (Mercado, 2012).

Con respecto a la variable análisis de datos Pérez (2023) manifiesta que el termino análisis de datos o en inglés *data analytics* referido a la recopilación de diferentes tipos de datos, comienza con la extracción, organización, análisis y termina con la transformación en datos numéricos, tablas gráficas mediante una información coherente. El análisis de datos, al igual que las habilidades y el conocimiento, contribuyen a una mayor competitividad en diferentes campos como la medicina, educación, política, economía, investigación entre otros. Fortalecen las

habilidades, transformándolas en información útil y tomar buenas decisiones. En ese sentido, los especialistas en análisis de datos deben tener conocimientos estadísticos, dominio de herramientas tecnológicas y comunicación efectiva con el fin de realizar trabajos de investigación con mejor precisión.

Para analizar datos cuantitativos tener en cuenta sus características como orden, representación numérica, escala, operación y analizar con los métodos estadísticos a nivel descriptivo e inferencial. Los resultados numéricos se interpretan de acuerdo con un contexto, por ejemplo, un mismo valor de presión arterial de un bebe es diferente al de una persona adulta (Hernández *et al.*, 2014).

García *et al.* (2021) en su estudio tuvo como fin analizar la representación de datos con intervención de las medidas de tendencia central es decir la media, la mediana y la moda en tres libros de texto de séptimo básico en Chile, la metodología empleada fue cualitativa, de tipo descriptiva, los resultados evidenciaron el predominio de la media a diferencia de la mediana y moda.

Para Gea y Roa (2014) al desarrollar el sentido de la correlación y regresión en nuestros estudiantes la preparación del docente es importante; debe dominar softwares estadísticos, saber diferenciar las variables de estudio y elegir métodos adecuados.

Solorzano *et al.* (2022) señalan que los estudios estadísticos en el campo de la investigación educativa proporcionan varios procedimientos para verificar las hipótesis estadísticas, normalmente se utilizan las pruebas chi-cuadrado, la prueba t-Student para muestras relacionadas, t para muestras independientes, la prueba z diferencia de dos muestras, la prueba Fisher, Rho Spearman, Pearson, Wilcoxon, U, Mann Whitney y entre otros cada una de ella se emplea en diferentes situaciones.

La presente investigación aporta un aspecto teórico y práctico sobre el análisis de datos cuantitativos para estudiantes de posgrado en sus trabajos de investigación usando la herramienta estadística denominada Stata, es importante también para el docente porque pone en práctica una nueva estrategia metodológica y permite fomentar un aprendizaje autónomo, interactivo y más activo en sus estudiantes.

Los objetivos de nuestro estudio fueron: 1. Analizar el impacto del aplicativo Stata en el análisis de datos cuantitativos en la asignatura de estadística en los estudiantes de la Escuela de Posgrado de la Universidad Enrique Guzmán y Valle de Lima. Como objetivos específicos: 2. Determinar el impacto preponderante del aplicativo Stata en el análisis de datos en la dimensión medidas de tendencia central y dispersión; 3. Precisar impacto preponderante del aplicativo Stata en el análisis de datos en la dimensión medidas de tendencia central y dispersión; 4. Conocer impacto significativo del aplicativo Stata en el análisis de datos dimensión pruebas de hipótesis.

2. Metodología

Hernández *et al.* (2014) enfatizan que un enfoque cuantitativo se hace uso de la recolección de datos para comprobar las hipótesis a través de cuantificación numérica, en ese sentido el estudio es de corte cuantitativo que busca regirse a la medición y calificación de los resultados.

Es necesario conocer el tipo de investigación para seleccionar con idoneidad el método, la técnica, los instrumentos y el tratamiento estadístico (Gerrero, 2022). Bajo este contexto, se ha desarrollado un estudio de tipo experimental aplicada debido a que se aplicó el software Stata para conocer su impacto en el análisis de datos cuantitativos en la asignatura de estadística para la investigación.

En nuestro trabajo de investigación el método fue cuantitativo-aplicado, en ese sentido, se obtuvo información significativa mediante un análisis estadístico antes y después de aplicar el software Stata en el análisis de datos estadísticos cuantitativos (Arias y Covinos, 2021). Al respecto, Hernández y Mendoza (2018) señalan que una investigación es aplicada porque busca en poner en práctica los conocimientos adquiridos para obtener nuevo conocimiento fundamentada en hallazgos tecnológicos de la investigación básica.

El estudio se desarrolló bajo el diseño cuasi experimental debido a que antes de poner en práctica los contenidos estadísticos de análisis de datos cuantitativos se eligió en forma intencional dos grupos previamente establecidos o son llamados también grupos intactos que tenían características comunes como: la misma asignatura, ciclo académico, los contenidos desarrollados y el nivel socioeconómico, a estos grupos se le denominó experimental (GE) y el otro grupo control (GC). En seguida presentamos el esquema de diseño.

	Preprueba		Posprueba
Grupo Experimental	O ₁	x	O ₃
Grupo Control	O ₂	---	O ₄

Donde: x uso del aplicativo Stata.

--- sin uso.

O₁, O₂, O₃ y O₄ Observaciones.

Para el estudio la población es finita y fue delimitada por 85 estudiantes de la Escuela de Posgrado de la mención doctorado en ciencias de la educación que llevan la asignatura de estadística aplicada al campo de la investigación y educacional.

En cuanto a la muestra que es un subconjunto de la población. Se determinó para el estudio mediante el método no probabilístico de forma intencional y arbitraria teniendo en cuenta el mismo nivel socioeconómico, la misma asignatura, el mismo ciclo y se desarrollaron los mismos contenidos (Ver tabla 1).

Tabla 1.

Determinación de la muestra

Grupos	Secciones	Ciclo académico	N° de estudiantes
Experimental	Doctorado A	primero	28
Control	Doctorado B	primero	28

Fuente: Elaboración propia (2024) a partir de la población.

La técnica de recojo de datos para esta investigación, fueron dos pruebas de conocimientos principales constituidas por 20 ítems tomando en cuenta los indicadores de estudio de la variable dependiente denominado análisis de datos cuantitativos en la asignatura de estadística aplicada a la investigación, para conocer los avances de niveles de competencias estadísticas utilizamos también algunas técnicas auxiliares como lista de cotejo y rúbricas. Antes de ser aplicadas a toda la muestra de estudio se determinaron su validez con el método de juicio de expertos y la concordancia de los ítems con el coeficiente de V. Aiken y su confiabilidad con el método dos mitades utilizando la fórmula de Spearman Brown a continuación a continuación muestro los resultados en las siguientes tablas.

Tabla 2.

Análisis de evolución de los expertos

Expertos	Preprueba	Posprueba	Decisión
Juez1	92%	92%	Si es aplicable
Juez2	93%	94%	aplicar
Juez3	91%	93%	aplicable
Juez4	91%	93%	aplicar
Total	92%	93%	Son aplicables

Fuente: Elaboración propia (2024) a partir de evaluación de jueces.

Se evidencia en la tabla 2 los resultados de validez de contenido de los instrumentos mediante jueces, obteniéndose para la preprueba un promedio en porcentaje de 92% y para la posprueba un 93%, por lo tanto, ambas pruebas si eran aplicables.

Tabla 3.

Relevancia de ítems con Aiken

Indicadores	Claridad	Congruencia	Contexto	V. Aiken
Preprueba	0,82	0,82	0,82	0,82
Posprueba	0,83	0,80	0,81	0,81
Total	0,83	0,81	0,82	0,82

Fuente: Elaboración propia (2024) a partir de datos obtenidos.

Los valores que se aprecia en la tabla3 indican la concordancia de los ítems medido mediante los indicadores claridad, congruencia y contexto utilizando la formula del coeficiente de V. Aiken, se observa para la Preprueba un coeficiente de (0,82) y para la Posprueba (0,81), estos valores nos indicaron que los ítems si eran concordantes de contenido en su totalidad. Por lo tanto, ningún ítem fue reformulado ni eliminado.

Tabla 4.

Análisis de confiabilidad con Spearman Brown

	Piloto	Ítems	Coefficiente Spearman Brown
Preprueba			
validados	10	20	0,87
excluidos	0	0	
Total	10	1	0,87
Posprueba			
validados	10	20	0,88
excluidos	0	0	
Total	10	21	0,88

Fuente: Elaboración propia (2024) a partir de datos producto de la confiabilidad.

Con respecto a la confiabilidad según Sampieri *et al.*, (2014) manifiestan que la confiabilidad es el grado en la cual su aplicación repetida a los mismos sujetos da resultados iguales. Para el estudio utilizamos el método de dos mitades con la fórmula de Spearman Brown que es aplicable a cuestionario dicotómicos. El procedimiento se procesó con ayuda del SPSS, a continuación, damos a conocer su fórmula y resultados:

(e1)

$$R = \frac{2 * r}{1 + r}$$

R= Spearman Brown
r = Pearson

En la tabla 4 se observa los resultados de confiabilidad procesados con una muestra piloto de 10 estudiantes, 20 ítems analizados y un valor del coeficiente Spearman Brown para la Preprueba de (0,87) y para la Posprueba un valor de (0,88) ambos resultados nos indicaron una muy alta confiabilidad, siendo coherentes y consistentes para su uso. En consecuencia, se demostró la validez y confiabilidad de los instrumentos para ser aplicadas a toda la muestra de estudio.

3. Resultados

En esta sección, se presentan los datos recogidos de manera organizada y sistemática, utilizando tablas, figuras, y análisis estadístico a nivel descriptivo e inferencial en base a la muestra y la recolección de datos. Los resultados se obtuvieron con apoyo del MS. Excel y SPSS, que a continuación, presentamos de manera objetiva.

Tabla 5.

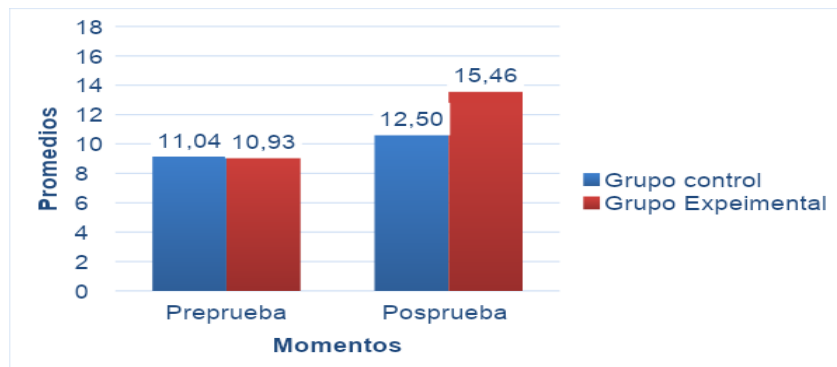
Análisis dimensión medidas de tendencia central y dispersión

	Preprueba	Posprueba
Grupo control	11,04	12,5
Grupo Experimental	10,93	15,46

Fuente: Elaboración propia (2024).

