Journal Title
XX(X):1–15
©The Author(s) 2020
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Explorative Visualization for Traffic Safety Using Adaptive Study Areas

Andreas S Berres[1], Haowen Xu[1], Sarah A Tennille[1], Joseph Severino [2], Srinath Ravulaparthy[3], Jibonananda Sanyal[1]

Abstract

The pressing need for improving traffic safety has become a societal concern in many cities around the world. Many traffic accidents are not occurring as stand-alone events but as consequences of other road incidents and hazards. To capture the traffic safety indications from a holistic aspect, this paper presents a suite of visualization techniques to explore large traffic safety datasets collected from different sources using adaptive study areas which include the whole region (Hamilton County) as well as smaller sub-areas. In the present study, these data source include (1) Hamilton county's 911 emergency response data, which includes traffic incidents as well as other types of incidents throughout the county, and (2) Tennessee crash data, which contains only vehicle crashes with more detail on the circumstances of each crash. We use both abstract and spatial visualization techniques to derive a better understanding of traffic safety patterns for different traffic participants in various urban environment. In addition to the entire region of Hamilton County, we examine safety on the highways, in the downtown area, and in a shopping district east of the city center. We are able to characterize incidents in the different areas, gain a better understanding of common incident patterns, and identify outliers in the data. Finally, we present a textured tile calendar to compare spatiotemporal patterns.

Keywords

safety visualization, incident visualization, spatiotemporal visualization

Introduction

Safety is an important aspect of any urban environment, and there is a large body of research focused on reducing safety threats. According to the National Highway Traffic Safety Administration, approximately 6.4 million vehicle accidents are reported in the US every year, and many these result in severe injuries or fatalities National Highway Traffic Safety Administration (2001). With the increasing availability of crash reports and in-vehicle sensor data, analysis of traffic incidents are frequently approached through visual-analytics techniques Sobral et al. (2019); Wang et al. (2010). In terms of advantages, visual-analytics, by employing visual channels to represent complex datasets and transforming various types of data into easy-to-understand visual representations Hansen and Johnson (2004), are able to incorporate human reasoning into an interactive visual interface, thereby combining machine intelligence with human intelligence Chen et al. (2015).

In the context of analyzing traffic incidents, data visualization can efficiently display the locations of traffic incidents over a region. This is especially critical not only for improving the rapid response of accidents but also for providing data-driven insights into potential factors (e.g., condition and state of the transportation network) that incur traffic incidents Wang et al. (2010). At the individual road level, several studies have focused on presenting potential crash risks on road locations using 3D GIS-based visualization Li et al. (2007), coordinated bar chart views Pack et al. (2009), and traffic origins visualization Anwar et al. (2014). At the regional level, previous studies have utilized map-based heat maps Hilton et al. (2011); Plug et al. (2011) and Kernel Density Functions Xie and Yan (2008) for identifying regions and road networks that are prone to traffic accidents.

Despite the amount of significant research efforts in the past, the analysis and visualization of traffic incidents should follow a multi-hazard paradigm Kappes et al. (2012); Komendantova et al. (2014), as the occurrence of many incidents are indirectly caused by other road hazards. Some incidents, such as vehicle crashes, are caused by traffic itself, and they frequently lead to (or are caused by) traffic jams. In addition, many other incidents, such as fires or floods, can still have an impact on traffic due to road blocks or a large number of incident response vehicles. In this regard, a more holistic approach that considers the interconnections among road incidents, as well as between traffic accidents and other hazards, should be developed.

In order to better characterize the impact of safety on traffic, we developed visualizations that examine different

[1]Computational Science and Engineering Division, Oak Ridge National Laboratory
[2]Computational Sciences Center, National Renewable Energy Laboratory
[3]Energy Analysis & Environmental Impacts Division, Lawrence Berkeley National Laboratory

Corresponding author:
Andreas Berres, Computational Science and Engineering Division, Oak Ridge National Laboratory, PO Box 2008, MS-6085, Oak Ridge, TN 37831, USA.
Email: berresas@ornl.gov

aspects of safety. Often, such visualizations are carried out at state scale or county scale. However, by examining smaller areas (e.g., a corridor or intersection), one can gain additional insights about local differences in on-road and off-road safety. In this paper, we demonstrate this by examining vehicle incidents and non-vehicle incidents in the Chattanooga Metropolitan Area. We present web-based adaptive visualizations at the county scale (Hamilton County, Tennessee), as well as for three regions of interest: downtown Chattanooga, a shopping district in the Shallowford Road area in East Chattanooga, and interstate/highway incidents.

Shopping centers have become major vehicular attractions in many urban areas by incurring a large volume of shopping trips, subsequently generating unfavorable impacts on local transportation systems by increasing the amount of traffic congestion and incidents Hasan et al. (2012); Fancello and Fadda (2002). This has created socioeconomic implications for optimizing both the transportation systems design and urban planning in the shopping center vicinity Kristoffersson et al. (2018). The traffic dynamics affected by these shopping center attraction trips entail a significant variability in traffic speed and volume at multiple spatial scales (e.g., intersection, corridor, and region) and are further complicated by the presence of traffic incidents. Therefore, it is difficult and challenging to explore this ubiquitous problem using traditional modeling-based approaches.

Overview and Data

The goals of the research presented in this paper include providing an explorative visualization of incidents in the Chattanooga Metropolitan Area. For some data, such as safety data, it makes sense to use the same visualizations for different subsets of the data to study different areas of interest. While it is relevant to know what the most common incident types or locations are for an entire region (such as a state or county), one can also gain additional insight by restricting the data to a smaller area, such as downtown or shopping districts.

In this work, we want to emphasize the merits of using adaptive scales for safety data analysis and visualization. We demonstrate the capability to select subsets of data using different criteria and present unique differences between three different study areas as well as the region at large.

In Section , we present abstract visualization techniques, such as bar charts, heatmap matrices, and a sunburst chart. They are interactive visualizations to survey the data aggregated by day of week, time of day, incident type, and different incident conditions.

In Section , we present various map-based visualization techniques which represent the data in a spatial format at different aggregation levels. Kernel Density Functions provide a high-level overview of the region's incident hot spots, individual points give an impression of the distribution of incidents throughout the different study areas, and clustering provides an in-between level of aggregation.

Study Areas

In addition to a regional view, we also provide different study areas which are of interest due to their various characteristics. We provide different methods to define study areas using polygon boundaries or a distance function in relation to a feature. In the following sections, we present three different study areas: the highway system, downtown Chattanooga, and a shopping district.

Highway System Chattanooga is located near the Tennessee-Georgia border at the intersection of several major highways. It is cited as one of the most congested highway sections for truck traffic in the United States, according to the American Transportation Research Institute (ATRI) (2018). Interstate Highway 24 (I-24) runs south from Nashville into North Georgia, before it meets US-27 near the Chattanooga city center. Just five miles further East, it meets I-75 which connects Knoxville and Atlanta. Due to the steep terrain and high traffic volume, this particular area experiences very high congestion levels for trucking and passenger traffic alike American Transportation Research Institute (ATRI) (2018), and it features different incident types when compared to local roads. North of the interstate junction, State Route 153 (SR-153) joins I-75.

The four highways considered for this study area are I-24, I-75, US-27, and SR-153. We define this study area to be a band that extends 100 meters on each side of the chosen highways within the city, as seen in Figure 1. The width of the band was determined such that it included incidents with imprecise coordinates while excluding nearby frontage roads.

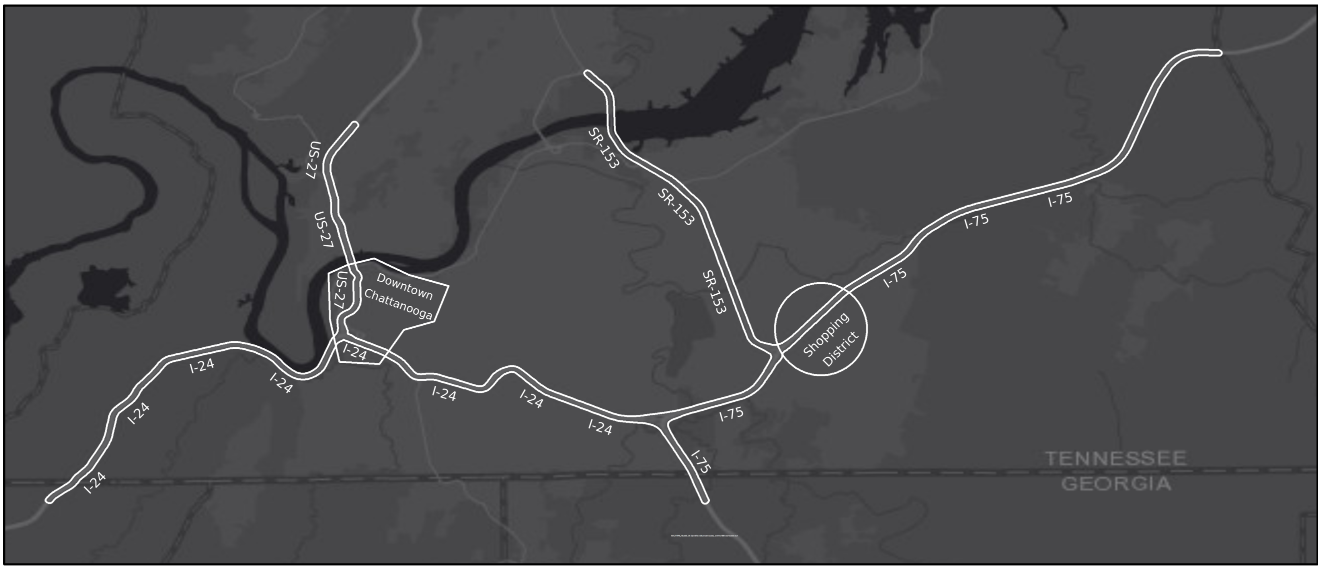
Downtown Chattanooga The downtown area of Chattanooga features a tight mesh of streets of different sizes. There are many businesses, such as offices, restaurants, and retail, as well as residential housing. As expected from such a diverse area, there is a high volume of traffic throughout the day which is comprised of personal vehicles, public transportation, bicycles, and pedestrians.

This study area encompasses all of downtown Chattanooga, including the campus of the University of Tennessee, Chattanooga, using a manually-created polygon created to reflect the unique local geography.

Shopping District This third study area is a shopping district in Chattanooga which provides unique traffic safety challenges. This popular commercial area is off of I-75, a little over two miles north of the I-24/I-75 junction, accessible by the Hamilton Place Boulevard and Shallowford Road exits. The area stretches from Igou Gap Road in the South to Shallowford Road in the North, along both sides of Gunbarrel Road and Shallowford Road. Containing a large shopping mall, restaurants, businesses, a strip mall, two large grocery stores, gas stations, a movie theatre, and more, this location presents unique safety concerns with a prevalence of high traffic volumes, pedestrians, and densely packed commercial buildings. Shallowford Road extends west, crossing over I-75 and intersecting Lee Highway, another major arterial in this area. This dense cluster of intersecting major roads poses unique challenges to traffic operations as well as traffic safety.

We define this region using a 2-mile radius centering around the Hamilton Place Mall. This captures some of the nearby highway traffic which may be affected by shopping traffic, and it also covers most of the area's businesses and hotels.

Figure 1. Overview of the Chattanooga Metropolitan Area, which highlights the study areas for this paper: downtown Chattanooga (irregular polygon) in the West, a shopping district in the Shallowford Road area (circle) in the East, and the major highways and interstates (branch-like structure) throughout the region.



Data Sources

We consider two data sources for this work: incident response reports from Hamilton County's 911 calls and Tennessee-wide vehicle crashes. The two datasets overlap in Hamilton County, but where 911 includes non-crash data, Tennessee crashes contains more details about the conditions and types of crashes. Therefore, both data sources add value to our analysis.

Hamilton County 911 Incidents Hamilton County provides 911 records of all incidents that require response teams throughout the county, classified by incident type for vehicle and non-vehicle incidents. These records include time stamps, locations, the number of response teams responding to the incident, and the type(s) and number of response teams, such as medical, fire, and law enforcement. The 911 incidents are publicly available in real-time, and we receive daily updates of historic data.

Tennessee Vehicle Crashes The Enhanced Tennessee Roadway Information Management System (E-TRIMS) is another source for highway incident data. While the 911 data only covers Hamilton county, E-TRIMS covers the entire state of Tennessee. However, this dataset is restricted to vehicle crashes, whereas Hamilton 911 data also includes non-vehicle incidents that require a response team. In addition to basic information like time and location, E-TRIMS offers additional information about crashes that is not available for the Hamilton County 911 data. The data includes:

- Outside circumstances, such as weather and light conditions;
- Details about the crash, including the manner of first collision (rear-end, sideswipe, etc.) and the first point of impact (another vehicle, guardrail face, etc.); and
- Additional safety factors, such as presence of hazardous materials, involvement of railroad or school buses, and whether it took place in a construction zone.

The vehicle crashes are updated weekly through a data portal.

Traffic Flow The Traffic Flow by Day of Week and Time dataset Sevigny (2020) is published by the City of Chattanooga Department of Transportation (CDOT). It provides vehicle counts and a traffic level rating from free flow traffic (0) to stand still traffic (4) for each day of the week at half-hour intervals. These vehicle counts are provided as a total for the entire county.

Road Network The Chattanooga Roadways dataset Chattanooga Department of Public Works (2020) is published by the Chattanooga Department of Public Works. It contains the road network geometry, as well as many attributes, such as road type, number of lanes, and base volumes overall, during morning and afternoon, and during off-peak hours. As these attributes are available for each link, it is possible to obtain more detailed information for each study area.

Bus Stops The CARTA Stops dataset Chattanooga Area Regional Transportation Authority (2020) is provided by the Chattanooga Area Regional Transportation Authority (CARTA). It contains the geolocation for each bus stop in the Chattanooga Area.

Base Map For all map-based visualizations, we use a dark gray canvas base map Esri and HERE and Garmin and OpenStreetMap contributors and the GIS user community (2020). This base map provides relevant geospatial context, such as roads and the nearby river. Due to the grayscale colorscheme, this base map is unobtrusive, and its dark colors provide a backdrop that highlights the visualizations.

Data Quality

While the Tennessee Vehicle Crashes dataset only includes one record for each incident, the Hamilton County 911 Incidents dataset can include multiple records if different entities were involved in an incident. For instance, for some vehicle crashes, there are three response teams: fire fighters, medical, and law enforcement. In this case, three records appear in the dataset.

To produce accurate aggregates of totals, we carry out two duplicate removal steps: In a first step, we run duplicate checks on a combination of date, time stamp, and incident type. This step removed 14,392 duplicate, which reduces the initial dataset size by 12.3% from 117,133 down to 102,741 incidents. Upon closer examination, we found that even after this step, some duplicates remain as some incidents are reported by multiple individuals. To remove these additional duplicates, we allow a tolerance of up to 5 minutes, and a distance of up to 0.1 miles between incidents to be considered identical, as long as the incident type matches. This second step removed another 274 duplicates, which constitutes a reduction of 0.3% compared to the first step, resulting in a final dataset size of 102,467 incidents.

Incident Types

We distinguish two different groups of incidents: *traffic incidents* and *collateral incidents*. The main criterion for this is classification is whether or not the majority of incidents of a given type happens in traffic or if their impact is more collateral in nature. We define these two categories as follows:

- *Traffic Incidents* are incidents which are directly caused by traffic or immediately affect traffic. This category includes all vehicle crashes, vehicle fires, entrapment, broken-down vehicles, and delayed vehicles.
- *Collateral Incidents* are incidents which are of a secondary nature but still have an impact on traffic. This category includes building fires, flooding, gas leaks, spills, and struck pedestrians.

Abstract Visualization

The main goal of abstract visualization is to provide information in an abstract way through graphics such as bar charts or heat maps. In this section, we present different types of abstract visualizations which we apply to the full region (Hamilton County) and to the three defined study areas.

Heatmap Matrices of Incidents by Day of Week

We visualize incidents by day of week and incident type. For this view, we choose a heatmap visualization. The colormap chosen ranges from a desaturated dark purple to a strongly saturated yellow, crossing through shades of blue, teal, and green with rising saturation. This creates a strong highlight for incident hot spots while also providing differentiation between areas of lower incident frequency. On the lower end, we fade out the colormap to white for very infrequent incidents to better blend in with the white background. With the resulting colormap, very frequent incidents are colored yellow or lime green, frequent incidents are colored medium green or teal, less frequent incidents are colored dark blue, and rare events are colored pale blue or white. For each cell, the numbers are printed with high contrast: for dark colors, we use white font and for light colors we use black font.

Heat Matrices for Traffic Incidents Figure 2 shows heatmaps of traffic incidents for the Hamilton County Region (top left), the highways (top right), downtown (bottom right) and the shopping district (bottom right).

Across most areas, traffic stops, i.e. law enforcement pulling over vehicles, constitute the overwhelming majority of incidents at about 64%. Throughout Hamilton County, Wednesday is the most frequent day for traffic stops, followed by Friday. Wednesday also has the highest overall number of reported incidents (17,368), followed by Friday (16,861). The lowest number of overall incidents (11,687) as well as traffic stops is, unsurprisingly, Sunday. The difference in traffic stops on Wednesdays is particularly striking for the shopping district, where drivers are almost twice as likely to be pulled over on a Wednesday than any other day of the week.

The one notable exception to traffic stops ranking highest is the shopping district, where crashes without injuries (41% of all incidents) are more frequent than traffic stops (38%). This underlines the importance of studying this particular area. In all other study areas, crashes without injuries rank second, with 18% of all incidents county-wide (or 66% of all traffic incidents).

Among traffic incidents (not including collateral incidents), crashes without injuries constitute 19% of county-wide incidents, with 4.5% of crashes with injuries, and 3% of crashes with unknown injuries.

County-wide, there are 4.3 times as many crashes without injuries as there are crashes with injuries. This rate is as low as 4.1 on highways, 7.7 in the shopping district, and 8.4 downtown.

The other incident types are much less common in the shopping district, downtown Chattanooga, and the region as a whole, with the exception of broken-down vehicles which constitute 15.5% of all traffic incidents on highways.

Heat Matrices for Collateral Incidents Figure 3 shows heatmaps of collateral incidents for Hamilton County (top left) and the three study areas. For the entire region, the most frequent collateral incident are residential fires, followed by struck pedestrians, brush fires, and gas leaks.

For the highway, the most frequent incident types are brush fires, spills, and struck pedestrians. Pedestrian incidents occurred more than twice as frequently on Thursdays as any other day. Brush fires were the most frequent on weekends, whereas spills were most frequent from Tuesday to Friday.

In downtown Chattanooga, the most common incident type is stuck pedestrians. There is an astonishingly large number of struck pedestrians on Wednesday. As downtown features by far the most residential land use out of the study areas, it is not surprising that apartment fires and residential building fires are very common. The high number of brush fires is less expected; however, there are natural areas within the selected downtown area.

