

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [2]: data=pd.read_csv(r"C:\Users\user\Desktop\Vicky\1_fiat500_VehicleSelection_Dataset.csv")
```

```
In [3]: data.shape
```

```
Out[3]: (1549, 11)
```

```
In [4]: data.head
```

```
Out[4]: <bound method NDFrame.head of
m previous_owners \
0      1.0 lounge      51.0      882.0  25000.0      1.0
1      2.0 pop      51.0      1186.0  32500.0      1.0
2      3.0 sport      74.0      4658.0  142228.0      1.0
3      4.0 lounge      51.0      2739.0  160000.0      1.0
4      5.0 pop      73.0      3074.0  106880.0      1.0
...      ...      ...      ...      ...      ...
1544 NaN NaN NaN NaN NaN NaN
1545 NaN NaN NaN NaN NaN NaN
1546 NaN NaN NaN NaN NaN NaN
1547 NaN NaN NaN NaN NaN NaN
1548 NaN NaN NaN NaN NaN NaN

      lat      lon      price  Unnamed: 9  Unnamed: 10
0      44.907242  8.611559868      8900      NaN      NaN
1      45.666359  12.24188995      8800      NaN      NaN
2      45.503300      11.41784      4200      NaN      NaN
3      40.633171  17.63460922      6000      NaN      NaN
4      41.903221  12.49565029      5700      NaN      NaN
...      ...      ...      ...      ...      ...
1544      NaN      length      5      NaN      NaN
1545      NaN      concat lonprice      NaN      NaN
1546      NaN  Null values      NO      NaN      NaN
1547      NaN      find      1      NaN      NaN
1548      NaN      search      1      NaN      NaN

[1549 rows x 11 columns]>
```

In [5]:

data.tail()

Out[5]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9
1544	NaN	NaN	NaN	NaN	NaN	NaN	NaN	length	5	NaN
1545	NaN	NaN	NaN	NaN	NaN	NaN	NaN	concat	lonprice	NaN
1546	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Null values	NO	NaN
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	find	1	NaN
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	search	1	NaN

In [6]:

data.describe()

Out[6]:

	ID	engine_power	age_in_days	km	previous_owners	lat	Unnamed: 9
count	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	0.
mean	769.500000	51.904421	1650.980494	53396.011704	1.123537	43.541361	Na
std	444.126671	3.988023	1289.522278	40046.830723	0.416423	2.133518	Na
min	1.000000	51.000000	366.000000	1232.000000	1.000000	36.855839	Na
25%	385.250000	51.000000	670.000000	20006.250000	1.000000	41.802990	Na
50%	769.500000	51.000000	1035.000000	39031.000000	1.000000	44.394096	Na
75%	1153.750000	51.000000	2616.000000	79667.750000	1.000000	45.467960	Na
max	1538.000000	77.000000	4658.000000	235000.000000	4.000000	46.795612	Na

```
In [7]: data.isna()
```

Out[7]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9
0	False	False	False	False	False	False	False	False	False	True
1	False	False	False	False	False	False	False	False	False	True
2	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	True
4	False	False	False	False	False	False	False	False	False	True
...	...	...	...	...	...	...	...	...	...	...
1544	True	True	True	True	True	True	True	False	False	True
1545	True	True	True	True	True	True	True	False	False	True
1546	True	True	True	True	True	True	True	False	False	True
1547	True	True	True	True	True	True	True	False	False	True
1548	True	True	True	True	True	True	True	False	False	True

1549 rows × 11 columns

```
In [8]: data.size
```

Out[8]: 17039

In [9]: data.isna

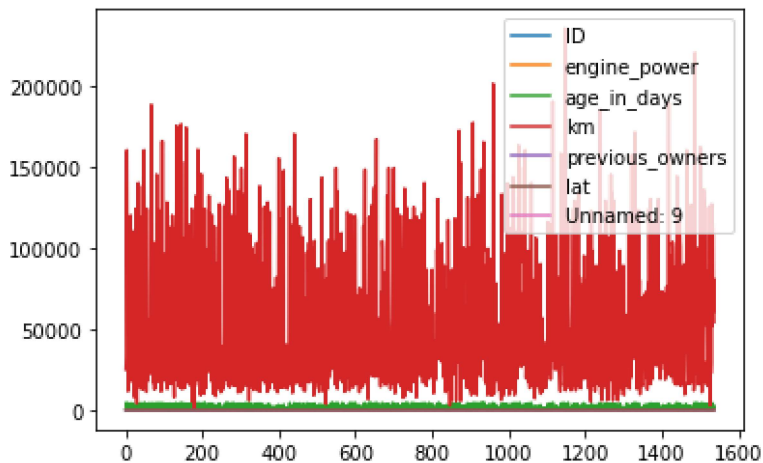
```
Out[9]: <bound method DataFrame.isna of
km  previous_owners \
0    1.0    lounge    51.0    882.0    25000.0    1.0
1    2.0    pop      51.0    1186.0    32500.0    1.0
2    3.0    sport    74.0    4658.0    142228.0    1.0
3    4.0    lounge    51.0    2739.0    160000.0    1.0
4    5.0    pop      73.0    3074.0    106880.0    1.0
...    ...    ...    ...    ...    ...    ...
1544 NaN    NaN    NaN    NaN    NaN    NaN    NaN
1545 NaN    NaN    NaN    NaN    NaN    NaN    NaN
1546 NaN    NaN    NaN    NaN    NaN    NaN    NaN
1547 NaN    NaN    NaN    NaN    NaN    NaN    NaN
1548 NaN    NaN    NaN    NaN    NaN    NaN    NaN

lat    lon    price  Unnamed: 9  Unnamed: 10
0    44.907242  8.611559868    8900    NaN    NaN
1    45.666359  12.24188995    8800    NaN    NaN
2    45.503300    11.41784    4200    NaN    NaN
3    40.633171  17.63460922    6000    NaN    NaN
4    41.903221  12.49565029    5700    NaN    NaN
...    ...    ...    ...    ...    ...
1544    NaN    length    5    NaN    NaN
1545    NaN    concat  lonprice    NaN    NaN
1546    NaN  Null values    NO    NaN    NaN
1547    NaN    find    1    NaN    NaN
1548    NaN    search    1    NaN    NaN
```

[1549 rows x 11 columns]>

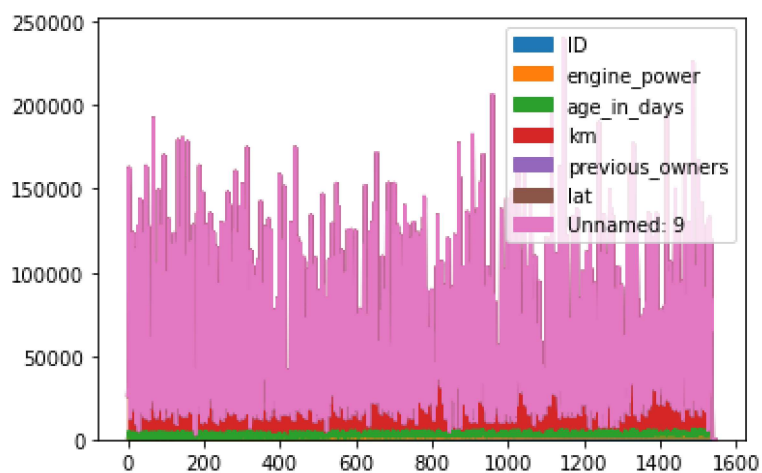
In [10]: data.plot.line()

Out[10]: <AxesSubplot:>



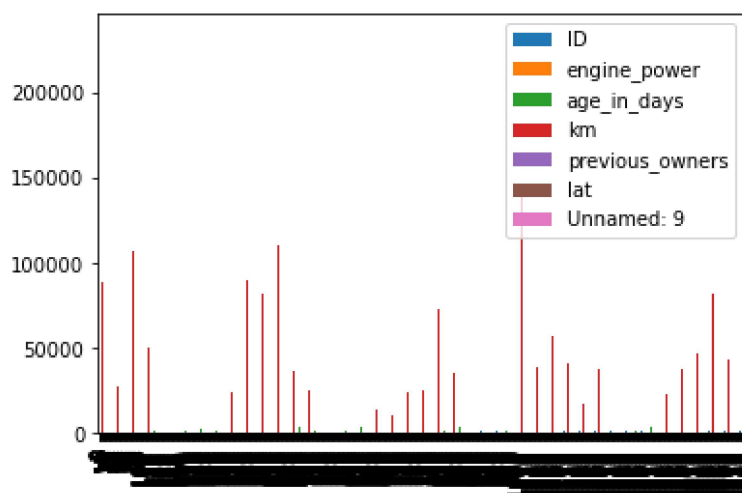
```
In [11]: data.plot.area()
```

```
Out[11]: <AxesSubplot:>
```



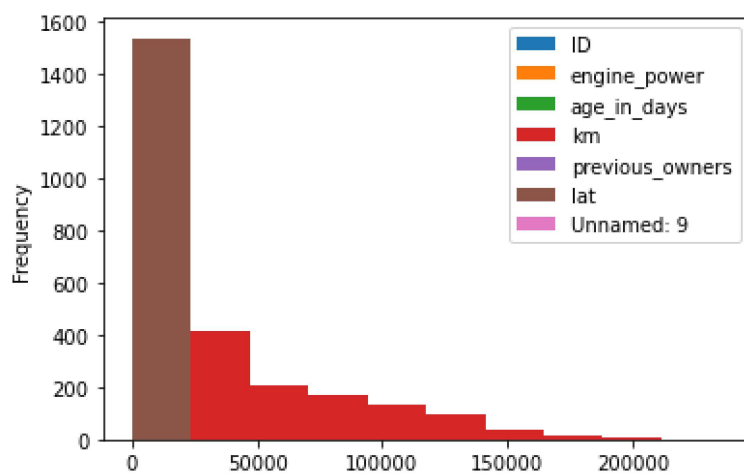
```
In [12]: data.plot.bar()
```

```
Out[12]: <AxesSubplot:>
```



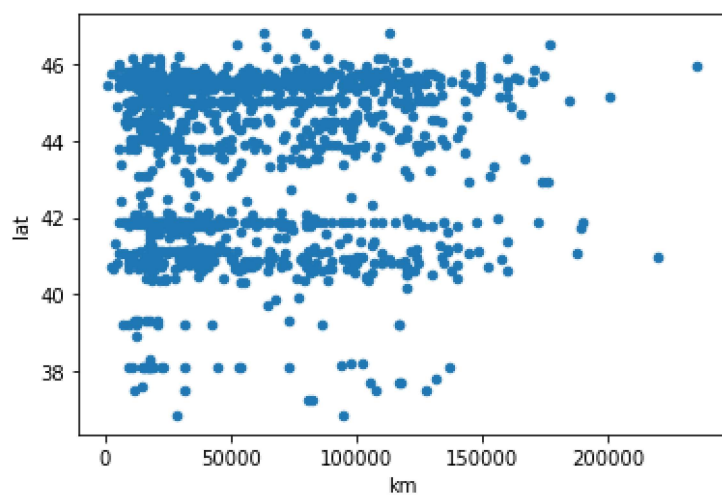
```
In [13]: data.plot.hist()
```

```
Out[13]: <AxesSubplot:ylabel='Frequency'>
```



```
In [14]: data.plot.scatter(x="km",y='lat')
```

```
Out[14]: <AxesSubplot:xlabel='km', ylabel='lat'>
```



In [15]:

data.fillna(value=5)

Out[15]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	pri
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868	89
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995	88
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784	42
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922	60
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.49565029	57
...	...	...	...	...	...	...	...	...	...
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	concat	lonpr
1546	5.0	5	5.0	5.0	5.0	5.0	5.000000	Null values	t
1547	5.0	5	5.0	5.0	5.0	5.0	5.000000	find	
1548	5.0	5	5.0	5.0	5.0	5.0	5.000000	search	

1549 rows × 11 columns

In [16]:

pd.isna(data)

Out[16]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9
0	False	False	False	False	False	False	False	False	False	True
1	False	False	False	False	False	False	False	False	False	True
2	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	True
4	False	False	False	False	False	False	False	False	False	True
...	...	...	...	...	...	...	...	...	...	...
1544	True	True	True	True	True	True	True	False	False	True
1545	True	True	True	True	True	True	True	False	False	True
1546	True	True	True	True	True	True	True	False	False	True
1547	True	True	True	True	True	True	True	False	False	True
1548	True	True	True	True	True	True	True	False	False	True

1549 rows × 11 columns

In [ ]:

