Package 'crimedata'

Title Access Crime Data from the Open Crime Database

Version 0.3.5

Description Gives convenient access to publicly available police-recorded open crime data from large cities in the United States that are included in the Crime Open Database <https://osf.io/zyaqn/>.

Depends R (>= 3.2.0)

License MIT + file LICENSE

Language en-US

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Suggests testthat, knitr, rmarkdown

Imports digest, dplyr, osfr, purrr, rlang, sf, stringr

URL http://pkgs.lesscrime.info/crimedata/, https://github.com/mpjashby/crimedata

BugReports https://github.com/mpjashby/crimedata/issues

VignetteBuilder knitr

NeedsCompilation no

Author Matthew Ashby [aut, cre, cph] (<https://orcid.org/0000-0003-4201-9239>)

Maintainer Matthew Ashby <matthew.ashby@ucl.ac.uk>

Repository CRAN

Date/Publication 2023-11-09 00:20:02 UTC

R topics documented:

---

block_geoid_to	*Convert Census Block GEOIDs*

Description

Convert the GEOID of a 2016 US Census block to the name or GEOID for the corresponding state, county, tract or block group.

Usage

```
block_geoid_to(geoid, to, name = FALSE)

block_geoid_to_state(geoid, name = TRUE)

block_geoid_to_county(geoid, name = TRUE)

block_geoid_to_tract(geoid)

block_geoid_to_block_group(geoid)
```

Arguments

geoid	A character vector of 15-digit US Census block GEOIDs.
to	One of "state", "county", "tract", "block group" or (as an alias) "blockgroup".
name	Should the function return the state/county name rather than FIPS code?

Details

For details of the format of US Census GEOIDs, see https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html.

Value

A character vector of GEOIDs or names.

Examples

```
block_geoid_to("360810443021005", to = "county", name = TRUE)
```

| crimedata | *crimedata: a package for accessing US city crime data* |

Description

Access incident-level crime data from the Open Crime Database

Crime Open Database

The Crime Open Database (CODE) is a service that makes it convenient to use crime data from multiple US cities in research on crime. All the data are available to use for free as long as you acknowledge the source of the data.

For more about CODE data, see https://osf.io/zyaqn/.

Accessing the data

To access CODE data, call get_crime_data. Data are returned as a 'tidy' tibble with each row corresponding to one recorded crime.

Chicago data license

This site provides applications using data that has been modified for use from its original source, https://www.chicago.gov/, the official website of the City of Chicago. The City of Chicago makes no claims as to the content, accuracy, timeliness, or completeness of any of the data provided at this site. The data provided at this site is subject to change at any time. It is understood that the data provided at this site is being used at one's own risk.

Author(s)

Maintainer: Matthew Ashby <matthew.ashby@ucl.ac.uk> (ORCID) [copyright holder]

See Also

Useful links:

- http://pkgs.lesscrime.info/crimedata/
- https://github.com/mpjashby/crimedata
- Report bugs at https://github.com/mpjashby/crimedata/issues

---

get_crime_data   *Get Data from the Open Crime Database*

Description

Retrieves data from the Open Crime Database for the specified years. Latitude and longitude are specified using the WGS 84 (EPSG:4326) co-ordinate reference system.

Usage

```
get_crime_data(
  years = NULL,
  cities = NULL,
  type = "sample",
  cache = TRUE,
  quiet = !interactive(),
  output = "tbl"
)
```

Arguments

years	A single integer or vector of integers specifying the years for which data should be retrieved. If NULL (the default), data for the most recent year will be returned.
cities	A character vector of city names for which data should be retrieved. Case insensitive. If NULL (the default), data for all available cities will be returned.
type	Either "sample" (the default), "core" or "extended".
cache	Should the result be cached and then re-used if the function is called again with the same arguments?
quiet	Should messages and warnings relating to data availability and processing be suppressed?
output	Should the data be returned as a tibble by specifying "tbl" (the default) or as a simple features (SF) object using WGS 84 by specifying "sf"?

Details

By default this function returns a one-percent sample of the 'core' data. This is the default to minimize accidentally requesting large files over a network.

Setting type = "core" retrieves the core fields (e.g. the type, co-ordinates and date/time of each offense) for each offense. The data retrieved by setting type = "extended" includes all available fields provided by the police department in each city. The extended data fields have not been harmonized across cities, so will require further cleaning before most types of analysis.

Requesting all data (more than 17 million rows) may lead to problems with memory capacity. Consider downloading smaller quantities of data (e.g. using type = "sample") for exploratory analysis.

Setting output = "sf" returns the data in simple features format by calling `sf::st_as_sf(..., crs = 4326, remove = FALSE)`

For more details see the help vignette: `vignette("introduction", package = "crimedata")`

Value

A tibble containing data from the Open Crime Database.

---

homicides15   *Homicides in nine cities in 2015*

Description

Dataset containing records of homicides in nine large US cities in 2015, obtained from the Crime Open Database.

Usage

`homicides15`

Format

A tibble with 1,922 rows and 15 variables:

uid an integer unique identifier for the offense

city_name name of the city in which the crime occurred

offense_code offense code, modified from the FBI NIBRS offense code

offense_type offense type name

date_single date (and, in most cases, time) of the offense

address approximate address of the offense*

longitude approximate longitude

latitude approximate latitude

location_type type of location*

location_category category of location type*

fips_state two-digit FIPS state code (possibly with leading zero)

fips_county three-digit FIPS county code (possibly with leading zero)

tract six-digit code for 2016 census tract

block_group one-digit code for 2016 census block group

block four-digit code for 2016 census block

Details

More details of the data format are available on the [Crime Open Database website](). Variables marked * are only available for some of the data, due to limitations in the data published by some cities.

The variables in this dataset mirror those obtained by calling get_crime_data(type = "core"), except that some fields have been removed because they are redundant (e.g. if they have the same value for all rows in this dataset).

Source

https://osf.io/zyaqn/

---

list_crime_data   *List Data Available in the Open Crime Database*

Description

Get a tibble showing what years of crime data are available from which cities in the Open Crime Database.

Usage

```
list_crime_data(quiet = !interactive())
```

Arguments

quiet   Should messages and warnings relating to data availability and processing be suppressed?

Value

A tibble

---

nycvehiclethefts   *Thefts of motor vehicles 2014 to 2017*

Description

Dataset containing records of thefts of motor vehicles in New York City from 2014 to 2017, obtained from the [Crime Open Database]().

Usage

```
nycvehiclethefts
```

Format

A tibble with 35,746 rows and 13 variables:

uid an integer unique identifier for the offense

date_single date (and, in most cases, time) half-way between the first and last possible dates at which the offense could have occurred

date_start first possible date (and, in most cases, time) at which the offense could have occurred

date_send last possible date (and, in most cases, time) at which the offense could have occurred

longitude approximate longitude

latitude approximate latitude

location_type type of location*

location_category category of location type*

fips_state two-digit FIPS state code (possibly with leading zero)

fips_county three-digit FIPS county code (possibly with leading zero)

tract six-digit code for 2016 census tract

block_group one-digit code for 2016 census block group

block four-digit code for 2016 census block

Details

More details of the data format are available on the [Crime Open Database website](). Variables marked * are only available for some of the data, due to limitations in the data published by some cities.

The variables in this dataset mirror those obtained by calling `get_crime_data(type = "core")`, except that some fields have been removed because they are redundant (e.g. if they have the same value for all rows in this dataset).

Source

<https://osf.io/zyaqn/>

Index