Exercise 2

1. Go here and download the "25 Portfolios Formed on Size and Book-to-Market (5 x 5)" (not the ex. Dividends or Daily series) and inspect the data file (open it up in Excel or something else).
2. What do you notice about the file?
3. Read the first series "Average Value Weighted Returns – Monthly" into a Pandas dataframe
4. Extract all the columns that say "LoBM" or "BM1." BM means Book-to-market! ME means "market equity". ME is measure how big or small a company's equity value is.
5. Find the average value of each of the 5 series
6. Using Matrix operations (you don't need for-loops for this, at least within the summation you see below. Check out NumPy as a tool for executing matrix operations) find the quantity

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

   Where x and y are each one of the 5 series (SMALL LoBM, ME2 BM1, etc). Do this for all pairs of the 5 series, *including* pairing each variable with itself. Store the results in a 5x5 matrix where each row represents a series and each column a series. See below for an example:

$$\mathbf{\Sigma} = \begin{array}{ccc} & x & y \\ x & \sigma_{xx} & \sigma_{xy} \\ y & \sigma_{yx} & \sigma_{yy} \end{array}$$

   Notice that $\sigma_{yx} = \sigma_{xy}$
7. What you've just found is called the (co)variance. It tells you how two different series move with each other. If covariance is positive, then when one of the variables (x) is high, so is the other one (y). Vice versa if covariance is negative. This quantity is fundamental to portfolio management as well as statistical analysis. Remember what we talked about the first day, that a lot of econometric analysis is understanding how one variable changes as a different one changes, ie how the 'covary.'

   Now say you have a matrix $\boldsymbol{b}_{(n+1)x1} = \begin{pmatrix} \mathbf{0}_{\boldsymbol{n\,x\,1}} \\ 1 \end{pmatrix}$ where $\boldsymbol{b}$ is an $(n+1)\ x\ 1$ vector. Say you also have $\boldsymbol{A} = \begin{pmatrix} 2\mathbf{\Sigma} & \mathbf{1}_{\boldsymbol{nx1}} \\ \mathbf{1}_{\boldsymbol{nx1}}^{\boldsymbol{T}} & 0 \end{pmatrix}$ where $\mathbf{\Sigma}$ comes from the operation in part e. I want you to find $\boldsymbol{A}^{-1}\boldsymbol{b}$**.** The first n elements of this would be the portfolio weights to create a minimum variance portfolio. That is the lowest risk possible (like we did last time with 2 assets, but this one has 5 assets, the 5 portfolios). A portfolio weight is the decimal percentage of the dollars invested in the portfolio that is invested in each asset (eg. $w_1 = 0.2$ means that 20% of the dollars are invested in asset 1. Etc.) What do you notice about the weights? What does this mean?
8. If the return in each month t is $r_{1,t}w_{1,t-1} + r_{2,t}w_{2,t-1} + \ldots$ for all the assets in the portfolio, find the portfolio return for the entire time period. Assume you rebalance the portfolio (meaning make sure the weights are equal to the values we calculated in 7) at the end of each month. Why is this necessary? What alternative rebalancing schemes might you follow and what are the advantages/disadvantages of each?
9. Plot the portfolio VALUE assuming $100 invested at the beginning of the time period
10. Now do the same thing for the natural log of the portfolio value, ie $\ln(portfolioValue_t)$ for all times t.