

# UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

Tecnologías de la información



## Extracción de Conocimiento en Bases de Datos

Elaboración de gráficas

**IDGS91N**

PROFESOR:

Enrique Mascote

Alumno:

Emanuel Chavira

Jorge Lopez

Raquel Almuina

Erick Maffiodo

29 de Noviembre de 2025

# Proyecto de Visualización de Datos

## 1. Introducción

El presente proyecto tiene como propósito **comprender el proceso de elaboración de gráficas** utilizando herramientas de visualización de datos. Debido a la falta de un proyecto asignado en el curso, se diseñó un **conjunto de datos ficticio** que simula las calificaciones de quince estudiantes de una asignatura de programación. Las variables registradas son la edad de cada estudiante y sus calificaciones en tres exámenes parciales. A partir de estas variables se elaboran diferentes tipos de gráficas para extraer conocimiento. La visualización de la información permite **representar de manera accesible tendencias, valores atípicos y patrones en los datos**.

## 2. Desarrollo

### 2.1 Descripción del conjunto de datos

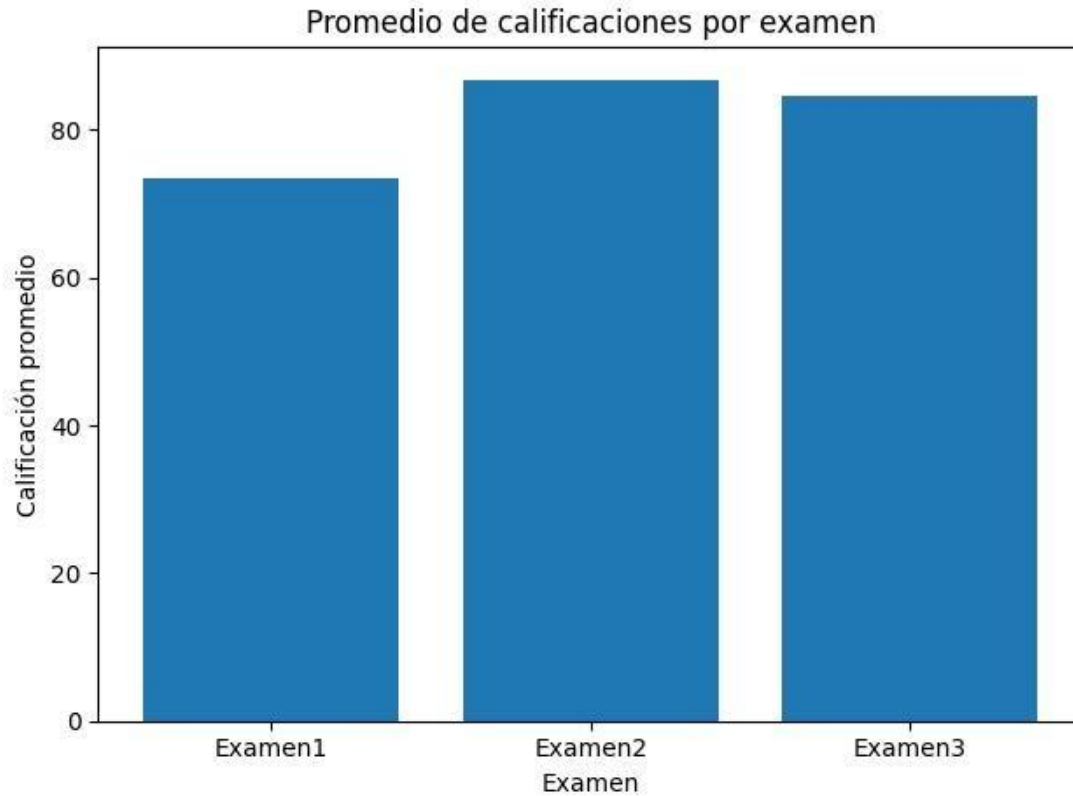
Se generó un dataset con 15 registros. Cada registro corresponde a un estudiante, identificando su edad y las calificaciones obtenidas en tres exámenes (Examen 1, Examen 2 y Examen 3). A partir de estas calificaciones se calculó el promedio final. Las medias de cada examen fueron de 73.4 (Examen 1), 86.8 (Examen 2) y 84.5 (Examen 3).

Los promedios finales oscilaron entre 76.3 y 89.3 puntos.

