

**Modelo Predictivo del Indicador de Reprobación de la Población Estudiantil  
de una Institución de Educación Superior (IES).**

**Failure Indicator's Predictive Model of the Student Population of a  
Higher Education Institution (HEI).**

Méndez Lambarén, José Luis\*, Salazar Ortíz, Ana María\*\*,  
Fernández Martínez, Oscar Ramón\*\*\*

\*Maestro en Ingeniería Industrial. Tecnológico Nacional de México, Campus Tapachula.  
Email: jos.mendez@tapachula.tecnm.mx, ORCID: <https://orcid.org/0009-0008-0842-2263>.

\*\*Maestra en Educación Holista. Tecnológico Nacional de México, Campus Tapachula.  
Email: ana.salazar@tapachula.tecnm.mx, ORCID: <https://orcid.org/0009-0000-5232-2739>.

\*\*\*Maestro en Ingeniería Aplicada. Tecnológico Nacional de México, Campus Tapachula.  
Email: osc.fernandez@tapachula.tecnm.mx, ORCID: <https://orcid.org/0009-0005-3724-3780>.

**Correo para recibir correspondencia:**

José Luis Méndez Lambarén  
[jos.mendez@tapachula.tecnm.mx](mailto:jos.mendez@tapachula.tecnm.mx)

## RESUMEN

**OBJETIVO:** Determinar el modelo de regresión lineal múltiple de la IES acorde a su retícula de los programas educativos (PE), así como analizar los factores que afectan al indicador de logro de las competencias de las asignaturas de los diferentes planes de estudio que oferta el TecNM Campus Tapachula.

**MATERIAL Y MÉTODO:** Es de carácter cualitativo y cuantitativo, que se basa en un modelo de regresión lineal múltiple en el cual se considera como variables el indicador de reprobación, programas educativos y categorías de conocimiento, acorde al Consejo de Acreditación de la Enseñanza de la Ingeniería (CACEI).

**RESULTADOS:** Las siete categorías de conocimiento del CACEI se consideran fundamentales para la formación de ingenieros competentes de acuerdo con los retos de los desafíos del siglo XXI. Cabe mencionar que, las categorías no son excluyentes entre sí, sino que se interrelacionan y se integran a lo largo del programa de estudios de ingeniería.

**CONCLUSIONES:** Se detectaron las acciones de vinculación que mayormente se implementan con el sector productivo y de servicios, que son: el Servicio Social y la Residencia Profesional o prácticas profesionales, seguido de la Promoción de Vacantes en las Bolsas de Trabajo de las IES.

**Palabras clave:** Análisis estadístico, Indicador, Reprobación, Python, Jupyter notebook.

## ABSTRACT

**OBJECTIVE:** To determine the multiple linear regression model of the HEI according to its grid of educational programs (EP), as well as analyzing the factors that affect the indicator of the subjects' competencies achievement of the different study plans offered by TecNM in Tapachula Campus.

**MATERIAL AND METHOD:** It is qualitative and quantitative in nature, based on a multiple linear regression model in which the failure indicator, educational programs and categories of knowledge are considered as variables, according to the Education Accreditation Council of Engineering (CACEI, by its acronym in Spanish).

**RESULTS:** The seven categories of knowledge of the CACEI are considered fundamental for the training of competent engineers in accordance with the challenges of the 21<sup>st</sup> century. It is worth mentioning that the categories are not mutually exclusive but are interrelated and integrated throughout the engineering study program.

**CONCLUSIONS:** The linking actions that are mostly implemented with the productive and service sector were detected, they are: Social Service and Professional Residency or professional internships, followed by the Promotion of Vacancies in the Job Boards of the HEI.

**KEYWORDS:** Statistical analysis, Indicator, Failure, Python, Jupyter notebook.

Las Instituciones de Educación Superior (IES) tienen como desafío ofertar una educación no solo de calidad, sino también con equidad, permitiendo con ello el desarrollo del individuo (Barajas, 2019), por lo que la calidad no solo se basa en la eficiencia, sino que además se debe considerar la pertinencia, cobertura, desempeño y eficacia; es decir, las IES deben trabajar para cumplir con los objetivos para lo que fueron creadas (Jaimes, et al., 2015).

En concordancia con la Constitución Política de los Estados Unidos Mexicanos, acorde al Art. 3., inciso i, que toda educación será de excelencia, entendiéndolo como el mejoramiento integral constante que promueve el máximo logro de aprendizaje de los educandos, para el desarrollo de su pensamiento crítico (Cámara, 2024).

Los autores Lorenzo y Zaragoza (2014) aseguran que la educación superior es el tercer Sistema Educativo Nacional Mexicano, el cual comprende el nivel de licenciatura y posgrado y no es de carácter obligatorio. Es por ello que, para mantener y/o incrementar la matrícula en el nivel superior, es de vital importancia que el Sistema Educativo cuente con toda la información sobre los indicadores básicos, tales como: ingreso, reprobación, deserción, eficiencia terminal y titulación; el análisis de estos indicadores permite al sistema educativo de nivel superior tener un contexto más detallado sobre el comportamiento de la matrícula escolar por carrera, por cohorte y por ciclo escolar, contribuyendo a la toma de decisiones estratégicas (Martínez, 2010).

Los estudios de cohortes derivan del seguimiento de los estudiantes que inician y terminan su actividad académica en un periodo determinado, de acuerdo con el Manual de Lineamientos del Tecnológico Nacional de México (TecNM) son de 9 a 12 semestres el período de duración del estudio de una licenciatura (Callejo, 2001).

El aseguramiento de la calidad de los diferentes programas de estudios de licenciatura del TecNM que se imparten en los diferentes institutos tecnológicos, solo el 42% son reconocidos por su buena calidad. Algunos de los retos son, incrementar el número de programas académicos de licenciatura reconocidos a nivel nacional e internacional, por su calidad (TecNM, 2019).

La intencionalidad de llevar a cabo los procesos de evaluación está orientada hacia la conducción de la mejora continua de los programas educativos y sus comunidades, en todas las IES en donde se operan procesos integrales de cambio institucional.

Por lo tanto, llevar a cabo un proceso de acreditación de un Programa Educativo (PE) con el objetivo de garantizar la calidad y su pertinencia, en el cual se busca que cumplan con estándares mínimos internacionales reconocidos para los programas de ingeniería y que se promueva en las

IES la cultura de mejora continua que incorpore tendencias nacionales e internacionales para la formación de ingenieros. El proceso de acreditación en México es de carácter voluntario, en este sentido el Consejo de Acreditación de la Enseñanza de la Ingeniería (CACEI) considera criterios con base a estándares internacionales aceptados por organismos acreditadores a los que pertenece el Washington Accord (CACEI, 2018).

El índice de eficiencia terminal en el periodo 2020-2023 de acuerdo con el INEGI para el Estado de Chiapas, ha incrementado un 7% en el nivel media superior (INEGI, 2023). Por lo tanto, se espera que los egresados de este nivel educativo comiencen sus estudios en un sistema educativo de nivel superior.

La reprobación de asignaturas es multifactorial, sin embargo, en este análisis se toma como referencia la dificultad de cada PE, derivado de que algunos contenidos académicos son más complejos o exigentes, por lo que pudieran tener tasas de reprobación más altas, especialmente si los estudiantes no están adecuadamente preparados o no cuentan con los recursos necesarios.

Por otra parte, el indicador de reprobación en las IES es un parámetro importante para evaluar el rendimiento académico y éxito estudiantil, por lo tanto, la cantidad de estudiantes inscritos en una asignatura puede tener un impacto en el proceso de enseñanza aprendizaje (Rosario, 2013).

Por lo anterior, es importante determinar si la variable investigada es descrita en términos de otras variables, la cual se llamará variable de respuesta o variable dependiente. Asimismo, se puede usar la variable en conjunción con otras variables para describir una variable respuesta dada, la cual se llamará predictor o variable independiente (López, 2015).

Según el modelo de regresión lineal múltiple, busca una función de regresión poblacional para predecir el valor de la variable dependiente, así como evaluar la influencia que tienen los predictores sobre ella (Escalante, 2012). Por lo tanto, la ecuación de regresión múltiple la definen como:

$$y = a_0 + a_1u_1 + a_2u_2 + a_3u_3 + a_4u_4 + a_5u_5 + a_nu_n$$

El objetivo del trabajo de investigación es determinar el modelo de regresión lineal múltiple de la IES acorde a su retícula de los Programas Educativos (PE), así como analizar los factores que afectan al indicador de logro de las competencias de las asignaturas de los diferentes planes de estudio que oferta el TecNM Campus Tapachula.

## Preguntas de investigación

1. ¿Qué correlación existe entre la variable dependiente con las variables independientes?
2. ¿Cuál es la varianza que existe en el modelo predictivo de investigación?
3. ¿Cuáles son los coeficientes de las variables independientes para la construcción del modelo predictivo?

Por otro lado, es importante analizar la información inicial para modelar y estructurar una base de datos, y dependiendo de la cantidad, permite acceder a un conjunto de datos extremadamente grandes y complejos que son difíciles de procesar utilizando métodos comunes de gestión. El uso del *Big Data* se ha caracterizado por las 3V's, las cuales son: volumen, variedad y velocidad; algunos otros autores han agregado otras V's como: valor, visualización y veracidad (Wintjen, 2020).

En los últimos años, el lenguaje de programación Python de alto nivel, ofrece características de legibilidad de código, lenguaje interpretado, múltiples paradigmas de programación, extensa biblioteca estándar con módulos y funciones para realizar una variedad de tareas; y con una comunidad activa de desarrolladores que contribuyen con recursos educativos (Hinojosa, 2016).

El uso de la librería pandas de Python, realiza un análisis especializado en el manejo y análisis de estructuras de datos, por lo que permite compilar información en formato CSV o Excel, a través de datos tabulares, generando de igual manera gráficos de alta calidad y análisis estadísticos (Galea, 2018).

