DAY-22

OCTOBER-08

LIBRARIES:

They are the collection of modules and functions to perform specific tasks.

Some of the libraries used in data field(data analyst):

1. pandas-data manipulation
2. numpy(numerical python)-multi dimensional arrays and mathematical operations.
3. matplot lib, seaborn, plotly- data visualization
4. statistics-for all statistical operations.

For data scientists:

1. sklearn(scikit learn)-machine learning
2. keras-deep learning and neural networks
3. tensorflow(google)-deep learning
4. pytorch(facebook)-deep learning

Pandas:

Import the libraries before using it.

Read the files:

1. d1=pd.read_excel(r"C:\Users\INDU PRIYA\OneDrive\Attachments\Desktop\MLP\emp_data.xlsx")

o/p:

	eid	ename	sal	dept
0	1	aa	1234	10
1	2	bb	2345	20
2	3	cc	3214	30
3	4	dd	5332	40

2. d2=pd.read_csv(r"C:\Users\INDU PRIYA\OneDrive\Attachments\Desktop\MLP\mtcars.csv")

o/p:

	manufacturer	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
5	Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
6	Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
7	Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
8	Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
9	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
10	Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
11	Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
12	Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
13	Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
14	Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
15	Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4

3. d3=pd.read_csv(r"C:\Users\INDU PRIYA\OneDrive\Attachments\Desktop\MLP\studata.txt")

	sid	sname	emialid	address	phoneno
0	1	priya	priya@gmail.com	dvg	9876
1	2	karthik	karthik@gmail.com	blr	12345
2	3	karan	NaN	msr	34567

4.d4=pd.read_csv(r"C:\Users\INDU PRIYA\OneDrive\Attachments\Desktop\MLP\chipotle.tsv",sep="\t")

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Chips and Fresh Tomato Salsa	NaN	$2.39
1	1	1	Izze	[Clementine]	$3.39
2	1	1	Nantucket Nectar	[Apple]	$3.39
3	1	1	Chips and Tomatillo-Green Chili Salsa	NaN	$2.39
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	$16.98
...	...	...	...	...	...
4617	1833	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Black Beans, Sour ...	$11.75
4618	1833	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Sour Cream, Cheese...	$11.75
4619	1834	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Pinto...	$11.25
4620	1834	1	Chicken Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	$8.75

- To display the contents of the file – d2.shape()
- To display only columns - d2.shape[0]
- To display first entries of the file(5 is the default entries)- d2.head()

	manufacturer	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2

- To display last entries of the file - d2.tail()

	manufacturer	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
27	Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.9	1	1	5	2
28	Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.5	0	1	5	4
29	Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.5	0	1	5	6
30	Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.6	0	1	5	8
31	Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.6	1	1	4	2

- To display the entries randomly-d2.sample(5)

	manufacturer	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
21	Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
25	Fiat X1-9	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
27	Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
20	Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2

- To display only columns – d2.columns()
- To display the data types – d2.dtypes

```
manufacturer     object
mpg             float64
cyl               int64
disp            float64
hp                int64
drat            float64
wt              float64
qsec            float64
vs                int64
am                int64
gear              int64
carb              int64
dtype: object
```

- how to check the null value:
  isna is used find null value. it show as true

```
manufacturer    0
mpg             0
cyl             0
disp            0
hp              0
drat            0
wt              0
qsec            0
vs              0
am              0
gear            0
carb            0
dtype: int64
```