**Laboratory Activity: Data Exploration and Preprocessing**

**Objective:**  
To learn how to explore and preprocess a dataset using Microsoft Excel.

**Dataset:**  
We'll use the Titanic dataset, which contains information about passengers on the Titanic and whether they survived.

**Download Link:** [Titanic Dataset CSV](https://github.com/datasciencedojo/datasets/raw/master/titanic.csv)

**Step-by-Step Instructions**

1. **Data Collection and Importing**

a. Download the Titanic dataset from the link above.

1. **Data Cleaning and Transformation**

a. **View the Data:** Check the first few rows of the dataset to see what it looks like.

b. **Check for Duplicates:** Ensure there are no duplicate rows in the dataset.

d. **Convert Categorical Data:** Change the 'Sex' column to numbers (0 for male, 1 for female).

1. **Handling Missing Data and Outliers**

a. **Missing Values:** Identify any missing values and fill them in with appropriate values (e.g., replace missing ages with the average age).

1. **Feature Selection and Engineering**

a. **Create New Features:** Add a new column called 'FamilySize' by combining 'SibSp' (siblings/spouses) and 'Parch' (parents/children).

1. **Data Exploration**

a. **Summary Statistics:** Get basic statistics (mean, median) for 'Age' and 'Fare'.

b. **Visualizations:** Plot histograms for 'Age' and 'Fare', and bar charts to show survival rates by 'Pclass' and 'Sex'.