Chapter 2 - Data Preparation Basics Segment 5 - Grouping and data aggregation

In [3]:
```
import numpy as np
import pandas as pd
from pandas import Series, DataFrame
```

Grouping data by column index

In [4]:
```
address = 'C:/Users/danal/Desktop/ExerciseFiles/Data/mtcars.csv'

cars = pd.read_csv(address)

cars.columns = ['car_names', 'mpg', 'cyl', 'disp', 'hp', 'drat', 'wt', 'qsec', 'v
cars.head()
```

Out[4]:

	car_names	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2

In [6]:
```
cars_groups = cars.groupby(cars['cyl'])
cars_groups.mean()
```

Out[6]:

cyl	mpg	disp	hp	drat	wt	qsec	vs	am	gea
4	26.663636	105.136364	82.636364	4.070909	2.285727	19.137273	0.909091	0.727273	4.09090
6	19.742857	183.314286	122.285714	3.585714	3.117143	17.977143	0.571429	0.428571	3.85714
8	15.100000	353.100000	209.214286	3.229286	3.999214	16.772143	0.000000	0.142857	3.28571