

# tigris: An R Package to Access and Work with Geographic Data from the US Census Bureau

by Kyle Walker

**Abstract** TIGER/Line shapefiles from the United States Census Bureau are commonly used for the mapping and analysis of US demographic trends. The **tigris** package provides a uniform interface for R users to download and work with these shapefiles. Functions in **tigris** allow R users to request Census geographic datasets using familiar geographic identifiers and return those datasets as objects of class "Spatial\*DataFrame". In turn, **tigris** ensures consistent and high-quality spatial data for R users' cartographic and spatial analysis projects that involve US Census data. This article provides an overview of the functionality of the **tigris** package, and concludes with an applied example of a geospatial workflow using data retrieved with **tigris**.

## Introduction

Analysis and visualization of geographic data are often core components of the analytical workflow for researchers and data scientists; as such, access to open and reliable geographic datasets are of paramount importance. The United States Census Bureau provides access to such data in the form of its TIGER/Line shapefile products (United States Census Bureau, 2016a). The files are extracts from the Census Bureau's Master Address File/Topologically Integrated Geographic Encoding and Referencing (TIGER) database, which in turn are released to the public as shapefiles, a common format for encoding geographic data as vectors (e.g. points, lines, and polygons). Available TIGER/Line shapefiles include all of the Census Bureau's areal enumeration units, such as states, counties, Census tracts, and Census blocks; transportation data such as roads and railways; and both linear and areal hydrography. The TIGER/Line files are updated annually, and include attributes that allow them to be joined with other tabular data, including demographic data products released by the Census Bureau.

The **tigris** package aims to simplify the process of working with these datasets for R users (Walker, 2016). With functions in **tigris**, R users can specify the data type and geography for which they would like to obtain geographic data, and return the corresponding TIGER/Line data as an R object of class "Spatial\*DataFrame". This article provides an overview of the **tigris** package, and gives examples that show how it can contribute to common geographic visualization and spatial analysis workflows in R. Examples in the article include a discussion of how **tigris** helps R users retrieve and work with data from the US Census Bureau, as well as an applied example of how **tigris** can fit within a common geospatial workflow in R, in which data from the United States Internal Revenue Service are visualized with both static and interactive cartography.

## Geographic data and Census visualization in R

The TIGER/Line files were first released by the US Census Bureau in ASCII format in 1989, and represented street centerline data for the entire United States. Since then, the Census Bureau has expanded the coverage of the TIGER/Line data, and transitioned the core format of the publicly-available files to the shapefile in 2007. TIGER/Line shapefiles include *boundary files*, which encompass the boundaries of governmental units or other areal units for which the Census Bureau tabulates data. This includes the core Census hierarchy of areal units from the Census block (analogous to a city block) to the entire United States, as well as common geographic entities such as city boundaries. The Census Bureau distinguishes between *legal entities*, which have official government standing, and *statistical entities*, which have no legal definition but are used for the tabulation of data. The Census Bureau also makes available shapefiles of *geographic features*, which include entities such as roads, rivers, and railroads. All TIGER/Line shapefiles are distributed in a geographic coordinate system using the North American Datum of 1983 (NAD83) (United States Census Bureau, 2014).

As the TIGER/Line datasets are available in shapefile format, they can be read into and translated to R objects by the **rgdal** package (Bivand et al., 2015). **rgdal** is an R interface to the open-sourced Geospatial Data Abstraction Library, or GDAL, an open-source translator that can convert between numerous common vector and raster spatial data formats (GDAL Development Team, 2015). When loaded into R, shapefiles will be represented as objects of class "Spatial\*" by the **sp** package (Bivand et al., 2013). Most Census datasets obtained by **tigris** will be loaded as objects of class

"SpatialPolygonsDataFrame" given that they represent Census areal entities; selected geographic features, such as roads, linear water features, and landmarks, may be represented as objects of class "SpatialLinesDataFrame" or "SpatialPointsDataFrame". "Spatial\*DataFrames" are R objects that represent spatial data as closely as possible to regular R data frames, yet also contain information about the feature geometry and coordinate system of the data (see [Bivand et al., 2013](#), for more information).

