**IFT 598 Data Visualization & Reporting for IT**

Hima sree Chalasani

Nikhil Jain Sanghami Padmaprabha

Venkata Bala Nagendra Srirangam

Project - Phase III Dashboard Implementation

Asmaa Elbadrawy

April23rd, 2023

**Section 1: The Dashboard**

**Map

Description automatically generated**

Graphical user interface

Description automatically generated

**Graphical user interface

Description automatically generated**

This dashboard helps us visualize the data from the US Accidents dataset. The dataset contains all information about an accident like the time of the accident, severity, location, road type, the road sign at which the accident occurred and others. This dashboard will show information that will help citizens to take calculated decisions, governments to make improvements and insurance companies to design policies.

**Section II: The Dataset**

This is the original dataset:

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

**The attributes in the dataset:**

“Attributes in Dataset

* ID: This is a unique identifier for each accident record.
* Source: This is the source of the accident report (e.g., mapquest, Bing, etc.).
* TMC: The Traffic Message Channel (TMC) code is a unique identifier assigned to road segments for traffic information exchange.
* Severity: This is a measure of the severity of the accident, ranging from 1 (least severe) to 4 (most severe).
* Start\_Time: The date and time at which the accident started.
* End\_Time: The date and time at which the accident ended.
* Start\_Lat: The latitude of the starting point of the accident.
* Start\_Lng: The longitude of the starting point of the accident.
* End\_Lat: The latitude of the ending point of the accident.
* End\_Lng: The longitude of the ending point of the accident.
* Distance(mi): The length of the road extent affected by the accident in miles.
* Description: A textual description of the accident.
* Number: The street number of the address where the accident occurred.
* Street: The street name of the address where the accident occurred.
* Side: The relative side of the street (Right/Left) in the direction of the address where the accident occurred.
* City: The city where the accident occurred.
* County: The county where the accident occurred.
* State: The state where the accident occurred.
* Zipcode: The postal code where the accident occurred.
* Country: The country where the accident occurred.
* Timezone: The timezone in which the accident occurred.
* Airport\_Code: The airport code nearest to the location where the accident occurred.
* Weather\_Timestamp: The time at which the weather conditions were observed.
* Temperature(F): The temperature in Fahrenheit.
* Wind\_Chill(F): The wind chill temperature in Fahrenheit.
* Humidity(%): The relative humidity.
* Pressure(in): The air pressure in inches.
* Visibility(mi): The visibility in miles.
* Wind\_Direction: The direction of the wind.
* Wind\_Speed(mph): The wind speed in miles per hour.
* Precipitation(in): The amount of precipitation in inches.
* Weather\_Condition: The current weather condition.
* Amenity: A boolean indicating whether there was an amenity (e.g., a restaurant, gas station) near the accident location.
* Bump: A boolean indicating whether there was a speed bump near the accident location.
* Crossing: A boolean indicating whether there was a crossing (e.g., a pedestrian crossing, a railroad crossing) near the accident location.
* Give\_Way: A boolean indicating whether there was a give way sign near the accident location.
* Junction: A boolean indicating whether there was a junction (e.g., an intersection, a roundabout) near the accident location.
* No\_Exit: A boolean indicating whether there was a no exit sign near the accident location.
* Railway: A boolean indicating whether there was a railway near the accident location.
* Roundabout: A boolean indicating whether there was a roundabout near the accident location.
* Station: A boolean indicating whether there was a station (e.g., a bus or train station) near the accident location.
* Stop: A boolean indicating whether there was a stop sign near the accident location.
* Traffic\_Calming: A boolean indicating whether there was a traffic calming measure (e.g., a speed bump, a roundabout) near the accident location.
* Traffic\_Signal: A boolean indicating whether there was a traffic signal (e.g., a traffic light) near the accident location.
* Turning\_Loop: A boolean indicating whether there was a turning loop near the accident location.
* Sunrise\_Sunset: Indicates whether the accident occurred before or after the sunrise or sunset.
* Civil\_Twilight: Indicates whether the accident occurred during civil twilight.
* Nautical\_Twilight: Indicates whether the accident occurred during nautical twilight. Nautical twilight is the period after civil twilight.
* Astronomical\_Twilight: Indicates whether the accident occurred during astronomical twilight.” [1]

**Pre-processing on the data:**

1. Extracted Start time hours as a separate column in Tableau so that it will be easy to visualize the hour-wise data.
2. Excluded certain weather conditions as much data was not available for certain weather conditions and grouped the common weather conditions together to avoid too much data.
3. Excluded the accident IDs column as it was not serving any purpose in our visualizations.
4. Removed the data column named precipitation as this info was not mentioned for all accidents.
5. Calculated the total time of an accident from start time and end time data provided.

**Section III: Dashboard Users**

The potential users of the dashboard could vary depending on the specific focus and purpose of the dashboard. However, some potential users of the US Accidents dashboard could be:

* **Transportation companies:** They can use the dashboard to analyse the frequency and severity of accidents on different routes and make data-driven decisions to improve the safety of their vehicles and drivers.
* **Government agencies:** They can use the dashboard to identify high-risk areas and take necessary measures to improve the road infrastructure, traffic management, and emergency response systems.
* **Insurance companies:** They can use the dashboard to evaluate the risk of insuring vehicles and drivers in different locations and set premiums accordingly.
* **Researchers and analysts:** They can use the dashboard to study the patterns and trends of accidents over time and across different regions and generate insights to inform public policy and transportation planning.
* **General public:** They can use the dashboard to stay informed about the latest traffic and weather conditions and plan their routes accordingly to avoid potential accidents and delays.

**Section IV: Questions**

1. What is the distribution of accidents across different states in the US? Which states have the highest and lowest numbers of accidents?
2. What is the distribution of accidents across different days of the week and the severity trends?
3. What is the distribution of accidents across different times of day?
4. How do weather conditions and other factors (such as road surface conditions or visibility) impact the likelihood of accidents occurring throughout the year?
5. What is the average time taken to clear an accident at different locations?
6. How does the frequency and severity of accidents vary across different traffic signs, such as bumps, crossing, junctions?
7. What is the year-wise severity trend of the accidents?
8. How does the twilight time affect the accident severity?
9. How does the severity of the accident affect the traffic congestion (distance of the road affected due to accident)?
10. How does the visibility in miles affect severity and number of accidents?

**Section V: Dashboard Plots**

**Plot 1: Accidents in each state**

Map

Description automatically generated

The above plot shows the number of accidents that have occurred in each state. This will help transportation companies, government agencies and insurance companies to take calculated decisions based on the state wise data. This plot has an additional filter to choose the state which will display the data for one particular state also, Severity level filter which allows the user to filter the accidents based on the severity level and also the year wise filter.

**Plot 2: Distribution of accidents across different days of the week**

Graphical user interface, application

Description automatically generated

The above plot shows the number of accidents that have occurred on each day of the week. This plot shows relationship between days of the week and the number of accidents. In addition to the day-wise accidents, we also have the severity control to check the severity of the accidents. We also have the state control to filter out the state wise data and the time of the day control. The plot successfully shows the accidents that have occurred over different days of the week.

**Plot 3: Distribution of accidents accross different times of the day**

Graphical user interface, application

Description automatically generated

The above plot shows the relationship between the number of accidents and the time of the day. This will help us figure out if the accidents are occurring more at night or day. We have divided the day into different parts as Early morning, evening, midnight, morning, night and noon for easy grouping and identification. This will help the authorities to find out the times of day when there are more accidents and will help them take additional measures.

**Plot 4: Impact of weather conditions on accidents**

Graphical user interface, text, application

Description automatically generated

The above plot shows the impact of weather conditions on accidents. This will help the dashboard users to figure out how harsh weather conditions impact the possibility of accidents and will help them stay cautious while driving during harsh weather conditions.

We also have the control to check the severity of accidents and filter out data for each weather condition.

**Plot 5: Time taken to clear an accidents**

Chart

Description automatically generated with low confidence

The above plot shows the time taken to clear the accident in each state. This will help the authorities to check the time taken to clear an accident in each location and also improve their response systems for quicker help. This visualization as a control to filter out the state which will help the state wide authorities to filter out the data corresponding to their state.

**Plot 6: Accidents across different traffic signs**Graphical user interface, application

Description automatically generated

The above visualization shows the accidents that occur at different traffic signs. We can see that there are a greater number of accidents at the junctions. This will help the riders to be cautious while driving at the accident-prone areas and it will also help the authorities in installing more signs to alert drivers and to ensure their safety. This also has the severity level control to know the severity and also the state control that will help the state level authorities.

**Plot 7: Year wise severity trend of the accidents**

Graphical user interface, application

Description automatically generated

The above plot shows relationship between the number of accidents that have occurred across different years and their severity level. This will help the authorities know if the accident cases are increasing year by year or if there is any improvement in reducing the number of accidents. This will help the authorities take additional measures to control the accidents in the coming years.

**Plot 8: Twilight time and severity of accidents**

Graphical user interface, application

Description automatically generated

This plot shows the number of accidents during the twilight times when brightness is increasing or decreasing. This will help the people stay extra cautious and it will also help the authorities regulate the traffic in risky areas during this time. We also have the severity level control to filter out the different severity levels.

**Plot 9: Severity with traffic congestion**Graphical user interface, text, application

Description automatically generated

The above graph shows us the distance of the road affected for the accidents and also it relates the severity of accidents with the distance of road affected with the accident. This will help us know how much the severity of accident will affect the traffic congestion. We can see the for most of the accidents the traffic that is affected is withing the 10 miles range. This will help the authorities in redirecting the traffic and also the riders in taking an alternative route.

**Plot 10: Visibility with severity levels**

Graphical user interface, text, application

Description automatically generated

The above plot shows how the visibility in miles will affect the number of accidents. This will help the users figure out if they should take any additional precautions to stay safe during lower visibility times. We also have the control to check the severity of accidents under different visibility conditions.

**Section VI: Dashboard Interactivity**

1. Severity level control from the ‘Severity’ data attribute which contains a dimension value of the severity to interact with the following visualizations:

Plot 1 – Accidents in each state

Plot 2: Distribution of accidents across different days of the week

Plot 4: Impact of weather conditions on accidents

Plot 6: Accidents across different traffic signs

Plot 8: Twilight time and severity of accidents

Plot 9: Severity with traffic congestion

Plot 10: Visibility with severity levels

1. Year control from the ‘Start Time’ data attribute which contains only years to filter out the year-wise data in the following visualizations:

Plot 1 – Accidents in each state

Plot 7: Year wise severity trend of the accidents

1. State level control which allows us to filter out the state-wise data. This control can be seen in the following visualizations:

Plot 1 – Accidents in each state

Plot 2: Distribution of accidents across different days of the week

Plot 5: Time taken to clear an accidents

Plot 6: Accidents across different traffic signs

1. We have the time control which will help us figure out the accidents that have occurred during different times of the day.

Plot 2: Distribution of accidents across different days of the week

1. We have control to filter out the weather condition which can be used to check the severity and accident details in each weather condition. This control can be seen in the following visualizations:

Plot 4: Impact of weather conditions on accidents

**References**

**Tableau Public Link –**

[https://public.tableau.com/authoring/USAccidentsUpdated/USAccidentsPatterns#1](file:///Users/nagendrasrirangam/Downloads/Dashboard%20Link)

**Mural -** <https://app.mural.co/t/greeshmasworkspace0156/m/greeshmasworkspace0156/1680555917958/0960ed43596c4be3fb8a716bdd2332f73c5e5a3d?sender=ufd0807b59c2de6a557262362>

**Dataset source –** [https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents](file:///Users/nagendrasrirangam/Downloads/Dataset%20Link)

Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. *“A Countrywide Traffic Accident Dataset.”*, arXiv preprint arXiv:1906.05409 (2019).

Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. *“Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights.”* In proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2019