

Import labary

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

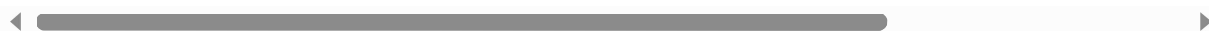
Import dataset

```
In [2]: data=pd.read_csv(r"c:\Users\user\Downloads\VehicleSelection.csv")
data
```

Out[2]:

	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



Print head first 10 rows

```
In [20]: data.head(10)
```

```
Out[20]:
```

	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

Print tail last 14 rows

```
In [21]: data.tail(14)
```

```
Out[21]:
```


	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon
1535	1536.0	pop	51.0	2223.0	60457.0	1.0	45.481541	9.4134791
1536	1537.0	lounge	51.0	2557.0	80750.0	1.0	45.000702	7.682270
1537	1538.0	pop	51.0	1766.0	54276.0	1.0	40.323410	17.568261
1538	NaN	NaN	NaN	NaN	NaN	NaN	NaN	s
1539	NaN	NaN	NaN	NaN	NaN	NaN	NaN	i
1540	NaN	NaN	NaN	NaN	NaN	NaN	NaN	cc
1541	NaN	NaN	NaN	NaN	NaN	NaN	NaN	cou
1542	NaN	NaN	NaN	NaN	NaN	NaN	NaN	su
1543	NaN	NaN	NaN	NaN	NaN	NaN	NaN	counta (emj
1544	NaN	NaN	NaN	NaN	NaN	NaN	NaN	len
1545	NaN	NaN	NaN	NaN	NaN	NaN	NaN	cor
1546	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Null val
1547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	i
1548	NaN	NaN	NaN	NaN	NaN	NaN	NaN	sea

To print statistical data

```
In [22]: data.describe()
```

Out[22]:

	ID	engine_power	age_in_days	km	previous_owners	lat	U
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	



To print rows and coloum

```
In [23]: np.shape(data)
```

Out[23]: (1549, 11)

To print no. of elements

```
In [24]: np.size(data)
```

Out[24]: 17039

To print missing values

```
In [25]: data.isna()
```

```
Out[25]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unit
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	
...	
1544	True	True	True	True	True	True	True	False	False	
1545	True	True	True	True	True	True	True	False	False	
1546	True	True	True	True	True	True	True	False	False	
1547	True	True	True	True	True	True	True	False	False	
1548	True	True	True	True	True	True	True	False	False	

1549 rows × 11 columns



Fill a value 60 in missing place

```
In [26]: data.fillna(value=60)
```

```
Out[26]:
```

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price	Unit
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.61155981		
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.2418891		
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.4171		
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.6346093		
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.4956503		
...		
1544	60.0	60	60.0	60.0	60.0	60.0	60.000000			length
1545	60.0	60	60.0	60.0	60.0	60.0	60.000000			concentration
1546	60.0	60	60.0	60.0	60.0	60.0	60.000000			Null value
1547	60.0	60	60.0	60.0	60.0	60.0	60.000000			fill
1548	60.0	60	60.0	60.0	60.0	60.0	60.000000			search

1549 rows × 11 columns



