Accident Analyzer:

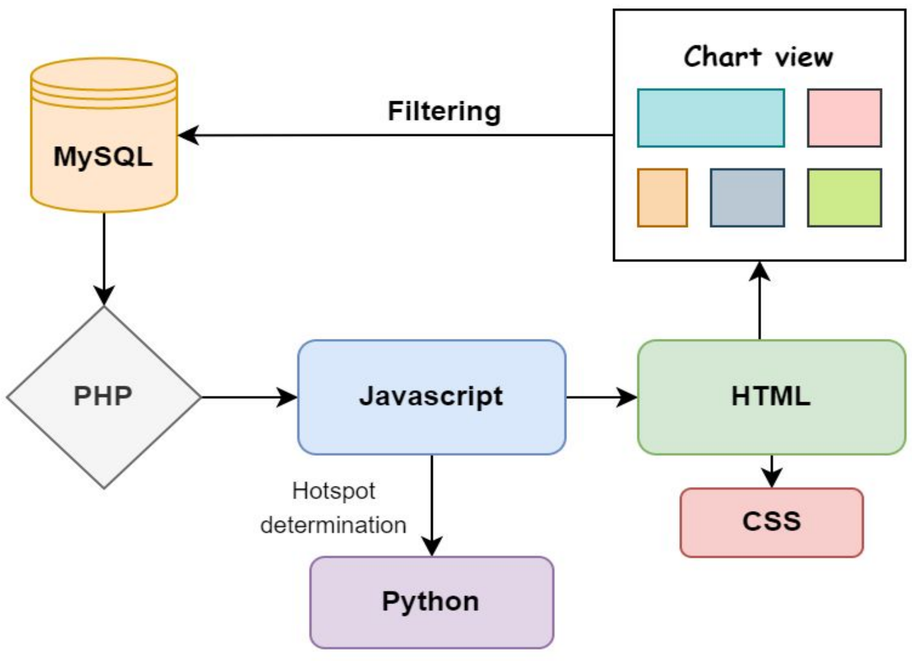
Understanding Vehicle Accident Patterns in the United States

Kyle Otstot, Shreyash, Jaswanth, Anudeep, Hruthik, Nitesh

Ira A. Fulton Schools of Engineering
Arizona State University

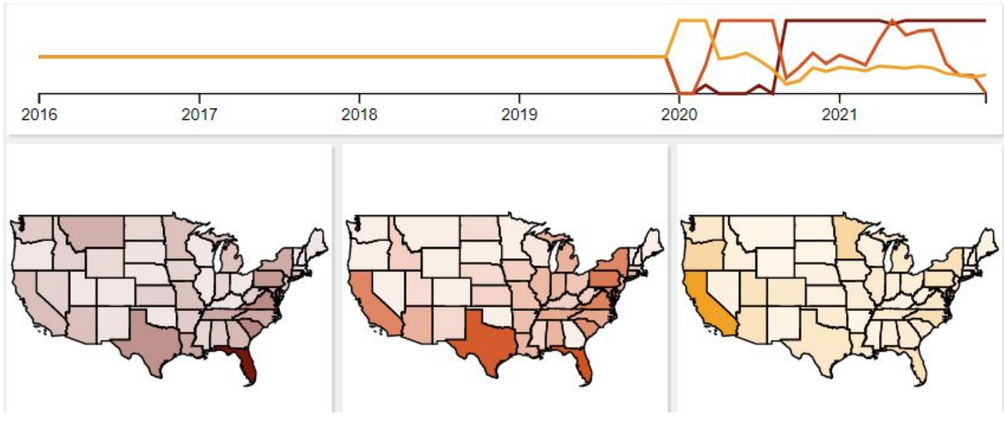
Introduction

- **Motivations** - To understand the impact of Accidents with the help of spatio-temporal data.
- The **Accident-Analyzer** system allows for users to identify local hotspots, visualize accident trends over time, and filter the data by key weather categories in real-time.
- **Dataset** - US Accident Dataset which is available on kaggle. The dataset contains 2.8 millions rows and we have used the following columns: Start_time, End_time, Distance, City, County, State, Weather_Condition, Sunrise_Sunset
- The visualization was primarily created with the *D3.js (v7)* and *Leaflet.js* JavaScript libraries, the dataset preprocessing was done using *Python*, and the data is stored/accessed via a *MySQL* database. Lastly, the hotspot computation is performed by a client-side Python library called Pyodide using Scikit Learn's non-negative matrix factorization (NMF) implementation.

