

Artículo original

Predicción del consumo energético mediante un modelo matemático de regresión lineal a partir de la producción diaria en la planta envasadora Q'Agua

Walter Jácome-Velez^[1]  Joselyn Vilela-Sabando^[1]  Mayerli Cedeño-Hormaza^[1]  Cristian Jácome-Duarte^[1] 

[1] Facultad de Ciencia de la Industria y Producción. Universidad Técnica Estatal de Quevedo (UTEQ). Quevedo, Ecuador.



Autor para correspondencia: wjacomev@uteq.edu.ec

Resumen

En las industrias, la eficiencia energética se ha convertido en un desafío, donde en las plantas envasadoras de agua, tienen altos costos de operación debido al consumo de electricidad. El objetivo de este artículo fue predecir el consumo de energía eléctrica de la planta envasadora de agua Q'Agua mediante un modelo de regresión lineal simple, donde la producción diaria es la variable independiente y los kWh la variable dependiente. Los datos se registraron mediante observación directa, escribiendo el total de unidades envasada, así como también el consumo diario de electricidad. En los resultados el coeficiente de determinación (R^2) que es del 99.27% mostraron una correlación positiva extraordinaria entre la producción y la demanda de energía. El análisis de Kolmogorov-Smirnov mostró desviaciones respecto a la normalidad en los residuos, aunque el modelo en general mostró ser significativo ($P < 0.05$). Este estudio ofrece una herramienta de predicción que optimice la gestión energética en las prácticas sostenibles de los procesos industriales.

Palabras Clave: *eficiencia energética; regresión lineal; producción; predicción; sostenibilidad.*

Article

Prediction of energy consumption using a linear regression mathematical model based on daily production at the Q'Agua bottling plant

Abstract

In industries, energy efficiency has become a challenge, where in water bottling plants, they have high operating costs due to electricity consumption. The objective of this article was to predict the electrical energy consumption of the Q'Agua water bottling plant using a simple linear regression model, where daily production is the independent variable and kWh the dependent variable. The data were recorded by direct observation, writing down the total units packaged as well as the daily electricity consumption. In the results, the coefficient of determination (R^2), which is 99.27%, showed an extraordinary positive correlation between energy production and demand. The Kolmogorov-Smirnov analysis showed deviations from normal in the residuals, although the overall model was significant ($P < 0.05$). This study offers a prediction tool that optimizes energy management in sustainable practices of industrial processes.

Keywords: *energy efficiency; linear regression; production; prediction; sustainability.*

1. Introducción

En la actualidad, la sostenibilidad y la eficiencia energética se han convertido en metas esenciales en la gestión industrial, la capacidad de prever el consumo de energía se ha vuelto fundamental para las organizaciones, en este contexto, las plantas envasadoras de agua enfrentan el desafío de operar bajo condiciones dinámicas que influyen directamente en su gasto energético (Poveda, 2007).

Dado que este factor repercute en los costos operativos y en la huella ambiental de las organizaciones, los modelos matemáticos surgen como herramientas para el análisis y la proyección de datos relacionados con su utilización (Gonçalves Fernandes et al., 2017).

Una de las metodologías utilizadas en este ámbito es la regresión lineal simple, el cual es un enfoque que permite establecer una relación entre variables independientes y dependientes a través de una función lineal.

Al comprender la relación entre la producción diaria y el consumo energético, las empresas pueden ajustar sus estrategias de operación, implementando medidas de eficiencia energética que respondan a la demanda proyectada.

En el caso de la planta envasadora de agua, la producción diaria puede ser considerada una variable independiente que influye directamente en el consumo energético de la planta. Mediante el uso de un modelo de regresión lineal, se puede analizar cómo las variaciones en la producción impactan el gasto energía diario (Hechavarria Pérez et al., 2015).

El software R-Studio se presenta como una herramienta integral para la implementación de este tipo de análisis, ya que nos proporciona un entorno intuitivo para realizar modelos estadísticos y predictivos, mediante la combinación de datos de producción diarios y consumo energético.

RStudio simplifica la tarea de limpiar y organizar datos, utilizar modelos estadísticos como la regresión lineal, crear gráficos para visualizar tendencias y evaluar la precisión de modelos predictivos. Estas capacidades mejoran la comprensión de cómo las variables afectan los resultados y garantizan información confiable para respaldar decisiones (Caselli Gismondi & Beltrán Canessa, 2025).

En esta investigación, se presentará un modelo matemático de regresión lineal aplicado a una planta envasadora de agua, donde se va a estudiar la metodología utilizada para la recolección y análisis de datos, así como los resultados obtenidos a partir de la implementación del modelo en R-Studio.

Se espera que este estudio contribuya a la mejora del rendimiento energético en la planta envasadora de agua y que también sirva de referencia para otras industrias que enfrentan desafíos similares. A medida que la conciencia crece sobre la sostenibilidad energética, la implementación de técnicas de análisis de datos, como la regresión lineal en programas como R-Studio, se vuelve cada vez más relevante.

La regresión lineal es una herramienta fundamental en el análisis estadístico que nos permite comprender y predecir cómo se comporta una variable dependiente en relación con una variable independiente, siempre que ambas sean continuas. La forma más básica, conocida como regresión lineal simple, establece esta conexión utilizando únicamente una variable predictiva y una función lineal (Espinosa Martínez, 2025).

2. Background

En su investigación (Arellano & Peña, 2020), desarrollaron dos modelos de análisis de regresión lineal con múltiples factores para pronosticar el consumo de agua potable en el sector residencial, basados en indicadores sociodemográficos, socioeconómicos, administrativos y de calidad del agua, así como en factores climáticos.

El primer modelo, con una precisión del 80,87 %, estima el consumo semestral per cápita a partir de 19 variables, mientras que el segundo, con una precisión del 38,88 %, pronostica el consumo mensual utilizando solo 6 variables. Ambos enfoques son herramientas útiles para garantizar una distribución racional de los recursos hídricos, siempre que la información se actualice periódicamente.

(Fernández Lizana, 2020), plantea R como una herramienta básica para el análisis y la presentación de datos en las ciencias sociales, resaltando sus ventajas frente a los programas estadísticos comerciales. También señala que R, al ser gratuito, versátil y de código abierto, se ha convertido en un lenguaje común en el campo de la estadística, capaz de ejecutar una amplia gama de procedimientos y crear gráficos de alta calidad. Por otra parte, destaca que su creciente aceptación en los ámbitos académico y de investigación lo convierte en una opción sólida y accesible para los científicos sociales, en consonancia con la filosofía del software libre.

En su estudio, (Fernando Cardona Madariaga et al., 2013), aplicaron un modelo de regresión lineal al análisis de la pobreza en Colombia, utilizando datos de 2010 y 2011 relativos a las trece ciudades más grandes del país. A partir de esta información, los autores determinaron que las tasas de crecimiento de la pobreza en 2011 se encontraban entre el 0,83 % y el 1,053 % por cada 1 % de crecimiento registrado en 2010.

El estudio permitió realizar pronósticos sobre la situación de la pobreza en estas ciudades y proponer posibles orientaciones de política pública destinadas a reducir los indicadores observados. Los investigadores subrayaron la importancia de ser conscientes de que los modelos de regresión, aunque son útiles para describir las interrelaciones y obtener estimaciones, no permiten establecer relaciones causales entre las variables analizadas.

Autores, (Avello Martínez & Losa Seisdedo, 2017), destacan la relevancia de utilizar herramientas estadísticas en la ciencia, centrándose en el papel de R como un software gratuito y de código abierto para el análisis de datos. Los autores subrayan que este entorno de programación no solo facilita la creación de modelos predictivos y el manejo de cálculos complejos, sino que también fomenta valores como la colaboración, la reproducibilidad y el acceso libre al conocimiento, resaltando las ventajas de su sintaxis amigable, su amplia comunidad de usuarios y las diversas interfaces desarrolladas, como R-Commander, RKWard y RStudio, que amplían sus capacidades y hacen su uso más sencillo. R es una alternativa robusta y sostenible frente a los programas estadísticos comerciales, lo que ayuda a promover una investigación científica más accesible e independiente.

En el artículo titulado "RStudio y su importancia en la educación", (Rivera Cruz, 2022), explora cómo RStudio se integra en la enseñanza de las matemáticas. Resalta su papel como una herramienta educativa que no solo facilita el aprendizaje, sino que también mejora la comprensión y aumenta la motivación de los estudiantes. El autor sostiene que, en un contexto donde los métodos de enseñanza están en constante evolución y las tecnologías de la información y la comunicación tienen un gran impacto, el uso de programas como RStudio puede enriquecer la resolución de problemas y transformar las prácticas educativas, fortaleciendo así las competencias básicas de los estudiantes.

La relevancia de este estudio radica en que (Rodríguez Mañay et al., 2015), utilizaron un modelo de regresión lineal para identificar los parámetros más adecuados para calcular la tasa estándar de los gastos generales de producción. Al analizar diferentes variables, como el tiempo de funcionamiento de las máquinas, el tamaño de los lotes de producción y las horas de mano de obra indirecta, concluyeron que el tiempo de funcionamiento de las máquinas era el factor más preciso para explicar estos costos. Esta herramienta resulta ser muy útil para los contadores, ya que se puede aplicar a cualquier tipo de empresa manufacturera y refuerza la importancia de las estadísticas en el análisis contable.

En la investigación de (De Lucas Coloma, 2018), utiliza regresión lineal, el cual ha demostrado que esta técnica es

muy útil para analizar cómo se relacionan distintas variables. Por ejemplo, un estudio sobre las emisiones de CO₂ en Ecuador mostró cómo el crecimiento de la población influye en las emisiones de combustible, evidenciando una relación directa.

Este enfoque demuestra que la regresión lineal permite identificar factores clave y predecir patrones a partir de datos confiables. De manera similar, en la industria, esta herramienta ayuda a anticipar el consumo de energía según la producción diaria, mejorando la eficiencia de los recursos y apoyando la toma de decisiones. Así, los métodos aplicados en estudios ambientales sirven como referencia para crear modelos predictivos en eficiencia energética (De Lucas Coloma, 2018).

3. Metodología

La metodología empleada en esta investigación es de observación directa y analítica, su estructura consta dos fases principales. La primera fase consiste en un registro diario de la producción total y de los kWh. Esta etapa permite contextualizar el estado actual del campo de estudio y evaluar las tendencias y patrones en el consumo de energía eléctrica. En la segunda fase, se lleva a cabo una actividad en un contexto no experimental para realizar un análisis cuantitativo.

Con la finalidad de estudiar este modelo matemático, se utilizan los datos tomados de una muestra real, extraídos del medidor de consumo energético que nos revela diariamente los kilovatios-hora (kWh) que se han consumido a lo largo de la producción diaria de envasado de botellas y la producción total de llenados de botellas es una jornada diaria de trabajo.

3.1 Modelo de Regresión lineal simple

$$y = \beta_0 + \beta_1 x + \varepsilon$$

β_0 y β_1 se conocen como los parámetros del modelo, y la letra griega épsilon (ε) es una variable aleatoria que se conoce como el error. El término del error da cuenta de la variabilidad de y que no puede ser explicada por la relación lineal entre "X" y "Y" (David R et al., 2008).

Las distancias entre los puntos y la recta de regresión se denominan residuos. Contienen la parte de la respuesta que no se explica mediante la ecuación de regresión, es decir, la diferencia entre el valor aproximado y el valor observado es el residuo.

En cualquier análisis de regresión, indicará que algunos puntos están más cerca de la línea y otros mucho más alejados, entre más cerca se encuentren los puntos a la línea, mejor será el ajuste entre la línea de regresión y el dato. Los

residuos permiten verificar la ecuación con el fin de demostrar que tanto se ajusta la línea a los datos (Carrasquilla-Batista et al., 2016).

3.2 Tipos de relación entre dos variables

En general, existen cuatro tipos posibles de relación entre variables (ver **Figura 1**), a saber: relación lineal directa, relación lineal inversa, relación no lineal directa y relación no lineal inversa, cuya estructura formal y funcional permite explicar objetivamente las actividades destinadas a decidir qué ecuación utilizar, qué ecuación se ajusta mejor a los datos y cómo validar la significación estadística de las predicciones realizadas (Fernando Cardona Madariaga et al., 2013).

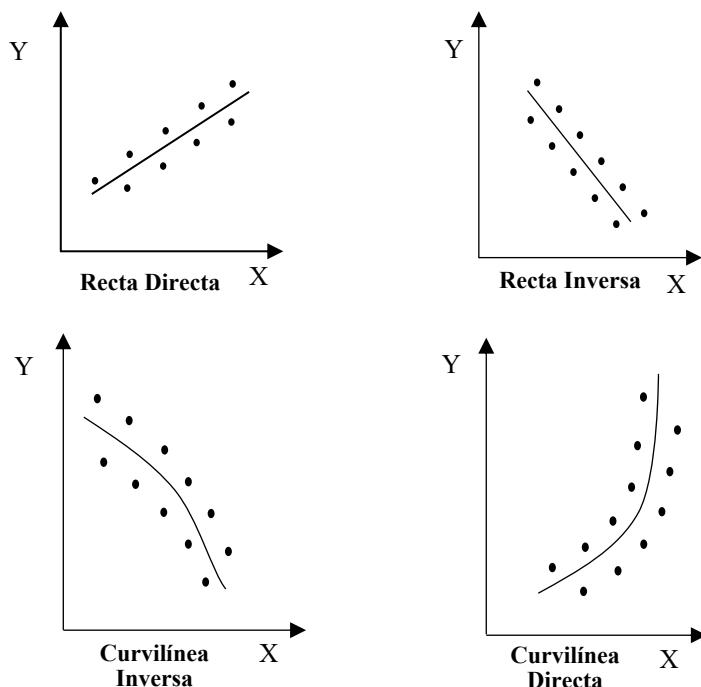


Figura 1. Tipos de relación entre dos variables

3.1 Procedimiento

El estudio del modelo matemático se llevó a cabo en la planta envasadora de Q'Agua, utilizando como unidades de medida los datos de producción diaria (número de botellas llenas al día) y el consumo de energía (en kWh). El propósito de este estudio fue desarrollar un modelo matemático que pudiera explicar y predecir el comportamiento de la demanda de electricidad en función del número de botellas llenas durante las horas de trabajo.

El desarrollo abarcó tres fases principales:

- *Recolección de datos:*

Los datos se recopilaron durante días laborables consecutivos. En este periodo se tuvo en cuenta la producción diaria de bidones llenos y el consumo eléctrico, mediante lecturas obtenidas directamente de los medidores de la fábrica como de la cantidad de bidones, por observación.

Los registros se verificaron con el número de bidones producidos y el nivel de consumo, lo que garantizó la trazabilidad y la eliminación de posibles inconsistencias o valores atípicos.

- *Ánálisis exploratorio*

El análisis exploratorio se centró en una primera evaluación de los datos recopilados, con el propósito de comprender su comportamiento e identificar posibles relaciones entre la producción y el consumo de electricidad. Se elaboró una serie gráficos para identificar tendencias lineales y efectuar cálculos de estadísticos descriptivos, tales como medias y gráficos. Este análisis desempeñó un papel crucial a la hora de determinar la validez del uso de modelos de regresión lineal como instrumento de predicción.

- *Modelado estadístico*

En el modelo estadístico, se utilizó una regresión lineal simple donde el consumo de energía se definió como la variable dependiente en kWh, mientras que la producción de bidones de agua actuó como la variable independiente.

Los parámetros del modelo se estimaron a través del método de mínimos cuadrados, y se revisaron las hipótesis de validez: linealidad, homocedasticidad, normalidad de los residuos e independencia temporal, además, se calcularon los intervalos de confianza para los coeficientes estimados. Este enfoque permitió crear un modelo predictivo que puede ser muy útil en la gestión energética (Moral Peláez, 2006).

Materiales

Los materiales empleados en esta investigación incluyeron principalmente medidores de energía, fuentes de información documental y software estadístico, que fueron esenciales para el análisis y el diseño metodológico de la investigación. Cada uno de estos elementos fue seleccionado con mucho cuidado para asegurar la precisión de los datos, la exactitud del análisis y la relevancia de los resultados.

Equipos de medición

Para la medición de los kilovatios-hora, se obtienen los datos de medidores digitales de energía eléctrica colocados en la planta, con registros diarios en kWh. Las mediciones se

realizaron siempre a la misma hora del día, lo que garantizó la uniformidad y comparabilidad de las series de datos.

Fuentes de información

Las principales fuentes de información corresponden a los informes diarios de producción de la planta y a los datos sobre el consumo energético. Los datos de producción se obtuvieron de los reportes oficiales del departamento operativo, mientras que los datos de consumo se obtuvieron del sistema digital de suministro energético.

Software

El procesamiento de datos y la creación de modelos estadísticos se realizaron utilizando R-Studio, un software especializado en análisis estadístico y visualización gráfica. En consecuencia, se utilizaron múltiples paquetes para facilitar la recopilación de datos, la generación de diagramas de dispersión y la estimación de regresiones lineales. El uso de R-Studio garantizó la fiabilidad analítica y la transparencia metodológica, lo que permitió la elaboración de informes y resultados claros. Esta herramienta desempeñó un papel crucial en la obtención de conclusiones sólidas basadas en pruebas empíricas.

Configuración de equipos

Para llevar a cabo el análisis de regresión lineal, se empleó tanto una computadora de escritorio como una laptop, ambos con la potencia necesaria para manejar datos y ejecutar aplicaciones estadísticas. Cada uno de estos equipos tenía instalado el programa R-Studio, lo que facilitó el desarrollo, la práctica y la aplicación del modelo matemático de manera efectiva.

La configuración de las máquinas proporcionó un entorno estable para manipular datos, ejecutar códigos, generar gráficos y obtener resultados precisos, lo que garantizó la confiabilidad y la reproducibilidad del análisis.

Recolección de datos

Para la recolección de datos se utilizó el método de observación directa, donde se elaboró una tabla de los valores tomados diariamente del consumo energético y de la producción diaria de bidones llenos.

Tabla 1: Recolección de datos

Días	Bidones llenos	kWh consumido
1	115	15
2	512	26
3	267	16
4	128	24
5	445	23
6	237	14
7	102	9
8	477	24
9	291	16
10	220	13
11	535	27
12	168	12
13	300	17
14	388	21
15	441	23
16	463	24
17	395	21
18	110	9
19	199	13
20	458	23
21	198	13
22	184	12
23	128	10
24	235	14
25	358	19
26	262	16
27	181	12
28	385	20
29	262	16
30	419	22
31	596	29
32	139	10
33	493	25
34	453	23
35	166	12
36	429	22
37	391	21
38	164	12
39	331	18
40	567	28
41	409	22
42	225	13
43	135	11
44	295	17
45	303	17
46	189	13
47	126	10
48	403	21
49	139	11
50	289	17
51	377	20
52	353	18
53	397	21
54	204	14
55	585	29
56	359	18
57	298	18
58	477	24
59	149	11
60	138	11
61	265	16
62	440	23
63	270	15
64	113	10
65	433	22

4. Resultados

4.1. Estadística descriptiva

Análisis de la variable producción

El histograma y la tabla de frecuencias (ver **Figura 2, Tabla 2, 3**) muestran cómo se distribuye la producción de bidones de agua en siete clases. El primer intervalo [102,173] presenta la mayor concentración de datos, con 16 observaciones que representan el 23,88 % del total. Esto indica una producción inicial elevada en comparación con otras áreas. La frecuencia tiende a disminuir a medida que aumenta el valor de producción, pero en los intervalos medios, como [384,455], se observan ligeros aumentos. El histograma muestra una distribución irregular con fluctuaciones que indican cierta dispersión en la producción, sin que se aprecie un patrón de simetría evidente.

Tabla 2: *Medidas de tendencia central de la variable producción*

Media	Mediana	Moda	Desviación Estándar
302.63	267	128	163.40

Tabla 3: *Tabla de frecuencia de la variable producción*

Clase	ni	Ni	F1	Fi acumulado
[102, 173)	16	16	0.2388	0.2388
[173, 243)	10	26	0.1492	0.3880
[243, 314)	12	28	0.1791	0.5671
[314, 384)	5	43	0.0746	0.6417
[384, 455)	14	57	0.2089	0.8507
[455, 525)	6	63	0.0895	0.9402
[525, 596)	4	67	0.0597	1

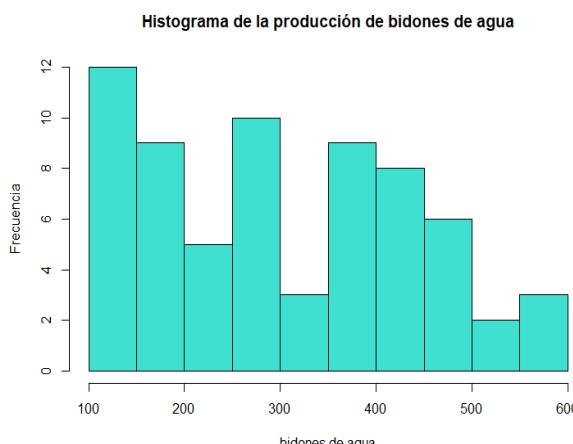


Figura 2. *Histograma de la variable producción.*

Al realizar la tabla de frecuencias y preparar el histograma del consumo energético en kWh (ver **Figura 3, Tabla 5, 6**), también podemos observar que las frecuencias más elevadas se hallan en el intervalo medio [12,4, 15,7] y en el intervalo [19, 22,4]; ambos intervalos han sido representados por 12 y 15 observaciones, que significa el 17,91 % y el 22,38 % respectivamente. Esto haría suponer que el consumo mayormente se halla en el intervalo medio, indicando por su parte que las costumbres de consumo son relativamente estables. En lo que se refiere a los extremos, tanto los valores mínimos [5,73, 9,07] como los valores máximos [25,7, 29] se observan con poca frecuencia, lo que significa que en el grupo de investigación no se generan frecuentemente consumos inusualmente altos o muy bajos.

Tabla 4: *Medidas de tendencia central de la variable kWh*

Media	Mediana	Moda	Desviación Estándar
18.45	16	22	5.75

Tabla 4. *Tabla de frecuencia de la variable kWh*

Clase	ni	Ni	F1	Fi acumulado
[5.73, 9.07)	3	3	0.0447	0.0447
[9.07, 12.4)	13	16	0.1940	0.2388
[12.4, 15.7)	10	26	0.1492	0.3880
[15.7, 19)	15	41	0.2238	0.6119
[19, 22.4)	12	53	0.1791	0.7910
[22.4, 25.7)	10	63	0.1492	0.9402
[25.7, 29)	4	67	0.0597	1

Histograma del consumo energético kWh

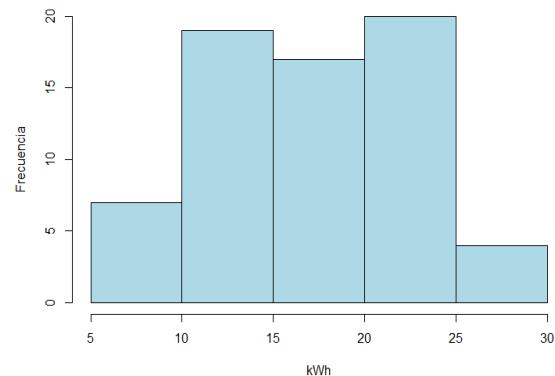


Figura 3. *Histograma de la variable kWh.*

Cada punto de la recta (ver **Figura 4**) representa una observación, es decir, cuánta más energía se consumió cuando se produjo una determinada cantidad de bidones llenos. El gráfico muestra que a medida que aumenta la producción, también aumenta el consumo de energía y su

relación es bastante continua y ordenada donde no se observan valores atípicos que se desvien de la tendencia.

Gráfico exploratorio de los KWh en base a la producción

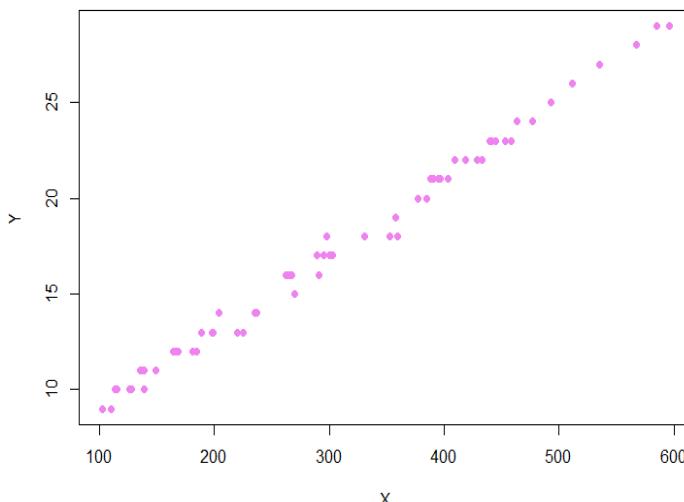


Figura 4: Diagrama realizado en R Studio

4.2. Ajuste del modelo

El modelo de regresión lineal que relaciona la variable Y con la variable X (ver **Figura 5**) muestra un ajuste significativo. El intercepto es 4.9592 y el coeficiente asociado a las horas de estudio es 0.040, lo cual indica que, por cada hora adicional de estudio, la nota final se incrementa en promedio en 0.075 puntos, manteniendo las demás condiciones constantes. El valor p- value es < 2e-16, lo cual evidencia que esta relación es significativa. El R2 es de 0.9927 lo que representa un 99.27% de la variabilidad de los kWh a partir de la producción diaria de bidones de agua, siendo un ajuste estable.

Gráfico exploratorio de los KWh en base a la producción

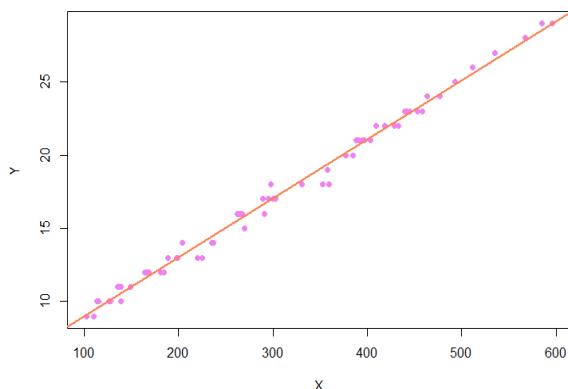


Figura 5: Ajuste del modelo realizado en RStudio

El error estándar de los residuos es de 0.4801 lo que significa que es bajo y permite confirmar que las predicciones de los modelos sólidos, estos resultados indican que la relación entre los kWh y la producción de bidones diarios es positiva y estadísticamente significativa.

4.3. Modelo lineal ajustado

$$y = \beta_0 + \beta_1 x + \varepsilon$$

$$y = 4.9592 + 0.0403x + \varepsilon$$

$$Y = 4.9592389 + 0.403070 * X + \varepsilon$$

Los resultados de la evaluación del modelo de residuos utilizando el test de Kolmogorov-Smirnov muestran un valor D = 0,095886 y un valor p = 0,5885, dado que estos resultados superan el nivel de significancia ($\alpha = 0,05$), se puede argumentar que los residuos no siguen la distribución normal. Por lo tanto, se rechaza la hipótesis nula y se acepta la hipótesis alternativa. Esto indica que los errores del modelo no siguen una distribución normal, lo que puede afectar a la validez de los determinados análisis estadísticos.

Los puntos del diagrama de dispersión y de los valores ajustados (ver **Figura 6**) se distribuyen aleatoriamente alrededor de la línea central cero, no se aprecian patrones claros o concentraciones que indiquen una tendencia al alzar o a la baja, demostrando que si existe homocedasticidad. Esto se debe a que la dispersión de los residuos se mantiene constante a lo largo de los valores previstos. Así mismo, no se observan asimetrías significativas ni acumulaciones excesivas lo cual nos indica que los errores están distribuidos uniformemente, lo que respalda la hipótesis de la normalidad de los residuos.

5. Discusión

El modelo de regresión lineal simple muestra una correlación positiva significativa entre los kWh y la producción diaria de bidones de agua, siendo el $R^2 = 99\%$ de la varianza en el rendimiento académico. El P-value = 2.2e-16 siendo menor a 0.05 nos indica que el modelo de regresión lineal simple es aceptable.

Los resultados de la prueba de Kolmogorov-Smirnov ($p = 0,5885$) muestran que el valor p supera el nivel de significación ($\alpha = 0,05$), siguiendo con este criterio, se rechaza la hipótesis nula, por lo que podemos decir que los residuos no siguen una distribución normal. La comprobación de la normalidad de los residuos es fundamental en el análisis estadístico que utiliza este modelo, ya que aumenta la fiabilidad de las conclusiones y los resultados.

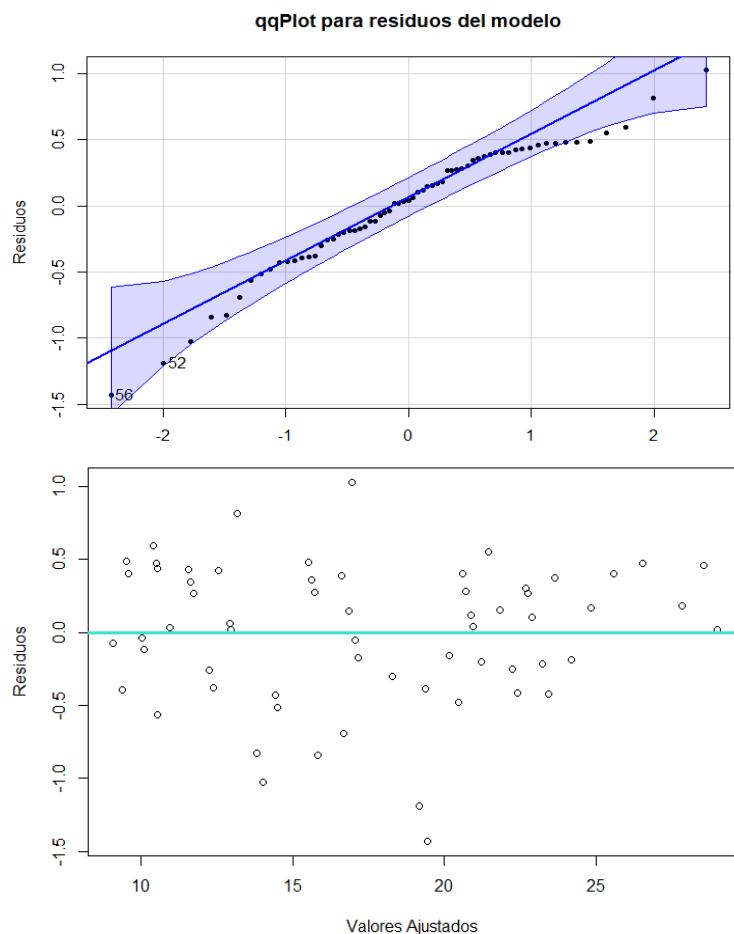


Figura 6: Plot para residuos del modelo.

6. Conclusiones

Los resultados del modelo de regresión lineal simple que combina la producción diaria de bidones de agua (variable independiente) y el consumo de energía en kilovatios·hora (variable dependiente), revela una fuerte correlación positiva entre estas dos variables, con un coeficiente de determinación R^2 del 99,27 %, lo que indica que la producción explica casi toda la varianza del consumo de energía. El modelo resultante $Y = 4,959 + 0,0403X$ es estadísticamente significativo (valor $p < 2e-16$). Aunque la prueba de Kolmogorov-Smirnov sugiere que los residuos no siguen una distribución normal ($p = 0,5885$), el análisis gráfico confirma la simetría para predecir la actividad de la planta.

Referencias

- Arellano, A., & Peña, D. (2020). Modelos de regresión lineal para predecir el consumo de agua potable. *Sinergia, Revista de Ciencia, Ingeniería y Tecnología*. 3(1). <https://novasinergia.unach.edu.ec/index.php/novasinergia/article/view/144/158>
- Avello Martínez, R., & Losa Seisdedo, A. (2017). El procesamiento estadístico con R en la investigación científica. *Medisur, Revista de Ciencias Médica Universidad Cien Fuegos*. 15(5) <http://medisur.sld.cu/index.php/medisur/article/view/3662>
- Beltrán Canessa, P. (2025). *Modelo Predictivo de agroexportaciones para la Región La Libertad aplicando Data Science con RStudio*. [Tesis de posgrado. Universidad Nacional del Santa]. Obtenido de

https://alicia.concytec.gob.pe/vufind/Record/UNSR_4ffbb166d948e629ce814a789be28f3e/Details

Cardona, D., González, J., Rivera, M., & Cárdenas Vallejo, E. (2013). Aplicación de la regresión lineal en un problema de pobreza. *Revista Interacción*, 12, 73-84. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=http://revistas.unilibre.edu.co/index.php/interaccion/article/download/2315/1767&ved=2ahUKEwj-7_7muLSSAxWCRjABHd_rHWAQFnoECBqQAO&usg=AOvVaw3154SfLlgWXv9XC5SX8kh3

