R Module 3 Rubric

Questions

1. Within your `data/` folder, open the file `NC_Counties.prj` with a program like Notepad or Notepad++.
   - What kind of information does this file contain?
   - Why is it important to include these "auxiliary" files?
   - What would happen if you forgot to include the `.prj` file?

This file contains projection metadata for the shapefile, such as datum, GCS, and units. These files are important to include as they provide the necessary information to display and project the data. Without this file, the system wouldn't know the projection system, so you wouldn't be able to do spatial analysis.

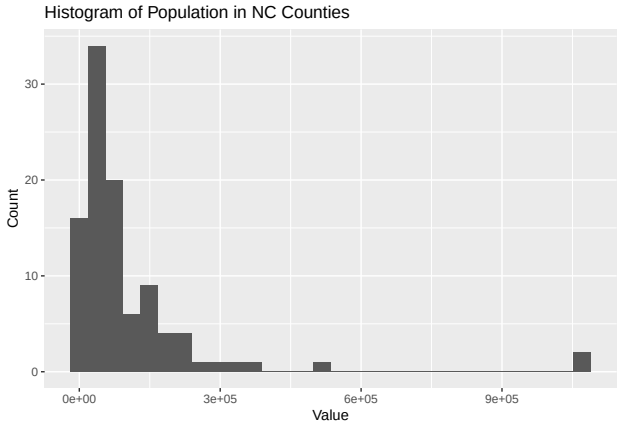
2. For the `US_States` layer, the `fill` argument stands on its own, while for `NC_Counties`, it's inside the `aes()` function. Why is this the case – what's the difference between these two layers?

In the case of `US_States`, we want to assign a single "background color", so we set the `fill` argument directly. However, with `NC_Counties`, we wish to "map" the values in `population` to the `fill` aesthetic, so we need to include it within the `aes()` call.

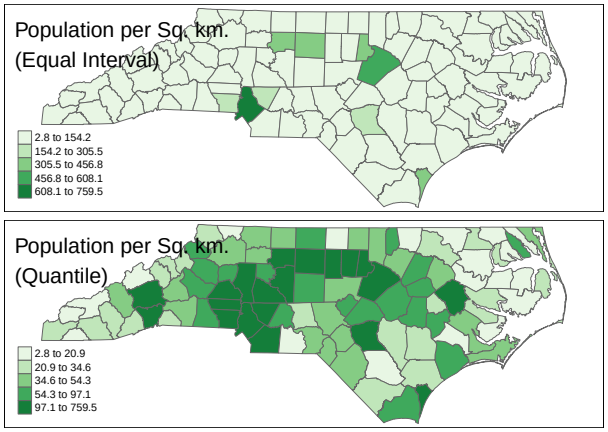
3. What's the purpose of the `coord_sf()` function? Use ?coord_sf() to view documentation and usage, and describe its arguments.

`coord_sf()` allows us to set parameters when visualizing sf objects. Some arguments include: `xmin` and `ymin`, which are limits on the "bounding box" of our map; `expand`, which ensures the data and axes don't overlap; `crs`, which sets the coordinate reference system, etc. ... This question is meant to get students to explore the documentation, rather than just rehash what's included in the module itself.

4. Using `ggplot2`, generate a histogram of the distribution of county population. Include axis labels, a title, and your name.



5. What type (continuous or discrete) does this distribution follow? Is the data normally distributed? If not, what kind of transformation can help "normalize" our data?
6. With a shapefile of your own (using ArcGIS, the `tigris` or `tidycensus` packages, or other), create a choropleth map using 5 classes, including a title and a legend. Create a map for both *quantile* and *equal-interval* classification, and briefly describe the difference in distribution between your two maps.



7. Repeat Question 6, using a different choice of classes.

Same basic maps as above, but with a different number of classes. Students should make 4 maps in total.