Smart and Safe Kenya Transport (smarTTrans)

Using technology, analytics and policy experiments to save lives and promote inclusive growth

Image by Nina Stock from Pixabay

Context

Road traffic crashes (RTCs) devastate the lives of victims and their families, while causing economic destruction in terms of property and lost earnings, especially for the poor. With 1.35 million deaths per year, RTCs are the primary cause of death among young people and the eighth leading cause of death globally (WHO 2018). They cause 50 million nonfatal injuries, imposing heavy burdens on health systems. The overall annual costs are estimated between 1% and 5% of GDP in developing countries (World Bank 2014). Without urgent attention to this issue, the Sustainable Development Goal (SDG) target 3.6 to halve RTC fatalities by 2020 will not be met (WHO 2018). Policy action must focus on Africa, where road traffic deaths are the highest (26.6 deaths per 100,000 people), and in contexts like Kenya, where these rates rank among the highest in the world. This is also where pedestrians make up 37% of RTC fatalities, compared to the 22% incidence worldwide (WHO 2018). In Nairobi, where the Smart and Safe Kenya Transport (smarTTrans) pilot takes place, walking is the main mode of commuting (Pendakur 2005), especially among the poor. Here, pedestrian deaths represent 71% of all RTC deaths reported in police crash records (2012–18).

Poor data, no analytics, and the lack of policy experimentation constrain the government's ability to develop policies and interventions to effectively regulate, monitor, and enforce road safety. In Kenya, the WHO estimates that the number of RTC fatalities is 4.5 times the figure in the official registry. The difference is due to lack of information on the location of crashes and their characteristics, on the dispatch of ambulances and the destination of the victims, on the results of emergency response at trauma centers, and on vital statistics about people who might die at home from complications. As a result, data analytics—which would provide policy makers with a better understanding


Image by Nina Stock from Pixabay

of the problem—are lacking. Solutions are imported from more developed contexts without adaptation to local conditions and factors, and thus have limited impact. Policy experiments that could arise from better data and analytics are limited. Developing analytical tools is of paramount importance if Kenya is to create context-appropriate solutions to road safety.

smarTTrans and Its Expected Policy Influence

The Kenyan government and the World Bank started the smarTTrans collaboration to develop data, analytics, and evidence to improve the country's road safety and transport policy. This joint effort is supported by the Kenyan ministries, departments, and agencies in the transport sector, as well as by development partners, and aims to test innovations using data, analytics, and experiments to support more effective transport policies. The first focus of this collaboration, chosen in 2018 by these actors, is to significantly improve road safety.

The strategy we adopted is to: (i) invest in innovative ways of collecting high-frequency, real-time data, (ii) understand the local constraints to road safety, and (iii) experiment with behavioral, infrastructure and institutional approaches to improve road safety. To this end, smarTTrans aims to (a) build a pilot interactive data platform for road safety and urban mobility, with particular attention to advancing a measurement framework that brings together data on crashes and victims; (b) identify blackspots (delimited locations with the highest number of crashes or crash outcomes); (c) identify risk factors and characteristics of blackspots; and (d) evaluate interventions to improve road safety and urban mobility.

An important innovation of this project stems from its systematic approach to building the data systems that can support the design and implementation of road safety policies in data-resource-constrained contexts and develop the tools that can be used for both a national scale-up and replications in other contexts. The project aims to directly inform government agencies such as the National Transport and Safety Authority and the National Police Service on interventions that can have high returns and how to target them in time and space to maximize effectiveness. The example of Kenya will be carefully documented to provide a roadmap on building innovative data systems in data-scarce settings to monitor and manage road safety policy, and how to use data to extract relevant information for more effective policy making. The tools and instruments developed in this pilot will be made available for deployment across the African continent by leveraging the portfolio of transport projects of the World Bank and other development agencies. Addressing the urgent need for data systems to facilitate appropriate planning and resource allocation is a necessary step to improve road safety.

smarTTrans Outputs: Road Safety

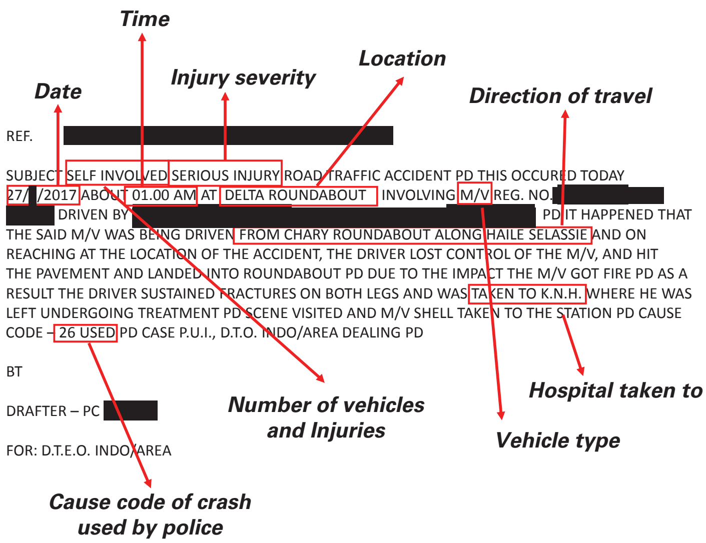
Nairobi Crash Map (Ongoing)

Using multiple data sources and machine learning algorithms, the smarTTrans team built the first georeferenced multiyear crash dataset and crash visualizations for the city of Nairobi. This constitutes the first step in building data and data systems leading to better road safety. The data allow for the identification of crash location and time, helping identify the riskiest locations based on actual crashes and crash outcomes. Having five or more years of data is critical to support any good analysis of road safety (US FHA 2017). The data include two types of sources:

1. Administrative: National Police Service paper-based crash reports between 2012 and 2018 for all 14 police stations in Nairobi (figure 1)
2. Crowdsourcing (contributions from a large online community): (i) Twitter handle Ma3Route (2012–19) (figure 2); (ii) Waze, a GPS navigation software app that provides navigation information but also collects crowdsourced data on crashes (2018–19); and (iii) Sendy, a motorcycle delivery phone platform (2018–19)

About 10,000 police and 30,000 crowdsourcing reports have been accessed. Paper reports from police, such as the example provided in figure 1, have been digitized, all relevant data coded into variables as demonstrated in the figure and the location of the crashes geolocated to enable maximum use of the data for informing policy. The smarTTrans team developed machine learning algorithms using natural language processing to extract relevant information from crowdsourcing reports to enable their geolocation (figure 2).

Some reports correspond to the same crash but many are unique, allowing for a better picture of all crashes occurring across the city and for learning about the types of biases that the different sources may have. For instance, 98% of the police reports contain crashes with a death or injury, indicating that focus is on attending and recording the most severe crashes. This also represents the best data to analyze crash outcomes, but it is currently collected using paper forms, limiting its use for timely monitoring and action. Crowdsourcing allow for real-time reporting, which facilitates timely monitoring and action. However, they also contain less standardized and verifiable details on crash outcomes, and decline late at night when many deadly crashes


REF. [redacted]

SUBJECT SELF INVOLVED SERIOUS INJURY ROAD TRAFFIC ACCIDENT PD THIS OCCURED TODAY 27/[redacted]/2017 ABOUT 01.00 AM AT DELTA ROUNDABOUT INVOLVING M/V REG. NO.[redacted] DRIVEN BY [redacted] PD IT HAPPENED THAT THE SAID M/V WAS BEING DRIVEN FROM CHARY ROUNDABOUT ALONG HAILE SELASSIE AND ON REACHING AT THE LOCATION OF THE ACCIDENT, THE DRIVER LOST CONTROL OF THE M/V, AND HIT THE PAVEMENT AND LANDED INTO ROUNDABOUT PD DUE TO THE IMPACT THE M/V GOT FIRE PD AS A RESULT THE DRIVER SUSTAINED FRACTURES ON BOTH LEGS AND WAS TAKEN TO K.N.H. WHERE HE WAS LEFT UNDERGOING TREATMENT PD SCENE VISITED AND M/V SHELL TAKEN TO THE STATION PD CAUSE CODE – 26 USED PD CASE P.U.I., D.T.O. INDO/AREA DEALING PD

BT

DRAFTER – PC [redacted]

FOR: D.T.E.O. INDO/AREA

Figure 1. Police crash report illustration

Note: This example police report demonstrates the relevant data that was extracted and coded as variables to produce a single dataset of all reports that can be analyzed to inform policy. All identifiable information has been removed and location information has been changed to provide an illustration while maintaining privacy.

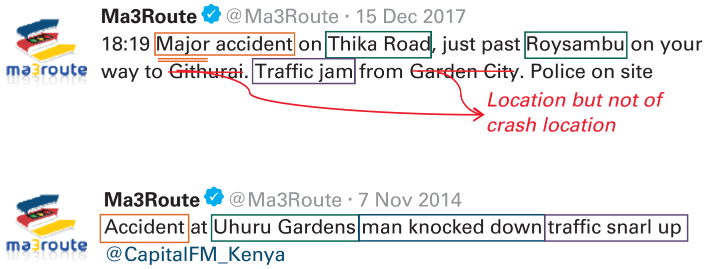


Figure 2. Tweet crash report illustration

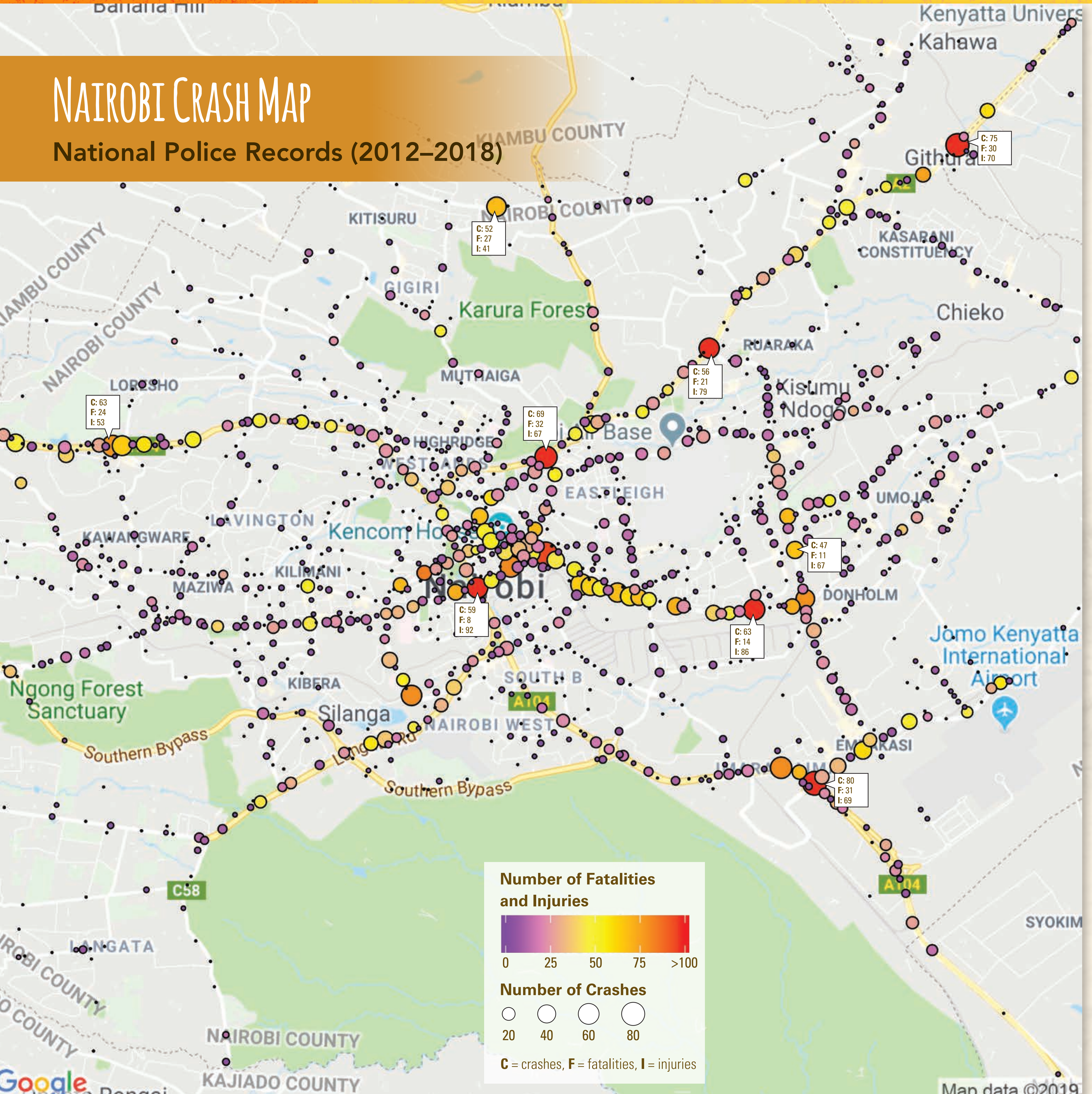
Note: These tweets demonstrate the type of crowdsourcing reports accessed. The machine learning algorithms developed by the team extract relevant information on location, as well as discard location variables that are irrelevant to enable unique geolocation of each report.

occur. The goal is a system that integrates multiple sources and produces analytics in real time to facilitate early identification of adverse events and risk factors, and to reduce reaction time, improving planning and response.

Using crowdsourcing is an inexpensive way to collect data at scale on all crashes across a city to complement the administrative data. Administrative and additional data collection have also allowed us to ground truth the crowdsourcing in order to develop better algorithms. Through this process, we have worked to improve the administrative data, exploring best-practices for obtaining digital information. These methods can be adapted for other contexts for effective policy action.

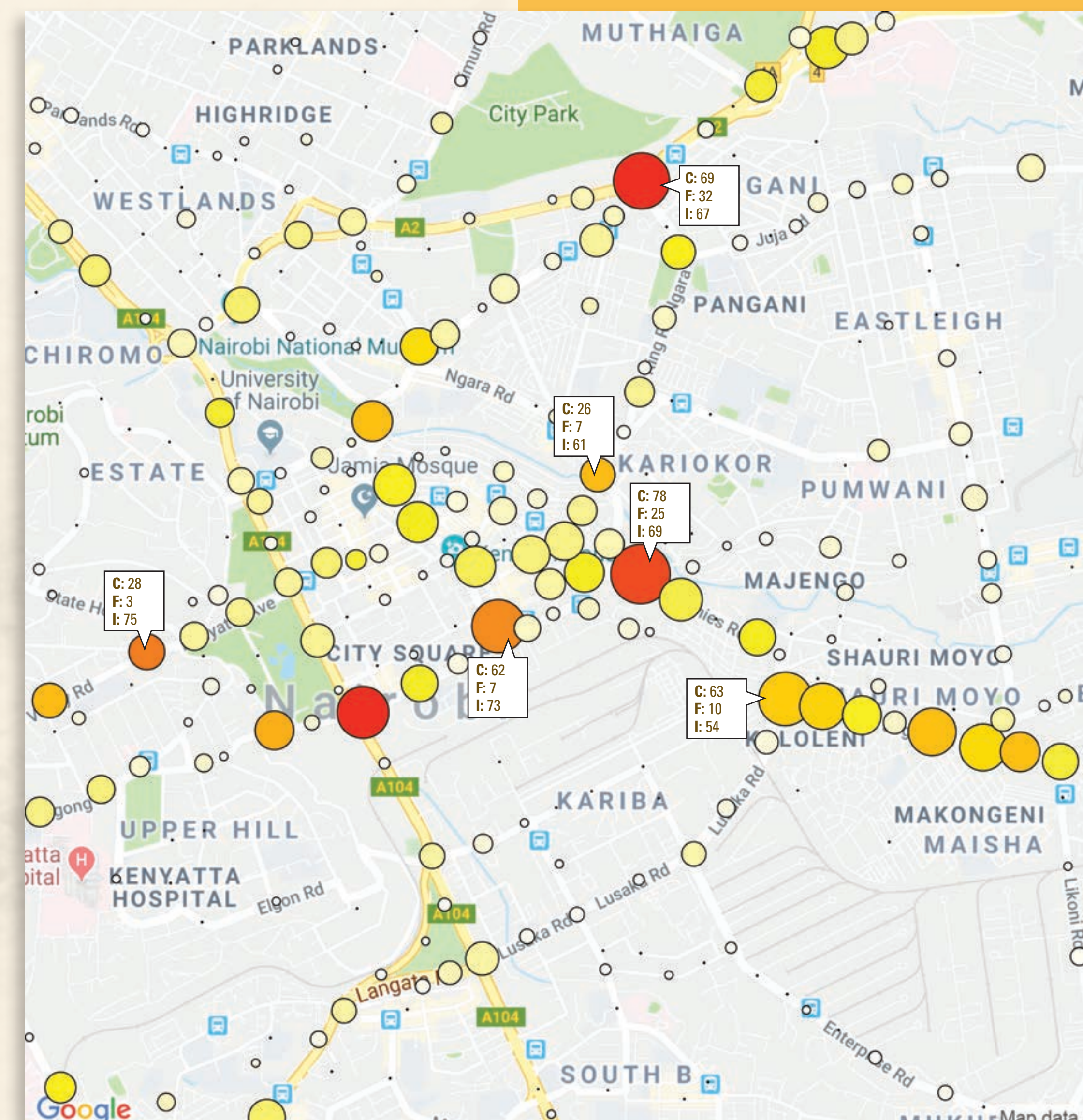
Nairobi Crash Map

National Police Records (2012–2018)



71% of road traffic deaths are **pedestrians**

Central Business District



200 **locations** (blackspots) represent **52%** of crashes and **55%** of deaths

Blackspots are locations where a high number of crashes, deaths or injuries have occurred in a delimited space and time.

Blackspot Identification (Ongoing)

Blackspots are crash locations or hazardous road locations that usually indicate that there are common causes for the crashes. In developed countries such as Australia and New Zealand, safety strategies aimed at risky locations have led to reductions in the concentration of crashes; most fatal and serious crashes now occur at locations on roads with no other injury crashes reported in the previous five years (Austroads 2015). In countries where crashes are concentrated, which is likely the case for the deadliest countries in terms of RTCs, the focus of road-safety analysis and policy should therefore be on blackspots first. The first step is the identification of such locations and classification of risk, which will depend on the context. Next, a diagnosis of the main risk factors and typology of blackspots, prioritization, and monitoring and evaluation of the effectiveness of different interventions is needed to maximize the benefits in scarce-resource contexts.

For Nairobi, six complete years of police crash reports were used from 2013 to 2018 to help identify these blackspots. The data includes 9,315 reports from the 14 National Police Stations across the city. These paper reports provide a description of where the crash occurred but no GPS to enable visualization on a map. Of these, 89% of crashes were geolocated using double data entry by local coders (the remaining reports lacked enough information to locate the crash). The smarTTrans team developed a clustering algorithm to identify blackspots. The preliminary analysis defines a blackspot in two steps. First, we identified those locations with more than 15 crashes within 300 meters during the analysis period, identifying 150 blackspots. A second step considered additional crash outcomes. Locations that contained less than 15 crashes but have either at least 21 injuries or at least 6 deaths were also included as blackspots, since these crash outcomes may be related to risk factors at those locations. This added 50 blackspots for a total of 200.

Risk Factors Associated with Blackspots (Ongoing)

A system of indicators was designed as part of this first smarTTRans pilot based on internationally recognized standards. They were discussed with the Kenyan government agencies to analyze risk factors at the 200 blackspots described above. The indicators aim to measure multiple dimensions around the road segments that compose the blackspots, and use multiple data sources, including onsite data surveys and several months of real-time crowdsourcing through Uber and Waze. Examples of indicators included and their source are as follows:

- Road physical attributes such as location of sidewalk (if any) relative to the road, quality of delineation, whether there are safety barriers along the road, or whether there are pedestrian crossings and whether they have clear markings on the road and signalization (survey following the International road Assessment Programme (iRAP) road coding indicators)
- Road user flows, including cars, pedestrians, motorbikes, and trucks (surveys and observation)
- Economic activity and land usage around the blackspots, such as schools, public transportation stops, and markets (web-scraping of landmarks from Google Maps)
- Compliance with speed limits based on big data built with average speed in time intervals over each day for several months (Uber and Waze real-time web-scraping around the blackspots)
- Compliance with select road safety practices, such as motorbike drivers' use of helmets and the use of pedestrian crossings (five-minute videos recorded at the blackspots)

This exercise will contribute to a structured set of indicators that can be used for crash blackspot assessments and road safety planning. They also aim to advance a framework to monitor and improve road safety in Kenya and similar contexts.

Pilots of Interventions to Improve Road Safety (Expected)

The results from the outputs just noted are the basis for discussions with Kenyan government agencies and all stakeholders regarding the types of interventions that should be tested in order to address some of the high-risk problems identified in Nairobi. Taking into consideration the constrained resources and in line with the Kenyan Road Safety Strategic Plan, this phase will continue improving the data platform, developing a low-cost data platform (systems to assess deaths and injuries outside of the crash site, which will improve the ability to assess the extent and costs of road crashes), and incorporating interventions that address the risk factors identified. This work will be defined with the Kenyan government to identify feasible interventions such as changes in road attributes (engineering interventions), road safety management, enforcement and behavioral interventions.

smarTTrans Early Findings

The first result from merging police and crowdsourced data is that there are at least 50% more crashes reported by bystanders than are reported in the police crash reports. Crowdsourcing crash data, in other words, can help us identify more of the universe of crashes, gain a better understanding of their characteristics, develop policies to address them as well as estimate the economic costs of RTCs.

Second, a statistical analysis of the 2013–2018 police reports reveals the following:

Finding 1: 71% of RTC deaths in the city of Nairobi covered by the main 14 Police stations are pedestrians. Pedestrian deaths reached 79% in 2014 and adjusted downward to 66% by 2017.

Finding 2: While crashes occur in 1400 different locations across the city, 200 locations represent 55% of the deaths and 53% of the injuries across the city. These locations represent 14% of all locations where severe crashes have occurred in the six years under review. This high concentration of crashes and negative crash outcomes is an opportunity to prioritize these locations for policy action and substantially remedy the loss of life in the city.

Finding 3: 35% of deaths occur within twenty meters of matatu stages. This staggering number is a call for action. Developing a better understanding of how to regulate and enforce matatu flows, and driver and pedestrian behavior, while at the same time improving the infrastructure at these sites will be required to guide policy action.

Finding 4: Deaths and injuries are concentrated between 5 a.m. and 8 a.m. and between 5 p.m. and 11 p.m., representing 53% (deaths) and 50% (injuries) of the total. The deadliest times for pedestrians are at night between 7 p.m. and midnight when 41% of pedestrian deaths occur. Implicating factors that need to be investigated further could include poor visibility, drunk driving and speed. Cross examining location and time of crashes could help develop a strategy for remedying street lighting and enforcement of speed and alcohol regulation.

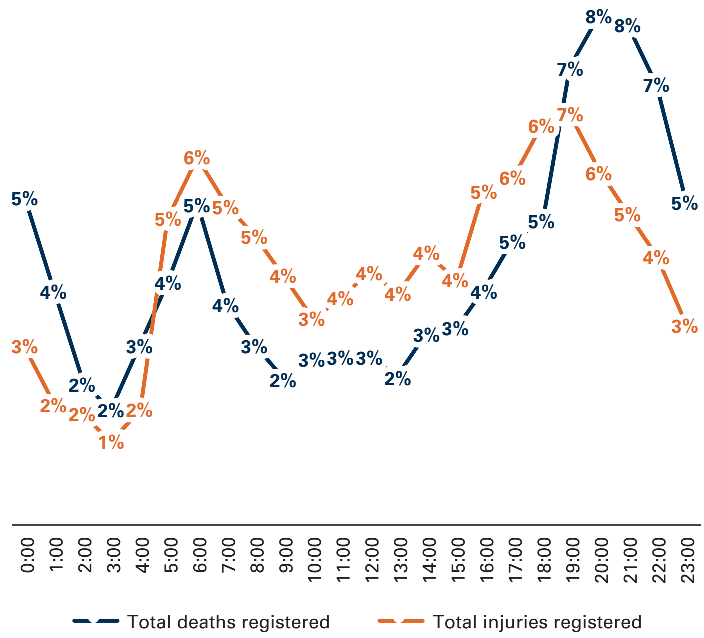


Figure 3. Deaths and Injuries Reported in Police Crash Reports by Time of Day, 2013–18

Moving Forward

The early results from the analysis of the data we collected and digitized show how critical it is to clearly identify the problems and prioritize policy interventions around them. This would avoid dissipating scarce fiscal resources and instead use them strategically to narrow down where policy action can be most effective in reducing mortality. Targeting 14% of locations where severe crashes have occurred and focusing enforcement efforts in the most dangerous hours of the day could potentially halve the number of deaths in the city. This will not only allow Kenya to meet its SDG targets but avoid enormous suffering and economic costs.

Moving forward, the team will conduct a more comprehensive analysis of both police and crowdsourced data as well as a detailed analysis of the survey conducted on the priority blackspots. The latter will be used to understand what makes these locations so dangerous and what might be the menu of interventions and investments we might want to test and implement to curb crashes and deaths.

smarTTrans Team and Funding

The research team comprises of Guadalupe Bedoya, Arianna Legovini, Sveta Milusheva, Sarah Williams, and Robert Marty. This work is in collaboration with the World Bank Transport Global Practice, the National Transport and Safety Authority, the National Police Service, and other Kenyan government agencies. Amy Dolinger, Meyhar Mohammed, and Robert Tenorio provided research assistance throughout the project. Elizabeth Resor developed the first geocoded crash map and provided field support. Funding was provided by the DIME Impact Evaluation to Development Impact ieConnect program, which has been funded with UK aid from the UK government, and the Transport Global Practice at the World Bank.

References

Austroads. 2015. *Guide to Road Safety*, part 8, "Treatment of Crash Locations." Sydney: Austroads. https://austroads.com.au/safety-and-design/road-safety/guide-to-road-safety.

Gonzales, Eric J., Celeste Chavis, Yuwei Li, and Carlos F. Daganzo. 2009. "Multimodal Transport Modeling for Nairobi, Kenya: Insights and Recommendations with an Evidence-Based Model." Berkeley, CA: UC Berkeley. https://escholarship.org/uc/item/6dv195p7.

Pendakur, V. Setty. 2005. "Non-Motorized Transport in African Cities. Lessons from Experience in Kenya and Tanzania." SSATP Working Paper 80. Africa Transport Policy Program, World Bank, Washington, DC. https://www.ssatp.org/en/publication/non-motorized-transport-african-cities-lessons-experience-kenya-and-tanzania.

US FHA (US Federal Highway Administration). 2017. *Road Safety Fundamentals: Concepts, Strategies, and Practices that Reduce Fatalities and Injuries on the Road*. Washington, DC: US FHA. https://rspcb.safety.fhwa.dot.gov/rsf/.

WHO (World Health Organization). 2010. "Data Systems: A Road Safety Manual for Decision-Makers and Practitioners." Geneva: WHO. https://www.who.int/roadsafety/projects/manuals/data/en/.

WHO. 2018. *Global Status Report on Road Safety 2018*. Geneva: WHO. https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.

World Bank. 2014. *Transport for Health: The Global Burden of Disease from Motorized Road Transport*. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/984261468327002120/Transport-for-health-the-global-burden-of-disease-from-motorized-road-transport.

IE CONNECT FOR IMPACT

ieConnect has over 30 ongoing impact evaluations across 19 different countries. The IEs focus on urban mobility, transport corridors, road safety, and rural roads sectors with thematic emphasis on gender, female economic empowerment, and fragile situations. From the ieConnect program we will learn how to improve the availability and quality of data that can be used for measuring the impact of transport projects and generate evidence that can be used to improve decision making for transport investments in the long-term. The ieConnect for Impact program is a collaboration between the World Bank's DIME group and the Transport Global Practice. This program has been funded with UK aid from the UK government.