Figura 1.

Análisis de promedios medidas de tendencia central y dispersión



Fuente: Elaboración propia (2024).

Según la tabla 5 y la figura 1 se evidencia las puntuaciones en promedios para la variable análisis datos cuantitativos en su dimensión medidas de tendencia central y de dispersión, es así para el momento de la Preprueba para el grupo experimental el promedio fue (10,93) y para el grupo control (11,04) es decir en esta etapa todavía no se aplicó el software Stata, Sin embargo después de su aplicación es decir en el Posprueba la tendencia en promedios fue favorable para el grupo experimental con (15,46) y para el control (12,50), estas ponderaciones nos indicó que el impacto del aplicativo Stata fue positivo en el proceso de enseñanza, contribuyendo a aún aprendizaje activo, más colaborativo, autónomo, interactivo, con una alta motivación y de fácil resolución de problemas respecto a medidas de tendencia central y dispersión en los estudiantes de posgrado de la Universidad Enrique Guzmán y Valle.

Tabla 6.

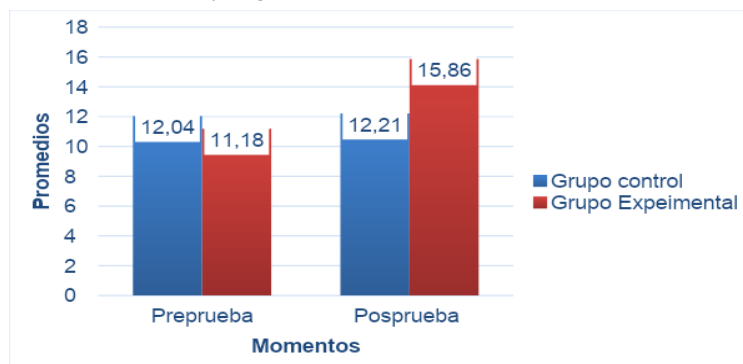
Análisis dimensión correlación y regresión

	Preprueba	Posprueba
Grupo control	12,04	12,21
Grupo Experimental	11,18	15,86

Fuente: Elaboración propia (2024) a partir de datos obtenidos de las pruebas.

Figura 2.

Análisis de promedios correlación y regresión



Fuente: Elaboración propia (2024) a partir de la tabla6.

En la tabla 6 y su respectiva figura 2 se aprecian los promedios de los grupos control y experimental en la dimensión correlación y regresión de la variable análisis de datos cuantitativos. En la Preprueba se observa un valor para grupo control de (12,04), en comparación al promedio del grupo experimental de (11,18), aparentemente con una pequeña ventaja para el grupo control, sin embargo, después de aplicar el aplicativo Stata los promedios se invierten con una puntuación para el grupo experimental en promedio de (15,86) a comparación del grupo control (12,21), estas evidencias nos indicaron que el impacto también fue muy positivo en la enseñanza aprendizaje de los estudiantes del grupo experimental, al resolver problemas, y desarrollar procedimientos en los temas tratados acerca de correlación y regresión estadística.

Tabla 7.

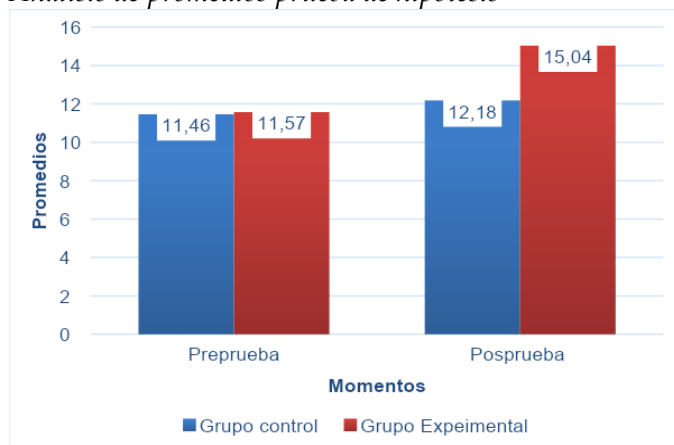
Análisis dimensión Prueba de hipótesis

	Preprueba	Posprueba
Grupo control	11,46	12,18
Grupo Experimental	11,57	15,04

Fuente: Elaboración propia (2024) a partir de los datos del análisis descriptivo.

Figura 3.

Análisis de promedios prueba de hipótesis



Fuente: Elaboración propia (2024) obtenidos a partir de la tabla7

En la tabla 7 y la figura 3 se aprecian los promedios de los dos momentos Preprueba y Posprueba para la dimensión prueba de hipótesis de la variable análisis de datos cuantitativos, para el primer momento es decir antes de la aplicación del Stata los promedios fueron para grupo control de (11,46) y para el grupo experimental (11,57), después de su aplicación el promedio para el grupo experimental fue (15,04) mayor al promedio del grupo control (12,18), estos valores obtenidos significaron que el impacto del aplicativo Stata mejoró los logros y niveles en el proceso de enseñanza aprendizaje resolviendo problemas sobre pruebas de hipótesis utilizando estadísticos paramétricos como no paramétricos en los estudiantes de posgrado que luego aplican en sus trabajos de investigación.

Tabla 8.

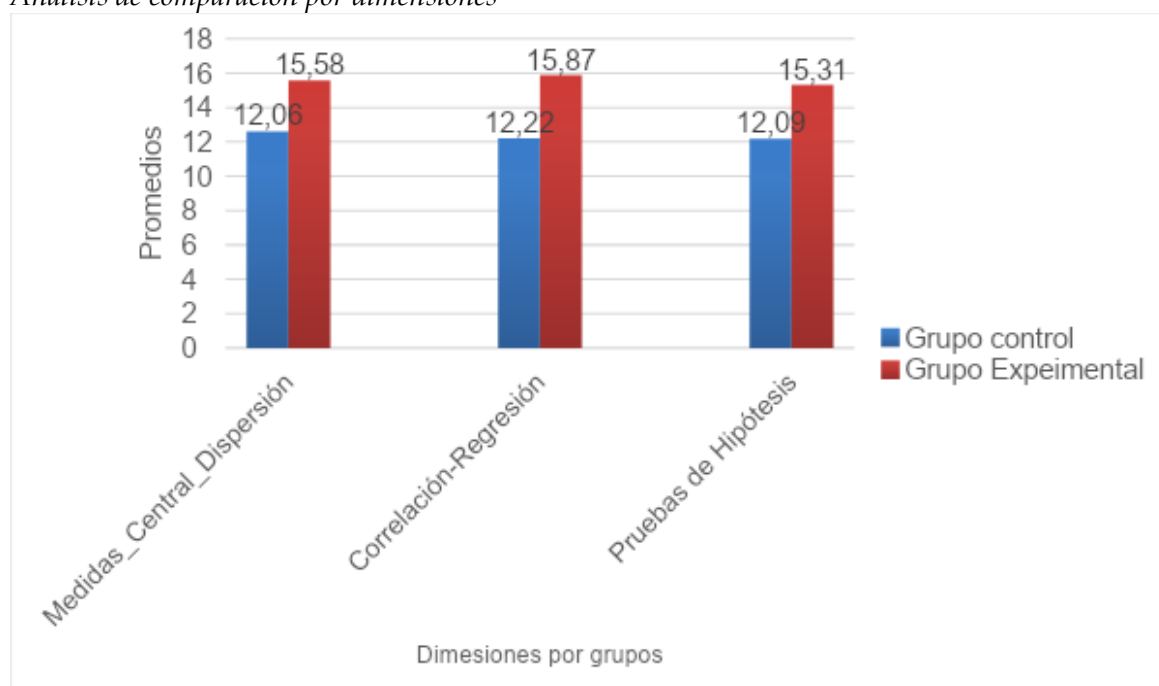
Análisis comparación de promedios por grupos

Grupos	Medidas de tendencia y dispersión	Correlación y regresión	y Prueba de hipótesis
Control	12,06	12,22	12,19
Experimenta l	15,58	15,58	15,58

Fuente: Elaboración propia (2024) a partir de los datos del análisis a nivel descriptivo.

Figura 4.

Análisis de comparación por dimensiones



Fuente: Elaboración propia (2024) a partir de la tabla 8.

En la tabla 8 y su figura 4 se evidencian las diferencias de promedios en las tres dimensiones medidas de tendencia central y dispersión, correlación y regresión y prueba de hipótesis, las mayores puntuaciones obtenidas en la Posprueba son para el grupo experimental con las puntuaciones (15,58; 15,87 y 15,31), mientras para el grupo control fueron (12,06, 12,22 y 12,19), estos resultados fueron significativos al aplicar el aplicativo Stata en el análisis de datos cuantitativos, contribuyendo un impacto positivo en la enseñanza de una estadística para la investigación científica, por otro lado, mejora el proceso de aprendizaje haciéndolo más activo y motivado en los estudiantes de posgrado.

Por otro lado, para conocer la distribución de los datos recogidos y elegir una fórmula apropiada para contrastar las hipótesis, recurrimos a una prueba de normalidad con el método numérico de Shapiro Wilk que a continuación se presenta los resultados.

Tabla 9.*Análisis prueba de normalidad*

	Estadístico	gl	Sig.
Preprueba	0,541	28	0,872
Posprueba	0,529	28	0,748

Fuente: Elaboración propia (2024) a partir de la prueba de normalidad.

Se planteó como hipótesis nula H_0 : Los datos no tienen una distribución normal y como alterna H_a : Los datos tiene una distribución normal. Si y solo si: sig. $\leq 0,05$ se rechazará la hipótesis nula, de lo contrario Si y solo si: sig. $> 0,05$ se aceptará la hipótesis nula

De acuerdo con la tabla9 se aprecia los resultados producto de la prueba de normalidad con Shapiro wilk. Se evidencia un valor estadístico para la Preprueba de (0,541) con 28 grados de libertad y un sig.-bilateral de (0,872) a diferencia de la Posprueba (0, 529) con 28 gl. y un p-valor de (0,748), ambos valores fueron mayores al valor de (0,05), de acuerdos con estos hallazgos se tomó la decisión de aceptar la hipótesis nula y concluir que los datos no provenían de una distribución normal, en consecuencia, para la prueba de hipótesis se utilizó un estadístico no paramétrico U. Mann Whitney.

A continuación, se evidencia los resultados de las hipótesis planteadas con el fin conocer e inferir su influencia e impacto del software Stata en el análisis de datos cuantitativos desarrollados en la asignatura de estadística. Para la hipótesis general se plantea como hipótesis alterna sí existe impacto significativo del aplicativo Stata en el análisis de datos cuantitativos, en la asignatura estadística en los alumnos de a nivel Posgrado y como nula no existe impacto significativo del aplicativo Stata en el análisis de datos cuantitativos, en la asignatura estadística en los alumnos de a nivel Posgrado, se estimó al 95% de confianza y 5% de significancia. A continuación, presento la formula y resultados.

Donde:

U=valor de U. Mann Whitney

Z= valor obtenido

n1 =muestra del GE

n2=muestra GC

$$z = \frac{U - \left(\frac{n_1 n_2}{2} \right)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Tabla 10.*Prueba con U. Mann Whitney impacto de Stata en el análisis de datos cuantitativos*

Rangos de la prueba			Estadísticos de prueba			
Métodos	n	Rango promedio	Suma de rangos	U. Mann-Whitney	Z- Calculado	Sig. asintótica (bilateral)
Con Stata	28	41,57	1164,00	26,000	-6,117	0,00
Sin Stata	28	15,43	432,00			

Fuente: Elaboración propia (2024) a partir del contraste de hipótesis con SPSS.

De la tabla 10 se evidencia los valores de z-calculado de U. Mann Whitney (-6,117) siendo superior al valor de z-crítico (-1.96), por otro lado, también se aprecia el valor del sig. (bilateral) (0,00) inferior al valor de significancia (0,05) razones suficientes para inferir estadísticamente que el impacto del aplicativo Stata en el análisis de datos cuantitativos fue muy significativo y positivo en la asignatura estadística en los alumnos de a nivel Posgrado de la Universidad Enrique Guzmán y Valle, Lima, Perú.

Referente a la hipótesis específica 1 se planteó como hipótesis verdadera si existe impacto preponderante del aplicativo Stata en el análisis de datos en la dimensión medidas de tendencia central y dispersión y como hipótesis nula su negación. Se estimó también al 95% de confianza y 5% de significancia y para comprobarlo se utilizó el mismo estadístico de prueba de U. Mann Whitney y se procesó con ayuda del software SPSS.

Tabla 11.

Prueba con U. Mann Whitney dimensión medidas de tendencia central y dispersión

Rangos de la prueba		Estadísticos de prueba				
Métodos	n	Rango promedio	Suma de rangos	U. Mann-Whitney	Z-Calculado	Sig. asintótica (bilateral)
Con Stata	28	42,18	1181,00	9,000	-6,323	0,00
Sin Stata	28	14,82	415,00			

Fuente: Elaboración propia (2024) a partir de los datos obtenidos de la prueba de hipótesis.

En la tabla 11 se aprecia la suma de rangos para la aplicación con Stata (1181,00) y para el grupo sin aplicación (415,00), en la dimensión medidas estadísticas, también se aprecia un valor de U. Mann Whitney (9,000) un valor de z-obtenido (-6,323) siendo mayor que el z-crítico (-1.96), además se aprecia que p-valor (0,00) es menor al valor de (0,05), razones para inferir que si se rechaza la hipótesis nula y concluir que si existe impacto preponderante del aplicativo Stata en el análisis de datos en la dimensión medidas de tendencia central y dispersión.

Con respecto a la hipótesis específica 2, también se estima al 95% de nivel de confianza y se planteó si existe impacto positivo del aplicativo Stata en el análisis de datos dimensión correlación y regresión y como hipótesis nula no existe impacto positivo del aplicativo Stata en el análisis de datos dimensión correlación y regresión. A continuación, muestro los resultados con ayuda del SPSS.

Tabla 12.

Prueba con U. Mann Whitney impacto de Stata dimensión correlación y regresión

Rangos de la prueba			Estadísticos de prueba			
Métodos	n	Rango promedio	Suma de rangos	U. Mann-Whitney	Z- Calculado	Sig. asintótica (bilateral)
Con Stata	28	42,48	1189,50	0.500	-6,495	0,00
Tradicional	28	14,52	406,50			

Fuente: Elaboración propia (2024).

Se evidencian los resultados del contraste de la hipótesis en la tabla 12 para la dimensión correlación y regresión donde el rango promedio para grupo experimental fue de (42,48) y para control fue de (14,52), por otro lado, se observa un valor del z-calculado de (-6,495) mayor al valor de z- tabla (1,96), razón suficiente para decidir rechazar la hipótesis nula y aceptar la alterna, concluyendo que sí existe impacto positivo del aplicativo Stata en el análisis de datos dimensión correlación y regresión en los estudiante de posgrado de la mención doctorado.

Con respecto a la hipótesis específica 3 se planteó como hipótesis verdadera si existe impacto significativo del software Stata en el análisis de datos dimensión pruebas de hipótesis y como hipótesis nula su negación, se realizó los cálculos utilizando la fórmula de U. Mann Whitney y con apoyo del software estadístico SPSS se obtuvo los resultados.

Tabla 13.

Prueba con U. Mann Whitney Impacto dimensión prueba de hipótesis

Rangos de la prueba			Estadísticos de prueba			
Métodos	n	Rango promedio	Suma de rangos	U. Mann-Whitney	Z- Calculado	Sig. asintótica (bilateral)
Con Stata	28	42,05	1177,50	12,500	-6,283	0,00
Sin aplicar	28	14,95	418,50			

Fuente: Elaboración propia (2024) a partir de los datos del contraste de hipótesis.

De la tabla 13 se evidencian los resultados producto de la prueba con U. Mann Whitney para el contraste de hipótesis 3, es así, se aprecia un valor de suma de rangos para el grupo experimental de (1177,50) y para el grupo control que no recibió el tratamiento, es decir no uso el software Stata obtuvo (418,50), también se evidencia un estadístico de U. Mann Whitney de (12,500) con un sig. asintótica bilateral de (0, 00) inferior al nivel de significancia (0,05). Por lo tanto, tomamos la decisión de rechazar también la hipótesis nula y aceptar la verdadera. Infiriendo estadísticamente que sí existe impacto significativo del software Stata en el análisis de datos dimensión pruebas de hipótesis en la asignatura estadística aplicada al campo de la investigación en los alumnos de Posgrado de la Universidad Enrique Guzmán y Valle, Lima, Perú.

4. Discusión

Con respecto a la validez de contenido de los instrumentos de investigación se recurrió a 5 jueces expertos en el tema quienes evaluaron y dieron su calificación promedio en porcentaje obteniéndose un promedio general de 92% y 93% para cada instrumento, luego para saber si los ítems eran adecuados en base a los indicadores se recurrió a la fórmula del V. Aiken y con ayuda del software Ms. Excel se obtuvo resultados de (0,82) y (0,83) respectivamente, estos valores nos indicaron que ningún ítems sería reformulado ni eliminado es decir cumplieron con una adecuada validez de contenido.

En cuanto a la confiabilidad de los instrumentos aplicados, se recurrió al método de dos mitades y con la fórmula de Spearman Brown que consistió en dividir los ítems en dos partes pares e impares y luego se procesó las puntuaciones obtenidas mediante una muestra piloto con dicha fórmula y con apoyo del software estadístico SPSS se procesaron y se evidenciaron los resultados para el instrumento denominado Preprueba fue (0,87) para el instrumento Posprueba (0,88), estas puntuaciones nos indicaron que los instrumentos si tenían muy alta confiabilidad es decir, eran consistentes y coherentes para su uso.

Con respecto al análisis descriptivo se crearon tablas y figuras con el fin de obtener una información inicial y luego interpretarlos en base a las dimensiones e indicadores de la variable pendiente, es decir, de modo que conoceríamos el comportamiento de la variable análisis de datos cuantitativo producto de su influencia del software Stata. Luego se realizó un análisis de datos a nivel inferencial donde se realizó una prueba de normalidad utilizando Shapiro Wilk, debido a que la muestra fue menor a 50 y con el fin de conocer la distribución de los datos, una vez conocida utilizamos una fórmula adecuada para la prueba, siendo el estadístico no paramétrico U. Mann Whitney, adecuado para la estrategia empleada, es decir un diseño cuasi experimental para comparar dos grupos independientes.

En cuanto a la comparación con el estudio con el de Hidalgo (2019), que empleo en los cálculos estadísticas un análisis cuantitativo de datos utilizando variables categóricas y numéricas para tal, utilizó diversas fórmulas estableciendo interrelaciones entre variables con fines explicativos y predictivos. Nuestro estudio es de tipo experimental donde se da a conocer el impacto del software Stata en el análisis de datos cuantitativos a nivel de estudiantes de posgrado, los resultados fueron muy significativos, mejoraron el nivel teórico y práctico de las competencias estadísticas en los alumnos con apoyo de las TIC, para el docente una nueva estrategia de enseñanza, impactando en el campo de la investigación científica.

5. Conclusiones

Se concluye al 95% de nivel de confianza que el impacto del software estadístico Stata sí influye positivamente en el análisis de datos cuantitativos, es decir, los estudiantes de posgrado mejoran sus niveles en sus capacidades y competencias estadísticas al resolver problemas relacionados al campo de la investigación, además de elevar su motivación los procedimientos fueron muy fácil al utilizar el software de manera rápida e interactiva.

Respecto al objetivo 1 se concluye que el impacto del STATA también fue muy significativo al emplear en cálculos de medidas de tendencia central y dispersión la rapidez del entorno y sus procedimientos elevaron sus niveles de resolver problemas y el aprendizaje fue más activo y motivado.

En cuanto al objetivo 2 se concluye también al 95% de nivel de confianza que la mayoría de los estudiantes del grupo experimental que utilizaron el software estadístico Stata en procesos estadísticos de correlación y regresión mejoraron sus niveles de aprendizaje en comparación al grupo de estudiantes que solo desarrollaron las clases en forma tradicional.

En cuanto al objetivo 3 se concluye también al 95% de nivel de confianza que más del 50% de los estudiantes del grupo experimental que utilizaron el software estadístico Stata en procesos de prueba de hipótesis mejoraron sus niveles de logro haciendo uso de diversas fórmulas paramétricas y no paramétricas en el campo de la investigación al realizar sus tesis y trabajos de investigación.

Finamente se concluye al 5% de significancia que al aplicar el software Stata los promedios obtenidos para el grupo experimental fueron superiores al promedio del grupo control, Es decir, su impacto del software fue positivo al realizar procedimientos estadísticos en campo de la investigación es estudiantes de posgrado de la Universidad Enrique Guzmán y Valle. Perú.

6. Referencias

- Arias, J. y Covinos, M. (2021). *Diseño y metodología de la investigación*. Repositorio CONCYTEC. <https://hdl.handle.net/20.500.12390/2260>
- Cisneros-Caicedo, A. J., Guevara-García, A. F., Urdánigo-Cedeño, J. J. y Garcés-Bravo, J. E. (2022). Técnicas e instrumentos para la recolección de datos que apoyan a la investigación científica en tiempo de pandemia. *Domino de las Ciencias*, 8(1), 1165-1185. <https://doi.org/10.23857/dc.v8i1.2546>
- Escobar, M., Fernández, E. y Bernardi, F. (2022). *Análisis de datos con Stata* (2ª ed.). Centro de investigaciones sociológicas. <https://bit.ly/4cFsMiG>
- García García, J. I., Imilpán Rivera, I., Díaz Levicoy, D. y Arredondo, E. H. (2021). Las medidas de tendencia central en libros de texto de séptimo básico de Chile. *Conrado*, 17(81), 261-268. <https://bit.ly/3OwXXSd>
- Gea, M., Batanero, C. y Roa, R. (2014). El sentido de la correlación y regresión. *Números*, 87, 25-35. <https://bit.ly/456m3ug>
- Guerrero Támara, V. (2022). Enfoque cuantitativo: taxonomía desde el nivel de profundidad de la búsqueda del conocimiento. *Llalliq*, 2(1), 13-27. <https://bit.ly/3zvFNwy>
- Hernández-Sampieri, R. y Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill Education. <https://bit.ly/4crNr9Q>
- Hernández-Sampieri, R., Fernández-Collado, C. y Baptista-Lucio, P. (2014). *Metodología de la investigación: Análisis de los datos cuantitativos*. McGraw-Hill Education. <https://bit.ly/3XXbl8I>
- Hidalgo, A. (2019). Técnicas estadísticas en el análisis cuantitativo de datos. *Revista Sigma*, 15(1), 28-44. <https://bit.ly/4crNnH8>

- Matzumura-Kasano, J., Gutiérrez-Crespo, H., Ruiz-Arias, A. y Huamán-Carhuas, L. (2021). Utilización de aplicativo Remind como herramienta de apoyo en la enseñanza de investigación básica. *Revistas de Investigación UNMSM*, 82(3), 9-21. <https://doi.org/10.15381/anales.v82i3.20470>
- Mercado Mercado, E., Fernández Macías, E. y Bernardi, F. (2012). *Análisis de datos con Stata. Cuadernos metodológicos*. Centro de Investigaciones Sociológicas. <https://bit.ly/3OFTf5I>
- Millones, T., Alencar, S. y Bueno, S. (2021). Dificultades en la enseñanza y el aprendizaje de la estadística: comparación de investigaciones de Perú y Brasil, 2009 a 2017. *Revista Research, Society and Development*, 10(12). <https://doi.org/10.33448/rsd-v10i12.19975>
- Perez-Leon, G. [gabriel-perez-leon]. (11 de septiembre de 2023). *Importancia del análisis de datos*. LinkedIn. <https://bit.ly/40EUrev>
- Sierra, J., Palmezano, Y. y Romero, B. (2022). Causas que determinan las dificultades de la incorporación de las TIC. *Aulas de clases*. <https://bit.ly/3XO8m2c>
- Solorzano, F. C., Apaza, H. A. C., Vilcanqui, B. M., Mercado, C. M. Z. y Polo, A. E. A. (2022). El uso de las pruebas de hipótesis en la investigación educativa. *Universidad y Sociedad*, 14(1), 701-711. <https://bit.ly/3RTvSqJ>
- Vásquez, C. y García, A. (2020). La educación estadística para el desarrollo sostenible en la formación del profesorado. *Profesorado. Revista de Currículum y Formación de Profesorado*, 24(3), 125-147. <https://doi.org/10.30827/profesorado.v24i3.15214>
- Vela Meléndez, L. y Guerrero Carrasco, G. E. (2020). *Guía para la construcción de modelos de regresión lineal clásico y modelos de elección binaria con STATA 15*. <https://bit.ly/3XPdVO7>

CONTRIBUCIONES DE AUTORES/AS, FINANCIACIÓN Y AGRADECIMIENTOS

Contribuciones de los/as autores/as:

Conceptualización: Caballero Cifuentes, Lolo; **Software:** Huamán Hurtado, Juan; **Validación:** Caballero Cifuentes, Lolo; **Análisis formal:** Caballero Cifuentes, Lolo; **Curación de datos:** Caballero Cifuentes, Lolo; **Redacción-Preparación del borrador original:** Caballero Cifuentes, Lolo; **Redacción-Re-visión y Edición:** Caballero Cifuentes, Lolo; **Visualización:** Quivio Cuno, Richard; **Supervisión:** Caballero Cifuentes, Lolo; **Administración de proyectos:** Caballero Cifuentes, Lolo. **Todos los/as autores/as han leído y aceptado la versión publicada del manuscrito:** Caballero Cifuentes, Lolo; Quivio Cuno, Richard; Gutiérrez Guadalupe, Sandra; Sito Justiniana, Luz; Jaime Andía, Mario; Alegre Huerta, Jesús; Huamán Hurtado, Juan.

Financiación: Esta investigación ha recibido financiación pública de parte de la universidad para el proceso de elaboración, más no para su publicación

Agradecimientos: El presente texto nace como producto de una investigación desarrollada por los docentes de la Facultad de ciencias de la Universidad Enrique Guzmán y Valle, Perú.

Conflicto de intereses: Los autores declaramos no tener ningún conflicto de interés.

