

STA404/504 HW9

Pei Wang

08/22/2022

Learning Objectives:

- Practice creating choropleth maps.
- Handle different data resources when creating maps.
- Combine dataset with join statement.
- Choose and create an effective visual for storytelling.

Assignment Description:

Please complete the following two tasks and submit:

- (1) A .png files saved to computer using “**ggsave()**” containing the plot created from question 1.
- (2) An. R file containing the starting code as well as the code you write to build the plots in Question 1. Please include necessary comments so that it’s reader friendly.
- (3) A .gif file by running the R object “my.animation” in “Animation example” R code, as asked in Question 2.

Please keep in mind the design principles we discussed in class, include necessary labels, titles, etc.

You may start your work on Question 1 using the R file “hw9start.R”. You may directly write your code on that file. After the choropleth plot is created, use **ggsave()** to save the image as a good resolution and sized .png file. (Example of **ggsave()** is provided in week 9, day 18, the French fries example R code) That is, you determine a resolution and DPI that you feel the plot looks good. It will take some trial and error to determine the dpi, width and height to get a nice cleaned .png file.

1.The USDA Food Environment Atlas is a governmental database on local food availability. See <https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads/>. One particular component is the presence of local agriculture. Farms that have direct to consumer sales are potentially of interest because they provide a source for purchasing food that was produced locally.

Please use the current version data (which is the excel data with name “Food Environment Atlas Data Download” from the website with a last uploaded date 9/10/2020) to recreate a choropleth map of county level counts of farms with direct sales. Make sure you fix the missing value problem for at least one county.

Below are some explanations and suggestions I have for creating this map:

- (1) You may check the spelling of the state and county names in the three provided datasets to figure out how to combine them together.
- (2) The fill mapping on the map can be changed to use a $\log_{10}()$ transformation on the counts.
- (3) The color of the numbers should be intuitive (darker color represent higher values).
- (4) The default map created may contain a lot of missing values. **You are required to fix the missing value problem for at least one county.** These missing values usually due to county names don’t match with each other. You may also fix missing values for more counties, but that is not required.

2.The animated chart is an important type of interactive displays. It displays several chart states one after the other. In this question, we will run a code that creates animation plot. First, please watch this video: <https://www.youtube.com/watch?v=...>

[//www.youtube.com/watch?v=Z8t4k0Q8e8Y&feature=youtu.be](https://www.youtube.com/watch?v=Z8t4k0Q8e8Y&feature=youtu.be) Then run the code: “Animation example” to create a similar animation plot in R. Then you can right click on the gif generated by running the “my.animation” R object, and save the gif to your computer. There are a few packages to be installed, in order for the gif to be properly created and displayed. The packages include but not limited to “**gganimate**”, “**gifski**”, “**av**”, etc.

Grading Rubric:

Question 1 (70 points): Data handling properly merging the datasets. The map is correct with appropriate aesthetic and design settings, appropriate scale (log-scale is recommended). The image created by **ggsave()** has reasonable options to save a high quality image. Code provided runs successfully without edit and will generate the same image as provided.

Question 2 (30 points): Run the provided code successfully to generate the .gif file required.