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

In [38]:
```
#Loading CSV
data=pd.read_csv('cars.csv')
data.head()
```

Out[38]:

	Brand	Price	Body	Mileage	EngineV	Engine Type	Registration	Year	Model
0	BMW	4200.0	sedan	277	2.0	Petrol	yes	1991	320
1	Mercedes-Benz	7900.0	van	427	2.9	Diesel	yes	1999	Sprinter 212
2	Mercedes-Benz	13300.0	sedan	358	5.0	Gas	yes	2003	S 500
3	Audi	23000.0	crossover	240	4.2	Petrol	yes	2007	Q7
4	Toyota	18300.0	crossover	120	2.0	Petrol	yes	2011	Rav 4

In [3]:
```
#Checking Null Values
data.isnull()
```

Out[3]:

	Brand	Price	Body	Mileage	EngineV	Engine Type	Registration	Year	Model
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
...	...	...	...	...	...	...	...	...	...
4340	False	False	False	False	False	False	False	False	False
4341	False	False	False	False	False	False	False	False	False
4342	False	False	False	False	False	False	False	False	False
4343	False	False	False	False	True	False	False	False	False
4344	False	False	False	False	False	False	False	False	False

4345 rows × 9 columns

In [4]:
```
#remove dupicate
data.duplicated()
```

Out[4]:
```
0       False
1       False
2       False
3       False
4       False
       ...
4340    False
4341    False
4342    False
4343    False
4344    False
Length: 4345, dtype: bool
```

In [7]:
```
#Reading specific columns using multy axex indexing
print (data.loc[[1,2,3,4,5,6,7,8,9,10],['Brand','Price']])
```

```
           Brand     Price
1   Mercedes-Benz    7900.0
2   Mercedes-Benz   13300.0
3            Audi   23000.0
4          Toyota   18300.0
5   Mercedes-Benz  199999.0
6             BMW    6100.0
7            Audi   14200.0
8         Renault   10799.0
9      Volkswagen    1400.0
10        Renault   11950.0
```

In [9]:
```
#Grouping Function
grouped=data.groupby('Year')
print(grouped.get_group(2014))
```

```
              Brand     Price       Body  Mileage  EngineV Engine Type  \
38             Audi   33900.0      other       14      1.8      Petrol
51          Renault    8250.0        van       66      1.5      Diesel
83       Volkswagen   13708.5      hatch       51      1.4      Petrol
99       Volkswagen   31500.0  crossover        1      2.0      Petrol
111             BMW   47000.0      other        5      2.0      Petrol
...             ...       ...        ...      ...      ...         ...
4291  Mercedes-Benz   28950.0      sedan       40      1.8      Petrol
4305            BMW       NaN  crossover        1      3.0      Diesel
4339         Toyota   17900.0      sedan       35      1.6      Petrol
4340  Mercedes-Benz  125000.0      sedan        9      3.0      Diesel
4343         Toyota   14200.0      sedan       31      NaN      Petrol

     Registration  Year    Model
38            yes  2014       TT
51            yes  2014   Kangoo
83            yes  2014     Polo
99            yes  2014   Tiguan
111           yes  2014       Z4
...           ...   ...      ...
4291          yes  2014  C-Class
4305          yes  2014       X6
4339          yes  2014  Corolla
4340          yes  2014    S 350
4343          yes  2014  Corolla

[167 rows x 9 columns]
```

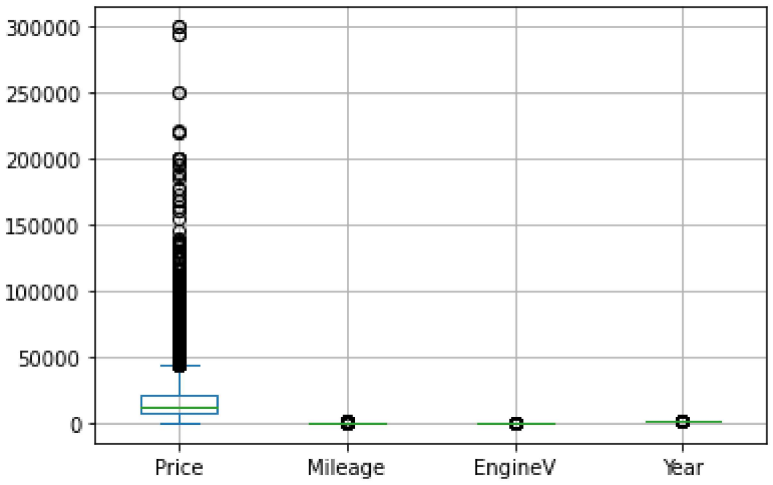
In [27]:
```
#data analysis
#mean
import statistics
print (np.mean(data.Mileage))
```

```
161.2372842347526
```

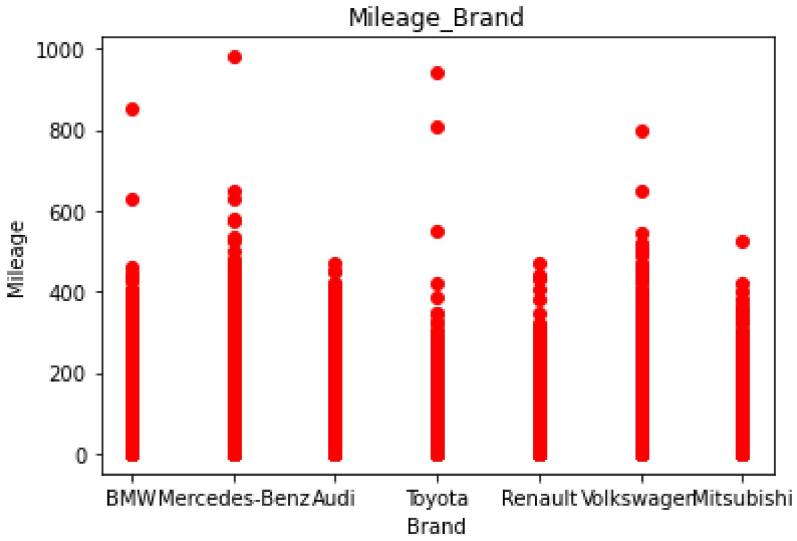
In [34]:
```
#mode
print(data.mode)
```

```
<bound method DataFrame.mode of              Brand     Price       Body  Mileage  En
gineV Engine Type  \
0              BMW    4200.0      sedan      277      2.0      Petrol
1     Mercedes-Benz    7900.0        van      427      2.9      Diesel
2     Mercedes-Benz   13300.0      sedan      358      5.0         Gas
3              Audi   23000.0  crossover      240      4.2      Petrol
4            Toyota   18300.0  crossover      120      2.0      Petrol
...             ...       ...        ...      ...      ...         ...
4340  Mercedes-Benz  125000.0      sedan        9      3.0      Diesel
4341            BMW    6500.0      sedan        1      3.5      Petrol
4342            BMW    8000.0      sedan      194      2.0      Petrol
4343         Toyota   14200.0      sedan       31      NaN      Petrol
4344     Volkswagen   13500.0        van      124      2.0      Diesel

     Registration  Year              Model
0             yes  1991                320
1             yes  1999       Sprinter 212
2             yes  2003              S 500
3             yes  2007                 Q7
4             yes  2011              Rav 4
...           ...   ...                ...
4340          yes  2014              S 350
4341          yes  1999                535
4342          yes  1985                520
4343          yes  2014            Corolla
4344          yes  2013  T5 (Transporter)

[4345 rows x 9 columns]>
```

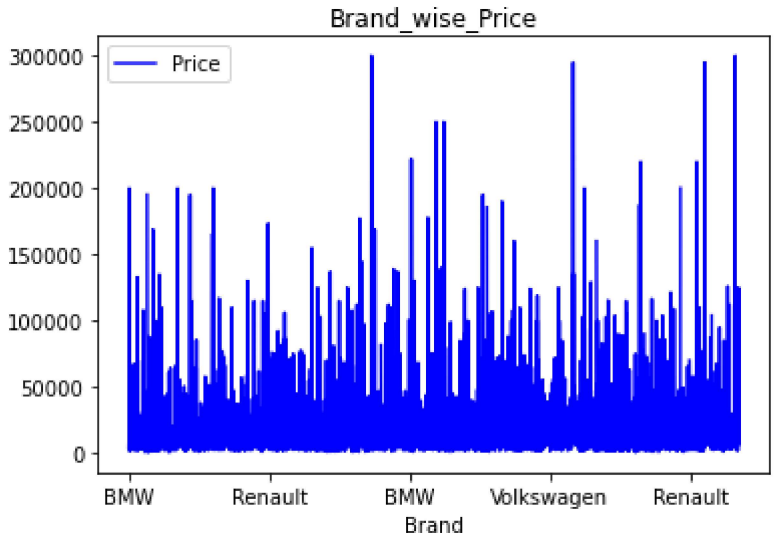
In [35]:
```
#Data Visualisation
#box plot
data.plot.box(grid='True')
mp.savefig('boxplot.png')
```



In [51]:
```
mp.scatter(data.Brand,data.Mileage,color='red')
mp.xlabel('Brand')
mp.ylabel('Mileage')
mp.title('Mileage_Brand')
mp.savefig('scatter.png')
```



In [49]:
```
#bar plot
data.plot(x='Brand',y='Price',color='blue')
mp.title('Brand_wise_Price')
mp.savefig('bar.png')
```



In [ ]: