Analyzing Traffic Crash Data

Presented By:
Huaqing Shen - Rakshith Manandi - Willy Ma

Introduction

- Fatality Analysis Reporting System **(FARS)** Dataset
  - Provided by: National Highway Traffic Safety Administration
  - **Yearly** data on **fatal traffic accidents**
    - Age of persons involved
    - Month/day/time of accident
    - Blood alcohol concentration
    - Etc.
- Goal:
  - Visualize Data → insights on **key factors** in crashes
  - Enable data-driven decision making

Data Collection/Pre-Processing

- Raw Data via FTP site
  - Holistic data extraction
  - .zip files per year
    - contains .csv data files
- FARS Query Tool
  - Must specify crash variables
    - Ex: Age, time of crash, vehicle, etc.
  - Allows for cross-tabulation
    - Ex: # of people killed vs vehicle type

Data Collection/Pre-Processing

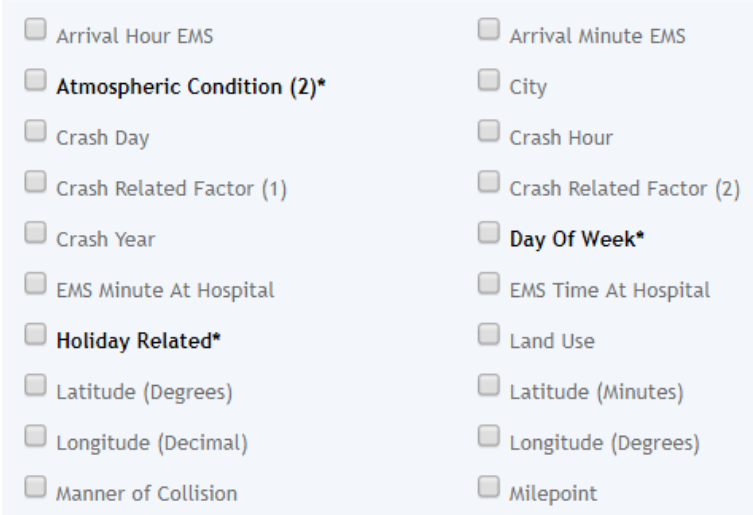
- Raw Data via FTP site

Index of /fars/

[parent directory]

Name	Size	Date Modified
1975/		12/14/11, 3:11:00 AM
1976/		12/14/11, 3:12:00 AM
1977/		12/14/11, 3:12:00 AM
1978/		12/14/11, 3:13:00 AM
1979/		10/9/09, 4:06:00 AM
1980/		10/13/09, 3:48:00 AM
1981/		10/13/09, 3:47:00 AM
1982/		10/7/10, 9:25:00 AM
1983/		10/7/10, 9:25:00 AM
1984/		10/7/10, 9:25:00 AM
1985/		10/7/10, 9:25:00 AM

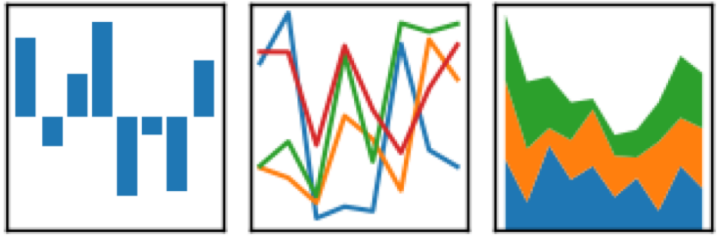
- FARS Query Tool



Tools

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



plotly

Bokeh

Diving Into Data Visualization

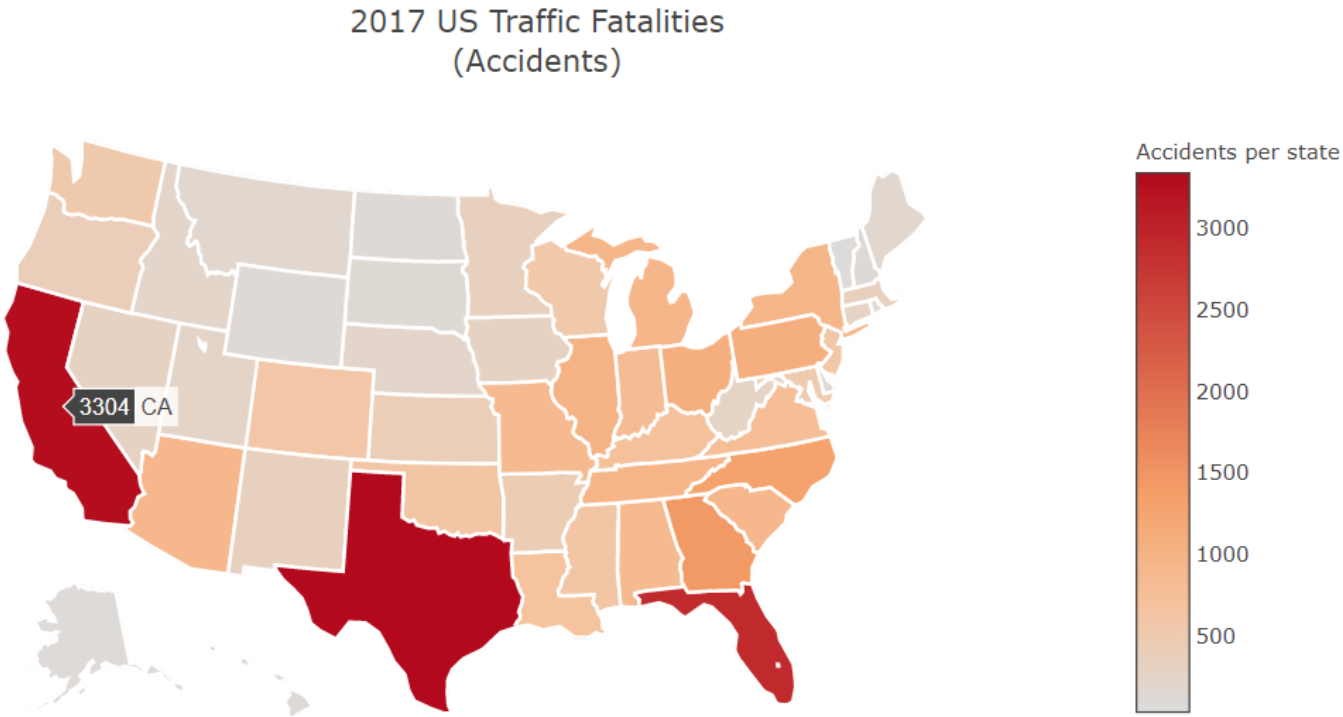


Figure 1. Total Accidents By State in 2017. Darker states like CA, TX, FL have a higher count

