The International Transport Energy Modeling (iTEM) Open Data Project
Rule Book

November 28, 2023

Contents

General Information

This documentation contains detailed information of the **iTEM Open Database**, a harmonized transport data set of historical values, 1970 - 2018. It aims to create transparency through two key features:

Open-Data: Assembling a comprehensive collection of publicly-available transportation data

Open-Code: All code and documentation will be publicly accessible and open for modification and extension.

The iTEM Open Database is comprised of individual datasets collected from public sources. Each dataset is downloaded, cleaned, and harmonised to the common region and technology definitions defined by the iTEM consortium `https://transportenergy.org`. For each dataset, we describe the name of the dataset, the web link to the original source, the web link to the cleaning script (in python), variables, and explain the data cleaning steps (which explains the data cleaning script in plain English).

Nomenclature

- Datasets are numbered by the order they were collected and processed. Names that start with **T** stand for *Transport*, names that start with **S** stand for *sociodemographic*, and names that start with **A** stand for *Analysis* in which variables are derived for validation purposes. An "_iTEM" identifier is added to the end of the variable names indicating that **iTEM** is the data source, meaning that these variables are calculated by iTEM for validation instead of being collected from any of the original sources.

Definitions of regions

Unless otherwise specified, all the ITEM regions are obtained from the following file: `https://github.com/transportenergy/metadata/blob/master/model/regions.yaml`. The ISO code of each country is obtained according to the library *PyCountry*. However, certain countries in the dataset do not have the exact names as those appearing in the library; therefore, the section *Country and ISO Code* indicates what name is used for the countries that are not found in *PyCountry*.

How Tos

Forthcoming

How to navigate the Open Data

The input file used in each script is located at `https://github.com/transportenergy/metadata/tree/master/historical/input`. Detailed instructions on how to generate the latest iTEM Open Database (the merged file of individual datasets) is forthcoming.

How to navigate the Open Code

The scripts for cleaning the data is located at `https://github.com/transportenergy/database/tree/master/item/historical/scripts`.

S012

Information

- Dataset name: De facto population (both sexes) in a country as of 1 July of the year indicated
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/S012.ipynb`

Source : United Nations `https://population.un.org/wpp/Download/Files/1_Indicators%20(Standard)/EXCEL_FILES/1_Population/WPP2019_POP_F01_1_TOTAL_POPULATION_BOTH_SEXES.xlsx`

Country and ISO Code The following name changes were performed:

- Venezuela (Bolivarian Republic of) →Venezuela
- Holy See →Holy See (Vatican City State)
- China, Taiwan Province of China →Taiwan
- State of Palestine →Palestine
- Wallis and Futuna Islands →Wallis and Futuna
- Saint Helena →Saint Helena, Ascension and Tristan da Cunha
- United States Virgin Islands →Virgin Islands, U.S.
- Iran (Islamic Republic of) →Iran, Islamic Republic of
- Dem. People's Republic of Korea →Korea, Democratic People's Republic of
- Democratic Republic of the Congo →Congo, The Democratic Republic of the
- China, Macao SAR →Macao
- Bolivia (Plurinational State of) →Bolivia
- Republic of Korea →Korea, Republic of
- China, Hong Kong SAR →Hong Kong
- Micronesia (Fed. States of) →Micronesia, Federated States of

The only country we could not assign an ISO code was *Channel Islands*

ITEM Region The following countries were not assigned an ITEM region:

1. *Channel Islands*
2. *Saint Martin (French part)*
3. *Sint Maarten (Dutch part)*
4. *South Sudan*
5. *Bonaire, Sint Eustatius and Saba*
6. *Saint Barthelemy*
7. *Curaçao*

Variable : *Population*.

Unit : 10^3 *people*.

Service : Null.

Mode : Null.

Vehicle Type : Null.

Technology : Null.

Fuel : Null.

T000

Information

- Dataset name: Passenger transport: Inland passenger transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T000.py`

Source : International Transport Forum `https://stats.oecd.org/index.aspx?queryid=79863`.

Country and ISO Code : The following name changes were performed:

- Montenegro, Republic of →Montenegro
- Bosnia-Herzegovina →Bosnia and Herzegovina
- Korea →Korea, Republic of
- Serbia, Republic of →Serbia

ITEM Region : All countries belong to an ITEM region.

