Plan Overview

A Data Management Plan created using DMP Tool

Title: Vehicle Ownership Data by Race in Brazil

Creator: Glaucia Pereira

Affiliation: Multiplicidade Mobilidade Urbana

Principal Investigator: Glaucia Pereira

Data Manager: Glaucia Pereira

Project Administrator: Glaucia Pereira

Funder: Digital Curation Centre (dcc.ac.uk)

Template: Digital Curation Centre

Project abstract:

This project presents secondary data from the National Family Budget Survey 2017 - 2018 (POF 2017 - 2018) of the Brazilian Institute of Geography and Statistics (IBGE). POF 2017- 2018 collected data on people's consumption habits in 57,920 dwellings sampled, equivalent to a total of 68,862,296 dwellings in Brazil. The sample represents Brazil's urban areas, large regions (North, Northeast, Midwest, Southeast, and South), states, metropolitan regions, and states' capitals. Also, it is representative of Brazil and large regions in rural areas. Datasets regarding residents, inventory, and dwelling were joined to analyze vehicle ownership by residents' race. The registered vehicles were automobiles, motorcycles, and bicycles. Each row in the dataset contains a sampled dwelling, sample weight, total residents, number of black, brown, white, indigenous, yellow, and missing race residents, variable "family_race" categorized as "only black and brown", "only white" and "other combinations", and weighted and straightforward quantity of car, motorcycle, and bicycle. Researchers can use the data to analyze racial inequalities regarding vehicle ownership, urban access, and urban mobility by location in Brazil. It also allows comparing inventoried fleet versus administrative data fleet of automobiles and motorcycles in Brazil. Lastly, in an unprecedented way, it will enable estimating the bicycle fleet, data with no administrative data in the country.

Start date: 01-04-2021

End date: 03-31-2021

Last modified: 07-01-2024

Copyright information:

Vehicle Ownership Data by Race in Brazil

Data Collection

What data will you collect or create?

We will analyze secondary data from the National Family Budget Survey 2017 - 2018 (POF 2017 - 2018) of the Brazilian Institute of Geography and Statistics (IBGE). POF 2017- 2018 collected data on people's consumption habits in 57,920 dwellings sampled. We will create a database that related family vehicles and people race. The database will be available in CSV e XLSX format.

How will the data be collected or created?

The primary database is available on the IBGE website.

We will join and label tables and variables to merge dwelling data with people data. The whole process will be handle by R language. The code is available for assurance validation.

POF 2017-2018 is a one-time survey, and we do not expect to have a new data version.

Documentation and Metadata

What documentation and metadata will accompany the data?

We will incorporate metadata in R code and in a Readme.MD file. It will contain the date of creation, right of use, DOI, and ISSN.

Ethics and Legal Compliance

How will you manage any ethical issues?

The primary data is already public in Brazil. The sample unit preserves the identity of the dwelling surveyed.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

The R code is licensed under MIT license. The database will be shared under Creative Commons CC BY-SA 4.0

Storage and Backup

How will the data be stored and backed up during the research?

We estimated only 120 Mb for this project. So it is not necessary extra storage services. We will work on GitHub and the final project will be available at Harvard Dataverse too.

All files are on https://github.com/Multiplicidademobilidade/Vehicle-Ownership

How will you manage access and security?

There is no need for access control or security demands.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All files must be kept in the long-term.

What is the long-term preservation plan for the dataset?

The long-term plan is to keep the GitHub folder intact and keep it on Harvard Dataverse too.

Data Sharing

How will you share the data?

The potential users will know about this project on Multiplicidade Mobilidade Urbana social media e further communications. The data was made available on February 2021 on GitHub and on Harvard Dataverse. Harvard Dataverse generated a persistent identifier, https://doi.org/10.7910/DVN/NIFSKE

Are any restrictions on data sharing required?

There are no restrictions on data sharing.

Responsibilities and Resources

Who will be responsible for data management?

Glaucia Pereira, founder and researcher at Multiplicidade Mobilidade Urbana, is the responsible person in charge.

What resources will you require to deliver your plan?

There are no extra resources required.