The Effects of The Nudge Theory on Injuries: A Case Study on New York State's Daily Non-Intersection Injury Rates in 2019

Sherien Hassan

May 2021

Introduction

Injuries that are sustained by pedestrians are often classified as accidents, but as we analyze traffic behaviors further, we are more likely to deviate away from the inference that these 'accidents' are random but rather, are consequences of poor infrastructure. According to The National Highway Traffic Safety Administration (NHTSA), 73.17% of pedestrian injuries took place at non-intersections in 2019 throughout the continental states, which is an improvement from 1997-2003 with a rate of 87%. Despite the decrease, we have not been able to conclude the underlying causes of such injuries. Across the continental United States, New York recorded the largest number of non-intersection pedestrian injuries in 2019 on both February 6th and September 8th. This justifies the selection of New York as our case study, analysis will be presented on both state and city-wide data. A potential response to address pedestrian injuries nationally can be adopted through the philosophical approach of Nudge Theory. The Nudge theory posits that through subliminal signalling, individuals are *nudged* into making preferrable choices that guarantee the safety of drivers and pedestrians alike. An application of this theory has been seen through New York City's Vision Zero initiative, a city wide effort to decrease pedestrian fatalities and injuries. This program indeed influenced New York's injury rates, which will be further looked at throughout this paper. Through an analysis of car crashes across the continental United States, this work aims to understand the environment surrounding these collisions and how Nudge Theory can help us create safer roads for people.

Hypothesis

The working hypothesis stipulates that we can put measures in place that can reduce pedestrian injuries through the analysis of accident prone sites using the Nudge theory, resulting in safer choices when crossing. Visualization tools and general linear models are employed to identify spatial temporal patterns in the NHTSA data. Computations are done using both R and Tableau.

Data Description

This data was extracted from The National Highway Traffic Safety Administration, specifically the FARS Analytical User System which contains datasets about car crashes across the United States. The two datasets that are of interest for this research are the Person and the Accident datasets. The latter of which consists of a single entry for each car crash recorded during the year. The dataset identifies many variables but for the purpose of this study the focus is on the variables of location, time, whether or not it was at an intersection and weather conditions. Location i.e. the latitude and longitudinal coordinates identifying the precise location of where the accident happened making it easier to group accidents by state down to the street. Time is presented in the dataset down to the minute but we are only interested in daily crashes for the year of 2019. Intersection Type, which identifies what type of intersection, if any, was the site of the accident. The final variable to be considered from this dataset is weather conditions as a constant value for our model. The Person dataset allows us to see the actual pedestrian injury count and the two datasets merge through the use of a data element number stated as ST_CASE, which assigns a unique identifier number to each crash. In order to ensure that the data obtained is comparable throughout the continental states, pedestrian injury rates are used instead of pedestrian injury count by using the 2019 US Census count. This allows us to offset the varying population rates for each respective state. Since it’s entirely possible that 1 single pedestrian injury occurred on a day which would result in miniscule numbers, that rate is also multiplied by one hundred million. For example, as we will see in our exploratory analysis of New York State (Figure 1), on December 6, 2019, there was an injury rate of 10%. This means two people died at a non-intersection in the state on that given day. This value of 10% is better to analyze than its respective decimal value.

Exploratory Analysis on case study: New York State

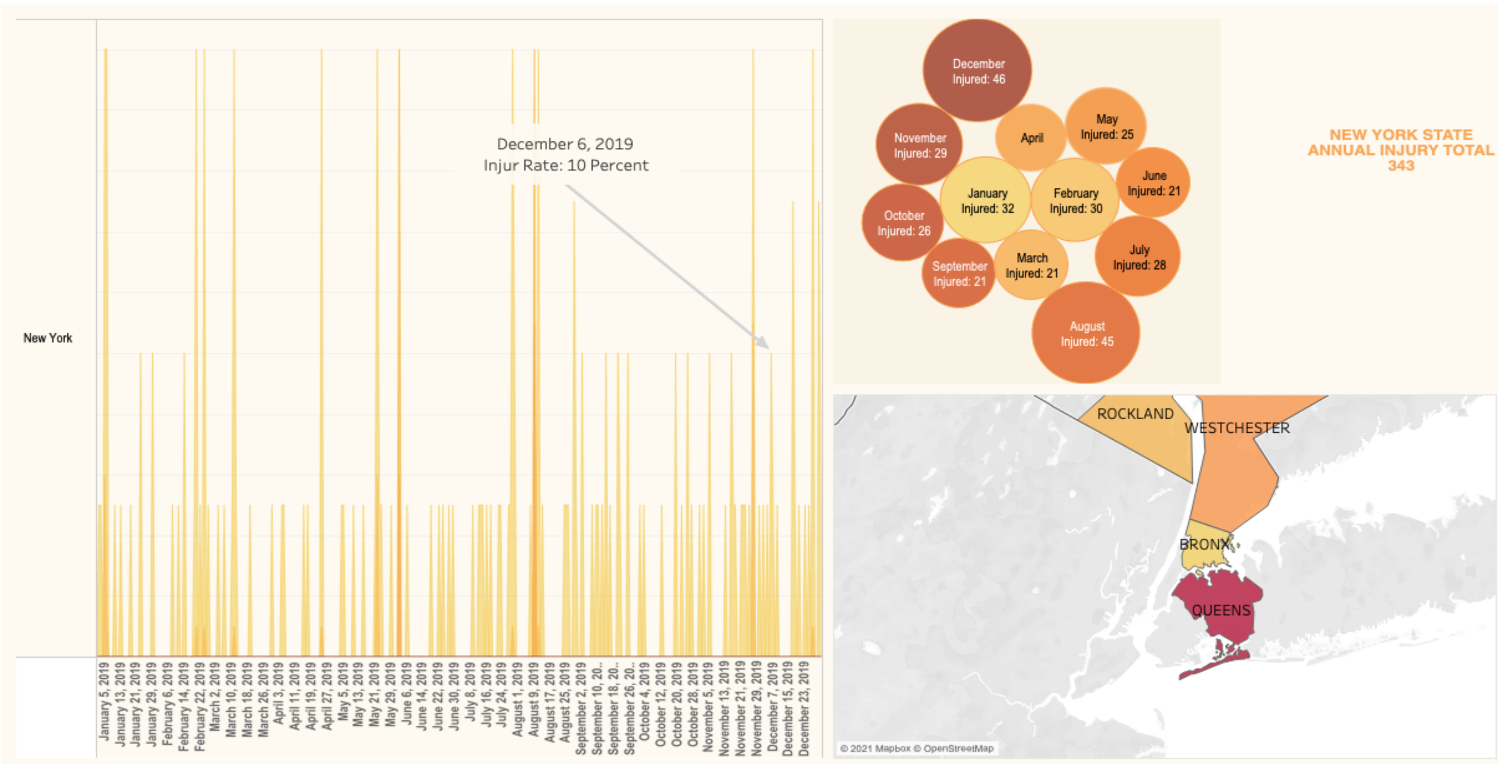

Figure 1: New York State's time series, pedestrian injury by month with its respective location

Figure 1 depicts New York States time series which does not show any form of trends, seasonality, or cyclic behavior and hence, no strong patterns can be deduced. There were several spikes that occured in random clusters throughout the year which are, as of the exploratory stage, unexplainable in terms of time alone. If the variable of space is factored with time however, it's possible to hypothesize that at a given point on the timeline, the structure of the street may not have been suitable to allow the pedestrian to make the safest choice. Since pedestrians are not given optimal options, there is an increased risk of injury which could help in explaining these spatiotemporal trends. Queens alone had a high of 38 injured in a single month, making it the leading county in New York State for pedestrian injuries. Therefore, deeming the spatiotemporal analysis of pedestrian injuries at non intersections in New York State important to analyze beyond the exploratory stage.

Model

A Multiple Linear Regression was chosen due to its ability to model and understand the association of each individual factor with respect to pedestrian injuries. In this case, $\hat{Y}$ is the expected or predicted pedestrian injuries of our dependent variables. Non-intersection and weather conditions: rain, snow, fog/smog/smock, cloudy, and not reported are distinct independent variables and their associated estimated regression coefficients indicate the value of $\hat{Y}$ when our independent variables are equal to zero. Since all of our independent variables are indicator variables (except for pedestrian injury count), non-intersection would be the change in $\hat{Y}$ relative to whether non-intersection is equal to zero or not equal to zero (e.g. 0 for intersection, 1 for non-intersection). The multiple linear regression model we will look at will be in terms of our case study, New York State, depicted below:

$$\hat{Y} = 0.0680556 + 0.0822929(\text{Non-intersection}) - 0.0115825(\text{Rain})$$
$$- 0.0273291(\text{Snow}) - 0.0308884(\text{Fog/Smog/Smoke})$$
$$- 0.0000741(\text{Cloudy}) + 0.0081270(\text{Not Reported})$$

Here, we have specific weather conditions because the ones presented have the most significance compared to the other potential weather conditions. The first thing to note is when this model was first conducted with a simple linear regression model, the association between being at a non-intersection and having there be a pedestrian injury there increased (0.072524 versus 0.0822929) after the adjustment for weather. We must be reminded that our regression coefficients are rates per one hundred million people, hence why our values may seem smaller than they actually would have been had they been a count of pedestrians instead. It's possible to conclude that being at a non-intersection remains statistically significant in its association with pedestrian injury where we have a p-value less than .01. Another noteworthy point is that the independent variables, being weather conditions and non-intersection, were limited since adding any other variables such as lighting assessment (i.e if the closest light fixture was working) would reveal the existence of collinearity between variables. This redundancy between predictor variables would cause for our regression model to become unstable and was therefore left out. We can however presume that non-intersection pedestrian injuries in New York still hold significance even with varying weather conditions.

Results

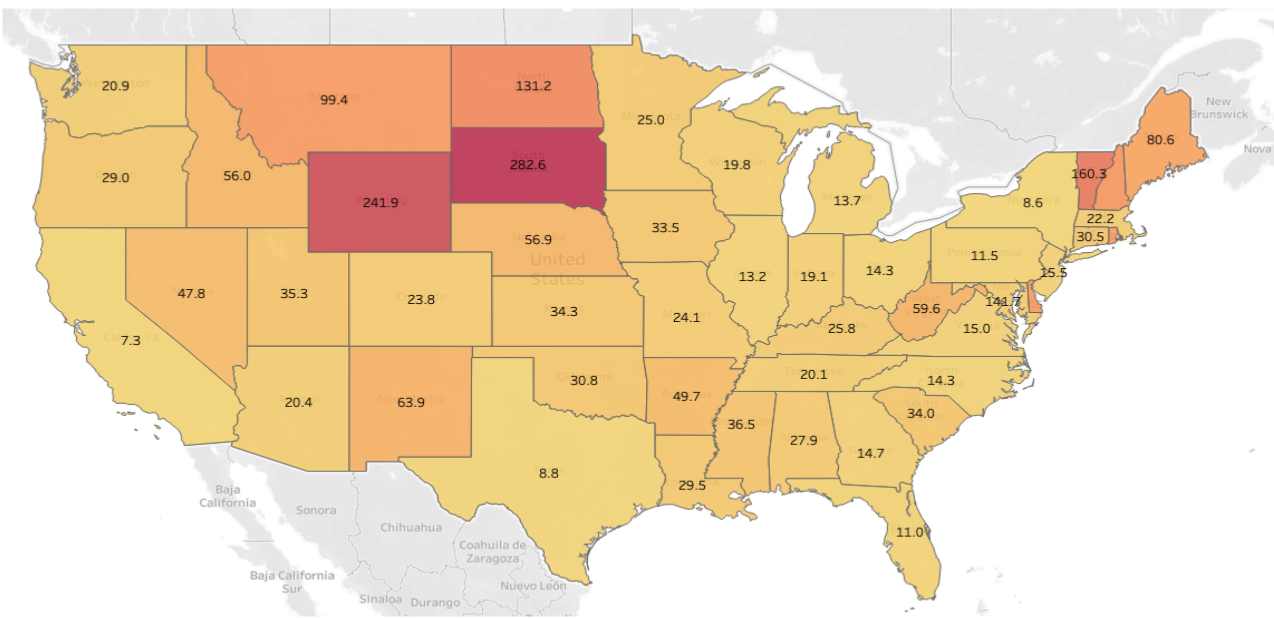

Figure 2: Non Intersection injury rates throughout the continental states

As shown above, Figure 2 depicts the injury rate per one hundred million people, categorized by state. The states showing the deeper red are states with higher pedestrian injury rates, while the lighter yellow signifies states with lower injury rates. New York States' value could be interpreted in the following way, when injuries occur at non-intersections, it occurs with an average rate of 8.6 percent more as compared to injuries at intersections. It's worth noting that this figure is with the inclusion of weather as a variable. Despite having an obviously influential variable such as weather in our multi-linear model, non-intersection injuries still remain significant. Now, New York states having a low injury rate was not expected when exploring the state on the exploratory level. It's important to highlight the importance of population statistics since it's a factor of our injury rate. New York's unexpected injury rate is largely due to its much larger population than Wyoming or South Dakota for example. New York's population is 19,453,561 compared to Wyoming which has 5,822,434 and South Dakota which has 884,659. Yet, the fact that New York's average rate is low does not necessarily mean it is not worth analyzing.

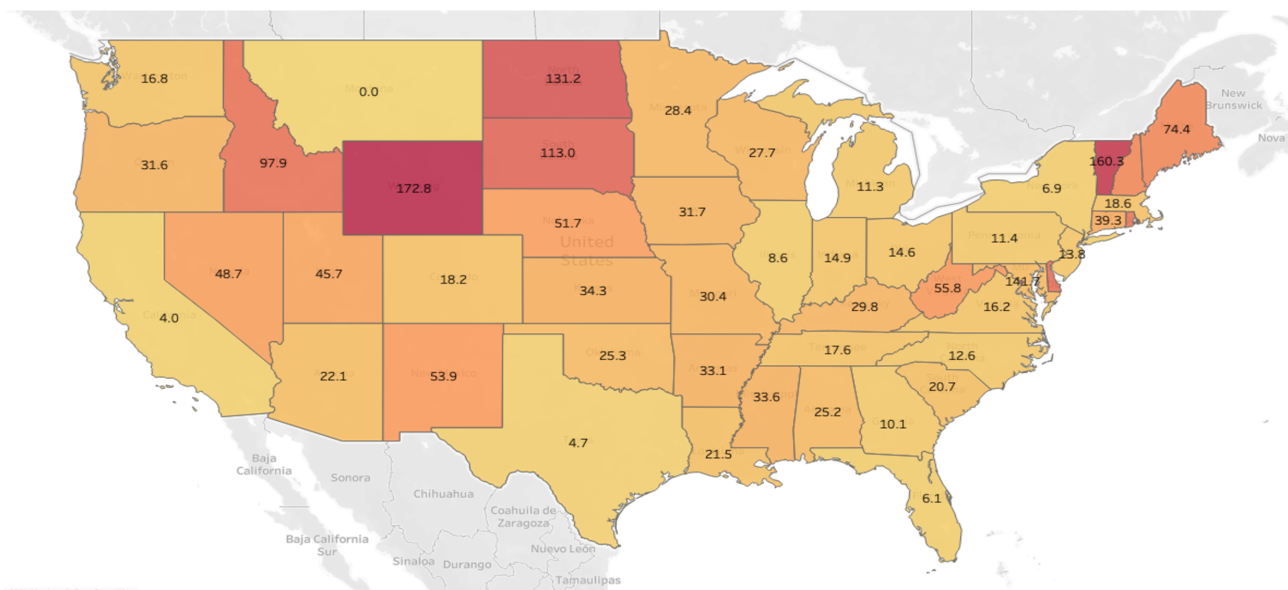

Figure 3: Intersection injury rates throughout the continental states

Figure 3 depicts injury rates at intersections, which shows improvements in almost all states when compared to injury rates at non-intersection. This implies that the analysis could be broadened to different sorts of street structures as the raw values are still high. One thing to note is Montana's pedestrian injury rate at intersections is zero which could be explained by how the state is defining non-intersection. Since each state is providing the raw data for injuries, it is possible that some states may inadvertently skew their figures. However, low average injury rates do not necessarily signify misrepresented data, at least not in the case of New York. New York's implementation of Vision Zero has been the closest practical application of the Nudge Theory and it has done a tremendous job in keeping injury rates relatively low.

