Elijah Lopez

Car Collisions in Toronto

Introduction

My reason for choosing the topic of car collisions/accidents is because I just got my G2 and was wondering what the most dangerous time on the road is. That is why my thesis question is: What time do car accidents happen the most? My hypothesis is that I think car accidents happen at night because darkness leads to limited vision.

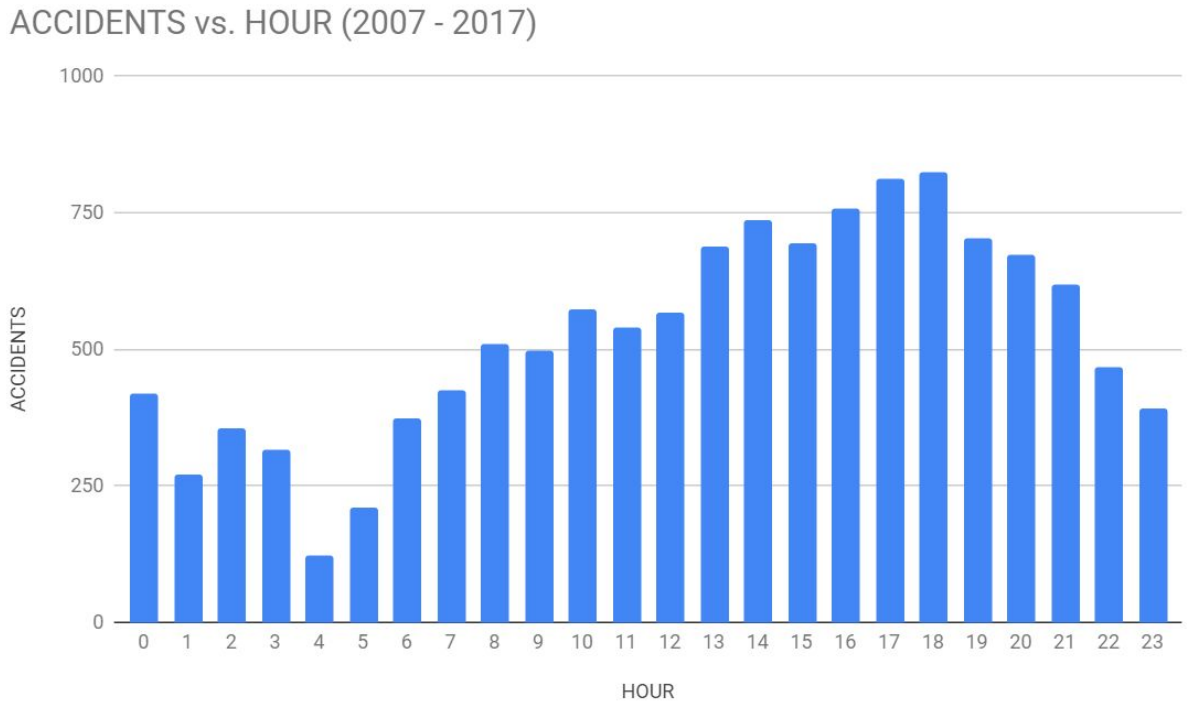
Background or History

This is an important question to investigate because of the prevalence of driving in the modern day. We drive everywhere so knowing what time is the safest or worst is very useful for when to pay even more attention to the road. The data I used includes all traffic collisions where a person was Killed or Seriously Injured.

Procedure

1. I acquired the secondary data from http://data.torontopolice.on.ca/datasets/ksi.
2. Programmed a script to make the data more useful (accidents/hour)
3. Analyzed the accidents/hour by aggregating the data from 2007 - 2017
4. Analyzed the accidents/hour also the total accidents per year

