**Proposal:**

Daily count of the number of Elon Musk’s tweets vs. the following day’s Tesla stock market price change (close - open) open to close.

**EXTRACT (Data Sources)**

* Tesla stock prices: <https://www.kaggle.com/rpaguirre/tesla-stock-price>
* Musk Tweets: <https://data.world/adamhelsinger/elon-musk-tweets-until-4-6-17>

**TRANSFORMATION**

* Tweets:
  + Formatting tweet dates from date &  time to date
  + Drop ‘text’ field
  + Count of tweets, grouped by date
* Stock:
  + Calculation of stock deltas (closing - opening)
  + Drop all other fields
* Joining:
  + Join tweet counts and stock deltas on dates (stocks\_date = tweet date +1)

**LOAD**

The team selected SQL to load the data transformation results. This decision was made as only one table was required to load the final results. The

The type of final production database to load the data into (relational or non-relational).

* Relational: SQL (“So we can perform joins and execute query statements!” -Rowena)

The final tables or collections that will be used in the production database.

* Table of daily Musk tweet counts and Tesla stock price deltas
* Scatter plots of stock prices, daily tweet count, and date

* Extract: your original data sources and how the data was formatted (CSV, JSON, MySQL, etc).
* Transform: what data cleaning or transformation was required.
* Load: the final database, tables/collections, and why this was chosen.

Please upload the report to Github and submit a link to Bootcampspot.