Abstract

This Report give in-depth analysis of dataset related to accidents and also provide some insightful observation.

Accident Analysis Report

ISYS 623

Given By:-Professor Hartzel

Date:-02-14-2023

Rishav Shrivastava

Shrivastavar@duq.edu

**A Deep Dive into the Causes of Road Accidents/Crashes**

**Introduction:**

This report or project main aim is to analyze the reason behind the Accidents in Pennsylvania. We analyzed the crash dataset, road dataset and flag dataset to find the possible reasons behind the accidents.

**Goal:**

The goal of this project is to conduct an in-depth analysis of road accident data to identify the main factors contributing to road accidents.

**Audience:**

This project is relevant for road safety organizations, road constructor , and also politician who are policy makers. There is some observation which is really interesting. The results of this project can be used to develop effective strategies to reduce road accidents and make the roads safer for everyone.

**Method:**

In this project, multiple data sets were analyzed and used to gather insights into the causes of accidents. When I saw the dataset, the first thing that came to my mind was to find the trend of accidents over the years. So, I merged different datasets of COMMVEH, ROADWAYS, and FLAG from different years(2010-2021) using the Pandas library. The data was in about 11 JSON files. SO, I combined these multiple JSON files into a single CSV file to streamline the data handling process. Then, I started the cleaning process of the data and merged different subsets of certain datasets according to my needs. I also removed columns that were not useful for a better understanding of the data. Finally, the data was loaded into Tableau for visualization and analysis.

1. The first visualization shows the number of accidents per year and reveals that the year with the most accidents was 2016. It also shows a sudden decrease in accidents/crashes in 2020, which may be due to the COVID-19 pandemic.

Chart, treemap chart

Description automatically generated

Figure 1

1. The second visualization reveals that most accidents occur on roads with a 25-mph speed limit. This is intriguing and could be due to driver overconfidence or a lack of good road structures or design.

Chart, line chart

Description automatically generated

Figure 2

1. The third visualization clearly shows that state highways have the most recorded accidents.

Chart, bar chart

Description automatically generated

Figure 3

1. To understand the reasons for such high accident rates on state highways, a bar graph comparing the average speed limit to road owners was plotted.

Chart, bar chart

Description automatically generated

Figure 4

1. This visualization was used to identify potential reasons for the accidents on state highway. It shows that state roads have the most icy roads, which may be one of the contributing factor.

Chart, bubble chart

Description automatically generated

Figure 5

1. This visualization I used to investigate other reasons for accidents and found that most accidents have no recorded reason, or the reason listed may not be the actual cause.

Chart, bubble chart

Description automatically generated

Figure 6

1. Lastly, to understand the most contributing factor to road accidents, I investigated whether it was alcohol or drinking and driving or Driver negligence. To my surprise, most accidents occur not because of alcohol but due to road conditions.

Graphical user interface, text, application, Word

Description automatically generated

Figure 7

**Key Findings:**

1. I compared the number of accidents (CRN) with the year in which they occurred and found that there was a sudden decrease in accidents in 2020.
2. I then analyzed the ROADWAY dataset and compared the CRN with the average speed limit. I discovered that most accidents happen when the speed limit is 25 mph.
3. Further analysis of the ROADWAY dataset showed that most accidents occurred on state highways.
4. I combined the ROAD and CRASH datasets and found that state highways are more likely to have ice on the road, which could explain why they have a higher rate of accidents.
5. I also analyzed the Flag dataset and found that most accidents happened for no apparent reason, i.e., no alcohol, distraction, hit, or negligence was involved.
6. And finally, I also found out that it is not the alcohol that is the root cause of the accident, it is the bad structure of the roads or No traffic rules that causing the most accidents

**Recommendations:**

1. State authorities should prioritize road maintenance and improvement on State Highways to reduce the amount of ice on the road during the winter months.
2. Further investigation is needed to determine the underlying cause of accidents that have no specific reason and also need to impose some strict rule to keep eye on Driver’s age.

**Conclusion:**

Finally, the analysis of different dataset of different year provides valuable insights into the causes of state highway accidents. We were also found a pattern and trends in the data by collecting, cleaning, and visualizing it. Key findings as well as Observation can be further studied and can be used to improve the condition in future.