

Road Safety in Numbers: Decoding Road Accident Patterns in Massachusetts

Meghaana Tummapudi, Vikramaditya Pabba, Harish Reddy Vuyyuru

IE6600 – Computation and Visualization

December 10th, 2023

I. Introduction

The arteries of commerce and daily life, our roads, and highways, are not just mere conduits for traffic but lifelines that connect the heartbeats of cities and towns. As vehicles traverse these paths, they carry more than passengers and goods; they bear the pulse of economic vitality and human interaction. However, this flow is not without its perils. Road accidents emerge as unintended by-products of this dynamic movement, reflecting a confluence of human, vehicular, and environmental factors. The data, meticulously recorded and made accessible by the Massachusetts government, encompasses a wide array of variables, from the severity of accidents to environmental conditions, capturing a detailed picture of the multifaceted nature of road incidents. Our motivation here is twofold: Leveraging this rich dataset, we crafted interactive dashboards that not only reflect historical patterns but also shine a light on potential areas for intervention and improvement. By examining the patterns and trends of road incidents, we endeavor to provide insights that can help steer the state toward safer roads. This analysis is not just about numbers; it's a proactive step towards safeguarding the community's well-being.

II. Problem Statement

This project seeks to extract insights from the Massachusetts Road Crash Dataset to inform road safety measures. Our aim is to distill complex data into actionable patterns for policy development and prevention strategies. We'll investigate the timing and causes of accidents in Massachusetts, striving to improve our understanding of accident trends for long-term community safety.

Hypothesis and Research questions we seek answers for:

- Regarding Temporal Patterns: Is there a significant increase in road accidents during certain months of the year or on specific days of the week, suggesting a temporal pattern that could be linked to factors such as weather conditions, holiday-related traffic, or seasonal behavioral changes?
- Regarding Age and Gender: Do age and gender play a decisive role in the likelihood of being involved in a road accident, indicating that specific demographic groups might require targeted road safety education and interventions?
- Regarding Weather Conditions: Are road accidents more prevalent during certain weather conditions, and if so, does this suggest a need for improved infrastructure or driver support systems during adverse weather?
- Regarding Road Surface Conditions: Does the condition of the road surface, such as wet or dry, have a statistically significant impact on the frequency of road accidents, warranting an examination of current road maintenance and drainage systems?
- Hypothesis Regarding Light Conditions: Are there discernible differences in the number of accidents that occur during daylight versus dark, suggesting that lighting conditions or driver behavior at night could be a factor in accident rates?

III. Data Description

Our dataset consists of 1.04 million rows and includes 40 columns, providing person-level crash details occurring between 2018 and 2023. This dataset is generated by extracting information from the Massachusetts government website's impact section, specifically from <https://massdot-impact-crashes-vhb.opendata.arcgis.com/>. To create a comprehensive dataset encompassing all years, we merged the individual CSV files containing crash details from each year (2018-2023)

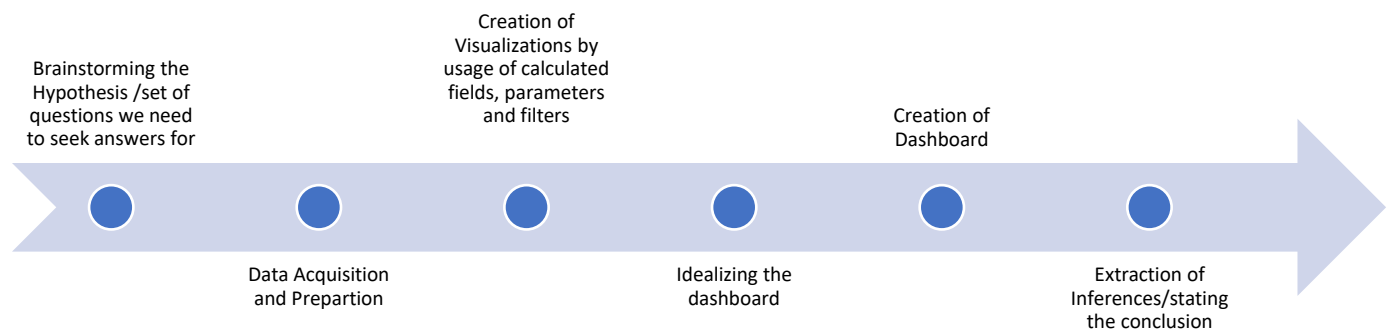


Unnamed: 0	Crash Number	City Town Name	Crash Date	Crash Severity	Crash Time	Crash Year	Max Injury Severity Reported	Number of Vehicles	Age of Driver - Youngest Known	...	Person Number	Age	Injury Type	Person Type
0	4476768	DRACUT	01/01/2018	Property damage only (none injured)	12:30 AM	2018.0	No injury	1.0	21-24	...	1.0	21.0	No injury	Driver
1	4477257	FRANKLIN	01/01/2018	Non-fatal injury	5:45 PM	2018.0	Non-fatal injury - Non-incapacitating	1.0	25-34	...	1.0	33.0	Non-fatal injury - Non-incapacitating	Driver
2	4477404	NORTH ATTLEBOROUGH	01/02/2018	Property damage only (none injured)	11:27 AM	2018.0	No injury	2.0	45-54	...	1.0	52.0	No injury	Driver
3	4477404	NORTH ATTLEBOROUGH	01/02/2018	Property damage only (none injured)	11:27 AM	2018.0	No injury	2.0	45-54	...	2.0	NaN	Not reported	Driver

This rich dataset allows for a deep and multifaceted analysis of road accidents, providing insights into patterns and trends that could inform and enhance road safety measures.

IV. Methodology

Steps followed to complete this project:



1. This data has been pre-processed and prepared in following steps:



Columns after cleaning –

```
df_cleaned.columns
Index(['Crash Number', 'City Town Name', 'Crash Date', 'Crash Severity',
      'Crash Time', 'Crash Year', 'Number of Vehicles', 'Light Conditions',
      'Road Surface Condition', 'Vehicle Configuration (All Vehicles)',
      'Weather Conditions', 'County Name', 'Hit and Run', 'Locality', 'X',
      'Y', 'Alcohol Suspected', 'Driver Age', 'Sex', 'Crash_severity'],
      dtype='object')
```

2. Creation of Visualizations:

- Bar Chart has been used to depict the frequency of accidents by year and by month to identify any obvious trends or patterns over time.
- Pie charts to represent the proportion of accidents by gender and Road surfaces.
- Heat maps to represent age groups and the accidents incurred by them.
- Geographic heat maps to illustrate the distribution and frequency of accidents by location, highlighting potential hotspots for accidents.
- Horizontal bar chart to illustrate the accidents by weather conditions.
- All crashes by vehicle types.

3. Use of Filters and Parameters to make the dashboard Interactive:

In our dashboard, we've integrated filters so users can filter accident data by severity, location, road conditions, weather, and lighting. This customization refines our analysis. Additionally, 'Current Year' and 'Previous Year' parameters enable year-on-year comparisons, aiding in identifying trends and assessing road safety performance.

The image shows a vertical sidebar of filters on the left and two year selection controls on the right. The filters are:

- Select Severity:** Dropdown menu with 'All' selected.
- City Town Name:** Dropdown menu with '(All)' selected.
- Road surface condition:** Dropdown menu with '(All)' selected.
- Weather conditions:** Dropdown menu with '(All)' selected.
- Light Conditions:** Dropdown menu with '(All)' selected.

On the right side:

- Current Year:** Dropdown menu with '2022' selected.
- Previous Year:** Dropdown menu with '2021' selected.

4. Idealizing the dashboard:



Final Dashboard Outputs:





V. Key Insights:

- There is a huge decrease in the accidents in the year of 2020 due to lockdown during Covid. When compared to previous years, there is a notable decrease in the current year (2023), which could be resonated with changes in traffic patterns or the implementation of road safety measures.
- High Number of accidents occurred in the months of August – November (this is due to changes in weather conditions such as fog during fall and occurrence of sleet).
- Friday afternoons (3PM – 6PM) had high crashes due to rush hours, while the fatal accidents occurred during late and midnights (11PM – 2AM).
- Though it might imply that Males incur high accidents than female, it would not convey any significant information – as number of male drivers could be higher than female drivers.
- The majority of accidents occur under clear weather conditions and on dry road surfaces, which might point to driver error or other factors being more significant than environmental conditions in causing accidents.
- However, there is still a substantial proportion of accidents happening during rain and on wet roads, reinforcing the need for caution and improved safety measures during adverse weather conditions.
- Passenger vehicles are involved in the majority of accidents, which is expected given their prevalence on the roads.
- More accidents occurred during Day light rather than in night which could imply that visibility factor alone is not a predominant variable in accident occurrences; other factors like the volume of traffic during daylight hours might be more influential.
- More crashes occurred during in the coastal cities. This could be used to identify accident hotspots and target safety improvements.

VI. Conclusion

The analysis of road accident trends in Massachusetts points to specific areas for safety improvements. Seasonal increases in accidents suggest a need for proactive traffic management and enhanced safety protocols during high-risk periods. Targeted educational campaigns may reduce the disproportionate accident involvement among young adults and males. Moreover, the prevalence of accidents in clear conditions indicates a potential oversight in driver awareness, emphasizing the need for continuous defensive driving education. Strategic infrastructure upgrades in accident hotspots could further bolster statewide road safety efforts.