

# Deena

20104016

## libraries

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

## importing data set

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

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lo
0	1	lounge	51	882	25000	1	44.907242	8.61156
1	2	pop	51	1186	32500	1	45.666359	12.24189
2	3	sport	74	4658	142228	1	45.503300	11.41784
3	4	lounge	51	2739	160000	1	40.633171	17.63460
4	5	pop	73	3074	106880	1	41.903221	12.49565
...	...	...	...	...	...	...	...	.
1533	1534	sport	51	3712	115280	1	45.069679	7.70492
1534	1535	lounge	74	3835	112000	1	45.845692	8.66687
1535	1536	pop	51	2223	60457	1	45.481541	9.41348
1536	1537	lounge	51	2557	80750	1	45.000702	7.68227
1537	1538	pop	51	1766	54276	1	40.323410	17.56827

mean median mode()

```
In [3]: df.mean()
```

```
Out[3]: ID                769.500000
engine_power            51.904421
age_in_days            1650.980494
km                   53396.011704
previous_owners         1.123537
lat                   43.541361
lon                   11.563428
price                 8576.003901
dtype: float64
```

```
In [4]: df.median()
```

```
Out[4]: ID                769.500000
engine_power            51.000000
age_in_days            1035.000000
km                   39031.000000
previous_owners         1.000000
lat                   44.394096
lon                   11.869260
price                 9000.000000
dtype: float64
```

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

```
Out[5]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
0	1	lounge	51.0	366.0	17000.0	1.0	41.903221	12.49565
1	2	NaN	NaN	790.0	NaN	NaN	NaN	NaN
2	3	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	4	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	5	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...
1533	1534	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1534	1535	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1535	1536	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1536	1537	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1537	1538	NaN	NaN	NaN	NaN	NaN	NaN	NaN

1538 rows × 9 columns



## describe ()

```
In [6]: df.describe()
```

Out[6]:

	ID	engine_power	age_in_days	km	previous_owners	lat	
count	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	1
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	
75%	1153.750000	51.000000	2616.000000	79667.750000	1.000000	45.467960	
max	1538.000000	77.000000	4658.000000	235000.000000	4.000000	46.795612	

sum()

```
In [7]: df.sum()
```

Out[7]:

ID	1183491
model	loungepopsportloungepoppoploungeloungesportspo...
engine_power	79829
age_in_days	2539208
km	82123066
previous_owners	1728
lat	66966.61372
lon	17784.55279
price	13189894
dtype: object	

cumsum ()

```
In [8]: a=df.head(10)
a.cumsum()
```

Out[8]:

	ID	model	engine_power	age_in_days	km	previ
0	1	lounge	51	882	25000	
1	3	loungepop	102	2068	57500	
2	6	loungepopsport	176	6726	199728	
3	10	loungepopsportlounge	227	9465	359728	
4	15	loungepopsportloungepop	300	12539	466608	
5	21	loungepopsportloungepoppop	374	16162	536833	
6	28	loungepopsportloungepoppoplounge	425	16893	548433	
7	36	loungepopsportloungepoppoplounge	476	18414	597509	
8	45	loungepopsportloungepoppoplounge	549	22463	673509	
9	55	loungepopsportloungepoppoplounge	600	26116	762509	

## min() and min()

```
In [9]: df.min()
```

```
Out[9]: ID          1
model        lounge
engine_power    51
age_in_days    366
km           1232
previous_owners    1
lat        36.855839
lon         7.2454
price        2500
dtype: object
```

```
In [10]: df.max()
```

```
Out[10]: ID          1538
model        sport
engine_power    77
age_in_days    4658
km          235000
previous_owners    4
lat        46.795612
lon        18.36552
price        11100
dtype: object
```

## count()

```
In [11]: df.count()
```

```
Out[11]: ID                1538  
         model             1538  
         engine_power      1538  
         age_in_days        1538  
         km                1538  
         previous_owners    1538  
         lat               1538  
         lon               1538  
         price             1538  
         dtype: int64
```

## Covariance

```
In [12]: from numpy import cov
```

```
In [13]: cov(df['engine_power'],df['age_in_days'])
```

```
Out[13]: array([[1.59043266e+01, 1.64148089e+03],  
                [1.64148089e+03, 1.66286770e+06]])
```

## pearsonr and spearmanr

```
In [14]: from scipy.stats import pearsonr  
         from scipy.stats import spearmanr
```

```
In [15]: pearsonr(a['engine_power'],a['age_in_days'])
```

```
Out[15]: (0.7507941685043964, 0.012332990872150272)
```

```
In [16]: spearmanr(df['engine_power'],df['age_in_days'])
```

```
Out[16]: SpearmanrResult(correlation=0.27353999528180917, pvalue=8.481828289120299e-2  
8)
```