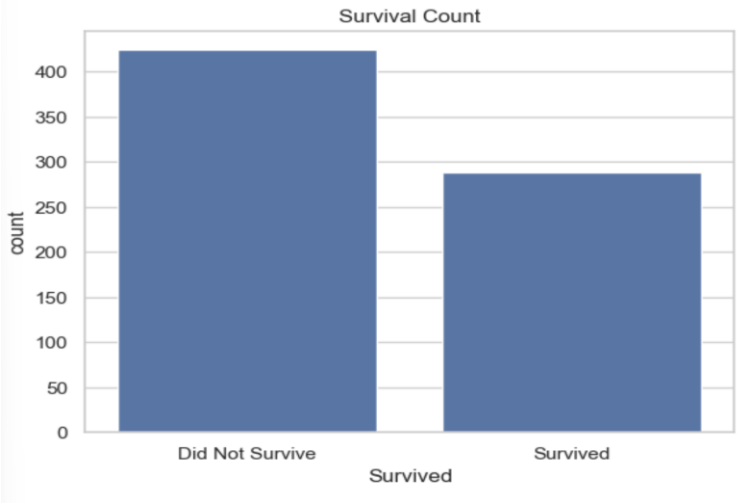
(Findings Based on Visual Analysis)

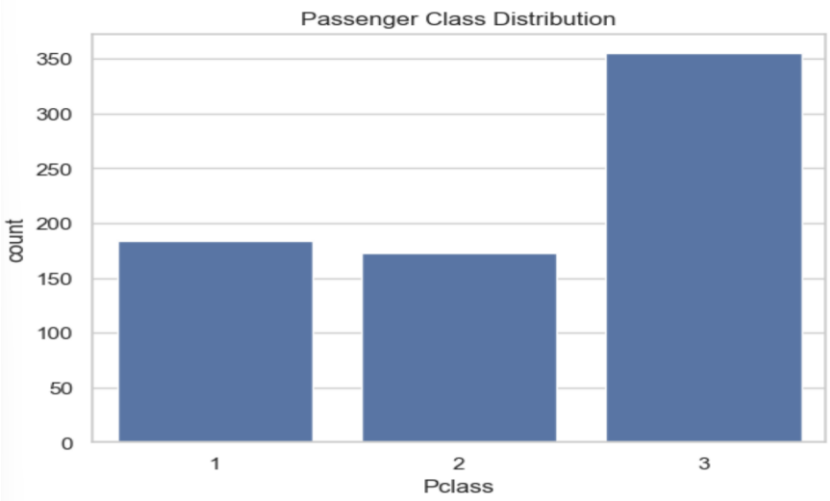
1. Survival (Target Variable)

- The dataset is imbalanced: around 62% passengers did not survive, while only 38% survived.
- This indicates the Titanic disaster had a high fatality rate overall.



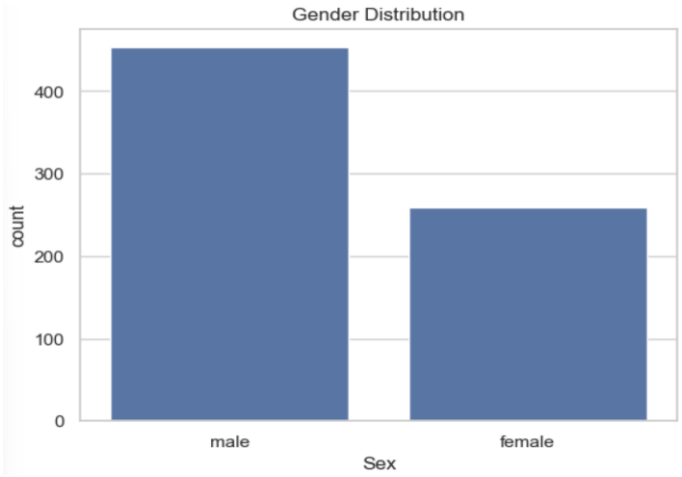
2. Pclass (Passenger Class)

- Most passengers belonged to 3rd class, followed by 1st and 2nd class.
- This suggests that the Titanic had more affordable tickets for the general public, resulting in more lower-class travelers.



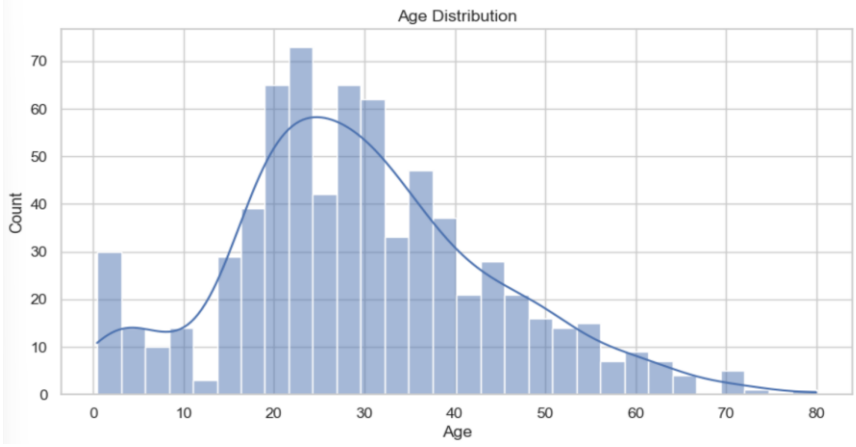
3. Sex

- The number of male passengers was significantly higher than females.
- This skew can influence overall survival statistics, as males had lower survival rates.



4. Age

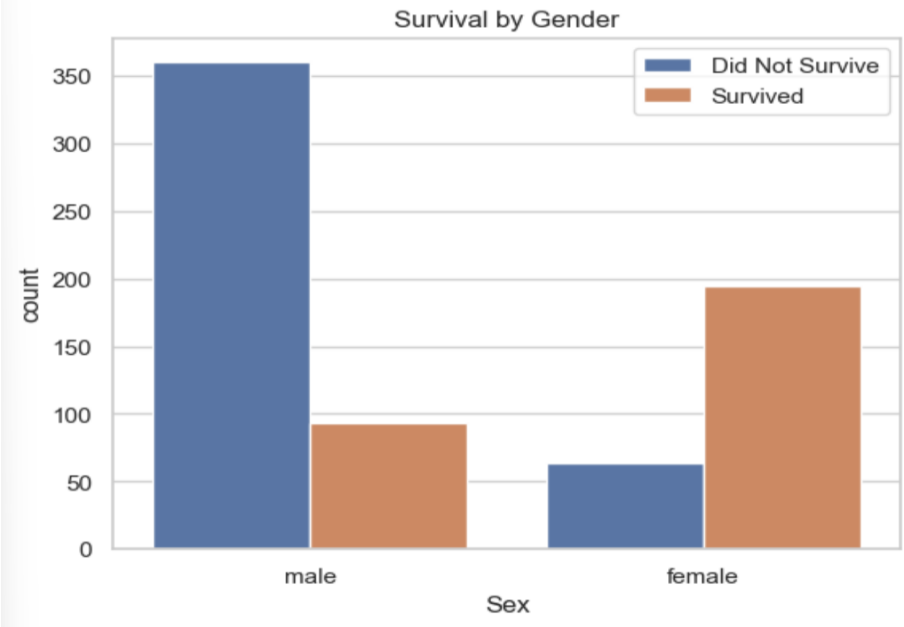
- The age distribution shows a right-skewed pattern, where:
    - Majority of passengers were between 20 and 40 years old.
    - There are fewer children and elderly passengers.
    - A small number of infants (age < 1) and seniors (age > 65) were on board.



5. Survival Rate by Gender

- From the scatterplot and count plots, it is evident that female passengers had a significantly higher survival rate compared to male passengers.

- This aligns with the historical narrative of "women and children first" being followed during evacuation.
- Conclusion: Gender played a vital role in determining survival.

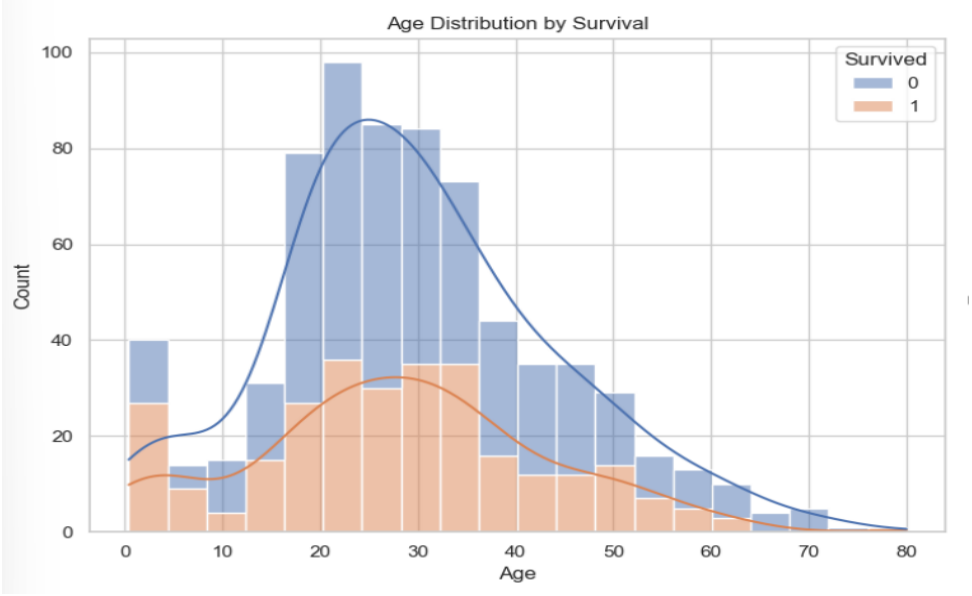


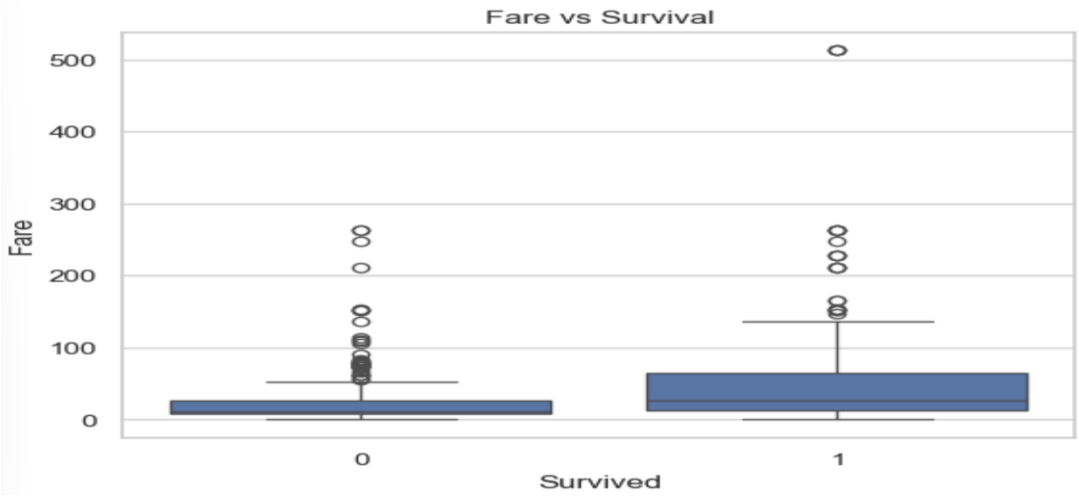
6. Survival Rate by Passenger Class (Pclass)

- The boxplot of Fare vs Pclass shows that 1st class passengers paid the highest fares and had a higher probability of survival.
- Passengers in 3rd class had the lowest survival rate, with most paying minimal fares.
- Conclusion: Higher socio-economic status (reflected by higher class) increased chances of survival.

