STA 404/504 Homework 5

Due Tuesday April 7th, by 11:59pm, submit through canvas

**Learning Objectives:**

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

**Assignment Description:**

Please complete the following two tasks and submit:

1. Two .png files containing the plots created from part 1 and 2 respectively.
2. A .R file containing the stating code as well as your code to build the plots in parts 1 and 2. Please include necessary question label so that it’s reader friendly.
3. 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 “Data Download” from the website with a last uploaded date 3/27/2018) to recreate a choropleth map of county level counts of farms with direct sales (shown below).

The code to start your work is provided in the R file “hw5start.R”. You can work on this question based on that code. After the plot is created, use ggsave() to save the image as a good resolution and sized png file. That is, you determine a resolution and DPI that you feel the plot looks good and allows for some zooming of the image.

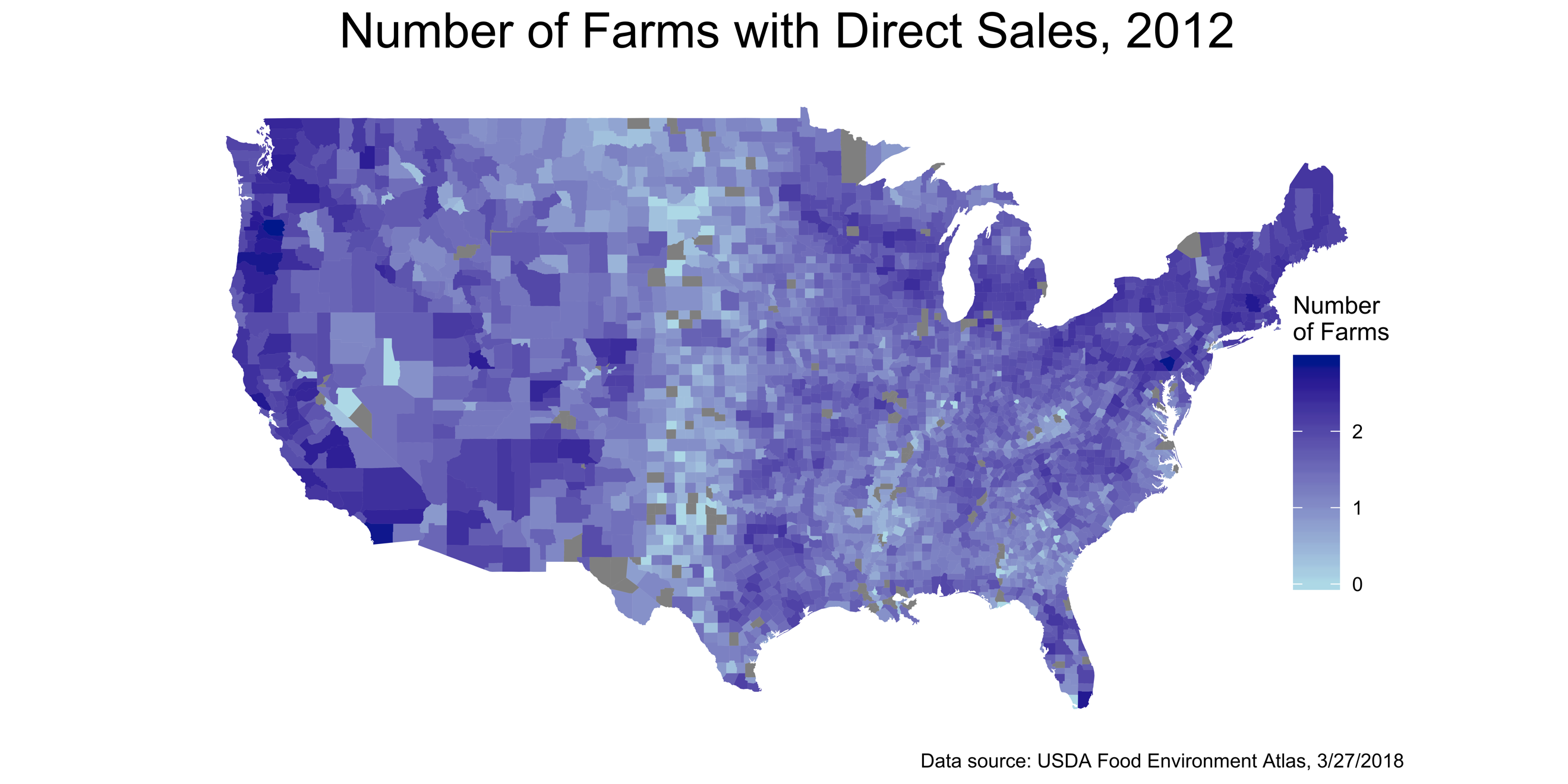
Below are some explanations and suggestions I have for creating this map:

(1) You like need to join the full names of the states onto the food data before it can then be merged onto the county outlines

(2) The fill mapping on the map is done using a log10() transformation on the counts

(3) You do not need to match the colors of the below image exactly, but something similar should be used

(4) It will take some trial and error to determine the dpi, width and height to get a nice cleaned png file.



1. COVID-19 data

The following github page contains the COVID-19 dailty reports and time series data.

<https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data>

Let’s use the daily reports data and create a choropleth map. In the daily reports data, the number of confirmed cases up to a certain date is summarized in a variable “Confirmed”. The dataset is recorded at the county level. Please use a daily report data no earlier than 3/29/2020 to create a choropleth map of the number of confirmed cases at the state level for the United States.

The way to read in the csv file from the website and other necessary code are available in the R file “hw5start.R”. Please include necessary legends, titles, colors, and keep in mind the design principals we discussed when creating the plot. After you finish it, use ggsave() to save the image as a good resolution and sized png file.

**Grading Rubric:**

Task 1 (15 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.

Task 2 (15 points): Data handling properly merging the datasets. The map is correct with appropriate aesthetic and design settings, appropriate scale, 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.