Homework 1 - due 10/8

Total market returns and the SP500

The SP500 index analyzed in class (`GSPC` ticker) is available since 1928 in *Yahoo Finance* (backfilled, since SP500 started only in 1957). Unfortunately, the index tracks only the price movements of its components and it does not reflect cash distributions (dividends), missing a substantial portion of total return. The SP500 Total Return index captures the overall performance of the stock market by reinvesting dividends, but it only goes as far back as 1988. Our goal is to backfill the SP500 Total Return index, and use it to analyze the historical performance of the US stock market.

a. Extract the data for the SP500 Total return index from *Yahoo Finance* up to 12/31/2023. Select only the *closing prices* of last day of each *month*. This code snipet gets you the daily data

```
library(tidyquant)
#tr_index holds daily total return index prices
sp500_tr_index <- tq_get("^SP500TR", get = "stock.prices",
    from = "1988-01-04", to = "2023-12-31") |>
      select(date, tr_index = close)
```

b. Download Robert Shiller's data containing *monthly* SP500 price and dividend data, used in his "Irrational Exuberance" book. Prices and dividends are in the "P" and "D" columns (you have to convert them to numeric with `as_numeric()`). The "Date" column has a non-standard format so you will have to fix it, e.g., using `str_replace`. To get you started,

```
library(readxl)
aux <- tempfile(fileextension = ".xls")
download.file(url = "http://www.econ.yale.edu/~shiller/data/ie_data.xls",
              destfile = aux, mode='wb')
rshiller <- read_excel(aux, sheet = "Data", skip = 7)
```

c. Shiller's data represents the 12-month moving sum of monthly dividends. To get the approximative monthly dividends, divide by 12, and then calculate the monthly total returns of the SP500 (Hint: $ret = (P + D/12)/lag(P) - 1$.)

d. Check whether the total return computed above is able to match the actual total return since 1988 which you obtained in a). Use the cumulative (product of) gross returns since 1988 using Shiller data and Yahoo Finance total return index to construct and compare the time series of compounded returns. Plot them on the same graph. Do they overlap perfectly? Report also the *Pearson* and *Spearman* correlations between the post 1988 total returns calculated from Shiller's data with actual returns those calculated from Yahoo Finance data.

e. For comparison, download the (undjusted) SP500 price index from Yahoo Finance, and plot in the same graph the cumulative unadjusted returns vs the total returns calculated using Shiller's index, from 1928 to 2023. Report the average and standard deviations of the *monthly* adjusted and unadjusted returns (from 1928 to 2023). The following code block gives you the daily prices of the SP500.

```
# pr_index contains unadjusted  SP500 index daily close prices
sp500_price_index <- tq_get("^GSPC", get = "stock.prices",
                            from = "1928-01-01", to = "2023-12-31") |>
  select(date, pr_index = close) |>
  drop_na() |>
```

f. French's data library is a great resource for returns on various portfolios. The precise definitions of the variables is spelled out there. R simplifies the access to this date through the `frenchdata` library. We will extract the monthly risk-free rate (`rf`) and monthly returns to a broad portfolios of stocks ("market return") in excess of the risk-free rate (`mre`). Calculate the SP500 monthly total excess returns (in excess of the risk-free rate), and plot the cumulative excess returns of the SP500 total index. Based on this chart, report the 5 longest historical downturns in US, from peak to trough. For each of them, report the duration in months needed to recoup the losses (for the stock market to regain the same level as when the downturn started). The following code snippet pulls the needed data from French's library.

```
library(frenchdata)
ff3 <- download_french_data("Fama/French 3 Factors")
ff3 <- ff3$subsets$data[[1]] |>
  mutate(
    date = floor_date(ymd(str_c(date, "01")), "month"),
    across(c(RF, `Mkt-RF`, SMB, HML), ~as.numeric(.) / 100),
    .keep = "none"
  ) |>
```

```
    rename_with(str_to_lower) |>
    rename(rme = `mkt-rf`) |>
    select(date,rf,rme) |>
    filter(date <= ymd("2023-12-31"))
```

g. Report the Pearson and Spearman correlations between excess returns on the SP500 total index and the broader French's excess market returns. Plot the time series of the *cumulative excess returns* for the two indices.

h. Based on the SP500 total excess returns, report the 3 longest drawdown periods (peak to through) in the US stock report. For each of them, report how long it took for the compounded excess return to reach the previous high watermark.