

In [1]:

```
import pandas as pd
import numpy as np
```

In [2]:

```
a=pd.read_csv(r"C:\Users\user\Downloads\fiat500_VehicleSelection_Dataset.csv")
```

To print 1st Five Rows

In [3]:

```
a.head(5)
```

Out[3]:

	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	8.61155986	11.4178
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.2418895	11.4178
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.4178	11.4178
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.6346092	11.4178
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.4956502	11.4178

To print last 5 rows

In [4]:

```
a.tail()
```

Out[4]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
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

To find shape

In [5]:

```
a.shape
```

Out[5]:

(1549, 11)

To find size

In [6]:

```
a.size
```

Out[6]:

17039

To describe

In [7]:

```
a.describe()
```

Out[7]:

	ID	engine_power	age_in_days	km	previous_owners	last_seen
count	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000	1538.000000
mean	769.500000	51.904421	1650.980494	53396.011704	1.123537	43.54136
std	444.126671	3.988023	1289.522278	40046.830723	0.416423	2.13351
min	1.000000	51.000000	366.000000	1232.000000	1.000000	36.85583
25%	385.250000	51.000000	670.000000	20006.250000	1.000000	41.80299
50%	769.500000	51.000000	1035.000000	39031.000000	1.000000	44.39409
75%	1153.750000	51.000000	2616.000000	79667.750000	1.000000	45.46796
max	1538.000000	77.000000	4658.000000	235000.000000	4.000000	46.79561

To fill the null value

In [8]:

```
a.fillna(value=3)
```

Out[8]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	
0	1.0	lounge	51.0	882.0	25000.0	1.0	44.907242	8.61156
1	2.0	pop	51.0	1186.0	32500.0	1.0	45.666359	12.2418
2	3.0	sport	74.0	4658.0	142228.0	1.0	45.503300	11.4
3	4.0	lounge	51.0	2739.0	160000.0	1.0	40.633171	17.6346
4	5.0	pop	73.0	3074.0	106880.0	1.0	41.903221	12.4956
...
1544	3.0	3	3.0	3.0	3.0	3.0	3.000000	l
1545	3.0	3	3.0	3.0	3.0	3.0	3.000000	c
1546	3.0	3	3.0	3.0	3.0	3.0	3.000000	Null v
1547	3.0	3	3.0	3.0	3.0	3.0	3.000000	
1548	3.0	3	3.0	3.0	3.0	3.0	3.000000	s

1549 rows × 11 columns



To drop the null value rows

In [9]:

```
a.dropna()
```

Out[9]:

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



To check the null values

In [10]:

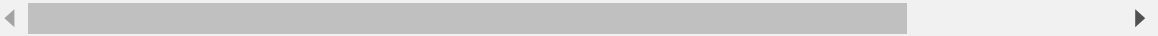
```
a.isna()
```

Out[10]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lon	price
	0	False	False	False	False	False	False	False	False
	1	False	False	False	False	False	False	False	False
	2	False	False	False	False	False	False	False	False
	3	False	False	False	False	False	False	False	False
	4	False	False	False	False	False	False	False	False

	1544	True	True	True	True	True	True	False	False
	1545	True	True	True	True	True	True	False	False
	1546	True	True	True	True	True	True	False	False
	1547	True	True	True	True	True	True	False	False
	1548	True	True	True	True	True	True	False	False

1549 rows × 11 columns



In [13]:

```
b=a[["age_in_days", "km"]]
b
```

Out[13]:

	age_in_days	km
0	882.0	25000.0
1	1186.0	32500.0
2	4658.0	142228.0
3	2739.0	160000.0
4	3074.0	106880.0
...
1544	NaN	NaN
1545	NaN	NaN
1546	NaN	NaN
1547	NaN	NaN
1548	NaN	NaN

1549 rows × 2 columns

In [14]:

```
conda install matplotlib
```

Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working... done

All requested packages already installed.

Note: you may need to restart the kernel to use updated packages.

==> WARNING: A newer version of conda exists. <==
current version: 4.10.1
latest version: 23.5.2

Please update conda by running

```
$ conda update -n base -c defaults conda
```

In [16]:

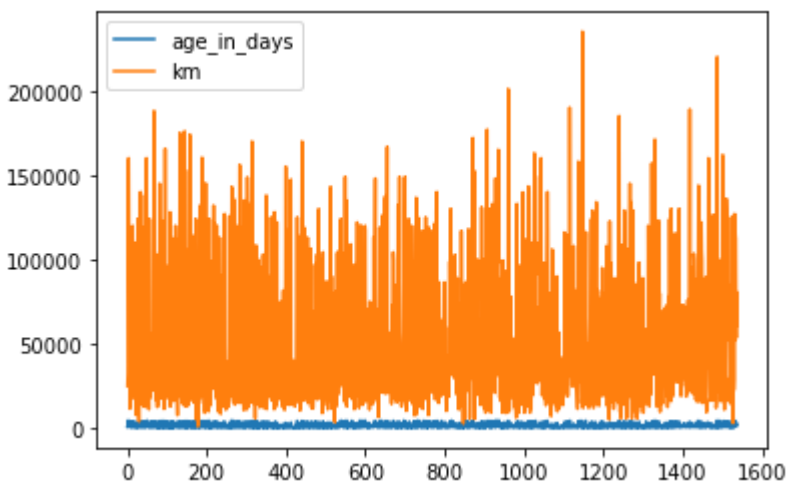
```
import matplotlib.pyplot as pp
```

In [17]:

```
b.plot.line()
```

Out[17]:

<AxesSubplot:>

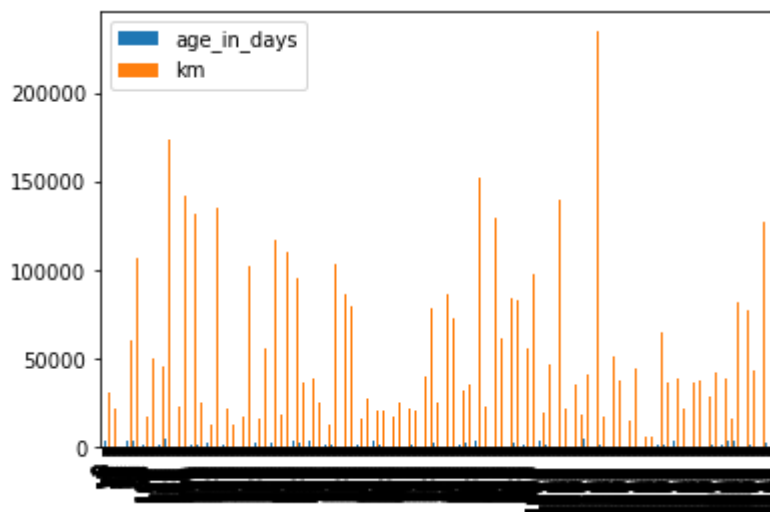


In [18]:

```
b.plot.bar()
```

Out[18]:

<AxesSubplot:>

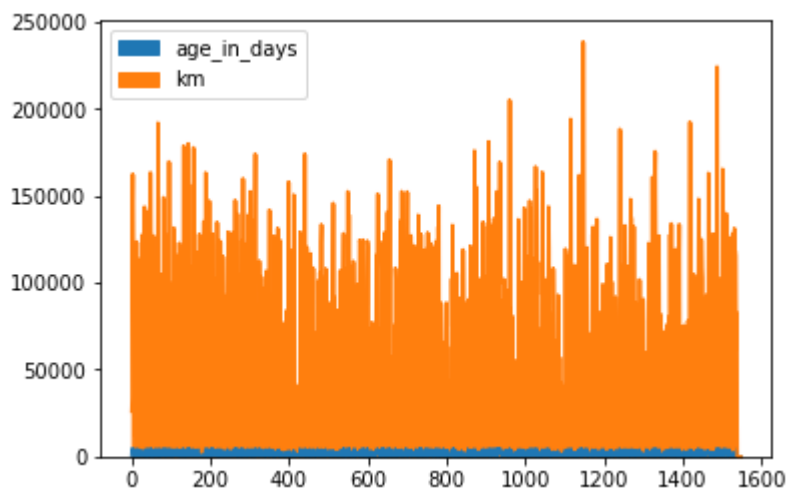


In [19]:

```
b.plot.area()
```

Out[19]:

<AxesSubplot:>

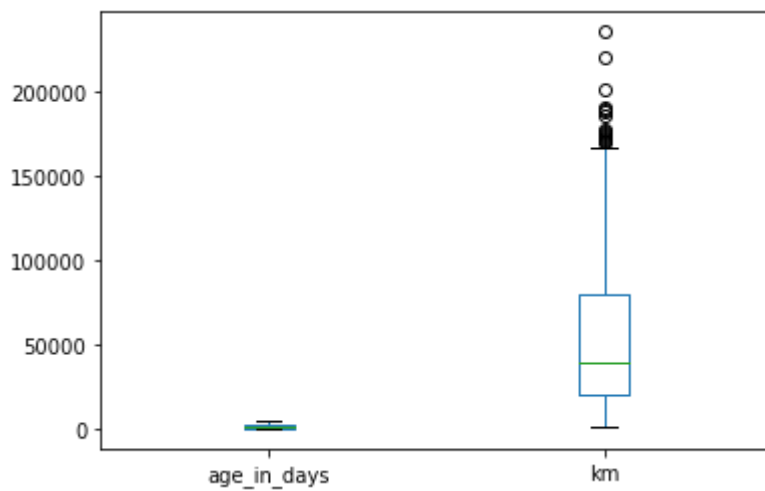


In [20]:

```
b.plot.box()
```

Out[20]:

<AxesSubplot:>

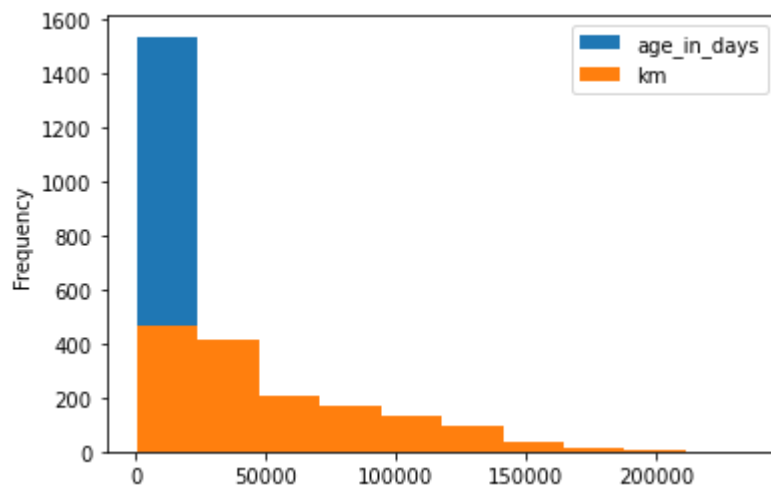


In [22]:

```
b.plot.hist()
```

Out[22]:

<AxesSubplot:ylabel='Frequency'>

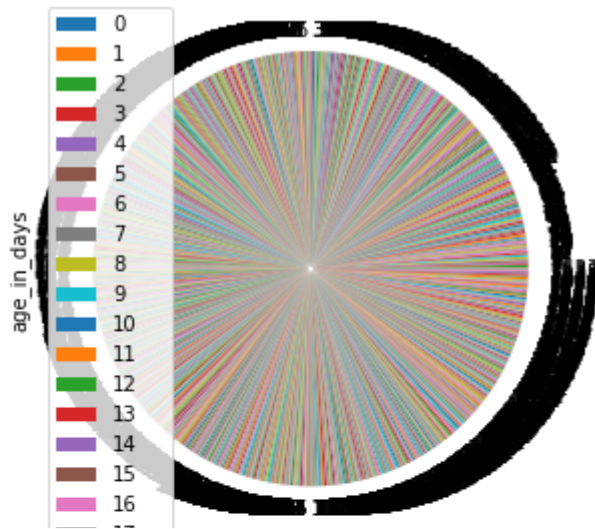


In [27]:

```
b.plot.pie(y='age_in_days',figsize=(5,5))
```

Out[27]:

<AxesSubplot:ylabel='age_in_days'>

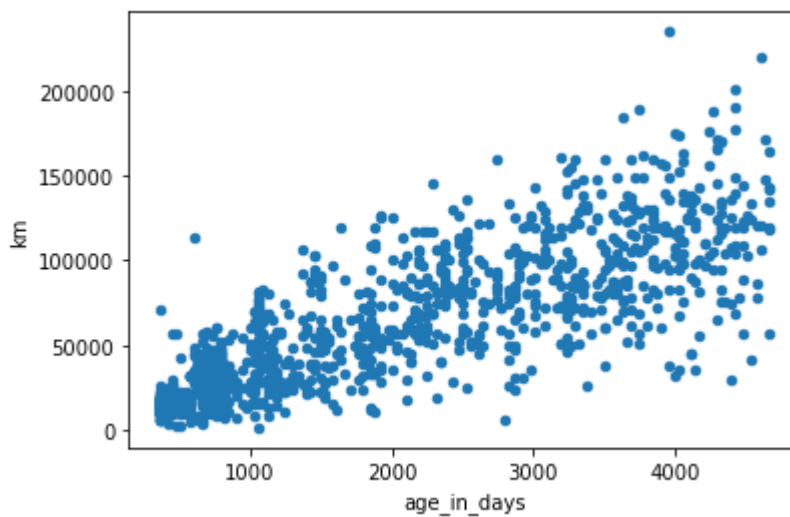


In [24]:

```
b.plot.scatter(x='age_in_days',y='km')
```

Out[24]:

<AxesSubplot:xlabel='age_in_days', ylabel='km'>



In []: