Using SEC EDGAR 10-k Data

1. Scrape the data
   1. Find the most recently published 10-k by entering the company’s ticker
   2. Break up the 10-k into easier parts to parse
      1. Get the filing summary of the 10-k
      2. Request/parse the filing summary
      3. Create a list to store all the individual components of the report (these components are all found in the filing summary)
         1. Create a dictionary to store all the different types of info about each component of the report in the list
      4. Now create a list to store all the statement URL’s
   3. Organize the financial statements. Let’s assume we want all the statements in one dataset (finished)
      1. Loop through each statement URL and create a dictionary that will store the different parts of the statement:
         1. Headers
         2. Sections
         3. Data
      2. Request the statement file content
         1. Find all the rows, figure out what type of row it is, parse the elements, and store in the statement file list
         2. Append all the statement data to the master list
2. Transform the data into a data frame (finished)
   1. Grab the proper components (headers and data)
   2. Use pandas to put the components into a dataframe
   3. Define the Index column, rename it, and we need to make sure to drop the old column once we reindex.
   4. Get rid of miscellaneous characters like ‘$’, etc
   5. Convert all the data from strings to float
   6. Change column headers
3. Output the data to excel
   1. Add a title row
   2. Format headers
   3. Format section headers
   4. Format data
4. Make the excel data manipulative
   1. Need code to systematically assign variables to certain cells
5. Do some analysis on the output data
   1. Profitability analysis
      1. Profit margin = Net Income / Sales Revenue
      2. Return on assets = Net Income / Average Total Assets
         1. Average Total Assets = (total assets of current year + total assets of previous year) / 2
      3. Return on Equity = Net Income / Average Stockholders’ Equity
         1. Average Stockholders’ Equity = (stockholders’ equity of current year + stockholders’ equity of previous year) / 2
   2. Liquidity analysis
      1. Working Capital = current assets – current liabilities
         1. If WC > 0, print: “This company has sufficient current resources to cover its liabilities.”
         2. If WC > 0, print: “This company does not have sufficient resources to cover its liabilities.” AND trigger some kind of alarm bell
      2. Current Ratio = current assets / current liabilities
         1. If CR > 1, print: “This company has CR the amount of dollars in assets for each dollar in current liabilities.”
         2. If CR < 1, print: “This company has CR the amount of dollars in assets for each dollar in current liabilities.” AND maybe want some other kind of reasoning.
   3. Solvency analysis
      1. Debt-to-equity ratio = Total Liabilities / Total Stockholders’ Equity
         1. Measure of leverage, an indicator of the relative size of financing from creditors versus financing from owners
         2. The higher the debt-to-equity ratio, the lower the solvency, and the greater the risk of default
      2. Interest Coverage Ratio = (Net Income + Interest Expense + Tax Expense) / Interest Payments
         1. Indicates how many times a company can cover its interest charges before taxes