AUTOR/ES:**Lolo Caballero Cifuentes:**

Universidad Nacional de Educación.

Doctor en Ciencias de la Educación, con Maestría en Problemas de Aprendizaje, Licenciado en la especialidad de Matemática e Informática por la Universidad Enrique Guzmán y Valle, Con Diplomado en Estadística Aplicada a la Investigación Científica. Con más de 18 años de experiencia como docente universitario de Pregrado y Posgrado en instituciones públicas y privadas. Mi línea de investigación está orientada a las Tecnologías de la Información y Comunicación, matemática, estadística y educación; realizo proyectos de investigación con fondos concursables y participo como docente en diversos diplomados de estadística aplicada a la investigación Científica. Dicto asignaturas en la Escuela de Postgrado en los niveles de Doctorado y Maestría: Estadística Aplicada a la Investigación Científica.

lcaballero@une.edu.pe

Índice H: 1

Orcid ID: <https://orcid.org/0000-0002-9764-1327>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=58567875000>

Richard Quivio Cuno:

Universidad Nacional de Educación.

Doctor en Ciencias de la Educación. Magister en Docencia Universitaria. Licenciado en Matemática e Informática. Profesor Principal a Dedicación Exclusiva de la Escuela Académico profesional de Matemática e Informática, adscrito al Departamento Académico de Matemática e Informática de la Universidad Nacional de Educación Enrique Guzmán y Valle, así como docente de la Escuela de Posgrado en las cátedras de Estadística aplicada a la Investigación e Investigación científica (Tesis I, II, III) en la sección Maestría y Doctorado. Tiene experiencia en el campo académico de Educación en Matemática, Informática, Estadística e Investigación, actuando principalmente en las siguientes áreas de investigación: Estrategias de aprendizaje en Matemática, Arquitectura de computadoras y desarrollo de software educativo, procesos estadísticos aplicados a la investigación.

rquivio@une.edu.pe

Orcid ID: <https://orcid.org/0000-0002-5986-3711>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=58096948700>

Sandra Gutiérrez Guadalupe:
Universidad Nacional de Educación.

Bachiller en Ciencias de la Educación (UNE EGyV-2006), título de Licenciado en Educación Matemática e Informática (UNE EGyV-2006); Maestría en Ciencias de la Educación, con mención en Educación Tecnológica (UNE EGyV-2008) y Doctorado en Ciencias de la Educación (UNE EGyV-2011). Actualmente es Profesor Principal a Dedicación Exclusiva de la Escuela Académico profesional de Matemática e Informática, adscrito al Departamento Académico de Matemática e Informática de la Universidad Nacional de Educación Enrique Guzmán y Valle, así como docente de la Escuela de Posgrado en la cátedra de Estadística aplicada a la Investigación en la sección Maestría.

sgutierrezg@une.edu.pe

Orcid ID: <https://orcid.org/0000-0003-4188-7448>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=58096868100>

Mario Jaime Andía:
Universidad Nacional de Educación.

Docente Auxiliar Tiempo completo de la Facultad de Ciencias en la especialidad de Matemática e Informática de la Universidad Nacional de Educación Enrique Guzmán y Valle. Doctor en Ciencias de la Educación y Magíster en la mención de Docencia Universitaria por la Universidad Cesar Vallejo. Docente en diversos cursos de Informática en Universidades públicas y privadas.

mjaime@une.edu.pe

Orcid ID: <https://orcid.org/0000-0002-9026-3060>

Jesús Alegre Huerta:
Universidad Nacional de Educación.

Docente Auxiliar Tiempo completo de la Facultad de Ciencias en la especialidad de Matemática e Informática de la Universidad Nacional de Educación Enrique Guzmán y Valle. Magíster en la mención de Educación Matemática por la Universidad Nacional de Educación Enrique Guzmán y Valle. Docente en diversos cursos de Informática en Universidades públicas y privadas.

jalegre@une.edu.pe

Orcid ID: <https://orcid.org/0009-0008-8944-6143>

Juan Huamán Hurtado:

Universidad Nacional de Educación.

Bachiller en Ciencias de la Educación, título de Licenciado en Educación Matemática Física; Maestría en Ciencias de la Educación, con mención en Docencia Universitaria y Doctorado en Ciencias de la Educación. Actualmente es Profesor Principal a Dedicación Exclusiva de la Escuela Académico profesional de Matemática e Informática, adscrito al Departamento Académico de Matemática e Informática de la Universidad Nacional de Educación Enrique Guzmán y Valle, así como docente de la Escuela de Posgrado en la cátedra de Estadística aplicada a la Investigación e informática en la sección Maestría.

jhuaman@une.edu.pe

Orcid ID: <https://orcid.org/0000-0003-1064-7062>

Luz Sito Justiniano:

Universidad Nacional de Educación.

Docente Principal de la UNE, adscrita a la Facultad de Educación Inicial, formo maestras de Educación Inicial hace 38 años, tiempo de servicio en la docencia universitaria, trabajé como docente contratada y en el 1985 nombrada hasta la fecha, hace 24 años docente principal de la UNE. Licenciada en Psicología, Educación Inicial y Educación Primaria; Doctora en Educación, Psicología y en Gobierno y Políticas Públicas, defensora de la primera infancia, los niños y sus familias, la mujer y los ancianos; por mi formación en educación, psicología, gerencia social y políticas públicas.

lsito@une.edu.pe

Orcid ID: <https://orcid.org/0000-0001-6212-5982>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57487177600>