**import pandas as pd**

**5/7/24, 6:42 PM Assignment 8 .ipynb - Colab**

**import numpy as np**

**import seaborn as sns**

**import matplotlib.pyplot as plt**

**data =pd.read\_csv('/content/Titanic-Dataset.csv') df=pd.DataFrame(data)**

**df**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Passengerld** | **Survived** | **Pclass** | **Name** | **Sex** | **Age SibSp Parch Ticket** | | | **Fa** |
| 0 | 1 | 0 | 3 | Braund, Mr. Owen | male | 22.0 | 1 | 0 A/5 21171 | 7.25 |
|  |  |  |  | Harris |  |  |  |  |  |
|  |  |  |  | Cumings,  Mrs. John |  |  |  |  |  |
| 1 | 2 | 1 | 1 | Bradley  female  (Florence | | 38.0 | 1 | 0 PC 17599 | 71.28 |
|  |  |  |  | Briggs |  |  |  |  |  |
|  |  |  |  | Th... |  |  |  |  |  |
| **2** | 3 | 1 | 3 | Heikkinen,  Miss. female  Laina | | 26.0 | 0 | STON/02.  0  3101282 | 7.92 |
|  |  |  |  | Futrelle,  Mrs. |  |  |  |  |  |
| **3** | 4 | 1 | 1 | Jacques  female  Heath | | 35.0 | 1 | 0 113803 | 53.1C |
|  |  |  |  | (Lily May |  |  |  |  |  |
|  |  |  |  | Peel) |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | ■ |
| **df.isnull()** |  |  |  |  |  |  |  |  |  |
|  | **Passengerld** | **Survived** | **Pclass** | **Name Sex** | **Age** | **SibSp** | **Parch** | **Ticket Fare** | **Cabin** |
| 0 | False | False | False | False False | False | False | False | False False | True |
| 1 | False | False | False | False False | False | False | False | False False | False |
| **2** | False | False | False | False False | False | False | False | False False | True |
| **3** | False | False | False | False False | False | False | False | False False | False |
| **4** | False | False | False | False False | False | False | False | False False | True |
| **886** | False | False | False | False False | False | False | False | False False | True |
| **887** | False | False | False | False False | False | False | False | False False | False |
| **888** | False | False | False | False False | True | False | False | False False | True |
| **889** | False | False | False | False False | False | False | False | False False | False |
| **890** | False | False | False | False False | False | False | False | False False | True |

**RA1 rows x 19 columns**

■

**df.head()**

**Passengerld Survived Pclass Name Sex Age SibSp Parch Ticket Fare**

Braund,

1 0 3 Mr. Owen male 22.0

Harris

2 1 1

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 0 | A/5 21171 | 7.2500 |
| 1 | 0 | PC 17599 | 71.2833 |

0

1

|  |  |
| --- | --- |
| Cumings, Mrs. John Bradley  (Florence  \_ . | female 38.0 |

■

**df.tail()**

[**https://colab.research.google.com/drive/15xrZ1oKiH**](https://colab.research.google.com/drive/15xrZ1oKiH) **P20apYxogJjS5P4fi8UJJ L\_#printMode=true 1/9**

**Passengerld Survived Pclass Name Sex Age SibSp Parch Ticket Fare C**

5/7/24, 6:42 PM Assignment 8 .ipynb - Colab

<https://colab.research.google.com/drive/15xrZ1oKiH> P20apYxogJjS5P4fi8UJJ L\_#printMode=true 2/9

Montvila,

**886** 887 0 2 Rev. male 27.0 0 0 211536 13.00

Juozas

|  |  |  |  |
| --- | --- | --- | --- |
| **887** 888 1 1 | Graham,  Miss.  Margaret  Edith | female 19.0 0 0 112053 30.00 |  |
| df.isnull().sum() |  |  | ■ |

Passengerld 0

Survived 0

Pclass 0

Name 0

Sex 0

Age 177

SibSp 0

Parch 0

Ticket 0

Fare 0

Cabin 687

Embarked 2

dtype: int64

df['Age']=df['Age'].fillna(np.mean(df['Age']))

df.isnull().sum()

Passengerld 0

Survived 0

Pclass 0

Name 0

Sex 0

Age 0

SibSp 0

Parch 0

Ticket 0

Fare 0

Cabin 687

Embarked 2

dtype: int64

df['Cabin']=df['Cabin'].fillna((df['Cabin'].mode()[0])) df['Embarked']=df['Embarked'].fillna((df['Embarked'].mode()[0])) df.isnull().sum()

Passengerld 0

Survived 0

Pclass 0

Name 0

Sex 0

Age 0

SibSp 0

Parch 0

Ticket 0

Fare 0

Cabin 0

Embarked 0

dtype: int64

sns.distplot(x=dfrAge'libins=10)

**5/7/24, 6:42 PM Assignment 8 .ipynb - Colab**

[**https://colab. research.google.com/drive/15xrZ1oKiH**](https://colab.research.google.com/drive/15xrZ1oKiH) **P20apYxogJjS5P4fi8UJJ L\_#printMode=true 3/9**

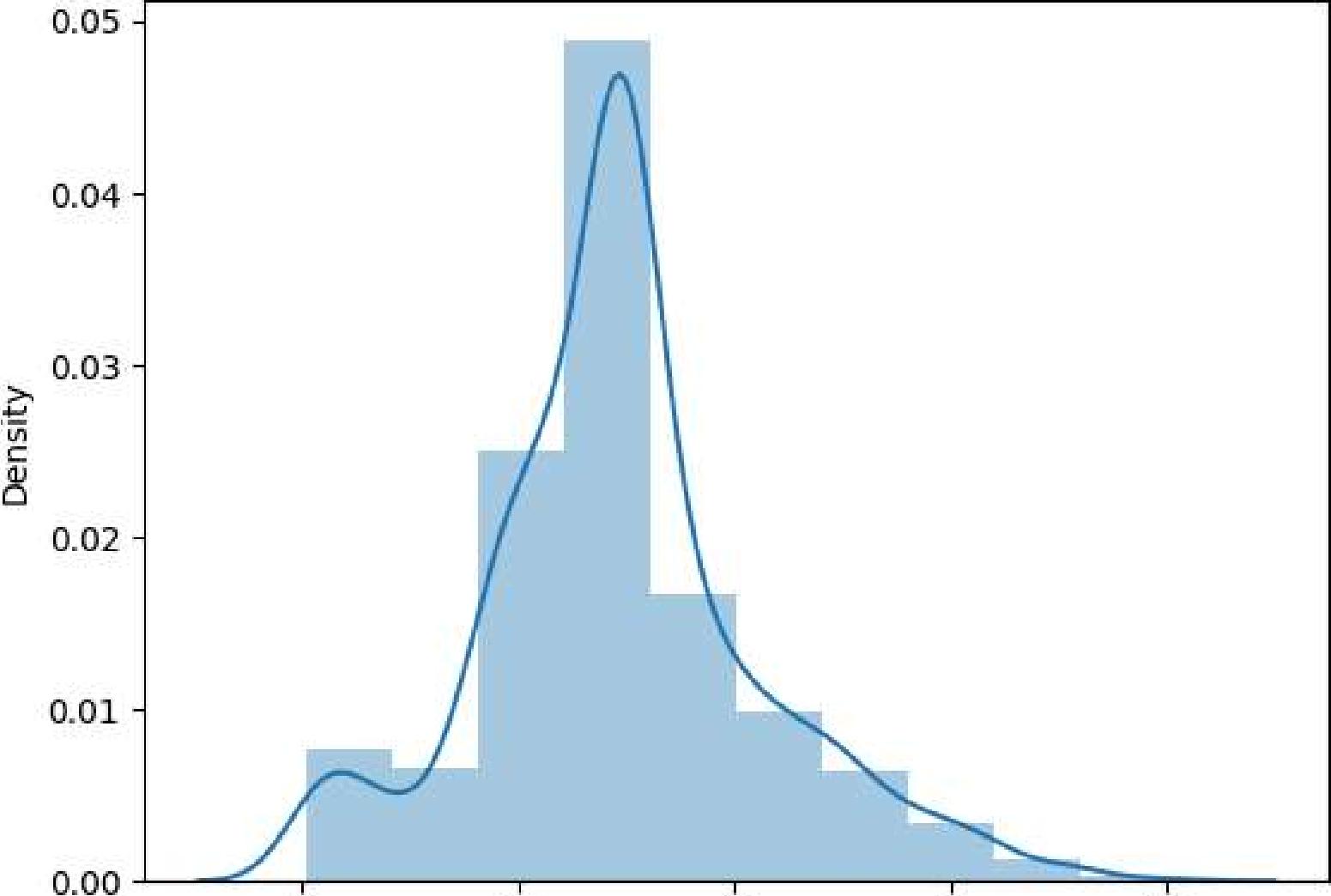
**4) <ipython-input-12-bc9510521c67>:1: UserWarning:**

**-distplot' is a deprecated function and will be removed in seaborn v0.14.0.**

**Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or -histplot- (an axes-level function for histograms).**

**For a guide to updating your code to use the new functions, please see** [**httuILLgist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751**](http://httuILLgist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751)

**sns.distplot(x=df['Age'],bins=10) <Axes: ylabel='Density'>**



|  |  |  |  |
| --- | --- | --- | --- |
| 20 | 40 | 60 | 80 |

**sns.jointplot(x=df['Age'],y=df['Fare']) <seaborn.axisgrid.JointGrid at 0x7997a5a74ac0>**

=■ *•*

500 -

400-

|  |  |
| --- | --- |
| **300 *-***  ***P-2*** | SI•CO • |

* **•** •
* **on** •

**200-**

* **P**
* **• IS IS** •
* **•**

**!a • V • S. •**

* **-•41.1.1.. •a •**

**100-**

**r**

**%a salodi**

**•• • . •••**

**G •**

**4n 614S%iF 1•ireel Wage' 0\_40 01 •** •
  
**.1•174% gip "al I 421) t •**

**0-**

**■•■**

0 lb 20 30 40 50, 6D 70 80

Age

**sns.rugplot(df['Fare'])**

<Axes: xlabel='Fare'>

5/7/24, 6:42 PM Assignment 8 .ipynb - Colab

* 0.06



0 06

0.04 -

0.02 -

0.00 -

* 0.02 -
* 0.04 -

|  |  |  |  |
| --- | --- | --- | --- |
| 0 | 100 200 | 300  Fare | 400 500 |

sns.barplot(x=df['Sex'],y=df['Age']) <Axes: xlabel='Sex', ylabel='Age'>

30- 25 - 20 - 15 - ID-5 - 0

male female

Sex

sns.countplot(x=df['Survived'])

<https://colab.research.google.com/drive/15xrZ1oKiH> P20apYxogJjS5P4fi8UJJ L\_#printMode=true 4/9

<Axes: xlabel='Survived', ylabel='count'>

5/7/24, 6:42 PM Assignment 8 .ipynb - Colab

500-

400-

c 300 -

=

8

200 -

100 -

0

1

Survived

<https://colab.research.google.com/drive/15xrZ1oKiH> P20apYxogJjS5P4fi8UJJ L\_#printMode=true 5/9

sns.countplot(x=df['Pclass'])

<Axes: xlabel='Pclass', ylabel='count'> 500

400-

300

a

200 -

100 -

**0**

|  |  |  |
| --- | --- | --- |
| **1** | 2  **Pclass** | **3** |

sns.boxplot(x=df['Sex'],y=df['Age'])