

Getting Started With the 10K_10Q_XML_Parser

10K_10Q_XML_Parser application will take as input the same CSV file that was used for the 10Q_10K_Downloader application. See example input sheet of tickers / CIK's.



Ticker_CIK_List.csv

Assumptions about file locations –

It is assumed the input sheet of tickers / CIK's is located in the current working directory.

It is assumed the XML files to be parsed for each ticker are located in ~/OutputXML/Ticker (which is the directory structure generated by the downloader app). So, there will be separate directory for set of files associated with a ticker.

For example, if current working directory is /usr/brian, the following directories would be used to read all the input XML files for AGNC and all the generated CSV files for AGNC would be written to this directory.

/usr/brian/OutputXML/AGNC

The extracted CSV files for a ticker are combined into a single large file with a name that starts with the ticker followed by '_Combined'.