Aaron Templeton
U0734119
DS 2500
Final Project Report

The Effectiveness of School Zones: An analysis of car accidents in school zones in the US

Introduction

Are school zones really safe? Do school zones prevent injury or accidents? We all think that school zones are safe due to their design and use, however, that may not always be the case. Although all school zones have reduced speeds, traffic calming lights, crossing guards and safe pick-up or drop-off zones, there is an alarming number of accidents, injuries and even deaths that occur each year in school zones. In order to determine how safe school zones are, and where we can improve, I have performed a statistical data analysis on car accidents in the US and school zones.

The Reality

Upon my initial research of this matter, I discovered that school zones are not as safe as I initially had thought. In fact, data has shown that approximately 100 or more children are killed in pedestrian accidents each year within school zones. Also, approximately 25,000 children are injured each year in a car accident inside a school zone. Children only account for 1 out 5 school zone victims. The statistics gathered from Safe Kids Organization (Safe Kids) and the National Highway Traffic Safety Administration (NHTSA) are concerning.

School zone safety has been studied for years by organizations like Safe Kids and NHTSA, however, the data that they have gathered still shows an alarming number of accidents occuring inside school zones. Certainly too many for any parent to worry about. In order to understand more about the school zone accidents and possible safety concerns, I decided to perform my own analysis on car accidents to determine my own statistics.

The Data

The data that has been analyzed in this report is all publicly available. The report consists of two datasets, which were cleaned, transformed and evaluated.

US Accidents

Link to data: https://www.kaggle.com/sobhanmoosavi/us-accidents

This is a countrywide car accident dataset, which covers 49 states of the United States. The accident data are collected from February 2016 to December 2019. Currently, there are about 3.0 million accident records in this dataset.

USA Public Schools

Link to data: https://www.kaggle.com/carlosaguayo/usa-public-schools

This Public Schools feature dataset is composed of all Public elementary and secondary education facilities in the United States.

Working with the Data

The analysis performed on the data was done in a Google Colab Jupyter Notebook available at the following link:
https://colab.research.google.com/drive/1K9gx56AYRW-rLNxj9y15qQsDkjs1-TqH

I first read in the USA Public Schools dataset. It is a very complete dataset and I ended up only having to remove a couple rows from that dataset that was missing some important data. I also dropped a few columns from the table that were not necessary for my analysis purposes.

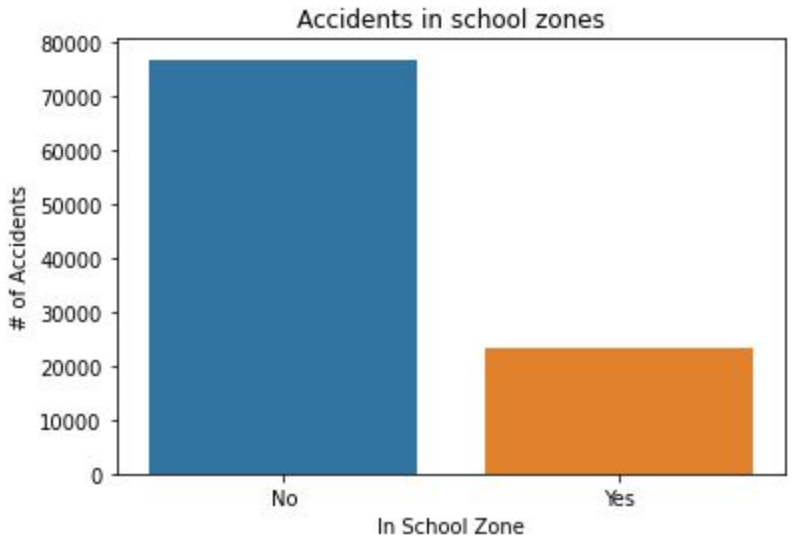
Working with the US Accidents dataset was a little bit harder. Very early on, I discovered that 3 million records of accidents was too much to work with. I ended up taking a random sample of 100,000 records from the dataset for my analysis.

After the dataset sample was ready, I had to do a few transformations in the dataset to fill in some missing values. I filled in zip codes and other empty values with consistent values in the dataset. I also dropped many columns in the table that were not necessary for the analysis I wanted to perform.

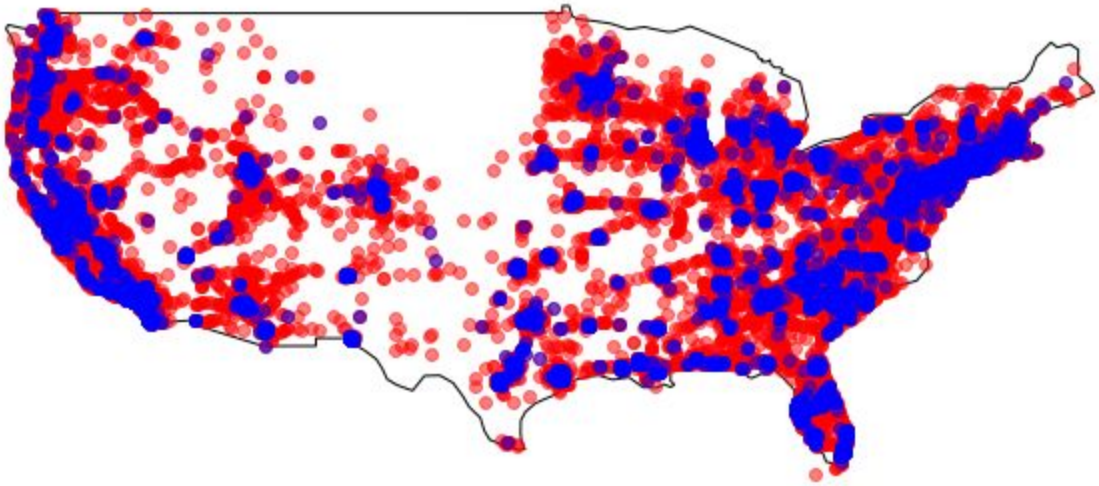
The most difficult part of the analysis was determining which accidents occured in school zones. In order to do this, I used the Geopy Python library which has a function to calculate the geodesic distance between two geographical points. The NHTSA suggests that the average size of a school zone is between ½ mile and 1 mile in radius. I used ½ mile to calculate the distances from the accidents and the schools.

Results

The results of the distance calculations on the accidents showed that out of 100,000 accidents in the US, 23,193 occuring in a school zone. We can infer that about 23% of all accidents within a school zone. Below is a chart representing this information.

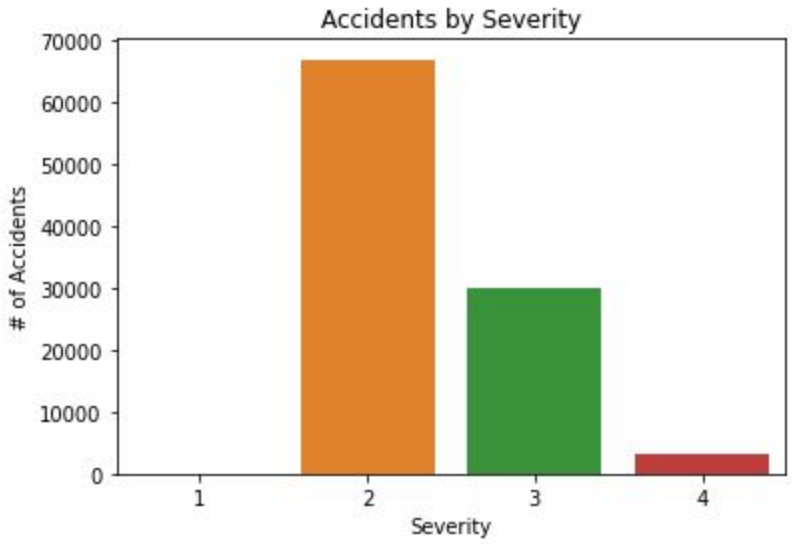


The map below shows the accidents in with a red point and the accidents in a school zone with a blue point. Interestingly enough, a lot of the accidents occur in and around big cities in the US.



Some other questions were also answered in our analysis. How do accidents in school zones differ in severity? Do other contributing factors affect the accidents in school zones (Crosswalks, roundabouts, Lights, etc.)? Does weather affect accidents in school zones?

In the US Accidents dataset, severity is measured on a scale from 1 to 4. 1 being least severe and 4 being the most severe accident. The study showed that the most common severity in an accident was between a 2 or 3. A chart below depicts this data.



I tested the data in relation to other contributing factors, specifically with crosswalks and stop signs, since they are commonly found in school zones. All of these variables are categorical variables and a chi-squared correlation test was performed to determine if they had an impact on accidents in school zones.

The results showed that the accident strongly correlates with crosswalks and stop signs, suggesting that there is a relation between accidents, crosswalks and stop signs.

I also performed a chi-squared in relation to weather. There is a categorical variable for weather in the US Accidents dataset which consists of the weather at the time of the accident. The result also showed that weather has a strong impact and correlation with accidents in and outside of school zones.

Conclusion

To conclude, we can say that school zones may not be as effective as we think they are. With about 23% of accidents occuring in school zones and approximately one-hundred children killed in school zone related accidents each year.

We can also conclude that crosswalks and stop signs need to be reevaluated, among other traffic safety designs that may not be as safe as we thought they initially were. Weather was also a contributing factor but we have no control over that.

School zones need to be improved. People driving in them need to be more cautious and children and teenages walking to and from school need to pay attention. The goal is to reduce the number of accidents in school zones. 23% of accidents is an alarming number and we should be more concerned about how we behave in school zones.