

Task 5

Exploratory Data Analysis

Used Colab for EDA



Presented by
Karthik Doguparthi

- **Insights:**
- Most passengers preferred Class 3
- Male and Female passengers choose Class 3

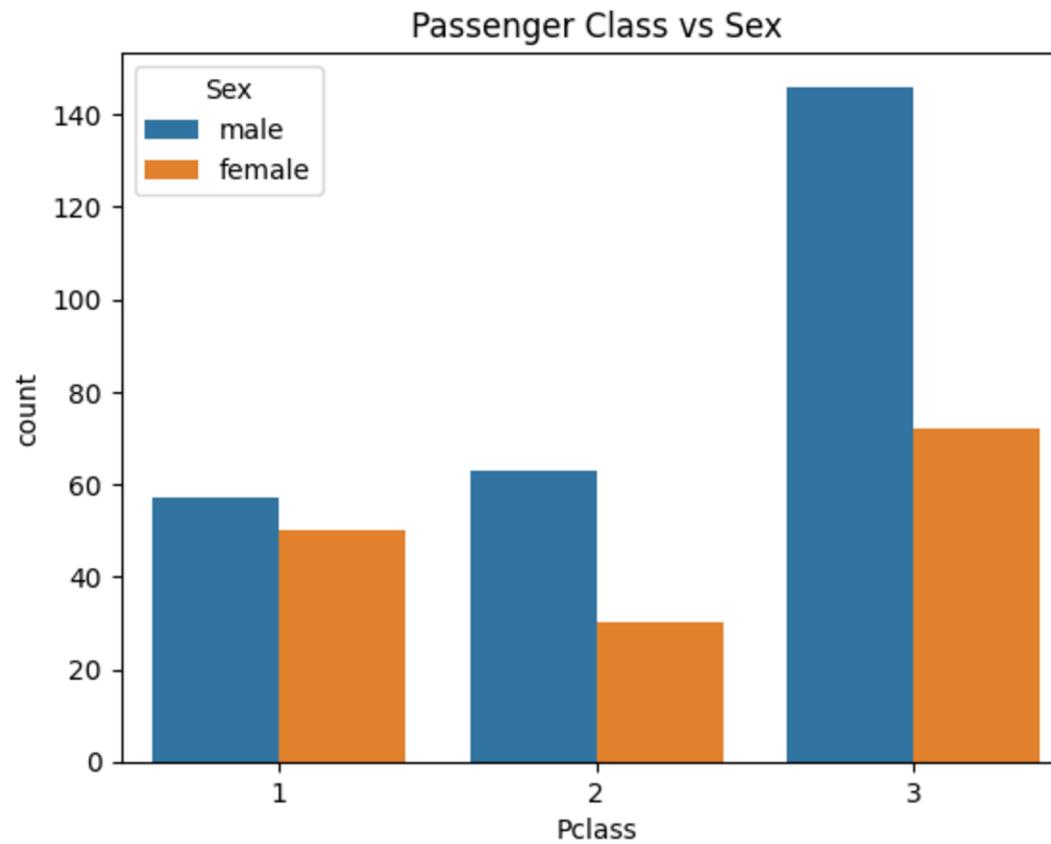


Chart 1

- **Insights:**
- Most people embarked from Southampton
- Queenstown has least embarkments
- Most males and females from Southampton, Least males and females form Queenstown.

