

# Package ‘xSub’

June 30, 2022

**Title** Cross-National Data on Sub-National Violence

**Version** 3.0.2

**Description** Tools to download and merge data files on sub-national conflict, violence and protests from <<http://www.x-sub.org>>.

**URL** <https://github.com/zhukovyuri/xSub>

**Depends** R (>= 3.3.2)

**Imports** countrycode, haven, RCurl

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**NeedsCompilation** no

**Author** Yuri Zhukov [aut, cre],  
Christian Davenport [aut],  
Nadiya Kostyuk [aut]

**Maintainer** Yuri Zhukov <[zhukov@umich.edu](mailto:zhukov@umich.edu)>

**Repository** CRAN

**Date/Publication** 2022-06-30 16:40:02 UTC

## R topics documented:

get_xSub . . . . .	2
get_xSub_multi . . . . .	4
info_xSub . . . . .	6
xSub_census_individual_raw . . . . .	7
xSub_census_individual_spatial . . . . .	8
xSub_census_multiple_raw . . . . .	9
xSub_census_multiple_spatial . . . . .	9

<b>Index</b>	<b>11</b>
--------------	-----------

---

 get\_xSub

*Get xSub file*


---

## Description

This function downloads individual files from [www.x-sub.org](http://www.x-sub.org). Function produces a data.frame, for the user's choice of data source, country, spatial and temporal units, and (optionally) writes this data.frame to disk, in multiple formats.

## Usage

```
get_xSub(
  data_source,
  sources_type = "individual",
  data_type = "spatial panel",
  country_iso3 = NULL,
  country_name = NULL,
  space_unit,
  time_unit,
  geo_window = "1 km",
  time_window = "1 day",
  dyad_type = "undirected",
  out_dir = getwd(),
  write_file = TRUE,
  write_format = "csv",
  verbose = FALSE
)
```

## Arguments

data_source	Name of data source. See <code>info_xSub()</code> for full list.
sources_type	Type of data sources ("individual" or "multiple"). Character string.
data_type	Type of dataset ("event" or "panel"). Character string.
country_iso3	Country code (ISO3). See <code>info_xSub()</code> for full list.
country_name	Country name. See <code>info_xSub()</code> for full list.
space_unit	Geographic level of analysis. Character string. Can be one of "adm0" (country), "adm1" (province), "adm2" (district), "priogrid" (grid cell), "clea" (electoral constituency). See <code>info_xSub(details=TRUE)</code> for availability by country.
time_unit	Temporal level of analysis. Character string. Can be one of "year", "month", "week", "day". See <code>info_xSub(details=TRUE)</code> for availability by country.
geo_window	Geographic window (if <code>source_type="multiple"</code> ). Could be either of "1 km" (default) or "5 km". Character string or vector.
time_window	Time window (if <code>source_type="multiple"</code> ). Could be either of "1 day" (default) or "2 day". Character string or vector.

dyad_type	Time window (if source_type="multiple"). Could be either of "undirected" (default) or "directed". Character string or vector.
out_dir	Path to directory where files will be saved.
write_file	Logical. If write_file=TRUE, selected file will be written to disk, at location specified by out_dir.
write_format	Output file format. Can be one of "csv" (comma-separated values, default), "R" (RData format, compatible with R statistical programming language), "STATA" (dta format, compatible with Stata 14).
verbose	Logical. When verbose=TRUE, file download progress is printed to console.

**See Also**

[info\\_xSub](#), [get\\_xSub\\_multi](#)

**Examples**

```
# Check which countries are available for ACLED
info_xSub(data_source="ACLED")

# Download ACLED data for Egypt, at country-year level
## Not run:
my_file <- get_xSub(data_source = "ACLED",country_iso3 = "EGY",
                    space_unit = "adm0",time_unit = "year")

## End(Not run)

# Download ACLED data for Egypt, at district-month level
## Not run:
my_file <- get_xSub(data_source = "ACLED",country_iso3 = "EGY",
                    space_unit = "adm2",time_unit = "month")

## End(Not run)

# With country name instead of ISO3 code
## Not run:
my_file <- get_xSub(data_source = "ACLED",country_name = "Egypt",
                    space_unit = "adm2",time_unit = "month")

## End(Not run)

## Not run:
# Download ACLED data for Egypt, event level
my_file <- get_xSub(data_source = "ACLED",country_iso3 = "EGY",
                    data_type = "event")

## End(Not run)

## Not run:
# Download multiple source data for Egypt, at province-month level
my_file <- get_xSub(sources_type = "multiple",country_iso3 = "EGY",
                    space_unit = "adm1",time_unit = "month", geo_window = "1 km",
```

```

        time_window = "1 day", dyad_type = "undirected")

## End(Not run)

```

---

```

get_xSub_multi          Get xSub files for multiple countries

```

---

### Description

This function downloads and merges mutiple country files from [www.x-sub.org](http://www.x-sub.org). Syntax is similar to `get_xSub()`.

### Usage

```

get_xSub_multi(
  data_source,
  sources_type = "individual",
  data_type = "spatial panel",
  country_iso3 = NULL,
  space_unit,
  time_unit,
  geo_window = "1 km",
  time_window = "1 day",
  dyad_type = "undirected",
  merge_files = TRUE,
  out_dir = getwd(),
  write_file = FALSE,
  write_format = "csv",
  verbose = FALSE
)

```

### Arguments

<code>data_source</code>	Name of data source. Character string. See <code>info_xSub()</code> for full list.
<code>sources_type</code>	Type of data sources ("individual" or "multiple"). Character string.
<code>data_type</code>	Type of dataset ("event" or "panel"). Character string.
<code>country_iso3</code>	Country codes (ISO3). Character string or vector. See <code>info_xSub()</code> for full list. If left blank, function will download all available countries for selected data source.
<code>space_unit</code>	Geographic level of analysis. Character string. Can be one of "adm0" (country), "adm1" (province), "adm2" (district), "priogrid" (grid cell), "clea" (electoral constituency). See <code>info_xSub(details=TRUE)</code> for availability by country.
<code>time_unit</code>	Temporal level of analysis. Character string. Can be one of "year", "month", "week", "day". See <code>info_xSub(details=TRUE)</code> for availability by country.
<code>geo_window</code>	Geographic window (if <code>source_type="multiple"</code> ). Could be either of "1 km" or "5 km". Character string or vector.

time_window	Time window (if source_type="multiple"). Could be either of "1 day" or "2 day". Character string or vector.
dyad_type	Time window (if source_type="multiple"). Could be either of "undirected" or "directed". Character string or vector.
merge_files	Logical. If merge_files=TRUE (default), function will combine individual country files into single data.frame, and write single file to disk. If merge_files=FALSE, function produces a list, and writes individual country files to disk separately.
out_dir	Path to directory where files will be saved. Character string.
write_file	Logical. If write_file=TRUE, selected file will be written to disk, at location specified by out_dir.
write_format	Output file format. Character string. Can be one of "csv" (comma-separated values, default), "R" (RData format, compatible with R statistical programming language), "STATA" (dta format, compatible with Stata 14).
verbose	Logical. When verbose=TRUE, file download progress is printed to console..

**See Also**

[info\\_xSub](#), [get\\_xSub](#)

[info\\_xSub](#), [get\\_xSub](#)

**Examples**

```
# Check which countries are available for GED
info_xSub(data_source="GED")

# Example with two countries
## Not run:
my_file <- get_xSub_multi(data_source = "PITF",country_iso3 = c("ALB","ARM"),
  space_unit = "adm0",time_unit = "year")

## End(Not run)

# Example with two countries
## Not run:
my_file <- get_xSub_multi(data_source = "GED",country_iso3 = c("EGY","AGO"),
  space_unit = "adm1",time_unit = "month")

## End(Not run)

# Example with two countries, multiple sources, event-level
## Not run:
my_file <- get_xSub_multi(sources_type = "multiple",data_type="event",country_iso3 = c("EGY","AGO"))

## End(Not run)

# Example with all countries (WARNING: this can take a long time to run)
## Not run:
my_file <- get_xSub_multi(data_source = "BeissingerProtest",country_iso3 = NULL,
  space_unit = "adm0",time_unit = "year")
```

```
## End(Not run)
```

---

```
info_xSub
```

```
Information on available xSub files
```

---

## Description

This function reports the availability of files on the `www.x-sub.org` server, and corresponding country codes and units of analysis. For additional info, see `www.x-sub.org/about/what-is-xsub`.

## Usage

```
info_xSub(
  details = FALSE,
  sources_type = "individual",
  data_type = "panel",
  data_source = NULL,
  country_iso3 = NULL,
  country_name = NULL,
  geo_window = NULL,
  time_window = NULL,
  dyad_type = NULL
)
```

## Arguments

<code>details</code>	Logical. If <code>details=TRUE</code> , function returns information on available units of analysis for each country.
<code>sources_type</code>	Type of data sources ("individual" or "multiple"). Character string.
<code>data_type</code>	Type of dataset ("event" or "panel"). Character string.
<code>data_source</code>	Subset results by data sources. Character string or vector.
<code>country_iso3</code>	Subset results by country codes (ISO3). Character string or vector.
<code>country_name</code>	Subset results by country name. Character string or vector.
<code>geo_window</code>	Geographic window (if <code>source_type="multiple"</code> ). Could be either of "1 km" or "5 km". Character string or vector.
<code>time_window</code>	Time window (if <code>source_type="multiple"</code> ). Could be either of "1 day" or "2 day". Character string or vector.
<code>dyad_type</code>	Time window (if <code>source_type="multiple"</code> ). Could be either of "undirected" or "directed". Character string or vector.

## See Also

[get\\_xSub](#), [get\\_xSub\\_multi](#)

**Examples**

```

# General info on data sources and countries
info_xSub()

# Available files for Pakistan
info_xSub(country_name = "Pakistan")

# Detailed info for Pakistan
info_xSub(details=TRUE, country_name = "Pakistan")

# Available files for SCAD data source
info_xSub(data_source = "SCAD")

# Available files for SCAD data source, event-level
info_xSub(data_source = "SCAD", data_type = "event")

# Multiple data sources, directed dyads
info_xSub(sources_type = "multiple", dyad_type = "directed")

# Multiple data sources, directed dyads, Russia
info_xSub(sources_type = "multiple", dyad_type = "directed", country_name = "Russia")

```

---

xSub\_census\_individual\_raw

*Census of individual-source event-level datasets in xSub (updated June 15, 2020)*

---

**Description**

A list of data sources and countries available for download. Used by info\_xSub()

**Usage**

```
xSub_census_individual_raw
```

**Format**

A list with 4 elements:

**level0\_bysource** Countries organized by data\_source. List object, where each sub-entry is also a list, containing entries for data\_source, country\_iso3, country\_name.

**level0\_bycountry** Data sources organized by country. List of data.frames, where each row is a country, with columns for country\_iso3, country\_name, data\_sources.

**level1** Detailed information on data sources, countries and spatial levels of analysis. data.frame, where each row is a source-country combination, with columns for data\_source, country\_iso3, country\_name, units.

**all\_countries** Vector of all country ISO3 codes. Used by get\_xSub\_multi.

**Source**

<http://www.x-sub.org/>

---

xSub\_census\_individual\_spatial

*Census of individual-source panel datasets in xSub (updated June 15, 2020)*

---

**Description**

A list of data sources, countries and levels of analysis available for download. Used by `info_xSub()`

**Usage**

`xSub_census_individual_spatial`

**Format**

A list with 6 elements:

**level0\_bysource** Countries organized by data\_source. List object, where each sub-entry is also a list, containing entries for data\_source, country\_iso3, country\_name.

**level0\_bycountry** Data sources organized by country. List of data.frames, where each row is a country, with columns for country\_iso3, country\_name, data\_sources.

**level1** Detailed information on data sources, countries and spatial levels of analysis. data.frame, where each row is a source-country combination, with columns for data\_source, country\_iso3, country\_name, space

**level2** Detailed information on data sources, countries, spatial and temporal levels of analysis. data.frame, where each row is a source-country-spatial unit combination, with columns for data\_source, country\_iso3, country\_name, space\_unit, time\_units.

**level3** File census. data.frame, where each row is a single file, with columns for file\_name, data\_source, country\_iso3, co

**all\_countries** Vector of all country ISO3 codes. Used by `get_xSub_multi`.

**Source**

<http://www.x-sub.org/>



---

 xSub\_census\_multiple\_raw

*Census of multiple-source event-level datasets in xSub (updated June 15, 2020)*

---

### Description

A list of data sources, countries and levels of analysis available for download. Used by `info_xSub()`

### Usage

`xSub_census_multiple_raw`

### Format

A list with 4 elements:

**level0\_bysource** Countries organized by `data_source`. List object, where each sub-entry is also a list, containing entries for `data_source`, `geo_window`, `time_window`, `dyad_type`, `country_iso3`, `country_name`.

**level0\_bycountry** Data sources organized by country. List of `data.frames`, where each row is a country, with columns for `country_iso3`, `country_name`, `geo_window`, `time_window`, `dyad_type`, `data_sources`.

**level1** Detailed information on data sources, countries and spatial levels of analysis. `data.frame`, where each row is a source-country combination, with columns for `data_source`, `geo_window`, `time_window`, `dyad_type`.

**all\_countries** Vector of all country ISO3 codes. Used by `get_xSub_multi`.

### Source

<http://www.x-sub.org/>

---

 xSub\_census\_multiple\_spatial

*Census of multiple-source panel datasets in xSub (updated June 15, 2020)*

---

### Description

A list of data sources, countries and levels of analysis available for download. Used by `info_xSub()`

### Usage

`xSub_census_multiple_spatial`

**Format**

A list with 6 elements:

**level0\_bysource** Countries organized by data\_source. List object, where each sub-entry is also a list, containing entries for data\_source,geo\_window,time\_window,dyad\_type,country\_iso3,country\_name.

**level0\_bycountry** Data sources organized by country. List of data.frames, where each row is a country, with columns for country\_iso3,country\_name,geo\_window,time\_window,dyad\_type,data\_sources.

**level1** Detailed information on data sources, countries and spatial levels of analysis. data.frame, where each row is a source-country combination, with columns for data\_source,geo\_window,time\_window,dyad\_type

**level2** Detailed information on data sources, countries, spatial and temporal levels of analysis. data.frame, where each row is a source-country-spatial unit combination, with columns for data\_source,geo\_window,time\_window,dyad\_type,country\_iso3,country\_name,space\_unit,time\_units.

**level3** File census. data.frame, where each row is a single file, with columns for file\_name,data\_source,geo\_window,time

**all\_countries** Vector of all country ISO3 codes. Used by get\_xSub\_multi.

**Source**

<http://www.x-sub.org/>

# Index

## \* datasets

- xSub\_census\_individual\_raw, [7](#)
- xSub\_census\_individual\_spatial, [8](#)
- xSub\_census\_multiple\_raw, [9](#)
- xSub\_census\_multiple\_spatial, [9](#)

get\_xSub, [2](#), [5](#), [6](#)

get\_xSub\_multi, [3](#), [4](#), [6](#)

info\_xSub, [3](#), [5](#), [6](#)

xSub\_census\_individual\_raw, [7](#)

xSub\_census\_individual\_spatial, [8](#)

xSub\_census\_multiple\_raw, [9](#)

xSub\_census\_multiple\_spatial, [9](#)