# Article 6.1

The file named Article6_1.xlsx contains data for US firms for the period 2011-2019. The variables included in the dataset are:

- Investment is the ratio of capital expenditure to total asset;
- Mtb is firm's market to book;
- Leverage is the amount of total debt to total asset;
- Cashflow is the cashflow of each firm to total asset;
- Cashtoasset is the ratio between cash and cash equivalents to total asset;
- Total Asset.

**QUESTIONS:**

1. The structure of the data is a panel. Sort the data by ID and time.
2. Generate a new variable "size". A good proxy for the size of the firm is the log of total asset. So, calculate "size" as the natural logarithm of total asset.
3. Some literature argues that if firms hold an amount of cash and cash equivalents not smaller than the 25% of their total assets, then they should be considered cash conservative. Construct a dummy variable, which takes the value of 1 if the value in the variable cashtoasset is equal or greater than 0.25 and 0 otherwise.
4. Calculate the summary statistics for the variables above. Comment the values obtained. Which is the probability of being cash conservative firms?
5. Derive the correlation matrix for the main variables.
6. Using the pivot table, calculate the average of investments for cash conservative and non-cash conservative firms.
7. Create a dummy variable which takes the value of 1 if size is equal or greater than its average (big size firms) and 0 otherwise (small size firms)
8. Using the pivot table, calculate the average of investments for different sizes and for cash and non-cash conservative firms.

9. Calculate the mean of the variable cashflow for each year. Plot the obtained values in an appropriate graph.

10. Sheet2 contains the variable “Sales” which is the ln of net sales, and it is employed as a proxy of firms’ revenues. Incorporate this variable into the main dataset (sheet “data”) using VLOOKUP function.

11. Estimate the following equation using OLS:

$$investment_{it} = \beta_0 + \beta_1 cashflow_{it} + \beta_2 mtb_{it} + \varepsilon_{it}$$