Package ‘covid19sf’

July 3, 2021

Title The Covid19 San Francisco Dataset
Version 0.1.1
Maintainer Rami Krispin <rami.krispin@gmail.com>
License MIT + file LICENSE
Encoding UTF-8
LazyData true
Depends R (&gt;= 2.10)
Imports devtools, dplyr, lubridate, magrittr, mapview, plotly, sf
Suggests testthat, knitr, rmarkdown
RoxygenNote 7.1.1
VignetteBuilder knitr
URL https://github.com/RamiKrispin/covid19sf
BugReports https://github.com/RamiKrispin/covid19sf/issues
NeedsCompilation no
Author Rami Krispin [aut, cre]
Repository CRAN
Date/Publication 2021-07-03 15:20:02 UTC

R topics documented:
covid19sf_age ......................................................... 2
covid19sf_demo ..................................................... 3
covid19sf_gender .................................................. 4
covid19sf_geo ....................................................... 5
covid19sf_homeless ............................................... 6
covid19sf_hospital ............................................... 7
covid19sf_age

San Francisco COVID-19 Cases Summarized by Age Group

Description
This dataset represents the COVID-19 positive confirmed cases by age group. Demographic and transmission data are based on information reported from case interviews, laboratories, and providers. This data may not be immediately available for recently reported cases and data will change to reflect as information becomes available. Cumulative counts of 5 or fewer are excluded from the dataset. More information about the data available here.

Usage
covid19sf_age

Format
An object class data.frame with 5 variables

specimen_collection_date date which case was recorded in YYYY-MM-DD format.
age_group case age group c("under 18", "18-30", "31-40", "41-50", "51-60", "71-80")
new_confirmed_cases Daily new confirmed cases
cumulative_confirmed_cases Cumulative number of confirmed cases
last_updated The table last update time in POSIX format

Details
The dataset contains the daily summary of covid19 cases in San Francisco by age group.

Source
San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data portal website.
**Description**

This dataset represents the COVID-19 positive confirmed cases by race and ethnicity. Demographic data are based on information reported from case interviews, laboratories, and providers. This data may not be immediately available for recently reported cases and data will change to reflect as information becomes available. Cumulative counts of 5 or fewer are excluded from the dataset. More information about the data available [here](#).

**Usage**

covid19sf_demo

**Format**

An object class data.frame with 5 variables

- **specimen_collection_date**: Date which case was recorded in YYYY-MM-DD format.
- **race_ethnicity**: The cases race/ethnicity
- **new_confirmed_cases**: Daily new confirmed cases
- **cumulative_confirmed_cases**: Cumulative confirmed cases
- **last_updated**: The table last update time in POSIX format

**Details**

The dataset contains the daily summary of covid19 cases in San Francisco by race/ethnicity group

**Source**

San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data portal [website](#).

**Examples**

data(covid19sf_demo)

head(covid19sf_demo)
San Francisco COVID-19 Cases Summarized by Gender

Description

This dataset represents the COVID-19 positive confirmed cases by gender. Demographic and transmission data are based on information reported from case interviews, laboratories, and providers. This data may not be immediately available for recently reported cases and data will change to reflect as information becomes available. Cumulative counts of 5 or fewer are excluded from the dataset. More information about the data available here.

Usage

covid19sf_gender

Format

An object class data.frame with 5 variables

- specimen_collection_date  Date which case was recorded in YYYY-MM-DD format.
- gender  The cases gender ("Female", "Male", "Trans Female", "Unknown")
- new_confirmed_cases  Total cases confirmed cases per date and gender category
- cumulative_confirmed_cases  Cumulative confirmed cases by category
- last_updated  The table last update time in POSIX format

Details

The dataset contains the daily summary of covid19 cases in San Francisco by gender

Source

San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data protal website.

Examples

data(covid19sf_gender)

head(covid19sf_gender)
San Francisco COVID-19 Cases and Deaths Summarized by Geography

Description
Medical provider confirmed COVID-19 cases and confirmed COVID-19 related deaths in San Francisco, CA aggregated by several different geographic areas and normalized by 2018 American Community Survey (ACS) 5-year estimates for population data to calculate rate per 10,000 residents. More information about the data available here.

Usage
covid19sf_geo

Format
An object class sf and data.frame with 8 variables.

| area_type | Area type, c("ZCTA", "Analysis Neighborhood", "Census Tract", "Citywide") |
| id | area id |
| count | The count of cases in the area |
| rate | The rate of cases in the area, calculated as (count/acs_population) * 10000 which is a rate per 10,000 residents |
| deaths | The number of cases in the area |
| acs_population | The population from the latest 5-year estimates from the American Community Survey (2014-2018)) |
| last_updated | Last update of the data in POSIXc format |
| geometry | The area polygon data |

Details
The dataset contains a summary of covid19 cases in San Francisco by geographic area.

Source
San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data protal website.

Examples
data(covid19sf_geo)
head(covid19sf_geo)
library(sf)
# Plotting SF Covid19 counts using base plot function
# Plotting by zip code
plot(covid19sf_geo[which(covid19sf_geo$area_type == "ZCTA"),
c("count", "geometry")],
main = "Covid19 Cases by ZIP Code")
# Plotting by neighborhood
plot(covid19sf_geo[which(covid19sf_geo$area_type ==
  "Analysis Neighborhood"),
c("count", "geometry")],
main = "Covid19 Cases by Neighborhood")
# Plotting by census tract
plot(covid19sf_geo[which(covid19sf_geo$area_type == "Census Tract"),
c("count", "geometry")],
main = "Covid19 Cases by Census Tract")
plot(covid19sf_geo[which(covid19sf_geo$area_type == "Census Tract"),
c("rate", "geometry")],
main = "Covid19 Cases Rate per 10,000 by Census Tract")

---

**covid19sf_homeless**  
*San Francisco COVID-19 Cases Summarized by Homelessness Status*

**Description**

This dataset represents the COVID-19 positive confirmed cases by homelessness. Demographic and transmission data are based on information reported from case interviews, laboratories, and providers. This data may not be immediately available for recently reported cases and data will change to reflect as information becomes available. Cumulative counts of 5 or fewer are excluded from the dataset. More information about the data available [here](#).

**Usage**

`covid19sf_homeless`

**Format**

An object class data.frame with 5 variables

- **specimen_collection_date** Date which case was recorded in YYYY-MM-DD format.
- **homelessness_status** The homelessness status, a single category variable c("Homeless")
- **new_confirmed_cases** Total cases confirmed cases per date
- **cumulative_confirmed_cases** Cumulative confirmed cases
- **last_updated** The table last update time in POSIX format

**Details**

The dataset contains the daily summary of covid19 cases in San Francisco by homelessness status
**covid19sf_hospital**

**Source**
San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data portal [website](#).

**Examples**
```r
data(covid19sf_homeless)
head(covid19sf_homeless)
```

---

**covid19sf_hospital  San Francisco COVID-19 Hospital Capacity**

**Description**
Data on daily hospital bed use and available capacity at San Francisco acute care hospitals from April 2020 onward. Long Term Care facilities (like Laguna Honda and Kentfield) are not included in this data as acute care patients cannot be admitted to these facilities. More information about the data available [here](#).

**Usage**
```
covid19sf_hospital
```

**Format**
An object class data.frame with 5 variables

- **hospital**  The hospital name, currently a single categorical variable, c("All SF Acute Hospitals")
- **date**  Date which the data was recorded in YYYY-MM-DD format
- **bed_type**  The bed type, c("Intensive Care Surge", "Acute Care", "Acute Care Surge", "Intensive Care")
- **status**  The bed category status, c("Available", "COVID-19 (Confirmed & Suspected)", "Other Patients")
- **count**  The bed count

**Details**
The dataset contains a summary of San Francisco hospital bed status

**Source**
San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data portal [website](#).
Examples
data(covid19sf_hospital)
head(covid19sf_hospital)

covid19sf_hospitalizations

San Francisco COVID-19 Hospitalizations

Description
Count of COVID+ patients admitted to the hospital. Patients who are hospitalized and test positive for COVID-19 may be admitted to an acute care bed (a regular hospital bed), or an intensive care unit (ICU) bed. This data shows the daily total count of COVID+ patients in these two bed types, and the data reflects totals from all San Francisco Hospitals. More information about the data available here.

Usage
covid19sf_hospitalizations

Format
An object class data.frame with 5 variables

- **reportdate**  date which case was recorded in YYYY-MM-DD format.
- **hospital** The hospital which patients were admitted, currently it labeled under "All SF Hospitals"
- **dphcategory** The type of hospitalization bed, either an acute care bed (a regular hospital bed), or an intensive care unit (ICU) bed
- **covidstatus** The patient diagnostic, either PUI (Patient Under Investigation) or COVID+ (positive case)
- **patientcount** Daily cases count

Details
Each record represents how many people were hospitalized on the date recorded in either an ICU bed or acute care bed (shown as Med/Surg under DPHCategory field)

Source
San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data protocol website.

Examples
data(covid19sf_hospitalizations)
head(covid19sf_hospitalizations)
Description
This dataset includes aggregate data on the type, status, population served, and individuals placed at each alternative housing site under contract with HSA. More information about the data available here.

Usage

```r
covid19sf_housing
```

Format

An object class data.frame with 8 variables

- `site_id` Site ID
- `status` The site status, c("Active", "In Preparation")
- `facility_type` The facility type, c("Hotel", "Safe Sleep", "Congregate", "RV")
- `site_type` The site type, c("SIP: COVID-Negative/Unknown", "I/Q", "SS: COVID-Negative/Unknown", "SIP: Post-COVID")
- `units_occupied` Number of units occupied per site
- `total_units` Total number of units available
- `population_covid_status` The population covid status, c("COVID Negative/Unknown", "COVID Positive", "Post-COVID")
- `date_updated` Date which data was updated in YYYY-MM-DD format)

Details

The dataset contains a summary of covid19 housing site in San Francisco by site, facility and covid19 status.

Source

San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data protal website.

Examples

```r
data(covid19sf_housing)
head(covid19sf_housing)
```
covid19sf_refresh  
*Refreshing the covid19sf Package Datasets*

**Description**

The function enables to keep the package datasets with most recent data available on the package main repository. The main repository is refreshed on a daily basis.

**Usage**

```r
covid19sf_refresh(force = FALSE)
```

**Arguments**

- `force`  
  A boolean, if set to TRUE will update the package if new data is available automatically

---

covid19sf_summary  
*San Francisco COVID-19 Cases Summarized by Date, Transmission and Case Disposition*

**Description**

This dataset represents the COVID-19 positive confirmed cases and deaths by day and transmission type. The transmission data are based on information reported from case interviews, laboratories, and providers. This data may not be immediately available for recently reported cases and data will change to reflect as information becomes available. More information about the data available [here](#).

**Usage**

```r
covid19sf_summary
```

**Format**

An object class data.frame with 5 variables

- **specimen_collection_date**  
  Date which case was recorded in YYYY-MM-DD format.

- **case_disposition**  
  The case disposition c("Confirmed", "Death")

- **transmission_category**  
  The case transmission category c("Community", "From Contact", "Unknown")

- **case_count**  
  Daily cases count

- **last_updated**  
  The table last update time in POSIX format
Details
The dataset contains the daily summary of covid19 cases in San Francisco by transmission and case disposition.

Source
San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data portal website.

Examples
```r
data(covid19sf_summary)
head(covid19sf_summary)

# Case disposition summary
table(covid19sf_summary$case_disposition)

# Transmission category
table(covid19sf_summary$transmission_category)

# Summary of case disposition and transmission category
table(covid19sf_summary$case_disposition, covid19sf_summary$transmission_category)
```

Description
Case information on COVID-19 Laboratory testing. This data includes a daily count of test results reported, and how many of those were positive, negative, and indeterminate. Reported tests include tests with a positive, negative or indeterminate result. Indeterminate results, which could not conclusively determine whether COVID-19 virus was present, are not included in the calculation of percent positive. Testing for the novel coronavirus is available through commercial, clinical, and hospital laboratories, as well as the SFDPH Public Health Laboratory. More information about the data available here.

Usage
covid19sf_tests

Format
An object class data.frame with 7 variables

`specimen_collection_date` date which case was recorded in YYYY-MM-DD format.
**tests**  Daily tests count
**pos**  Number of positive cases
**pct**  Percentage of positive cases
**neg**  Number of negative cases
**indeterminate**  Number of indeterminate cases
**last_updated**  The table last update time in POSIX format

**Details**
A daily COVID-19 testing results report

**Source**
San Francisco, Department of Public Health - Population Health Division through San Francisco Opne Data protal website.

**Examples**
```r
data(covid19sf_tests)
head(covid19sf_tests)
```

---

### covid19sf_test_loc  
*San Francisco COVID-19 Testing Locations*

**Description**
A list of testing locations including address and coordinates for mapping. More information about the data available [here](#)

**Usage**
covid19sf_test_loc

**Format**
An object class sf and data.frame with 17 variables

- **id**  Location ID
- **medical_home**  Medical home
- **name**  The medical name
- **address**  The medical address
- **phone_number**  The medical phone number
- **phone_number_formatted**  The medical phone number formatted
- **testing_hours**  The medical testing hours
The dataset contains the San Francisco testing location information.

**Source**
San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data portal website.

**Examples**
```r
data(covid19sf_test_loc)
head(covid19sf_test_loc)
```

**Description**
This dataset represents doses of COVID-19 vaccine administered in California to residents of San Francisco. The data is broken down by multiple demographic slices. The three dose types are counted separately, i.e. (1) first doses administered as a part of a two-dose vaccination, (2) second doses administered as part of a two-dose vaccination, and (3) single-dose vaccines administered.

**Usage**
covid19sf_vaccine_demo
Format

An object class data.frame with 15 variables

overall_segment  Segment (universe) of analysis. Unique combination of administering_provider_type, age_group, and demographic_group. Filter to a single option to derive meaningful totals.

administering_provider_type  Providers included in a given overall_segment. Two possible values: 'All' (including SF DPH) or 'DPH Only'

age_group  Age range included in a given overall_segment

demographic_group  Type of demographic group included in a given overall_segment (e.g. Age, Race/Ethnicity)

demographic_subgroup  Specific demographic group counted in a given record (e.g. 16-24, Asian)

demographic_subgroup_sort_order  Numeric sort order for all demographic_subgroups. Convenient for maintaining consistent ordering across multiple data visualizations.

total_1st_doses  Total number of first doses administered

total_2nd_doses  Total number of second doses administered

total_single_doses  Total number of single dose vaccines administered

total_recipients  Total number of unique vaccine recipients

total_series_completed  Total number of individuals fully vaccinated (those having received the second dose of a two-dose vaccine or one dose of a single-dose vaccine)

subgroup_population  2018 5-year American Community Survey population estimates for given DEMOGRAPHIC_SUBGROUP

age_group_population  2018 5-year American Community Survey population estimates for overall AGE_GROUP

data_as_of  Timestamp for last update date in source system

data_loaded_at  Timestamp when the record (row) was most recently updated in Socrata

Details

The dataset contains a summary of COVID-19 vaccine doses given to San Franciscans by demographics

Source

San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data portal website.

Examples

data(covid19sf_vaccine_demo)

head(covid19sf_vaccine_demo)
COVID-19 Vaccine Doses Given to San Franciscans by Demographics Over Time

Description

This dataset represents doses of COVID-19 vaccine administered in California to San Francisco residents over time. The data is broken down by multiple demographic slices. The three dose types are counted separately, i.e. (1) first doses administered as a part of a two-dose vaccination, (2) second doses administered as part of a two-dose vaccination, and (3) single-dose vaccines administered.

Usage

covid19sf_vaccine_demo_ts

Format

An object class data.frame with 21 variables

date_administered Date vaccination administered

overall_segment Segment (universe) of analysis. Unique combination of administering_provider_type, age_group, and demographic_group. Filter to a single option to derive meaningful totals.

administering_provider_type Providers included in a given overall_segment. Two possible values: 'All' (including SF DPH) or 'DPH Only'

age_group Age range included in a given overall_segment

demographic_group Type of demographic group included in a given overall_segment (e.g. Age, Race/Ethnicity)

demographic_subgroup Specific demographic group counted in a given record (e.g. 16-24, Asian)

demographic_subgroup_sort_order Numeric sort order for all demographic_subgroup. Convenient for maintaining consistent ordering across multiple data visualizations.

new_1st_doses Count of 1st doses administered for vaccines that take two doses to complete

new_2nd_doses Count of 2nd doses administered for vaccines that take two doses to complete

new_single_doses Count of doses administered for vaccines that take one dose to complete

new_series_completed Count of individuals newly fully vaccinated on a given day (given the 2nd dose of a two-dose vaccine or one dose of a single dose vaccine)

new_recipients Count of individuals vaccinated (with any dose) for the first time according to CA’s records

cumulative_1st_doses Cumulative total of 1st doses administered for vaccines that take two doses to complete

cumulative_2nd_doses Cumulative total of 2nd doses administered for vaccines that take two doses to complete
**cumulative_single_doses** Cumulative total of doses administered for vaccines that take one dose to complete

**cumulative_series_completed** Cumulative total individuals fully vaccinated (given the 2nd dose of a two-dose vaccine or one dose of a single dose vaccine)

**cumulative_recipients** Cumulative total individuals vaccinated (with any dose) according to CA’s records

**subgroup_population** American Community Survey population estimates for given demographic_subgroup

**age_group_population** American Community Survey population estimates for overall age_group

**data_as_of** Timestamp for last update date in source system

**data_loaded_at** Timestamp when the record (row) was most recently updated here in the Open Data Portal

**Details**

The dataset contains a time series of COVID-19 vaccine doses given to San Franciscans by demographics

**Source**

San Francisco, Department of Public Health - Population Health Division through San Francisco Open Data portal website.

**Examples**

```r
data(covid19sf_vaccine_demo_ts)

head(covid19sf_vaccine_demo_ts)
```
Index

* COVID19
  - covid19sf_age, 2
  - covid19sf_demo, 3
  - covid19sf_gender, 4
  - covid19sf_geo, 5
  - covid19sf_homeless, 6
  - covid19sf_hospital, 7
  - covid19sf_hospitalizations, 8
  - covid19sf_housing, 9
  - covid19sf_summary, 10
  - covid19sf_test_loc, 12
  - covid19sf_tests, 11
  - covid19sf_vaccine_demo, 13
  - covid19sf_vaccine_demo_ts, 15

* age
  - covid19sf_age, 2

* bed
  - covid19sf_hospital, 7

* datasets
  - covid19sf_age, 2
  - covid19sf_demo, 3
  - covid19sf_gender, 4
  - covid19sf_geo, 5
  - covid19sf_homeless, 6
  - covid19sf_hospital, 7
  - covid19sf_hospitalizations, 8
  - covid19sf_housing, 9
  - covid19sf_summary, 10
  - covid19sf_test_loc, 12
  - covid19sf_tests, 11
  - covid19sf_vaccine_demo, 13
  - covid19sf_vaccine_demo_ts, 15

* demographic
  - covid19sf_demo, 3

* ethnicity
  - covid19sf_demo, 3

* gender
  - covid19sf_gender, 4

* geo
  - covid19sf_geo, 5

* homeless
  - covid19sf_homeless, 6

* hospital
  - covid19sf_hospital, 7
  - covid19sf_hospitalizations, 8

* housing
  - covid19sf_housing, 9

* map
  - covid19sf_geo, 5

* medical
  - covid19sf_test_loc, 12

* race
  - covid19sf_demo, 3

* series
  - covid19sf_vaccine_demo_ts, 15

* summary
  - covid19sf_age, 2
  - covid19sf_demo, 3
  - covid19sf_gender, 4
  - covid19sf_geo, 5
  - covid19sf_homeless, 6
  - covid19sf_hospital, 7
  - covid19sf_hospitalizations, 8
  - covid19sf_housing, 9
  - covid19sf_summary, 10
  - covid19sf_test_loc, 12
  - covid19sf_tests, 11
  - covid19sf_vaccine_demo, 13

* testing
  - covid19sf_test_loc, 12

* tests
  - covid19sf_tests, 11

* time
  - covid19sf_tests, 11

* vaccine
  - covid19sf_vaccine_demo, 13
  - covid19sf_vaccine_demo_ts, 15

* covid19sf_age, 2
covid19sf_demo, 3
covid19sf_gender, 4
covid19sf_geo, 5
covid19sf_homeless, 6
covid19sf_hospital, 7
covid19sf_hospitalizations, 8
covid19sf_housing, 9
covid19sf_refresh, 10
covid19sf_summary, 10
covid19sf_test_loc, 12
covid19sf_tests, 11
covid19sf_vaccine_demo, 13
covid19sf_vaccine_demo_ts, 15