Open-ended Capstone Step 4: Data Exploration

I’ve reviewed the data from all sources and the data seem homogenous in each column. Therefore, did not see any columns that need cleaning or exception handing. The columns that are not necessary have been removed and only the columns listed in the ER diagram will be loaded into the database. The ticker symbol which represents each company will be used to join data from multiple sources. Since the data/table from different sources will be joined often, view tables may be used to make the querying easier.

One thing I need to figure out is how to efficiently store a twitter record that has multiple ticker mentions. I see two options for this. Load the multiple ticker symbols as a one long string or load a tweet multiple time but with unique ticker symbol.