



```
In [106... import pandas as pd
```

```
In [107... import pandas as pd ##importing numpy , import is key word , as is alias , r
print(pd.__version__)
```

2.3.0

```
In [108... %pip install pandas
import pandas as pd
```

Requirement already satisfied: pandas in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (2.3.0)Note: you may need to restart the kernel to use updated packages.

Requirement already satisfied: numpy>=1.23.2 in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (from pandas) (2.0.0)

Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (from pandas) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (from pandas) (2025.2)

Requirement already satisfied: tzdata>=2022.7 in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (from pandas) (2025.2)

Requirement already satisfied: six>=1.5 in c:\users\saini\appdata\local\programs\python\python311\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)

[notice] A new release of pip is available: 23.2.1 -> 25.1.1

[notice] To update, run: python.exe -m pip install --upgrade pip

```
In [109... # duplicated
# Null
# nan
# balance Dataset

# maths
# encodding>>>>>>>>>
# one hot encoding
# label encoding nblapping
# string
```

```
In [110... import pandas as pd
```

```
In [111... df = pd.read_csv("used_bikes.csv")
df
```

Out[111...

	bike_name	price	city	kms_driven	owner	age	power	brai
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	T
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Roy Enfie
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Trium
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	T
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yama
...
32643	Hero Passion Pro 100cc	39000.0	Delhi	22000.0	First Owner	4.0	100.0	He
32644	TVS Apache RTR 180cc	30000.0	Karnal	6639.0	First Owner	9.0	180.0	T
32645	Bajaj Avenger Street 220	60000.0	Delhi	20373.0	First Owner	6.0	220.0	Ba
32646	Hero Super Splendor 125cc	15600.0	Jaipur	84186.0	First Owner	16.0	125.0	He
32647	Bajaj Pulsar 150cc	22000.0	Pune	60857.0	First Owner	13.0	150.0	Ba

32648 rows × 8 columns

In [112...

```
df.head()
```

Out[112...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Royal Enfield
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triumph
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamaha

In [113...

df.tail()

Out[113...

	bike_name	price	city	kms_driven	owner	age	power	brand
32643	Hero Passion Pro 100cc	39000.0	Delhi	22000.0	First Owner	4.0	100.0	Hero
32644	TVS Apache RTR 180cc	30000.0	Karnal	6639.0	First Owner	9.0	180.0	TVS
32645	Bajaj Avenger Street 220	60000.0	Delhi	20373.0	First Owner	6.0	220.0	Bajaj
32646	Hero Super Splendor 125cc	15600.0	Jaipur	84186.0	First Owner	16.0	125.0	Hero
32647	Bajaj Pulsar 150cc	22000.0	Pune	60857.0	First Owner	13.0	150.0	Bajaj

In [114...

df.head(9)

Out[114...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Royal Enfield
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triumph
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamaha
5	Yamaha FZs 150cc	53499.0	Delhi	25000.0	First Owner	6.0	150.0	Yamaha
6	Honda CB Hornet 160R ABS DLX	85000.0	Delhi	8200.0	First Owner	3.0	160.0	Honda
7	Hero Splendor Plus Self Alloy 100cc	45000.0	Delhi	12645.0	First Owner	3.0	100.0	Hero
8	Royal Enfield Thunderbird X 350cc	145000.0	Bangalore	9190.0	First Owner	3.0	350.0	Royal Enfield

In [115...

```
df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32648 entries, 0 to 32647
Data columns (total 8 columns):
#   Column      Non-Null Count  Dtype
---  -
0   bike_name    32648 non-null  object
1   price        32648 non-null  float64
2   city         32648 non-null  object
3   kms_driven   32648 non-null  float64
4   owner        32648 non-null  object
5   age          32648 non-null  float64
6   power        32648 non-null  float64
7   brand        32648 non-null  object
dtypes: float64(4), object(4)
memory usage: 2.0+ MB

```

In [116... `df.describe()`

```

Out[116...
           price    kms_driven    age    power
count  3.264800e+04  32648.000000  32648.000000  32648.000000
mean    6.829542e+04  26344.625184    8.048211   213.511302
std     9.071860e+04  22208.527695    4.031700   134.428868
min     4.400000e+03    1.000000    1.000000   100.000000
25%     2.500000e+04  12000.000000    5.000000   150.000000
50%     4.300000e+04  20373.000000    7.000000   150.000000
75%     8.000000e+04  35000.000000   10.000000   220.000000
max     1.900000e+06  750000.000000   63.000000  1800.000000

```

In [117... `df.dtypes`

```

Out[117... bike_name    object
price      float64
city       object
kms_driven float64
owner      object
age        float64
power      float64
brand      object
dtype: object

```

In [118... `df.duplicated()`

```

Out[118... 0      False
          1      False
          2      False
          3      False
          4      False
          ...
        32643    True
        32644    True
        32645    True
        32646    True
        32647    True
Length: 32648, dtype: bool

```

```
In [119... df.duplicated().sum()
```

```
Out[119... np.int64(25324)
```

```
In [120... df.isnull()
```

```

Out[120...
      bike_name  price  city  kms_driven  owner  age  power  brand
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
...          ...    ...    ...          ...    ...    ...    ...
32643      False  False  False      False  False  False  False  False
32644      False  False  False      False  False  False  False  False
32645      False  False  False      False  False  False  False  False
32646      False  False  False      False  False  False  False  False
32647      False  False  False      False  False  False  False  False

```

32648 rows × 8 columns

```
In [121... df.isnull().sum()
```

```

Out[121... bike_name    0
           price      0
           city      0
           kms_driven  0
           owner      0
           age        0
           power      0
           brand      0
           dtype: int64

```

```
In [122... df.drop_duplicates(inplace=True) ###inplace == true
df
```

```
Out[122...
```

	bike_name	price	city	kms_driven	owner	age	power	bran
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TV
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Roy Enfie
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triump
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TV
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamal
...	
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	He
9369	Bajaj Avenger 220cc	35000.0	Bangalore	60000.0	First Owner	9.0	220.0	Baj
9370	Harley-Davidson Street 750 ABS	450000.0	Jodhpur	3430.0	First Owner	4.0	750.0	Harle Davidsc
9371	Bajaj Dominar 400 ABS	139000.0	Hyderabad	21300.0	First Owner	4.0	400.0	Baj
9372	Bajaj Avenger Street 220	80000.0	Hyderabad	7127.0	First Owner	5.0	220.0	Baj

7324 rows x 8 columns

```
In [123... df.duplicated().sum()
```

```
Out[123... np.int64(0)
```

```
In [124... # df.dropna() ###remove null data values from data set
df
```

Out [124...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Royal Enfield
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triumph
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamaha
...
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	Hero
9369	Bajaj Avenger 220cc	35000.0	Bangalore	60000.0	First Owner	9.0	220.0	Bajaj
9370	Harley-Davidson Street 750 ABS	450000.0	Jodhpur	3430.0	First Owner	4.0	750.0	Harley Davidson
9371	Bajaj Dominar 400 ABS	139000.0	Hyderabad	21300.0	First Owner	4.0	400.0	Bajaj
9372	Bajaj Avenger Street 220	80000.0	Hyderabad	7127.0	First Owner	5.0	220.0	Bajaj

7324 rows × 8 columns

In [125...

```
df['brand']
```



```
Out[125... 0          TVS
1      Royal Enfield
2          Triumph
3          TVS
4          Yamaha
...
9362         Hero
9369         Bajaj
9370  Harley-Davidson
9371         Bajaj
9372         Bajaj
Name: brand, Length: 7324, dtype: object
```

```
In [126... df['brand'].unique()
```

```
Out[126... array(['TVS', 'Royal Enfield', 'Triumph', 'Yamaha', 'Honda', 'Hero',
      'Bajaj', 'Suzuki', 'Benelli', 'KTM', 'Mahindra', 'Kawasaki',
      'Ducati', 'Hyosung', 'Harley-Davidson', 'Jawa', 'BMW', 'Indian',
      'Rajdoot', 'LML', 'Yezdi', 'MV', 'Ideal'], dtype=object)
```

```
In [127... df['brand'].nunique()
```

```
Out[127... 23
```

```
In [128... df['brand'].value_counts()
```

```
Out[128... brand
Bajaj          2081
Royal Enfield  1346
Hero           1142
Honda          676
Yamaha         651
TVS            481
KTM            375
Suzuki         203
Harley-Davidson  91
Kawasaki        61
Hyosung         53
Mahindra        50
Benelli         46
Triumph         21
Ducati          20
BMW             10
Jawa            7
Indian          3
MV              3
Rajdoot         1
LML             1
Yezdi           1
Ideal           1
Name: count, dtype: int64
```

```
In [129... df
```

Out[129...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Royal Enfield
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triumph
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamaha
...
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	Hero
9369	Bajaj Avenger 220cc	35000.0	Bangalore	60000.0	First Owner	9.0	220.0	Bajaj
9370	Harley-Davidson Street 750 ABS	450000.0	Jodhpur	3430.0	First Owner	4.0	750.0	Harley Davidson
9371	Bajaj Dominar 400 ABS	139000.0	Hyderabad	21300.0	First Owner	4.0	400.0	Bajaj
9372	Bajaj Avenger Street 220	80000.0	Hyderabad	7127.0	First Owner	5.0	220.0	Bajaj

7324 rows × 8 columns

In [130...

```
# bhaya mere ko
# tvs
# kms 50000
# price = 50000
# owner = 1st
```

In [131...

```
tv_df = df[df['brand'] == 'TVS']
tv_df
```

Out[131...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
52	TVS Apache RTR 160cc	60000.0	Mumbai	30000.0	First Owner	5.0	160.0	TVS
114	TVS Apache RTR 160 4V Disc	69900.0	Delhi	8700.0	First Owner	3.0	160.0	TVS
130	TVS Phoenix Disc 125cc	21500.0	Barasat	10500.0	First Owner	5.0	125.0	TVS
...
9247	TVS Apache RTR 160cc Rear Disc	70000.0	Ghaziabad	4116.0	First Owner	3.0	160.0	TVS
9307	TVS Apache RTR 160cc	30000.0	Alibag	30000.0	First Owner	10.0	160.0	TVS
9312	TVS Apache RTR 200 4V FI	65450.0	Delhi	9238.0	First Owner	3.0	200.0	TVS
9320	TVS Apache 150cc	20000.0	Hissar	84916.0	First Owner	14.0	150.0	TVS
9322	TVS Radeon 110cc Drum SBT	58000.0	Delhi	4020.0	First Owner	2.0	110.0	TVS

481 rows × 8 columns

In [132...

```
kms_df = tvs_df[tvs_df['kms_driven'] <= 50000]
kms_df
```

Out[132...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TVS
52	TVS Apache RTR 160cc	60000.0	Mumbai	30000.0	First Owner	5.0	160.0	TVS
114	TVS Apache RTR 160 4V Disc	69900.0	Delhi	8700.0	First Owner	3.0	160.0	TVS
130	TVS Phoenix Disc 125cc	21500.0	Barasat	10500.0	First Owner	5.0	125.0	TVS
...
9192	TVS Apache RTR 160cc	60000.0	Hyderabad	20482.0	Second Owner	4.0	160.0	TVS
9247	TVS Apache RTR 160cc Rear Disc	70000.0	Ghaziabad	4116.0	First Owner	3.0	160.0	TVS
9307	TVS Apache RTR 160cc	30000.0	Alibag	30000.0	First Owner	10.0	160.0	TVS
9312	TVS Apache RTR 200 4V FI	65450.0	Delhi	9238.0	First Owner	3.0	200.0	TVS
9322	TVS Radeon 110cc Drum SBT	58000.0	Delhi	4020.0	First Owner	2.0	110.0	TVS

446 rows × 8 columns

In [133...

```
price_df = kms_df[kms_df['price'] <= 50000]
price_df
```

Out[133...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
130	TVS Phoenix Disc 125cc	21500.0	Barasat	10500.0	First Owner	5.0	125.0	TVS
131	TVS Apache RTR 160cc	40000.0	Delhi	20000.0	First Owner	6.0	160.0	TVS
215	TVS Star City 110cc	28000.0	Bangalore	28428.0	First Owner	9.0	110.0	TVS
235	TVS Sport 100cc	28000.0	Mandi	36000.0	First Owner	5.0	100.0	TVS
...
9143	TVS Star City 110cc	15000.0	Chennai	39888.0	First Owner	12.0	110.0	TVS
9155	TVS Star City 110cc	14000.0	Chennai	17602.0	First Owner	13.0	110.0	TVS
9157	TVS Star City 110cc	32000.0	Jaipur	17870.0	First Owner	7.0	110.0	TVS
9158	TVS MAX 4R 110cc	18000.0	Chennai	13673.0	First Owner	7.0	110.0	TVS
9307	TVS Apache RTR 160cc	30000.0	Alibag	30000.0	First Owner	10.0	160.0	TVS

221 rows × 8 columns

In [134...

```
owner_df = price_df[price_df['owner'] == "First Owner"]
owner_df
```

Out[134...

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TVS
130	TVS Phoenix Disc 125cc	21500.0	Barasat	10500.0	First Owner	5.0	125.0	TVS
131	TVS Apache RTR 160cc	40000.0	Delhi	20000.0	First Owner	6.0	160.0	TVS
215	TVS Star City 110cc	28000.0	Bangalore	28428.0	First Owner	9.0	110.0	TVS
235	TVS Sport 100cc	28000.0	Mandi	36000.0	First Owner	5.0	100.0	TVS
...
9143	TVS Star City 110cc	15000.0	Chennai	39888.0	First Owner	12.0	110.0	TVS
9155	TVS Star City 110cc	14000.0	Chennai	17602.0	First Owner	13.0	110.0	TVS
9157	TVS Star City 110cc	32000.0	Jaipur	17870.0	First Owner	7.0	110.0	TVS
9158	TVS MAX 4R 110cc	18000.0	Chennai	13673.0	First Owner	7.0	110.0	TVS
9307	TVS Apache RTR 160cc	30000.0	Alibag	30000.0	First Owner	10.0	160.0	TVS

200 rows × 8 columns

In [135...

```
# brand = "hero"
# price = under 70k
# age =<2 years
# first owner
# power 150cc
```

In [136...

```
Hero_df = df[df['brand'] == "Hero"]
Hero_df
```

Out[136...

	bike_name	price	city	kms_driven	owner	age	power	brand
7	Hero Splendor Plus Self Alloy 100cc	45000.0	Delhi	12645.0	First Owner	3.0	100.0	Hero
22	Hero Splendor iSmart Plus IBS 110cc	46500.0	Delhi	3500.0	First Owner	2.0	110.0	Hero
26	Hero Super Splendor 125cc	20000.0	Ahmedabad	29305.0	First Owner	16.0	125.0	Hero
48	Hero Hunk 150cc	37000.0	Mumbai	10800.0	First Owner	8.0	150.0	Hero
66	Hero CD Deluxe 100cc	12200.0	Agra	46643.0	First Owner	14.0	100.0	Hero
...
9316	Hero Glamour Fi 125cc	37000.0	Delhi	28478.0	First Owner	5.0	125.0	Hero
9339	Hero Splendor Plus 100cc	11400.0	Gurgaon	20000.0	Second Owner	17.0	100.0	Hero
9341	Hero CD Deluxe 100cc	25000.0	Sidhi	11122.0	First Owner	11.0	100.0	Hero
9343	Hero Passion Plus 100cc	24000.0	Hyderabad	68000.0	First Owner	14.0	100.0	Hero
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	Hero

1142 rows × 8 columns

In [137...

```
price_df = Hero_df[Hero_df['price'] <=70000]
price_df
```

Out[137...

	bike_name	price	city	kms_driven	owner	age	power	brand
7	Hero Splendor Plus Self Alloy 100cc	45000.0	Delhi	12645.0	First Owner	3.0	100.0	Hero
22	Hero Splendor iSmart Plus IBS 110cc	46500.0	Delhi	3500.0	First Owner	2.0	110.0	Hero
26	Hero Super Splendor 125cc	20000.0	Ahmedabad	29305.0	First Owner	16.0	125.0	Hero
48	Hero Hunk 150cc	37000.0	Mumbai	10800.0	First Owner	8.0	150.0	Hero
66	Hero CD Deluxe 100cc	12200.0	Agra	46643.0	First Owner	14.0	100.0	Hero
...
9316	Hero Glamour Fi 125cc	37000.0	Delhi	28478.0	First Owner	5.0	125.0	Hero
9339	Hero Splendor Plus 100cc	11400.0	Gurgaon	20000.0	Second Owner	17.0	100.0	Hero
9341	Hero CD Deluxe 100cc	25000.0	Sidhi	11122.0	First Owner	11.0	100.0	Hero
9343	Hero Passion Plus 100cc	24000.0	Hyderabad	68000.0	First Owner	14.0	100.0	Hero
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	Hero

1125 rows × 8 columns

In [138...

```
age_df = price_df[price_df['age'] <= 2]
age_df
```


Out[138...

	bike_name	price	city	kms_driven	owner	age	power	brand
22	Hero Splendor iSmart Plus IBS 110cc	46500.0	Delhi	3500.0	First Owner	2.0	110.0	Hero
121	Hero Splendor iSmart Plus IBS 110cc	48672.0	Ernakulam	608.0	First Owner	2.0	110.0	Hero
172	Hero Super Splendor 125cc	43778.0	Ernakulam	948.0	First Owner	2.0	125.0	Hero
176	Hero Splendor Plus Kick Alloy 100cc	48500.0	Delhi	5500.0	First Owner	2.0	100.0	Hero
197	Hero HF Deluxe i3s iBS 100cc	60000.0	Mumbai	4000.0	First Owner	2.0	100.0	Hero
328	Hero Passion Xpro 110cc	58000.0	Delhi	7758.0	First Owner	2.0	110.0	Hero
419	Hero Passion Pro 110cc Drum	56500.0	Delhi	10000.0	First Owner	2.0	110.0	Hero
492	Hero Splendor Plus IBS i3S 100cc	53300.0	Delhi	10000.0	First Owner	2.0	100.0	Hero
505	Hero Splendor Plus IBS i3S 100cc	65000.0	Delhi	2600.0	First Owner	2.0	100.0	Hero
708	Hero Splendor Plus Self Alloy 100cc	51300.0	Palakkad	2800.0	First Owner	2.0	100.0	Hero
787	Hero Xtreme 200R	66000.0	Faridabad	1112.0	First Owner	2.0	200.0	Hero
2479	Hero Splendor Plus IBS i3S 100cc	47167.0	Delhi	4664.0	First Owner	2.0	100.0	Hero
2631	Hero Splendor iSmart Plus	50000.0	Delhi	11584.0	First Owner	2.0	110.0	Hero

	bike_name	price	city	kms_driven	owner	age	power	brand
	IBS 110cc							
2654	Hero Passion Pro i3S Alloy 100cc	60000.0	Ghaziabad	1877.0	First Owner	2.0	100.0	Hero
2952	Hero Passion Pro i3S Alloy 100cc IBS	62000.0	Mumbai	1800.0	First Owner	2.0	100.0	Hero
3029	Hero HF Deluxe Eco 100cc	55000.0	Gurgaon	2309.0	First Owner	2.0	100.0	Hero
3298	Hero Splendor Plus IBS i3S 100cc	56000.0	Delhi	1745.0	First Owner	2.0	100.0	Hero
3882	Hero Splendor Plus Self Alloy 100cc	55000.0	Faridabad	4195.0	First Owner	2.0	100.0	Hero
4055	Hero Splendor Plus Self Alloy 100cc	55000.0	Gurgaon	5238.0	First Owner	2.0	100.0	Hero
4512	Hero HF Deluxe Eco 100cc	41000.0	Faridabad	4544.0	First Owner	2.0	100.0	Hero
5933	Hero Super Splendor 125cc	57000.0	Noida	3200.0	First Owner	2.0	125.0	Hero
6613	Hero HF Deluxe self Alloy 100cc	39500.0	Faridabad	5448.0	First Owner	2.0	100.0	Hero
6981	Hero Passion Pro 110cc Drum	56000.0	Delhi	1168.0	First Owner	2.0	110.0	Hero
7381	Hero HF Deluxe Eco 100cc	46000.0	Faridabad	11309.0	First Owner	2.0	100.0	Hero

```
In [139... owner_df= age_df[age_df['owner'] == "First Owner"]
owner_df
```

Out[139...

	bike_name	price	city	kms_driven	owner	age	power	brand
22	Hero Splendor iSmart Plus IBS 110cc	46500.0	Delhi	3500.0	First Owner	2.0	110.0	Hero
121	Hero Splendor iSmart Plus IBS 110cc	48672.0	Ernakulam	608.0	First Owner	2.0	110.0	Hero
172	Hero Super Splendor 125cc	43778.0	Ernakulam	948.0	First Owner	2.0	125.0	Hero
176	Hero Splendor Plus Kick Alloy 100cc	48500.0	Delhi	5500.0	First Owner	2.0	100.0	Hero
197	Hero HF Deluxe i3s iBS 100cc	60000.0	Mumbai	4000.0	First Owner	2.0	100.0	Hero
328	Hero Passion Xpro 110cc	58000.0	Delhi	7758.0	First Owner	2.0	110.0	Hero
419	Hero Passion Pro 110cc Drum	56500.0	Delhi	10000.0	First Owner	2.0	110.0	Hero
492	Hero Splendor Plus IBS i3S 100cc	53300.0	Delhi	10000.0	First Owner	2.0	100.0	Hero
505	Hero Splendor Plus IBS i3S 100cc	65000.0	Delhi	2600.0	First Owner	2.0	100.0	Hero
708	Hero Splendor Plus Self Alloy 100cc	51300.0	Palakkad	2800.0	First Owner	2.0	100.0	Hero
787	Hero Xtreme 200R	66000.0	Faridabad	1112.0	First Owner	2.0	200.0	Hero
2479	Hero Splendor Plus IBS i3S 100cc	47167.0	Delhi	4664.0	First Owner	2.0	100.0	Hero
2631	Hero Splendor iSmart Plus	50000.0	Delhi	11584.0	First Owner	2.0	110.0	Hero

	bike_name	price	city	kms_driven	owner	age	power	brand
	IBS 110cc							
2654	Hero Passion Pro i3S Alloy 100cc	60000.0	Ghaziabad	1877.0	First Owner	2.0	100.0	Hero
2952	Hero Passion Pro i3S Alloy 100cc IBS	62000.0	Mumbai	1800.0	First Owner	2.0	100.0	Hero
3029	Hero HF Deluxe Eco 100cc	55000.0	Gurgaon	2309.0	First Owner	2.0	100.0	Hero
3298	Hero Splendor Plus IBS i3S 100cc	56000.0	Delhi	1745.0	First Owner	2.0	100.0	Hero
3882	Hero Splendor Plus Self Alloy 100cc	55000.0	Faridabad	4195.0	First Owner	2.0	100.0	Hero
4055	Hero Splendor Plus Self Alloy 100cc	55000.0	Gurgaon	5238.0	First Owner	2.0	100.0	Hero
4512	Hero HF Deluxe Eco 100cc	41000.0	Faridabad	4544.0	First Owner	2.0	100.0	Hero
5933	Hero Super Splendor 125cc	57000.0	Noida	3200.0	First Owner	2.0	125.0	Hero
6613	Hero HF Deluxe self Alloy 100cc	39500.0	Faridabad	5448.0	First Owner	2.0	100.0	Hero
6981	Hero Passion Pro 110cc Drum	56000.0	Delhi	1168.0	First Owner	2.0	110.0	Hero
7381	Hero HF Deluxe Eco 100cc	46000.0	Faridabad	11309.0	First Owner	2.0	100.0	Hero

```
In [140... power_df = owner_df[owner_df['power'] == 150]
power_df
```

```
Out[140... bike_name price city kms_driven owner age power brand
```

In [141...

df

Out[141...

	bike_name	price	city	kms_driven	owner	age	power	bran
0	TVS Star City Plus Dual Tone 110cc	35000.0	Ahmedabad	17654.0	First Owner	3.0	110.0	TV
1	Royal Enfield Classic 350cc	119900.0	Delhi	11000.0	First Owner	4.0	350.0	Roy Enfie
2	Triumph Daytona 675R	600000.0	Delhi	110.0	First Owner	8.0	675.0	Triump
3	TVS Apache RTR 180cc	65000.0	Bangalore	16329.0	First Owner	4.0	180.0	TV
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000.0	Bangalore	10000.0	First Owner	3.0	150.0	Yamat
...	
9362	Hero Hunk Rear Disc 150cc	25000.0	Delhi	48587.0	First Owner	8.0	150.0	He
9369	Bajaj Avenger 220cc	35000.0	Bangalore	60000.0	First Owner	9.0	220.0	Baj
9370	Harley-Davidson Street 750 ABS	450000.0	Jodhpur	3430.0	First Owner	4.0	750.0	Harle Davidsc
9371	Bajaj Dominar 400 ABS	139000.0	Hyderabad	21300.0	First Owner	4.0	400.0	Baj
9372	Bajaj Avenger Street 220	80000.0	Hyderabad	7127.0	First Owner	5.0	220.0	Baj

7324 rows × 8 columns

In [142...

df.drop(['bike_name','city'],axis=1,inplace=True)

In [143...

df

Out[143...

	price	kms_driven	owner	age	power	brand
0	35000.0	17654.0	First Owner	3.0	110.0	TVS
1	119900.0	11000.0	First Owner	4.0	350.0	Royal Enfield
2	600000.0	110.0	First Owner	8.0	675.0	Triumph
3	65000.0	16329.0	First Owner	4.0	180.0	TVS
4	80000.0	10000.0	First Owner	3.0	150.0	Yamaha
...
9362	25000.0	48587.0	First Owner	8.0	150.0	Hero
9369	35000.0	60000.0	First Owner	9.0	220.0	Bajaj
9370	450000.0	3430.0	First Owner	4.0	750.0	Harley-Davidson
9371	139000.0	21300.0	First Owner	4.0	400.0	Bajaj
9372	80000.0	7127.0	First Owner	5.0	220.0	Bajaj

7324 rows × 6 columns

In [144...

```
owner_dict = {
    'First Owner':1,
    'Second Owner':2,
    'Third Owner':3,
    'Fourth Owner or more':4
}
```

In [145...

owner_dict

Out[145...

```
{'First Owner': 1,
 'Second Owner': 2,
 'Third Owner': 3,
 'Fourth Owner or more': 4}
```

In [146...

```
df['owner'] = df['owner'].map(owner_dict)
```

In [147...

df

Out[147...

	price	kms_driven	owner	age	power	brand
0	35000.0	17654.0	1.0	3.0	110.0	TVS
1	119900.0	11000.0	1.0	4.0	350.0	Royal Enfield
2	600000.0	110.0	1.0	8.0	675.0	Triumph
3	65000.0	16329.0	1.0	4.0	180.0	TVS
4	80000.0	10000.0	1.0	3.0	150.0	Yamaha
...
9362	25000.0	48587.0	1.0	8.0	150.0	Hero
9369	35000.0	60000.0	1.0	9.0	220.0	Bajaj
9370	450000.0	3430.0	1.0	4.0	750.0	Harley-Davidson
9371	139000.0	21300.0	1.0	4.0	400.0	Bajaj
9372	80000.0	7127.0	1.0	5.0	220.0	Bajaj

7324 rows × 6 columns

In [148...

`df.info()`

```
<class 'pandas.core.frame.DataFrame'>
Index: 7324 entries, 0 to 9372
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  -
0   price       7324 non-null   float64
1   kms_driven  7324 non-null   float64
2   owner       7314 non-null   float64
3   age         7324 non-null   float64
4   power       7324 non-null   float64
5   brand       7324 non-null   object
dtypes: float64(5), object(1)
memory usage: 400.5+ KB
```

In [149...

```
brand_dict = {
    'TVS':1,
    'Royal Enfield':2,
    'Triumph':3,
    'Yamaha':4,
    'Honda':5,
    'Hero':6,
    'Bajaj':7,
    'Suzuki':8,
    'Benelli':9,
    'KTM':10,
    'Mahindra':11,
    'Kawasaki':12,
    'Ducati':13,
```

```
'Hyosung':14,  
'Harley-Davidson':15,  
'Jawa':16,  
'BMW':17,  
'Indian':18,  
'Rajdoot':19,  
'LML':20,  
'Yezdi':21,  
'MV':22,  
'Ideal':23  
}
```

In [150... brand_dict

```
Out[150... {'TVS': 1,  
            'Royal Enfield': 2,  
            'Triumph': 3,  
            'Yamaha': 4,  
            'Honda': 5,  
            'Hero': 6,  
            'Bajaj': 7,  
            'Suzuki': 8,  
            'Benelli': 9,  
            'KTM': 10,  
            'Mahindra': 11,  
            'Kawasaki': 12,  
            'Ducati': 13,  
            'Hyosung': 14,  
            'Harley-Davidson': 15,  
            'Jawa': 16,  
            'BMW': 17,  
            'Indian': 18,  
            'Rajdoot': 19,  
            'LML': 20,  
            'Yezdi': 21,  
            'MV': 22,  
            'Ideal': 23}
```

In []: # df['brand']=df['brand'].map(brand_dict)

In [104... df

Out[104...

	price	kms_driven	owner	age	power	brand
0	35000.0	17654.0	1.0	3.0	110.0	1
1	119900.0	11000.0	1.0	4.0	350.0	2
2	600000.0	110.0	1.0	8.0	675.0	3
3	65000.0	16329.0	1.0	4.0	180.0	1
4	80000.0	10000.0	1.0	3.0	150.0	4
...
9362	25000.0	48587.0	1.0	8.0	150.0	6
9369	35000.0	60000.0	1.0	9.0	220.0	7
9370	450000.0	3430.0	1.0	4.0	750.0	15
9371	139000.0	21300.0	1.0	4.0	400.0	7
9372	80000.0	7127.0	1.0	5.0	220.0	7

7324 rows × 6 columns

In [105...

```
df.dtypes
```

Out[105...

```
price      float64
kms_driven float64
owner      float64
age        float64
power      float64
brand      int64
dtype: object
```

In [151...

```
cat_col=df.select_dtypes(include='object')
```

In [152...

```
cat_col
```

Out[152...

	brand
0	TVS
1	Royal Enfield
2	Triumph
3	TVS
4	Yamaha
...	...
9362	Hero
9369	Bajaj
9370	Harley-Davidson
9371	Bajaj
9372	Bajaj

7324 rows × 1 columns

In [153...

```
num_col=df.select_dtypes(exclude='object')
```

In [154...

```
num_col
```

Out[154...

	price	kms_driven	owner	age	power
0	35000.0	17654.0	1.0	3.0	110.0
1	119900.0	11000.0	1.0	4.0	350.0
2	600000.0	110.0	1.0	8.0	675.0
3	65000.0	16329.0	1.0	4.0	180.0
4	80000.0	10000.0	1.0	3.0	150.0
...
9362	25000.0	48587.0	1.0	8.0	150.0
9369	35000.0	60000.0	1.0	9.0	220.0
9370	450000.0	3430.0	1.0	4.0	750.0
9371	139000.0	21300.0	1.0	4.0	400.0
9372	80000.0	7127.0	1.0	5.0	220.0

7324 rows × 5 columns

In []: