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 CDC (2020), the top cause of death for US adults aged 1-54 is traffic accidents. Understanding more about 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 a variety of 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. From our analysis, we find 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 serve to better inform drivers and lawmakers who seek to improve road safety. We also present information about future research that may be done in the study area.**

***Index Terms*—keyword 1; keyword 2**

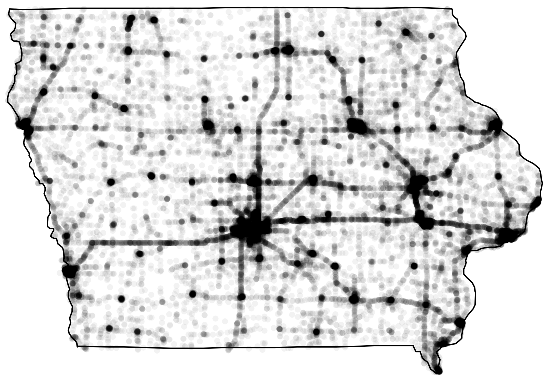

Fig. 1. Locations of all crashes in Iowa in 2021. Each crash is drawn with an opacity of only 5% to provide an x-ray of high-density regions in the spatial distribution. Only when twenty crashes happen in the same location, do we see a fully saturated black dot on the map. The map paints a picture of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

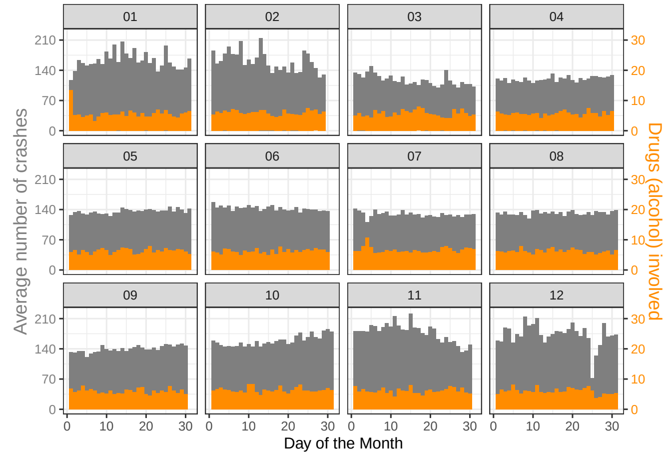

Fig. 2. Average number of crashes by day of the year.

I. Introduction

According to the Iowa Department of Transportation (2022), there are over 50,000 crashes per year 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 correlations between crashes and driving conditions to gain a better understanding of how to make the road a safer place.

The data for this project is collected by the Iowa Department of Transportation and made publicly available at https://icat.iowadot.gov/. Each row contains data for every recorded vehicle crash since January 2009 and is updated monthly. For each crash, a large number of variables are 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.

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 larger numbers of crashes.

In our project we are interested in the two of the biggest factors contributing to crashes: alcohol/drug involvement and crashes with animals. For that, we are looking at how

II. First overview

Between Jan 1 2009 and the end of September in 2022 a total of 728,442 crashes in Iowa were reported. Figure 2 shows the average number of crashes in Iowa by day of year. We see that during spring and summer months there are on average fewer crashes than during the rest of the year. From November to February, we see an increased variability in the number of crashes – this is likely due to bad weather days with large number of crashes in some of the 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 in comparison to the rest of the year (shown in orange). The fact that New Year's day has triple the number of drug-involved crashes as any other day is particularly scary because January first has the lowest amount of overall crashes in the month of January. We will further investigate the relationship between alcohol and crashes in section IV.

Figure 3 shows bar charts of the average number of car crashes 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 to a single huge peak with the evening commuting time in November. Looking at crashes involving animals we see that this third peak is strongly correlated with the time that animals are hit in the evening. We are going to look at this phenomenon more closer in section III.

III. Sunrise/Sunset Analysis

Our initial data analysis shows a potential correlation between animal-related crashes and the dawn and dusk hours. We wanted to delve deeper into this idea to see what exactly could be causing this spike in animal-related crashes.

A. Data Extraction

Because evening and morning times are variable throughout the year, we felt that it was appropriate to look at when the sun is rising and setting. Because this data was not included in the provided data from https://icat.iowadot.gov/, we scraped

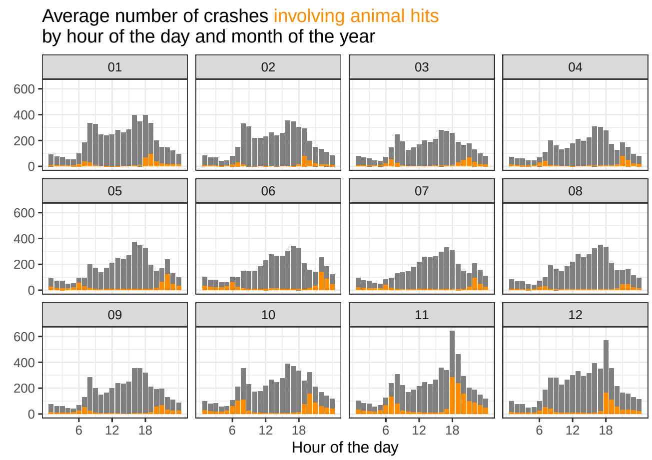

Fig. 3. Average number of crashes by hour of the day and month of the year. Crashes involving animal hits are shown in orange.

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 table of the data collected for sunrise and sunset times.

the data from Time and AS (1995--). The data that was sraped includes sunrise and sunset times for Ames, Iowa in 2020. The reasoning behind using Ames was that it is the nearest major town to the geographical center of Iowa. This would limit 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 were limited to under three minutes. The year 2020 was chosen because it is the most recent leap year and would provide data for February 29th.

Figure 4 shows an example table from Time and AS (1995--) that was used for the data collection.

B. Sunrise/Sunset Analysis

After joining the two data sets, we used sunrise and sunset times to calculate the number of minutes each crash was away from when the sun set or rose. We then compared this to the number of animal-related crashes in Iowa.

Figure 5 shows the minutes from sunrise compared to number of animal-related crashes, with a line denoting when the sunrise time is:

This clearly appears that there is a small spike, beginning approximately 100 minutes before the sun rises. However, after the sun rises, the number of animal-related crashes reduces drastically. One explanation for this is that an increased number of drivers are on the road during the rush hour which often falls just before sunrise. Due to low visibility which

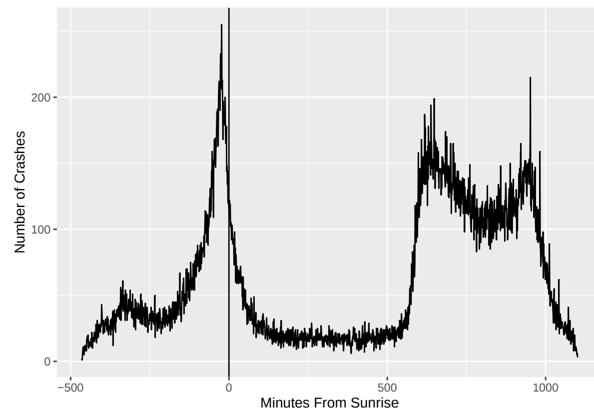

Fig. 5. All crashes shown by the difference between crash time and sunrise.

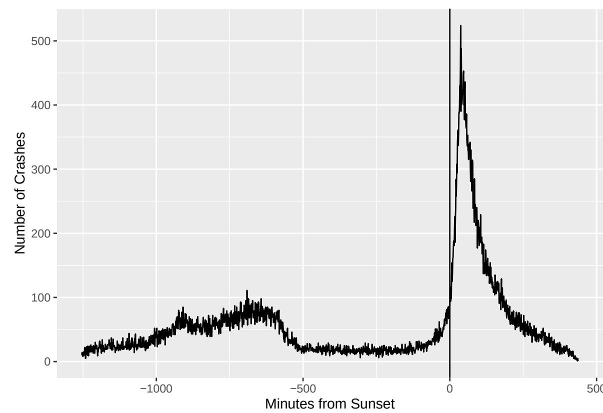

Fig. 6. All crashes shown by the difference between crash time and sunset

decreases reaction time, drivers may be more likely to cause a vehicular accident.

There also seems to be a noticeable spike approximately 750 minutes after sunrise time. This presumably corresponds to the sun setting, so we also decided to plot animal-related crashes in relation to when the sun sets. Figure 6 plots how many minutes away from sunset there are when a crash happens, with a line denoting where the sun sets.

6 shows that there is a large increase in animal-related crashes directly after sunset. 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 driver sometimes fail to turn on their headlights. This may be a reason why there is such a strong correlation between sunset time and a spike in car crashes.

C. Daylight Analysis

One point of note is looking into how the length of the day affects car crashes, regardless of whether or not drugs or alcohol are present. 7shows the number of crashes based on the length of the day. A line of best fit was added to the graph using LOESS (Locally Weighted Scatterplot Smoothing).

While there is more variability in earlier months, there is very clearly a downward trend, showing that there is a

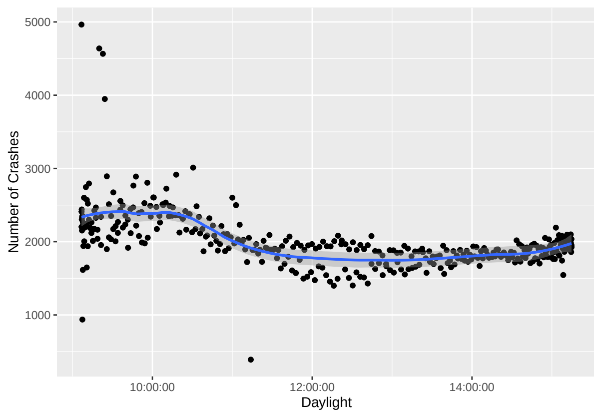

Fig. 7. A graph of crashes by length of the day, with a line fit using LOESS smoothing

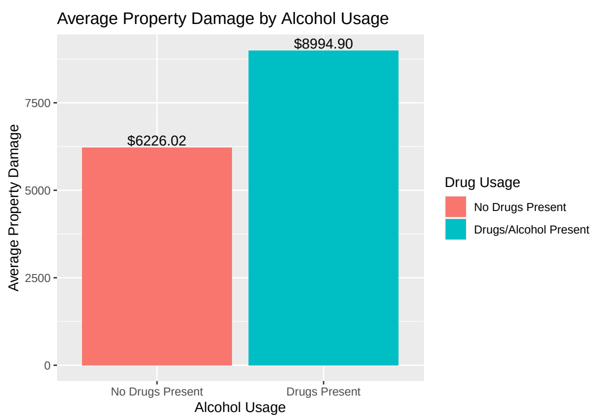

Fig. 8. A plot of the differnce between average property damage among sober drivers and drivers under the influence of drugs and alcohol

correlation to fewer daylight hours and more crashes. This is likely due to colder weather causing poor road conditions in the winter months, as well as an increase in drivers during dusk and dawn. As stated earlier, dusk and dawn prove to be some of the most dangerous times to drive.

IV. Alcohol Analysis

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

A. Property Damage

We found that although crashes while under the influence are less common, they cause almost $3000 more damage, on average. 8shows the difference in property damage between drunk and sober crashes.

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 fatality rate.

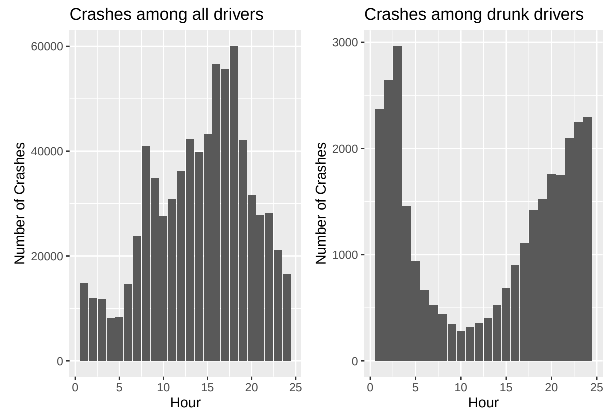

Fig. 9. Crashes by hour shown across all drivers(left) and inebriated drivers (right)

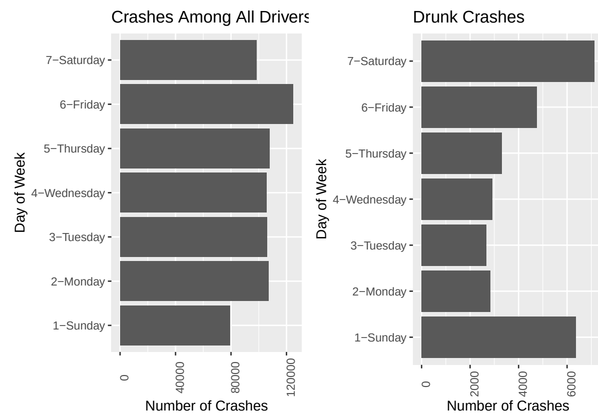

Fig. 10. Crashes among all days shown among all drivers (left) and drivers operating under the influence (right)

B. Time of Day

Next we looked at when crashes happened among drivers. Similar to our analysis in section II, we found that there were spikes in crashes at sunrise and sunset hours. However, as shown in 9crashes are inverted among drunk drivers, with most crashes happening between 11pm and 3 am.

Next, we decided to look at drinking by day of the week. We expected to see a large spike in drunk crashes on the weekends, especially Saturday. 10 shows a comparison between crashes among all drivers and those under the influence of drugs.

Among all drivers, Friday is the most common day for car crashes. One explanation for this is that more people are on the road due to weekend travel. Drunk drivers, however, tend to crash more often on Sundays and Saturdays. This is likely due to drinking on Friday and Saturday night which leads to crashes early on Saturday and Sunday morning.

V. Conclusion

Based on our data, we found sufficient evidence that the dawn and dusk hours cause a significant increase in animal-related crashes due to increased animal activity, the adjustment of driver's eyes, and failure to adequately turn on headlights. We also found that alcohol and drugs significantly impair a driver's ability, leading to a drastic increase in 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 being inebriated while operating a motor vehicle. Those efforts should be continued as they are important to keeping the roads safe.

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 changes in elevation, and differences 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 would however, 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, as well as inform research into safety measures that can be applied to mitigate the adverse effects of dusk driving.

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

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.