Justin - Introduction and Cleanup of CSV data queries (importedDatatosqllite)

Yahia – 3 main graphs (boxplot, CTG, YoY)

Madhu – Dashboard, Leaflet and changes in the filter

Diane – HTML and Flask

Things to talk about

* Different types of graphs and their data shown (story)
* Acquisition and cleanup of data (code)
* App.py code (code)
* Main.js code (code)
* Html code (code)

1. Different types of graphs and their data shown (story)

# The Story of Traffic Violations

We wanted to look into the traffic violations for a specific area over time, to see if traffic stops are getting less frequent or more frequent. Traffic violations such as not wearing seatbelt, property damage, personal damage, speeding were all found in a CSV from Kaggle called “Traffic Violations in USA”. However, the CSV contained to much data to represent and needed cleaning up. Our hypothesis was that the number of violations would be going up, but year over year data shows an improving trend of violations going down.

# Clean Up Process

From the CSV we found several regions of data based off of the State column, and found it originally started with Geolocation, Gender columns, along with specific violations as our dataset to analyze. However the geolocation data set ran into 490k unique hits and needed further refinement. Our solution to clean this up was dropping the Geolocation column from the data and used another field Districts to group the rest of the data by.

# Flask Root

To push the data into our HTML, we used Flask if

Traffic Violations and Year over Year growth (YoY)

Story: we looked at the 2013-2016 (4 years of data), and noticed a steady increase in 2013-2015, and then a drastic decline in 2016 for most of the districts and violation types.

Code: graph created in plotly, using 2 y axis’s (left y axis is Traffic Violation Counts, and right side is YoY growth).

From the CSV we are using

1. Year

2. District

3. Violation ()

4. Citation (warning)

5. Impact (personal damage and property damage)

We display the Traffic Violation of YoY (year over year) growth as static

The heaviest year in terms of violations by district is 2014, then we see a significant decrease in traffic violations after this year, which brought us to the question, why?

Our hypothesis for this are use of Red light cameras (cameras that pick up when you run a red light). An increase in speed detection sensors that tell you how fast you are going. Increase in speed traps (where a patrol car rests on side of the road pointing a radar gun). None of the data from our CSV contains the answer.