Data Curation Guidebook - The Social Life of Neighborhoods

Author: Francine Stephens

Contents

Preface

Welcome to The Social Life of Neighborhoods! As the TA, you will be guiding students through the process of using and synthesizing spatial and quantitative data to create maps for their neighborhood assignments. The TA of this course also plays the important role of the *data curator*. When the TA prepares and provides the necessary data for the neighborhood assignments, the students are able to focus more on analyzing the data and drawing connections to the course readings to make sense of the patterns emerging in and around their neighborhoods. This makes for a smoother class experience for the students, TA, and professor!

The responsibility of curating data for the class over multiple weeks of the quarter may seem a little daunting, but the good news is that you do not have to come up with a workflow from scratch. This guidebook along with the R scripts stored in the linked GitHub folder will walk you through the steps to automate a substantial amount of the work.[1] The guidebook is organized in a series of chapters based on the course assignment and the data the students need from you to complete their assignment in ArcGIS Online. A brief description of each chapter is provided below:

1. **Chapter 1** covers Neighborhood Assignment #1 and the crucial steps to complete after students have submitted this assignment and before they begin their first mapping assignment.
2. **Chapter 2** covers how to create the ArcGIS Online class page and students' neighborhood boundaries from the census tracts they submitted in Assignment #1 for the Neighborhood Organizations assignment.
3. **Chapter 3** covers the Theories of Crime/Social Disorder Assignment, which involves collecting and preparing the crime data.
4. **Chapter 4** covers the Segregation assignment and how to prepare data for the cities and metro areas/regions that contain the neighborhoods chosen by students.

[1] Some weeks, e.g., Crime Data, will require a little more manual data work on your part because the data are not coming from a single provider like the U.S. Census, and have different formats and attributes that you will have to standardize.

5. **Chapter 5** covers the Gentrification assignment and how to prepare data for the city that contains the neighborhoods chosen by students.
6. **Chapter 6** covers the Globalization assignment and how to prepare the immigration data for the students' neighborhoods and the city that contains their chosen neighborhood.[2]

Each chapter will explain the key materials and steps that you need to know to curate the data, including the **data sources** needed for the assignment; a description of the **key components of the R script**; the specific **parameters that need to updated in the script** to run for your class; and the **data/shapefiles that the script will output** for you to upload to the ArcGIS Online class page for the students to use.

Before You Start

Before diving into Chapter 1, download the code and boilerplate files from Francine's public GitHub repository to your computer. The data and scripts within the GitHub repository folder will be referenced in the following chapters. *You do not need to set up a GitHub account/page* to run any of the scripts or save the exported data.

[2]This assignment was scratched for the 2021 Spring Quarter, so the data collection and preparation steps did not get a "real" test run, but the script and steps are still included in this guidebook to serve as boilerplate code and documentation for future TAs to build upon as needed.

Chapter 1

Neighborhood-Census Geographies Crosswalk

This chapter is fundamental to all of the data that you will curate during the quarter. It explains how to create the neighborhood-census geographies crosswalk, which is a master spreadsheet that links each student's neighborhood to census geographies that will be used later in the course, such as census tracts, cities, and metropolitan statistical areas. This crosswalk file is for you and your future work, not the students, so you do not need to worry about sharing the crosswalk with students in the class.

When should you curate this data? After students have submitted assignment 1 - Neighborhood selection - and before Assignemtn #3 - the first mapping assignment in ArcGIS online.

What do you need to get started?

How to Prepare the Crosswalk

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
```

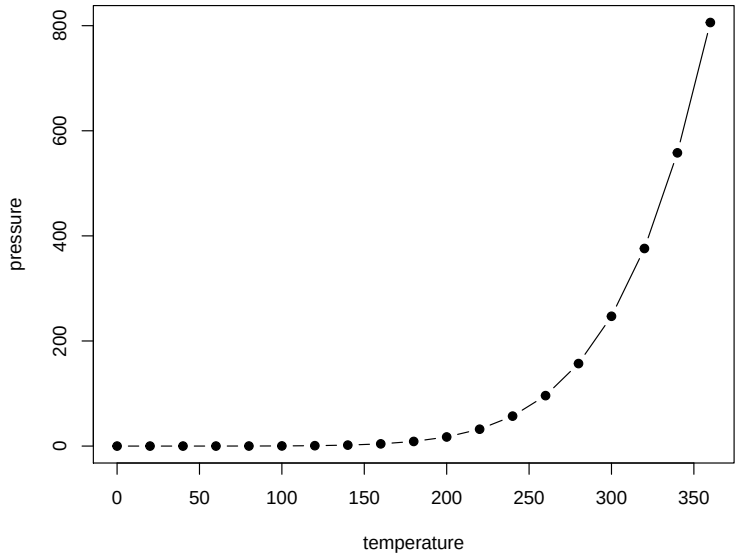

Figure 1.1: Here is a nice figure!

```
  booktabs = TRUE
)
```

Table 1.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 2

Neighborhood Boundaries

Assignment 2 or 3 (CHECK). Create neighborhood boundaries and export. Upload to ArcGIS Online.

Chapter 3

Crime Data

How to obtain and prepare crime data.

Chapter 4

Segregation Data

Some *significant* applications are demonstrated in this chapter.

4.1 Example one

4.2 Example two

Chapter 5

Gentrification Data

We have finished a nice book.

Chapter 6

Immigration Data