Econometrics II [...]

https://github.com/minyoungrho/Econometrics2 (https://github.com/minyoungrho/Econometrics2)

Econometrics

Let's look at some data

In [4]:
```
using CSV
using DataFrames
begin
    data = DataFrame(CSV.File("../data/qpm.csv"; datarow=1))
    first(data,10)
end
```

Out[4]: 10 rows × 3 columns

	Column1	Column2	Column3
	Float64	Float64	Float64
1	59.7417	39.9898	0.300883
2	59.987	39.9811	0.437111
3	59.4584	40.5011	0.545672
4	59.5963	39.8678	0.574567
5	60.7715	40.399	0.604033
6	59.7878	39.4423	0.629612
7	61.1404	41.0072	0.724331
8	59.7776	40.7407	0.742034
9	60.9996	40.5388	0.813239
10	62.0706	40.9443	0.822884

What are these data? How were these data points generated?

A theoretical (economic) model, also known as a data generateing procee (DGP), is a key ingredient to assign **causal relationships**.

The variables we were looking at are:

- Quantity (q)
- Price (p)
- Income (m)

The data was generated using the following economic theory.

Economic theory tells us that the quantity of a good that consumers will purchase (the demand function) is something like:

$$q = d(p, m, Z)$$

where

- q is the quantity demanded
- p is the price of the good
- m is the income
- Z is other variables that may affect demand The supply of the good to the market is the aggregation of the firms' supply functions which looks something like:

$$q = s(p, V)$$

- q is quantity supplied
- V is other variables that may affect supply

This is the basic economic model of supply and demand: q and p are determined in the market equilibrium, given by the intersection of the two curves.

- These two variables are determined jointly by the model and are called the **endogenous variables**.
- Income (m) is not determined by this model, or its value is determined independently of q and p, and is called **exogenous variables**.
- m causes p and q; p and q do not cause m; p and q have a joint causal relationship

The model is essentially a theoretical construct up to now. Throughout this course, we will attempt to quantify these theoretical relationships more precisely. For example,

- Model and estimate functional forms of s and d
- Divide Z into components that are observable and non-observable

For example, OLS

$$q_i = \alpha_1 + \alpha_2 p_i + \alpha_3 m_i + \epsilon_i$$
$$q_j = \beta_1 + \beta_2 p_j + \epsilon_j$$

- the functions d and s have been specified (as a linear function, remember OLS?)
- the parameters are in place and constant across consumers and firms
- there exist an (additively) unobservable component which make up the difference between the realized demand/supply (a.k.a. data) and our model
- $E[\epsilon_i] = 0$ and $E[m_i \epsilon_i] = 0$

In this course, we will generalize and study estimation of any structural (economic) models. Let us first focus on extreme estimators.