Collaborative Sciences Center for ROAD SAFETY

FINAL REPORT

An enhanced systemic approach to road safety

September 2019

Offer Grembek, PhD
Co-Director
Safe Transportation and Education Center
University of California, Berkeley

Joy Pasquet
Graduate Student Researcher
Safe Transportation and Education Center
University of California, Berkeley

Catalina Vanoli
Gradiate Student Researcher
Safe Transportation and Education Center
University of California, Berkeley

THE UNIVERSITY OF TENNESSEE KNOXVILLE · Berkeley UNIVERSITY OF CALIFORNIA · FAU FLORIDA ATLANTIC UNIVERSITY · Duke UNIVERSITY · THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

www.roadsafety.unc.edu

U.S. DOT Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgement of Sponsorship

This project was supported by the Collaborative Sciences Center for Road Safety, www.roadsafety.unc.edu, a U.S. Department of Transportation National University Transportation Center promoting safety.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. CSCRS-R2	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle: An Enhanced Systemic Approach to Road Safety		**5. Report Date** September 2019 **6. Performing Organization Code**
7. Author(s) Offer Grembek, PhD https://orcid.org/0000-0003-1869-9457 Joy Pasquet Catalina Vanoli		**8. Performing Organization Report No.**
9. Performing Organization Name and Address University of California, Berkeley, Safe Transportation Research & Education Center, Berkeley, CA		**10. Work Unit No.** **11. Contract or Grant No.** Collaborative Sciences Center forRoad Safety (Grant #:69A3551747113)
12. Sponsoring Agency Name and Address		**13. Type of Report and Period Covered** Final Report (March 2017 –September 2019) **14. Sponsoring Agency Code** -

15. Supplementary Notes

Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. Abstract

The traditional, transportation safety management approach involves the identification of crash hotspots, in which public agencies prioritize locations eligible for safety improvements based on historical collision concentrations. This report presents an enhanced systemic approach which consists of targeting blanket improvements at sites across a road network based on specific roadway features that are associated with a particular crash type. The systemic approach uses historical crash data to identify the type of roadways that suffer from recurring safety concerns, designating it as reactive to data, but also provides a mechanism to make improvements to sites that have not experienced many—or any—crashes in a proactive manner. The systemic approach is a flexible, data-driven methodology that aims to identify recurring safety concerns by identifying the crash profiles that are associated with certain roadway features. The analysis takes the form of a transparent systemic crash matrix that shows what types of crashes occur on what types of facilities, with rows representing crash characteristics and columns corresponding to facility types. Using such a matrix provides agencies with a snapshot of systemic problems within their networks, which is both easy to assemble and to interpret, thus overcoming potential barriers to changes in road safety management due to limited institutional capacity or financial means. The framework set by the systemic approach is flexible enough to allow agencies with varying degrees of data availability to implement it—regardless of the level of performance their data management systems—and with different safety priorities.

17. Key Words Traffic Safety; Systemic Approach; Safe Systems; Vision Zero; Crashes; Countermeasures; Transportation Safety Management.		**18. Distribution Statement** No restrictions. This document is available through the Collaborative Sciences Center for Road Safety (roadsafety.unc.edu), Chapel Hill, NC.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 51	**22. Price**

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Abstract

The traditional, transportation safety management approach involves the identification of crash hotspots, in which public agencies prioritize locations eligible for safety improvements based on historical collision concentrations. This report presents an enhanced systemic approach which consists of targeting blanket improvements at sites across a road network based on specific roadway features that are associated with a particular crash type. The systemic approach uses historical crash data to identify the type of roadways that suffer from recurring safety concerns, designating it as reactive to data, but also provides a mechanism to make improvements to sites that have not experienced many—or any—crashes in a proactive manner. The systemic approach is a flexible, data-driven methodology that aims to identify recurring safety concerns by identifying the crash profiles that are associated with certain roadway features. The analysis takes the form of a transparent systemic crash matrix that shows what types of crashes occur on what types of facilities, with rows representing crash characteristics and columns corresponding to facility types. Using such a matrix provides agencies with a snapshot of systemic problems within their networks, which is both easy to assemble and to interpret, thus overcoming potential barriers to changes in road safety management due to limited institutional capacity or financial means. The framework set by the systemic approach allows agencies with varying degrees of data availability to implement it—regardless of the level of performance their data management systems—and with different safety priorities.

Executive Summary

The systemic approach is a desirable methodology to transition from existing practices in road safety to approaches such as safe systems that fundamentally change the way road systems are conceived and designed to ensure that none of their users are severely injured. At a time when local and national governments throughout the world are committing to Vision Zero to curb traffic deaths and severe injuries, it represents an opportunity to smoothly divert from reactive approaches while building on the existing expertise of establishes road safety stakeholders. The systemic approach stands at a midpoint between traditional hotspot and corridor approaches, and emerging fully proactive approaches such as safe systems, which is why it constitutes an exceptional opportunity for road safety agencies. Engaging in ambitious comprehensive approaches such as safe systems may seem out of reach to some of them, especially in a financially-constrained environment or if there is a lack of interagency coordination. The systemic approach represents a first step in the right direction, which is ready for implementation with existing tools and yet promises more safety benefits than pure reactive schemes.

The systemic approach is a flexible, data-driven methodology that aims to identify recurring safety concerns within a road network, by identifying the crash profiles that are associated with certain roadway features. The analysis takes the form of a transparent systemic crash matrix that shows what types of crashes occur on what types of facilities, with rows representing crash characteristics and columns corresponding to facility types. Using such a matrix provides agencies with a snapshot of systemic problems within their networks, which is both easy to assemble and to interpret, thus overcoming potential barriers to changes in road safety management due to limited institutional capacity or financial means. How cells get populated is left to the discretion of the agency implementing the safety approach—the matrix can display crash counts, crash frequency, crash density or other safety indicators—however, regardless, these cells define crash profiles and contain aggregate information on the crashes that occurred. The cells with the highest value constitute so-called *systemic hotspots*, i.e., systemic challenges on the roadway network where a particular crash profile is consistently associated with a particular type of road infrastructure, which then allows to target blanket improvements across an entire facility type.

The framework set by the systemic approach allows agencies with varying degrees of data availability to implement it—regardless of the level of performance their data management systems—and with different safety priorities. It is not specific to a particular database and lays out the main steps to be undertaken by any agency that wishes to conduct systemic improvements on its road network. Adequate data collection on past collisions and detailed infrastructure features, followed by well-documented data cleaning, ensures that the systemic matrix is generated with the right information. There is no single answer as to which matrix structure should be adopted, as the decision-making on how acceptable the structure of a systemic matrix is inherently data specific. This process both relies on road safety expertise to evaluate different options and understand the implications of certain crash profiles, but also on quantitative comparisons between structures in order to choose the final one. The validity of the outcome of the systemic approach should therefore not be a concern for agencies. Additionally, the way by which the systemic matrix gets populated determines what priorities will be captured by the resulting systemic hotspots, which lets each agency choose their own priorities (in terms of modes affected by collisions, the severity of the crashes targeted, and other factors).

Developing a systemic crash matrix is only the first step on the way to systemic road safety. Once systemic hotspots are identified, engineering countermeasures can be selected to address their crash profiles across the corresponding facility types. Non-engineering countermeasures, such as educational and enforcement opportunities should also be investigated as various means to gradually achieving truly safe systems. In light of all the opportunities for future research that could build on or support the further improvement of the systemic approach, road safety stakeholders should consider forging new innovative partnerships to best implement this approach to flag high-priority systemic safety concerns, and complement this approach with other initiatives that would help reduce traffic deaths and injuries in the near future.

Table of Contents

An enhanced systemic approach to road safety 1
U.S. DOT Disclaimer 2
Acknowledgement of Sponsorship 2
Abstract 4
Executive Summary 4
Table of Contents 6
List of Figures 6
List of Tables 6
Introduction 8
Background 9
Literature Review 9
Beginnings of the systemic approach 9
Measuring safety using the systemic approach: choosing the right safety indicators 10
The systemic matrix scheme 12
Methodology for building a systemic matrix 12
Data collection requirements 12
Crash data 14

List of Figures

Figure 1: Continuum that ranges from fully reactive to truly proactive approaches

Figure 2. Quantitative quantile criteria for systemic hotspots

Figure 3. Quantitative quantile criteria for systemic hotspots

Figure A-1: Left portion of the non-PDO auto-only matrix for intersection crashes in California (2010-2011)

Figure A-2: Right portion of the non-PDO auto-only matrix for intersection crashes in California (2010-2011)

Figure A-3: Left portion of the pedestrian-involved matrix for intersection crashes in California (2010-2011)

Figure A-4: Right portion of the pedestrian-involved matrix for intersection crashes in California (2010-2011)

Figure A-5: Bicycle-involved matrix for intersection crashes in California (2010-2011)

List of Tables

Table 1. Minimum crash data requirements for the systemic matrix.

Table 2. Minimum infrastructure data requirements for the systemic matrix.

Table 3. Advantages and disadvantages of using crash counts or ratios in the systemic matrix

Table 4. Eligible variables for the systemic matrix rows

Table 5. Eligible variables for the systemic matrix columns

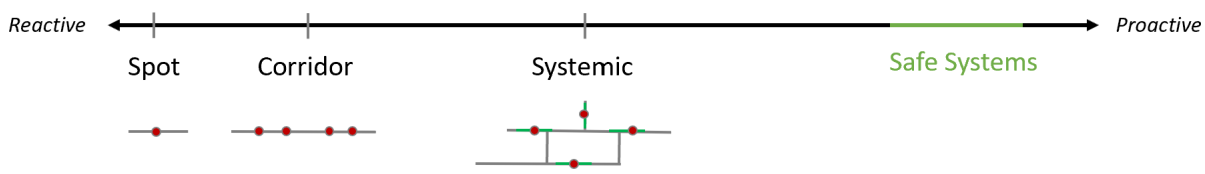
Table 6. Quality assessment of the matrix structure

Introduction

Transportation safety professionals strive to build safe systems within which no road user can be severely or fatality injured. To design such a system, it is necessary to utilize all the core protective opportunities provided by the system. When insufficient efforts are dedicated to preventing such catastrophic outcomes, road crashes continue to occur. A total of 34,247 fatal crashes and were reported in the United States in 2017 (NHTSA, 2019). Globally, there are over 1.3 million fatalities annually (Welle et al, 2018). Each of these crashes typically involves a series of undesirable events that result in an instantaneous and violent transfer of kinetic energy occurring in a specific time-space location. The fact that these preventable outcomes occur in specific sites, warrants investigation of what went wrong and the changes that need to be implemented to prevent future crashes at the same location. While such spot efforts are a necessity and can save lives, such an approach addresses the problem in a reactive manner and on a very small scale. Moreover, the shortage of data to sufficiently reconstruct the events that led to the crash cannot facilitate considerations across all the core protective opportunities.

Current road safety management practices can be assessed using a continuum that ranges from fully reactive to truly proactive approaches as shown in Figure 1 below.

Figure 1: Continuum that ranges from fully reactive to truly proactive approaches



The traditional, dominant approach involves the identification of safety hotspots, in which public agencies prioritize locations eligible for safety improvements based on historical collision concentrations. This focus constitutes a reactive approach, where the possibility of a safety improvement for a specific site is tied to the previous occurrence of crashes at that location. Corridor approaches are slightly less reactive since they result in safety improvements along an entire corridor, which include some high collision hotspots as well as locations with lower crash concentrations. Conversely, new approaches such as “Vision Zero” and “safe systems” that attempt to prevent the very possibility of severe crashes anticipate the occurrence of crashes and target improvements at locations regardless of their historical collision profile (OECD, 2008). The systemic approach is found between these two extremes, involving both reactive and proactive components. This approach consists of targeting blanket improvements at sites across a road network based on specific roadway features that are associated with a particular crash type. The systemic approach uses historical crash data to identify the type of roadways that suffer from recurring safety concerns, designating it as a partly reactive approach. However, the fact that it provides a mechanism to also make improvements to sites that have not experienced many—or any—crashes designates it as a partly *proactive* approach. The systemic approach is typically used in parallel to the hotspot approach and is considered a complement rather than an alternative.

The goal of the present study was to develop a method for agencies to conduct systemic safety analysis across modes at a scalable area, using a matrix structure to model these relationships between infrastructure features and crash characteristics. The main outcomes include a method for developing a systemic safety matrix across different transportation modes, along with a set of possible engineering and non-engineering countermeasures;

and a sample systemic matrix resulting from that method, which could be used to identify systemic safety concerns that would benefit from blanket improvements across a desired area to support system-wide safety improvements.

The structure of the present report is as follows: Section 1 establishes the background of the systemic safety approach, by first framing the broader problem of how to quantify road safety, then sets the academic context for the systemic approach through a literature review, followed by an update of the state of implementation of this approach worldwide and in the United States in particular, and finally an introduction of the practical framework applied by the systemic approach in this study, i.e., the systemic matrix. Section 2 investigates the actual methodology for building a systemic safety matrix, starting with the definition of data needs, then listing the steps to prepare the available data for matrix compilation, based on prior experience with HSIS data, followed by an explanation of the methodology to generate systemic matrices and assessment of the quality of their structure, with examples of safety matrices for the state of California, and lastly a summary of the limitations to the methodology. Finally, the concluding section summarizes the main outcomes of the methodology as well as its shortcomings and identify opportunities for further research.

Background

Literature Review

Beginnings of the systemic approach

The systemic approach to road safety originated at the intersection of two distinct strategies of road safety management that have emerged over the past two decades—the traditional, reactive approach, and the proactive approach.

The first is the traditional approach, for which sites with a higher than expected occurrence of crashes are identified. Appropriate countermeasures are then adopted for these specific sites, which are commonly called hotspots. Whether the high-crash locations are isolated from one another in a *spot approach* or considered along corridors with recurring safety concerns in a *corridor approach*, both schemes utilize a *reactive* rationale. This type of approach is problematic because addressing safety issues requires waiting for a crash occurrence, and in addition because underreporting issues can negatively impact the accuracy of the data upon which safety analyses are based. Indeed, not all crashes are reported, and not all those that are reported can be found in a single database—police, hospitals, road administrations, and insurance companies each have their own reporting systems and cannot individually capture all previous crashes (PIARC, 2013). Furthermore, in some contexts these high-risk location approaches are no longer relevant due to the increase of crashes at locations with fewer crashes (SWOV, 2010; SWOV, 1982).

The second is the *proactive* approach, of which the most emblematic program is *Vision Zero*. This approach was first introduced in Sweden in 1997, when it was passed into the national legislation. Vision Zero maintains that no loss of life is acceptable for users of the transportation system and assigns the responsibility for traffic deaths and permanent injuries on the designer of the system. From this perspective, human error is considered, and the system's features should make it impossible under any circumstances for anyone to be killed or severely injured as a result of road traffic. This approach has been proven to generate quite satisfying results in Sweden, with traffic deaths having plummeted since the new policy was implemented. The Swedish Vision Zero program has been translated into the *safe systems approach*, which according to involves "building a system in which people cannot be fatally or severely injured on despite human error." (Jobs et al, 2016a).

The systemic approach is found between these two extremes. Defined by the Federal Highway Administration (FHWA) as making "improvement[s] that [are] widely implemented based on high-risk roadway features that are correlated with particular crash types," the systemic approach intersects reactive and proactive strategies. Indeed, it uses historical crash data to target road facilities that have experienced higher incidences of crashes. However, it goes beyond identifying clusters of crashes, as it does not consider specific locations, but rather high-risk road features, and ultimately would also apply countermeasures to low or no-crash sites.

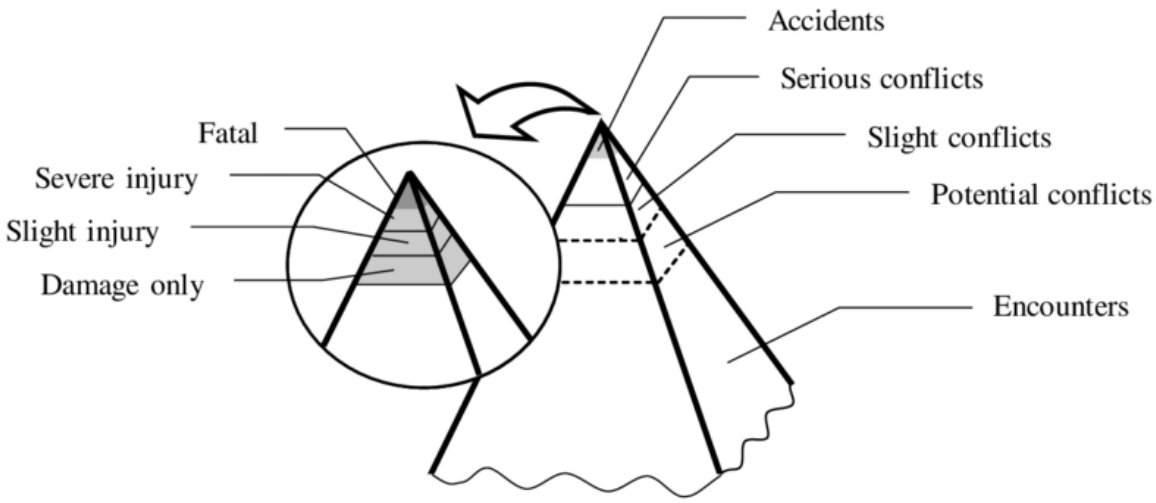
The rationale behind the adoption of this approach is that transportation agencies moved away from approaches trying to address all levels of crash severity and chose to focus on reducing the occurrence of the most severe crashes (PIARC, 2013). Considering the low density and wide distribution of such crashes over the road network—in 2013, 53% of fatal crashes within the federal road network were located in rural areas—adopting a traditional hotspot approach would not efficiently identify potential safety investments. In addition, adopting a *systematic* approach, that is, implementing countermeasures across the entire network, is not realistic in a budget-constrained environment. The systemic approach therefore appears to be best suited to address the occurrence of severe crashes across road networks.

Measuring safety using the systemic approach: choosing the right safety indicators

Ultimately, the systemic approach is about improving road safety, by better identifying safety needs—that is, by better spotting unsafe features of the road network. But measuring road safety is not an easy task because the concept of safety itself is hard to define, though the term is broadly used among both experts and the public (SWOV, 1994). The core problem lies in the fact that safety problems are brought to light when unsafe situations occur in the form of crashes and subsequent injury or fatality. This is why Ezra Hauer felt the need to state at the beginning of his *Observational before-after studies in road safety* that "road safety is manifest in the occurrence of accidents and their harm," as opposed to the subjective feeling of security (Hauer, 1997). Crashes are manifestly correlated with road "unsafety," and crash counts have therefore been widely used as metrics for road safety, especially by policy makers because figures of road fatalities and injuries make a stronger case for road safety than complex measurements.

However, using crash data as a direct measure of road safety has its caveats. The overarching goal of safety analysis as formulated by Leonard Evans is to "examine factors associated with crashes with the aim of identifying those that can be changed by countermeasures (or interventions) to enhance future safety" (Evans, 1991). This implies the need for large sample sizes for significant statistical observations, which is not always possible when it comes to crash data. Additionally, relying solely on crash data ignores the fact that crashes themselves are a result of the emergence of hazardous situations—some of which resulted in a crash, while the others do not. This distinction is essential, because it recognizes an essential dilemma: what do we consider to be a safer system, a lower number of crashes or a lower risk of getting into a crash?

Figure 2: "Safety pyramid" (adopted from Hydén, 1987)



Traffic events can be represented as a continuum of situations in pyramidal layers (see Fig. 2), whose volumes corresponds to an event's frequency (Hyden, 1987; Hauer, 1997; Tarko, 2012). The connection between these events and road "unsafety" make the case for the use of surrogate measures of safety, which: (i) are correlated with the occurrence of crashes, and (ii) capture the effects of safety countermeasures (Hauer, 1997; Gettman and Head, 2003; Tarko et al., 2009; Tarko, 2012). These two features make surrogate measures of road safety valuable because they deepen the understanding of factors leading to failure mechanisms in the road system. Furthermore, as shown in Fig. 2, more frequent events are easier to measure, which would call for a wider reliance on surrogate measures. Still, as mentioned by Hauer, using surrogate measures to quantify safety "rests on the observation that where there is smoke, there is fire." Such an assumption, true or not, reinforces the fundamental link between road safety and crashes. Nonetheless, some nuanced interpretation should be made between fluctuating crash counts and the permanent idea of the safety of a road entity: Facilities with zero past occurrences of crashes should not be considered perfectly safe, since only roads with no traffic at all have a zero chance of a crash.

Recognizing this last point calls for taking into account levels of traffic when measuring whether a traffic facility is safe. The busier a roadway is, the more likely it is, all else being equal, that vehicles will collide. Therefore, some studies have relied on crash exposure rather than crash frequency to measure road safety. And furthermore, why should agencies worry about fixing facilities that are not predominant within their network? This is the concern that could be addressed by relying on a third road safety indicator: crash density—the ratio between the number of crashes and the "size" of the road network, whether in terms of the mileage of roadways or the number of intersections. These last two indicators are quite similar, in the sense that they can be respectively seen as activity-based measures of exposure and an infrastructure-based one. The Dutch Institute for Road Safety Research summarizes the concept of exposure measures as capturing a unit amount of risk—a unit that can express duration, distance, population, expected number of encounters, or other factors, depending on the intended use of the measure—mainly useful for making situations comparable (SWOV, 1994; SWOV, 2002). While they are in no way the only indicators for measuring safety on a road network, they share important advantages over more elaborate ones, the first one being their ease of calculation based on crash data from police reports and basic infrastructure data. Depending on data availability, it may be relevant for some agencies to consider mixed safety indicators that would go beyond the dichotomy between crash numbers and rates. An example is the ratio between crash frequency and vehicle-miles traveled, a combined infrastructure- and activity-based exposure rate

that takes into account both the length of the road network and the traffic flow on the infrastructure. Many other safety indicators could be built, some of them more direct measures of safety, some more surrogate measures. Ultimately, the purpose of the present study is not to outline in a definitive manner *the* right way of measuring road safety, and *the* unique safety indicator to be used when implementing the systemic approach. Each and every indicator responds to different safety concerns, and choosing one over the others constitutes nothing less than a political choice. It is therefore the responsibility of each safety agency to decide which indicator is most appropriate in identifying systemic safety concerns. In this study, for illustrative purposes, the level of safety of a traffic facility is measured as the number of crashes, by kind and severity, that occurred on this facility during a specified period—considering that crash counts constitute the most directly available information on crashes based on any police reports, regardless of the local jurisdiction.

The systemic matrix scheme

At the core of the proposed approach is an easy-to-interpret systemic crash matrix that shows what types of crashes occur on what types of facilities. Matrix rows represent crash types, while columns correspond to facility types. The cells of the matrix are referred to as crash profiles and include aggregate information on crashes that occurred for each crash profile. The way in which this information is aggregated depends on the chosen safety indicator—in the case of crash frequencies, each cell contains the number of type X crashes that occurred on type α roadways. The cells with the highest value represent *systemic hotspots*, which are systemic challenges on the roadway network in which a particular crash profile is consistently associated with a particular type of road infrastructure.

Using such a matrix provides agencies with a snapshot of any systemic problems on their networks that are both easy to assemble and to interpret. The advantage of this scheme is that it is compatible with the data-driven rationale of the systemic approach, offering enough flexibility to allow agencies with varying degrees of data availability to implement it. The approach mainly expands on two previous initiatives in the United States: FHWA's Systemic Safety Project Selection Tool, and California's Systemic Pedestrian Safety Analysis. Both approaches involved building a matrix, the rows and columns of which were determined to best illustrate the infrastructure-related dynamics behind road collisions. The FHWA tool has been regularly used to guide road safety analyses across the nation and to help prioritize locations. The process developed by FHWA starts with the identification of focus crash types and facility types based on crash data and infrastructure information. This principle was adopted by the Californian analysis, in which the crash matrix uses columns representing locational characteristics understood to influence the collisions and based on data availability, and rows corresponding to crash types, understood as primary collision factors and behaviors thought to influence the crash. The following section will guide road safety professionals and researchers through the process of creating a systemic safety matrix. While it builds on findings from a methodology elaborated with data from the Highway Safety Information System (HSIS), this data-driven process was considered to be generalizable to other data sources and road networks beyond the seven HSIS member states, and the following will provide guidance on how to assess various matrix structures and select the most appropriate one.

Methodology for building a systemic matrix

Data collection requirements

Building a systemic matrix requires the acquisition of historical data on road collisions as well as infrastructural characteristics—ideally of the entire road network, but at the very least of the crash locations. This section describes the various aspects to consider when collecting data for a systemic crash matrix.

