Summary of Findings – Titanic Dataset (EDA)

1. **Data Overview:**
- The dataset contains passenger details such as age, class, fare, sex, and survival status.
- Missing values were found in age, embarked, embark_town, and deck.
- Missing values were handled by imputing median/mode, and the deck column was dropped due to excessive nulls.

2. **Univariate Analysis:**
- Most passengers were in 3rd class.
- Majority of passengers were male adults.
- Age distribution shows most passengers were between 20–40 years old.
- Fare distribution was right-skewed — few passengers paid very high fares.

3. **Bivariate Analysis:**
- Survival rates were higher for females and 1st-class passengers.
- Higher fare passengers had better survival chances.
- Adult males and 3rd-class passengers had lower survival probabilities.

4. **Correlation Insights:**
- Survival is positively correlated with fare (0.26) and negatively correlated with pclass (-0.34) and adult_male (-0.56).
- SibSp and Parch are positively correlated (0.41), showing families often traveled together.

5. **Overall Insights:**
- Socioeconomic status and gender played a key role in survival.
- Passengers traveling in 1st class and paying higher fares had better chances of survival.
- Women and children had a higher survival rate than adult males.