

TAREA 9. Usando los datos de este [paper](#), repetir el análisis de [Otro de Estad. \(jupyter nb\)](#) en la medida que se pueda analizar (expliquen porque no en su caso o utilicen alguno de los procedimientos de los otros tutoriales). Expliquen que variables eligen de analizar y por qué.

1. Formalismo

- **Ecuaciones:** Como el análisis explora la relación entre dos variables mediante una correlación, el formalismo se basa en la fórmula del coeficiente de correlación de Pearson:

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \cdot \sum(Y - \bar{Y})^2}}$$

donde r es el coeficiente de correlación, X y Y son las variables, y \bar{X} y \bar{Y} son las medias de cada variable.

- **Variables:** Se eligen Variable1 y Variable2. La elección de variables se basa en el interés de analizar si existe una relación lineal significativa entre ambas.
- **Limitaciones:** Este análisis asume que las variables tienen una distribución aproximadamente normal y que existe una relación lineal. Si los datos no cumplen estos supuestos, el coeficiente de Pearson puede no ser el más adecuado.

2. Algoritmos

- **Método:** El análisis utiliza el **coeficiente de correlación de Pearson** para medir la relación entre dos variables continuas. Este método es adecuado para relaciones lineales en datos normalmente distribuidos.
- **Modificaciones posibles:** Si los datos no son normales, un análisis alternativo podría ser el coeficiente de correlación de Spearman, que es adecuado para datos no paramétricos y relaciones no lineales.

4. Resultados

- **Descripción de Resultados:**
 - El código calcula la media y la desviación estándar para ambas variables, lo que permite entender su dispersión y valor central.
 - El gráfico de dispersión visualiza si existe una relación lineal entre Variable1 y Variable2.
 - La prueba de Pearson calcula el coeficiente de correlación y el valor p, que indica la fuerza y la significancia de la relación lineal.
- **Interpretación:**
 - Si el coeficiente de correlación es cercano a +1 o -1, esto indica una relación fuerte. Un valor cercano a 0 indica una relación débil.
 - El valor p permite determinar si la correlación es significativa (usualmente, un valor p < 0.05 es considerado significativo).

5. Análisis Crítico

- **Aprendizaje:** Este análisis ayuda a comprender la relación entre dos variables y a interpretar la fuerza y significancia de dicha relación mediante el coeficiente de Pearson.

- **Fiabilidad de los resultados:** La fiabilidad depende de los supuestos de normalidad y linealidad de los datos. Este análisis es confiable si estos supuestos se cumplen; de lo contrario, puede que no refleje adecuadamente la relación entre las variables.
- **Mejoras posibles:**
 - Verificar la normalidad de los datos (por ejemplo, con una prueba de Shapiro-Wilk) para asegurar la idoneidad del coeficiente de Pearson.
 - Si los datos no son normales, considerar el uso de una correlación de Spearman.
 - Explorar la relación usando modelos de regresión, que podrían proporcionar más detalles sobre la relación.