



ISSN: 2675-682X

DOSSIER TEMÁTICO

“LA GESTIÓN ORGANIZACIONAL DESDE LA INNOVACIÓN Y LA VISIÓN INTEGRAL”

ANÁLISIS ESTADÍSTICO PARA CONOCER LA RELACIÓN ASIGNATURA- BAJA EN LOS ESTUDIANTES DE PRIMER AÑO DE LA CARRERA INGENIERÍA INDUSTRIAL DE LA UNIVERSIDAD DE PINAR DEL RÍO “HERMANOS SAIZ MONTES DE OCA”

STATISTICAL ANALYSIS TO KNOW THE SUBJECT-LOW RELATIONSHIP IN FIRST-YEAR STUDENTS OF THE INDUSTRIAL ENGINEERING CAREER AT THE UNIVERSITY OF PINAR DEL RÍO “HERMANOS SAIZ MONTES DE OCA”

ANÁLISE ESTATÍSTICA PARA CONHECER A RELAÇÃO SUJEITO-BAIXO EM ALUNOS DO PRIMEIRO ANO DA CARREIRA DE ENGENHARIA INDUSTRIAL DA UNIVERSIDADE DE PINAR DEL RÍO “HERMANOS SAIZ MONTES DE OCA”

Lianet Meléndez Breijo ¹

Carlos Jesús Cruz Bracho ²

Resumen: En el presente trabajo se hace un análisis estadístico de los resultados obtenidos en las asignaturas, por los estudiantes de primer año de la carrera Ingeniería Industrial en cinco cursos académicos consecutivos. El objetivo del mismo está en conocer qué asignatura es la que más influye en que el estudiante tienda a ser baja. Para ello se hizo uso de métodos estadísticos (análisis clúster, prueba chi-cuadrada, regresión logística binaria y la prueba de Hosmer-Lemeshow) que se utilizaron para determinar la confiabilidad de los instrumentos aplicados, y para la obtención, procesamiento y análisis de los datos obtenidos.

Palabras claves: Análisis estadístico, Baja, Asignaturas.

Abstract: In the present work, a statistical analysis is made of the results obtained in the subjects, by the first-year students of the Industrial Engineering career in five consecutive academic years. The objective of the same is to know which subject is the one that most influences the student's tendency to be low. For this, statistical methods were used (cluster

¹ Máster en Dirección. Universidad de Pinar del Río Hermanos Saiz Monte de Oca (UPR), Pinar del Río, Cuba. Docente de la Facultad de Ciencias Técnicas.

Email: lianetmelendezbreijo@gmail.com

ORCID: <https://orcid.org/0000-0002-2149-6151>

² Máster en Ingeniería Industrial y Sistemas. Universidad de Pinar del Río Hermanos Saiz Monte de Oca (UPR), Pinar del Río, Cuba. Docente de la Facultad de Ciencias Técnicas.

Email:

ORCID: <https://orcid.org/0000-0002-2219-848X>

analysis, chi-square test, binary logistic regression and the Hosmer-Lemeshow test) that were used to determine the reliability of the applied instruments, and to obtain, process and analyze the results. data obtained.

Keywords: Statistical analysis, Low, Subjects.

Resumo: No presente trabalho é feita uma análise estatística dos resultados obtidos nas disciplinas, pelos alunos do primeiro ano da carreira de Engenharia Industrial em cinco anos lectivos consecutivos. O objetivo do mesmo é saber qual disciplina é a que mais influencia a tendência do aluno a ser baixo. Para isso, foram utilizados métodos estatísticos (análise de agrupamento, teste do qui-quadrado, regressão logística binária e teste de Hosmer-Lemeshow) que serviram para determinar a confiabilidade dos instrumentos aplicados e para obter, processar e analisar os resultados.

Palavras-chave: Análise estatística, Retirada, Sujeitos.

1. Introducción

Problema que se ha resuelto de acuerdo con los objetivos del trabajo.

