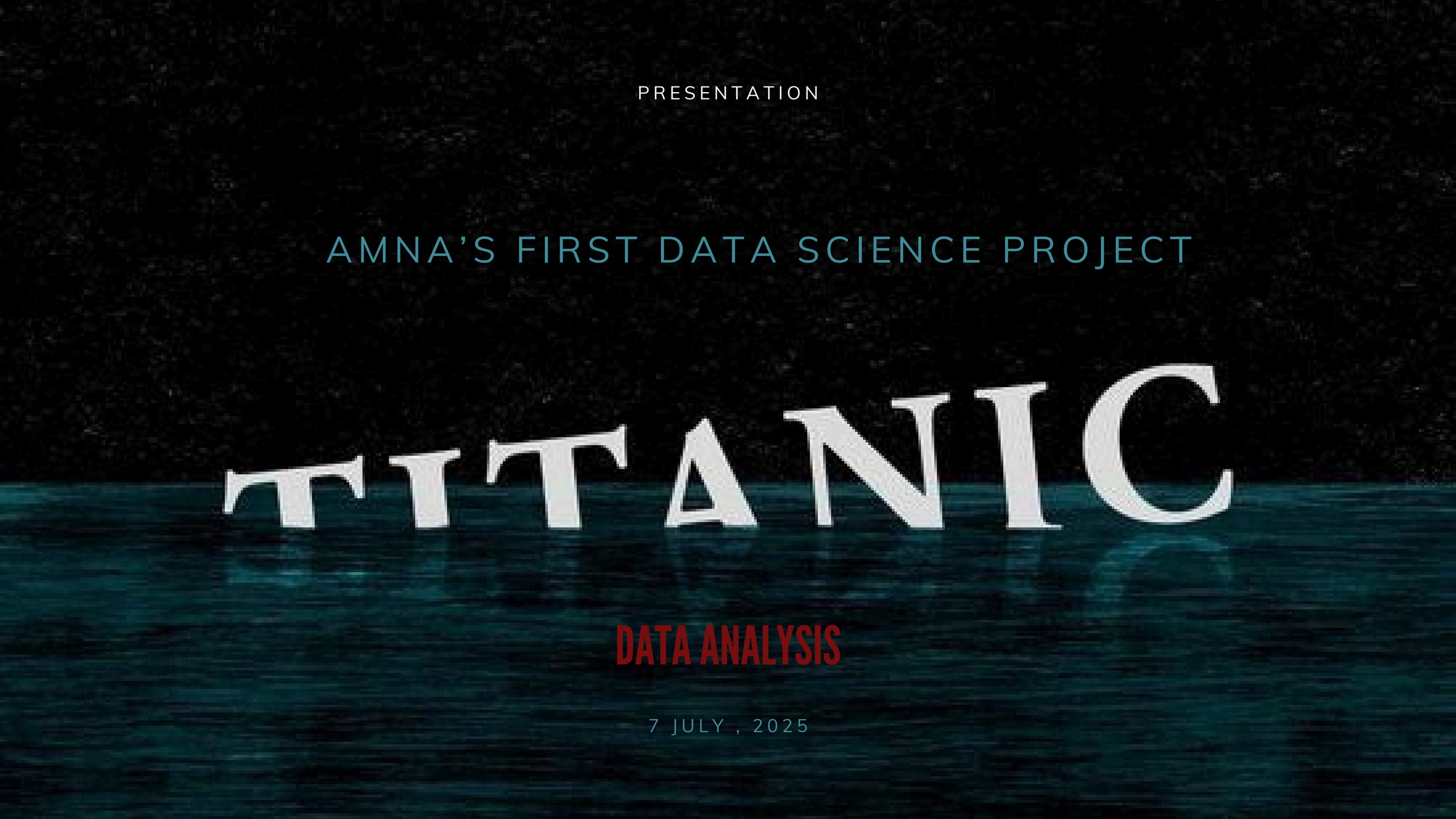
PRESENTATION

AMNA'S FIRST DATA SCIENCE PROJECT



DATA ANALYSIS

7 JULY , 2025

TITANIC DATASET

- INTRODUCTION :

