**Department of Computer Science and Engineering** **Data Science**

**Academic Year:** 2024-2025 **Name of Student:** Diya Thakkar

**Semester:** VI **Student ID:** 22107040

**Class / Branch:** TE CSE (DS) **Date of Performance:**3-2-25 **Subject:** ML Lab **Date of Submission:** 3-2-25

**Name of Instructor:** Prof.Ujwala Pagare

**Experiment No. 3**

**Aim:**- To implement Logistic Regression using Python

**Program:-**

**import numpy as np**

**import** pandas **as** pd

**import** matplotlib.pyplot **as** plt

**import** seaborn **as** sns

***# load the data from csv file to Pandas DataFrame***

titanic\_data **=** pd.read\_csv('/content/Titanic dataset.csv')

***# printing the first 5 rows of the dataframe***

titanic\_data.head()

*# number of rows and Columns*

titanic\_data.shape

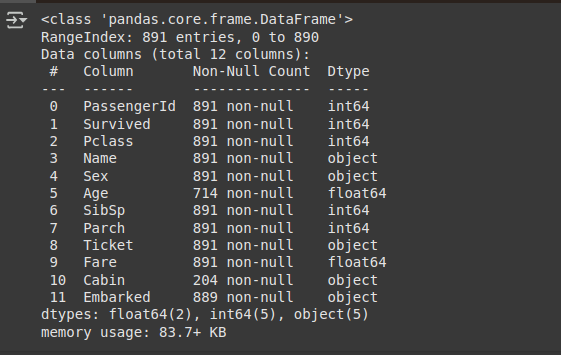
**Output:**



***# getting some informations about the data***

titanic\_data.info()

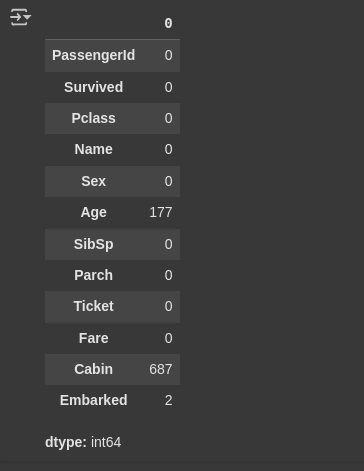
**Output:**



*# check the number of missing values in each column*

titanic\_data.isnull().**sum**()

**Output:**

****

*# drop the "Cabin" column from the dataframe*

titanic\_data**=**titanic\_data.drop(columns**=**'Cabin',axis**=1**)

*# replacing the missing values in "Age" column with mean value*

titanic\_data['Age'].fillna(titanic\_data['Age'].mean(),inplace**=True**)

*# Finding the mode value of the “Embarked” column as it will have* *occurred the maximum number of times*

**print**(titanic\_data['Embarked'].mode())

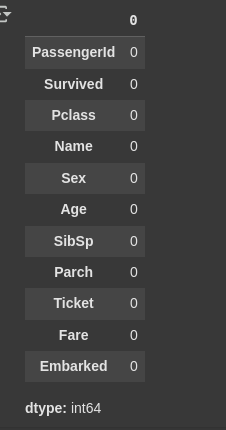
**print**(titanic\_data['Embarked'].mode()[**0**])

******

*#Replacing the missing values in the “Embarked” column with mode value*

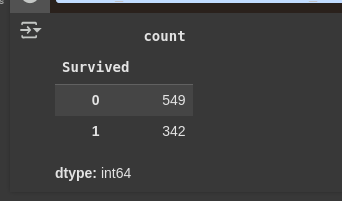
titanic\_data['Embarked'].fillna(titanic\_data['Embarked'].mode()[**0**],inplace**=True**)

titanic\_data.isnull().**sum**()



titanic\_data.describe()

titanic\_data['Survived'].value\_counts()

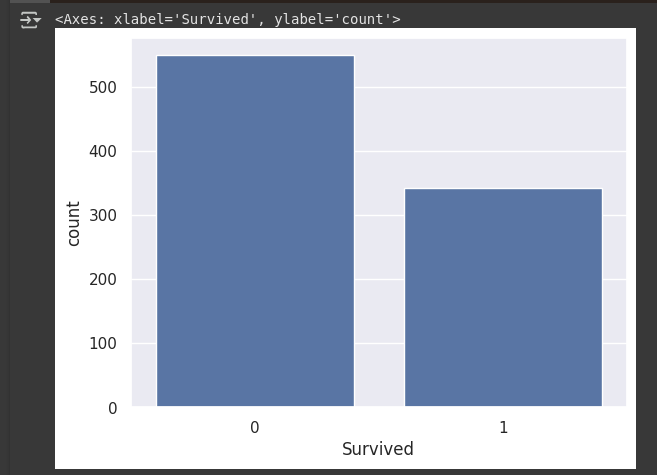
****

*# It will load seaborn's default theme and color palette to the session.*

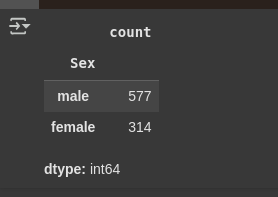
sns.**set**()

*# making a count plot for "Survived" column*

sns.countplot(x**=**'Survived',data**=**titanic\_data)

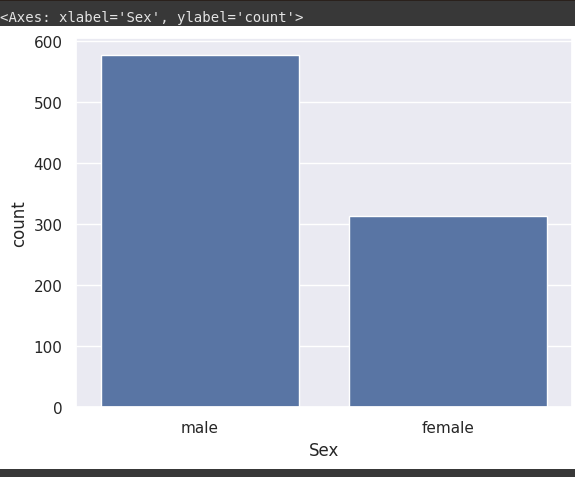


titanic\_data['Sex'].value\_counts()

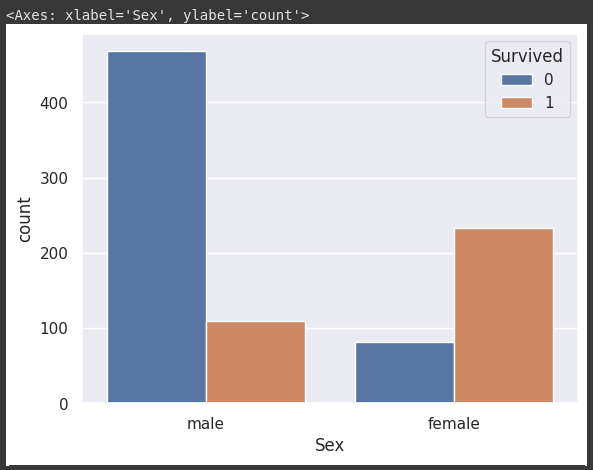


*# making a count plot for "Sex" column*

sns.countplot(x**=**'Sex', data**=**titanic\_data)

*# number of survivors Gender wise*

sns.countplot(x**=**'Sex', hue**=**'Survived', data**=**titanic\_data)

*# making a count plot for "Pclass" column*

sns.countplot(x**=**'Pclass', data**=**titanic\_data)

