Transport Starter Data Kit: Historical socio-transport data for Kenya

Authors

Naomi Tan[1,2], Robert Ambunda[3], Nikola Medimorec[3] Angel Cortez[3], Agustina Krapp[3], Erin Maxwell[1], John Harrison[1], Mark Howells[1,2]

Affiliations

1. Centre for Sustainable Transitions: Energy, Environment and Resilience, Loughborough University
2. Centre for Environmental Policy, Imperial College London
3. SLOCAT Partnership on Sustainable, Low Carbon Transport

Abstract

Data on transport activity is an important element for the development of national transport decarbonisation strategies. By having freight and passenger transport information, the impacts on vehicle and fuel consumption changes from replacing internal combustion engine vehicles with electric vehicles can be calculated. The development of a national decarbonisation strategy requires significant efforts. However, access to data is often a barrier to starting transport system modelling in developing countries, thereby causing delays. This article provides data that can be used to support a model for Kenya, which may act as a starting point for further model development and scenario analysis. The data are collected entirely from publicly available and accessible sources, focusing on national reports, statistical yearbooks, and academia.

Keywords

U4RIA

Transport data

Transport modelling

MAED

Kenya

Specifications Table

Subject	Transport
Specific subject area	Transport Data
Type of data	Tables Graphs
How data were acquired	Literature survey (databases and reports from international organisations; journal articles)
Data format	Raw and analysed
Parameters for data collection	Data collected based on inputs required to create an energy system model for Kenya
Description of data collection	Data were collected from the websites, annual reports and databases of international organisations, as well as from academic articles and existing modelling databases.
Data source location	Not applicable
Data accessibility	With the article and in a repository. Repository name: Zenodo. Direct URL to data: https://doi.org/10.5281/zenodo.6539210

Value of the Data

- The data can be used to develop national transport demand models to inform national investment outlooks and decarbonisation strategies.
- The data are useful for country analysts, policy makers, and the broader scientific community, as a zero-order starting point for model development.
- This data could be used to examine a range of possible transport pathways, in addition to the examples given in this study, to provide further insights into the evolution of Kenya's transport system.
- The data can be used both for conducting an analysis of transport activity and emissions, but also for capacity building activities.
- The data can be used as a call to action in addressing transport data gaps and establishing parameters for data collection to improve the consistency of transport-climate research in these countries.

1. Data Description

The data provided in this paper can be used to support the development of a transport model for Kenya. The data provided were collected from publicly available sources, including statistical yearbooks, transport ministry reports, statistics from national authorities and affiliated research institutions, academia, and journal articles. Global datasets (primarily from the World Bank) were only consulted if severe data gaps existed. The dataset includes parameters on passenger and freight transport activity, disaggregated by transport mode (road, rail, aviation, etc.) and geographic scale (inter-city or inner-city), if available. The dataset also covers the size of the vehicle fleet, disaggregated by vehicle types. The data coverage and subtypes vary among the parameters. The overall ambition is to include the most recent available year(s).

Item	*Description of Content*
Figure 1	A graph showing total population (million people), as well as the share of urban and rural population in Kenya.
Figure 2	A graph showing total GDP (million USD in 2015), as well as the share of the different sectors contributing to GDP in Kenya: agriculture, construction, mining, manufacturing, service, and energy.
Table 1	A table showing passenger transport activity in Kenya for the most recent year data was available. The data are curated from national statistics agencies or other government-affiliated agencies.
Table 2	An additional table showing passenger transport activity in Kenya based on UN DESA Statistics Division data (see explanation below). The data feature information for 2018.
Table 3	A table showing freight transport activity in Kenya for the most recent year data was available.
Table 4	An additional table showing freight transport activity in Kenya based on UN DESA Statistics Division data (see explanation below). The data feature information for 2018.
Tables 5 and 6	Tables showing vehicle fleet data in Kenya for the most recent year data were available.

For the parameters on passenger and freight transport activity, an additional dataset was included in Table 2 and Table 4. The UN DESA Statistics Division modelled passenger activity and freight activity for every country in support of SDG Indicator 9.1.2[1]. Passenger activity data provide information for road, rail, and air transport. Freight data cover road, rail and inland water, and aviation. The passenger-km and tonnes-km data originate from the Open SDG Data Hub. In this dataset, only the data for International Transport Forum (ITF) (representing mostly OECD countries) and UNECE countries (mostly European countries) are based on national reporting. For non-ITF/UNECE countries, the data are estimated using the ITF model, which uses several covariates such as GDP, population, and transport network coverage. A description of

[1] Freight: https://www.sdg.org/datasets/undesa::indicator-9-1-2-freight-volume-by-mode-of-transport-tonne-kilometres/about ;
Passenger: https://www.sdg.org/datasets/undesa::indicator-9-1-2-passenger-volume-passenger-kilometres-by-mode-of-transport/about

the model can be found in the ITF Transport Outlook 2017.

