

# Análisis Exploratorio del Titanic

Nazario Bustos Eusebio

## Introducción

Este documento presenta un análisis exploratorio del conjunto de datos del Titanic. Utilizamos herramientas del paquete Tidyverse para manipular y visualizar los datos con el fin de extraer información relevante sobre los pasajeros y sus características relacionadas con la supervivencia.

## Resumen

Después de analizar los datos del Titanic, descubrimos que los pasajeros de clase alta (Upper Class) tenían una mayor tasa de supervivencia en comparación con aquellos de clase media y baja, especialmente entre las mujeres. El género jugó un papel importante, ya que una mayor proporción de mujeres sobrevivió en todas las clases. La edad también fue un factor relevante: los sobrevivientes tienden a ser personas más jóvenes, aunque la distribución varía entre las diferentes clases sociales.

Además, los pasajeros con boletos más costosos (Fare más altos) tenían más probabilidades de sobrevivir, lo cual podría estar relacionado con la clase socioeconómica. El puerto de embarque, sin embargo, no mostró un patrón claro respecto a la supervivencia.

## Carga de datos y preparación

```
```{r setup, include=FALSE} # Cargamos las librerías necesarias library(tidyverse)
```

## Importamos los datos

```
titanic <- read.csv("Titanicv2.csv")
```

## Selección y limpieza de datos

```
titanic_clean <- titanic %>% select(Survived, Pclass, Sex, Age, Fare, Embarked) %>%  
mutate(Survived = factor(Survived, levels = c(0, 1), labels = c("No", "Yes")), Pclass =  
factor(Pclass, levels = c(1, 2, 3), labels = c("Upper Class", "Middle Class", "Lower  
Class")), Sex = as.factor(Sex))
```