Data Handling and Visualization

MGMT 675, AI-Assisted Financial Analysis

Kerry Back, JGSB, Rice University

Outline from Today

- Work with metrics.xlsx and tickers.xlsx.
    - Download from the course website.
    - Upload in Julius.
- Get online data from various sources.

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 the ticker column
- 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

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, ...)

Federal Reserve Economic Data

- Ask Julius to use the pandas-datareader to get crude oil prices from FRED.
- 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.

Visualization

- Distributions
    - Histograms, density plots, box plots, pie charts
- Bivariate
    - Line plots, scatter plots, scatter plots with regression lines
- 3D
    - 3D plots, contour plots

Formats

- jpegs, pngs, and pdf images
    - Tip: ask Julius to use the seaborn whitegrid style or other seaborn style.
- html
    - add hover data
    - animations
    - use plotly and ask Julius to save as html
    - Tip: ask Julius to use the plotly_white template or other plotly template

Features

- Set linewidth, linestyle, point size, colors
- Set axis labels and title and font sizes
- Control legend location
- Annotate points

Examples

- Ask Julius to create a bar chart of total marketcap by sector. Ask Julius to save as a jpeg.
- Ask Julius to use seaborn to create a box plot of marketcap by sector. Ask Julius to save as a png.
- Ask Julius to use plotly to create a box plot of marketcap by sector and to include the ticker in the hover data. Ask Julius to save as html.
- Ask Julius to calculate the daily percent changes in the crude oil price and to create a filled density plot using seaborn.