Variable : *Passenger Activity.*

Unit : The unit is changed from *Passenger-kilometres, Millions* to 10^9 *passenger-km / yr.*

Service : *Passenger.*

Mode

- The mode for *Total inland passenger transport* is *All.*
- The mode for *Rail passenger transport* is *Rail.*
- The mode for *Road passenger transport by passenger cars* is *Road.*
- The mode for *Road passenger transport by buses and coaches* is *Road.*

Vehicle Type

- The Vehicle Type for *Total inland passenger transport* is *All.*
- The Vehicle Type for *Rail passenger transport* is *All.*
- The Vehicle Type for *Road passenger transport by passenger cars* is *LDV.*
- The Vehicle Type for *Road passenger transport by buses and coaches* is *Bus.*

Technology : *All.*

Fuel : *All.*

Data Cleaning

- The variable *Road Passenger Transport* is the sum of *Road passenger transport by passenger cars* and *Road passenger transport by buses and coaches*. In other words, Mode *Road* Vehicle Type *All* is the sum of Mode *Road* Vehicle Type *LDV* and Mode *Road* Vehicle Type *Bus*.
- There are 22 countries that have missing data for *Road passenger transport by passenger cars* or *Road passenger transport by buses and coaches* for certain years (we call it "problematic time periods" below), therefore the total sum *Road Passenger Transport* is incorrectly reported. Below are the rules on how we handle these cases:
  - Albania: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) & *Road passenger transport by buses and coaches* (Mode *Road* Vehicle Type *Bus*) during the problematic time periods.
  - Armenia: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Azerbaijan: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Belarus: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Bulgaria: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Canada: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Russian Federation: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - Switzerland: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.
  - United States: Remove *Road passenger transport* (Mode *Road* Vehicle Type *All*) during the problematic time periods.

T001

Information

- Dataset name: Costal Transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T001.py`

Source : International Transport Forum `https://stats.oecd.org/Index.aspx?DataSetCode=ITF_GOODS_TRANSPORT#`.

Country and ISO Code The following name changes were performed:

- Montenegro, Republic of →Montenegro
- Korea →Korea, Republic of
- Serbia, Republic of →Serbia

ITEM Region All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Tonnes-Kilometer* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : *Shipping.*

Vehicle Type : *Coastal.*

Technology : *All.*

Fuel : *All.*

T002

Information

- Dataset name: Container Transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T002.ipynb`

Source : International Transport Forum `https://stats.oecd.org/Index.aspx?DataSetCode=ITF_GOODS_TRANSPORT#`.

Country and ISO Code The following name changes were performed:

- Korea →Korea, Republic of

ITEM Region : All countries belong to an ITEM region.

Variable : The variable is set to either *Freight (TEU)* or *Freight (Weight)*.

Unit Since there are two varables, their corresponding unit is the following:

- *Freight (TEU)* is *Number*
- *Freight (Weight)* is 10^3 *tonne / yr*

Service : *Freight*.

Mode

- The mode for *Rail containers transport (TEU)* is *Rail*
- The mode for *Maritime containers transport (weight)* is *Shipping*

Vehicle Type : *Container*.

Technology : *All*.

Fuel : *All*.

T003

Information

- Dataset name: Inland Freight Transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T003.py`

Source : International Transport Forum `https://stats.oecd.org/Index.aspx?DataSetCode=ITF_GOODS_TRANSPORT#`.

Country and ISO Code The following name changes were performed:

- Montenegro, Republic of →Montenegro
- Bosnia-Herzegovina →Bosnia and Herzegovina
- Korea →Korea, Republic of
- Serbia, Republic of →Serbia

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity*.

s ectionUnit: The unit is changed from *Million Tonnes-kilometers* to 10^9 *tonne-km / yr*.

Service

- The service for *Road freight transport on own account* is *Freight*.
- The service for *Inland waterways freight transport* is *Freight*.
- The service for *Rail freight transport* is *Freight*.
- The service for *Road freight transport* is *Freight*.
- The service for *Road freight transport for hire and reward* is *Freight*.
- The service for *Total inland freight transport* is *Freight*.
- The service for *Pipelines transport* is *Pipeline*.