The first step is to identify the matrix categories that are to be built with the systemic method. As indicated in the previous section, not all crashes will be considered at the same time, but only crashes of a certain type, within a certain timeframe. These categories are dependent on the type of road network being studied. Indeed, in a data-driven approach, it is essential that each category include enough crashes to lead to meaningful, generalizable conclusions. This implies that matrix categories cannot be too restrictive if too few crash situations are captured by them. Examples of matrix categories include:

- Collisions that led to property damage only (PDO), that is, crashes in which no one was injured or killed
- Collisions that led to physical injuries or death (i.e., excluding PDOs), severe and fatal collisions only
- Collisions that only involve motorized vehicles
- Collisions that include at least one pedestrian
- Collisions that include at least one bicycle

These categories illustrate various concerns in classifying road crashes. First, the type of damage considered needs to be specified: Is there a focus on human casualties? On property damage? On all types of damage caused? The inclusion or exclusion of PDOs relates to the severity of the collision. Considering the overall issue of the allocation of scarce financial resources to road safety improvements and the recent interest in Vision Zero and safe systems initiatives—including severity in the systemic analysis of road safety—helps put the focus on designing road networks in which no traffic fatalities or severe injuries could occur. Second, mentioning the party involved is important, because distinguishing between the transport modes involved in collisions is more realistic regarding the dynamics specific to each collision. A common way of defining the party type of a collision is as follows: (i) pedestrian crashes refer to collisions in which at least one pedestrian was involved in a collision; (ii) bicycle crashes refer to collisions in which at least one bicyclist was involved in a collision; and (iii) auto collisions refer to crashes in which no pedestrian nor bicycle were involved, i.e., only those including motorized vehicles.

The decision about whether to consider the above mentioned matrix categories individually or to combine some of them is ultimately up to the agency undertaking the systemic approach based on the specific outcome pursued in this endeavor (e.g., preventing the occurrence of severe or fatal collisions, addressing the occurrence of PDOs, increasing pedestrian safety on the network) and the available amount of historical data falling under the selected matrix categories.

As mentioned in the introduction, the use of HSIS data was central to how the enhanced systemic approach to road safety was initiated. Each of the subsequent subsections will detail guidelines for determining which data elements are needed to create a systemic matrix, followed by the decisions made regarding the development of the systemic matrices with HSIS data. Regarding the above mentioned matrix categories, four were retained:

i. Auto collisions involving only PDOs (excludes injury, fatal, pedestrian and bicycle-involved collisions)
ii. Non-PDO auto collisions (includes injury, fatal collisions; excludes pedestrian and bicycle-involved collisions)
iii. Pedestrian-involved collisions (PDOs and non-PDOs)
iv. Bicycle-involved collisions (PDOs and non-PDOs).

Subdivisions based on party type (auto, pedestrian or bicycle) lead to a drastic reduction in the number of crashes when it comes to those involving pedestrians and bicyclists. The corresponding small sample size is a downside, as small changes can skew the data. It was therefore decided to include all collisions in the matrices (iii) and (iv) to allow for a conservative analysis based on enough data points. As for auto-only collisions, they were divided into two matrix categories: one for PDOs only, and the other for injury and fatal crashes only.

Crash data

The singularity of the systemic approach resides in both its reactive and proactiveness. The use of historical crash data to identify systemic road safety challenges is at the center of the approach. While not all characteristics of a specific crash are used to identify systemic hotspots, it is important that agencies undertaking the systemic approach be aware of the data requirements that this approach requires. The collection of the necessary crash data should be comprehensive to provide a truthful picture of the network crash profiles, thus enabling an accurate pairing between crash profiles and roadway types. The following describes minimum crash data requirements.

As emphasized before, crash data can be derived from a host of data sources, and thus can take multiple forms. There is no fixed standard regarding the structure of crash databases, even within the United States. The data used to develop the systemic approach came from the HSIS database, which provides information on crashes and the corresponding roadways for seven different states. The HSIS database provided all of the data elements listed in the table below, among others, for five years, from 2010 to 2014.

As established by the World Health Organization (WHO) in 2010, there are minimum data requirements regarding crash characteristics and consequences. Depending on the type of reports generated, additional variables can be derived from the initial ones derived directly from police crash reports, which can save time and effort for those creating the systemic matrix.

Table 1. Minimum crash data requirements for the systemic matrix.

	Variables	Description	Data type	Data values	Comments
C1	Crash identifier	Unique identifier within a given year that identifies a particular crash.	Numeric or character string	0123456789	Value usually assigned by the police, as the first entity recording the incident at the crash scene.
C2	Crash date	Date on which the crash occurred.	Numeric	DDMMYYY	Useful for seasonal comparisons and time series analyses, among others.
C3	Crash time	Time at which the crash occurred, using the 24-hour clock format.	Numeric	HHMM	Useful for comparisons between periods (e.g., AM, PM, nighttime).
C4	Crash location	Exact location at which the crash occurred.	Character string		Various referencing methods are possible and include: (1) latitude/longitude coordinates; (2) linear referencing system; (3) link-node system. Ideally, a combination of GPS coordinates with the route name or another designation is desired to best relate geographic coordinates to roadway elements listed in the road infrastructure directory. If not available, the crash location should at the very least document the street or road name, a reference point, and the distance and direction from that reference point. The accuracy of the crash location documentation is critical for the identification and implementation of engineering countermeasures on crash sites.
C5	Crash type	Other party or object that led to the injury or damage-producing event of the crash.	Categorical	Moving vehicle; parked vehicle; pedestrian; bicycle; fixed object; non-fixed obstacle; animal; train; no object; etc.	Collisions can include more than one event. However, the main triggering element of the collision should be listed and is key to identifying countermeasures.

C6	Primary collision factor	Principal cause of the collision.	Categorical	Alcohol; failure to yield; improper turn; following too closely; speeding; etc.	Similarly, there may be multiple factors at play in a single crash. Knowing the primary cause is key to identifying countermeasures.
C7	Violation code	If applicable, legal code of the traffic violation that led to the crash.	Categorical	22107	Provides more flexibility in the grouping of crashes by traffic violation types (e.g., control violation) than the standard primary collision factor (C6) categories listed above, at the discretion of the matrix developer. Provides more details on the specific primary causes of a collision.
C8	Impact type	Manner in which the motorized vehicle(s) involved initially collided with another vehicle, object or person.	Categorical	Single-vehicle crash; rear-end; head-on; Sideswipe; broadside; etc.	Useful for suggesting the trajectory of the vehicles involved in the collision.
C9	Movement prior to the collision	Type of movement of the primary vehicle preceding the first impact.	Categorical	Proceeding straight; left turn; right turn; U turn; backwards; changing lanes; unknown; etc.	Useful for suggesting the trajectory of the vehicles involved in the collision.
C10	Number of parties involved	Number of parties involved, including motorized and non-motorized vehicles	Numeric		Informs on the overall scale of the crash.
C11	Party type	Type of parties involved in the crash, in addition to the motorized vehicle(s).	Categorical	Auto-involved; pedestrian-involved; bicycle-involved.	Informs on the involvement of non-motorized individuals in the collision. Considering that some collisions may involve vehicles, bikes and pedestrians, they would be flagged as both pedestrian and bike-involved, and thus included in more than one of the matrix categories listed in the previous section.

C12	Crash severity	Most severe injury of any person involved.	Categorical	Fatal; severe injury; slight injury; property damage only.	Facilitates the grouping of crashes by severity level, thus enabling different policy focuses (e.g., reducing traffic deaths and severe injuries).
C13	Number of fatalities	Number of deaths resulting from the crash.	Numeric		Count includes all vehicles and individuals involved in the crash.
C14	Number of non-fatal injuries	Number of non-fatal injuries resulting from the crash.	Numeric		Count includes all vehicles and individuals involved in the crash.
C15	Weather conditions	Prevailing atmospheric conditions at the crash location, at the time of the crash.	Categorical	Clear; rain; snow; fog; strong winds; unknown; etc.	Unveils potential causes of vision impairment or challenging conditions of the road pavement surface.
C16	Light conditions	Level of natural and artificial light at the crash location, at the time of the crash.	Categorical	Daylight; dusk; dark; dark with streetlights; unknown; etc.	Unveils potential issues of visibility.

The way collisions are recorded dramatically affects the quality of the information provided. Substantial variations in the methodology from one year to another, especially if poorly documented, can significantly decrease the reliability of the information provided on past collisions. Underreporting is also a problem when it comes to the accuracy of the crash data that is being used. While this could be addressed by combining the main crash data source with other up-to-date and accessible databases, incompatibilities in databases can constitute challenges preventing such a consolidation.

A key factor when collecting crash data from a specific database is to decide on the number of years of past collisions that should be included in the systemic matrix. A common time period for analysis is a five-year collision data window. Increasing that number of years represents one way of mitigating the lack of data points, which is critical in a data-driven process as generalizability is key. However, the caveat of settling for too many years of data is that road infrastructure is not permanent. There should therefore be a balance between maximizing data points and considering potential roadway network changes, or travel pattern modifications evolutions. Assuming that the database structure is conservative/solid, this could be solved by ensuring the appropriate pairing of crashes with roadway characteristics that were prevailing at the time of the crash. This requires regular internal updates to the roadway elements of the database. The structure of the HSIS database provided yearly information for road infrastructure, as presented in the following section.

Infrastructure data

The systemic approach links crash profiles and infrastructure types to unveil linkages between specific types of crashes and specific features of roadways, thus allowing the implementation of blanket improvement across an entire facility type. Infrastructural elements at the location of a collision are therefore central to the development of a systemic matrix. The following describes minimum infrastructure data requirements.

In the systemic matrix, columns represent locational attributes of the infrastructure that help predict the likelihood of the occurrence of a crash. Similar to crash data, infrastructure data can be found in multiple databases under various forms. The HSIS database provides yearly information on road infrastructure in addition to crash data for its eight member states (California, Illinois, Maine, Michigan, Minnesota, Ohio, Utah, and Washington) and provides guidance on how to link both.

Using the systemic approach for an extensive, diverse road network can lead to the temptation to try to be too exhaustive in describing facility types in the systemic matrix columns. A roadway can be described using its number of lanes in each direction, the presence of traffic controls, traffic volumes, presence of a median, of a crosswalk, speed limit, as well as many other attributes. However, the more roadway characteristics are included in the matrix columns, the more the matrix expands. This leads to a much larger number of cells, thus lowering significantly the "size" of systemic hotspots by spreading road crashes among a greater number of cells. As a consequence, interventions on the road network following the identification of relatively small systemic hotspots would have a smaller scale, thus reducing the impact of the systemic approach. It is therefore recommended to thoughtfully select the attributes that will be used to describe the roadway infrastructure.

The table below lists the minimum infrastructure data requirements to allow the purposeful selection of the final set of columns for the systemic matrix. Not all matrix categories require the same information for infrastructure: quite simply, some crashes cannot take place on some roads because access is limited to certain users. Therefore, the table differentiates between matrix categories[1] as well as broad location categories: intersections and road segments. Depending on which road network is being studied, this location categorization can vary. For example, when using the HSIS data system, one can differentiate between intersections and state highways; while when

[1] The table only comprises 3 of the 4 matrix categories as the distinction between PDOs and non-PDOs does not affect what infrastructure variables are relevant for car-only crashes.

using urban data, one can distinguish the footprint of intersections from intersection influence areas and mid-block areas. To allow for a consistent analysis and smooth application of corresponding countermeasures, the categorization of locations is best conducted by splitting road segments so as to obtain a uniform cross-section: road segments, intersections, and horizontal curves. Consequently, for the present project, separate matrices were built with HSIS data for each roadway environment. In order for the columns to be catered to the specific environment type and for a more focused data collection approach for the records falling into each category, it was decided to create two locational categories: (i) intersections; and (ii) state highways. Quite simply, collisions occurring within 250 feet from the center of the intersection were flagged as intersection crashes, while crashes beyond were flagged as highway crashes.

Unless otherwise specified, the collection of the variables detailed below is valid for every party type, and explains the reasoning behind these minimum requirements, by referring to the HSIS study case.

Number and type of lanes:

Regardless of the parties involved in a crash, the characteristics of the lanes on a road affect the potential movements made by motorized and non-motorized vehicles. It is therefore important to collect information on the number of lanes as well as their type, namely, whether there are turning pockets. Some of these variables could be combined if deemed appropriate, but the following will detail the reasoning around why the list shown above was chosen.

- I1: The number of lanes of the primary road matters for all party types and all crash locations.
- I2: The number of lanes of the secondary road is unnecessary outside of the intersection footprint. With HSIS data, considering that the intersection area extends 250 feet beyond the center of the intersection, it is assumed that the characteristics of the secondary road are not influencing vehicles anymore once they enter the highway area.
- I3: Knowing the number of through lanes in each direction is valuable information, as it shows how important the traffic around the vehicle could be. It is only interesting for the primary road, for conciseness concerns—in other words, to keep the systemic matrix dimensions within reasonable boundaries, as mentioned earlier.
- I4: Similarly, knowing how much traffic the vehicle is facing is extremely valuable, especially if there is no physical separation between the two directions.
- I5-I8: The presence of turning lanes can influence the way the vehicle is driving, in particular in terms of potential changes of lanes or turning movements, which may result in hazardous situations. Turning lanes are only to be found in the immediate vicinity of an intersection, and therefore the collection of the corresponding data is only limited to the intersection area. If no detail is available, it is also acceptable to combine this information as the presence and number of turning lanes, instead of differentiating between left and right-turn lanes.

Median:

- I9: The presence of a median and its type matter for all crashes outside of the intersection area. This is justified by the fact that within the footprint of an intersection, there can be no median.

Speed:

- I10: The speed limit on primary roads matters both on intersections and highways.
- I11: The speed limit on secondary roads only within the intersection area, for the same reason leading to the exclusion of information on secondary roads beyond the intersection for other variables.

Traffic control:

- I12: The presence and type of intersection controls only matters for intersection area, by definition. Indeed, outside of it, the characteristics of the intersection (including the existing of traffic controls) do not influence drivers.

Dedicated ways:

- I13: The presence of a bike lane is only relevant for bike-involved crashes. Additionally, considering bike lanes are not necessarily marked within intersections, they do not need to be included in data on highway segments. Another reason is that the to/from direction of a cyclist in the intersection footprint cannot be confirmed through HSIS data, nor whether those intersection legs have a marked bike lane.
- I14: The presence of a sidewalk, similarly, ought to be collected for pedestrian crashes only (as it is only relevant to crashes involving pedestrians), and not within the intersection area (because the sidewalk ends there).
- I15: The presence and type of crosswalk was deemed important for pedestrian crashes only, and all locations.

Volume counts:

Statistically speaking, a highly frequented roadway is more likely to have higher crash counts. It is thus indicated to consider traffic counts, with various volume category breaks. There are various ways of measuring said traffic volumes. The HSIS comprises two of them: the average annual daily traffic (AADT) and vehicle-miles traveled.

- I16: Traffic volumes along the primary road matter for all party types, both within the intersection area and on highway segments.
- I17: Traffic volumes along the secondary road, on the other hand, matter for all party types, but for crashes within the intersection area only—for the same reason leading to the exclusion of information on secondary roads beyond the intersection for other variables.

Other roadway environments may call for additional attributes (e.g., in urban settings, the presence of on-street parking). The list above does not have the ambition of being exhaustive, but providing guidance to agencies on how what infrastructure data to target when thinking about developing a systemic matrix.

Table 2. Minimum infrastructure data requirements for the systemic matrix.

Attributes		Variables	**A)** Automobile-Only		**B)** Pedestrian-Involved		**C)** Bicycle-Involved	
			Intersection	Highway	Intersection	Highway	Intersection	Highway
Number and type of lanes	I1	Number of through lanes – both directions – primary road	x	x	x	x	x	x
	I2	Number of through lanes – both directions – secondary road	x		x		x	
	I3	Number of through lanes – direction of party 1 – primary road	x	x	x	x	x	x
	I4	Number of through lanes – reverse direction of party 1 – primary road	x	x	x	x	x	x
	I5	Number of left turn lanes – primary road	x		x		x	
	I6	Number of right turn lanes – primary road	x		x		x	
	I7	Number of left turn lanes – secondary road	x		x		x	
	I8	Number of right turn lanes – secondary road	x		x		x	
Median	I9	Presence and type of median		x		x		x
Speed	I10	Posted speed limit – primary road	x	x	x	x	x	x
	I11	Posted speed limit – secondary road	x		x		x	
Traffic control	I12	Presence and type of intersection control	x		x		x	
Dedicated ways	I13	Presence and type of bike lane						x
	I14	Presence and type of sidewalk				x		
	I15	Presence and type of crosswalk			x	x		
Traffic counts	I16	Traffic volumes along primary road	x	x	x	x	x	x
	I17	Traffic volumes along secondary road	x		x		x	

Data needs that are specific to non-frequency safety indicators

Regardless of the structure of the crash and infrastructure data at hand, crash frequency constitutes the most straightforward safety indicator because it does not require the combination of infrastructure information—such as the number of roads falling under a certain facility type, or the traffic volumes of a certain facility—to the raw number of collisions, as required by the two crash ratios mentioned in the previous section.

Populating the systemic matrix with crash exposure instead of crash frequency requires dividing the number of collisions that fall under a specific crash profile and that occurred in a specific facility type by the traffic levels experienced by that facility. There are some nuances in this calculation depending on which location category is considered. For intersections, there is an option of either taking into account the traffic flow on the primary road only, or a combination of both the primary and secondary traffic flows. For highway segments (and for intersections, secondarily), if the data is available, there is an option of counting traffic flows on both directions either jointly or separately.

One important challenge arising from the calculation of exposure is that it imposes the inclusion of traffic counts as one of the column attributes, otherwise, facilities with different volumes would belong to the same facility and call for the use of volume averages, which would defeat the purpose of illustrating the singularities of facility types. Second, when it comes to pedestrian or bicycle-involved collisions, relying solely on vehicle traffic counts only addresses part of the problem: a comprehensive approach would require taking into account jointly auto traffic volumes and pedestrian (respectively bicycle) traffic volumes. Nevertheless, this ambition is thwarted by the low availability of volume counts for non-motorized transport modes.

When it comes to crash density, the matrix cells should contain the ratio between the number of collisions falling under the corresponding crash profile and facility type, and the number of corresponding facilities on the network being studied, in the case of intersections—or the total combined length of said facilities on the entire network, in the case of road segments. This implies that unlike frequency and exposure, the use of density rates does not allow direct comparisons between intersection and roadway segments, since the denominator is different, which can be an issue for an agency willing to consider the systemic safety challenges of its network as a whole. Another challenge posed by this indicator is one related to data availability: the database in use may not contain information on the entire road network under examination. If it only comprises information on the characteristics of facilities that experienced collisions during the study period, then the resulting density rate will not be representative of the entire network.

In addition to these challenges that complicate the use of non-frequency indicators, the respective intrinsic advantages and disadvantages of these alternate ways of identifying systemic hotspots are summarized in the Table 3 below.

Table 3. Advantages and disadvantages of using crash counts or ratios in the systemic matrix

Indicator	Pros	Cons
Frequency	Ease of access: data directly available from police reports. Representativity: "real" number of hazardous situations, thus providing an indirect measure of the number of people affected by the safety hazards of the road network. Powerful for external communications and advocacy.	Not representative of risk levels: disconnected from the activity levels of the infrastructure. Not representative of risk levels: disconnected from the extension of the road network.
Exposure	Representativity: measure of the "riskiness" of facility types from an individual's perspective, by showing the odds of getting in a crash based on traffic levels. Interesting for external communications and advocacy: supports a new narrative about systemic risk instead.	Only partially representative of risk levels: disconnected from the extension of the road network. Constraining: requires a certain column structure, by necessitating the inclusion of traffic volumes as infrastructure attributes in the systemic matrix. Not adapted to all party types: only taking into account the flow of cars, regardless of pedestrian and bicycle activity levels. Data availability: information on volume counts for all transport modes are rarely available. Not representative of the layout of car lanes: only taking into account the overall flow of vehicles, not the two opposite flows that are meeting on streets that are not one-way. Not representative of the number of people affected by the safety hazards of the road network. Not adequate to measure the effect of safety treatments, due to the non-linear relationship between exposure rate and traffic flow.
Density	Representativity: measure of the "riskiness of facility types from an agency's perspective, by showing the expected number of crashes on a network of a certain size.	Only partially representative of risk levels: disconnected from the activity levels of the infrastructure. Not representative of the number of people affected by the safety hazards of the road network. Location dichotomy: denominator specific to the crash location category (intersection v. roadway segment). Limited usefulness for external communications and advocacy: not adequate to spot facility types with the most crashes nor with the most impact on individuals. Sensitivity: substantially affected by random variations in the case of "small" facility classes (few intersections, or little mileage). Data availability: need for information on the characteristics of the entire network, not just on crash locations.

Data cleaning

Adequate data collection has to be followed by thoughtful data cleaning to ensure the validity of the conclusions coming from the systemic matrix. Data quality is best guaranteed by accurate, complex, and standard data (PIARC, 2013). This means that not only should the data include all the minimum features mentioned in the

previous section, but also follow standard definitions. However, even for a database such as HSIS that constitutes a repository of standardized data across its member states, definitions vary from one state to the other, and even over time. Considering that implementing the systemic approach involves relying on data from several years, agencies creating systemic matrices should request the documentation that was in effect for each of these years, to avoid leaving changes in definitions unnoticed. A simple example of this risk is that of violation codes: with legislation being specific to each state and evolving over time, infractions that led to a collision in one year may not be considered infractions for a different year. This could fundamentally change how some collisions were flagged with a certain primary causing factor. Another example is how crash locations are flagged: Do all states define intersection crashes within 250 feet of the center of the intersection? It is important to consider consistency within the data, and to be meticulous when it comes to the definitions of categorical variables.

Following is a description of the process undertaken with the HSIS data to provide readers with an idea of the challenges involved with this database, and the types of potential challenges other databases might also present. The cleaning of HSIS data was undertaken in Python but could be conducted using other programming languages. The research team started with California data to capitalize on the availability of data from other sources, for comparison purposes if necessary.

I. Once the data request fulfilled by HSIS, the first step was to convert the SAS data into Excel files.
II. HSIS data is available in separate files, the numbers for which vary from one state to the other. For California, crash data is available in two files, one with information on each collision, and the second with separate information on each vehicle involved in the crash[2]. Infrastructure data is also divided in two: intersections and roads. Each of these files required cleaning, which followed the same steps, as will be detailed below.
III. For better tracking of the cleaning process on the raw data, summary files were created:
    - A text file summarizing the total number of files provided by HSIS (for all states) and their basic characteristics: types of files, empty ones, similar variables, etc. By looking at the number of records in each file (i.e., number of collisions, vehicles involved, roads, intersections) and their evolution over time, it was possible to observe issues in the data, visible through aberrant jumps in the data. This indicated either coding errors, reporting issues, or undocumented changes in the scope of the data collection or in core definitions.
    - A workbook with one spreadsheet per state, listing the files obtained and their variables, for direct comparison of the information that was collected over time. Looking at the consistency in variables showed that for some states (namely, Maine, North Carolina, and Washington), they changed from one year to the other. More specifically, it permitted flagging of: (a) some variables were absent from any other file within the same state; (b) some variables were absent from other years within the same state, but were present in other files for that state; (c) some variables were not collected at all for a period of time within the same state. Once these variables were identified, going through the states' codebooks to check what the variables stood for, whether they were needed, and if so, how they could be linked to other years' variables, could justify either dropping or simply ignoring them.
    - A mastertable of variables for collisions from all states and all years, which showed the challenges of applying the systemic approach to several states in the hope to be able to make comparisons. Indeed, it appeared that depending on the state, many categorical variables for the same concept had a different classification or even underlying definition.
IV. After having explored the data in depth, the different files with the relevant yearly crash and infrastructure data were linked, resulting in two sets of data for each state and year: first, crashes that

[2] Other states also have a file recording information on each occupant of the vehicles involved in a collision.

occurred at intersections, combined with the corresponding infrastructure data; and second, crashes that occurred on road segments, combined with the corresponding infrastructure data. If the previous step led to the conclusion that all the years[3] of data were of good quality, these files could be combined to include several years for each state.

V. The following pertains to more specific cleaning, and therefore need not be conducted simultaneously for several states. From now on, the present report will focus on the case of California, which was studied in more details than other HSIS member states. After verifying that the combined crash and infrastructure data met the minimum requirements listed in the previous section, the existence of potential duplicates was verified, in which case the corresponding columns would be consolidated. After this step, it was time to think about reformatting some variables, and the opportunity to create new ones that would increase the conciseness of the systemic matrix (e.g., using existing categorical variables with numerous values, such as violation codes, to create new ones with larger, consolidated categories).

The resulting clean data is a table with a combination of crash and infrastructure data in a specific state, for a certain time period, matrix category, and location type (intersection or roadway segment), ready for use in the systemic matrix.

Matrix generation

Finding acceptable structures for the systemic matrix

Generating a systemic matrix for a given matrix category is a data-driven process, based on numerous successive trial-and-error iterations with the years of crash records available. It is important to note that a different dataset (even for the same road network but with different years of collisions) may lead to a different arbitration between options for selecting the matrices' rows and columns. Defining rows and columns, therefore, takes place concurrently with populating the matrix. Identifying these rows and columns is equivalent to defining the crash profiles and creating the framework for the systemic matrix. As mentioned in the data collection section, this crucial step should consider the tradeoffs between the desire to include as many crash profiles as possible and the need for a compact and legible matrix structure. Judging whether a particular structure is fit for the dataset under study requires finding just the right balance between the personal logic of the matrix developer and some objective measure of fitness. Adding a variable in the rows, deleting one in the columns, every single choice on the variables to be included affects not only the size of the matrix, but more importantly, the story told by the crash profiles they define, as detailed in the case studies to follow.

Rows and columns are defined separately, in no particular order: starting with one or the other does not affect the final matrix structure. The following will therefore detail each consecutively—starting with the definition of rows—though they follow the same overall logic. To illustrate how crashes are influenced by the built environment, the rows of the matrix need to represent crash dynamics. These dynamics are specific to each transport mode (or party type) involved, as crashes between a car and a pedestrian show many dissimilarities with crashes between two cars. The row arrangement therefore needs to be tailored to each above-mentioned matrix category. As

[3] This was not the case for California, for which it appeared that the number of collisions on the state highway system (SHS) as reported by the HSIS database fluctuated significantly between 2010 and 2014, with the proportion of non-SHS crashes soaring from 1% in 2010 to as much as 83% in 2014. This may have indicated an erroneous data import from TASAS to the HSIS database, with numerous non-coded values. The non-reliability of the crash data led the research team to only perform its analyses on data from 2011, considering that the main purpose of the research was to develop the systemic approach, and not necessarily to perform the approach on the most recent data. Considering the extent of the Californian road network and the scale of collisions happening on that network, relying on one-year data was still acceptable.

emphasized before, unlike mere crash counts for hotspots, systemic matrix tells a story—a story about the entire road network. What the systemic approach intends to unveil is the underlying causes of typical collisions, so that their causes can be addressed in a comprehensive way and their future occurrence be prevented on all suspect road locations. The primary cause of a collision (variable C6 in Table 1) allows "explanation" of its occurrence better/more concisely than a long combination of its individual characteristics, especially in the perspective of keeping the number of rows reasonable. However, with a database such as HSIS, the values listed in the corresponding categorical variable can vary substantially from one state to the other, as can the definitions of these values (e.g., failures to yield may be defined differently in Washington state and California). Crashes are therefore best described by the underlying violation code that were assigned to them by the reporting officer that registered the collision.

Another way to be concise is to group categories into larger overarching categories. For example, several violation codes may correspond to different instances of the same violation type, such as a failure to yield under various circumstances. Ultimately, this new categorization is equivalent to manually redefining "primary collision factors." In that case, defining these large categories is up to the discretion of the systemic matrix developer. Additionally, focusing on the most represented crash types is also acceptable to obtain a systemic matrix as compact as possible. As a consequence, an option would be to choose a certain threshold for the share of crashes to be covered by the newly-defined violation types (e.g., defining violation types that encompass the top 80% violation codes in the crash data), thus categorizing the remaining collisions into a default violation category named "Other" In order to allow for meaningful comparisons when trying to describe the dynamics behind any particular hotspot, it is recommended to homogenize the resulting violation crash types across matrix categories. This means (i) retrieving the violation codes corresponding to the collision data under study, (ii) ordering it in a descending way, (iii) setting aside as many codes as needed to reach the defined threshold, and finally (iv) defining a limited, logical list of violation groupings based on this subset of the recorded violation codes. It is recommended to limit the "Other" category to a maximum percentage of crashes for each location type (intersection v. roadway segment) and matrix category (respectively, auto PDOs, auto non-PDOs, bike, pedestrian). For example, if the initial criteria to define new violation types is that at least 80% of all crashes, regardless of location or matrix category, should be covered by the new violation categories, another way to put that criteria is that no more than 20% of all crashes should fall within the "Other" violation category. But beyond the larger crash dataset, a potential additional criterion would be to have no more than 20% of crashes per location and party type fall within the "Other" violation category. This ensures that none of the developed systemic matrices will have an overrepresentation of "exceptional" crashes in their rows, which would not be informative when trying to define systemic hotspot profiles. The resulting mode-specific violation groupings for 2010 and 2011 intersection crashes in California are laid out in the three tables in Appendix A.

A convenient way of making the above-mentioned judgments and "playing" with the use of various categories or variables in a flexible way, and sort the data to the user's convenience, is to build the systemic matrix with a Pivot Table in Microsoft Excel. Factors such as the party type, the year or the severity can be used as basic filters. Additional ones may include the classification of the area (e.g., urban, interurban or rural), or some other broad classifications (e.g., the differentiation between freeways and other roads) if the implementing agency is only interested in a subset of its network. As mentioned earlier, collision characteristics are used as table rows, while infrastructure characteristics are used as columns. Values are counts of collisions.

The following subsections detail for each party type which variables were chosen and on what grounds, for intersection collisions that occurred in California between 2010 and 2011. As emphasized before, these structures are by no means the only valid ones. They only correspond to "acceptable," meaningful matrix structures that fit the data well and told a story about systemic safety hazards on the Californian state highway system at the time. The next section will detail more systematic, objective way that provide guidance on how to compare different

matrix structures in a quantitative way. However, ultimately, the systemic matrix is nothing but a decision-making tool to inform agencies about the flaws of their road network and the potential improvements they could make in order to improve safety outcomes for some subsets of the population (drivers, bicyclists, and/or pedestrians).

Case study: Variable section by party type for intersection collisions in California between 2010 and 2011

Some of the variables listed below were excluded from case study for various reasons. However, the reasoning applied in the following table does not aim to dismiss any of these core variables in a definitive way, but rather to simplify the work of the systemic matrix developer. Indeed, these matrices matter to understand the crash data. But not all of them are needed to build a comprehensive and yet compact systemic matrix. Additionally, the sub-selection resulting from this table is a general one, and then needs to be tested against actual collision data, for each matrix category, in order to identify the final row structure.

Table 4. Eligible variables for the systemic matrix rows

	Variables	Selected?	Reason
C1	Crash identifier	No	Used only for the counts populating the matrix cells.
C2	Crash date	No	Not a direct describer of the collision causes or dynamics.
C3	Crash time	No	Secondary describer of the context of a collision: knowing at what time of the day a collision either attempts to capture lighting conditions (which ought to be covered by a different variable) or the level of road traffic (captured in the infrastructure data).
C4	Crash location	No	Used only for allocating crashes to a certain location category (in the case of HSIS, intersection versus roadway segment).
C5	Crash type	No	Secondary describer of the parties (fixed or not) involved in a collision. Describing the type of crash in terms of the triggering element can be captured separately with the party type and primary collision factor variables.
C6	Primary collision factor	Yes	The primary cause acts as the best describer of the dynamics of a crash,and is key to identifying countermeasures.
C7	Violation code	Yes	Relying on violation codes rather than standard, pre-defined primary collision factors provides more flexibility in the grouping of crashes by traffic violation types to fit the specificities of the data (e.g., crash data from a rural environment may require different violation categories compared to violations that could only happen in urban contexts), at the discretion of the matrix developer.
C8	Impact type	Yes	Useful for suggesting the trajectory of the vehicles involved in the collision.
C9	Movement prior to the collision	No	Indirectly captured by the type of impact, which suggests where each party involved was coming from at the time of the collision.

C10	Number of parties involved	No	Though it is useful to inform on the overall scale of the crash, considering that most crashes involve two parties, the lack of information on crashes of a different scale (especially in a case where few years of crash data are available) would greatly increase the number of rows without creating meaningful subdivisions that would allow to identify systemic issues.
C11	Party type	No	This information determines which matrix is being built, i.e., it plays the role of a filter of the collision data that is displayed in the systemic matrix. It is therefore not used as a row.
C12	Crash severity	No	This information can be included in the matrix regardless of its structure, by filtering the collision data by severity instead of using it as a row.
C13	Number of fatalities	No	Count includes all vehicles and individuals involved in the crash.
C14	Number of non-fatal injuries	No	Count includes all vehicles and individuals involved in the crash.
C15	Weather conditions	No	Unveils potential causes of vision impairment or challenging conditions of the road pavement surface.
C16	Light conditions	No	Unveils potential issues of visibility.

Collision records were used for each matrix category to establish whether the violation type was the leading determinant behind the collision or if additional describers were needed. Violations were grouped as follows: alcohol; control violation; failure to yield; following too closely; improper turn; other improper driving; other than driving; pedestrian violation; falling asleep; speeding; and others. As explained earlier, the same violation categorization was kept for all matrices, in order to allow better easier comparisons and a straightforward determination of engineering countermeasures.

In the case of crashes involving cars only at intersections, the predominant violation type was by far failure to yield, followed by control violations. When looking for ways to single out systemic hotspots, the best strategy is to try to break down the categories with the most weight, until the matrix reaches a critical size and no variable addition fundamentally changes the balance between predominant categories. As mentioned in the table above, the type of impact—that is, whether the crash was a broadside, rear-end, or sideswipe—is another key describer of the conditions of a collision. When paired with the violation type, it allows differentiation between meaningful crash dynamics such as a broadside crash due to a “failure to yield” or a “control violation”, or a rear-end crash due to “unsafe speeds”. The row structure was thus limited to two variables. The type of impact was used as the top layer because it spread out the collisions in a more satisfactory way than if it had been the violation type first.

When it comes to pedestrian intersection crashes, considering the reduced overall number of collisions captured with only two years of data, breaking down every violation type with another layer did not make sense for intersection crashes. Had this not been the case, since the most represented violation category, failure to yield, still largely outweighed the other ones (followed by pedestrian violations), it could have been broken down with the action of the

pedestrian involved in the crash. But considering that this additional layer did not fundamentally change the significant imbalance that remained between the top two violation categories and the other rows in terms of crash records, increasing the size of the matrix without helping to single out systemic issues did not make sense.

Lastly, for bicycle-involved collisions at intersections, considering that cars and bicycles share the road, the story behind a particular crash changes dramatically depending on which party is at fault: different countermeasures will be taken if in a crash with a failure to yield, the bike was at fault, or if it was the car, regardless of the location. This is the reason why it was decided to use the party at fault as the first layer of rows, and then add the violation type as the second explanatory factor. This information is readily available in the HSIS database for California, and uses categories defined by the state DOT.

Automobile only

For collisions involving cars only, the same matrix structure was maintained whether these were property-damage only (PDO) or injury and fatal crashes. Consequently, the matrix structure for this mode only varied by location. Another reason why the same structure was kept for PDO and non-PDO auto crashes is that having the same rows and columns allows more direct comparisons between these two types of crashes and illustrates the ways in which some features of the road network are more likely to lead to severe collisions, while others may lead to many crashes that in fact present no real threat to people's safety. The following table lists the roadway characteristics variables that were collected and those that were retained in the final matrix structure, for auto-only intersection crashes.

Table 5. Eligible variables for the systemic matrix columns

Variable category	ID	Variable	Selected?	Reason
	I-1	Number of through lanes – both directions – primary road	Yes	The set of variables regarding the number of lanes on the primary and secondary roads were redefined to allow a more compact structure for the matrix columns. It was decided to combine them into a single variable that would illustrate the number of lanes of both roads at once, in a format *[number of lanes of road A]+[number of lanes of road B]*. It is worthwhile noting that this variable is symmetric, as the collision data showed that there were no meaningful differences in the types of crashes that occurred on one branch of the intersection or the other.
	I-2	Number of through lanes – both directions – secondary road	Yes	
	I-3	Number of through lanes – direction of party 1 – primary road	No	

	I-4	Number of through lanes – reverse direction of party 1 – primary road	No	
	I-5	Number of left turn lanes – primary road	No	For compacity purposes, no information on turning lanes was added on top of the total number of lanes as they only spread out the columns without breaking down collisions in a meaningful way.
	I-6	Number of right turn lanes – primary road	No	
	I-7	Number of left turn lanes – secondary road	No	
	I-8	Number of right turn lanes – secondary road	No	
Speed	I-10	Posted speed limit – primary road	No	As the number of lanes and traffic volumes already efficiently capture how busy traffic can be on the roadway infrastructure, retaining speed limits would be useless.
	I-11	Posted speed limit – secondary road	No	
Traffic control	I-12	Presence and type of intersection control	Yes	The presence and type of traffic control was used as the very first layer of roadway characteristics for the intersection footprint, as it determines fundamentally the behavior of vehicles interacting in the intersection. Based on observed similarities collision data, some groupings were made, and eventually this variable distinguished between intersections with timed traffic signals, all-way stops, two-way stops, yield signs, unsignalized intersections, and others (e.g., roundabouts). Considering that some types of signalization are logically associated with specific traffic volume levels (e.g., it is logical to have timed signals or four-way stop signs with high-volume roads), it was preferable to find some additional way of breaking down crashes than only including information on traffic controls and volumes.
Volume counts	I-16	Traffic volumes along primary road	Yes	For car crashes, the exposure of vehicles to potential hazards can be translated into the traffic volumes on the road network: driving on a highly-frequented road presents more risks of crashing into a car than driving on an empty road. Auto collision data showed that

				distinguishing “high-volume” primary roads was sufficient, with a cut-off taken at 50,000 AADT. This threshold allows for a better match with countermeasures in terms of traffic volumes (low vs high volume), by using the FHWA convention, which classifies high-volume roads as those whose AADT is greater than 50,000.
	I-17	Traffic volumes along secondary road	Yes	Similarly, for secondary roads traffic volumes were grouped into two categories, only with a different threshold: the final variable differentiated between “low-volume” roads, i.e., under the FHWA threshold of a 400 AADT, and other ones. For cross-street roads, it did not make sense to look at high volumes because there are few high-volume roads that intersect with other high-volume roads on the SHS.

Pedestrian-involved

Variable category	ID	Variable	Selected?	Reason
Number and type of lanes	I-1	Number of through lanes – both directions – primary road	Yes	The set of variables regarding the number of lanes on the primary and secondary roads were redefined to allow a more compact structure for the matrix columns. It was decided to combine them into a single variable that would illustrate the number of lanes of both roads at once, in a format *[number of lanes of road A]+[number of lanes of road B]*. It is worthwhile noting that this variable is symmetric, as the collision data showed that there were no meaningful differences in the types of crashes that occurred on one branch of the intersection or the other.
	I-2	Number of through lanes – both directions – secondary road	Yes	
	I-3	Number of through lanes – direction of party 1 – primary road	No	
	I-4	Number of through lanes – reverse direction of party 1 – primary road	No	
	I-5	Number of left turn lanes – primary road	No	

	I-6	Number of right turn lanes – primary road	No	For compacity purposes, no information on turning lanes was added on top of the total number of lanes as they only spread out the columns without breaking down collisions in a meaningful way.
	I-7	Number of left turn lanes – secondary road	No	
	I-8	Number of right turn lanes – secondary road	No	
Speed	I-10	Posted speed limit – primary road	No	As the number of lanes and traffic volumes already efficiently capture how busy traffic can be on the roadway infrastructure, retaining speed limits would be useless.
	I-11	Posted speed limit – secondary road	No	
Traffic control	I-12	Presence and type of intersection control	Yes	The presence and type of traffic control was used as the very first layer of roadway characteristics for pedestrians, as it determines fundamentally the behavior of the parties interacting in the intersection, and in particular who is at fault. Based on observed similarities collision data, some groupings were made, and eventually this variable distinguished between intersections with traffic signals, all-way stops, two-way stops, yield signs, unsignalized intersections, and others (including unsignalized intersections, roundabouts, etc.).
Dedicated way	I-15	Presence and type of crosswalk	No	Though the presence of dedicated ways such as crosswalks may seem crucial at first when analyzing pedestrian crashes, it was not included in the HSIS database as an infrastructural characteristic.
Volume counts	I-16	Traffic volumes along primary road	Yes	For car crashes, the exposure of vehicles to potential hazards can be translated into the traffic volumes on the road network: driving on a highly frequented road presents more risks of crashing into a car than driving on an empty road. Pedestrian collision data showed that distinguishing “high-volume” primary roads was sufficient, with a cut-off taken again at the FTA threshold of 50,000 AADT.
	I-17	Traffic volumes along secondary road	No	Since the overall number of pedestrian crashes is relatively low due to the limited timeframe of the collision data, breaking down more the columns would have watered down too much potential systemic hotspots and complicated their identification.

Bicycle-involved

Variable category	ID	Variable	Selected?	Reason
Number and type of lanes	I-1	Number of through lanes – both directions – primary road	No	The set of variables regarding the number of lanes on the primary and secondary roads were redefined to allow a more compact structure for the matrix columns. It was decided to combine them into a single variable that would illustrate the number of lanes of both roads at once, in a format *[number of lanes of road A]+[number of lanes of road B]*. It is worthwhile noting that this variable is symmetric, as the collision data showed that there were no meaningful differences in the types of crashes that occurred on one branch of the intersection or the other.
	I-2	Number of through lanes – both directions – secondary road	Yes	
	I-3	Number of through lanes – direction of party 1 – primary road	No	
	I-4	Number of through lanes – reverse direction of party 1 – primary road	Yes	
	I-5	Number of left turn lanes – primary road	No	For compacity purposes, no information on turning lanes was added on top of the total number of lanes as they only spread out the columns without breaking down collisions in a meaningful way, especially because very few violations were related to improper turns.
	I-6	Number of right turn lanes – primary road	No	
	I-7	Number of left turn lanes – secondary road	No	
	I-8	Number of right turn lanes – secondary road	No	
Speed	I-10	Posted speed limit – primary road	No	

	I-11	Posted speed limit – secondary road	No	As the number of lanes already somewhat captures how busy traffic can be on the roadway infrastructure and that speeding represents a minimal share of the violation types involved in bike collisions, retaining speed limits would be useless.
Traffic control	I-12	Presence and type of intersection control	Yes	The presence and type of traffic control was used as the very first layer of roadway characteristics for pedestrians, as it determines fundamentally the behavior of the parties interacting in the intersection, and in particular who is at fault. Based on observed similarities collision data, some groupings were made, and eventually this variable distinguished between intersections with traffic signals, all-way stops, two-way stops, yield signs, unsignalized intersections, and others (including unsignalized intersections, roundabouts, etc.).
Volume counts	I-16	Traffic volumes along primary road	No	Since the overall number of bike crashes in that location is relatively low due to the limited timeframe of the collision data, breaking down more of the columns would have watered down too much potential systemic hotspots and complicated their identification. It was therefore decided to leave out traffic counts, assuming that the intersection geometry would help get an idea of the level of activity of the roadways.
	I-17	Traffic volumes along secondary road	No	

The same logic can be followed for crashes on roadway segments, and other collision datasets (e.g., different years, different state). It is important to note that a different dataset might have led to different arbitration between options for selecting the matrices' rows.

What matters is the end result rather than the logic by which how each option for the row and the column structure is considered, and then either selected or left out.

Quality assessment of the matrix structure: data-driven decision-making based on structural indices and the need to make some arbitrary choices

All the row structures outlined above were chosen after successive iterations, based on the two years of crash records available. The product of this process can be a unique systemic matrix structure or multiple matrices for a specific dataset. Afterward, it is up to the decision-maker to determine which matrix structure is best aligned with its operational goals and means: which matrix will lead to the actual implementation of engineering countermeasures on the identified systemic hotspots? However, it is important to keep in mind that the systemic approach is data driven. Therefore, while the previous section outlined decisions mostly based on professional judgment and expertise, the present section describes some tools that can support a data-driven decision-making process when looking for an acceptable systemic matrix structure.

First, regardless of its size, it is useful to keep in mind the possible extreme forms that a systemic matrix can take. On one hand, it can be almost empty, with all crashes being clustered in a single hotspot, i.e., a unique pair of collision characteristics and infrastructure characteristics. On the other hand, it can be almost full, with each crash profile having occurred at least once during the period of study. Both cases are to be avoided, as they prevent the identification of an appropriate number of systemic hotspots. Having too few means that a couple crash profiles with high numbers of collisions will be the only systemic hotspots. In that case, the crash profile appears to be too general, meaningless, and defining the numerous corresponding countermeasures will be too burdensome and expensive. Conversely, having too many systemic hotspots, that is, if the matrix is too balanced between crash profiles, too many cells will have similar numbers of crashes, meaning that too many infrastructure types will require engineering countermeasures, which is not realistic financially. In summary, finding the appropriate balance in terms of matrix structure is key to the success of the implementation of the systemic approach, which culminates with countermeasures and ought to be feasible.

What this implies, is that ultimately, the share of empty cells in a systemic matrix does not matter. It is only important that it does not fall into one extreme or the other. As shown in Table 6, the actual share of blank cells for the 2010-2011 matrices for California described in the case study stand relatively high, above 75%. As emphasized throughout this report, the focus of the systemic matrix is cells with the highest weights in terms of numbers of collisions, not crash profiles with only few occurrences. This is why the kurtosis of the matrix constitutes an interesting index to aid decision-making between different matrix structures. Indeed, the kurtosis indicates the extent to which a distribution is peaked or flat. Knowing that a normal distribution has a kurtosis of 3, matrices with an overall kurtosis superior to 3 have systemic hotspots that significantly stand out. Conversely, a kurtosis below 3 would indicate a less than ideal matrix structure, where collisions are too evenly distributed. Though this index validates in an absolute manner the structure of a systemic matrix, since kurtosis is non-linear, it does not directly measure the advantages of switching from one structure to the other. For example, for the matrix of bike crashes at intersections from the case study, when determining whether to include a third layer of rows, such as the violation codes, one could build the two matrices and look at the kurtosis of each option. In that case, the scale of the relative increase of the kurtosis would not speak to the “impact” of each choice on the systemic hotspots. Doubling the kurtosis does not divide by two the number of systemic hotspots, nor increase their weight by two. In short, the kurtosis only indicates how much a particular distribution of crash profiles peaks; it only says whether the matrix structure is acceptable as is, not whether it is inherently better than another “acceptable” matrix structure.

The emergence of systemic hotspots is best enabled by the presence of peaks in rows and columns as well. This is indicated by the row and column totals, and uses their respective kurtoses as informative indicators as well. Logically, the row kurtosis of an almost empty matrix will be greater than 3: since almost all of its crashes will be contained in a single hotspot, the row on which this hotspot is located will also contain the wide majority of the

crashes. The same goes for the columns' kurtoses. On the contrary, an evenly distributed matrix will have balanced rows and columns, respectively, and therefore both kurtoses will stand below 3.

Using these three indices (the overall kurtosis, the column kurtosis, and the row kurtosis) confirms that the structures chosen for auto, pedestrian and bicycle crashes respectively in the case study are all acceptable.

Another useful and straightforward index is the size of the matrix. It allows assessment of the legibility of the matrix, which shall remain in reasonable proportions to allow better navigation between crash profiles. However, the compactness of a systemic matrix is very relative, since it depends greatly on the number of variables included in the structure, and the number of categories that these include. Sometimes, variables are just binary (e.g., presence of a crosswalk), while others need to be broken down in many categories to be comprehensive (e.g., types of primary collision factors).

Finally, the last quantitative index is the ratio of the cell maximum (i.e., the number of collisions pertaining to the first systemic hotspot) to the 95th percentile. This index goes beyond the kurtosis in that it not only indicates how acute the peak of the distribution is, but it also allows more fathomable comparisons between two distributions, since it is a percentage. Additionally, the higher the ratio, the easier it is to set the threshold for hotspot identification, since the top five percent of the crash profiles stand out so much.

Table 6. Quality assessment of the matrix structure

Matrix	Blank	Even	Random	CA1011_I_V	CA1011_I_P	CA1011_I_B
Share of blank cells	98%	36%	47%	86%	70%	78%
Overall kurtosis	4.00	-2.04	2.75	138.76	17.74	18.06
Column kurtosis	4.00	2.78	-1.05	16.87	31.45	3.88
Row kurtosis	4.00	8.00	0.64	21.15	4.28	9.60
Max to 95th pctile	18%	0%	46%	1543%	189%	175%
Number of rows	8	8	8	57	8	28
Number of columns	20	20	20	106	36	27
Table size	160	160	160	6042	288	756
Mean 95th pctile	5427	37	416	201	40	23

Using the above-mentioned indices does not disqualify the case-by-case decision-making detailed previously. The two processes are complementary and should both be applied when considering how many variables should be included in the systemic matrix, which ones, and in what order. This thoughtful approach allows the emergence of systemic hotspots, which then call for another decision to be made: How should systemic hotspots be defined? Should there be a fixed cut-off number of collisions? The next section describes how to identify systemic hotspots once the structure of the systemic matrix has been finalized.

Identification of systemic hotspots

Developing the quantitative criteria for systemic hotspots

Once the systemic crash matrices are generated, systemic hotspots are identified, using quantiles. For vehicle matrices the criteria for high-priority systemic hotspots is matrix cells with crashes counts that are above the 99.5% percentile. For pedestrian and bicycle matrices, a criterion of 90% percentile is applied (due to the fact that the number of cells in the pedestrian and bicycle-involved matrices is smaller relative to the vehicle ones). For both criteria, the percentile is rounded down to prevent situations in which a systemic hotspot was missed due to a fraction of a crash.

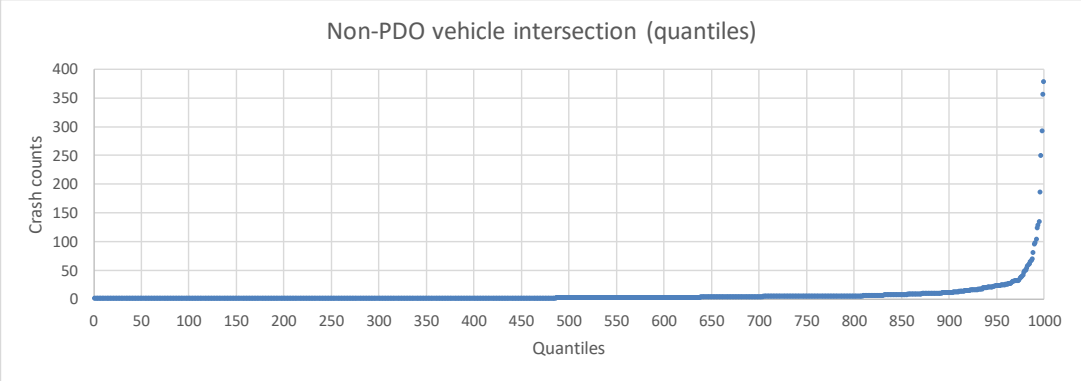
High-priority systemic hotspots are defined as matrix cells that require the attention of an agency, and represent the primary output of the systemic matrices. Considering the material constraints experienced by public agencies in charge of road infrastructure, there is a need for a metric that can efficiently and reliably alert these agencies of systemic safety concerns. If the metric is too restrictive, it can miss valuable safety-improving opportunities. However, if the metric is too inclusive, it can reduce the agency's ability to respond effectively. In light of this, the trade-off between the desire to have an inclusive list and the efficiencies of a restrictive list needs to be taken into account.

Descriptive statistics and data visualizations were used to assess several approaches for criteria-setting. This includes average-based confidence intervals, signal-to-noise ratios, triangular distributions, and quantiles. The quantile method was determined to provide the best fit across the different types of matrices. This is partly driven by the empirical distributions across matrices. More specifically, although the data in each of the matrices is zero-inflated, the behavior at the upper extremities varied quite a bit and a quantile-based method provided the most transparent and consistent outcome.

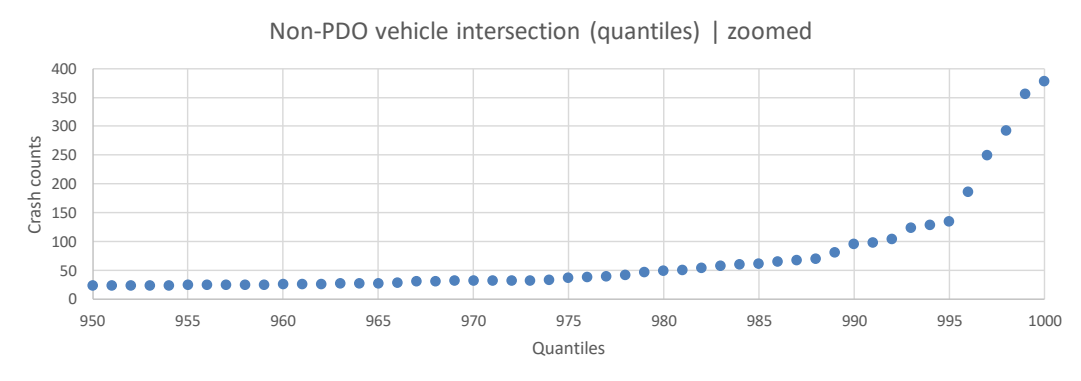
To determine the actual criteria, the data for each matrix was broken down to 1000 quantiles. The data was then plotted and reviewed to look for change-points. Figure 3 below includes charts with the quantiles for non-PDO vehicle-only (a,b), pedestrian (c), and bicycle (d) crashes at intersections. The quantiles are shown from left to right. At the far left is the 0.001 percentile, and at the far right is the 99.9 percentile. For non-PDO auto-only crashes (Figure 3a), the following plot (Figure 3b) represents a zoomed-in version of the upper part of the data, between the 95th and the 99.9th percentile. Using 1000 quantiles identifies a criterion that is at a lower resolution than the plotted quantiles. After reviewing the data at a threshold of 99.5 percentile is strongly indicated by the two auto-only plots. The plots for pedestrian intersection crashes, and bicycle intersection crashes respectively did not demonstrate a clear inflection point because the datasets were much smaller, but at a resolution of 0.01 percentile, a criterion of 90th percentile emerged as a reasonable threshold.

Figures 3a-d. quantitative quantile criteria for systemic hotspots

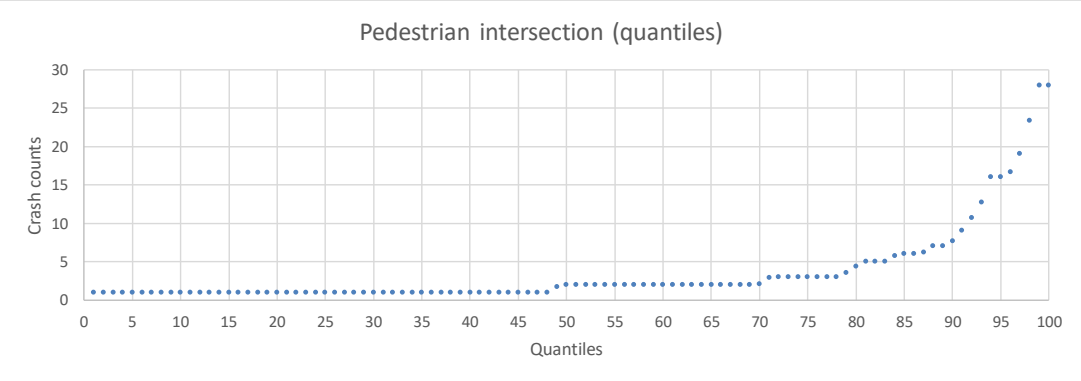
A.



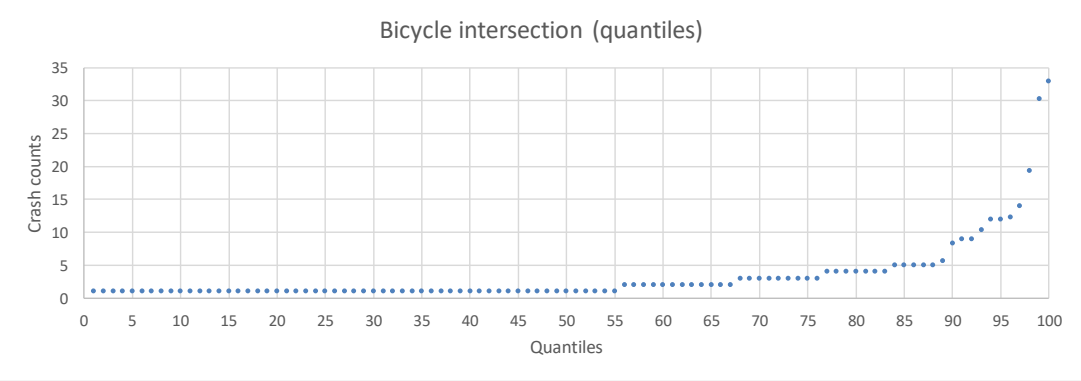
B.



C.



D.



Prioritizing among systemic hotspots

Once the high-priority crash profiles are identified, they are considered systemic hotspots and are labeled as first, second, and third priority. The priorities are determined by the order of each systemic hotspot within a column: the top systemic hotspot in a column is labeled as first priority, the second highest is labeled as second priority, and any additional systemic hotspots are labeled as third priority. The final ranking of the systemic hotspots is by descending order (in the number of crashes) of all of the first-priority systemic hotspots, followed by descending order of all of the second-order priority ones, and ends with all of the third-priority hotspots, by descending order. Finally, the total number of crashes for an entire facility type (or column) is, used as tie-breakers between hotspots with the same crash counts: the hotspot corresponding to the "most dangerous" facility type will come first. If the tied crash profiles pertain to the same facility type, then it is the row totals that will help to decide between them.

Auto-only collisions, non-PDO, intersection: (99.5th percentile is 134.00 crashes and the systemic hotspot threshold is 134).

Crashes	Crash type	Roadway type	Priority	Tie-breakers
378	Broadside; Failure to yield	2-way stop signs; 4+2: <=50,000 AADT; >400 AADT	1	n/a
353	Broadside; Failure to yield	2-way stop signs; 2+2: <=50,000 AADT; >400 AADT	2	n/a
272	Broadside; Control violation	Timed signals; 4+2; <=50,000 AADT; >400 AADT	3	n/a
235	Broadside; Failure to yield	2-way stop signs; 2+2: <=50,000 AADT; <=400 AADT	4	n/a
137	Broadside; Failure to yield	2-way stop signs; 4+2: <=50,000 AADT; <=400 AADT	5	n/a

Pedestrian collisions, intersection: (90th percentile is 7.70 crashes and the systemic hotspot threshold is 8).

Crashes	Crash type	Roadway type	Priority	Tie-breakers
28	Failure to yield	2-way stop signs; 4+2; <=50,000 AADT	1	59
28	Failure to yield	Timed signals; 4+2; <=50,000 AADT	2	56
16	Pedestrian violation	2-way stop signs; 2+2; <=50,000 AADT	3	41
16	Failure to yield	Timed signals; 4+4; <=50,000 AADT	4	28
10	Failure to yield	2-way stop signs; 6+2; <=50,000 AADT	5	n/a
8	Failure to yield	Timed signals; 6+4; <=50,000 AADT	6	n/a
21 (2nd)	Pedestrian violation	2-way stop signs; 4+2; <=50,000 AADT	7	n/a
17 (2nd)	Pedestrian violation	Timed signals; 4+2; <=50,000 AADT	8	n/a
12 (2nd)	Failure to yield	2-way stop signs; 2+2; <=50,000 AADT	9	n/a

Bicycle-involved collisions, intersection: (90th percentile is 8.30 crashes and the systemic hotspot threshold is 8 crashes).

Crashes	Crash type	Roadway type	Priority	Tie-breakers
33	Bicycle at fault; Other improper driving	2-way stop signs; 4+2	1	153 (row)
29	Bicycle at fault; Other improper driving	Timed signals; 4+2	2	107
21	Vehicle at fault; Failure to yield	2-way stop signs; 2+2	3	78
16	Bicycle at fault; Other improper driving	Timed signals; 4+4	4	58
12	Bicycle at fault; Other improper driving	Timed signals; 6+2	5	33
11	Bicycle at fault; Other improper driving	Timed signals; 6+4	6	
33 (2nd)	Vehicle at fault; Failure to yield	2-way stop signs; 4+2	7	86 (row)
14 (2nd)	Bicycle at fault; Failure to yield	2-way stop signs; 4+2	8	
13 (2nd)	Bicycle at fault; Other improper driving	2-way stop signs; 2+2	9	78
12 (2nd)	Bicycle at fault; Failure to yield	Timed signals; 4+2	10	51 (row)
12 (2nd)	Bicycle at fault; Control violation	Timed signals; 4+2	11	44 (row)
10 (2nd)	Vehicle at fault; Failure to yield	Timed signals; 4+2	12	107
9 (2nd)	Vehicle at fault; Improper turn	Timed signals; 4+2	13	107
9 (2nd)	Bicycle at fault; Improper turn	2-way stop signs; 2+2	14	78
9 (2nd)	Bicycle at fault; Control violation	Timed signals; 4+4	15	58 | 44 (row)
9 (2nd)	Vehicle at fault; Improper turn	Timed signals; 4+4	16	58 | 36 (row)
8 (2nd)	Other party at fault; Other violation type	2-way stop signs; 4+2	17	

This logic allows a first-level systemic hotspot to be ranked above a second-level hotspot with more crashes, and it is established to provide more opportunities to develop systemic improvements across multiple facility types. Indeed, a first-level hotspot will be the first type of facility that will benefit from engineering countermeasures to prevent the occurrence of certain crash types. By treating this facility type, it is possible that some co-benefits will result and also reduce the occurrence of different types of crashes for this facility (i.e., lower the number of crashes in different cells of the same column). In light of this, this approach provides more opportunities to solve systemic concerns across the road network in a more comprehensive way than focusing on a specific facility type.

