

Mapping data

Andrew Ba Tran

Chapter 5

Spatial analysis

Consider where you are sitting. Think of the businesses you're next to, the school district border you fall in, the socioeconomic makeup of your neighbors. Think about how that might change for the next person in the list for this class. Does it snow more often there or do they live in a more moderate climate? How far away from a grocery store are they compared to you? Is your quality of life better or worse? What opportunities are you missing out on that others might enjoy have based on where you are located?

When you're able to analyze data in the framework of geography, you might be able to see patterns that you wouldn't see otherwise.

There's something tangible about visualizing data with maps. We recognize where we fit in the borders. It tells us who we are based on where we've been. And there's more location data than ever before thanks to GPS-enable devices like phones, cameras, and vehicles.

Most spatial analysis is conducted with graphical programs like the (very expensive) ArcGIS and (free) QGIS. Recently, GIS work can be done online with Google Fusion Tables and Carto.

There's a steep learning curve in using R to programmatically analyze and visualize spatial data. But the perks include approaching your work in a customizable, transparent, and reproducible way.

Some recent work done with spatial analysis include:

- The Geographic Divide of Oscar Films - The Pudding
- BuzzFeed 311 calls increase in gentrifying neighborhoods - BuzzFeed
- Murder with Impunity - Washington Post
- Spies in the Skies - BuzzFeed

In this chapter, we'll learn how to create static maps and interactive maps with geolocated data. We'll be using the more-recent packages in the R GIS community: **simple features** and **leaflet**.

We'll access the Census API to download data and use what we've already learned to join that data to shapefiles to make choropleth maps— both interactive and static.

We'll do some digging into traffic stop data to find which neighborhoods are most racially profiled by police.

Goals

- Static maps
- Mapping with `sf` and `ggplot2`
- Geolocating data
- Getting latitude and longitude from addresses
- Traffic stops case study
- Looking for racial profiling in traffic stops
- Interactive maps with Leaflet
- Locator maps and geolocating addresses
- Interactive choropleth maps

- With Starbucks and Census data

Files and folders

The repo for this class is on GitHub, but can be easily downloaded to your desktop with the following commands:

```
install.packages("usethis")
```

```
usethis::use_course("https://github.com/r-journalism/learn-chapter-5/archive/master.zip")
```

Readings

- Behind the dialect map interactive: How an intern created the New York Times' most popular piece of content in 2013 - knight lab
 - Visual history of the U.S. Census - City Lab
 - BuzzFeed 311 calls increase in gentrifying neighborhoods - BuzzFeed | GH
 - Murder with Impunity - Washington Post
 - Spies in the Skies - BuzzFeed | GH
 - America is more diverse than ever but still segregated - Washington Post
 - Geographic divide of Oscar movies - The Pudding | GH
-

Test yourself

There are links to exercise what you've learned spread through out this section.

It's possible to run these files locally to test yourself if you've downloaded the files for the chapter as instructed above.

Make sure your project directory is correct and then run these lines in the console:

```
install.packages("learnr")  
install.packages("rmarkdown")  
install.packages("tidyverse")  
install.packages("sf")  
install.packages("leaflet")
```

and then

```
rmarkdown::run("chapter-5/index.Rmd")
```