DATA QUALITY REPORT

Data Source

The data been worked on is financial times series data set, to be explored for insight extraction as well as model prediction. The data sets include; company information, stock prices, price, and market indices.

Data overview

Upon loading the data sets, we can view the various columns in each set and their respective data types. The following was observed;

a) DATA TYPES MISMATCH OVERVIEW

- Ticker column appear as object rather than strings
- Date column also appears as object rather than a date datatype
- ipo_date column is also an object rather than a date datatype¶

b) MISSING VALUES OVERVIEW

- Stock prices data set, had presence of missing values. All other data sets were free of missing values.

c) DUPLICATES OVERVIEW

- There were no observed duplicates in the data sets

DATA ISSUES ADDRESSED

- Data type corrections needed = Ticker and date columns would be corrected to ensure they reflect their true data types (i.e string and date type)
- Missing dates discovered: - i.e either applying Forward fill (ffill) which carries last observation forward (common for financial data). or - Backward fill (bfill) method which fills with next available value.

Data was cleaned and ready for further use.