The Nudge theory proposes indirect suggestions and positive reinforcements as a way to influence the behavior and decision making of individuals. As mentioned previously, the most efficient way of reducing pedestrian injury rates is by improving infrastructure. A street that is built in a particular manner is able to suggest specific crosswalks, the absence of which allows pedestrians to undertake the quickest way, which is sometimes straight through traffic. If we take a look at New York State, their non-intersection rate was low compared to other states and this is actually due to the Nudge Theory. Specifically, it is due to the unintentional application of the Nudge Theory, presented through New York Cities Vision Zero initiative. Vision Zero was created with the aim to achieve a zero fatality rate across the city and one of their main projects was changing street structure to indirectly suggest to pedestrians the safest way to cross the street. An example of a project vision zero worked on is depicted below.

BEFORE

AFTER

Figure 4: Queens renovation to Great St 65th pl with a pedestrian Island

During the exploratory analysis, it was shown that Queens county was one of the most problematic counties in the state, with 38 pedestrian injuries in a given month. In figure 4, we see a successful attempt, proposed by New York Cities vision zero initiative, to decrease pedestrian injuries and fatalities on Great St 65th pl which had an all time high of 18 pedestrian fatalities in a single day. Their solution was to add a pedestrian island which made those crossing the street less likely to cross on a red light, for the island partitioned a longer than average street into shorter ones. This also decreased the time pedestrians were sharing a road with cars which also benefited the pedestrian and therefore, resulted in New York having one of the lowest pedestrian injury rates across the continental states. The Vision Zero program was undertaken by New York

City in 2014, meaning that over the course of the following 5 years where data is collected pushed the number of car crashes down and should be extrapolated to a nationwide implementation.

Conclusion

The Continental United States has spent decades optimizing its roads and transportation systems to cater to cars, rather than pedestrians. This shows in as early as the exploratory stage of our research on daily pedestrian injuries at non-intersections in 2019. Specifically, New York state was analyzed due to having recorded the largest number of non-intersection pedestrian injuries on both February 6th and September 8th. Through a model selection process, a multi-linear model was chosen with variables varying from non-intersection injury rates to the various weather conditions that could have affected such values. In result, when injuries occur at non-intersections, it occurs with an average rate of 8.6 percent more as compared to injuries at intersections in New York State. Of course, this spatiotemporal issue may not seem to have an obvious fix at first, but through a philosophical notion of nudging, presented by the Nudge theory, we see that indeed there is a solution to such injury rates. One example was presented through Vision Zero, a city wide effort in New York to decrease pedestrian injuries and fatalities which had tremendous results. All in all, city planners must begin to think like a pedestrian rather than like a car to decrease rates statewide and make streets pedestrians safe.

Work Cited

Parvareh, Maryam, et al. “Assessment and Prediction of Road Accident Injuries Trend Using Time-Series Models in Kurdistan.” *Burns & Trauma*, BioMed Central, 9 Mar. 2018, burnstrauma.biomedcentral.com/articles/10.1186/s41038-018-0111-6.

“Vision Zero: Engineering Initiatives .” *Vision Zero*, www1.nyc.gov/content/visionzero/pages/engineering.

Pedestrian Safety Action Plan Vision Zero: Queens , www.nyc.gov/html/dot/downloads/pdf/ped-safety-action-plan-queens.pdf.

“Explore Census Data.” *Population*, www.census.gov/topics/population.html.

“NCSA Tools, Publications, and Data.” *NCSA | Tools, Publications, and Data*, cdan.nhtsa.gov/.