



US data: `{tidycensus}` and `{tigris}` packages



US data:
`{tidycensus}` and
`{tigris}` packages

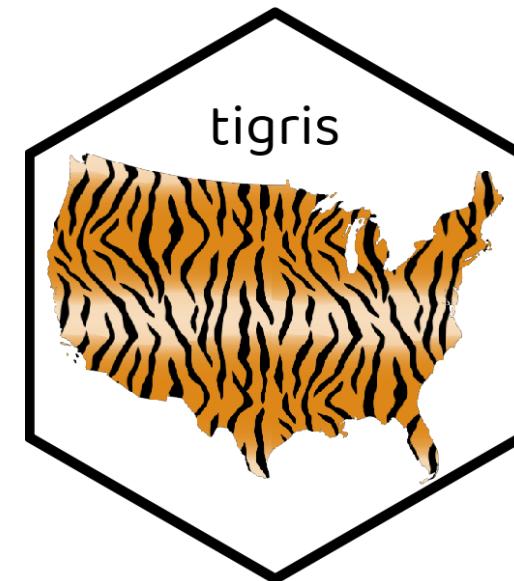


TIGER & {tigris}

The US Census Bureau is responsible for maintaining shapefiles and other geospatial datasets for contiguous US and its territories.

The collated dataset is called TIGER (Topologically Integrated Geographic Encoding and Referencing).

The `{tigris}` package provides programmatic access to **current and historic** components of TIGER.

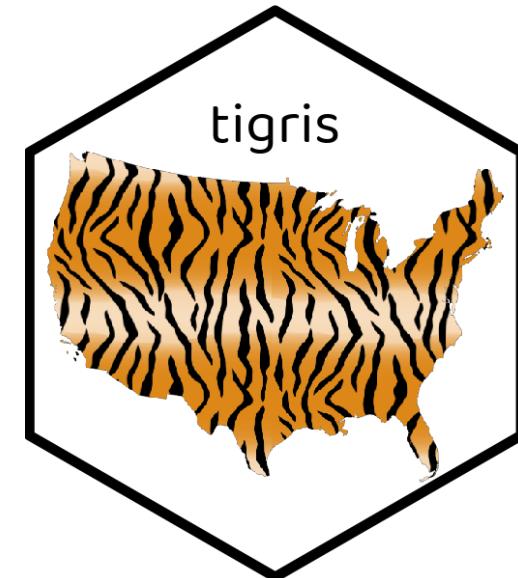




{tigris}

The `{tigris}` doesn't contain any datasets, they must be **downloaded** after the package is loaded.

Think carefully about how detailed a shapefile you require - for many functions the default `{sf}` objects are **large**.





(RStudio Coding Slide)



tigris::__(cb = TRUE)

It's a common mistake to forget to set cb = TRUE

```
states(resolution = "500k") %>%  
  filter(NAME == "Washington") %>%  
  mapview()
```

```
states(resolution = "500k", cb = TRUE) %>%  
  filter(NAME == "Washington") %>%  
  mapview()
```



(RStudio Coding Slide)



{tidycensus}

{tidycensus} provides programmatic access to two data sources:

The [Decennial US Census](#) which currently includes data for 1990, 2000 and 2010.

```
get_decennial(geography = "county",
               variables = "P005004",
               state = "VT")
```

The [5-year American Community Survey](#) (ACS) which currently includes data for 2009 through 2018.

```
get_acs(geography = "county",
         variables = "B19013_001",
         state = "VT")
```



{tidycensus} API

Before we can use the `{tidycensus}` package we must register for a **free** API key from the census bureau by following this workflow:

1. Obtain a key from the signup page: api.census.gov/data/key_signup.html
2. Check your email and click on the **activate your key** link

You've attempted to validate an unknown key. If it has been more than 48 hours since you submitted your request for this API key then the request has been removed from the system. Please request a new key and activate it within 48 hours.

3. Add your Census API key to your `.Renvironment` using `census_api_key()`

```
census_api_key("your-key", install = TRUE)
```



Searching for census variables

{tidycensus} uses **variable codes** to access data from the API.

Use `load_variable()` to find the variable code for the data you're interested in.

The process varies for the Census and ACS datasets:

Use `dataset = "sf1"` for **Decennial Census** data.

```
vars_census_2010 <-  
  load_variables(2010,  
                 dataset = "sf1")
```

Use `dataset = "acs5"` for **ACS** data.

```
vars_acs_2010 <-  
  load_variables(2018,  
                 dataset = "acs5")
```



(RStudio Coding Slide)



My Turn

I'm going to create a choropleth visualising the % of households that are rented in each state.





(RStudio Coding Slide)



Your Turn

Produce the same choropleth as I did but obtain data from the 2015 American Community Survey (ACS) dataset.

1. Use `load_variables()` to find the variables in the 2015 American Community Survey.
2. Use `get_acs()` to extract these two variables.
3. Use `pivot_wider()` to prepare the data for joining with the US states shapefiles from `{tigris}`.
4. Join the Census and shapefiles datasets with `left_join()` and visualise with `mapview()`