VISION ZERØ
CITY OF PHILADELPHIA

Vision Zero High Injury Network Methodology

VISION ZERO DATASETS:

Date released: October 2, 2017

1. HIGH INJURY NETWORK

   https://www.opendataphilly.org/dataset/high-injury-network

2. STREET CENTERLINES FOR HIGH INJURY NETWORKS

   https://www.opendataphilly.org/dataset/street-centerlines-for-high-injury-network

3. PENNDOT CRASH DATA

   Available through Open Data Philly

DATA SETS USED:

1. Reportable crashes in Philadelphia from 2012-2016, available as open data through PennDOT

2. Street Centerline geographic layer, maintained by the Philadelphia Streets Department, available as open data through Open Data Philly

3. Street Intersections geographic layer, maintained by the Philadelphia Streets Department, available as open data through Open Data Philly.

DATA CLEANING:

PennDOT provides crash data in tabular format, with Latitude and Longitude coordinates of the crash as reported by officials at the scene of the crash. Using ArcGIS Desktop:

- Plotted the X,Y locations of each crash point

- Removed crashes without valid coordinates

Removed crashes that were either on the interstates or not on a street:

- Acquired PennDOT's State Roads layer to get the widths of the interstates. This introduces some uncertainty since it does not capture variations in the width of a street over a single stretch of road.

- Identified and removed all crashes that were within the Interstate roadway or ramps, including overpasses. It is possible that a crash within an overpass occurred on the local road below, but this was not possible to distinguish so these were always disregarded as likely interstate crashes. We accounted for the uncertainty in the street widths by manually inspecting any crashes that occurred within double the reported width against aerial imagery captured between March 1 and April 30, 2016.

Now that we had a set of all the crashes on local streets, we wanted to be able to identify the street and/or intersection where the crash occurred.

- Removed all street segments with "EXPY" (expressway) in their street name or type. Also removed all segments of functional class "Non-travelable" or "Walking Connector" or indicated as "Stricken" which means it no longer exists.

- Identified the nearest street to each remaining crash and the distance using the Near function in ArcGIS. Manually inspected all crashes farther than 30 feet from the closest street and removed crashes located in parking lots, back alleys, or back driveways. The threshold of 30 feet was derived from the distribution of distances among crashes which identified what constituted an outlier. No fatalities occurred among these removed crashes.

- PennDOT's crash data has an "Intersection Type" flag that indicates where on a given block the accident took place. We considered any crash marked as occurring at a type of intersection, a crossover, or a railroad crossing to be occurring "at an intersection".

- Additionally identified all crashes within 35 feet of an intersection point or within 5 feet of a crash flagged by PennDOT as occurring at an intersection. Manually inspected all crashes indicated as Intersection crashes by PennDOT but located more than 30 feet from the nearest intersection and identified those not located near an intersection identified by the Intersections layer as non-intersection crashes.

- For all these intersection crashes, we used the same Near function in ArcGIS to identify the nearest intersection and the distance to it.

- We go on to alter the Street and Intersection layers to optimize them for crash analysis. However, the crashes need to be associated to the street infrastructure first because the unaltered layer better represents the physical structure of the streets.

The Street and Intersection layers are maintained by the Streets Department for several purposes, especially for vehicle routing, and in some cases a single street might be broken into segments for each lane or a complex intersection may be represented by multiple intersection points. If these are not aggregated, they produce misleading results, so to accurately capture the distribution of crashes in Philadelphia, we needed to alter these layers. These layers have fields that can be used to correct for these issues.

- Each street segment or intersection is assigned a unique ID. We created fields that would represent the aggregated ID after making our adjustments.

- The Street Centerline layer contains a field "MULTI_REP" that indicates where a street is represented by duplicate segments. Note that this field does not group the related segments. The Intersections layer contains a "NODE_ID" field that is unique for each point, and an "INT_ID" field that groups duplicate nodes. By aggregating intersections by INT_ID, we identified those intersections containing multiple nodes.

- Identifying any street segments where MULTI_REP > 0, these segments were manually grouped and assigned an aggregated ID to represent a single segment per city block (or section of highway).

- Removed all segments marked as duplicates except for the first. The aggregated ID that was assigned to each removed segment links it to its respective segment that was not removed.

- Identified intersections that intersected a removed street segment. These were visually inspected to determine if they should be removed (for example, nodes occurring at an intersection of more than four lines could intersect a removed street without being a duplicate). Removed intersections were likewise assigned an aggregated ID linking them to the intersection that remained.

- Intersections identified as duplicates by the Streets Department but that did not intersect a removed street were de-aggregated and would not be considered duplicates for our purposes.

- Removed all duplicate nodes except the one that intersects the remaining aggregated street.

- Used map topology to identify intersections not located on a remaining street segment or that intersected a removed street. Manually snapped each intersection to the appropriate street segment. Placement was determined by moving the intersection perpendicular to the street segment.

- Used map topology to identify dangling street segments. Manually inspected each dangle and if appropriate modified the geometry of the segment to extend and snap the endpoint to the corresponding intersection. Care was taken to ensure continuity of streets where adjustments had interrupted them, as well as to not create new continuous segments.

Now that our Street Centerline and Street Intersections layers were optimized for crash analysis, we wanted to be able to examine streets as corridors. The Complete Streets plan already identified streets by the following types:

- Auto-Oriented Commercial/Industrial
- City Neighborhood Street
- Civic/Ceremonial Street
- High-Volume Pedestrian
- Local (Catch-All)
- Lower Density Residential
- Park Road
- Scenic Drive
- Shared Narrow
- Urban Arterial
- Walkable Commercial Corridors
- Other

These types of categorizations are accompanied by street improvement recommendations and requirements. We decided to define corridors based on these types.We first joined our optimized Streets layer to the Complete Streets layer by the street segment ID in order to get the Complete Streets Type for each segment.

