

# Package ‘congress’

May 18, 2022

**Title** A Tool for the Congress Data

**Version** 1.0

**Author** Caleb Lucas (<https://caleblucas.com/>)

**Description** A tool that imports, subsets, and exports the Congress Data dataset assembled by Matt Grossmann, Caleb Lucas, Josh McCrain, and Ian Ostrander. The Congress Data contains approximately 800 variables concerning all US congressional districts and covers 1789-2021. Users with only a basic understanding of R can subset this data across multiple dimensions, export their search results, create custom aggregations of the dataset, identify the citations associated with their searches, and more.

**License** GPL (>= 3)

**Encoding** UTF-8

**LazyData** true

**Language** en-US

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.1.2

**Imports** stringr,  
dplyr,  
readr,  
rlang,  
tidyselect,  
congressData

**Remotes** caleblucas/congressData

**VignetteBuilder** knitr

## R topics documented:

aggregate_cong_dat . . . . .	2
get_congress_data . . . . .	3
get_var_info . . . . .	3
<b>Index</b>	<b>5</b>

---

aggregate\_cong\_dat      *Load Congress Data into the R environment*

---

## Description

aggregate\_cong\_dat loads either a full or subsetting version of the full Congress Data dataset into the R environment as a dataframe.

## Usage

```
aggregate_cong_dat(
  states = NULL,
  related_to = NULL,
  sessions = NULL,
  census_nonperc_vars = "Mean",
  census_perc_vars = "Mean",
  bill_vars = "Mean",
  com_vars = "Mean",
  else_vars = "Mean"
)
```

## Arguments

states	Default is NULL. If left blank, returns all states. Takes a string or vector of strings of state names.
related_to	Default is "". Provide a string to search a variable's name, short/long descriptions from the codebook, and its citation for non-exact matches of a search term. For example, searching 'tax' will return variables with words like 'taxes' and 'taxable' in any of those columns.
sessions	Default is NULL. If left blank, returns all sessions. Input can be a single session (e.g. 50) or a two sessions that represent the first and last that you want in the outputted dataframe (such as c(1, 20)).
census_nonperc_vars	Default is 'Mean'. Method to aggregate from member-years to member-sessions. One of "Mean", "Sum", or "First". Requires a string.
census_perc_vars	Default is 'Mean'. Method to aggregate from member-years to member-sessions. One of "Mean", "Sum", or "First". Requires a string.
bill_vars	Default is 'Mean'. Method to aggregate from member-years to member-sessions. One of "Mean", "Sum", or "First". Requires a string.
com_vars	Default is 'Mean'. Method to aggregate from member-years to member-sessions. One of "Mean", "Sum", or "First". Requires a string.
else_vars	Default is 'Mean'. Method to aggregate from member-years to member-sessions. One of "Mean", "Sum", or "First". Requires a string.

---

get_congress_data	<i>Load Congress Data into the R environment</i>
-------------------	--------------------------------------------------

---

### Description

get\_congress\_data loads either a full or subsetting version of the full Congress Data dataset into the R environment as a dataframe.

### Usage

```
get_cong_data(states = NULL, related_to = "", years = NULL)
```

### Arguments

states	Default is NULL. If left blank, returns all states. Takes a string or vector of strings of state names.
related_to	Default is "". Provide a string to search a variable's name, short/long descriptions from the codebook, and its citation for non-exact matches of a search term. For example, searching 'tax' will return variables with words like 'taxes' and 'taxable' in any of those columns.
years	Default is NULL. If left blank, returns all years Input can be a single year (e.g. 2001) or a two year that represent the first and last that you want in the outputted dataframe (such as c(1989,20001)).

---

get_var_info	<i>Get information regarding the Congress Data variables</i>
--------------	--------------------------------------------------------------

---

### Description

get\_var\_info retrieves information regarding variables in the Congress Data dataset. The information available includes: the years each variable is observed in the data; a short and long description of each variable; and the source and citation/s for each variable. Citations are available in both bibtex and plain text.

### Usage

```
get_var_info(var_names = NULL, related_to = NULL)
```

### Arguments

var_names	Default is NULL. Takes a character string. If left blank the function does not subset by variable name. Note that this searches variable names, it does not subset by the specific string you provide. So "over65" will return a variable called 'over65' along with another called 'percent_over65'.
related_to	Default is NULL. Takes a character string. If the user supplies a character string, the function searches the other relevant fields (variable name, short/long description, and source) for string matches. Not case sensitive.

**Details**

Specifying no arguments returns all the information for all the variables in the Congress Data dataset.

**Examples**

```
# returns all variable information
get_var_info()

# match var names that contain "over65" - note this returns multiple variables
get_var_info(var_names = "over65")

get_var_info(related_to = c("pop", "femal"))
```

# Index

`aggregate_cong_dat`, [2](#)

`get_cong_data (get_congress_data)`, [3](#)

`get_congress_data`, [3](#)

`get_var_info`, [3](#)