Diving Into Data Visualization

- Population size can skew data

US States - Ranked by Population 2018

Rank	State	2018 Population
1	**California**	39,776,830
2	**Texas**	28,704,330
3	**Florida**	21,312,211

Figure 2. The darkest states account for nearly 30% of US population[1]

[1]: US States – Ranked by Population 2018. (2017-11-09). Retrieved 2018-12-04, from http://worldpopulationreview.com/states/

Diving Into Data Visualization

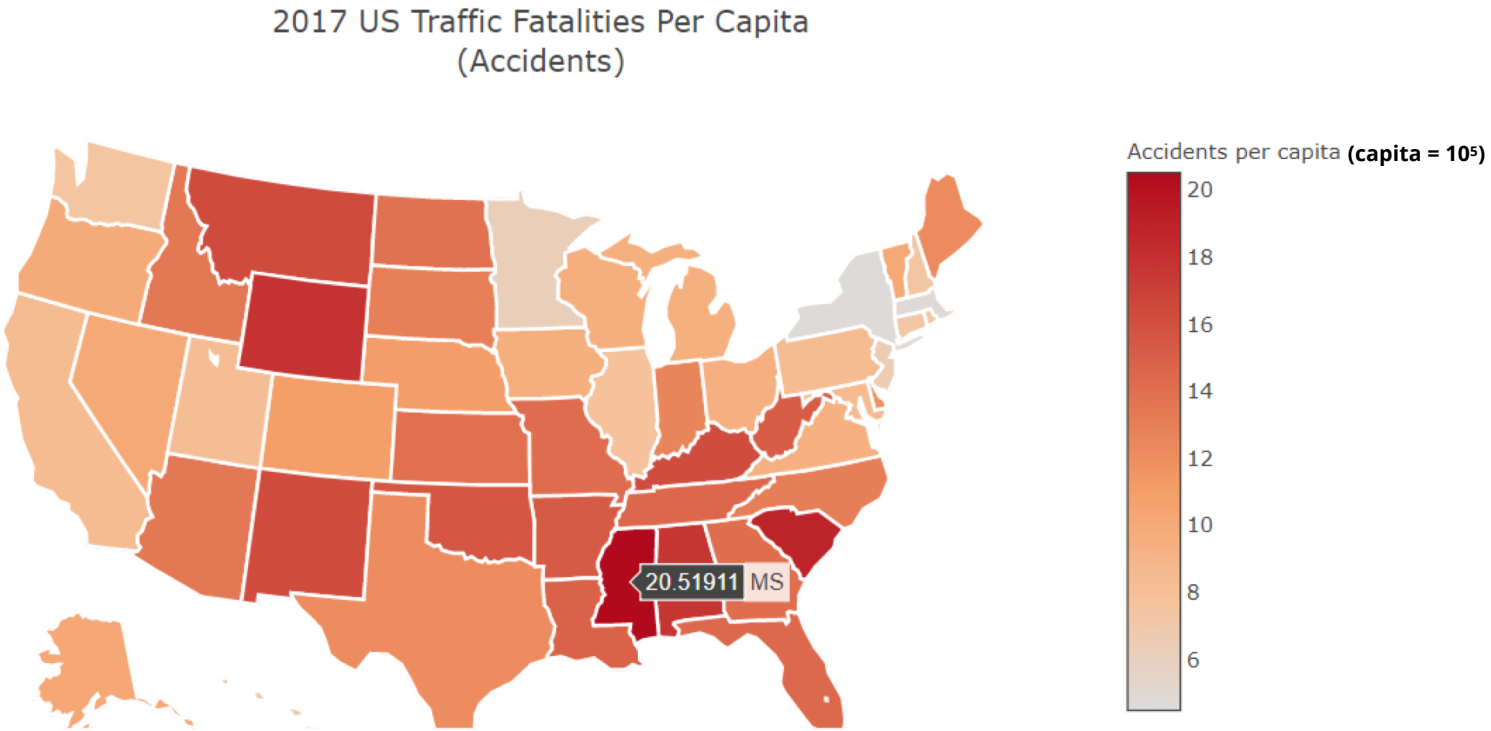


Figure 3. Total Accidents Per Capita By State in 2017. Darker states like MS, AL, WY, have a higher count

Diving Into Data Visualization

- Observations:
  - **Ranking changed when compared per capita**
    - Mississippi has the **most crashes**
    - Most populated states are not leading in crashes
    - Top states: MS, AL, WY, MT...
- Goal:
  - Understand and explain trend in data
  - Produce meaningful insight

Diving Into Data Visualization

- Findings from research[2]:
  - Mississippi is the 4th largest rural state
  - Rural roads are more unforgiving
    - Higher speed limits than city streets
    - Hospital locations are sparse

- Next steps in analysis:
  - Use insight gained from publication to analyze the proper variables

[2]: Campbell, L. (2018). ***Why Mississippi is the deadliest place to drive a car | Mississippi Today***. [online] Mississippi Today. Available at: https://mississippitoday.org/2018/07/03/why-mississippi-is-the-deadliest-place-to-drive-a-car/ [Accessed 5 Dec. 2018].