However, the streets categorized as type Other needed to be categorized, so for any such street segment, the Type was defined as the functional Class of the street as defined by the street centerline, which includes the following categories:

- City Boundary
- Collector Residential
- Driveway
- Expressway
- High-Speed Ramp
- Local Residential
- Low-Speed Ramp
- Major Arterial
- Minor Arterial
- New Road
- Non-Travel
- Traffic-Controlled Crosswalk
- Walkway Connector

Additionally, the Complete Streets plan was created a few years ago, but the Street Centerline is updated every month.
So, there were a number of street segments in our optimized layer that had no street type assigned by the Complete Streets layer.

- To identify the appropriate street type, we identified the adjacent streets to each segment and if a segment was bounded on both sides by streets with that street's name and the same Complete Streets type, the segment was automatically assigned that Type. Segments that were only adjacent to a street with its shared name on one side were also assigned the Type of that adjacent segment. This accounted for majority of segments. For the remaining uncategorized segments, they were manually inspected to determine which of the two types assigned to the adjacent streets should be assigned to that segment.

- Then we dissolved the Street Segments layer by the full street name and the complete streets type, making sure to not create multi-part (non-contiguous) features.

- Finally, we use these aggregated streets and intersections to create single-lane and single-node-intersection crash corridors network.

- Then we matched segments to their respective corridor using a "within" spatial join. Now every street segment has a defined Corridor ID.

- Five segments did not join within any corridor due to alterations made to correct duplicates and artifacts in the ArcGIS dissolve process. Each segment was examined manually and in all cases, the segment was split across two adjacent corridors with the same name and Complete Street type. Those 5 corridors were merged and other segments updated to reflect the new corridor IDs.

- Intersections were matched to their respective corridors using a one-to-many "intersect" spatial join. Since each intersection can be part of multiple corridors, each intersection cannot be assigned a single Corridor ID.

DATA ANALYSIS:

Now, each crash is geo-located to several levels of granularity:

- Original street segment from the official Street_Centerline

- Aggregated street segment after removing duplicate lanes

- Original intersection node from the Street Intersections

- Aggregated intersection after removing duplicate nodes

In order to calculate the frequency of crashes at any of these levels, we need to count up the number of incidents per the respective unit of the street network. However, in order to represent a corridor, rather than just combining the segments within the corridor, we want to also include all the crashes at intersections within the corridor. Any crash at an intersection potentially reflects on all the streets meeting at that intersection.

- Each crash was identified as either an "intersection" or "mid-block" crash. Aggregated counts were calculated for each intersection and for each segment (counting only mid-block crashes toward the street segment's total and only intersection crashes for an intersection).

- A one-to-many spatial join of intersections to corridors allowed us to aggregate counts of crashes per corridor occurring at intersections. Each segment is associated with only one corridor, so using the corridor ID we could aggregate the total count of mid-block crashes occurring within a corridor. Finally by combining these two counts we get the total number of crashes along a given corridor.

- Note that because intersections get counted towards each corridor (and by definition of an intersection, therefore within at least two corridors), the total number of crashes on all corridors will be greater than the actual total number of crashes and should not be added up to get the total incidence of crashes.

For this analysis, we were interested in assessing crashes that resulted in either fatality or significant injury (KSI = killed or significantly injured), with special emphasis on crashes that involved either a pedestrian or biker fatality or significant injury.

- Each crash record contains fields that designate any Fatality, Major Injury, Pedestrian Fatality, Pedestrian Major Injury, Bike Fatality, and Bike Major Injury. The degree of injury is assessed by officers at the crash scene and can additionally be categorized as Moderate Injury, Minor Injury, Injury of Unknown Severity, and Unknown injury status.

- Each crash was flagged for each of the above conditions and the counts were aggregated up in the same way as described above for total crashes.
- Now we have a network of crash frequencies. We then needed to prioritize the most dangerous sections for this network.
- After evaluating the methods used by various peer cities, we determined that a weighted measure of KSI per mile was the best metric for prioritization to generate a High Injury Network.

The score was calculated using the following process:

- For each corridor or intersection, a new field was created to record this metric.
- Pedestrian and Bicycle KSI crashes were weighted by a factor of 1.25 in order to highlight these crashes. Compared to vehicular crashes, pedestrian and bike crashes are significantly more likely to result in death or serious injury. This factor reflects the degree of higher risk.

For corridors, additional calculation was done to normalize scores by corridor length and to reduce occurrence of outliers by using a measure of KSI per mile to determine high risk segments.

- Each corridor score was divided by the corridor's length in miles.
- All corridors shorter than 1000 ft or with one or fewer KSI crashes were assigned a score of zero to avoid having the rankings skewed too much by the corridor's length.
- Note that all KSI are events of crashes, not counts of individuals involved (e.g., a crash in which 4 people were killed or a crash in which 1 person was seriously injured get counted as 1 KSI).

HIGH INJURY NETWORK CRITERIA:

Using the network with score for each corridor, we filtered out the roads with score higher than 3. Then we optimized for addressing 50 percent of all KSI in the city.

- By using MS Excel, we created a table with streets in descending order of their KSI per mile score.
- We added a table of cumulative total of KSI and road length to this table.
- The target of 50 percent KSI was achieved at 2.2 KSI per mile using this calculation.
- Addressing 50 percent of all KSI within the network, we covered only 12 percent of road length of the city.
- Few shorter segments were added manually to complete the corridors based on classifications of those corridors as per Complete streets.