Karyna Steele
Exam 1 Content Summary

So far in this course we have mostly gone deeper into topics that we learned in Stat 112, such as data wrangling and visualizations. We recreated NYT visualizations by exploring geom functions that could give us more interesting plots that have a lot of information, as opposed to just a simple bar or scatter plot. Some of the newer functions include geom_linerange() for a temperature plot, geom_area() for a precipitation plot, and geom_dumbbell() for a labor union plot. Another topic we have much deeper explored is spatial visualization.

In Stat 112, we learned about choropleth maps and how to work with shapefiles, and this semester so far we have done so much more. One of the newer concepts we learned about is coordinate reference systems (CRS), which provides a way to measure and locate points on the three dimensional earth on a two dimensional map. Different geographic coordinate reference systems provide different fits to the earth's ellipsoid shape, so we need to be mindful of what CRS underlies the data we want to use. We learned about more functions in the sf package, such as st_crs() which can help with the task just mentioned. We can also transform a dataset's CRS with st_transform() if we want to join two shapefiles, using st_join(), that may be using two different systems. We also used st_read() to read in the shapefile data that we want to visualize, st_crop() to crop the margins of a shapefile to a specific area that we are interested in, and st_bbox() to determine the coordinate region of a shapefile. For visualizing spatial data, we heavily relied on geom_sf().

We have also gone more in depth about data wrangling techniques, going further than the functions we learned in Stat 112. Specifically, we dove deeper into the stringr package, learning how to use regular expressions to find patterns in strings. We defined some core stringr function that we should keep in mind for the rest of the course. Some other non-content related things that I have also learned to keep in mind are how to manage my files, open new projects in RStudio, use github and generate tokens, and publish a website.