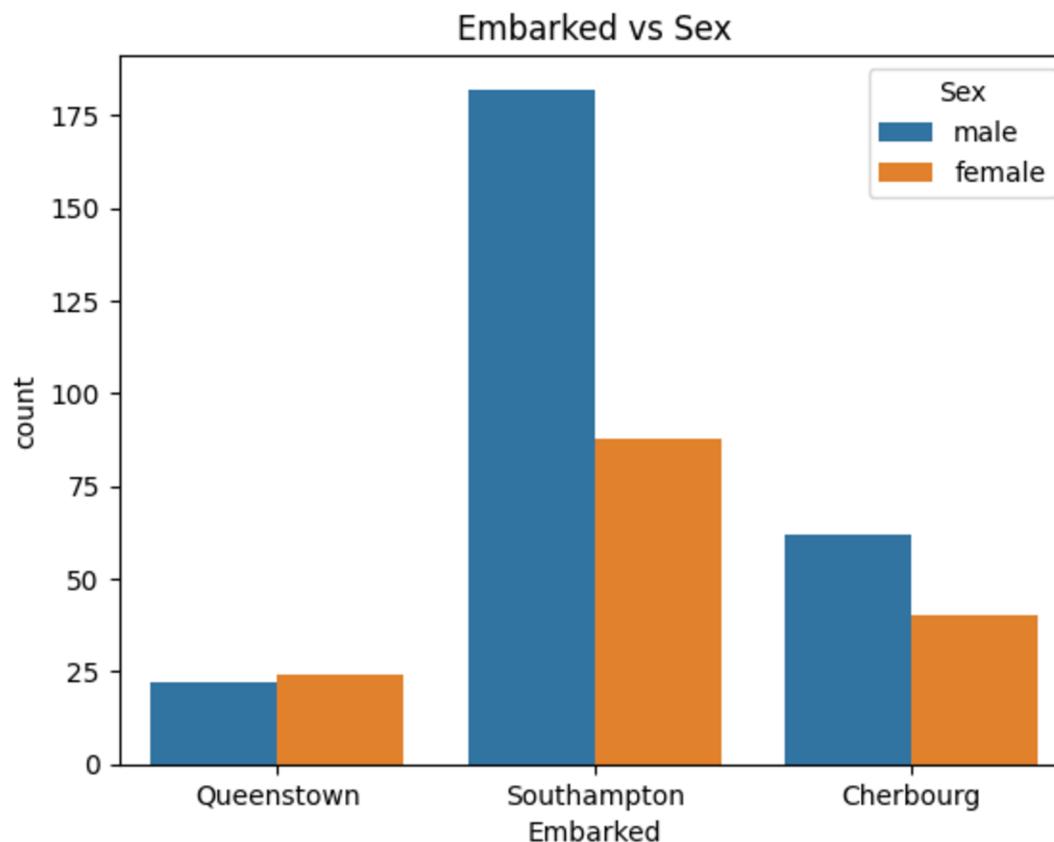


Chart 2

- **Insight:**
- Most male and female prefer Class 3 and from Southampton
- Least male prefer Class 1 from Queenstown.
- Least females prefer Class 2 from Queenstown.

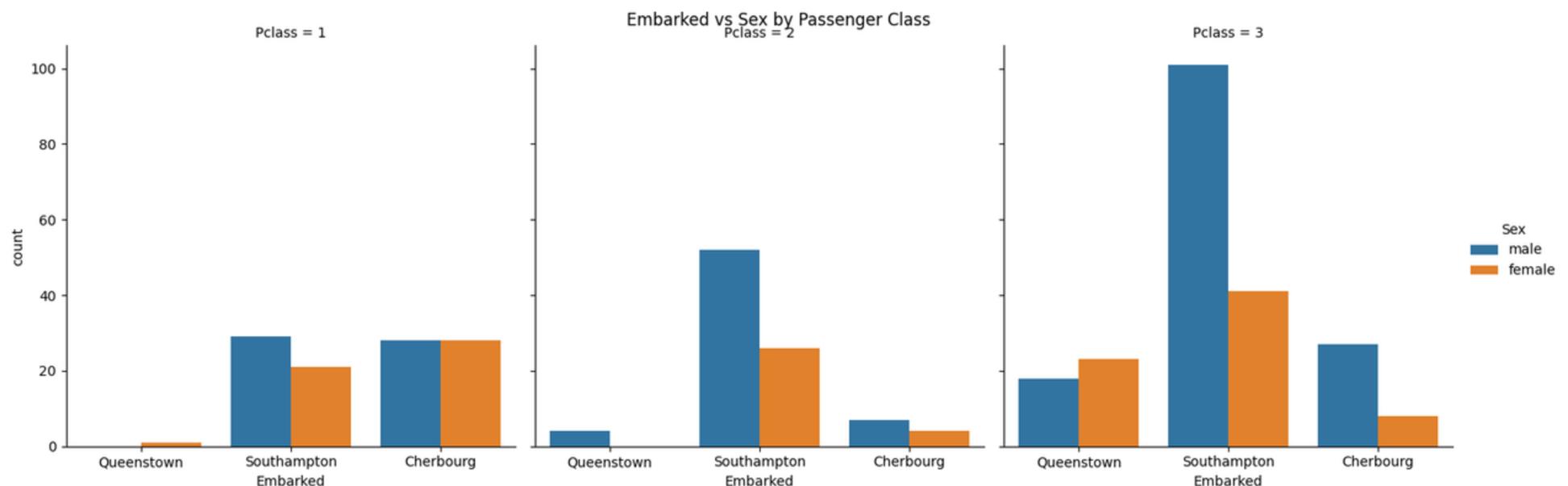


Chart 3

- **Insight:**
- Most male and female prefer Class 3 and from Southampton
- Least male prefer Class 1 from Queenstown.
- Least females prefer Class 2 from Queenstown.

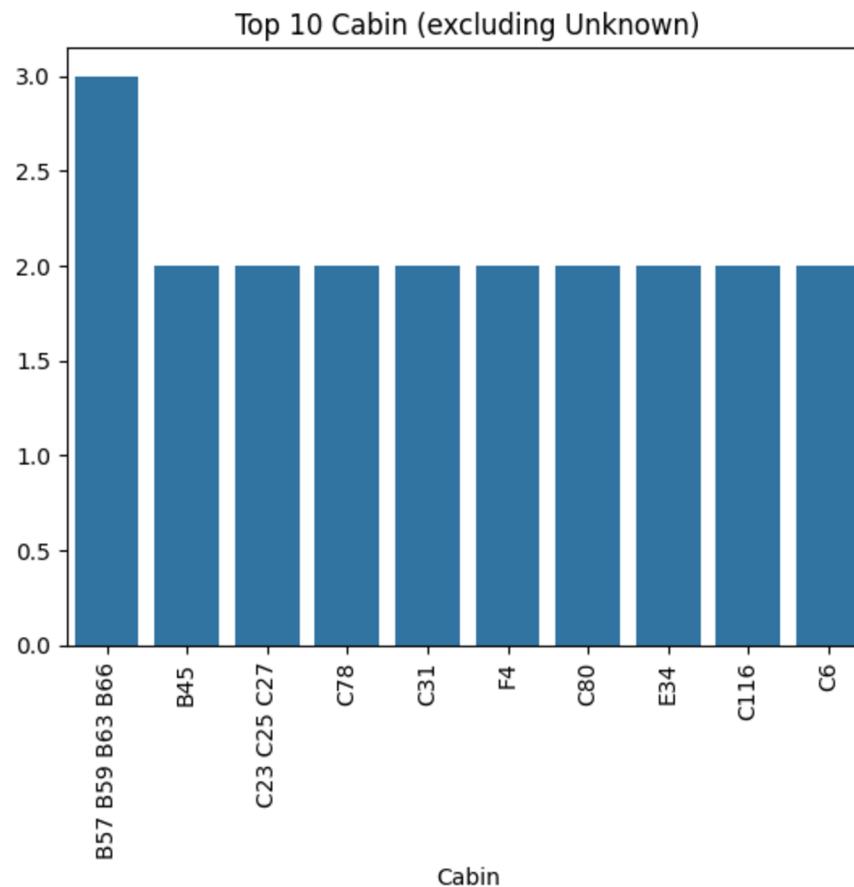


Chart 4

- **Insights:**
- Most Survived are Males and count over 250
- Least survived are Females and count over 150

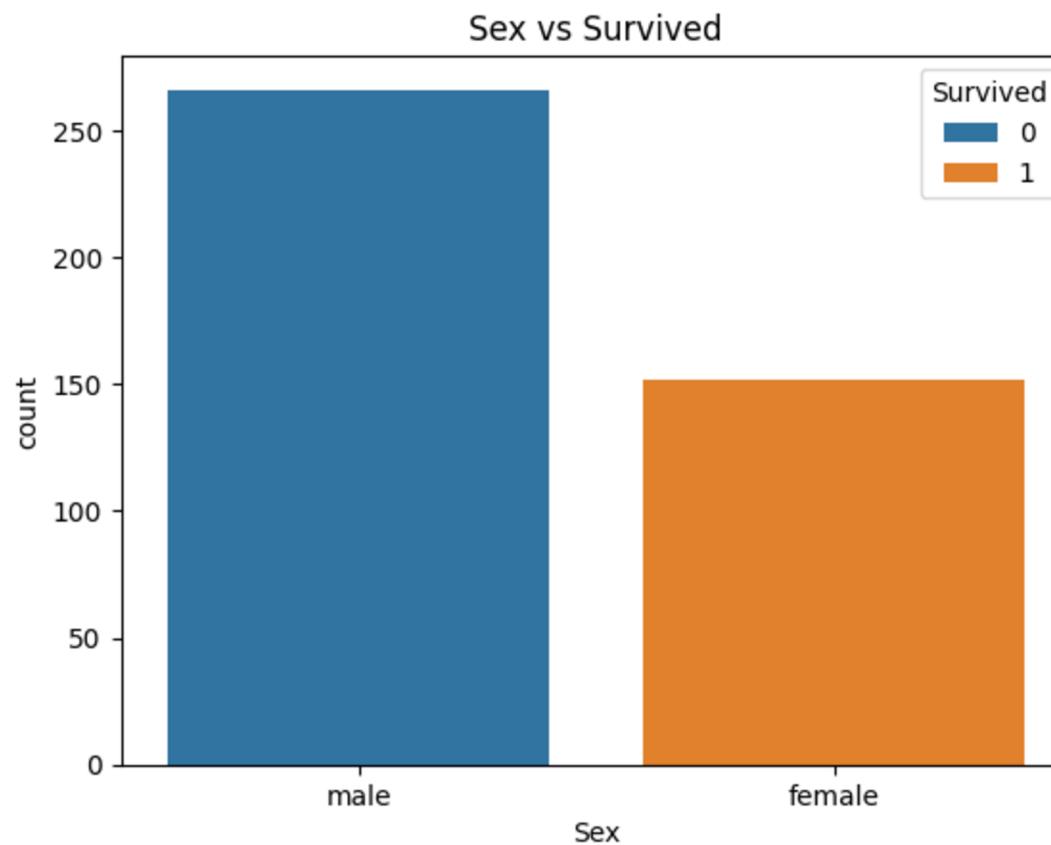


Chart 5

- **Insights:**
- Most sold price of the Titatic is in the range -0.512 to 51.233
- Most tickets sold are from Male and over 240
- Most tickets sold from Women under the same category is over 100.

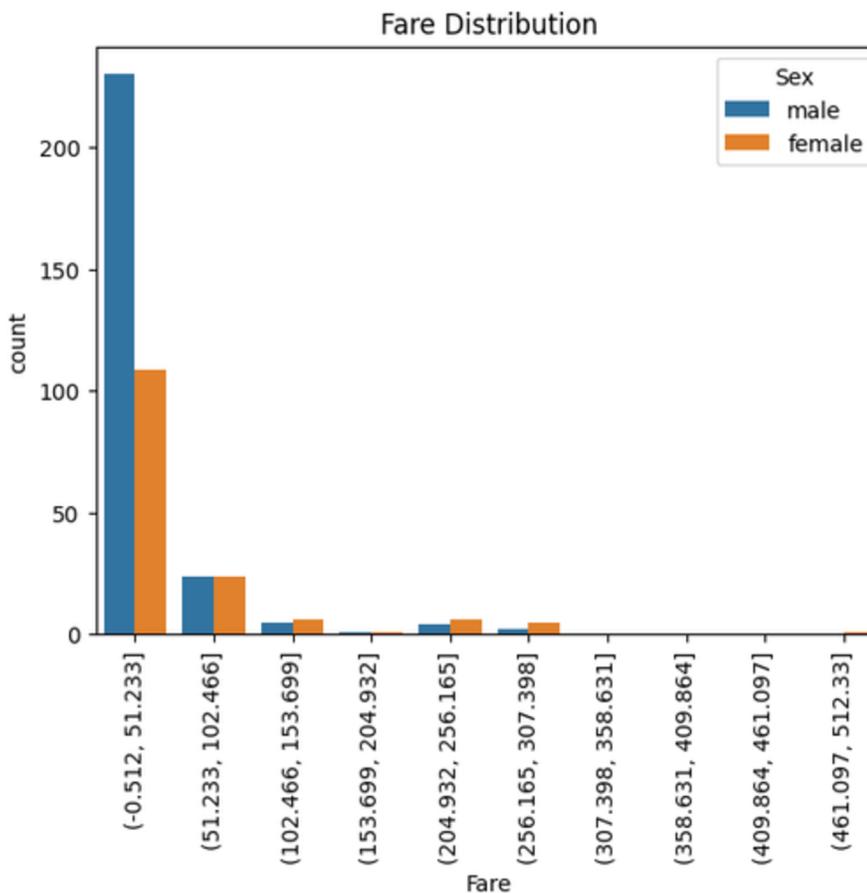


Chart 6

- **Insights:**
- Most people travelled Alone
- Few people travelled with atleast one sibling or spouse.

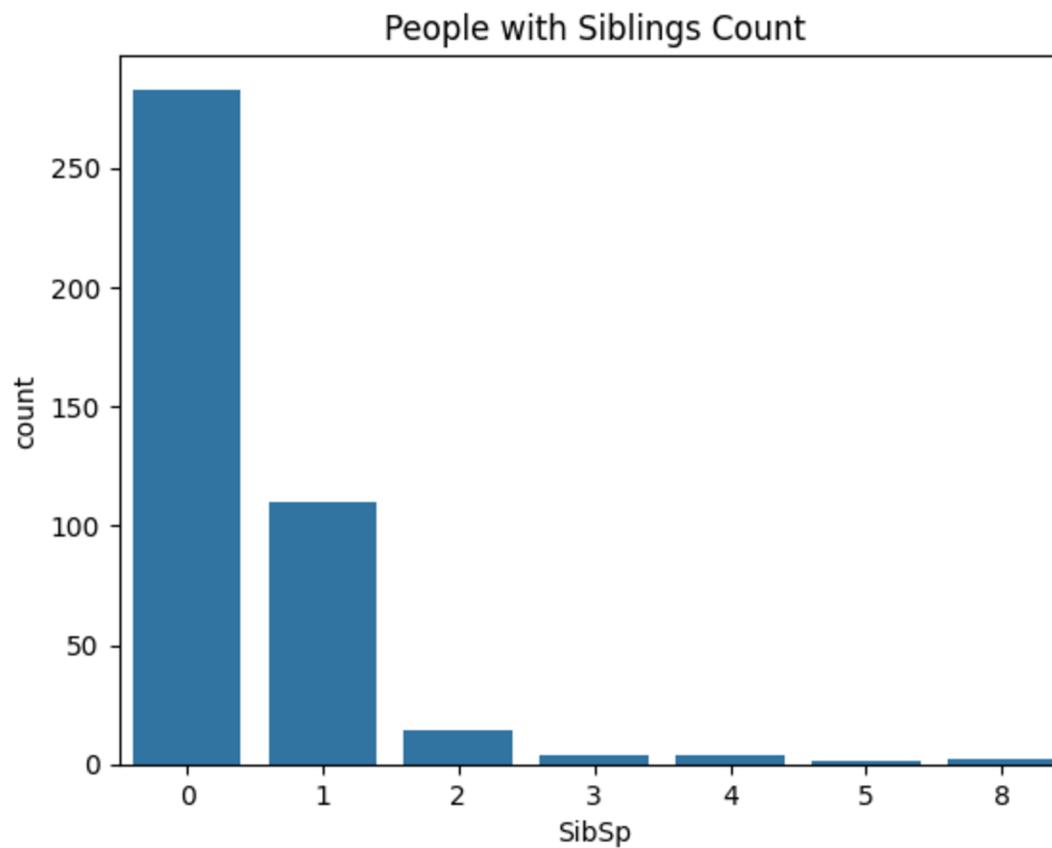


Chart 7

- **Insights:**
- Most people travelling alone are Males.
- Female count travelling alone is around 100

