# **Project Proposal Document:**

**Project Name: Analysis for Future FEMA Funding**

**Team: GENIUS TEAM ☺**

Mary Brown  
Pournima Joshi  
Jacqueline McBean-Blake  
Billy Martinez  
Brendan Law

**Proposal Date:** 11/2/2019  
**Due Date:** 11/6/19

**Statement of work**:

We will analyze US natural disaster data to see frequency , types, strength/impact and location, to compare to FEMA disaster declarations to give insight on forecasting for FEMA budget planning.

**Data sources**: (include details and why pick these data sources?)

1. US\_ZIP\_CODE\_LAT\_LONG (XLS file, will convert to a reference table with us city, state, zip, latitude and longitude)

<https://www.zip-codes.com/account_database.asp>

1. FEMA Disaster Declaration (xls file, 1953-2019 will convert to raw data table and then pull into main table)

<https://www.fema.gov/media-library/assets/documents/28318>

1. Earthquake data (csv file, 1965-2016 with date, latitude, longitude, type, strength/size)

<https://www.kaggle.com/usgs/earthquake-database#database.csv>

1. Tornados data (csv file 1950-2018 with date, state, latitude, longitude, type, strength/size):

[https://www.spc.noaa.gov/wcm/data/1950-2018\_actual\_tornadoes.csv](https://www.spc.noaa.gov/wcm/data/1950-2018_actual_tornadoes.csv" \t "_blank)

1. Wildfires (csv file 1992-2008 with date, latitude, longitude, type,strength/size)

<https://catalog.data.gov/dataset/fire-program-analysis-fire-occurrence-database-feature-layer>

**Proposed ETL:**

1. We will load each data source into individual raw data tables (Earthquake, Tornado, Wildfire, FEMA, Location and Source)
2. We will create a source table that identifies the source tag and the URL
3. We will create a location reference table based on the US\_ZIP\_CODE\_LAT\_LONG input file
4. We will transform the individual weather events from individual tables (Earthquake, Tornado, Wildfire and FEMA) into a uniform table (NATURAL\_DISASTER). For this we will transform the data:
   1. to use uniform properties (date, city, state, strength, type)
   2. to tag each record with SOURCE tags AND
   3. to limit data to US weather events.

**Proposed Final Schema:** PostgreSQL

**Approved by:**

**Final Project Report (11/6/19 or 11/7/19)**

At about 8 PM, 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.

Present 3-4 minutes on the project discussing some pain points and how did you resolve them. Only one student from each team should present. 1 min for any Q&A to the class.