## MATERIAL Y MÉTODO

El área de estudio de la presente investigación son los factores que impactan en el proceso de enseñanza-aprendizaje para el logro de las competencias de las asignaturas de los programas educativos que ofrece el TecNM Campus Tapachula.

Esta investigación es de carácter cualitativo y cuantitativo, debido a que se sostiene que existe relación entre los aspectos cualitativos de los diferentes programas educación y áreas del conocimiento, con respecto a los resultados en los niveles de competencias adquiridas en las asignaturas, que se expresan de manera cuantitativa.

El fenómeno observado en cuanto a los niveles de competencias adquiridas en las asignaturas, ha sido descrito a través de tres principales técnicas:

**Foros de discusión.** Son categorizados por áreas del conocimiento que marca el INEGI: Artes y humanidades; Ciencias sociales, Administración y Derecho; Ciencias naturales, exactas y de la computación; Ingeniería, manufactura y construcción; Agronomía y veterinaria; Salud y Educación (INEGI, 2016) los cuales se realizaron como mecanismo exploratorio para el análisis de la articulación de los planes y programas de estudio de IES con sectores productivos y de servicios en el Estado de Chiapas, en el cual se recuperó información sobre la percepción de los empleadores con respecto a los desempeños de egresados de las IES. Por otro lado, para el caso de los programas educativos objeto del presente estudio, se recuperaron datos en las salas de Ciencias Sociales, Administración y Derecho; Ciencias Naturales, Exactas y de la Computación; Ingeniería, Manufactura y Construcción.

**Estudio de caso:** El Comité para la Autoevaluación de los Programas Educativos a fin de alcanzar la acreditación, realiza el análisis de los criterios del Marco de Referencia 2018 del CACEI, especialmente en el criterio 2. Estudiantes, el cual establece los aspectos metodológicos sobre reprobación, rezago y deserción. En dicho análisis se han recuperado datos de cinco cohortes generacionales, que han permitido un análisis inductivo sobre las posibles causas especialmente del índice de reprobación

**Entrevista con estudiantes:** Con la implementación del Programa Institucional de Tutorías se realizan entrevistas al 100% de la matrícula estudiantil, obteniendo datos para identificar los obstáculos que se presentan durante la trayectoria académica de cada estudiante, como son: bajos niveles de desempeño, repetición, rezago y fracaso estudiantil, deserción, abandono y baja eficiencia terminal. Cuando se detectan circunstancias relacionadas con el desempeño académico, se canaliza para brindar asesorías mediante un programa. En este caso, la constante es en asignaturas clasificadas en el área del conocimiento de las Ciencias Básicas.

Las tres fuentes de información cualitativa permiten la reflexión sobre el proceso de enseñanza-aprendizaje, identificando las fortalezas, pero también las oportunidades de mejorarlo.

Esta investigación es de carácter cualitativo y cuantitativo, se basa en realizar un modelo de regresión lineal múltiple en el cual se considera las siguientes variables:

### **Variable dependiente**

- Indicador de reprobación. Variable numérica recuperada de la cantidad de estudiantes reprobados por asignatura/programa educativo con la totalidad de matrícula estudiantil, en un periodo de cinco cohortes.

## Variable independiente

- Cantidad de estudiantes inscritos. Variable numérica recuperada de la cantidad de estudiantes inscritos por asignatura/programa educativo con la totalidad de matrícula estudiantil, en un periodo de cinco cohortes.
- Programa educativo. Variable cualitativa (de atributo) que describe el perfil de egreso relacionado a los seis programas educativos de Ingenierías: Civil, Industrial, Electromecánica, Sistemas Computacionales, Química y Gestión Empresarial.
- Categorías de CACEI. Variable cualitativa (de atributos) que describe las áreas del conocimiento para la organización curricular de los programas educativos de ingeniería; Ciencias Básicas, Ciencias de la Ingeniería, Ingeniería aplicada, Ciencias Sociales y Humanidades; y Ciencias económico-administrativas.

## Hipótesis:

1. La complejidad de los productos evaluables según la organización reticular del CACEI para evidenciar que se alcanza la competencia, en cada área del conocimiento impacta en el indicador de reprobación.
2. La complejidad del PE en lo que respecta al logro de las competencias específicas puede tener un impacto significativo en el indicador de reprobación derivado del nivel de dificultad académica, competencias previas como requisitos, carga de trabajo y estilos de aprendizaje.

