Problem Set 1
MFE 431: Quantitative Asset Management
Professor: Bernard Herskovic
Student: Xiahao Wang
Name of whom I discussed this problem set with: Yu Yue

Question 1

Summary:

The resulting dataset:

```
##       Year Month  Stock_Vw_Ret  Stock_Ew_Ret Stock_lag_MV
##    1: 1926     1  0.0002478675  0.0234050475     26829.38
##    2: 1926     2 -0.0339951421 -0.0544178487     27017.14
##    3: 1926     3 -0.0649615280 -0.0967514747     26093.78
##    4: 1926     4  0.0369713089  0.0323434148     24323.52
##    5: 1926     5  0.0119837079  0.0007612809     25152.03
##   ---
## 1112: 2018     8  0.0360297665  0.0342473575  28766451.29
## 1113: 2018     9  0.0020913824 -0.0160036557  29697876.16
## 1114: 2018    10 -0.0748784718 -0.1021241763  29595694.90
## 1115: 2018    11  0.0186908801 -0.0030372859  27240633.09
## 1116: 2018    12 -0.0936117434 -0.1292800866  27587151.57
```

1. **Universe of stocks:**

Following the the Ken French Procedure, I restrict the sample to common shares(share code 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:**

- Handle Missing Data denoted by the CRSP: CRSP denotes missing return and delisting return data by certain number. Exclude these data from the dataset.

```
# Filter out missing ret and dlret data
 crsp_monthly <- crsp_monthly[!RET %in% c(-99,-88,-77,-66,-55,-44)]
 crsp_monthly <- crsp_monthly[!DLRET %in% c(-99,-88,-77,-66,-55,-44)]
```

- Handle missing data after return and market cap calculation: knock out all the NAs from price , return, cum-dividend returns and lagged total market capitalisation:

```
# remove NAs
 crsp_monthly <- crsp_monthly[PRC != "NA"]
 crsp_monthly <- crsp_monthly[RET != "NA"]
 crsp_monthly <- crsp_monthly[cumDivRet != "NA"]
 crsp_monthly <- crsp_monthly[mktCapLagged != "NA"]
```

3. **Delisting return calculation:**

To calculate the cum-dividend return, let return be cum-dividend return if delisting return is missing.

If delisting return is not missing and return is missing, let delisting return be just cum-dividend return.

If both are not missing, use the formula from lecture notes: (1 + DLRET)(1 + RET) - 1 to get the return.

Delisting return that are not numeric are converted to NAs and removed later on.

code:

```
# calculates the cum-Dividend returns
crsp_monthly[, `:=`(cumDivRet, ifelse(is.na(DLRET),RET,ifelse(is.na(RET), DLRET, (1+DLRET)*(1+RET)-1 ))
```

4. **Market Capitalisation calculation:** There are some price that are negative, I take the absolute value for both price and shares outstanding and multiply them together to get the market capitalisation. Also I shift the market cap by 1 period for each firm.

```
# calculate the marketcap
crsp_monthly[, `:=`(ME, abs(PRC) * abs(SHROUT))]

#lag the market cap of each firm
crsp_monthly[, mktCapLagged := shift(ME), by=c("PERMNO")]
```

5. **Portfolio Weights:**

Value-weight the cum-dividend return by lagged market cap: $Weight_i = \frac{MarketCap_{i-1}}{TotalMarketCap_{i-1}}$

For equal weight, just take the mean of returns for each period.

```
 # value weight portfolio, this might be wrong, let's see
  valueweight <- crsp_monthly[,list(Stock_Vw_Ret = weighted.mean(cumDivRet, mktCapLagged, na.rm = TRUE)
                                    by=list(Year, Month)]
  # equal weight portfolio
  equalweight <- crsp_monthly[,list(Stock_Ew_Ret = mean(cumDivRet)),by=list(Year, Month)]
```

6. **Sample Period:** As I need to lag the market capitalisation while acquring data from Jan 1926 to Dec 2018, I have obtained data starting from Dec 1925 so that I have data for Jan 1926.

7. **Definition of portfolio Weights:**

As explained in part 5, portfolio return is the sum of all weighted return for that period.

$$portfolioReturn_i = \Sigma_{i=1}^{n} w_i * r_i$$

Where n is the number of stocks at that time.

Question 2

Sample Data:

- Download Fama French 3 fator from the website
- Divide both Excess Return and Risk Free by 100 as the numbers are in percentage
- Truncated result from Qn1 to start from July 1929

Annualized Mean:

I use the monthly mean return and multiply by 12 to get annual data

Annualized Standard Deviation:

I use the monthly standard deviation return and multiply by $\sqrt{12}$ to get annualized data

Sharpe Ratio:

$$SharpeRatio = \frac{AnnualizedMean}{AnnualizedStandardDeviation}$$

Skewness and Kurtosis:

Using library called `moments` to obtain the statistics.

Summary:

```
##              Stats EstimatedExMktRet ActualExMktRet
## 1: Annualized Mean        0.07811358     0.07808973
## 2:   Annualized SD        0.18476993     0.18477282
## 3:    Sharpe Ratio        0.42276130     0.42262563
## 4:        Skewness        0.18430365     0.18425557
## 5: Excess Kurtosis        7.85094617     7.85049201
```

Question 3

Sample Data:

Truncated result from Qn1 to start from July 1929

Find: * Correlation between estimated excess return and actual excess return * Maximum absolute difference between the two returns

Summary:

The correlation is almost 1 and, but the difference between the replicated portfolio and the one from French website is not zero.

This could be due to the high precision used in the replicated portfolio (10 decimal places) where the data from Fama French has precision up to only 3 digits for excess return.

Another reason could be due to the change in historical data from CRSP:http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/changes_crsp.html

Quote from the website: Please note: CRSP has recently completed the Pre62 Daily Data Series Project. The addition of these new daily data results in changes to month-end prices and to dividend ex-dates. These changes have resulted in many small changes to historical returns on my website.

```
##    Correlation Max difference
##    0.999998323    0.002060508
```