Mode

- The mode for *Road freight transport* is *Road*.
- The mode for *Road freight transport for hire and reward* is *Road*.
- The mode for *Road freight transport on own account* is *Road*.
- The mode for *Rail freight transport* is *Rail*.
- The mode for *Pipelines transport* is *Pipeline*.
- The mode for *Inland waterways freight transport* is *Shipping*.
- The mode for *Total inland freight transport* is *Inland*.

We also created a new Mode called *Inland (exl. Pipeline)*, which is the result represent the sum of all services except *Pipeline*.

Vehicle Type

- The vehicle type for *Road freight transport* is *All*.
- The vehicle type for *Road freight transport for hire and reward* is *For Hire and Reward*.
- The vehicle type for *Road freight transport on own account* is *For Own Account*.
- The vehicle type for *Rail freight transport* is *All*.
- The vehicle type for *Pipelines transport* is *Pipeline*.
- The vehicle type for *Inland waterways freight transport* is *Inland*.
- The vehicle type for *Total inland freight transport* is *All*.
- The vehicle type for *Inland (exl. Pipeline)* is *All*.

Technology : *All*.

Fuel : *All*.

T004

Information

- Dataset name: New Road Vehicle Registrations by Vehicle Category and Fuel Type
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T004.py`

Source : United Nations Economic Commission for Europe `https://datasource.kapsarc.org/explore/dataset/new-road-vehicle-registrations-by-vehicle-category-and-fuel-type/export/?disjunctive.country_name&disjunctive.date&disjunctive.frequency&disjunctive.fuel_type_name&disjunctive.type_of_vehicle_name`.

Country and ISO Code The following name changes were performed:

- The former Yugoslav Republic of Macedonia →North Macedonia

ITEM Region : All countries belong to an ITEM region.

Variable : *Sales (New Vehicles).*

Unit : 10^6*vehicle / yr.*

Service

- The service for *New lorries (vehicle wt over 3500 kg)* is *Freight*
- The service for *New road tractors* is *Freight*
- The service for *New passenger cars* is *Passenger*
- The service for *New motor coaches, buses and trolley buses* is *Freight*
- The service for *New light goods vehicles* is *Freight*

Mode : *Road.*

Vehicle Type

- The Vehicle Type for *New lorries (vehicle wt over 3500 kg)* is *Heavy Truck*
- The Vehicle Type for *New road tractors* is *Medium Truck*
- The Vehicle Type for *New passenger cars* is *LDV*
- The Vehicle Type for *New motor coaches, buses and trolley buses* is *Bus*
- The Vehicle Type for *New light goods vehicles* is *Light Truck*

Technology

- The Technology for *LPG* is *Conventional*
- The Technology for *Compressed natural gas (CNG)* is *Natural Gas Vehicle*
- The Technology for *Liquefied natural gas (LNG)* is *Natural Gas Vehicle*
- The Technology for *Bioethanol* is *Conventional*
- The Technology for *Bi-fuel vehicles* is *Conventional*
- The Technology for *Biodiesel* is *Conventional*
- The Technology for *Diesel (excluding hybrids)* is *Conventional*
- The Technology for *Hybrid electric-diesel* is *Conventional*
- The Technology for *Hybrid electric-petrol* is *Conventional*
- The Technology for *Diesel* is *Conventional*
- The Technology for *Petrol* is *Conventional*
- The Technology for *Petrol (excluding hybrids)* is *Conventional*
- The Technology for *Plug-in hybrid diesel-electric* is *PHEV*
- The Technology for *Plug-in hybrid petrol-electric* is *PHEV*
- The Technology for *Hydrogen and fuel cells* is *Fuel Cell*
- The Technology for *Electricity* is *BEV*
- The Technology for *Total* is *All*
- The Technology for *Alternative (total)* is *Alternative*