En la actualidad, los docentes tienen la responsabilidad de orientar a los estudiantes para que exploren y desarrollen nuevos conocimientos facilitando el acceso de los estudiantes a los recursos, lo que ayuda a promover el desarrollo de habilidades y destrezas. De igual forma, a partir de la relación entre los conocimientos previos y la nueva información, las habilidades necesarias para comprender, aclarar e interpretar la información mediante el uso correcto de herramientas digitales para lograr un aprendizaje significativo. Niño Morante, *et.al* (2022)

En la Carrera Ingeniería Industrial los estudiantes de primer año presentan mucha tendencia a ser baja al finalizar el curso escolar. Esto es debido a un grupo de asignaturas que inciden directamente en la desmotivación por continuar en la carrera. Con la realización de esta investigación se pretende conocer cuál o cuáles son esas asignaturas que provocan dicho fenómeno, para así tomar medidas de perfeccionamiento docente.

Análisis estadístico a los estudiantes de primer año de la carrera Ingeniería Industrial de la Universidad de Pinar del Río “Hermanos Saiz Montes de Oca”

Se realizó un análisis estadístico donde se recopilaron los datos de 210

estudiantes, todos de primer año académico de la carrera ingeniería industrial, para conocer cuáles son las asignaturas que más están afectando sus índices académicos y llevándolos a abandonar la carrera.

Estos estudiantes están organizados en los cursos 2017-2018, 2018-2019, 2019-2020, 2021 y 2022. Se agregó una variable curso que toma un valor ordinal que se muestra entre paréntesis.

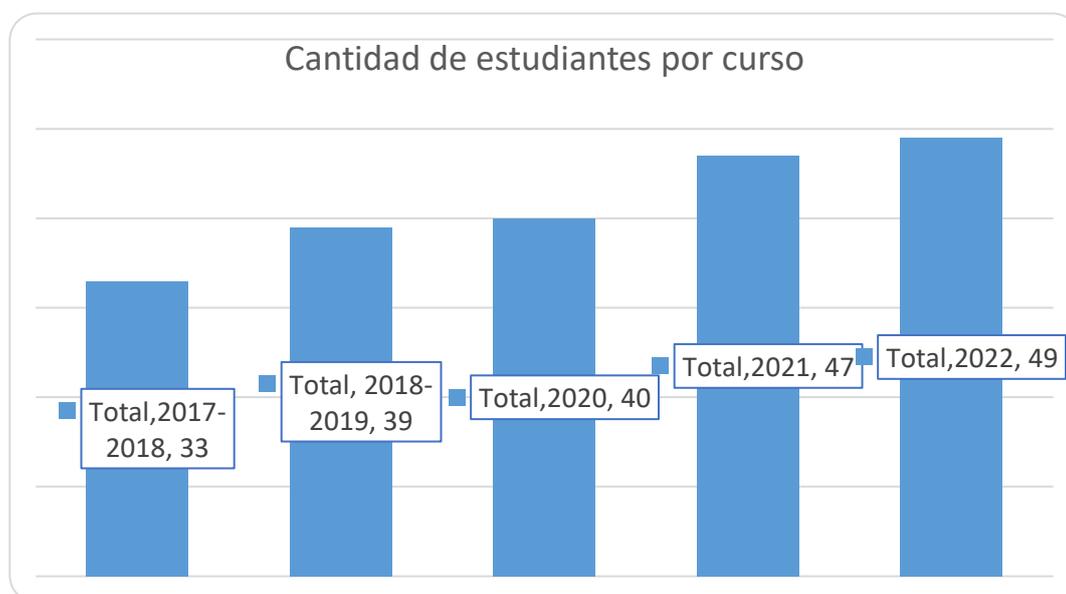
En el primer año académico se cursan 16 asignaturas, para este análisis se tomaron las notas de las asignaturas del primer semestre:

- Matemática I (M1)
- Álgebra Lineal (AL)
- Filosofía y Sociedad (FS)
- Introducción a la Informática (IINF)
- Historia de Cuba (HC)
- Inglés I (I1)
- Introducción a la ingeniería (IING)
- Educación Física I (EF1)

La matrícula de los estudiantes en cada curso se muestra a continuación:

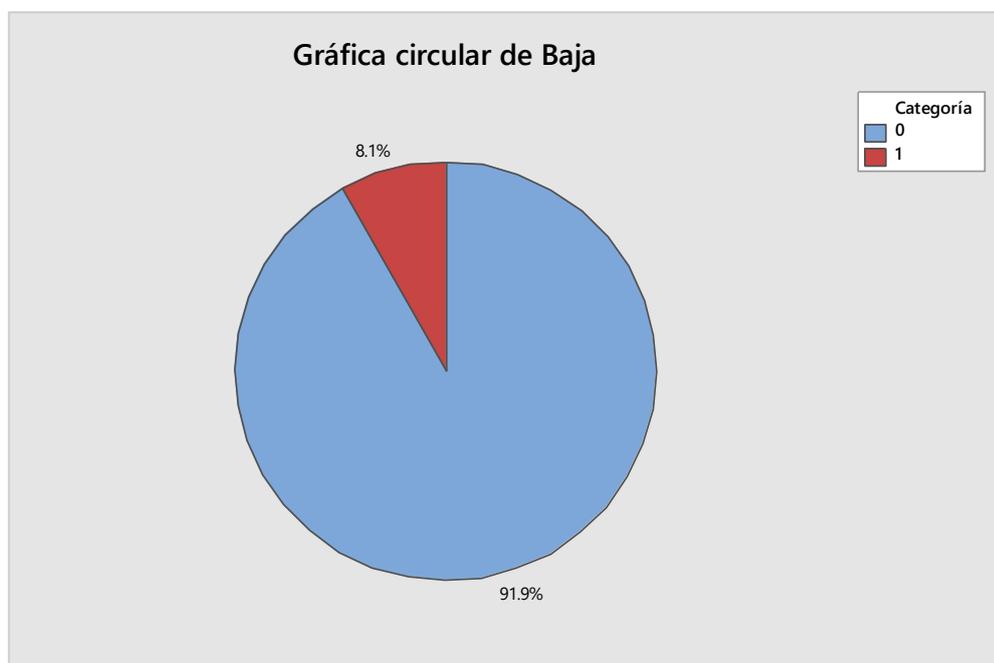
Se puede ver que la matrícula aumenta con cada curso que comienza.

Gráfica 1: Cantidad de estudiantes por curso escolar



Se creó una variable llamada baja, tomando valor 0 cuando ese estudiante no es baja y 1 cuando el estudiante presenta baja y no cursa el segundo semestre.

Gráfica 2: Porcentaje de baja en el primer semestre



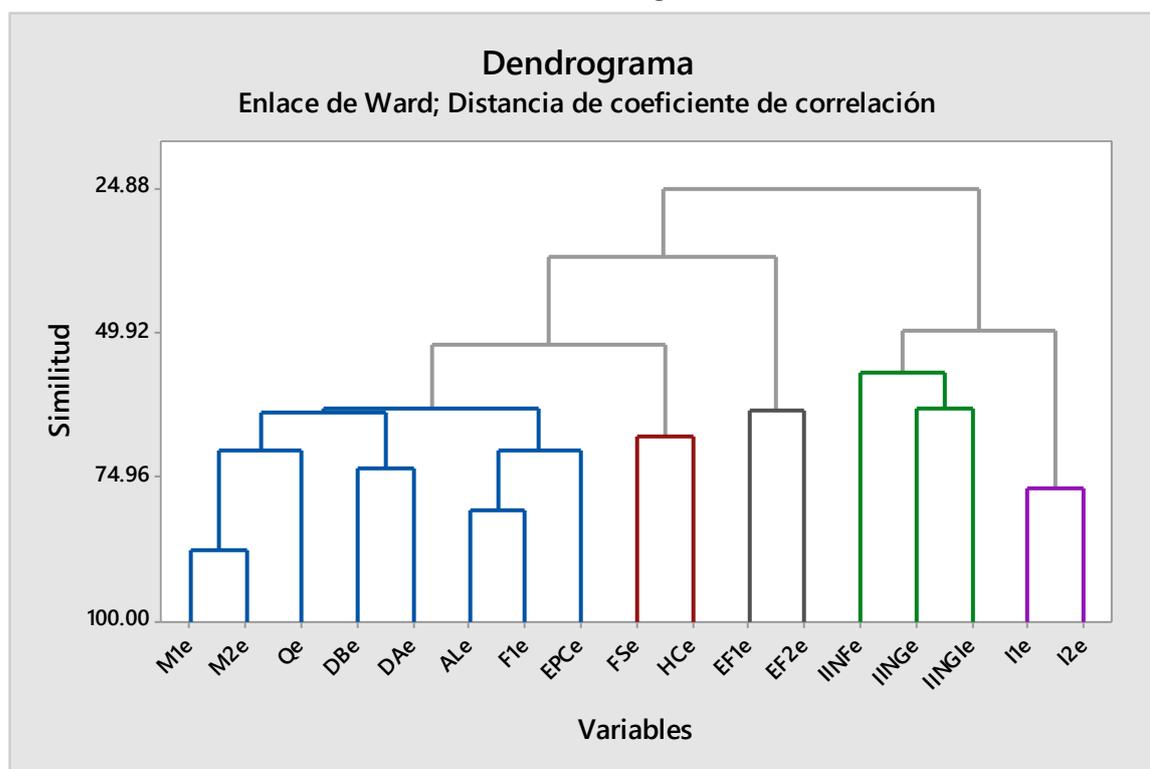
El 8.1 % de la matrícula fue baja luego de cursar el primer semestre.

Gráfica 3: Relación de notas del primer semestre



Las notas de los estudiantes durante todo primer año se normalizaron aplicándosele una transformación, luego se realizó un análisis clúster con el objetivo de evaluar la relación entre las asignaturas a través de sus notas. El resultado es el siguiente:

Gráfica 4: Relación asignaturas-notas



Para el análisis que se realizará más adelante es válido conocer la cercanía entre las notas de Filosofía e Historia de Cuba y de Introducción a la Informática con Introducción a la Ingeniería.

De los 210 estudiantes que inicialmente se tomaron para este análisis, muchos tenían notas faltantes, como resultados de convalidaciones, otros estudiantes no tenían nota en algunas asignaturas.

Se seleccionaron todos los estudiantes que tenían una nota en cada una de las asignaturas a analizar y las notas se transformaron en una variable dicotómica, la nota 2 se transformó en S (Suspendo) y las otras notas en A (Aprobado), un conteo realizado a las notas arrojó el resultado siguiente:

Cuenta de variables discretas: Matemática; Álgebra; Filosofía; Informática; Historia; ...

Matemática	Conteo	Porcentaje	Álgebra	Conteo	Porcentaje
A	126	66.32	A	146	76.84
S	64	33.68	S	44	23.16
N=	190		N=	190	

Filosofía	Conteo	Porcentaje	Informática	Conteo	Porcentaje
A	181	95.26	A	189	99.47
S	9	4.74	S	1	0.53
N=	190		N=	190	

Historia	Conteo	Porcentaje	Ingles	Conteo	Porcentaje	Ing	Conteo	Porcentaje
A	185	97.37	A	189	99.47	A	188	98.95
S	5	2.63	S	1	0.53	S	2	1.05
N=	190		N=	190		N=	190	

Educación Física	Conteo	Porcentaje	Baja	Conteo	Porcentaje
A	187	98.42	0	178	93.68
S	3	1.58	1	12	6.32
N=	190		N=	190	

Se puede ver como las asignaturas, matemática y álgebra son las que tienen la mayor cantidad de suspensos (un 33 %y 23% respectivamente) y las asignaturas introducción a la informática e introducción a la ingeniería son las que menos suspensos tienen. Solo el 6% de los estudiantes fue baja luego de cursar todo el primer semestre.

A continuación, se analiza la relación de cada una de las variables con la variable Baja, para eso se realizan pruebas de χ^2 dando los siguientes resultados:

Prueba chi-cuadrada para asociación: Baja; Matemática

Filas: Baja Columnas: Matemática

	A	S	Todo
0	124	54	178
	118.04	59.96	
1	2	10	12
	7.96	4.04	
Todo	126	64	190
Contenido de la celda:			Conteo
			Conteo esperado

Chi-cuadrada de Pearson = 14.135; GL = 1; Valor p = 0.000

Chi-cuadrada de la tasa de verosimilitud = 13.501; DF = 1; Valor p = 0.000

Prueba chi-cuadrada para asociación: Baja; Álgebra

Filas: Baja Columnas: Álgebra

	A	S	Todo
0	144	34	178
	136.78	41.22	
1	2	10	12
	9.22	2.78	
Todo	146	44	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 26.065; GL = 1; Valor p = 0.000

Chi-cuadrada de la tasa de verosimilitud = 21.218; DF = 1; Valor p = 0.000

Prueba chi-cuadrada para asociación: Baja; Filosofía

Filas: Baja Columnas: Filosofía

	A	S	Todo
0	174	4	178
	169.568	8.432	
1	7	5	12
	11.432	0.568	
Todo	181	9	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 38.713; GL = 1

Chi-cuadrada de la tasa de verosimilitud = 17.889; GL = 1

Prueba chi-cuadrada para asociación: Baja; Informática

Filas: Baja Columnas: Informática

	A	S	Todo
0	178	0	178
	177.063	0.937	
1	11	1	12
	11.937	0.063	
Todo	189	1	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 14.912; GL = 1

Chi-cuadrada de la tasa de verosimilitud = 5.605; GL = 1

Prueba chi-cuadrada para asociación: Baja; Historia

Filas: Baja Columnas: Historia

	A	S	Todo
0	177	1	178
	173.316	4.684	
1	8	4	12
	11.684	0.316	
Todo	185	5	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 47.120; GL = 1

Chi-cuadrada de la tasa de verosimilitud = 18.609; GL = 1

Prueba chi-cuadrada para asociación: Baja; Inglés

Filas: Baja Columnas: Inglés

	A	S	Todo
0	178	0	178
	177.063	0.937	
1	11	1	12
	11.937	0.063	
Todo	189	1	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 14.912; GL = 1

Chi-cuadrada de la tasa de verosimilitud= 5.605; GL = 1

Prueba chi-cuadrada para asociación: Baja; Introducción a la Ingeniería

Filas: Baja Columnas: Intro. Ing

	A	S	Todo
0	178	0	178
	176.126	1.874	
1	10	2	12
	11.874	0.126	
Todo	188	2	190
Contenido de la celda:		Conteo	
		Conteo esperado	

Chi-cuadrada de Pearson = 29.982; GL = 1

Chi-cuadrada de la tasa de verosimilitud = 11.381; GL = 1

Prueba chi-cuadrada para asociación: Baja; Educación Física

Filas: Baja Columnas: Educación Física

	A	S	Todo
0	177	1	178
	175.189	2.811	
1	10	2	12
	11.811	0.189	
Todo	187	3	190
Contenido de la celda:		Conteo	
		Conteo esperado	

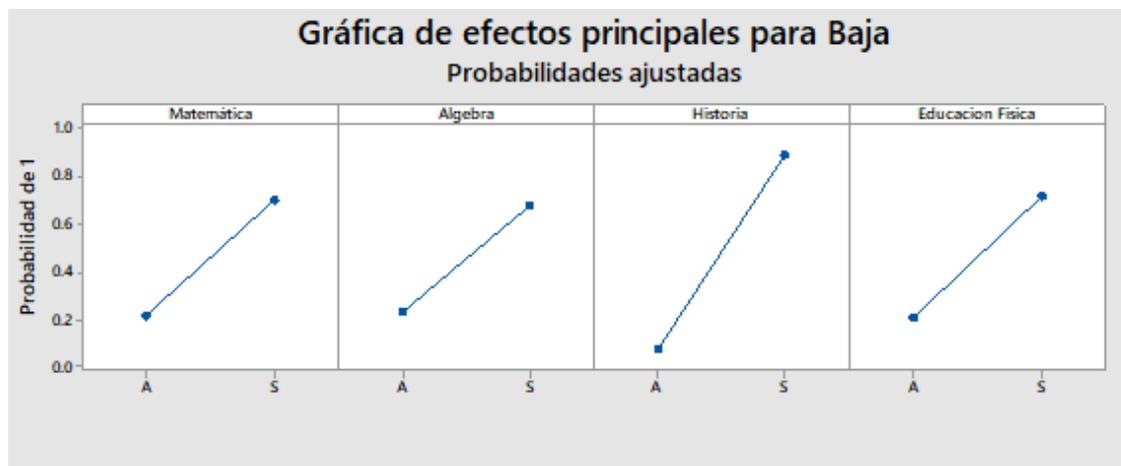
Chi-cuadrada de Pearson = 18.763; GL = 1

Chi-cuadrada de la tasa de verosimilitud = 7.671; GL = 1

Estas pruebas no tuvieron los resultados deseados ya que solo se tiene certeza que matemática y algebra se relacionan con la variable baja.

Un gráfico de efectos principales para la variable Baja nos brinda un poco de información respecto a la interacción de las variables independientes

Gráfica 5: Efectos principales para la baja



No se encuentran interacciones entre las variables, por tanto, no es necesario buscar su efecto en el modelo.

Luego de esto, se procederá a aplicar una regresión logística, para esto se fijó el método paso a paso con un alfa de entrada y salida de 0.10.

Regresión logística binaria: Baja vs. Matemática; Álgebra; Filosofía; Informática; ...

Método

Función de enlace Logit

Codificación de predictores categóricos (1; 0)

Filas utilizadas 190

Selección de términos escalonada

α a entrar = 0.1; α a retirar = 0.1

Información de respuesta

Variable Valor Conteo

Baja	1	12	(Evento)
	0	178	
	Total	190	

Tabla de desviaciones

	Desv.	Media		
Fuente	GL ajust.	ajust.	Chi-cuadrada	Valor
P				
Regresión	4	41.471	10.3677	41.47
0.000				
Matemática	1	4.318	4.3180	4.32
0.038				
Algebra	1	4.626	4.6262	4.63
0.031				
Historia	1	15.076	15.0759	15.08
0.000				
Educación Física	1	3.229	3.2294	3.23
0.072				
Error	185	48.046	0.2597	
Total	189	89.516		

Resumen del modelo

R-cuad.			
R-cuad. de (ajust) de			
desviación	desviación	AIC	
46.33%	41.86%	58.05	

La ecuación no tiene un buen ajuste, que podría deberse a la calidad de los datos

Coefficientes

EE del			
Término	Coef	coef.	VIF
Constante	-5.66	1.27	
Matemática			
S	2.13	1.19	1.25
Algebra			

S	1.926	0.967	1.05
Historia			
S	4.53	1.37	1.22
Educación Física			
S	2.30	1.32	1.05

El VIF tiene un buen valor, lo que me dice que la relación entre las variables independientes no está afectando los resultados del modelo

Ecuación de regresión

$$P(1) = \frac{\exp(Y')}{1 + \exp(Y')}$$

$$Y' = -5.66 + 0.0 \text{ Matemática_A} + 2.13 \text{ Matemática_S} + 0.0 \text{ Álgebra_A} + 1.926 \text{ Álgebra_S}$$

$$+ 0.0 \text{ Historia_A} + 4.53 \text{ Historia_S} + 0.0 \text{ Educación Física A} + 2.30 \text{ Educación Física S}$$

La ecuación devuelta por el procedimiento se analiza fácilmente considerando siempre quien se toma como evento. En este caso, 1, significa que el estudiante es baja, por tanto el signo + de los coeficientes me informa que las probabilidades son a ser baja, y mientras mayor es el coeficiente, la transformación logit me devolverá que la probabilidad de ser baja es mayor.

Dicho esto, se puede observar que estar suspenso en matemática, algebra y educación física convierten en probable el hecho de que un estudiante sea baja, sin embargo, quien más aporta a ello es historia

Pruebas de bondad del ajuste

Prueba	GL	Chi-cuadrada	Valor p
Desviación	185	48.05	1.000
Pearson	185	108.62	1.000
Hosmer-Lemeshow	2	0.51	0.775

La prueba de Hosmer-Lemeshow es un método para estudiar la bondad de ajuste del modelo de regresión logística que consiste en comparar los valores previstos (esperados) por el modelo con los valores realmente observados. Ambas

distribuciones, esperada y observada, se contrastan mediante una prueba de Chi-2. La hipótesis nula del test de Hosmer-Lemeshow es que no hay diferencias entre los valores observados y los valores pronosticados (el rechazo este test indicaría que el modelo no está bien ajustado).

En este caso se utilizaron todos los datos disponibles, por esta razón el valor de p en la prueba de desviación y la Pearson da 1, y tomando el estadígrafo de Hosmer-Lemeshow tiene un buen ajuste también.