### 2.2 Elaboración de gráficas

A continuación se presentan las gráficas elaboradas con *Python* y la biblioteca **Matplotlib**, así como una interpretación de los resultados.

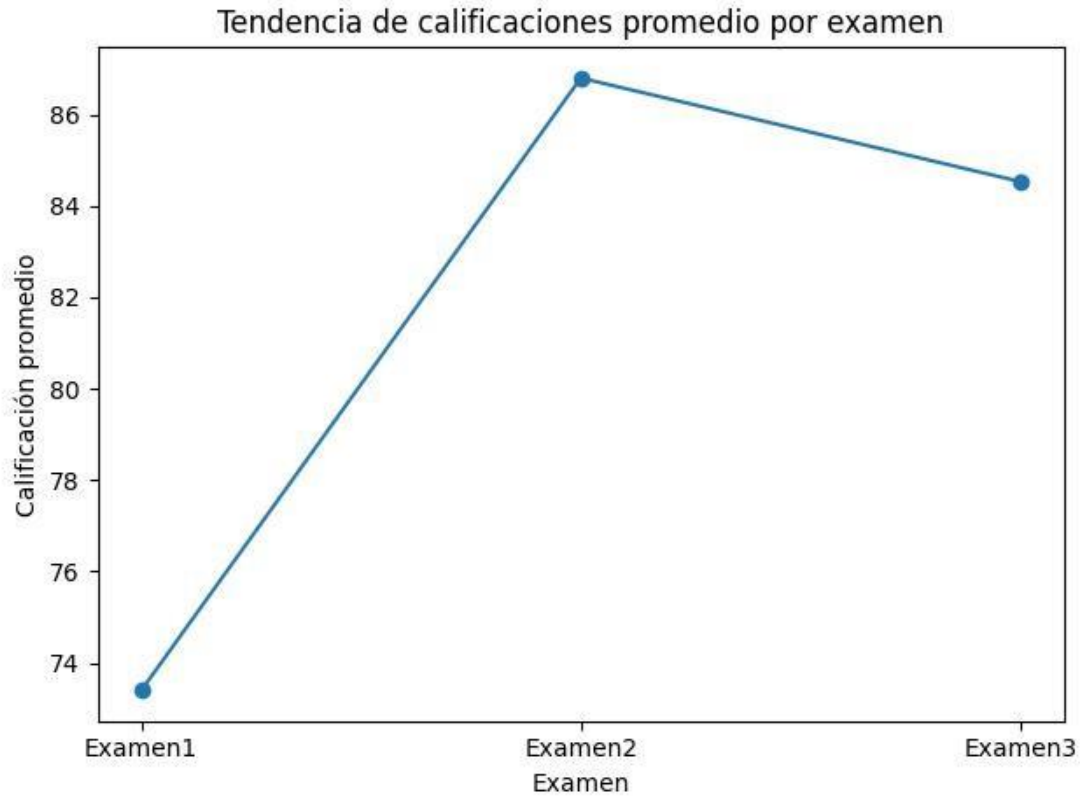
*Gráfico de barras*



#### *Promedios por examen*

El gráfico de barras muestra la media de las calificaciones obtenidas en cada examen. Las barras son adecuadas para **comparar categorías discretas**, ya que cada barra representa un examen y su altura es proporcional al promedio alcanzado. Se observa que el **Examen 2** registró la media más alta (86.8), mientras que el **Examen 1** presentó la media más baja (73.4). Esta visualización permite identificar rápidamente qué evaluación fue más desafiante para los estudiantes.

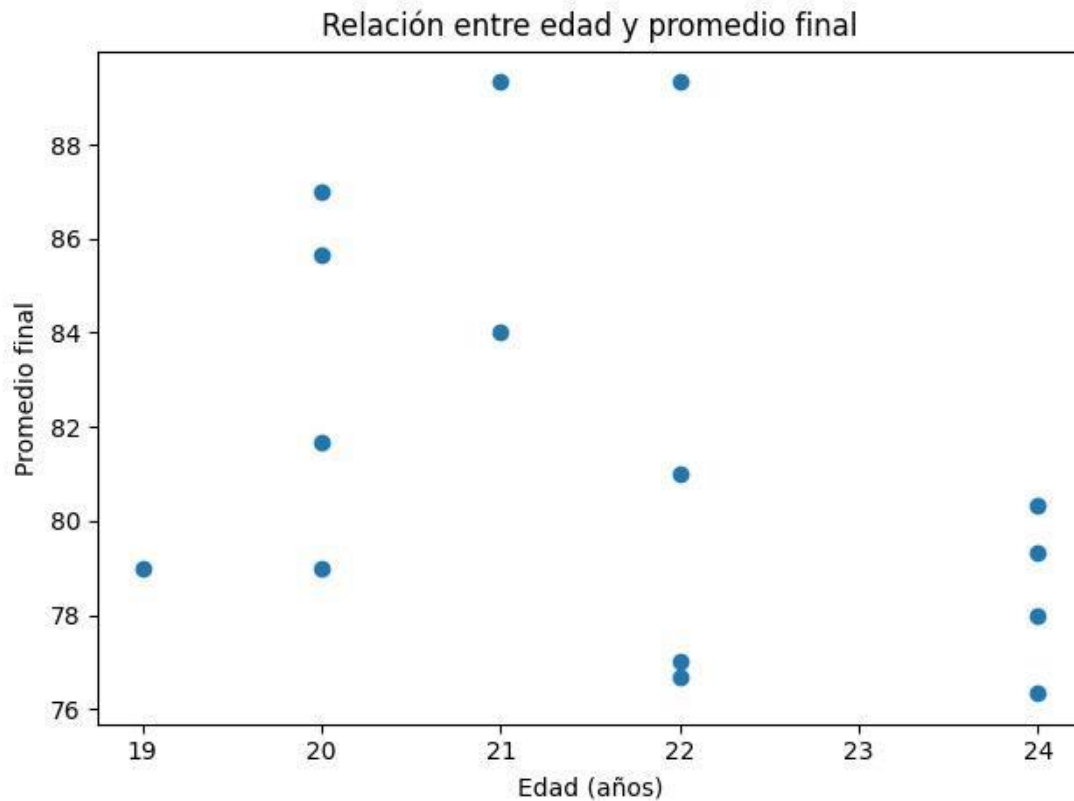
#### *Gráfico de líneas*



#### *Tendencia de promedios*

El gráfico de líneas representa las mismas medias de calificaciones, conectando los puntos para reflejar la **tendencia a lo largo de los exámenes**. Los gráficos de líneas permiten visualizar cómo una variable continua cambia en un intervalo de tiempo o secuencia ordinal, facilitando la detección de **tendencias y patrones**. Se evidencia una mejora significativa del Examen 1 al Examen 2 y una leve disminución en el Examen 3. Esta gráfica refuerza la percepción de que los estudiantes se adaptaron y mejoraron antes del segundo examen.