Fuel

- The Fuel for *LPG* is *Liquid - Fossil*
- The Fuel for *Compressed natural gas (CNG)* is *Natural gas*
- The Fuel for *Liquefied natural gas (LNG)* is *Natural gas*
- The Fuel for *Bioethanol* is *Liquid-Bio*
- The Fuel for *Bi-fuel vehicles* is *Liquid-Bio*
- The Fuel for *Biodiesel* is *Liquid-Bio*
- The Fuel for *Diesel (excluding hybrids)* is *Liquid - Fossil*
- The Fuel for *Hybrid electric-diesel* is *Liquid - Fossil*
- The Fuel for *Hybrid electric-petrol* is *Liquid - Fossil*
- The Fuel for *Diesel* is *Liquid - Fossil*
- The Fuel for *Petrol* is *Liquid - Fossil*
- The Fuel for *Petrol (excluding hybrids)* is *Liquid - Fossil*
- The Fuel for *Plug-in hybrid diesel-electric* is *Electricity*
- The Fuel for *Plug-in hybrid petrol-electric* is *Electricity*
- The Fuel for *Hydrogen and fuel cells* is *Hydrogen*
- The Fuel for *Electricity* is *Electricity*
- The Fuel for *Total* is *All*
- The Fuel for *Alternative (total)* is *Alternative*

T005

Information

- Dataset name: Direct CO2 emissions from global (and regional) transport scenarios
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T005.ipynb`

Source : Joint Research Center `https://edgar.jrc.ec.europa.eu/overview.php?v=50_GHG`.

Country and ISO Code The following name changes were performed:

- Swaziland →Eswatini
- Saint Helena →Saint Helena, Ascension and Tristan da Cunha
- Libyan Arab Jamahiriya →Libya
- Congo_the Democratic Republic of the →Congo, The Democratic Republic of the
- Reunion →Réunion
- Int. Aviation →World
- Int. Shipping →World
- Virgin Islands_British →Virgin Islands, British
- Cote d'Ivoire →Côte d'Ivoire
- Taiwan_Province of China →Taiwan, Province of China
- Cape Verde →Cabo Verde
- Tanzania_United Republic of →Tanzania, United Republic of
- The former Yugoslav Republic of Macedonia →North Macedonia

ITEM Region To the following countries we assigned the ITEM region manually:

- Serbia and Montenegro →SCG
- World →WLD
- Netherlands Antilles →ANT

Variable : *CO2 Emission (ttw).*

Unit : 10^6 *t CO2 / yr.*

Service : *All.*

Mode The mapping done for countries is the following:

- The mode for *Railways* is *Rail*
- The mode for *Road Transportation* is *Road*
- The mode for *Civil Aviation* is *Air*
- The mode for *Other Transportation* is *Other*
- The mode for *Water-borne Navigation* is *Shipping*

The mapping done for the Int. Aviation country is the following:

- The mode for *Civil Aviation* is *Domestic Aviation*

The mapping done for the Int. Shipping country is the following:

- The mode for *Water-Borne Navigation* is *Domestic Shipping*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

T006

Information

- Dataset name: Modal split of freight transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T006.ipynb`

Source : Eurostat `https://datasource.kapsarc.org/explore/dataset/modal-split-of-freight-transport/information/?disjunctive.date&disjunctive.frequency&disjunctive.geo_name&disjunctive.measure_name&disjunctive.tra_mode_name`.

Country and ISO Code The following name changes were performed:

- European Union (current composition) →EU28

ITEM Region To the following countries, the ITEM region was assigned manually as follows:

- European Union (current composition) →EU-28

Variable : *Freight Activity.*

Unit : *% tonne-kilometres / yr.*

Service : *Freight.*

Mode

- The mode for *Railways* is *Rail*
- The mode for *Roads* is *Road*
- The mode for *Inland waterways* is *Shipping*

Vehicle Type

- The mode for *Railways* is *All*
- The mode for *Roads* is *All*
- The mode for *Inland waterways* is *Inland Waterway*

Technology : *All.*

Fuel : *All.*

T007

Information

- Dataset name: Modal split of passenger transport
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T007.ipynb`

Source : Eurostat `https://datasource.kapsarc.org/explore/dataset/modal-split-of-passenger-transport/export/?disjunctive.date&disjunctive.frequency&disjunctive.geo_name&disjunctive.measure_name&disjunctive.vehicle_name`.

