

Basic Analysis using numpy and pandas

vehicle dataset

to import library

```
In [2]: import numpy as np
import pandas as pd
```

To import dataset

```
In [5]: d=pd.read_csv(r"C:\Users\user\Downloads\fiat500_VehicleSelection_Dataset (3).csv")
d
```

Out[5]:

| | 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 | 8.6115598 |
| 1 | 2.0 | pop | 51.0 | 1186.0 | 32500.0 | 1.0 | 45.666359 | 12.241889 |
| 2 | 3.0 | sport | 74.0 | 4658.0 | 142228.0 | 1.0 | 45.503300 | 11.417 |
| 3 | 4.0 | lounge | 51.0 | 2739.0 | 160000.0 | 1.0 | 40.633171 | 17.634609 |
| 4 | 5.0 | pop | 73.0 | 3074.0 | 106880.0 | 1.0 | 41.903221 | 12.495650 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1544 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | len... |
| 1545 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | con... |
| 1546 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | Null val... |
| 1547 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | fi... |
| 1548 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | sear... |

1549 rows × 11 columns



In [6]:

d.head(10)

Out[6]:

| | 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 |
| 5 | 6.0 | pop | 74.0 | 3623.0 | 70225.0 | 1.0 | 45.000702 | 7.68227005 |
| 6 | 7.0 | lounge | 51.0 | 731.0 | 11600.0 | 1.0 | 44.907242 | 8.611559868 |
| 7 | 8.0 | lounge | 51.0 | 1521.0 | 49076.0 | 1.0 | 41.903221 | 12.49565029 |
| 8 | 9.0 | sport | 73.0 | 4049.0 | 76000.0 | 1.0 | 45.548000 | 11.54946995 |
| 9 | 10.0 | sport | 51.0 | 3653.0 | 89000.0 | 1.0 | 45.438301 | 10.99170017 |

In [7]:

d.tail(10)

Out[7]:

| | ID | model | engine_power | age_in_days | km | previous_owners | lat | lon | price |
|------|-----|-------|--------------|-------------|-----|-----------------|-----|-----------------------|-------------|
| 1539 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | avg | 8576.003901 |
| 1540 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | count | 1538 |
| 1541 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | countif | 371 |
| 1542 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | sumif | 4017825 |
| 1543 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | counta (not empty) | 1538 |
| 1544 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | length | 5 |
| 1545 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | concat | lonprice |
| 1546 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | Null values | NC |
| 1547 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | find | 1 |
| 1548 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | search | 1 |

In [8]: `d.describe()`

Out[8]:

| | ID | engine_power | age_in_days | km | previous_owners | lat | lon |
|--------------|-------------|--------------|-------------|---------------|-----------------|-------------|-------------|
| count | 1538.000000 | 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 | 10.834262 |
| std | 444.126671 | 3.988023 | 1289.522278 | 40046.830723 | 0.416423 | 2.133518 | 1.629921 |
| min | 1.000000 | 51.000000 | 366.000000 | 1232.000000 | 1.000000 | 36.855839 | 10.834262 |
| 25% | 385.250000 | 51.000000 | 670.000000 | 20006.250000 | 1.000000 | 41.802990 | 10.834262 |
| 50% | 769.500000 | 51.000000 | 1035.000000 | 39031.000000 | 1.000000 | 44.394096 | 10.834262 |
| 75% | 1153.750000 | 51.000000 | 2616.000000 | 79667.750000 | 1.000000 | 45.467960 | 10.834262 |
| max | 1538.000000 | 77.000000 | 4658.000000 | 235000.000000 | 4.000000 | 46.795612 | 10.834262 |

In [9]: `np.shape(d)`

Out[9]: (1549, 11)

To get the missing value

In [10]: `d.isna()`

Out[10]:

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

1549 rows × 11 columns

```
In [11]: d.fillna(value=5)
```

Out[11]:

| | ID | model | engine_power | age_in_days | km | previous_owners | lat | lo |
|------|-----|--------|--------------|-------------|----------|-----------------|-----------|------------|
| 0 | 1.0 | lounge | 51.0 | 882.0 | 25000.0 | 1.0 | 44.907242 | 8.61155986 |
| 1 | 2.0 | pop | 51.0 | 1186.0 | 32500.0 | 1.0 | 45.666359 | 12.2418899 |
| 2 | 3.0 | sport | 74.0 | 4658.0 | 142228.0 | 1.0 | 45.503300 | 11.4178 |
| 3 | 4.0 | lounge | 51.0 | 2739.0 | 160000.0 | 1.0 | 40.633171 | 17.6346092 |
| 4 | 5.0 | pop | 73.0 | 3074.0 | 106880.0 | 1.0 | 41.903221 | 12.4956502 |
| ... | ... | ... | ... | ... | ... | ... | ... | . |
| 1544 | 5.0 | 5 | 5.0 | 5.0 | 5.0 | 5.0 | 5.000000 | length |
| 1545 | 5.0 | 5 | 5.0 | 5.0 | 5.0 | 5.0 | 5.000000 | conce |
| 1546 | 5.0 | 5 | 5.0 | 5.0 | 5.0 | 5.0 | 5.000000 | Null value |
| 1547 | 5.0 | 5 | 5.0 | 5.0 | 5.0 | 5.0 | 5.000000 | fin |
| 1548 | 5.0 | 5 | 5.0 | 5.0 | 5.0 | 5.0 | 5.000000 | searc |

1549 rows × 11 columns



```
In [ ]:
```