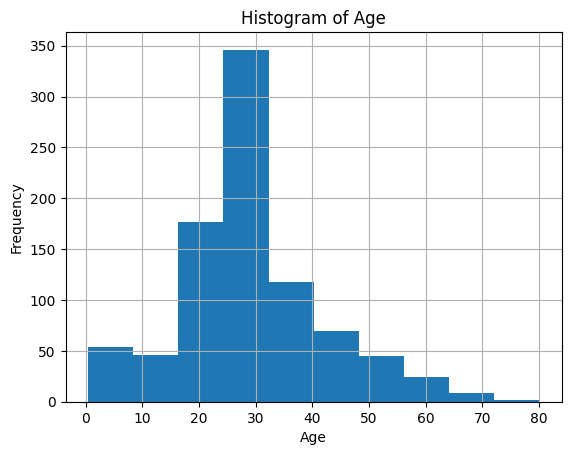
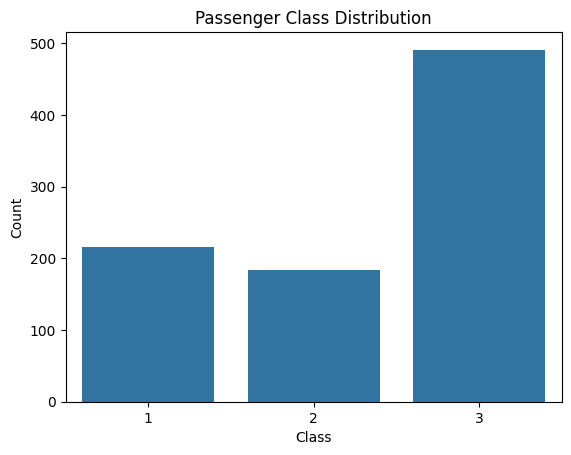
EDA on Titanic dataset

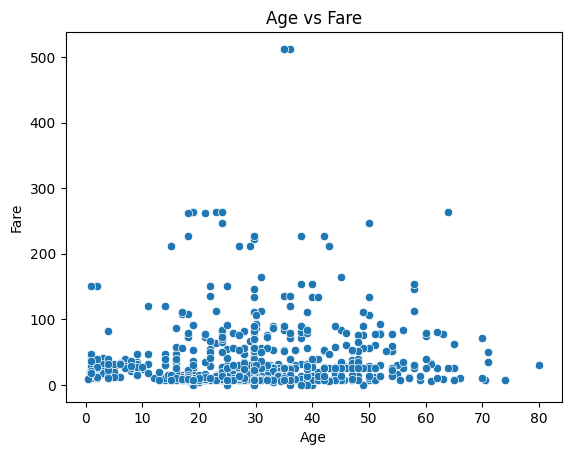
The Dataset ‘Titanic’ includes the details of the passengers boarded. The details include their name, gender, age, salary, etc.



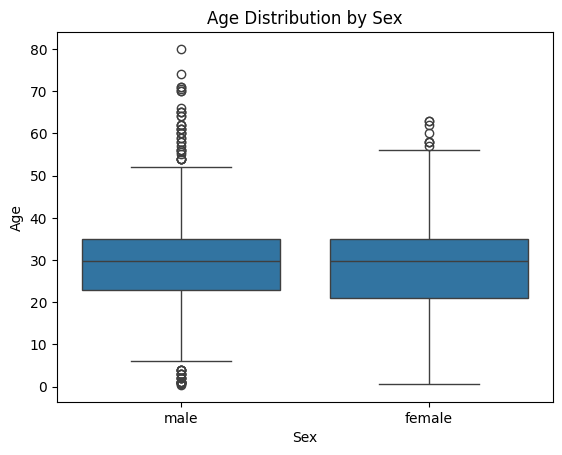
The histogram of age shows that most Titanic passengers were young adults, with the highest frequency between ages 20 and 30. There is a noticeable right skew, indicating fewer older passengers. Very few were above 60, and a moderate number were children under 10. This age distribution hints at a predominantly younger crowd aboard the ship.



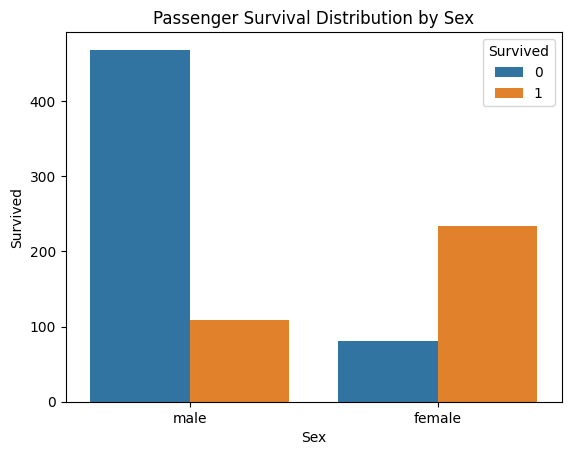
The bar chart shows that most Titanic passengers traveled in 3rd class, with nearly 500 individuals, compared to about 200 each in 1st and 2nd class. This suggests a majority of lower-income travelers onboard. The stark difference highlights the class disparity, which could have played a significant role in survival outcomes.



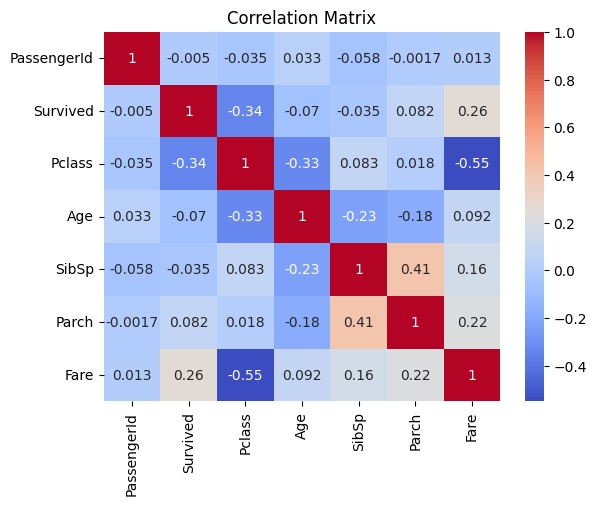
The scatter plot of Age vs Fare shows no strong correlation—passengers of all ages paid a wide range of fares. Most fares cluster under 100, with a few high-paying outliers (above 500) likely representing wealthy first-class passengers. The plot confirms that fare was more influenced by class than age. So yeah, a 20-year-old could either be ballin’ in first class or broke in steerage—age didn’t dictate the ticket price.



The boxplot shows that the age distribution for males and females on the Titanic was quite similar, with both genders having a median age around 28–30 years. However, males had slightly more outliers on the higher end—some over 70—while females had fewer extreme values. Both distributions are moderately spread out, but male ages show more variation overall. Bottom line: passengers of all ages and genders were on board, but the age spread for men was a bit more diverse.



This chart makes it crystal clear: gender was a major factor in survival. A huge number of males died (over 450), while only a small fraction survived. In contrast, more females survived than died, reflecting the "women and children first" evacuation protocol. The survival rate for females was dramatically higher—this wasn’t just luck; it was policy in action.



The correlation matrix reveals some key Titanic insights:

* Survival is moderately negatively correlated with Pclass (-0.34), meaning higher-class passengers (1st class) were more likely to survive.
* Fare shows a positive correlation with survival (0.26), which aligns—rich folks had better odds.
* Age has a weak negative correlation with survival (-0.07), hinting that younger passengers might have had a slight edge.
* Pclass is strongly negatively correlated with Fare (-0.55)—the higher the class, the more you paid (shocking, right?).