Carrasquilla-Batista, A., Chacón-Rodríguez, A., Núñez-Montero, K., Gómez-Espinoza, O., Valverde-Cerdas, J., & Guerrero-Barrantes, M. (2016). Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal. *Revista Tecnología En Marcha*, 29(8), 33. <https://doi.org/10.18845/tm.v29i8.2983>

David R, A., Dennis J, S., & Thomas A, W. (2008). Estadística Para Administración Y Economía. Pearson Education. 7ma Edición. Obtenido de: <https://profefily.com/wp-content/uploads/2017/12/Estadística-para-administración-y-economía-Richard-I.-Levin.pdf>

De Lucas Coloma, L. (2018). Análisis estadístico de las emisiones de co2 por el consumo de combustible gaseoso en ecuador mediante la aplicación de la regresión lineal simple. *Revista Caribeña de Ciencias Sociales*. www.eumed.net/rev/caribe/2018/07/combustible-gaseoso-ecuador.html

Espinosa Martínez, V. (2025). Regresión lineal simple. *Vida Científica Boletín Científico de la Escuela Preparatoria No. 4. 13 (26)36-37.* <https://repository.uah.edu.mx/revistas/index.php/prepa4/article/view/14751/12306>

Fernández Lizana, M. (2020). Ventajas de R como herramienta para el Análisis y Visualización de datos en Ciencias Sociales. *Revista Científica de la UCSA*. 7(2). <https://revista.ucsa.ct.edu.py/ojs/index.php/ucsa/article/view/30/30>

Gonçalves Fernandes, A. C., Moreira Cavalcanti, L. F., Siqueira Campos Barros, M. L., & Da Costa Oliveira, F. M. (2017). Análises Descritivas e Microbiológicas das Águas Minerais Envasadas e Comercializadas na Região Metropolitana e Recife-PE. *Ciência E Natura*, 39(2), 272-284. <https://doi.org/10.5902/2179460X23622>

Hechavarría Pérez, J., Coello Velázquez, A., Robles Proenza, D., & Menéndez Aguado, J. (2015). Eficiencia energética en la trituración por impactos en la planta de zeolitas de San Andrés (Holguín, Cuba). *Dyna*, 82(193), 93-97. <https://doi.org/10.15446/dyna.v82n193.46085>

Moral Peláez, I. (2006). Modelos de regresión: lineal simple y regresión logística.

Poveda, M. (2007). Eficiencia energética: recurso no aprovechado. Obtenido de: www.olade.org

Rivera Cruz, M. A. (2022). RStudio y su importancia en la educación. *Revista CENCI*, 1(3). <https://doi.org/10.59142/cenci.v1i3.13>

Rodríguez Mañay, L., Saltos Chacán, M., & Muñoz Moreta, E. (2015). Cálculo de la tasa predeterminada de los costos indirectos de fabricación aplicados utilizando un modelo de regresión lineal. *Revista Publicando*, 2(5). <https://revistapublicando.org/revista/index.php/crv/article/view/87>

Contribución de los autores (CRediT)

Jácome-Vélez, W.: Conceptualización, Curación de contenidos y datos, Análisis formal de datos, Adquisición de fondos, Investigación, Metodología, Recursos materiales, Software, Redacción – borrador original, Redacción – revisión y edición. **Vilela-Sabando, J.:** Conceptualización, Curación de contenidos y datos, Análisis formal de datos, Adquisición de fondos, Investigación, Metodología, Recursos materiales, Software, Redacción – borrador original, Redacción – revisión y edición. **Cedeño-Hormaza, M.:** Conceptualización, Metodología, Administración de proyecto, Supervisión, Validación, Visualización, Redacción – revisión y edición. **Jácome-Duarte, C.:** Administración de proyecto, Supervisión, Validación, Visualización, Redacción – revisión y edición.

Todos los autores han leído y aceptado la versión publicada del manuscrito.

Disponibilidad de datos

Los datos que respaldan los hallazgos de este estudio están disponibles a solicitud razonable al autor de correspondencia.

Conflictos de intereses

Los autores han declarado que no existe conflicto de intereses en esta obra.

Declaración sobre el uso de IA generativa y tecnologías asistidas por IA en el proceso de redacción

El manuscrito no contiene una declaración específica sobre el uso de herramientas de inteligencia artificial durante su elaboración. La autoría y responsabilidad del contenido corresponden íntegramente a los autores.

Nota del Editor

Descargo de responsabilidad: Los datos, declaraciones, opiniones contenidas en el documento son responsabilidad únicamente de los autores y no de la *Revista Científica FINIBUS – Ingeniería, Industria y Arquitectura*. La Revista



y sus editores renuncian a toda responsabilidad por daño a persona o propiedad resultante de los métodos, instrucciones, producto o idea mencionado en el contenido.



Derechos de autor 2026.

Esta obra está bajo una licencia:
Internacional Creative Commons
Atribución-NoComercial-CompartirIgual
.4.0



Revista Científica *FINIBUS* - ISSN: 2737-6451.