# *Transport Starter Data Kit: Historical socio-transport data for Malaysia*

**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 Malaysia, 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

Malaysia

## Specifications Table

| | |
|---|---|
| **Subject** | Transport |
| **Specific subject area** | Transport Data |
| **Type of data** | Tables<br>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 Malaysia |
| **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.6539954 |

**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 Malaysia'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 Malaysia. 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 Malaysia. |
| Figure 2 | A graph showing total GDP (million USD in 2015), as well as the share of the different sectors contributing to GDP in Malaysia: agriculture, construction, mining, manufacturing, service, and energy. |
| Table 1 | A table showing passenger transport activity in Malaysia 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 Malaysia 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 Malaysia for the most recent year data was available. |
| Table 4 | An additional table showing freight transport activity in Malaysia based on UN DESA Statistics Division data (see explanation below). The data feature information for 2018. |
| Table 5 | Tables showing vehicle fleet data in Malaysia 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 the SDG Indicator 9.1.2[1]. The passenger activity provides information for road, rail, and air transport. Freight data covers the 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 is estimated using the ITF model, which uses several covariates such as GDP, population, and transport network coverage. A description of 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 Malaysia.

[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

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

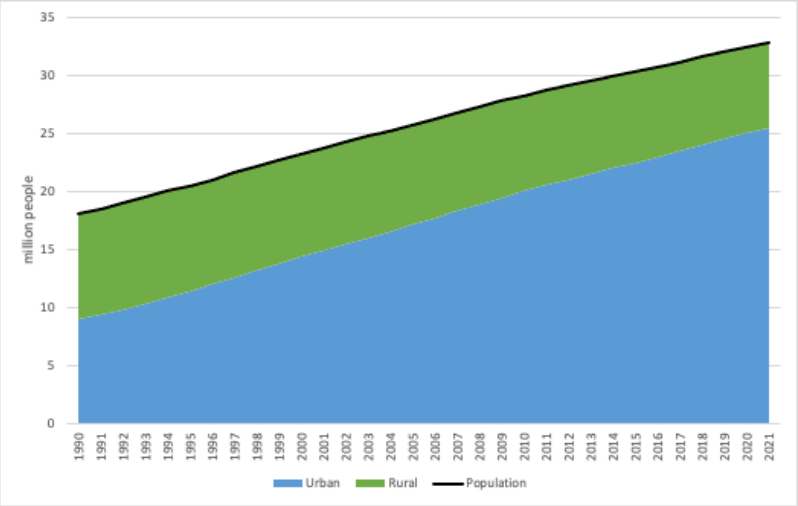


**Figure 1: Total population (million people) disaggregated by urban and rural in Malaysia**

## 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 Malaysia.

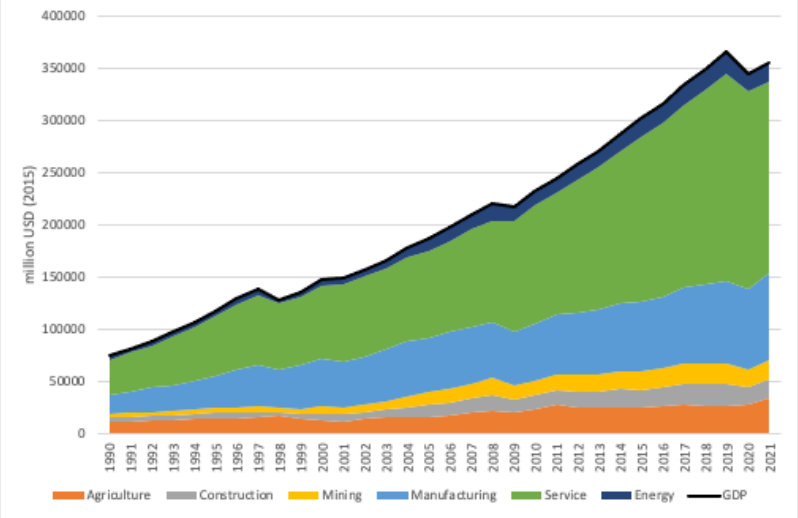


**Figure 2: Total GDP (million USD in 2015) disaggregated by share in Malaysia**

## 1.3 Passenger transport activity

Information on passenger transport activity in Malaysia is only captured through rail transport. Rail passenger activity recorded 196 million passenger-km in 2018 (Table 1). The data covers 2002 to 2018. A

strong decrease (83%) in rail passenger activity has been recorded during the period. The data was provided by the Asian Development Bank's Asian Transport Outlook Database.