For the shopping district, the most frequent incidents are struck pedestrians, as one might expect from an area with a lot of parking lots and therefore many pedestrians. In addition, there were two very noticeable events: 5 gas leaks and 7 flooding reports on Wednesdays. Further investigation revealed that on July 31, 2019, Chattanooga experienced

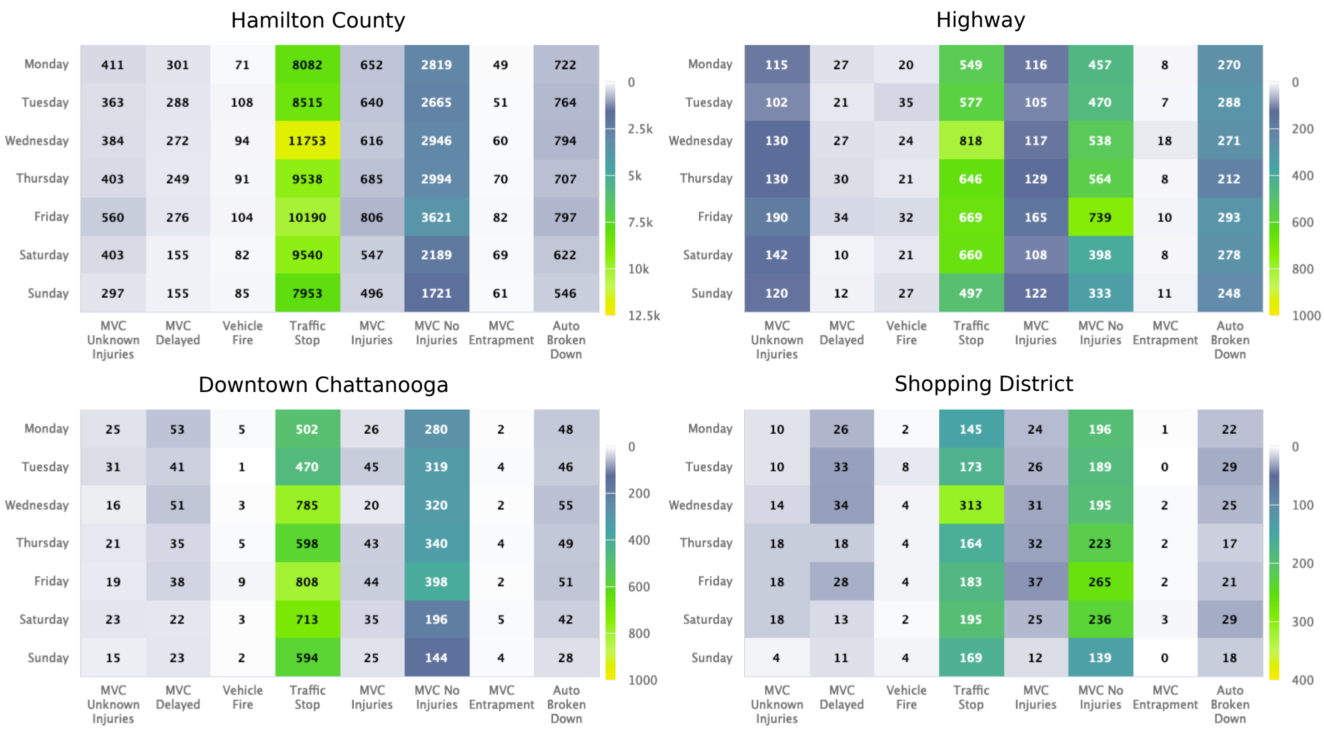


Figure 2. Heatmaps for 911 *traffic* incidents in Hamilton County (top left), for the highways (top right), downtown (bottom left), and the shopping district (bottom right). The data ranges from October 1, 2018, to September 30, 2019. Traffic stops are the most common type of traffic incident, followed by crashes without injuries.

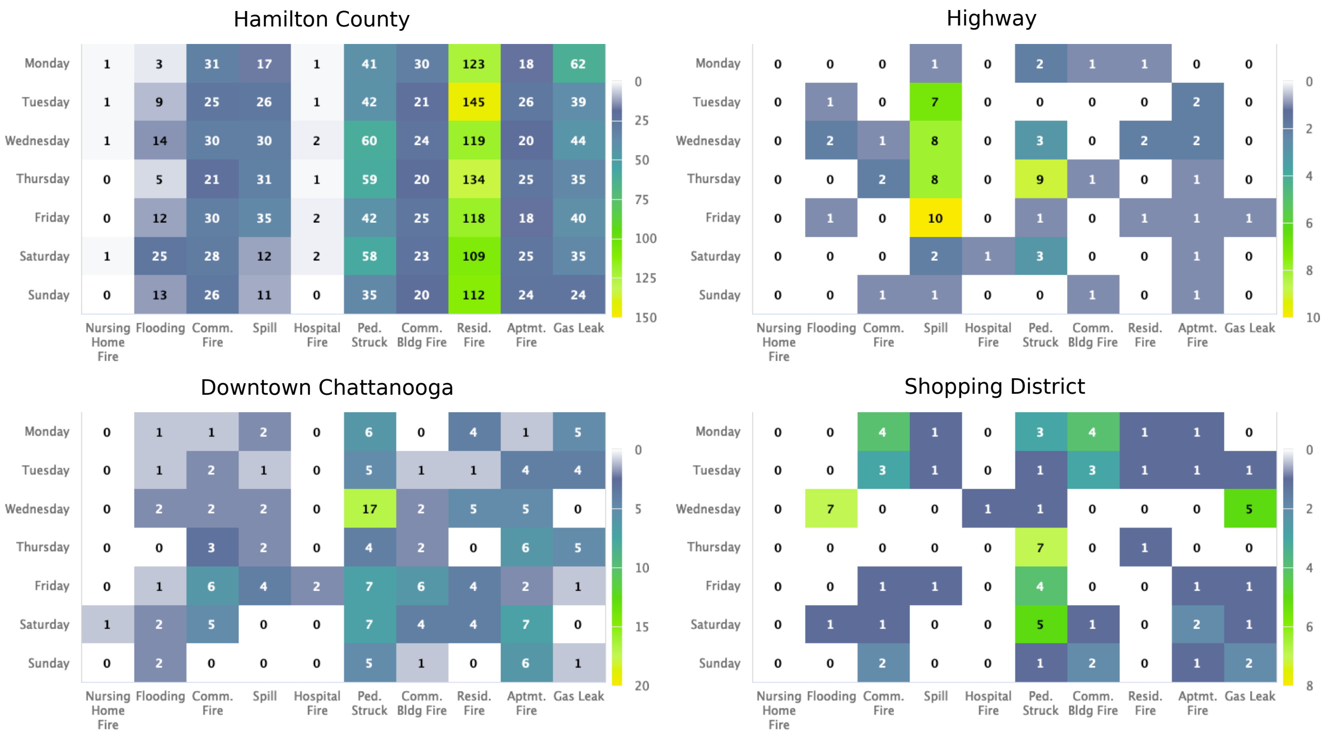


Figure 3. Heatmaps for 911 *collateral* incidents (including struck pedestrians) in Hamilton County (top left), for the highways (top right), downtown (bottom left), and the shopping district (bottom right). The data ranges from October 1, 2018 to September 30, 2019. The most frequent collateral incident type varies by region.

unusually heavy rain, which resulted in flooding in the mall Pare (2019). 5 out of 7 incident responses for flooding in the shopping district occurred on this day. We found that the gas leaks were distributed throughout the year and spread out across the region of interest, so the high numbers for Wednesdays are likely a coincidence.

911 Incidents by Time of Day

For further exploration of the incidents, we provide bar charts. We separate traffic and collateral incidents and choose colors from a color family that makes sense for the incident type (e.g. reds for fires, blue for flooding), and we reuse the same colors for traffic incidents that were used on the overview map. Table 1 lists the hours assigned to each time of day.

Time of Day	Hours
Early Morning	Midnight – 6 a.m.
Morning	6 a.m. – 10 a.m.
Mid-Day	10 a.m. – 3 p.m.
Evening	3 p.m. – 7 p.m.
Night	7 p.m. – Midnight

Table 1. Timespan for each defined time of day. Morning and evening are the peak traffic timeframes.

911 Traffic Incidents

Figure 4 provides a bar chart of traffic incidents for the region (top left) and the three study areas. A large number of incidents happen between 10 a.m. and midnight, with fewer incidents in the early morning and morning hours. Downtown Chattanooga has a comparatively high number of incidents in the early morning hours, while the shopping district has particularly low numbers, presumably because most businesses are closed during these hours.

On highways, the number of incidents is more consistent throughout the day than for the other two study areas. The relative number of vehicle crashes and broken-down vehicles is much higher on highways than it is for the entire county, and highways are the only study area with a substantial number of vehicle fires and entrapments. The shopping district is the only other study area that shows a noticeable number of these incident types, and upon closer examination (see Figure 10), the majority of these incidents happened on the stretch of I-75 which crosses through this area.

For the downtown traffic incidents, the most noticeable feature are the traffic stops, which peak very visibly at night and in the early morning hours and plummet in the morning. Vehicle crashes are very rare in the early morning and most common mid-day and in the evening.

Compared with the region-wide data, the shopping district has increased numbers of accidents without injuries during the main shopping hours (mid-day and evening), which almost vanish in the early morning. All other incident types are similar. Traffic stops are more consistent throughout the day in this area than in any other one.

911 Collateral Incidents

Figure 5 provides a bar chart of collateral incidents for the region (left) and the three study areas. The majority of incidents happen between 10 a.m. and midnight, with relatively few incidents in the early morning and morning hours.

On highways, the most distinctive feature is the much higher number of spills compared to all other collateral incidents. Struck pedestrians constitute another very distinctive feature, as their distribution throughout the day is unbalanced: there are a lot of incident during mid-day and night hours, but few in the early morning and morning and none in the evening. Numbers for various types of fires are very low and only incidental as there are some buildings that fall within the buffer around the highways.

Downtown, the most prominent incident type are struck pedestrians. Out of all areas, downtown has the most consistent numbers independent of the time of day, however they peak in the evening. Furthermore, there are a much larger proportion of apartment fires in this area than in any other area.

For the shopping district, there are three features that are significantly different from the county-wide data. First of all, there are more commercial building fires, which is expected in a highly commercialized area. Second, there is a high number of flooding incidents in the evening, which we have also identified through the heatmap to be on a Wednesday. The two charts combined are unusual enough to justify an investigation into news articles. Finally, there are many more events for struck pedestrians in this area than in the region at large, in particular during the night and early morning. Upon further research, we found that the majority of these incidents was reported for addresses of businesses with parking lots.

According to the Governors Highway Safety Association, 33% of pedestrian fatalities occur on local streets, and 75% occur when it is dark Retting and Consulting (2017). A study by (Lin et al. 2017) confirms the same finding that dark surroundings are one of the contributing factors to pedestrian fatalities. Furthermore, they examined land use and found that discount grocery stores and restaurants carry a higher-than-average risk to pedestrians. The data for the shopping district confirms with both of these results: dark surroundings are clearly reflected in Figure 5. We will take a closer look at incident locations in Section . Maps of collateral incident locations can be found in Figure 11.

Normalized Views

In order to account for the variation in traffic volumes throughout each day, and between different days of the week, we provide an optional normalization step for these charts. To normalize the incident counts, we use a Traffic Flow by Day of Week and Time dataset Sevigny (2020) and the base volumes from the road network Chattanooga Department of Public Works (2020). To maintain easy-to-read and interpret numbers, we scale the normalized traffic up by a correction factor to the achieve same magnitude.

For the heat maps, we first compute the fraction f_h of traffic for each day of the week as $f_h = \frac{\text{week day volumes}}{\text{all volumes}}$. Then, we multiply each day's incident counts by f_h and the correction factor $c_h = 7$ (number of week days). Similarly, we compute the fraction f_b of traffic by time of day as $f_b = \frac{\text{week day volumes}}{\text{all volumes}}$. Then, we multiply the incident counts by time of day by f_b and the correction factor $c_b = 5$ (number of segments in time of day). Figure 6 shows a comparison of heatmaps and bar charts before (left) and after (right) normalization. The normalized charts show the incident counts that would be expected based on the traffic volumes for the given day of week (heatmap) or time of day (bar chart).

A comparison of traffic stops between the two heatmaps highlights how unusually high the number of reported traffic stops on Wednesdays is: based on traffic volumes one would expect Fridays to feature 1.5 times as many traffic stops as Wednesdays. The other incident types behave roughly as expected, as one can see from the very similar color striping patterns. The only collateral incident with markedly different

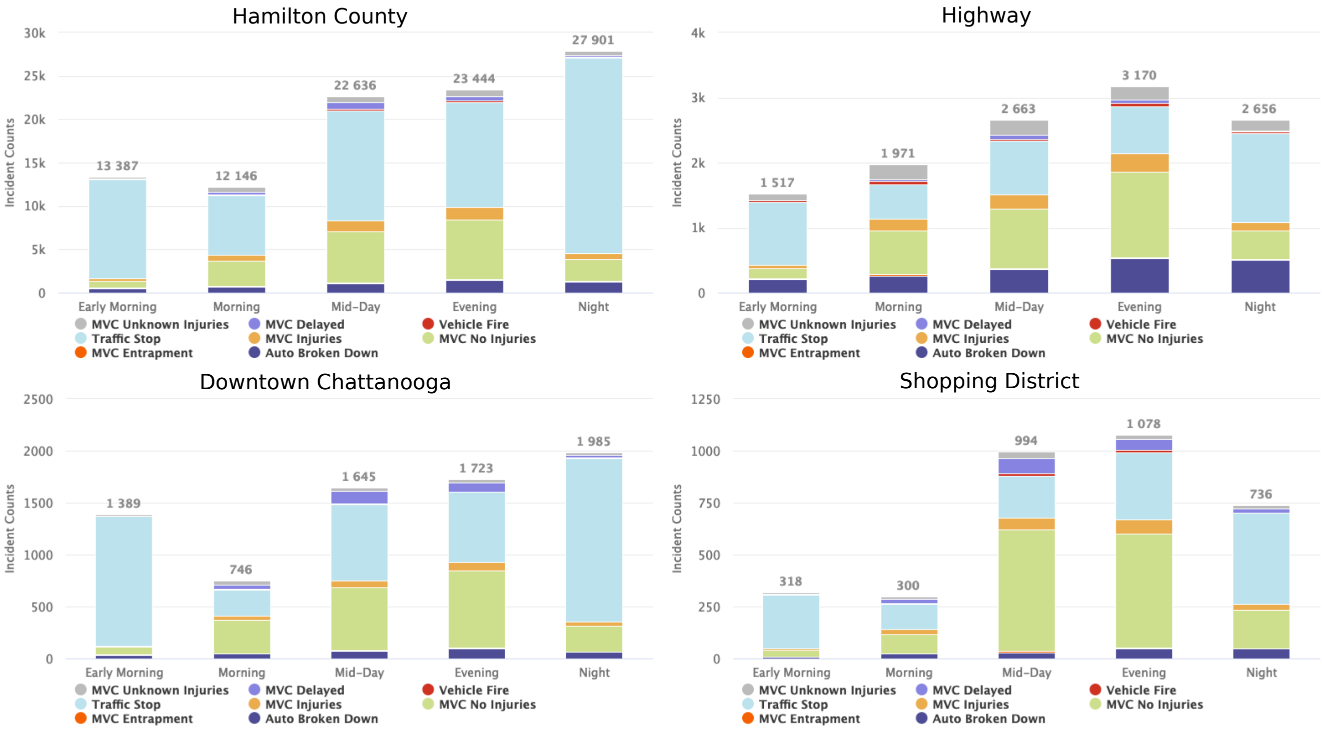


Figure 4. Bar charts for 911 *traffic* incidents in Hamilton County (top left), for the highways (top right), downtown (bottom left), and the shopping district (bottom right). The data ranges from October 1, 2018 to September 30, 2019. There are large differences in incident distributions and numbers depending on time of day and area.

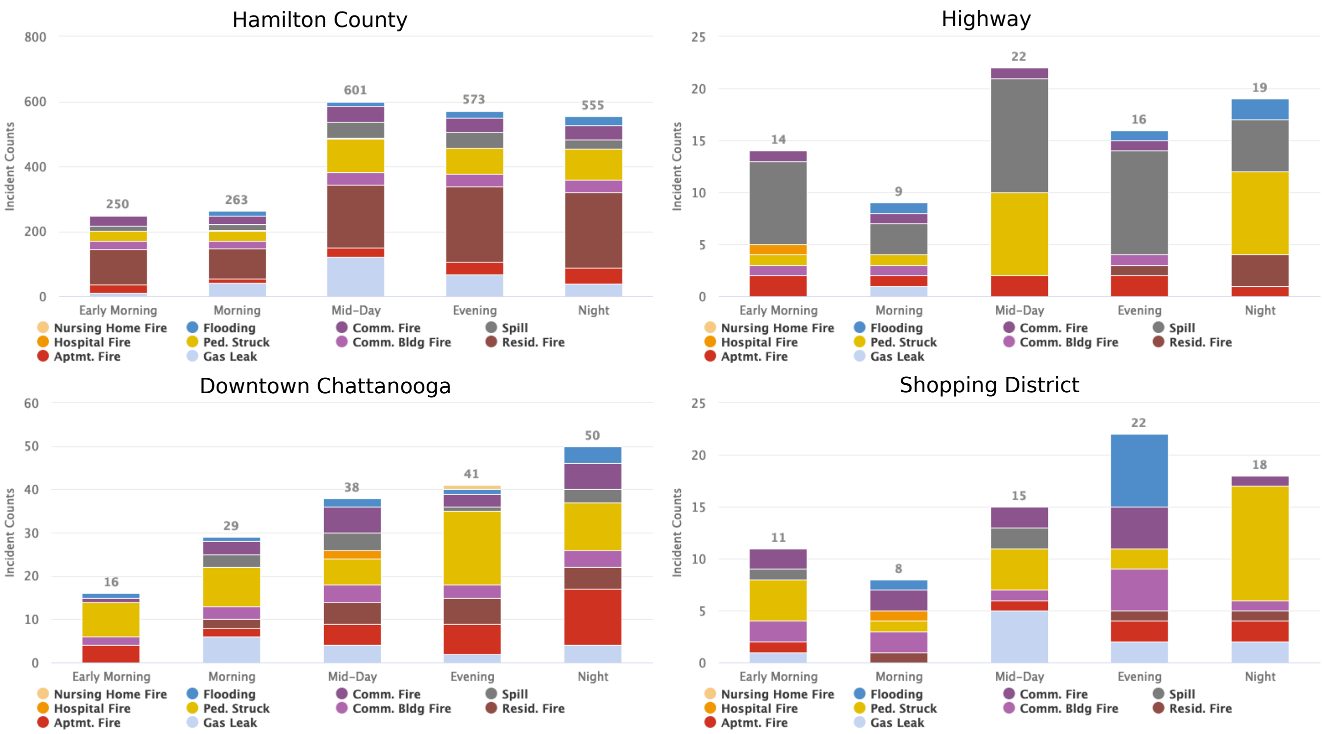


Figure 5. Bar charts for 911 *collateral* incidents in Hamilton County (top left), for the highways (top right), downtown (bottom left), and the shopping district (bottom right). The data ranges from October 1, 2018 to September 30, 2019. The distribution of incident types varies strongly between different areas.

behavior are residential fires, which are likely not related to traffic volumes.

The comparison of bar charts proves a much bigger difference between normalized and original count data. Based on traffic volumes, there should be almost no traffic incidents in the early morning and morning, and much fewer at night. This is mostly true for early morning incidents and nights, with the exception of traffic stops which are represented disproportionately. However, these are also the time frames during which driving under the influence is most likely to occur, which may explain this pattern.

Figure 6. Comparison of incident counts (left column) with normalized counts (right column) for 911 in the entire county. The heatmaps and bar charts highlight the expected incident counts based on traffic volumes (right) compared to the actual incident counts (left).

Tennessee Crashes by Traffic Condition

To represent different aspects of any given incident, we employ a sequence-sunburst chart Rodden (2019), an interactive multivariate visualization technique. Figure 7 presents four different views of a sunburst chart for the occurrence of traffic incidents and their associated attributes within a user-defined date range. By adopting a radial layout (shown in the middle), the sunburst chart is capable of visualizing the percentage of incidents that are hierarchically categorized by time of day when incident occurred (inner ring), type of crash (middle ring), and weather condition during the incident occurrence (outer ring). These categories are color-coded using a color ramp display in the legend (left bar). When a user uses their mouse cursor to highlight a specific arc component in any ring, which represent a composite category of incidents (e.g., property damage occurred in midday during a rainfall event), the name and percentage of that incident category are then displayed in the middle of the chart, as shown in the bottom row of Figure 7.

There is a separate set of colors for each ring. The general intent was to have darker, more saturated colors for more difficult or harmful circumstances and incidents (dark, rainy, injuries, etc.), and lighter or less saturated colors for

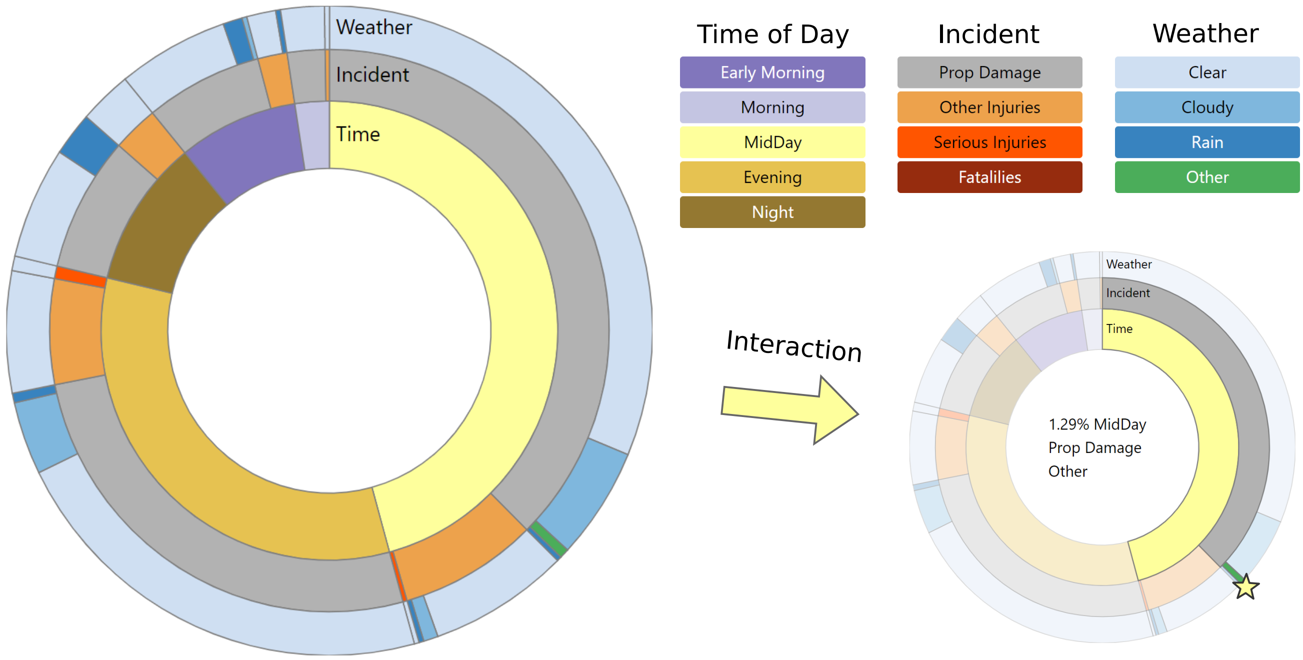


Figure 7. Multivariate visualization of Tennessee vehicle crash data for Hamilton County from October 1, 2018, to September 30, 2019. Each ring represents a different variable: time of day (inner), type of incident (middle), and weather (outer). The legend at the top provides colors for all possible values in the dataset. Interaction is demonstrated in the bottom right part of this Figure: upon mouse-over of one of the ring segments (marked by a star).

less harmful circumstances or events (clear sky, daytime, property damage, etc.).

- For the inner ring (Time of Day), we used a colormap from purple (beginning after midnight) to brown (ending at midnight). The colors are lighter during daylight hours and darker at night and in the early morning.
- For the middle ring (Type of Crash), we used the same colors as for the incidents plotted on the map.
- For the outer ring (Weather) we used a light blue for clear skies, dark blue for rain, and green, which is dissimilar, for “other” weather circumstances, which could include fog, snow, flooding, or other less common weather events.

For the given year of data, almost half of the recorded vehicle crashes happened during mid-day, and mid-day and evening crashes combined constitute about 80% of the overall incidents in the region.

Spatial Visualization

The main goal of spatial visualization is to provide context and location to the data. In this section, we present different types of spatial visualizations which work at different aggregation levels.

Incident Overview Maps

We have created different maps with different purposes. For a quick overview of hot spots, we have an overview map with a Kernel Density Function representing incident frequency.

The 911 dataset is based on the street address of each incident. For a large area such as a mall and its parking lot, multiple incidents will be plotted on a single point. To address this shortcoming and better represent the number of accidents, we disperse these stacked incidents around their original location. The incidents were dispersed with a 30 foot minimum distance between one another in order to be visually identified as unique events. An expanded dispersal pattern was used as it maintains the existing general spatial pattern when possible and defaults to encircling the original location of the stacked incidents with dispersed points when no pattern is present.

Incidents with Kernel Density Function We created a map of all vehicle-related 911 incident responses in the area, which is displayed in Figure 8. On top of a base layer, we plot a kernel density function (KDF) for the distribution of accidents. The colormap is very similar to the one used for the heatmap matrices.

The Kernel Density Function (KDF) map was created by integrating the open layers web map engine and the heatmap.js library developed by Patrick Wied Heatmap.js (2020). The KDF applies equal weight to individual incidents and provides a visual interface that allows users to explore the visualization with different radius and weights of each incident. The maps in Figure 8 provide a great first impression of hot spots; however, they cannot give an accurate impression of individual incidents. We address this by providing maps for the different incident types.

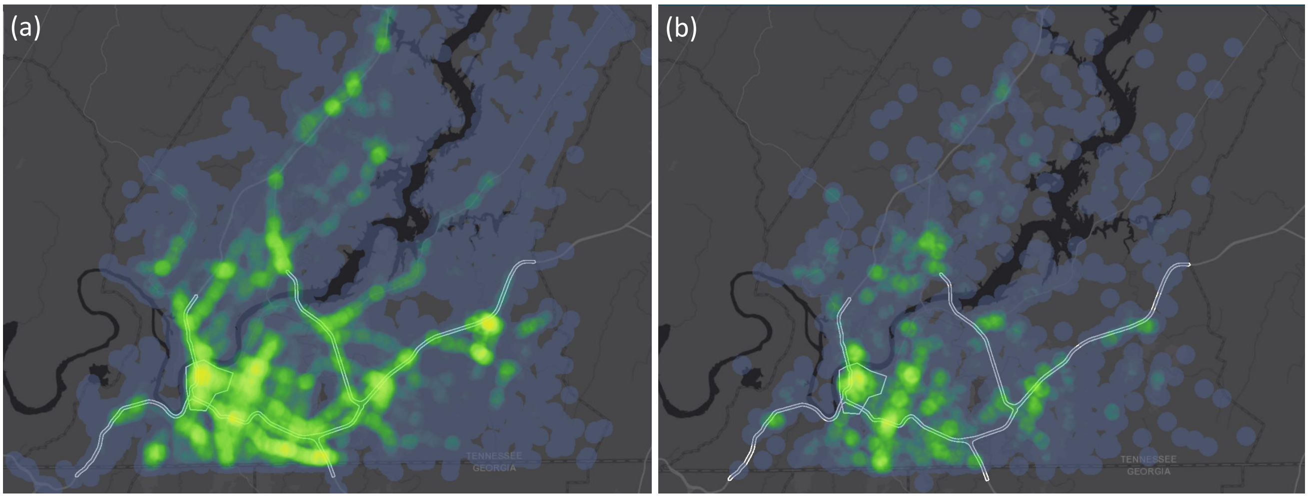


Figure 8. Overview maps for 911 traffic incidents (left) and collateral incidents (right) from October 1, 2018 to September 30, 2019. The map provides a semi-transparent KDF rendered over a base map. For additional context, the three study regions are rendered alongside the KDF.