Several R packages provide access to Census geographic and demographic data. The **UScensus2000** (no longer on CRAN) and **UScensus2010** packages by Zack Almquist allow for access to several geographic datasets for the 2000 and 2010 Censuses, including blocks, Census tracts, and counties ([Almquist, 2010](#)). These datasets can also be linked to demographic data from the 2000 and 2010 Censuses, which are stored in related, external packages. The **USABoundaries** package similarly provides access to some Census geographic boundary files such as zip code tabulation areas (ZCTAs) and counties; it also makes available historical boundary files dating back to 1629 ([Mullen, 2015](#)). Another R package, **choroplethr**, wraps **ggplot2** ([Wickham, 2009](#)) to map data from the US Census Bureau's American Community Survey aggregated to common Census geographies ([Lamstein and Johnson, 2015](#)). The purpose of the **tigris** package is to help R users work with US Census Bureau geographic data by granting direct access to the Census shapefiles via a simple, uniform interface. Further, as **tigris** interfaces directly with Census Bureau data stores, it ensures access to high-quality and up-to-date geographic data for R projects.

## Core functionality of tigris

The core functionality of **tigris** consists of a series of functions, each corresponding to a single Census Bureau geography of interest, that grant access to geographic data from the US Census Bureau. **tigris** allows R users to obtain both the core TIGER/Line shapefiles as well as the Census Bureau's Cartographic Boundary Files. Cartographic Boundary Files, following the [United States Census Bureau \(2015\)](#), "are simplified representations of selected geographic areas from the Census Bureau's MAF/TIGER geographic database" ([United States Census Bureau, 2015](#)).

To download geographic data using **tigris**, the R user calls the function corresponding to the desired geography. For example, to obtain a "SpatialPolygonsDataFrame" of US states from the TIGER/Line dataset, the user calls the `states()` function in **tigris**, which can then be plotted with the `plot()` function from the **sp** package, which is loaded by **tigris** automatically:

```
library(tigris)
us_states <- states()
plot(us_states)
```



**Figure 1:** Basic plot of US states retrieved from the Census TIGER/Line database.

The `states()` function call instructs **tigris** to fetch a TIGER/Line shapefile from the US Census Bureau that represents the boundaries of the 50 US states, the District of Columbia, and US territories. **tigris** then uses **rgdal** to load the data into the user's R session as an object of class "Spatial\*DataFrame". Many functions in **tigris** have a parameter, `cb`, that if set to `TRUE` will direct **tigris** to load a cartographic boundary file instead. Cartographic boundary files default to a simplified resolution of 1:500,000; in some cases, as with states, resolutions of 1:5 million and 1:20 million are available. For example, an R user could specify the following modifications to the `states()` function, and retrieve a simplified dataset.

```
us_states_20m <- states(cb = TRUE, resolution = "20m")
ri <- us_states[us_states$NAME == "Rhode Island", ]
ri_20m <- us_states_20m[us_states_20m$NAME == "Rhode Island", ]
plot(ri)
plot(ri_20m, border = "red", add = TRUE)
```



**Figure 2:** Difference between default TIGER/Line and 1:20 million cartographic boundary outlines of Rhode Island.

The plot illustrates some of the differences between the TIGER/Line shapefiles and the cartographic boundary files, in this instance for the state of Rhode Island. The TIGER/Line shapefiles are the most detailed datasets in interior areas, and represent the legal boundaries for coastal areas which extend three miles beyond the shoreline. The Cartographic Boundary Files have less detail in interior areas, but are clipped to the shoreline of the United States, which may be preferable for thematic mapping but can introduce additional detail for coastal features. A full list of the geographic datasets available through **tigris** is found in Table 1; datasets with an asterisk are available as both TIGER/Line and cartographic boundary files.

When Census data are available for download at sub-national levels, they are referenced by their Federal Information Processing Standard (FIPS) codes, which are codes that uniquely identify geographic entities in the Census database. When applicable, **tigris** uses smart state and county lookup to simplify the process of data acquisition for R users. This allows users to obtain data by supplying the name or postal code of the desired state – along with the name of the desired county, when applicable – rather than their FIPS codes. In the following example, the R user fetches roads data for Kalawao County Hawaii, the smallest county in the United States by area, located on the northern coast of the island of Moloka'i.

```
kw_roads <- roads("HI", "Kalawao")
plot(kw_roads)
```

While many Census shapefiles correspond to these common geographic identifiers in the United States, not all datasets are identifiable in this way. A good example is the Zip Code Tabulation

Family	Functions
General area functions	nation*; divisions*; regions*; states*; counties*; tracts*; block_groups*; blocks; places*; pumas*; school_districts; county_subdivisions*; zctas*
Legislative district functions	congressional_districts*; state_legislative_districts*; voting_districts (2012 only)
Water functions	area_water; linear_water; coastline
Metro area functions	core_based_statistical_areas*; combined_statistical_areas*; metro_divisions; new_england*; urban_areas*
Transportation functions	primary_roads; primary_secondary_roads; roads; rails
Native/tribal geometries functions	native_areas*; alaska_native_regional_corporations*; tribal_block_groups; tribal_census_tracts; tribal_subdivisions_national
Other	landmarks; military

**Table 1:** Functions available in the tigris package. Functions denoted with an asterisk are also available as cartographic boundary files.

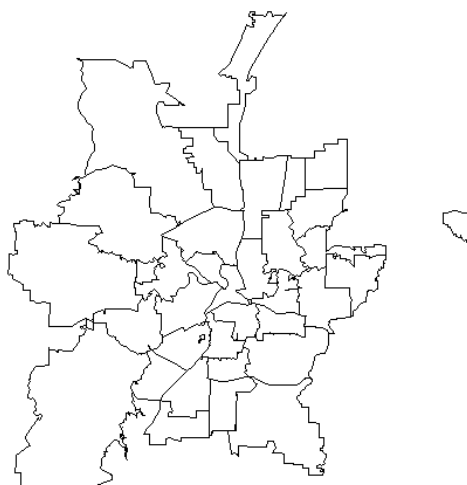


**Figure 3:** Roads in Kalawao County, Hawaii.

Area (ZCTA), a geographic dataset developed by the US Census Bureau to approximate zip codes, postal codes used by the United States Postal Service. Social data in the United States are commonly distributed at the zip code level, including an example later in this article; however, zip codes themselves are not coherent geographic entities, and change frequently. ZCTAs, then, function as proxies for zip codes, and are built from Census blocks in which a plurality of addresses on a given block have a given zip code (United States Census Bureau, 2016b).

ZCTAs commonly cross county lines and even cross state lines in certain instances; as such, the US Census Bureau only makes the entire ZCTA dataset of over 33,000 zip codes available for download. Often, analysts will not need all of these ZCTAs for a given project. **tigris** allows users to subset ZCTAs on load with the `starts_with` parameter, which accepts a vector of strings that contains the beginning digits of the ZCTAs that the analyst wants to load into R. The example below retrieves ZCTAs in the area around Fort Worth, Texas.

```
fw_zips <- zctas(cb = TRUE, starts_with = "761")
plot(fw_zips)
```



**Figure 4:** Zip Code Tabulation Areas that start with "761" (near Fort Worth, Texas).

When **tigris** downloads Census shapefiles to the R user's computer, it uses the **rappdirs** package to cache the downloads for future access (Ratnakumar et al., 2014). In turn, once the R user has downloaded the Census geographic data, **tigris** will know where to look for it and will not need to re-download. To turn off this behavior, a **tigris** user can set `options(tigris_use_cache = FALSE)` after loading the package; this will direct **tigris** to download shapefiles to a temporary directory on the user's computer instead, and load data into R from there.

The Census Bureau releases updated TIGER/Line shapefiles every year, and these yearly updates are available to **tigris** users. **tigris** defaults to the 2015 shapefiles, which at the time of this writing is the most recent year available. However, **tigris** users can supply a different year to a **tigris** function as a named argument to obtain data for a different year; for example, `year = 2014` in the function call will fetch TIGER/Line shapefiles or cartographic boundary files from 2014. Additionally, R users can set this as a global option in their R session by entering the command `options(tigris_year = 2014)`.

## Data manipulation with **tigris**

The primary utility of the **tigris** package is for consistent and quick data access for R users with a minimum of code. As **tigris** loads objects of class "Spatial\*" from the **sp** package, data analysis and visualization using the acquired data can be handled with R's suite of cartographic and spatial analysis packages, which will be addressed later in the article. However, **tigris** does include two

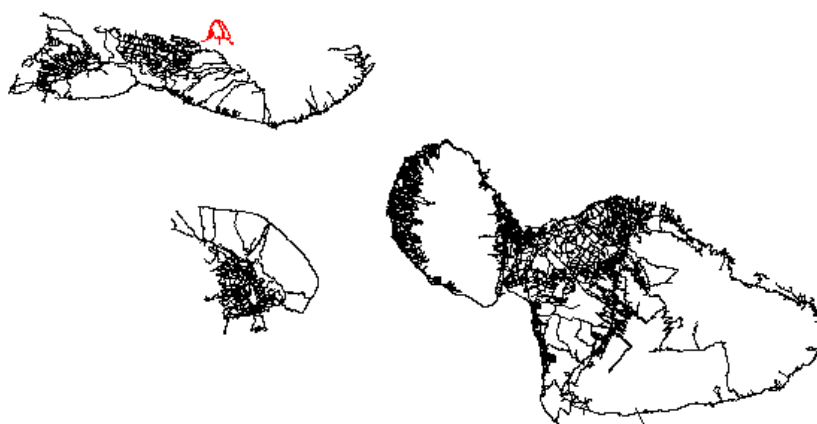
functions, `rbind_tigris()` and `geo_join()`, to assist with common operations when working with Census Bureau geographic data: combining datasets with one another, or merging them to tabular data.

Some Census shapefiles, like roads, are only available by county; however, an R user may want a roads dataset that represents multiple counties. **tigris** has built-in functionality to handle these circumstances. Data loaded into R by **tigris** functions are assigned a special "tigris" attribute that identifies the type of geographic data represented by the object. This attribute can be checked with the function `tigris_type()`:

```
> tigris_type(kw_roads)
[1] "road"
```

Objects with the same "tigris" attributes can then be combined into a single object using the function `rbind_tigris()`. In the example below, the user loads data for Maui County, which comprises the remainder of the island of Moloka'i as well as the islands of Maui and Lana'i. As the roads data do not explicitly contain information about counties, an identifying "county" column is specified in the example below. The data are then plotted and colored by county; Kalawao County is colored red in the figure.

```
maui_roads <- roads("HI", "Maui")
kw_roads$county <- "Kalawao"
maui_roads$county <- "Maui"
maui_kw_roads <- rbind_tigris(kw_roads, maui_roads)
plot(maui_kw_roads, col = c("red", "black")[as.factor(maui_kw_roads$county)])
```

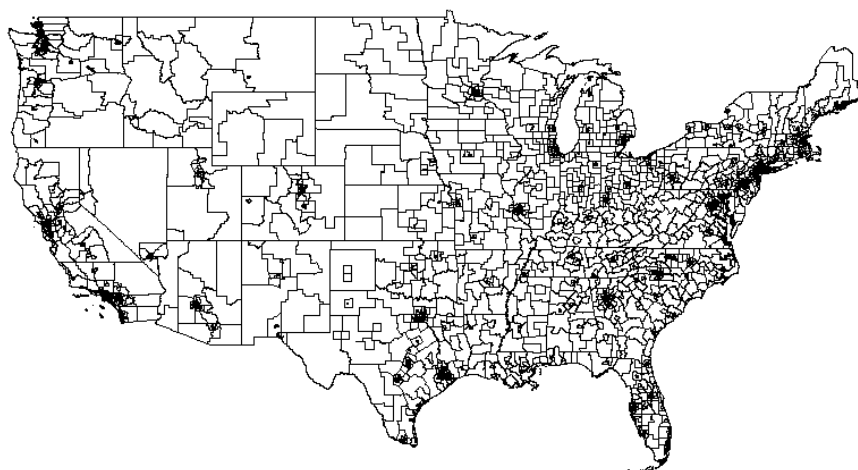


**Figure 5:** Roads in Maui County and Kalawao County, Hawaii.

The `rbind_tigris()` function also accepts a list of **sp** objects with the same "tigris" attributes. This is particularly useful in the event that an analyst needs a dataset that covers the United States, but is only available at sub-national levels. An example of this is the Public Use Microdata Area (PUMA), the Census geography at which individual-level microdata samples are associated. PUMAs are available by state in **tigris** via the `pumas()` function; however, an analyst will commonly want PUMA geography for the entire United States to facilitate national-level analyses. In this example, `rbind_tigris()` can be used with `lapply()` to fetch PUMA datasets for each state and then combine them into a dataset covering the continental United States.

**tigris** includes a built-in data frame named `fips_codes` that used to match state and county names with Census FIPS codes in the various functions in the package. It can also be used to generate vectors of codes to be passed to **tigris** functions as in this example, so that the analyst does not have to generate the full list of state codes by hand.

```
us_states <- unique(fips_codes$state)[1:51]
continental_states <- us_states[!us_states %in% c("AK", "HI")]
us_pumas <- rbind_tigris(
  lapply(
    continental_states, function(x) {
      pumas(state = x, cb = TRUE)
    }
  )
)
plot(us_pumas)
```



**Figure 6:** Public Use Microdata Areas (PUMAs) for the continental United States, generated with `rbind_tigris`.

The above code directs R to iterate through the state codes for the continental United States, fetching PUMA geography for each state and storing it in a list which `rbind_tigris()` then combines into a continental PUMA dataset.

