US Traffic Fatalities Analysis

Save lives, prevent injuries, reduce vehicle-related crashes

Team Members: Deepshikha Purwar

Meetu

Tarun Sharma

DATA DESCRIPTION AND BUSINESS PROBLEM

Dataset: **accidents.csv from National Highway Traffic Safety Administration.

Business Problem: To analyze the factors that affect the likelihood of the vehicle crashes and take preventive measures

End Users: Police department, Traffic safety offices

**Source: https://www.kaggle.com/nhtsa/2015-traffic-fatalities

Data Description:

- US traffic fatalities data, including district of columbia and puerto rico, in the year of 2015 created and maintained by national centre of statistics and analysis(NCSA). NCSA is the component of National Highway Traffic Safety Administration(NHTSA).

- This dataset has details of motor vehicle crash resulting in the death of occupant/non-occupant of vehicle within 30 days of crash.

- 17 Datasets in total related to accident details, weather conditions, vehicle information including VIN, vehicle damage area etc. but we will focus on **Accident.csv** dataset

Accident.csv:

Number of columns = 52

Number of Rows = 32166

Main Attributes of Accident Data File:

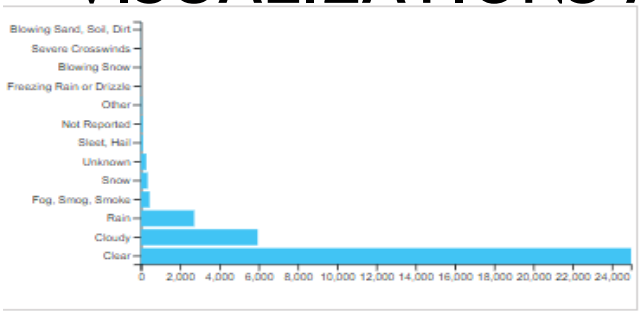
- **General Accident Information: Fatals, Drunk_dr**
- **Geolocation attributes: Latitude, Longitude, State**
- **Time of crash: Month, Day , Day_Week, Year, Hour, Minute**
- **Atmospheric Conditions: Weather, Lgt_Cond**

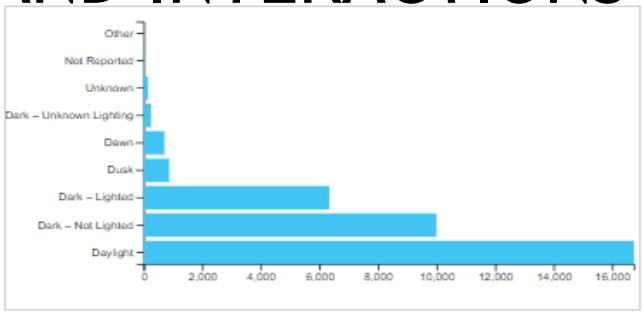
DATA PREPROCESSING

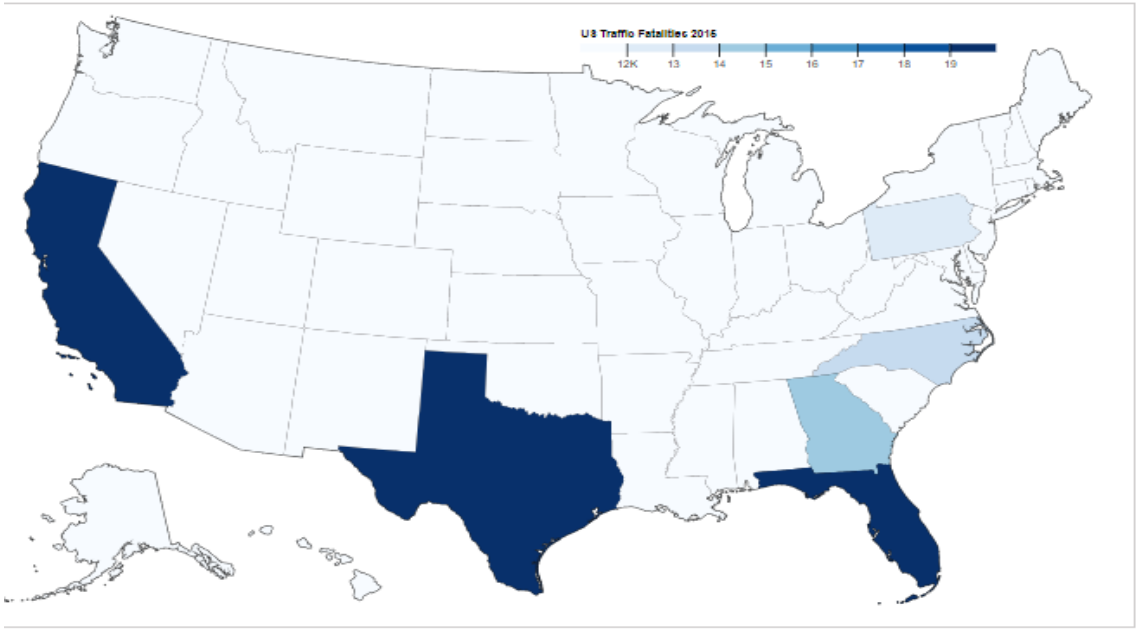
- Removed Missing Data
- Converted codes of below attributes to actual names using FARS user manual.

State
Route
Day_Week
Weather
Lgt_condition

VISUALIZATIONS AND INTERACTIONS







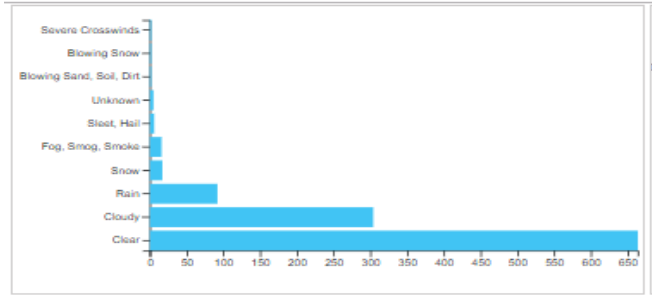
We have included 3 visualizations to achieve the following 3 different tasks –
Visualization 1: Bar chart that depicts relationship between weather and the fatalities.
Task: Identify if there is trend between weather and the accidents.
Observation: Surprisingly most fatalities occur in clear weather.

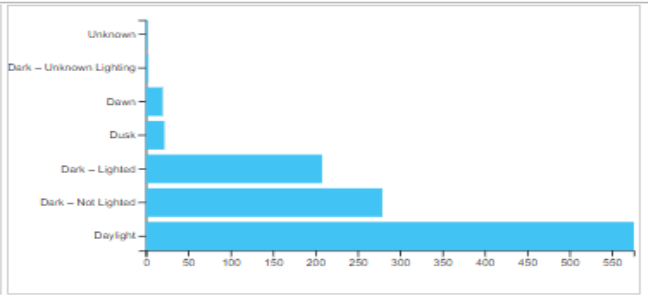
Visualization 2: Bar chart that depicts relationship between light condition and the fatalities.
Task: How light condition and number of accidents are related?
Observation: Surprisingly most fatalities occur in full light conditions.

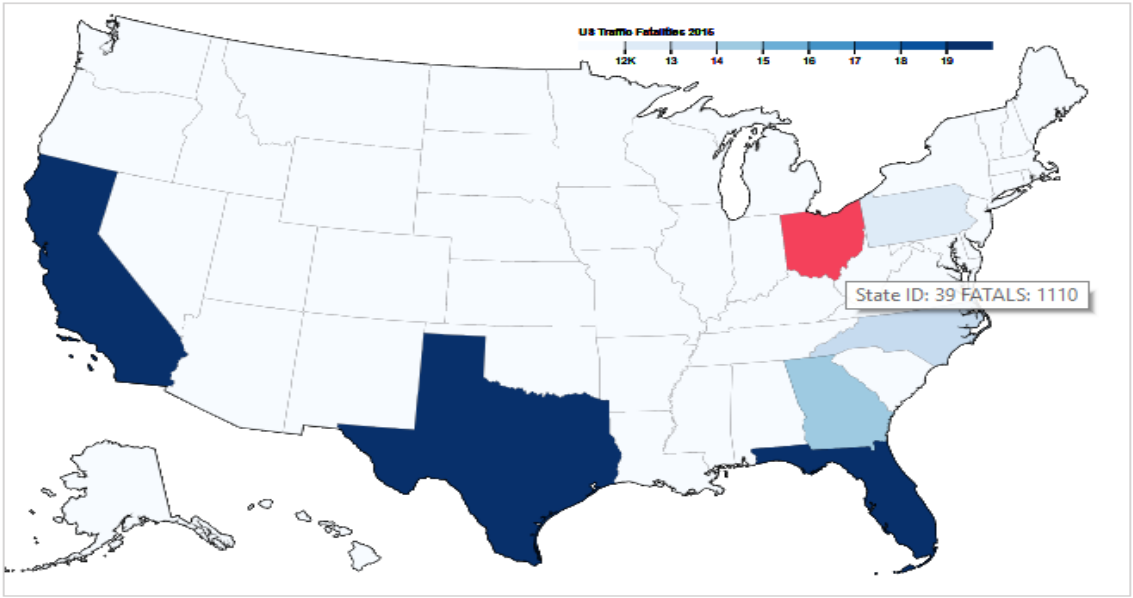
Visualization 3: Geographic Map that shows different states data related to fatalities
Task: To show different states persistent to traffic accidents
Observation: Different states are shown with fatalities in the label

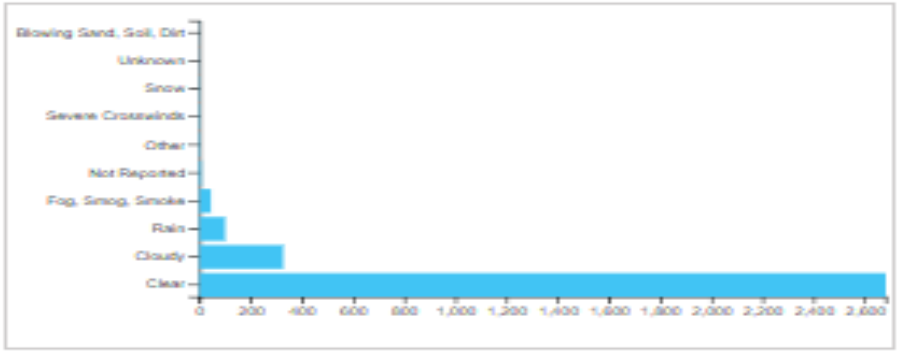
Interaction: All these 3 visuals interact via D3. On hovering cursor to a particular state, the data related to number of fatalities is displayed as label, and the corresponding weather condition and light condition related data is displayed in the bar graphs, shown above the map. Hovering over the bars in the bar charts colors the associated states in the map according to the color legend.

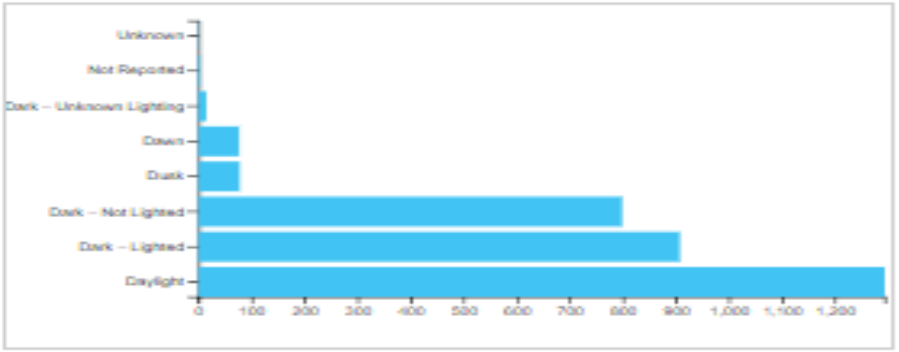
EXAMPLE SHOWING HOW VISUALIZATION IS WORKING

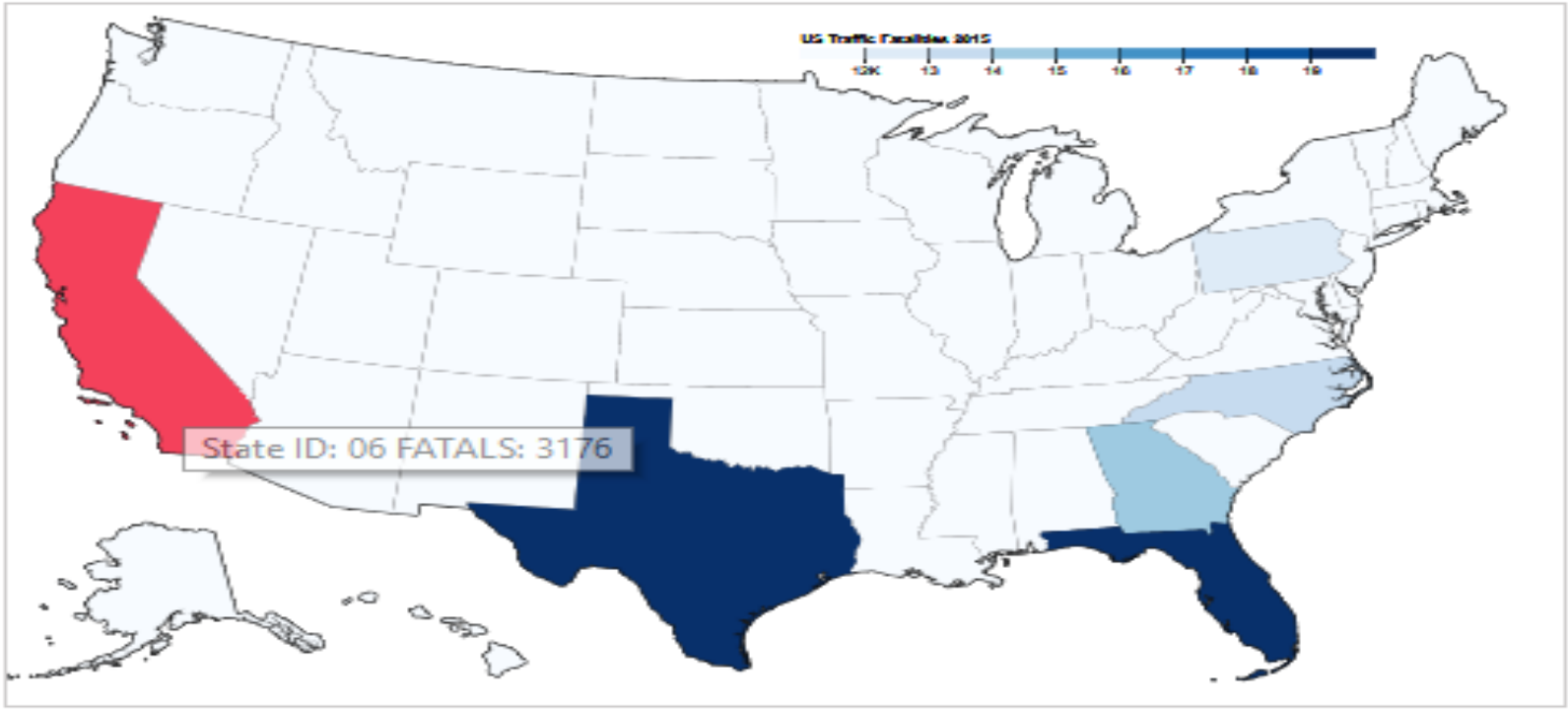







Blowing Sand, Soil, Dirt
Unknown
Snow
Severe Crosswinds
Other
Not Reported
Fog, Smog, Smoke
Rain
Cloudy
Clear
0
200
400
600
800
1,000
1,200
1,400
1,600
1,800
2,000
2,200
2,400
2,600


Unknown
Not Reported
Dark – Unknown Lighting
Dawn
Dusk
Dark – Not Lighted
Dark – Lighted
Daylight
0
100
200
300
400
500
600
700
800
900
1,000
1,100
1,200


US Traffic Fatalities 2015
12K
13
14
15
16
17
18
19
State ID: 06 FATALS: 3176