Diving Into Data Visualization

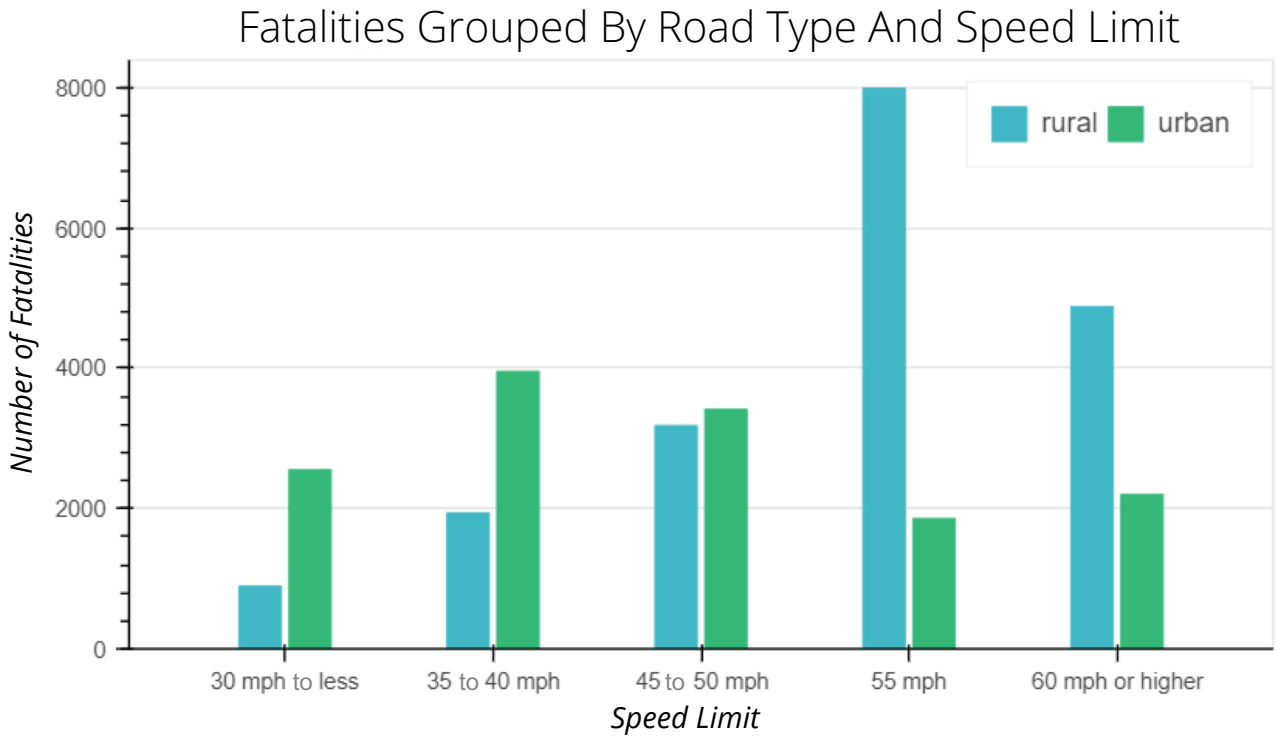


Figure 4. Plot visualizing fatalities, speed limit of accidents, and road type.

Diving Into Data Visualization

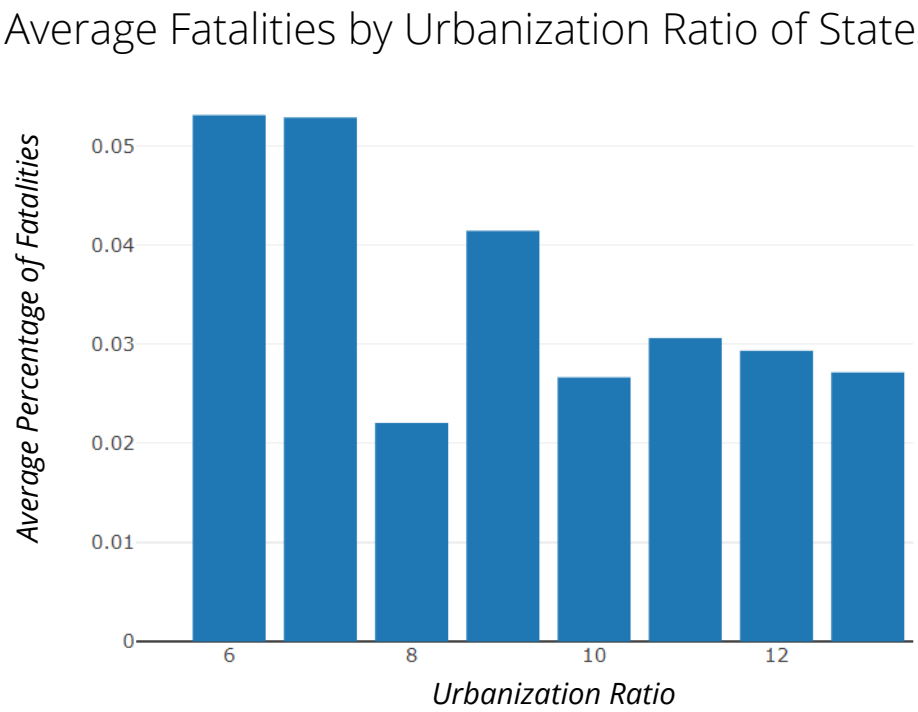


Figure 5. Plot relating state urbanization to fatalities.