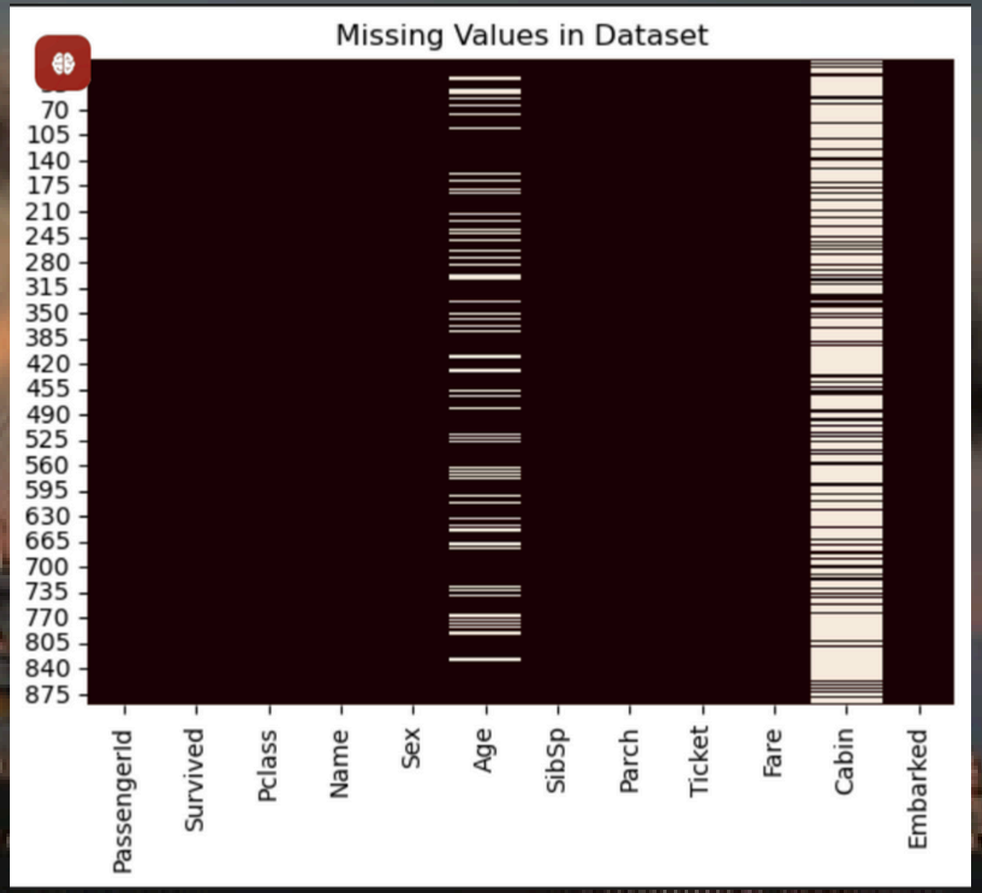
This project analyzes the famous Titanic dataset which contains detailed information about passengers, such as age, gender, class, and survival status.
The goal is to gain insights and practice data analysis skills using real-world data.

- OBJECTIVE :

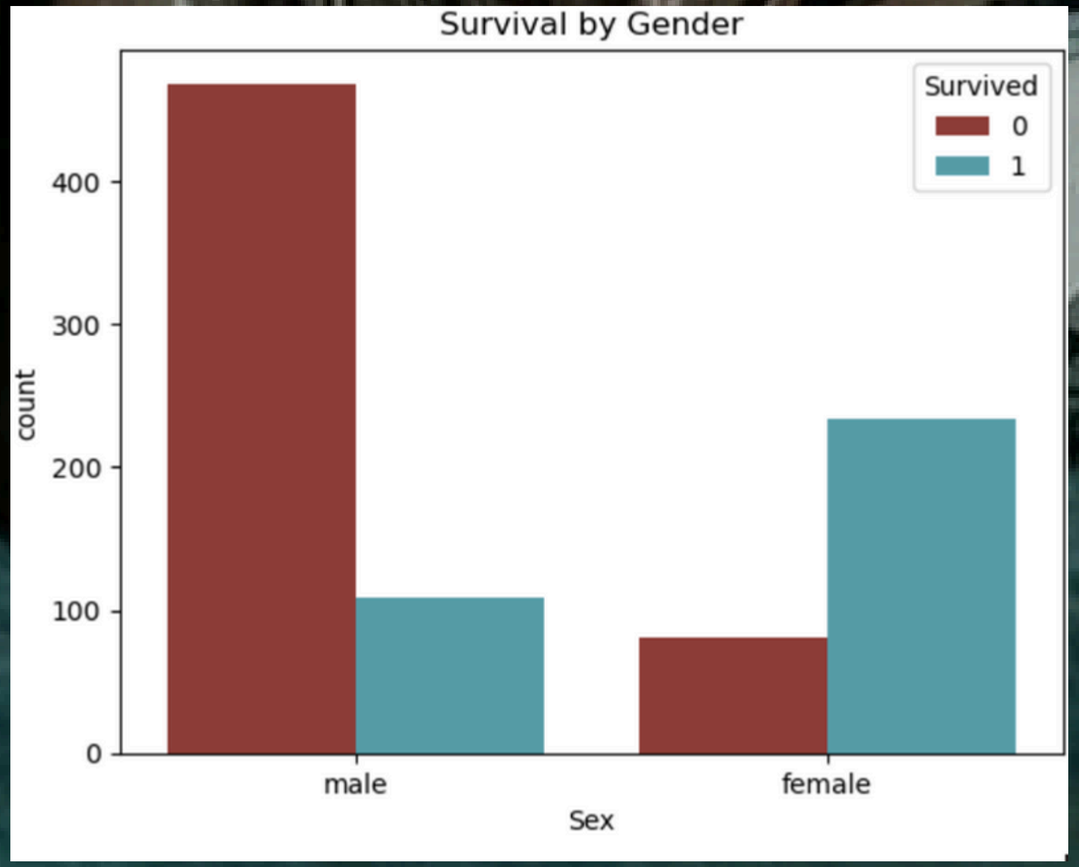
The objective is to explore which factors affected passengers' chances of survival, such as age, gender, and class.
This helps understand patterns in the data and practice data cleaning, filtering, and visualization.

