

Titanic Classification

Draft saved

File Edit View Run Add-ons Help

  Run All 

Save Version

Data Preprocessing

```
[51]: # Number of rows and columns
      titanic_data.shape
```

```
[51...] (891, 12)
```



```
# Get information about the data
titanic_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column          Non-Null Count  Dtype
---  -
 0   PassengerId     891 non-null    int64
 1   Survived        891 non-null    int64
 2   Pclass         891 non-null    int64
 3   Name            891 non-null    object
 4   Sex             891 non-null    object
 5   Age             714 non-null    float64
 6   SibSp           891 non-null    int64
 7   Parch           891 non-null    int64
 8   Ticket          891 non-null    object
 9   Fare            891 non-null    float64
10   Cabin           204 non-null    object
11   Embarked        889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

[+ Code](#)[+ Markdown](#)

```
[53]: # To check the number of missing values
      titanic_data.isnull().sum()
```

```
[53...] PassengerId     0
      Survived        0
      Pclass         0
      Name            0
      Sex             0
      Age             177
```

 Draft Session (0m)