In [1]:

```
import pandas as pd
```

In [4]:

```
data=pd.read_csv("/home/palcement/Downloads/fiat500.csv")
```

In [5]:

```
data.describe()
```

Out[5]:

	ID	engine_power	age_in_days	km	previous_owners	lat
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

In [6]:

```
data.head()
```

Out[6]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	pri
0	1	lounge	51	882	25000	1	44.907242	8.611560	89(
1	2	pop	51	1186	32500	1	45.666359	12.241890	88(
2	3	sport	74	4658	142228	1	45.503300	11.417840	42(
3	4	lounge	51	2739	160000	1	40.633171	17.634609	60(
4	5	pop	73	3074	106880	1	41.903221	12.495650	57(

In [7]:

```
data.tail()
```

Out[7]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
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

In [8]:

```
data.head(10)
```

Out[8]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	pr
0	1	lounge	51	882	25000	1	44.907242	8.611560	8
1	2	pop	51	1186	32500	1	45.666359	12.241890	8
2	3	sport	74	4658	142228	1	45.503300	11.417840	4
3	4	lounge	51	2739	160000	1	40.633171	17.634609	6
4	5	pop	73	3074	106880	1	41.903221	12.495650	5
5	6	pop	74	3623	70225	1	45.000702	7.682270	7
6	7	lounge	51	731	11600	1	44.907242	8.611560	10
7	8	lounge	51	1521	49076	1	41.903221	12.495650	9
8	9	sport	73	4049	76000	1	45.548000	11.549470	5
9	10	sport	51	3653	89000	1	45.438301	10.991700	6

In [9]:

```
data ['previous_owners'].unique()
```

Out[9]:

```
array([1, 2, 3, 4])
```

In [10]:

```
list(data.columns)
```

Out[10]:

```
['ID',
 'model',
 'engine_power',
 'age_in_days',
 'km',
 'previous_owners',
 'lat',
 'lon',
 'price']
```

In [12]:

```
data.groupby(['model']).count()
```

Out[12]:

model	ID	engine_power	age_in_days	km	previous_owners	lat	lon	price
lounge	1094	1094	1094	1094	1094	1094	1094	1094
pop	358	358	358	358	358	358	358	358
sport	86	86	86	86	86	86	86	86

In [14]:

```
data.groupby(['previous_owners']).count()
```

Out[14]:

previous_owners	ID	model	engine_power	age_in_days	km	lat	lon	price
1	1389	1389	1389	1389	1389	1389	1389	1389
2	117	117	117	117	117	117	117	117
3	23	23	23	23	23	23	23	23
4	9	9	9	9	9	9	9	9

In [15]:

```
data.groupby(['model']).count()
```

Out[15]:

model	ID	engine_power	age_in_days	km	previous_owners	lat	lon	price
lounge	1094	1094	1094	1094	1094	1094	1094	1094
pop	358	358	358	358	358	358	358	358
sport	86	86	86	86	86	86	86	86

In [17]:

```
data1=data.drop(['lat','ID'],axis=1)
```

In [18]:

```
data1.head()
```

Out[18]:

	model	engine_power	age_in_days	km	previous_owners	lon	price
0	lounge	51	882	25000	1	8.611560	8900
1	pop	51	1186	32500	1	12.241890	8800
2	sport	74	4658	142228	1	11.417840	4200
3	lounge	51	2739	160000	1	17.634609	6000
4	pop	73	3074	106880	1	12.495650	5700

In [19]:

```
data['price'].sum()
```

Out[19]:

13189894

In [23]:

```
data2=data.loc[(data.model=='lounge')]
data2
```

Out[23]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lo
0	1	lounge	51	882	25000	1	44.907242	8.61156
3	4	lounge	51	2739	160000	1	40.633171	17.63460
6	7	lounge	51	731	11600	1	44.907242	8.61156
7	8	lounge	51	1521	49076	1	41.903221	12.49565
11	12	lounge	51	366	17500	1	45.069679	7.70492
...	...	...	...	...	...	...	...	.
1528	1529	lounge	51	2861	126000	1	43.841980	10.51531
1529	1530	lounge	51	731	22551	1	38.122070	13.36112
1530	1531	lounge	51	670	29000	1	45.764648	8.99450
1534	1535	lounge	74	3835	112000	1	45.845692	8.66687
1536	1537	lounge	51	2557	80750	1	45.000702	7.68227

1094 rows × 9 columns

In [24]:

```
data3=data.loc[(data.km<=20000)]
data3
```

Out[24]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
6	7	lounge	51	731	11600	1	44.907242	8.61156
11	12	lounge	51	366	17500	1	45.069679	7.70492
12	13	lounge	51	456	18450	1	45.426571	11.78813
20	21	lounge	51	397	19037	1	45.707249	11.47760
22	23	lounge	51	1035	8000	1	44.506088	12.04417
...	...	...	...	...	...	...	...	...
1517	1518	pop	51	366	16100	1	44.692520	10.10396
1518	1519	lounge	51	397	16053	1	38.122070	13.36112
1520	1521	lounge	51	1035	15000	1	41.903221	12.49565
1522	1523	lounge	51	366	14618	1	45.707249	11.47760
1527	1528	pop	51	517	3000	1	40.748241	14.52835

385 rows × 9 columns

In [25]:

```
data2=data.loc[(data.model=='pop')]
data2
```

Out[25]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lor
1	2	pop	51	1186	32500	1	45.666359	12.24189(
4	5	pop	73	3074	106880	1	41.903221	12.49565(
5	6	pop	74	3623	70225	1	45.000702	7.68227(
10	11	pop	51	790	43286	1	40.871429	14.43896(
13	14	pop	51	3835	120000	1	40.531590	17.43615(
...	...	...	...	...	...	...	...	..
1524	1525	pop	51	2192	53300	1	40.609531	14.98093(
1527	1528	pop	51	517	3000	1	40.748241	14.52835(
1532	1533	pop	51	1917	52008	1	45.548000	11.54947(
1535	1536	pop	51	2223	60457	1	45.481541	9.41348(
1537	1538	pop	51	1766	54276	1	40.323410	17.56827(

358 rows × 9 columns

In [26]:

```
data3=data.loc[(data.model=='pop')]
data3
```

Out[26]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lor
1	2	pop	51	1186	32500	1	45.666359	12.24189(
4	5	pop	73	3074	106880	1	41.903221	12.49565(
5	6	pop	74	3623	70225	1	45.000702	7.68227(
10	11	pop	51	790	43286	1	40.871429	14.43896(
13	14	pop	51	3835	120000	1	40.531590	17.43615(
...	...	...	...	...	...	...	...	..
1524	1525	pop	51	2192	53300	1	40.609531	14.98093(
1527	1528	pop	51	517	3000	1	40.748241	14.52835(
1532	1533	pop	51	1917	52008	1	45.548000	11.54947(
1535	1536	pop	51	2223	60457	1	45.481541	9.41348(
1537	1538	pop	51	1766	54276	1	40.323410	17.56827(

358 rows × 9 columns