**Tabla 1**

*ID de variables del programa educativo*

IDVariables	Programa Educativo
IC	Ingeniería Civil
IID	Ingeniería Industrial
IEM	Ingeniería Electromecánica
ISC	Ingeniería en Sistemas Computacionales
IGE	Ingeniería en Gestión Empresarial
IQ	Ingeniería Química

*Nota.* Esta tabla muestra la identificación (ID) de las variables de cada PE acorde de oferta académica de la IES. Adaptado del mensaje de bienvenida del TecNM Campus Tapachula ([https://tapachula.tecnm.mx/?page\\_id=7](https://tapachula.tecnm.mx/?page_id=7))

**Tabla 2**  
*ID de Variables de Organización reticular de CACEI*

IDVariables	Organización Reticular CACEI
CB	Ciencias Básicas
CEA	Ciencias Económico-Administrativas
CI	Ciencias de Ingeniería
CSyH	Ciencias Sociales y Humanidades
DI	Diseño de Ingeniería
IA	Ingeniería Aplicada

*Nota.* Esta tabla muestra la identificación de la organización reticular de CACEI. Adaptado de Marco de Referencia 2018 del CACEI en el Contexto Internacional (Ingenierías). Pp.129-132 (<http://cacei.org.mx/nv/nv02/nv0208.html>)

### Preparación de los datos

Para validar el modelo regresión lineal múltiple, es fundamental realizar la preparación de los datos (Bases de Datos), especialmente cuando consideran tanto variables cualitativas como variables cuantitativas.

Derivado de que las variables independientes describen los atributos asociados al PE y a la de organización reticular de CACEI, se crean variables binarias (0,1), es decir variables *dummy*, dado que el modelo de regresión lineal se procesa con caracteres numéricos.

**Tabla 3**  
*ID de Variables programa educativo (PE) conversión a variable dummy*

IDVariables	Variable Dummy
IC	10000
IID	01000
IEM	00100
ISC	00010
IGE	00001
IQ	00000

*Nota.* Esta tabla muestra la identificación de las variables de los PE a variables *dummy*. Elaboración propia.

**Tabla 4**

*ID de Variables de organización curricular conversión a variable dummy*

IDVariables	Variable Dummy
CB	10000
CEA	01000
CI	00100
CSyH	00010
DI	00001
IA	00000

*Nota. Esta tabla muestra la identificación de las variables de la organización curricular a variables dummy. Elaboración propia.*

## RESULTADOS

Las siete categorías de conocimiento del CACEI se consideran fundamentales para la formación de ingenieros competentes de acuerdo con los retos de los desafíos del siglo XXI. Cabe mencionar que, las categorías no son excluyentes entre sí, sino que se interrelacionan y se integran a lo largo del programa de estudios de Ingeniería.

En la Tabla 5, se muestra el coeficiente de correlación múltiple, siendo este de 0.5795, lo que representa la relación de la variable de reprobación con sus predictoras; asimismo, se encuentra el coeficiente de determinación de bondad y ajuste, en el cual ambos están en promedio de 0.30, lo que indica la relación entre las variables predictoras, lo que indica el 30.99% con respecto a la variable de reprobación. De igual manera, se observa el error con un valor de 6.46 y las observaciones con 294, es por ello, que el modelo de regresión lineal multivariable explica una cantidad moderada de la variabilidad de la variable dependiente, por lo que los factores que no están siendo considerados por el modelo, también pueden estar influyendo en la variable dependiente.

**Tabla 5**

*Datos estadísticos del modelo de regresión lineal múltiple*

<i>Estadísticas de la regresión</i>	
Coeficiente de correlación múltiple	0.579518329
Coeficiente de determinación R <sup>2</sup>	0.335841493
R <sup>2</sup> ajustado	0.309934601
Error típico	6.464521944
Observaciones	294

*Nota. Muestra la estadística del modelo de regresión planteado. Elaboración propia.*

En la Tabla 6, se muestra el análisis de varianza, en el cual el valor crítico de F es menor al 5%, por lo que se propone el modelo de regresión con la variable dependiente de reprobación con sus variables predictoras.

**Tabla 6**  
*Datos del análisis de varianza*

Análisis de varianza					
Items	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	11	5959.153181	541.7411983	12.96340341	6.69734E-20
Residuos	282	11784.7924	41.79004396		
<b>Total</b>	<b>293</b>	<b>17743.94558</b>			

Nota: Muestra el análisis de varianza del modelo de regresión planteado. Elaboración propia.

En la Tabla 7, se muestran los coeficientes de correlación del modelo de regresión múltiple que considera la variable dependiente de reprobación con sus variables predictoras del programa educativo y la organización reticular de CACEI. Dado que, la significancia de las variables independientes con respecto a la reprobación se estima cuando el valor de la probabilidad es menor de 5%, es decir, conforme el valor sea más pequeño mayor significancia. En este sentido, se puede descartar que la variable independiente del PE de Ingeniería Industrial tenga significancia con el indicador de reprobación.

**Tabla 7**  
*Datos estadísticos de coeficientes de PE y organización reticular de CACEI*

ID	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
Intercepción	-2.704665029	1.343363609	-2.013352909	0.045027455	-5.348957943	-0.060372115	-5.348957943	-0.060372115
IC	6.674403912	1.559717259	4.279239633	2.57186E-05	3.604237897	9.744569927	3.604237897	9.744569927
IID	-0.279603545	1.423110029	-0.196473596	0.844381012	-3.080870261	2.521663171	-3.080870261	2.521663171
IEM	1.394050772	1.356863556	1.027406747	0.305109415	-1.276815598	4.064917142	-1.276815598	4.064917142
ISC	3.068081317	1.420237758	2.160258942	0.031595022	0.272468412	5.863694222	0.272468412	5.863694222
IGE	1.205786099	1.464621029	0.823275151	0.411047207	-1.677191364	4.088763563	-1.677191364	4.088763563
CB	1.988758391	1.314245893	1.513231582	0.131340746	-0.598218864	4.575735646	-0.598218864	4.575735646
CEA	2.899215599	1.632634298	1.775789963	0.07684588	-0.314481182	6.11291238	-0.314481182	6.11291238
CI	1.62631315	1.167989623	1.392403766	0.164897115	-0.672771524	3.925397824	-0.672771524	3.925397824
CSyH	1.664962367	1.740353256	0.956680698	0.339547862	-1.760769693	5.090694427	-1.760769693	5.090694427
DI	1.92784958	1.174443025	1.641501153	0.10180816	-0.383938048	4.239637208	-0.383938048	4.239637208
Inscritos	0.104056503	0.016273097	6.394388421	6.66728E-10	0.072024345	0.13608866	0.072024345	0.13608866

Nota. Muestra información estadística de los coeficientes del PE y organización reticular de CACEI.

De lo anterior, se deriva el modelo de regresión lineal múltiple, tomando como referencia los coeficientes de los PE y la organización reticular de CACEI, por lo que se define como:

$$y = -2.70 + 6.67 * IC_i - 0.28 * IID_i + 1.39 * IEM_i + 3.07 * ISC_i + 1.21 * IGE_i + 0 * IQ_i + 1.99 * CB_i + 2.90 * CEA_i + 1.63 * CI_i + 1.66 * CSyH_i + 1.93 * DI_i + 0 * IA_i + 0.10 * inscritos_i$$

Por otro lado, se ejecutó el presente modelo incluyendo las variables independientes de sexo (F: femenino y M: Masculino), sin embargo, en la tabla de coeficientes de la regresión lineal se detectó que para la variable F (sexo femenino), el coeficiente fue igual 0, lo que podría significar que esta variable no contribuye al modelo, es por ello que se determinó excluir variable dicha variable del modelo propuesto.

**Tabla 8**  
*Datos estadísticos del coeficiente de la variable F*

ID	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
F	0	0	65535	#¡NUM!	0	0	0	0

*Nota.* Esta tabla muestra información estadística de coeficiente de la variable de sexo femenino. Elaboración propia.

Adicionalmente, se ejecutó el modelo de regresión lineal múltiple utilizando el lenguaje de programación de Python y Jupyter Notebook (Python, 2023), a fin de comprobar los resultados recuperados a partir del modelo de regresión con Microsoft Excel. En la Figura 1 se muestra el lenguaje de programación utilizando Python y Jupyter, tomando como referencia el análisis de modelo de regresión lineal múltiple desarrollado con Excel, es por ello que se observa en la Figura 2 el modelo de la ecuación de regresión múltiple, en donde se indica la variable dependiente y los coeficientes de las variables predictoras que permiten dar certidumbre en la correlación del indicador de reprobación con el PE y la organización reticular del CACEI.

**Figura 1**  
*Captura de pantalla de la programación con Python y Jupyter Notebook utilizando la librería Pandas para el análisis y modelado de la información*

```
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import pandas as pd

# Cargar los datos (asegúrate de tener los datos cargados en 'datos')
datos = pd.read_excel('base.xlsx')

# Definir las variables independientes (X) y dependiente (y)
X = datos[['Inscritos', 'IC', 'IID', 'IEM', 'ISC', 'IGE', 'CB', 'CEA', 'CI', 'CSyH', 'DI']]
y = datos['Reprobados']

# Agregar la constante al conjunto de datos (intercepto)
X = sm.add_constant(X)

# Ajustar el modelo de regresión
modelo = sm.OLS(y, X).fit()

# Obtener los valores predichos por el modelo
y_pred = modelo.predict()

# Crear un gráfico de dispersión
plt.figure(figsize=(10, 8))
sns.scatterplot(x=datos['Inscritos'], y=datos['Reprobados'], label='Real')
sns.scatterplot(x=datos['Inscritos'], y=y_pred, label='Predicción')

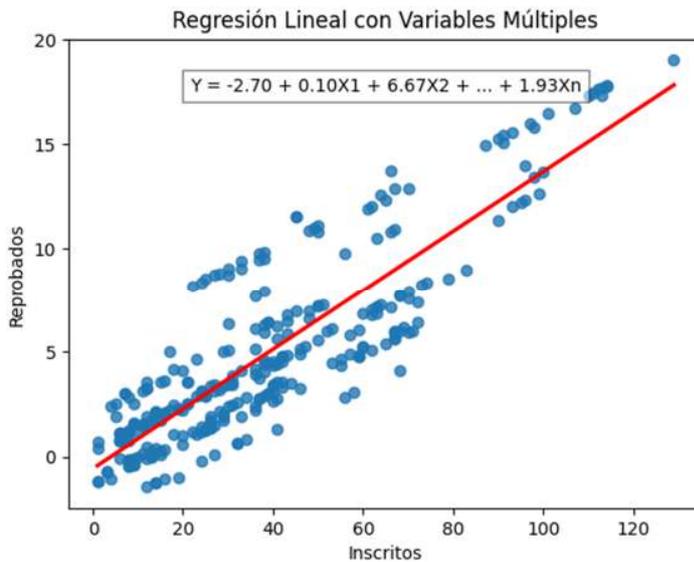
# Agregar la línea de regresión
sns.regplot(x=datos['Inscritos'], y=y_pred, scatter=False, color='red', label='Línea de Regresión')

# Configurar el gráfico
plt.title('Modelo de Regresión Lineal con Múltiples Variables')
plt.xlabel('Inscritos')
plt.ylabel('Reprobados')
plt.legend()
plt.show()
```

*Nota.* La captura de pantalla, muestra información del código de programación de Python y Jupyter Notebook del modelo predictivo.

### Figura 2

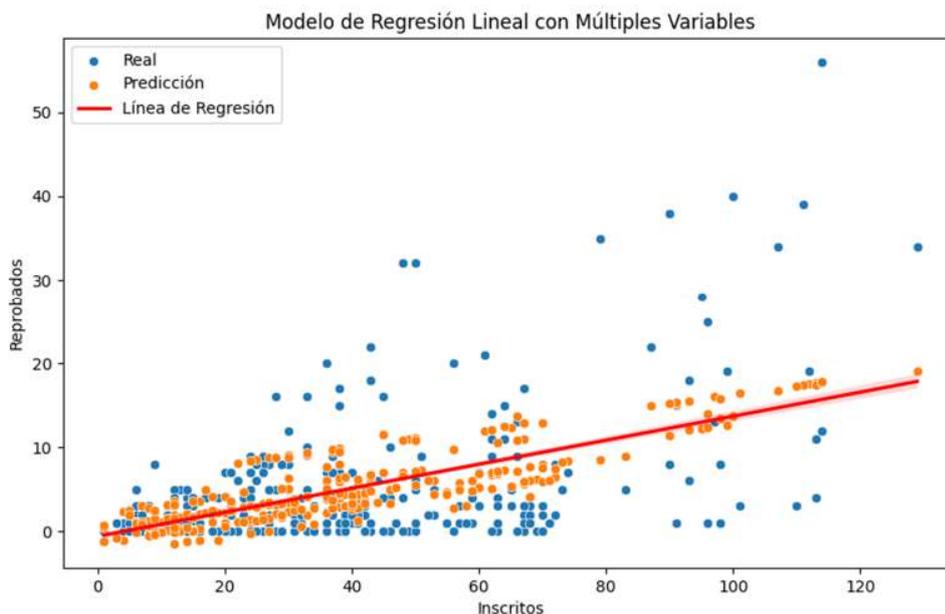
Captura de pantalla del modelo de regresión lineal múltiple con Python y Jupyter Notebook



Nota. La captura de pantalla, muestra el modelo de regresión lineal con variables múltiples del software Python y Jupyter Notebook del modelo predictivo.

### Figura 3

Captura de pantalla de gráfico del modelo de regresión lineal múltiple con Python y Jupyter Notebook.



Nota. La captura de pantalla, muestra el modelo de regresión lineal con múltiples variables del software Python y Jupyter Notebook del modelo predictivo, en donde se indica la tendencia de regresión, así como el ajuste del modelo real con su predicción.

En la Figura 3, se observa el modelo de regresión lineal con múltiples variables en el cual se indica la línea de regresión, así como los datos de predicción del modelo ajustado a la ecuación de regresión.

El modelo predictivo fue probado mediante el software de Python, a través de su entorno Jupyter Notebook, ya que cuenta con varias bibliotecas en un entorno de trabajo para el análisis de datos utilizando Pandas, además de exportar gráficos para una mejor interpretación de los resultados. Por lo tanto, para el desarrollo de la siguiente investigación se tiene que preparar, seleccionar, modelar y validar utilizando el método estadístico de regresión lineal multivariable en el entorno de programación de Jupyter Notebook.

En el desarrollo de un modelo de regresión lineal es importante considerar que los residuales se distribuyan normalmente, para considerar que las conclusiones derivadas del modelo sean válidas. En este sentido, los resultados que se muestran en la siguiente figura 4 muestran que Prob(JB): 1.80 e-170, es decir, el modelo presentado muestra que hay evidencia estadística significativa en contra del planteamiento de la hipótesis nula de normalidad de los residuales.

**Figura 4**  
Resultados de la regresión utilizando el lenguaje de programación de Python

OLS Regression Results						
Dep. Variable:	Reprobados	R-squared:	0.336			
Model:	OLS	Adj. R-squared:	0.310			
Method:	Least Squares	F-statistic:	12.96			
Date:	Sun, 19 Nov 2023	Prob (F-statistic):	6.70e-20			
Time:	14:58:17	Log-Likelihood:	-959.74			
No. Observations:	294	AIC:	1943.			
Df Residuals:	282	BIC:	1988.			
Df Model:	11					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-2.7047	1.343	-2.013	0.045	-5.349	-0.060
Inscritos	0.1041	0.016	6.394	0.000	0.072	0.136
IC	6.6744	1.560	4.279	0.000	3.604	9.745
IID	-0.2796	1.423	-0.196	0.844	-3.081	2.522
IEM	1.3941	1.357	1.027	0.305	-1.277	4.065
ISC	3.0681	1.420	2.160	0.032	0.272	5.864
IGE	1.2058	1.465	0.823	0.411	-1.677	4.089
CB	1.9888	1.314	1.513	0.131	-0.598	4.576
CEA	2.8992	1.633	1.776	0.077	-0.314	6.113
CI	1.6263	1.168	1.392	0.165	-0.673	3.925
CSyH	1.6650	1.740	0.957	0.340	-1.761	5.091
DI	1.9278	1.174	1.642	0.102	-0.384	4.240
Omnibus:	131.403	Durbin-Watson:	1.896			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	781.707			
Skew:	1.738	Prob(JB):	1.80e-170			
Kurtosis:	10.192	Cond. No.	353.			

Nota. Esta figura muestra los resultados estadísticos del modelo de regresión lineal multivariable.

De igual manera, el resultado de Prob (F-statistic):  $6.70 \times 10^{-20}$  indica que el modelo de regresión lineal multivariable en su conjunto es estadísticamente significativo, por lo que al menos una de las variables independientes tiene un efecto significativo en la variable dependiente.

## **CONCLUSIONES**

En la presente investigación se detecta que las acciones de vinculación que mayormente se implementan con el sector productivo y de servicios son el Servicio Social y la Residencia Profesional o prácticas profesionales, seguido de la Promoción de Vacantes en las Bolsas de Trabajo de las IES. Aunque la mayoría de las organizaciones reconocen la formalización de la vinculación a través de convenios de colaboración con las IES, en particular el sector salud expresa que la normatividad aplicable en esta instancia dificulta su concertación, sin embargo, la colaboración se realiza.

Dado que existen acciones de vinculación con baja implementación, es recomendable diseñar un catálogo de programas, esquemas y/o acciones de vinculación general, que las IES pueden ejecutar con las organizaciones a través de estudiantes, egresados y docentes, a fin de ser difundido ampliamente y de forma sistemática con el sector productivo y de servicios.

Otro de los puntos que se encontraron para mejorar es la preocupación por la falta de acompañamiento y seguimiento puntual de parte de las IES durante las estadías, así como el involucramiento de los docentes y su actualización práctica, por lo que se considera de gran importancia fortalecer estos aspectos para garantizar se cumplan los objetivos de cada uno de los programas o proyectos a desarrollar.

El modelo muestra que existe correlación entre las variables, sin embargo, el valor de este coeficiente es bajo para considerar que los factores analizados son significativos para mitigar el indicador de reprobación es consecuencia de estos.

En este sentido, se debe considerar un modelo más complejo, que integre otras variables o reemplace algunas de las analizadas. Para efectos de futuros estudios se recomienda contar con información estadística de los siguientes factores:

1. Calidad de la enseñanza, en donde se considere la metodología y los recursos educativos disponibles.
2. Apoyo académico como son las tutorías y las asesorías, con el objeto de aumentar el desempeño de los estudiantes.

3. Factores socioeconómicos, en el cual se aborde los desafíos para enfrentar la falta de recursos financieros, así como el acceso a materiales educativos.
4. Cambio de carrera o desinterés, en donde se considere las dificultades académicas debido a la elección incorrecta de la carrera o la falta de interés en el campo de aplicación de la carrera.
5. Políticas institucionales; en este último punto se puede asociar los criterios de evaluación, así como el sistema de calificación que puede afectar el indicador de reprobación.

Adicionalmente, a la implementación de estrategias como la orientación basada en competencias, la transparencia en la información de los programas educativos y la consejería a lo largo de toda la trayectoria académica, destaca la preocupación por el desarrollo integral de los estudiantes.

Por último, abordar las etapas críticas desde la admisión hasta el logro académico, la institución no solo busca mejorar sus indicadores educativos, sino que también se esfuerza por proporcionar un apoyo personalizado que atienda los desafíos específicos que los estudiantes pueden enfrentar en su camino educativo. La propuesta refleja una comprensión profunda de los factores que influyen en el rendimiento académico y establece un enfoque preventivo para abordar estos desafíos.

## REFERENCIAS

- Barajas, C. M. (2019). Gestión del cambio: el nuevo desafío para mejorar la calidad de la educación superior. *Revista de investigación*, 43(98). Universidad Pedagógica Experimental Libertador. <https://www.redalyc.org/journal/3761/376168604012/html/>
- Callejo, J. (2001). Estudio de Cohorte de Estudiantes de la UNED: una aproximación al análisis de abandono. *Revista Iberoamericana de Educación a Distancia*, 33-70. <https://dialnet.unirioja.es/servlet/articulo?codigo=1704699>
- Consejo de Acreditación de la Enseñanza de la Ingeniería A.C. [CACEI], (2018). Marco de Referencia 2018 en el Contexto Internacional. [http://cacei.org.mx/nv/nvdocs/marco\\_ing\\_2018.pdf](http://cacei.org.mx/nv/nvdocs/marco_ing_2018.pdf)
- Cámara de Diputados del H. Congreso de la Unión. (2024). Constitución Política de los Estados Unidos Mexicanos. <https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf>
- Escalante, R. D. (2012). Manual de aplicación del modelo de regresión lineal múltiple con correcciones de especificación, usos de STATA 9.0, STATA 10.0, EVIEWS 5.0, SSPS 11.0: (ed.). B - EUMED. <https://www.eumed.net/libros-gratis/2010c/720/index.htm>
- Galea, A. (2018). Beginning Data Science with Python and Jupyter: Use Powerful Industry-standard Tools Within Jupyter and the Python Ecosystem to Unlock New, Actionable

Insights from Your Data. Reino Unido: Packt Publishing.  
[https://www.google.com.mx/books/edition/Beginning\\_Data\\_Science\\_with\\_Python\\_and\\_J/L4lfDwAAQBAJ?hl=es-419&gbpv=1&dq=.+Beginning+Data+Science+with+Python+and+Jupyter:+Use+Powerful+Industry-standard+Tools+Within&printsec=frontcover](https://www.google.com.mx/books/edition/Beginning_Data_Science_with_Python_and_J/L4lfDwAAQBAJ?hl=es-419&gbpv=1&dq=.+Beginning+Data+Science+with+Python+and+Jupyter:+Use+Powerful+Industry-standard+Tools+Within&printsec=frontcover)

Hinojosa, Á. P. (2016). Python Paso a paso. España: RA-MA Editorial.

Instituto Nacional de Estadística y Geografía (2023). Eficiencia Terminal por Entidad Federativa según nivel educativo, ciclos escolares seleccionados 2020-2023.  
[https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Educacion\\_Educacion\\_12\\_987437b7-5398-4067-8b19-a1cd2ec36b7b](https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Educacion_Educacion_12_987437b7-5398-4067-8b19-a1cd2ec36b7b)

Instituto Nacional de Estadística y Geografía (2016). Clasificación mexicana de planes de estudio por campos de formación académica 2016. Educación superior y media superior.  
[https://www.inegi.org.mx/contenidos/productos/prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/702825086664.pdf](https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825086664.pdf)

Jaimes, R. N., Cardoso, J. D., y Bobadilla, B. S. (2015). La educación superior en México, una demanda con compromiso social. *Iberoamericana para la Investigación y el Desarrollo Educativo*, 5(10). <https://www.redalyc.org/pdf/4981/498150318011.pdf>

López, J. I. (2015). Análisis de regresión aplicado: teoría y práctica. Primera edición. Universidad Juárez Autónoma de Tabasco.  
<https://pcientificas.ujat.mx/index.php/pcientificas/catalog/view/93/82/325>

Lorenzo, O., y Zaragoza, J. E. (2014). Educación Media y Superior en México: análisis teórico de la realidad actual. Dedicada. *Revista de Educação e Humanidades*.  
<https://dialnet.unirioja.es/servlet/articulo?codigo=4733974>

Martínez, F. (2010). Los indicadores como herramientas para la evaluación de la calidad de los sistemas educativos. *Sinéctica*, (35), 1-17.  
[http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-109X2010000200004&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2010000200004&lng=es&tlng=es).

Tecnológico Nacional de México (2019). Programa de Desarrollo Institucional PDI 2019-2024.  
[https://www.tecnm.mx/menu/conocenos/PDI-TecNM-2019-2024\\_2oct2020.pdf](https://www.tecnm.mx/menu/conocenos/PDI-TecNM-2019-2024_2oct2020.pdf)

Python Software Foundation. (2023). Download Python 3.12.2. The latest version for window, Linux/ UNIX, macOS, other. <https://www.python.org/downloads/>

Rosario, V. M. (2013). La acreditación de la educación superior en Iberoamérica: La gestión de la calidad de los programas educativos. tensiones, desencuentros, conflictos y resultados. Red de Académicos de Iberoamérica A.C.  
[https://www.researchgate.net/publication/308171592\\_La\\_gestion\\_de\\_la\\_calidad\\_de\\_los\\_programas\\_educativos\\_Tensiones\\_desencuentros\\_conflictos\\_y\\_resultados](https://www.researchgate.net/publication/308171592_La_gestion_de_la_calidad_de_los_programas_educativos_Tensiones_desencuentros_conflictos_y_resultados)

Wintjen, M. (2020). Practical data analysis using jupyter notebook: learn how to speak the language of data by extracting useful and actionable insights using python. India: Packt Publishing. [https://www.google.com.mx/books/edition/Practical\\_Data\\_Analysis\\_Using\\_Jupyter\\_No/tqTsDwAAQBAJ?hl=es-419&gbpv=0](https://www.google.com.mx/books/edition/Practical_Data_Analysis_Using_Jupyter_No/tqTsDwAAQBAJ?hl=es-419&gbpv=0)