

# Trabajo de Datos

Informe de Análisis de Datos – Curso 2025/26

**Prof. Rubén Oncala**  
ruben.oncalamesa@ceu.es  
Web: [droncala.github.io/TemaX](https://droncala.github.io/TemaX)  
YouTube: @droncalaTemaX

---

Entrega antes del 6 de Noviembre del 2025 en un solo PDF por email o por la web de TemaX.

## Descripción General

---

El objetivo de este proyecto es que cada alumno realice un informe estadístico completo basado en una base de datos real.

- **Selección de datos:** Cada estudiante elegirá una categoría distinta y podrá seleccionar libremente la fuente de datos (ejemplos de variables: grado de miopía, edad, sexo, peso, etc.).
- **Referencias:** Es obligatorio incluir en el informe la referencia explícita a la fuente de los datos.
- **Ejemplo:** Se puede consultar un ejemplo de estructura de base de datos sobre miopía en el archivo: `rpartsey/MYDPIA.csv`.

## Formato de Entrega

---

El trabajo consta de dos partes indivisibles:

1. **Informe Escrito:** Debe ser un documento breve y conciso (**máximo 4 páginas**) que contenga los gráficos, las tablas estadísticas y la interpretación de los mismos.
2. **Defensa Oral:** Una exposición de aproximadamente **10 minutos** donde se presentará una interpretación resumida de los hallazgos más relevantes.

## Criterios de Calificación

---

La calificación final (sobre 10 puntos) evaluará conjuntamente la calidad técnica del informe escrito y la claridad expositiva en la defensa oral.

## Contenido en el Informe

---

El **Informe Escrito** deberá incluir de forma orientativa los siguientes análisis:

### Análisis Estadístico

- 1. Descripción de los datos:** Definición clara de la base de datos y descripción de las variables elegidas.
- 2. Análisis Gráfico:** Histograma general y segmentado por grupos de una de las variables principales.
- 3. Estadística Descriptiva:** Cálculo e interpretación de:
  - Medidas de tendencia central y dispersión: Media, Varianza, Cuartiles.
  - Coeficiente de Variación (CV).
  - Medidas de forma: Asimetría ( $g_1$ ) y Curtosis ( $g_2$ ), comparándolos con el histograma.
- 4. Comparativa de Grupos:** Diagrama de caja y bigotes (boxplot) por grupos, incluyendo el cálculo de los valores- $p$  correspondientes para verificar diferencias significativas.
- 5. Correlación:** Estudio de la relación entre dos variables cuantitativas mediante diagrama de dispersión y coeficientes de correlación.
- 6. Modelos de Ajuste (Regresión):**
  - Ecuación de la curva y parámetros obtenidos.
  - Bondad de ajuste: Coeficiente de determinación ( $R^2$ ).
- 7. Discusión:** Interpretación crítica de los resultados, limitaciones del modelo y obtención de predicciones coherentes basadas en la regresión.
- 8. Inferencia Estadística:** Cálculo de intervalos de confianza y comparación de medias.
- 9. Contraste de Hipótesis:** Planteamiento del contraste, obtención del valor- $p$  y conclusión estadística final.