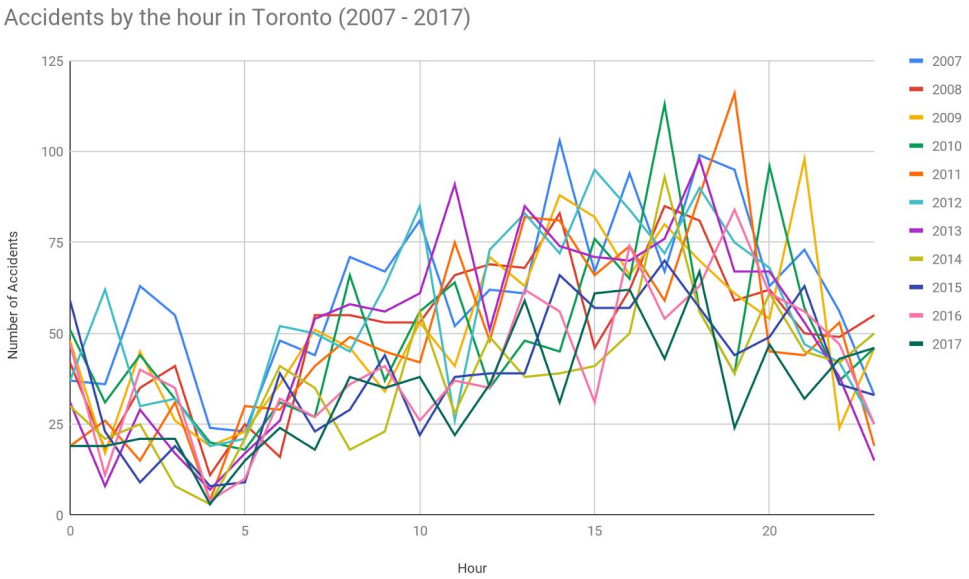
One-Variable Analysis



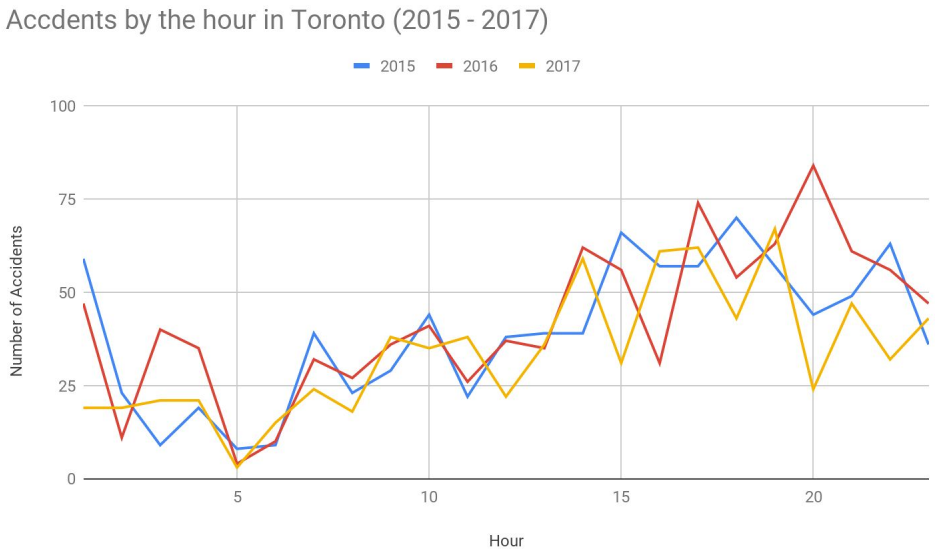
Mean	523.1666667	**Mean hour**	13.2 → 1 - 2	**Range**	703
Median	525.5	**Median hour**	14 → 2 - 3	**Q1**	388.25
Mode	N/A	**Mode hour**	18 → 6 - 7	**Q3**	689.25
Bottom Boundary	-63.25			**IQR**	301
Upper Boundary	1140.75			**Standard Dev**	192.8172522
Outliers	N/A			**Variance**	37178.49275

The histogram is skewed left; the mean (13.2) is less than the median (14) which is less than the mode (18). The safest time to drive in the last 10 years was from 4 am - 5 am, while the most dangerous time was from 6 pm - 7 pm. This closely aligns with my hypothesis of night time being the most dangerous time on the road because 6 pm - 7 pm is dark during the winter. My findings indicate that the most dangerous times are generally during the evenings from 4 pm - 7 pm. This is probably due to people coming back from work at those times into already congested roads rather than in the morning when people come onto less congested roads.

Two-Variable Analysis



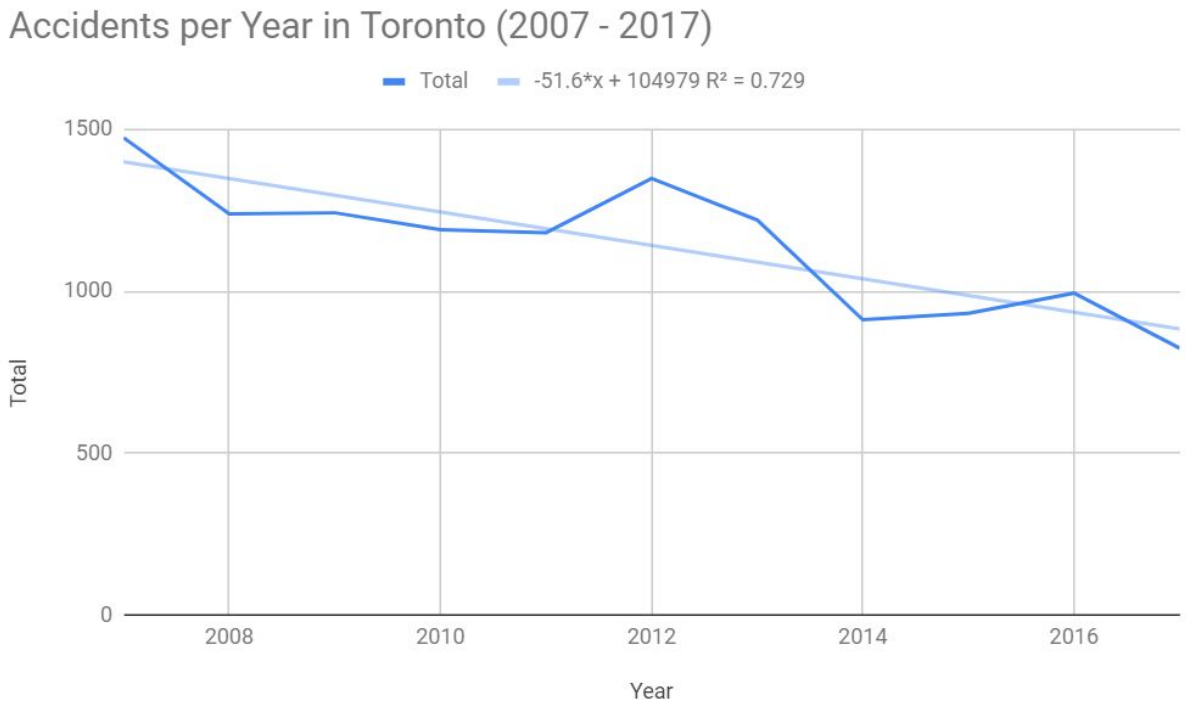
Elijah Lopez



Year	Peak	Hour	Total	Mode	17	1700 - 1800
2007	103	14	1474	Min	14	1400 - 1500
2008	85	17	1239	Max	21	2100 - 2200
2009	98	21	1242	Range	8	1400 - 2200
2010	113	17	1190	Standard Deviation	1.92	
2011	116	19	1181	Mean	17.45	1750
2012	95	15	1348	Median	17	1700 - 1800
2013	98	18	1220			
2014	93	17	912			
2015	70	17	932			
2016	84	19	994			
2017	67	18	824			

From these two charts and the table, we can see that the peak number of accidents varies every year but has a mode and median of 17 meaning that the common time interval for accidents from the last 10 years was from 5 pm - 6 pm. The Peak times in the last 10 years were usually in the evening from 2 pm to 10 pm, a large range but only because of 2007 and 2009.

Most of the years had peak time intervals from 3:00 pm to 7:00 pm. This is right on with what bactrack.com said: “The NHTSA (National Highway Traffic Safety Administration) reports that most accidents occur during "rush hour," between 3 p.m. and 6 p.m.” The peak hour goes up and then down by the years and sometimes stays the same, so a reasonable prediction for peak hours in 2018 is 6 pm to 8 pm.



This lovely line graph was made using the data in the table on the previous page.

From the table and also the graph right above, it is evident that the accidents per year are decreasing every year. This is a coincidental type of relationship because it's not the years that is affecting the number of accidents but another factor. This factor may be efforts from the Toronto police to reduce the number of bad drivers on the road (better technology, more arrests) or the city’s fines on speeding and breaking driving laws. Or it could mean that people are just getting better at driving which is unlikely because of the sheer amount of drivers in Toronto.

The equation of the trend line is y = -51.6x + 104979 and the coefficient of correlation is -0.854 indicating a strong negative linear relationship. My prediction for the number of accidents in 2018 using the equation is 832.

More Discussion

The data I used was collected by a government entity, Toronto police, and was also based on facts rather than from just people's opinions, making the data very resistant to any bias. The analysis that I did on the data though, is susceptible to being manipulated and presented in a manner that will appear as though this is the case for all cities in Ontario or Canada. That is simply out of bounds and would require data from a lot more cities than just Toronto and would also need to include all collisions since the data I used guarantees only collisions that involved a person killed or seriously injured. The analysis of the data is very useful for drivers commuting near Toronto because they will be able to change their driving schedule to keep themselves safer (and maybe even get to work faster).

Conclusion

In conclusion, my analysis of the Toronto car collision data was very insightful and answered my question: "what time do car accidents happen the most?" I was able to answer my question on a 10 year (2007 - 2017) span basis and also on a year by year basis giving out the peak hour interval for collisions. For 10 years, it was 6 pm to 7 pm and for 2017 it was also 6 pm to 7 pm. I additionally used the data to analyze the number of accidents per year. The results were eye opening because there was a strong negative linear correlation between the year and the number of accidents; the trend was a decrease in accidents every year. Further possible investigations include analyzing the location of these accidents, analyzing collision data from other cities in the same way, and analyzing a different field of this data such as the cause ("DRIVERACT" Column).

Bibliography

Otchy, C. (2015, June 30). The Most Dangerous Times on the Road. Retrieved from https://www.bactrack.com/blogs/expert-center/35042821-the-most-dangerous-times-on-the-road

(n.d.). Retrieved from http://data.torontopolice.on.ca/datasets/ksi