

Package ‘metrosp’

May 8, 2026

Type Package

Title 'São Paulo' Metro Passenger Demand Data

Version 1.0.0

Description Provides passenger demand data for the 'São Paulo' metro system, covering 2012 to 2025. Datasets include monthly passenger entries and transported counts by line, average weekday station entries, daily station entries, and spatial geometries for metro and commuter train lines and stations.

License MIT + file LICENSE

Encoding UTF-8

Language en-US

LazyData true

Depends R (>= 3.5.0)

Suggests dplyr, ggplot2, knitr, quarto, rmarkdown, sf, testthat (>= 3.0.0)

Config/testthat/edition 3

RoxygenNote 7.3.3

URL <https://github.com/viniciusoike/metrosp>,
<https://viniciusoike.github.io/metrosp/>

BugReports <https://github.com/viniciusoike/metrosp/issues>

VignetteBuilder quarto

NeedsCompilation no

Author Vinicius Oike [aut, cre]

Maintainer Vinicius Oike <viniciusoike@gmail.com>

Repository CRAN

Date/Publication 2026-05-05 19:00:02 UTC

Contents

lines	2
metro_colors	3
metro_lines	4
passengers_entrance	4
passengers_transported	5
stations	6
station_averages	7
station_daily	8

Index	10
--------------	-----------

lines	<i>Metro and Train Line Routes</i>
-------	------------------------------------

Description

Spatial line geometries for Sao Paulo metro (METRO SP) and commuter train (CPTM) lines, including both currently operating lines and planned future expansions.

Usage

lines

Format

An sf data frame with LINESTRING geometry (CRS: WGS84 / EPSG:4326) and the following columns:

line_number Official line number (integer).

line_name_pt Portuguese color name of the line (character).

line_name English color name of the line (character).

company_name Operating company name (character).

type Either "metro" (METRO SP) or "train" (CPTM) (character).

status Either "current" (operating) or "future" (planned expansion) (character).

geometry Line route geometry (sfc_LINESTRING).

Details

Requires the **sf** package to work with spatial features. The distinction between types isn't always consistent, but we follow GeoSampa's classification. Broadly speaking, the "metro" runs undergrounds as a subway, and "train" runs above grounds as a commuter rail (although there are exceptions)

Source

GeoSampa, Prefeitura de Sao Paulo. <https://geosampa.prefeitura.sp.gov.br/>

See Also

[stations](#) for station point locations.

metro_colors

Metro SP Official Line Colors

Description

A named character vector of official hex color codes for the six metro lines operated by METRO SP (Lines 1-3, 5, 15) and ViaMobilidade Line 4.

Usage

```
metro_colors
```

Format

A named character vector of length 6. Names are English color names; values are hex color codes:

Blue Line 1 — "#171796"

Green Line 2 — "#007A5E"

Red Line 3 — "#ED2E38"

Yellow Line 4 — "#FFD525"

Lilac Line 5 — "#874ABF"

Silver Line 15 — "#8F8F8C"

Details

Colors follow the official METRO SP and ViaMobilidade branding. Only the six currently operating metro lines are included; CPTM train lines and planned future lines (e.g., Line 6 Orange, Line 17 Gold) are not covered.

See Also

[metro_lines](#) for the full line reference table.

`metro_lines`*Metro SP Line Reference Table*

Description

A reference tibble mapping metro line numbers to their Portuguese and English color names. Covers all METRO SP and ViaMobilidade lines including planned future lines and the network total.

Usage`metro_lines`**Format**

A tibble with 13 rows and 3 columns:

line_number Official line number (integer). Includes 1, 2, 3, 4, 5, 6, 15, 16, 17, 19, 20, 22, and 99 (network total).

line_name_pt Portuguese color name of the line (character).

line_name English color name of the line (character).

Details

This dataset serves as a dimension/lookup table for joining line names onto passenger and station datasets. Not all lines have passenger data — some (e.g., Lines 6, 16, 17) are planned future lines with only spatial geometry available in [lines](#).