Drawing Insight From Data Trends

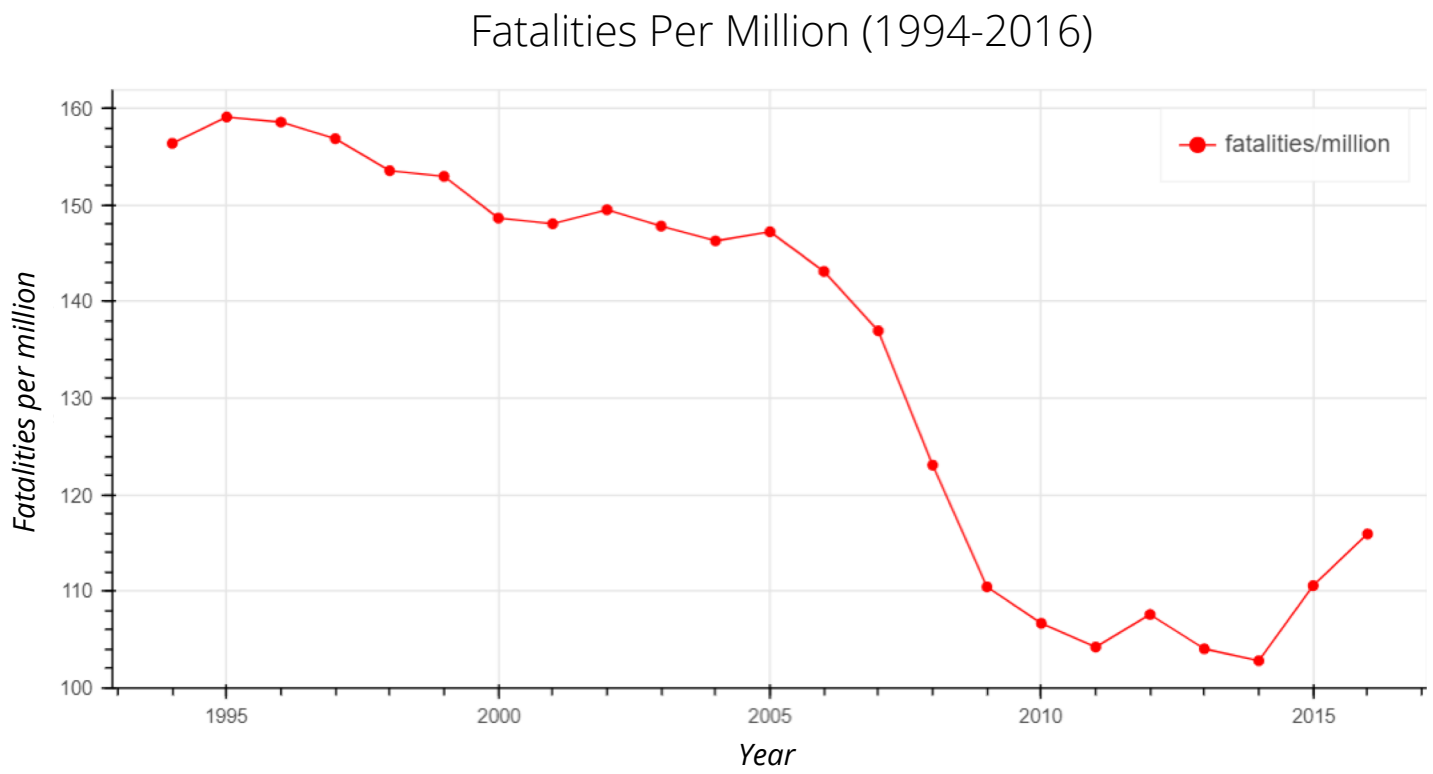


Figure 6. Plot showing downward trend of fatalities over time.

Drawing Insight From Data Trends

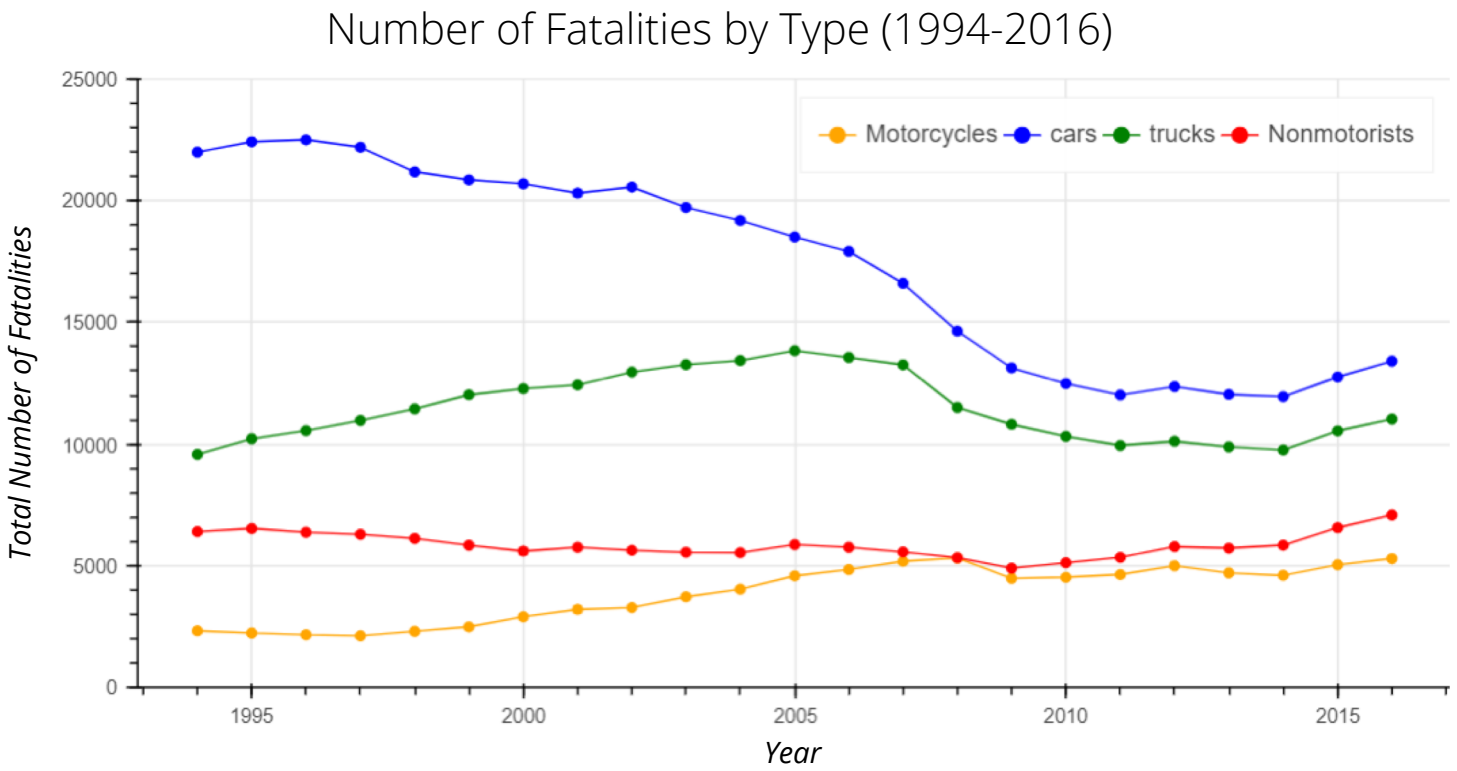


Figure 7. Plot visualizing the trend of fatalities over time by type

Drawing Insight From Data Trends

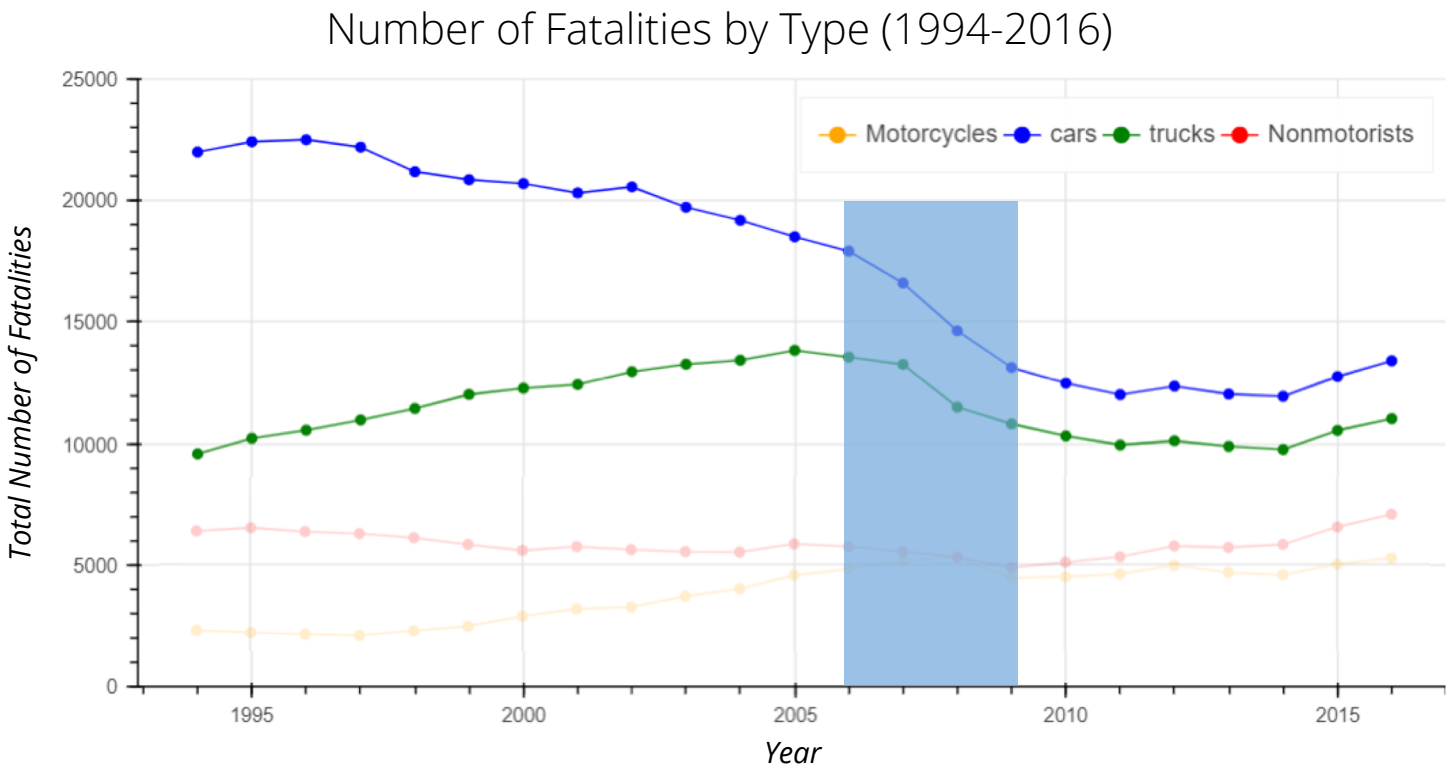


Figure 7. Plot visualizing the trend of fatalities over time by type

Drawing Insight From Data Trends

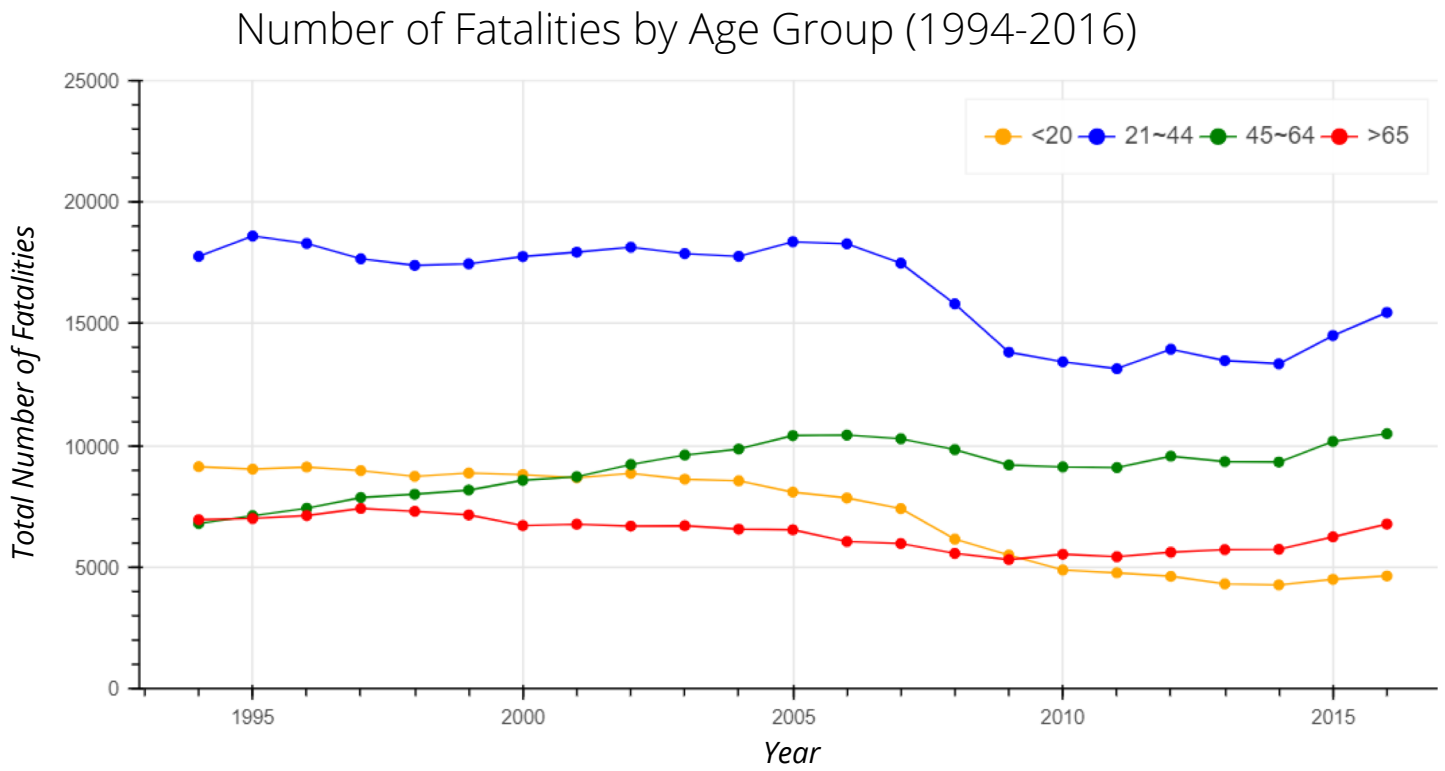


Figure 8. Plot showing trend of fatalities by age group.

Drawing Insight From Data Trends

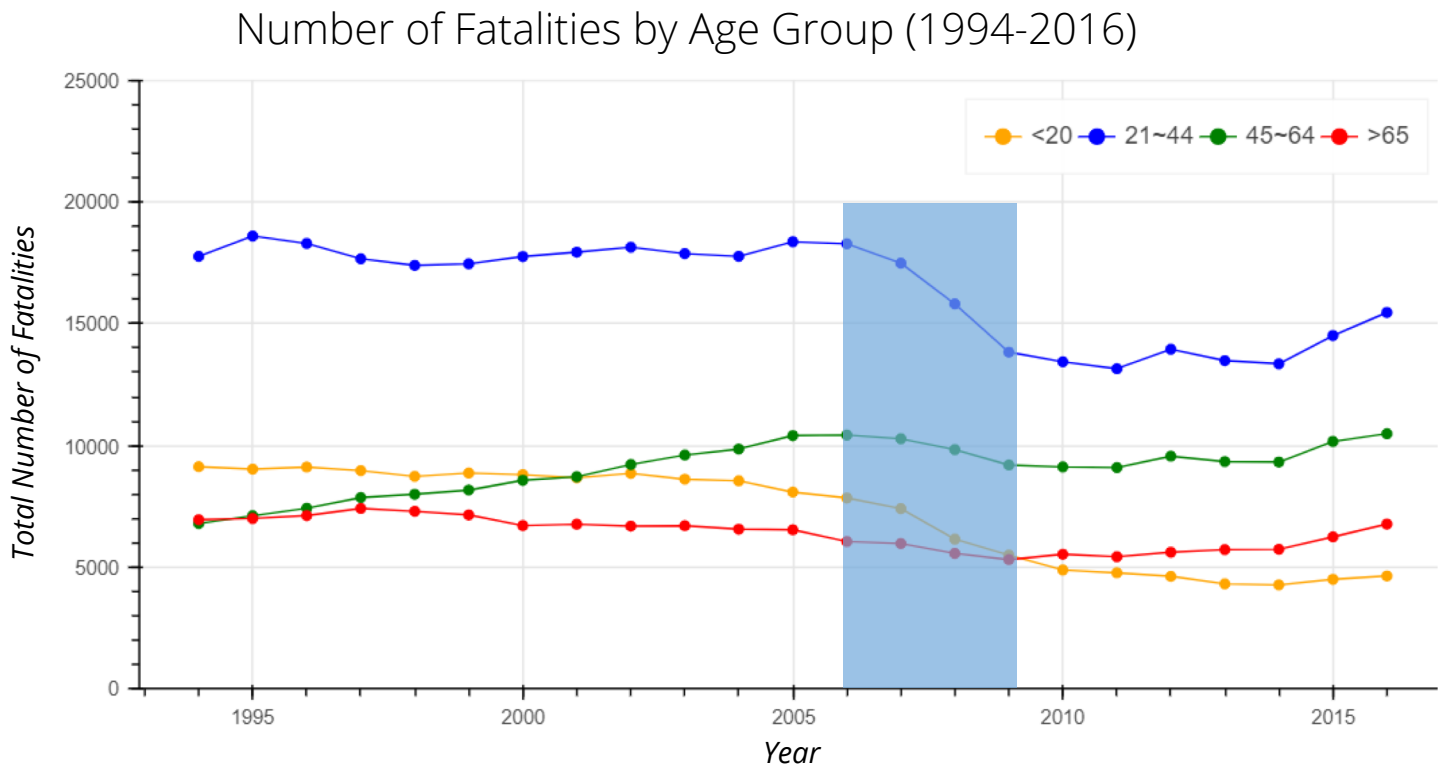


Figure 8. Plot showing trend of fatalities by age group.

Drawing Insight From Data Trends

- Large decrease in fatalities in 2007-2009
  - Why? Could be many factors
    - Better technology, education, rideshare, etc.
    - We researched other publications
- Interesting results[3]:
  - Advancements in vehicle safety technology
  - Trend common in recession (historically)
  - Emerging laws and regulations contribute to decrease

[3]: An analysis of the significant decline in motor vehicle traffic fatalities in 2008
McKay M.P., Thoma T., Kahn C., Gotschall C.S.
(2011) ***Annals of Emergency Medicine***

Exploring by Hypothesizing

- So far:
  - Started by visualizing data
    - Asking questions lead us to insight
- Next:
  - Start by asking questions
    - Visualizing data guides us to insight
- Hypothesis: Time correlates to crash frequency

