

# Kenya - Tweet IDs From Ma3Route 2012-2020

**World Bank**

Report generated on: March 24, 2021

Visit our data catalog at: <https://microdata.worldbank.org/index.php>

## Overview

### Identification

ID NUMBER  
KEN\_2012-2020\_TIDMA3\_v01\_M

### Version

VERSION DESCRIPTION  
-v01

### Overview

#### ABSTRACT

The purpose of the dataset is identify tweets from the @Ma3Route twitter handles that report road traffic crash reports and to identify the location of the crashes. Using the Twitter API, tweets were scraped from Ma3Route, which is a mobile/web/SMS platform that crowdsources transport data and provides users with information on on road traffic crash reports as well as traffic, matatu directions, and driving reports. This dataset provides the tweet IDs of tweets from Ma3Route from May 2012 until July 2020.

KIND OF DATA  
Observation data/ratings [obs]

UNITS OF ANALYSIS  
Road traffic crash reports

### Coverage

GEOGRAPHIC COVERAGE  
Kenya (primarily Nairobi)

UNIVERSE  
Tweets from Ma3Route

### Producers and Sponsors

#### PRIMARY INVESTIGATOR(S)

Name	Affiliation
World Bank	

### Metadata Production

#### METADATA PRODUCED BY

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

DATE OF METADATA PRODUCTION

2020-11-23

DDI DOCUMENT VERSION  
Version 01 (March 2021)

DDI DOCUMENT ID  
DDI\_KEN\_2012-2020\_TIDMA3\_v01\_M\_WB

## Sampling

No content available

## Questionnaires

No content available

## Data Collection

### Data Collection Dates

---

Start	End	Cycle
2012-05-01	2020-07-30	N/A

### Data Collection Mode

---

Other [oth]

### Data Collection Notes

---

Using the Twitter API, tweets were scraped from the twitter handle @Ma3Route. Ma3Route is a mobile/web/SMS platform that crowdsources transport data and provides users with information on traffic, road traffic crash (RTC), matatu directions and driving reports. Users post RTC or traffic information to Ma3Route, where Ma3Route then publishes the post on Twitter. Tweets were obtained in order to identify tweets that reported RTC. Tweets from May 2012 to July 2020 were scraped. Additional information on the data is provided in additional documents found under the 'Documentation' tab.

## Data Processing

No content available

## Data Appraisal

No content available



## File Description

## Variable List

**ma3route\_tweetids\_2012**

Content Tweet IDs 2012

Cases 4451

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V13	tweet_id	tweet_id	contin	numeric	

## ma3route\_tweetids\_2013

Content Tweet IDs 2013

Cases 27548

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

## Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V12	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2014**

Content Tweet IDs 2014

Cases 125704

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V11	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2015**

Content Tweet IDs 2015

Cases 231224

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V10	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2016**

Content Tweet IDs 2016

Cases 153003

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V18	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2017**

Content Tweet IDs 2017

Cases 93331

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V17	tweet_id	tweet_id	contin	numeric	



**ma3route\_tweetids\_2018**

Content Tweet IDs 2018

Cases 81570

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V16	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2019**

Content Tweet IDs 2019

Cases 93741

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V15	tweet_id	tweet_id	contin	numeric	

**ma3route\_tweetids\_2020**

Content Tweet IDs 2020

Cases 43030

Variable(s) 1

Structure Type:  
Keys: ()

Version

Producer

Missing Data

**Variables**

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V14	tweet_id	tweet_id	contin	numeric	



# tweet\_id (tweet\_id)

## File: ma3route\_tweetids\_2012

### Overview

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 2.04076177386783e+17-2.85873982866072e+17

Valid cases: 4451  
Invalid: 0  
Minimum: 20407600000000000  
Maximum: 28587400000000000

tweet\_id (tweet\_id)

File: ma3route\_tweetids\_2013

**Overview**

Type: Continuous

Format: numeric

Width: 11

Decimals: 2

Range: 2.8598722017375e+17-4.18148161983488e+17

Valid cases: 27548

Invalid: 0

Minimum: 285987000000000000

Maximum: 418147999999999926

tweet\_id (tweet\_id)

File: ma3route\_tweetids\_2014

**Overview**

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 4.18237961390404e+17-5.50436459392418e+17

Valid cases: 125704  
Invalid: 0  
Minimum: 418238000000000000  
Maximum: 550436000000000000

tweet\_id (tweet\_id)

File: ma3route\_tweetids\_2015

**Overview**

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 5.50481549703934e+17-6.82705040289563e+17

Valid cases: 231224  
Invalid: 0  
Minimum: 550482000000000268  
Maximum: 682705000000000000



tweet\_id (tweet\_id)

File: ma3route\_tweetids\_2016

#### Overview

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 6.82794423504122e+17-8.15345649700913e+17

Valid cases: 153003  
Invalid: 0  
Minimum: 682794000000000000  
Maximum: 8153460000000000268

tweet\_id (tweet\_id)

File: ma3route\_tweetids\_2017

**Overview**

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 8.15377609211007e+17-9.4769178260675e+17

Valid cases: 93331  
Invalid: 0  
Minimum: 815378000000000000  
Maximum: 947692000000000000

# tweet\_id (tweet\_id)

## File: ma3route\_tweetids\_2018

### Overview

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 9.58889716035699e+17-1.07988897448168e+18

Valid cases: 81570  
Invalid: 0  
Minimum: 958890000000000426  
Maximum: 1079890000000000000

# tweet\_id (tweet\_id)

## File: ma3route\_tweetids\_2019

### Overview

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 1.079908855189e+18-1.2121260008317e+18

Valid cases: 93741  
Invalid: 0  
Minimum: 1079909999999999684  
Maximum: 1212130000000000000

# tweet\_id (tweet\_id)

## File: ma3route\_tweetids\_2020

### Overview

Type: Continuous  
Format: numeric  
Width: 11  
Decimals: 2  
Range: 1.21216815345273e+18-1.28528374818823e+18

Valid cases: 43030  
Invalid: 0  
Minimum: 1212170000000000000  
Maximum: 1285280000000000000

# Documentation

## Technical documents

### Documentation - Tweet IDs from Ma3Route 2012-2020

---

Title Documentation - Tweet IDs from Ma3Route 2012-2020  
Author(s) World Bank  
Country Kenya  
Language English  
Filename documentation.pdf

---