Problem Set 1: Market Portfolio

Professor Bernard Herskovic
bernardherskovic@anderson.ucla.edu
Office: Room C-413

Due by April 11 at 4:00pm. This is an individual assignment, but you can discuss it with your classmates. If you discuss with other classmates, indicate their names in your write-up. Please submit all coding as well as a separate write-up. Explain the procedure and your answers clearly in the write-up. Code must be formatted as instructed in order to receive a grade. Use CCLE to submit your answers.

You should submit two files:

- .R **PS1_YourStudentID** (for example, PS1_012345678.R), with **only** the functions described below
- .pdf **PS1_YourStudentID** (for example, PS1_012345678.pdf), with answers to the questions written below

1. Construct the value-weighted market return using CRSP data,[1] replicating the market return time series available in Kenneth French website.[2] Also calculate the equal-weighted market return, and the total market capitalization. Your sample should be from February 1926 to December 2015, at a monthly frequency.

- Function name: **PS1_Q1**
  - Input
    - data.table **CRSP_Stocks**, with columns: PERMNO, date, SHRCD, EXCHCD, RET, RETX, DLRET, DLRETX, PRC, SHROUT
      - This should be the data as pulled from WRDS, with one exception. Format the date column as a Date type.
  - Output
    - data.table, with each row corresponding to a unique year and month, with columns
      - Note: Returns should be formatted in decimal proportion (not percent).

[1] Available at WRDS: `https://wrds-web.wharton.upenn.edu/wrds/`. You will have to go to CRSP > Annual Update > Stock / Security Files > CRSP Monthly Stock.

[2] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

Variable Name	Variable type	Variable description
Year	Integer	Year
Month	Integer	Month
Stock_lag_MV	Numeric	Total market value the previous month (in millions)
Stock_Ew_Ret	Numeric	Equal-weighted returns
Stock_Vw_Ret	Numeric	Value-weighted returns

- Hints:
  - Do not forget to include delisting returns (DLRET).
  - Use cum-dividend returns (RET and DLRET).
  - Do not forget to restrict your sample by exchange type (EXCHCD) and share codes (SHRCD).
  - Construct market equity by multiplying prices (PRC) by number of shares outstanding (SHROUT).
  - Check the CRSP manuals available. Do not forget to fix missing values, especially for returns and prices.

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 average, annualized volatility, annualized Sharpe ratio, skewness, and excess kurtosis. Annualized values should be calculated geometrically. You should be comparing between July 1926 to December 2015, at a monthly frequency.

- Function name: **PS1_Q2**
  - Inputs
    - data.table **Monthly_CRSP_Stocks**, the output of **PS1_Q1**
    - data.table **FF_mkt**, with columns

Variable Name	Variable type
Year	Integer
Month	Integer
Market_minus_Rf	Numeric
SMB	Numeric
HML	Numeric
Rf	Numeric

      - Note: Returns should be formatted in decimal proportion (not percent).
  - Output
    - 5×2 numeric matrix. Rows: Annualized Mean, Annualized Standard Deviation, Sharpe Ratio, Skewness, and Excess Kurtosis. Columns: Estimated FF Market, Actual FF Market.

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 2015, at a monthly frequency.

- Function name: **PS1_Q3**
  - Inputs
    - data.table **Monthly_CRSP_Stocks**, the output of **PS1_Q1**
    - data.table **FF_mkt** as described in **PS1_Q2**
  - Output
    - Vector of length two. Correlation between time series and maximum absolute difference between two time series.

- Hint: the correlation should be close to one, and the moments reported for both time series should be very very similar (maybe the same?).