**UNIVERSIDAD TECNOLÓGICA DE PANAMÁ**

**FACULTAD DE INGENIERÍA DE SISTEMAS COMPUTACIONALES**

**DEPARTAMENTO DE SISTEMAS DE INFORMACIÓN, CONTROL Y EVALUACIÓN**

**CARRERA LICENCIATURA EN INGENIERÍA DE SOFTWARE**

**MÉTODOS NUMÉRICOS PARA INGENIEROS**

**Laboratorio #3**

**Pertenece a:**

**González, Emanuel** **8-1022-2376**

**Grupo:**

**1SF134**

**Profesor:**

**Ing. Johel Batista**

**I SEMESTRE, 2025**

# **Resumen ejecutivo**

Este informe presenta el análisis de tres modelos de regresión lineal simple aplicados a bases de datos educativas internacionales, con el fin de examinar relaciones entre variables clave para el aprendizaje escolar. Se calcularon parámetros básicos como pendiente, intercepto y coeficiente de determinación (R²).

Los resultados evidencian que el índice socioeconómico (ESCS) tiene un impacto significativo en el rendimiento en lectura, con un R² aproximado de 0.71, reflejando una fuerte desigualdad educativa. Por otro lado, la relación entre Matemática y Ciencias mostró una pendiente moderada y un R² bajo (R≈ 0.17), señalando que otros factores influyen en el aprendizaje científico más allá de la matemática.

En estudiantes de tercer grado (ERCE 2023), se encontró una alta correlación entre Lectura y Matemática (R² ≈ 0.75), aunque con una pendiente inferior a uno, indicando que la mejora en lectura impulsa el aprendizaje matemático, pero no explica por completo su desarrollo.

Estos hallazgos permiten distinguir relaciones causales directas y vínculos curriculares complejos, así como las limitaciones del modelo estadístico empleado. Basados en esta evidencia, se sugieren intervenciones pedagógicas, ajustes curriculares y estrategias focalizadas para fortalecer los procesos de enseñanza y aprendizaje.

Este análisis demuestra que la regresión lineal es una herramienta valiosa para comprender dinámicas educativas y orientar políticas basadas en datos, integrando aspectos tanto cognitivos como socioemocionales y contextuales.

**Índice**

[**Resumen ejecutivo** 2](#_Toc204186776)

[**Introducción** 4](#_Toc204186777)

[**Fundamento teórico** 5](#_Toc204186778)

[**Datos y Metodología** 8](#_Toc204186779)

[**Presentación de los Datos** 8](#_Toc204186780)

[**Fases Metodológicas** 8](#_Toc204186781)

[**Resultados** 10](#_Toc204186782)

[**Problema 1 — PISA 2022: ESCS → Lectura (85 escuelas)** 10](#_Toc204186783)

[**Tabla de datos de PISA 2022: ESCS → Lectura (85 escuelas)** 10](#_Toc204186784)

[**Cálculos relacionados** 12](#_Toc204186785)

[**Medidas estadísticas** 14](#_Toc204186786)

[**Visualización** 14](#_Toc204186787)

[**Problema 2 — PISA 2022: Matemática → Ciencias (80 estudiantes)** 15](#_Toc204186788)

[**Tabla de datos de PISA 2022: Matemática → Ciencias (80 estudiantes)** 15](#_Toc204186789)

[**Cálculos relacionados** 17](#_Toc204186790)

[**Medidas estadísticas** 18](#_Toc204186791)

[**Visualización** 18](#_Toc204186792)

[**Problema 3 — ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)** 20](#_Toc204186793)

[**Tabla de datos de ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)** 20](#_Toc204186794)

[**Cálculos relacionados** 21](#_Toc204186795)

[**Medidas estadísticas** 23](#_Toc204186796)

[**Visualización** 23](#_Toc204186797)

[**Discusión** 24](#_Toc204186798)

[**Problema 1 — PISA 2022: ESCS → Lectura (85 escuelas)** 24](#_Toc204186799)

[**Problema 2 — PISA 2022: Matemática → Ciencias (80 estudiantes)** 29](#_Toc204186800)

[**Problema 3 — ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)** 34](#_Toc204186801)

[**Contraste entre los problemas** 39](#_Toc204186802)

[**Conclusión** 41](#_Toc204186803)

[**Referencias Bibliográficas** 42](#_Toc204186804)

[**Apéndice** 44](#_Toc204186805)

# **Introducción**

El presente informe tiene como objetivo analizar la relación entre distintas variables educativas a través de la aplicación del modelo de regresión lineal simple, una técnica estadística que permite estimar cómo el cambio en una variable predictora afecta al comportamiento de una variable de respuesta. En el ámbito educativo, esta herramienta resulta especialmente útil para identificar patrones de rendimiento, evaluar posibles desigualdades y orientar estrategias de mejora en función de evidencia cuantitativa.

A lo largo del informe, se abordan tres estudios de caso construidos a partir de bases de datos reales provenientes de evaluaciones internacionales ampliamente reconocidas:

* PISA 2022: Se analiza el impacto del índice socioeconómico (ESCS) en los puntajes obtenidos en lectura, con el propósito de comprender hasta qué punto las condiciones sociales y económicas influyen en el desarrollo de competencias lectoras.
* PISA 2022: Se explora la relación entre el rendimiento en Matemática y en Ciencias, observando si las habilidades cuantitativas se trasladan de forma directa al razonamiento científico.
* ERCE 2023: Se estudia la asociación entre los resultados en Lectura y en Matemática en estudiantes de tercer grado de primaria, permitiendo observar cómo interactúan estas competencias en las primeras etapas de la escolaridad, especialmente en un contexto post-pandemia.

En cada problema se calculan e interpretan los elementos clave del modelo: la pendiente (β1​), el intercepto (β0​), el coeficiente de determinación (R2), la Suma de Cuadrados del Error (SSE) y la Suma Total de Cuadrados (SST). Sin embargo, el propósito del análisis no se limita a la obtención de estos valores numéricos; también se busca interpretar sus implicaciones pedagógicas, identificar patrones que reflejen la desigualdad educativa, y proponer acciones o intervenciones que contribuyan a un aprendizaje más equitativo y eficaz.

Asimismo, este trabajo permite reflexionar sobre cómo las condiciones escolares, sociales y emocionales pueden influir en el aprendizaje, revelando que la estadística no solo mide, sino que también orienta. Al contrastar los tres problemas, se pone en evidencia que las relaciones entre variables educativas no siempre son lineales, proporcionales ni homogéneas, por lo que el análisis requiere una lectura crítica y contextualizada.

En definitiva, este informe busca no solo validar el uso de herramientas estadísticas en la educación, sino también demostrar su potencial para generar conocimiento aplicable, guiar decisiones docentes y apoyar la construcción de políticas públicas más justas y basadas en evidencia.

# **Fundamento teórico**

**Regresión Lineal Simple**

La regresión lineal simple es un método estadístico fundamental que sirve para entender y cuantificar la relación entre dos variables numéricas. Su propósito principal es establecer una ecuación que permita predecir el valor de una variable (llamada variable dependiente) a partir del valor de otra variable (conocida como variable independiente).

**¿En qué consiste la regresión lineal simple?**

Esta técnica asume que existe una correlación de tipo lineal entre dos variables. El resultado se refleja en una línea recta que mejor se ajusta a los datos observados, haciendo posible anticipar el comportamiento de una de las variables según los cambios de la otra. La ecuación que la representa tiene la forma:

En esta ecuación, cada componente tiene un significado clave para el análisis:

**Pendiente (β₁)**

La pendiente representa el cambio promedio en la variable dependiente por cada unidad de aumento en la variable independiente. Su fórmula es:

En términos prácticos, si se analiza la relación entre el índice socioeconómico y el puntaje en lectura, una pendiente positiva indicaría que, a mayor índice socioeconómico, mayor rendimiento lector, en promedio.

**Intercepto (β₀)**

El intercepto es el valor estimado de la variable dependiente cuando la variable independiente es igual a cero. Se calcula como:

Aunque en muchos contextos educativos un valor de x=0x = 0x=0 no tenga interpretación directa, el intercepto permite ubicar correctamente la recta de regresión en el plano cartesiano.

**Estimación (​)**

La ecuación estimada de la regresión es:

Esta fórmula se usa para predecir valores de y (por ejemplo, puntaje en lectura) a partir de un valor dado de x (por ejemplo, índice ESCS).

**SSE (Suma de Cuadrados del Error)**

El SSE mide la diferencia total entre los valores reales observados y los valores estimados por el modelo:

Un valor bajo de SSE indica que el modelo ajusta bien a los datos, mientras que un valor alto sugiere que hay mucha variabilidad no explicada.

**SST (Suma Total de Cuadrados)**

La SST representa la variabilidad total de los valores reales de y con respecto a su media:

Se usa como referencia para comparar cuánto del total de la variabilidad ha sido explicada por el modelo.

**Coeficiente de determinación (R²)**

El coeficiente de determinación expresa la proporción de la varianza total de la variable dependiente que es explicada por el modelo:

El valor de R2 oscila entre 0 y 1. Un valor cercano a 1 indica que el modelo explica gran parte de la variabilidad de los datos, mientras que un valor cercano a 0 indica que el modelo tiene bajo poder explicativo.

En resumen, el proceso inicia con la selecciónde las variables a analizar. Se define una variable independiente x aquella que se supone explica o influye y una variable dependiente y aquella que se quiere predecir o entender. Por ejemplo, se puede estudiar cómo el índice socioeconómico (ESCS) incide en los puntajes de lectura o matemática.

Luego, se organizan los datos recolectados en una tabla que incluya las columnas x, y, x2 y xy. A partir de esta tabla, se calculan las sumatorias necesarias: ∑x, ∑y, ∑x2, ∑xy, así como la cantidad de observaciones n.

Con estas sumatorias, se aplican las fórmulas de mínimos cuadrados para determinar los coeficientes del modelo. La pendiente β1 ​ se calcula como:

y el intercepto β0​ se obtiene mediante:

Con ambos coeficientes, se construye la ecuación estimada de regresión:

Posteriormente, para cada valor de x se puede calcular el valor estimado ​, el cual representa la predicción del modelo. A partir de estas estimaciones, se procede a evaluar la calidad del ajuste, calculando primero la Suma Total de Cuadrados (SST), que representa la variación total de y respecto a su media, y luego la Suma de los Cuadrados del Error (SSE), que mide cuánto se desvían los valores estimados de los valores reales.

Finalmente, se calcula el coeficiente de determinación R2, que expresa la proporción de la variabilidad en y que es explicada por el modelo:

Un valor de R2 cercano a 1 indica que el modelo es altamente explicativo, mientras que un valor bajo sugiere una relación débil entre las variables.

# **Datos y Metodología**

## **Presentación de los Datos**

El presente estudio se fundamenta en el análisis de tres conjuntos de datos, cada uno representando distintas dimensiones del rendimiento académico de estudiantes panameños, obtenidos a partir de dos relevantes evaluaciones internacionales: PISA 2022 y ERCE 2023. La selección de estos conjuntos responde al interés de explorar relaciones clave en el ámbito educativo panameño, con miras a contribuir al diagnóstico y la toma de decisiones sustentadas en evidencia empírica.

* **Primer conjunto de datos:** Incluye 85 observaciones correspondientes a distintas escuelas panameñas, donde se busca analizar la relación entre el índice socioeconómico (ESCS) y los puntajes obtenidos en lectura por parte de los estudiantes.
* **Segundo conjunto de datos:** Compuesto por 80 estudiantes seleccionados, examina el vínculo entre los resultados en matemática y ciencia, permitiendo identificar posibles patrones de desempeño transversal en estas áreas.
* **Tercer conjunto de datos:** Integra a 70 estudiantes de tercer grado evaluados en el marco de ERCE 2023, centrando el análisis en la relación existente entre habilidades lectoras y competencias matemáticas en los primeros años de formación escolar.

