**Project Report**

At the end of the week, your team will submit a Final Report that describes the following:

* **E**xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).
* **T**ransform: what data cleaning or transformation was required.
* **L**oad: the final database, tables/collections, and why this was chosen.

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

Datasets:

Traffic Fatalities for 2015 and 2016 by State and City (Geographic Locator Code)

https://www.kaggle.com/usdot/nhtsa-traffic-fatalities - BigQuery

Geographic Locator Codes - .xlsx

https://www.gsa.gov/reference/geographic-locator-codes/glcs-for-the-us-and-us-territories

Historical Weather Data (including 2015-16) for 30 cities. -.csv

https://www.kaggle.com/selfishgene/historical-hourly-weather-data/data

Extract:

Extract the 30 city names from csv

Extract their respective geographic locator codes if available from .xlsx

Use BigQuery to extract traffic fatalities that

occur in those cities

occured during the years for which we have weather data.

Task Division:

D: Use BigQuery to extract the data

T: Get the trimmed city list using the geographic locator codes and historical weather data

Transform:

Fatalities:

City Code

Atmospheric Conditions 1-3

Datetime

Number of Fatalities

Weather:

City Code

Datetime

Weather Description (fog/mist/haze, rain/thunderstorm, snow, no precipitation)

-parse description into one of the above 4 cats.

Temperature

Humidity

Pressure

Wind Speed

Location:

State

City Name

City Code

Task Division:

T: Weather

D: Fatalities

Load:

Postgresql RDBMS

Three tables:

fatalities

weather

locations