7. Age Distribution and Survival

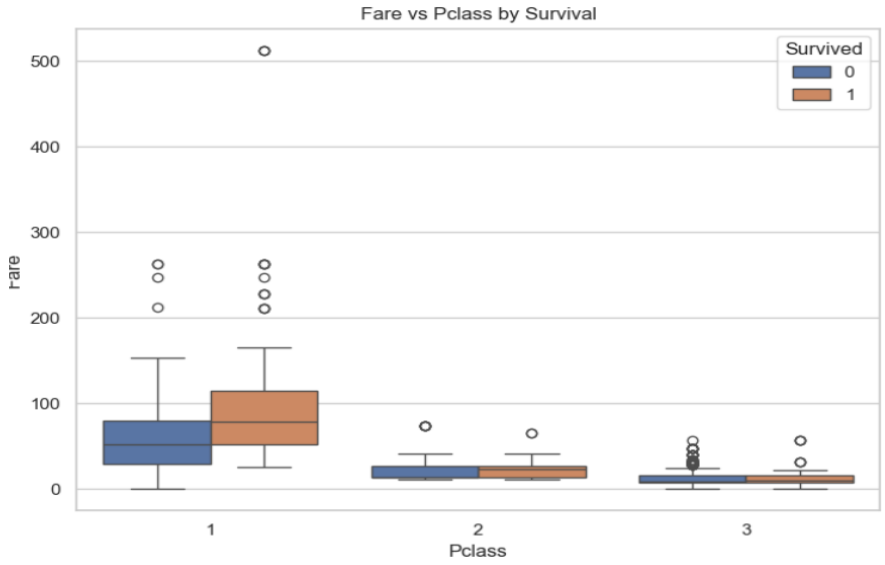
- The age histogram segmented by survival shows that children (age < 10) had relatively good survival chances.
- Most survivors were concentrated between ages 20 to 40.
- However, some older passengers in 1st class also survived, likely due to priority evacuation.
- Conclusion: Age was moderately correlated with survival, especially for children and young adults.





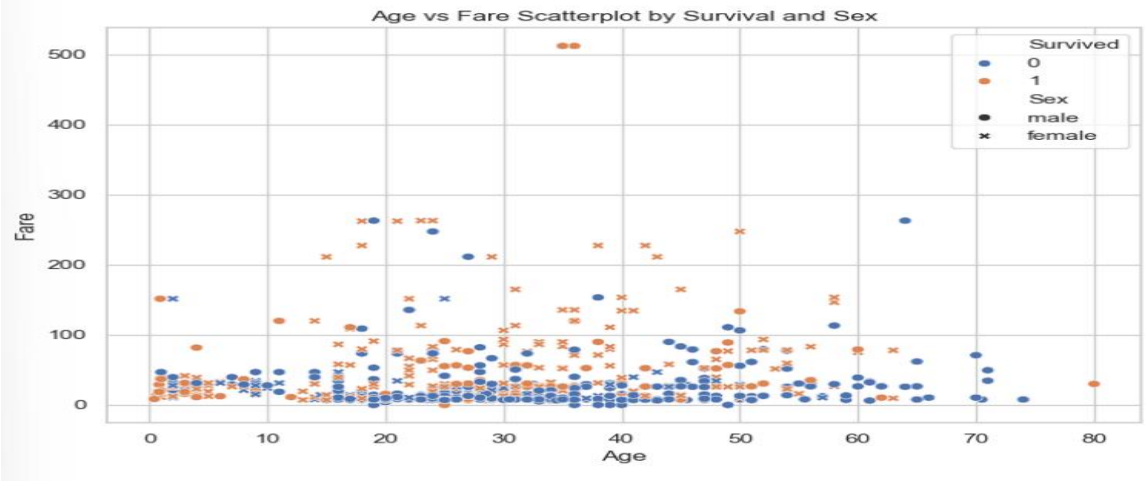
7. Distribution of Fare by Survival (Boxplot)

- The boxplot comparing Fare between survivors and non-survivors indicates that the median fare was higher among survivors.
- There are also several high-fare outliers who survived, indicating possible influence of VIP treatment or cabin location.
- Conclusion: Survival is somewhat positively correlated with fare.



8. Age vs Fare Scatterplot Analysis

- The scatterplot between Age and Fare, segmented by gender and survival, provides a deep dive:
    - Young females paying higher fares had a very high survival rate.
    - Elderly males paying lower fares had the least survival rate.
- Conclusion: The interaction between age, gender, and fare is important in survival prediction.



9. Fare

- The fare distribution is highly right-skewed:
    - Most passengers paid less than 50 (currency units).
    - A few passengers paid extremely high fares (> 200), indicating VIP or suite-class tickets.