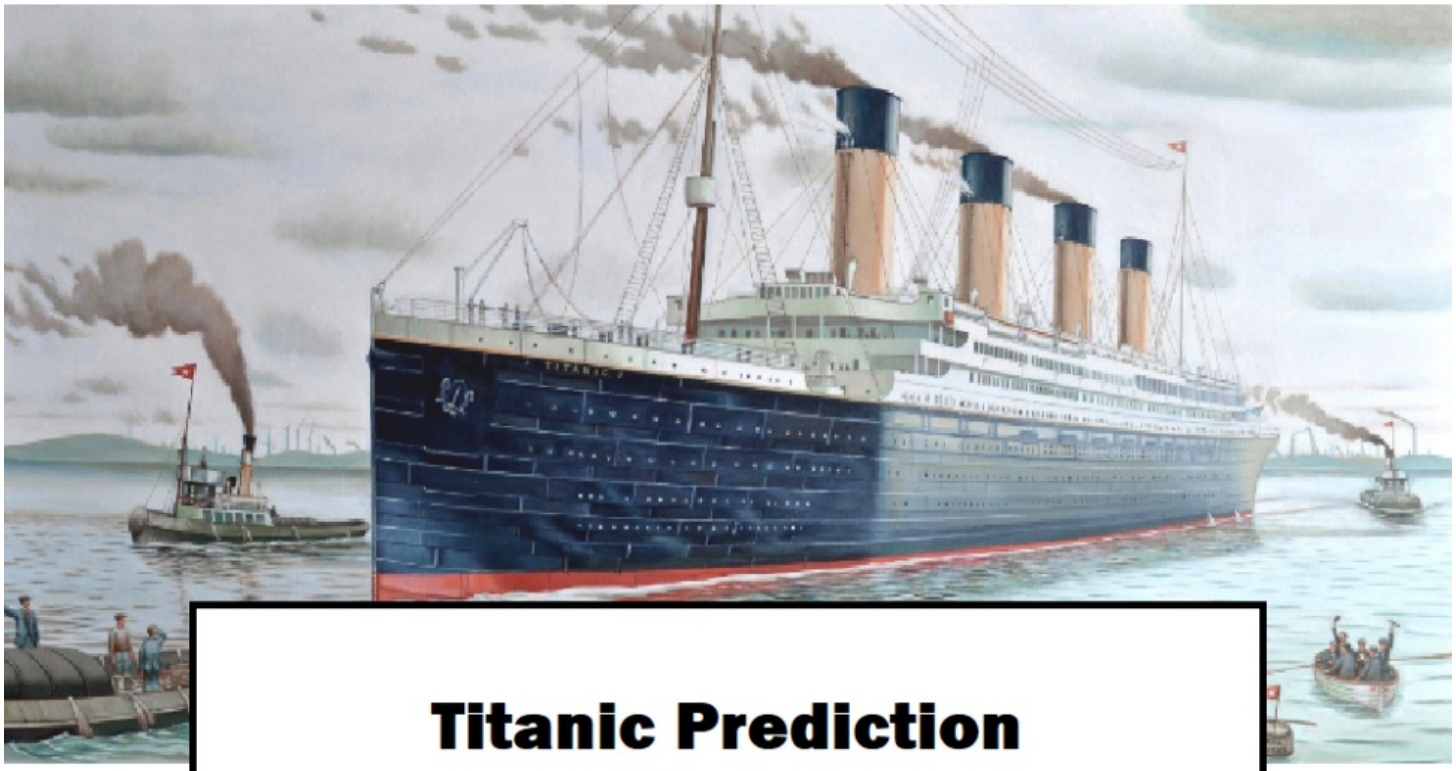
Presentation

Overview	Project Goal & Flow	Heatmap Correlations	Relative Correlation	Classification Models



Titanic Prediction Model

Presentation

Overview	Project Goal & Flow	Heatmap Correlations	Relative Correlation	Classification Models

Project Goal:

Find who is more likely to survive.

