

Importing pandas and numpy

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

Reading the data

```
In [2]: kcd=pd.read_csv("/home/placement/Downloads/fiat500.csv")  
kcd.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1538 entries, 0 to 1537  
Data columns (total 9 columns):  
#   Column                Non-Null Count  Dtype  
---  -  
0   ID                    1538 non-null   int64  
1   model                 1538 non-null   object  
2   engine_power          1538 non-null   int64  
3   age_in_days           1538 non-null   int64  
4   km                    1538 non-null   int64  
5   previous_owners       1538 non-null   int64  
6   lat                   1538 non-null   float64  
7   lon                   1538 non-null   float64  
8   price                 1538 non-null   int64  
dtypes: float64(2), int64(6), object(1)  
memory usage: 108.3+ KB
```

```
In [3]: kcd.head(10)
```

```
Out[3]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1	lounge	51	882	25000	1	44.907242	8.611560	8900
1	2	pop	51	1186	32500	1	45.666359	12.241890	8800
2	3	sport	74	4658	142228	1	45.503300	11.417840	4200
3	4	lounge	51	2739	160000	1	40.633171	17.634609	6000
4	5	pop	73	3074	106880	1	41.903221	12.495650	5700
5	6	pop	74	3623	70225	1	45.000702	7.682270	7900
6	7	lounge	51	731	11600	1	44.907242	8.611560	10750
7	8	lounge	51	1521	49076	1	41.903221	12.495650	9190
8	9	sport	73	4049	76000	1	45.548000	11.549470	5600
9	10	sport	51	3653	89000	1	45.438301	10.991700	6000

```
In [4]: list(kcd)
```

```
Out[4]: ['ID',  
         'model',  
         'engine_power',  
         'age_in_days',  
         'km',  
         'previous_owners',  
         'lat',  
         'lon',  
         'price']
```

Removing model column from table

In [5]: `k=kcd.drop('model',axis=1)`
`k`

Out[5]:

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1	51	882	25000	1	44.907242	8.611560	8900
1	2	51	1186	32500	1	45.666359	12.241890	8800
2	3	74	4658	142228	1	45.503300	11.417840	4200
3	4	51	2739	160000	1	40.633171	17.634609	6000
4	5	73	3074	106880	1	41.903221	12.495650	5700
...
1533	1534	51	3712	115280	1	45.069679	7.704920	5200
1534	1535	74	3835	112000	1	45.845692	8.666870	4600
1535	1536	51	2223	60457	1	45.481541	9.413480	7500
1536	1537	51	2557	80750	1	45.000702	7.682270	5990
1537	1538	51	1766	54276	1	40.323410	17.568270	7900

1538 rows × 8 columns

```
In [6]: kcd.isnull()
```

Out[6]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...
1533	False	False	False	False	False	False	False	False	False
1534	False	False	False	False	False	False	False	False	False
1535	False	False	False	False	False	False	False	False	False
1536	False	False	False	False	False	False	False	False	False
1537	False	False	False	False	False	False	False	False	False

1538 rows × 9 columns

In [7]: `kcd.notnull()`

Out[7]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	True	True	True	True	True	True	True	True	True
1	True	True	True	True	True	True	True	True	True
2	True	True	True	True	True	True	True	True	True
3	True	True	True	True	True	True	True	True	True
4	True	True	True	True	True	True	True	True	True
...
1533	True	True	True	True	True	True	True	True	True
1534	True	True	True	True	True	True	True	True	True
1535	True	True	True	True	True	True	True	True	True
1536	True	True	True	True	True	True	True	True	True
1537	True	True	True	True	True	True	True	True	True

1538 rows × 9 columns

In [8]: `kcd.count()`

Out[8]:

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 [9]: kcd.groupby(["previous_owners"]).count()
```

```
Out[9]:
```

	ID	model	engine_power	age_in_days	km	lat	lon	price
previous_owners								
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 [10]: kcd["price"].sum()
```

```
Out[10]: 13189894
```

Correlating the values in the table

```
In [11]: cor=k.corr()  
cor
```

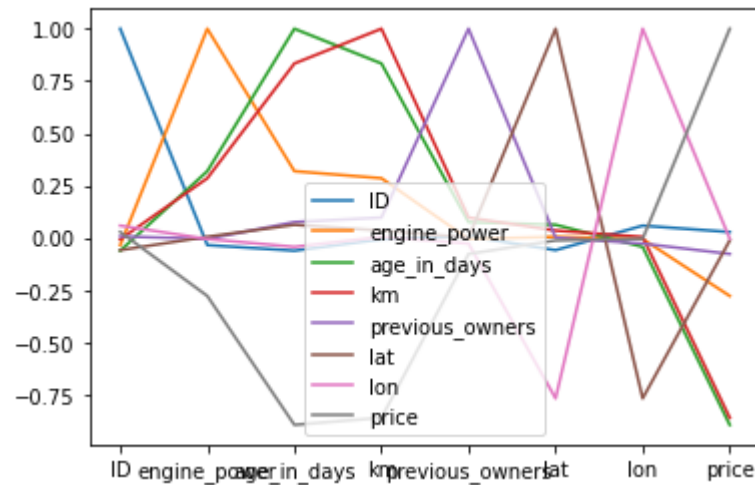
```
Out[11]:
```

	ID	engine_power	age_in_days	km	previous_owners	lat	lon	price
ID	1.000000	-0.034059	-0.060753	-0.006537	0.007803	-0.058207	0.058941	0.028516
engine_power	-0.034059	1.000000	0.319190	0.285495	-0.005030	0.005721	-0.005032	-0.277235
age_in_days	-0.060753	0.319190	1.000000	0.833890	0.075775	0.062982	-0.042667	-0.893328
km	-0.006537	0.285495	0.833890	1.000000	0.097539	0.035519	0.004839	-0.859373
previous_owners	0.007803	-0.005030	0.075775	0.097539	1.000000	0.001697	-0.026836	-0.076274
lat	-0.058207	0.005721	0.062982	0.035519	0.001697	1.000000	-0.766646	-0.011733
lon	0.058941	-0.005032	-0.042667	0.004839	-0.026836	-0.766646	1.000000	-0.003541
price	0.028516	-0.277235	-0.893328	-0.859373	-0.076274	-0.011733	-0.003541	1.000000

Plot the graph for correlation

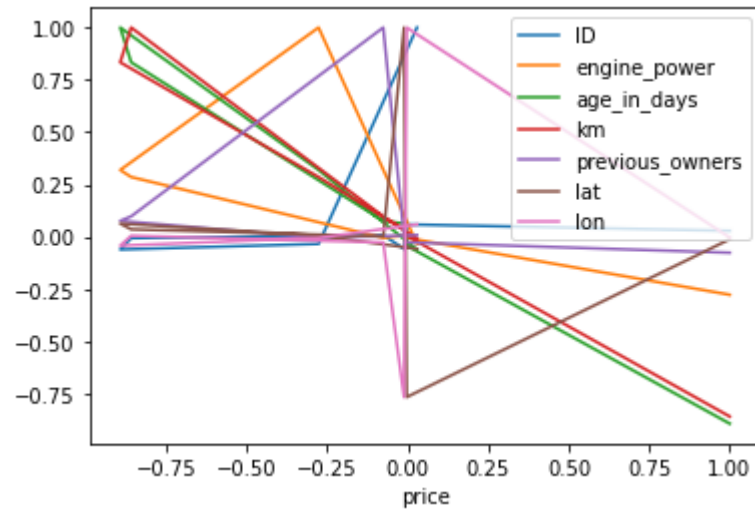
In [12]: `cor.plot()`

Out[12]: `<Axes: >`



```
In [13]: cor.plot("price")
```

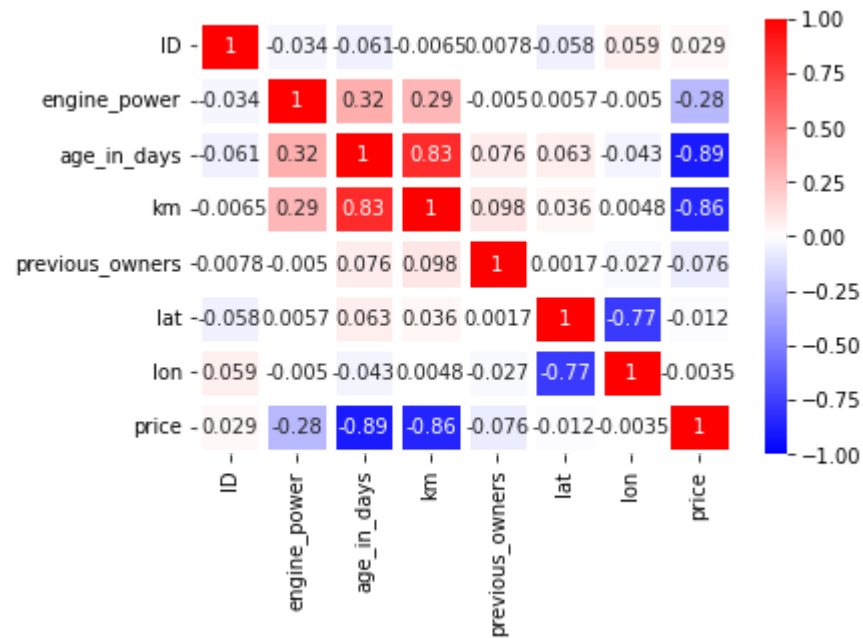
```
Out[13]: <Axes: xlabel='price'>
```



plotting correlating graph

```
In [14]: import seaborn as sb
sb.heatmap(cor,vmax=1,vmin=-1,annot=True,linewidth=5,cmap='bwr')
```

Out[14]: <Axes: >



removing unwanted columns from table

```
In [15]: det=kcd.drop(['ID', 'engine_power', 'model', 'age_in_days', 'previous_owners', 'lat', 'lon'],axis=1)
det
```

Out[15]:

	km	price
0	25000	8900
1	32500	8800
2	142228	4200
3	160000	6000
4	106880	5700
...
1533	115280	5200
1534	112000	4600
1535	60457	7500
1536	80750	5990
1537	54276	7900

1538 rows × 2 columns

```
In [16]: det1=kcd.drop(['ID','model','lat','lon','previous_owners','engine_power','price'],axis=1)
det1
```

Out[16]:

	age_in_days	km
0	882	25000
1	1186	32500
2	4658	142228
3	2739	160000
4	3074	106880
...
1533	3712	115280
1534	3835	112000
1535	2223	60457
1536	2557	80750
1537	1766	54276

1538 rows × 2 columns

converting strings into integers

```
In [17]: kcd['model']=kcd['model'].map({'lounge':1,'pop':2,'sport':3})
kcd
```

Out[17]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1	1	51	882	25000	1	44.907242	8.611560	8900
1	2	2	51	1186	32500	1	45.666359	12.241890	8800
2	3	3	74	4658	142228	1	45.503300	11.417840	4200
3	4	1	51	2739	160000	1	40.633171	17.634609	6000
4	5	2	73	3074	106880	1	41.903221	12.495650	5700
...
1533	1534	3	51	3712	115280	1	45.069679	7.704920	5200
1534	1535	1	74	3835	112000	1	45.845692	8.666870	4600
1535	1536	2	51	2223	60457	1	45.481541	9.413480	7500
1536	1537	1	51	2557	80750	1	45.000702	7.682270	5990
1537	1538	2	51	1766	54276	1	40.323410	17.568270	7900

1538 rows × 9 columns