Problem Set 1

Name: Xiangui (Jasmine) Mei

Names of whom I discussed this problem set with: Jiaqi (Johnny) Li

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 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:

- **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).

- **Return calculation and missing returns:**

  In CRSP, for each stock i at month t, I have:

  Holding period returns: $r_{i,t}^{h}$ and delisting returns $r_{i,t}^{d}$

  Then I used cum-dividend total returns:

$$r_{i,t} = \begin{cases} r_{i,t}^{h} & if\ r_{i,t}^{d}\ is\ missing \\ r_{i,t}^{d} & if\ r_{i,t}^{h}\ is\ missing \\ \left(1 + r_{i,t}^{h}\right)\left(1 + r_{i,t}^{d}\right) - 1 & if\ both\ not\ missing \end{cases}$$

  For missing returns, which refers to both holding period returns and delisting returns are missing, I deleted them.

- **Market Capitalization calculation:**

  Market cap was calculated using the absolute value of price and shares outstanding.

- **Value Weighted Portfolio weights:**

  First of all, I calculated market cap as: $me_{i,t} = price \times shares\ outstanding$

  Then, by firm (PERMNO), I derived lagged (shift one year) market portfolio weights as:

$$w_{i,t}^{mkt} = \frac{me_{i,t-1}}{\sum_i me_{i,t-1}}$$

- **Equal Weighted Portfolio weights:**

  First of all, by date, I derived the number of stocks q, the weight for each stock is $1/q$ for specified time.

- **Market Return:**

  For value-weighted return, by date:

$$R_{mkt\_v}^{t} = \sum_i w_{i,t}^{mkt} r_{i,t} = \frac{me_{i,t-1}}{\sum_i me_{i,t-1}} r_{i,t}$$

  For equal-weighted return, by date:

$$R_{mkt\_e}^{t} = \sum_i \frac{1}{q} r_{i,t}$$

- **Sample period:**

  January 1926 to December 2018, at a monthly frequency

- **Result:**

```
> Monthly_CRSP_Stocks
            date Year Month Stock_lag_MV Stock_Ew_Ret Stock_Vw_Ret
   1: 1926-02-27 1926    02     26789658      13.8429 -0.033336937
   2: 1926-03-31 1926    03     25996474      13.8429 -0.064637870
   3: 1926-04-30 1926    04     24057518      13.8429  0.036766396
   4: 1926-05-28 1926    05     25235880      13.8429  0.012319261
   5: 1926-06-30 1926    06     25047346      13.8429  0.053492172
  ---
1111: 2018-08-31 2018    08  28709224535      13.8429  0.036007378
1112: 2018-09-28 2018    09  29665669755      13.8429  0.002114236
1113: 2018-10-31 2018    10  29575303491      13.8429 -0.074831140
1114: 2018-11-30 2018    11  27211251201      13.8429  0.018654681
1115: 2018-12-31 2018    12  27576149950      13.8429 -0.093613564
```

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 Table 1 below, I reported 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

	Replication	French's
Annualized_Mean	0.0780	0.0781
Annualized_Std	0.1847	0.1848
Annualized_Sharpe_Ratio	0.4226	0.4229
Skewness	0.1877	0.1843
excess_kurtosis	7.8817	7.8505

From question, we have a times series of value-weighted market returns, namely $\{r_t\}_{t=1}^{T}$. Let the market return from French's website be given by $\{r_t^F\}_{t=1}^{T}$. I compute these statistics as follows:

- **Sample period:**

  January 1926 to December 2018, at a monthly frequency

- **Excess Skewness**

  With the build in function moments, I calculate excess skewness of rt from the monthly time series directly (no annualization, and no logs) using the full sample.

- **Kurtosis**

  With the build in function moments, I calculate kurtosis from the monthly time series directly (no annualization, and no logs) using the full sample.

- **Annualized Mean:**

  I calculate excess market return from the monthly time series directly using the formula:

$$R_{excess}^{t} = R_{mkt\_v}^{t} - \frac{R_f}{100}$$

  Then I calculate annualized return

$$R_{annulized} = R_{excess}^{t} * 12$$

- **Annualized Standard Deviation:**

I calculate annualized standard deviation from the monthly time series directly using the formula:

$$Std_{annulized} = Std_{monthly} * \sqrt{12}$$

- **Sharpe Ratio:**

  I calculate annualized Sharpe ratio as:

$$Sharpe\ Ratio = \frac{R_{annulized}}{Std_{annulized}}$$

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.

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.

Table 2

```
======================
Correlation  0.99998780
max_abs_diff 0.00292205
----------------------
```

I limit the sample to be between July 1926 and December 2018. The correlation between replication portfolio and French's portfolio is almost 1. The difference between the replicated portfolio and the one from French's website is not zero, but it is negligible. The small difference may be because the sample data set doesn't specifically manage the change in exchanges after a delist happens. Moreover, as a result of different assumptions in dealing with the data in return, delisting return and share prices for instance, there might be different interpretation of letter "C, the result might be different as well.