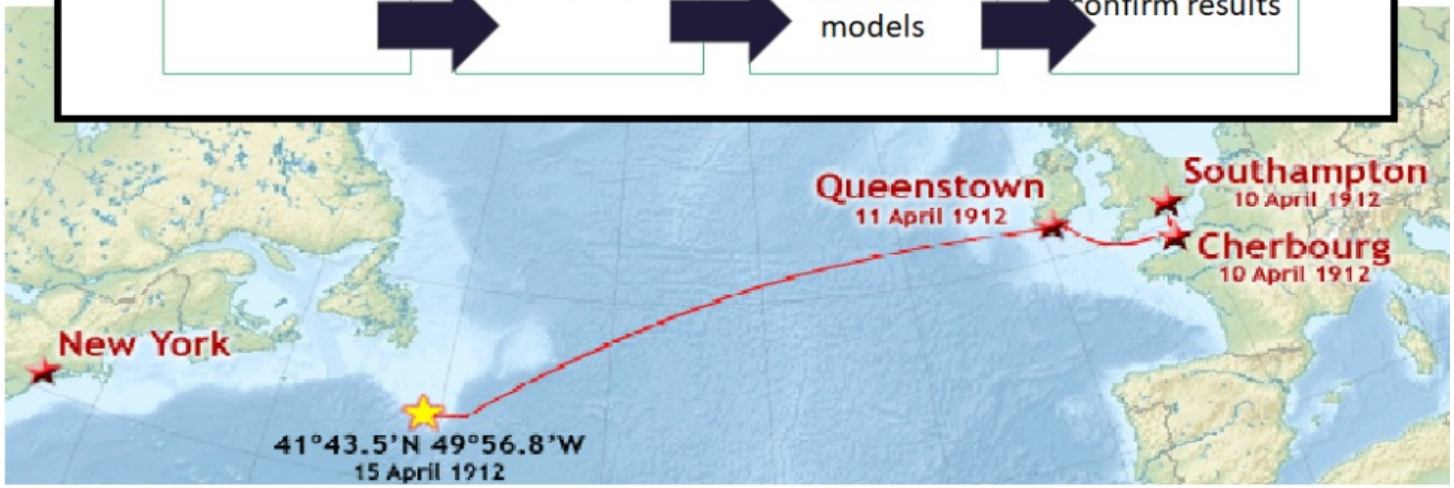
Project Flow

Collect data: Download Titanic Dataset from Kaggle

Clean and review data: Clean and Extract Relevant Data

Create an EDA and model: Created a Heat Map to find correlations and build models