Map of Traffic Stops As traffic stops make up over 60% of the total incidents, they would clutter the incidents map and are therefore featured separately. Figure 9 shows all traffic stops in the three study areas.

On the highways, there are some major clusters of traffic stops: US-27 and SR-153 both have much larger numbers of traffic stops than either of the two interstates. Near downtown, there are far fewer traffic stops on US-27, which may be due to ongoing construction and a lack of shoulders to stop on. I-24 has hot spots near exit 181, which has a lot of nearby businesses, and between exits 183 and the junction with I-75. I-75 has several clusters as well, including one by the Georgia border, one by the shopping district (exits 4-5), and near exit 9, which connects to several large warehouses.

For downtown, the largest cluster of traffic stops is in the section that features multiple restaurants and bars, especially along Market Street, one of the main arterials in that area. There are a number of other clusters in an area with a lot of apartment housing, near the baseball stadium, and in two small areas south of I-24, one of which is a school zone.

In the shopping district, there are a lot of traffic stops on Shallowford Road between Lee Highway and I-75, but there are many other hot spots in other areas of Shallowford Road and Gunbarrel Road. One can also see the very distinct cluster of traffic stops on I-75 near exit 4A.

Map of Traffic Incidents Speed plays an important role in fatality of crashes: at a speed of 17 mph, the risk of serious injury to a pedestrian is 10%. Less than doubling the speed to 33 mph increases the risk to 50% Fischer et al. (2015). This is reflected in the visualization: Figure 10 presents all traffic incidents in the different study areas.

The density of incidents without injuries (light green) is very high in areas such as the shopping district's parking lots as well as a stretch of US-27 west of downtown Chattanooga which has been under construction with a lower speed limit for the entire year of data Tennessee Department of Transportation (2019).

Incident severity is more mixed on busy local roads like Shallowford Road in the shopping district or Market Street in downtown Chattanooga. On highways, a much larger percentage of incidents involves injuries. The highest number of vehicle entrapments and vehicle fires happen in areas where highway traffic is generally faster, such as I-75.

In downtown Chattanooga, there is one intersection which has an unusually high number of incidents with entrapment. This intersection is adjacent to a hospital, which could have led to more distracted and rushed driving in this area.

Overall, the biggest hot spots for traffic incidents are the junction of I-24 and I-75, the construction zone on US-27, and the entire stretch of I-24 between US-27 and I-75.

Throughout the entire region, there are many broken-down vehicles. This is very apparent in the highway sections that have lower numbers of crashes.

Map of Collateral Incidents Collateral incidents along highways are relatively rare. In addition to various building fires that are close enough to the highway to fall within the buffer, there are a handful of spills, several struck pedestrians, and some road flooding. The spills are more concentrated on the busy stretch of I-24 than elsewhere. For pedestrians, the most dangerous areas are on the highway itself.

Walkable areas with many restaurants or bars and areas with nearby bus stops or other points of interest, such as hospitals, have high numbers of struck pedestrians. Many of these areas have traffic lights with pedestrian phases, where increased pedestrian traffic may be expected. Downtown, out of 51 total incidents of struck pedestrians, 36 incidents (71%) happened within about 0.1 miles (downtown city block size) of a bus stop. In the mall area, 15 out of 22 (68%) of struck pedestrian incidents happened in parking lots.

Finally, there is an unusually high number of apartment fires west of downtown. Though this is an area that does indeed have a lot of apartment housing, a detailed look at the source data shows a wide spread of dates and times for these incidents.

In the left part of the Figure 12, we visualize speed limits and volumes for each road in the study area Chattanooga Department of Public Works (2020). Speed is represented by a colormap from red ($\leq$ 25 mph) to purple ($\geq$ 50 mph). One can clearly see the interstate and state route which run

Figure 9. Traffic stops along highways (top), in downtown Chattanooga (bottom left), and in the shopping district (bottom right) from October 1, 2018 to September 30, 2019. To better represent the number of incidents, we scatter the individual incidents around their locations. Review note: Figure in separate file due to size.

Figure 10. Traffic incidents along highways (top), in downtown Chattanooga (bottom left), and in the shopping district (bottom right) from October 1, 2018, to September 30, 2019. To better represent the number of incidents, we scatter the individual incidents around their locations. Review note: Figure in separate file due to size.

Figure 11. Collateral incidents along highways (top), in downtown Chattanooga (bottom left), and in the shopping district (bottom right) from October 1, 2018, to September 30, 2019. To better represent the number of incidents, we scatter the individual incidents around their locations.

through the city (teal, blue and purple). The walkable part of downtown has speed limits at 35 or below as seen by the yellow and orange lines. On the outskirts of this area, speed limits increase to 40 miles per hour. Volume is represented by line thickness. For this purpose we used the baseline volumes for each link, and we applied a non-linear scaling function $s = \sqrt{0.1 \cdot v}$. As one can see, the roads that were labeled as "busy, relatively fast local traffic" in Figure 10 show similar behavior. They are among the boldest lines on the map (indicating high volume, which is only surpassed by state route and highway traffic), and they are rendered bright yellow (speed limits of 40-45 mph).

In the right part of Figure 12, we highlight incidents of struck pedestrians which happened in the vicinity of bus stops (gold) compared to those who did not (lavender). To gain this insight, we applied a distance function of about one city block (0.1 miles) around each bus stop Chattanooga Area Regional Transportation Authority (2020). As one can see when comparing this with the Figure on the left, and with Figure 11, the proportion of struck pedestrians is highest in walkable areas with many bus stops, despite the rather low speed limits.

Interactive Clustering of Motor Vehicle Crashes Geo-visualizations are commonly used to display the locations of traffic incidents Kmet et al. (2019); Galka (2013). However, many existing accidents mapping techniques are limited in that they are only effective in geospatially presenting incident reports at a smaller spatial scale (e.g., corridor and intersection) or temporal scale (e.g., daily and monthly). These visualizations may become cluttered as the size of the incident dataset increases. When a traffic incident map becomes cluttered, it is difficult to extract useful information, especially for a regional scale visualization that displays car crashes for a full year. Many past studies utilize pattern abstraction (e.g., heatmap and kernel density) to resolve the visual-cluttering challenge and to visualize traffic incident locations for a large area at a longer time-span (e.g., a full year). However, these applications are not adaptive enough to also show the incident locations at finer spatial resolution levels Hilton et al. (2011); Plug et al. (2011); Xie and Yan (2008). In this regard, there is a pressing need for developing adaptive mapping techniques to explore both the spatial and temporal patterns in large datasets from the regional scale to the intersection scale. In order to find a middle ground between purely aggregated visuals, such as the KDF in Figure 8, and plotting individual points, such as in Figures 8, 9, and 10, we devised an adaptive geo-visualization that integrates interactive cluster markers with pie charts to display the locations and types of traffic incidents at multiple spatial scales. At the regional level, our technique groups the locations of traffic incidents into clusters based on map zoom levels, distances between individual incident locations, and generates pie charts to represent the composition of incident types within different clusters. Figure 11 demonstrate our visual interface which provides a regional scale overview of the Chattanooga metropolitan area. In the side bar, we provide a control panel that enables users to change (1) the radius to scale pie chart sizes, (2) the time period of incidents (currently four quarters from October 2018 to September 2019), and (3) the basemap layer. A legend of incident types is also included to show the color codes used in the pie charts. As shown in the map, the biggest clusters occur downtown, at the I-24/I-75 interstate junction, and at the shopping district study area. All other clusters are spread out along the highway system. Most incidents on local streets are represented as individual points on the map.

A user can hover over each pie chart cluster to see the bounding box of its pertaining incidents (depicted as the white boundary in the left map in Figure 13) and exact counts for each incident type in the pie chart through a modified map legend. In this setting, our level-of-detail technique preserves the location and properties of individual accidents while hiding this excessive information at the regional level to avoid visual cluttering. As the user zooms in to a smaller spatial scale (e.g., corridor or intersection) through the web map, the interactive visualization displays either (1) the individual locations of accidents color coded by incident types or (2) pie charts that represent a cluster of incidents too dense to be displayed at the present zoom level.

Figure 13 illustrates the user workflow to interact with this visualization. In the left portion of the Figure, we show the user interface: we select a date range and base map (top left), and use the interactive legend (bottom left) to select what incident types to show. Check boxes enable a quick selection of all traffic or collateral incidents. In addition, clicking on specific incident types can activate or deactivate them on demand. In this example, we have deactivated Gas Leaks, as seen by the near-transparent color. In the map view, we have selected a cluster near the I-24 and I-75 junction. By hovering over the pie chart, we are able to visualize a bounding polygon of all incidents in this cluster. When zooming into this area further, the selected pie chart breaks up into multiple smaller pie charts to give a higher spatial