Cada base de datos será organizada en hojas de cálculo independientes, lo que facilitará su gestión, procesamiento y el control riguroso de la integridad de la información.

## **Fases Metodológicas**

La metodología propuesta se estructura en una serie de etapas sucesivas diseñadas para garantizar un análisis robusto y replicable:

1. **Organización y exploración inicial de los datos:**

Se procederá a una inspección detallada de cada conjunto, verificando estrictamente la cantidad de observaciones y la naturaleza de las variables involucradas. Este paso es crucial para asegurar la confiabilidad de los resultados y la pertinencia del análisis.

1. **Construcción de la tabla de análisis:**

Para cada conjunto, se confeccionará una tabla exhaustiva que incluya, al menos, las siguientes columnas:

* + Valores observados de X y Y (variables independientes y dependientes).
  + Cuadrados de X (x2).
  + Producto de X y Y (*xy*).
  + Esta estructura permitirá realizar los cálculos subsiguientes de manera ordenada y transparente.

1. **Cálculo manual de los coeficientes del modelo:**

Se aplicará el método de los mínimos cuadrados para obtener los parámetros fundamentales de la regresión lineal simple:

* + **Pendiente (*β*1):** Representa la variación promedio de la variable dependiente ante un cambio unitario en la variable independiente.
  + **Intercepto (*β*0):** Indica el valor estimado de Y cuando X es igual a cero.
  + **Coeficiente de determinación (R2):** Expresa el grado de ajuste del modelo a los datos observados.

Los resultados se reportarán redondeados a cinco cifras decimales, asegurando precisión y rigurosidad.

1. **Representación gráfica de los datos:**

Cada caso será ilustrado mediante un diagrama de dispersión en el que se sobrepondrá la recta de regresión ajustada. Las gráficas contendrán títulos descriptivos, ejes rotulados apropiadamente y se anotará el valor correspondiente de R2, facilitando la interpretación visual de la relación entre las variables.

1. **Diagnóstico del modelo ajustado:**

Se llevará a cabo una evaluación visual crítica para corroborar los principales supuestos del modelo de regresión lineal simple:

* + Linealidad de la relación.
  + Homocedasticidad (constancia de la varianza de los residuos).
  + Ausencia de outliers significativos que puedan distorsionar los resultados.  
    Esta etapa contribuye a la validez y credibilidad de las conclusiones derivadas.

1. **Interpretación cuantitativa y pedagógica de los resultados:**

Se interpretarán los coeficientes obtenidos, subrayando el significado estadístico y pedagógico de la pendiente, el grado de ajuste y el alcance de los hallazgos en el contexto educativo panameño. Esta reflexión permitirá traducir los resultados en insumos pertinentes para la mejora educativa.

1. **Comparación entre los tres casos analizados:**

Se contrastarán los resultados, considerando las diferencias en contexto, población y variables seleccionadas. El objetivo es extraer conclusiones más amplias sobre el sistema educativo nacional y generar propuestas de intervención o mejora fundamentadas en la evidencia recopilada.

# **Resultados**

## **Problema 1 — PISA 2022: ESCS → Lectura (85 escuelas)**

### **Tabla de datos de PISA 2022: ESCS → Lectura (85 escuelas)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **x** | **y** | **x2** | **y2** | **xy** |
| -1.19 | 355.2 | 1.41610 | 126167.04 | -422.68800 |
| -0.631 | 356.7 | 0.39816 | 127234.89 | -225.07770 |
| 0.859 | 392.2 | 0.73788 | 153820.84 | 336.89980 |
| -0.125 | 382.8 | 0.01562 | 146535.84 | -47.85000 |
| 0.528 | 388.5 | 0.27878 | 150932.25 | 205.12800 |
| 0.219 | 398.8 | 0.04796 | 159041.44 | 87.33720 |
| -0.873 | 366.1 | 0.76213 | 134029.21 | -319.60530 |
| 0.188 | 389.2 | 0.03534 | 151476.64 | 73.16960 |
| -0.585 | 367.7 | 0.34222 | 135203.29 | -215.10450 |
| -0.59 | 385.2 | 0.34810 | 148379.04 | -227.26800 |
| -0.213 | 361.5 | 0.04537 | 130682.25 | -76.99950 |
| -1.673 | 346.2 | 2.79893 | 119854.44 | -579.19260 |
| 0.773 | 379.8 | 0.59753 | 144248.04 | 293.58540 |
| -0.904 | 373.6 | 0.81722 | 139576.96 | -337.73440 |
| 0.6 | 401.0 | 0.36000 | 160801 | 240.60000 |
| -0.177 | 370.6 | 0.03133 | 137344.36 | -65.59620 |
| 1.079 | 419.3 | 1.16424 | 175812.49 | 452.42470 |
| -0.894 | 360.3 | 0.79924 | 129816.09 | -322.10820 |
| -0.581 | 352.9 | 0.33756 | 124538.41 | -205.03490 |
| 0.004 | 376.7 | 0.00002 | 141902.89 | 1.50680 |
| -2.287 | 343.2 | 5.23037 | 117786.24 | -784.89840 |
| 0.445 | 385.2 | 0.19802 | 148379.04 | 171.41400 |
| 1.087 | 426.8 | 1.18157 | 182158.24 | 463.93160 |
| 0.714 | 403.1 | 0.50980 | 162489.61 | 287.81340 |
| 0.379 | 407.6 | 0.14364 | 166137.76 | 154.48040 |
| -0.431 | 388.9 | 0.18576 | 151243.21 | -167.61590 |
| 0.854 | 375.1 | 0.72932 | 140700.01 | 320.33540 |
| 0.667 | 387.5 | 0.44489 | 150156.25 | 258.46250 |
| 0.053 | 348.2 | 0.00281 | 121243.24 | 18.45460 |
| -0.295 | 366.9 | 0.08702 | 134615.61 | -108.23550 |
| -0.626 | 382.1 | 0.39188 | 146000.41 | -239.19460 |
| -1.407 | 344.6 | 1.97965 | 118749.16 | -484.85220 |
| 0.804 | 386.8 | 0.64642 | 149614.24 | 310.98720 |
| -2.255 | 343.3 | 5.08502 | 117854.89 | -774.14150 |
| -0.633 | 378.1 | 0.40069 | 142959.61 | -239.33730 |
| -0.152 | 373.0 | 0.02310 | 139129 | -56.69600 |
| 0.474 | 385.5 | 0.22468 | 148610.25 | 182.72700 |
| 1.285 | 412.8 | 1.65122 | 170403.84 | 530.44800 |
| 0.594 | 393.7 | 0.35284 | 154999.69 | 233.85780 |
| -0.562 | 366.5 | 0.31584 | 134322.25 | -205.97300 |
| 0.355 | 390.2 | 0.12602 | 152256.04 | 138.52100 |
| -1.435 | 355.1 | 2.05922 | 126096.01 | -509.56850 |
| 0.987 | 400.8 | 0.97417 | 160640.64 | 395.58960 |
| -0.441 | 396.1 | 0.19448 | 156895.21 | -174.68010 |
| -0.906 | 360.0 | 0.82084 | 129600 | -326.16000 |
| -0.875 | 357.8 | 0.76562 | 128020.84 | -313.07500 |
| -0.355 | 390.6 | 0.12602 | 152568.36 | -138.66300 |
| -0.654 | 359.8 | 0.42772 | 129456.04 | -235.30920 |
| 1.761 | 432.3 | 3.10112 | 186883.29 | 761.28030 |
| -0.946 | 360.6 | 0.89492 | 130032.36 | -341.12760 |
| -0.271 | 376.0 | 0.07344 | 141376 | -101.89600 |
| -0.275 | 383.9 | 0.07563 | 147379.21 | -105.57250 |
| -0.275 | 389.6 | 0.07563 | 151788.16 | -107.14000 |
| -0.25 | 387.5 | 0.06250 | 150156.25 | -96.87500 |
| -0.733 | 364.4 | 0.53729 | 132787.36 | -267.10520 |
| -0.825 | 360.8 | 0.68062 | 130176.64 | -297.66000 |
| -1.076 | 364.8 | 1.15778 | 133079.04 | -392.52480 |
| -1.639 | 352.1 | 2.68632 | 123974.41 | -577.09190 |
| -0.105 | 370.4 | 0.01102 | 137196.16 | -38.89200 |
| 0.586 | 395.5 | 0.34340 | 156420.25 | 231.76300 |
| -0.789 | 377.7 | 0.62252 | 142657.29 | -298.00530 |
| -0.803 | 367.5 | 0.64481 | 135056.25 | -295.10250 |
| -0.585 | 388.2 | 0.34222 | 150699.24 | -227.09700 |
| -0.715 | 372.1 | 0.51122 | 138458.41 | -266.05150 |
| -1.593 | 349.5 | 2.53765 | 122150.25 | -556.75350 |
| 0.929 | 397.9 | 0.86304 | 158324.41 | 369.64910 |
| 0.095 | 375.0 | 0.00902 | 140625 | 35.62500 |
| -0.941 | 350.7 | 0.88548 | 122990.49 | -330.00870 |
| -0.033 | 385.1 | 0.00109 | 148302.01 | -12.70830 |
| -0.695 | 357.3 | 0.48302 | 127663.29 | -248.32350 |
| -0.49 | 366.9 | 0.24010 | 134615.61 | -179.78100 |
| 0.028 | 379.4 | 0.00078 | 143944.36 | 10.62320 |
| 0.558 | 394.7 | 0.31136 | 155788.09 | 220.24260 |
| 1.068 | 404.8 | 1.14062 | 163863.04 | 432.32640 |
| 1.234 | 402.9 | 1.52276 | 162328.41 | 497.17860 |
| -0.524 | 347.4 | 0.27458 | 120686.76 | -182.03760 |
| -0.75 | 364.0 | 0.56250 | 132496 | -273.00000 |
| -0.21 | 366.9 | 0.04410 | 134615.61 | -77.04900 |
| -0.184 | 375.3 | 0.03386 | 140850.09 | -69.05520 |
| -0.961 | 374.6 | 0.92352 | 140325.16 | -359.99060 |
| -0.858 | 374.2 | 0.73616 | 140025.64 | -321.06360 |
| 0.432 | 390.4 | 0.18662 | 152412.16 | 168.65280 |
| 1.178 | 407.3 | 1.38768 | 165893.29 | 479.79940 |
| -0.504 | 353.1 | 0.25402 | 124679.61 | -177.96240 |
| -0.883 | 366.4 | 0.77969 | 134248.96 | -323.53120 |

### **Cálculos relacionados**

Se trabajó un conjunto de 85 observaciones, donde se estudió la relación entre el índice socioeconómico (ESCS) y los puntajes obtenidos en lectura (Reading\_score). A continuación, se presentan los valores calculados:

* Sumatorias:
* Cálculo de la pendiente (β₁):
* Cálculo del intercepto (β0):
* Cálculo de la estimación (​):
* Cálculo de SSE (Suma de Cuadrados del Error):
* Cálculo de SST (Suma Total de Cuadrados):
* Cálculo de Coeficiente de determinación (R2):

### **Medidas estadísticas**

Antes de aplicar el modelo de regresión, se calcularon medidas estadísticas básicas para las variables ESCS y Reading\_score. Estos valores permiten entender la dispersión y tendencia central de los datos analizados:

|  |  |  |
| --- | --- | --- |
| **Medida estadística** | **ESCS (x)** | **Reading score (y)** |
| Media | -0.22172 | 377.18588 |
| Mediana | -0.27500 | 375.30000 |
| Moda | -0.58500 | 366.90000 |
| Varianza | 0.68371 | 387.50623 |

### **Visualización**

A continuación, se presenta la gráfica de dispersión entre el índice socioeconómico (ESCS) y el puntaje en lectura, junto con la recta ajustada por mínimos cuadrados y el valor de R2 calculado mediante regresión lineal:

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

La gráfica muestra una tendencia general ascendente entre el índice socioeconómico (ESCS) y el puntaje en lectura, lo que concuerda con el valor de β1=19.97, indicando que por cada punto adicional en ESCS, el rendimiento en lectura aumenta aproximadamente 20 puntos, en promedio.

El valor de R2=0.71 indica que el modelo explica un 71% de la variabilidad observada, lo cual sugiere un ajuste bastante bueno. La mayoría de los puntos se concentran alrededor de la recta, con algunas excepciones que podrían considerarse leves outliers.

Visualmente, no se observa una dispersión creciente o decreciente, lo que indica que el supuesto de homocedasticidad parece cumplirse. Este comportamiento sugiere que el ESCS puede ser un predictor razonablemente confiable del rendimiento lector en la muestra evaluada.

## **Problema 2 — PISA 2022: Matemática → Ciencias (80 estudiantes)**

### **Tabla de datos de PISA 2022: Matemática → Ciencias (80 estudiantes)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **x** | **y** | **x2** | **y2** | **xy** |
| 365.0 | 393.0 | 133225.0 | 154449 | 143445.0 |
| 354.0 | 355.0 | 125316.0 | 126025 | 125670.0 |
| 363.0 | 413.0 | 131769.0 | 170569 | 149919.0 |
| 357.0 | 406.0 | 127449.0 | 164836 | 144942.0 |
| 368.0 | 407.0 | 135424.0 | 165649 | 149776.0 |
| 369.0 | 397.0 | 136161.0 | 157609 | 146493.0 |
| 356.0 | 417.0 | 126736.0 | 173889 | 148452.0 |
| 372.0 | 382.0 | 138384.0 | 145924 | 142104.0 |
| 358.0 | 386.0 | 128164.0 | 148996 | 138188.0 |
| 376.0 | 391.0 | 141376.0 | 152881 | 147016.0 |
| 356.0 | 400.0 | 126736.0 | 160000 | 142400.0 |
| 361.0 | 392.0 | 130321.0 | 153664 | 141512.0 |
| 362.0 | 429.0 | 131044.0 | 184041 | 155298.0 |
| 354.0 | 393.0 | 125316.0 | 154449 | 139122.0 |
| 346.0 | 384.0 | 119716.0 | 147456 | 132864.0 |
| 354.0 | 401.0 | 125316.0 | 160801 | 141954.0 |
| 369.0 | 395.0 | 136161.0 | 156025 | 145755.0 |
| 357.0 | 381.0 | 127449.0 | 145161 | 136017.0 |
| 357.0 | 403.0 | 127449.0 | 162409 | 143871.0 |
| 372.0 | 401.0 | 138384.0 | 160801 | 149172.0 |
| 364.0 | 415.0 | 132496.0 | 172225 | 151060.0 |
| 366.0 | 407.0 | 133956.0 | 165649 | 148962.0 |
| 367.0 | 405.0 | 134689.0 | 164025 | 148635.0 |
| 353.0 | 393.0 | 124609.0 | 154449 | 138729.0 |
| 361.0 | 398.0 | 130321.0 | 158404 | 143678.0 |
| 361.0 | 405.0 | 130321.0 | 164025 | 146205.0 |
| 359.0 | 391.0 | 128881.0 | 152881 | 140369.0 |
| 351.0 | 364.0 | 123201.0 | 132496 | 127764.0 |
| 364.0 | 402.0 | 132496.0 | 161604 | 146328.0 |
| 347.0 | 381.0 | 120409.0 | 145161 | 132207.0 |
| 353.0 | 373.0 | 124609.0 | 139129 | 131669.0 |
| 354.0 | 404.0 | 125316.0 | 163216 | 143016.0 |
| 357.0 | 390.0 | 127449.0 | 152100 | 139230.0 |
| 367.0 | 379.0 | 134689.0 | 143641 | 139093.0 |
| 373.0 | 408.0 | 139129.0 | 166464 | 152184.0 |
| 374.0 | 400.0 | 139876.0 | 160000 | 149600.0 |
| 366.0 | 379.0 | 133956.0 | 143641 | 138714.0 |
| 364.0 | 405.0 | 132496.0 | 164025 | 147420.0 |
| 341.0 | 368.0 | 116281.0 | 135424 | 125488.0 |
| 369.0 | 415.0 | 136161.0 | 172225 | 153135.0 |
| 381.0 | 402.0 | 145161.0 | 161604 | 153162.0 |
| 364.0 | 415.0 | 132496.0 | 172225 | 151060.0 |
| 362.0 | 405.0 | 131044.0 | 164025 | 146610.0 |
| 366.0 | 404.0 | 133956.0 | 163216 | 147864.0 |
| 366.0 | 392.0 | 133956.0 | 153664 | 143472.0 |
| 361.0 | 412.0 | 130321.0 | 169744 | 148732.0 |
| 357.0 | 403.0 | 127449.0 | 162409 | 143871.0 |
| 363.0 | 395.0 | 131769.0 | 156025 | 143385.0 |
| 349.0 | 397.0 | 121801.0 | 157609 | 138553.0 |
| 345.0 | 372.0 | 119025.0 | 138384 | 128340.0 |
| 371.0 | 389.0 | 137641.0 | 151321 | 144319.0 |
| 356.0 | 394.0 | 126736.0 | 155236 | 140264.0 |
| 366.0 | 404.0 | 133956.0 | 163216 | 147864.0 |
| 372.0 | 414.0 | 138384.0 | 171396 | 154008.0 |
| 370.0 | 407.0 | 136900.0 | 165649 | 150590.0 |
| 356.0 | 412.0 | 126736.0 | 169744 | 146672.0 |
| 363.0 | 382.0 | 131769.0 | 145924 | 138666.0 |
| 351.0 | 398.0 | 123201.0 | 158404 | 139698.0 |
| 359.0 | 400.0 | 128881.0 | 160000 | 143600.0 |
| 368.0 | 424.0 | 135424.0 | 179776 | 156032.0 |
| 360.0 | 400.0 | 129600.0 | 160000 | 144000.0 |
| 337.0 | 384.0 | 113569.0 | 147456 | 129408.0 |
| 342.0 | 408.0 | 116964.0 | 166464 | 139536.0 |
| 354.0 | 404.0 | 125316.0 | 163216 | 143016.0 |
| 371.0 | 415.0 | 137641.0 | 172225 | 153965.0 |
| 362.0 | 394.0 | 131044.0 | 155236 | 142628.0 |
| 361.0 | 401.0 | 130321.0 | 160801 | 144761.0 |
| 349.0 | 366.0 | 121801.0 | 133956 | 127734.0 |
| 361.0 | 387.0 | 130321.0 | 149769 | 139707.0 |
| 342.0 | 392.0 | 116964.0 | 153664 | 134064.0 |
| 363.0 | 375.0 | 131769.0 | 140625 | 136125.0 |
| 369.0 | 407.0 | 136161.0 | 165649 | 150183.0 |
| 353.0 | 393.0 | 124609.0 | 154449 | 138729.0 |
| 369.0 | 410.0 | 136161.0 | 168100 | 151290.0 |
| 354.0 | 370.0 | 125316.0 | 136900 | 130980.0 |
| 361.0 | 371.0 | 130321.0 | 137641 | 133931.0 |
| 359.0 | 413.0 | 128881.0 | 170569 | 148267.0 |
| 361.0 | 400.0 | 130321.0 | 160000 | 144400.0 |
| 372.0 | 414.0 | 138384.0 | 171396 | 154008.0 |
| 367.0 | 426.0 | 134689.0 | 181476 | 156342.0 |

### **Cálculos relacionados**

Se trabajó un conjunto de 80 observaciones, donde se estudió la relación entre los puntajes obtenidos en matemáticas (Math\_score) y los puntajes obtenidos en ciencias (Science\_Score). A continuación, se presentan los valores calculados:

* Sumatorias:
* Cálculo de la pendiente (β₁):
* Cálculo del intercepto (β0):
* Cálculo de la estimación (​):
* Cálculo de SSE (Suma de Cuadrados del Error):
* Cálculo de SST (Suma Total de Cuadrados):
* Cálculo de Coeficiente de determinación (R2):

### **Medidas estadísticas**

Antes de aplicar el modelo de regresión, se calcularon medidas estadísticas básicas para las variables Math\_score y Science\_score. Estos valores permiten entender la dispersión y tendencia central de los datos analizados:

|  |  |  |
| --- | --- | --- |
| **Medida estadística** | **Math\_score (x)** | **Science\_score (y)** |
| Media | 360.62500 | 396.98750 |
| Mediana | 361.00000 | 400.00000 |
| Moda | 361.00000 | 400.00000 |
| Varianza | 75.50316 | 223.48085 |

### **Visualización**

A continuación, se presenta la gráfica de dispersión entre el puntaje obtenido en matemáticas (Math\_score) y el puntaje obtenido en ciencias (Science\_score), junto con la recta ajustada por mínimos cuadrados y el valor de R2 calculado mediante regresión lineal:

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

La gráfica de dispersión generada muestra una ligera tendencia positiva entre los puntajes de matemática y los puntajes de ciencias. A pesar de que la pendiente obtenida en el modelo de regresión lineal es positiva (β1​=0.70458), lo cual indica que a mayor puntaje en matemática se espera un mayor puntaje en ciencias, la relación observada es débil.

Esto se confirma por el valor del coeficiente de determinación R2=0.16772, que indica que solo el 16.77% de la variabilidad en los puntajes de ciencias puede explicarse por los puntajes de matemática. El 83.23% restante de la variación podría deberse a otros factores no incluidos en este modelo.

Visualmente, los puntos aparecen bastante dispersos en torno a la recta de regresión. Muchos de ellos se alejan significativamente de la línea estimada, lo que sugiere que el modelo no ajusta bien a los datos. A pesar de la dirección positiva, el patrón de los datos presenta una gran variabilidad vertical en todos los rangos de x, lo que también puede indicar una violación del supuesto de homocedasticidad.

Además, se observa una acumulación de puntos alrededor de ciertos valores discretos de los puntajes (por ejemplo, 361 y 400), lo cual coincide con la moda y mediana reportadas, y podría estar relacionado con la naturaleza estandarizada o redondeada de los puntajes de la prueba.