Exploring by Hypothesizing

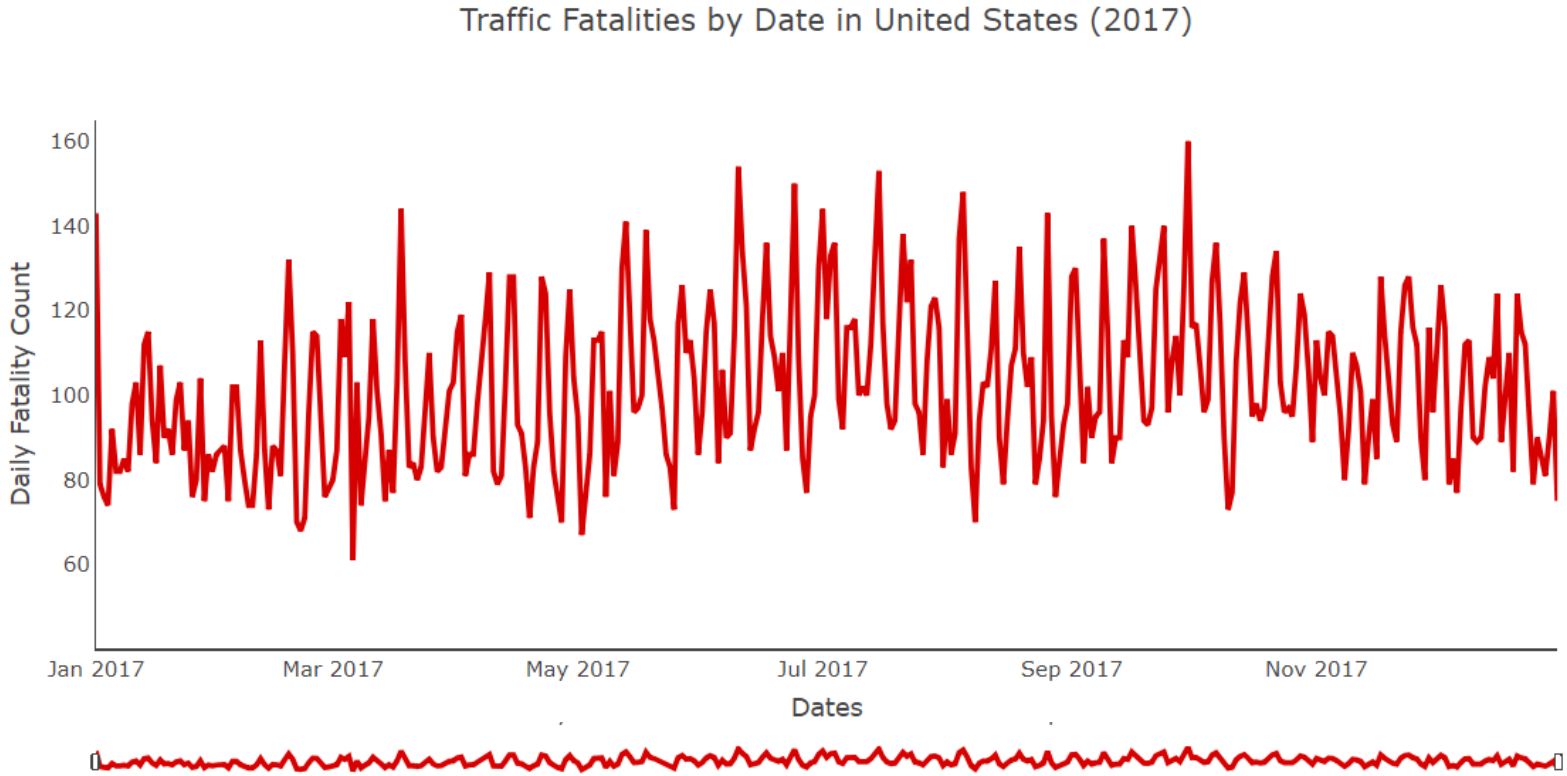


Figure 9. Interactive plot of daily fatalities over time created from plotly

Exploring by Hypothesizing

- Zooming in on the peaks

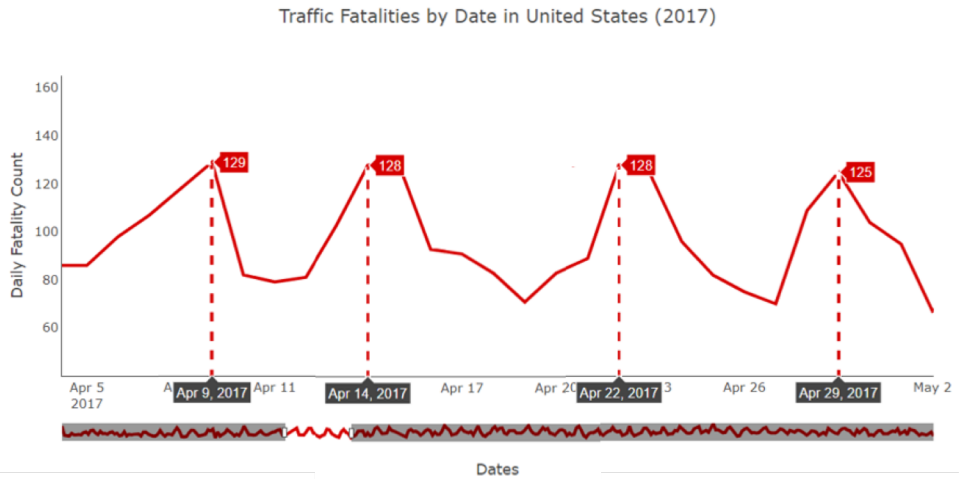


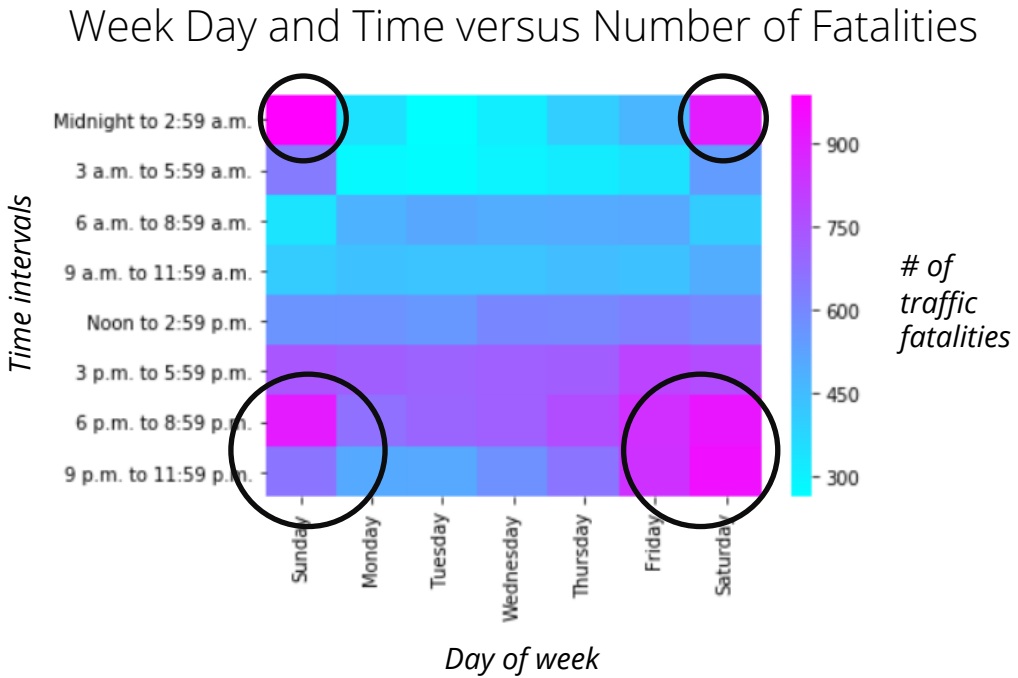
Figure 10. Zoomed-in plot created using the interactivity from plotly With cursors on peaks

Figure 11. Calendar with dates corresponding to peaks

- Correlation to day of week (i.e. weekends)

Exploring by Hypothesizing

- One level deeper
- Peaks during:
  - Weekends
  - 9 p.m. to Midnight
  - Midnight to 3 a.m.



- Last call?
  - Most bars close at 2 a.m.
  - Alcohol could be a key factor

Figure 12. Heatmap of time and day versus fatalities shows peaks during late night weekends

Exploring by Hypothesizing

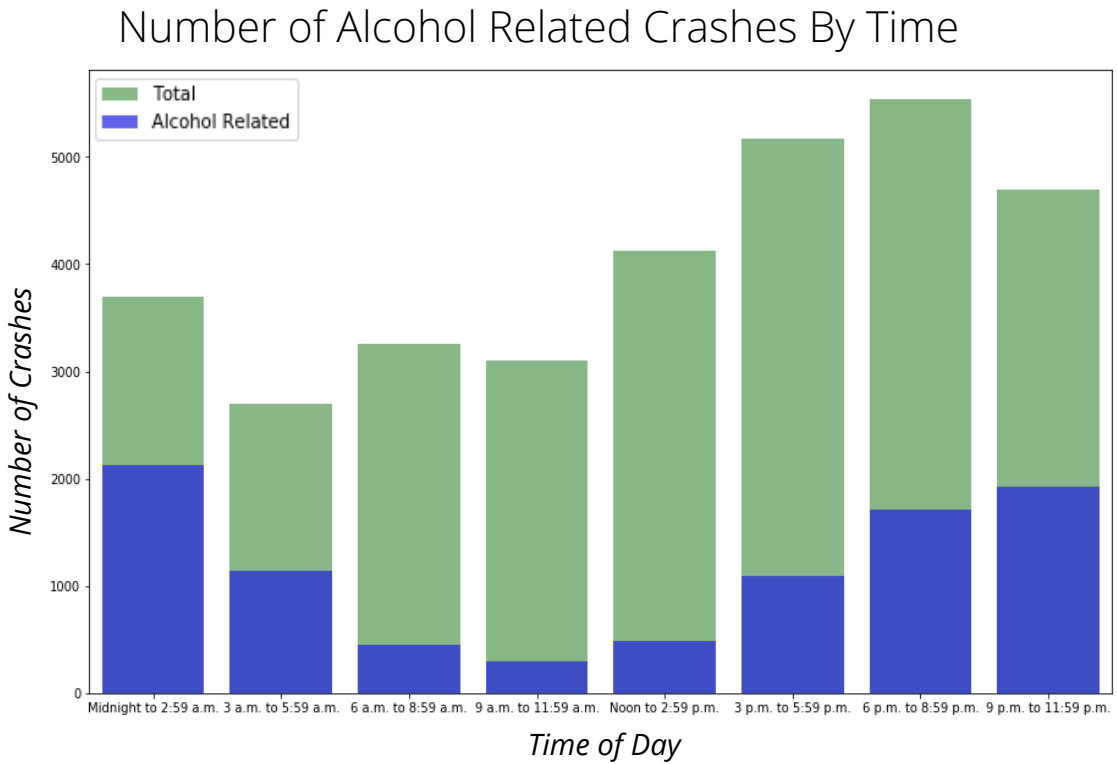


Figure 13. Plot of number of total and alcohol related crashes by time interval

Exploring by Hypothesizing

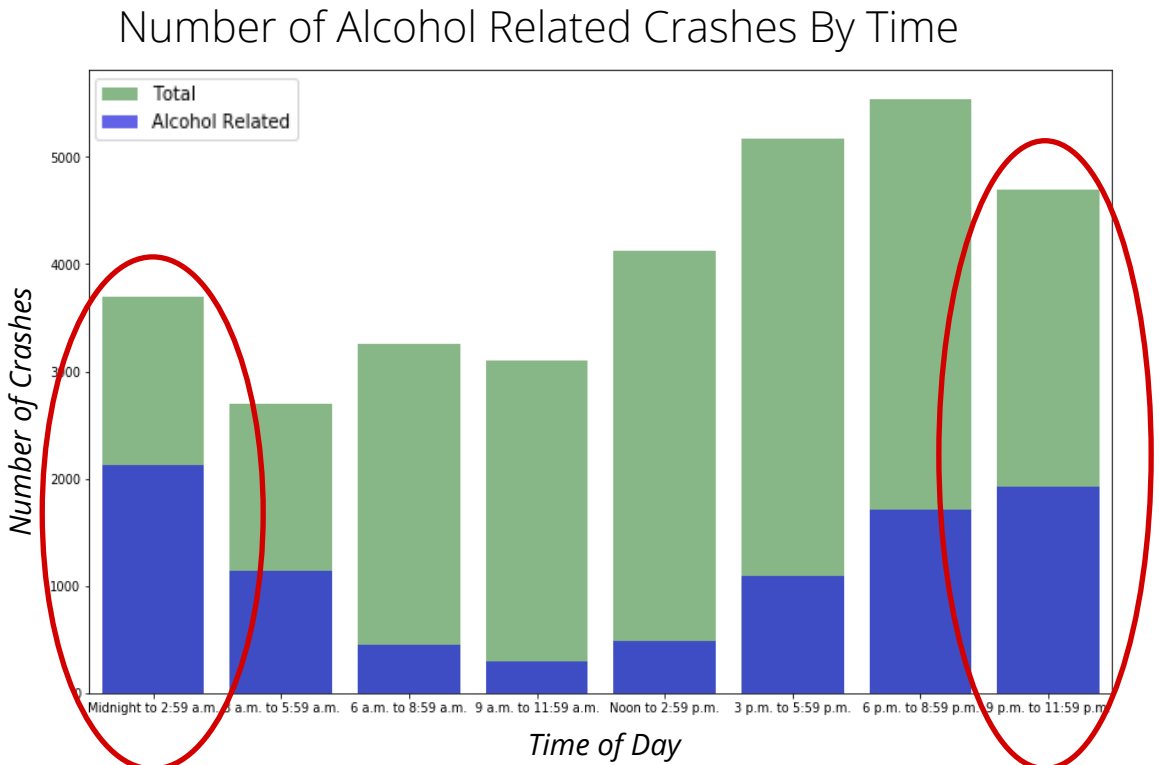


Figure 13. Plot of number of total and alcohol related crashes by time interval

Exploring by Hypothesizing

- Recap:
  - Hypothesis: “Time correlates with crashes”
    - Analyzed crashes over a year
    - Zoomed into intervals
      - Looked at weekends
      - Saw a trend with time
      - Hypothesized a relationship to bar closures
        - Lead us to analyzing alcohol related crashes as a key factor

Conclusion

- We were able to gather meaningful insight by
  1) Visualizing data
  2) Narrowing in on features to find key factors
     - Research related literature

- Key factors found:
  1) Ruralness of roads
  2) Economic state contributes as a factor
  3) Bar closure and alcohol is a key feature

Thanks For Listening

Sources

[1]: US States – Ranked by Population 2018. (2017-11-09). Retrieved 2018-12-04, from http://worldpopulationreview.com/states/

[2]: Campbell, L. (2018). *Why Mississippi is the deadliest place to drive a car | Mississippi Today*. [online] Mississippi Today. Available at: https://mississippitoday.org/2018/07/03/why-mississippi-is-the-deadliest-place-to-drive-a-car/ [Accessed 5 Dec. 2018].

[3]: An analysis of the significant decline in motor vehicle traffic fatalities in 2008 McKay M.P., Thoma T., Kahn C., Gotschall C.S.(2011) *Annals of Emergency Medicine*