

# Taller 1 de Business Analytics

Tema: Limpieza de datos en R  
Profesor: Eduard F. Martínez González

Agosto 2025

## Instrucciones generales

- Cada taller debe resolverse en R, utilizando preferiblemente un script o archivo RMark-down bien documentado.
- Incluya comentarios explicativos en su código.
- El puntaje total de cada taller es 100 puntos.
- Se evaluará la correcta interpretación de los resultados, no únicamente la ejecución del código.

## Taller 1: Descriptiva y limpieza de datos

1. (30 pts) Cargue la base de datos en R y use la librería `skimr` para realizar una descripción general. Reporte:
  - Número de variables y observaciones.
  - Tipos de variables presentes.
2. (30 pts) Seleccione una variable numérica (por ejemplo, ventas o número de empleados) e interprete sus principales estadísticas:
  - Media, mediana y percentiles (25 y 75).
  - ¿Qué nos dice la distribución de esta variable acerca de las empresas?
3. (15 pts) Normalice los nombres de las variables usando la librería `janitor::clean_names` para eliminar espacios y mayúsculas.
4. (25 pts) Filtre la base de datos para conservar únicamente las observaciones con ventas positivas. Indique cuántas observaciones quedan en el nuevo objeto.

## Taller 2: Descriptiva y manipulación con dplyr

1. (15 pts) Cargue la base de datos y normalice los nombres de las variables con `clean_names`.
2. (20 pts) Realice un análisis descriptivo con `skimr::skim` e interprete al menos una variable numérica en detalle (percentiles, media, distribución). Comente qué tan representativos son los valores centrales.
3. (15 pts) Filtre los datos para quedarse únicamente con las empresas que tienen ventas positivas y al menos 2 sedes. Indique cuántas observaciones cumplen estas condiciones.
4. (15 pts) Cree una nueva variable que calcule las ventas por empleado de cada empresa. Interprete el rango de valores que obtiene.
5. (20 pts) Usando `dplyr`, agrupe los datos por año y calcule:
  - Número promedio de accidentes por empresa.
  - Ventas totales y promedio por año.
6. (15 pts) Genere una tabla resumen con las tres empresas con mayor número de empleados en 2025.