



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA



Desarrollo de una Interfaz Gráfica de Usuario en Matlab
para Potenciar el aprendizaje de Regresión Lineal en la
Educación Universitaria

TESIS

Para optar el título profesional de
Licenciado(a) en Estadística

Autoras:

Bach. Est. Pozo Machado Brigge Kimberley

Bach. Est. Silva Oblitas Ricardo Junior

Asesor:

Mg. Alpañil Santisteban Víctor Teofilo

LAMBAYEQUE – PERÚ

2026

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA

Los firmantes, por la presente certifican que han leído y recomiendan a la Facultad de Ciencias Físicas y Matemáticas la aceptación de la tesis titulada **Desarrollo de una Interfaz Gráfica de Usuario en Matlab para Potenciar el aprendizaje de Regresión Lineal en la Educación Universitaria**, presentada por las bachilleres en Estadística, Pozo Machado Brigge Kimberley y Silva Oblitas Ricardo Junior, en el cumplimiento parcial de los requisitos necesarios para la obtención del título profesional de licenciados en Estadística.



Dr. Lic. Est. Alvarado Castillo Wilder Ángel

Presidente Jurado de Tesis



Dr. Lic. Est. Benavides Campos Grimaldo Dermalí

Secretario Jurado de Tesis



M.Sc. Lic. Est. Macalopú Inga Hipólito

Vocal Jurado de Tesis

Fecha de Defensa: 30 de julio de 2026

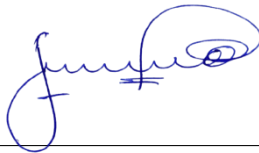
UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA

Desarrollo de una Interfaz Gráfica de Usuario en Matlab
para Potenciar el aprendizaje de Regresión Lineal en la
Educación Universitaria



Bach. Est. Pozo Machado Brigge Kimberley

Autora



Bach. Est. Silva Oblitas Ricardo Junior

Autor



Mg. Albañil Santisteban Víctor Teofilo

Asesor

Lambayeque – Perú

Julio-2026



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DECANATO

Ciudad Universitaria - Lambayeque

LICENCIADA - RESOLUCIÓN DEL CONSEJO DIRECTIVO N° 015 -2023-SUNEDU / CD

UNIVERSIDAD NACIONAL
SECRETARÍA GENERAL
PEDRO RUIZ GALLO
0095



ACTA DE SUSTENTACIÓN N° 055.-2026.-D/FACFyM

Siendo las 12:00 a.m. del día 30 de julio del 2026, se reunieron los miembros del jurado evaluador de la Tesis titulada:

"Desarrollo de una Interfaz Gráfica de Usuario en Matlab para Potenciar el aprendizaje de Regresión Lineal en la Educación Universitaria"

Designados por Resolución N° 174-2025 - D/FACFyM de fecha 21 de febrero 2025

Con la finalidad de evaluar y calificar la sustentación de la tesis antes mencionada, conformada por los siguientes docentes:

Dr. Lic. Est. Alvaro do Castillo Wilder Angel Presidente
Dr. Lic. Est. Genaides Campa Germal b Dermal Secretario
MSc. Lic. Est. Macalopu Hipolito Vocal

La tesis fue asesorada por (el) (la) Mg. Albaril Santisteban Victor Trifilo, nombrado por Resolución N° 174-2025 - D/FACFyM de fecha 21 de febrero 2025

El Acto de Sustentación fue autorizado por Resolución N° 546-2026 - D/FACFyM de fecha 24 de julio 2026

La Tesis fue presentada y sustentada por (el) (los) Bachiller (es): Rozo Machado Briggie Kimberley y Silva Obhitas Ricardo Junior y tuvo una duración de 30 minutos.

Después de la sustentación, y absueltas las preguntas y observaciones de los miembros del jurado se procedió a la calificación respectiva, otorgándole el Calificativo de DIECINUEVE (19) en la escala vigesimal, mención (MUY BUENO).

Por lo que queda(n) apto(s) para obtener el Título Profesional de Licenciado en Estadística de acuerdo con la Ley Universitaria 30220 y la normatividad vigente de la Facultad de Ciencias Físicas y Matemáticas y la Universidad Nacional Pedro Ruiz Gallo.

Siendo las 12:50 p.m. se dio por concluido el presente acto académico, dándose conformidad al presente acto con la firma de los miembros del jurado.

Dr. Wilder Angel Alvaro do Castillo
Presidente

MSc. Hipolito Macalopu Inga
Vocal

Dr. Germal b Dermal Genaides Campa
Secretario

Mg. Albaril Santisteban Victor Trifilo
Asesor

CONSTANCIA DE VERIFICACIÓN DE ORIGINALIDAD

Yo, Mg. Albañil Santisteban Víctor Teofilo, usuario revisor de la tesis titulada: **Desarrollo de una Interfaz Gráfica de Usuario en Matlab para Potenciar el aprendizaje de Regresión Lineal en la Educación Universitaria**, cuyos autores Pozo Machado Brigge Kimberley, identificado con D.N.I. N°. 73001591 y Silva Oblitas Ricardo Junior, identificado con D.N.I. N°. 71413151, ha arrojado un porcentaje de similitud de **18%**, verificable en el Resumen de Reporte automatizado de similitudes que se acompaña.

El suscrito analizó dicho reporte y concluyó que cada una de las coincidencias detectadas dentro del porcentaje de similitud permitido no constituyen plagio y que el documento cumple con la integridad científica y con las normas para el uso de citas y referencias establecidas en los protocolos respectivos.

Se cumple con adjuntar el Recibo Digital a efectos de la trazabilidad respectiva del proceso.

Lambayeque, 13 de julio de 2026



Mg. Albañil Santisteban Víctor Teofilo

DNI: 17633734

Asesor

Desarrollo de una Interfaz Gráfica de Usuario en Matlab para Potenciar el aprendizaje de Regresión Lineal en la Educación Universitaria

INFORME DE ORIGINALIDAD

18%

INDICE DE SIMILITUD

17%

FUENTES DE INTERNET

7%

PUBLICACIONES

8%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	repositorio.unprg.edu.pe	Fuente de Internet	5%
2	repositorio.ucv.edu.pe	Fuente de Internet	1%
3	Submitted to Universidad Cesar Vallejo	Trabajo del estudiante	1%
4	Submitted to Universidad Nacional Pedro Ruiz Gallo	Trabajo del estudiante	1%
5	Submitted to University of California, Los Angeles	Trabajo del estudiante	1%
6	archive.org	Fuente de Internet	<1%
7	repositorio.uns.edu.pe	Fuente de Internet	<1%
8	repositorio.unc.edu.pe	Fuente de Internet	<1%
9	repositorio.usil.edu.pe	Fuente de Internet	<1%
10	www.unprg.edu.pe	Fuente de Internet	<1%
11	Submitted to Unviersidad de Granada	Trabajo del estudiante	<1%
12	idoc.pub	Fuente de Internet	<1%
13	repository.poltekkes-malang.ac.id	Fuente de Internet	<1%
14	www.studocu.com	Fuente de Internet	<1%



Mg. Albañil Santisteban Victor Teofilo
Departamento de Estadística
DNI: 17633734
Asesor

Submitted to Docentes



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por **Turnitin**. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Pozo Machado Brigge Kimberley, Silva Oblitas Ricardo Junior
Título del ejercicio: Quick Submit
Título de la entrega: Desarrollo de una Interfaz Gráfica de Usuario en Matlab para ...
Nombre del archivo: TESIS_BRIGIT-RICARDO-2026.pdf
Tamaño del archivo: 3.41M
Total de páginas: 153
Total de palabras: 24,564
Total de caracteres: 129,068
Fecha de entrega: 11-ago-2026 09:51a. m. (UTC-0500)
Identificador de la entrega: 3011149754



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA



Desarrollo de una Interfaz Gráfica de Usuario en Matlab
para Potenciar el aprendizaje de Regresión Lineal en la
Educación Universitaria

TESIS

Para optar el título profesional de
Licenciado(a) en Estadística

Autoras:

Bach. Est. Pozo Machado Brigge Kimberley
Bach. Est. Silva Oblitas Ricardo Junior

Asesor:

Mg. Albañil Santisteban Victor Teofilo

LAMBAYEQUE - PERÚ
2026

Mg. Albañil Santisteban Victor Teofilo
Departamento de Estadística
DNI: 17633734
Asesor

Agradecimientos

Agradezco a Dios por permitirme llegar hasta aquí y darme fuerzas para seguir adelante.

Agradezco a mi familia por su amor, comprensión y apoyo durante todo este proceso. A mis asesores y docentes, por su guía, sus observaciones y su acompañamiento académico, que ayudaron a mejorar esta investigación. Agradezco también a mis amistades y compañeros, por la ayuda brindada, la motivación y el respaldo en cada etapa. Por último, agradezco a todas las personas que aportaron con su tiempo y colaboración para hacer posible este trabajo.

Brigge

Agradezco a Dios por la vida y la fortaleza para culminar esta etapa. Mi agradecimiento especial a mis padres y familia por su apoyo constante, paciencia y ánimo en los momentos difíciles. A mis docentes y asesores, por su orientación académica, sus correcciones y el tiempo dedicado para mejorar este trabajo. También agradezco a mis compañeros y amistades, por el acompañamiento y las palabras de aliento durante el proceso. Finalmente, agradezco a todas las personas que directa o indirectamente aportaron para que esta investigación se hiciera realidad.

Ricardo

Dedicatoria

Dedico este trabajo a mi familia, por su cariño y apoyo incondicional. A mis padres, por enseñarme a no rendirme. A mis docentes, por su paciencia y enseñanza. Y a mí misma, por la perseverancia para continuar incluso cuando fue difícil. Este logro representa esfuerzo, disciplina y esperanza.

Brigge

Dedico este trabajo a mi familia, por ser mi motor y sostén en cada paso. A mis padres, por su ejemplo y sacrificio. A mis docentes, por compartir sus conocimientos y guiarme con responsabilidad. Y a todas las personas que creyeron en mí cuando el camino se hizo pesado; este logro también les pertenece.

Ricardo

Índice

Agradecimientos	I
Dedicatoria	II
Índice	III
Índice de tablas	VI
Índice de figuras	VII
Resumen	VIII
Abstract	IX
Introducción	X
I. DISEÑO TEÓRICO	1
1.1 Antecedentes de la Investigación	1
1.1.1 A nivel internacional	1
1.1.2 A nivel nacional	4
1.2 Base Teórica	5
1.2.1 Regresión lineal simple	5
1.2.1.1 Medias	6
1.2.1.2 Suma de cuadrados y producto cruzado (misma notación)	6
1.2.1.3 SCE (Suma de Cuadrados del Error)	7
1.2.1.4 Teorema principal: fórmula de la pendiente y del intercepto	7
1.2.1.5 Sumas de cuadrados: STC , SCR , SCE	10
1.2.1.6 Teorema de descomposición: $STC = SCR + SCE$	11
1.2.1.7 Coeficiente de determinación	12
1.2.1.8 Error cuadrático medio y su raíz	13
	III

1.2.1.9 Supuestos del modelo de regresión lineal	14
1.2.2 Regresión lineal múltiple	17
1.2.3 Notación de sumatorias (para construir la matriz)	17
1.2.4 Forma matricial: sistema lineal para estimar $b_0, b_1, \dots, b_k$	18
1.2.5 Variable 1: Desarrollo de una Interfaz Gráfica de Usuario en MATLAB	19
1.2.6 Variable 2: Aprendizaje de Regresión Lineal	23
1.3 Bases Conceptuales	28
1.3.1 Regresión Lineal	28
1.3.2 Regresión Lineal Simple	28
1.3.3 Regresión Lineal Múltiple	28
1.3.4 Educación Universitaria	28
1.3.5 Interfaz Gráfica de Usuario (GUI)	28
1.3.6 MATLAB:	29
1.3.7 Aprendizaje Activo:	29
1.3.8 Tecnologías Educativas:	29
1.3.9 Variable independiente	29
1.3.10 Variable dependiente	30
1.3.11 Operacionalización de variables	30
II. DISEÑO METODOLÓGICO	32
2.1 Tipo de Investigación	32
2.2 Método de Investigación	32
2.3 Diseño de Contrastación	33
Procedimiento de aplicación	34
2.4 Población, Muestra y Muestreo	35
2.5 Técnicas, Instrumentos, Equipos y Materiales de Recolección de Datos	35
2.6 Procesamiento y Análisis de Datos	36

III. RESULTADOS	39
3.1 Interfaz Gráfica de Usuario para la regresión lineal simple	39
3.1.1 Aplicación de la regresión lineal simple	46
3.2 Interfaz Gráfica de Usuario para la regresión lineal múltiple	56
3.2.1 Aplicación de la regresión lineal múltiple	57
3.3 Validación de la interfaz gráfica con estudiantes de Ingeniería Civil	65
3.3.1 Resultados del pretest de regresión lineal	65
3.3.2 Resultados del postest de regresión lineal	66
3.3.3 Comparación de los resultados del pretest y postest	67
3.3.4 Prueba de normalidad	67
3.3.5 Prueba t de Student para muestras relacionadas	68
IV. DISCUSIÓN DE RESULTADOS	69
V. CONCLUSIONES	74
RECOMENDACIONES	75
Referencias Bibliográficas	76
Anexos	79

Índice de tablas

1.	Operacionalización de variables.	31
2.	Técnicas e instrumentos de recolección de datos	36
3.	Población de estudiantes y ventas trimestrales en 10 restaurantes	46
4.	Tabla de cálculos para la regresión lineal	47
5.	Calificación de estadística, calificación del examen y clases perdidas ($n = 12$)	57
6.	Valores observados, estimados y errores del modelo de regresión ($n = 12$) .	57
7.	Resultados descriptivos del pretest de regresión lineal	66
8.	Resultados descriptivos del posttest de regresión lineal	66
9.	Comparación de resultados del pretest y posttest	67
10.	Prueba de normalidad de las diferencias entre pretest y posttest	67
11.	Prueba t de Student para muestras relacionadas	68
12.	Resultados del Pretest - Ingeniería Civil	128
13.	Resultados del Posttest - Ingeniería Civil.	133
14.	Resumen de las notas obtenidas por los estudiantes en el pretest y posttest .	134
15.	Escala de valoración	135
16.	Cuestionario de usabilidad de la interfaz gráfica	135
17.	Resultados del cuestionario de usabilidad de la interfaz gráfica.	138

Índice de figuras

1.	<i>Ingreso de la palabra guide.</i>	39
2.	<i>Entorno de la GUI.</i>	39
3.	<i>Se colocó el nombre de regresión lineal simple.</i>	40
4.	<i>Guardar.</i>	40
5.	<i>Nombre de la GUI.</i>	41
6.	<i>Entorno de la GUI.</i>	41
7.	<i>Ingreso de 14 paneles.</i>	42
8.	<i>Ingreso de 6 pushbutton.</i>	42
9.	<i>Se ingresó 5 edit text.</i>	43
10.	<i>Ingresar 4 tablas.</i>	43
11.	<i>Se ingresó un axes para mostrar la gráfica de la ecuación lineal.</i>	44
12.	<i>Se colocó 3 static text.</i>	44
13.	<i>Diseño de la (GUI).</i>	45
14.	<i>Diseño de la GUI de la regresión lineal simple.</i>	45
15.	<i>GUI de la regresión lineal simple.</i>	46
16.	<i>Solución mediante la GUI de regresión simple.</i>	52
17.	<i>Diseño final de la GUI de la regresión lineal múltiple.</i>	56
18.	<i>GUI de la regresión lineal múltiple.</i>	56
19.	<i>Solución mediante la GUI de regresión múltiple.</i>	62

Resumen

La investigación tuvo como objetivo desarrollar una interfaz gráfica de usuario en MATLAB para potenciar el aprendizaje de regresión lineal en la educación universitaria. La propuesta tecnológica comprendió el diseño y programación de dos módulos: regresión lineal simple y regresión lineal múltiple. La interfaz fue elaborada en MATLAB R2021a e incorporó funciones para ingresar datos, generar tablas de cálculo, estimar parámetros, visualizar ecuaciones, elaborar gráficos, calcular valores estimados, residuos y métricas de ajuste. La investigación presentó un enfoque cuantitativo, fue de tipo aplicada y empleó un diseño preexperimental de un solo grupo con pretest y posttest para validar la propuesta en estudiantes de Ingeniería Civil. La aplicación de la interfaz permitió que los participantes desarrollaran actividades de regresión lineal simple y múltiple, interpretaran la pendiente, el intercepto, el coeficiente de determinación, los residuos y las predicciones obtenidas mediante el modelo. Asimismo, se aplicó un cuestionario de usabilidad para conocer la percepción de los estudiantes sobre el diseño, funcionalidad, interactividad y utilidad de la GUI. Los resultados permitieron comprobar el funcionamiento correcto de la interfaz mediante la comparación de los cálculos con procedimientos manuales y funciones internas de MATLAB. Además, la comparación entre los resultados del pretest y posttest permitió identificar cambios en el aprendizaje de regresión lineal después del uso de la herramienta. Se concluyó que la interfaz gráfica de usuario en MATLAB constituyó un recurso didáctico y tecnológico que apoyó el aprendizaje de regresión lineal en el contexto de la educación universitaria.

Palabras clave: Interfaz gráfica de usuario, MATLAB, regresión lineal, aprendizaje, educación universitaria.

Abstract

The aim of this research was to develop a graphical user interface in MATLAB to enhance the learning of linear regression in higher education. The technological proposal included the design and programming of two modules: simple linear regression and multiple linear regression. The interface was developed in MATLAB R2021a and incorporated functions for data entry, calculation tables, parameter estimation, equation display, graph generation, fitted values, residuals, and goodness-of-fit measures. The study followed a quantitative approach, was applied in nature, and used a one-group pre-experimental design with pre-test and posttest to validate the proposal with Civil Engineering students. The use of the interface allowed participants to carry out activities related to simple and multiple linear regression, interpret the slope, intercept, coefficient of determination, residuals, and predictions produced by the model. In addition, a usability questionnaire was applied to identify students' perceptions regarding the design, functionality, interactivity, and usefulness of the GUI. The results confirmed the correct functioning of the interface through comparisons with manual procedures and MATLAB built-in functions. Furthermore, the comparison between pretest and posttest results made it possible to identify changes in linear regression learning after using the tool. It was concluded that the MATLAB graphical user interface constituted a didactic and technological resource that supported the learning of linear regression in higher education.

Keywords: Graphical user interface, MATLAB, linear regression, learning, higher education.

Introducción

La regresión lineal constituye un contenido utilizado para analizar datos, estudiar relaciones entre variables y realizar predicciones en distintas carreras universitarias. Sin embargo, su enseñanza presenta dificultades en diversos contextos educativos. En muchos casos, se desarrolla mediante procedimientos tradicionales centrados en fórmulas y operaciones mecánicas, lo que puede dificultar la comprensión del significado de los parámetros y resultados obtenidos. Aunque existen recursos tecnológicos para apoyar su enseñanza, su incorporación todavía es limitada, por lo que algunos estudiantes presentan dificultades para relacionar los procedimientos estadísticos con situaciones aplicadas.

En el Perú, esta situación se relaciona con la limitada incorporación de recursos tecnológicos en algunos espacios universitarios y con la necesidad de disponer de herramientas adaptadas al proceso de aprendizaje. Los estudiantes pueden presentar dificultades para interpretar la ecuación de regresión, el coeficiente de determinación, los valores estimados, los residuos y las medidas de error. Además, programas como Python, R o SPSS requieren conocimientos previos para su manejo, lo que puede dificultar su utilización por estudiantes que se encuentran en las primeras etapas de su formación universitaria.

En Lambayeque, particularmente en la Universidad Nacional Pedro Ruiz Gallo, estas dificultades también pueden presentarse durante el aprendizaje de contenidos estadísticos. Frente a esta situación, el desarrollo de una interfaz gráfica en MATLAB constituye una alternativa tecnológica que permite organizar los procedimientos de regresión lineal y mostrar de forma secuencial los cálculos realizados. Mediante esta herramienta, el estudiante puede ingresar datos, observar los resultados obtenidos, interpretar los parámetros del modelo y realizar predicciones sin ejecutar directamente instrucciones de programación.

Por esta razón, la investigación se orientó a responder la siguiente pregunta: ¿De qué manera el desarrollo de una interfaz gráfica de usuario en MATLAB potencia el aprendizaje de regresión lineal en la educación universitaria?

Se planteó como hipótesis que la aplicación de la interfaz gráfica de usuario en MATLAB mejora significativamente el aprendizaje de regresión lineal en los estudiantes de Ingenie-

ría Civil. El objetivo general fue desarrollar una interfaz gráfica de usuario en MATLAB para potenciar el aprendizaje de regresión lineal en la educación universitaria. Los objetivos específicos se orientaron a analizar e incorporar los fundamentos de la regresión lineal simple y múltiple, diseñar y validar técnicamente la interfaz gráfica y evaluar su aporte al aprendizaje de regresión lineal en estudiantes universitarios.

El aporte de la investigación consistió en el desarrollo de interfaces gráficas en MATLAB R2021a para trabajar con regresión lineal simple y múltiple. Estas incorporaron opciones para ingresar o cargar datos, generar tablas de cálculo, estimar los coeficientes del modelo, obtener valores estimados y residuos, calcular sumas de cuadrados y determinar medidas de ajuste como el coeficiente de determinación, el error cuadrático medio y la raíz del error cuadrático medio. Asimismo, permitieron realizar predicciones a partir de nuevos valores de entrada.

El funcionamiento técnico de la interfaz fue comprobado mediante la comparación de los resultados obtenidos con procedimientos manuales y funciones internas de MATLAB. Se verificó la consistencia de los coeficientes estimados, valores ajustados, residuos, medidas de ajuste y predicciones. Además, la interfaz fue convertida en un ejecutable independiente, permitiendo su utilización sin requerir la instalación completa de MATLAB R2021a.

La interfaz gráfica fue aplicada a 20 estudiantes del segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo. Para evaluar su aporte al aprendizaje, se aplicó un pretest antes de la intervención y un posttest después del uso de la herramienta. El promedio pasó de 11.15 puntos en el pretest a 17.15 puntos en el posttest, mostrando una diferencia de 6.00 puntos. La prueba (t) de Student para muestras relacionadas determinó que la diferencia fue estadísticamente significativa (($t=10.589$), ($gl=19$), ($p<0.001$)).

De esta manera, la investigación permitió integrar los fundamentos de la regresión lineal simple y múltiple en una herramienta gráfica, comprobar su funcionamiento técnico y evaluar su aplicación en estudiantes universitarios. Los resultados obtenidos mostraron que la interfaz desarrollada en MATLAB contribuyó a mejorar el aprendizaje de regresión lineal en los estudiantes participantes.

La tesis se organizó en 5 capítulos. El primer capítulo presentó el diseño teórico, los antecedentes, la base teórica y la operacionalización de variables. El segundo capítulo desarrolló el diseño metodológico. El tercer capítulo presentó los resultados. El cuarto capítulo expuso la discusión. Finalmente, se presentó las conclusiones y recomendaciones de la investigación.

I. DISEÑO TEÓRICO

1.1 Antecedentes de la Investigación

1.1.1 *A nivel internacional*

Tan et al. (2024) desarrollaron una interfaz gráfica (GUI) en MATLAB para apoyar la comprensión de la ley de Coulomb y la ley de Gauss mediante visualizaciones 3D y un juego serio tipo “snake”. La interfaz mostró distribuciones de carga, fuerzas eléctricas y flujo eléctrico, y permitió modificar parámetros definidos por el usuario para observar cambios en tiempo real. Se incluyeron simulaciones para distintas configuraciones de cargas en la ley de Coulomb y casos de la ley de Gauss con carga puntual, línea de carga infinita y plano cargado. Además, el juego incorporó preguntas para reforzar el contenido. Los hallazgos indicaron que, al interactuar con la GUI y el juego, el estudiantado comprendió mejor los conceptos y participó de manera más activa. Se concluyó que la herramienta facilitó el abordaje de nociones abstractas de electrostática al presentar representaciones visuales y actividades interactivas.

Osorio et al. (2023) estudiaron cómo el uso de actividades generadoras de modelos (AGM) se relacionó con el aprendizaje de correlación y regresión lineal, y describieron dificultades durante su aplicación. La investigación fue cualitativa, con enfoque de ingeniería didáctica, y participaron nueve docentes en formación de Colombia. Se diseñó una AGM con datos sobre calentamiento global, presentada como si proviniera de un portal de noticias ficticio. La intervención se organizó en seis fases: presentación del problema, trabajo individual, discusión grupal, consolidación, institucionalización del contenido y elaboración de un informe final. Se revisaron respuestas, representaciones gráficas y ajustes realizados a los modelos. Al inicio, el grupo trabajó sin planificación, usó representaciones poco útiles y no vinculó bien las variables; luego, con el avance del ciclo de modelación, mejoró sus propuestas, estableció correlaciones, construyó regresiones y calculó coeficientes de ajuste. Se observó exploración, análisis crítico y autoevaluación durante el proceso. Se concluyó que las AGM apoyaron el aprendizaje contextualizado de correlación y regresión lineal.

Vargas et al. (2023) identificaron los significados que futuros docentes de matemáticas atribuyeron a la educación estadística y describieron limitaciones en su formación. Se aplicó un diseño descriptivo cualitativo con 10 estudiantes de licenciatura en matemáticas del Caribe colombiano (20 a 25 años). Se usó un cuestionario de preguntas abiertas y se analizó con contenido temático. El grupo indicó que la estadística se vinculó con pensamiento crítico y toma de decisiones; aun así, reportó carencias en profundidad de contenidos, predominio de enseñanza transmisiva y ausencia de contexto aplicado. También se reportó poco uso de proyectos y herramientas tecnológicas. Se concluyó que la formación docente revisada necesitó ajustes curriculares, con metodologías basadas en proyectos e integración de tecnología.

Momox y Alonso (2023) desarrollaron una GUI en MATLAB para controlar instrumentos electrónicos en laboratorios de ingeniería de pregrado y, a la vez, enseñar procesamiento digital de señales (DSP) aplicado a biosignales. Se capacitó a estudiantes de primer año para programar una interfaz que permitió interactuar de forma remota con un osciloscopio y un generador de funciones. Se realizaron actividades de control, adquisición y análisis de señales, además de sesiones sobre análisis en el dominio del tiempo y de la frecuencia. Como producto, la GUI automatizó tareas de adquisición y procesamiento de datos. El estudiantado comprendió conceptos como el emparejamiento de impedancias y su efecto en las mediciones, y aplicó técnicas de DSP a señales biomédicas. Se concluyó que la experiencia integró teoría y práctica y acercó el uso de software a procesos de automatización.

Corica y Ferrari (2022) analizaron cómo se enseñó correlación y regresión en la formación de profesores de matemática en Argentina, desde la Teoría Antropológica de lo Didáctico. El estudio fue cualitativo, tipo caso, y empleó revisión documental de medios didácticos usados por cuatro docentes de Institutos de Formación Docente en la provincia de Buenos Aires. Se revisaron textos, materiales sugeridos y tareas propuestas, observando su relación con el aprendizaje de correlación y regresión. Los resultados mostraron tareas puntuales, rígidas y aisladas, con poco espacio para integrar conocimientos o interpretar resultados. También se hallaron ejercicios con datos ficticios o descontextualizados, con baja orientación hacia

discusión y reflexión. Se concluyó que se requirió mejorar la metodología, incorporando datos reales, debate e interpretación, y conectando estos temas con otras áreas de la matemática.

Ogbonnaya (2022) exploró el uso de MATLAB para enseñar estadística y comparó su utilidad frente a programas como R y SPSS, considerando especialmente a estudiantes con formación en ingeniería. Se diseñó un curso introductorio de análisis de datos en MATLAB para participantes con poca experiencia en el software y conocimientos básicos de estadística. Se trabajó con carga de datos, objetos y tipos de datos, estadística descriptiva y regresión. También se usaron herramientas interactivas, como MATLAB Grader, para evaluar respuestas y entregar retroalimentación inmediata. Los resultados indicaron mejor comprensión del análisis estadístico en el grupo, en especial por el uso de funciones interactivas y opciones para compartir código en línea. Parte de los participantes indicó que se necesitó más tiempo para práctica de análisis avanzados y menos para sintaxis básica. Se concluyó que MATLAB funcionó bien para enseñar estadística en perfiles afines a ingeniería y se sugirió avanzar hacia cursos con mayor nivel, incluyendo temas como aprendizaje automático.

Telles (2022) examinó desafíos y reflexiones sobre la enseñanza de la estadística en la educación superior dentro de la sociedad del dato, mediante una revisión documental cualitativa. Se analizaron propuestas pedagógicas e investigaciones, comparando enfoques tradicionales con alternativas orientadas al contexto. El estudio indicó que la enseñanza centrada en algoritmos y procedimientos no promovió con fuerza el razonamiento estadístico ni el uso crítico de datos en situaciones reales. También se planteó la necesidad de enfoques interdisciplinarios basados en proyectos, donde el estudiantado comprendiera el contexto de los datos y aplicara técnicas de manera situada. Se concluyó que la alfabetización estadística tomó mayor peso en la actualidad y que la enseñanza universitaria necesitó mayor práctica, contextualización e integración con procesos de investigación.

Pinargote et al. (2021) diseñaron un modelo de regresión lineal múltiple para predecir el rendimiento académico de estudiantes de Institutos Superiores Tecnológicos Públicos en Ecuador durante la nueva normalidad, considerando factores sociodemográficos, socioeconó-

micos y motivacionales. La metodología fue cuantitativa, con diseño descriptivo, inferencial y correlacional. Se recolectó información de 280 estudiantes mediante un cuestionario basado en el CEAM, incluyendo variables como tiempo de estudio, motivación, esfuerzo y contexto socioeconómico. El análisis por regresión múltiple mostró que el tiempo de estudio tuvo el mayor peso en la predicción, seguido por motivación interna y necesidad de reconocimiento. Además, 59.31 % reportó que la pandemia dificultó la continuidad educativa y 75.86 % señaló conocer a alguien que abandonó los estudios por esa situación. Se concluyó que el desempeño se asoció principalmente con dedicación al estudio y aspectos motivacionales, y se planteó fortalecer apoyo académico y acceso a herramientas educativas.

1.1.2 A nivel nacional

Salcedo (2024) evaluó el uso de TIC en el proceso de enseñanza-aprendizaje de estadística descriptiva en estudiantes del segundo ciclo del IESTP “Andrés Avelino Cáceres Dorregaray” (Huancayo) durante 2021. El estudio fue cuantitativo y pre-experimental con 30 estudiantes. Se aplicaron pruebas de entrada y salida y se analizaron resultados con la prueba t para muestras relacionadas. Se reportó $p < 0.05$, mostrando diferencias entre pretest y posttest luego del trabajo con TIC. Se concluyó que el desempeño del grupo aumentó tras la intervención y se recomendó continuar integrando recursos tecnológicos en la enseñanza.

García (2022) diseñó una propuesta metodológica para mejorar el aprendizaje de regresión lineal simple en estudiantes de segundo ciclo de Marketing en un instituto superior privado de Lima. El estudio fue cualitativo, aplicado y no experimental, con tres docentes y 15 estudiantes seleccionados por conveniencia. Se usaron entrevistas semiestructuradas, observación de clases, encuestas y pruebas pedagógicas; además, la propuesta se sometió a criterio de expertos. Los resultados mostraron bajo dominio del tema y baja motivación en el grupo estudiantil. En el caso docente, se observó trabajo poco dinámico, con poca colaboración y participación, y un tratamiento más mecánico que conceptual, con escasa conexión a situaciones reales. Se concluyó que la propuesta se orientó a métodos activos (aprendizaje basado en problemas y proyectos) y uso de TIC para aumentar participación

y comprensión.

Montoya (2022) analizó la relación entre regresión lineal y precisión del pronóstico de demanda de productos hidrobiológicos congelados en el Perú. Se trabajó con series temporales de 1980 a 2018 y se consideró la demanda como variable dependiente, junto con indicadores macroeconómicos como variables independientes. Se separaron datos en dos grupos: 1980–2013 para construir y validar modelos, y 2014–2018 para evaluar el pronóstico mediante el Error Medio Absoluto Porcentual (EMAP), usando Gretl. Los resultados mostraron que la precisión no dependió directamente de la longitud de la serie, aunque rangos de 13 a 16 años dieron mejores pronósticos. También se halló que, al incluir más variables independientes, la predicción mejoró, y que los modelos validados superaron a los no validados. Se concluyó que la regresión lineal aportó a mejorar estimaciones para toma de decisiones en el sector.

Ramos (2019) revisó retos y oportunidades de la educación estadística universitaria, considerando actitudes estudiantiles y metodologías de enseñanza. El estudio se basó en revisión bibliográfica y documental. Se identificaron actitudes desfavorables hacia la estadística, asociadas a enfoques tradicionales centrados en exposición teórica y a la poca conexión con situaciones reales. También se describieron alternativas como aprendizaje basado en proyectos y uso de herramientas tecnológicas para aumentar comprensión y motivación. Se concluyó que existieron dificultades por actitudes y enfoques tradicionales, y se propuso adoptar estrategias que promuevan razonamiento y pensamiento estadístico.

1.2 Base Teórica

1.2.1 Regresión lineal simple

Montgomery et al. (2021) mencionó:

Definición 1 (Modelo de regresión lineal simple). *Se observa un conjunto de n pares de datos: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. La idea de la **regresión lineal simple** es aproximar la relación entre X e Y mediante una recta dada por:*

$$\hat{y} = b_0 + b_1x$$

donde:

- b_0 es el **intercepto**, es decir, el valor estimado de Y cuando $x = 0$.
- b_1 es la **pendiente**, es decir, el cambio estimado en Y por cada unidad que aumenta X .
- $\hat{y}$ es el **valor estimado** (ajustado) de Y para un valor dado de x .

1.2.1.1 Medias

Definición 2 (Media de x). $\bar{x}$ es el promedio de los valores de x :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Definición 3 (Media de y). $\bar{y}$ es el promedio de los valores de y :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

Observación 1. Estas medias son el centro de los datos y sirven para medir desviaciones y productos cruzados.

1.2.1.2 Suma de cuadrados y producto cruzado (misma notación)

Montgomery et al. (2021) mencionó:

Definición 4 (Suma de cuadrados: S_{xx}). S_{xx} mide cuánta variación hay en X alrededor de su media:

$$S_{xx} = \sum (x_i - \bar{x})^2$$

Definición 5 (Producto cruzado: S_{xy}). S_{xy} mide cómo varían conjuntamente X e Y alrededor de sus medias:

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

Observación 2. Interpretación práctica en una tabla de cálculo:

- Las columnas $(x - \bar{x})$ y $(y - \bar{y})$ se multiplican fila por fila para formar $(x - \bar{x})(y - \bar{y})$.
- Luego se suman todas esas filas y se obtiene S_{xy} .
- Lo mismo con $(x - \bar{x})^2$ para obtener S_{xx} .

1.2.1.3 SCE (Suma de Cuadrados del Error)

En esta notación:

$$SCE = \sum (y_i - \hat{y}_i)^2$$

Definición 6 (Residuo y error).

$$\hat{y}_i = b_0 + b_1 x_i \quad (\text{valor estimado})$$

$$e_i = y_i - \hat{y}_i \quad (\text{residuo})$$

Entonces:

$$SCE = \sum e_i^2$$

Observación 3. Objetivo de mínimos cuadrados: elegir b_0 y b_1 para que SCE sea lo más pequeña posible.

1.2.1.4 Teorema principal: fórmula de la pendiente y del intercepto

James et al. (2021) mencionó:

Teorema 1 (Intercepto (fórmula de b_0)). Sea

$$SCE(b_0, b_1) = \sum_{i=1}^n (y_i - (b_0 + b_1 x_i))^2, \quad e_i = y_i - (b_0 + b_1 x_i).$$

Si (b_0, b_1) minimiza $SCE(b_0, b_1)$, entonces

$$b_0 = \bar{y} - b_1 \bar{x}.$$

Demostración. Minimizar SCE requiere que sus derivadas parciales se anulen.

Paso 1: derivar respecto a b_0 e igualar a cero.

$$\frac{\partial SCE}{\partial b_0} = \sum_{i=1}^n 2(y_i - b_0 - b_1 x_i)(-1) = 0 \implies \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0.$$

Como $e_i = y_i - (b_0 + b_1 x_i)$, se obtiene el lema:

Lema 1.

$$\sum_{i=1}^n e_i = 0.$$

Ahora se reordena:

$$\sum_{i=1}^n y_i = nb_0 + b_1 \sum_{i=1}^n x_i.$$

Dividiendo entre n :

$$\bar{y} = b_0 + b_1 \bar{x}.$$

Finalmente,

$$b_0 = \bar{y} - b_1 \bar{x}.$$

□

Teorema 2 (Pendiente (fórmula de b_1)). *Sea*

$$SCE(b_0, b_1) = \sum_{i=1}^n (y_i - (b_0 + b_1 x_i))^2, \quad e_i = y_i - (b_0 + b_1 x_i).$$

Supóngase además que

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 > 0.$$

Si (b_0, b_1) minimiza $SCE(b_0, b_1)$, entonces

$$b_1 = \frac{S_{xy}}{S_{xx}}, \quad S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

Demostración. **Paso 1: derivar respecto a b_1 e igualar a cero.**

$$\frac{\partial SCE}{\partial b_1} = \sum_{i=1}^n 2(y_i - b_0 - b_1 x_i)(-x_i) = 0 \implies \sum_{i=1}^n x_i(y_i - b_0 - b_1 x_i) = 0.$$

Como $e_i = y_i - (b_0 + b_1 x_i)$, se obtiene el lema:

Lema 2.

$$\sum_{i=1}^n x_i e_i = 0.$$

Paso 2: sustituir $b_0 = \bar{y} - b_1 \bar{x}$. Del teorema anterior:

$$b_0 = \bar{y} - b_1 \bar{x}.$$

Sustituyendo en $\sum x_i(y_i - b_0 - b_1 x_i) = 0$:

$$\sum_{i=1}^n x_i(y_i - (\bar{y} - b_1 \bar{x}) - b_1 x_i) = 0.$$

Simplificando el paréntesis:

$$y_i - \bar{y} + b_1 \bar{x} - b_1 x_i = (y_i - \bar{y}) - b_1(x_i - \bar{x}),$$

por tanto:

$$\sum_{i=1}^n x_i [(y_i - \bar{y}) - b_1(x_i - \bar{x})] = 0.$$

Paso 3: pasar a la forma centrada. Se usa $x_i = (x_i - \bar{x}) + \bar{x}$:

$$\sum_{i=1}^n [(x_i - \bar{x}) + \bar{x}] [(y_i - \bar{y}) - b_1(x_i - \bar{x})] = 0.$$

Al expandir y usar

$$\sum_{i=1}^n (y_i - \bar{y}) = 0 \quad \text{y} \quad \sum_{i=1}^n (x_i - \bar{x}) = 0,$$

se cancelan los términos con $\bar{x}$ y queda:

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) - b_1 \sum_{i=1}^n (x_i - \bar{x})^2 = 0.$$

Luego:

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = b_1 \sum_{i=1}^n (x_i - \bar{x})^2.$$

Reconociendo definiciones:

$$S_{xy} = b_1 S_{xx}.$$

Como $S_{xx} > 0$, se concluye:

$$b_1 = \frac{S_{xy}}{S_{xx}}.$$

□

Observación 4. *Estas son exactamente las fórmulas mostradas en la simbología indicada.*

Con b_0 y b_1 queda determinado el modelo:

$$\hat{y} = b_0 + b_1 x.$$

1.2.1.5 Sumas de cuadrados: STC , SCR , SCE

James et al. (2021) mencionó:

Definición 7 (Suma Total de Cuadrados).

$$STC = \sum (y_i - \bar{y})^2$$

Mide la variación total de Y respecto a su media.

Definición 8 (Suma de Cuadrados de la Regresión).

$$SCR = \sum (\hat{y}_i - \bar{y})^2$$

Mide cuánta variación de Y explica la recta ajustada.

Definición 9 (Suma de Cuadrados del Error).

$$SCE = \sum (y_i - \hat{y}_i)^2$$

Mide la variación que queda sin explicar (errores o residuos).

1.2.1.6 Teorema de descomposición: $STC = SCR + SCE$

James et al. (2021) mencionó:

Teorema 3 (Descomposición de la variación). Sea $\{(x_i, y_i)\}_{i=1}^n$ con $n \geq 2$ y sea la recta de regresión ajustada por mínimos cuadrados

$$\hat{y}_i = b_0 + b_1 x_i, \quad e_i = y_i - \hat{y}_i.$$

Defínanse

$$STC = \sum_{i=1}^n (y_i - \bar{y})^2, \quad SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2.$$

Entonces se cumple la identidad

$$STC = SCR + SCE.$$

Demostración. **Paso 1: descomposición puntual.** Para cada i se tiene

$$y_i - \bar{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})$$

Paso 2: elevar al cuadrado.

$$(y_i - \bar{y})^2 = (y_i - \hat{y}_i)^2 + (\hat{y}_i - \bar{y})^2 + 2(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}).$$

Paso 3: sumar para $i = 1, \dots, n$.

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}).$$

Reconociendo definiciones:

$$STC = SCE + SCR + 2 \sum_{i=1}^n e_i(\hat{y}_i - \bar{y}).$$

Paso 4: anular el término cruzado (ortogonalidad). Como $\hat{y}_i = b_0 + b_1 x_i$ y, por la propiedad de mínimos cuadrados con intercepto,

$$\sum_{i=1}^n e_i = 0 \quad \text{y} \quad \sum_{i=1}^n x_i e_i = 0,$$

se obtiene:

$$\sum_{i=1}^n e_i(\hat{y}_i - \bar{y}) = \sum_{i=1}^n e_i(b_0 + b_1 x_i - \bar{y}) = (b_0 - \bar{y}) \sum_{i=1}^n e_i + b_1 \sum_{i=1}^n x_i e_i = 0.$$

Por tanto, el término cruzado es cero:

$$\sum_{i=1}^n e_i(\hat{y}_i - \bar{y}) = 0.$$

Paso 5: conclusión. Sustituyendo en la igualdad del Paso 3:

$$STC = SCE + SCR.$$

□

1.2.1.7 Coeficiente de determinación

James et al. (2021) mencionó:

Definición 10 (Coeficiente de determinación).

$$R^2 = \frac{SCR}{STC}$$

Interpretación:

- Si R^2 está cerca de 1, la recta explica gran parte de la variación de Y .
- Si R^2 está cerca de 0, la recta explica muy poco de la variación de Y .

Corolario 1 (Forma equivalente de R^2). Como $STC = SCR + SCE$, entonces:

$$R^2 = 1 - \frac{SCE}{STC}$$

Esto muestra que R^2 aumenta cuando SCE disminuye, es decir, cuando el ajuste mejora.

1.2.1.8 Error cuadrático medio y su raíz

Definición 11 (Error cuadrático medio).

$$MSE = \frac{SCE}{n - 2}$$

Observación 5. ¿Por qué $n - 2$? Porque al ajustar la recta se estiman dos cantidades (b_0 y b_1), y eso reduce los grados de libertad disponibles para medir el error.

Definición 12 (Raíz del MSE).

$$RMSE = \sqrt{MSE}$$

Interpretación:

- $RMSE$ es una medida del tamaño típico del error en las mismas unidades de Y .
- Mientras menor sea el $RMSE$, mejor el ajuste (en promedio).

1.2.1.9 Supuestos del modelo de regresión lineal

Para garantizar que los resultados obtenidos mediante la interfaz gráfica de usuario en MATLAB sean estadísticamente válidos, se consideraron los principales supuestos del modelo de regresión lineal simple y múltiple. Estos supuestos permiten verificar que el modelo ajustado represente adecuadamente la relación entre las variables y que las conclusiones derivadas del análisis no dependan únicamente del cálculo automático realizado por la herramienta.

1. Linealidad

El supuesto de linealidad establece que la relación entre la variable dependiente y las variables independientes debe aproximarse mediante una función lineal. En la regresión lineal simple, este supuesto puede evaluarse mediante el diagrama de dispersión entre X e Y , observando si los puntos presentaban una tendencia aproximadamente recta. En la regresión lineal múltiple, se revisaron gráficos parciales y la relación entre los valores observados y los valores estimados.

Matemáticamente, el modelo lineal simple se expresa como:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

y el modelo lineal múltiple como:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i$$

donde Y_i representa la variable dependiente, X_i o X_{ki} representan las variables independientes, β_0 es el intercepto, β_j son los coeficientes del modelo y ε_i representa el error aleatorio.

2. Independencia de los errores

El supuesto de independencia indica que los errores o residuos del modelo no deben estar correlacionados entre sí. En esta investigación, este supuesto se comprobó mediante el análisis gráfico de los residuos y, cuando correspondió, mediante el estadístico de Durbin-Watson. Un valor cercano a 2 indicó ausencia de autocorrelación considerable entre los residuos.

El residuo se definió como:

$$e_i = y_i - \hat{y}_i$$

donde y_i es el valor observado y $\hat{y}_i$ es el valor estimado por el modelo.

3. Homoscedasticidad

La homoscedasticidad señala que la variabilidad de los residuos debe mantenerse aproximadamente constante para todos los valores estimados. Este supuesto se verificó mediante el gráfico de residuos frente a valores ajustados. Cuando los residuos se distribuyeron de manera aleatoria alrededor de cero, sin formar patrones de abanico o curvaturas marcadas, se consideró que el supuesto se cumplía.

También se pudo contrastar este supuesto mediante pruebas estadísticas como Breusch-Pagan, especialmente cuando se trabajó con bases de datos de mayor tamaño.

4. Normalidad de los residuos

El supuesto de normalidad establece que los errores del modelo deben aproximarse a una distribución normal. Este supuesto fue revisado mediante histogramas de residuos, gráficos Q-Q y la prueba de Shapiro-Wilk, según el tamaño de la muestra. Si el valor de significancia fue mayor que 0,05, se consideró que no existía evidencia suficiente para rechazar la normalidad de los residuos.

$$H_0 : \text{Los residuos siguen una distribución normal}$$

H_1 : Los residuos no siguen una distribución normal

5. Ausencia de multicolinealidad

En el caso de la regresión lineal múltiple, se consideró el supuesto de ausencia de multicolinealidad. Este supuesto indica que las variables independientes no deben presentar relaciones lineales demasiado fuertes entre ellas, ya que esto puede afectar la estabilidad de los coeficientes estimados.

La multicolinealidad se evaluó mediante el Factor de Inflación de la Varianza, conocido como VIF:

$$VIF_j = \frac{1}{1 - R_j^2}$$

donde R_j^2 corresponde al coeficiente de determinación obtenido al explicar una variable independiente mediante las demás variables independientes del modelo. Se consideró aceptable un valor de $VIF < 10$, aunque valores menores a 5 indicaron una mejor estabilidad del modelo.

6. Comprobación de los supuestos en la interfaz gráfica

La interfaz gráfica desarrollada en MATLAB permitió calcular los valores ajustados, residuos, errores y métricas de ajuste del modelo. Asimismo, facilitó la visualización de gráficos de dispersión, recta de regresión, residuos y comparación entre valores observados y estimados. Estas salidas permitieron comprobar de manera práctica los supuestos de linealidad, homoscedasticidad, normalidad e independencia de los errores, reforzando el uso estadístico correcto de la regresión lineal dentro del proceso de aprendizaje.

1.2.2 Regresión lineal múltiple

Definición 13 (Modelo general). Sea n el número de observaciones y k el número de variables explicativas. Para cada observación $i = 1, \dots, n$ se dispone de:

$$(x_{i1}, x_{i2}, \dots, x_{ik}, y_i).$$

El modelo de regresión lineal múltiple es:

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_kx_{ik} + e_i, \quad i = 1, \dots, n.$$

- b_0 es el intercepto.
- b_j ($j = 1, \dots, k$) representa el efecto parcial de X_j sobre Y manteniendo constantes los demás predictores.
- e_i es el error asociado a la observación i .

1.2.3 Notación de sumatorias (para construir la matriz)

Definición 14 (Sumatorias generales). Se definen:

- $\sum y = \sum_{i=1}^n y_i$.
- Para cada predictor X_j :

$$\sum x_j = \sum_{i=1}^n x_{ij}, \quad \sum x_j y = \sum_{i=1}^n x_{ij} y_i.$$

- Para pares de predictores X_r y X_s :

$$\sum x_r x_s = \sum_{i=1}^n x_{ir} x_{is}, \quad r, s = 1, \dots, k.$$

En particular, $\sum x_j^2 = \sum x_j x_j$.

1.2.4 Forma matricial: sistema lineal para estimar $b_0, b_1, \dots, b_k$

Definición 15 (Sistema de ecuaciones normales (general)). *La estimación por mínimos cuadrados conduce al sistema:*

$$\mathbf{A} \mathbf{b} = \mathbf{d},$$

donde:

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} \sum y \\ \sum x_1 y \\ \vdots \\ \sum x_k y \end{pmatrix},$$

y la matriz $\mathbf{A}$ (de tamaño $(k+1) \times (k+1)$) es:

$$\mathbf{A} = \begin{pmatrix} n & \sum x_1 & \sum x_2 & \cdots & \sum x_k \\ \sum x_1 & \sum x_1 x_1 & \sum x_1 x_2 & \cdots & \sum x_1 x_k \\ \sum x_2 & \sum x_2 x_1 & \sum x_2 x_2 & \cdots & \sum x_2 x_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_k & \sum x_k x_1 & \sum x_k x_2 & \cdots & \sum x_k x_k \end{pmatrix}.$$

- La primera fila/columna incorpora el intercepto (por eso aparece n y las sumas $\sum x_j$).
- Las entradas $\sum x_r x_s$ son productos cruzados entre predictores (incluye cuadrados cuando $r = s$).
- El vector $\mathbf{d}$ reúne las sumas cruzadas predictor-respuesta $\sum x_j y$ y la suma $\sum y$.

Observación 6. Si $\det(\mathbf{A}) \neq 0$ (equivalentemente, $\mathbf{A}$ es invertible), el sistema tiene solución única:

$$\hat{\mathbf{b}} = \mathbf{A}^{-1} \mathbf{d} \implies (\hat{b}_0, \hat{b}_1, \dots, \hat{b}_k).$$

1.2.5 Variable 1: Desarrollo de una Interfaz Gráfica de Usuario en MATLAB

El desarrollo de una interfaz gráfica de usuario requiere considerar las necesidades, tareas y condiciones de uso de quienes interactuarán con la herramienta. Por ello, el diseño debe organizarse de forma clara y pasar por procesos de revisión que permitan mejorar su funcionamiento.

La calidad de una herramienta computacional depende de procesos de verificación y prueba. Estos procedimientos ayudan a comprobar que los cálculos realizados por el sistema sean correctos y que los resultados entregados sean confiables (ISTQB, 2024; Winters et al., 2020).

MATLAB, mediante App Designer y sus herramientas de compilación, permite crear aplicaciones interactivas con tablas, controles, gráficos y salidas visuales. Además, facilita la generación de ejecutables, lo que permite usar la aplicación en equipos que no tienen instalado MATLAB completo (MathWorks, 2024).

1. Dimensión: Diseño y estructura de la interfaz

El diseño de la interfaz debe responder a las tareas que realiza el usuario. Una pantalla organizada ayuda a reducir confusiones y permite que el estudiante ubique con mayor facilidad las funciones principales.

Los principios de interacción señalan que la comunicación entre el usuario y el sistema debe ser clara, ordenada y consistente. También se debe permitir que el usuario mantenga control sobre las acciones que realiza dentro de la aplicación (ISO, 2020).

La accesibilidad busca que la interfaz pueda ser comprendida y utilizada por distintos usuarios. Para ello, se requiere una navegación sencilla, opciones visibles y elementos que no generen barreras durante el uso.

■ Indicadores:

- **Organización intuitiva:** La organización de la interfaz se relaciona con la forma en que el usuario logra cumplir sus tareas dentro de un contex-

to determinado, considerando la facilidad de uso, el tiempo empleado y la satisfacción obtenida (ISO, 2018).

- **Claridad en la navegación:** La navegación clara permite que el usuario comprenda las opciones disponibles, avance sin confusión y reconozca el efecto de cada acción realizada.
- **Accesibilidad de funciones principales:** La accesibilidad de funciones se refiere a la disposición de herramientas visibles y comprensibles, de modo que el usuario pueda operar la interfaz con menos dificultad.

2. Dimensión: Funcionalidades técnicas

La enseñanza de la regresión lineal requiere que el estudiante observe cómo se estiman los parámetros del modelo y cómo se interpretan los resultados obtenidos. Por ello, la herramienta debe permitir el cálculo correcto de coeficientes, valores estimados, residuos y medidas de ajuste (Montgomery et al., 2021).

El estudio actual de la regresión lineal se apoya en el uso de datos, el ajuste de modelos y la interpretación de resultados mediante recursos computacionales. Estas herramientas ayudan a trabajar con situaciones reales y a representar mejor la relación entre variables (Fahrmeir et al., 2021; James et al., 2021).

App Designer permite integrar botones, cuadros de entrada, tablas y gráficos dentro de una misma aplicación. Esta característica facilita la presentación de ecuaciones, cálculos y representaciones visuales del modelo de regresión (MathWorks, 2024a).

■ Indicadores:

- **Precisión en modelos:** La precisión del modelo se relaciona con la correcta estimación de los parámetros y con la coherencia entre los resultados calculados y los datos ingresados (Montgomery et al., 2021).
- **Visualización de ecuaciones:** La visualización de la ecuación permite que el estudiante identifique la relación estimada entre las variables y comprenda

el significado de los coeficientes del modelo (Fahrmeir et al., 2021).

- **Gráficos de ajuste:** Los gráficos de ajuste permiten observar la cercanía entre los datos reales y los valores estimados, lo que ayuda a interpretar el comportamiento del modelo (James et al., 2021).

3. Dimensión: Interactividad y retroalimentación

La interactividad permite que el usuario modifique datos, ejecute procedimientos y observe los cambios generados por sus acciones. En una interfaz educativa, esta característica ayuda a que el estudiante participe de manera activa durante el proceso de aprendizaje.

Los principios de interacción plantean que el sistema debe responder de forma clara a las acciones del usuario. De esta manera, quien utiliza la herramienta puede reconocer si los datos fueron procesados correctamente y qué resultado se obtuvo.

App Designer permite programar acciones mediante *callbacks*, actualizar tablas y modificar gráficos cuando cambian los datos ingresados. Esta posibilidad favorece una retroalimentación inmediata durante el uso de la interfaz (MathWorks, 2024a).

■ Indicadores:

- **Resultados en tiempo real:** Los resultados en tiempo real permiten que el usuario observe de inmediato las salidas generadas por el sistema después de ingresar o modificar datos.
- **Respuesta a modificaciones:** La respuesta a modificaciones se refiere a la capacidad de la interfaz para actualizar cálculos, tablas y gráficos cuando el usuario cambia la información de entrada.
- **Representación visual de cálculos:** La representación visual de cálculos permite mostrar los procedimientos y resultados mediante tablas, gráficos y valores numéricos actualizados dentro de la aplicación (MathWorks, 2024a).

4. Dimensión: Validación funcional

La validación funcional permite comprobar si la interfaz cumple con las tareas para las que fue diseñada. Para ello, se realizan pruebas con datos conocidos y se comparan los resultados generados por el sistema con cálculos manuales o procedimientos ya aceptados.

Las pruebas de software ayudan a identificar fallas, revisar el comportamiento de la aplicación y confirmar que los resultados obtenidos coincidan con lo esperado (ISTQB, 2024).

En herramientas que realizan cálculos numéricos, la revisión funcional permite reducir errores de programación y aumentar la confianza en los resultados mostrados al usuario (Winters et al., 2020).

■ **Indicadores:**

- **Pruebas con datos:** Las pruebas con datos consisten en ingresar información previamente conocida para comparar los resultados producidos por la interfaz con resultados esperados (ISTQB, 2024).
- **Precisión de cálculos:** La precisión de cálculos se observa cuando los coeficientes, predicciones, residuos y medidas de ajuste obtenidos por la interfaz coinciden con los procedimientos de referencia (Winters et al., 2020).
- **Funcionamiento de la interfaz:** El funcionamiento de la interfaz se relaciona con la capacidad del usuario para utilizar sus opciones, ejecutar los procesos y obtener resultados sin interrupciones durante la actividad (ISO, 2018).

5. Dimensión: Distribución ejecutable

MATLAB permite compilar aplicaciones para convertirlas en programas ejecutables. Esta opción facilita que la interfaz pueda instalarse o ejecutarse en otros equipos sin necesidad de abrir el entorno completo de MATLAB (MathWorks, 2024b).

Las aplicaciones compiladas requieren un entorno de ejecución o *runtime*, el cual

contiene los componentes necesarios para que el programa funcione fuera de MATLAB (MathWorks, 2024c).

La distribución mediante ejecutable permite que la herramienta sea usada en espacios educativos donde no todos los equipos cuentan con MATLAB instalado o con licencia activa.

■ **Indicadores:**

- **Ejecutable independiente:** El ejecutable independiente corresponde a una aplicación compilada que puede abrirse como programa propio después del proceso de empaquetado (MathWorks, 2024b).
- **Funcionalidades integradas:** Las funcionalidades integradas son los módulos, botones, tablas, gráficos y procesos de cálculo incluidos dentro de la aplicación compilada.
- **Uso sin MATLAB instalado:** El uso sin MATLAB instalado se logra mediante el *runtime*, que permite ejecutar la aplicación en equipos que no tienen la versión completa del software (MathWorks, 2024c).

1.2.6 Variable 2: Aprendizaje de Regresión Lineal

La regresión lineal se enseña como modelado de relaciones entre variables y requiere interpretación adecuada de parámetros y ajuste (Montgomery et al., 2021).

La enseñanza moderna de regresión integra procedimientos, interpretación y apoyo computacional para trabajar con datos reales y comprender el modelo (Fahrmeir et al., 2021; James et al., 2021).

El uso de herramientas interactivas puede favorecer el aprendizaje al permitir experimentar con datos y observar cambios en resultados y gráficos (MathWorks, 2024a).

1. Dimensión: Comprensión conceptual

Los textos de regresión enfatizan comprender el significado de intercepto, pendiente y supuestos, más allá del cálculo mecánico (Montgomery et al., 2021).

La regresión aplicada propone interpretar el modelo en contexto, revisar el ajuste y comunicar conclusiones a partir de datos (Fahrmeir et al., 2021).

El aprendizaje mejora cuando el estudiante conecta fórmulas con representación gráfica y lectura de resultados en software, consolidando conceptos (James et al., 2021).

- **Indicadores**

- **Claridad de regresión simple/múltiple:** Comprender regresión implica entender cómo el modelo representa una relación estadística entre variables y qué significan sus parámetros (Montgomery et al., 2021).
- **Formulación de ecuaciones:** La comprensión conceptual se refleja en explicar resultados del ajuste y su sentido en el problema real (Fahrmeir et al., 2021).
- **Análisis de resultados:** La comprensión incluye interpretar tablas, métricas y gráficos de ajuste para justificar conclusiones (James et al., 2021).

2. Dimensión: Habilidad procedimental

Resolver regresión requiere dominar procedimientos: cálculo/estimación, sustitución de datos y obtención de métricas (Montgomery et al., 2021).

La práctica con datasets y software desarrolla destrezas para ejecutar pasos de análisis y verificar resultados (James et al., 2021).

La fluidez procedimental se asocia con aplicar procedimientos con precisión y flexibilidad, evitando errores en pasos (NCTM, 2023).

- **Indicadores**

- **Guía en resolución:** Habilidad procedimental es ejecutar correctamente los pasos necesarios para estimar y evaluar un modelo (Montgomery et al., 2021).

- **ejemplos interactivos:** Se entiende como aplicar procedimientos con precisión, eficiencia y flexibilidad, adaptándolos a diferentes ejercicios (NCTM, 2023).
- **simulaciones:** La guía interactiva se considera pertinente cuando ayuda a realizar pasos y verificar resultados con apoyo de cálculos computacionales (James et al., 2021).

3. Dimensión: Validación de resultados

La regresión exige validar resultados mediante diagnóstico básico y revisión del ajuste para evitar interpretaciones erróneas (Montgomery et al., 2021).

La comparación con software o cálculo alternativo permite verificar consistencia de parámetros, predicciones y métricas (Fahrmeir et al., 2021).

La validación se fortalece cuando el estudiante interpreta residuales, ajuste y supuestos con apoyo de herramientas computacionales (James et al., 2021).

■ Indicadores

- **Precisión:** Validar resultados implica comprobar que el modelo y sus estimaciones sean coherentes con los datos y supuestos (Montgomery et al., 2021).
- **Método manual/software:** La validación incluye contrastar resultados con métodos alternativos o herramientas para detectar discrepancias (Fahrmeir et al., 2021).
- **Consistencia de cálculos:** Comparar y visualizar resultados favorece detectar errores de procedimiento e interpretación (James et al., 2021).

4. Dimensión: Simplicidad de uso

La simplicidad se relaciona con usabilidad: lograr objetivos con eficiencia y satisfacción, sin barreras innecesarias (ISO, 2018).

El diseño centrado en el usuario prioriza que la herramienta se adapte a estudiantes y docentes, con funciones claras y accesibles.

La accesibilidad recomienda criterios para reducir dificultades de uso y facilitar interacción para más usuarios.

- **Indicadores**

- **Facilidad de ingreso de datos:** La facilidad de uso se evidencia cuando el usuario completa tareas con efectividad y eficiencia, y se siente satisfecho (ISO, 2018).
- **Visualización:** Simplicidad significa diseño alineado a necesidades del usuario y probado con usuarios para reducir confusión.
- **Baja dificultad técnica:** Reducir dificultad técnica implica aplicar recomendaciones que hagan la interacción más comprensible y operable.

5. Dimensión: Evaluación funcional de ejemplos

Los ejemplos apoyan el aprendizaje porque permiten ver el procedimiento aplicado en casos concretos y luego transferirlo a nuevos ejercicios (James et al., 2021).

Una evaluación funcional exige que los ejemplos generen resultados correctos y coherentes con el modelo implementado (Montgomery et al., 2021).

La herramienta debe permitir ejecutar ejemplos de forma controlada y verificable, aplicando criterios básicos de prueba/validación.

- **Indicadores**

- **Número/calidad de ejemplos:** Un ejemplo funcional se define como un caso con datos, modelo y resultados interpretables que reflejan el método enseñado (James et al., 2021).
- **Variedad de aplicaciones:** La calidad del ejemplo se observa cuando parámetros y métricas corresponden a un ajuste correcto y explicable (Montgomery et al., 2021).

- **Coherencia de resultados:** La evaluación funcional exige casos de prueba que confirmen resultados esperados y detecten fallos (ISTQB, 2024).

1.3 Bases Conceptuales

1.3.1 Regresión Lineal

Es un método estadístico que modela la relación entre una variable dependiente y una o más variables independientes mediante una ecuación de tipo lineal. Se utiliza en diversas áreas para hacer predicciones y análisis de tendencias basados en datos observados.

1.3.2 Regresión Lineal Simple

La regresión lineal simple es un modelo estadístico que describe la relación entre una variable dependiente y una única variable independiente mediante una ecuación de tipo lineal. Su objetivo es predecir o explicar cómo cambia la variable dependiente en función de la variable independiente.

1.3.3 Regresión Lineal Múltiple

La regresión lineal múltiple es un modelo estadístico que analiza la relación entre una variable dependiente y dos o más variables independientes. Se utiliza para predecir o explicar la influencia conjunta de múltiples factores sobre la variable dependiente.

1.3.4 Educación Universitaria

La educación universitaria es el nivel de enseñanza superior que se imparte en universidades e institutos de educación superior, donde se forman profesionales en distintas áreas del conocimiento. Se caracteriza por un enfoque teórico-práctico que permite el desarrollo de competencias académicas, investigativas y profesionales.

1.3.5 Interfaz Gráfica de Usuario (GUI)

Una interfaz gráfica de usuario (GUI) es una herramienta que permite a los usuarios interactuar con un programa de manera visual y amigable, en lugar de utilizar comandos textuales. En el contexto de MATLAB, una GUI permite al usuario ejecutar cálculos estadísticos y visualizar resultados sin necesidad de escribir códigos directamente, facilitando el proceso de aprendizaje al ofrecer una plataforma interactiva (MathWorks, 2024a).

1.3.6 MATLAB:

MATLAB es un entorno de programación y un lenguaje de cálculo numérico que facilita el análisis de datos, el desarrollo de algoritmos y la creación de modelos. Es ampliamente utilizado en la educación e investigación, especialmente en áreas como matemáticas, ingeniería y ciencias aplicadas. Su capacidad para crear GUIs lo convierte en una herramienta poderosa para desarrollar aplicaciones educativas (MathWorks, 2024a).

1.3.7 Aprendizaje Activo:

El aprendizaje activo es un enfoque pedagógico que involucra a los estudiantes en actividades que fomentan la participación directa, en lugar de ser receptores pasivos de información. En el contexto de esta investigación, el uso de una GUI interactiva permite a los estudiantes manipular datos y realizar cálculos estadísticos, lo que refuerza su comprensión a través de la práctica directa.

1.3.8 Tecnologías Educativas:

Las tecnologías educativas se refieren a las herramientas tecnológicas que facilitan el proceso de enseñanza-aprendizaje. En este contexto, el uso de MATLAB y su capacidad para crear GUIs es una tecnología educativa que permite a los estudiantes interactuar con los datos y realizar análisis estadísticos de manera práctica y visual.

1.3.9 Variable independiente

Desarrollo de una Interfaz Gráfica de Usuario en MATLAB.

- **Definición conceptual:** Es el proceso de diseño y construcción de una plataforma interactiva en MATLAB que permite la visualización y manipulación de datos, con el propósito de facilitar el análisis y comprensión de la regresión lineal.
- **Definición operacional:** Se medirá en términos de sus características técnicas, funcionalidad, interfaz de usuario, herramientas de análisis de regresión lineal y facilidad

de uso. La evaluación se realizará mediante pruebas de funcionalidad y validaciones de la usabilidad del sistema.

1.3.10 Variable dependiente

Aprendizaje de Regresión Lineal

- **Definición conceptual:** Hace referencia a la adquisición y consolidación de conocimientos sobre regresión lineal, incluyendo su teoría, aplicación práctica y resolución de problemas matemáticos.
- **Definición operacional:** Se analizará a través de simulaciones, resolución de ejercicios en la interfaz y generación de reportes automáticos sobre la regresión lineal.

1.3.11 Operacionalización de variables

Tabla 1

Operacionalización de variables.

Variable	Definición Operacional	Dimensiones	Indicadores
Desarrollo de una Interfaz Gráfica de Usuario en MATLAB	Implementación de una interfaz gráfica en MATLAB que facilite el aprendizaje de regresión lineal mediante funcionalidades interactivas validadas a través de pruebas técnicas y ejemplos.	Diseño y estructura de la interfaz	Organización intuitiva de los elementos visuales, claridad en la navegación, accesibilidad de las funciones principales.
		Funcionalidades técnicas	Precisión en la representación de modelos de regresión lineal (simple y múltiple), visualización de ecuaciones y gráficos de ajuste de datos.
		Interactividad y retroalimentación	Capacidad de la interfaz para mostrar resultados en tiempo real, responder a modificaciones de datos y representar visualmente los cálculos de regresión.
		Validación funcional	Pruebas realizadas con conjuntos de datos y ejemplos para comprobar la precisión de los cálculos y la efectividad de la interfaz.
		Distribución ejecutable	Generación de un ejecutable independiente con todas las funcionalidades integradas y sin la necesidad de MATLAB instalado.
Aprendizaje de Regresión Lineal	Evaluación del impacto de la interfaz gráfica en MATLAB como herramienta para la enseñanza y comprensión de la regresión lineal en la educación universitaria.	Comprensión conceptual	Claridad en la representación de la regresión lineal simple y múltiple, formulación de ecuaciones y análisis de resultados.
		Habilidad procedimental	Capacidad de la interfaz para guiar en la resolución de problemas de regresión lineal con ejemplos interactivos y simulaciones.
		Validación de resultados	Precisión de los cálculos obtenidos con la interfaz en comparación con métodos manuales o software estadísticos.
		Simplicidad de uso	Facilidad para ingresar datos, visualizar resultados y aplicar conceptos de regresión sin dificultad técnica.
		Evaluación funcional de ejemplos	Número y calidad de ejemplos desarrollados dentro de la interfaz para ilustrar diferentes aplicaciones de la regresión lineal.

II. DISEÑO METODOLÓGICO

2.1 Tipo de Investigación

La investigación presentó un enfoque cuantitativo, debido a que se recopilaron y analizaron datos numéricos obtenidos mediante el pretest, posttest y cuestionario de usabilidad aplicado a los estudiantes del segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

Fue de tipo aplicada, porque se desarrolló una solución tecnológica concreta: una interfaz gráfica de usuario en MATLAB orientada a apoyar el aprendizaje de regresión lineal en la educación universitaria. La investigación no se limitó al diseño de la herramienta, sino que también consideró su aplicación con estudiantes para verificar su funcionamiento, usabilidad y aporte al aprendizaje.

Asimismo, presentó alcance descriptivo y explicativo. Fue descriptiva porque se detallaron las características, componentes, módulos y funciones de la interfaz gráfica. Fue explicativa porque se analizó el cambio producido en el aprendizaje de regresión lineal mediante la comparación de los resultados obtenidos antes y después de la aplicación de la GUI (Hernández & Mendoza, 2018).

2.2 Método de Investigación

Se empleó el método hipotético-deductivo, debido a que se formuló la hipótesis de que la interfaz gráfica de usuario en MATLAB potenciaría el aprendizaje de regresión lineal en la educación universitaria. A partir de esta hipótesis se recopilaron datos mediante un pretest y un posttest, los cuales permitieron contrastar los resultados obtenidos por los estudiantes antes y después de utilizar la herramienta.

También se empleó el método tecnológico aplicado, ya que se diseñó, programó, probó y validó una interfaz gráfica de usuario en MATLAB para resolver problemas de regresión lineal simple y múltiple. La interfaz permitió automatizar cálculos, representar datos mediante gráficos y facilitar la interpretación de los resultados estadísticos (Hernández & Mendoza, 2018).

2.3 Diseño de Contrastación

La investigación empleó un diseño preexperimental de un solo grupo con pretest y posttest. Este diseño permitió evaluar el aprendizaje de regresión lineal antes y después de la aplicación de la interfaz gráfica de usuario en MATLAB a los estudiantes de segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

El esquema del diseño fue el siguiente:

$$G \quad O_1 \quad X \quad O_2$$

Donde:

- G : Grupo de estudiantes de Ingeniería Civil seleccionados para la validación de la interfaz gráfica.
- O_1 : Aplicación del pretest de aprendizaje de regresión lineal.
- X : Aplicación de la interfaz gráfica de usuario en MATLAB durante el desarrollo de actividades de regresión lineal simple y múltiple.
- O_2 : Aplicación del posttest de aprendizaje de regresión lineal.

El proceso de contrastación comprendió dos componentes. El primero correspondió al desarrollo tecnológico de la interfaz gráfica, el cual incluyó el diseño de los módulos, la programación de los algoritmos de regresión lineal, la generación de tablas, gráficos, métricas de ajuste y predicciones. El segundo componente correspondió a la validación educativa de la herramienta mediante su aplicación a estudiantes de segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

La validación técnica se realizó mediante la comparación de los resultados producidos por la interfaz con procedimientos manuales y funciones internas de MATLAB. La validación educativa se realizó mediante la comparación de los resultados obtenidos en el pretest y posttest. Asimismo, se aplicó un cuestionario de usabilidad después de la intervención para conocer la percepción de los estudiantes sobre la interfaz gráfica.

Procedimiento de aplicación

El procedimiento de aplicación de la investigación se desarrolló en las siguientes etapas:

1. Se coordinó la aplicación de la investigación con los estudiantes de segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo seleccionados como grupo de validación.
2. Se aplicó el pretest para identificar el nivel inicial de aprendizaje de regresión lineal de los participantes.
3. Se presentó la interfaz gráfica de usuario en MATLAB y se explicó el funcionamiento de sus módulos principales.
4. Se desarrollaron actividades relacionadas con regresión lineal simple, incluyendo ingreso de datos, cálculo de medias, estimación de pendiente e intercepto, elaboración de gráficos e interpretación de R^2 .
5. Se desarrollaron actividades relacionadas con regresión lineal múltiple, incluyendo ingreso de variables predictoras, cálculo de coeficientes, análisis de valores estimados, residuos y predicciones.
6. Los estudiantes resolvieron ejercicios utilizando la interfaz gráfica como recurso de apoyo para el aprendizaje.
7. Se aplicó el posttest para identificar el nivel de aprendizaje después de la utilización de la GUI.
8. Se aplicó el cuestionario de usabilidad para conocer la percepción de los estudiantes sobre el diseño, funcionalidad, interactividad y utilidad de la interfaz.
9. Se procesaron los resultados obtenidos para realizar el análisis descriptivo e inferencial correspondiente.

2.4 Población, Muestra y Muestreo

La población objetivo estuvo conformada por estudiantes universitarios que cursan asignaturas vinculadas con estadística, análisis de datos, métodos cuantitativos o investigación aplicada, en las cuales se requiere comprender y aplicar procedimientos de regresión lineal.

Debido a que la interfaz gráfica de usuario en MATLAB fue diseñada como una herramienta de apoyo para la educación universitaria, su uso puede extenderse a estudiantes de diferentes escuelas profesionales que desarrollen contenidos de regresión lineal dentro de sus asignaturas de estadística.

Para validar el funcionamiento técnico, la usabilidad y el aporte de la interfaz al aprendizaje de regresión lineal, se trabajó con un grupo de aplicación conformado por 20 estudiantes del segundo ciclo 2026-I de la Escuela Profesional de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

La muestra estuvo constituida por los mismos 20 estudiantes de Ingeniería Civil, quienes participaron en la aplicación del pretest, en las actividades desarrolladas con la interfaz gráfica, en el postest y en el cuestionario de usabilidad.

Se empleó un muestreo no probabilístico por conveniencia, debido a que se seleccionó un grupo accesible de estudiantes que participó en la validación de la propuesta tecnológica. Por ello, los resultados obtenidos permitieron valorar el funcionamiento y aporte de la interfaz en el grupo evaluado, sin pretender generalizarlos automáticamente a todos los estudiantes universitarios.

2.5 Técnicas, Instrumentos, Equipos y Materiales de Recolección de Datos

Para la recolección de datos se emplearon técnicas de evaluación pedagógica, encuesta y validación funcional. Los instrumentos permitieron recoger información sobre el aprendizaje de regresión lineal, la percepción de usabilidad de la interfaz y el funcionamiento técnico de la herramienta.

Tabla 2*Técnicas e instrumentos de recolección de datos*

Técnica	Instrumento	Finalidad
Evaluación pedagógica	Pretest de regresión lineal	Identificar el nivel inicial de aprendizaje de regresión lineal antes de utilizar la interfaz gráfica.
Intervención educativa	Sesiones de aprendizaje con la GUI	Aplicar la interfaz gráfica de usuario en MATLAB en actividades de regresión lineal simple y múltiple.
Evaluación pedagógica	Postest de regresión lineal	Identificar el nivel de aprendizaje de regresión lineal después de utilizar la interfaz gráfica.
Encuesta	Cuestionario de usabilidad	Conocer la percepción de los estudiantes sobre el diseño, funcionalidad, interactividad, utilidad y satisfacción con la GUI.
Validación funcional	Ficha de validación funcional	Verificar el funcionamiento de ingreso de datos, cálculos, tablas, gráficos, predicciones y ejecutable de la interfaz.
Juicio de expertos	Matriz de validación	Evaluar la pertinencia, claridad y coherencia de los instrumentos empleados en la investigación.

Los equipos empleados fueron computadoras o laptops con acceso a la interfaz gráfica desarrollada en MATLAB R2021a. También se emplearon hojas de aplicación, pretest, postest, cuestionarios de usabilidad y materiales de apoyo para el desarrollo de las actividades.

2.6 Procesamiento y Análisis de Datos

- **Procesamiento de Datos:** La interfaz procesa los datos ingresados por el usuario (manual o mediante carga de archivo), realiza validaciones básicas (formato numérico, tamaño de muestras, presencia de datos faltantes) y ejecuta los algoritmos programados para estimar el modelo de regresión lineal. Como salida, genera parámetros del modelo, predicciones, residuales, sumas de cuadrados y métricas de ajuste.
- **Análisis de Datos:** Se realiza un análisis comparativo entre los resultados obtenidos mediante la interfaz y los obtenidos por métodos manuales (en ejercicios controlados)

o por software estadístico. El análisis considera precisión numérica de coeficientes, coherencia del ajuste, interpretación de R^2 y evaluación mediante gráficos (dispersión y residuales), observando además cómo la interfaz facilita la comprensión de los conceptos y procedimientos de la regresión lineal.

- **Validación de Resultados:** Se desarrolla un proceso iterativo de validación técnica, verificando que la interfaz entregue resultados correctos y consistentes (coeficientes, $\hat{y}$, residuales, SST/SSR/SSE, R^2 , y gráficos asociados). Se evalúa también la claridad de la presentación visual y su utilidad pedagógica para que el estudiante identifique patrones, ajuste y posibles desviaciones del modelo.
- **Creación del Ejecutable:** Una vez verificado el correcto funcionamiento de la interfaz gráfica de usuario desarrollada en MATLAB, se procede a generar su versión ejecutable utilizando MATLAB Compiler. Esto permite que la aplicación sea utilizada por estudiantes, docentes e investigadores sin necesidad de tener MATLAB instalado. El ejecutable puede ser distribuido dentro del entorno educativo, favoreciendo accesibilidad, portabilidad y uso práctico en sesiones de aprendizaje y reforzamiento académico.

Análisis inferencial del aprendizaje

Para evaluar el aporte de la interfaz gráfica de usuario en MATLAB al aprendizaje de regresión lineal, se compararon los puntajes obtenidos por los estudiantes en el pretest y posttest.

Primero, se evaluó la normalidad de las diferencias entre ambos puntajes mediante la prueba de Shapiro-Wilk. Cuando el valor de significancia fue mayor que 0.05, se aplicó la prueba t de Student para muestras relacionadas. Cuando el valor de significancia fue menor o igual que 0.05, correspondió aplicar la prueba no paramétrica de rangos con signo de Wilcoxon.

En la presente investigación, la prueba de Shapiro-Wilk mostró un valor de significancia mayor que 0.05, por lo cual se utilizó la prueba t de Student para muestras relacionadas.

El nivel de significancia establecido fue: $\alpha = 0.05$.

La hipótesis nula planteó que la aplicación de la interfaz gráfica de usuario en MATLAB no produjo una mejora significativa en el aprendizaje de regresión lineal. La hipótesis de investigación planteó que la aplicación de la interfaz gráfica de usuario en MATLAB mejoró significativamente el aprendizaje de regresión lineal en los estudiantes participantes.

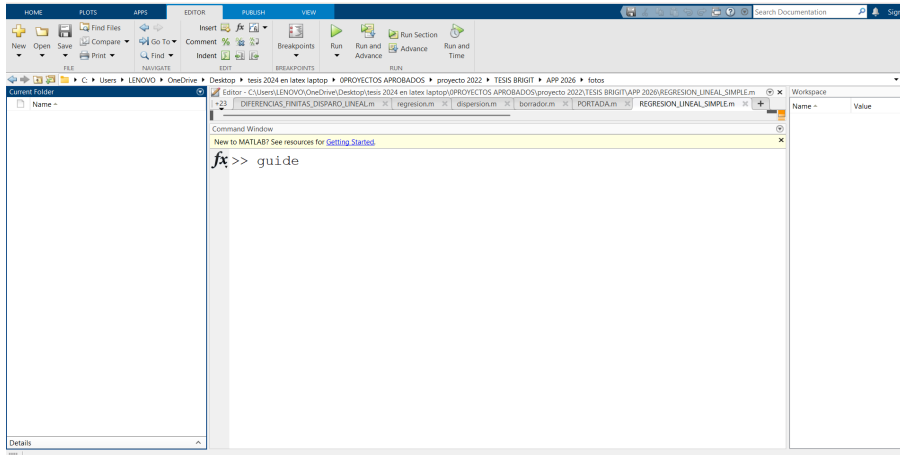
III. RESULTADOS

3.1 Interfaz Gráfica de Usuario para la regresión lineal simple

Paso 1: La Figura (1) muestra la palabra guide, lo cual va a permitir ingresar al diseño de la interfaz gráfica de usuario (GUI).

Figura 1

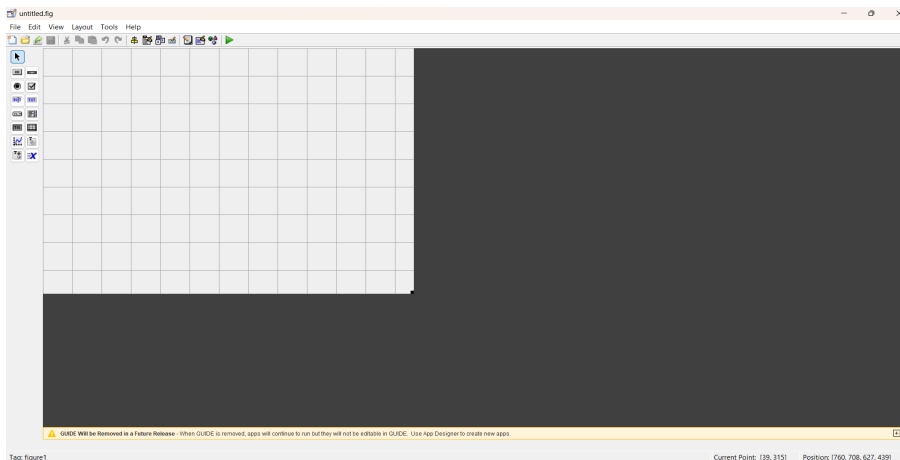
Ingreso de la palabra guide.



Paso 2: La Figura (2) muestra el entorno de diseño de la GUI en blanco.

Figura 2

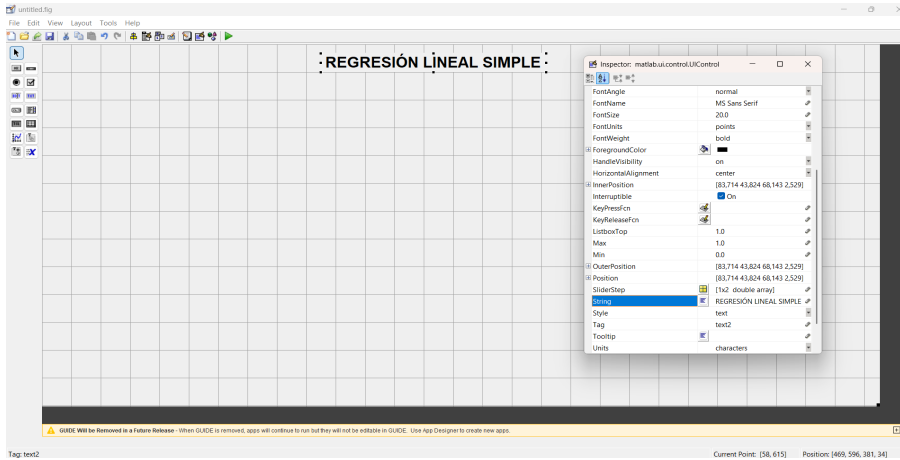
Entorno de la GUI.



Paso 3: La Figura (3), se mostró el uso de un static text para colocar el nombre de REGRESIÓN LINEAL SIMPLE.

Figura 3

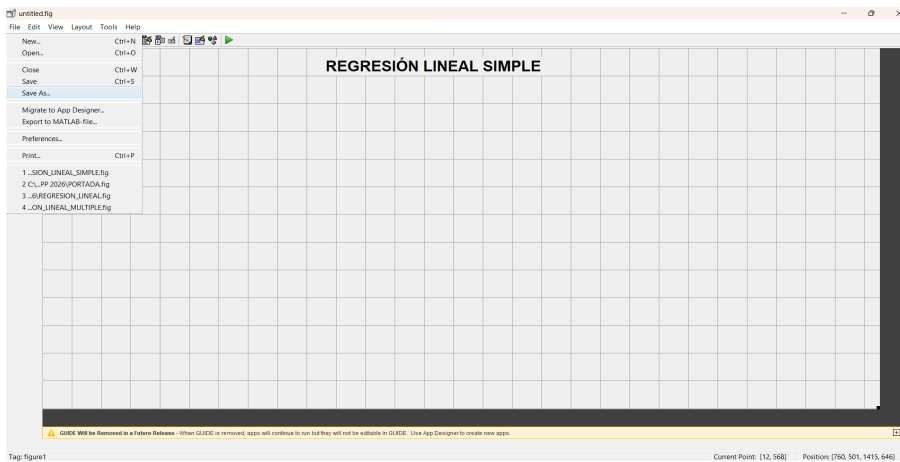
Se colocó el nombre de regresión lineal simple.



Paso 4: La Figura (4) muestra como se guarda la GUI.

Figura 4

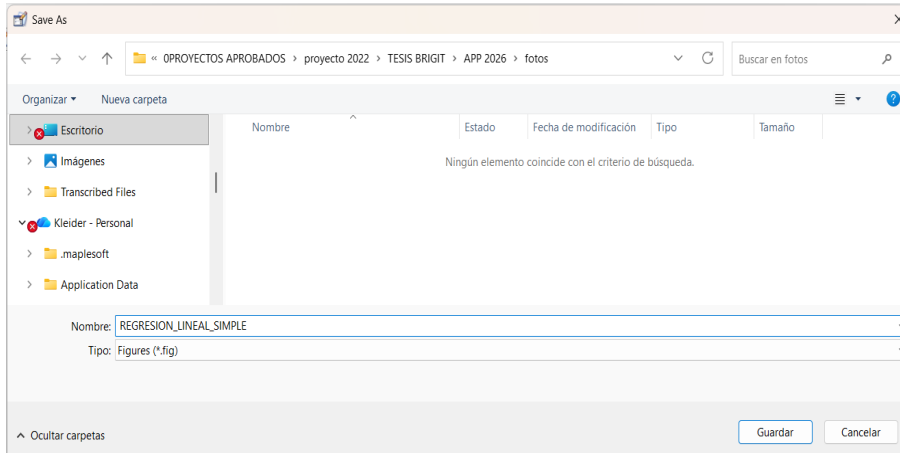
Guardar.



Paso 5: La Figura (5) muestra que se guardó con el nombre de REGRESION LINEAL SIMPLE.

Figura 5

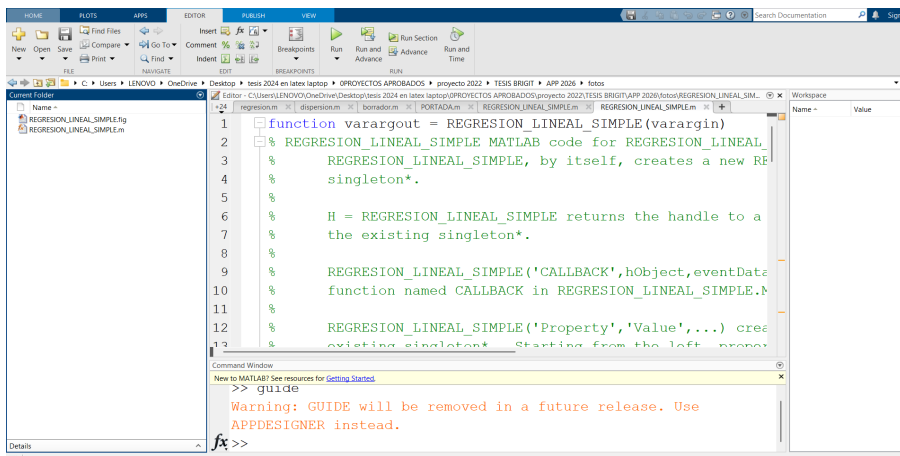
Nombre de la GUI.



Paso 6: La Figura (6) muestra que se generaron 2 archivos, uno con la extensión .fig y el otro con la extensión .m.

Figura 6

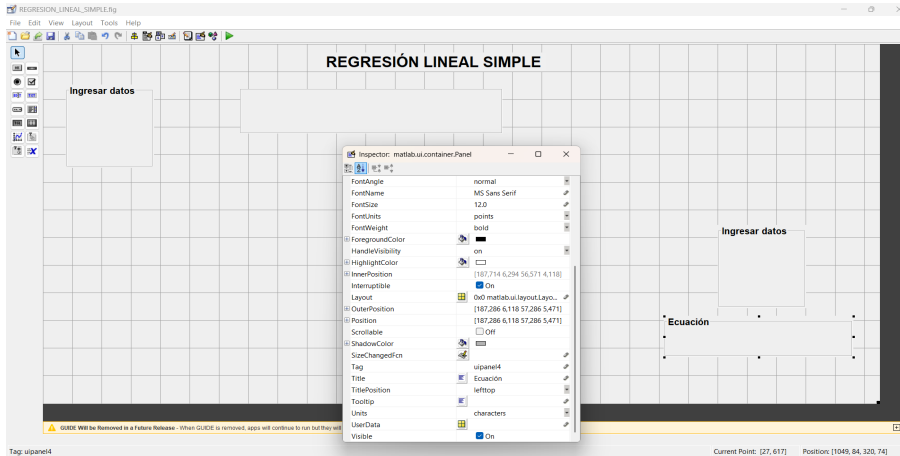
Entorno de la GUI.



Paso 7: La Figura (7) muestra que se han ingresado 4 paneles para encapsular los static text y edit text.

Figura 7

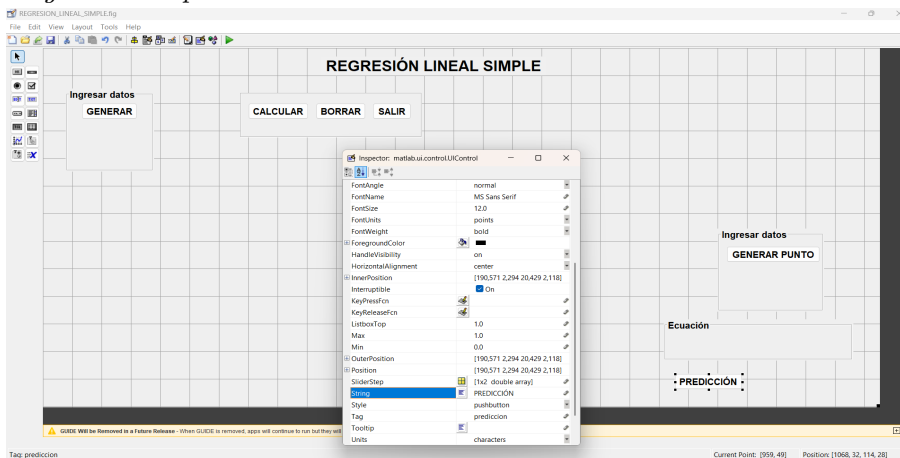
Ingreso de 4 paneles.



Paso 8: La Figura (8) muestra que se han ingresado 6 botones, para poder GENERAR, CALCULAR, BORRAR, SALIR, GENERAR PUNTO, PREDICCIÓN.

Figura 8

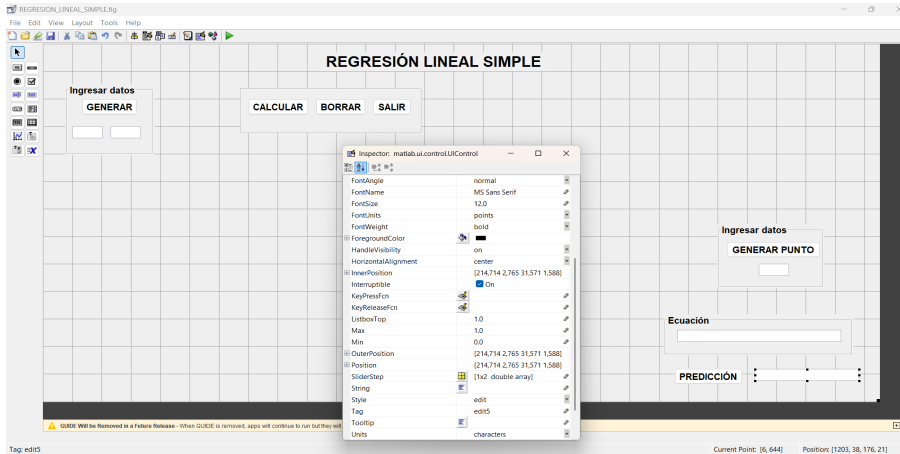
Ingreso de 6 pushbutton.



Paso 9: La Figura (9) muestra que se han ingresado 5 edit text, para poder ingresar los valores de filas y columnas, también permitan mostrar la ecuación y el valor predictivo.

Figura 9

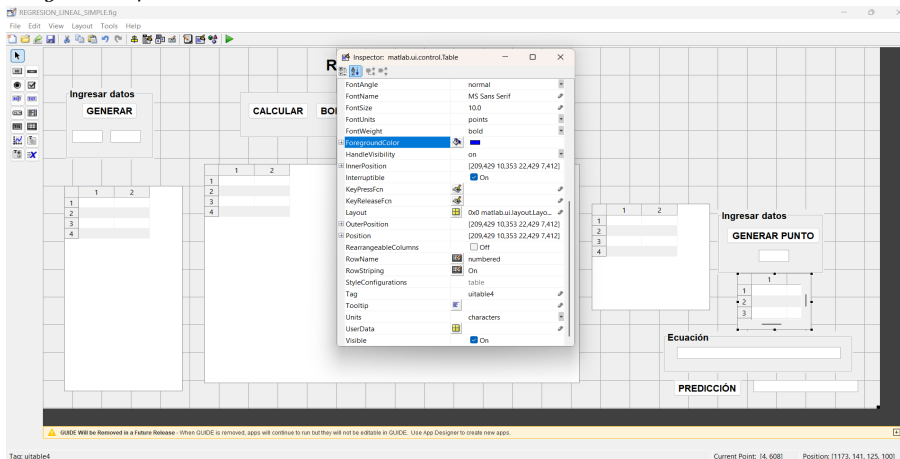
Se ingresó 5 edit text.



Paso 10: La Figura (10) muestra que se ingresó 4 uitables, para cargar y mostrar los datos.

Figura 10

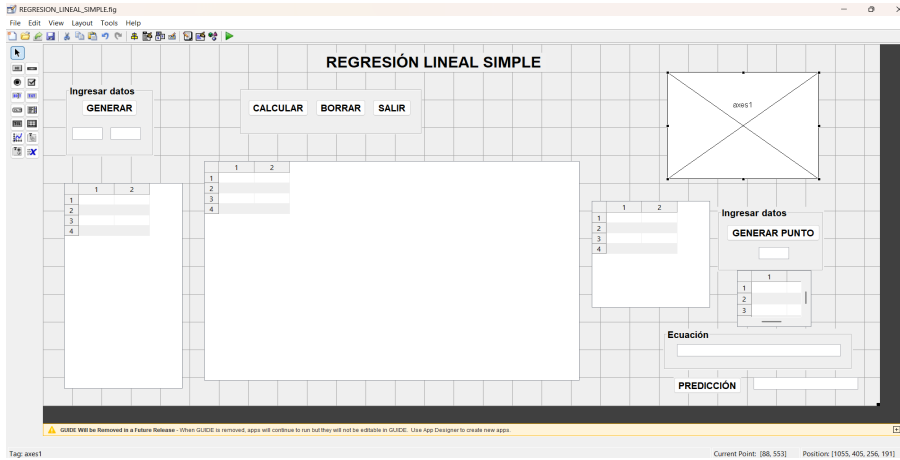
Ingresar 4 tablas.



Paso 11: La Figura (11) muestra que se ingresó un axes para mostrar la ecuación lineal.

Figura 11

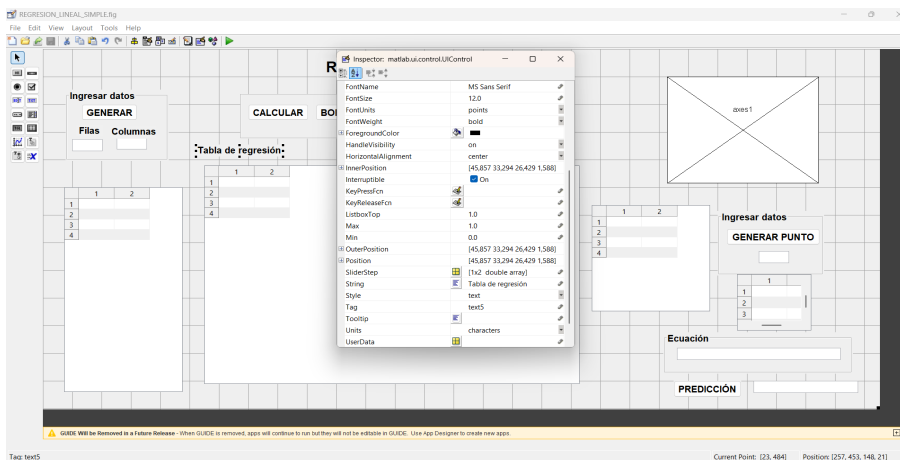
Se ingresó un axes para mostrar la gráfica de la ecuación lineal.



Paso 12: La Figura (12) muestra que se ingresó 3 static text, para colocar el nombre de filas, columnas y tabla de regresión.

Figura 12

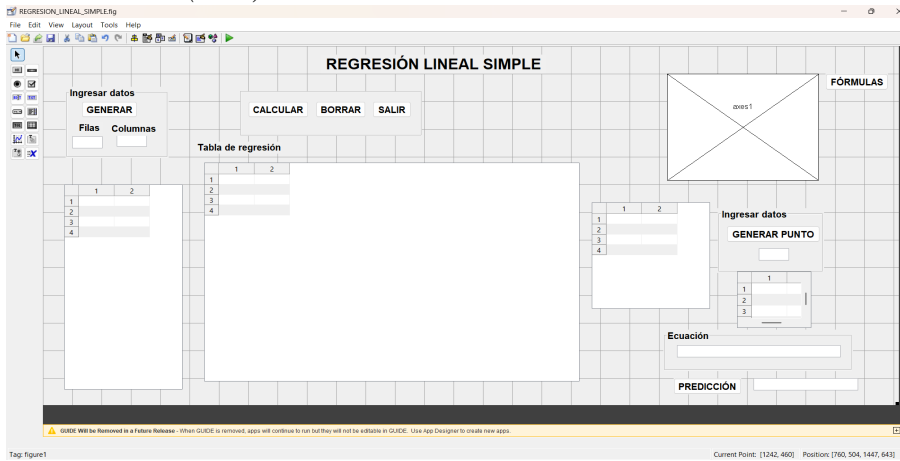
Se colocó 3 static text.



Paso 13: La Figura (13) muestra el diseño de la regresión lineal simple.

Figura 13

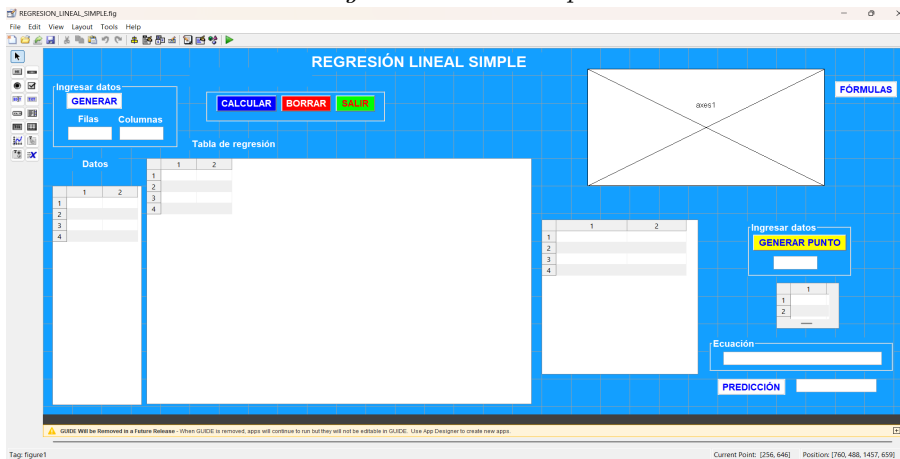
Diseño de la (GUI).



Paso 14: La Figura (14) muestra el diseño de la GUI de la regresión lineal simple.

Figura 14

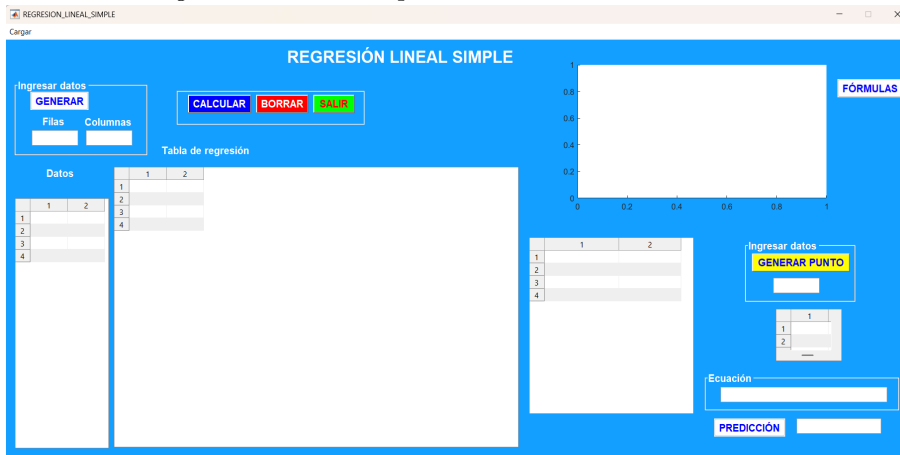
Diseño de la GUI de la regresión lineal simple.



Paso 15: La Figura (15) muestra el diseño final de la GUI.

Figura 15

GUI de la regresión lineal simple.



3.1.1 Aplicación de la regresión lineal simple

Una cadena de restaurantes de comida italiana tiene locales en todo el país. Han notado que los mejores resultados se dan en los que están cerca de universidades. Por eso, los gerentes piensan que las ventas de cada trimestre dependen del número de estudiantes: si la universidad tiene muchos alumnos, el restaurante vende más; si tiene pocos, vende menos.

Tabla 3

Población de estudiantes y ventas trimestrales en 10 restaurantes

Restaurante	Población de estudiantes (miles)	Ventas trimestrales (miles de \$)
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202

Solución

Tabla 4

Tabla de cálculos para la regresión lineal

i	x	y	$(x - \bar{x})$	$(y - \bar{y})$	$(x - \bar{x})(y - \bar{y})$	$(x - \bar{x})^2$	$\hat{Y} = b_0 + b_1x$	$e = y - \hat{Y}$	e^2
1	2	58	-12	-72	864	144	70	-12	144
2	6	105	-8	-25	200	64	90	15	225
3	8	88	-6	-42	252	36	100	-12	144
4	8	118	-6	-12	72	36	100	18	324
5	12	117	-2	-13	26	4	120	-3	9
6	16	137	2	7	14	4	140	-3	9
7	20	157	6	27	162	36	160	-3	9
8	20	169	6	39	234	36	160	9	81
9	22	149	8	19	152	64	170	-21	441
10	26	202	12	72	864	144	190	12	144
$\bar{x}, \bar{y}$	14	130							
Σ					2840	568			1530

1. Datos y tamaño de muestra

Se consideran los datos:

$$x = \{2, 6, 8, 8, 12, 16, 20, 20, 22, 26\}, \quad y = \{58, 105, 88, 118, 117, 137, 157, 169, 149, 202\}, \quad n = 10.$$

2. Cálculo de las medias

■ Media de x :

$$\sum x_i = 2 + 6 + 8 + 8 + 12 + 16 + 20 + 20 + 22 + 26 = 140$$

$$\bar{x} = \frac{1}{n} \sum x_i = \frac{140}{10} = 14$$

■ Media de y :

$$\sum y_i = 58 + 105 + 88 + 118 + 117 + 137 + 157 + 169 + 149 + 202 = 1300$$

$$\bar{y} = \frac{1}{n} \sum y_i = \frac{1300}{10} = 130$$

3. Suma de cuadrados S_{xx}

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$$

$$S_{xx} = (-12)^2 + (-8)^2 + (-6)^2 + (-6)^2 + (-2)^2 + (2)^2 + (6)^2 + (6)^2 + (8)^2 + (12)^2$$

$$S_{xx} = 144 + 64 + 36 + 36 + 4 + 4 + 36 + 36 + 64 + 144 = 568$$

$$\boxed{S_{xx} = 568}$$

4. Producto cruzado S_{xy}

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$S_{xy} = 864 + 200 + 252 + 72 + 26 + 14 + 162 + 234 + 152 + 864 = 2840$$

$$\boxed{S_{xy} = 2840}$$

5. Cálculo de la pendiente b_1

$$b_1 = \frac{S_{xy}}{S_{xx}} = \frac{2840}{568} = 5$$

$$\boxed{b_1 = 5}$$

6. Cálculo del intercepto b_0

$$b_0 = \bar{y} - b_1\bar{x} = 130 - 5(14) = 130 - 70 = 60$$

$$\boxed{b_0 = 60}$$

7. Ecuación del modelo de regresión lineal simple

$$\hat{y} = b_0 + b_1x$$

$$\boxed{\hat{y} = 60 + 5x}$$

8. Predicciones $\hat{y}_i$, errores e_i y suma de cuadrados del error SCE

- Para cada dato:

$$\hat{y}_i = 60 + 5x_i, \quad e_i = y_i - \hat{y}_i, \quad e_i^2 = (y_i - \hat{y}_i)^2$$

- La suma de cuadrados del error:

$$SCE = \sum_{i=1}^n e_i^2$$

$$SCE = 144 + 225 + 144 + 324 + 9 + 9 + 9 + 81 + 441 + 144 = 1530$$

$$\boxed{SCE = 1530}$$

9. Suma total de cuadrados (STC)

$$STC = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$STC = (-72)^2 + (-25)^2 + (-42)^2 + (-12)^2 + (-13)^2 + (7)^2 + (27)^2 + (39)^2 + (19)^2 + (72)^2$$

$$STC = 5184 + 625 + 1764 + 144 + 169 + 49 + 729 + 1521 + 361 + 5184 = 15730$$

$$\boxed{STC = 15730}$$

10. Suma de cuadrados de la regresión (SCR)

$$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$SCR = (70-130)^2 + (90-130)^2 + (100-130)^2 + (100-130)^2 + (120-130)^2 + (140-130)^2$$

$$+ (160-130)^2 + (160-130)^2 + (170-130)^2 + (190-130)^2$$

$$SCR = 3600 + 1600 + 900 + 900 + 100 + 100 + 900 + 900 + 1600 + 3600 = 14200$$

$$\boxed{SCR = 14200}$$

11. Verificación de la descomposición

$$STC = SCR + SCE \Rightarrow 15730 = 14200 + 1530$$

12. Coeficiente de determinación R^2

$$R^2 = \frac{SCR}{STC} = \frac{14200}{15730} \approx 0.9027$$

$$\boxed{R^2 \approx 0.9027 \text{ (90.27\%)}}$$

13. Error cuadrático medio (MSE)

$$MSE = \frac{SCE}{n-2} = \frac{1530}{10-2} = \frac{1530}{8} = 191.25$$

$$\boxed{MSE = 191.25}$$

14. Raíz del MSE (RMSE)

$$RMSE = \sqrt{MSE} = \sqrt{191.25} \approx 13.83$$

$$\boxed{RMSE \approx 13.83}$$

Del procedimiento realizado se concluye que **existe una relación lineal positiva y fuerte** entre la **población estudiantil** cercana al campus (x) y las **ventas trimestrales** del restaurante (y). En consecuencia, a mayor cantidad de estudiantes, mayores ventas, y el modelo obtenido permite realizar predicciones con buen nivel de ajuste.

■ Modelo estimado:

$$\hat{y} = 60 + 5x.$$

- **Interpretación de la pendiente:**

Como x está expresado en miles de estudiantes y y en miles de dólares, la pendiente $b_1 = 5$ indica que por cada **1,000 estudiantes adicionales**, las ventas trimestrales aumentan en promedio **5,000 dólares**. Equivalentemente, esto representa aproximadamente **5 dólares por estudiante**.

- **Ejemplo de predicción:**

Para un campus con 26,000 estudiantes ($x = 26$):

$$\hat{y} = 60 + 5(26) = 190,$$

lo cual significa ventas trimestrales pronosticadas de:

$$190 \text{ (miles de dólares)} = 190,000 \text{ dólares.}$$

- **Calidad del ajuste:**

El coeficiente de determinación obtenido fue:

$$R^2 = \frac{SCR}{STC} = \frac{14200}{15730} \approx 0.9027 \text{ (90.27 \%)}.$$

Por tanto, **el 90.27 % de la variabilidad** de las ventas se explica por la relación lineal con la población estudiantil, mientras que el $\approx 9.73\%$ restante se debe a otros factores y al error aleatorio.

- **Fuerza de la relación:**

El coeficiente de correlación fue:

$$r = +\sqrt{R^2} = +\sqrt{0.9027} = 0.9501,$$

lo cual evidencia una **correlación lineal positiva fuerte** entre x e y .

■ **Error típico de predicción:**

Con $SCE = 1530$ y $n = 10$, se obtuvo:

$$MSE = \frac{SCE}{n-2} = \frac{1530}{8} = 191.25, \quad RMSE = \sqrt{MSE} \approx 13.83.$$

Dado que y está en miles de dólares, el error típico es aproximadamente:

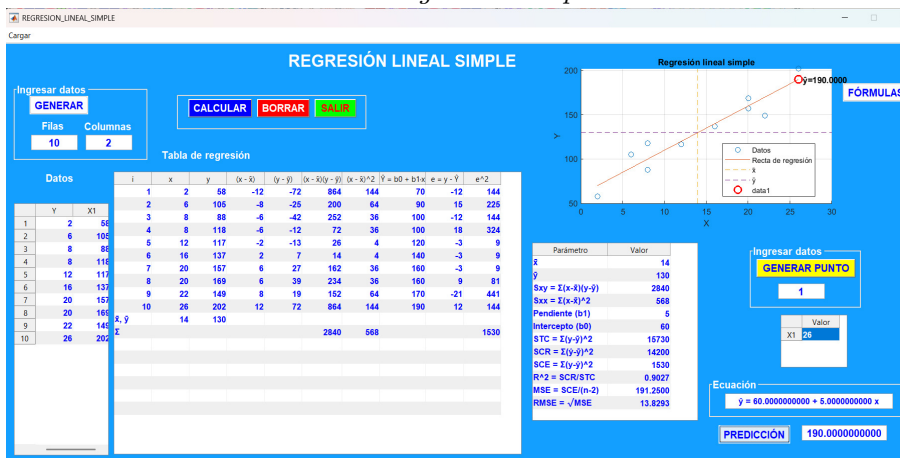
$$RMSE \approx 13.83 \text{ (miles)} \approx 13830 \text{ dólares.}$$

Esto significa que, en promedio, las predicciones pueden diferir del valor real en alrededor de ± 13830 dólares por trimestre.

En síntesis, el análisis confirma que la población estudiantil es un **factor determinante** para las ventas trimestrales de los restaurantes cercanos a universidades, y el modelo lineal obtenido resulta **útil y confiable para pronosticar** las ventas en función del tamaño del campus.

Figura 16

Solución mediante la GUI de regresión simple.



La Figura (16) muestra una interfaz gráfica desarrollada en *MATLAB* para aplicar el modelo de **regresión lineal simple**. La aplicación permite ingresar un conjunto de datos (x, y) , ejecutar el cálculo de los parámetros del modelo, visualizar el ajuste en un gráfico

y obtener predicciones para nuevos valores de x . En este caso, se trabaja con $n = 10$ observaciones.

La interfaz se organiza en bloques funcionales:

- **Ingreso de datos:** se define el número de *filas* y *columnas* (en la imagen: 10 filas y 2 columnas) y se presiona el botón **GENERAR** para crear la tabla de entrada.
- **Botones de control:**
 - **CALCULAR:** ejecuta el procedimiento de regresión lineal.
 - **BORRAR:** limpia los valores ingresados y los resultados obtenidos.
 - **SALIR:** cierra la aplicación.
- **Tabla de datos:** permite ingresar los pares (x, y) .
- **Tabla de regresión:** presenta el desarrollo de los cálculos por observación:

$$(x - \bar{x}), \quad (y - \bar{y}), \quad (x - \bar{x})(y - \bar{y}), \quad (x - \bar{x})^2, \quad \hat{y} = b_0 + b_1x, \quad e = y - \hat{y}, \quad e^2.$$

- **Panel de parámetros:** resume los valores estadísticos calculados.
- **Gráfico:** muestra los datos observados y la recta de regresión, incluyendo referencias de $\bar{x}$ y $\bar{y}$, y el punto de predicción.

Del panel *Parámetro-Valor* se observan los siguientes resultados:

- Medias:

$$\bar{x} = 14, \quad \bar{y} = 130.$$

- Sumatorias para el ajuste:

$$S_{xy} = \sum (x - \bar{x})(y - \bar{y}) = 2840, \quad S_{xx} = \sum (x - \bar{x})^2 = 568.$$

- Parámetros del modelo:

$$b_1 = \frac{S_{xy}}{S_{xx}} = \frac{2840}{568} = 5, \quad b_0 = \bar{y} - b_1\bar{x} = 130 - 5(14) = 60.$$

- Ecuación de la regresión lineal simple:

$$\hat{y} = 60 + 5x.$$

- Sumas de cuadrados:

$$STC = \sum (y - \bar{y})^2 = 15730, \quad SCR = \sum (\hat{y} - \bar{y})^2 = 14200, \quad SCE = \sum (y - \hat{y})^2 = 1530.$$

- Coeficiente de determinación:

$$R^2 = \frac{SCR}{STC} = \frac{14200}{15730} = 0,9027.$$

- Medidas de error:

$$MSE = \frac{SCE}{n-2} = \frac{1530}{8} = 191,25, \quad RMSE = \sqrt{MSE} = \sqrt{191,25} \approx 13,8293.$$

A partir de los resultados se interpreta lo siguiente:

- La pendiente $b_1 = 5$ indica una **relación positiva**: cuando x aumenta, el valor estimado $\hat{y}$ también aumenta.
- El coeficiente $R^2 = 0,9027$ muestra que **aproximadamente el 90.27 %** de la variabilidad de y es explicada por el modelo lineal en función de x . Esto sugiere un **ajuste alto** y una relación lineal fuerte.
- El valor $RMSE \approx 13,8293$ representa el **error típico de predicción** del modelo en las mismas unidades de y . Es decir, en promedio, las predicciones pueden diferir del

valor real en alrededor de $\pm 13,8293$.

En el panel de predicción se ingresa el valor $X_1 = 26$. La aplicación evalúa el modelo:

$$\hat{y} = 60 + 5(26) = 190,$$

por lo cual se obtiene una **predicción** de:

$$\hat{y} = 190.$$

En la gráfica, este resultado se observa como un punto resaltado y etiquetado alrededor de $(26, 190)$, mostrando coherencia entre el cálculo numérico y la visualización.

La interfaz gráfica permite integrar en un solo entorno:

- el **ingreso** de datos,
- el **cálculo automático** del modelo lineal $\hat{y} = b_0 + b_1x$,
- la **visualización** del ajuste mediante la recta de regresión,
- la **evaluación del ajuste** con R^2 , MSE y $RMSE$,
- y la **predicción** inmediata para nuevos valores de x .

Por ello, la herramienta resulta adecuada como apoyo didáctico, ya que muestra tanto el procedimiento numérico (tablas) como la interpretación gráfica del modelo.

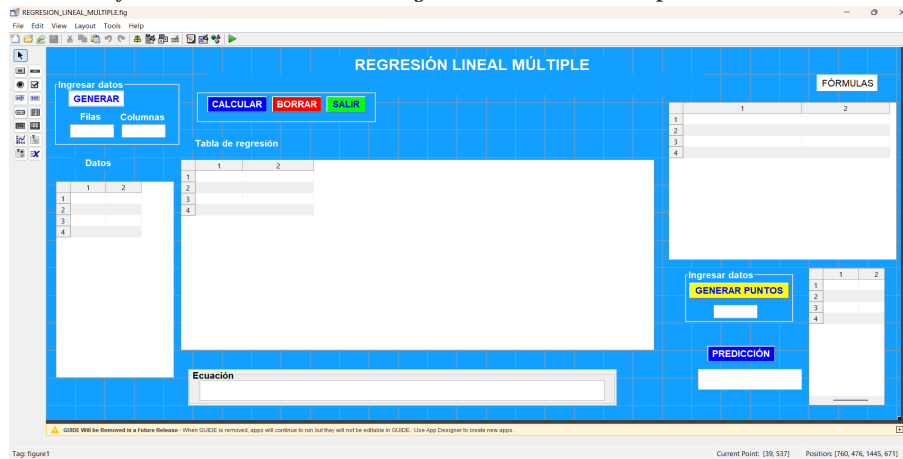
3.2 Interfaz Gráfica de Usuario para la regresión lineal múltiple

Para poder realizar el diseño de interfaz gráfica de usuario para la regresión lineal múltiple, se debe realizar los mismos 15 pasos mencionados anteriormente.

Paso 16: La Figura (17) muestra el diseño final de la GUI para la regresión lineal múltiple.

Figura 17

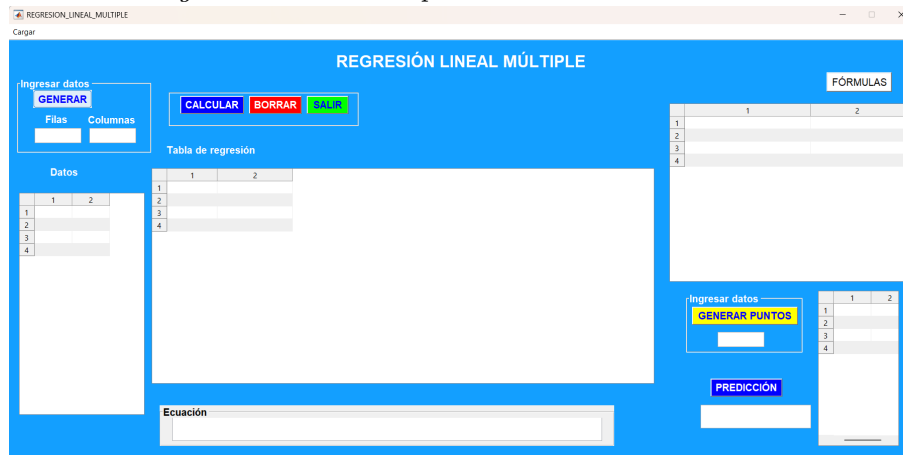
Diseño final de la GUI de la regresión lineal múltiple.



Paso 17: La Figura (18) muestra la GUI para la regresión lineal múltiple.

Figura 18

GUI de la regresión lineal múltiple.



3.2.1 Aplicación de la regresión lineal múltiple

Se hizo un estudio con 12 estudiantes para analizar cómo influyen la nota del examen y la cantidad de clases que faltan en la calificación final del curso de estadística. Los datos completos se muestran a continuación.

Tabla 5

Calificación de estadística, calificación del examen y clases perdidas ($n = 12$)

Estudiante	Calificación de estadística	Calificación del examen	Clases perdidas
1	85	65	1
2	74	50	7
3	76	55	5
4	90	65	2
5	85	55	6
6	87	70	3
7	94	65	2
8	98	70	5
9	81	55	4
10	91	70	3
11	76	50	1
12	74	55	4

Solución

Tabla 6

Valores observados, estimados y errores del modelo de regresión ($n = 12$)

i	X_1	X_2	Y	$\hat{Y}$	$e = Y - \hat{Y}$	e^2
1	65	1	85	87.7400	-2.7400	7.5079
2	50	7	74	75.6204	-1.6204	2.6256
3	55	5	76	79.6603	-3.6603	13.3975
4	65	2	90	88.0243	1.9757	3.9034
5	55	6	85	79.9445	5.0555	25.5580
6	70	3	87	92.9169	-5.9169	35.0102
7	65	2	94	88.0243	5.9757	35.7090
8	70	5	98	93.4854	4.5146	20.3813
9	55	4	81	79.3760	1.6240	2.6373
10	70	3	91	92.9169	-1.9169	3.6747
11	50	1	76	73.9149	2.0851	4.3478
12	55	4	74	79.3760	-5.3760	28.9015
Promedio	60.4167	3.5833	84.2500			
Σ	725.0000	43.0000	1011.0000			183.6541

Se estudia la relación entre la calificación final de Estadística (Y) y dos variables:

$$X_1 = \text{calificación del examen}, \quad X_2 = \text{clases perdidas}.$$

El modelo es:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2.$$

1. Parámetros del estudio

$$n = 12, \quad p = 2.$$

2. Sumatorias obtenidas de los datos

$$\sum Y = 1011, \quad \sum X_1 = 725, \quad \sum X_2 = 43$$

$$\sum X_1^2 = 44475, \quad \sum X_2^2 = 195, \quad \sum (X_1X_2) = 2540$$

$$\sum (X_1Y) = 61685, \quad \sum (X_2Y) = 3581.$$

Promedio de Y :

$$\bar{Y} = \frac{\sum Y}{n} = \frac{1011}{12} = 84.25.$$

3. Matriz de ecuaciones normales (forma matricial)

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \\ \sum X_2 & \sum X_1X_2 & \sum X_2^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum (X_1Y) \\ \sum (X_2Y) \end{bmatrix}$$

Sustituyendo:

$$\begin{bmatrix} 12 & 725 & 43 \\ 725 & 44475 & 2540 \\ 43 & 2540 & 195 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1011 \\ 61685 \\ 3581 \end{bmatrix}$$

Sistema equivalente:

$$\begin{cases} 12b_0 + 725b_1 + 43b_2 = 1011 \\ 725b_0 + 44475b_1 + 2540b_2 = 61685 \\ 43b_0 + 2540b_1 + 195b_2 = 3581 \end{cases}$$

4. **Despeje de b_0** De la primera ecuación:

$$b_0 = \frac{1011 - 725b_1 - 43b_2}{12}.$$

5. **Reducción a 2 ecuaciones con 2 incógnitas** Sustituyendo b_0 en la 2da y 3ra ecuación y simplificando:

$$\begin{cases} 8075b_1 - 695b_2 = 7245 \\ -695b_1 + 491b_2 = -501 \end{cases}$$

6. **Cálculo de b_1 y b_2 (Regla de Cramer)** Determinante principal:

$$D = \begin{vmatrix} 8075 & -695 \\ -695 & 491 \end{vmatrix} = 8075(491) - (-695)(-695) = 8075(491) - 695^2$$

$$D = 3966625 - 483025 = 3481800.$$

Determinante para b_1 :

$$D_1 = \begin{vmatrix} 7245 & -695 \\ -501 & 491 \end{vmatrix} = 7245(491) - (-695)(-501) = 7245(491) - 695(501)$$

$$D_1 = 3557295 - 348195 = 3209100.$$

Determinante para b_2 :

$$D_2 = \begin{vmatrix} 8075 & 7245 \\ -695 & -501 \end{vmatrix} = 8075(-501) - 7245(-695)$$

$$D_2 = -4045575 + 5035275 = 989700.$$

Coeficientes:

$$b_1 = \frac{D_1}{D} = \frac{3209100}{3481800} = 0.9216784422, \quad b_2 = \frac{D_2}{D} = \frac{989700}{3481800} = 0.2842495261.$$

7. Cálculo del intercepto b_0

$$b_0 = \frac{1011 - 725(0.9216784422) - 43(0.2842495261)}{12} = 27.5466999828.$$

8. Ecuación final estimada

$$\hat{Y} = 27.5466999828 + 0.9216784422 X_1 + 0.2842495261 X_2$$

9. Cálculo de SST (suma total de cuadrados) Se usa:

$$SST = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}.$$

Con:

$$\sum Y_i^2 = 85905, \quad \sum Y_i = 1011, \quad n = 12,$$

$$SST = 85905 - \frac{1011^2}{12} = 85905 - \frac{1022121}{12} = 85905 - 85176.75 = 728.25.$$

$$SST = 728.25$$

10. **Cálculo de SSE (suma de cuadrados del error)** Por definición:

$$SSE = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2.$$

De la tabla de residuos:

$$SSE = 183.6540582457$$

11. **Cálculo de SSR (suma de cuadrados de la regresión)**

$$SSR = SST - SSE = 728.25 - 183.6540582457 = 544.5959417543.$$

$$SSR = 544.5959417543$$

12. **Cálculo de R^2 y R**

$$R^2 = \frac{SSR}{SST} = \frac{544.5959417543}{728.25} = 0.7478145441.$$

$$R^2 = 0.7478145441$$

$$R = \sqrt{R^2} = \sqrt{0.7478145441} = 0.8647627097.$$

$$R = 0.8647627097$$

13. **Cálculo de R^2 ajustado**

$$R_{aj}^2 = 1 - \frac{SSE/(n - p - 1)}{SST/(n - 1)}.$$

$$R_{aj}^2 = 1 - \frac{183.6540582457/(12 - 2 - 1)}{728.25/(12 - 1)} = 1 - \frac{183.6540582457/9}{728.25/11}.$$

$$R_{aj}^2 = 1 - \frac{20.4060064717}{66.2045454545} = 0.6917733317.$$

$$R_{aj}^2 = 0.6917733317$$

14. Cálculo de MSE y RMSE

$$MSE = \frac{SSE}{n - p - 1} = \frac{183.6540582457}{9} = 20.4060064717.$$

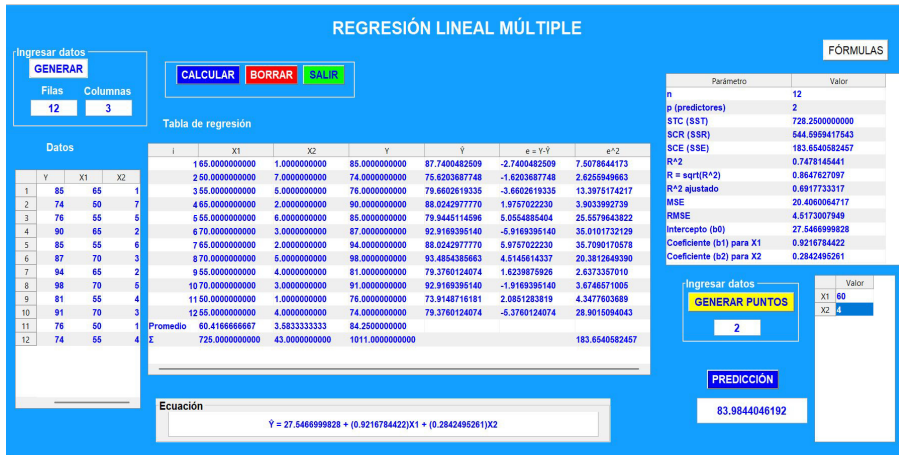
$$MSE = 20.4060064717$$

$$RMSE = \sqrt{MSE} = \sqrt{20.4060064717} = 4.5173007949.$$

$$RMSE = 4.5173007949$$

Figura 19

Solución mediante la GUI de regresión múltiple.



La Figura (19) muestra una interfaz gráfica desarrollada para ejecutar el procedimiento de **regresión lineal múltiple**, donde la variable dependiente es la **calificación del curso de Estadística (Y)**, y las variables independientes son la **calificación del examen (X₁)** y el **número de clases perdidas (X₂)**. La interfaz permite ingresar los datos, calcular el modelo, visualizar resultados numéricos y realizar predicciones de manera automática.

1. Ingreso y organización de datos.

En el panel izquierdo se observa la sección de *Ingresar datos*, donde se define el tamaño de la matriz (12 filas y 3 columnas), correspondiente a los $n = 12$ estudiantes analizados y a las tres variables registradas (Y, X₁, X₂). Posteriormente, los datos

son cargados en una tabla, lo que permite que el usuario verifique rápidamente cada observación antes de procesarla.

2. Cálculo y generación de la tabla de regresión.

En la parte central se presenta la *Tabla de regresión*, que resume para cada estudiante: los valores de X_1 y X_2 , el valor real de Y , el valor estimado $\hat{Y}$, el residuo $e = Y - \hat{Y}$ y el error cuadrático e^2 . Esta salida es clave para evaluar el ajuste del modelo, debido a que permite identificar qué observaciones presentan mayores diferencias entre el valor observado y el valor estimado.

3. Estimación del modelo y ecuación final.

En la sección inferior se muestra la ecuación estimada por mínimos cuadrados:

$$\hat{Y} = 27.5466999828 + 0.9216784422X_1 + 0.2842495261X_2.$$

El intercepto $b_0 = 27.5466999828$ representa el valor esperado de Y cuando $X_1 = 0$ y $X_2 = 0$. El coeficiente $b_1 = 0.9216784422$ indica que, manteniendo constante X_2 , un aumento de una unidad en la calificación del examen incrementa en promedio la calificación de Estadística en aproximadamente 0.92 puntos. Asimismo, el coeficiente $b_2 = 0.2842495261$ sugiere que, manteniendo constante X_1 , cada clase perdida aumenta en promedio el valor estimado de Y en aproximadamente 0.28 puntos, lo cual refleja el comportamiento del conjunto de datos utilizado para el ajuste.

4. Medidas globales de ajuste del modelo.

En el panel derecho se muestran los indicadores principales del ajuste:

$$SST = 728.25, \quad SSR = 544.5959417543, \quad SSE = 183.6540582457.$$

Estos resultados evidencian que la variación total de la variable dependiente (SST) se descompone en variación explicada por el modelo (SSR) y variación no explicada

o error (SSE). Además, el coeficiente de determinación

$$R^2 = 0.7478145441$$

indica que aproximadamente el 74.78 % de la variabilidad de Y es explicada conjuntamente por X_1 y X_2 . El coeficiente de correlación múltiple

$$R = 0.8647627097$$

muestra una asociación positiva fuerte entre los valores observados y estimados. Por otro lado, el valor de R^2 ajustado

$$R_{aj}^2 = 0.6917733317$$

corrige el ajuste considerando el tamaño de muestra y el número de predictores, lo cual confirma que el modelo mantiene un nivel adecuado de explicación.

5. Error promedio de predicción.

La interfaz reporta:

$$MSE = 20.4060064717, \quad RMSE = 4.5173007949.$$

El $RMSE$ expresa el error típico de predicción en unidades de la variable Y , por lo que se interpreta como que, en promedio, las predicciones del modelo pueden desviarse alrededor de 4.52 puntos respecto a las calificaciones reales.

6. Predicción con nuevos valores.

En el módulo *Generar puntos* se ingresan valores de prueba $X_1 = 60$ y $X_2 = 4$, obteniéndose la predicción:

$$\hat{Y} = 83.9844046192.$$

Esto demuestra que la interfaz no solo estima el modelo, sino que también permite utilizarlo para pronosticar calificaciones en función de valores nuevos de las variables independientes.

En síntesis, la Figura (19) evidencia que la interfaz integra el flujo completo del análisis: ingreso de datos, cálculo de coeficientes, evaluación del ajuste mediante medidas estadísticas y generación de predicciones. Esto facilita la comprensión del procedimiento de regresión lineal múltiple y permite automatizar cálculos que, de forma manual, requieren múltiples operaciones algebraicas y estadísticas.

3.3 Validación de la interfaz gráfica con estudiantes de Ingeniería Civil

La interfaz gráfica de usuario en MATLAB fue aplicada a 20 estudiantes del segundo ciclo 2026-I de la Escuela Profesional de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo. La aplicación se realizó con la finalidad de validar el funcionamiento de la herramienta y evaluar su aporte al aprendizaje de regresión lineal.

Durante la intervención, los estudiantes utilizaron los módulos de regresión lineal simple y múltiple para ingresar datos, generar tablas de cálculo, observar gráficos, interpretar la pendiente, el intercepto, el coeficiente de determinación, los residuos y realizar predicciones. Antes de utilizar la interfaz se aplicó un pretest y, después de la intervención, se aplicó un postest.

Los resultados individuales obtenidos por los 20 estudiantes en el pretest y postest se presentan en el Anexo 1A y Anexo 1B, respectivamente.

3.3.1 Resultados del pretest de regresión lineal

El pretest permitió identificar el nivel inicial de aprendizaje de regresión lineal de los 20 estudiantes de Ingeniería Civil antes de utilizar la interfaz gráfica de usuario en MATLAB.

Tabla 7*Resultados descriptivos del pretest de regresión lineal*

Estadístico	Valor
Número de estudiantes	20
Promedio	11.15
Desviación estándar	1.95
Puntaje mínimo	7
Puntaje máximo	14

En la Tabla 7 se observa que los estudiantes obtuvieron un promedio de 11.15 puntos en el pretest. El puntaje mínimo fue 7 y el puntaje máximo fue 14. Estos resultados mostraron que, antes de utilizar la interfaz gráfica, los estudiantes presentaban dificultades en la comprensión e interpretación de algunos elementos de la regresión lineal, tales como la pendiente, el intercepto, el coeficiente de determinación, los residuos y las predicciones.

3.3.2 Resultados del postest de regresión lineal

El postest permitió evaluar el nivel de aprendizaje de regresión lineal después de la aplicación de la interfaz gráfica de usuario en MATLAB.

Tabla 8*Resultados descriptivos del postest de regresión lineal*

Estadístico	Valor
Número de estudiantes	20
Promedio	17.15
Desviación estándar	1.98
Puntaje mínimo	14
Puntaje máximo	20

En la Tabla 8 se observa que los estudiantes obtuvieron un promedio de 17.15 puntos después de utilizar la interfaz gráfica. El puntaje mínimo fue 14 y el puntaje máximo fue 20. Estos resultados evidenciaron un mayor dominio de los contenidos de regresión lineal en comparación con la evaluación inicial.

3.3.3 Comparación de los resultados del pretest y posttest

Tabla 9

Comparación de resultados del pretest y posttest

Indicador	Pretest	Posttest	Diferencia
Promedio	11.15	17.15	6.00
Desviación estándar	1.95	1.98	0.03
Puntaje mínimo	7	14	7
Puntaje máximo	14	20	6

La Tabla 9 muestra que el promedio del posttest fue superior al promedio del pretest. Los estudiantes pasaron de un promedio de 11.15 puntos a 17.15 puntos, lo cual representó un incremento de 6.00 puntos.

Asimismo, el puntaje mínimo aumentó de 7 a 14 puntos y el puntaje máximo aumentó de 14 a 20 puntos. Estos resultados permitieron observar una mejora general en el aprendizaje de regresión lineal después de la aplicación de la interfaz gráfica de usuario en MATLAB.

3.3.4 Prueba de normalidad

Antes de realizar el contraste de hipótesis, se evaluó la normalidad de las diferencias entre los puntajes del pretest y posttest mediante la prueba de Shapiro–Wilk.

Tabla 10

Prueba de normalidad de las diferencias entre pretest y posttest

Prueba	Estadístico	Sig.
Shapiro–Wilk	0.983	0.965

En la Tabla 10 se observa que el valor de significancia fue 0.965, valor mayor que 0.05. Por tanto, se consideró que las diferencias entre los puntajes del pretest y posttest presentaron una distribución aproximadamente normal.

En consecuencia, se aplicó la prueba t de Student para muestras relacionadas para contrastar la hipótesis de investigación.

3.3.5 Prueba *t* de Student para muestras relacionadas

Tabla 11

*Prueba *t* de Student para muestras relacionadas*

Comparación	Media de diferencia	<i>t</i>	gl	Sig.
Postest – Pretest	6.00	10.589	19	<0.001

La Tabla 11 muestra que el valor de significancia fue menor que 0.05 ($p < 0.001$). Por tanto, se rechazó la hipótesis nula y se aceptó la hipótesis de investigación.

En consecuencia, se concluyó que la aplicación de la interfaz gráfica de usuario en MATLAB mejoró significativamente el aprendizaje de regresión lineal en los 20 estudiantes del segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

Resultados del cuestionario de usabilidad

Los resultados del cuestionario de usabilidad mostraron una valoración favorable de los estudiantes respecto a la interfaz gráfica de usuario en MATLAB. Las respuestas se concentraron principalmente en las categorías “De acuerdo” y “Totalmente de acuerdo”, correspondientes a los valores 4 y 5 de la escala de valoración.

Los estudiantes manifestaron una percepción positiva sobre el diseño, facilidad de uso, funcionalidad, interactividad y utilidad de la interfaz para el aprendizaje de regresión lineal.

IV. DISCUSIÓN DE RESULTADOS

Discusión del objetivo específico 1: Analizar el modelo de regresión lineal simple y múltiple

El primer objetivo específico estuvo orientado a analizar los fundamentos de la regresión lineal simple y múltiple que fueron incorporados en la interfaz gráfica de usuario. Los resultados del desarrollo mostraron que la herramienta integró los principales componentes del análisis de regresión: ingreso de datos, cálculo de medias, estimación de la pendiente e intercepto, obtención de valores estimados, análisis de residuos, cálculo de sumas de cuadrados, coeficiente de determinación y error cuadrático medio.

Este resultado se relacionó con lo señalado por Corica y Ferrari (2022), quienes identificaron que la enseñanza de correlación y regresión suele presentarse mediante tareas aisladas, con poca oportunidad para interpretar resultados y relacionarlos con situaciones concretas. De forma similar, Telles (2022) indicó que una enseñanza centrada únicamente en algoritmos y procedimientos no favorece suficientemente el razonamiento estadístico ni el análisis crítico de los datos.

En respuesta a esta necesidad, la interfaz gráfica no se limitó a mostrar la ecuación final del modelo, sino que presentó el proceso completo de construcción e interpretación de la regresión lineal. Los estudiantes pudieron observar la relación entre los datos ingresados, la ecuación estimada, los valores ajustados, los residuos y las medidas globales de ajuste. Esto permitió que el tratamiento de la regresión lineal se vinculara con la comprensión de los resultados y no únicamente con la aplicación mecánica de fórmulas.

Asimismo, el uso de MATLAB resultó pertinente porque permitió integrar cálculo, visualización y análisis estadístico dentro de un mismo entorno. Este planteamiento coincidió con Ogbonnaya (2022), quien destacó que MATLAB puede apoyar el aprendizaje de estadística en estudiantes de áreas afines a ingeniería cuando se utiliza para combinar procedimientos computacionales con interpretación de resultados.

Discusión del objetivo específico 2: Diseñar y programar una interfaz gráfica de usuario en MATLAB

El segundo objetivo específico consistió en diseñar y programar una interfaz gráfica de usuario en MATLAB para el ingreso de datos, cálculo de parámetros, visualización de gráficos e interpretación de resultados de regresión lineal.

Como resultado, se desarrollaron dos módulos: uno para regresión lineal simple y otro para regresión lineal múltiple. La interfaz permitió ingresar datos, generar tablas de cálculo, estimar coeficientes, visualizar ecuaciones, representar gráficamente los resultados, calcular residuos, determinar medidas de ajuste y generar predicciones con nuevos valores de entrada.

Este resultado guardó relación con Tan et al. (2024), quienes señalaron que las interfaces gráficas basadas en MATLAB favorecen la comprensión de conceptos abstractos cuando incorporan interacción y representación visual. También coincidió con Momox y Alonso (2023), quienes encontraron que las interfaces desarrolladas en MATLAB permitieron integrar teoría y práctica, facilitando el aprendizaje mediante procesos de automatización y análisis.

La estructura organizada de la interfaz permitió separar el proceso de trabajo en etapas: ingreso de información, cálculo de parámetros, presentación de resultados, visualización gráfica y predicción. Esta organización facilitó que los estudiantes siguieran la secuencia lógica del análisis de regresión lineal.

Además, la presentación simultánea de los valores observados, valores estimados, residuos y errores cuadrados permitió que el estudiante reconociera cómo se evaluaba el ajuste del modelo. El módulo de predicción permitió relacionar la regresión con situaciones aplicadas, debido a que mostró la utilidad del modelo para estimar valores futuros o desconocidos a partir de nuevas variables de entrada.

Por tanto, el diseño de la interfaz no respondió solamente a una necesidad técnica de automatizar cálculos, sino también a una necesidad didáctica de organizar el proceso

de aprendizaje de regresión lineal de manera visual, secuencial e interactiva.

Discusión del objetivo específico 3: Evaluar el aporte de la interfaz gráfica al aprendizaje de regresión lineal

El tercer objetivo específico fue evaluar el aporte de la interfaz gráfica de usuario en MATLAB al aprendizaje de regresión lineal mediante su aplicación a 20 estudiantes del segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo.

Los resultados mostraron que el promedio del pretest fue de 11.15 puntos, mientras que el promedio del posttest fue de 17.15 puntos. Esto representó un incremento de 6.00 puntos después de la aplicación de la interfaz gráfica. Asimismo, el puntaje mínimo aumentó de 7 a 14 puntos y el puntaje máximo pasó de 14 a 20 puntos.

Estos resultados evidenciaron que, después de utilizar la interfaz gráfica, los estudiantes lograron un mejor desempeño en contenidos relacionados con la regresión lineal. La mejora observada se vinculó con la posibilidad de visualizar el procedimiento completo, interpretar la pendiente e intercepto, reconocer el coeficiente de determinación, analizar residuos y realizar predicciones.

Antes de aplicar la prueba de hipótesis, se evaluó la normalidad de las diferencias entre los puntajes de pretest y posttest mediante la prueba de Shapiro–Wilk. El resultado obtenido fue $p = 0,965$, valor mayor que 0,05, por lo que se consideró que las diferencias presentaron una distribución aproximadamente normal.

Posteriormente, se aplicó la prueba t de Student para muestras relacionadas. El resultado mostró una diferencia media de 6.00 puntos, con un valor de $t = 10,589$, $gl = 19$ y un nivel de significancia menor que 0,001. Debido a que el valor de significancia fue menor que 0,05, se rechazó la hipótesis nula y se aceptó la hipótesis de investigación.

Por tanto, se concluyó que la aplicación de la interfaz gráfica de usuario en MATLAB mejoró significativamente el aprendizaje de regresión lineal en los estudiantes de Ingeniería Civil participantes.

Este resultado coincidió con Salcedo (2024), quien encontró diferencias significativas entre el pretest y posttest después de aplicar tecnologías de la información y comunicación en la enseñanza de estadística. Del mismo modo, García (2022) señaló que las dificultades en el aprendizaje de regresión lineal se relacionaron con metodologías poco dinámicas, baja participación estudiantil y escasa conexión entre los procedimientos y situaciones aplicadas.

En el presente estudio, la interfaz gráfica aportó una alternativa para disminuir estas dificultades, debido a que permitió que los estudiantes trabajaran directamente con datos, observaran los resultados del modelo y relacionaran los cálculos con gráficos e interpretaciones estadísticas. Sin embargo, debido a que se empleó un diseño preexperimental con un solo grupo, los resultados deben interpretarse dentro del contexto de los 20 estudiantes participantes y no generalizarse automáticamente a toda la población universitaria.

Discusión del objetivo específico 4: Validar el funcionamiento técnico de la interfaz gráfica y generar un ejecutable independiente

El cuarto objetivo específico estuvo orientado a validar el funcionamiento técnico de la interfaz gráfica mediante la comparación de sus resultados con procedimientos manuales y funciones internas de MATLAB. Además, se consideró la conversión de la herramienta en un ejecutable independiente para facilitar su uso en contextos académicos.

Los resultados mostraron que la interfaz permitió calcular correctamente los coeficientes de regresión lineal simple y múltiple, valores estimados, residuos, sumas de cuadrados, coeficiente de determinación, error cuadrático medio y raíz del error cuadrático medio. Los resultados obtenidos mediante la GUI fueron contrastados con cálculos manuales y con funciones internas de MATLAB, verificándose consistencia en los valores generados.

Este resultado coincidió con los planteamientos de Ogbonnaya (2022), quien indicó

que MATLAB puede ser utilizado como herramienta para enseñar análisis estadístico, debido a que permite integrar programación, cálculo y representación de resultados. Asimismo, se relacionó con Vargas et al. (2023), quienes señalaron la necesidad de incorporar recursos tecnológicos y actividades prácticas para fortalecer la formación estadística de los estudiantes.

La conversión de la interfaz en un ejecutable independiente representó un aporte adicional, debido a que permitió reducir la dependencia de tener instalado MATLAB en cada equipo. Esta característica facilitó la distribución de la herramienta y amplió las posibilidades de uso en actividades académicas relacionadas con regresión lineal.

En ese sentido, la interfaz gráfica no quedó limitada a un prototipo de programación, sino que se convirtió en un recurso utilizable para actividades de enseñanza y aprendizaje en el contexto universitario. Los resultados del cuestionario de usabilidad mostraron que los estudiantes valoraron favorablemente el diseño, la funcionalidad, la interactividad y la utilidad de la interfaz gráfica para el aprendizaje de regresión lineal.

V. CONCLUSIONES

1. Se analizaron e incorporaron en la interfaz gráfica los principales fundamentos de la regresión lineal simple y múltiple, incluyendo pendiente, intercepto, valores estimados, residuos, sumas de cuadrados, coeficiente de determinación, error cuadrático medio y raíz del error cuadrático medio. Estos elementos permitieron presentar de manera ordenada el procedimiento de estimación e interpretación del modelo de regresión lineal.
2. Se diseñó y programó una interfaz gráfica de usuario en MATLAB R2021a para el aprendizaje de regresión lineal en la educación universitaria. La herramienta integró módulos de regresión lineal simple y múltiple, ingreso de datos, tablas de cálculo, representación gráfica, estimación de parámetros, métricas de ajuste y predicciones. Su funcionamiento técnico fue validado mediante la comparación con procedimientos manuales y funciones internas de MATLAB, obteniéndose resultados consistentes. Además, su conversión en un ejecutable independiente permitió utilizarla sin requerir la instalación de MATLAB R2021a.
3. La aplicación de la interfaz gráfica en 20 estudiantes del segundo ciclo 2026-I de Ingeniería Civil de la Universidad Nacional Pedro Ruiz Gallo mostró una mejora en el aprendizaje de regresión lineal. El promedio aumentó de 11.15 puntos en el pretest a 17.15 puntos en el posttest, con una diferencia de 6.00 puntos. La prueba (t) de Student para muestras relacionadas confirmó que esta diferencia fue estadísticamente significativa (($t=10.589$), ($gl=19$), ($p<0.001$)), por lo que se concluyó que el uso de la interfaz gráfica en MATLAB mejoró significativamente el aprendizaje de regresión lineal en los estudiantes participantes.

RECOMENDACIONES

1. Nuestra interfaz gráfica de usuario en MATLAB se debe utilizar como recurso de apoyo en asignaturas universitarias relacionadas con estadística, análisis de datos, métodos cuantitativos e investigación aplicada, debido a que facilita la visualización y comprensión de los procedimientos de regresión lineal simple y múltiple.
2. Es importante capacitar a los docentes en el manejo y uso de esta interfaz, de modo que puedan incorporarla como recurso tecnológico durante sus clases y fortalecer el proceso de enseñanza de la regresión lineal.

Referencias Bibliográficas

- Burden, R. & Faires, J. (2001). *Análisis Numérico*. 7ª ed. Cengage.
- Kharab, A. & Guenther, R. (2012). *An introduction to numerical methods a Matlab approach*. CRC Pres, New York, USA.
- Cox, F., González, D., Magreñán, Á., & Orcos, L. (2022). *Enseñanza de estadística descriptiva mediante el uso de simuladores y laboratorios virtuales en la etapa universitaria. Bordón*. Revista de Pedagogía, 74(4). <https://n9.cl/j3x3k>
- Corica, A. & Ferrari, C. (2022). enseñanza de correlación y regresión en la formación de profesores de Matemática: Un estudio exploratorio. *Didacticae: Revista de Investigación en Didácticas Específicas*, 11, 131-152. <https://n9.cl/1f820>
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2021). *Regression: Models, Methods and Applications* (2nd ed.). Springer.
- Garcia, A. (2022). *Propuesta metodológica para el análisis de regresión lineal simple en los estudiantes de la carrera de marketing de un instituto superior privado de Lima* [Tesis de Maestría, Universidad San Ignacio de Loyola]. <https://n9.cl/eabfi>
- International Organization for Standardization. (2018). *ISO 9241-11:2018 Ergonomics of human-system interaction?Part 11: Usability: Definitions and concepts*. ISO.
- ISTQB. (2024). *Certified Tester Foundation Level (CTFL) Syllabus*. International Software Testing Qualifications Board.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R* (2nd ed.). Springer.
- MathWorks. (2024a). *App Designer* (Documentation). The MathWorks, Inc.
- MathWorks. (2024b). *MATLAB Compiler* (Documentation). The MathWorks, Inc.
- MathWorks. (2024c). *MATLAB Runtime* (Documentation). The MathWorks, Inc.
- Montgomery, D., Peck, E., & Vining, G. (2021). *Introduction to Linear Regression Analysis* (6th ed.). Wiley.

- Momox, E., & Alonso, L. (2023). *A MATLAB GUI for Engineering Education in the Undergraduate Laboratory*. International Journal of Information and Education Technology, 13(5), 861-866. <https://n9.cl/2a1xk>
- Montoya, I. (2022). *El análisis de regresión lineal y la precisión del pronóstico de demanda de productos hidrobiológicos congelados en el Perú* [Tesis de Maestría, Universidad Nacional Federico Villarreal]. <https://n9.cl/3uw4uh>
- Ogbonnaya, C. (2022). *Teaching Statistics With MATLAB: How and Why?*. <https://doi.org/10.52041/iase.icots11.T10E2>
- Osorio, E., Inzunza, S., & Ward, S. (2023). Modelación estadística para el aprendizaje de la correlación y regresión lineal. PNA, 17(3), 295-321. <https://n9.cl/mgo10>
- Pinargote, E., Falconi, G., Mosquera, G., Bolaños, G., & Erraes, S. (2021). *Diseño de un Modelo de Regresión Lineal Múltiple Para Predecir el Rendimiento de Estudiantes de Institutos Superiores Tecnológicos Públicos Frente a la Nueva Normalidad*. ResearchGate, 17(2), 30-55. <https://n9.cl/8oza9>
- Ramos, L. (2019). La educación estadística en el nivel universitario: Retos y oportunidades. Revista Digital de Investigación en Docencia Universitaria, 13(2), 67-82. <https://doi.org/10.19083/ridu.2019.1081>
- Salcedo, M. (2024). *TIC en la enseñanza aprendizaje de la estadística descriptiva en estudiantes de un instituto tecnológico de Huancayo* [Tesis de Licenciatura, Universidad Nacional de Huancavelica]. <https://n9.cl/lfa0u>
- Tan, X., Chan, K., & Lim, S. (2024). *Development of a MATLAB-based graphical user interface to illustrate Coulomb's law and Gauss's law*. Computer Applications in Engineering Education, n/a(n/a), e22759. <https://doi.org/10.1002/cae.22759>
- Telles, F. (2022). *La enseñanza de la estadística en el contexto de la sociedad del dato: Desafíos y reflexiones*. Journal de Ciencias Sociales, 1(18), Article 18. <https://doi.org/10.18682/jcs.vi18.4338>

- Vargas, L., Ávila, J., Escorcía, K., & Molina, C. (2023). *Importancia de la enseñanza de la estadística y limitaciones en su formación desde los significados de los profesores de matemáticas en formación*. PANORAMA, 17(32). <https://www.redalyc.org/journal/3439/343975993007/html/>
- Vides, S., Barros, J., & Triana, G. (2021). *Modelo para el proceso de enseñanza aprendizaje de la asignatura estadística en estudiantes de ingeniería de la Universidad Popular del Cesar (Colombia)*. Espacios, 42(03), 10-26. <https://n9.cl/s7z6y>

Anexos

Programa para regresión lineal simple

```
function varargout = REGRESION_LINEAL_SIMPLE(varargin)

% Last Modified by GUIDE v2.5 06-Jan-2026 17:51:29

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',   gui_Singleton, ...
                  'gui_OpeningFcn',   @REGRESION_LINEAL_SIMPLE_OpeningFcn, ...
                  'gui_OutputFcn',    @REGRESION_LINEAL_SIMPLE_OutputFcn, ...
                  'gui_LayoutFcn',    [] , ...
                  'gui_Callback',     []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State,
                                         varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT
```

```

% -----
function excel_2_Callback(hObject, eventdata, handles)

[archivo, ruta] = uigetfile({'*.xlsx;*.xls;*.csv',
'Archivos (*.xlsx,*.xls,*.csv)'}, ...
'Seleccione archivo con datos');

if isequal(archivo,0), return; end
fullpath = fullfile(ruta, archivo);

% ===== Leer archivo =====
try
    M = readmatrix(fullpath);
catch
    try
        M = xlsread(fullpath);
    catch
        errordlg('No se pudo leer el archivo.','Error');
        return;
    end
end

if isempty(M)
    errordlg('El archivo está vacío.','Error');
    return;
end

```

```

% Quitar filas completamente vacías (NaN)
rowsValid = ~all(isnan(M),2);
M = M(rowsValid,:);

if isempty(M) || size(M,2) < 2
    errordlg('El archivo debe tener al menos 2 columnas:
        [Y, X...].','Error');
    return;
end

fil = size(M,1);
col = size(M,2);

% ===== Actualizar edit1 y edit2 =====
if isfield(handles,'edit1') && ishghandle(handles.edit1)
    set(handles.edit1,'String', num2str(fil));
end
if isfield(handles,'edit2') && ishghandle(handles.edit2)
    set(handles.edit2,'String', num2str(col));
end

p = col - 1;
if isfield(handles,'edit4') && ishghandle(handles.edit4)
    set(handles.edit4,'String', num2str(p));
end

```

```

% ===== Pasar datos a uitable1 =====

set(handles.uitable1,'Data', num2cell(M));
set(handles.uitable1,'ColumnEditable', true(1,col));

% ? NO DESAPARECE la enumeración: fuerza 1..fil
set(handles.uitable1,'RowName', arrayfun(@num2str, 1:fil,
'UniformOutput', false));

set(handles.uitable1,'ColumnFormat', repmat({'numeric'},
1,col));

% ColumnName: Y, X1..Xp
cn = cell(1,col);
cn{1} = 'Y';
for k = 2:col
    cn{k} = sprintf('X%d', k-1);
end
set(handles.uitable1,'ColumnName', cn);

set(handles.uitable1,'Visible','on');

guidata(hObject, handles);

% --- Executes on button press in calcular.
function calcular_Callback(hObject, eventdata, handles)

if ~isfield(handles,'uitable1') ||
~ishandle(handles.uitable1)

```

```

        errordlg('No existe uitable1 (tabla de
        entrada).','Error'); return;
end
if ~isfield(handles,'uitable2') ||
~ishghandle(handles.uitable2)
    errordlg('No existe uitable2 (tabla de
    salida).','Error'); return;
end
if ~isfield(handles,'uitable3') ||
~ishghandle(handles.uitable3)
    errordlg('No existe uitable3 (tabla de
    coeficientes).','Error'); return;
end

% ===== 1) Leer datos desde uitable1 =====
raw = get(handles.uitable1,'Data');
if isempty(raw) || size(raw,2) < 2
    errordlg('uitable1 debe tener 2 columnas: X y Y.','Datos
    insuficientes');
    return;
end

n0 = size(raw,1);
X = nan(n0,1); Y = nan(n0,1);

for k = 1:n0
    vx = raw{k,1}; vy = raw{k,2};

```

```

    if isnumeric(vx), X(k) = vx;
    elseif ischar(vx) || isstring(vx), X(k) = str2double(vx);
    end

    if isnumeric(vy), Y(k) = vy;
    elseif ischar(vy) || isstring(vy), Y(k) = str2double(vy);
    end
end

% Limpiar filas vacías
valid = ~(isnan(X) | isnan(Y));
X = X(valid); Y = Y(valid);

if numel(X) < 2
    error(dlg('Ingresa al menos 2 pares (X,Y) válidos en
    uitable1.','Datos insuficientes'));
    return;
end

X = X(:); Y = Y(:);
n = numel(X);

% ===== 2) Cálculos de mínimos cuadrados =====
xbar = mean(X);
ybar = mean(Y);

dx = X - xbar;
dy = Y - ybar;

```

```

dxdy = dx .* dy;
dx2   = dx.^2;

Sxy = sum(dxdy);
Sxx = sum(dx2);

% ===== 3) Modelo =====
b1 = Sxy / Sxx;           % pendiente
b0 = ybar - b1*xbar;      % intercepto

Yhat = b0 + b1*X;         % ?
res   = Y - Yhat;         % e

% ===== 4) STC / SCR / SCE + métricas =====
STC = sum((Y - ybar).^2);  % ?(y-?)^2
SCE = sum(res.^2);         % ?(y-?)^2
SCR = sum((Yhat - ybar).^2); % ?(??-?)^2

if STC == 0
    R2 = NaN;
else
    R2 = SCR / STC;
end

if n > 2
    MSE = SCE / (n - 2);
    RMSE = sqrt(MSE);
else

```



```

        MSE    = NaN;
        RMSE   = NaN;

end

% ===== 5) Llenar uitable2 (extendida) =====
data2 = cell(n + 2 + 6, 10);

for i = 1:n
    data2{i,1} = i;
    data2{i,2} = X(i);
    data2{i,3} = Y(i);
    data2{i,4} = dx(i);
    data2{i,5} = dy(i);
    data2{i,6} = dxdy(i);
    data2{i,7} = dx2(i);
    data2{i,8} = Yhat(i);
    data2{i,9} = res(i);
    data2{i,10} = res(i)^2;
end

% Fila medias
r = n+1;
data2{r,1} = 'x?', '?';
data2{r,2} = xbar;
data2{r,3} = ybar;

% Fila sumas
r = n+2;

```

```

data2{r,1} = '?';
data2{r,6} = Sxy;
data2{r,7} = Sxx;
data2{r,10} = SCE;

set(handles.uitable2,'Data', data2);
set(handles.uitable2,'RowName', []);
set(handles.uitable2,'ColumnEditable', false(1,10));
set(handles.uitable2,'ColumnName', { ...
    'i','x','y', ...
    '(x - x?)','(y - ?)', ...
    '(x - x?)(y - ?)','(x - x?)^2', ...
    '? = b0 + b1·x','e = y - ?','e^2' ...
});

% ===== 6) Llenar uitable3 (SIN ecuación) =====
coefData = {
    'x?', xbar
    '?', ybar
    'Sxy = sum(x-xbar)(y-?)', Sxy
    'Sxx = sum(x-xbar)^2', Sxx
    'Pendiente (b1)', b1
    'Intercepto (b0)', b0
    'STC = sum(y-ybar)^2', STC
    'SCR = sum(yp-ybar)^2', SCR
    'SCE = sum(y-yp)^2', SCE
    'R^2 = SCR/STC', R2
    'MSE = SCE/(n-2)', MSE

```

```

        'RMSE = ?MSE',          RMSE
    };

    set(handles.uitable3,'Data', coefData);
    set(handles.uitable3,'RowName', []);
    set(handles.uitable3,'ColumnName', {'Parámetro','Valor'});
    set(handles.uitable3,'ColumnEditable', [false false]);

    % ===== 7) Ecuación en edit3 (NO en uitable3) =====
    if isfield(handles,'edit3') && ishghandle(handles.edit3)
        ecuacion = sprintf('? = %.10f + %.10f x', b0, b1);
        set(handles.edit3,'String', ecuacion);
    end

    % ===== 8) Gráfica en axes1 =====
    if isfield(handles,'axes1') && ishghandle(handles.axes1)
        axes(handles.axes1); %#ok<LAXES>
        cla; grid on; hold on;

        plot(X, Y, 'o');

        Xs = sort(X);
        Yline = b0 + b1*Xs;
        plot(Xs, Yline, '-');

        xl = xlim; yl = ylim;
        plot([xbar xbar], yl, '--');
        plot(xl, [ybar ybar], '--');
    end

```

```

        xlabel('X'); ylabel('Y');

        title('Regresión lineal simple');

        legend('Datos','Recta de
regresión','x?','?','Location','best');

        hold off;
end

guidata(hObject, handles);

% --- Executes on button press in borrar.
function borrar_Callback(hObject, eventdata, handles)

function salir_Callback(hObject, eventdata, handles)

% --- Executes on button press in generar.
function generar_Callback(hObject, eventdata, handles)

fil = str2double(get(handles.edit1,'String')); % filas
col = str2double(get(handles.edit2,'String')); % columnas

% Validaciones

```

```

if isnan(fil) || fil < 1 || mod(fil,1) ~= 0
    errordlg('Ingrese un número entero válido de filas en
    edit1.','Error');
    return;
end

if isnan(col) || col < 1 || mod(col,1) ~= 0
    errordlg('Ingrese un número entero válido de columnas en
    edit2.','Error');
    return;
end

% Crear tabla vacía editable
set(handles.uitable1,'Data', cell(fil, col));
set(handles.uitable1,'ColumnEditable', true(1, col));
set(handles.uitable1,'RowName', []);

% ColumnName (opcional: si col=2, pon X y Y)
if col == 2
    set(handles.uitable1,'ColumnName', {'X','Y'});
else
    % nombres genéricos C1, C2, ...
    cn = arrayfun(@(k) sprintf('C%d',k), 1:col,
    'UniformOutput', false);
    set(handles.uitable1,'ColumnName', cn);
end

% ColumnFormat: una por columna

```

```

fmt = repmat({'numeric'}, 1, col);
set(handles.uitable1,'ColumnFormat', fmt);

set(handles.uitable1,'Visible','on');

guidata(hObject, handles);

% --- Executes on button press in prediccion.
function prediccion_Callback(hObject, eventdata, handles)

if ~isfield(handles,'uitable3') ||
~ishghandle(handles.uitable3)
    errordlg('No existe uitable3 (coeficientes).','Error');
    return;
end
if ~isfield(handles,'uitable4') ||
~ishghandle(handles.uitable4)
    errordlg('No existe uitable4 (valores X para
predicción).','Error'); return;
end
if ~isfield(handles,'edit5') || ~ishghandle(handles.edit5)

```

```

        errordlg('No existe edit5 (salida y_pred).','Error');
        return;
end

% (Opcional) donde se mostrará la ecuación con reemplazo
hasEdit4 = isfield(handles,'edit4') &&
ishghandle(handles.edit4);

% ===== 1) Leer coeficientes desde uitable3 =====
T = get(handles.uitable3,'Data');
if isempty(T) || size(T,2) < 2
    errordlg('uitable3 debe tener 2 columnas: [Parámetro,
    Valor].','Error');
    return;
end

% Convertidor robusto a número (acepta coma decimal)
toNum = @(v) ( ...
    (isnumeric(v) && ~isempty(v)) * double(v) + ...
    (~(isnumeric(v) && ~isempty(v))) *
    str2double(strrep(string(v),',','.')) ...
);

param = string(T(:,1));

% --- b0 (Intercepto) buscando "(b0)" o "Intercepto"
idx_b0 = find(~cellfun('isempty', regexp(param, '\
(\s*b0\s*\)', 'once'))), 1);

```

```

if isempty(idx_b0)
    idx_b0 = find(contains(param, "Intercepto", "IgnoreCase",
        true), 1);
end
if isempty(idx_b0)
    errordlg('No se encontró "Intercepto (b0)" en
        uitable3.', 'Error'); return;
end

b0 = toNum(T{idx_b0,2});
if isnan(b0)
    errordlg('El valor del Intercepto (b0) en uitable3 no es
        numérico.', 'Error');
    return;
end

% ===== 2) Leer X desde uitable4 =====
Xraw = get(handles.uitable4, 'Data');
if isempty(Xraw)
    errordlg('Ingrese valores en uitable4.', 'Error'); return;
end

% p filas = X1..Xp
p = size(Xraw,1);

X = nan(p,1);
for j = 1:p
    v = Xraw{j,1};

```



```

    if isnumeric(v)
        X(j) = double(v);
    elseif ischar(v) || isstring(v)
        X(j) = str2double(strrep(string(v),',',' ','.'));
    else
        X(j) = NaN;
    end
end

if any(isnan(X))
    error('En uitable4 hay valores vacíos o no
    numéricos.','Error');
    return;
end

% ===== 3) Leer b1..bp desde uitable3 =====
b = nan(p,1);

for j = 1:p
    patron = sprintf('\\(\\s*b%d\\s*\\)', j);
    idx = find(~cellfun('isempty', regexp(param, patron,
    'once'))), 1);

    % Alternativa por si no está "(b#)": busca "Xj"
    if isempty(idx)
        patron2 = sprintf('X%d', j);
        idx = find(contains(param, patron2, 'IgnoreCase',
        true), 1);
    end
end

```

```

end

if isempty(idx)
    errordlg(sprintf('No se encontró el coeficiente b%d
(X%d) en uitable3.', j, j), 'Error');
    return;
end

b(j) = toNum(T{idx,2});
if isnan(b(j))
    errordlg(sprintf('El coeficiente b%d en uitable3 no
es numérico.', j), 'Error');
    return;
end
end

% ===== 4) Calcular predicción =====
y_pred = b0 + sum(b .* X);

% ===== 5) Mostrar resultado en edit5 (10 decimales, sin
notación científica) =====
set(handles.edit5, 'String', sprintf('%.10f', y_pred));

% ===== 6) Mostrar punto predicho en axes1 =====
if isfield(handles, 'axes1') && ishghandle(handles.axes1)

    % Usamos X1 como eje horizontal para ubicar el punto
    x_plot = X(1);

```

```

axes(handles.axes1); %#ok<LAXES>

hold on; grid on;

% Dibujar punto predicho (rojo)
plot(x_plot, y_pred, 'ro', 'MarkerSize', 9, 'LineWidth',
2);

% Etiqueta opcional
text(x_plot, y_pred, sprintf(' ?=%.4f', y_pred), ...
'FontSize', 10, 'FontWeight', 'bold');

hold off;

end

guidata(hObject, handles);

% --- Executes on button press in pushbutton6.
function pushbutton6_Callback(hObject, eventdata, handles)

% ===== 0) Verificar componentes =====
if ~isfield(handles, 'uitable1') ||
~ishandle(handles.uitable1)
    errordlg('No existe uitable1.', 'Error'); return;
end

```

```

if ~isfield(handles,'edit1') || ~ishandle(handles.edit1)
    errordlg('No existe edit1 (filas).','Error'); return;
end

if ~isfield(handles,'edit2') || ~ishandle(handles.edit2)
    errordlg('No existe edit2 (columnas).','Error'); return;
end

% ===== 1) Leer filas y columnas =====
fil = str2double(get(handles.edit1,'String')); % filas
col = str2double(get(handles.edit2,'String')); % columnas totales = (Y + X.

% ===== 2) Validaciones =====
if isnan(fil) || fil < 1 || mod(fil,1) ~= 0
    errordlg('Ingrese un número entero válido de filas en
    edit1.','Error');
    return;
end

if isnan(col) || col < 2 || mod(col,1) ~= 0
    errordlg('Ingrese un número entero válido de columnas en
    edit2 (mínimo 2: Y y al menos un X).','Error');
    return;
end

% ===== 3) NO BORRAR datos: preservar lo que ya existe =====
oldData = get(handles.uitable1,'Data');
if isempty(oldData)
    oldData = cell(0,0);
end

```

```

% Convertir a celda si vino como double
if isnumeric(oldData)
    oldData = num2cell(oldData);
end

[oldFil, oldCol] = size(oldData);

% Crear nueva matriz celda del tamaño solicitado
newData = cell(fil, col);

% Copiar lo que exista sin perderlo (zona común)
minFil = min(fil, oldFil);
minCol = min(col, oldCol);

if minFil > 0 && minCol > 0
    newData(1:minFil, 1:minCol) = oldData(1:minFil,
    1:minCol);
end

% Asignar data preservada
set(handles.uitable1,'Data', newData);

% ===== 4) Propiedades de edición/formato =====
set(handles.uitable1,'ColumnEditable', true(1, col));

% ? MOSTRAR NÚMEROS DE FILA (1..fil)
set(handles.uitable1,'RowName', cellstr(num2str((1:fil)')));

```

```

set(handles.uitable1,'ColumnFormat', repmat({'numeric'}, 1,
col));

% ===== 5) ColumnName: primera Y, resto X1.. =====
cn = cell(1,col);
cn{1} = 'Y';
for k = 2:col
    cn{k} = sprintf('X%d', k-1);
end
set(handles.uitable1,'ColumnName', cn);
set(handles.uitable1,'Visible','on');

p = col - 1;    % igual se calcula p para crear uitable4

% ===== 7) Activar/crear uitable4 (1 columna, filas = p,
RowName = X1..Xp) =====
if isfield(handles,'uitable4') &&
ishghandle(handles.uitable4)

    old4 = get(handles.uitable4,'Data');
    if isempty(old4)
        old4 = cell(0,1);
    end
    if isnumeric(old4)
        old4 = num2cell(old4);
    end
end

```

```

end

% La tabla debe ser p filas x 1 columna
new4 = cell(p,1);
min4 = min(p, size(old4,1));
if min4 > 0
    new4(1:min4,1) = old4(1:min4,1);
end

rowNames4 = arrayfun(@(k) sprintf('X%d',k), 1:p,
'UniformOutput', false);

set(handles.uitable4,'Data', new4);
set(handles.uitable4,'RowName', rowNames4);      % X1,
X2, ...
set(handles.uitable4,'ColumnName', {'Valor'});    % 1 sola
columna
set(handles.uitable4,'ColumnEditable', true(1,1));
set(handles.uitable4,'ColumnFormat', {'numeric'});
set(handles.uitable4,'Visible','on');
end

guidata(hObject, handles);

function formulas_Callback(hObject, eventdata, handles)

```

regression

Programa para regresión lineal múltiple

```
function varargout = REGRESION_LINEAL_MULTIPLE(varargin)

% Last Modified by GUIDE v2.5 05-Jan-2026 18:45:33

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',   gui_Singleton, ...
                  'gui_OpeningFcn',   @REGRESION_LINEAL_MULTIPLE_OpeningFcn, ...
                  'gui_OutputFcn',    @REGRESION_LINEAL_MULTIPLE_OutputFcn, ...
                  'gui_LayoutFcn',    [] , ...
                  'gui_Callback',     []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State,
                                         varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT
```

```

% -----
function excel_2_Callback(hObject, eventdata, handles)

[archivo, ruta] = uigetfile({'*.xlsx;*.xls;*.csv','Archivos
(*.xlsx,*.xls,*.csv)'} , ...
                           'Seleccione archivo con datos');

if isequal(archivo,0), return; end
fullpath = fullfile(ruta, archivo);

% Leer archivo
try
    M = readmatrix(fullpath);
catch
    try
        M = xlsread(fullpath);
    catch
        errordlg('No se pudo leer el archivo.','Error');
        return;
    end
end

if isempty(M)
    errordlg('El archivo está vacío.','Error');
    return;
end

% Quitar filas completamente vacías (NaN)

```

```

rowsValid = ~all(isnan(M),2);
M = M(rowsValid,:);

fil = size(M,1);
col = size(M,2);

% Actualizar edit1 y edit2 automáticamente
set(handles.edit1,'String', num2str(fil));
set(handles.edit2,'String', num2str(col));

% Pasar datos a uitable1 (como celdas para edición)
set(handles.uitable1,'Data', num2cell(M));
set(handles.uitable1,'ColumnEditable', true(1,col));

% ? MOSTRAR NUMERACIÓN DE FILAS (NO OCULTAR)
set(handles.uitable1,'RowName', cellstr(num2str((1:fil)')));
% <- 1,2,3,4...

% Formato numérico
set(handles.uitable1,'ColumnFormat', repmat({'numeric'},
1,col));

% Nombres de columnas: 1ra Y, resto X1..Xp
cn = cell(1,col);
cn{1} = 'Y';
for k = 2:col
    cn{k} = sprintf('X%d', k-1);
end

```

```

set(handles.uitable1,'ColumnName', cn);

set(handles.uitable1,'Visible','on');

% (Opcional) llenar edit5 con # variables X (p = col-1)
p = col - 1;
if isfield(handles,'edit5') && ishghandle(handles.edit5)
    set(handles.edit5,'String', num2str(p));
end

guidata(hObject, handles);
% --- Executes on button press in calcular.

function calcular_Callback(hObject, eventdata, handles)

if ~isfield(handles,'uitable1') ||
~ishghandle(handles.uitable1)
    errordlg('No existe uitable1 (tabla de
    entrada).','Error'); return;
end
if ~isfield(handles,'uitable2') ||
~ishghandle(handles.uitable2)
    errordlg('No existe uitable2 (tabla de
    salida).','Error'); return;
end
if ~isfield(handles,'uitable3') ||
~ishghandle(handles.uitable3)

```

```

        errordlg('No existe uitable3 (tabla de
        coeficientes).','Error'); return;
end

% ===== 1) Leer datos desde uitable1 =====
raw = get(handles.uitable1,'Data');

if isempty(raw) || size(raw,2) < 2
    errordlg('uitable1 debe tener al menos 2 columnas: [Y,
    X...].','Datos insuficientes');
    return;
end

[n0, m] = size(raw);          % m = 1 + p
p = m - 1;                    % nº predictores

% Convertir todo a numérico (admite celdas con números o
texto)
A = nan(n0, m);
for i = 1:n0
    for j = 1:m
        v = raw{i,j};
        if isnumeric(v)
            A(i,j) = v;
        elseif ischar(v) || isstring(v)
            A(i,j) = str2double(v);
        else
            A(i,j) = NaN;
        end
    end
end

```

```

        end
    end
end

% Quitar filas con NaN en cualquier columna
valid = all(~isnan(A), 2);
A = A(valid,:);

if size(A,1) < (p + 2)
    errordlg(sprintf('Datos insuficientes: se requiere n >=
        p+2.\nActualmente n=%d y p=%d.', size(A,1), p), ...
        'Datos insuficientes');
    return;
end

% ===== 2) Separar Y y X (Y es la 1ra columna) =====
Y = A(:,1);
X = A(:,2:end);
n = size(X,1);

% Matriz de diseño con intercepto
X1 = [ones(n,1), X];

% ===== 3) Estimar coeficientes beta (mínimos cuadrados)
=====
r = rank(X1);
if r < size(X1,2)
    warndlg('Advertencia: colinealidad (X''X singular o casi

```

```

        singular). Se usará pinv().', 'Colinealidad');
        beta = pinv(X1) * Y;
    else
        beta = (X1' * X1) \ (X1' * Y);
    end

    Yhat = X1 * beta;
    e     = Y - Yhat;

    % ===== 4) Descomposición de la variación =====
    ybar = mean(Y);

    STC = sum((Y - ybar).^2);           % SST
    SCE = sum(e.^2);                   % SSE
    SCR = sum((Yhat - ybar).^2);        % SSR

    % ===== 5) Métricas =====
    if STC == 0
        R2 = NaN;
    else
        R2 = 1 - (SCE/STC);             % = SCR/STC
    end

    % Coeficiente de correlación múltiple (como la foto): R =
    sqrt(R^2)
    if isnan(R2)
        R_pos = NaN;
    else

```

```

        R_pos = sqrt(max(R2,0));
end

gl_total = n - 1;
gl_err   = n - p - 1;

if gl_err <= 0
    MSE = NaN;
    RMSE = NaN;
    R2_adj = NaN;
else
    MSE = SCE / gl_err;
    RMSE = sqrt(MSE);

    if STC == 0
        R2_adj = NaN;
    else
        R2_adj = 1 - ((SCE/gl_err) / (STC/gl_total));
    end
end

end

verif = STC - (SCR + SCE); % debería ser ~0

% ===== 6) Formato a 10 decimales (para uitable) =====
fmt10 = @(v) sprintf('%.10f', v);

% ===== 7) Llenar uitable2 (tabla completa) + PROMEDIO y ?
=====

```



```

% Columnas: i | X1..Xp | Y | ? | e | e^2

% Promedios y sumas (solo lo que estás mostrando)
meanX = mean(X,1);
sumX   = sum(X,1);

meanY = mean(Y);
sumY   = sum(Y);

sumE2 = sum(e.^2);

% Crear tabla con 2 filas extra
data2 = cell(n+2, 1 + p + 4);

% Filas de datos
for i = 1:n
    data2{i,1} = i;
    for j = 1:p
        data2{i,1+j} = fmt10(X(i,j));
    end
    data2{i,1+p+1} = fmt10(Y(i));
    data2{i,1+p+2} = fmt10(Yhat(i));
    data2{i,1+p+3} = fmt10(e(i));
    data2{i,1+p+4} = fmt10(e(i)^2);
end

% Fila Promedio (solo Y? y X?)
data2{n+1,1} = 'Promedio';

```

```

for j = 1:p
    data2{n+1,1+j} = fmt10(meanX(j));
end
data2{n+1,1+p+1} = fmt10(meanY);

% Fila ? (sumas que estás mostrando)
data2{n+2,1} = '?';
for j = 1:p
    data2{n+2,1+j} = fmt10(sumX(j));
end
data2{n+2,1+p+1} = fmt10(sumY);
data2{n+2,1+p+4} = fmt10(sumE2);

% Encabezados
colNames = cell(1, 1+p+4);
colNames{1} = 'i';
for j = 1:p
    colNames{1+j} = sprintf('X%d', j);
end
colNames{1+p+1} = 'Y';
colNames{1+p+2} = '?';
colNames{1+p+3} = 'e = Y-?';
colNames{1+p+4} = 'e^2';

set(handles.uitable2,'Data',data2);
set(handles.uitable2,'RowName',[ ]);
set(handles.uitable2,'ColumnName',colNames);
set(handles.uitable2,'ColumnEditable',false(1,numel(colNames)));

```

```

% ===== 8) Llenar uitable3 (coeficientes + métricas) SIN
ecuación =====
coefData = {};

coefData(end+1,:) = {'n', num2str(n)};
coefData(end+1,:) = {'p (predictores)', num2str(p)};

coefData(end+1,:) = {'STC (SST)', fmt10(STC)};
coefData(end+1,:) = {'SCR (SSR)', fmt10(SCR)};
coefData(end+1,:) = {'SCE (SSE)', fmt10(SCE)};

coefData(end+1,:) = {'R^2',          fmt10(R2)};
coefData(end+1,:) = {'R = sqrt(R^2)', fmt10(R_pos)};
coefData(end+1,:) = {'R^2 ajustado',  fmt10(R2_adj)};
coefData(end+1,:) = {'MSE',          fmt10(MSE)};
coefData(end+1,:) = {'RMSE',         fmt10(RMSE)};
%coefData(end+1,:) = {'Verif: STC-(SCR+SCE)', fmt10(verif)};
% opcional

coefData(end+1,:) = {'Intercepto (b0)', fmt10(beta(1))};
for j = 1:p
    coefData(end+1,:) = {sprintf('Coeficiente (b%d) para
    X%d', j, j), fmt10(beta(1+j))};
end

set(handles.uitable3, 'Data', coefData);
set(handles.uitable3, 'RowName', []);

```

```

set(handles.uitable3,'ColumnName',{'Parámetro','Valor'});
set(handles.uitable3,'ColumnEditable',[false false]);

% ===== 9) Ecuación en edit3 (NO en uitable3) =====
eq = sprintf('? = %.10f', beta(1));
for j = 1:p
    eq = [eq, sprintf(' + (%.10f)X%d', beta(1+j), j)];
    %#ok<AGROW>
end

if isfield(handles,'edit3') && ishghandle(handles.edit3)
    set(handles.edit3,'String', eq);
end

% ===== 10) Gráfica en axes1: Y vs ? =====
if isfield(handles,'axes1') && ishghandle(handles.axes1)
    axes(handles.axes1); %#ok<LAXES>
    cla; grid on; hold on;

    plot(Y, Yhat, 'o');

    mn = min([Y; Yhat]);
    mx = max([Y; Yhat]);
    plot([mn mx],[mn mx], '-'); % línea ideal

    xlabel('Y real');
    ylabel('? predicho');
    title(sprintf('Regresión lineal múltiple: R^2=%.4f, R=%.

