

# Heamnath

## 20104028

## Basic Analysis using Numpy and Pandas

### Importing libraries

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

### importing datasets

```
In [2]: df=pd.read_csv("fiat.csv")
df
```

```
Out[2]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.49565029
...	...	...	...	...	...	...	...	...
1544	NaN	NaN	NaN	NaN	NaN	NaN	NaN	length
1545	NaN	NaN	NaN	NaN	NaN	NaN	NaN	concat
1546	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Null values
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	find
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	search

1549 rows × 11 columns

### To display first 10 rows

```
In [3]: df.head(10)
```

```
Out[3]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868	890
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995	880
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784	420
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922	600
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.49565029	570
5	6.0	pop	74.0	3623.0	70225.0	1.0	45.000702	7.68227005	790
6	7.0	lounge	51.0	731.0	11600.0	1.0	44.907242	8.611559868	1075
7	8.0	lounge	51.0	1521.0	49076.0	1.0	41.903221	12.49565029	919
8	9.0	sport	73.0	4049.0	76000.0	1.0	45.548000	11.54946995	560
9	10.0	sport	51.0	2652.0	80000.0	1.0	45.428201	10.99170017	600

To display last 5 rows

```
In [4]: df.tail(5)
```

```
Out[4]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnan
1544	NaN	NaN	NaN	NaN	NaN	NaN	NaN	length	5	I
1545	NaN	NaN	NaN	NaN	NaN	NaN	NaN	concat	lonprice	I
1546	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Null values	NO	I
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	find	1	I
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	search	1	I

Statistical Summary

```
In [5]: df.describe()
```

```
Out[5]:
```

	ID	engine_power	age_in_days	km	previous_owners	lat	Unnan
count	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	
mean	769.500000	51.904421	1650.980494	53396.011704	1.123537	43.541361	↑
std	444.126671	3.988023	1289.522278	40046.830723	0.416423	2.133518	↑
min	1.000000	51.000000	366.000000	1232.000000	1.000000	36.855839	↑
25%	385.250000	51.000000	670.000000	20006.250000	1.000000	41.802990	↑
50%	769.500000	51.000000	1035.000000	39031.000000	1.000000	44.394096	↑

	ID	engine_power	age_in_days	km	previous_owners	lat	Unnamed: 0
75%	1153.750000	51.000000	3616.000000	70667.750000	1.000000	45.467060	1

## To print no of rows and columns

```
In [6]: df.shape
```

```
Out[6]: (1549, 11)
```

## To print total no of elements

```
In [7]: df.size
```

```
Out[7]: 17039
```

## To find the null value

```
In [8]: df.isna()
```

```
Out[8]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 0
0	False	False	False	False	False	False	False	False	False	True
1	False	False	False	False	False	False	False	False	False	True
2	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	True
4	False	False	False	False	False	False	False	False	False	True
...	...	...	...	...	...	...	...	...	...	...
1544	True	True	True	True	True	True	True	False	False	True
1545	True	True	True	True	True	True	True	False	False	True
1546	True	True	True	True	True	True	True	False	False	True
1547	True	True	True	True	True	True	True	False	False	True
1548	True	True	True	True	True	True	True	False	False	True

1549 rows × 11 columns

## To fill the missing value

```
In [9]: df.fillna(value=0)
```

Out[9]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.49565029
...	...	...	...	...	...	...	...	...
1544	0.0	0	0.0	0.0	0.0	0.0	0.000000	length
1545	0.0	0	0.0	0.0	0.0	0.0	0.000000	concat
1546	0.0	0	0.0	0.0	0.0	0.0	0.000000	Null values
1547	0.0	0	0.0	0.0	0.0	0.0	0.000000	find
1548	0.0	0	0.0	0.0	0.0	0.0	0.000000	search

1549 rows × 11 columns