

22/07/23 tony3

```
In [2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as pp
```

```
In [76]: x=pd.read_csv(r"C:\Users\user\Downloads\1_fiat500_VehicleSelection_Dataset.csv")
x
```

Out[76]:

	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	leng
1545	NaN	NaN	NaN	NaN	NaN	NaN	NaN	conc
1546	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Null valu
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	fi
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	sear

1549 rows × 11 columns

```
In [78]: x. dtypes
```

Out[78]:

ID	float64
model	object
engine_power	float64
age_in_days	float64
km	float64
previous_owners	float64
lat	float64
lon	object
price	object
Unnamed: 9	float64
Unnamed: 10	object
dtype:	object

In [79]: `x.dtypes`

```
Out[79]: ID                float64
model                object
engine_power         float64
age_in_days          float64
km                   float64
previous_owners       float64
lat                  float64
lon                  object
price                object
Unnamed: 9            float64
Unnamed: 10           object
dtype: object
```

In [80]: `x.tail()`

```
Out[80]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Un
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	NO	
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	find	1	
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	search	1	

In [81]: `x.columns`

```
Out[81]: Index(['ID', 'model', 'engine_power', 'age_in_days', 'km', 'previous_owners',
               'lat', 'lon', 'price', 'Unnamed: 9', 'Unnamed: 10'],
              dtype='object')
```

In [82]: `x.index`

```
Out[82]: RangeIndex(start=0, stop=1549, step=1)
```

In [83]:

x.describe()

Out[83]:

	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	8.611559868
std	444.126671	3.988023	1289.522278	40046.830723	0.416423	2.133518	8.611559868
min	1.000000	51.000000	366.000000	1232.000000	1.000000	36.855839	8.611559868
25%	385.250000	51.000000	670.000000	20006.250000	1.000000	41.802990	8.611559868
50%	769.500000	51.000000	1035.000000	39031.000000	1.000000	44.394096	8.611559868
75%	1153.750000	51.000000	2616.000000	79667.750000	1.000000	45.467960	8.611559868
max	1538.000000	77.000000	4658.000000	235000.000000	4.000000	46.795612	8.611559868

In [85]:

x["ID"]

Out[85]:

0	1.0
1	2.0
2	3.0
3	4.0
4	5.0
...	
1544	NaN
1545	NaN
1546	NaN
1547	NaN
1548	NaN

Name: ID, Length: 1549, dtype: float64

In [86]:

x.iloc[0:2]

Out[86]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868	86000.0
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995	86000.0

In [87]:

x.loc[0:3]

Out[87]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868	12.24188995
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995	11.41784
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784	17.63460922
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922	5.000000

In [88]:

x.loc["Country":"Freedom"]

Out[88]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9	Unnamed: 10

In [89]:

x[x["km"]<=20]

Out[89]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9	Unnamed: 10

In [90]:

x.fillna(value=5)

Out[90]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9	Unnamed: 10
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.611559868	12.24188995	11.41784	17.63460922
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.24188995	11.41784	17.63460922	5.000000
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.41784	17.63460922	5.000000	5.000000
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.63460922	5.000000	5.000000	5.000000
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.49565021	5.000000	5.000000	5.000000
...
1544	5.0	5	5.0	5.0	5.0	5.0	5.000000	length	5.000000	5.000000	5.000000
1545	5.0	5	5.0	5.0	5.0	5.0	5.000000	conca	5.000000	5.000000	5.000000
1546	5.0	5	5.0	5.0	5.0	5.0	5.000000	Null value	5.000000	5.000000	5.000000
1547	5.0	5	5.0	5.0	5.0	5.0	5.000000	fini	5.000000	5.000000	5.000000
1548	5.0	5	5.0	5.0	5.0	5.0	5.000000	search	5.000000	5.000000	5.000000

1549 rows × 11 columns

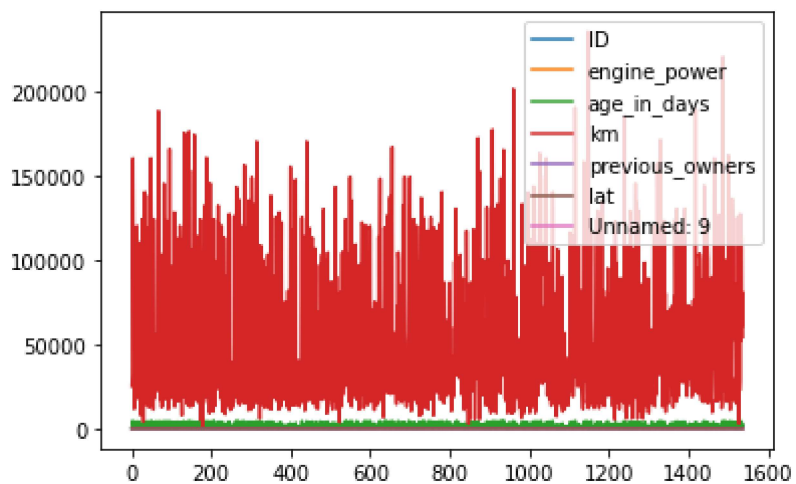
```
In [91]: x.dropna()
```

```
Out[91]:
```

ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unnamed: 9
----	-------	--------------	-------------	----	-----------------	-----	-----	-------	------------

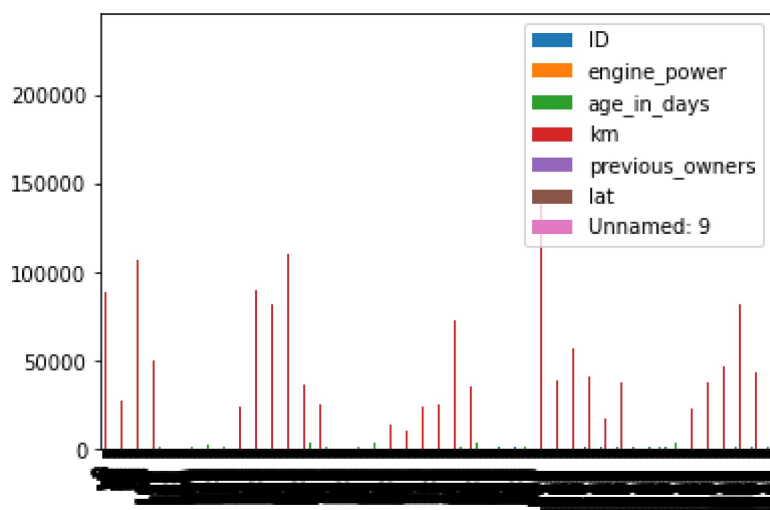
```
In [92]: x.plot.line()
```

```
Out[92]: <AxesSubplot:>
```



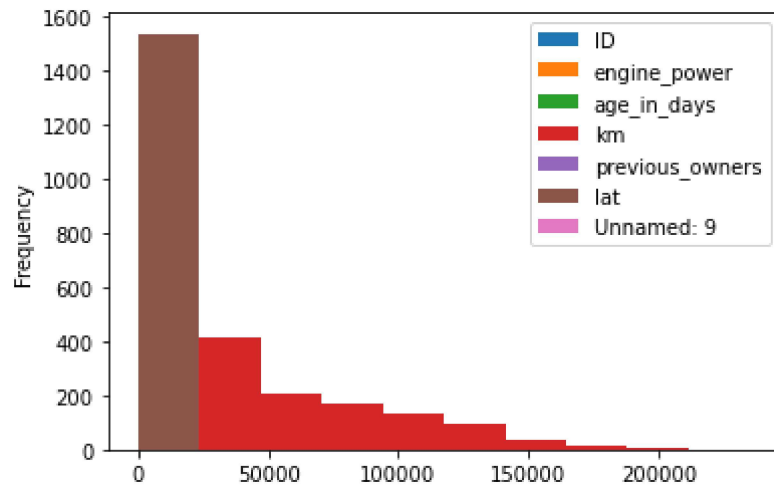
```
In [93]: x.plot.bar()
```

```
Out[93]: <AxesSubplot:>
```

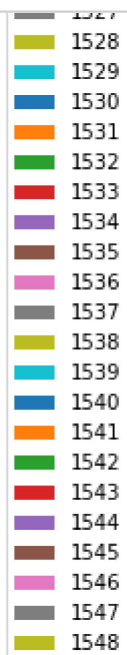


```
In [94]: x.plot.hist()
```

```
Out[94]: <AxesSubplot:ylabel='Frequency'>
```

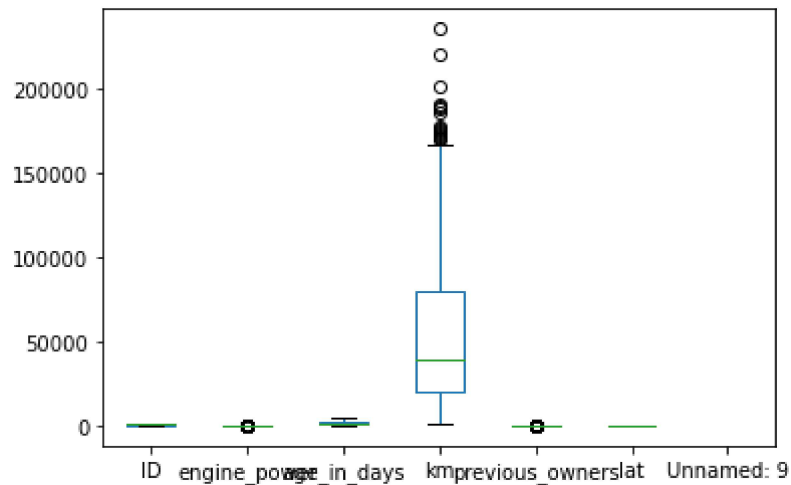


```
In [98]: x.plot.pie(y='lat')
```



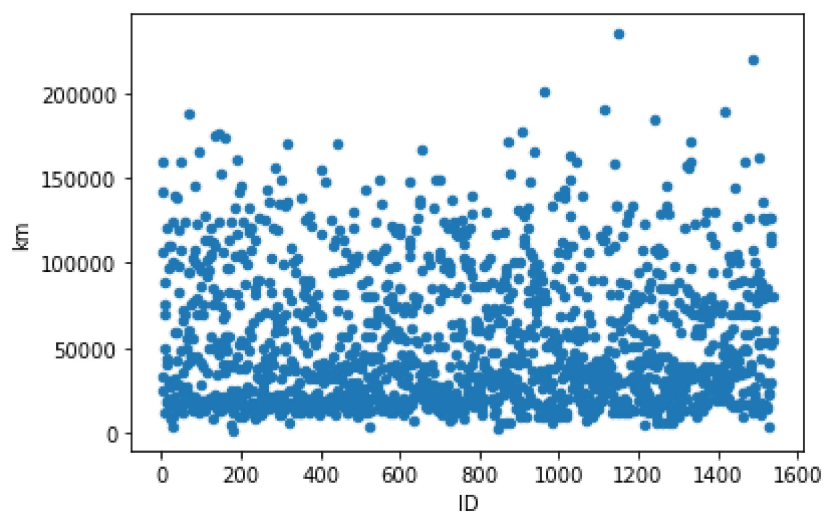
```
In [101]: x.plot.box()
```

```
Out[101]: <AxesSubplot:>
```



```
In [100]: x.plot.scatter(x='ID',y='km')
```

```
Out[100]: <AxesSubplot:xlabel='ID', ylabel='km'>
```



```
In [ ]:
```

```
In [ ]:
```