MGMT 675

AI-ASSISTED FINANCIAL ANALYSIS

RICE | BUSINESS
Jones Graduate School of Business

DATA HANDLING

TOPICS

- Merge
- Filter
- Aggregate by group
- Transform

DATASETS

- metrics.xlsx and tickers.xlsx.
  - Download from the course website.
  - Upload in Julius.
- Online data from various sources.

TICKERS DATA

- Ask Julius to read tickers.xlsx.
- Ask Julius how many rows are in the data set and what the column names are.
- Ask Julius what the unique values are in the category and sector columns.

FILTER

- Ask Julius to filter the tickers data set on the category column to “Domestic Common Stock” and “Domestic Common Stock Primary Class.” Ask Julius to call this data frame common_stock_tickers.
- Ask Julius how many rows are in the common_stock_tickers.
- Ask Julius to show the first few rows.

METRICS DATA

- Ask Julius to read metrics.xlsx.
- Ask Julius how many rows are in the data set and what the column names are.
- Ask Julius what the unique values are in the date column.

MERGE

- Tell Julius to do an inner merge of common_stock_tickers with the metrics data on the ticker column.
- Ask Julius to show the first few rows.

SORT

- Ask Julius to sort the merged data set on date and to show the first few rows.
- Ask Julius to sort on ticker and date and to show the first few rows.
- Ask Julius to save the data set as an Excel file.

AGGREGATE BY GROUP

- Ask Julius to compute the mean marketcap by date.
- Ask Julius to compute the mean marketcap by sector on the date 2024-03-08.
- Ask Julius to compute the mean marketcap grouped by (date, sector) and to present the results as a two-dimensional table.
- Ask Julius to count the number of tickers grouped by (date, sector) and to present the results as a two-dimensional table.

MORE AGGREGATION

- Ask Julius to compute the percent change in marketcap by ticker.
- Ask Julius to compute the total marketcap by sector on each date and then to compute the percent change by sector.
- Ask Julius to compute the percent of firms for which pe < 0 grouped by (sector, scalemarketcap) and to present the results as a two-dimensional table.

TRANSFORM

- Ask Julius to create a new variable that is 1 if pe > 0 and 0 otherwise (a dummy variable).
- Ask Julius to group by date and rank by marketcap in descending order at each date and add the ranks as a new column in the data set.
- Ask Julius to group by (date, sector) and calculate the excess of pb over the median pb and add this variable to the data set.

ONLINE DATA

YAHOO FINANCE

- 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, …)
- Can get with yfinance library

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.

CAVEAT

On ex-dividend days, the percent change in the adjusted closing price is

$$\frac{P_t}{P_{t-1} - D_t} - 1$$

rather than what we might prefer:

$$\frac{P_t + D_t}{P_{t-1}} - 1$$

but this is a minor issue (small difference 4 days a year).

MONTHLY, ANNUAL, ... RETURNS

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.

EXAMPLE

- Ask Julius to run pip install yfinance==0.1.70 and to import yfinance as yf
- Ask Julius to use yf.download to get adjusted closing prices for SPY for the longest history available.
- Ask Julius to downsample the prices to end-of-month.
- Ask Julius to compute monthly returns as the percent change in the downsampled prices.

FEDERAL RESERVE ECONOMIC DATA

- Ask Julius to use the pandas datareader to get the history of crude oil prices from FRED.
- Ask Julius to get the history of inflation rates from FRED.

KEN FRENCH'S DATA LIBRARY

- Ask Julius to use pandas datareader to get the Fama-French factors from Ken French's data library.
- Ask Julius to list the datasets on Ken French's data library and to show the entire list in the chat (you may have to tell Julius to use get_available_datasets).
- Ask Julius to get the 48 industry returns from Ken French's data library.

SCRAPING

- 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.