# Homework Set 1

## Problem 1

In this problem you will make use of the S&P 500 index data, found here:
`http://finance.yahoo.com/q/hp?s=^GSPC+Historical+Prices`.
You can also search for "s&p 500 yahoo finance" on Google, click the top link and then click "Historical Prices" on the left of the page. Click "Download to Spreadsheet" to download the data.

Load the last column "Adj. Close" of the data into your favorite software. Note that the data is formatted so that the most recent price is at the top; you may need to reverse the prices to put them in chronological order.

1. Generate a plot of the closing prices of the S&P 500 index.

2. Convert the close prices $P_t$ into logarithmic returns: $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$
   Make a plot of the log returns. Why do we use log returns?

3. Draw a set of 1000 data points from a Gaussian distribution with mean $\mu = 100$ and standard deviation $\sigma = 1$. Consider this new set of data points to be the closing prices of the index. Make a plot of the log returns. What is different compared to your previous result?

## Problem 2

The Zipf's law can be expressed as $r^b f = a$ where $r$ is the rank, $f$ is the frequency and $a, b$ are constants. Remember that is convenient to use log-log when plotting the Zipf's law.

1. Give some examples where the Zipf's law is applicable.

2. Choose a country you care about (yours!) and make a Zipf population-rank plot for the largest N cities. Increase N until the plot is no more linear, and speculate why the plot ceases to be linear.

3. Rank the top N stocks of your favorite stock exchange based on their capitalization. Using a reasonable value for N, create a log-log plot of their capitalization versus their rank.

## Bonus problem

Watch this video on YouTube and try to apply Benford's law in the population of the countries of the world.

https://www.youtube.com/watch?v=XXjlR2OK1kM

A way to exrtact the leading digit $d$ of a number $x$ is

$$d = floor\left(\frac{|x|}{10^{floor(\log_{10}|x|)}}\right)$$

where "floor" means rounding toward negative infinity. For example, $floor(5.78) = 5$.

**Guidelines**

You are free to submit an electronic or a handwritten copy of your homework. Please don't forget to attach your graphs and your code in whatever software you used. If you have ANY question please do not hesitate to ask me. I will be happy to help you.
Have fun!

Contact: Alexandros Kyrtsos, akyrtsos@bu.edu