**ASSIGNMENT-4**

**Questions 1:**

How to import pandas and check the version?

**Ans**: # importing pandas as pd

import pandas as pd

# Check the version of the dependencies

pd.show\_versions()

**Questions** 2: How to create a series from a numpy array?

**Ans**:

# import pandas as pd

import pandas as pd

# import numpy as np

import numpy as np

# numpy array

arr = np.array(['G','E','E','K','S','F','O','R','G','E','E','K','S'])

# forming series

s = pd.Series(arr)

# output

print(s)

**Questions** 3:How to convert the index of a series into a column of a dataframe?

Ans:

|  |
| --- |
| # importing the pandas library as pd  import pandas as pd      # Creating the dataframe df  df = pd.DataFrame({'Roll Number': ['20CSE29', '20CSE49', '20CSE36', '20CSE44'],                     'Name': ['Amelia', 'Sam', 'Dean', 'Jessica'],                     'Marks In Percentage': [97, 90, 70, 82],                     'Grade': ['A', 'A', 'C', 'B'],                     'Subject': ['Physics', 'Physics', 'Physics', 'Physics']})    # Printing the dataframe  Df  **Questions 4**:  Write the code to list all the datasets available in seaborn library.  Load the 'mpg' dataset  Note: mpg dataset will be read from seaborn module in the manner sir has already shown(provided in the  materials folder) |

**Ans**:

The following command will help you import Pandas −

# Pandas for managing datasets

import pandas as pd

Now, let us import the Matplotlib library, which helps us customize our plots.

# Matplotlib for additional customization

from matplotlib import pyplot as plt

We will import the Seaborn library with the following command −

# Seaborn for plotting and styling

import seaborn as sb

Importing Datasets

We have imported the required libraries. In this section, we will understand how to import the required datasets.

Seaborn comes with a few important datasets in the library. When Seaborn is installed, the datasets download automatically.

You can use any of these datasets for your learning. With the help of the following function you can load the required dataset

load\_dataset()

Importing Data as Pandas DataFrame

In this section, we will import a dataset. This dataset loads as Pandas DataFrame by default. If there is any function in the Pandas DataFrame, it works on this DataFrame.

The following line of code will help you import the dataset −

# Seaborn for plotting and styling

import seaborn as sb

df = sb.load\_dataset('tips')

print df.head()

**Questions** 5:

Which country origin cars are a part of this dataset?   
Makesort

text\_formatModelsort

text\_formatTypesort

text\_formatOriginsort

text\_formatDriveTrainsort

text\_formatMSRPsort

text\_formatInvoicesort

grid\_3x3EngineSizesort

grid\_3x3Cylinderssort

grid\_3x3Horsepowersort

Toyota7%

Chevrolet6%

Other (373)87%

###### 425

unique values

Sedan61%

SUV14%

Other (106)25%

Asia37%

USA34%

Other (123)29%

Front53%

Rear26%

Other (92)21%

###### 410

unique values

###### 425

unique values

1.3

8.3

3

12

73

500

Acura

MDX

SUV

Asia

All

$36,945

$33,337

3.5

6

265

Acura

RSX Type S 2dr

Sedan

Asia

Front

$23,820

$21,761

2

4

200

Acura

TSX 4dr

Sedan

Asia

Front

$26,990

$24,647

2.4

4

200

Acura

TL 4dr

Sedan

Asia

Front

$33,195

$30,299

3.2

6

270

Acura

3.5 RL 4dr

Sedan

Asia

Front

$43,755

$39,014

3.5

6

225

Acura

3.5 RL w/Navigation 4dr

Sedan

Asia

Front

$46,100

$41,100

3.5

6

225

Acura

NSX coupe 2dr manual S

Sports

Asia

Rear

$89,765

$79,978

3.2

6

290

Audi

A4 1.8T 4dr

Sedan

Europe

Front

$25,940

$23,508

1.8

4

170

Audi

A41.8T convertible 2dr

Sedan

Europe

Front

$35,940

$32,506

1.8

4

170

**Questions** 6:

Extract the part of the dataframe which contains cars belonging to 'usa'

**Ans**: df = pd.read\_csv(sales\_data, index\_col='month')

df