Importing libraries

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

In [3]:
```
df=pd.read_csv(r'C:\Users\user\Desktop\fiat500_VehicleSelection_Dataset (1).csv
```

Find mean, median, mode and describe

In [4]:
```
print(df.mean())
```

```
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 [5]:
```
print(df.median())
```

```
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 [6]:
```
print(df.mode())
```

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

            lat       lon    price
0     41.903221  12.49565  10500.0
1           NaN       NaN      NaN
2           NaN       NaN      NaN
3           NaN       NaN      NaN
4           NaN       NaN      NaN
...         ...       ...      ...
1533        NaN       NaN      NaN
1534        NaN       NaN      NaN
1535        NaN       NaN      NaN
1536        NaN       NaN      NaN
1537        NaN       NaN      NaN

[1538 rows x 9 columns]
```

In [7]:
```
print(df.describe())
```

```
                ID  engine_power  age_in_days             km  previous_owners
\
count  1538.000000   1538.000000  1538.000000    1538.000000      1538.000000
mean    769.500000     51.904421  1650.980494   53396.011704         1.123537
std     444.126671      3.988023  1289.522278   40046.830723         0.416423
min       1.000000     51.000000   366.000000    1232.000000         1.000000
25%     385.250000     51.000000   670.000000   20006.250000         1.000000
50%     769.500000     51.000000  1035.000000   39031.000000         1.000000
75%    1153.750000     51.000000  2616.000000   79667.750000         1.000000
max    1538.000000     77.000000  4658.000000  235000.000000         4.000000

               lat          lon         price
count  1538.000000  1538.000000   1538.000000
mean     43.541361    11.563428   8576.003901
std       2.133518     2.328190   1939.958641
min      36.855839     7.245400   2500.000000
25%      41.802990     9.505090   7122.500000
50%      44.394096    11.869260   9000.000000
75%      45.467960    12.769040  10000.000000
max      46.795612    18.365520  11100.000000
```

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

In [8]:
```
print(df.sum())
```

```
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 [9]:
```
print(df.cumsum())
```

```
           ID                                                model  \
0           1                                               lounge
1           3                                            loungepop
2           6                                       loungepopsport
3          10                                  loungepopsportlounge
4          15                             loungepopsportloungepop
...       ...                                                  ...
1533  1177345  loungepopsportloungepoppoploungeloungesportspo...
1534  1178880  loungepopsportloungepoppoploungeloungesportspo...
1535  1180416  loungepopsportloungepoppoploungeloungesportspo...
1536  1181953  loungepopsportloungepoppoploungeloungesportspo...
1537  1183491  loungepopsportloungepoppoploungeloungesportspo...

      engine_power  age_in_days        km  previous_owners           lat  \
0               51          882     25000                1     44.907242
1              102         2068     57500                2     90.573601
2              176         6726    199728                3    136.076900
3              227         9465    359728                4    176.710072
4              300        12539    466608                5    218.613293
...            ...          ...       ...              ...           ...
1533         79602      2528827  81815583             1724  66789.962376
1534         79676      2532662  81927583             1725  66835.808068
1535         79727      2534885  81988040             1726  66881.289608
1536         79778      2537442  82068790             1727  66926.290310
1537         79829      2539208  82123066             1728  66966.613720

                lon     price
0         8.611560      8900
1        20.853450     17700
2        32.271290     21900
3        49.905899     27900
4        62.401549     33600
...            ...       ...
1533  17741.221900  13163904
1534  17749.888771  13168504
1535  17759.302250  13176004
1536  17766.984520  13181994
1537  17784.552790  13189894

[1538 rows x 9 columns]
```

In [10]:
```
print(df.count())
```

```
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 [11]:
```
print(df.min())
```

```
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 [12]:
```
print(df.max())
```

```
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 [13]:
```
from numpy import cov
```

In [20]:
```
cov(df['km'],df['ID'])
```

Out[20]:
```
array([[ 1.60374865e+09, -1.16261338e+05],
       [-1.16261338e+05,  1.97248500e+05]])
```

In [22]:
```
from scipy.stats import pearsonr
from scipy.stats import spearmanr
```

In [23]:
```
spearmanr(df['km'],df['ID'])
```

Out[23]:
```
SpearmanrResult(correlation=0.024271367253139893, pvalue=0.3414902951355948)
```

In [ ]:
```

```