## **Problema 3 — ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)**

### **Tabla de datos de ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **x** | **y** | **x2** | **y2** | **xy** |
| 697 | 720 | 485809 | 518400 | 501840 |
| 675 | 675 | 455625 | 455625 | 455625 |
| 692 | 689 | 478864 | 474721 | 476788 |
| 703 | 701 | 494209 | 491401 | 492803 |
| 691 | 695 | 477481 | 483025 | 480245 |
| 692 | 703 | 478864 | 494209 | 486476 |
| 689 | 718 | 474721 | 515524 | 494702 |
| 710 | 688 | 504100 | 473344 | 488480 |
| 707 | 726 | 499849 | 527076 | 513282 |
| 677 | 686 | 458329 | 470596 | 464422 |
| 714 | 663 | 509796 | 439569 | 473382 |
| 687 | 690 | 471969 | 476100 | 474030 |
| 715 | 710 | 511225 | 504100 | 507650 |
| 682 | 703 | 465124 | 494209 | 479446 |
| 706 | 720 | 498436 | 518400 | 508320 |
| 689 | 698 | 474721 | 487204 | 480922 |
| 693 | 686 | 480249 | 470596 | 475398 |
| 704 | 725 | 495616 | 525625 | 510400 |
| 688 | 685 | 473344 | 469225 | 471280 |
| 698 | 682 | 487204 | 465124 | 476036 |
| 698 | 692 | 487204 | 478864 | 483016 |
| 707 | 700 | 499849 | 490000 | 494900 |
| 709 | 709 | 502681 | 502681 | 502681 |
| 714 | 706 | 509796 | 498436 | 504084 |
| 698 | 697 | 487204 | 485809 | 486506 |
| 699 | 718 | 488601 | 515524 | 501882 |
| 689 | 712 | 474721 | 506944 | 490568 |
| 687 | 706 | 471969 | 498436 | 485022 |
| 682 | 688 | 465124 | 473344 | 469216 |
| 693 | 688 | 480249 | 473344 | 476784 |
| 697 | 699 | 485809 | 488601 | 487203 |
| 695 | 699 | 483025 | 488601 | 485805 |
| 701 | 721 | 491401 | 519841 | 505421 |
| 681 | 688 | 463761 | 473344 | 468528 |
| 712 | 691 | 506944 | 477481 | 491992 |
| 698 | 689 | 487204 | 474721 | 480922 |
| 706 | 741 | 498436 | 549081 | 523146 |
| 698 | 687 | 487204 | 471969 | 479526 |
| 686 | 673 | 470596 | 452929 | 461678 |
| 692 | 694 | 478864 | 481636 | 480248 |
| 662 | 712 | 438244 | 506944 | 471344 |
| 697 | 705 | 485809 | 497025 | 491385 |
| 694 | 707 | 481636 | 499849 | 490658 |
| 707 | 681 | 499849 | 463761 | 481467 |
| 677 | 663 | 458329 | 439569 | 448851 |
| 694 | 701 | 481636 | 491401 | 486494 |
| 692 | 692 | 478864 | 478864 | 478864 |
| 682 | 701 | 465124 | 491401 | 478082 |
| 680 | 672 | 462400 | 451584 | 456960 |
| 683 | 713 | 466489 | 508369 | 486979 |
| 705 | 733 | 497025 | 537289 | 516765 |
| 705 | 680 | 497025 | 462400 | 479400 |
| 681 | 697 | 463761 | 485809 | 474657 |
| 693 | 696 | 480249 | 484416 | 482328 |
| 700 | 731 | 490000 | 534361 | 511700 |
| 698 | 708 | 487204 | 501264 | 494184 |
| 680 | 709 | 462400 | 502681 | 482120 |
| 689 | 696 | 474721 | 484416 | 479544 |
| 681 | 698 | 463761 | 487204 | 475338 |
| 705 | 747 | 497025 | 558009 | 526635 |
| 721 | 756 | 519841 | 571536 | 545076 |
| 695 | 671 | 483025 | 450241 | 466345 |
| 674 | 695 | 454276 | 483025 | 468430 |
| 703 | 711 | 494209 | 505521 | 499833 |
| 711 | 693 | 505521 | 480249 | 492723 |
| 710 | 698 | 504100 | 487204 | 495580 |
| 697 | 706 | 485809 | 498436 | 492082 |
| 700 | 722 | 490000 | 521284 | 505400 |
| 696 | 712 | 484416 | 506944 | 495552 |
| 702 | 699 | 492804 | 488601 | 490698 |

### **Cálculos relacionados**

Se trabajó un conjunto de 70 observaciones, donde se estudió la relación entre los puntajes obtenidos en lectura (Reading\_Score) y los puntajes obtenidos en matemáticas (Math\_score). A continuación, se presentan los valores calculados:

* Sumatorias:
* Cálculo de la pendiente (β₁):
* Cálculo del intercepto (β0):
* Cálculo de la estimación (​):
* Cálculo de SSE (Suma de Cuadrados del Error):
* Cálculo de SST (Suma Total de Cuadrados):
* Cálculo de Coeficiente de determinación (R2):

### **Medidas estadísticas**

Antes de aplicar el modelo de regresión, se calcularon medidas estadísticas básicas para las variables Reading\_score y Math\_score. Estos valores permiten entender la dispersión y tendencia central de los datos analizados:

|  |  |  |
| --- | --- | --- |
| **Medida estadística** | **Reading\_score (x)** | **Math\_score (y)** |
| Media | 695.21429 | 700.94286 |
| Mediana | 696.50000 | 699.00000 |
| Moda | 698.00000 | 688.00000 |
| Varianza | 132.25776 | 331.21408 |

### **Visualización**

A continuación, se presenta la gráfica de dispersión entre el puntaje obtenido en lectura (Reading\_score) y el puntaje obtenido en matemáticas (Math\_score), junto con la recta ajustada por mínimos cuadrados y el valor de R2 calculado mediante regresión lineal:

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

La gráfica de dispersión generada muestra una tendencia positiva débil entre los puntajes obtenidos en lectura (variable independiente) y los puntajes obtenidos en matemática (variable dependiente). Aunque la recta de regresión tiene una pendiente positiva β1​=0.51994, lo que indica que a medida que aumenta el rendimiento en lectura también tiende a aumentar el rendimiento en matemática, el ajuste general del modelo es bastante bajo.

El coeficiente de determinación obtenido es R2=0.10795, lo que significa que solo el 10.80% de la variabilidad en los puntajes de matemática puede explicarse por los puntajes de lectura. Esto sugiere que la relación lineal entre ambas variables, si bien existe, es muy débil dentro de este conjunto de datos.

Visualmente, los puntos se encuentran altamente dispersos en torno a la recta, con varias observaciones muy alejadas de la línea de regresión. Esta dispersión ocurre en todo el rango de puntajes, sin concentrarse en ningún extremo particular, lo que sugiere una pobre capacidad predictiva del modelo.

Además, al observar las medidas estadísticas, se nota que los puntajes de lectura están más concentrados (varianza ≈ 132.26) en comparación con los de matemática (varianza ≈ 331.21), lo que podría contribuir a una menor relación lineal y a un mayor grado de variabilidad no explicada.

# **Discusión**

## **Problema 1 — PISA 2022: ESCS → Lectura (85 escuelas)**

1. **Interpretación de β1​: ¿Cuánto sube el puntaje de lectura por cada 0?1 pts en ESCS?**

La pendiente del modelo de regresión lineal simple (β1=19.97484) constituye un parámetro fundamental para interpretar la relación entre el índice socioeconómico (ESCS) y el puntaje de lectura. Específicamente, este coeficiente indica que, en promedio, por cada incremento unitario en el índice ESCS de un estudiante, se prevé un aumento de aproximadamente 19.97 puntos en su puntaje de lectura.

Cabe destacar que, en situaciones reales, las variaciones en el índice ESCS suelen ser menores a una unidad. Por ejemplo, una mejora de 0.1 puntos en dicho índice se traduciría, según el modelo, en un incremento estimado de 1.997 puntos en el puntaje de lectura (aproximadamente 2 puntos), evidenciando la proporcionalidad de la relación estimada por la pendiente.

El puntaje estimado en lectura subiría 2 puntos aproximadamente.

Aunque el aumento numérico pueda parecer modesto cuando se considera a nivel individual, resulta relevante desde una perspectiva colectiva. Mejoras graduales en el índice ESCS, producto de políticas públicas orientadas al fortalecimiento del entorno socioeconómico (como inversiones en alimentación escolar, mejora de la infraestructura educativa o el fortalecimiento del apoyo familiar), tienen el potencial de traducirse en avances significativos en el desempeño académico de amplios grupos estudiantiles.

1. **Varianza no explicada: Tres factores escolares no capturados por ESCS que podrían influir en lectura**

Si bien el índice ESCS constituye una síntesis representativa de las condiciones socioeconómicas de los estudiantes incluyendo variables como el ingreso familiar, el nivel educativo de los padres y el acceso a recursos en el hogar, su capacidad explicativa del desempeño en lectura resulta limitada cuando se analizan los resultados a partir de un modelo de regresión lineal simple. En efecto, aproximadamente el 29% de la variabilidad observada en los puntajes de lectura queda sin ser explicada por este indicador, lo que evidencia la intervención de otros elementos contextuales y pedagógicos de igual o mayor relevancia.

A continuación, se presentan tres factores escolares fundamentales, no contemplados directamente por el índice ESCS, que pueden incidir significativamente en el desarrollo de las competencias lectoras:

* **Calidad Docente**

La calidad de la enseñanza constituye uno de los pilares fundamentales en la formación lectora de los estudiantes. El perfil académico del docente, su experiencia profesional, y sobre todo las estrategias pedagógicas que adopta, son determinantes para el progreso lector. Aquellos docentes que demuestran dominio en estrategias de lectura comprensiva, que personalizan sus prácticas según las necesidades y niveles de sus alumnos, y que promueven activamente el pensamiento crítico mediante los textos, contribuyen notablemente a la mejora del rendimiento lector, aun en contextos de desventaja socioeconómica. Es pertinente destacar que la calidad docente es reconocida internacionalmente como uno de los mayores predictores del aprovechamiento escolar.

* **Acceso y Uso de Bibliotecas Escolares**

La mera disponibilidad de una biblioteca escolar no garantiza, por sí sola, el mejoramiento de las habilidades lectoras. El impacto real proviene del uso efectivo de este recurso. Instituciones que fomentan la lectura diaria, facilitan préstamos de libros, organizan actividades de comprensión lectora y promueven clubes de lectura generan entornos estimulantes para el desarrollo del hábito lector. Estas prácticas, independientes en muchos casos del nivel socioeconómico de los estudiantes, pueden atenuar desventajas iniciales y despertar el interés genuino por la lectura.

* **Gestión Institucional y Liderazgo Escolar**

La organización y el liderazgo dentro de un centro educativo desempeñan un papel directo en el aprendizaje de los estudiantes. Escuelas que cuentan con una dirección clara, expectativas académicas definidas, metas compartidas y un clima institucional positivo suelen mostrar mejores resultados académicos. Por el contrario, la desorganización, la baja motivación del profesorado o la falta de visión pedagógica inciden negativamente en los logros estudiantiles, incluso en contextos con condiciones socioeconómicas favorables.

En síntesis, aunque el índice ESCS constituye una variable poderosa para describir parte del entorno de los estudiantes, resulta insuficiente para explicar a cabalidad la complejidad del proceso de aprendizaje lector. La consideración de factores escolares como la calidad docente, el uso efectivo de bibliotecas y la gestión institucional, entre otros posibles determinantes, se torna indispensable al interpretar los resultados académicos y al diseñar políticas y estrategias orientadas a la mejora educativa. El abordaje integral de estos elementos permitirá una comprensión más profunda y una intervención más efectiva sobre las trayectorias escolares y el desarrollo de las competencias lectoras.

1. **Intervención: Programa piloto para 10 escuelas de bajo ESCS**

**Objetivo general:**Implementar un programa integral de lectura en 10 escuelas con niveles bajos de ESCS, con el fin de mejorar al menos 15 puntos en el puntaje PISA de lectura en el transcurso de un año escolar.

* **Denominación y duración**

Nombre del programa: La Lectura Nos Une  
Duración: 12 meses (correspondientes a un ciclo escolar completo)

* **Fundamentación**

El programa “Leer para Crecer” está concebido para intervenir de manera directa sobre factores escolares trascendentales que, si bien no se encuentran reflejados en el índice ESCS, tienen un potencial comprobado para mitigar los efectos desfavorables de la vulnerabilidad socioeconómica en el aprendizaje. La iniciativa integra metodologías pedagógicas innovadoras, provisión de recursos bibliográficos y estrategias sistemáticas de apoyo docente, con el propósito de robustecer la formación lectora en entornos escolares diversos.

* **Componentes Principales**

1. Lectura Guiada Diaria

* Implementación de una sesión diaria de una hora enfocada en la lectura guiada, bajo la dirección del docente titular.
* Las actividades se centrarán en la comprensión lectora mediante técnicas como la lectura compartida, formulación de inferencias, realización de preguntas abiertas y establecimiento de conexiones entre el texto y la experiencia vital de los estudiantes.
* Se busca, a través de la práctica constante, propiciar el desarrollo de la fluidez lectora, ampliar el vocabulario y profundizar la comprensión textual.

1. Talleres Mensuales para el Cuerpo Docente

* Realización de encuentros mensuales de formación docente, enfocados en el fortalecimiento de métodos eficaces para la instrucción en lectura.
* Los contenidos abordarán la lectura en voz alta, la aplicación de estrategias metacognitivas, la diferenciación pedagógica según el nivel lector de los alumnos y el uso de la evaluación formativa.
* Este componente apunta a elevar la calidad de la enseñanza, haciendo hincapié en instituciones situadas en contextos de mayor vulnerabilidad.

1. Biblioteca de Aula Móvil

* Dotación de cada aula con una biblioteca móvil compuesta por al menos 150 libros cuidadosamente seleccionados y adaptados a la diversidad y nivel de los estudiantes.
* Los fondos bibliográficos serán organizados considerando niveles de dificultad, géneros y temáticas de interés, garantizando así tanto la pertinencia como la equidad en el acceso al material.
* Se busca fomentar el disfrute y el hábito lector autónomo, así como fortalecer la permanencia del recurso bibliográfico en el entorno escolar cotidiano.

1. Seguimiento Individualizado
2. Realización de evaluaciones diagnósticas de lectura al inicio, a mitad y al término del programa.
3. Con base en los resultados, se elaborarán planes de apoyo y refuerzo individualizados para los estudiantes que presenten mayores rezagos, en coordinación entre el docente del aula y el personal de apoyo especializado.
4. Este proceso garantiza la atención oportuna y la adaptabilidad de la intervención según las necesidades detectadas.

* **Estimación Presupuestaria del Piloto**

| **Concepto** | **Costo por escuela** | **Total (10 escuelas)** |
| --- | --- | --- |
| Materiales didácticos | $2,000 | $20,000 |
| Capacitación docente | $3,000 | $30,000 |
| Evaluaciones y seguimiento | $500 | $5,000 |
| Total estimado | 3,500 | $55,000 |

El presupuesto propuesto está calculado para una fase piloto en 10 escuelas, y ha sido estructurado con criterios de escalabilidad, con miras a una eventual expansión a nivel regional o nacional, sujeta a la evaluación de los resultados obtenidos.

* **Indicadores Clave de Desempeño (KPI)**
* KPI 1 – Rendimiento Académico:

Se establece como meta que al menos el 70% de los estudiantes participantes incremente en 15 puntos o más su evaluación de lectura, conforme a pruebas estandarizadas administradas antes y después de la intervención.

* KPI 2 – Fomento del Hábito Lector:

Lograr un aumento del 30% en el promedio de libros leídos por estudiante durante el año escolar, según los registros de préstamo y los diarios individuales de lectura.

* KPI 3 – Comprensión Lectora:

Obtener una mejora mínima del 20% en los indicadores de comprensión lectora en evaluaciones internas específicas, cubriendo dimensiones como inferencia, interpretación y reflexión crítica.

El programa “La Lectura Nos Une” representa una respuesta integral a los desafíos educativos asociados al contexto socioeconómico, priorizando la calidad pedagógica, el acceso igualitario a materiales y el monitoreo personalizado como ejes de transformación educativa. La formalización de sus componentes, su presupuesto detallado y la definición clara de indicadores de éxito sientan las bases para una implementación rigurosa, monitoreable y susceptible de réplica en escenarios de mayor escala.

## **Problema 2 — PISA 2022: Matemática → Ciencias (80 estudiantes)**

1. **Outliers: Estudiantes ±2 desviaciones estándar fuera de la recta de regresión**

En el contexto de un análisis de regresión lineal aplicado a los puntajes de Matemática (variable independiente, x) y Ciencias (variable dependiente, y), los valores atípicos (outliers) se definen como aquellos estudiantes cuyo rendimiento en Ciencias se distancia notablemente de la predicción obtenida a partir de su desempeño en Matemática. Técnica y estadísticamente, se consideran atípicos aquellos casos cuyo residuo, es decir, la diferencia entre el valor observado (yi) y el valor estimado por el modelo ()—supera en dos desviaciones estándar, ya sea por encima o por debajo del promedio de los residuos.

La identificación de estos puntos no solo altera la visualización de la recta de regresión, sino que también puede influir de manera significativa en la pendiente del modelo y en el coeficiente de determinación (R2), afectando con ello la precisión y validez de las conclusiones estadísticas.

**Razones Académicas para la Aparición de Outliers**

El surgimiento de outliers no implica necesariamente un error de medición o anomalía casual, sino que suele ser reflejo de realidades pedagógicas complejas y multifactoriales. A continuación, se desarrollan las principales razones académicas que pueden explicar la presencia de estos casos atípicos:

* **Estudiantes con Fortalezas Específicas en Ciencias**
* Presentan puntajes marcadamente altos en Ciencias aun teniendo un rendimiento bajo o promedio en Matemática.
* Sus fortalezas pueden estar asociadas con intereses personales orientados a la exploración y comprensión del mundo natural, así como una participación en actividades extracurriculares (clubes de ciencia, ferias científicas).
* Un alto nivel de competencia en lectura comprensiva favorece la interpretación y resolución de situaciones científicas, influyendo positivamente en el desempeño en Ciencias, incluso más que en el área matemática.
* Estos casos evidencian que el aprendizaje académico no es lineal ni depende exclusivamente de una sola competencia.
* **Dominio Mecánico en Matemática sin Transferencia a Ciencias**
* Algunos estudiantes pueden mostrar elevados puntajes en Matemática producto de un dominio mecánico de algoritmos y cálculos rutinarios, sin lograr transferir estas habilidades a contextos científicos.
* Frecuentemente carecen de:
  + Habilidades para la aplicación contextual y el análisis de fenómenos naturales.
  + Capacidad para establecer relaciones entre fórmulas matemáticas y situaciones concretas en Ciencias.
  + Competencia en la lectura crítica e interpretación de gráficos o representaciones científicas.
* Este tipo de disociación revela que el desarrollo matemático no necesariamente se traduce en un desempeño equivalente en Ciencias.
* **Incidencia de Factores Personales, Emocionales o de Entorno**
* El bajo rendimiento en Ciencias puede encontrarse influido por variables no necesariamente cognitivas, tales como:
  + - Motivación escasa o desinterés hacia la materia.
    - Ansiedad o baja autoestima académica vinculada a la asignatura.
    - Problemas personales o familiares que repercuten negativamente en la concentración y el desempeño escolar.
    - Dificultades frente a ciertos formatos de evaluación, como las preguntas abiertas o la experimentalidad exigida en Ciencias.
* Incluso estudiantes con buenas habilidades matemáticas se ven afectados cuando no logran establecer una conexión significativa con los contenidos científicos.

**Importancia de la Identificación y Análisis de Outliers**

La detección de valores atípicos reviste una importancia que trasciende lo estadístico. Su análisis posibilita a los docentes y equipos de gestión educativa:

* Comprender en mayor profundidad la diversidad y heterogeneidad de los perfiles estudiantiles.
* Diseñar intervenciones pedagógicas más pertinentes y ajustadas a las verdaderas necesidades del alumnado.
* Revisar críticamente las estrategias de enseñanza y fortalecer las conexiones curriculares entre áreas como Matemática y Ciencias.

En definitiva, los outliers deben interpretarse como manifestaciones de la complejidad inherente al proceso de aprendizaje humano. Lejos de considerarse simples “errores del modelo”, constituyen puntos clave para la reflexión y la mejora pedagógica y curricular.

1. Pendiente vs planes de estudio: ¿Es coherente la mejora en Ciencias sugerida por β1​?

La pendiente estimada en el modelo de regresión lineal, β1=0.70458, representa la tasa esperada de variación en el rendimiento en Ciencias asociada a un incremento unitario en los puntajes de Matemática. En términos prácticos, una mejora de 10 puntos en Matemática implica, según el modelo, un aumento aproximado de:

Δy=β1⋅Δx=0.70458

Esto indica que se prevé una mejora cercana a 7.05 puntos en Ciencias. Dicho valor refleja una relación positiva entre ambas disciplinas; sin embargo, esta relación no es completamente proporcional ni completa. La magnitud moderada de la pendiente es congruente con el marco educativo, dado que Matemática y Ciencias comparten competencias fundamentales, aunque cada una posee características y demandas propias.

**Competencias Compartidas que Respaldan la Relación Observada**

* **Razonamiento lógico:** Ambas asignaturas requieren la aplicación de un pensamiento estructurado, la capacidad de deducir e inferir conclusiones fundamentadas.
* **Lectura e interpretación de gráficos, tablas y fórmulas:** Las ciencias naturales incorporan abundantemente elementos cuantitativos que el estudiante debe analizar y comprender adecuadamente.
* **Resolución de problemas:** En Matemática y Ciencias se utilizan metodologías para plantear y solucionar problemas, formular hipótesis y verificar resultados experimentales o teóricos.

Estas competencias comunes facilitan que el desempeño en una disciplina influya positivamente, aunque indirectamente, en el rendimiento de la otra.

A pesar de la existencia de esta relación positiva, el coeficiente de determinación R2≈0.1677indica que solo alrededor del 16.8% de la variabilidad en los puntajes de Ciencias es explicada por los resultados en Matemática. Este hallazgo evidencia que, aunque Matemática contribuye al desempeño científico, la mayor parte del aprendizaje en Ciencias depende de factores adicionales que el modelo estadístico no capta.

Esta observación tiene implicaciones pedagógicas y curriculares relevantes. El aprendizaje en Ciencias requiere, además de competencias matemáticas, el desarrollo de habilidades como:

* **Comprensión lectora técnica y científica:** Interpretación y análisis crítico de textos especializados.
* **Habilidades prácticas:** Capacidades vinculadas a la realización de experimentos, observación directa y manejo de material científico.
* **Argumentación basada en evidencia:** Construcción de razonamientos fundamentados en datos empíricos.
* **Comunicación científica:** Elaboración de hipótesis, informes y explicaciones escritas coherentes y precisas.

Por tanto, resulta insuficiente trasladar automáticamente estrategias pedagógicas de Matemática al área de Ciencias con la expectativa de obtener un impacto equivalente. Cada disciplina demanda enfoques metodológicos específicos, capaces de atender sus particularidades y promover un aprendizaje integral y significativo.