```

```

    4f', R2, R_pos));

    legend('Datos','Línea ideal','Location','best');

    hold off;

end

% ===== 11) (Opcional) Mostrar fórmula en axes2 =====
if isfield(handles,'axes2') && ishghandle(handles.axes2)
    axes(handles.axes2);
    cla; axis off;
    text(0.05, 0.60, '$$R=\sqrt{\text{Coeficiente de determinaci}\acute{o}n}=\sqrt{R^2}$$', ...
        'Interpreter','latex','FontSize',16,'Color',[0 0 1]);
    text(0.05, 0.20, sprintf('R = %.10f', R_pos), ...
        'FontSize',14,'Color',[0 0 1]);
end

guidata(hObject, handles);

% --- Executes on button press in prediccion.
function prediccion_Callback(hObject, eventdata, handles)

if ~isfield(handles,'uitable3') ||
~ishghandle(handles.uitable3)
    errordlg('No existe uitable3 (coeficientes).','Error');
    return;
end

```

```

if ~isfield(handles,'uitable4') ||
~ishghandle(handles.uitable4)
    errordlg('No existe uitable4 (valores de X).','Error');
    return;
end
if ~isfield(handles,'edit4') || ~ishghandle(handles.edit4)
    errordlg('No existe edit4 (salida).','Error'); return;
end

% ===== 1) Leer coeficientes desde uitable3 =====
t3 = get(handles.uitable3,'Data');
if isempty(t3) || size(t3,2) < 2
    errordlg('uitable3 debe tener 2 columnas: Parámetro |
Valor.','Error'); return;
end

% Obtener b0 y bj
b0 = NaN;
bjMap =
containers.Map('KeyType','double','ValueType','double'); % j
-> bj

for i = 1:size(t3,1)
    param = t3{i,1};
    val    = t3{i,2};

    if isstring(param), param = char(param); end
    if isempty(param), continue; end

```

```

% valor numérico
if isnumeric(val)
    v = val;
elseif ischar(val) || isstring(val)
    v = str2double(val);
else
    v = NaN;
end
if isnan(v), continue; end

% Intercepto (b0)
if contains(lower(param),'intercepto') ||
contains(lower(param),'(b0)')
    b0 = v;
    continue;
end

% Coeficiente (b#) para X#
tok = regexp(char(param), 'Coeficiente\s*\s*(b(\d+)\s*)\s*para\s*X(\d+)\s*', 'tokens', 'once');
if ~isempty(tok)
    j = str2double(tok{2}); % índice de Xj
    bjMap(j) = v;          % bj
end
end

if isnan(b0)

```

```

        errordlg('No se encontró "Intercepto (b0)" en
        uitable3.', 'Error'); return;
end
if bjMap.Count < 1
    errordlg('No se encontraron coeficientes b1, b2, ... en
    uitable3.', 'Error'); return;
end

p = bjMap.Count; % número de X

% ===== 2) Leer X desde uitable4 =====
raw4 = get(handles.uitable4, 'Data');
if isempty(raw4) || size(raw4, 2) < 1
    errordlg('uitable4 debe tener 1 columna con valores de
    X.', 'Error'); return;
end
if size(raw4, 1) < p
    errordlg(sprintf('uitable4 debe tener al menos %d filas
    (X1..X%d).', p, p), 'Error'); return;
end

x = nan(p, 1);
for j = 1:p
    vv = raw4{j, 1};
    if isnumeric(vv)
        x(j) = vv;
    elseif ischar(vv) || isstring(vv)
        x(j) = str2double(vv);
    end
end

```



```

        else
            x(j) = NaN;
        end
    end
end

if any(isnan(x))
    errordlg('Complete todos los valores X en
    uitable4.','Datos incompletos');
    return;
end

% ===== 3) Calcular y_pred =====
y_pred = b0;
for j = 1:p
    if ~isKey(bjMap, j)
        errordlg(sprintf('Falta el coeficiente para X%d en
        uitable3.',' j'),'Error');
        return;
    end
    y_pred = y_pred + bjMap(j)*x(j);
end

% ===== 4) Mostrar SOLO y_pred en edit4 (10 decimales, sin
notación científica) =====
set(handles.edit4,'String', sprintf('%.10f', y_pred));

guidata(hObject, handles);

```

```

% --- Executes on button press in genpuntos.
function genpuntos_Callback(hObject, eventdata, handles)

if ~isfield(handles,'uitable1') ||
~ishghandle(handles.uitable1)
    errordlg('No existe uitable1.','Error'); return;
end
if ~isfield(handles,'edit1') || ~ishghandle(handles.edit1)
    errordlg('No existe edit1 (filas).','Error'); return;
end
if ~isfield(handles,'edit2') || ~ishghandle(handles.edit2)
    errordlg('No existe edit2 (columnas).','Error'); return;
end

% ===== 1) Leer filas y columnas =====
fil = str2double(get(handles.edit1,'String')); % filas
col = str2double(get(handles.edit2,'String')); % columnas
totales = (Y + X...)

% ===== 2) Validaciones =====
if isnan(fil) || fil < 1 || mod(fil,1) ~= 0
    errordlg('Ingrese un número entero válido de filas en
    edit1.','Error');
    return;
end
if isnan(col) || col < 2 || mod(col,1) ~= 0

```

```

        errordlg('Ingrese un número entero válido de columnas en
        edit2 (mínimo 2: Y y al menos un X).','Error');
        return;
    end

% ===== 3) NO BORRAR datos: preservar lo que ya existe =====
oldData = get(handles.uitable1,'Data');
if isempty(oldData)
    oldData = cell(0,0);
end

% Convertir a celda si vino como double
if isnumeric(oldData)
    oldData = num2cell(oldData);
end

[oldFil, oldCol] = size(oldData);

% Crear nueva matriz celda del tamaño solicitado
newData = cell(fil, col);

% Copiar lo que exista sin perderlo (zona común)
minFil = min(fil, oldFil);
minCol = min(col, oldCol);

if minFil > 0 && minCol > 0
    newData(1:minFil, 1:minCol) = oldData(1:minFil,
    1:minCol);

```

```

end

% Asignar data preservada
set(handles.uitable1,'Data', newData);

% ===== 4) Propiedades de edición/formato =====
set(handles.uitable1,'ColumnEditable', true(1, col));

% ? MOSTRAR NÚMEROS DE FILA (1..fil)
set(handles.uitable1,'RowName', cellstr(num2str((1:fil)')));

set(handles.uitable1,'ColumnFormat', repmat({'numeric'}, 1,
col));

% ===== 5) ColumnName: primera Y, resto X1.. =====
cn = cell(1,col);
cn{1} = 'Y';
for k = 2:col
    cn{k} = sprintf('X%d', k-1);
end
set(handles.uitable1,'ColumnName', cn);
set(handles.uitable1,'Visible','on');

% ===== 6) Auto-llenar edit5 con p = col-1 =====
p = col - 1;
if isfield(handles,'edit5') && ishghandle(handles.edit5)
    set(handles.edit5,'String', num2str(p));
end

```

```

% ===== 7) Activar/crear uitable4 (1 columna, filas = p,
RowName = X1..Xp) =====
if isfield(handles,'uitable4') &&
ishghandle(handles.uitable4)

    % Preservar datos previos en uitable4 (si existieran) sin
    borrarlos
    old4 = get(handles.uitable4,'Data');
    if isempty(old4)
        old4 = cell(0,1);
    end
    if isnumeric(old4)
        old4 = num2cell(old4);
    end

    % La tabla debe ser p filas x 1 columna
    new4 = cell(p,1);
    min4 = min(p, size(old4,1));
    if min4 > 0
        new4(1:min4,1) = old4(1:min4,1);
    end

    rowNames4 = arrayfun(@(k) sprintf('X%d',k), 1:p,
    'UniformOutput', false);

    set(handles.uitable4,'Data', new4);
    set(handles.uitable4,'RowName', rowNames4);           % X1,

```

```

X2, ...

set(handles.uitable4,'ColumnName', {'Valor'});    % 1 sola
columna

set(handles.uitable4,'ColumnEditable', true(1,1));
set(handles.uitable4,'ColumnFormat', {'numeric'});
set(handles.uitable4,'Visible','on');

end

guidata(hObject, handles);

```

ANEXO 1A. PRETEST SOBRE REGRESIÓN LINEAL

Título de la investigación: Desarrollo de una Interfaz Gráfica de Usuario en MATLAB para Potenciar el Aprendizaje de Regresión Lineal en la Educación Universitaria.

Objetivo: Medir el nivel inicial de aprendizaje de regresión lineal antes del uso de la interfaz gráfica de usuario en MATLAB.

Momento de aplicación: Antes de la intervención.

Condición de aplicación: El estudiante resolverá la prueba con los conocimientos adquiridos mediante enseñanza tradicional, sin utilizar la interfaz gráfica.

Duración: 40 minutos.

Puntaje total: 20 puntos.

Instrucciones: Lea atentamente cada pregunta. Marque la alternativa correcta o desarrolle el procedimiento solicitado.

I. Comprensión conceptual

1. ¿Qué estudia la regresión lineal?
 - a) La comparación de dos promedios.
 - b) La relación entre una variable dependiente y una o más variables independientes.
 - c) La clasificación de datos cualitativos.
 - d) La elaboración de tablas de frecuencia.
2. ¿Qué representa la variable dependiente en un modelo de regresión lineal?
 - a) La variable que se desea explicar o predecir.
 - b) La variable que siempre toma el valor cero.
 - c) La variable que no participa en el modelo.
 - d) La variable que solo se usa para ordenar datos.
3. En el modelo $\hat{y} = b_0 + b_1x$, ¿qué representa b_0 ?

- a) La pendiente de la recta.
 - b) El intercepto del modelo.
 - c) El residuo del modelo.
 - d) El coeficiente de determinación.
4. En el modelo $\hat{y} = b_0 + b_1x$, ¿qué representa b_1 ?
- a) El intercepto del modelo.
 - b) La pendiente del modelo.
 - c) El error total.
 - d) El tamaño de la muestra.
5. ¿Cuál es la principal diferencia entre regresión lineal simple y regresión lineal múltiple?
- a) La regresión simple usa una variable dependiente y la múltiple no.
 - b) La regresión simple usa una variable independiente y la múltiple usa dos o más variables independientes.
 - c) La regresión múltiple solo utiliza variables cualitativas.
 - d) La regresión simple no permite realizar predicciones.

II. Aplicación procedimental

Considere los siguientes datos:

x	1	2	3	4	5
y	2	4	5	4	6

6. ¿Cuál es la media aritmética de los valores de x ?
7. ¿Cuál es la media aritmética de los valores de y ?

8. ¿Cuál es el valor de $S_{xx} = \sum (x_i - \bar{x})^2$?
9. ¿Cuál es el valor de $S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$?
10. ¿Cuál es la pendiente del modelo, si se utiliza la fórmula $b_1 = \frac{S_{xy}}{S_{xx}}$?
11. ¿Cuál es el intercepto del modelo, si se utiliza la fórmula $b_0 = \bar{y} - b_1\bar{x}$?
12. ¿Cuál es la ecuación estimada de regresión lineal simple?

III. Interpretación de resultados

13. Si el modelo obtenido fuera $\hat{y} = 1,8 + 0,8x$, ¿qué significa la pendiente 0,8?
14. Si el coeficiente de determinación fuera $R^2 = 0,85$, ¿cómo se interpretaría este resultado?
15. Si el valor observado es $y_i = 10$ y el valor estimado es $\hat{y}_i = 8$, ¿cuál sería el residuo?
16. Si un residuo es positivo, ¿qué significa?
 - a) Que el valor observado es mayor que el valor estimado.
 - b) Que el valor observado es menor que el valor estimado.
 - c) Que el modelo no tiene error.
 - d) Que la pendiente es igual a cero.

IV. Supuestos del modelo

17. ¿Qué indica el supuesto de linealidad en un modelo de regresión lineal?
- a) Que los datos deben presentar una relación aproximadamente lineal.
 - b) Que todos los datos deben ser iguales.
 - c) Que la variable dependiente debe ser cualitativa.
 - d) Que no deben existir residuos.
18. ¿Cómo se puede revisar la homocedasticidad de los residuos?
- a) Mediante el gráfico de residuos frente a valores ajustados.
 - b) Mediante una tabla de frecuencias.
 - c) Mediante el tamaño de la muestra.
 - d) Mediante la suma de los valores de x .
19. ¿Qué herramienta puede utilizarse para revisar la normalidad de los residuos?
- a) Histograma, gráfico Q-Q o prueba de Shapiro-Wilk.
 - b) Gráfico circular.
 - c) Tabla de multiplicar.
 - d) Diagrama de barras de frecuencias absolutas.
20. En regresión lineal múltiple, ¿qué permite detectar el factor de inflación de la varianza, conocido como VIF?
- a) La posible multicolinealidad entre variables independientes.
 - b) La media aritmética de la variable dependiente.
 - c) La moda de los residuos.
 - d) El valor máximo de la variable dependiente.

Tabla 12

Resultados del Pretest - Ingeniería Civil

Estudiante	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	NOTA
1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	13
2	0	1	0	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	14
3	1	0	1	0	0	0	1	0	0	0	0	1	0	1	1	0	0	1	1	0	8
4	1	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	1	0	1	0	7
5	0	1	0	0	1	0	0	1	1	1	0	1	1	1	0	1	1	0	0	0	10
6	1	1	1	1	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	0	10
7	1	1	0	1	0	0	0	0	1	1	1	1	1	0	0	1	1	1	0	1	12
8	0	0	0	0	1	1	1	0	0	1	0	0	0	1	1	0	1	0	1	0	8
9	1	0	1	1	1	1	0	1	0	1	1	0	0	0	0	0	1	1	1	1	12
10	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1	0	0	1	0	0	12
11	1	0	1	0	0	0	0	1	1	1	1	1	1	1	0	1	0	1	1	0	12
12	1	0	0	1	1	1	0	0	1	0	0	1	1	0	1	0	1	1	1	0	11
13	0	1	1	1	1	0	1	0	1	1	0	0	1	1	0	0	0	0	1	1	11
14	0	0	1	1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	0	11
15	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	1	1	1	0	1	12
16	1	0	0	1	1	0	1	1	0	0	1	0	0	1	1	0	1	0	0	1	10
17	0	1	1	1	0	0	1	0	1	0	1	1	0	0	0	0	1	1	1	0	10
18	1	0	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	14
19	1	1	1	0	1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	13
20	1	0	1	1	0	1	0	0	1	1	1	1	0	0	1	1	1	1	0	1	13

ANEXO 1B. POSTEST SOBRE REGRESIÓN LINEAL

Título de la investigación: Desarrollo de una Interfaz Gráfica de Usuario en MATLAB para Potenciar el Aprendizaje de Regresión Lineal en la Educación Universitaria.

Objetivo: Medir el nivel de aprendizaje de regresión lineal después del uso de la interfaz gráfica de usuario en MATLAB.

Momento de aplicación: Después de la intervención.

Condición de aplicación: El estudiante resolverá la prueba después de haber utilizado la interfaz gráfica de usuario en MATLAB como recurso didáctico.

Duración: 40 minutos.

Puntaje total: 20 puntos.

Instrucciones: Lea atentamente cada pregunta. Marque la alternativa correcta o desarrolle el procedimiento solicitado.

I. Comprensión conceptual

1. ¿Para qué se utiliza principalmente la regresión lineal?
 - a) Para analizar la relación entre variables y realizar predicciones.
 - b) Para ordenar datos alfabéticamente.
 - c) Para eliminar variables independientes.
 - d) Para reemplazar todas las técnicas estadísticas.
2. ¿Qué representa la variable independiente en un modelo de regresión lineal?
 - a) La variable que ayuda a explicar o predecir el comportamiento de otra variable.
 - b) La variable que siempre depende del error.
 - c) La variable que no se utiliza en el modelo.
 - d) La variable que siempre tiene valor constante.
3. En el modelo $\hat{y} = b_0 + b_1x$, ¿qué representa $\hat{y}$?

- a) El valor observado de la variable independiente.
 - b) El valor estimado o predicho de la variable dependiente.
 - c) El error total del modelo.
 - d) El tamaño de la muestra.
4. ¿Cómo se calcula un residuo en regresión lineal?
- a) $e_i = y_i - \hat{y}_i$
 - b) $e_i = x_i + \bar{x}$
 - c) $e_i = b_0 + b_1$
 - d) $e_i = R^2 + 1$
5. ¿Cuándo se utiliza una regresión lineal múltiple?
- a) Cuando se desea explicar una variable dependiente usando dos o más variables independientes.
 - b) Cuando solo existe una variable independiente.
 - c) Cuando no se desea realizar predicciones.
 - d) Cuando solo se trabaja con variables cualitativas.

II. Aplicación procedimental

Considere los siguientes datos:

x	2	4	6	8	10
y	3	5	7	9	11

6. ¿Cuál es la media aritmética de los valores de x ?
7. ¿Cuál es la media aritmética de los valores de y ?

8. ¿Cuál es el valor de $S_{xx} = \sum (x_i - \bar{x})^2$?
9. ¿Cuál es el valor de $S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$?
10. ¿Cuál es la pendiente del modelo, si se utiliza la fórmula $b_1 = \frac{S_{xy}}{S_{xx}}$?
11. ¿Cuál es el intercepto del modelo, si se utiliza la fórmula $b_0 = \bar{y} - b_1\bar{x}$?
12. ¿Cuál es la ecuación estimada de regresión lineal simple?

III. Interpretación de resultados

13. Si el modelo obtenido fuera $\hat{y} = 1 + 1x$, ¿qué significa la pendiente 1?
14. Si el coeficiente de determinación fuera $R^2 = 0,92$, ¿cómo se interpretaría este resultado?
15. Si el valor observado es $y_i = 12$ y el valor estimado es $\hat{y}_i = 15$, ¿cuál sería el residuo?
16. Si un residuo es negativo, ¿qué significa?
 - a) Que el valor observado es menor que el valor estimado.
 - b) Que el valor observado es mayor que el valor estimado.
 - c) Que no existe error en el modelo.
 - d) Que la variable independiente no existe.

IV. Supuestos del modelo

17. ¿Qué significa la independencia de los errores en un modelo de regresión lineal?
- a) Que los residuos no deben estar correlacionados entre sí.
 - b) Que todos los datos deben ser iguales.
 - c) Que la pendiente siempre debe ser cero.
 - d) Que la variable dependiente no debe medirse.
18. ¿Cuándo se considera que existe homocedasticidad?
- a) Cuando los residuos presentan una dispersión aproximadamente constante.
 - b) Cuando los residuos forman una curva perfecta.
 - c) Cuando los datos son únicamente cualitativos.
 - d) Cuando el modelo no tiene variables independientes.
19. ¿Por qué es importante revisar la normalidad de los residuos?
- a) Porque permite respaldar pruebas estadísticas sobre el modelo.
 - b) Porque elimina la necesidad de interpretar resultados.
 - c) Porque convierte todos los datos en porcentajes.
 - d) Porque reemplaza el cálculo de la pendiente.
20. ¿Qué indica un valor alto de VIF en una regresión lineal múltiple?
- a) Posible multicolinealidad entre variables independientes.
 - b) Ausencia total de error.
 - c) Que el modelo siempre es perfecto.
 - d) Que no existe variable dependiente.

Tabla 13

Resultados del Postest - Ingeniería Civil.

Estudiante	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	Nota
E01	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	1	1	1	16
E02	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	17
E03	0	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	0	1	14
E04	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	18
E05	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	18
E06	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	1	1	0	1	15
E07	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	19
E08	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	16
E09	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	17
E10	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	19
E11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20
E12	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	17
E13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20
E14	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	0	15
E15	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	16
E16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20
E17	1	0	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	15
E18	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	18
E19	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	19
E20	1	0	0	1	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	14

Tabla 14*Resumen de las notas obtenidas por los estudiantes en el pretest y postest*

Estudiante	Pretest	Postest	Diferencia
E01	13	16	3
E02	14	17	3
E03	8	14	6
E04	7	18	11
E05	10	18	8
E06	10	15	5
E07	12	19	7
E08	8	16	8
E09	12	17	5
E10	12	19	7
E11	12	20	8
E12	11	17	6
E13	11	20	9
E14	11	15	4
E15	12	16	4
E16	10	20	10
E17	10	15	5
E18	14	18	4
E19	13	19	6
E20	13	14	1
Promedio	11.15	17.15	6.00

ANEXO 2. CUESTIONARIO DE USABILIDAD DE LA INTERFAZ GRÁFICA

Título de la investigación: Desarrollo de una Interfaz Gráfica de Usuario en MATLAB para Potenciar el Aprendizaje de Regresión Lineal en la Educación Universitaria.

Objetivo del instrumento: Evaluar la percepción de los estudiantes sobre la usabilidad de la interfaz gráfica de usuario en MATLAB después de haberla utilizado como recurso didáctico.

Tipo de instrumento: Cuestionario con escala de Likert.

Momento de aplicación: Después del uso de la interfaz gráfica y después de aplicado el postest.

Instrucciones: Marque con una X la alternativa que mejor represente su opinión sobre la interfaz gráfica utilizada.

Tabla 15

Escala de valoración

Valor	Categoría
1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

Tabla 16

Cuestionario de usabilidad de la interfaz gráfica

N.º	Ítem	1	2	3	4	5
Dimensión 1: Diseño y estructura de la interfaz						
1	La interfaz presenta una organización clara de botones, tablas y gráficos.					
2	La distribución de los elementos facilita el ingreso de datos.					

N.º	Ítem	1	2	3	4	5
3	Los nombres de los botones y opciones son comprensibles.					
4	La interfaz permite navegar sin confusión durante el proceso de cálculo.					

Dimensión 2: Funcionalidad técnica

5	La interfaz calcula correctamente los parámetros de regresión lineal simple.					
6	La interfaz calcula correctamente los parámetros de regresión lineal múltiple.					
7	Las tablas generadas ayudan a revisar el procedimiento de cálculo.					
8	Los gráficos generados representan adecuadamente el ajuste del modelo.					

Dimensión 3: Interactividad y retroalimentación

9	La interfaz muestra resultados de manera rápida después de ingresar los datos.					
10	La interfaz permite modificar los datos y observar nuevos resultados.					
11	Los mensajes de error o advertencia ayudan a corregir datos mal ingresados.					
12	La interfaz permite visualizar de manera clara los valores estimados y residuos.					

Dimensión 4: Utilidad para el aprendizaje

13	La interfaz ayudó a comprender el significado de la pendiente e intercepto.					
----	---	--	--	--	--	--

N.º	Ítem	1	2	3	4	5
14	La interfaz ayudó a interpretar el coeficiente de determinación R^2 .					
15	La interfaz facilitó la comprensión de los residuos del modelo.					
16	La interfaz ayudó a relacionar los cálculos con la gráfica de regresión.					

Dimensión 5: Satisfacción general

17	Considero que la interfaz es útil para aprender regresión lineal.					
18	Me gustaría utilizar esta interfaz en otras actividades del curso.					
19	La interfaz reduce la dificultad para resolver ejercicios de regresión lineal.					
20	Recomendaría el uso de esta interfaz a otros estudiantes.					

Tabla 17

Resultados del cuestionario de usabilidad de la interfaz gráfica.

Estudiante	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	Total
E1	5	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	5	5	4	5	94
E2	4	5	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	4	5	5	94
E3	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	5	4	94
E4	5	5	5	4	5	4	5	5	5	4	5	4	5	5	4	5	5	4	5	5	95
E5	4	5	4	5	5	4	5	4	5	5	4	5	4	5	5	4	5	4	5	5	93
E6	5	5	4	5	4	5	5	4	5	4	5	5	4	5	4	5	5	4	5	5	94
E7	5	4	5	5	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	94
E8	4	5	5	4	5	5	4	5	4	5	5	4	5	4	5	5	4	5	5	4	93
E9	5	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	4	5	5	4	94
E10	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	5	4	5	5	94
E11	4	5	5	5	4	5	4	5	5	4	5	5	4	5	5	4	5	5	4	5	94
E12	5	5	4	5	5	4	5	4	5	5	4	5	5	4	5	5	4	5	4	5	94
E13	5	4	5	5	4	5	5	4	5	4	5	5	4	5	5	4	5	4	5	5	94
E14	4	5	5	4	5	4	5	5	4	5	5	4	5	5	4	5	4	5	5	4	93
E15	5	5	5	4	5	5	4	5	4	5	5	4	5	4	5	5	4	5	5	4	94
E16	5	4	5	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	5	94
E17	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	4	5	5	4	5	93
E18	5	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	4	5	5	94
E19	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	5	4	5	5	94
E20	4	5	5	4	5	4	5	5	4	5	5	4	5	4	5	5	4	5	5	4	93