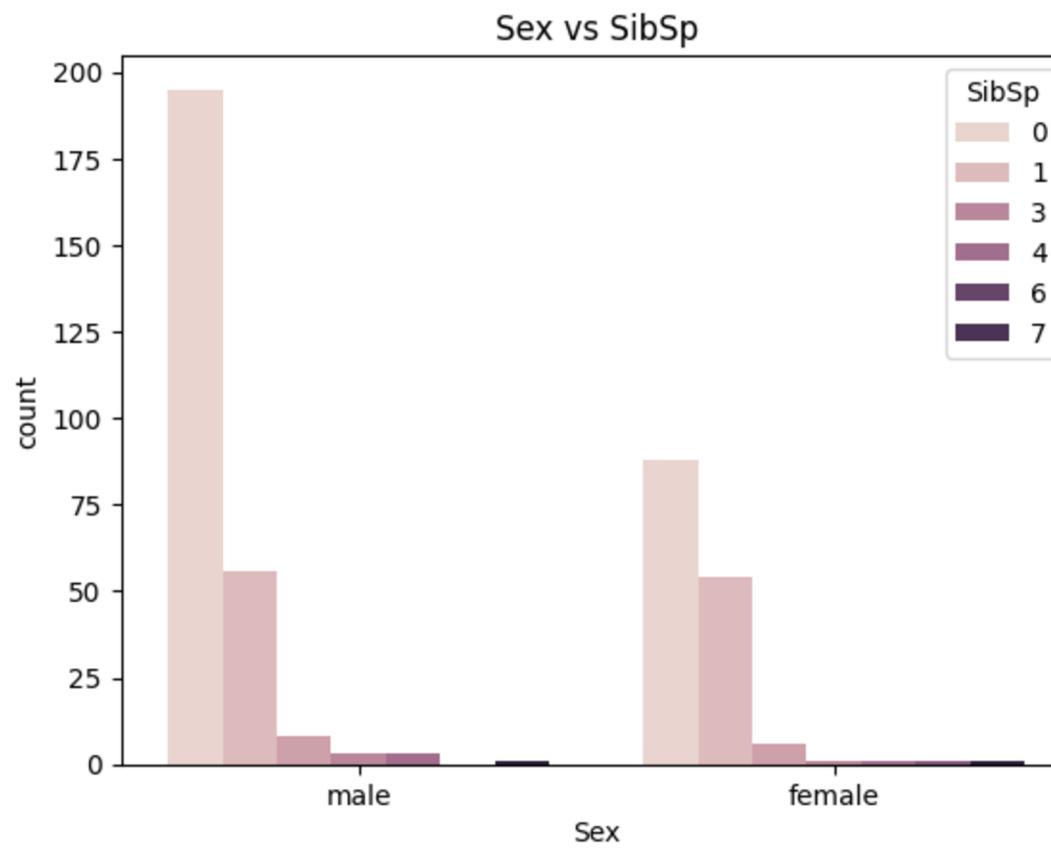


Chart 8

- **Insights:**
- Most people travelled alone.
- Few people travelled with atleast with parent or child.

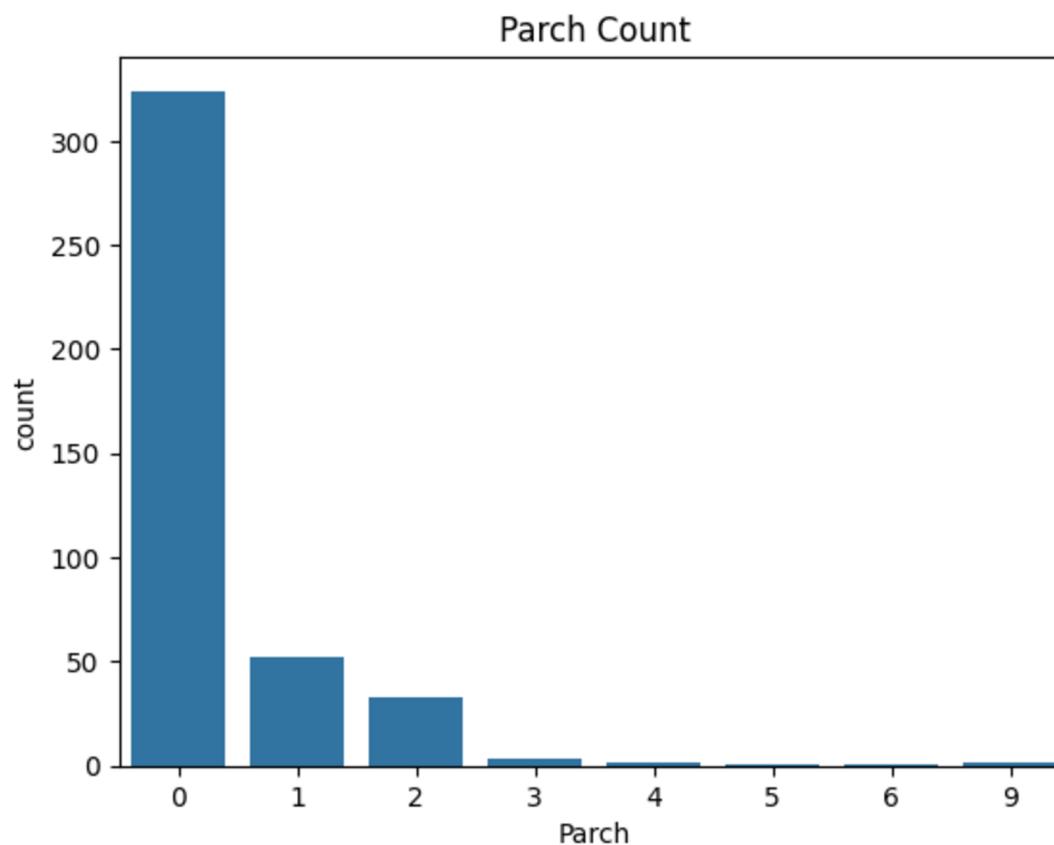


Chart 9

- **Insights:**
- Most people who travelled alone are Males
- Female count who travelled alone is around 100.
- Female lean towards having someone with them.

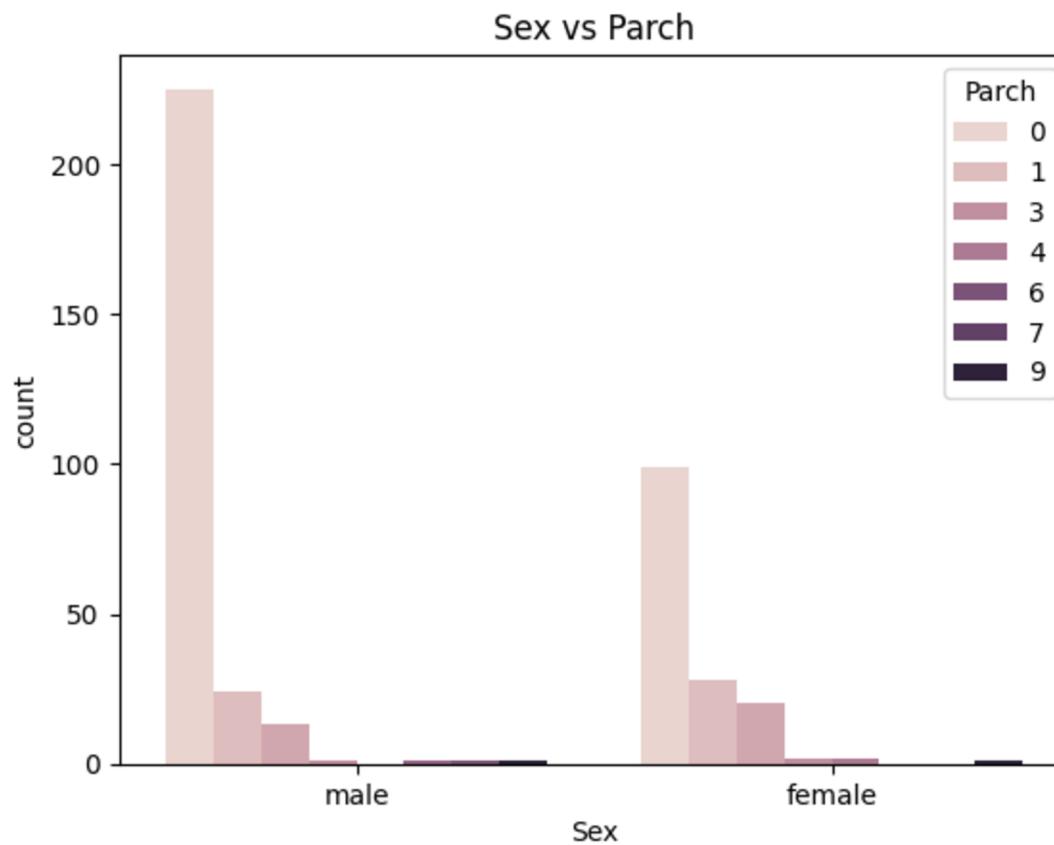


Chart 10

- **Insights:**
- Most travelled age group is 27.
- 22, 24, 21 years show similar travel results.

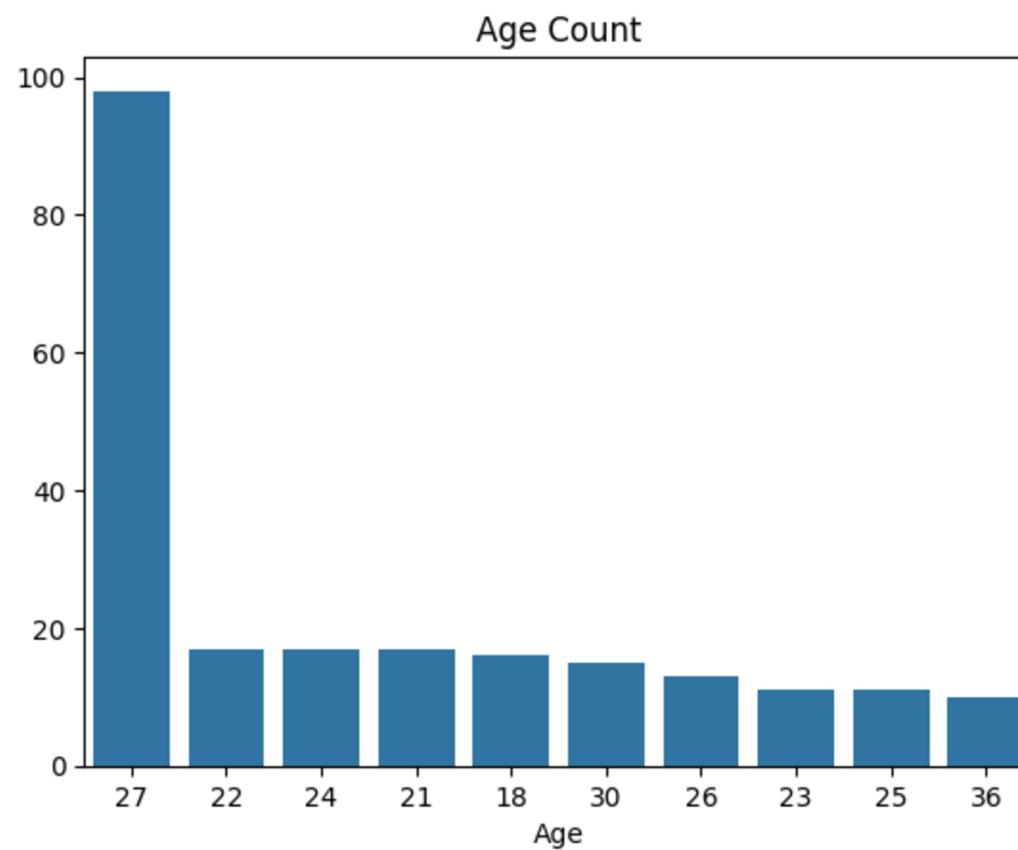


Chart 11

- **Insights:**
- Most data is scattered towards Age 20-40.
- Most fare is scattered from 0 to 100.

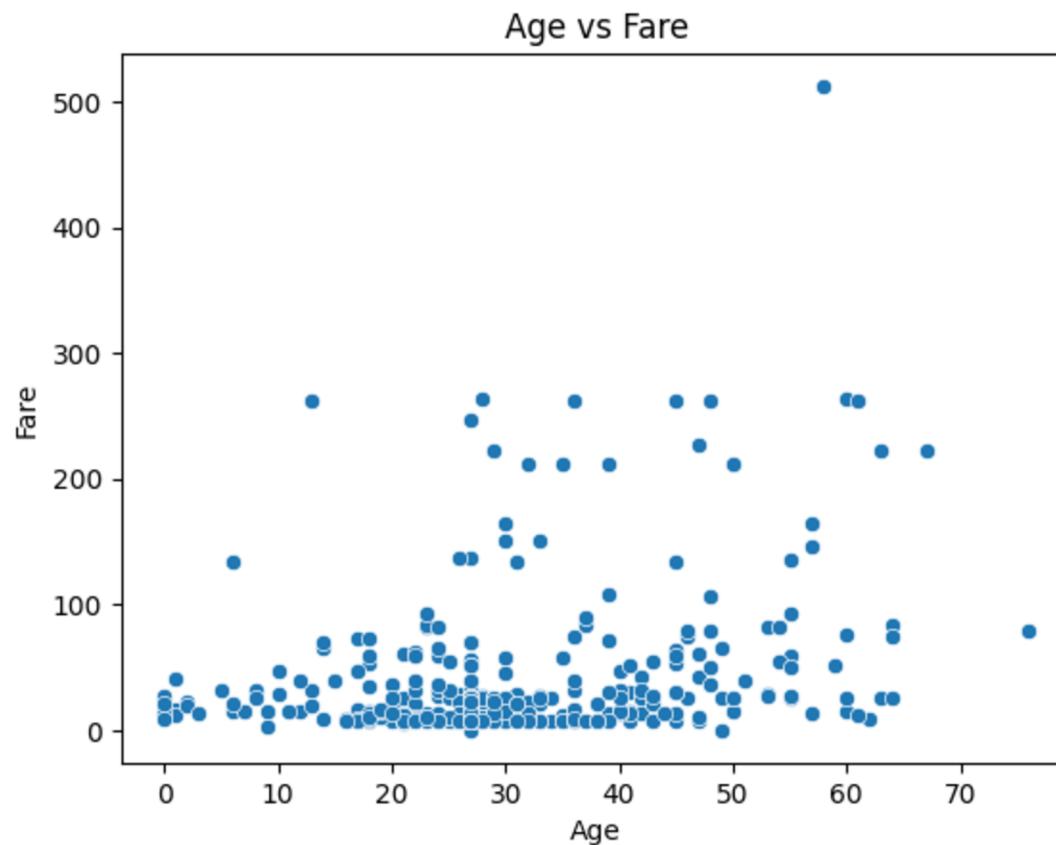


Chart 11

- **Insights:**
- Passengers who paid higher fares generally had a higher chance of survival.
- Survival rates are noticeably higher in 1st class compared to 3rd class.

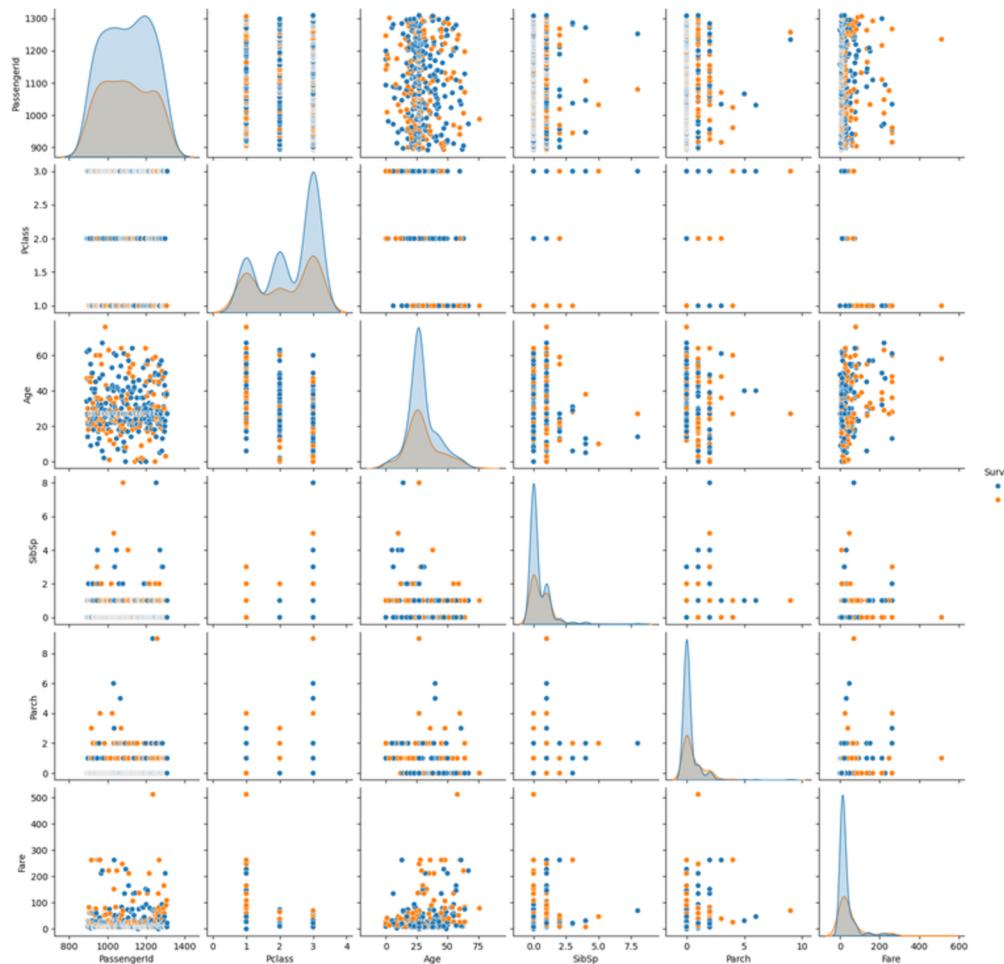


Chart 11

Business Insights

- Passenger Class and Embarkation: Most passengers were in Class 3 and embarked from Southampton.
- Missing Values: The 'Age', 'Fare', and 'Cabin' columns had missing values, which were handled by imputation (median for Age and Fare, 'Unknown' for Cabin).
- Cabin Information: A large number of cabin values were missing, and among the recorded cabins, several had only a few passengers.
- Travel Companions: A significant majority of passengers traveled alone, with fewer traveling with siblings/spouses or parents/children.
- Age Distribution: The most frequent age among passengers was 27.
- Fare Distribution: Most tickets were in the lower fare range.
- Survival:
- The pairplot suggests that passengers who paid higher fares and were in higher classes (Pclass 1) had a higher chance of survival.
- The 'Sex vs Survived' chart indicates that a higher number of males did not survive compared to females.