Compare & evaluate results: Compare models to confirm results

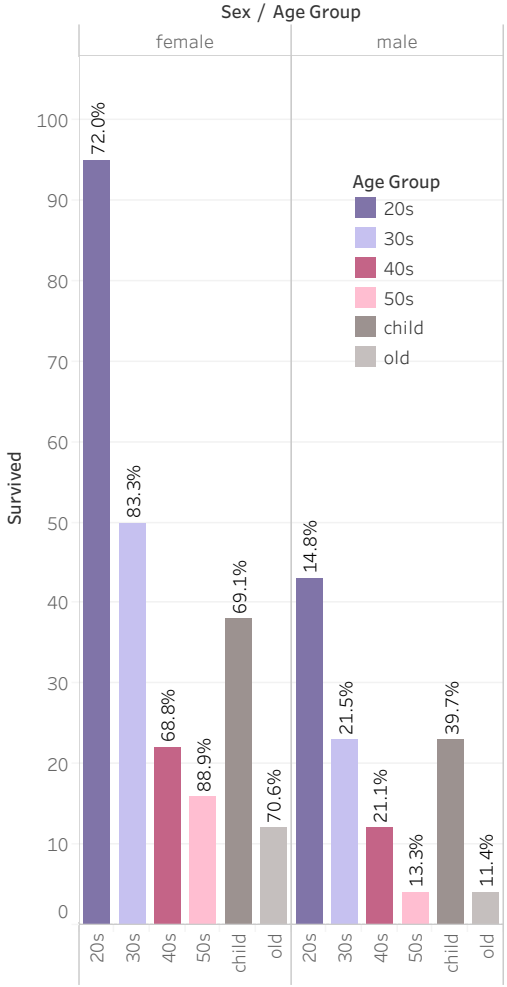


Presentation

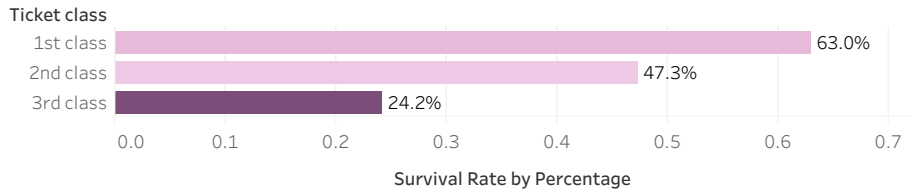
Heatmap Correlations

There is a correlation in survival based on both Sex and Ticket Class. Women had a much higher survival rate than men, which would make sense, as typically women have priority over men. Ticket class also has a high correlation to survival. 1st class tickets have a significantly higher rate of survival than 3rd class.

Survival Rate by Male and Female



Survival Rate by Ticket Class

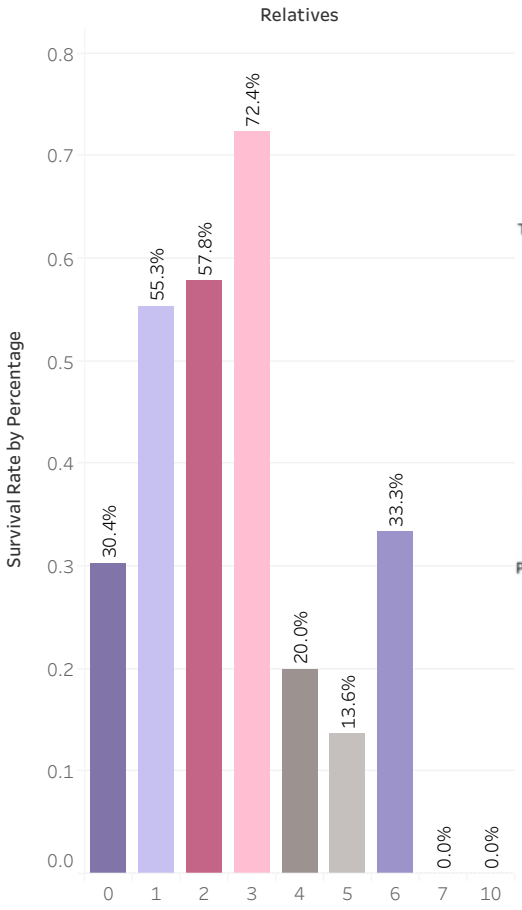


Presentation

Relative Correlation

In addition to the correlation between Sex and Ticket Class, there was a relative correlation between Sibling-Spouse and Parent-Child. I investgated further and noticed that passengers who had 3 or fewer relatives had a better chance of survival. Also, if you had more than 7 relatives onboard, there was no chance of survival.

Survival Rate by Number of Relatives



Presentation

Classification Models

For this dataset, I tried to use a **Linear Model**, however, the outputs were skewed. I realized classification models would work better for this dataset.

I also used a **Logistical Regression Model** to confirm accuracy. I used Survived as my dependent variable and Ticket class, Sex, Age, Relatives, and Embarked as my independent variables. Based on the data, this model provided a 79% accuracy score, which is good, however, does leave room for improvement.

However, I also used the **Random Forest model** on the Titanic dataset to predict survival based on three key factors: ticket class, sex, and number of relatives. First-class passengers are more likely to survive due to their wealth and access to lifeboats. Female passengers are more likely to survive due to the priority given to women and children. Passengers with fewer relatives are more likely to survive due to their care from other passengers. Age and embarked are also correlated with survival, but to a lesser extent.