**ETL Project**

Proposal presentation (July 27, 2019)

**Topic**: Effect of Clinical Trials Outcomes on Stock Market Price

* Questions to answer
  + What phase of clinical trial has the maximum effect on stock price?
  + Does condition (disease state) being studied have different effects on the stock market price?
* Data sources used to answer the questions
  + Clinicaltrials.gov
    - Extracted clinical trials data completed in the USA
    - Obtained following data:
      * Title
      * Status
      * Condition
      * Sponsor/Collaborators
      * Phases
      * Funded by
      * Start date
      * Completion date
  + Kaggle – Huge stock market database
  + Still looking for more recent stock market data for
    - SNP500 companies (63 healthcare companies)
    - Nasdaq or nyse for looking at small companies (look for api)
* Plan to break-up the tasks

|  |  |
| --- | --- |
| Task | Leader |
| Data Collection | IV, PH, BG, RM |
| Data Cleaning | BG |
| Draw ERD diagram | IV, PH, BG, RM |
| Place it in Database (SQL) | RM |
| Pandas analysis | PH |
| HTML/CSS | IV |

Final project should include:

Data directory, Schema for final data (if need be, resources folder)