

Package ‘tidyprf’

August 3, 2026

Title Tidy Access to Brazilian Federal Highway Police ('PRF') Data

Version 0.1.1

Description Download and read Brazilian Federal Highway Police ('PRF') open data on traffic accidents by person, by occurrence, and traffic violations
<<https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-da-prf>>.
Data are retrieved on demand as 'Parquet' files distributed via 'GitHub' releases and cached locally for reuse.

License MIT + file LICENSE

URL <https://github.com/bonijoao/tidyprf>,
<https://bonijoao.github.io/tidyprf/>

BugReports <https://github.com/bonijoao/tidyprf/issues>

Depends R (>= 4.1)

Imports arrow, cli, dplyr, fs, httr2, purrr, rlang, tibble

Suggests httptest2, knitr, rmarkdown, testthat (>= 3.0.0), withr

VignetteBuilder knitr

Config/testthat/edition 3

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 7.3.3

NeedsCompilation no

Author João Paulo Assis Bonifácio [aut, cre] (ORCID:
<<https://orcid.org/0009-0009-2335-8391>>),
Geraldo Magela da Cruz Pereira [aut] (ORCID:
<<https://orcid.org/0000-0001-6280-4870>>),
Pedro Mambelli Fernandes [aut]

Maintainer João Paulo Assis Bonifácio <jpab.27@hotmail.com>

Repository CRAN

Date/Publication 2026-08-03 17:50:02 UTC

Contents

codebook	2
get_accidents	3
get_crashes	3
get_violations	4
info_accidents	5
info_crashes	5
info_violations	6
prf_cache	6
prf_cache_clear	7
prf_years	7
Index	9

codebook	<i>Variable codebook for tidyprf datasets</i>
----------	---

Description

Bilingual (English/Portuguese) variable descriptions for all three PRF datasets. Used internally by [info_accidents\(\)](#), [info_crashes\(\)](#), and [info_violations\(\)](#).

Usage

codebook

Format

A tibble with 97 rows and 5 columns:

dataset Dataset name: "accidents", "crashes", or "violations"

variable Variable name as returned by the `get_*`() functions

type R type abbreviation (chr, int, dbl, date)

description_en Description in English

description_pt Description in Portuguese

Source

Compiled from the official PRF variable dictionaries published at <https://www.gov.br/prf/pt-br/aceso-a-informacao/dados-abertos/dados-abertos-da-prf>.

get_accidents	<i>Get PRF accident data by person</i>
---------------	--

Description

Downloads and returns accident records at the person level (one row per person involved in each accident). Files are cached after first download.

Usage

```
get_accidents(year, uf = NULL, br = NULL, severity = NULL)
```

Arguments

year	Integer or vector of integers. Year(s) to fetch (2007-2026).
uf	Character or NULL. State abbreviation(s), e.g. "SP". NULL = all.
br	Integer or NULL. Federal highway number(s), e.g. 101. NULL = all.
severity	Character or NULL. "fatal", "injured", or "no_victims". NULL = all.

Value

A tibble.

Examples

```
## Not run:
get_accidents(2023)
get_accidents(2020:2023, uf = "SP", severity = "fatal")
df <- get_accidents(c(2019:2021, 2023), br = 101)

## End(Not run)
```

get_crashes	<i>Get PRF accident data by occurrence</i>
-------------	--

Description

Downloads and returns accident records at the occurrence level (one row per accident, with aggregated person and vehicle counts). Files are cached after first download.

Usage

```
get_crashes(year, uf = NULL, br = NULL, severity = NULL)
```

Arguments

year	Integer or vector of integers. Year(s) to fetch (2007-2026).
uf	Character or NULL. State abbreviation(s). NULL = all.
br	Integer or NULL. Federal highway number(s). NULL = all.
severity	Character or NULL. "fatal", "injured", or "no_victims". NULL = all.

Value

A tibble.

Examples

```
## Not run:
get_crashes(2023)
get_crashes(2020:2023, uf = c("SP", "RJ"))

## End(Not run)
```

get_violations	<i>Get PRF infraction data</i>
----------------	--------------------------------

Description

Downloads PRF infraction records and returns an Arrow query. Files are 100-300 MB; use `dplyr::filter()` before `dplyr::collect()` to avoid loading the full dataset into memory.

Usage

```
get_violations(year, uf = NULL, br = NULL)
```

Arguments

year	Integer or vector of integers. Available: 2019-2020, 2022-2026.
uf	Character or NULL. State of infraction (uf_infracao). NULL = all.
br	Integer or NULL. Federal highway number (num_br_infracao). NULL = all.

Value

An Arrow query. Call `dplyr::collect()` to load into a tibble.

Examples

```
## Not run:
# Filter before collecting to avoid loading 243 MB into RAM
get_violations(2024, uf = "SP") |>
  dplyr::filter(cod_infracao == "55412") |>
  dplyr::collect()

## End(Not run)
```

info_accidents	<i>Variable descriptions for get_accidents()</i>
----------------	--

Description

Variable descriptions for get_accidents()

Usage

```
info_accidents(lang = "en")
```

Arguments

lang "en" (default) or "pt".

Value

A tibble with columns variable, type, description.

Examples

```
info_accidents()  
info_accidents(lang = "pt")
```

info_crashes	<i>Variable descriptions for get_crashes()</i>
--------------	--

Description

Variable descriptions for get_crashes()

Usage

```
info_crashes(lang = "en")
```

Arguments

lang "en" (default) or "pt".

Value

A tibble with columns variable, type, description.

Examples

```
info_crashes()
```

info_violations	<i>Variable descriptions for get_violations()</i>
-----------------	---

Description

Variable descriptions for get_violations()

Usage

```
info_violations(lang = "en")
```

Arguments

lang "en" (default) or "pt".

Value

A tibble with columns variable, type, description.

Examples

```
info_violations(lang = "pt")
```

prf_cache	<i>Show locally cached PRF data files</i>
-----------	---

Description

Show locally cached PRF data files

Usage

```
prf_cache(dataset = NULL)
```

Arguments

dataset Character or NULL. One of "accidents", "crashes", "violations". NULL (default) shows all.

Value

A tibble with columns dataset, year, size_mb, cached_at.

Examples

```
prf_cache()
```

prf_cache_clear	<i>Delete locally cached PRF data files</i>
-----------------	---

Description

Delete locally cached PRF data files

Usage

```
prf_cache_clear(dataset = NULL, year = NULL)
```

Arguments

dataset	Character or NULL. Dataset to clear. NULL = all.
year	Integer or NULL. Year(s) to clear. NULL = all years.

Value

Invisible NULL.

Examples

```
## Not run:  
prf_cache_clear("violations", year = 2024)  
prf_cache_clear()  
  
## End(Not run)
```

prf_years	<i>Show years available in the PRF catalog</i>
-----------	--

Description

Show years available in the PRF catalog

Usage

```
prf_years(dataset = NULL)
```

Arguments

dataset	Character or NULL. One of "accidents", "crashes", "violations". NULL (default) returns all three.
---------	---

Value

A tibble with columns dataset, year, rows, size_mb.

Examples

```
## Not run:  
prf_years()  
prf_years("accidents")  
  
## End(Not run)
```

Index

* datasets

codebook, [2](#)

codebook, [2](#)

dplyr::collect(), [4](#)

dplyr::filter(), [4](#)

get_accidents, [3](#)

get_crashes, [3](#)

get_violations, [4](#)

info_accidents, [5](#)

info_accidents(), [2](#)

info_crashes, [5](#)

info_crashes(), [2](#)

info_violations, [6](#)

info_violations(), [2](#)

prf_cache, [6](#)

prf_cache_clear, [7](#)

prf_years, [7](#)