

EXECUTIVE SUMMARY

ANÁLISIS DEL DATASET TITANIC

INTRODUCCIÓN

En esta actividad trabajé con el conjunto de datos del Titanic con el objetivo de extraer y comunicar de manera clara los factores más relevantes que influyeron en la supervivencia de los pasajeros. Para ello, utilicé técnicas de visualización de datos que me permitieron construir una narrativa visual clara y comprensible.

METODOLOGÍA

Utilicé el conjunto de datos `titanic_train` del paquete `titanic` en R. Traté los datos con herramientas del paquete `tidyverse` y generé las visualizaciones con `ggplot2`. Antes del análisis, realicé una limpieza básica del dataset, eliminando las filas con valores perdidos en variables clave como la edad (`Age`) y el puerto de embarque (`Embarked`).

RESULTADOS Y VISUALIZACIONES

Gráfico 1: Supervivencia general

Para comenzar, analicé la proporción general de pasajeros que sobrevivieron frente a los que no.

Conclusión: Aproximadamente el 38% de los pasajeros lograron sobrevivir al hundimiento.

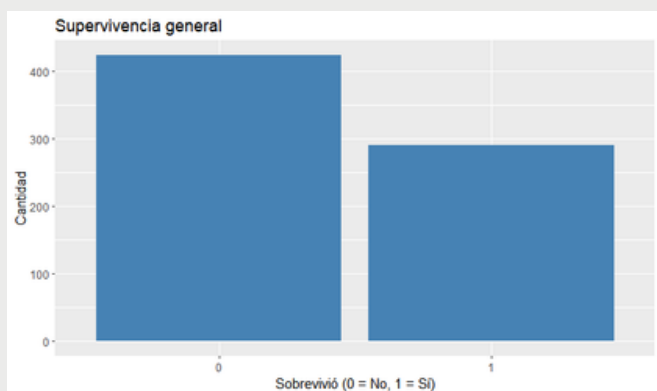


Gráfico 2: Supervivencia por sexo

Exploré la relación entre el sexo y la probabilidad de supervivencia.

Conclusión: Las mujeres sobrevivieron en una proporción mucho mayor que los hombres. Más del 70% de las mujeres sobrevivieron, frente a menos del 20% de los hombres.



Gráfico 3: Supervivencia por clase y sexo

Decidí cruzar la clase social con el sexo para entender mejor las diferencias de supervivencia.

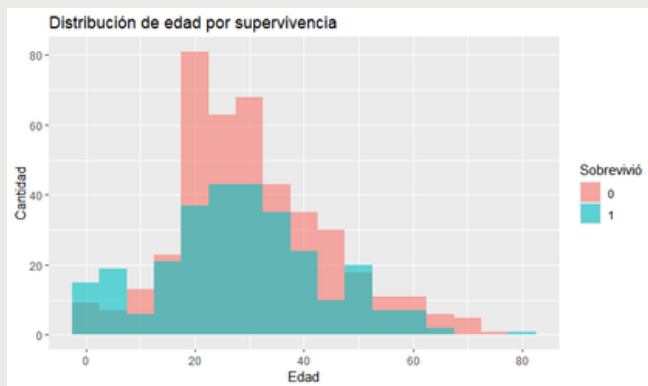
Conclusión: Las mujeres de primera clase presentaron la mayor tasa de supervivencia. Por el contrario, los hombres de tercera clase fueron el grupo más vulnerable.



Gráfico 4: Distribución de edad por supervivencia

También analicé la edad como factor en la supervivencia.

Conclusión: Observé que los niños tenían una mayor tasa de supervivencia, especialmente en comparación con adultos jóvenes y mayores en clases más bajas.



CONCLUSIONES

A partir del análisis pude concluir que los factores más relevantes en la probabilidad de supervivencia fueron la clase social, el sexo y la edad. Las visualizaciones me permitieron comunicar estos patrones de forma efectiva, facilitando la interpretación de los datos. Considero que este tipo de herramientas visuales son fundamentales para tomar decisiones informadas en contextos reales.

REFLEXIÓN FINAL

Realizar este análisis me permitió aplicar conocimientos sobre limpieza de datos, uso de ggplot2 y narrativa visual. La brevedad del informe supuso un reto adicional: priorizar la información más relevante me ayudó a comprender en profundidad los datos. En futuras actividades, me gustaría complementar este análisis explorando modelos predictivos o técnicas más avanzadas de visualización.

ANEXO: VISUALIZACIÓN INTERACTIVA CON SHINY

Objetivo del Anexo

Con el fin de complementar el análisis anterior, desarrollé una versión interactiva de las visualizaciones utilizando el paquete shiny. Esta versión permite al usuario explorar dinámicamente los datos y observar las mismas relaciones analizadas previamente, pero con mayor flexibilidad.

Comentario Personal

Al desarrollar esta versión interactiva con Shiny, pude observar cómo cambia la experiencia del usuario al explorar los datos. Esta herramienta aporta valor cuando se busca que el usuario final participe activamente en el análisis, mientras que los gráficos estáticos siguen siendo la mejor opción para informes impresos o PDF formales.