





Al final de la clase de hoy

- Estaremos en capacidad de identificar la variable objetivo y relacionarla en un análisis con la pregunta de negocio.
- 2. Identificar el nivel de relación entre dos variables numéricas.





La historia detrás de los datos

Existen relaciones entre las variables. Esas relaciones pueden ser:

- Causa efecto: Independiente ==> Dependiente
- Correlacionales: Se mueven juntas pero no se afectan.

En EDA somos detectives estudiando esas relaciones entre las variables.



Algunos ejemplos

- La publicidad de un producto afecta en el número de ventas del mismo producto.
- Las notas de los estudiantes dependen del estudio, constancia y disciplina de cada uno.
- El precio de un producto afecta la cantidad de demanda.

Variables dependientes e independientes

Relaciones que no son estructurales:

 Cuando Anne Hathaway se ganó el premio Oscar, subieron los precios de las acciones en Berkshire Hathaway.



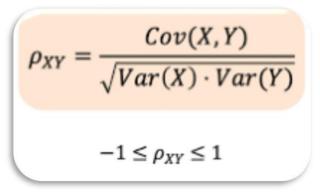
¿Qué es la correlación?

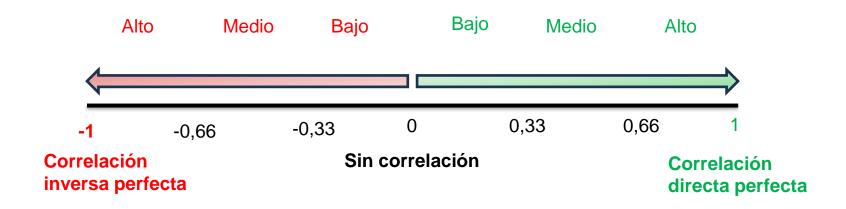
- La correlación es una medida para evaluar si dos variables están relacionadas linealmente.
- La medida cuantifica la intensidad de la relación.

Describimos la correlación mediante una medida llamada coeficiente de correlación, que puede tomar valores de -1 a 1.



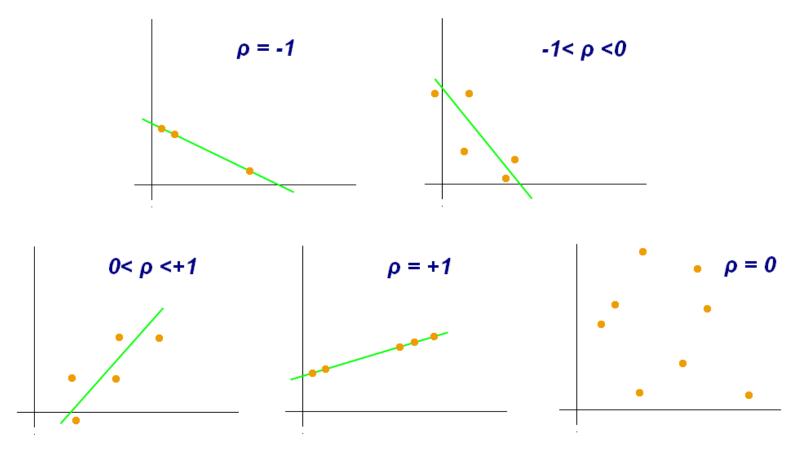
Coeficiente de Pearson





^{*}Aplica únicamente para variables numéricas.





Diagramas de dispersión con diferentes valores en el coeficiente de correlación



Con todo esto...

- 1. Entendemos el concepto de correlación.
- 2. Podemos medir la correlación entre dos variables.

¡Con estas lentes vamos al Colab a jugar a ser detectives!



¡Gracias!

Aprendiendo juntos a lo largo de la vida



educacioncontinua.uniandes.edu.co

Síguenos: EdcoUniandes f 🞯 庙 💟 🖸

