New Jersey Safety and Health Outcomes (NJ-SHO) Data Warehouse Technical Documentation

https://njsho.chop.edu/

Last Updated: 05/03/2024

Table of Contents

Description of NJ-SHO Data Warehouse

To address the lack of comprehensive sources of traffic safety data, we began developing the New Jersey Safety and Health Outcomes (NJ-SHO) Data Warehouse in 2011. Our goal was to use a public health lens to catalyze the field's ability to address critical, high-priority issues related to traffic safety and injury prevention. We utilized rigorous data science techniques to link data from numerous complex administrative databases in New Jersey (Curry et al., 2021). The Warehouse includes (1) driver licensing histories, (2) traffic-related citations and license suspensions, (3) police-reported motor vehicle crashes, (4) birth certificates, (5) death certificates, (6) hospital discharges (emergency department, inpatient, and outpatient), (7) electronic health records for all NJ patients in the CHOP network, (8) census tract-level community indicators (using geocoded residential addresses), (9) trauma registry records (to be integrated by FY 2026), and (10) emergency medical services (EMS) data (to be integrated by FY 2026).

The warehouse currently contains linked longitudinal data for 24 million individuals over a 17-year period (2004-2020). Details about crash-involved individuals—including age, race and ethnicity, sex, and residential address—were derived from all linked sources in the NJ-SHO warehouse. The extensive information we now have on different populations of road users, including demographics, residential neighborhoods, crash involvement, injury profiles, and traffic-related behaviors, allows researchers to examine how to keep road users safe using a variety of approaches.

The NJ-SHO Data Warehouse includes several unique features:

(1) More complete identification of injuries. Research indicates a substantial proportion of injuries are missed when crash report or hospital discharge data are analyzed in isolation, in particular among pedestrians and bicyclists. The NJ-SHO Warehouse already contains linked crash report and hospital discharge data; we plan to integrate trauma registry and EMS data into the NJ-SHO Data Warehouse in the future. Thus, we have a more complete picture of events that contribute to an injury, as well as the nature and severity of an injury (Lombardi et al., 2022).

(2) Unique ability to locate at-risk communities. The NJ-SHO Data Warehouse includes both geocoded crash locations, which enable us to identify dangerous roads and intersections, and geocoded residential addresses for individuals involved in crashes. Thus, we are capable of capturing community safety both by the location of a traffic safety event ("In which communities do events happen?", i.e., an urban planning lens) and by where crash-involved drivers, bicyclists, pedestrians, and other affected individuals live ("What communities experience the burden of these events?", i.e., a public health lens). Among the more than 20

million distinct addresses we processed for geocoding, 93% were successfully geocoded to an address point or street address; in all, 97% were geocoded to a ZIP code or more precise location.

(3) More complete data to identify community-level disparities and inequities. As described above, we have geocoded residential addresses to the census tract level for most individuals included in the NJ-SHO Data Warehouse. This has allowed us to incorporate community-level indicators of equity (e.g., county or census tract), such as those available from American Community Survey 5-year estimates (e.g., median household income, education, and population density), as well as more novel ways to estimate disparities, such as Community Resilience Estimates (CRE) or Social Vulnerability Index (SVI).

(4) Individual-level race and ethnicity data. Using these integrated data, we can characterize race and ethnicity for almost all crash-involved individuals. If an individual linked to a known race and ethnicity value from another data source, such as hospital discharges or vital statistics, we assigned that value to the individual. If the individual did not have a known race and ethnicity, we used an algorithm called Bayesian Improved Surname Geocoding (BISG) to estimate their probability of being in 6 mutually exclusive groups: non-Hispanic American Indian/Alaska Native, non-Hispanic Asian/Pacific Islander, non-Hispanic Black, Hispanic, non-Hispanic multiracial, and non-Hispanic White (Sartin et al., 2021).

Detailed Description of Data Sources

Motor Vehicle Crash Reports

Data source: New Jersey Department of Transportation
Years: 2004-2019
Record Information: ~8.8M drivers, ~3M passengers, ~137K pedestrians/bicyclists
Description: Motor vehicle crash data are collected by law enforcement officers investigating a crash using the New Jersey crash report (NJTR-1). A crash is required to be reported if it resulted in injury to or death of any person or damage to property of any one person in excess of $500. These data include information about crashes, crash-involved vehicles, and crash-involved drivers and passengers, as well as crash-involved pedestrians and bicyclists. These data also provide information on crash-contributing factors, injury information for crash-involved individuals, and road and environmental conditions. The most recent requirements for filing a crash report are detailed in the New Jersey NJTR-1 Crash Report Manual, revised January 1, 2023, edition 2.0.

Decoded Vehicle Identification Number (VIN) Information

Data source: National Highway Traffic Safety Administration, Product Information Catalog and Vehicle Listing (vPIC)
Years: 2004-2019
Record Information: Varies by year; 2019 data: 92% of vehicles
Description: Information obtained from vPIC includes VIN specifications received from vehicle manufactures for model years 1981 and forward. The first 11 characters of crash-involved vehicle VINs were used to decode information about the vehicle, such as vehicle type, manufacturer, make, model, trim, weight, engine specifications, and safety specifications.

Birth Certificates

Data source: New Jersey Department of Health, Vital Statistics and Registry
Years: 1979-2019
Record Information: ~4.6M births
Description: Birth certificate data were obtained for all births occurring in New Jersey. These data include information such as birth location, date of birth, and birthweight, as well as age and other characteristics of the mother.

Death Certificates

Data source: New Jersey Department of Health, Vital Statistics and Registry
Years: 2004-2019
Record Information: ~1.2M deaths
Description: Death certificate data were obtained for all deaths occurring in New Jersey. These data include information such as cause of death, death location, date of death, and other characteristics of the decedent.

Pediatric Electronic Health Records

Data source: Children's Hospital of Philadelphia (CHOP), electronic health records (EHR)
Years: 2005-2020
Record Information: ~540K patients
Description: CHOP EHR data were obtained for all pediatric patients who met all of the following criteria: 1) were born in 1987 or later, 2) had a New Jersey address at any point in time, and 3) had at least one visit to any CHOP location in New Jersey or in Pennsylvania. These data include information such as visit dates, visit types, locations, diagnoses, and prescribed medications, in addition to patient demographic characteristics.

Hospital Discharge Records

Data source: New Jersey Department of Health, Discharge Data Collection System
Years: 2004-2019
Record Information: ~76M visits
Description: Hospital discharge data include information on all New Jersey inpatient hospitalizations, outpatient and same-day surgery visits, and emergency department (ED) visits. These data are derived from hospital uniform billing information. They include information on demographic characteristics of patients, admission and discharge dates, diagnosis and procedure codes, and billing information.

Driver Licensing Records

Data source: New Jersey Motor Vehicle Commission
Years: 2004-2020
Record Information: ~11.3M drivers
Description: These data were obtained for every driver who had a New Jersey basic/auto (class D) driver's license at some point during 2004-2020. They include information such as license type, valid period for a learner's permit, valid period for a probationary driver's license, and license expiration.

Traffic Citations and License Suspensions

Data source: New Jersey Administration of the Courts via New Jersey Motor Vehicle Commission
Years: 2004-2020
Record Information: ~92M events
Description: Data on traffic events include information such as event date, type of traffic citations, and license suspensions.

Geographic Indicators – Population Counts

Data source: U.S. Census Bureau, New Jersey Department of Labor and Workforce Development
Years: 2004-2020
Record Information: Geographic indicators are assigned to individuals based on geocoded residential address
Description: Census-year enumeration and intercensal population estimates data stratified by age, sex, and race and ethnicity were obtained for the state, for 21 counties, and for all census tracts for both 2010 and 2020 U.S. Census boundaries.

Geographic Indicators – Equity Measures

Data source: Various
Years: 2004-2020
Record Information: Geographic indicators are assigned to individuals based on geocoded residential address
Description: Census tract-level socioeconomic indicators, such as median household income, were obtained from the U.S. Census American Community Survey 5-year estimates. Various equity indicator measures were also identified at the county- and census tract-level, including Index of Concentration at the Extremes (ICE), Child Opportunity Index 2.0 (COI), Social Vulnerability Index (SVI), Community Resilience Estimates (CRE), and Disadvantaged Community Index (DCI).

Methods for Data Warehouse Development

Data Integration

The NJ-SHO Data Warehouse was constructed by first conducting a probabilistic linkage and then a hierarchical deterministic linkage. Details about our initial process to construct the warehouse can be found in Curry et al., 2021. Here we describe the methods and evaluation of our second iteration, which refined and improved upon the initial iteration.

We used LinkSolv 9.0 (2015 Strategic Matching, Inc.) for the probabilistic linkage. Briefly, LinkSolv uses Bayes' rule to calculate posterior probabilities of a true match between two records based on agreements (within a specified tolerance) and disagreements (outside the specified tolerance) between examined data elements. Comparisons across multiple data elements result in the generation of a match probability, or the likelihood that the pair is a true match. Match probabilities incorporate both the discriminating power of data elements (agreement on common values have less impact than agreement on rare values) and their reliability (disagreement on data elements thought to be less error-prone provides more evidence against a match than disagreement on data elements thought to be more error-prone). A full linkage process involves several passes, each of which brings together pairs of records with exactly the same values on selected criteria (join criteria, also commonly called block criteria) and subsequently evaluates those pairs based on additional criteria (match criteria). Match criteria are the same for each pass, but join criteria differ, thereby ensuring that disagreement on a single data element will not prevent the identification of a true match.

To prepare data for the probabilistic linkage, we combined records from all of the sources into one file so that we could execute a single file match. Then, using an iterative process, we developed and executed a linkage algorithm that ultimately consisted of two passes. We used

two criteria to control the quality of our process: (1) we rejected any pair of records with a match probability < 0.60 and (2) we selected 0.01 as the highest acceptable threshold for the false match rate. To determine the false match rate, we first calculated the false match probability for each pair as 1 minus the match probability. Then we ranked all matched pairs from the lowest to highest false match probability. The false match rate was then calculated iteratively as the sum of the false match probabilities for the ranked pairs divided by the number of pairs. Matched pairs were included in the calculation, one at a time in ranked order, until either all pairs were added or the false match rate was 0.01, whichever occurred first. The linkage algorithm identified all records that pertained to a single individual and combined them into a set. Importantly, records in each data source were linked independently of all other data sources (e.g., birth records were linked to CHOP EHR records regardless of driver license status; crash-involved driver records were linked to other crash-involved drivers even if the individual did not appear in any other source). Additionally, using a single file match method allowed us to maximize all information and connections. For instance, an individual may have had sufficient matching information to connect record A to B and record B to C, but not record A to C; because of the single file methodology, records A, B, and C were identified as a single individual or set.

Records that have full date of birth (and other required criteria) were processed through the probabilistic linkage. Records that do not have full date of birth but had first and last name and either age or street address were processed through the hierarchical deterministic process using SAS software, version 9.4 (SAS Institute Inc., Cary, NC, USA). The majority of such records were crash-involved passengers; many of those records only have age at the time of the crash captured and not full date of birth. The next largest category, at 13%, was license records that had month and year of birth but not day. The deterministic linkage process connected these records to each other or individuals from the probabilistic process, using multiple permutations of linkage criteria.

We concluded that our linkage yielded high-quality results based on multiple evaluations. First, we assessed the match probability for each pair of records accepted into a set (that is, belonging to single person). Overall, the median match probability was 0.9999972 and the interquartile range was 0.9992871-1.0000000. Second, the lowest match probability for any two records within a set was 0.99 or higher for 80.9% of sets and 0.90 or higher for 93.5% of sets. Third, based on a sample of records matched in the deterministic process that we hand reviewed, we estimated the overall true match proportion was 93%. We also estimated the false non-match proportion by taking a sample of unmatched records, finding the five most likely matches for them, and determining by hand review whether a match should have been made. Our estimated false non-match proportion was 8%. Last, we examined the number of individuals who had more than one record from a source expected to have only one record per individual (i.e., birth, license, EHR, and death data sources). The proportions of individuals with more than one record were very low: 0.09% with >1 birth record, 0.01% with >1 CHOP EHR record, 0.10% with >1 license

record, <0.01% with >1 death record. Overall, 0.2% (n=38,232) of all individuals had more than one record that should be unique.

Geocoding Residential Addresses

We geocoded the residential addresses of all licensed New Jersey drivers and all crash-involved drivers as well as each individual's most recent New Jersey residential address from other sources (when there was one). Records were prepared for geocoding the value for state was New Jersey or unknown and either 1) street, city, and ZIP code were populated; 2) street and city were populated (and ZIP was null); 3) street and ZIP were populated (and city was null); or 4) ZIP was populated (and street and city were null). Group 4 was included to geocode to the ZIP code level, although they could not be geocoded to the level of street address or point. Crash-involved driver records that did not meet this threshold generally belonged to parked/driverless or hit-and-run vehicles. After standard pre-processing cleaning steps, we conducted the geocoding process within the automated geocoding engine in ArcGIS 10.5 (Esri, Redlands, CA). The default geocoding options were used and include spelling sensitivity, minimum candidate score, and minimum match score. Unmatched addresses that remained after automated geocoding were reviewed and manipulated manually, then reprocessed and geocoded. Geocoding results were compared against Google Maps and included coordinate values (latitude and longitude).

Overall, 20 million distinct addresses were processed for geocoding; among those 93% were successfully geocoded to an address point or street address. For those records, their coordinate values were mapped to both 2010 and 2020 U.S. Census tract boundaries. When including addresses geocoded to the ZIP code level (or a more precise unit), 97% of those processed were geocoded. Among the 24 million people in the NJ-SHO data warehouse, over 18 million (77%) had at least one address processed. Among those 18 million, 95% had at least one address that was geocoded to an address point or street address. Among those who did not have at least one address that could be processed, 99% had a record in only one source. Those sources most commonly were crash-involved driver (36%, which includes hit-and-run and parked vehicles) or crash-involved passengers (35%, which includes passengers from crashes prior to 2009, when identifiers were not collected for passengers), followed by hospital records (18%).

Geocoding Crash Locations

We received geocoded latitude and longitude coordinates for crash locations via the New Jersey Department of Transportation. For crashes in the years 2004 through 2019, 79% have location coordinates. These locations have been mapped to census tract using both 2010 and 2020 U.S. Census tract boundaries.

Crash Distance Calculation

For crashes in the years 2004 through 2017, we calculated the distance between the crash location and the residential address of the driver(s), if the driver's address was geocoded to the point address or street address (i.e., not to the ZIP code level). We calculated both the network and Euclidean distances between the residential address and the crash location. The network distance was calculated using the Network Analyst extension in ArcGIS Pro 2.4.0 with the 2018 Esri Business Analyst North American routing street dataset. The route solver calculated drive distances (in miles) and time (in minutes) along the road network for each pair of driver and crash locations. Each solution is solved for the quickest route using the posted speed limit and restriction (e.g., one-way streets) within the routing street dataset. The solver also used the hierarchical road characterization with preferences made towards primary roads (highways and freeways). For the 7.7 million crash-involved drivers in the years 2004 through 2017, 4.7 million (61%) had both crash location and residential address available and consequently have distance calculated. We will replicate this process for crashes years 2018 and 2019.

Injury Classification

We categorized crash-involved individuals as injured using both crash report and hospital discharge data. For additional information and an application of this classification, see Lombardi et al. 2021.

An individual was considered injured based on the crash report if their physical condition was noted by the investigating officer as "fatal injury" ("killed" for years 2004–2017), "suspected serious injury" ("incapacitated" for years 2004–2017), "suspected minor injury" ("moderate injury" for years 2004–2017). These are equivalent to K, A, and B, respectively, in the KABCO scale.

For individuals involved in crashes from 2016 through 2019, we also classified them as injured based on hospital discharge data.

Hospital discharge data from inpatient, outpatient, and emergency department visits includes admission date and International Classification of Diseases 10th Revision, Clinical Modification (ICD-10-CM) diagnosis codes. We applied the 2019 ICD-10-CM external cause of injury matrix to all codes for each hospital discharge record to identify crash involvement. Those with an external cause of injury of "motor vehicle-traffic" according to the matrix were considered to have been involved in a crash. Among crash-involved individuals identified through these codes, we examined all ICD-10-CM diagnosis codes included in records within the three days of the crash to classify whether they were injured. Crash-involved individuals with an injury-related diagnosis code within the ICD-10-CM chapters "S" and "T" (except frostbite, poisoning, toxic effects,

unspecified/other external causes, and complications of surgical or medical care) were considered injured.

Race and Ethnicity

In New Jersey, race and ethnicity of crash-involved individuals are not collected on the crash report. Some individuals within the NJ-SHO have race and ethnicity information available via linked hospital, birth, and death data, as well as electronic health records for CHOP health care network patients who were ever residents of NJ. For individuals without this information, we applied a rigorous algorithm—utilizing a procedure called Bayesian Improved Surname Geocoding (BISG)—to estimate their probability of belonging to different race and ethnicity groups. Briefly, BISG combines information on last names with a US Census block group of residential addresses to estimate a posterior probability of membership in each of six mutually exclusive racial and ethnic categories (non-Hispanic White, Hispanic, non-Hispanic Black, non-Hispanic Asian or Pacific Islander, non-Hispanic multiracial, and non-Hispanic American Indian or Alaska Native). The sum of the six probabilities generated by the BISG algorithm equals 1; individuals are not assigned a specific race/ethnicity value. We rigorously validated this approach and our application of it to the NJ-SHO Data Warehouse, which increased the proportion of available race and ethnicity information among crash-involved drivers from 77% to 99% (Sartin et al. 2021).

Strengths of the Data Warehouse

Population Based. This resource provides a population-based perspective of the health and injury status of residents of New Jersey. That is, each integrated data source contains all individuals (including New Jersey residents and non-New Jersey residents) who experienced the events included in that data source. For example, the motor vehicle crash report data has information all police-reported crashes in New Jersey, which include all crash-involved vehicles and all crash-involved drivers, passengers, and other road users.

Rates. By having population-based data, we can count all events that occur in New Jersey and construct rates, which account for the size of the underlying population. Rates allow us to compare the magnitude of the health impact among groups of different sizes. For instance, we can compare the pedestrian crash rate who live in two different-sized counties to determine the relative difference between those two counties. Similarly, we can calculate annual pedestrian crash rates for New Jersey in order to examine trends over time, even if the size of population of New Jersey changes during that time period.

Ability to Use Quasi-Induced Exposure Methods. Because we have population-based crash data, we can use quasi-induced exposure (QIE) methods to estimate population prevalence of specific driving conditions or behaviors (Curry 2017). We can use the fundamental QIE assumption that non-responsible (not at-fault) drivers in clean multi-vehicle crashes (with one and only one responsible driver) are "randomly selected" by responsible drivers from the population of road users at the time and space of the crash. Thus, the characteristics of these non-responsible crash-involved drivers reasonably represent the characteristics of the underlying driving population.

Comprehensive Data Integration. All data sources included in the Data Warehouse are linked with all other data sources. We did not require linkage to a specific data source. This comprehensive data integration method allows the as much data for an individual from as many sources as possible to be used for research. We also have the ability to link additional data from other sources in the future.

Reduced Missingness. A single data source may be limited by missing data for important data elements. By linking multiple data sources, we have been able to leverage data from these sources to reduce the amount of missing data. In particular, we have been able to characterize race and ethnicity for crash-involved individuals, despite that information is not collected on crash reports, by linking crash reports with hospital discharge, birth, and death data (Sartin et al., 2021). Further, we have linked crash and hospital discharge data to better describe the severity and variety of injuries among crash-involved road users (Lombardi et al., 2022).

Ability to Geocode Residence. We were able to obtain residential address from multiple data sources at multiple time points, thus can identify an individual's most recent residential community compared to events of interest. By identifying communities in which impacted individuals live, we can begin to understand how neighborhood can influence health and injury. Approximately 17 million individuals have at least one residential address that was geocoded to an address point or street address.

Longitudinal Data. Using crash reports in isolation limits our understanding of the crash to events occurring just prior to, during, and immediately after the crash. The integrated data in the Data Warehouse allow us to expand this understanding from the minutes proximal to a crash to potentially the lifespan of individuals impacted by a crash. We can characterize pre-crash circumstances, like license status, or demographics that may influence crash risk. Further, we can examine post-crash consequences more in depth, such as repeat crash risk, injury severity, or death.

Highly Reliable Linkage. We have high confidence that the data integration was conducted validly. We thoroughly evaluated the quality of the linkage at each step of the process. The median match probability of accepted probabilistic matches was 0.999. The estimated true

match proportion for deterministic matches was 93% and the estimated false match proportion was 8%.

Limitations of the Data Warehouse

Limited Generalizability. Results using data from the Data Warehouse may have limited generalizability to populations beyond New Jersey. New Jersey has the highest population density of any U.S. state and is highly urbanized. In addition, New Jersey has the oldest licensing age in the U.S.; residents must be at least 17 to be eligible for a probationary (restricted) driver's license. Thus research results may not be applicable to states with different geographic and demographic characteristics.

Data Lag. It takes time for data to be reported to the agencies managing these data included in the Data Warehouse, and then these agencies need to compile and prepare data prior to release. Once we receive the data, we need additional time to harmonize, link, and prepare the data for analyses. Consequently, it can be several years from when events (e.g., crashes, hospitalizations) occur to when those events appear in the Data Warehouse.

Migration Out of New Jersey. For individuals with a New Jersey residence, we do not know if and when a person moved out of New Jersey. For instance, we can only remove drivers from the enumeration of the number of licensed drivers (used to calculate driver crash rates) when their license expires as the licensing data does indicate if they moved out of state. Out migration may have a greater impact on certain population groups that are more mobile, like adolescents and young adults who may move out of state for college.

Definition of a Reportable Crash. In New Jersey, a motor vehicle crash is required to be reported if it meets the following criteria: if it resulted in injury to or death of any person or damage to property of any one person in excess of $500. The specification of a reportable crash differs by state, including different thresholds for personal injury or property damage.

Humans Collect Data. Ultimately, all data sources included in the Data Warehouse are collected by humans, who are fallible. Each data source has its own limitations regarding the reliability and validity of its data. We believe that the data in the Data Warehouse are of very high quality, particularly because these data sources have been extensively used for both administrative purposes and research studies. For instance, crash reports are completed by the responding law enforcement officer, who may make mistakes and record inaccurate or incomplete data, despite receiving comprehensive training on the requirements to consistently and accurately record their investigation of the crash. Further, information can only recorded for variables and values that are present and/or required by the data collection system. As an example, the

decoded VIN data available through NHTSA's vPIC includes only data voluntarily submitted by vehicle manufacturers.

Data Stewardship

Legal agreements. All NJ-SHO Center activities are bound by legal agreements (e.g., Data Use Agreement, Memorandum of Agreement) between CHOP and data owners, which establish approved uses of these data as well as stringent security measures, including data transfer, storage, sharing, and release. Within the United States, release of traffic safety data is supported by the 1994 federal Driver's Privacy Protection Act, which exempts restrictions on release of data when used for research purposes; data accessibility varies by state. Partnerships between the NJ-SHO Center and external collaborators may also require Collaborative Research Agreements (CRA) that describe authorized use of these data, depending on the nature of the partnership.

Approval by Institutional Review Boards. Linkage and research activities have been reviewed and approved by the Institutional Review Boards (IRBs) at CHOP and Rowan University, the IRB agent for the New Jersey Department of Health. These approvals govern what type of data we have access to, the research questions we may pursue with the data, and who has access to data. Additionally, anyone who has access to NJ-SHO data must remain current on research, ethics, and compliance training.

Data Security. Datasets containing protected health information (e.g., name, street address, driver license number) are stored on a secure drive and accessible only to research staff who are integral to the linkage process. Once linked, individuals and their records are given new, random identification numbers that replace original identification numbers. Analytic datasets are stripped of direct identifiers and original identification numbers; these datasets are stored on a separate drive that also has limited access. Individual-level data are only accessible by members of our research team or workforce approved by the IRBs to work with data. Only aggregate data will be released to the public, such as through our Dashboard and in publications. No personal identifiers, protected health information, or otherwise identifiable data will be released. Additionally, small counts (generally counts of 10 or fewer) will be redacted or suppressed from aggregate data to avoid the possibility that individuals can be re-identified. This may mean redacting or suppressing additional counts if not doing so would allow small values to be derived. Note that values of zero are not subject to this policy.

Partnerships with External Collaborators. Potential partnerships with external academic and non-academic collaborators will be evaluated on an individual basis. The NJ-SHO Center team and potential collaborators will meet to discuss the parameters of the partnership, which could range from custom data summaries or data analyses conducted by the Center team to providing

access to select individual-level data for collaborators to analyze (subject to data stewardship requirements described above).

Overview of the NJ-SHO Data Dashboard

The NJ-SHO Data Dashboard is a free interactive dashboard that presents aggregate data on the traffic safety experience of New Jersey residents (https://njsho.chop.edu/data/data-dashboard). The Dashboard is powered by the NJ-SHO Data Warehouse and offers users the ability to compare transportation safety and injury metrics over time, by community, and by population characteristics. The Dashboard offers 10 views which provide crash information for drivers, pedestrians, and bicyclists in each New Jersey county. Each view can be downloaded as a PDF or image for easy incorporation into reports and presentations. More information about how to use the dashboard can be found on the NJ-SHO Center website: https://njsho.chop.edu/data/how-use-dashboard.

Suggested Reference for this Document

If you use information from this document for presentations, publications, or other reports, please acknowledge it as a source using the citation below. We also request that you provide the title and full citation for any publications, research reports, or educational materials making use of data or documentation from the NJ-SHO Center for Integrated Data by emailing us at njsho@chop.edu.

New Jersey Safety and Health Outcomes Center for Integrated Data: Technical Documentation. https://njsho.chop.edu/data/about-data. Accessed on [Month, Day, Year].