Verification of Crowdsourced Crash Reports Survey 2018

World Bank Group

report_generated_on: January 5, 2026

visit_data_catalog_at: http://catalog.ihsn.org/

Identification

SURVEY ID NUMBER
KEN_2018_VCCR_v01_M

TITLE
Verification of Crowdsourced Crash Reports Survey 2018

ABBREVIATION OR ACRONYM
VCCR 2018

COUNTRY

Name	Country code
Kenya	KEN

ABSTRACT
The purpose of the 2018 Verification of Crowdsourced Crash Reports survey is to physically verify road traffic crash (RTC) reports that were reported by bystanders in Nairobi, Kenya using Twitter/Ma3Route. Ma3Route is a mobile/web/SMS platform that crowdsources transport data and provides users with information on traffic, RTCs, matatu directions and driving reports. Users post RTC or traffic information to Ma3Route, where Ma3Route then publishes the post on Twitter. In this survey, we sought to verify the accuracy of crowdsourced information as a source of RTC data.

KIND OF DATA
--

UNIT OF ANALYSIS
Road traffic crash reports

Version

VERSION DESCRIPTION
- v01

Coverage

GEOGRAPHIC COVERAGE
Crowdsourced road traffic crash (RTC) reports were verified in Nairobi, Kenya

UNIVERSE
Road traffic crash reports in Nairobi, Kenya

Producers and sponsors

PRIMARY INVESTIGATORS

Name
World Bank Group

Data collection

DATES OF DATA COLLECTION

Start	End

2018-07-21	2018-09-01

DATA COLLECTION MODE
Other [oth]

DATA COLLECTION NOTES
The process by which data for the 2018 Verification of Crowdsourced Crash Reports Survey was collected is decribed below:

1. An algorithm scraped tweets from @Ma3Route in real time and used a machine learning algorithm to detect RTC related tweets. When the algorithm detected a RTC tweet, the tweet was sent to a field coordinator for the project. The field coordinator for the project confirmed whether the tweet reported a crash with sufficient location information and, if so, made a request on the Sendy Ltd platform for a driver to go to the RTC location. Sendy Ltd is a Kenyan on-demand logistics startup that offers a marketplace for deliveries and logistics services using a mobile application that connects clients to motorcycle riders and vehicles.
2. Sendy drivers that were a part of this survey (15 drivers) were alerted of the request. A Sendy driver could choose to accept the request, then would travel to the location.
3. Once a Sendy driver arrived at the RTC location, they filled out a short survey verifying the RTC—if there was no RTC nearby, they were instructed to ask a passer-by if one occurred recently. Sendy drivers were instructed to not go near the crash, and to not interfere with any first responders.

DATA COLLECTORS

Name
Sendy Ltd

Access policy

CONTACTS

Name	Affiliation	Email
Robert Marty	World Bank	rmarty@worldbank.org

CONFIDENTIALITY
Users of the data agree to keep confidential all data contained in these datasets and to make no attempt to identify, trace or contact any individual whose data is included in these datasets.

ACCESS CONDITIONS
Licensed access

CITATION REQUIREMENTS
Use of the dataset must be acknowledged using a citation which would include:
- the Identification of the Primary Investigator
- the title of the survey (including country, acronym and year of implementation)
- the survey reference number
- the source and date of download

Example,

World Bank Group. Kenya - Verification of Crowdsourced Crash Reports (VCCR) Survey 2018. Ref. KEN_2018_VCCR_v01_M. Dataset downloaded from [url] on [date].

Disclaimer and copyrights

DISCLAIMER
The user of the data acknowledges that the original collector of the data, the authorized distributor of the data, and the relevant funding agency bear no responsibility for use of the data or for interpretations or inferences based upon such uses.

Metadata production

DDI DOCUMENT ID
DDI_KEN_2018_VCCR_v01_M_WB

PRODUCERS

Name	Abbreviation	Affiliation	Role
Development Economics Data Group	DECDG	World Bank	Documentation of the study

DATE OF METADATA PRODUCTION
2020-11-20

DDI DOCUMENT VERSION
Version 01 (November 2020)

data_dictionary

Data file	Cases	variables
crash_verification	73	4

Data file: crash_verification

Cases:	73
variables:	4

variables

ID	Name	Label	Question
V1	observe_crash	Do you observe a crash at the specified location?	Do you observe a crash at the specified location?
V2	why_no_crash	Why is there no crash?	Why is there no crash?
V3	datetime_tweet	Date/time of tweet (East Africa time)	Date/time of tweet (East Africa time)
V4	suggested_landmark	Suggested landmark to travel to, provided by algorithm	Suggested landmark to travel to, provided by algorithm

total: 4

OBSERVE_CRASH: Do you observe a crash at the specified location?

Data file: crash_verification

Overview

Valid: 73 Invalid: 0
Type: Discrete Decimal: 0 Width: 1 Range: 0 - 1 Format: Numeric

Questions and instructions

LITERAL QUESTION
Do you observe a crash at the specified location?

CATEGORIES

Value	Category	Cases	
0	No	49	67.1%
1	Yes	24	32.9%

QUESTION POST TEXT
If No, go to next question.

WHY_NO_CRASH: Why is there no crash?

Data file: crash_verification

Overview

Valid: 49 Invalid: 24
Type: Discrete Decimal: 0 Width: 1 Range: 1 - 5 Format: Numeric

Questions and instructions

LITERAL QUESTION
Why is there no crash?

CATEGORIES

Value	Category	Cases	
1	Crash Cleared	42	85.7%
2	Never a Crash	6	12.2%
3	Cannot Reach	0	0%
4	Crash Nearby	1	2%
5	Other	0	0%
Sysmiss		24	

DATETIME_TWEET: Date/time of tweet (East Africa time)

Data file: crash_verification

Overview

Valid: 73 Invalid: 0 Minimum: 1847793509000 Maximum: 1851412044000 Mean: 1849372936041.1
Type: Continuous Decimal: 0 Width: 13 Range: 1847793509000 - 1851412044000 Format: Numeric

Questions and instructions

LITERAL QUESTION
Date/time of tweet (East Africa time)

SUGGESTED_LANDMARK: Suggested landmark to travel to, provided by algorithm

Data file: crash_verification

Overview

Valid: 73 Invalid: 0
Type: Discrete Width: 17 Range: - Format: character

Questions and instructions

LITERAL QUESTION
Suggested landmark to travel to, provided by algorithm

CATEGORIES

Value	Category	Cases	
NO LANDMARK		9	12.3%
airport junction		1	1.4%
airtel		2	2.7%
allsops		1	1.4%
alsops		1	1.4%
avic		1	1.4%
bangladesh		1	1.4%
belle vue		2	2.7%
cabanas		1	1.4%
capital centre		1	1.4%
city cabanas		1	1.4%
commercial		1	1.4%
foot bridge		1	1.4%
gari		1	1.4%
gateway mall		4	5.5%
general motors		1	1.4%
gigiri		1	1.4%
githurai		1	1.4%
gm		1	1.4%

gsu		1	1.4%
hamza		1	1.4%
hilton garden		1	1.4%
hilton garden inn		1	1.4%
imara daima		1	1.4%
integrity centre		1	1.4%
james gicuru		1	1.4%
kangemi		1	1.4%
kapa		2	2.7%
karura forest		1	1.4%
kibera		1	1.4%
koja roundabout		1	1.4%
langata		1	1.4%
libra house		1	1.4%
mathare hospital		1	1.4%
mombasa cement		1	1.4%
muthaiga square		1	1.4%
muthangari		1	1.4%
nairobi school		1	1.4%
outering		2	2.7%
panari hotel		1	1.4%
panari sky		1	1.4%
pangani		2	2.7%
prison		1	1.4%
roysambu		2	2.7%
sabaki		1	1.4%
sacco		2	2.7%
safaricom		2	2.7%
standard		1	1.4%
stella		1	1.4%
uchumi		1	1.4%
uhuru		1	1.4%
utalii		1	1.4%
utawala		1	1.4%
valley		1	1.4%

study_resources

questionnaires

Documentation - 2018 Verification of Crowdsourced Crash Reports Survey

title	Documentation - 2018 Verification of Crowdsourced Crash Reports Survey
authors	World Bank
country	Kenya
filename	documentation.pdf