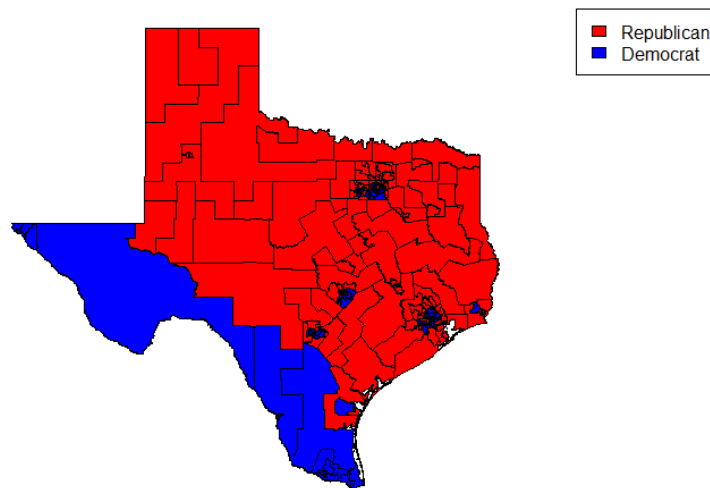
The other data manipulation function in **tigris**, `geo_join()`, is designed to assist with the common but sometimes-messy process of merging tabular data to US Census Bureau shapefiles. Such joined data can then be used for statistical mapping, such as a choropleth map that shows variation in an attribute by the shading of polygons.

In the example below, an analyst uses functions in **tigris** to help create a choropleth map that shows how the areas represented by legislators in the Texas State House of Representatives, the lower house of the Texas state legislature, vary by political party. By convention in the United States, areas represented by members of the Republican Party are shaded in red, and areas represented by members of the Democratic Party are shaded in blue.

To accomplish this, the analyst loads in a CSV containing information on party representation in Texas by legislative district, and uses the `state_legislative_districts()` function to retrieve boundaries for the legislative districts. The two datasets can then be joined with `geo_join()`. The first argument in the `geo_join()` call represents the object of class "Spatial\*DataFrame"; the second argument represents a regular R data frame. The third and fourth arguments specify the columns in the spatial data frame and regular data frame, respectively, to be used to match the rows; if the names of these columns are the same, that name can be passed as a named argument to the `by` parameter,

which is unused here. Once the two datasets are joined, they can be visualized with **sp** plotting functions.

```
df <- read.csv("http://personal.tcu.edu/kylewalker/data/txlege.csv",
  stringsAsFactors = FALSE)
districts <- state_legislative_districts("TX", house = "lower", cb = TRUE)
txlege <- geo_join(districts, df, "NAME", "District")
txlege$color <- ifelse(txlege$Party == "R", "red", "blue")
plot(txlege, col = txlege$color)
legend("topright", legend = c("Republican", "Democrat"),
  fill = c("red", "blue"))
```



**Figure 7:** State legislative districts for the Texas State House of Representatives, colored by the party affiliations of their representatives. Data derived from *The Texas Tribune*, <https://www.texastribune.org/directory/>

As the plot illustrates, Democrats in Texas tend to represent areas in and around major cities such as Dallas, Houston, and Austin, as well as areas along the United States-Mexico border. Republicans, on the other hand, tend to represent rural areas in addition to suburban areas on the edges of metropolitan areas in the state.

## Analytic visualization in R using data obtained with **tigris**

To this point, this paper has employed simplified examples to demonstrate the functionality of **tigris**; the following scenario combines these examples into an applied analytic and visualization workflow. The goal here is to show how **tigris** fits in with a broader spatial analysis workflow in R. R has a plethora of packages available for geographic visualization and spatial analysis; for analysts working with United States geographies, **tigris** can contribute to the analytic process by providing ample data access with a minimum of code, and without having to retrieve datasets outside of R.

This example demonstrates how to create metropolitan area maps of taxation data from the United States Internal Revenue Service (IRS), which are made available at the zip code level ([United States Internal Revenue Service, 2015](#)). As discussed earlier, zip codes are not physical areas but rather designations given by the United States Postal Service (USPS) to guide mail routes; as such, ZCTAs will be used instead, and accessed with the `zctas` function.

As ZCTAs do not have a clear correspondence between their boundaries and those of other Census units, ZCTA boundaries will commonly cross those of metropolitan areas, which are county-based. However, **tigris** provides programmatic access to metropolitan area boundaries as well, which in turn



can be used to identify intersecting ZCTAs through spatial overlay with the **sp** package. The resultant spatial data can then be merged to data from the IRS and visualized.

Such a workflow could resemble the following. An analyst reads in raw data from the IRS website as an R data frame, and uses the **dplyr** package (Wickham and Francois, 2015) to subset the data frame and identify the average total income reported to the IRS by zip code in thousands of dollars in 2013, assigning it to the variable `df`. In the original IRS dataset, `A02650` represents the aggregate total income reported to the IRS by zip code in thousands of dollars, and `N02650` represents the number of tax returns that reported total income in that zip code.

```
library(dplyr)
library(stringr)
library(readr)

# Read in the IRS data
zip_data <- "https://www.irs.gov/pub/irs-soi/13zpallnoagi.csv"
df <- read_csv(zip_data) %>%
  mutate(zip_str = str_pad(as.character(ZIPCODE), width = 5,
                           side = "left", pad = "0"),
         incpr = A02650 / N02650) %>%
  select(zip_str, incpr)
```

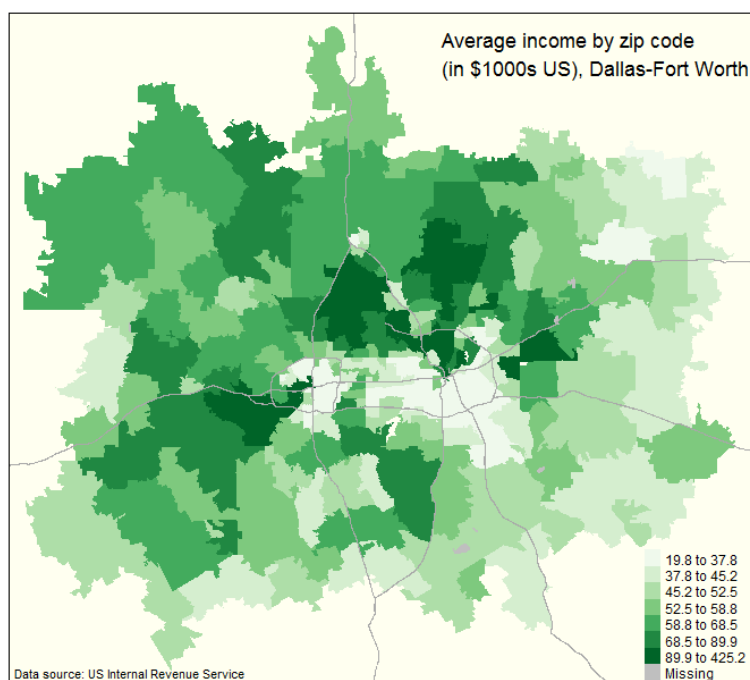
The analyst then defines a function that will leverage **tigris** to read in Census ZCTA and metropolitan area datasets as objects of class "SpatialPolygonsDataFrame", and return the ZCTAs that intersect a given metropolitan area as defined by the analyst.

```
library(tigris)
library(sp)

# Write function to get ZCTAs for a given metro
get_zips <- function(metro_name) {
  zips <- zctas(cb = TRUE)
  metros <- core_based_statistical_areas(cb = TRUE)
  # Subset for specific metro area
  # (be careful with duplicate cities like "Washington")
  my_metro <- metros[grepl(sprintf("^%s", metro_name),
                           metros$NAME, ignore.case = TRUE), ]
  # Find all ZCTAs that intersect the metro boundary
  metro_zips <- over(my_metro, zips, returnList = TRUE)[[1]]
  my_zips <- zips[zips$ZCTA5CE10 %in% metro_zips$ZCTA5CE10, ]
  # Return those ZCTAs
  return(my_zips)
}
```

The analyst can then fetch ZCTA geography for a given metropolitan area, which in this example will be Dallas-Fort Worth, Texas, and merge the IRS income data to it with `geo_join()`. For visualization, this example uses the **tmap** package (Tennekes, 2015), an excellent option for creating high-quality cartographic products in R, to create a choropleth map. To provide spatial reference to the Census tracts on the map, major roads obtained with the `primary_roads()` function in **tigris** are added to the map as well.

```
library(tmap)
rds <- primary_roads()
dfw <- get_zips("Dallas")
dfw_merged <- geo_join(dfw, df, "ZCTA5CE10", "zip_str")
tm_shape(dfw_merged, projection = "+init=epsg:26914") +
  tm_fill("incpr", style = "quantile", n = 7, palette = "Greens", title = "") +
  tm_shape(rds, projection = "+init=epsg:26914") +
  tm_lines(col = "darkgrey") +
  tm_layout(bg.color = "ivory",
            title = "Average income by zip code \n(in $1000s US), Dallas-Fort Worth",
            title.position = c("right", "top"), title.size = 1.1,
            legend.position = c(0.85, 0), legend.text.size = 0.75,
            legend.width = 0.2) +
  tm_credits("Data source: US Internal Revenue Service",
            position = c(0.002, 0.002))
```



**Figure 8:** Map of average reported total income to the Internal Revenue Service by zip code for the Dallas-Fort Worth, Texas metropolitan area in 2013, created with tmap.

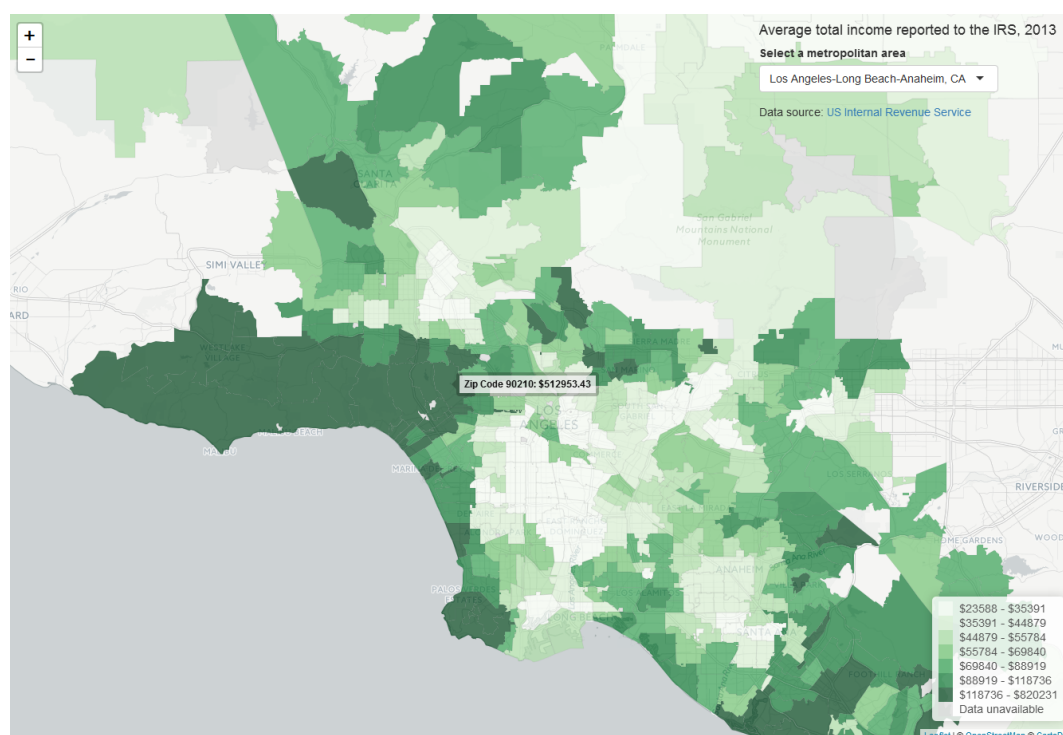
Given that both geographic data obtained through **tigris** and the IRS data are available for the entire country, an R developer could extend this example and create a web application that generates interactive income maps based on user input with the **shiny** package (Chang et al., 2016). Below is an example of such an application, which is viewable online at <http://walkerke.shinyapps.io/tigris-zip-income>; the code for the application can be viewed at <http://github.com/walkerke/tigris-zip-income>.

In the application, the user selects a metropolitan area from the drop-down menu, instructing the Shiny server to generate an interactive choropleth map of average reported total income by zip code from the IRS, as in the above example, but in this instance using the **leaflet** package (Cheng and Xie, 2015). The application uses the same process described above for the static map to subset the data; in this instance, however, the Shiny server makes these computations on-the-fly in response to user input. In the figure, the Los Angeles, California metropolitan area is selected; however, all metropolitan areas in the United States are available to users of the application. While both of these cartographic examples require considerable R infrastructure to process the data and ultimately create the visualizations, **tigris** plays a key role in each by providing direct access to reliable spatial data programmatically from R.

## Conclusion

This paper has summarized the functionality of the **tigris** package for retrieving and working with shapefiles from the United States Census Bureau. Access to high-quality spatial data is essential for the geospatial analyst, but can be difficult to access. For R users working on projects that can benefit from United States Census Bureau data, **tigris** provides direct access to the Census Bureau's TIGER/Line and cartographic boundary files using a simple and consistent API. In turn, tasks such as looking up FIPS codes to identify the correct datasets to download or combining several Census datasets are reduced to a few lines of R code in **tigris**.

More significantly, the utility of **tigris** is exemplified when included in a larger geospatial project that incorporates US Census data, such as the static and interactive maps of IRS data included in this article. These examples illustrate some clear advantages that R has over traditional desktop Geographic Information Systems software for geographic analysis and visualization. To create an interactive application showing IRS data by zip code as in Figure 9, a GIS analyst would traditionally have to search out and download the data from the web; load it into a desktop GIS for merging and calculating new columns; publish the data to a server; and build the application with a web mapping client and web application framework, often requiring several software applications. As shown in



**Figure 9:** Interactive Leaflet map of average reported total income to the Internal Revenue Service by zip code for the Los Angeles, California metropolitan area in 2013, built in Shiny.

this article, this entire process now can take place inside of an R script, helping ensure quality and reproducibility. For projects that require US Census Bureau geographic data, **tigris** aims to fit well within these types of workflows.

## Acknowledgments

I am indebted to Bob Rudis, whose key contributions to the **tigris** package were essential in improving its functionality. I would also like to thank Eli Knaap, who encouraged me to write the package; Hadley Wickham for writing *R Packages* and the **devtools** package (Wickham, 2015); (Wickham and Chang, 2015), which helped me significantly as I developed **tigris**; and Roger Bivand as well as the three anonymous reviewers for providing useful feedback on the manuscript.

## Bibliography

- Z. Almquist. US Census spatial and demographic data in R: The UScensus2000 suite of packages. *Journal of Statistical Software*, 37(6):1–31, 2010. [p232]
- R. Bivand, T. Keitt, and B. Rowlingson. *rgdal: Bindings for the Geospatial Data Abstraction Library*, 2015. URL <https://CRAN.R-project.org/package=rgdal>. R package version 1.1-1. [p231]
- R. S. Bivand, E. Pebesma, and V. Gomez-Rubio. *Applied spatial data analysis with R, Second edition*. Springer, New York, 2013. URL <http://www.asdar-book.org/>. [p231, 232]
- W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *shiny: Web Application Framework for R*, 2016. URL <https://CRAN.R-project.org/package=shiny>. R package version 0.13.2. [p240]
- J. Cheng and Y. Xie. *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library*, 2015. URL <http://rstudio.github.io/leaflet/>. R package version 1.0.1.9002. [p240]
- GDAL Development Team. *GDAL - Geospatial Data Abstraction Library, Version 2.0.1*. Open Source Geospatial Foundation, 2015. URL <http://www.gdal.org>. [p231]
- A. Lamstein and B. P. Johnson. *choroplethr: Simplify the Creation of Choropleth Maps in R*, 2015. URL <https://CRAN.R-project.org/package=choroplethr>. R package version 3.3.0. [p232]

- L. Mullen. *USAboundaries: Historical and Contemporary Boundaries of the United States of America*, 2015. URL <https://github.com/ropensci/USAboundaries>. R package version 0.1.1.9001. [p232]
- S. Ratnakumar, T. Mick, and T. Davis. *rappdirs: Application directories: determine where to save data, caches and logs.*, 2014. URL <https://CRAN.R-project.org/package=rappdirs>. R package version 0.3. [p235]
- M. Tennekes. *tmap: Thematic Maps*, 2015. URL <https://CRAN.R-project.org/package=tmap>. R package version 1.0. [p239]
- United States Census Bureau. TIGER/Line shapefiles: Technical documentation. Report, 2014. [p231]
- United States Census Bureau. Cartographic boundary file description, 2015. URL [https://www.census.gov/geo/maps-data/data/cbf/cbf\\_description.html](https://www.census.gov/geo/maps-data/data/cbf/cbf_description.html). Accessed: 2016-02-27. [p232]
- United States Census Bureau. Tiger/line® shapefiles and tiger/line® files, 2016a. URL <https://www.census.gov/geo/maps-data/data/tiger-line.html>. Accessed: 2016-05-22. [p231]
- United States Census Bureau. Zip code™ tabulation areas (zctas™), 2016b. URL <https://www.census.gov/geo/reference/zctas.html>. Accessed: 2016-05-22. [p235]
- United States Internal Revenue Service. Soi tax stats - individual income tax statistics - 2013 zip code data (soi), 2015. URL <https://www.irs.gov/uac/soi-tax-stats-individual-income-tax-statistics-2013-zip-code-data-soi>. Accessed: 2016-05-19. [p238]
- K. Walker. *tigris: Load Census TIGER/Line Shapefiles into R*, 2016. R package version 0.3. [p231]
- H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL <http://had.co.nz/ggplot2/book>. [p232]
- H. Wickham. *R Packages*. O'Reilly Media, Sebastopol, CA, 2015. ISBN 978-1491910597. URL <http://r-pkgs.had.co.nz/>. [p241]
- H. Wickham and W. Chang. *devtools: Tools to Make Developing R Packages Easier*, 2015. URL <https://CRAN.R-project.org/package=devtools>. R package version 1.8.0. [p241]
- H. Wickham and R. Francois. *dplyr: A Grammar of Data Manipulation*, 2015. URL <https://CRAN.R-project.org/package=dplyr>. R package version 0.4.3. [p239]

Kyle Walker  
Texas Christian University  
2850 S University Dr  
Fort Worth, TX 76109  
[kyle.walker@tcu.edu](mailto:kyle.walker@tcu.edu)