Problem Set 1

Name: Redmond Xia
Names of whom I discussed this problem set with:

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 2019, at a monthly frequency.

Before Calculating the portfolio time series, I conduct a series of data cleaning as part of my PS Q1 function. Next, I describe my data cleaning process and their respective assumptions:

1. **Universe of stocks:** Following Ken French procedure, I restrict the sample to common shares (share codes 10 and 11) and to securities traded in the New York Stock Exchange, American Stock Exchange, or the Nasdaq Stock Exchange (exchange codes 1, 2, and 3).

2. **Missing returns:** If both delisting return and holding period returns are missing then I remove the record.

3. **Delisting return calculation:** If the record only has a return but no delisted return, then return is used. If the record only has delisted return but no return, then the delisted return is used. We calculate the cum-dividend returns as $Return = (1 + RET) * (1 + DLRET) - 1$.

4. **Market Capitalization Calculation:** First, CRSP uses negative price as the average bid-ask spread. I used take them as market price by using its positive magnitude. Then, Market Capitalization is equivalent to $MktCap = PRC * SHROU$, which is price times number of shares. Then we lag the market cap by one month. If I find a missing lag market cap, the record is removed.

5. **Portfolio weights:** For the equallty weighted returns, each stock has the same weight, i.e. $EquallyWeighted = \frac{1}{N}$ where N is the number of stock for that month. In the value weighted returns, we weight each stock by the lagged market cap, i.e. $ValueWeighted = \frac{LagMktCap_i}{\sum_{i=1}^{N} LagMKtCap_i}$

6. **Sample period:** To get report the returns, we download the data starting from December 1925. This is because we use the lag methods to calculate the returns. Thus, the output is from January 1926 to December 2019 by monthly frequency.

7. **Definition of portfolio weights:** The definition of portfolio return is always the same, $PortfolioReturn = \sum_{i=1}^{n} w_i * return_i$. The difference is between equally weighted and value weighted is the w_i, which is defined in 5.

Displaying the first 15 months of returns:

Year	Month	Stock_lag_MV	Stock_Ew_Ret	Stock_Vw_Ret
1926	1	26.82938	0.0234050	0.0002479
1926	2	27.01723	-0.0543068	-0.0339950
1926	3	26.12723	-0.0964609	-0.0649615
1926	4	24.32352	0.0323434	0.0369713
1926	5	25.26952	0.0019658	0.0122619
1926	6	25.49971	0.0510312	0.0545785
1926	7	26.44901	0.0130330	0.0318134
1926	8	27.31789	0.0310209	0.0289125
1926	9	28.34824	-0.0064526	0.0059102
1926	10	28.15444	-0.0348373	-0.0292401

Year	Month	Stock_lag_MV	Stock_Ew_Ret	Stock_Vw_Ret
1926	11	27.25417	0.0244611	0.0284409
1926	12	28.27267	0.0243672	0.0289811
1927	1	29.60879	0.0120127	0.0018995
1927	2	29.74728	0.0518418	0.0442972
1927	3	31.08431	-0.0193360	0.0042588

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 2019, at a monthly frequency.

The summary The summary statistics are in Table 1 below. I report the following five statistics: annualized mean, annualized standard deviation, annualized sharpe ratio, skewness, and excess kurtosis. In Column 1, I report the statistics for the replicated value-weighted market portfolio of stocks calculated in the previous question. In Column 2, I report the statistics for the value-weighted market portfolio of stocks from Ken French's website

Table 1: Summary Statistics

	Estimated FF Market Excess Return	Actual FF Market Excess Return
Annualized Mean	0.0802	0.0800
Annualized Standard Deviation	0.1842	0.1844
Annualized Sharpe Ratio	0.4354	0.4338
Skewness	0.1821	0.1763
Excess Kurtosis	7.8146	7.8159

From question, we have a time series of value-weighted market returns, namely $\{r_t\}_{t=1}^{T}$

1. **Sample period:** Monthly from July 1926 to December 2019.

2. **Excess Skewness:** To calculate excess skewness, I used the library fBasics function skewness for the monthly time series, with no annualization/logs, using the full sample

3. **Kurtosis**: To calcualte excess kurtosis, I used the library fBasics function kurtosis (method = 'excess'), with no annualization/logs, using the full sample

4. **Annualized Mean**: I took the arithmetic mean $\sum_{i=1}^{n} \frac{Ret_i}{N}$ of the excess returns for both method. To annualized, I took the arithmetic approach by multiplying that number by 12.

5. **Annualized Standard Deviation**: I took the R base standard deviation function for the returns and annualized it by multiplying $\sqrt{12}$

6. **Sharpe Ratio**: Sharpe ratio is definited by $SharpeRatio = \frac{AnnualizedMean}{AnnualizedStandardDeviation}$

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 2019, at a monthly frequency.

My report uses the results from question 2 and finds the correlation based on the R base correlation function. Then I find the largest difference in the value weighted excess returns to FF excess returns. You can find the summary in the table below

In Table 2 below, I report the time-series correlation between the replicated value-weighted market portfolio of stocks and the value-weighted market portfolio of stocks from Ken French's website. I also report the maximum difference between the two series. I limit the sample to be between July 1926 and December 2019.

The finding shows that the correlation is very close to 1. This means that they almost match each other. The biggest difference in absolute excess returns is roughly 19 basis point. A neglible number shows that they almost exactly match each other. However, The difference between the replicated portfolio and the one from French's Website is not zero. This could be due to computer precision error or that FF model doesn't capture 100% of the replicated portfolio. It captures up to 99.99% which is close to perfect.

Table 2: Correlation and maximum difference

1. Correlation	0.9999943
2. Maximum Absolute Difference	0.0019391