Working with Data

MGMT 675, AI-Assisted Financial Analysis

Kerry Back, JGSB, Rice University

Julius Inputs and Outputs

- Julius will read data that we've uploaded, execute code generated by ChatGPT, and provide links to outputs.
- Can output images, Excel tables, csv files, ...
- Julius can also read some online data.

Running Code Elsewhere

- We can copy code from Julius and paste into a Jupyter notebook - for example, in Google Colab.
- We can save Colab notebooks to Google Drive and then reload them into Colab and run them whenever we want.
- Or, we could copy code into notebooks on our own machines.

Today

- Work with metrics.csv and tickers.csv.
    - Download from the course website.
    - Upload in Julius.
- Get online data from various sources.
- Work only in Julius (use colab later)

metrics and tickers

We will illustrate the following basic data steps.

- Merge
- Filter
- Sort
- Aggregate by group
- Transform

Merge

- Ask Julius to merge the datasets on ticker
- Ask Julius to describe the merged data.
    - How many rows are there?
    - What are the column names?
    - What are the unique values in the category column?
    - What are the unique values in the sector column?
    - What are the unique values in the scalemarketcap column?
    - Show the head of the data frame.

Filter

- Ask Julius to filter on the category column to "Domestic Common Stock" and "Domestic Common Stock Primary Class." Ask Julius to call this data frame common_stock.
- Ask Julius to create a copy of the common_stock data frame that contains only rows for which pe>0.
- Ask Julius to create a copy of the common_stock data frame that contains only rows for which marketcap is above the median marketcap.

Sort

- Ask Julius to sort on marketcap in descending order and to show the head of the data frame.

Aggregate by group

- Ask Julius to describe marketcap.
- Ask Julius to compute the mean marketcap by sector.
- Ask Julius to compute the number of firms by sector.
- Ask Julius to compute the total marketcap by sector.
- Ask Julius to compute the mean pe grouped by (sector, scalemarketcap) and to display the results as a two-dimensional table.
- Ask Julius to recreate the table using only rows for which pe > 0.
- Ask Julius to compute the percent of firms for which pe < 0 by sector.

Transform

- Ask Julius to create a new variable equal to the rank of marketcap in descending order.
- Ask Julius to create a new variable that is 1 if pe>0 and 0 otherwise (a dummy variable).
- Ask Julius to create a new variable equal to the excess of pb over the median sector pb.

Online data

- Julius will normally get some online data without complaining.
    - For example, it will get data from Yahoo Finance and Federal Reserve Economic Data (FRED).
- At other times, Julius will say it has no access to external websites and can only advise.
    - In those cases, it will still produce code that you can run elsewhere.
    - But, it will not be available to debug the code for you.
    - This happens, for example, with the SEC's Edgar site.

Yahoo's Adjusted Closing Prices

- Yahoo's adjusted closing prices are adjusted for splits and dividends.
- The percent change in the adjusted closing price is the daily close-to-close return including dividends (on ex-dividend dates).
- If we want returns at a different frequency, for example annual returns, then we can either
    - compound the daily returns, or
    - downsample the adjusted closing prices to annual data and compute the percent change of the downsampled data.

Other Yahoo Finance Data

- Daily open, high, low, close, adjusted close, volume
- Income statement, balance sheet, and statement of cash flows for past 5 years
- Current market option data (bid, ask, last price, open interest, implied volatility, ...)

Returns

- Ask Julius to download adjusted closing prices for SPY starting in 1990 (it won't actually start that early).
- Ask Julius to compute daily returns as the percent change in the adjusted closing price and to produce a boxplot of the daily returns and a link to download the boxplot.
- Ask Julius to downsample the adjusted closing prices to the last day of the year.
- Ask Julius to compute annual returns as the percent change in the downsampled adjusted closing price and to produce a boxplot of the annual returns and a link to the boxplot.
- Ask Julius to recreate the boxplot of daily returns using plotly and to include the date in the hover data. Ask Julius to save the boxplot as an html file and to provide a link.

Federal Reserve Economic Data

- Ask Julius to use the pandas-datareader to get crude oil prices from FRED.
- Ask Julius to plot the prices.
- Ask Julius to get 10-year Treasury yields from FRED.
- Ask Julius to get the inflation rate from FRED.

Other Data

- Ask Julius to find the constituents of the S&P 100.
- When Julius provides a link, ask Julius to read the table at the link.