

SOC 4650/5650: Lab-06 - Weather Events in the St. Louis Area

Christopher Prener, Ph.D.

March 5th, 2018

Directions

Using data from the MetroWx directory in CourseData, you will create a geodatabase and make some edits to a shapefile. Your entire project folder system, including data and notebook output, should be uploaded to GitHub by Monday, March 12th at 5:00pm.

Part 1: Analysis Development (Review from Lectures 01 and 02)

The goal of this section is to create a self contained project directory with all of the data, code, map documents, results, and documentation a project needs.

1. In your course folder system, find the Labs/Lab-05 subdirectory and add all necessary subfolders.
2. Open ArcCatalog and add all relevant data to the data/ folder within your assignment. Save these data in a subfolder called rawData/ Remember to remove any unneeded data before making a commit.
3. Open RStudio and add an R project, notebook, and the README file to your assignment. Make sure these are stored in the appropriate places in your assignment.
4. The README should contain addition details on both the source of your data and how you have modified different elements. Update this as you progress through the assignment.¹

¹ Think of this as writing a letter to your future self - what information would you need to create this process?

Part 2: Create a Geodatabase

The goal of this section is to be able to create a geodatabase of severe weather data using ArcCatalog.

5. In ArcCatalog, create a new geodatabase in the data/ folder of your assignment.
6. Import the four of the five shapefiles from the rawData/ subfolder into your new geodatabase. Do not import the METRO_WX_Hail.shp.

7. Create a shapefile from the new Tornados feature class in your new geodatabase. Save it to a new subfolder within data/ called cleanedData/ as METRO_WX_TornadosClean.shp.
8. Update your README.md file with notes about the changes you have made.

Part 3: Clean up the Hail Data

The goal of this section is the practice working with dates and strings in R.

9. In your R notebook, import the hail data from data/rawData/.
10. Remove the following variables from your shapefile: yr, mo, and dy.
11. In a pipeline, complete the following steps:
 - (a) Convert the date variable from its current string format to a date formatted variable, then
 - (b) Use the new date-formatted variable to create a variable called year with only the appropriate data, then
 - (c) Use the new date-formatted variable to create a variable called month with only the appropriate data, then
 - (d) Use the new date-formatted variable to create a variable called day with only the appropriate data.
12. Edit the variable st so that, if an observation is M0 it now is "State of Missouri;". All other observations (those representing Illinois) should be converted to "State of Illinois;".
13. In a pipeline, complete the following steps:
 - (a) Remove the semi-colon from the variable st,² then
 - (b) If the st variable contains the word "Missouri", change the text to "M0", otherwise change "Illinois" to "IL",³ then
 - (c) Change the st variable so that it is in all uppercase, then
 - (d) Create a state abbreviation variable that is "M0" if the third word of st is "M0", otherwise set the state abbreviation to "IL" if the third word of st is "IL".⁴ For the purposes of this assignment not assume that all other observations are for Illinois - be explicit in your code! Then
 - (e) Use the str_detect() function to create a binary variable that is TRUE if the state is "IL" and is FALSE otherwise, then

² Hint: Use str_replace() to do this!

³ Hint: Use str_replace() to do this!

⁴ Hint: Use word() to help do this!

- (f) Use the `str_detect()` function to create a binary variable that is TRUE if the state is "MO" and is FALSE otherwise.
- 14. Finally, subset your data so that hail events before 2000 are in a separate data frame from hail events that occurred in 2000 or more recently.
- 15. Save each of these subset data frames to shapefiles in `data/cleanedData/`.
- 16. Update your `README.md` file with notes about the changes you have made.

Part 4: Add New Data to Your Geodatabase

The goal of this section is bring your data full circle. We've made changes in ArcCatalog, imported some data into R, and then cleaned it and exported it. Now we want to again use ArcCatalog to bring these data into your geodatabase.

- 17. Import both of your subset hail shapefiles from `data/cleanedData/` to your geodatabase that you created in Part 2.
- 18. Update your `README.md` file with notes about the changes you have made.