1.1 Population

Population data including total population, population growth, and split by rural or urban was gathered from The World Bank Open Data platform[2]. Figure 1 displays the total population disaggregated by urban and rural in Kenya.

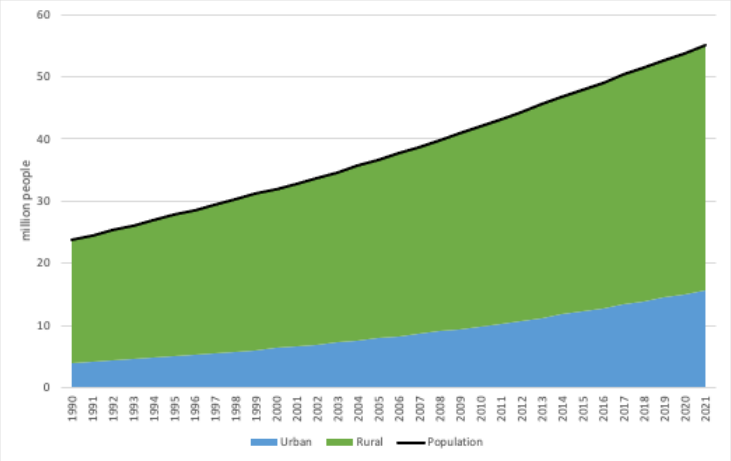
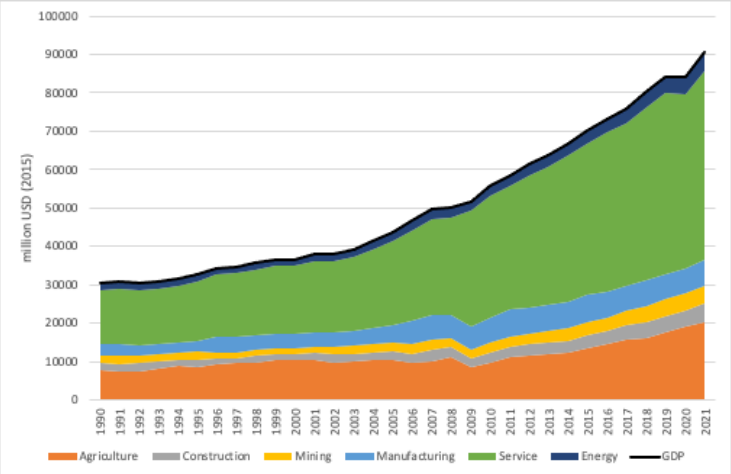


Figure 1: Total population (million people) disaggregated by urban and rural in Kenya

1.2 Gross domestic product (GDP)

GDP data including total GDP, GDP growth, and GDP share by sector (agriculture, manufacturing, service) was collected from The World Bank Open Data platform[2]. Where data was not available, data processing was done. Figure 2 shows the total GDP, as well as the share by sector, in Kenya.



[2] https://data.worldbank.org/

Figure 2: Total GDP (million USD in 2015) disaggregated by share in Kenya

1.3 Passenger transport activity

Kenya has only official statistics for rail passenger transport activity, covering 2016 to 2020. An additional source provides a data point for road passenger transport activity, but it is based on assumptions. Rail records 424 million passenger-km in 2020, while road records 8629 million passenger-km in 2015.

Table 1: Recorded passenger transport activity (million passenger-km) in Kenya

Mode	2015	2016	2017	2018	2019	2020
Rail		113	317	916	892	424
Road	8629					

Source:

- Rail: KNBS, 2021, Statistical Abstract 2021, https://www.knbs.or.ke/download/statistical-abstract-2021/
- Road: GIZ TraCS, 2018, Greenhouse gas emissions from the transport sector: Mitigation options for Kenya, https://www.changing-transport.org/wp-content/uploads/2018_GIZ_INFRAS_Transport_Mitigation_Options_Kenya.pdf

A 2006 report by the Japan International Cooperation Agency provides rail passenger transport activity from 1996 to 2002, indicating that rail had 297 million passenger-km in 2002.[3]

According to the UN DESA modelled data, it is estimated that the passenger activity in Kenya is 64190 million passenger-km in total in 2018. The large majority of passenger activity is conducted through road transport with 52485 million passenger-km for the same year.

Table 2: Modelled passenger transport activity (million passenger-km) in Kenya

Mode	2018
Aviation	11658.4098
Rail	46.932679
Road	52485.07434

1.4 Freight transport activity

Information on freight activity for Kenya has been only retrieved for rail. 2488 million tonnes-km were transported by rail in 2020. The data is curated by the Kenya National Bureau of Statistics. Data points cover 2016 to 2020.

Table 3: Freight transport activity (million tonnes-km) in Kenya

Mode	2016	2017	2018	2019	2020

[3] JICA, 2006, The Study on Master Plan for Urban Transport in the Nairobi Metropolitan Area, https://openjicareport.jica.go.jp/pdf/11823093_03.pdf

Rail	1141	857	1675	2334	2488

Source: KNBS, 2021, Statistical Abstract 2021, https://www.knbs.or.ke/download/statistical-abstract-2021/

While rail passenger activity saw a strong decrease, probably due to the COVID-19 pandemic, freight transport activity continued to grow in 2020. Kenya provides very recent information on rail passenger and freight activity. However, there is no information from national authorities or academia about freight transport activity by other transport modes.

The UN DESA modelled data estimates a road freight activity of 8855 million tonnes-km, 5506 million tonnes-km for rail, and 295 million tonnes-km for aviation in 2018. The estimated value for rail is twice as high as the actual numbers provided in Table 3.

Table 4: Modelled freight transport activity (million tonnes-km) in Kenya

Mode	2018
Aviation	294.965258
Rail	5506.228229
Road	8855.190131

1.5 Vehicle fleets

There was no data on the total vehicle fleet size in Kenya. The only available statistics were the new registrations per year by the Kenya National Bureau of Statistics. There were 346729 new vehicles registered in 2020. Official records cover data from 2016 to 2020. 60% more new vehicles were registered in 2020 than in 2016. The number of new motor- and autocycle registrations doubled between 2016 and 2020.

Table 5: Vehicle registrations in Kenya

Mode	2016	2017	2018	2019	2020
Motor and autocycles	123539	191601	195253	217425	252601
Motor cars	58613	66698	74683	82483	65716
Utilities, Panel, Pick-ups, etc.	12722	9866	11220	10189	6065
Lorries, trucks, and heavy vans	9632	7460	6514	6518	6476
Other motor vehicles	4096	3563	5659	5651	11505
Trailers	2829	1953	2083	1639	2382
Buses and Mini-buses	2284	1531	1877	3271	1984
Total	213715	282672	297289	327176	346729

Source: KNBS, 2021, Statistical Abstract 2021, https://www.knbs.or.ke/download/statistical-abstract-2021/

Data on the total vehicle fleet was retrieved in a study by GIZ Advancing Transport Climate Strategies (TraCS) programme. It provides fleet numbers for 2015 and 2017 (Table 6). In 2017, a total of 1.6 million road vehicles

were in operation. Motorcycles represent the largest number of vehicles with 762807 vehicles, followed by passenger cars with 626896 vehicles. While the overall vehicle fleet grew by 29% between 2015 and 2017, motorcycles recorded a growth of 41%, confirming the statistics shown in Table 5.

Table 6: Vehicle fleet numbers in Kenya

Mode	2015	2017
Passenger car	532406	626896
Light commercial vehicles	119230	142922
Buses	57792	75202
Heavy goods vehicles	10041	13070
Motorcycles	539768	762807
Total vehicles	1259237	1620897

Source: GIZ TraCS, 2019, Updated Transport Data in Kenya 2018, https://www.changing-transport.org/wp-content/uploads/2019_Updated-transport-data-in-Kenya.pdf

2. Methodology

The focus is on national data for passenger activity (passenger-km), freight activity (tonnes-km) and modes of transport (number of vehicles). The priority is to collect data released by national governments, government-affiliated organisations, or country-specific studies. The research identifies the most recent available data and any data available from 1990 onwards. The priority was for any data after 2010, because transport is a very dynamic growth sector and anything before 2010 adds limited value to understanding the current real-world situation.

Desk research is the main data collection approach for the Transport Starter Data Kits. The desk research examined annual yearbooks, transport statistics, country reporting, and any national statistical portals. Websites of the national government, transport ministries, statistical institutes and other related authorities were examined. Only when severe data gaps exist, global datasets are consulted. In some cases, World Bank data[4] on rail passenger and rail freight is included.

Each Transport Data for Starter Data Kit set contains an additional dataset, which is sourced from the United Nations Department of Economic and Social Affairs (UN DESA) Statistics Division. It is included as a secondary priority because this dataset is the result of a modelling exercise and covers every country. The UN DESA modelled passenger activity and freight activity has the purpose to support the Sustainable Development Goal Indicator 9.1.2[5]. The passenger activity provides information for road, rail, and air

[4] Rail passenger data: World Bank, 2022, Railways, passengers carried (million passenger-km), https://data.worldbank.org/indicator/IS.RRS.PASG.KM; rail freight data: World Bank, 2022, Railways, goods transported (million ton-km), https://data.worldbank.org/indicator/IS.RRS.GOOD.MT.K6

[5] UN DESA, 2021, Indicator 9.1.2: Freight volume by mode of transport (tonne kilometres): https://www.sdg.org/datasets/undesa::indicator-9-1-2-freight-volume-by-mode-of-transport-tonne-kilometres/about ;
UN DESA, 2021, Indicator 9.1.2: Passenger volume (passenger kilometres) by mode of transport:

transport. Freight data covers the road, rail and inland water, and aviation. The passenger-km and tonnes-km data originate from the Open Sustainable Development Goals (SDG) Data Hub. In the UN DESA dataset, only the data for countries participating in the International Transport Forum (ITF) (representing mostly member countries of the Organisation for Economic Co-operation and Development (OECD)) and the United Nations Economic Commission for Europe (UNECE) (mostly European countries) are based on national reporting. For non-ITF/UNECE countries, data are estimated using the ITF model, which uses several covariates such as gross domestic product, population, and transport network coverage. A description of the model can be found in the ITF Transport Outlook 2017[6]. The UN DESA dataset is included in the Transport Data for Starter Data Kits as additional tables to fill in the incomplete picture that most countries present. The UN DESA modelled data is less accurate and it shall only be regarded as offering the wider picture of transport activity in the country.

The collected data have been shared with a group of relevant SLOCAT partners to validate and explore any additional sources. The SLOCAT partners were selected based on their actions to lead projects in the region and their involvement in data-focused knowledge products or projects. The consultation involved ten anonymous organisations.

Desk research is an approach that limits the research to material available on the internet, accessible through search engines and linked to government and statistical institutes' websites. However, this does not pose a major limitation to obtaining data. Nearly every country has functional websites for statistics and transport authorities. In a few cases, websites are not well maintained, resulting in missing or broken hyperlinks to reports. By using services that provide access to archived websites, some of these broken pages can be retrieved. The collected information has been shared with partners and no additional information has been received.

While over 1,500 languages are spoken across Sub-Saharan Africa, government datasets are generally published in a smaller subset of languages including English, French, Portuguese and others. Nonetheless, language is not a barrier to navigating through the material and identifying the relevant parameters. The involved team members can navigate through reports in such languages. If needed, automatic translation tools were used.

Due to missing values in the country's historical GDP data, extrapolation between available years was done by the authors to address this. The World Bank's data platform provided GDP share by sector for agriculture, manufacturing, and services. However, GDP share by construction, mining, and energy was also needed to align the data structure with the MAED tool. To address the lack of data available for these sectors, the authors assumed that construction, mining, manufacturing, and energy all fall within the industry sector. Thus, to obtain data for the three remaining sectors, the remaining percentage after considering agriculture, manufacturing, and services from The World Bank's data platform, was divided by three. It is therefore assumed that the GDP share of the construction, mining, and energy sectors are the same.

https://www.sdg.org/datasets/undesa::indicator-9-1-2-passenger-volume-passenger-kilometres-by-mode-of-transport/about

[6] ITF, 2017, ITF Transport Outlook 2017, https://www.itf-oecd.org/transport-outlook-2017

3. Ethics Statement

Not applicable

4. CRediT Author Statement

Naomi Tan: Investigation, Conceptualisation, Methodology; Data Collection; Visualization, Writing and Editing; **Robert Ambunda:** Data Collection; Investigation; Writing and Editing; **Nikola Medimorec:** Conceptualisation; Methodology; Data Collection; Investigation; Writing, Review & Editing; Supervision; **Angel Cortez:** Data Collection; **Agustina Krapp:** Data Collection; **Erin Maxwell:** Data Collection; **John Harrison:** Supervision; **Mark Howells:** Supervision

Acknowledgements

We would like to acknowledge the SLOCAT Partnership on Sustainable, Low Carbon Transport who helped make this and future iterations possible. We would also like to acknowledge the International Road Federation (IRF) and the International Union of Railways (UIC) for providing us with these data. The data are extracted from IRF World Road Statistics (WRS) and their use is subject to copyright and specific Terms and Conditions available on the WRS website. More WRS data are available for free on its Data Warehouse www.worldroadstatistics.org. Likewise, data was extracted from the UIC Statistics Rail Information System and Analyses (Railisa) and more can be found on its online tool https://uic-stats.uic.org/

Funding

As well as support in kind provided by the employers of the authors of this note, we also acknowledge core funding from the Climate Compatible Growth Program (#CCG) of the UK's Foreign Development and Commonwealth Office (FCDO). The views expressed in this paper do not necessarily reflect the UK government's official policies.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.