1. **Plan de mejora curricular: Contenidos integrados y medición de impacto**

**Introducción a la Propuesta**

El programa "Matemática en Contexto Científico" tiene como objetivo fundamental superar la enseñanza compartimentada de las disciplinas, promoviendo una integración explícita de los contenidos matemáticos dentro del currículo de Ciencias. El propósito es que los estudiantes perciban las matemáticas no como una asignatura aislada o puramente abstracta, sino como una herramienta esencial y aplicada para la comprensión, análisis y modelado de fenómenos científicos reales [1](https://editorialeidec.com/wp-content/uploads/2022/08/CAPITULO-XVII-INTEGRALES-DEFINIDAS-E-INDEFINIDAS-USANDO-GEOGEBRA.pdf). Este enfoque didáctico busca establecer conexiones significativas, partiendo de situaciones científicas concretas —como experimentos, fenómenos físicos o simulaciones de laboratorio— que requieran la aplicación de herramientas matemáticas para su resolución, interpretación o modelización.

**Ajustes Curriculares Sugeridos**

La implementación de este programa implicará una serie de ajustes en el diseño curricular, orientados a la transversalidad y aplicación práctica de los conocimientos:

* + **Álgebra Básica Aplicada a Física y Química:**

Se diseñarán actividades donde los estudiantes manipulen fórmulas científicas relevantes (e.g., densidad = masa / volumen, velocidad = distancia / tiempo). El enfoque será el despeje de incógnitas, la sustitución de variables y la resolución algebraica de ecuaciones en contextos científicos. Esta aproximación permitirá a los estudiantes comprender que el álgebra es un lenguaje para describir relaciones entre magnitudes en la naturaleza, más allá de la mera resolución de ejercicios abstractos. Un ejemplo concreto es la resolución de problemas de diluciones en química mediante la aplicación de proporciones o ecuaciones simples.

* **Gráficas y Funciones a Partir de Datos Experimentales:**

En lugar de la enseñanza abstracta de la graficación, se fomentará que los estudiantes recolecten datos reales de experimentos (e.g., temperatura, masa, tiempo) para construir tablas y gráficos. Posteriormente, analizarán las tendencias y comportamientos funcionales que emergen de estos datos. Por ejemplo, los estudiantes podrían graficar la variación de la temperatura al calentar un líquido y relacionar la curva resultante con conceptos científicos como el calor específico o los cambios de estado. El aprendizaje de la matemática es propicio para el análisis de actividades cognitivas como la comprensión de textos y el razonamiento [1](https://editorialeidec.com/wp-content/uploads/2022/08/CAPITULO-XVII-INTEGRALES-DEFINIDAS-E-INDEFINIDAS-USANDO-GEOGEBRA.pdf).

* **Probabilidad y Estadística en Biología y Ecología:**

Se utilizarán modelos de probabilidad para analizar situaciones biológicas como la herencia genética (a través del cuadro de Punnett), las tasas de natalidad y mortalidad de poblaciones, o el riesgo de extinción de especies. Asimismo, se incorporará el trabajo con estadísticas de poblaciones o fenómenos ambientales. Un ejemplo ilustrativo es la estimación de la probabilidad de que un organismo exprese un rasgo genético recesivo, combinando los principios de Mendel con el cálculo probabilístico.

**Medición del Impacto del Plan**

Para evaluar la efectividad y el impacto de este plan de mejora curricular, se implementará una estrategia de medición multifacética:

* **Pruebas Diagnósticas y Post-Intervención:**

Se aplicarán evaluaciones estandarizadas en Ciencias antes y después de la implementación del programa. Estas pruebas incluirán componentes específicos que requieran el razonamiento matemático en contextos científicos, permitiendo cuantificar los progresos de los estudiantes.

* **Evaluaciones Integradas por Competencias:**

Se desarrollarán instrumentos de evaluación que trasciendan la división por asignaturas, planteando desafíos interdisciplinarios que integren y requieran la aplicación de conceptos tanto de Matemática como de Ciencias. Esto permitirá medir la capacidad de transferencia de aprendizajes entre ambas áreas. Un ejemplo sería una guía experimental que exija al estudiante recolectar datos, analizarlos matemáticamente y, finalmente, derivar conclusiones científicas fundamentadas.

* **Observación de Aula y Rúbricas de Desempeño:**

Se emplearán herramientas de evaluación cualitativa, tales como rúbricas, listas de cotejo y protocolos de observación. Estas permitirán valorar aspectos como la capacidad del estudiante para justificar resultados numéricos con argumentos científicos, la habilidad para argumentar con evidencia empírica y la claridad en la representación de información a través de tablas y gráficos.

**Resultado Esperado**

Mediante la implementación de este plan integrado, se anticipa que los estudiantes desarrollarán una comprensión más profunda, cohesionada y aplicable del conocimiento escolar. Al establecer conexiones directas entre los contenidos matemáticos y las situaciones del mundo real, se espera observar no solo una mejora en los resultados académicos en ambas disciplinas, sino también un aumento significativo en la motivación y el sentido de propósito del aprendizaje. Adicionalmente, esta estrategia busca reducir la brecha en el rendimiento entre Ciencias y Matemática que ha sido identificada por el modelo de regresión.

## **Problema 3 — ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)**

1. **Efecto post-pandemia**: **Con R2≈0.75, argumente por que la brecha Lectura-Matemática pudo haberse estrechado/ensanchado.**

El análisis del modelo de regresión aplicado al conjunto de datos revela un coeficiente de determinación elevado, R2≈0.75, lo que indica que aproximadamente el 75% de la variabilidad en los puntajes de Matemática puede ser explicada por los puntajes en Lectura. Esta fuerte asociación estadística entre ambas áreas adquiere particular relevancia en el contexto educativo posterior a la pandemia.

No obstante, la existencia de esta correlación significativa no implica que el desarrollo de las competencias en Lectura y Matemática haya sido simétrico o equivalente durante dicho periodo. La pandemia generó impactos diferenciales en el aprendizaje de ambas áreas, influenciados por una compleja interacción de factores pedagógicos, sociales y familiares. Esta situación puede analizarse desde dos perspectivas contrapuestas:

**Estrechamiento de la Brecha: Convergencia en los Desempeños**

Durante el período de educación virtual, los estudiantes enfrentaron condiciones adversas, tales como:

* Limitado acceso a internet y dispositivos tecnológicos, además de la necesidad de compartirlos en el hogar.
* Reducción del contacto directo y continuo con docentes.
* Cambios sustanciales en la dinámica familiar que afectaron rutinas de estudio.

Estas circunstancias impactaron negativamente tanto el aprendizaje en Lectura como en Matemática debido a:

* La pérdida de tiempo valioso de instrucción presencial, esencial para la adquisición progresiva y guiada de habilidades en ambas disciplinas.
* La ausencia de retroalimentación inmediata —determinante para corregir errores de comprensión lectora y para la resolución efectiva de problemas matemáticos.
* La disminución de rutinas escolares estructuradas, necesarias para sostener un aprendizaje sistemático y acumulativo.

Como consecuencia, la brecha entre el desempeño en Lectura y Matemática pudo haberse reducido no por una mejora simultánea, sino por un deterioro paralelo en ambas áreas, manteniendo así la correlación lineal observada.

**Ensanchamiento de la Brecha: Impacto Diferencial en las Competencias**

Contrariamente, es plausible que la pandemia haya afectado de manera desigual el progreso en las competencias lectoras y matemáticas:

* **Competencia lectora:** Al ser una habilidad con múltiples oportunidades de ejercicio fuera del ámbito estrictamente escolar —mediante el acceso a libros en el hogar, el consumo de contenidos digitales, las interacciones en redes sociales y la exposición a textos cotidianos como subtítulos en televisión—, algunos estudiantes pudieron continuar desarrollando o incluso fortalecer su comprensión lectora durante el confinamiento.
* **Competencia matemática:** Esta área requiere una secuencia didáctica más estructurada, instrucción directa, práctica constante y supervisión inmediata. La interrupción del contacto con el docente, la ausencia de ejercicios sistemáticos y evaluaciones frecuentes provocaron retrocesos significativos en la adquisición y consolidación de habilidades numéricas.

Por lo tanto, en este escenario, estudiantes que mantuvieron o elevaron su nivel en Lectura no necesariamente reflejaron avances equivalentes en Matemática, generándose un ensanchamiento de la brecha entre ambas áreas.

El análisis estadístico combinado con la interpretación pedagógica sugiere que, si bien la fuerte relación entre Lectura y Matemática persiste, la pandemia pudo haber incidido en la amplitud de la brecha de rendimiento entre estas competencias de manera heterogénea. Esta realidad exige reflexionar sobre intervenciones diferenciadas y direccionadas, que reconozcan las especificidades de cada área y atiendan las necesidades particulares de recuperación y fortalecimiento post-pandemia.

1. **Pendiente < 1: Interprete pedagógicamente que los incrementos en Lectura se traduzcan en incrementos proporcionalmente menores en Matemática.**

El análisis del modelo de regresión lineal aplicado al conjunto de datos del ERCE 2023 arrojó una pendiente estimada de β1=0.51994. Este coeficiente indica que, en promedio, por cada punto adicional obtenido en Lectura, se espera un incremento aproximado de 0.52 puntos en Matemática. Si bien la relación entre ambas variables es claramente positiva, el hecho de que la pendiente sea inferior a uno sugiere que las mejoras en una competencia no se traducen en incrementos proporcionales en la otra.

**Implicaciones Pedagógicas del Resultado**

Este hallazgo pone en evidencia que, aunque existe una interrelación entre el desarrollo de habilidades lectoras y el desempeño matemático, el dominio en Matemática no depende exclusivamente de la comprensión lectora. La Lectura desempeña un rol fundamental como competencia transversal, por ejemplo, en la interpretación de enunciados, la lectura de gráficos o el seguimiento de instrucciones. Sin embargo, la naturaleza específica de la matemática requiere la incorporación de habilidades cognitivas adicionales y especializadas, tales como:

* **Razonamiento abstracto:** Capacidad para manipular símbolos, fórmulas y operaciones en un nivel no concreto.
* **Visualización espacial:** Fundamental para áreas como la geometría y la representación gráfica.
* **Cálculo numérico:** Dominio de operaciones básicas, manejo de fracciones y números decimales.
* **Pensamiento lógico y estructurado:** Necesario para seguir procedimientos metodológicos y justificar soluciones formales.

Por tanto, aunque el fortalecimiento de la competencia lectora facilite el acceso y comprensión de problemas matemáticos, ello no garantiza en sí mismo que el estudiante posea las herramientas necesarias para su resolución adecuada. Esta circunstancia explica la pendiente menor a la unidad: la Lectura impulsa el aprendizaje matemático, pero no lo determina en su totalidad.

**Estrategias Pedagógicas Recomendadas en Función del Hallazgo**

La evidencia de un efecto parcial de la Lectura sobre Matemática refuerza la necesidad de adoptar un enfoque pedagógico dual e integrado, que contemple:

* **Fomento de la comprensión lectora en contextos matemáticos:** Mediante la enseñanza de habilidades específicas para identificar términos clave en problemas, descomponer enunciados complejos y elaborar esquemas visuales que faciliten la interpretación de situaciones matemáticas.
* **Fortalecimiento de habilidades lógico-matemáticas específicas:** A través de actividades de práctica estructurada, resolución sistemática de ejercicios, incorporación de juegos didácticos y andamiaje pedagógico orientado a la solución de problemas complejos.

La pendiente inferior a uno constituye un recordatorio estadístico de que no todas las habilidades escolares mantienen una conexión homogénea ni proporcional. Aunque Lectura y Matemática presentan puntos de convergencia, cada disciplina posee trayectorias de aprendizaje particulares que requieren intervenciones pedagógicas diferenciadas, diseñadas para potenciar sus competencias específicas y, al mismo tiempo, favorecer políticas educativas integrales y complementarias.

1. **Intervenciones socioemocionales: Liste dos estrategias que potencien ambas áreas simultáneamente.**

El desarrollo de habilidades cognitivas en Lectura y Matemática depende no únicamente de factores académicos, sino que está profundamente influenciado por componentes socioemocionales que determinan la manera en que los estudiantes afrontan desafíos, procesan información y se relacionan con el aprendizaje. Aspectos tales como la autoconfianza, el manejo de la ansiedad, la perseverancia, el trabajo colaborativo y la motivación intrínseca constituyen variables cruciales para el éxito académico en ambas disciplinas.

Por esta razón, incorporar estrategias socioemocionales de manera deliberada en el ámbito escolar puede generar un impacto positivo de carácter transversal, mejorando tanto el rendimiento lector como el desempeño matemático. A continuación, se describen dos intervenciones concretas, junto con la articulación de sus efectos cruzados:

* **Círculos de Confianza y Lectura Compartida**

Al inicio de la jornada escolar se establece una rutina diaria o semanal en la que los estudiantes leen en voz alta fragmentos breves de textos, cuentos o frases con contenido motivacional. Posteriormente, se abre un espacio de diálogo en el que cada estudiante comparte su interpretación del texto y expresa sus emociones respecto al contenido leído. Esta dinámica se desarrolla en formato de círculo, garantizando que todos los participantes tengan oportunidad de hablar y ser escuchados en un ambiente de respeto y confianza.

Objetivos Pedagógicos:

* Fortalecer la fluidez lectora y fomentar una comprensión emotiva y profunda del texto.
* Promover la expresión auténtica de emociones, contribuyendo a la reducción de tensiones que puedan obstaculizar el aprendizaje.
* Desarrollar habilidades socioemocionales como la empatía y la escucha activa, elementos fundamentales para la convivencia armónica dentro del salón de clases.

Impacto en Lectura:

Esta intervención mejora significativamente la capacidad interpretativa, la entonación y la seguridad al momento de leer en voz alta, transformando la lectura en una experiencia significativa y no meramente mecánica.

Impacto en Matemática:

Aunque no se interviene directamente en contenidos matemáticos, el fortalecimiento de la autoconfianza y la comodidad para expresar ideas puede traducirse en una disminución de la ansiedad asociada a la resolución de problemas matemáticos, especialmente en contextos evaluativos.

2. Resolución Colaborativa de Problemas Contextualizados

Se plantean situaciones contextualizadas de la vida cotidiana que requieren tanto comprensión lectora como razonamiento matemático. Ejemplos incluyen planificar un almuerzo con un presupuesto limitado, calcular distancias empleando un mapa o coordinar la organización de un evento escolar. Los estudiantes trabajan en grupos pequeños para interpretar la situación, identificar datos relevantes, plantear una solución viable y explicar sus decisiones al grupo.

Objetivos Pedagógicos:

* Integrar la lectura comprensiva con el razonamiento numérico en escenarios funcionales y significativos.
* Fomentar habilidades sociales como la cooperación, el respeto por la diversidad de ideas y la resolución constructiva de conflictos.
* Promover la tolerancia a la frustración y el manejo adecuado de emociones ante desacuerdos o errores que puedan surgir durante el trabajo colaborativo.

Impacto en Lectura:

Se fortalece la habilidad para extraer información específica, realizar inferencias adecuadas y justificar decisiones fundamentadas en textos o instrucciones escritas.

Impacto en Matemática:

Además de ejecutar cálculos, el estudiante desarrolla un razonamiento lógico aplicado a problemas reales, adquiriendo un sentido de propósito mayor. La interacción con compañeros permite la exposición a diversas estrategias cognitivas, ampliando así la flexibilidad mental.

Las intervenciones socioemocionales descritas no demandan recursos materiales significativos, pero requieren de una intención pedagógica clara y consistencia en su implementación. Debido a su naturaleza transversal y contextualizada, estas prácticas enriquecen simultáneamente el aprendizaje en Lectura y Matemática, al tiempo que:

* Mejoran el clima y la convivencia escolar.
* Incrementan la motivación intrínseca del estudiantado.
* Desarrollan habilidades esenciales para la vida, como la comunicación efectiva, la empatía y la autorregulación emocional.

En conjunto, estas estrategias contribuyen a la promoción de aprendizajes integrales, sostenibles y equitativos, especialmente en contextos donde las brechas educativas presentan una mayor vulnerabilidad a la ampliación.

## **Contraste entre los problemas**

Los tres modelos de regresión lineal desarrollados en el presente estudio abordan relaciones diferenciadas dentro del ámbito educativo, lo que posibilita un contraste no solo en términos de la fuerza estadística de las asociaciones analizadas, sino también en cuanto a su significado pedagógico y contextual. El análisis integrado de estos modelos revela diferencias sustanciales que enriquecen la comprensión acerca de cómo diversos factores se vinculan con los procesos de aprendizaje.

**Problema 1: Índice Socioeconómico (ESCS) y Desempeño en Lectura (PISA 2022)**

Este modelo evidenció la pendiente más elevada entre los tres casos estudiados, con un valor estimado de β1≈19.97. Esto implica que variaciones relativamente pequeñas en el índice socioeconómico del estudiante se asocian a cambios significativos en su rendimiento en lectura. Complementariamente, el coeficiente de determinación R2≈0.71 indica que más del 70% de la variabilidad en los puntajes de lectura queda explicada por dicho índice, posicionando este modelo como el más robusto en términos predictivos.

Desde una perspectiva pedagógica, este resultado subraya la profunda influencia que el contexto socioeconómico ejerce sobre el desempeño escolar. No solo pone de manifiesto una brecha en la igualdad de oportunidades educativas, sino que también enfatiza la relevancia de implementar políticas de compensación y apoyo específicas para estudiantes en situación de vulnerabilidad, dado que las condiciones familiares y sociales continúan siendo factores determinantes en el desarrollo de competencias lectoras.

**Problema 2: Desempeño en Matemática y Ciencias (PISA 2022)**

En este análisis, la pendiente registrada fue más moderada, aproximadamente β1≈0.70, y el coeficiente de determinación considerablemente inferior, con R2≈0.17. Estos valores evidencian que, aunque existe una relación estadística entre los puntajes en Matemática y Ciencias, el rendimiento en Matemática explica únicamente un 17% de la variabilidad en Ciencias, sugiriendo la presencia de otros factores influyentes significativos en esta última.

Desde el ámbito pedagógico, esta diferenciación resalta que, pese a la compartición de procesos cognitivos tales como la interpretación de datos, el razonamiento lógico y el análisis gráfico, el aprendizaje en Ciencias demanda además competencias específicas, como la formulación de hipótesis, el trabajo experimental y la escritura científica. Por ende, la mejora en Matemática no constituye una garantía automática de progreso en Ciencias, señalando la necesidad de enfoques curriculares integrados que conjunta y coherentemente articulen la enseñanza matemática en el contexto científico, en lugar de abordarlas de manera aislada.

**Problema 3: Competencias en Lectura y Matemática en Estudiantes de 3. ° Grado (ERCE 2023)**

Este modelo mostró una pendiente menor, estimada en β1≈0.52, al tiempo que exhibió el coeficiente de determinación más alto, con R2≈0.75. En términos prácticos, esto indica que las variaciones en los puntajes de Lectura explican una elevada proporción de la variabilidad en los puntajes matemáticos; sin embargo, el incremento en competencias lectoras no se traduce proporcionalmente en incrementos equivalentes en Matemática.

Este resultado sugiere que, en niveles iniciales como el tercer grado, la lectura constituye una habilidad fundamental que sustenta el aprendizaje matemático, dado que la comprensión de enunciados, la interpretación de instrucciones y el análisis de problemas verbales dependen en gran medida del desarrollo lector. No obstante, la pendiente inferior a la unidad evidencia que el aprendizaje matemático requiere también de procesos particulares, incluyendo cálculo numérico, razonamiento formal, resolución algorítmica y pensamiento abstracto.

Adicionalmente, este caso refleja el posible impacto de la pandemia sobre las trayectorias educativas, puesto que numerosas cohortes estudiantiles experimentaron alteraciones en su desarrollo académico tanto en Lectura como en Matemática. La elevada correlación sugiere que los avances o retrocesos en una competencia se reflejan en la otra de manera casi inmediata, lo cual pone de manifiesto una relación de interdependencia más que una causalidad directa.

La comparación entre estos tres escenarios evidencia que la naturaleza y magnitud de las relaciones entre variables educativas presentan variaciones sustanciales, condicionadas por el contexto, el nivel escolar y las competencias específicas evaluadas. Mientras que el índice socioeconómico ejerce una influencia marcada y directa sobre el rendimiento en lectura, las relaciones entre disciplinas como las que se observan entre Matemática y Ciencias o entre Lectura y Matemática dependen en mayor medida de factores asociados al currículo, la pedagogía y el componente socioemocional.

Este panorama reafirma la relevancia de complementar los análisis estadísticos con una interpretación crítica y contextualizada, que no solo cuantifique las asociaciones observadas, sino que también oriente el diseño e implementación de acciones educativas precisas, equitativas y adaptadas a las realidades particulares de los estudiantes.

# **Conclusión**

El análisis de regresión lineal efectuado sobre los tres conjuntos de datos educativos permitió identificar relaciones estadísticas significativas entre distintas variables asociadas al desempeño estudiantil, proporcionando evidencia robusta tanto para la comprensión cuantitativa como para la orientación de decisiones pedagógicas informadas.

Problema 1 (PISA 2022: ESCS → Lectura): Se observó una pendiente positiva considerable, estimada en β1≈19.97, junto con un coeficiente de determinación moderado, R2≈0.71. Estos resultados evidencian que el índice socioeconómico tiene un impacto directo y sustantivo sobre los puntajes en lectura. Sin embargo, el modelo también revela la existencia de factores escolares y personales no capturados, lo que subraya la necesidad de intervenciones integrales y contextualizadas para mitigar la desigualdad educativa.

Problema 2 (PISA 2022: Matemática → Ciencias): La relación entre el rendimiento en Matemática y Ciencias fue moderadamente débil, con una pendiente aproximada de β1≈0.70 y un coeficiente de determinación bajo, R2≈0.17. Esto indica que, si bien existe cierta coherencia entre ambas áreas, la predicción del desempeño en Ciencias a partir de Matemática no es robusta. Este hallazgo sugiere que el aprendizaje científico demanda habilidades adicionales, tales como la interpretación de fenómenos naturales, el razonamiento experimental y la comprensión lectora en contextos científicos. En consecuencia, se propuso un ajuste curricular que integre contenidos matemáticos inmersos en situaciones científicas, con el objetivo de fortalecer simultáneamente ambas áreas del conocimiento.

Problema 3 (ERCE 2023: Lectura → Matemática, tercer grado): Se identificó una fuerte correlación entre las competencias lectoras y matemáticas en estudiantes de primer ciclo, evidenciado por un coeficiente de determinación alto, R2≈0.75, y una pendiente inferior a la unidad, β1≈0.52. Esto sugiere que, aunque el desarrollo lector impulsa el aprendizaje matemático, no es el único factor determinante. La situación post-pandemia podría haber incrementado la interdependencia entre las trayectorias académicas en estas áreas, enfatizando la importancia de estrategias pedagógicas y socioemocionales que promuevan un aprendizaje integrado y holístico.

En su conjunto, estos resultados demuestran que la regresión lineal constituye una herramienta analítica valiosa no solo para modelar las relaciones entre variables educativas, sino también para fomentar una reflexión profunda sobre las interconexiones entre las distintas dimensiones del aprendizaje. Además, estos hallazgos ponen de relieve la importancia de desarrollar políticas educativas y programas escolares fundamentados en evidencia empírica, que aborden simultáneamente los aspectos cognitivos y socioemocionales, y que consideren las particularidades contextuales de la población estudiantil.

