

# Informe de Análisis Exploratorio de **mtcars**

Melani Forsythe Matos  
Daniela Guerrero Álvarez  
Rubén Martínez Rojas

El conjunto de datos **Titanic** es un conjunto clásico que contiene información sobre los pasajeros del Titanic, y se utiliza frecuentemente en análisis de supervivencia. A continuación, se describen las variables incluidas en este conjunto de datos, su significado, tipo de escala, y si son discretas o continuas:

#### 1. **PassengerId**

- **Descripción:** ID único para cada pasajero.
- **Escala:** Cuantitativa Discreta.
- **Significado:** Un identificador numérico para cada pasajero en la base de datos.

#### 2. **Class**

- **Descripción:** Clase del pasajero (1st = Primera, 2nd = Segunda, 3rd = Tercera, Crew = Trabajadores).
- **Escala:** Cualitativa Ordinal.
- **Significado:** Representa la clase socioeconómica del pasajero, con 1 siendo la clase más alta y 3 la más baja, y los trabajadores.

#### 3. **Age**

- **Descripción:** Si el pasajero es Adulto o Niño.
- **Escala:** La edad normalmente es una variable cuantitativa, pero en este caso está dividido por grupos etáreos, Adulto y Niño, por tanto se clasifica como Cualitativa Nominal.
- **Significado:**

#### 4. **Survived**

- **Descripción:** Indicador de supervivencia (No, Sí).
- **Escala:** La variable es Cualitativa Nominal, pero podría convertirse a Cuantitativa Discreta (Binaria), asignándole 1 a Sí y 0 a no, o viceversa.
- **Significado:** Indica si el pasajero sobrevivió al hundimiento del Titanic.

#### 5. **Sex**

- **Descripción:** Género del pasajero.
- **Escala:** Cualitativa Nominal.
- **Significado:** Indica el sexo del pasajero (hombre o mujer).

#### 6. **Freq**

- **Descripción:** Número de pasajeros.
- **Escala:** Cuantitativa Discreta.
- **Significado:** Indica la la **frecuencia** o el **número de pasajeros** que caen en cada combinación de las otras variables (**Class**, **Sex**, **Age**, **Survived**).

**Ejemplo:**

- **Class:** La clase del pasajero (**1st**, **2nd**, **3rd**, **Crew**).
- **Sex:** El género del pasajero (**Male**, **Female**).
- **Age:** El grupo de edad del pasajero (**Child**, **Adult**).
- **Survived:** Si el pasajero sobrevivió o no (**No**, **Yes**).
- **Freq:** El número de pasajeros que pertenecen a esa combinación de **Class**, **Sex**, **Age**, y **Survived**.

Supongamos que tienes la siguiente fila en el dataset:

Class	Sex	Age	Survived	Freq
1st	Male	Adult	Yes	57

Esto significa que **57 pasajeros adultos de sexo masculino en primera clase sobrevivieron**.

La columna **Freq** te permite ver cuántas personas están representadas por cada combinación de las otras variables, en lugar de tener una fila separada para cada pasajero individual.