REPORT ON TASK 5

FINDINGS FROM BASIC EXPLORATION:

- The dataset contains passenger details such as age, class, fare, gender, etc.
- Survived is binary (0 = did not survive, 1 = survived).
- There are missing values in some columns (e.g., Age, Cabin, Embarked).

OBSERVATIONS FROM HEATMAP:

- Age has several missing values.
- Cabin has many missing entries, indicating incomplete records for passenger cabin numbers.
- Embarked has a few missing values.

OBSERVATION FROM HISTOGRAM:

- Age distribution is right-skewed with most passengers between 20–40 years.
- Fare distribution is right-skewed with a few very high values.

OBSERVATION FROM COUNTPLOT(Survival Count):

- More passengers did not survive (0) compared to those who survived (1).
- Around 2/3 of the passengers died.

OBSERVATION FROM COUNTPLOT(Passenger Class Count):

- Majority of passengers traveled in 3rd class.
- 1st class had the fewest passengers, but higher survival rates will be explored later.

OBSERVATION FROM COUNTPLOT(Survival by Gender):

- Females had a much higher survival rate than males.
- The number of males who died is significantly larger than any other group.

OBSERVATION FROM COUNTPLOT(Survival by Passenger Class):

- 1st class passengers had the highest survival rate.

- 3rd class passengers had the lowest survival rate.

- Class appears to be strongly related to survival chances.

OBSERVATION FROM BOXPLOT (Age distribution by Survival):

- Median age of survivors is slightly lower than that of non-survivors.

- Young children have a higher chance of survival.

- Fare and Pclass are inversely related (lower class = lower fare).

- Higher fare passengers show a greater survival tendency.

- Age does not have a strong separation between survivors and non-survivors in scatterplots.

OBSERVATION FROM CORRELATION HEAT MAP:

- Survived has a positive correlation with Fare and a negative correlation with Pclass.

- Age has little correlation with survival.

- Pclass and Fare are strongly correlated.

FINAL SUMMARY OF EDA:

- The Titanic dataset clearly shows that survival chances were influenced by gender, passenger class, and fare.

- Females had much higher survival rates than males.

- 1st class passengers were prioritized during rescue, followed by 2nd and then 3rd class.

- Higher fare (linked to better class) increased survival probability.

- Younger passengers, especially children, had higher survival chances.

- Missing data is mainly in Age, Cabin, and Embarked columns.