# **Referencias Bibliográficas**

[1] OECD, PISA 2022 Results – Panama Factsheet, París: OCDE, 2023. [En línea]. Disponible: <https://www.oecd.org/pisa>

[2] UNESCO-OREALC y Ministerio de Educación de Panamá, Resumen ejecutivo ERCE pospandemia Panamá, Ciudad de Panamá, 2025.

[3] FUDESPA y JUXLAE, Un sexenio perdido: análisis de la política educativa panameña 2019–2024, Panamá, 2025.

[4] J. Batista, Método de Mínimos Cuadrados, Universidad Tecnológica de Panamá, s.f. [Material de clase en PDF].

[5] UNESCO, Education in a Post-COVID World: Nine Ideas for Public Action, París: UNESCO, 2024.

[6] L. R. Ríos y M. E. Caballero, "Desigualdad y aprendizaje escolar en América Latina: una mirada desde las pruebas ERCE", Revista Panameña de Investigación Educativa, vol. 7, no. 2, pp. 23–41, 2025.

[7] G. Vergara y M. Gómez, "Factores asociados al rendimiento en ciencias en PISA: análisis comparado regional", Revista de Evaluación y Currículo, vol. 9, no. 1, pp. 52–69, 2025.

[8] J. H. Stock y M. W. Watson, *Introducción a la econometría*, 4.ª ed., Pearson, 2019.

[9] D. C. Montgomery, E. A. Peck y G. G. Vining, *Introducción al análisis de regresión lineal*, 5.ª ed., Wiley, 2012.

[10] UNESCO, *Estudio Regional Comparativo y Explicativo (ERCE 2019): Informe de resultados*, Oficina Regional de Educación para América Latina y el Caribe (OREALC/UNESCO), Santiago, 2021.

[11] L. Sirin, “Estatus socioeconómico y rendimiento académico: una revisión meta-analítica de la investigación”, *Review of Educational Research*, vol. 75, n.º 3, pp. 417–453, otoño 2005.

[12] E. Hanushek y L. Woessmann, “El rol de la calidad educativa en el crecimiento económico”, *Documento de Trabajo de Investigación del Banco Mundial n.º 4122*, 2007. [En línea]. Disponible: https://doi.org/10.1596/1813-9450-4122

[13] M. Kuhfeld et al., “Aprendizaje durante la COVID-19: Hallazgos iniciales sobre el rendimiento y el crecimiento en lectura y matemáticas”, *Informe de investigación de NWEA*, 2020. [En línea]. Disponible: <https://www.nwea.org/research/publication/learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth/>

[14] B. Fraillon, J. Ainley, W. Schulz, T. Friedman y D. Gebhardt, *Prepararse para la vida en un mundo digital: Informe internacional del Estudio Internacional sobre Alfabetización Informática y en Tecnología de la Información y la Comunicación 2018 de la IEA*, Springer, 2019.

# **Apéndice**

**Código** **Problema 1 — PISA 2022: ESCS → Lectura (85 escuelas)**

import numpy as np

import matplotlib.pyplot as plt

# Conjunto de datos PISA 2022 – 85 observaciones

# x: Índice socioeconómico (ESCS)

# y: Puntaje en Lectura (Reading\_score)

x = np.array([

    -1.19, -0.631, 0.859, -0.125, 0.528, 0.219, -0.873, 0.188, -0.585, -0.59,

    -0.213, -1.673, 0.773, -0.904, 0.6, -0.177, 1.079, -0.894, -0.581, 0.004,

    -2.287, 0.445, 1.087, 0.714, 0.379, -0.431, 0.854, 0.667, 0.053, -0.295,

    -0.626, -1.407, 0.804, -2.255, -0.633, -0.152, 0.474, 1.285, 0.594, -0.562,

    0.355, -1.435, 0.987, -0.441, -0.906, -0.875, -0.355, -0.654, 1.761, -0.946,

    -0.271, -0.275, -0.275, -0.25, -0.733, -0.825, -1.076, -1.639, -0.105, 0.586,

    -0.789, -0.803, -0.585, -0.715, -1.593, 0.929, 0.095, -0.941, -0.033, -0.695,

    -0.49, 0.028, 0.558, 1.068, 1.234, -0.524, -0.75, -0.21, -0.184, -0.961,

    -0.858, 0.432, 1.178, -0.504, -0.883

])

y = np.array([

    355.2, 356.7, 392.2, 382.8, 388.5, 398.8, 366.1, 389.2, 367.7, 385.2,

    361.5, 346.2, 379.8, 373.6, 401.0, 370.6, 419.3, 360.3, 352.9, 376.7,

    343.2, 385.2, 426.8, 403.1, 407.6, 388.9, 375.1, 387.5, 348.2, 366.9,

    382.1, 344.6, 386.8, 343.3, 378.1, 373.0, 385.5, 412.8, 393.7, 366.5,

    390.2, 355.1, 400.8, 396.1, 360.0, 357.8, 390.6, 359.8, 432.3, 360.6,

    376.0, 383.9, 389.6, 387.5, 364.4, 360.8, 364.8, 352.1, 370.4, 395.5,

    377.7, 367.5, 388.2, 372.1, 349.5, 397.9, 375.0, 350.7, 385.1, 357.3,

    366.9, 379.4, 394.7, 404.8, 402.9, 347.4, 364.0, 366.9, 375.3, 374.6,

    374.2, 390.4, 407.3, 353.1, 366.4

])

# Cálculo de la regresión lineal

beta1, beta0 = np.polyfit(x, y, 1)

# Predicciones

y\_pred = beta1 \* x + beta0

# R² (coeficiente de determinación)

ss\_res = np.sum((y - y\_pred) \*\* 2)

ss\_tot = np.sum((y - np.mean(y)) \*\* 2)

r2 = 1 - (ss\_res / ss\_tot)

# Resultados

print(f"Pendiente (β1): {beta1:.5f}")

print(f"Intercepto (β0): {beta0:.5f}")

print(f"R²: {r2:.5f}")

# Gráfico

plt.figure(figsize=(8, 5))

plt.scatter(x, y, color='#008080', label='Observaciones')

plt.plot(x, y\_pred, color='#CD5C5C', label=f'Recta de regresión: y = {beta1:.5f}x + {beta0:.5f}')

plt.title(f'Relación entre ESCS y Lectura – PISA 2022\n$R^2$ = {r2:.5f}')

plt.xlabel('Índice socioeconómico (ESCS)')

plt.ylabel('Puntaje en Lectura')

plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show()

**Código Problema 2 — PISA 2022: Matemática → Ciencias (80 estudiantes)**

import numpy as np

import matplotlib.pyplot as plt

# Nuevos datos

x = np.array([

    365, 354, 363, 357, 368, 369, 356, 372, 358, 376,

    356, 361, 362, 354, 346, 354, 369, 357, 357, 372,

    364, 366, 367, 353, 361, 361, 359, 351, 364, 347,

    353, 354, 357, 367, 373, 374, 366, 364, 341, 369,

    381, 364, 362, 366, 366, 361, 357, 363, 349, 345,

    371, 356, 366, 372, 370, 356, 363, 351, 359, 368,

    360, 337, 342, 354, 371, 362, 361, 349, 361, 342,

    363, 369, 353, 369, 354, 361, 359, 361, 372, 367

])

y = np.array([

    393, 355, 413, 406, 407, 397, 417, 382, 386, 391,

    400, 392, 429, 393, 384, 401, 395, 381, 403, 401,

    415, 407, 405, 393, 398, 405, 391, 364, 402, 381,

    373, 404, 390, 379, 408, 400, 379, 405, 368, 415,

    402, 415, 405, 404, 392, 412, 403, 395, 397, 372,

    389, 394, 404, 414, 407, 412, 382, 398, 400, 424,

    400, 384, 408, 404, 415, 394, 401, 366, 387, 392,

    375, 407, 393, 410, 370, 371, 413, 400, 414, 426

])

# Calcular la regresión lineal

beta1, beta0 = np.polyfit(x, y, 1)

# Predecir valores

y\_pred = beta1 \* x + beta0

# Calcular R²

ss\_res = np.sum((y - y\_pred) \*\* 2)

ss\_tot = np.sum((y - np.mean(y)) \*\* 2)

r2 = 1 - (ss\_res / ss\_tot)

# Mostrar resultados

print(f"Pendiente (β1): {beta1:.5f}")

print(f"Intercepto (β0): {beta0:.5f}")

print(f"R²: {r2:.5f}")

# Graficar

plt.figure(figsize=(8, 5))

plt.scatter(x, y, color='#008080', label='Datos reales')

plt.plot(x, y\_pred, color='#CD5C5C', label=f'Recta de regresión  y = {beta1:.5f}x + {beta0:.5f}')

plt.title(f'Puntaje en Matemáticas vs Puntaje en Ciencias – PISA 2022\n$R^2$ = {r2:.5f}')

plt.xlabel('Puntaje en Matemáticas')

plt.ylabel('Puntaje en Ciencias')

plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show()

**Código Problema 3 — ERCE 2023: Lectura → Matemática, 3. ° grado (70 estudiantes)**

import numpy as np

import matplotlib.pyplot as plt

# Nuevos datos

x = np.array([

    697, 675, 692, 703, 691, 692, 689, 710, 707, 677,

    714, 687, 715, 682, 706, 689, 693, 704, 688, 698,

    698, 707, 709, 714, 698, 699, 689, 687, 682, 693,

    697, 695, 701, 681, 712, 698, 706, 698, 686, 692,

    662, 697, 694, 707, 677, 694, 692, 682, 680, 683,

    705, 705, 681, 693, 700, 698, 680, 689, 681, 705,

    721, 695, 674, 703, 711, 710, 697, 700, 696, 702

])

y = np.array([

    720, 675, 689, 701, 695, 703, 718, 688, 726, 686,

    663, 690, 710, 703, 720, 698, 686, 725, 685, 682,

    692, 700, 709, 706, 697, 718, 712, 706, 688, 688,

    699, 699, 721, 688, 691, 689, 741, 687, 673, 694,

    712, 705, 707, 681, 663, 701, 692, 701, 672, 713,

    733, 680, 697, 696, 731, 708, 709, 696, 698, 747,

    756, 671, 695, 711, 693, 698, 706, 722, 712, 699

])

# Calcular la regresión lineal

beta1, beta0 = np.polyfit(x, y, 1)

# Predecir valores

y\_pred = beta1 \* x + beta0

# Calcular R²

ss\_res = np.sum((y - y\_pred) \*\* 2)

ss\_tot = np.sum((y - np.mean(y)) \*\* 2)

r2 = 1 - (ss\_res / ss\_tot)

# Mostrar resultados

print(f"Pendiente (β1): {beta1:.5f}")

print(f"Intercepto (β0): {beta0:.5f}")

print(f"R²: {r2:.5f}")

# Graficar

plt.figure(figsize=(8, 5))

plt.scatter(x, y, color='#008080', label='Datos reales')

plt.plot(x, y\_pred, color='#CD5C5C', label=f'Recta de regresión  y = {beta1:.5f}x + {beta0:.5f}')

plt.title(f'Puntaje de Lectura vs Puntaje de Matemáticas – ERCE 2023\n$R^2$ = {r2:.5f}')

plt.xlabel('Puntaje en Lectura')

plt.ylabel('Puntaje en Matemáticas')

plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show()