In the case of Malaysia, the passenger activity misses to report any information for other transport modes.

**Table 1: Reported passenger transport activity (million passenger-km) in Malaysia**

| Mode | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rail | 1138 | 1031 | 1152 | 1195 | 1248 | 1317 | 1386 | 1526 | 1532 | 1428 | 1218 | 1099 | 638 | 426 | 291 | 202 | 196 |

Source: ADB, 2021, Asian Transport Outlook Database, https://data.adb.org/dataset/asian-transport-outlook-database

According to the UN DESA modelled data, it is estimated that the passenger activity in Malaysia is 328490 million passenger-km in 2018. Rail covers 52% of the passenger transport activity and even aviation (with 113966 million passenger-km in 2018) is higher than roads (44426 million passenger-km).

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

| Mode | 2018 |
|---|---|
| Aviation | 113966.3432 |
| Rail | 170097.289 |
| Road | 44426.21047 |

## 1.4 Freight transport activity

Information on freight activity for Malaysia has been only retrieved for rail. 1233 million tonnes-km were achieved by rail in 2018. The oldest record is from 2002 when the activity level is at 1073 million tonnes-km. The activity level experienced a growth between 2006 and 2014 but decreased again until 2018. The data is curated by the Asian Development Bank's Asian Transport Outlook Database.

**Table 3: Recorded freight transport activity (million tonnes-km) in Malaysia**

| Mode | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rail | 1073 | 887 | 1017 | 1178 | 1337 | 1356 | 1351 | 1384 | 1483 | 1536 | 1565 | 1762 | 1742 | 1475 | 1350 | 1234 | 1233 |

Source: ADB, 2021, Asian Transport Outlook Database, https://data.adb.org/dataset/asian-transport-outlook-database

According to the UN DESA modelled data, freight activity for 2018 in Malaysia is 112777 million tonnes-km. The dataset assumes that roads are the main freight transport activity. 92% of freight transport activity is curated by roads.

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

| Mode | 2018 |
|---|---|

| | |
|---|---|
| Aviation | 1404.408 |
| Rail | 7239.084 |
| Road | 104133.4 |

## 1.5 Vehicle fleets

Malaysia has a total of 31.2 million road vehicles in 2019. Official records cover information from 2002 to 2019. Data on buses is only reported until 2013. Road vehicle types covered in detail are two-wheelers, light-duty vehicles, buses, and freight vehicles. During this period, the total road vehicle fleet grew by 160%. The largest growth was recorded for light-duty vehicles. The vehicle fleet data is sourced from the Asian Development Bank's Asian Transport Outlook Database.

**Table 5: Vehicle fleet numbers in Malaysia**

| Mode | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|---|---|
| Two-wheelers | 9441907 | 9949774 | 10559370 | 11087878 | 11629265 | 11871696 | 12677220 | 13173069 | 13725918 | 14322129 |
| Light-duty vehicles | 9217881 | 9433490 | 10064851 | 10689450 | 11199910 | 12094790 | 13000358 | 13582636 | 14189693 | 14762708 |
| Buses | 69149 | 71784 | 73536 | 62784 | - | - | - | - | - | - |
| Freight vehicles | 966177 | 1034660 | 1075402 | 1116167 | 1159872 | 1197987 | 1190664 | 1223396 | 1262064 | 1296431 |
| Total road vehicles | 20188565 | 21304818 | 22616566 | 23819256 | 25101192 | 26301952 | 27613259 | 28738176 | 29956525 | 31214871 |

Source: ADB, 2021, Asian Transport Outlook Database, https://data.adb.org/dataset/asian-transport-outlook-database

# 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[3] on rail passenger and rail freight is included.

[3] 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

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[4]. The passenger activity provides information for road, rail, and air 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[5]. 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.

Government datasets are generally published in local languages. For this study, the relevant information is accessed through the Asian Transport Outlook Database[6] which is available in English, thus no language issues were encountered.

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

(million ton-km), https://data.worldbank.org/indicator/IS.RRS.GOOD.MT.K6

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

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

[6] ADB, 2021, Asian Transport Outlook Database, https://data.adb.org/dataset/asian-transport-outlook-database

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.

## 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.