GIS in R: files that make up a shapefile

Learning Objectives

After completing this tutorial, you will be able to:

- Be able to list the 3 core files associated with a shapefile.

What you need

You will need a computer with internet access to complete this lesson.

If you have not already downloaded the week 4 data, please do so now. Download Week 3 Data (~250 MB){:data-proofer-ignore=" .btn }

One Dataset - Many Files

While text files often are self contained (one CSV) is composed of one unique file, many spatial formats are composed of several files. A shapefile is created by 3 or more files, all of which must retain the same NAME and be stored in the same file directory, in order for you to be able to work with them.

Shapefile Structure

There are 3 key files associated with any and all shapefiles:

- **`.shp`:** the file that contains the geometry for all features.
- **`.shx`:** the file that indexes the geometry.
- **`.dbf`:** the file that stores feature attributes in a tabular format.

These files need to have the **same name** and to be stored in the same directory (folder) to open properly in a GIS, `R` or `Python` tool.

Sometimes, a shapefile will have other associated files including:

- `.prj`: the file that contains information on projection format including the coordinate system and projection information. It is a plain text file describing the projection using well-known text (WKT) format.
- `.sbn` and `.sbx`: the files that are a spatial index of the features.
- `.shp.xml`: the file that is the geospatial metadata in XML format, (e.g. ISO 19115 or XML format).

Data Management - Sharing Shapefiles

When you work with a shapefile, you must keep all of the key associated file types together. And when you share a shapefile with a colleague, it is important to zip up all of these files into one package before you send it to them!

Additional resources:

- Intro to Working With Vector Data in R - Data Carpentry / NEON series.