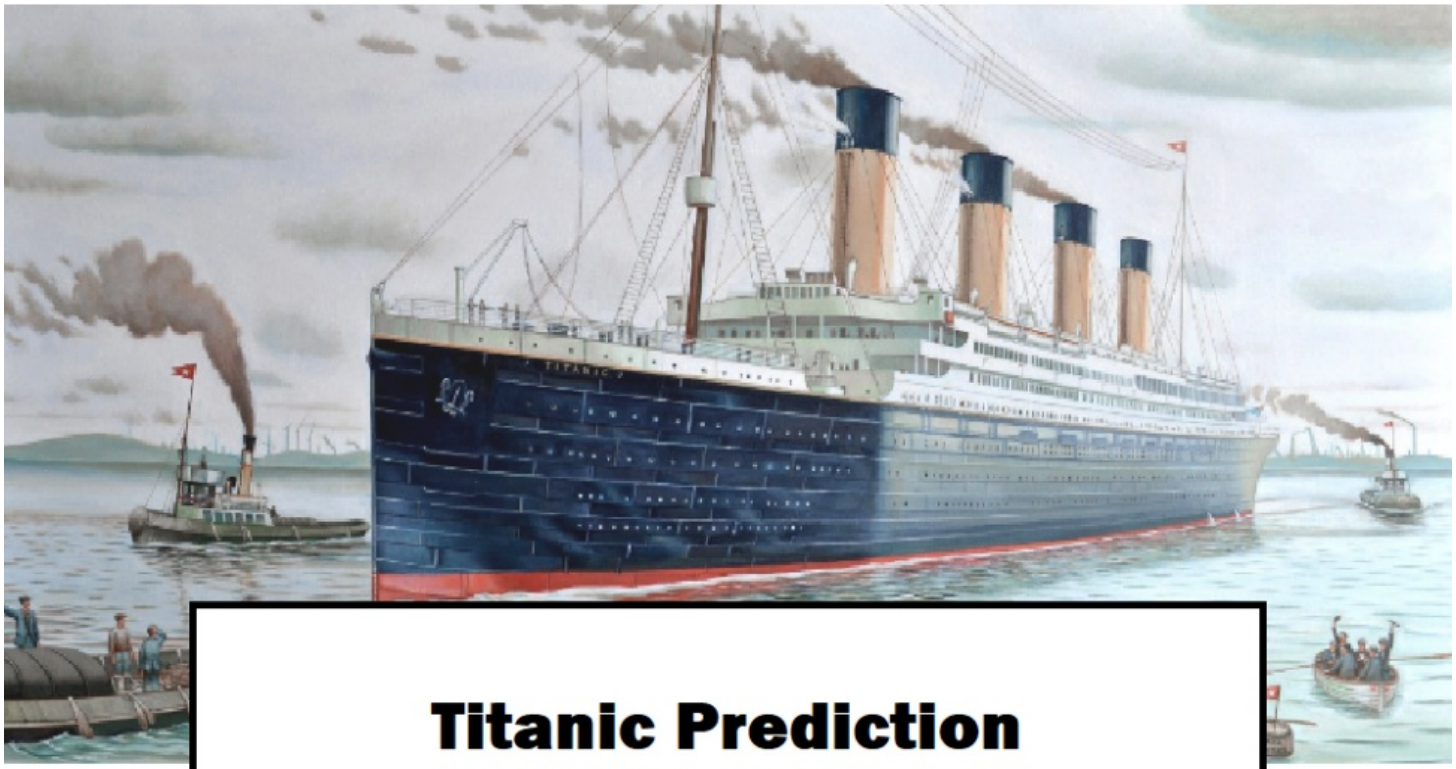
Presentation

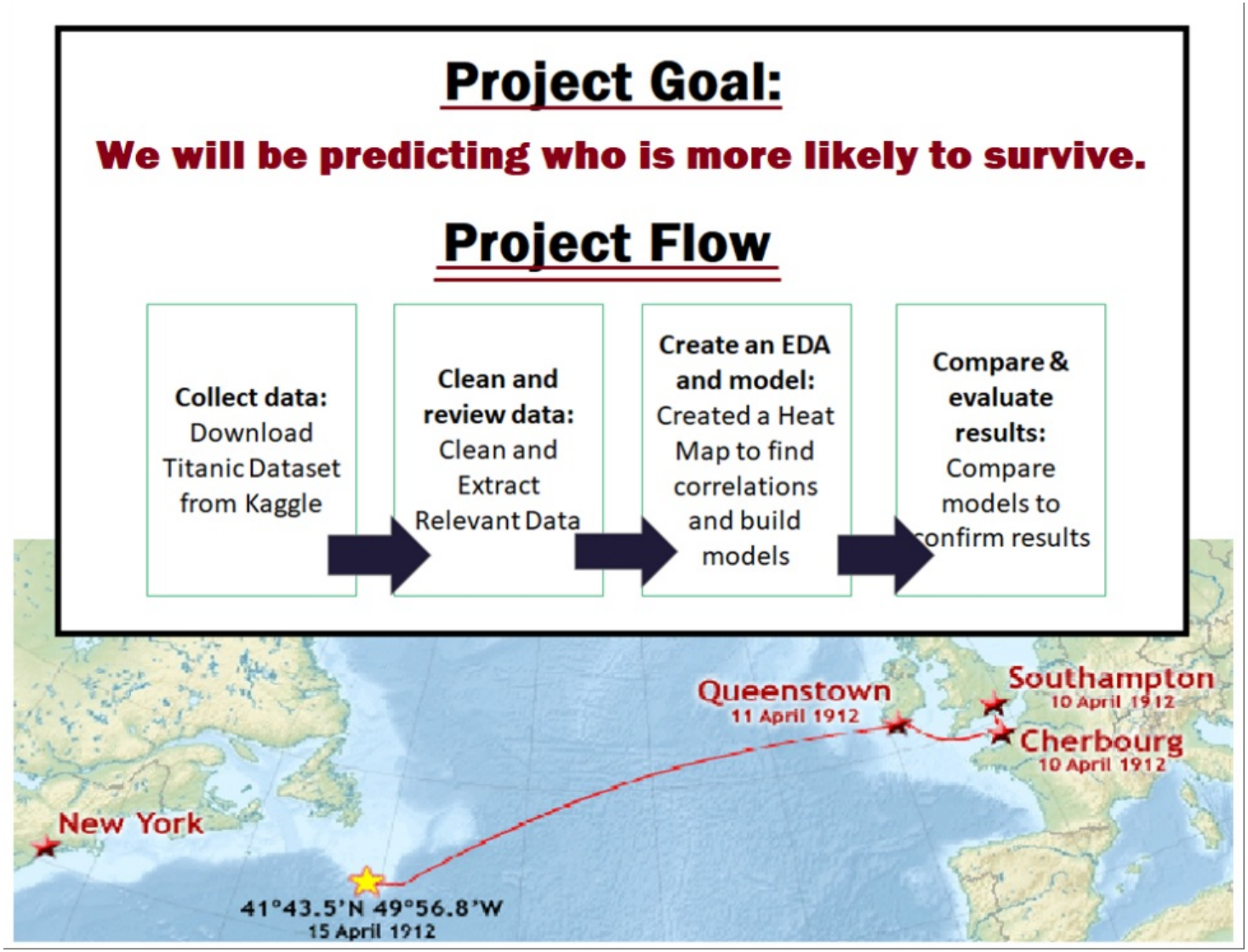
Overview	Project Goal & Flow	Heatmap Correlations	Relative Correlation	Classification Models



Titanic Prediction Model

Presentation

Overview	Project Goal & Flow	Heatmap Correlations	Relative Correlation	Classification Models

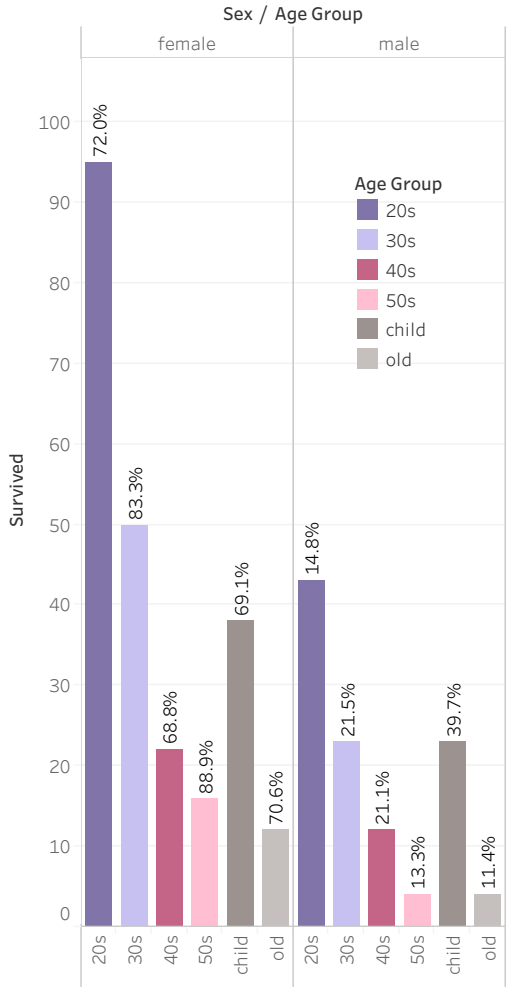


Presentation

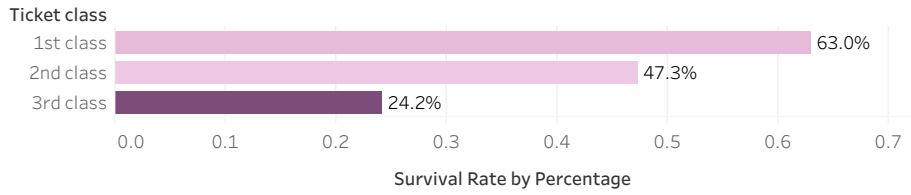
Heatmap Correlations

Based on this heatmap we can see there is a high correlation in both the Ticket class and Sex features.

Survival Rate by Male and Female



Survival Rate by Ticket Class

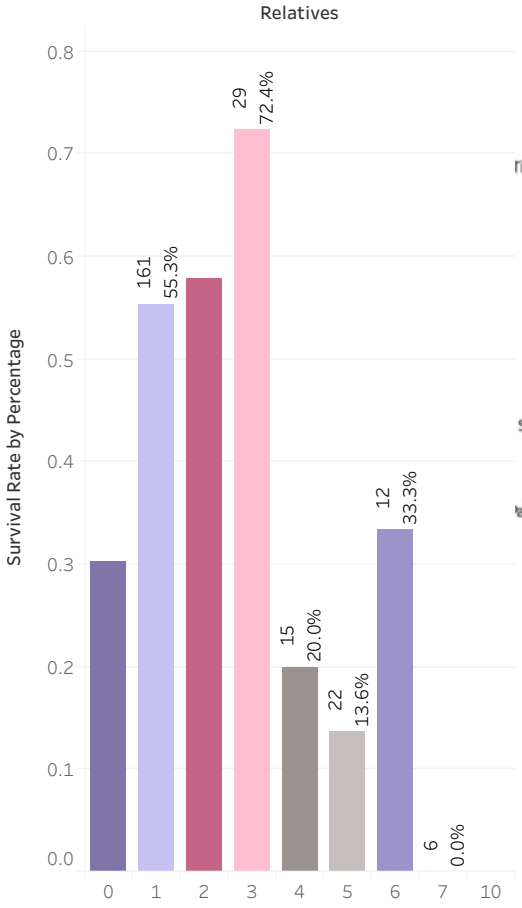


Presentation

Relative Correlation

In addition to the correlation between Sex and Ticket Class, there was a relative correlation between Sibling-Spouse and Parent-Child.

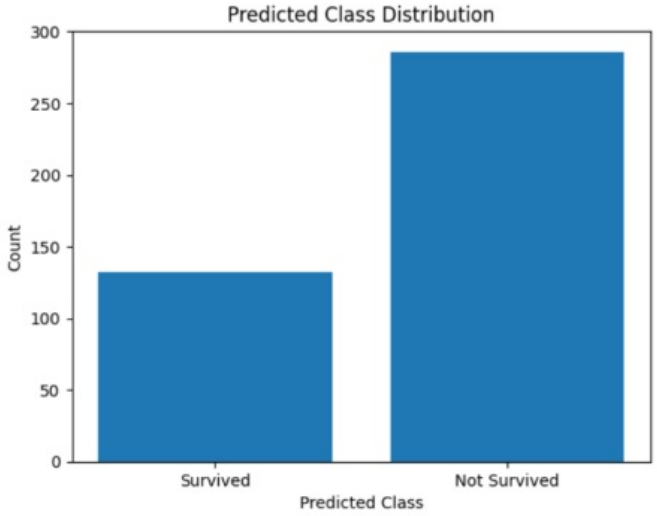
Survival Rate by Number of Relatives

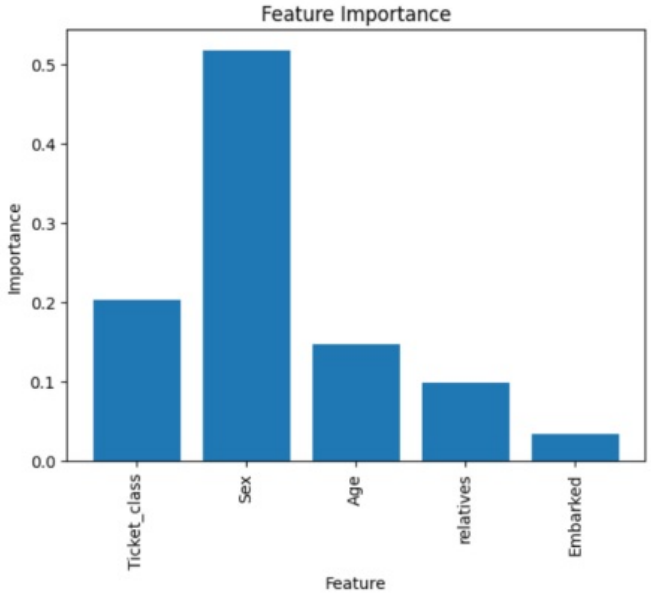


Presentation

Overview	Project Goal & Flow	Heatmap Correlations	Relative Correlation	Classification Models

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.

I also used the **Random Forest model** on the Titanic dataset to predict survival based on the same variables. The model has noted Ticket class and Sex are Feature Importance, which is similar to the heatmap correlation.

Therefore, based on our data, a 1st class woman travelling with 3 or fewer relatives would be most likely to survive.