```
In [27]: import matplotlib.pyplot as pp
```

```
In [28]: dd=data[['km','price']]
dd
```

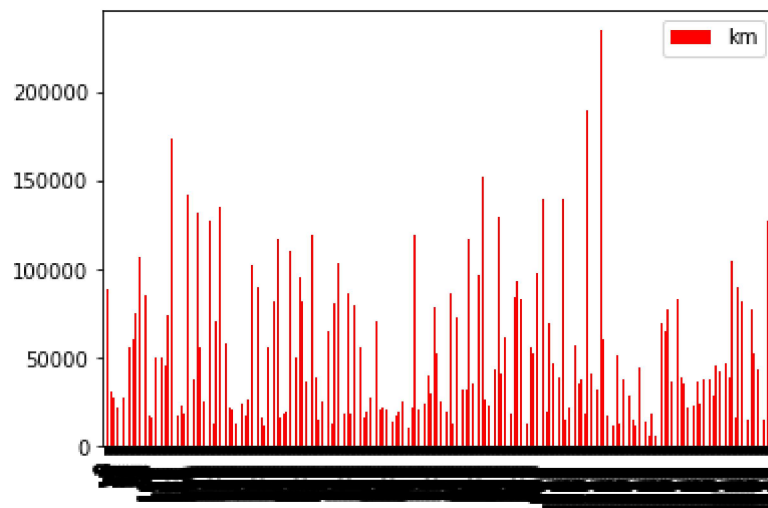
Out[28]:

	km	price
0	25000.0	8900
1	32500.0	8800
2	142228.0	4200
3	160000.0	6000
4	106880.0	5700
...
1544	NaN	5
1545	NaN	lonprice
1546	NaN	NO
1547	NaN	1
1548	NaN	1

1549 rows × 2 columns

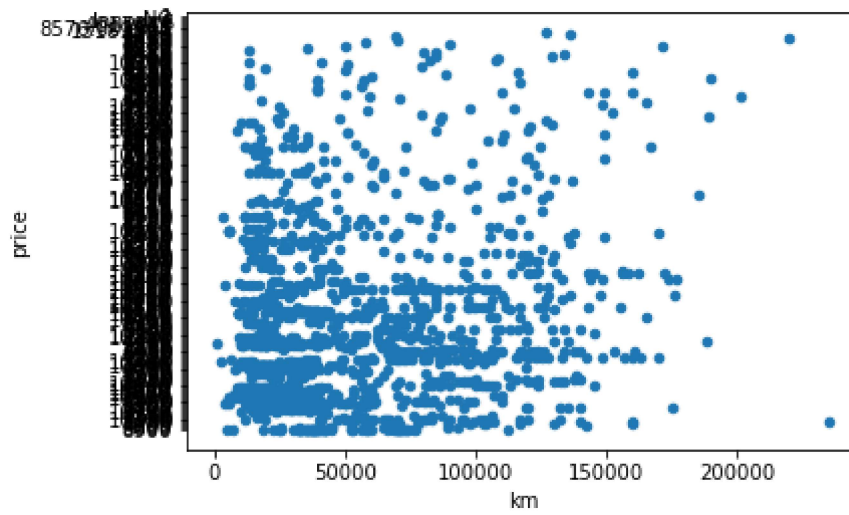
```
In [29]: dd.plot.bar(color='r')
```

Out[29]: <AxesSubplot:>



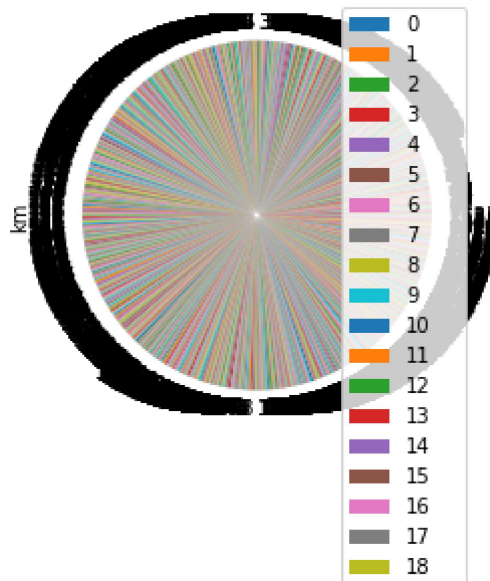
```
In [30]: dd.plot.scatter(x='km',y='price')
```

```
Out[30]: <AxesSubplot:xlabel='km', ylabel='price'>
```



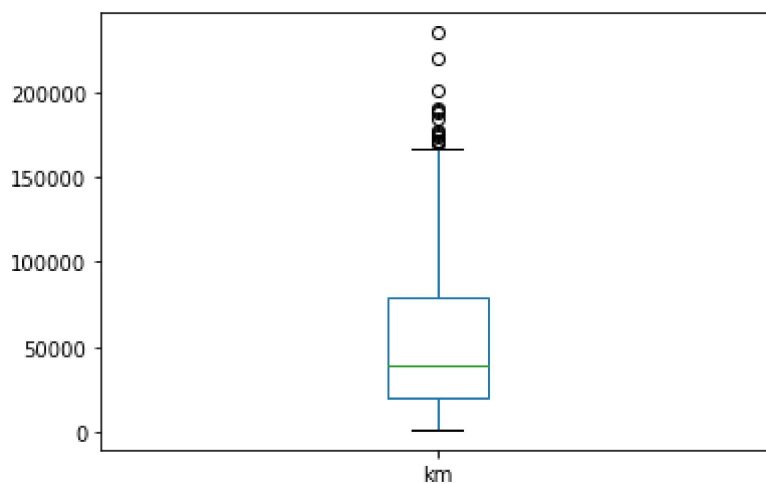
```
In [31]: dd.plot.pie(y='km')
```

```
Out[31]: <AxesSubplot:ylabel='km'>
```



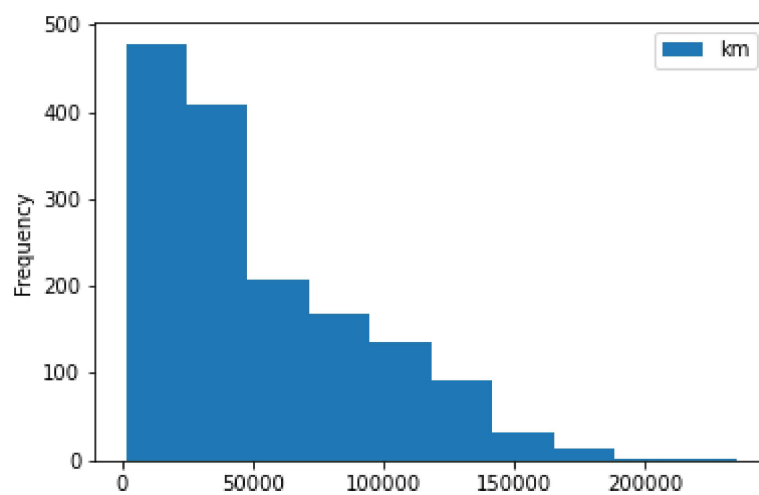
```
In [32]: dd.plot.box()
```

