What data needed and where was it found?

Data sources - 3

SEC Website - FTD data

IEX Api - for stock chart information, primarily for close price and volume

Yahoo Finance API - for general stock ticker info, primarly the Outstanding Share and Current Float stats

Challenges included needing to read and sort the FTD data from the SEC. Text file was too large for python reader, had to manually copy/paste into Excel, save as CSV format, and then could use Pandas to import file into a dataframe, clean the data, split the columns, and parse the file to find exactly what was needed, before appending the parsed data to a dataframe, and then repeating for next CSV file in the path directory.

Each file was searched and parsed for FTD data based on stock ticker – was having issues with Symbols matching to similar but wrong symbols (for example, GME to GMED) so instead changed search parameters to CUSIP number so we had a guaranteed unique identifier to find the data we were looking for.

IEX was used to lookup historical data and price history for stock prices. Was matched and merged against FTD data from SEC site.