Analysis of Vehicular Crashes in Iowa

Nathan Rethwisch
Data Science
Iowa State University
Ames, IA 50011
nreth@iastate.edu

Zachary Swayne
Statistics
Iowa State University
Ames, IA 50011
zswayne@iastate.edu

***Abstract*—According to the Center of Disease Control and Prevention (2020), traffic accidents are the top cause of death for US adults aged 1-54. Understanding what causes these accidents and how to prevent them is imperative to keeping vehicular operators and passengers safe. Our analysis uses public car crash data from the Iowa Department of Transportation. This data was collected from January 2009 until September 2022 and includes information about various accident-related statistics. We focus on investigating the link between animal-related car accidents and daylight, specifically sunrise and sunset time. We also investigate the impacts of driving under the influence of drugs and alcohol on the severity of vehicular accidents. Our analysis finds several links between sunrise/sunset times and crashes in Iowa. Furthermore, we were able to connect high rates of property damage and fatalities to the influence of drugs and alcohol. These findings can better inform drivers and lawmakers who seeming to improve road safety. We also present information about future research that may be done in the study area.**

***Index Terms*—data linkage; exploratory data analysis; data visualization; vehiculuar accidents; Iowa vehicular collisions**

I. Introduction

According to the Iowa Department of Transportation (2022), there are over 50,000 crashes annually in Iowa alone. These crashes cause millions of dollars in property damage and, unfortunately, the loss of life, with over 300 people dying in vehicular crashes each year. It is important to gain a greater understanding of the causes of these crashes to better create prevention strategies and protect the drivers on the road. By informing drivers of potential hazardous practices, they will be better prepared and encouraged to follow safe driving practices. The goal of this report is to explore some of the relationships between crashes and driving conditions to gain a better understanding of how to make the road a safer place.

The Iowa Department of Transportation collects data on every reported crash and makes this information publicly available at https://icat.iowadot.gov/. For each crash, a large number of variables is collected: date, time and location of the crash; the number of vehicles and passengers involved; crash severity measured in property damage, number of injuries, or fatalities; contributing factors such as weather and road conditions are reported, and whether any of the drivers were driving under the influence.

The Iowa Department of Transportation collects data on every reported crash and makes this information publicly available at https://icat.iowadot.gov/. For each crash, a large number of variables is collected: date, time and location of the crash, the number of vehicles and passengers involved; crash severity measured in property damage, number of injuries, or fatalities; contributing factors such as weather and road conditions are reported, and whether or not the driver was driving under the influence.

This report is based on crashes reported between Jan 1, 2009 and the end of September 2022. Figure 1 shows a map of all location of crashes in Iowa in 2021. We see that crashes happen everywhere, but there are certain highways and cities that contribute to higher numbers of crashes.

In our project, we are interested in two of the biggest factors contributing to crashes: alcohol/drug involvement and crashes with animals. In section II, we will present a short overview of the data and motivate our two lines of investigation. In section III, we explore the relationship between car crashes and daylight hours and demonstrate the underlying relationship to crashes with animals. In section IV, we close with a discussion of the effects of alcohol involvement on crashes.

II. First overview

Between Jan 1, 2009, and the end of Septembe 2022, a total of 728,442 crashes in Iowa were reported. Figure 2 shows the average number of crashes in Iowa by day of the year. We see that there are fewer crashes during spring and summer months (on average about 15% fewer) than during the rest of the year. Besides better weather during the summer months, more daylight hours provide better lighting conditions for drivers (see also section III C). From November to February, we see an increased variability in the number of crashes – this is likely due to bad weather days with a large number of crashes in some years. There is a clear holiday effect in the number of crashes: New Year's day, the Fourth of July, Thanksgiving, Christmas day and the day after have a much-reduced number of crashes. However, both New Year's day and the Fourth of July have a very large number of drunk crashes compared to the rest of the year (shown in orange). On New Year's Day, almost triple the number of drug-involved crashes happen compared to any regular day. We will further investigate the relationship between alcohol and crashes in section IV.

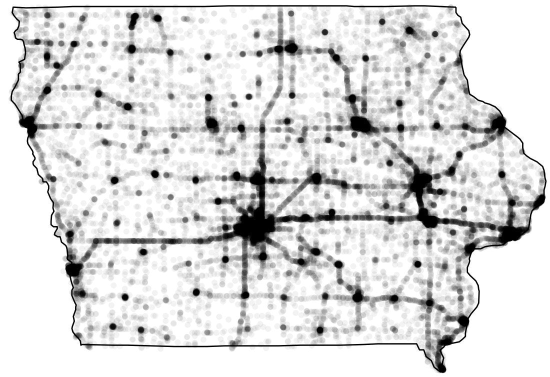


Fig. 1. Locations of all crashes in Iowa in 2021. Each crash is represented by a dot with an opacity of only 5% to provide an x-ray of high-density regions in the spatial distribution. Only when twenty crashes happen in the same location, do we see a fully saturated black dot on the map. The map paints a picture of Iowa's population centers and its network of roads.

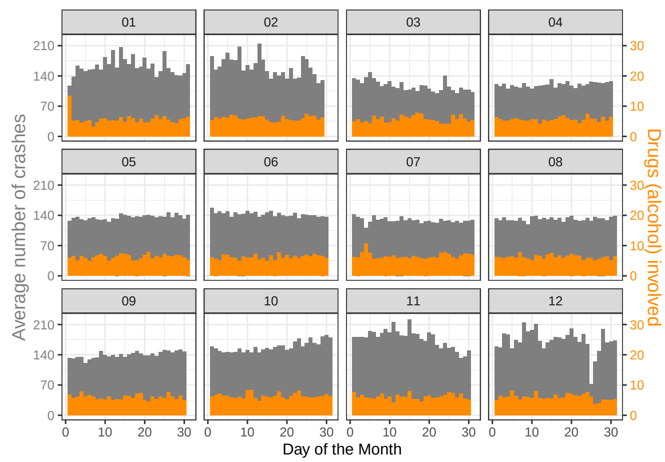


Fig. 2. Average number of crashes by day of the year.

Figure 3 shows bar charts of the average number of car crashes by hour over the course of a day throughout the months of the year. Initially, we see two prominent peaks in the number of crashes around the times of morning and evening commutes. However, in late spring, a third peak appears at a later time that only disappears by September and merges into a single huge peak with the evening commuting time in November. Looking at crashes involving animals, we see that this third peak strongly correlates with the time that animals are hit in the evening. We look at this phenomenon more

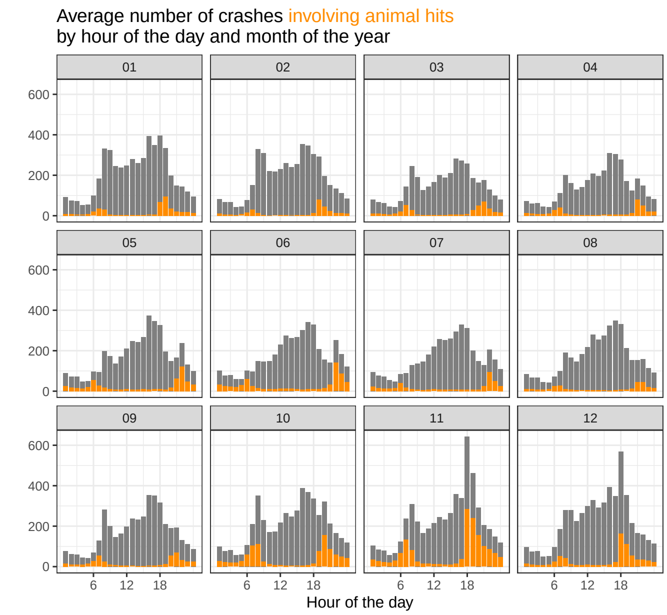


Fig. 3. Average number of crashes by hour of the day and month of the year. Crashes involving animal hits are shown in orange.

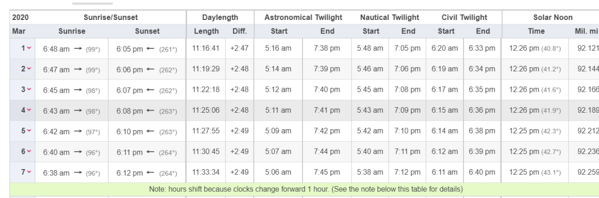
closely in section III.

III. Sunrise/Sunset Analysis

The initial data overview uncovered a potential connection between animal-related crashes and the dawn and dusk hours. In this section, we dig deeper into this idea to investigate the potential causes of the two daily spikes in animal-related crashes.

A. Data Extraction

While the time of day controls human behavior, sunrise and sunset regulate the daily routines of animals. Times of sunrise and sunset depend on one's location and the time of the year. Light conditions, such as dusk and dawn, are included as part of the accident report, however, their change over time made it hard for us to come to a single conclusion. Therefore, we thought it more appropriate for an analysis to transform the time of an accident with respect to when the sun is rising and setting. The accident report did not include sunset and sunrise times. As an alternative data source, we used the data portal by Time and AS (1995-). Figure 4 shows an example of the available data for the first week of March 2020 for Ames, IA. Rather than scraping all locations over time, we focused only on Ames. Ames is Iowa's major town closest to the state's geographical center. Making Ames the reference point limits the variation of sunrise and sunset times based on location. For approximately every 70 miles, there is a one-minute change in sunrise and sunset times. By using Ames, these discrepancies are limited to under three minutes. We chose to scrape data for 2020 because it is the most recent leap year and therefore provides data for February 29th (see appendix for details).



2020	Sunrise/Sunset		Daylength		Astronomical Twilight		Nautical Twilight		Civil Twilight		Solar Noon	
Mar	Sunrise	Sunset	Length	Diff.	Start	End	Start	End	Start	End	Time	Mil. mi
1	6:48 am → (99°)	6:05 pm ← (261°)	11:16:41	+2:47	5:16 am	7:38 pm	5:48 am	7:05 pm	6:20 am	6:33 pm	12:26 pm (40.8°)	92.121
2	6:47 am → (99°)	6:06 pm ← (262°)	11:19:29	+2:48	5:14 am	7:39 pm	5:46 am	7:06 pm	6:19 am	6:34 pm	12:26 pm (41.2°)	92.144
3	6:45 am → (98°)	6:07 pm ← (262°)	11:22:18	+2:48	5:12 am	7:40 pm	5:45 am	7:08 pm	6:17 am	6:35 pm	12:26 pm (41.6°)	92.166
4	6:43 am → (98°)	6:08 pm ← (263°)	11:25:06	+2:48	5:11 am	7:41 pm	5:43 am	7:09 pm	6:15 am	6:36 pm	12:26 pm (41.9°)	92.189
5	6:42 am → (97°)	6:10 pm ← (263°)	11:27:55	+2:49	5:09 am	7:42 pm	5:42 am	7:10 pm	6:14 am	6:38 pm	12:25 pm (42.3°)	92.212
6	6:40 am → (96°)	6:11 pm ← (264°)	11:30:45	+2:49	5:07 am	7:44 pm	5:40 am	7:11 pm	6:12 am	6:39 pm	12:25 pm (42.7°)	92.236
7	6:38 am → (96°)	6:12 pm ← (264°)	11:33:34	+2:49	5:06 am	7:45 pm	5:38 am	7:12 pm	6:11 am	6:40 pm	12:25 pm (43.1°)	92.259
Note: hours shift because clocks change forward 1 hour. (See the note below this table for details)												

Fig. 4. Example of data on sunrise and sunset times provided by timeand-date.com.

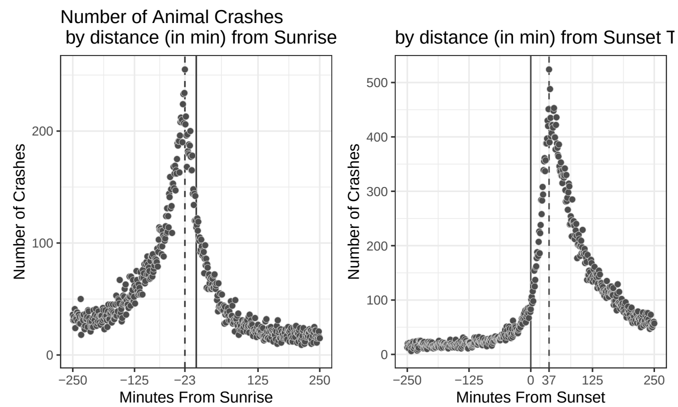


Fig. 5. A comparison of animal-related crashes with minutes away from sunrise (left) and minutes away from sunset (right)

B. Sunrise/Sunset Analysis

Sunrise and sunset times are incorporated into the crash records and allow us to calculate for each crash the distance in minutes from sunset or sunrise. Figure 5 gives an overview of the number of animal-related crashes in Iowa given time from sunrise (left) and time from sunset (right). The vertical lines denote the time of sunrise and sunset. The hashed lines are drawn through the times when animal hits spike. A gradual increase in crashes begins approximately 100 minutes before the sun rises. Once the sun is up, the number of animal-related crashes reduces drastically.

The plot on the right of Figure 5 shows the reverse phenomenon - crashes start to spike *after* the sun has set. According to the Council (2020), driving at dusk is extremely dangerous, as one's eyes take time to adjust to the relative darkness, shadows hide animals and road features, and drivers sometimes fail to turn on their headlights. This may be why there is such a strong correlation between sunset time and a spike in car crashes.

C. Daylight Analysis

Figure 6 shows a summary of the average number of crashes by day by the length of the day measured in daylight hours. A LOESS (Locally Weighted Scatterplot Smoothing) fit is added as a line to the scatterplot to summarize the relationship between length of day and the number of car crashes.

While there is more variability in earlier months (as seen before in Figure 2), there is a very clear downward trend, showing that the average number of crashes decreases with

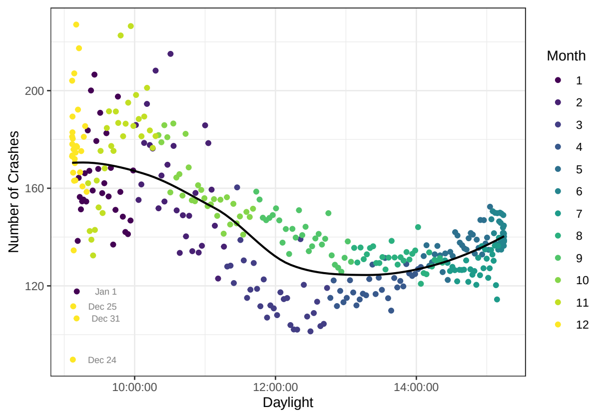


Fig. 6. Crashes by length of the day, with a line fit using LOESS smoothing

more daylight. However, for the long summer days, we see an increase in the number of crashes again – most likely due to the additional opportunities that longer days provide. The difference in crashes between March and September is interesting: while the amount of daylight hours is about the same (12 hours), there are, on average, fewer crashes in March than in September. This is most likely to increased animal activity during fall months, particularly concerning deer.

IV. Alcohol Analysis

For the next part of our analysis, we look into the relationship between alcohol/drug usage and car crashes. The initial overview showed that certain days of the year have significant spikes in alcohol-related accidents. Our exploration focuses on the difference alcohol use makes compared to driving sober.

A. Property Damage

Although crashes involving drunk drivers are less common, they cause about 45% or $3000 more damage on average. We also looked into the average number of fatalities with and without the use of alcohol. While sober, 1 in 127 crashes (0.79%) result in a fatality. However, when drugs or alcohol are present, fatalities happen in 1 of every 20 crashes (5%). The presence of drugs and alcohol leads to an unprecedented rise in the fatality rate.

B. Hour of the Day and Day of the Week

Figure 7 (left) shows the familiar pattern on average number of crashes by the hour of the day. This pattern changes considerably when one of the drivers is drunk. Most DUI crashes happen at night between 11 pm and 3 am.

Figure 8 shows the average number of crashes by day of the week (in grey) and compares that to the number of crashes under the influence of drugs (in orange).

Fridays are, on average, the day with the highest number of car crashes. One explanation for this trend is that more people are on the road due to weekend travel. However, DUI crashes occur at a much higher rate on Saturdays and Sundays. This is likely due to drinking on Friday and Saturday night leading to crashes early on Saturday and Sunday morning.

