

## Problem Set 1

**Name: Abdul Moiz (105078804)**

### Question 1:

*Construct the value-weighted market return using CRSP data, replicating the market return time series available in Kenneth French website. Also calculate the equal-weighted market return, and the lagged total market capitalization. Your output should be from January 1926 to December 2018, at a monthly frequency.*

Before starting with the procedure, I performed data cleaning ensuring that I don't encounter any issues during the process. My assumptions are as follows:

1. **Universe of Stocks:** As instructed in the homework, I extracted the full data and did not pre filter by any share or exchange code. The data taken was from January 1926 to December 2018.
2. **Missing Returns:** Places where holding returns were missing, I took delisting returns and vice versa in case Delisting returns were missing. However, when both of them were available, I used compounded the returns. In cases where both were missing I dropped them because I didn't have much information as to why there were missing.
3. **Delisting Return Calculation:** As explained above, there were included in the cumulative return if available.
4. **Market Capitalization Return:** For every month and PERMNO, I calculated the lagged market return for onward calculation of weight as it is the correct approach because we are rebalancing portfolio based on weights from previous month.
5. **Portfolio weights:** Value weights were calculated by dividing the lagged market cap of individual stock by total lagged market cap giving us the weight to be applied in current month.
6. **Sample Period:** I took sample from January 1926 to December 2018.
7. **Definition of Portfolio Weights:** Weights of portfolio were based on lagged market cap. Because if we were to take weights of current month, then our returns would be exaggerated naturally. In order to avoid that, you apply weight to current month based on market cap from last month.

**Question 2:**

*Using the risk-free rate of return from French's website, report the following moments of the market excess returns for both time series (4 decimal digits): annualized return, annualized volatility, annualized Sharpe ratio, skewness, and excess kurtosis. Annualized values should be calculated geometrically. You should be comparing between July 1926 to December 2018, at a monthly frequency.*

The summary statistics are in the following table. I have reported Annualized Mean, Standard Deviation, Sharpe Ratio, Skewness and Kurtosis for both my replicated portfolio and French's portfolio from his website. As evident, both figures are approximately very similar.

**Table 1: Summary Statistics**

	<b>Replication</b>	<b>French's</b>
	(1)	(2)
Annualized Mean	0.07794116	0.07922266
Annualized Standard Deviation	0.18367781	0.18480412
Annualized Sharpe Ratio	0.12249533	0.12375056
Excess Skewness	0.18415275	0.18116666
Kurtosis	8.08995895	7.80693396

From Question 1, we had the time series of value weighted returns which I used to compare with the French's portfolio. Further assumptions are as follows:

1. **Sample Period:** January 1926 to December 2018
2. **Excess Skewness:** I calculated using full sample of returns without logs and annualization.
3. **Kurtosis:** I calculated using full sample of returns without logs and annualization.
4. **Annualized mean:** I annualized by multiplying monthly average mean by 12 as instructed in the slides.
5. **Annualized Standard Deviation:** I annualized by multiplying monthly average mean by square root of 12 as instructed in the slides.
6. **Sharpe Ratio:** I calculated Sharpe ratio by dividing annualized mean on excess returns by annualized standard deviation.

**Question 3:**

*Report (up to 8 decimal digits) the correlation between your time series and French's time series, and the maximum absolute difference between the two time series. It is zero? If not, justify whether the difference is economically negligible or not. What are the reasons a nonzero difference? You should be comparing between July 1926 to December 2018, at a monthly frequency.*

After carrying out the steps in question 1 and 2, I compared my replicated portfolio with french's portfolio by calculating correlation and maximum absolute difference between the two datasets.

Results are as follows:

Table 2: Correlation and Maximum Abs. Difference

<b>Correlation</b> (1)	<b>Maximum Absolute Difference</b> (2)
0.9974199	0.09343033

The correlation is more than 99.7% signifying strong relationship between both methods with maximum absolute difference of only 0.09.