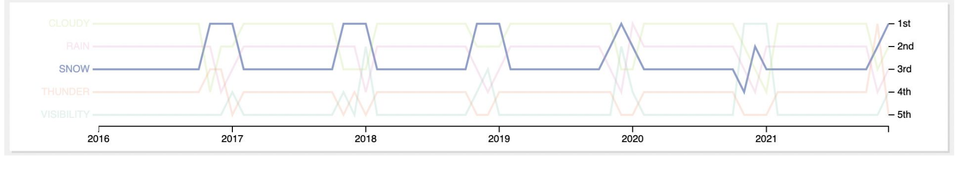


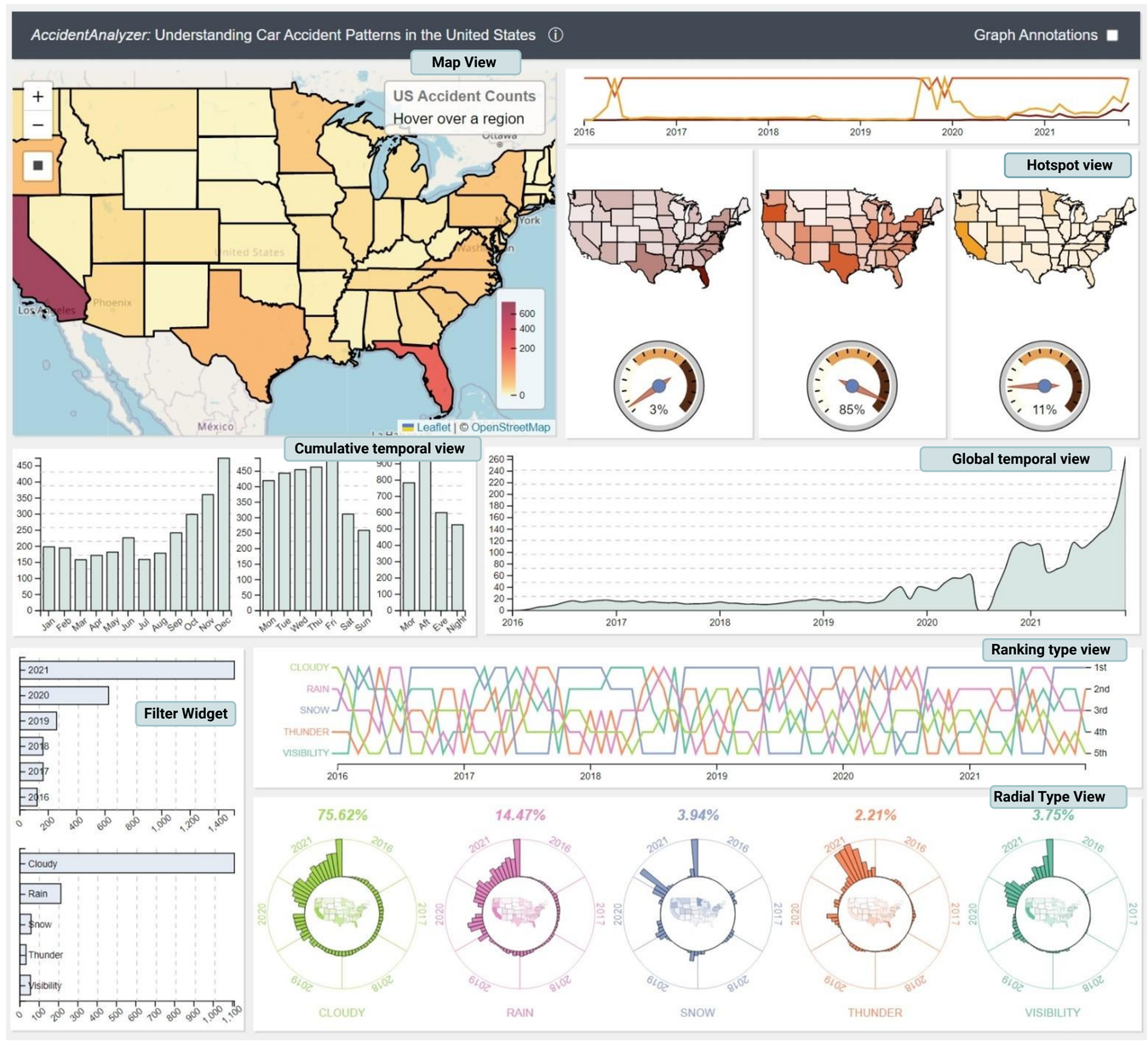
Case Studies

1. **Vehicle Accidents before/after COVID-19**



2. **Accident Severity during Winter in New York**





System Description

- **Map View:** A geographical map combined with a choropleth map that color-codes the number of accidents reported at each site in the region.
- **Hotspot View:** A collection of multiple maps with the spatial distribution of each hotspot. The spatial accident distribution is color-coded similarly to the map view.
- **Global Temporal View:** An area chart that provides an overview of the number of accidents reported over the whole time period.
- **Cumulative Temporal View:** a bar chart displaying the number of accidents by month, day, and period of day.
- **Ranking Type View:**A collecting of polylines that encode the relationship between weather condition, time, and accident severity
- **Radial Type View:**a collection of bar charts with a radial layout. Each chart represents a different weather condition, and the number on top shows the percentage of each weather condition in the dataset.
- **Filter Widget:**a collection of histograms over time, severity, and weather conditions, included for the user to quickly filter the data by the most relevant attributes.

Extension

- **Graph Annotation** - we implemented a checkbox feature which when selected will add a hovering tooltip to the interface
- **Temporal Hotspot View** - this view breaks down the temporal distribution of the hotspots,where each line represents a hotspot
- Hotspots are computed in real time and based on the current selected filters