ipea Instituto de Pesquisa Econômica Aplicada

Título do capítulo	SECTION 3 **PUBLIC TRANSPORT DATA**
Autor(es)	Rafael H. M. Pereira Daniel Herszenhut
DOI	DOI: http://dx.doi.org/10.38116/9786556350653section3

Título do livro	**Introduction to Urban Accessibility: a practical guide with R**
Organizadores(as)	Rafael H. M. Pereira Daniel Herszenhut
Volume	1
Série	-
Cidade	Rio de Janeiro
Editora	Instituto de Pesquisa Econômica Aplicada (Ipea)
Ano	2023
Edição	1a
ISBN	9786556350653
DOI	DOI: http://dx.doi.org/10.38116/9786556350653

© Instituto de Pesquisa Econômica Aplicada – ipea 2023

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos). Acesse: http://www.ipea.gov.br/portal/publicacoes

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento e Orçamento.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

SECTION 3

PUBLIC TRANSPORT DATA

The purpose of this section is: i) to introduce the GTFS public transport data specification; ii) to show where to download GTFS data for Brazilian cities and for other cities around the globe; and iii) to show how to manipulate and analyze GTFS data using R.

Public transport data is a key element of transport planning in general, and of accessibility analyses in particular. To be used with confidence, this data needs to be reliable and of simple inspection and interpretation.

To meet these criteria, transport agencies, decision makers and researchers have been trying to use data structured according to open and collaborative specifications – that is, whose format is decided upon by a community of different actors, including data producers (e.g. public transport agencies) and consumers (e.g. researchers and software developers). Although an open specification does not necessarily improve the quality and reliability of the data it describes, it brings many advantages that promote knowledge-sharing and transparency of analyses and applications that depend on it – factors that can substantially improve data quality and reliability.

The widespread usage of a standard data format to represent public transport systems promotes the development of computational tools and softwares that analyze and make use of this data, which helps creating a space in which actors from different cities and countries can learn from and support each other. Thus, an application developed by a Brazilian transport agency can easily be used by a researcher in the United States, a Japanese developer or another transport agency in South Africa – as long as, of course, they organize their data in the same format. Moreover, the more widely this format is used, the greater the reliability of the specification itself, as multiple actors tend to expand their ability to use, interpret and inspect this data.

The open and collaborative data specification most widely used in public transport planning and operation is the GTFS format, short for General Transit Feed Specification. As shown in chapter 3, GTFS feeds are also important pieces of data when estimating urban accessibility levels by public transport. In this section, we will learn more about GTFS data, how it is structured and how to work with it in R.