Conclusions

Findings on the systemic approach and main takeaways on its limitations

The systemic approach stands at a midpoint between fully reactive approaches to road safety and the safe systems approach. It consists of identifying recurring safety concerns on a road network, by finding the crash profiles that are associated with certain roadway features, with the goal of implementing blanket improvements across an entire facility type. The methodology detailed in the present report, and illustrated by a case study for the state of California, outlines the main steps to be undertaken by any agency that wishes to undertake systemic improvements to its road network. Using a matrix format reduces the barriers to the implementation of the systemic approach that agencies with limited institutional capacity and financial means might face due to software limitations, and enables a smooth data-driven process within the grasp of any practitioner, as long as the key steps recommended here are followed. The process begins with adequate data collection—since the systemic approach relies on information on past collisions, quality historical crash data needs to be collected, and paired with detailed infrastructure data allowing the tailored definition of facility types. After a meticulous and well-documented data cleaning comes the matrix generation. The structure of the matrix requires the separate review of options for the rows and for the columns, respectively. The inclusion or the rejection of each and every variable as a row layer or a column layer needs to be justified on a case-by-case basis, but there are also informative indices that aid the decision-making on how acceptable the structure of a systemic matrix is. Finally, based on the crash counts that populated the matrix, priorities are identified in terms of systemic hotspots, which are singled out and ranked according to mode-specific rules.

The challenges of the systemic approach are related to its core principle—being a data-driven endeavor. The absence of hard rules makes challenging each decision on the characteristics of the resulting systemic matrices, and complicates the assessment of which results are “acceptable”—or say, make sense. Additionally, the quality of the analysis conducted is highly dependent on the quality of the collision and infrastructure data available. Deficiencies in the original database may not only delay a systemic safety study, but can also completely invalidate its results, either because not enough data is available to develop a satisfying matrix structure and identify systemic hotspots, or because the data is poorly documented and does not allow the definition of new crash or infrastructure categories or other elements. In summary, data is key to the success of the systemic approach, and a robust data collection and management plan should be one of the first concerns for anyone attempting to create a systemic safety matrix.

Opportunities for further research on systemic safety

Identifying systemic hotspots is not the final step of the systemic approach. It marks the transition between the evaluation of systemic safety concerns on the road network and their actual solving through countermeasures. Such countermeasures may comprise engineering safety treatments, applied across entire facility types in order to prevent the occurrence of a specific high-priority crash profile, but also non-engineering efforts. Due to various challenges that emerged over the course of the present project regarding the data needs for developing the

systemic matrix, these aspects fell out of the scope of the definition of the systemic approach. These areas of research, though not explored as part of the present study, remain important topics for further exploration. For example, regarding non-engineering countermeasures through enforcement, beyond the way in which police officers enforce the vehicle code, studying in more depth the content of the code itself would be a complementary and yet disconnected (because somewhat similar to a legislative review) research effort. Although legislation varies from state to state, studying each state's driver's handbook and the structure of the driving code may give insights as to which problematic driving behaviors already identified by a systemic matrix are addressed in the code, or if there is a gap. This effort may very well use a matrix approach again, where for each facility type and crash type, the cells would be populated with the number of driving code items (if any) that cover the corresponding problematic behavior, or if there is an opportunity to add new elements to the legislation.

Another example of a research topic that stems from the core systemic approach methodology of the present study is non-engineering countermeasures through education. Due to the multitude of stakeholders already implementing educational countermeasures related to road safety, be it institutional ones (e.g., at the federal, state, local, or in schools) or not (e.g., associations), identifying such actions already in place would prove a complicated and very place-specific endeavor. However, there would be an opportunity to build on the systemic matrix to develop learning modules that would be targeted at priority safety concerns identified in the matrix. From a long-term perspective, this approach could also benefit from partnerships with other stakeholders to see if their respective programs already tackle the identified issues, and, if not, help them integrate these in their training agenda.

One last research opportunity to explore in the future would be the issue of under-reporting in the context of the systemic approach. Indeed, in addition to the numerous problems that come with only relying on a database such as HSIS, not all crashes are reported to the police. This means that a portion of the collisions that happened in the past are not captured by the systemic matrix as is, which implies the need for additional research to identify existing databases covering some of these non-police reported crashes (e.g., data from hospitals) and how to connect them to traditional data from the police. If these additional databases do not exist, the steps needed to increase reporting in a way that could be directly used to serve the systemic approach should be investigated (i.e., with adequate data standards and reporting processes).

Appendix A

A-1: Left portion of the non-PDO auto-only matrix for intersection crashes in California (2010-2011)

CALIFORNIA	INFRASTI																											
	Timed Signals																											
	2+0	2+1	2+2		3+0	3+2			3+3	4+1	4+2			4+3			4+4		5+2				5+3		5+4		
	<= 50000	<= 50000	<= 50000		<= 50000	<= 50000		> 50000	<= 50000	<= 50000	<= 50000			> 50000	<= 50000	> 50000	<= 50000		> 50000	<= 50000	> 50000		<= 50000	> 50000	<= 50000		> 50000
Collision	> 400	> 400	<= 400	> 400	> 400	<= 400	> 400	> 400	> 400	> 400	<= 400	> 400	(blank)	> 400	<= 400	> 400	> 400	<= 400	> 400	> 400	<= 400	> 400	<= 400	> 400	> 400	> 400	> 400	> 400
A																												
Alcohol				4			1					7							3									
Control Violation			2	8			1	1		1		20						1	9	2		1					1	1
Failure to yield	1		1	15			3				2	49	1	1		4		4	20	1		2					3	1
Improper Turn							1					9							4									
Other Improper Driving				5			1					7							1									
Other than driving				1								2							1									
Speeding												8				1			1									
Others				1								5							3								2	
B																												
Alcohol												2				1			2									
Control Violation		1		2			2				2	11							9								1	
Failure to yield				1			1				1	6						1	5									
Improper Turn	1					5	1					9				1			9	1								1
Other Improper Driving		1		4		2						13							3									1
Other than driving																												
Speeding										1		1													1			
Others				1			1					5							1								1	
C																												
Alcohol				4							1	17			1				9				1					
Control Violation							1					1							1									
Failure to yield				1								1							1			1						
Following too closely		2		5							1	13		1		2		1	11	3		1					1	
Improper Turn				1								1							1									
Other Improper Driving			2	7		1					4	16		1		3		1	11	2	1							
Other than driving												1																
Speeding		2	5	27	1	2	10		2	3	4	104		9	1	3	4		67	9		4	1				5	
Others												1							4	1								1
D																												
Alcohol			1	7		1	1		1	1	6	23		1		1		1	14	3		2				1	2	
Control Violation	7	8	15	71		2	36	5	17	7	15	272		7	4	24	4	5	126	20	1	9	2		4		19	2
Failure to yield	1	2	4	23			4				5	129		6		9	1	3	61	3		4		2	1		8	2
Following too closely																												
Improper Turn	1		1	7			3					16		1					9			1					2	
Other Improper Driving		1		4			1					11		2	1	1			7					1				3
Other than driving		1		1								2							2	1							1	
Speeding			2	5								7							5									
Others			1	5			2	1		1	3	31				2	1	2	20			2			1		1	1
E																												
Alcohol				4			1				1	5				1			5									
Control Violation												1																
Failure to yield				1														1										
Improper Turn				1			1					5			1				2									
Other Improper Driving				2								1																
Other than driving												2				1												
Speeding			1	2			2					4		2		1			2			1						
Others				1			1					2																
F																												
Alcohol				1								1							4									
Control Violation																			1								1	
Failure to yield												2							1									
Following too closely																												
Improper Turn	1		1	2								1										1						
Other Improper Driving				1								1		1														
Other than driving																												
Speeding												3							1									
Others								1											1									
Other				7			2		1		2	19	1			4			4								2	
Grand Total	12	18	36	232	1	13	77	8	21	14	47	846	2	32	8	59	10	20	441	46	2	29	4	3	6	1	3	50	13

A-2: Right portion of the non-PDO auto-only matrix for intersection crashes in California (2010-2011)

| | | |

A-3: Left portion of the pedestrian-involved matrix for intersection crashes in California (2010-2011)

CALIFORNIA	INFRASTI																		
	Timed Signals																		
	2+2	3+2	3+3	4+0	4+2		4+3	4+4		5+2	5+4		6+2		6+3	6+4		6+6		8+4
Collision	<= 50000	<= 50000	<= 50000	<= 50000	<= 50000	> 50000	<= 50000	<= 50000	> 50000	<= 50000	<= 50000	> 50000	<= 50000	> 50000	<= 50000	<= 50000	> 50000	<= 50000	> 50000	> 50000
Alcohol					1															
Control Violation								1			1		1	1	1	2		1		
Failure to yield	6	2	2	1	28		2	16	3	2	7		7		2	8	5	3		1
Improper Turn					3			1					1			1				
Other Improper Driving					2			2							1	1				
Other than driving								1					1							
Pedestrian Violation	1				17	2	1	4	2	1		1	6	1	3	6	1	3		1
Speeding					2						1		1							1
Others	1				3			3					3			1			2	
Grand Total	**8**	**2**	**2**	**1**	**56**	**2**	**3**	**28**	**5**	**3**	**9**	**1**	**20**	**2**	**7**	**19**	**6**	**7**	**2**	**3**

A-4: Right portion of the pedestrian-involved matrix for intersection crashes in California (2010-2011)

4 way Sto	2 way Stop signs												Yield signs	No controls		Grand Total
2+2	2+2	3+2	4+1	4+2		5+0	5+2	5+4	6+2		6+3	8+2	2+2	2+1	2+2	
<= 50000	<= 50000	<= 50000	<= 50000	<= 50000	> 50000	<= 50000	<= 50000	> 50000	<= 50000	> 50000	<= 50000	> 50000	<= 50000	<= 50000	<= 50000	
	2															3
	2			2					2			1				15
1	12	5	1	28	1		1	1	10	1		1			1	158
	2			2												10
	2			1					1							10
																2
1	16	5		21		1								1	1	96
	3			3							1		1			13
	2	1		2	1											19
2	**41**	**11**	**1**	**59**	**2**	**1**	**1**	**1**	**13**	**1**	**1**	**2**	**1**	**1**	**2**	**326**

A-5: Bicycle-involved matrix for intersection crashes in California (2010-2011)

CALIFORNIA	INFRASTI																												
	Timed Signals																	4 way Stop signs		2 way Stop signs							No controls	Others	Grand Total
Collision	2+2	3+2	3+3	4+2	4+3	4+4	5+2	5+4	6+2	6+3	6+4	6+5	6+6	8+2	8+3	8+4	2+2	4+2	2+2	3+2	4+2	4+4	5+2	6+2	6+3	8+2	4+1		
B																													
Alcohol	1			2		2													5		3			1		1			15
Control Violation	1	1		12		9	1	1	4	2	3								4	1	5								44
Failure to yield	1			12		4			1		3					4			4	1	14			6			1		51
Improper Turn			1	5		1					3						1		9		5								25
Other Improper Driving	5	1		29	1	16			12	1	11	3		1	1	1	1	1	13	4	33	1	3	12	1	2			153
Other than driving						1			1										3		2								7
Pedestrian Violation				1		1																							2
Speeding	1			4					2	1	1		1						2		4								16
Others				3		1					1					1			1		5		1						13
Other																													
Control Violation	1			1		1			1										1										5
Failure to yield				2	1																1		1					1	6
Improper Turn	1			1		1							1						3		1			2					10
Other Improper Driving				1		2			2			1																	6
Other than driving																			4		1								5
Speeding				1		2		1			1													1					6
Others	1			5	1	2	1	1			1		1						3		8		1						25
V																													
Alcohol				1		1															1								3
Control Violation				2					2		1								2		1		1	1					10
Failure to yield	3	1		10		3	1	1	3		5								21	2	33			2				1	86
Following too closely				1																									1
Improper Turn	1	3		9		9			1		4						1	1	2		4			1					36
Other Improper Driving	1			5		2		1	3	1			4							1	1		1			1			21
Pedestrian Violation									1																				1
Speeding													1						1									1	3
Others											1													1					2
Grand Total	17	6	1	107	3	58	3	5	33	5	35	4	8	1	1	6	3	2	78	9	122	1	8	27	1	4	1	3	552

Collaborative Sciences Center for
ROAD SAFETY

730 Martin Luther King Jr. Blvd.
Suite 300
Chapel Hill, NC 27599-3430
info@roadsafety.unc.edu

www.roadsafety.unc.edu