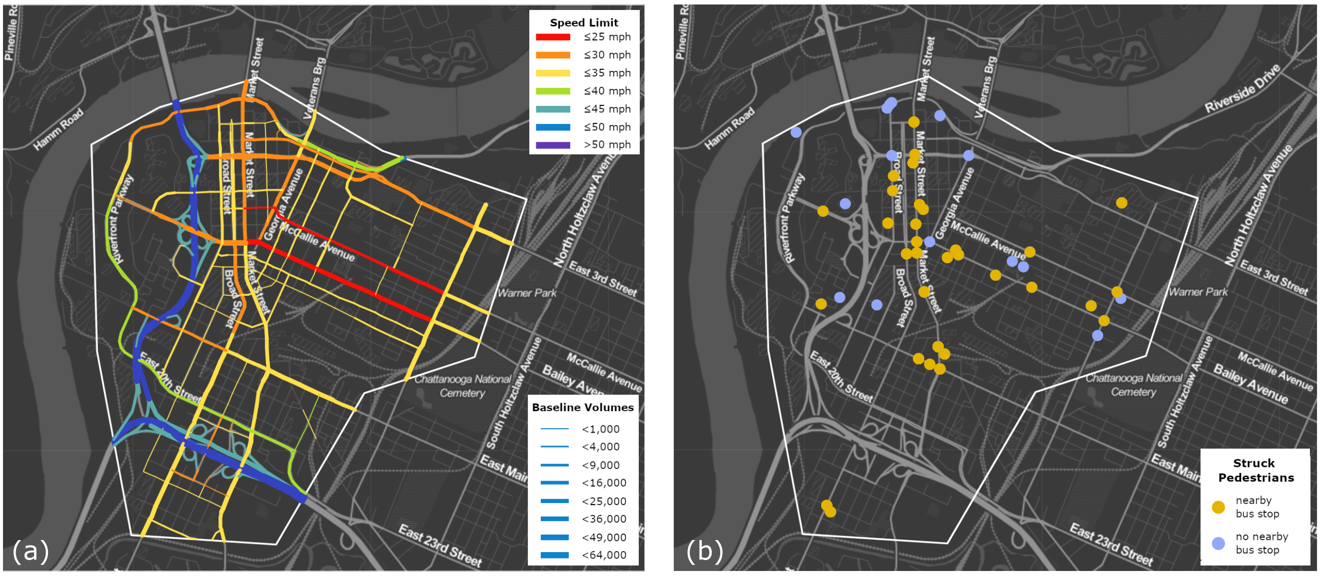


Figure 12. Additional visualizations of downtown. The visualization on the left displays speeds (color) and volumes (line thickness) for each road segment. The visualization on the right displays incidents of struck pedestrians near bus stops (gold) and other incidents of struck pedestrians (lavender).

level of detail, as shown in the right portion of the Figure. At maximum zoom level, users can see individual incident locations, as demonstrated in the close-up.

Through the visual interface, it can be concluded that viewing the interstate junction up close enables more insights about traffic safety in this area. While all crashes on I-75 south of the junction are incidents without injuries, this is not true for the crashes on I-24, which include several cases of entrapment, making this highway the one with most severe incidents in this area. In the Northeast section of the map, there are two clusters of incidents near SR-153, both consisting entirely of crashes with injuries. This area is therefore one of the two more dangerous sections of highway in the region.

Textured Tile Calendar

Finally, we would like to introduce the textured tile *calendar view* as a tool that provides an combination of a time series view and a geospatial view, which is presented in Figure 14. For each day of data, we create a textured tile, which contains an abstracted geospatial view of incident counts. These tiles are arranged in the shape of a calendar with one column per week. Between months, there is an offset to improve navigation. For each tile, we aggregate incidents in Hamilton County to a 7×7 grid. We then render the aggregate as a texture on one tile for each day. The date for a tile is displayed when hovering over a tile. With the calendar view, on can get a sense of patterns in the data. For instance, in the given example, Sundays and Mondays have very few incidents in December and January.

In addition to providing a high-level overview of incidents, this calendar view also serves as a navigational tool to further explore the underlying data. When a user selects a date, a large version of the tile appears in a *tile view*, with additional details (incident counts) for each cell of the grid. The color range (from muted dark blue to bright yellow/orange) is scaled to accommodate the maximum number of incidents for the full year of cell-wise aggregates. Low incident counts are a blue, incident counts are yellow, with values in between shown in various shades of teal, green, and lime. Cells without incidents (or without data) are displayed in gray. Examples of this view are presented in Figure 15.

For the calendar view users can pick the dataset they would like to explore from a drop-down menu. We have included the full 911 and E-TRIMS data, as well as extracts by incident type. Furthermore, the tile view includes an option for pair-wise comparisons of this data for a selected date. Figure 15 demonstrates the comparison between an extract of traffic incident data from the 911 Report and Tennessee Crash data for August 21, 2019. In most cells of the tile, we can see that the 911 dataset has more incidents than the Tennessee Crash dataset, however, in a few cases, the opposite is true (e.g. for the lime green 40 vs the medium green 39).

Conclusion

We have demonstrated a variety of different approaches to visualizing traffic safety data at adaptive scale for the Chattanooga Metropolitan Area. Through studying both abstract and spatial visualizations, we have been able to obtain insights about incident hot spots and pedestrian safety risks. We have furthermore discovered unusual incidents such as flooding through heavy rainfall, which caused a large number of incident reports within a short time frame.

Presently, most of the abstract visualizations are integrated in a web platform, a *digital twin* for traffic; however, the majority of these visualizations rely on static data. We plan to provide a selection of regions which users can choose from interactively, and we will furthermore include options for date range selection to facilitate the kind of insights we have been able to obtain by manually changing the data.

Finally, some of the spatial visualizations are currently prototypes which we are hoping to integrate into the web platform soon. This integration will include the same kind of

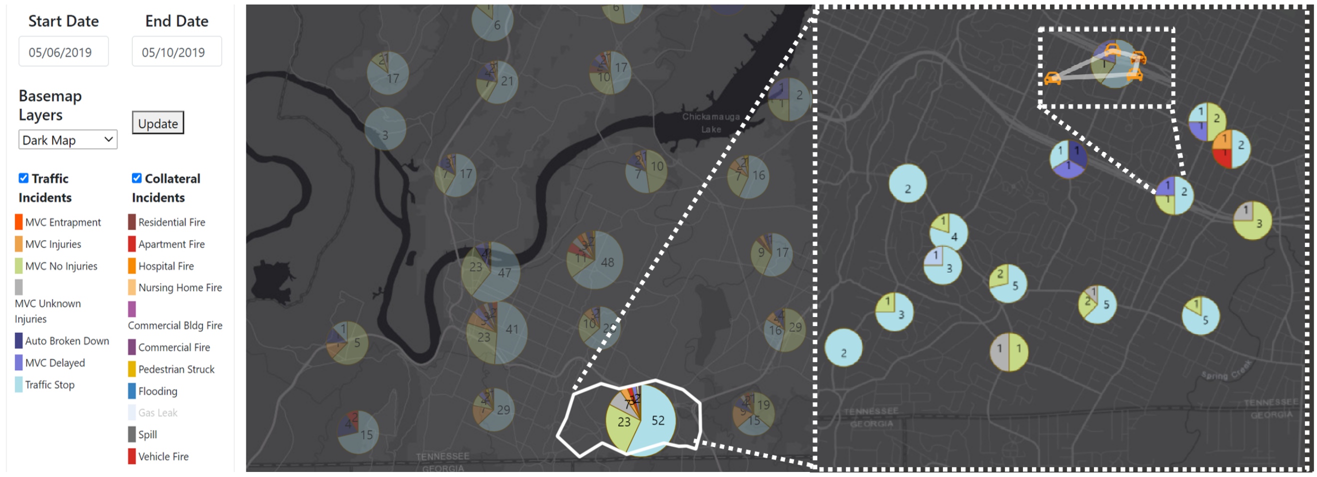


Figure 13. Clustering of vehicle crashes from July 1, 2019, to September 30, 2019. The left image demonstrates interactions: when hovering over a cluster, all other clusters become transparent. Additionally, the individual numbers of incidents per type are provided in the legend, and the bounding polygon is rendered on top. In the right image, we show a close-up of incidents that occurred near the interstate junction of I-24 and I-75.

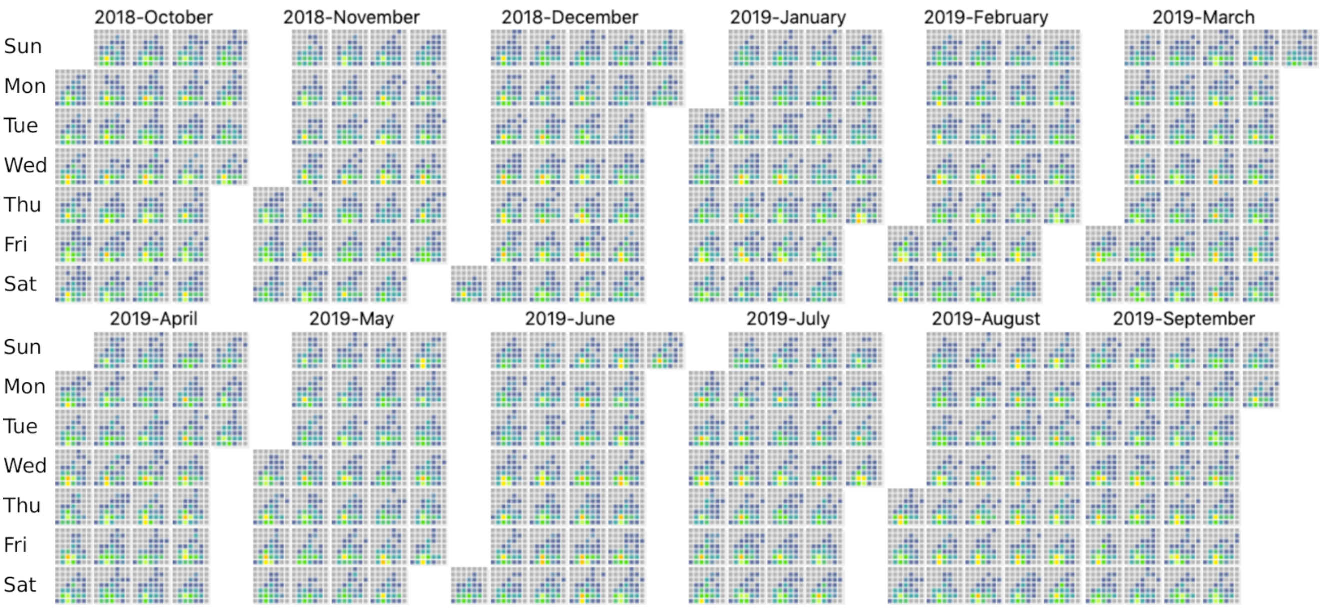


Figure 14. The calendar view shows the incident distribution of a selected dataset over time. Each day's incidents are displayed as a tile on the calendar.

