In [1]:
```python
import numpy as np
import pandas as pd
```

In [2]:
```python
df = pd.DataFrame(np.arange(1,26))
```

In [3]:
```python
df
```

Out[3]:

	0
0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11
11	12
12	13
13	14
14	15
15	16
16	17
17	18
18	19
19	20
20	21
21	22
22	23
23	24
24	25

In [4]:
```python
df = pd.DataFrame(np.linspace(1, 20, num=50))
```

```
In [5]: df
```

Out[5]:

	0
0	1.000000
1	1.387755
2	1.775510
3	2.163265
4	2.551020
5	2.938776
6	3.326531
7	3.714286
8	4.102041
9	4.489796
10	4.877551
11	5.265306
12	5.653061
13	6.040816
14	6.428571
15	6.816327
16	7.204082
17	7.591837
18	7.979592
19	8.367347
20	8.755102
21	9.142857
22	9.530612
23	9.918367
24	10.306122
25	10.693878
26	11.081633
27	11.469388
28	11.857143
29	12.244898
30	12.632653
31	13.020408
32	13.408163
33	13.795918
34	14.183673

	0
35	14.571429
36	14.959184
37	15.346939
38	15.734694
39	16.122449
40	16.510204
41	16.897959
42	17.285714
43	17.673469
44	18.061224
45	18.448980
46	18.836735
47	19.224490
48	19.612245
49	20.000000

In [6]:
```python
my_dict = {
    'a':np.arange(1,11),
    'b':np.linspace(1,50, num=10),
    'c':np.ones(10),
    'd':np.zeros(10),
    'e':np.empty(10),
    'f':np.linspace(1,100, num=10),
    'g':np.arange(4,41,4)
}
```

In [7]:
```python
df = pd.DataFrame(my_dict)
```

In [8]:
```python
print(df)
```

```
    a          b    c    d             e      f   g
0   1   1.000000  1.0  0.0  3.952525e-323    1.0   4
1   2   6.444444  1.0  0.0   0.000000e+00   12.0   8
2   3  11.888889  1.0  0.0  7.862656e-312   23.0  12
3   4  17.333333  1.0  0.0  7.862656e-312   34.0  16
4   5  22.777778  1.0  0.0  7.862656e-312   45.0  20
5   6  28.222222  1.0  0.0  7.862656e-312   56.0  24
6   7  33.666667  1.0  0.0  7.862656e-312   67.0  28
7   8  39.111111  1.0  0.0  7.862656e-312   78.0  32
8   9  44.555556  1.0  0.0  7.862656e-312   89.0  36
9  10  50.000000  1.0  0.0  7.862679e-312  100.0  40
```

In [11]:
```python
# help(np.linspace(1,20,num=10))
```

In [12]:
```python
df = pd.DataFrame(np.array([[10,20,30],[40,50,60],[70,80,90]]), columns=['c1',
'c2','c3'])
```

In [13]:
```python
df
```

Out[13]:

	c1	c2	c3
0	10	20	30
1	40	50	60
2	70	80	90

In [14]:
```python
df.loc[1]
```

Out[14]:
```
c1     40
c2     50
c3     60
Name: 1, dtype: int32
```

In [15]:
```python
df.loc[0]
```

Out[15]:
```
c1     10
c2     20
c3     30
Name: 0, dtype: int32
```

In [17]:
```python
df['c1']
```

Out[17]:
```
0     10
1     40
2     70
Name: c1, dtype: int32
```

In [18]:
```python
df['c2']
```

Out[18]:
```
0     20
1     50
2     80
Name: c2, dtype: int32
```

In [19]:
```python
df['Total'] = df['c1'] + df['c2']
```

In [20]:
```python
df
```

Out[20]:

	c1	c2	c3	Total
0	10	20	30	30
1	40	50	60	90
2	70	80	90	150

In [21]:
```python
my_data = {
    'Name_of_Company':['TATA','MAHINDRA','TESLA','FORD','TOYATO','BMW'],
    'Unit_cars':[10000,20000,15000,5000,7000,4000],
    'Sales':[6000,12000,7500,2500,3500,2000],
    'Prod_Cost':[600000,600000,600000,600000,600000,600000],
    'Price':[800000,900000,1000000,800000,600000,1100000]
}
df = pd.DataFrame(my_data)
df
```

Out[21]:

	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price
0	TATA	10000	6000	600000	800000
1	MAHINDRA	20000	12000	600000	900000
2	TESLA	15000	7500	600000	1000000
3	FORD	5000	2500	600000	800000
4	TOYATO	7000	3500	600000	600000
5	BMW	4000	2000	600000	1100000

In [22]:
```python
df['Profit'] = df['Price'] - df['Prod_Cost']
```

In [23]:
```python
df
```

Out[23]:

	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit
0	TATA	10000	6000	600000	800000	200000
1	MAHINDRA	20000	12000	600000	900000	300000
2	TESLA	15000	7500	600000	1000000	400000
3	FORD	5000	2500	600000	800000	200000
4	TOYATO	7000	3500	600000	600000	0
5	BMW	4000	2000	600000	1100000	500000

In [25]:
```python
df['Percent']= df['Profit'] / df['Prod_Cost'] * 100
```

In [26]: `df`

Out[26]:

	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit	Percent
0	TATA	10000	6000	600000	800000	200000	33.333333
1	MAHINDRA	20000	12000	600000	900000	300000	50.000000
2	TESLA	15000	7500	600000	1000000	400000	66.666667
3	FORD	5000	2500	600000	800000	200000	33.333333
4	TOYATO	7000	3500	600000	600000	0	0.000000
5	BMW	4000	2000	600000	1100000	500000	83.333333

In [27]: `neq_df = df.to_csv('car_data.csv')`

In [28]: `df1 = pd.read_csv('car_data.csv')`

In [29]: `df1.head()`

Out[29]:

	Unnamed: 0	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit	Percent
0	0	TATA	10000	6000	600000	800000	200000	33.333333
1	1	MAHINDRA	20000	12000	600000	900000	300000	50.000000
2	2	TESLA	15000	7500	600000	1000000	400000	66.666667
3	3	FORD	5000	2500	600000	800000	200000	33.333333
4	4	TOYATO	7000	3500	600000	600000	0	0.000000

In [30]: `df1.tail()`

Out[30]:

	Unnamed: 0	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit	Percent
1	1	MAHINDRA	20000	12000	600000	900000	300000	50.000000
2	2	TESLA	15000	7500	600000	1000000	400000	66.666667
3	3	FORD	5000	2500	600000	800000	200000	33.333333
4	4	TOYATO	7000	3500	600000	600000	0	0.000000
5	5	BMW	4000	2000	600000	1100000	500000	83.333333

In [31]: `df1.head(2)`

Out[31]:

	Unnamed: 0	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit	Percent
0	0	TATA	10000	6000	600000	800000	200000	33.333333
1	1	MAHINDRA	20000	12000	600000	900000	300000	50.000000

In [34]: `new_json = df.to_json('car_data1.json')`

In [35]: ```python
df2 = pd.read_json('car_data1.json')
```

In [36]: ```python
df2
```

Out[36]:

	Name_of_Company	Unit_cars	Sales	Prod_Cost	Price	Profit	Percent
0	TATA	10000	6000	600000	800000	200000	33.333333
1	MAHINDRA	20000	12000	600000	900000	300000	50.000000
2	TESLA	15000	7500	600000	1000000	400000	66.666667
3	FORD	5000	2500	600000	800000	200000	33.333333
4	TOYATO	7000	3500	600000	600000	0	0.000000
5	BMW	4000	2000	600000	1100000	500000	83.333333

In [37]: ```python
new_html = df.to_html('car_data.html')
```

In [38]: ```python
df = pd.read_html('car_data.html')
```

In [39]: ```python
df
```

Out[39]:
```
[Unnamed: 0 Name_of_Company Unit_cars Sales Prod_Cost   Price Profit \
0   0   TATA   10000   6000   600000   800000 200000
1   1   MAHINDRA   20000 12000   600000   900000 300000
2   2   TESLA   15000   7500   600000 1000000 400000
3   3   FORD   5000   2500   600000   800000 200000
4   4   TOYATO   7000   3500   600000   600000   0
5   5   BMW   4000   2000   600000 1100000 500000

     Percent
0 33.333333
1 50.000000
2 66.666667
3 33.333333
4   0.000000
5 83.333333]
```

In [40]: ```python
df = pd.read_csv(r'C:/Users/shyam.DESKTOP-3EM50G2/Desktop/Data_Science_Batch_
V/Datasets/1_Regression/2_Multivariate_Linear_Regression/autos_dataset.csv')
```

In [41]: `df.head()`

Out[41]:

	symboling	normalized-losses	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base
0	3	?	alfa-romero	gas	std	two	convertible	rwd	front	88.6
1	3	?	alfa-romero	gas	std	two	convertible	rwd	front	88.6
2	1	?	alfa-romero	gas	std	two	hatchback	rwd	front	94.5
3	2	164	audi	gas	std	four	sedan	fwd	front	99.8
4	2	164	audi	gas	std	four	sedan	4wd	front	99.4

5 rows × 26 columns

In [42]: `df.tail()`

Out[42]:

	symboling	normalized-losses	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	
200	-1	95	volvo	gas	std	four	sedan	rwd	front	109.1	
201	-1	95	volvo	gas	turbo	four	sedan	rwd	front	109.1	
202	-1	95	volvo	gas	std	four	sedan	rwd	front	109.1	
203	-1	95	volvo	diesel	turbo	four	sedan	rwd	front	109.1	
204	-1	95	volvo	gas	turbo	four	sedan	rwd	front	109.1	

5 rows × 26 columns

In [43]: `df.shape`

Out[43]: (205, 26)

In [ ]: