**Participants**: Jonah Brennan, Ron Javor, Cassie Liu

**Objective**

We sought to create a database information that a tourist may use to describe Chicago safety within a given month of travel. We used recent data covering calendar years 2017-2018 to insure current accuracy of the data. The date resolution was initially planned to be daily, but was moved to monthly

The databases that were extracted were;

* Weather – format input as a JSON file, downloaded from openweather.com. The data was purchased specifically for Chicago spanning 2017-2018 on a daily basis.
* Crime – format input as a CSV from the Chicago PD website. The data was free, and covered all incidents for Chicago spanning 2017-2018 on a daily basis. We focused on 3 specific violent crimes that are most likely to affect a tourist: Robbery, Assault and Sexual Assault
* Traffic Accidents – format as a {} spanning 2017-2018 on a daily basis.
* Visitors – (tbd on the source and details on this extraction due to medical hospitalization by one of our team members). The data was likely extracted from the Chicago Bureau of Tourism and the cleaned data was ultimately provided as a CSV file that tallied number of visitors on a monthly basis, spanning 2017-2018.

The majority of cleaning and merging was performed with Pandas dataframes. The key challenge in this activity was getting the dates in each dataset to the same format “YYYY-MM”. The merges were all performed with Dates as the reference.

We chose Mongo DB as the database format to load the cleaned and merged (transformed) data into. Mongo DB was chosen for speed and simplicity as compared to SQL databases which can be complex and time-consuming to construct. Additionally, the transformed data was also output as a JSON file for additional analysis as desired.

The HTML page …