In [6]:

```
import numpy as np
import pandas as pd
from numpy import mean,std
import matplotlib.pyplot as pp
from numpy import cov
from scipy.stats import pearsonr
from scipy.stats import spearmanr
```

In [8]:

```
a=pd.read_csv(r"C:\Users\user\Downloads\fiat500_VehicleSelection_Dataset (1).csv")
a
```

Out[8]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	l
0	1	lounge	51	882	25000	1	44.907242	8.6115
1	2	pop	51	1186	32500	1	45.666359	12.2418
2	3	sport	74	4658	142228	1	45.503300	11.4178
3	4	lounge	51	2739	160000	1	40.633171	17.6346
4	5	pop	73	3074	106880	1	41.903221	12.4956
...	...	...	...	...	...	...	...	
1533	1534	sport	51	3712	115280	1	45.069679	7.7049
1534	1535	lounge	74	3835	112000	1	45.845692	8.6668
1535	1536	pop	51	2223	60457	1	45.481541	9.4134
1536	1537	lounge	51	2557	80750	1	45.000702	7.6822
1537	1538	pop	51	1766	54276	1	40.323410	17.5682

1538 rows × 9 columns

a) Find mean, median, mode and describe

In [5]:

```
a.mean()
```

Out[5]:

```
Unnamed: 0           2.771250e+07
fare_amount          1.135996e+01
pickup_longitude    -7.252764e+01
pickup_latitude      3.993589e+01
dropoff_longitude   -7.252529e+01
dropoff_latitude     3.992389e+01
passenger_count      1.684535e+00
dtype: float64
```

In [9]:

```
y.median()
```

Out[9]:

```
Unnamed: 0           2.774550e+07
fare_amount          8.500000e+00
pickup_longitude    -7.398182e+01
pickup_latitude      4.075259e+01
dropoff_longitude   -7.398009e+01
dropoff_latitude     4.075304e+01
passenger_count      1.000000e+00
dtype: float64
```

In [10]:

```
a.mode()
```

Out[10]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lo
0	1	lounge	51.0	366.0	17000.0	1.0	41.903221	12.4956
1	2	NaN	NaN	790.0	NaN	NaN	NaN	Na
2	3	NaN	NaN	NaN	NaN	NaN	NaN	Na
3	4	NaN	NaN	NaN	NaN	NaN	NaN	Na
4	5	NaN	NaN	NaN	NaN	NaN	NaN	Na
...	...	...	...	...	...	...	...	
1533	1534	NaN	NaN	NaN	NaN	NaN	NaN	Na
1534	1535	NaN	NaN	NaN	NaN	NaN	NaN	Na
1535	1536	NaN	NaN	NaN	NaN	NaN	NaN	Na
1536	1537	NaN	NaN	NaN	NaN	NaN	NaN	Na
1537	1538	NaN	NaN	NaN	NaN	NaN	NaN	Na

1538 rows × 9 columns

In [11]:

```
a.describe()
```

Out[11]:

	ID	engine_power	age_in_days	km	previous_owners	la
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
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

b) Find sum(), cumsum(), count, min and max values

In [12]:

```
a.sum()
```

Out[12]:

```
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
```

In [13]:

```
a.cumsum()
```

Out[13]:

	ID	model	engine_power	age_in_days	
0	1	lounge	51	882	
1	3	loungepop	102	2068	
2	6	loungepopsport	176	6726	
3	10	loungepopsportlounge	227	9465	
4	15	loungepopsportloungepop	300	12539	
...	...	...	...	...	
1533	1177345	loungepopsportloungepoppoploungeloungesportspo...	79602	2528827	8
1534	1178880	loungepopsportloungepoppoploungeloungesportspo...	79676	2532662	8
1535	1180416	loungepopsportloungepoppoploungeloungesportspo...	79727	2534885	8
1536	1181953	loungepopsportloungepoppoploungeloungesportspo...	79778	2537442	8
1537	1183491	loungepopsportloungepoppoploungeloungesportspo...	79829	2539208	8

1538 rows × 9 columns

In [14]:

```
a.count()
```

Out[14]:

```
ID                1538
model             1538
engine_power      1538
age_in_days       1538
km                1538
previous_owners   1538
lat               1538
lon               1538
price             1538
dtype: int64
```

In [15]:

```
y.count()
```

Out[15]:

```
Unnamed: 0           200000
key                  200000
fare_amount          200000
pickup_datetime      200000
pickup_longitude     200000
pickup_latitude      200000
dropoff_longitude    199999
dropoff_latitude     199999
passenger_count      200000
dtype: int64
```

In [16]:

```
a.max()
```

Out[16]:

```
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
```

c) Find covariance and correlation (spearman and pearsons)

In [18]:

```
d1=a["engine_power"]
d2=a["age_in_days"]
cov(d1,d2)
```

Out[18]:

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

In [19]:

```
pearsonr(d1,d2)
```

Out[19]:

```
(0.3191900466644252, 9.164638156450997e-38)
```

In [20]:

```
spearmanr(d1,d2)
```

Out[20]:

```
SpearmanrResult(correlation=0.27353999528180917, pvalue=8.481828289120299e
-28)
```

In [ ]:

```

```