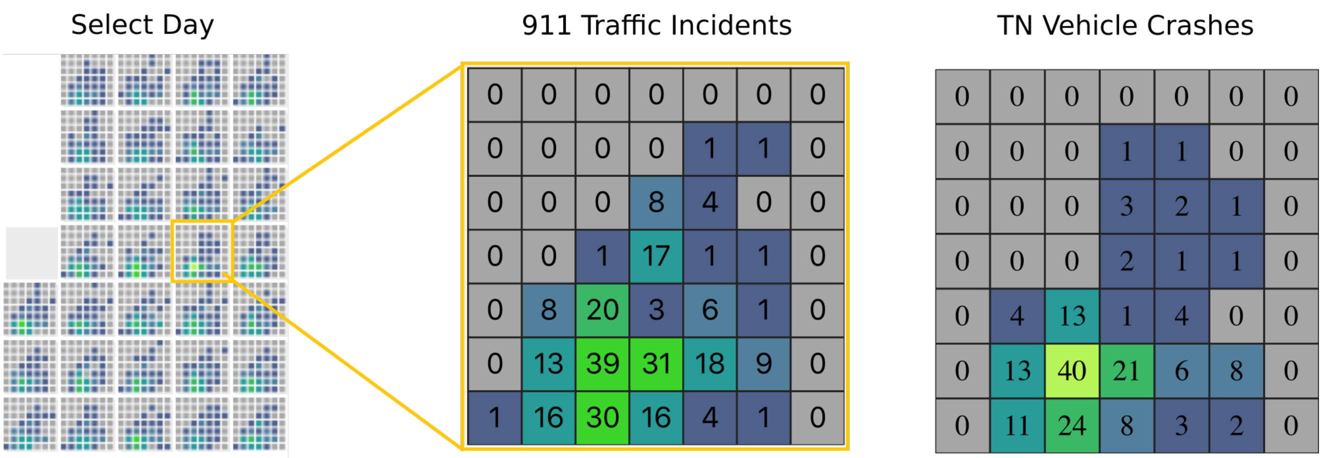


Figure 15. When the tile for a day is selected in the calendar view, a detailed version of the tile is displayed next to the calendar. The user can then choose to compare this detailed view with another dataset that is available for the same day. In this example, we have selected August 21, 2019 (left), and we have chosen to compare crashes from the Hamilton County 911 dataset (middle) with corresponding crashes from the Tennessee Vehicle Crashes dataset (right).

region selection as the abstract visualizations, with boundary polygons to restrict the query to incidents from a specific area. Furthermore, we would like to integrate filters to enable users to view specific incident types they are interested in. This would be a useful feature to support data exploration.

Acknowledgements

We thank the Chattanooga Department of Transportation for their guidance on choosing an arterial as well as providing traffic flow data, Tennessee Department of Transportation for providing the Tennessee crash data, and Hamilton County 911 for providing incident response data, Chattanooga Department of Public Works for providing the road network, and the Chattanooga Area Regional Transportation Authority for providing bus stop locations. Furthermore, we we would like to thank the Department of Energy Vehicle Technologies office for funding this work. Finally, we would like to thank Chieh (Ross) Wang and Tim LaClair for their feedback on this manuscript.

References

American Transportation Research Institute (ATRI) (2018) 2018 top 100 truck bottleneck list. Technical Report 2018, American Transportation Research Institute.

Anwar A, Nagel T and Ratti C (2014) Traffic origins: A simple visualization technique to support traffic incident analysis. In: *2014 IEEE Pacific Visualization Symposium*. IEEE, pp. 316–319.

Chattanooga Area Regional Transportation Authority (2020) Carta. https://www.chattadata.org/Transportation/CARTA-Stops/ebrc-i5v8.

Chattanooga Department of Public Works (2020) Chattanooga roadways. https://www.chattadata.org/dataset/Chattanooga-Roadways/mw3f-d2mz.

Chen W, Guo F and Wang FY (2015) A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems* 16: 2970–2984. DOI:10.1109/TITS.2015.2436897.

Esri and HERE and Garmin and OpenStreetMap contributors and the GIS user community (2020) Dark gray canvas reference. https://services.arcgisonline.com/ArcGIS/rest/services/Canvas/World_Dark_Gray_Reference/MapServer.

Fancello G and Fadda P (2002) Shopping centre traffic impact: A traffic forecasting model for medium sized towns in italy. ISBN 978-0-7844-0630-4, pp. 650–657. DOI:10.1061/40630(255)92.

Fischer P, Craft R, Harper V, Joyce L, Marshall JW, McPherson T, Thomas I and Williams AF (2015) Everyone walks. understanding & addressing pedestrian safety. https://www.ghsa.org/sites/default/files/2016-12/FINAL_sfped_0.pdf. Accessed November 7, 2019.

Galka M (2013) Mapping 10 years of fatal traffic accidents. *New York: Metrocosm* .

Hansen CD and Johnson CR (2004) *The Visualization Handbook*. Elsevier.

Hasan MR, Uddin M, Ahmed I, Das P, Uddin M and Hasan T (2012) A comprehensive study on trip attraction rates of shopping centers in dhanmondi area. *International Journal of Civil & Environmental Engineering* 12.

Heatmapjs (2020) Patrick wied. https://www.patrick-wied.at/static/heatmapjs/.

Hilton BN, Horan TA, Burkhard R and Schooley B (2011) Saferoadmaps: Communication of location and density of traffic fatalities through spatial visualization and heat map analysis. *Information Visualization* 10(1): 82–96.

Kappes MS, Keiler M, von Elverfeldt K and Glade T (2012) Challenges of analyzing multi-hazard risk: a review. *Natural hazards* 64(2): 1925–1958.

Kmet R, Dvorak Z and Kvet M (2019) Map of traffic accidents. *Transportation Research Procedia* 40: 1418 – 1425. DOI:https://doi.org/10.1016/j.trpro.2019.07.196. URL http://www.sciencedirect.com/science/article/pii/S2352146519303667. TRANSCOM 2019 13th International Scientific Conference on Sustainable, Modern and Safe Transport.

Komendantova N, Mrzyglocki R, Mignan A, Khazai B, Wenzel F, Patt A and Fleming K (2014) Multi-hazard and multi-risk decision-support tools as a part of participatory risk governance: Feedback from civil protection stakeholders. *International Journal of disaster risk reduction* 8: 50–67.

Kristoffersson I, Daly A and Algers S (2018) Modelling the attraction of travel to shopping destinations in large-scale modelling. *Transport Policy* 68(C): 52–62. DOI:10.1016/j.tranpol.2018.04. URL https://ideas.repec.org/a/eee/trapol/v68y2018icp52-62.html.

Li L, Zhu L and Sui DZ (2007) A gis-based bayesian approach for analyzing spatial–temporal patterns of intra-city motor vehicle crashes. *Journal of Transport Geography* 15(4): 274–285.

Lin PS, Kourtellis A, Zhang Y, Guo R, Bialkowska-Jelinska E et al. (2017) Application of demographic analysis to pedestrian safety. Technical report, University of South Florida. Center for Urban Transportation Research.

National Highway Traffic Safety Administration (2001) Traffic safety facts, 2000: Overview. Technical report, DOT HS 809 329. Washington, DC: Department of Transportation, National Highway Traffic Safety Administration.

Pack ML, Wongsuphasawat K, VanDaniker M and Filippova D (2009) Ice–visual analytics for transportation incident datasets. In: *2009 IEEE International Conference on Information Reuse & Integration*. IEEE, pp. 200–205.

Pare M (2019) Flooding from wednesday storm closes food court in hamilton place mall in chattanooga. https://www.timesfreepress.com/news/local/story/2019/jul/31/flooding-closes-food-court-hamilton-place-mall-chat 500195/. Accessed November 21, 2019.

Plug C, Xia JC and Caulfield C (2011) Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention* 43(6): 1937–1946.

Retting R and Consulting SS (2017) Pedestrian traffic fatalities by state. https://www.ghsa.org/sites/default/files/2018-02/pedestrians18.pdf. Accessed November 7, 2019.

Rodden K (2019) Sequences sunburst. https://gist.github.com/kerryrodden/7090426. Accessed November 21, 2019.

Sevigny A (2020) Traffic flow by day of week and time. https://www.chattadata.org/Transportation/Traffic-Flow-by-Day-of-Week-and-Time/

v868-x5mt/.

Sobral T, Galvão T and Borges J (2019) Visualization of urban mobility data from intelligent transportation systems. *Sensors* 19(2): 332.

Tennessee Department of Transportation (2019) Us 27 reconstruction in downtown chattanooga including widening the olgiati bridge over the tennessee river. https://www.tn.gov/tdot/projects/region-2/us-27-reconstruction-in-downtown-chattanooga.html. Accessed November 21, 2019.

Wang J, Shen Y and Khattak AM (2010) Visualization of traffic accidents.

Xie Z and Yan J (2008) Kernel density estimation of traffic accidents in a network space. *Computers, environment and urban systems* 32(5): 396–406.