DATA CLEANING STEPS

- Dropped irrelevant columns: PassengerId, Name, Ticket, Cabin
- Filled missing values in Age with mean
- Filled missing Embarked values with most frequent value
- Renamed Sex column to Gender for clarity
- These steps prepared the dataset for proper analysis and visualization.



EXPLORATORY DATA ANALYSIS (EDA)



An exploratory data analysis was conducted to better understand the characteristics of Titanic passengers. The most important findings are:

- The majority of passengers were male, and most of them did not survive.
- Passengers in the first class had a significantly higher survival rate.
- Women and children had a higher chance of survival.
- The most common embarkation point was Southampton.

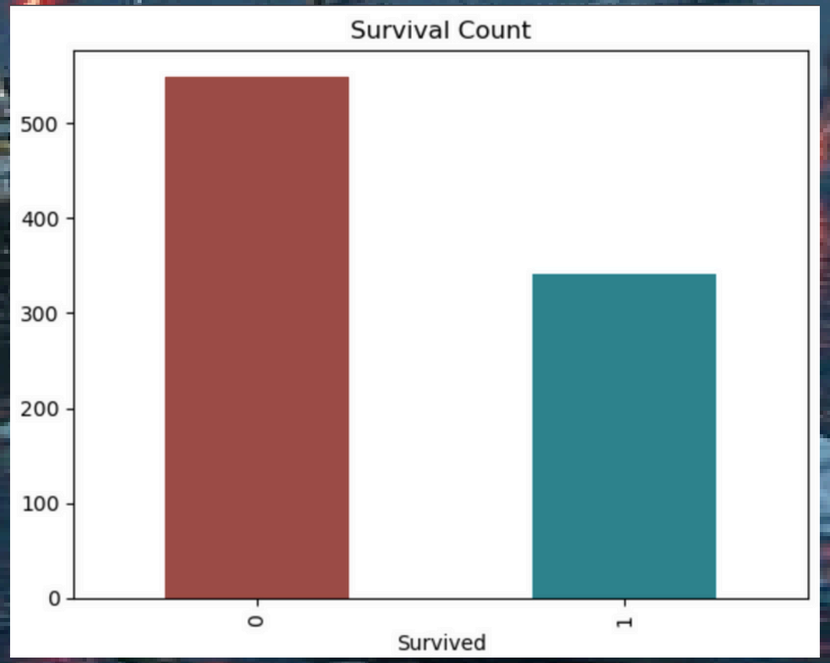
The following chart shows the distribution of passengers by gender and survival status.

INSIGHTS & CONCLUSIONS

Based on the analysis of the Titanic dataset, we discovered important patterns and insights:

- Gender was a strong factor in survival – females had a higher survival rate.
- First-class passengers had better chances of survival compared to other classes.
- Younger passengers were slightly more likely to survive.
- Most passengers embarked from Southampton, and their survival rates varied.

These findings help us understand how different factors influenced survival during the Titanic disaster.



RECOMMENDATIONS

Based on the insights derived from the Titanic dataset, we recommend the following:

- In future safety designs, ensure women and children have prioritized access to lifeboats.
- Improve the distribution of survival resources across all passenger classes.
- Enhance data collection during travel to allow better analysis and emergency response.

These suggestions could improve survival strategies in similar disaster scenarios.

REFERENCES

- Titanic Dataset from Kaggle
- Python & Pandas official documentation
- Matplotlib & Seaborn documentation
- Personal analysis and interpretation

CONCLUSION

This was my very first Data Science project.
Through analyzing the Titanic dataset, I explored how survival was affected by gender, age, and class.
The findings show how data can uncover hidden patterns and support fair, thoughtful decisions.
This beginner-friendly project helped me apply Python and data analysis tools like Pandas, Matplotlib, and Seaborn.
It marks the beginning of my journey toward becoming a professional Data Scientist.

LET'S CONNECT!

Thanks for reading my first project
I'm just getting started, and I'm excited to learn and grow in the world of Data Science.

Let's connect on GitHub!
📎 https://github.com/Am9AI

WE WANT TO SAY

THANK YOU

FOR YOUR ATTENTION