

Exercise #1

This report presents an analysis of the Titanic dataset, aiming to explore key attributes and draw insights regarding the passengers' demographics and survival rates. The analysis was conducted using Python, with libraries such as Pandas, Seaborn, and Matplotlib.

Dataset Overview:

The Titanic dataset contains information about passengers, including their demographics, ticket class, embarkation point, and survival status. The dataset consists of 891 rows and 12 columns.

Data Exploration:

1. Data Summary:

- The dataset includes both numerical and categorical columns.
- There are missing values in some columns, which may need further handling.

2. Gender Distribution:

- The analysis reveals that the dataset has a gender distribution, with male passengers (approximately 65%) outnumbering female passengers (approximately 35%).

3. Survival by Gender:

- When considering survival rates by gender, it is observed that a significantly higher number of females survived compared to males, indicating a strong association between gender and survival.

```
In [39]: titanic[['Sex', 'Survived']].groupby('Sex').count()
```

```
Out[39]:
```

Survived	
Sex	
female	314
male	577

```
In [55]: titanic[['Pclass', 'Survived']].groupby(['Pclass']).count()
```

```
Out[55]:
```

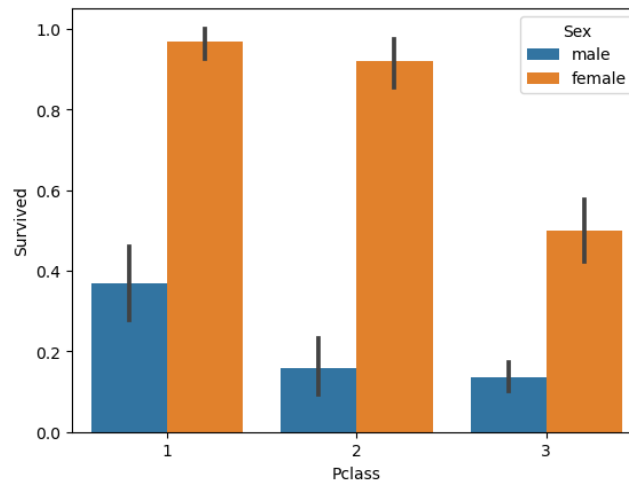
Survived	
Pclass	
1	216
2	184
3	491

4. Passenger Class Distribution:

- The passenger class distribution shows that the majority of passengers are in the third class, followed by the first and second classes.

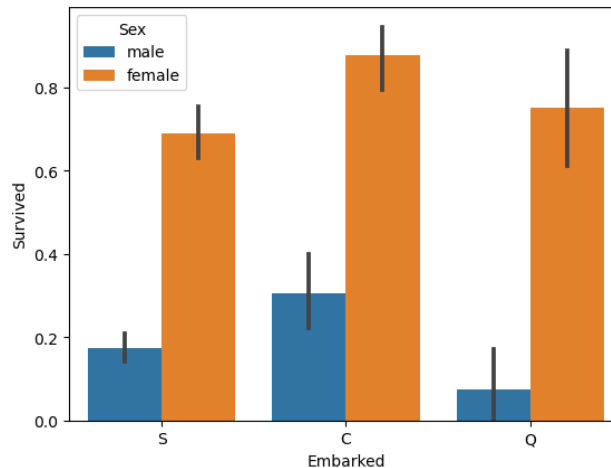
5. Survival by Passenger Class:

- Survival rates vary significantly across passenger classes. Passengers in the first class have the highest survival rate, followed by the second class, while the third class has the lowest survival rate.



6. Survival by Embarkation Point:

- Survival rates also appear to vary by embarkation point. Passengers who embarked from Cherbourg (C) have a relatively higher survival rate compared to those who embarked from Southampton (S) and Queenstown (Q).



Data Visualization:

1. Survival by Passenger Class and Gender:

- A bar plot illustrates that first-class female passengers had the highest survival rate, emphasizing the importance of both gender and class in survival.

2. Survival by Embarkation Point and Gender:

- Another bar plot indicates that females who embarked from Cherbourg (C) had a higher survival rate compared to other embarkation points.

Conclusion:

The analysis of the Titanic dataset revealed several important insights:

- Gender played a significant role in survival, with females having a higher chance of survival than males.
- Passenger class also had a substantial impact on survival, with first-class passengers having the highest survival rate.
- The embarkation point showed variations in survival rates, with passengers from Cherbourg having a relatively higher survival rate.