# Elevate Labs Data Analyst Internship

## Task 5 : Exploratory Data Analysis (EDA)

**📊 Summary of Findings — Titanic EDA**

1. **Passenger Demographics**

   o The majority of passengers were male.

   o Most passengers were in the 20–40 age range.

   o Passenger classes distribution: 3rd class had the most passengers, followed by 1st and 2nd class.

2. **Survival Insights**

   o Females had a much higher survival rate compared to males.

   o First-class passengers had the highest survival rate, followed by 2nd class, with 3rd class having the lowest.

   o Children (under 10 years old) had higher survival rates than adults.

3. **Fare & Survival**

   o Higher ticket prices (fares) were linked to higher survival probability.

   o First-class passengers generally paid more, which correlates with their higher survival rate.

4. **Correlation Observations**

   o Positive correlation between Fare and Survival (higher fare → more likely survived).

   o Strong negative correlation between Pclass and Fare (lower class number = higher fare).

   o Gender had a strong correlation with survival (being female increased survival odds).

5. **Other Observations**

   o Outliers in Fare show a few passengers paid extremely high amounts, mostly in first class.

   o Age distribution is relatively uniform for both survivors and non-survivors except for a small peak for young children among survivors.

   o Large families (higher SibSp or Parch) generally had lower survival rates compared to small families.