Exercise 5

We are going to use our new regression skills to evaluate the performance of Berkshire Hathaway, Warren Buffet's company. It's a publicly traded company so we can download its returns. They are in the "Berkshire.txt" data file linked in the assignment.

1. As always, INSPECT YOUR DATA!!! I shouldn't have to tell you to do this, but on the off chance some of you haven't gotten in the habit…INSPECT YOUR DATA!!! Do your data have the same periods? Are they in the same units? Are there other oddities with the file? Are there column headers? Etc.
2. Throughout this assignment, we are looking at an asset's *excess* return, which is what the return was *in excess* of the risk-free rate (basically treasury bills)

   $$r_{it}^e = r_{it} - r_{ft}$$

   Where r_{ft} is the risk-free rate at time t. This value is in the factor return data.
   We can think about decomposing a firm's excess return into a *factor* based component, and an idiosyncratic component. The factor component is saying a company's return is based on the fact that the firm is exposed to some risks, and those risks are priced by the market, and the idiosyncratic piece is just randomness, effectively. That is

   $$r_{it}^e = FactorReturn_{it} + IdiosyncraticReturn_{it}$$

   The idiosyncratic component itself will then be decomposed into a piece of constant outperformance, α_i, and a truly unexpected idiosyncratic piece, u_{it}

   $$r_{it}^e = \alpha_i + FactorReturn_{it} + u_{it}$$

   We can use these ideas to evaluate the performance of a stock (and therefore it's manager).

   Load both the factor returns and the Berkshire returns. Check the sample periods, and make sure they line up. IE restrict the dates on the larger sample to match the dates of the smaller one! Also, CHECK YOUR UNITS!

   Find Berkshire's average excess return. This gives another return decomposition:

   $$r_{it}^e = E(r_{it}^e) + u_{it}$$

   Which is basically saying when we take the average of a sample of excess returns, we are finding an estimate of the stocks *risk premium*. Effectively, we are taking noisy data (the observed r_{it}^e) and extracting the *signal* (the $E(r_{it}^e)$) and leaving behind the rest. Of course, this is just an *estimate* of the risk premium, and therefore we want to understand the uncertainty around this estimate (ie the standard error!). The standard error of the mean is

   $se(\bar{x}) = \frac{SD(x)}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}$, since we are working with excess returns, $se(r^e) = \frac{\sigma_{r^e}}{\sqrt{n}}$.

This is the variance *of the sample mean!* Not the variance of the sample. We can see that as samples get larger, we get more precise estimates of the sample mean. This makes sense, since as samples get larger, we get *closer to sampling the entire population!* This is called the law of large numbers. We can now perform a *t-test* on the sample mean. In this case, the t-distribution has $n-1$ degrees of freedom (which is potentially important for getting the correct critical/p- values). Subtracting 1 happens because we estimated one component (the mean). The t-test checks the formal hypotheses:

$$H_0: E(r_{it}^e) = 0 \text{ against the alternative } H_a: E(r_{it}^e) \neq 0$$

Perform the t-test and report the t-statistic as well as a 95% confidence interval and a p-value. Is Berkshire's average excess return significant at the 1% level? Create a figure with the correct t-distribution, the test statistic you computed above, as well as shading the parts of the figure where the t-statistic would be significant.

Repeat the above for a test of $\boldsymbol{H_0: E(r_{it}^e) = 1\%}$

3. Now let's extend the analysis beyond just average excess returns. We are specifically going to decompose the risk premium into components related to the factors from the FF3 data. We want to estimate parameters for the model

$$r_{it}^e = \alpha_i + \beta_i r_{mt}^e + s_i SMB_t + h_i HML_t + u_{it}$$

This is *one example* of a factor model for asset returns, that basically is saying there are different sources of risk in the economy (market risk measured by the excess return on the market, r_{mt}^e, risk measured by SMB_t, and a *different* risk measured by HML_t). We have good reason to think that investors require compensation in the form of risk-premium for risk exposures, so these factors should help explain returns, on average. The coefficient labels have been chosen to be consistent with standard practice in Finance.

Perform 3 regressions using the matrix estimation procedure we discussed in class. DO NOT use built in packages at this point. The regressions should be for the 3 models

$$r_{it}^e = \alpha_i + \beta_i r_{mt}^e + u_{it} \rightarrow \text{Capital Asset Pricing Model}$$
$$r_{it}^e = \alpha_i + \beta_i r_{mt}^e + s_i SMB_t + u_{it} \rightarrow \text{2-factor model}$$
$$r_{it}^e = \alpha_i + \beta_i r_{mt}^e + s_i SMB_t + h_i HML_t + u_{it} \rightarrow \text{Fama-French 3 factor}$$

You can calculate the variances of the parameter estimates as the diagonal elements of $Var(\hat{\beta}) = \sigma^2(X'X)^{-1}$ and therefore the standard errors would be the square root of these elements. **Report the results *in a useful format*, marking any significant coefficients with an asterisk. How do your α's change? What does**

this tell you about the simple CAPM model relative to the more complicated factor structure of the Fama French 3 factor model?
Now re-estimate the standard errors using *heteroskedasticity robust* standard errors, where

$$var(\hat{\beta}) = (X'X)^{-1}(X'diag(u_{it}^2)X)(X'X)^{-1}$$

Where $diag(u_{it}^2)$ is a matrix with the diagonal elements equal to u_{it}^2 and otherwise equal to zero. Report results in a similar way to before. **Do any of your conclusions change? If so, explain. What else do these results tell us about Berkshires performance, and therefore Buffet's stock picking?**

4. Now let's evaluate how much of Berkshire's performance can be explained by this factor model. Calculate the full FF3 model's R^2. **What percentage of Berkshire's return is explained by the model? Where is the rest of the return coming from?** Hopefully you've realized that adding an explanatory variable will ALWAYS increase the R^2 relative to the a same model with fewer explanatory variables. This motivates the idea of an *adjusted* R^2 which only increases if the new variable improves the model "enough" (which means the t-statistic on the new coefficient is greater than 1). **Compute the regular and adjusted R^2 for the 2-factor model and FF3 models from above. Did adding a variable actually improve the model?**

5. It can be useful to understand where most of the variance is being explained. A variable could have a significant coefficient *and still not explain much of the variance in the dependent variable!* We are going to do a variance decomposition to figure out which independent variables explain the most variance in Berkshire's return. Specifically, we are going to use the relationship

$$var(r_{it}^e) = \hat{\beta}\Sigma_X\hat{\beta}' + var(\hat{u}_t) \Rightarrow 1 = \frac{\hat{\beta}\Sigma_X\hat{\beta}'}{var(r_{it}^e)} + \frac{var(\hat{u}_t)}{var(r_{it}^e)}$$

Where Σ_X is the covariance matrix of the independent variables. $\hat{\beta}\Sigma_X\hat{\beta}'$is a scalar, but you can decompose it into component pieces as $\frac{\hat{\beta}\hat{\beta}'\Sigma_X}{var(r_{it}^e)}$. The diagonal elements of this resulting matrix tell you what percentage of the variance is being explained by each corresponding variable. (note: if you sum the elements of this matrix, you get R^2 → check this!). **What factor has the biggest impact on the model's ability to explain variance in r_{it}^e? Compare this with your coefficient estimates on all the factors. Is Buffet compensated for factor returns without actually having his return strongly vary with those factors?**

6. Re-estimate the models using a standard python package (like statsmodels). Give the F-test statistic and p-value for a test of overall model significance for all 3 models. **Which are significant and which aren't (if any)?** Test the joint hypothesis $H_0: s_{BRK} = 0, h_{BRK} = 0$ against the alternative $H_a: SMB\ \boldsymbol{or}\ HML$ are significant in explaining BRK's returns. **What do you conclude? Does the extended model add value?**

7. Lastly, let's take a look at BRK outperformance over time. First, create a benchmark portfolio that mimics BRK's exposures. Specifically, calculate
$$r_{bmk,t} = (1 - \beta_{BRK})r_f + \beta_{BRK}(r_{mt}^e) + s_{BRK}\ SMB_t + h_{BRK} SMB_t$$
Why is $(\mathbf{1} - \boldsymbol{\beta}_{\boldsymbol{BRK}})$ multiplying the risk-free rate? Now calculate the future value factors and calculate the value of \$1 invested in BRK, the benchmark, and the market portfolios starting at the beginning of the sample for each point in time. Plot the natural log of these portfolio values across time. **What do you notice about the performance of Berkshire pre-2000 compared to the market and benchmark portfolios vs. post-2000? What happened to BRK's α over time? Why might this be the case?**