Eesha Awais

22108255

PFAI Lab Task

```
[1]  import pandas as pd

[2]  import numpy as np

[3]  ser = pd.Series() #empty series

[5]  print('pandas series:', ser)

     pandas series: Series([], dtype: object)

[7]  data = np.array(['g', 'e', 'e', 'k', 's'])

[8]  ser = pd.Series(data)

[9]  print('pandas series:\n', ser) #series with data

     pandas series:
      0    g
     1    e
     2    e
     3    k
     4    s
     dtype: object

[11] df = pd.DataFrame()

[12] print(df)

     Empty DataFrame
     Columns: []
     Index: []

[13] list = ['geeks', 'for', 'geeks']
     df = pd.DataFrame(list)

[14] print(df)

            0
     0  geeks
     1    for
     2  geeks
```

```
[65] dict = {'Brand' : ['maruti', 'renault', 'hyundai', 'tata', 'toyota', 'maruti', 'hyundai', 'tata', 'toyota'],
             'Year' : [2012, 2014, 2011, 2012, 2014, 2011, 2012, 2014, 2011],
             'Kms' : [50000, 30000, 60000, 25000, 10000, 46000, 31000, 45000, 12000],
             'city' : ['gurgaon', 'delhi', 'mumbai', 'delhi', 'mumbai', 'delhi', 'mumbai', 'chennai', 'ghaziabad'],
             'mileage' : [28, 27, 25, 26, 27, 28, 24, 29, 25]}
```

```
[66] df = pd.DataFrame(dict)
```

```
display(df)
```

	Brand	Year	Kms	city	mileage
0	maruti	2012	50000	gurgaon	28
1	renault	2014	30000	delhi	27
2	hyundai	2011	60000	mumbai	25
3	tata	2012	25000	delhi	26
4	toyota	2014	10000	mumbai	27
5	maruti	2011	46000	delhi	28
6	hyundai	2012	31000	mumbai	24
7	tata	2014	45000	chennai	29
8	toyota	2011	12000	ghaziabad	25

```
[68] #loc : filter data based on some criteria
     display(df.loc[(df['Brand'] == 'maruti') &(df['mileage'] >25)])
```

	Brand	Year	Kms	city	mileage
0	maruti	2012	50000	gurgaon	28
5	maruti	2011	46000	delhi	28

```
[69] display(df.loc[2:5]) #from 2nd to 5th index, in iloc  till 5-1 = 4th is considered
```

	Brand	Year	Kms	city	mileage
2	hyundai	2011	60000	mumbai	25
3	tata	2012	25000	delhi	26
4	toyota	2014	10000	mumbai	27
5	maruti	2011	46000	delhi	28

```
[89] df.loc[df['Year'] == 2012, 'mileage'] = 20
```

```
df
```

	Brand	Year	Kms	city	mileage
0	maruti	2012	50000	gurgaon	20
1	renault	2014	30000	delhi	27
2	hyundai	2011	60000	mumbai	25
3	tata	2012	25000	delhi	20
4	toyota	2014	10000	mumbai	27
5	maruti	2011	46000	delhi	28
6	hyundai	2012	31000	mumbai	20
7	tata	2014	45000	chennai	29
8	toyota	2011	12000	ghaziabad	25

Next steps: Generate code with df View recommended p

```
[91] df.loc[df['Kms'] >15000, 'city'] = 'delhi'
```

```
[92] df
```

	Brand	Year	Kms	city	mileage
0	maruti	2012	50000	delhi	20
1	renault	2014	30000	delhi	27
2	hyundai	2011	60000	delhi	25
3	tata	2012	25000	delhi	20
4	toyota	2014	10000	mumbai	27
5	maruti	2011	46000	delhi	28
6	hyundai	2012	31000	delhi	20
7	tata	2014	45000	delhi	29
8	toyota	2011	12000	ghaziabad	25

```
[94] #iloc, known indexes
     display(df.iloc[[0,2 ,4,7]])
```

	Brand	Year	Kms	city	mileage
0	maruti	2012	50000	delhi	20
2	hyundai	2011	60000	delhi	25
4	toyota	2014	10000	mumbai	27
7	tata	2014	45000	delhi	29

```
display(df.iloc[ 1 : 5, 2:5]) #[rows, columns]
```

	Kms	city	mileage
1	30000	delhi	27
2	60000	delhi	25
3	25000	delhi	20
4	10000	mumbai	27

```
[101] import pandas as pd
      import numpy as np
```

```
[114] df = pd.read_csv("cars.csv")
```

```
df.head()
```

	type	price	mpg_city	drive_train	passengers	weight
0	small	15.9	25	front	5	2705
1	midsize	33.9	18	front	5	3560
2	midsize	37.7	19	front	6	3405
3	midsize	30.0	22	rear	4	3640
4	midsize	15.7	22	front	6	2880

Convert this

Next steps: Generate code with df View recommended plots

```
[116] df.loc[df['mpg_city'] > 20, 'price'] = 10000
```

```
display(df.loc[(df['price'] > 30) & (df['passengers'] ==6)]) #and operator
```

	type	price	mpg_city	drive_train	passengers	weight
2	midsize	37.7	19	front	6	3405
4	midsize	10000.0	22	front	6	2880
8	large	34.7	16	front	6	3620
10	midsize	10000.0	21	front	6	3195
16	midsize	10000.0	21	front	6	3080
31	midsize	34.3	17	front	6	3695
32	large	36.1	18	rear	6	4055

```
df.loc[df['drive_train'] == 'rear', 'type'] = 'midsize' #update
```

```
[120] df.head()
```

	type	price	mpg_city	drive_train	passengers	weight
0	small	10000.0	25	front	5	2705
1	midsize	33.9	18	front	5	3560
2	midsize	37.7	19	front	6	3405
3	midsize	10000.0	22	rear	4	3640
4	midsize	10000.0	22	front	6	2880

Next steps: Generate code with df View recommended plots