In [1]:
```
import pandas as pd
import numpy as np
df = pd.read_csv("1.csv", keep_default_na=True)
```

In [2]:
```
#Q1
df.tail(5)
```

Out[2]:

| | index | company | body-style | wheel-base | length | engine-type | num-of-cylinders | horsepower | average-mileage | price |
|---|---|---|---|---|---|---|---|---|---|---|
| **56** | 81 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 85 | 27 | 7975.0 |
| **57** | 82 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 52 | 37 | 7995.0 |
| **58** | 86 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 100 | 26 | 9995.0 |
| **59** | 87 | volvo | sedan | 104.3 | 188.8 | ohc | four | 114 | 23 | 12940.0 |
| **60** | 88 | volvo | wagon | 104.3 | 188.8 | ohc | four | 114 | 23 | 13415.0 |

In [3]:
```
#Q2
df = pd.read_csv("1.csv", keep_default_na=False)
df=df.replace(r'^\s*$', np.nan, regex=True)
df.to_csv("1.csv",index=False)
```

In [4]:
```
#Q3
df = pd.read_csv("1.csv")
print(df[df["company"]=="bmw"])
```

```
    index company body-style  wheel-base  length engine-type num-of-cylinders  \
7       9     bmw      sedan        101.2   176.8         ohc             four
8      10     bmw      sedan        101.2   176.8         ohc             four
9      11     bmw      sedan        101.2   176.8         ohc              six
10     13     bmw      sedan        103.5   189.0         ohc              six
11     14     bmw      sedan        103.5   193.8         ohc              six
12     15     bmw      sedan        110.0   197.0         ohc              six

    horsepower  average-mileage    price
7          101               23  16430.0
8          101               23  16925.0
9          121               21  20970.0
10         182               16  30760.0
11         182               16  41315.0
12         182               15  36880.0
```

In [5]:
```
#Q4
print(df.groupby('company').size())
```

```
company
alfa-romero      3
audi             4
bmw              6
chevrolet        3
dodge            2
honda            3
isuzu            3
jaguar           3
mazda            5
mercedes-benz    4
mitsubishi       4
nissan           5
porsche          3
toyota           7
volkswagen       4
```

```
volvo            2
dtype: int64
```

In [6]:
```
#Q5
print(df.groupby("company")['price'].max())
```

```
company
alfa-romero      16500.0
audi             18920.0
bmw              41315.0
chevrolet         6575.0
dodge             6377.0
honda            12945.0
isuzu             6785.0
jaguar           36000.0
mazda            18344.0
mercedes-benz    45400.0
mitsubishi        8189.0
nissan           13499.0
porsche          37028.0
toyota           15750.0
volkswagen        9995.0
volvo            13415.0
Name: price, dtype: float64
```

In [7]:
```
#Q6
print(df.groupby('company')['average-mileage'].mean())
```

```
company
alfa-romero      20.333333
audi             20.000000
bmw              19.000000
chevrolet        41.000000
dodge            31.000000
honda            26.333333
isuzu            33.333333
jaguar           14.333333
mazda            28.000000
mercedes-benz    18.000000
mitsubishi       29.500000
nissan           31.400000
porsche          17.000000
toyota           28.714286
volkswagen       31.750000
volvo            23.000000
Name: average-mileage, dtype: float64
```

In [8]:
```
#Q7
Car_Price = {'Company': ['Toyota', 'Honda', 'BMV', 'Audi'], 'Price': [23845, 17995,
135925, 71400]}
Car_Horsepower = {'Company': ['Toyota', 'Honda', 'BMV', 'Audi'], 'horsepower': [141,
182, 160]}
df1 = pd.DataFrame(Car_Price, columns = ['Company', 'Price'])
df2 = pd.DataFrame(Car_Horsepower, columns = ['Company', 'horsepower'])
df3 = df1.set_index('Company').join(df2.set_index('Company'))
df3
```

Out[8]:

| Company | Price | horsepower |
|---|---|---|
| **Toyota** | 23845 | 141 |
| **Honda** | 17995 | 80 |
| **BMV** | 135925 | 182 |
| **Audi** | 71400 | 160 |

# Extra Questions

Q8 Sort alphabetically as per the car names
Q9 Extract the max price of car by each company and append it to a different data file, max_price.csv

In [9]:
```
#Q8
df = pd.read_csv("1.csv", keep_default_na=True)
df.sort_values(by='company')
```

Out[9]:

| | index | company | body-style | wheel-base | length | engine-type | num-of-cylinders | horsepower | average-mileage | pri |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | alfa-romero | convertible | 88.6 | 168.8 | dohc | four | 111 | 21 | 13495 |
| **1** | 1 | alfa-romero | convertible | 88.6 | 168.8 | dohc | four | 111 | 21 | 16500 |
| **2** | 2 | alfa-romero | hatchback | 94.5 | 171.2 | ohcv | six | 154 | 19 | 16500 |
| **3** | 3 | audi | sedan | 99.8 | 176.6 | ohc | four | 102 | 24 | 13950 |
| **4** | 4 | audi | sedan | 99.4 | 176.6 | ohc | five | 115 | 18 | 17450 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **55** | 80 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 52 | 37 | 7775 |
| **56** | 81 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 85 | 27 | 7975 |
| **58** | 86 | volkswagen | sedan | 97.3 | 171.7 | ohc | four | 100 | 26 | 9995 |
| **59** | 87 | volvo | sedan | 104.3 | 188.8 | ohc | four | 114 | 23 | 12940 |
| **60** | 88 | volvo | wagon | 104.3 | 188.8 | ohc | four | 114 | 23 | 13415 |

61 rows × 10 columns

In [10]:
```
#Q9
df1=df.groupby("company")['price'].max()
df1.to_csv("max_price.csv",index=False)
```

In [ ]: