# Package ‘covid19sf’

October 12, 2022

**Title**  The Covid19 San Francisco Dataset  
**Version**  0.1.2  
**Maintainer**  Rami Krispin <rami.krispin@gmail.com>  
**License**  MIT + file LICENSE  
**Encoding**  UTF-8  
**LazyData**  true  
**Depends**  R (>= 2.10)  
**Imports**  devtools, dplyr, lubridate, magrittr, mapview, plotly, sf  
**Suggests**  testthat, knitr, rmarkdown  
**RoxygenNote**  7.1.2  
**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**  2021-12-19 08:20:02 UTC  

## R topics documented:

<table>
<thead>
<tr>
<th>Function</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>covid19sf_geo</td>
<td>2</td>
</tr>
<tr>
<td>covid19sf_hospital</td>
<td>3</td>
</tr>
<tr>
<td>covid19sf_hospitalizations</td>
<td>4</td>
</tr>
<tr>
<td>covid19sf_housing</td>
<td>5</td>
</tr>
<tr>
<td>covid19sf_population</td>
<td>6</td>
</tr>
<tr>
<td>covid19sf_refresh</td>
<td>7</td>
</tr>
</tbody>
</table>
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

```r
data(covid19sf_geo)
head(covid19sf_geo)
library(sf)
# Plotting SF Covid19 counts using base plot function
plot(covid19sf_geo[which(covid19sf_geo$area_type == "Census Tract"),
                    c("rate", "geometry")],
     main = "Covid19 Cases Rate per 10,000 by Census Tract")
```

---

**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**

```r
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 protal [website](#).
covid19sf_hospitalizations

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 protal website.

Examples

data(covid19sf_hospitalizations)

head(covid19sf_hospitalizations)
covid19sf_housing

San Francisco COVID-19 Alternative Housing Sites

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
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 Open Data portal website.

Examples

```r
data(covid19sf_housing)
head(covid19sf_housing)
```
## Description
This dataset shows San Francisco COVID-19 cases by population characteristics and by specimen collection date. Cases are included on the date the positive test was collected. Population characteristics are subgroups, or demographic cross-sections, like age, race, or gender. The City tracks how cases have been distributed among different subgroups. This information can reveal trends and disparities among groups. Data is lagged by five days, meaning the most recent specimen collection date included is 5 days prior to today. Tests take time to process and report, so more recent data is less reliable. More details available [here](#).

## Usage
```r
covid19sf_population
```

## Format
An object class `data.frame` with 7 variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>specimen_collection_date</td>
<td>Date which case was recorded in YYYY-MM-DD format.</td>
</tr>
<tr>
<td>characteristic_type</td>
<td>Overall topic area for a given population characteristic. These are subgroups or demographic cross-sections, like age</td>
</tr>
<tr>
<td>characteristic_group</td>
<td>Each group or category within a characteristic type or topic area. ex 0-4 yrs, 5-10 yrs</td>
</tr>
<tr>
<td>characteristic_group_sort_order</td>
<td>Sort order of characteristic group to aid in visualizing data</td>
</tr>
<tr>
<td>new_cases</td>
<td>Cases are counted as confirmed on the date of specimen collection after a positive lab test result</td>
</tr>
<tr>
<td>cumulative_cases</td>
<td>Cumulative Cases</td>
</tr>
<tr>
<td>population_estimate</td>
<td>Population estimate for a given characteristic type and characteristic group</td>
</tr>
</tbody>
</table>

## Details
The dataset contains a summary of COVID-19 cases overtime by population characteristics

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

## Examples
```r
data(covid19sf_population)
head(covid19sf_population)
```
**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_tests**  
*San Francisco COVID-19 Tests*

**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**

```r
covid19sf_tests
```

**Format**

An object class data.frame with 6 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
Details

A daily COVID-19 testing results report

Source

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

Examples

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>Location ID</td>
</tr>
<tr>
<td>medical_home</td>
<td>Medical home</td>
</tr>
<tr>
<td>name</td>
<td>The medical name</td>
</tr>
<tr>
<td>address</td>
<td>The medical address</td>
</tr>
<tr>
<td>phone_number</td>
<td>The medical phone number</td>
</tr>
<tr>
<td>phone_number_formatted</td>
<td>The medical phone number formatted</td>
</tr>
<tr>
<td>testing_hours</td>
<td>The medical testing hours</td>
</tr>
<tr>
<td>popup_orPermanent</td>
<td>The medical testing type, c(&quot;Permanent&quot;, &quot;Pop-Up&quot;)</td>
</tr>
<tr>
<td>location_type</td>
<td>The medical location type, c(&quot;Private&quot;, &quot;Public&quot;)</td>
</tr>
<tr>
<td>eligibility</td>
<td>Eligibility information for accessing testing at this location</td>
</tr>
<tr>
<td>cta_text</td>
<td>The call to action used for the web map</td>
</tr>
<tr>
<td>cta_link</td>
<td>The call to action link for the button on the web map</td>
</tr>
<tr>
<td>sample_collection_method</td>
<td>The method for collecting samples at the lab</td>
</tr>
<tr>
<td>lap</td>
<td>The lab name</td>
</tr>
<tr>
<td>latitude</td>
<td>The medical latitude point</td>
</tr>
<tr>
<td>longitude</td>
<td>The medical longitude point</td>
</tr>
<tr>
<td>geometry</td>
<td>The medical geometry details</td>
</tr>
</tbody>
</table>
COVID-19 Vaccine Doses Given to San Franciscans by Demographics

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. Here

NOTE: This dataset is no longer supported and will be deprecated on the next release (v0.1.3). The covid19sf_population dataset is an alternative for covid19sf_vaccine_demo.

Usage

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**

```r
data(covid19sf_vaccine_demo)
head(covid19sf_vaccine_demo)
```

**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 19 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

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)
```

---

**covid19sf_vaccine_geo**  
*San Francisco COVID-19 Vaccines Given to San Franciscans by Geography*

**Description**
This dataset represents the COVID-19 vaccinations given to SF residents summarized by the geographic region of their residential address. All vaccines given to SF residents are included, no matter where the vaccination took place (the vaccine may have been administered in San Francisco or outside of San Francisco). Data provides counts for people who have received at least one dose and people who have completed a vaccine series. A vaccine series is complete after an individual has received both doses of a two-dose vaccine or one dose of a one-dose vaccine. More information about the data available [here](https://example.com).

**Usage**
covid19sf_vaccine_geo

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

- **id**  area id
- **area_type**  Area type, c("Analysis Neighborhood", "Summary")
- **count_vaccinated_by_dph**  Count of residents in the given geographic region who have received at least one dose administered by DPH
- **count_vaccinated**  Count of residents in the given geographic region who have received at least one dose regardless of who administered the vaccine
- **count_series_completed**  Count of residents in the given geographic region who have completed a vaccine series
- **acs_population**  2019 5-year American Community Survey population estimate for the given geographic region (all ages)
- **percent_pop_series_completed**  The total count of population that have completed a vaccine series by population estimate (acs_population)
- **last_updated**  Last update of the data in POSIXc format
- **geometry**  The area polygon data
Details

The dataset contains a summary of covid19 vaccination in San Francisco by neighborhood.

Source

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

Examples

```r
data(covid19sf_vaccine_geo)
head(covid19sf_vaccine_geo)
library(sf)
library(dplyr)

df <- covid19sf_vaccine_geo %>% filter(area_type == "Analysis Neighborhood")
plot(df[, c("percent_pop_series_completed", "geometry")],
     main = "San Francisco - Percentage of Fully Vaccinated Population")
```
Index

* COVID19
  - covid19sf_geo, 2
  - covid19sf_hospital, 3
  - covid19sf_hospitalizations, 4
  - covid19sf_housing, 5
  - covid19sf_population, 6
  - covid19sf_test_loc, 8
  - covid19sf_tests, 7
  - covid19sf_vaccine_demo, 9
  - covid19sf_vaccine_demo_ts, 10
  - covid19sf_vaccine_geo, 12

* bed
  - covid19sf_hospital, 3

* datasets
  - covid19sf_geo, 2
  - covid19sf_hospital, 3
  - covid19sf_hospitalizations, 4
  - covid19sf_housing, 5
  - covid19sf_population, 6
  - covid19sf_test_loc, 8
  - covid19sf_tests, 7
  - covid19sf_vaccine_demo, 9
  - covid19sf_vaccine_demo_ts, 10
  - covid19sf_vaccine_geo, 12

* demographic
  - covid19sf_population, 6

* geo
  - covid19sf_geo, 2
  - covid19sf_vaccine_geo, 12

* hospital
  - covid19sf_hospital, 3
  - covid19sf_hospitalizations, 4

* housing
  - covid19sf_housing, 5

* map
  - covid19sf_geo, 2
  - covid19sf_vaccine_geo, 12

* medical
  - covid19sf_test_loc, 8

* series
  - covid19sf_vaccine_demo_ts, 10

* summary
  - covid19sf_geo, 2
  - covid19sf_hospital, 3
  - covid19sf_hospitalizations, 4
  - covid19sf_housing, 5
  - covid19sf_population, 6
  - covid19sf_test_loc, 8
  - covid19sf_tests, 7
  - covid19sf_vaccine_demo, 9
  - covid19sf_vaccine_geo, 12

* testing
  - covid19sf_test_loc, 8

* tests
  - covid19sf_tests, 7

* time
  - covid19sf_vaccine_demo_ts, 10

* vaccine
  - covid19sf_population, 6
  - covid19sf_vaccine_demo, 9
  - covid19sf_vaccine_demo_ts, 10
  - covid19sf_vaccine_geo, 12

  - covid19sf_geo, 2
  - covid19sf_hospital, 3
  - covid19sf_hospitalizations, 4
  - covid19sf_housing, 5
  - covid19sf_population, 6
  - covid19sf_test_loc, 8
  - covid19sf_tests, 7
  - covid19sf_vaccine_demo, 9
  - covid19sf_vaccine_demo_ts, 10
  - covid19sf_vaccine_geo, 12