Analysis of U.S. Car Accident

Team Members: Tianchen Wang, Zitao Cheng, Yu Li

Introduction

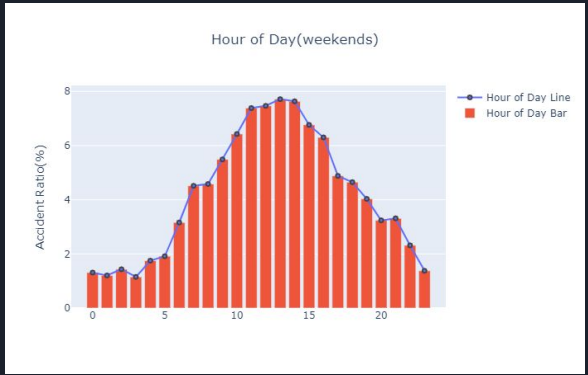
- This project goes over the dataset about the U.S. car accident between 2016 and 2020. It basically has these attributes: temperature, precipitation, wind speed, pressure, humidity, etc. To play around these attributes gets the relationship along with the occurrence rate of a car accident in the United States.
- Whether the occurrence rate is related to the day or night.
- Whether the time zone would affect the car accident.
- What range of visibility would cause a car accident more frequently?
- Which time period has the highest occurrence rate of a car accident in a day?

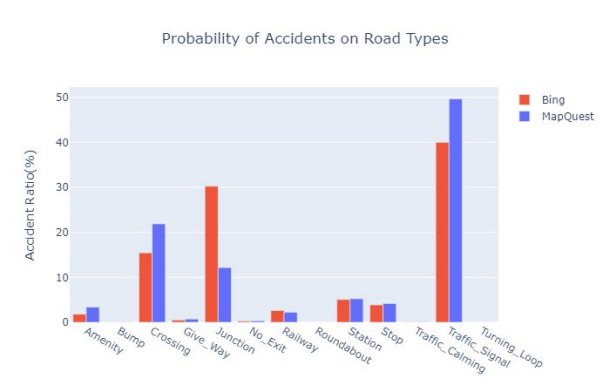
What changes have been made

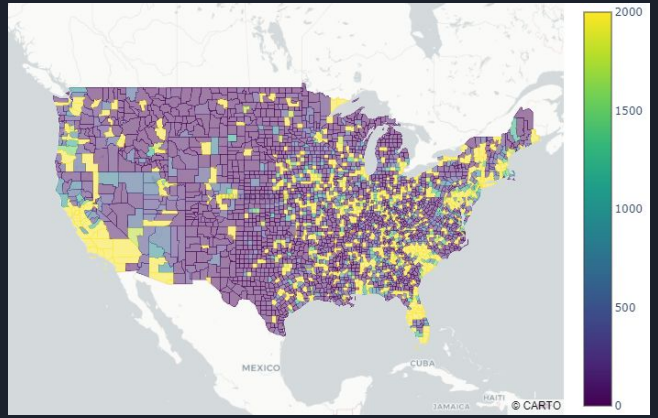
- Use plotly visualization API to visualize our data
- Adjust the structure of the report
  - Introduction of Data set and Data selection
  - Adding own graph analysis on the report
  - Answered provided questions
- Correct mistakes made in the past
  - Do Data management instead of Data Reduction

What have been accomplished so far

1. Data Cleaning (fill a mean or mode value in missing place)
2. Data Selection (select useful attributes)
3. We plot several graphs based on the data we mined







What remains to be done

- A figure that will review the relationship between accident severity and sunrise or sunset
- A histogram that show accident rates at different weather conditions.
- Evaluation
- Find more authoritative articles help with our evaluation