**FBE 551 Homework #2**

**Due on 9/19/2023 by end of day**

This is an individual homework assignment. You will need to use the SPY.csv, ^IRX.csv, and OAKMX.csv data files in the problem sets folder. You will also need to download and read in your own data file for returns on the ARKK (ARK Innovation ETF). Please be sure to name this file ‘ARKK.csv’.

Carry out the following calculations in a single Jupyter notebook. Make sure that I will be able to run it. For example, read in the CSV files from the current directory. I.e., do not write the full path of the file in your code. Do not have your code change directories to a pathway that is probably not valid on my computer. This will allow me to run your notebook in any directory on my own PC.

Please use markdown cells to help me understand what you are doing. (I know it may be obvious, but I want to see that you know how to use markdown.) Clearly highlight what parts of your notebook are used to answer each question below.

**Question 1**

Using the full SPY dataset, compute daily close-to-close returns on SPY. What is the average return on SPY over the full sample, during the first trading day of the month, and during the third-to-the-last trading day of the month?

**Question 2**

Create a new dataframe that includes all dates on which there is price data available for SPY. Add the T-Bill yield (the Close column of ^IRX) to this dataframe. Note that ^IRX is missing on a number of dates that are in SPY, such as 2016-10-10. You can fill in these missing values using the following command:

spy['Tbill yield filled'] = spy['Tbill yield'].fillna(method='ffill')

You should take a look at how this works and think about why it is probably ok.

Compute T-bill returns following the instructions at the bottom of this sheet. What are the mean and standard deviation of excess SPY returns, defined as the SPY return minus the T-bill return?

**Question 3**

What is the maximum drawdown of OAKMX? (Hint: use adjusted closing prices for portfolio values and then the cummax() method to get the prior highs.)

**Question 4**

What is the downside deviation of OAKMX? As a benchmark return, use the T-Bill return.

**Question 5**

What is the average excess return on OAKMX when the previous Adjusted Closing price is above its 200-day moving average? What about when it is below the moving average? (Hint: use the rolling() method to get the moving average.) Are these means statistically significant?

**Question 6**

From finance.yahoo.com, download the monthly historical financial data for ticker ARKK, which is the ARK Innovation ETF. Save this as a csv file and read it into your program. Construct a dataframe which includes the monthly returns for this ETF over the months for June 2018-June 2023. This should include a sample of 60 monthly returns. Add to this dataframe the Fama-French 3-factor data from the file ‘FFdata.csv’.

1. What has been the average annual return, average annual excess return, and average annual standard deviation for the ARKK ETF?
2. What has been the Sharpe ratio for ARKK over this period?
3. Using just the Mkt-rf factor, compute the CAPM beta for the ARKK fund. What has been the alpha for ARKK over this period? Has it been statistically significant?
4. Run the full Fama-French 3-factor regression on the ARKK fund. What are the factor loadings? What has been the alpha under the 3-factor model? Is it statistically significant?
5. In a markdown cell, explain why the alpha from the CAPM model is different from the alpha from the FF 3-factor model. Explain how the factor loadings from the 3-factor model regression reflect the type of stocks that ARKK tends to hold in its portfolio.

**Computing T-bill returns from ^IRX.csv**

This file contains annualized Treasury Bill yields. These can be transformed into approximate T-Bill returns by taking the first lag, dividing by 36500 (365 to convert annual to daily × 100 to convert percentage to decimal), and multiplying by the number of calendar days in the holding period, i.e.

RTB(t) = IRX(t-1) × N(t) / 36500

For example, if day t is a Monday and day t-1 is a Friday, then there are three calendar days in the holding period starting on Friday and ending on Monday, implying N(t)=3. For a regular weekday, N(t)=1. For a three-day weekend, N(t)=4.

Computing N(t) will likely be the hardest part of this calculation. Just try your best on this!