Conclusiones

A partir del análisis estadístico realizado a los estudiantes de primer año de la carrera Ingeniería Industrial, se pudo determinar que estar suspenso en matemática, algebra y educación física, convierten en probable el hecho de que un estudiante sea baja, sin embargo, quien más aporta a ello es historia.

Aporte científico personal de cada autor

Lianet Meléndez Breijo (60%): Líder de la investigación. Concibió la idea original del mismo. Estableció de conjunto con otros autores los protocolos de trabajo. Participó en el procesamiento de datos y descripción de resultados. Autor de artículos y ponencias en eventos científicos.

Carlos Jesús Cruz Bracho (40%): Participó en el diseño de la elaboración de la idea original de la investigación, también colaboró en la descripción de los resultados obtenidos. Autor de artículos y ponencias en eventos científicos.

Bibliografía

Rendón-Macías ME, Villasís-Keever MÁ, Miranda-Novales MG. Estadística descriptiva. Rev Alerg Mex. 2016;63(4):397-407. Disponible en: <http://revistaalergia.mx/ojs/index.php/ram/article/view/230>

Villasís-Keever MA, Miranda-Novales MG. El protocolo de investigación II: los diseños de estudio para investigación clínica. Rev Alerg Mex. 2016;63(1):80-90.

Disponible en: <http://revistaalergia.mx/ojs/index.php/ram/article/view/163>

Villasís-Keever MA, Miranda-Navales MG. El protocolo de investigación IV: las variables de estudio. Rev Alerg Mex. 2016;63(3):303-310. Disponible en: <http://revistaalergia.mx/ojs/index.php/ram/article/view/199>

Gómez-Gómez N, Danglot-Banck C, Vega-Franco L. Cómo seleccionar una prueba estadística (primera de dos partes). Rev Mex Pediatr. 2013;80(1):30-34. Disponible en: <http://www.medigraphic.com/pdfs/pediat/sp2013/sp131g.pdf>

Gómez-Gómez N, Danglot-Banck C, Vega-Franco L. Cómo seleccionar una prueba estadística (segunda parte). Rev Mex Pediatr 2013;80(2):81-85. Disponible en: <http://www.medigraphic.com/pdfs/pediat/sp-2013/sp132g.pdf>

Shankar S, Singh R. Demystifying statistics: How to choose a statistical test? Indian J Rheumatol. 2014;9(2):77-81. DOI: <https://doi.org/10.1016/j.injr.2014.04.002>

Talavera JO, Rivas-Ruiz R. Investigación clínica IV. Pertinencia de la prueba estadística. Rev Med Inst Mex Seguro Soc. 2011;49(4):401-405. Disponible en: <http://www.redalyc.org/articulo.oa?id=457745503010>

Sato Y, Goshō M, Nagashima K, Takahashi S, Ware JH, Laird NM. Statistical methods in the journal - An update. N Engl J Med. 2017;376(11):1086-1087.

Fajardo-Gutiérrez A. Medición en epidemiología: prevalencia, incidencia, riesgo, medidas de impacto. Rev Alerg Mex. 2017;64(1):109-120. Disponible en: <http://revistaalergia.mx/ojs/index.php/ram/article/view/252>

Berea-Baltierra R, Rivas-Ruiz R, Pérez-Rodríguez M, Palacios-Cruz L, Moreno J, Talavera JO. Investigación clínica XX. Del juicio clínico a la regresión logística múltiple. Rev Med Inst Mex Seguro Soc. 2014;52:192-197. Disponible en: https://www.cmu.org.mx/media/cms_page_media/57/39-INVESTIGACION-CLINICA-DEL-JUICIOCLINICO-A-LA-REGRESION.pdf

Niño Morante, N.R, Uceda Bazán, M.N., Fernández Otoyá, F.A., García González, M. (2022). Estrategias didácticas para promover el aprendizaje significativo dirigido a estudiantes universitarios. Revista Mendive Vol. 20 No. 4 (octubre-diciembre). pp. 1297-1309 2022 Disponible en: <https://mendive.upr.edu.cu/index.php/MendiveUPR/article/view/3090> ISSN. 1815-7696 RNPS 2057.

Manuscrito recibido el: 20 de marzo 2023.

Aprobado: 16 de abril de 2023.

Publicada: 15 de mayo de 2023.