```
Out[32]: <AxesSubplot:>
```



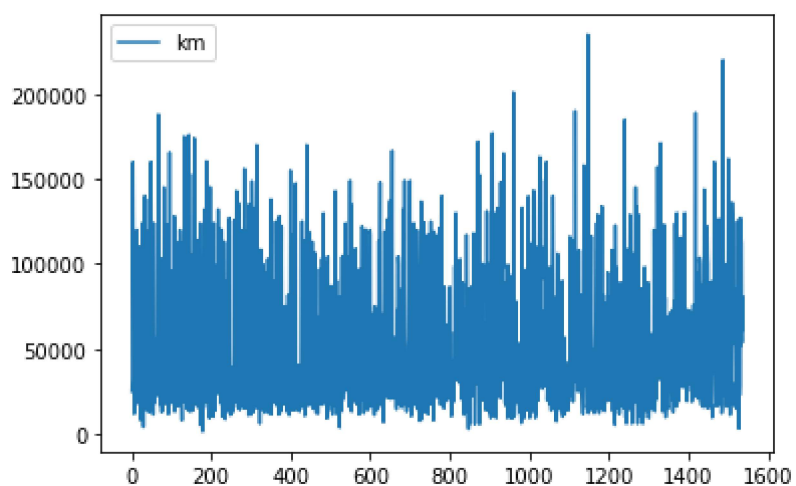
```
In [33]: dd.plot.hist()
```

```
Out[33]: <AxesSubplot:ylabel='Frequency'>
```



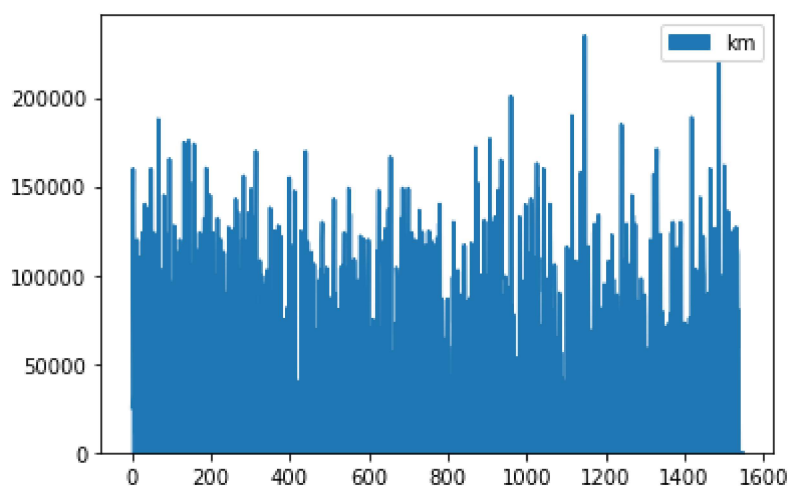
```
In [36]: dd.plot.line()
```

```
Out[36]: <AxesSubplot:>
```



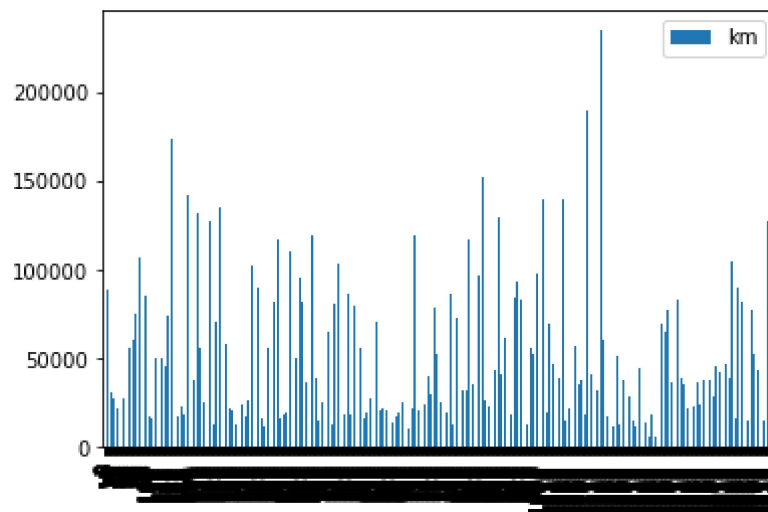
```
In [34]: dd.plot.area()
```

```
Out[34]: <AxesSubplot:>
```




```
In [35]: dd.plot.bar()
```

```
Out[35]: <AxesSubplot:>
```



```
In [ ]:
```