Country and ISO Code The following name changes were performed:

- European Union (28 countries) →EU28
- European Union (27 countries) →EU27
- The former Yugoslav Republic of Macedonia →North Macedonia

ITEM Region To the following countries, the ITEM region was assigned manually as follows:

- EU27 →EU-27
- EU28 →EU-28

Variable : *Passenger Activity.*

Unit : *% in total inland passenger-km / yr.*

Service : *Passenger.*

Mode

- The mode for *Trains* is *Rail*
- The mode for *Passenger cars* is *Road*
- The mode for *Motor coaches, buses and trolley buses* is *Road*

Vehicle Type

- The vehicle type for *Trains* is *All*
- The vehicle type for *Passenger cars* is *LDV*
- The vehicle type for *Motor coaches, buses and trolley buses* is *Bus*

Technology : *All.*

Fuel : *All.*

T008

Information

- Dataset name: Passenger Road Vehicle Fleet and rate per thousand inhabitants by Vehicle Category
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T008.ipynb`

Source : United Nations Economic Commission for Europe `https://datasource.kapsarc.org/explore/dataset/passenger-road-vehicle-fleet-and-rate-per-thousand-inhabitants-by-vehicle-catego/information/?disjunctive.country_name&disjunctive.date&disjunctive.frequency&disjunctive.measurement_name&disjunctive.vehicle_category_name`.

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Stock*.

Unit For the given variable there are two units:

- 10^6 *vehicle*.
- *Vehicles per 1000 inhabitants*.

Service : *Passenger*.

Mode : *Road*.

Vehicle Type

- The Vehicle Type for *Special purpose vehicles* is *Special purpose vehicles*
- The Vehicle Type for *Passenger cars* is *LDV*
- The Vehicle Type for *Trams* is *Trams*
- The Vehicle Type for *Motorcycles* is *Motorcycles*
- The Vehicle Type for *Motor coaches, buses and trolley bus* is *Bus*
- The Vehicle Type for *Mopeds* is *Mopeds*

Technology : *All*.

Fuel : *All*.

T009

Information

- Dataset name: Passenger Car
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T009.py`

Source

International Organization of Motor Vehicle Manufacturers
`http://www.oica.net/category/vehicles-in-use/`

Country and ISO Code

The following name changes were performed:

- RUSSIA →Russian Federation
- SYRIA →Syrian Arab Republic
- IRAN →Iran, Islamic Republic of
- BOSNIA →Bosnia and Herzegovina
- HONG-KONG →Hong Kong
- IVORY COAST →Côte d'Ivoire
- BRUNEI →Brunei Darussalam
- MOLDAVIA →Moldova, Republic of
- SOUTH KOREA →Korea, Republic of
- CONGO KINSHASA →Congo, The Democratic Republic of the
- PALESTINE →Palestine, State of
- MACEDONIA →North Macedonia

ITEM Region

All countries belong to an ITEM region.

Variable

The variable is set to *Stock*.

Unit

The unit is 10^6 *vehicle*.

Service

The service corresponds to *Passenger*.

Mode

The mode for the given service is *Road*

Vehicle Type : *All*.

Technology : *All*.

Fuel : *All*.

T010

Information

- Dataset name: Commercial Vehicle
- Link to cleaning script: `https://github.com/transportenergy/database/blob/master/item/historical/scripts/T010.ipynb`

Source

International Organization of Motor Vehicle Manufacturers
`http://www.oica.net/category/vehicles-in-use/`

Country and ISO Code

The following name changes were performed:

- RUSSIA →Russian Federation
- SYRIA →Syrian Arab Republic
- IRAN →Iran, Islamic Republic of
- BOSNIA →Bosnia and Herzegovina
- HONG-KONG →Hong Kong
- IVORY COAST →Côte d'Ivoire
- BRUNEI →Brunei Darussalam
- MOLDAVIA →Moldova, Republic of
- SOUTH KOREA →Korea, Republic of
- CONGO KINSHASA →Congo, The Democratic Republic of the
- PALESTINE →Palestine, State of
- MACEDONIA →North Macedonia

The only country we could not assign an ISO code was Azerbaidjan.

ITEM Region

All countries belong to an ITEM region, except for Azerbaidjan.

Variable

The variable is set to *Stock*.

Unit

The unit is 10^6 *vehicle*.

Service

The service corresponds to *Freight*.

Mode

The mode for the given service is *Road*

Vehicle Type : *All*.

Technology : *All*.

Fuel : *All*.

T011

Information

- Dataset name: Aviation Total Passenger Kilometers
- Link to cleaning script: `https://github.com/linero-tech/iteminternship/blob/main/code/dima/T011_TAS-PAT-017(1).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-PAT-017(1) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Passenger Activity.*

Unit : The unit is changed from *Million passenger kilometers* to 10^9 *passenger-km / yr.*

Service : *Aviation.*

Mode : *Aviation.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

Data Cleaning The following column names are removed from the datasets

- Remarks
- Source (2022-04)
- Source (2021-10)
- Source (2023-03)

T012

Information

- Dataset name: Freight Transport - Tonne-km for Railways
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/dima/T012_TAS-FRA-005(2).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-FRA-005(2) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Million tonne kilometers* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : *Rail.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

T013

Information

- Dataset name: Passengers Kilometer Travel - Roads
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/dima/T013_TAS-PAG-005(2).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-PAG-005(2) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Passenger Activity.*

Unit : The unit is changed from *Million passenger kilometers* to 10^9 *passenger-km / yr.*

Service : *Passenger.*

Mode : *Road.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

T014

Information

- Dataset name: LDV Sales
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/dima/T014_TAS-VEP-005(2).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-VEP-005(2) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Sales (New Vehicles)*.

Unit : The unit is changed from *Number* to 10^6 *vehicle / yr*

Service : *Passenger*.

Mode : *Road*.

Vehicle Type : *LDV*.

Technology : *All*.

Fuel : *All*.

T015

Information

- Dataset name: Freight Transport - Tonne-km for Railways
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/hanna/T015_TAS-FRA-005(3).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-FRA-005(3) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Million tonne kilometers* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : *Rail.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

Data Cleaning The following column names are removed from the datasets

- Remarks
- Source (2022-04)
- Source (2021-10)
- Source (2023-03)

T016

Information

- Dataset name: LDV Sales
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/hanna/T016_TAS-VEP-005(1).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-VEP-005(1) dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Sales (New Vehicles)*.

Unit : The unit is changed from *Number* to 10^6 *vehicle/year*.

Service : *Passenger*.

Mode : *Road*.

Vehicle Type : *LDV*.

Technology : *All*.

Fuel : *All*.

T017

Information

- Dataset name: Freight Vehicle Registration
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/hanna/T017_TAS-VEP-020.ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-VEP-020 dataset downloaded on November, 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Stock*.

Unit : The unit is changed from *Number* to 10^6 *vehicle/year*.

Service : *Freight*.

Mode : *Road*.

Vehicle Type : *All*.

Technology : *All*.

Fuel : *All*.

T018

Information

- Freight Transport - Tonne-km for Roads
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/serah/T018_TAS-FRA-004(2).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-FRA-004(2) dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Million tonne kilometers* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : *Road.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

Data Cleaning The following column names are removed from the datasets

- Remarks
- Source (2022-04)
- Source (2021-10)
- Source (2023-03)

T019

Information

- Dataset name: Freight Transport - Tonne-km for Aviation (Domestic)
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/serah/T019_TAS-FRA-007(3).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-FRA-007(3)) dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Million tonne kilometers* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : The mode is changed from *Aviation* to *Aviation (Domestic).*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

Data Cleaning The following column names are removed from the datasets

- Remarks
- Source (2022-04)
- Source (2021-10)
- Source (2023-03)

T020

Information

- Dataset name: Vehicle registration (Bus)
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/serah/T020_TAS-VEP-018.ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-VEP-018 dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Stock*.

Unit : The unit is changed from *Number*to 10^6 *vehicle*.

Service : *Passenger*.

Mode : *Road*.

Vehicle Type : *Bus*.

Technology : *All*.

Fuel : *All*.

Data Cleaning The following column names are removed from the datasets

- Remarks
- Source (2022-04)
- Source (2021-10)
- Source (2023-03)

T021

Information

- Dataset name: Passengers Kilometer Travel - Roads
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/sandra/T021_TAS_PAG_005(3).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-PAG-005(3) dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.
subsection*Variable: *Passenger Activety.*

Unit : The unit is changed from *Million passenger kilometers* to 10^9 *passenger-km / yr.*

Service : *Passenger.*

Mode : *Road.*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

T022

Information

- Dataset name: Vehicle registration (LDV)
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/sandra/T022_TAS_VEP_017.ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-VEP-017 dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Stock*.

Unit : The unit is changed from Number to 10^6 *vehicle*.

Service : *Passenger*.

Mode : *Road*.

Vehicle Type : *LDV*.

Technology : *All*.

Fuel : *All*.

T023

Information

- Dataset name: Freight Transport - Tonne-km for Aviation (Domestic)
- Link to cleaning script: `https://github.com/linero-tech/item-internship/blob/main/code/sandra/T023_TAS_FRA_007(2).ipynb`

Source : Asian Transport Outlook Database `https://data.adb.org/dataset/asian-transport-outlook-database`(ATO2023 TAS-FRA-007(2) dataset downloaded on November 06, 2023)).

Country and ISO Code : All countries have an assigned ISO code.

ITEM Region : All countries belong to an ITEM region.

Variable : *Freight Activity.*

Unit : The unit is changed from *Million tonne kilometers* to 10^9 *tonne-km / yr.*

Service : *Freight.*

Mode : The mode is changed from *Aviation* to *Aviation (Domestic).*

Vehicle Type : *All.*

Technology : *All.*

Fuel : *All.*

T024

Information

- Dataset name: Trends in global fuel economy of new vehicles: 2005 - 2022
- Link to cleaning script: ??

Source : Cazzola, P., Paoli, L., & Teter, J. (2023). Trends in global fuel economy of new vehicles: 2005 - 2022 [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10148349.

Country and ISO Code :

ITEM Region : All countries belong to an ITEM region.

Variable

- "registrations" = *Stock*
- "specific_energy_cosumption_l_100km" = *Fuel Economy*

Unit

- *Stock*: 10^6 *vehicle*
- *Fuel Economy*: *l per 100km*

Service : *Passenger*.

Mode : *Road*.

Vehicle Type : *LDV*.

Technology

- *All*: include the following powertrain: ev, hv, ice, phev, fcv, mhv, unclassified.
- *BEV*: include the following powertrain: ev.
- *Conventional*: include the following powertrain: ice, hv, mhv.
- *Fuel Cell*: include the following powertrain: fcv.
- *PHEV*: include the following powertrain: phev.

Fuel

- When *Technology* = *All*, *Fuel* = *All*.
- When *Technology* = *BEV*, *Fuel* = *Electricity*.
- When *Technology* = *Conventional*, *Fuel* = *Liquid*.
- When *Technology* = *Fuel Cell*, *Fuel* = *Hydrogen*.
- When *Technology* = *PHEV*, *Fuel* = *Liquid+Electricity*.

Data Cleaning

Calculation of Stock The variable `stock` represents the total number of vehicles within each country, summed by vehicle class and by technology class across all segments, and then normalized to millions of vehicles. The calculation of `stock` is formalized as:

$$\begin{aligned} stock = \frac{1}{10^6}\Bigg(& \sum_{\substack{\text{vehicle class} \\ \text{technology class}}} \text{registration}_{\text{large car}} \\ & + \text{registration}_{\text{large suv}} \\ & + \text{registration}_{\text{lcv}} \\ & + \text{registration}_{\text{medium car}} \\ & + \text{registration}_{\text{small car}} \\ & + \text{registration}_{\text{small suv}} \\ & + \text{registration}_{\text{unclassified}} \Bigg) \end{aligned} \tag{1}$$

Where the sum accumulates the total registrations for each combination of vehicle and technology class across all segments. The final stock value is expressed in units of millions of vehicles.

Calculation of Fuel Economy The variable *Fuel Economy* quantifies the average energy efficiency of vehicles within a country, differentiated by *Technology* type ("powertrain"). It is calculated as a weighted sum of the variable `specific_energy_consumption_l_100km`, with the weighting given by the `registration` variable for each segment. The calculation is formalized as:

$$\text{fuel economy}_{\text{technology type}} = \frac{\sum_{i=1}^{n}(\text{specific\_energy\_consumption\_l\_100km}_i \times \text{registration}_i)}{\sum_{i=1}^{n} \text{registration}_i} \tag{2}$$

Where:

- $\texttt{specific\_energy\_consumption\_l\_100km}_i$ is the specific energy consumption for the i-th segment, measured in liters per 100 kilometers,
- $\texttt{registration}_i$ is the number of registered vehicles for the i-th segment,
- The summation runs over all segments within the dataset for a given technology type.

This results in `fuel economy` being expressed as the average liters per 100 kilometers (l/100km) for each technology type within a country, providing an indicator of the overall fuel efficiency.