See Also

[metro_colors](#) for official hex color codes, [lines](#) for spatial line geometries.

`passengers_entrance`*Passengers Entering Metro SP Stations by Line*

Description

Monthly count of passengers entering Sao Paulo metro stations, aggregated by metro line. Data covers October 2017 through 2025, sourced from the METRO SP transparency portal.

Usage`passengers_entrance`

Format

A data frame with the following columns:

date First day of the month (Date).

line_number Metro line number: 1, 2, 3, 4, 5, 15, or 99 for network total (integer).

metric_abb Abbreviated metric code (character). One of: "total", "mdu", "msa", "mdo", "max".

value Passenger count in thousands (numeric).

metric Measurement type in Portuguese (character). One of: "Total", "Media dos Dias Uteis", "Media dos Sabados", "Media dos Domingos", "Maxima Diaria".

line_name English name of the metro line (character).

line_name_pt Portuguese name of the metro line (character).

year Calendar year (integer).

Details

Lines 1, 2, 3, and 15 come from the METRO transparency portal (2017-2025). Lines 4 (Amarela) and 5 (Lilas) come from the Insper Dataverse source (2020-2025). The network total (line_number = 99) may not be available for all years.

Values represent thousands of passengers (e.g., a value of 900 means 900,000 passengers).

Metrics:

- total: Total passengers in the month
- mdu: Average on business days (Media dos Dias Uteis)
- msa: Average on Saturdays (Media dos Sabados)
- mdo: Average on Sundays (Media dos Domingos)
- max: Daily maximum (Maxima Diaria)

Source

Companhia do Metropolitano de Sao Paulo (METRO). <https://transparencia.metrosp.com.br/dataset/demanda>

passengers_transported

Passengers Transported by Metro SP Line

Description

Monthly count of passengers transported by Sao Paulo metro, aggregated by metro line. Data covers October 2017 through 2025, sourced from the METRO SP transparency portal.

Usage

passengers_transported

Format

A data frame with the following columns:

date First day of the month (Date).

line_number Metro line number: 1, 2, 3, 5, 15, or 99 for network total (integer).

metric_abb Abbreviated metric code (character).

value Passenger count in thousands (numeric).

metric Measurement type in Portuguese (character).

line_name English name of the metro line (character).

line_name_pt Portuguese name of the metro line (character).

year Calendar year (integer).

Details

Lines 1, 2, 3, and 15 come from the METRO transparency portal (2017-2025). Line 5 (Lilas) is available October 2017 - December 2019 only. Line 4 (Amarela) is not available in this dataset (the Dataverse source does not include transported data for Lines 4/5). The network total (line_number = 99) may not be available for all years.

Values represent thousands of passengers (e.g., a value of 900 means 900,000 passengers).

Metrics:

- total: Total passengers in the month
- mdu: Average on business days (Media dos Dias Uteis)
- msa: Average on Saturdays (Media dos Sabados)
- mdo: Average on Sundays (Media dos Domingos)
- max: Daily maximum (Maxima Diaria)

Source

Companhia do Metropolitano de Sao Paulo (METRO). <https://transparencia.metrosp.com.br/dataset/demanda>

stations

Metro and Train Station Locations

Description

Spatial point locations for Sao Paulo metro (METRO SP) and commuter train (CPTM) stations, including both currently operating stations and planned future stations.

Usage

stations

Format

An sf data frame with POINT geometry (CRS: WGS84 / EPSG:4326) and the following columns:

station_name Station name in title case (character).

line_number Line number the station belongs to (integer).

line_name_pt Portuguese color name of the line (character).

line_name English color name of the line (character).

company_name Operating company name (character).

type Either "metro" (METRO SP) or "train" (CPTM) (character).

status Either "current" (operating) or "future" (planned expansion) (character).

geometry Station location (sfc_POINT).

Details

Requires the `sf` package to work with spatial features. The distinction between types isn't always consistent, but we follow GeoSampa's classification. Broadly speaking, the "metro" runs undergrounds as a subway, and "train" runs above grounds as a commuter rail (although there are exceptions)

Source

GeoSampa, Prefeitura de Sao Paulo. <https://geosampa.prefeitura.sp.gov.br/>

See Also

[lines](#) for line route geometries, [station_averages](#) for passenger data by station.

station_averages	<i>Average Weekday Passenger Entries by Station</i>
------------------	---

Description

Monthly average of weekday (business day) passenger entries for each station in the Sao Paulo metro system. Data covers October 2017 through 2025, sourced from the METRO SP transparency portal.

Usage

station_averages

Format

A data frame with the following columns:

date First day of the month (Date).

line_number Metro line number (integer).

station_name Name of the metro station (character).

avg_passenger Average weekday passenger entries in thousands (numeric).

line_name English name of the metro line (character).

line_name_pt Portuguese name of the metro line (character).

year Calendar year (integer).

Details

Only the weekday average (mdu) metric is available at the station level. For line-level data with all 5 metrics, see [passengers_entrance](#).

Station coverage varies by line:

- Line 1 (Azul/Blue): 23 stations
- Line 2 (Verde/Green): 14 stations
- Line 3 (Vermelha/Red): 18 stations
- Line 4 (Amarela/Yellow): available 2020-2025 (Insper Dataverse source)
- Line 5 (Lilas/Lilac): October 2017 - December 2019 (METRO portal) and 2020-2025 (Insper Dataverse source)
- Line 15 (Prata/Silver): 10 stations in 2020, 11 from 2021 onward

Source

Companhia do Metropolitano de Sao Paulo (METRO). <https://transparencia.metrosp.com.br/dataset/demanda>

station_daily

Daily Passenger Entries by Metro SP Station

Description

Daily passenger entries at each station in the Sao Paulo metro system. Data covers 2020 through 2025. Lines 1, 2, 3, and 15 come from the METRO SP transparency portal; Lines 4 and 5 come from the Insper Dataverse source.

Usage

station_daily

Format

A data frame with the following columns:

date Date of observation (Date).

line_number Metro line number: 1, 2, 3, 4, 5, or 15 (integer).

station_name Full station name (character).

passengers Daily passenger entries in thousands (numeric).

line_name English name of the metro line (character).

line_name_pt Portuguese name of the metro line (character).

station_code Three-letter station abbreviation used internally by METRO (character). NA for Lines 4 and 5 (Dataverse source).

year Calendar year (integer).

Details

Values represent thousands of passengers (e.g., a value of 50 means 50,000 passengers entering the station that day).

Station coverage by line:

- Line 1 (Azul/Blue): 23 stations
- Line 2 (Verde/Green): 14 stations
- Line 3 (Vermelha/Red): 18 stations
- Line 4 (Amarela/Yellow): available 2020-2025 (Insper Dataverse); station_code is NA
- Line 5 (Lilas/Lilac): available 2020-2025 (Insper Dataverse); station_code is NA
- Line 15 (Prata/Silver): 10 stations in 2020, 11 from 2021 onward (Jardim Colonial added)

Some stations appear on multiple lines (e.g., Ana Rosa on Lines 1 and 2, Paraiso on Lines 1 and 2, Se on Lines 1 and 3). These are recorded separately for each line.

Source

Companhia do Metropolitano de Sao Paulo (METRO). <https://transparencia.metrosp.com.br/dataset/demanda>

See Also

[station_averages](#) for monthly weekday averages, [passengers_entrance](#) for monthly line-level totals.

Index

* datasets

- lines, [2](#)
- metro_colors, [3](#)
- metro_lines, [4](#)
- passengers_entrance, [4](#)
- passengers_transported, [5](#)
- station_averages, [7](#)
- station_daily, [8](#)
- stations, [6](#)

lines, [2](#), [4](#), [7](#)

metro_colors, [3](#), [4](#)
metro_lines, [3](#), [4](#)

passengers_entrance, [4](#), [8](#), [9](#)
passengers_transported, [5](#)

station_averages, [7](#), [7](#), [9](#)
station_daily, [8](#)
stations, [3](#), [6](#)