Examining Bike and Pedestrian Harms in Chicago  
 Author: Cole von Glahn   
 Data: Sourced from Chicago Data Portal   
   
Cities are, by definition, compact human environments. Providing services that are constrained by   
distance traveled is very efficient in cities. As we continue to grapple with human-driven climate change,   
it is essential that we leverage this trait to help individuals and companies reduce their carbon   
footprints. These are the kinds of reductions that are achievable on an individual scale. That granularity   
empowers the individual and makes strides where the necessary largescale changes are stalled by   
consensus and collective decision-making processes.

To take advantage of this trait, however, requires the means of transit to be efficient. Car centered   
transportation systems are energy, space, and time inefficient within urban centers. While public transit is an integral piece of relieving automobile transportation inefficiencies, they are not individualized responses. Biking, scootering, skating, walking, and other personal, human-driven transit methods are. However, safety concerns form a significant barrier for populations that may otherwise be willing (if not excited) to enjoy the autonomy, exercise, speed, and environmental efficiency of these methods.

This series of visualizations sheds light on these dangers. Understanding the extent and location of   
danger to cyclists and pedestrians reveals the amount of work to be done. Providing better protected   
lanes, fewer car-legal streets, and more wraparound services for human-driven commuters can  
reduce conjoined harms to our populations and our planet.

Pulling from Chicago’s Open Data Portal, These visualizations employed crash data focusing on the involved parties (cyclists, pedestrians, drivers) the locations of incidents by street and coordinate, the time the crash occurred, the cause of the crash, and the severity of the incident. Information regarding the parties was well-coded and easily accessible in the dataset. However, it is widely accepted that crashes involving cars and cyclists are undercounted due to differences in insurance circumstances and civil or criminal accountability relative to car-on-car accidents. The other items required significantly more wrangling and translating to make useful for visualization.

Location data is provided in two manners. Streets are easily shown in the data without any need to process them further. Coordinate points and polygon matching python libraries were used to determine neighborhood data based on geoms provided by a separate Open Data Portal dataset. In spite of, Chicago’s organization of civil and political activity at the neighborhood-level, neighborhood definitions and names differ widely from dataset to dataset. The dataset used was the most granular that the author could find, preferring specificity in certain locations to the recognizability of larger geographic blocks.

Time was given in hours and translated to categories in order to provide more meaningful information to the reader. These categories are defined as: Morning 5am – 10am, Afternoon 11am – 5pm, Evening 5pm – 9pm, and Night 10pm – 4am. The break points were designed to identify the impact of commute times and lighting conditions. There was lighting condition data separately available, but was highly subjective and implied more specificity than the data actually offered.

Each crash was coded with a primary and – if appropriate – secondary cause. Causal claims in these visualizations were derived only from the primary cause field. This field provided a wide and subjective range of causes with immense granularity for certain reasons (differentiating general cell phone use from texting from navigation, for example) and generic breadth for others (weather). To improve categorical significance and readability, these categories were binned into driver skill, environmental, and other. The other category contains all data for incidents with causes reported as indeterminable or not applicable. No further definition was provided for these latter two subcategories.

Severity was reported in the data as a four-point scale ranging from “Reported, not evident” to “Fatal” and then counted by number of such injuries per incident. It is important to note that these reports only take on-site information into account, and only for those incidents where the police were called to the scene. If there was no police involvement, or the individuals’ injuries developed or worsened after the fact, such data was not available. It is therefore likely that the severity of car-on-bike and car-on-pedestrian accidents is higher than what is reported in these visualizations. A 0-1 continuous scale was applied to convert injury types and counts into overall accident severity. Fatalities were counted at a 1:1 ratio, scaling down at 25% increments for each of the three less severe outcomes. For example, an accident with a single death and 2 reported but not evident injuries would receive a severity score of 1 + .25 + .25 = 1.5.

Why are you interested in the topic?

Cities are, by definition, compact human environments. Providing services that are constrained by   
distance traveled is very efficient in cities. As we continue to grapple with human-driven climate change,   
it is essential that we leverage this trait to help individuals and companies reduce their carbon   
footprints. Intra-city transportation is the kind of reduction that is achievable on an individual scale. That granularity empowers the individual to make strides where the necessary largescale changes are stalled by consensus and collective decision-making processes.   
To take advantage of this trait, however, requires the means of transit to be efficient. Car centered   
transportation systems are energy, space, and time inefficient within urban centers. While public transit is an integral piece of relieving automobile transportation inefficiencies, they are not individualized responses. Biking, scootering, skating, walking, and other personal, human-driven transit methods are. However, safety concerns form a significant barrier for populations that may otherwise be willing (if not excited) to enjoy the autonomy, exercise, speed, and environmental efficiency of these methods.

On a personal level, I am a multi-modal commuter who would bike every day were it not for our infrastructural failures. Over the summer, I was involved in an ugly hit and (attempted) run while biking on Clark Street. The confrontation that followed crystallized the need to protect bikers from a car culture that sees their safety as incidental to drivers’ convenience.

This series of visualizations sheds light on these dangers. Understanding the extent and location of   
danger to cyclists and pedestrians reveals the amount of work to be done. Providing better protected   
lanes, fewer car-legal streets, and more wraparound services for human-driven commuters can  
reduce conjoined harms to our populations and our planet.

I had originally hoped to shed light on the relationship between bike lane infrastructure and biker safety by comparing multiple cities. However, the variance in features and subjectivity of coding choices made those comparisons loose, and I found myself connecting dots on tenuous grounds. Scaling back to Chicago, I then chose to eliminate the bike-lane specific portion of the analysis in favor of a more general view of our current state of play. Through conversations with administrators at the Open Data Portal, it was clear that bike lane data is poorly maintained and bears little resemblance to the reality of bike infrastructure in the city. Furthermore, I was unable to get an adequate explanation of what the features in their bike lane dataset corresponded to in reality. Fortunately, the [Traffic Crashes - Crashes](https://data.cityofchicago.org/Transportation/Traffic-Crashes-Crashes/85ca-t3if) dataset on the Chicago Open Data Portal had robust information on crashes in general, detailed injury data, and excellent coding for victims to focus on car-on-other type accidents.

How did you define your story? Any pivots that you made.

Is any data missing? Do you have concerns about  data gaps?

Any ethical consideration in the story telling or disclosures you would like to make about choices you made while visualizing?

Why did you choose the charts you chose? How do they help tell the story?

Any coding challenges or takeaways? Etc.