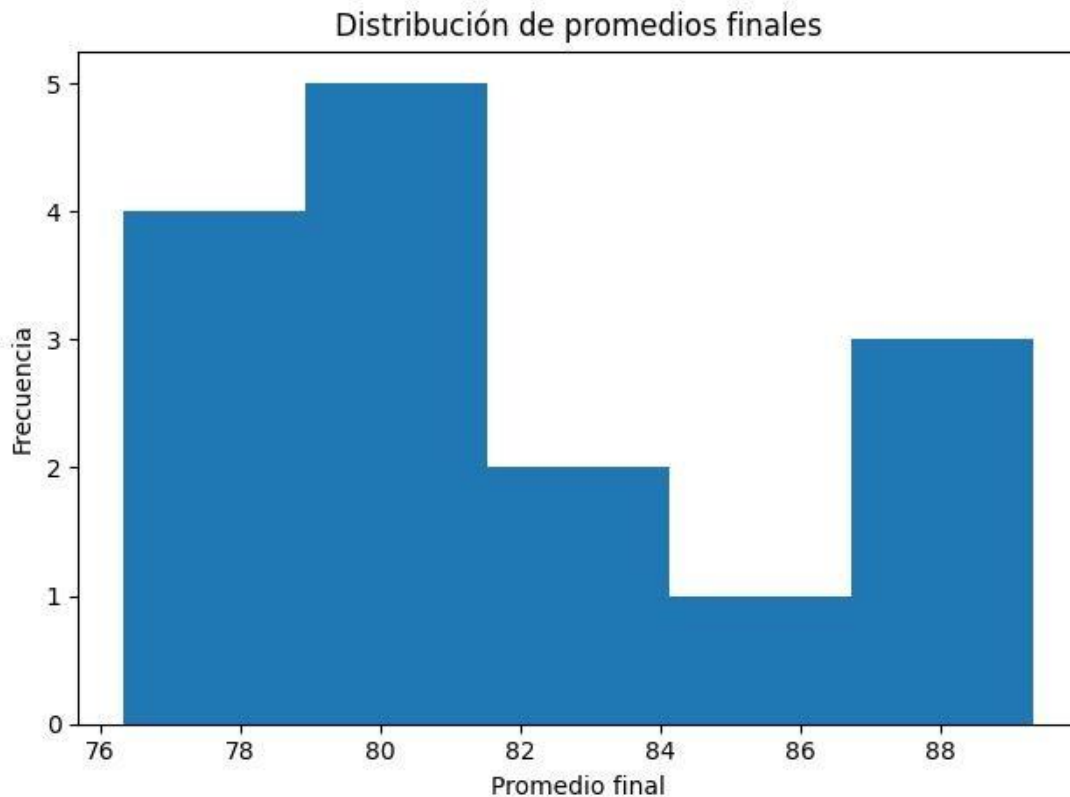
#### *Diagrama de dispersión*



#### *Edad vs. promedio*

El diagrama de dispersión presenta la relación entre la **edad de los estudiantes** y su **promedio final**. Este tipo de gráfico consiste en pares de puntos que permiten identificar posibles relaciones entre dos variables. Si los puntos se agrupan en una diagonal ascendente, se sugiere una correlación positiva; si forman una diagonal descendente, indica correlación negativa. En nuestro caso, la correlación entre edad y promedio es moderadamente negativa ( $-0.38$ ), lo que sugiere que los estudiantes más jóvenes obtuvieron ligeramente mejores resultados. Esta herramienta ayuda a explorar posibles dependencias sin implicar causalidad.

#### *Histograma de promedios*



### *Distribución de promedios*

El histograma muestra la **distribución** de los promedios finales. A diferencia del gráfico de barras que compara categorías, un histograma agrupa los datos continuos en intervalos o *bins* para visualizar su frecuencia. La mayoría de los promedios se concentran entre 78 y 88 puntos, con un rango relativamente estrecho. Esta visualización permite evaluar la dispersión de las calificaciones y detectar valores atípicos.

### **3. Conclusiones**

A través de la elaboración y análisis de múltiples gráficas se logró sintetizar el comportamiento de las calificaciones ficticias. La comparación de promedios mediante el **gráfico de barras** permitió identificar que el **Examen 2** fue el de mejor desempeño y el **Examen 1** el de menor rendimiento. El **gráfico de líneas** mostró una tendencia ascendente entre el primer y el segundo examen, lo que sugiere una adaptación al método de evaluación, y una leve caída posterior. El **diagrama de dispersión** reveló una correlación

inversa entre la edad y el rendimiento promedio, aunque no se trata de una relación causal. Finalmente, el **histograma** evidenció que la mayoría de los promedios se concentran en un rango alto, con poca dispersión. Estos hallazgos demuestran cómo la visualización de datos **facilita la extracción de conocimiento** al hacer visibles patrones y tendencias ocultas en tablas numéricas.

## Bibliografía

- Guía de visualización de datos para principiantes: definición, ejemplos y recursos de aprendizaje.* (s/f). Tableau. Recuperado el 29 de noviembre de 2025, de <https://www.tableau.com/es-mx/learn/articles/data-visualization>
- Hernández, J., Carmen, O. M., Lizandra, J., Minaya, N., Carlos, C., & Aranda, M. (s/f). *Extracción y Visualización de Conocimiento de Bases de Datos Médicas.* Upv.es. Recuperado el 29 de noviembre de 2025, de <http://josephorallo.webs.upv.es/escrits/Kdd-ACTA.pdf>
- Rezende, S. O., Oliveira, R. B. T., Felix, L. C. M., & Rocha, C. A. J. (s/f). *Visualization for Knowledge Discovery in database.* Witpress.com. Recuperado el 29 de noviembre de 2025, de <https://www.witpress.com/Secure/elibrary/papers/DATA98/DATA98006FU.pdf>
- (S/f). Gencat.cat. Recuperado el 29 de noviembre de 2025, de [https://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio\\_dades/guia\\_visualitzacio\\_es.pdf](https://atenciociudadana.gencat.cat/web/.content/manuals/visualitzacio_dades/guia_visualitzacio_es.pdf)