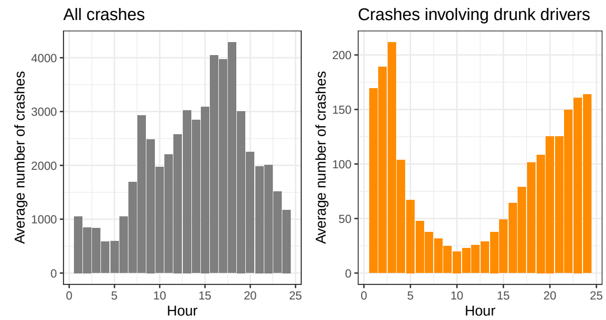


Fig. 7. Crashes by hour shown across all drivers (left) and inebriated drivers (right)

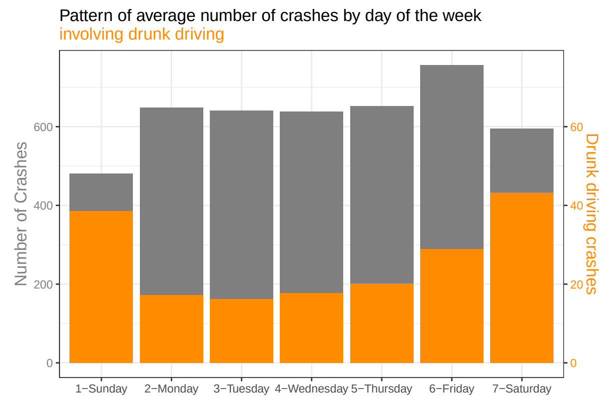


Fig. 8. Average number of crashes by day of the week (grey). The subset of drunk crashes is shown in orange. Note that drunk crashes are shown on a different scale to emphasize the weekly pattern.

V. Conclusion

Based on our data, we found sufficient evidence that dawn and dusk hours cause a significant increase in animal-related crashes due to increased animal activity, the adjustment of a driver's eyes, and failure to turn on headlights. We also found that alcohol and drugs significantly impair a driver's ability, drastically increasing fatalities and property damage. These crashes are concentrated at night and during the weekends.

From these findings, we recommend that drivers be further informed about the dangers of driving immediately after sunset. Furthermore, vehicle manufacturers can take precautions to ensure lights turn on at a certain darkness level. Our analysis also supports ongoing efforts to educate drivers about the dangers of intoxicated driving.

Some limitations of our analysis include the lack of exact positional data for all sunrise and sunset times. Although we reduced variability in our model as much as possible, we did not account for elevation changes, and differences in location can account for up to three minutes of error when analyzing sunset and sunrise times.

Regardless, these limitations provide minimal challenges to the results of our findings. We recommend an analysis of rush hour traffic and its effect on the prevalence of vehicular accidents. Overall, our findings may motivate drug and alcohol prevention programs and inform research into safety measures that can mitigate the adverse effects of dusk driving.

This report has been prepared in Rmarkdown. All results, figures, and code are publicly available from our GitHub repository at BLINDED.

Appendix

A. Data Cleaning/Processing

Data cleaning included changing empty strings and illogical values into NA values. Day, month, and year values were extracted using R's lubridate package. Latitude and longitude columns were also created from the 'Position' column to further explore where crashes happen in Iowa.

Information scraped from Time and AS (1995-) included hours of daylight, sunrise and sunset time, and date. Each month of the year was contained in a separate url which was a challenge of scraping the data. Furthermore, conditional cases of time changes presented additional challenges.

The main data set has a variable titled "Drug or Alcohol" with eight different levels. However, only two of these levels signify that substances were not involved. Because of this, we created a helper variable titled "Drug_Usage" that is TRUE when there are substances involved and FALSE when there is none present.

References

Center of Disease Control and Prevention. 2020. “Road Traffic Injuries and Deaths—a Global Problem.” https://www.cdc.gov/injury/features/global-road-safety/index.html.

Council, Delaware Safety. 2020. “Safe Driving Dusk and Dawn.” https://delawaresafety.org/resources/Documents/Safety%20Documents/Safe%20Driving%20-%20%20Dusk%20and%20Dawn.pdf.

Iowa Department of Transportation. 2022. “Iowa Motor Vehicle Crashes - 1925 to 2020.” https://iowadot.gov/mvd/stats/crashhistory.pdf.

Time, and Date AS. 1995-. “Ames, Iowa, USA - Sunrise, Sunset, and Daylength in 2020.” https://www.timeanddate.com/sun/usa/ames.