Problem Set 3 - PPOL 6802 Spring 2024

Alex Lundry

2024-02-17

Problem Set #3

Here is where you will provide the answers to problem set 3. I will provide you with the basic "buckets" to put everything in to make the R Markdown creation process as straightforward as possible. But remember that if you run into any issues, ChatGPT can be extremely helpful!

Loading the Data

To complete this problem set, you will need to download the necessary data from Github. You can do that by running the code below:

```r
library(tidyverse)

d1 <- read_csv("https://raw.githubusercontent.com/alexlundry/ppol6802_spring24/main/problem_sets/dataset

d2 <- read_csv("https://raw.githubusercontent.com/cphalpert/census-regions/master/us%20census%20bureau%2

d3 <- left_join(d1, d2, by = "State")
```

Question 1 - Bar Chart with Summary Data

Using ggplot2, make a bar chart that shows the average level of education (pct_bachelors) for counties located in each census Division. Ensure that bars are properly sorted in order and appropriately labelled. You must also have meaningful and accurate title and axis labels.

Bonus (not required): Color the bars by their Census Region. (+1 point)

Question 2 - Grouped boxplot

Using ggplot2, create a horizontal boxplot of `pct_soc_sec` that is grouped by census Region. Ensure that the boxplots are meaningfully sorted. You must also have a meaningful and accurate title and axis labels. Color the boxplots by Region, but make sure there is NOT a legend displayed.
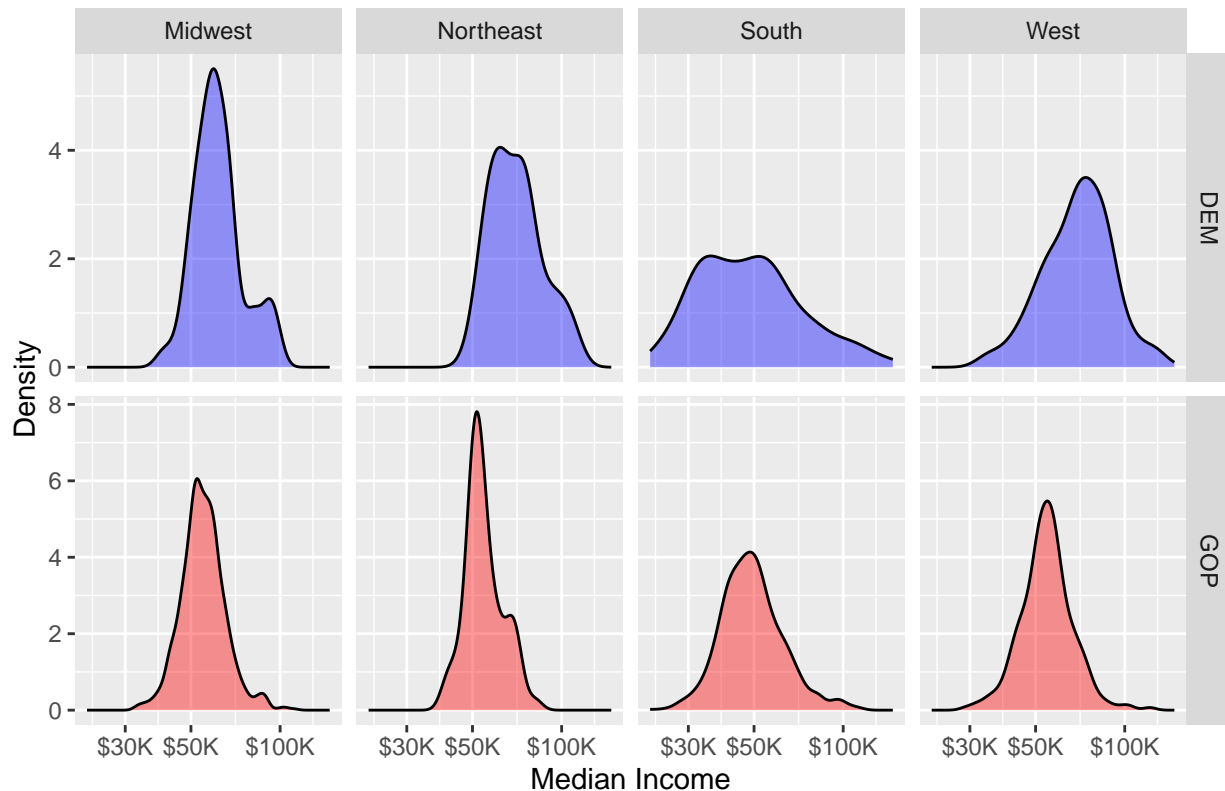
Question 3 - Plot Reproduction

Using ggplot2, reproduce the plot below. To help you do this, here are a few key notes:

- I have filtered out rows where `winner` is NA.
- I am visualizing `med_income` and coloring by `winner`

- These are density plots with transparency = 0.4
- The plots are faceted by two variables, Region as the columns, and winner as the rows.
- The scales of the X axis are fixed, and the scale is logged.
- The X axis labels are formatted for displaying dollars, and the results are in thousands with the "K" suffix.
- There is a title and axis labels.
- There is no legend.

Question 4 - Research Question

Oh no! Your try-hard problematic Gen X boss that wants to be a Gen Z just texted you:

> "Angela and Kevin over in Accounting asked me if we could whip up a chart showing how each census division contributes to the overall population of a census region. He said it was something about an IRS audit?? ROFL! I told him I'd put my best people on it, but he said something about"plausible deniability"?!? "¯*(:)/¯*" Get that to me ASAP - I've got an Orange Theory class in an hour."

What do you send your boss? (don't overthink this - the chart can be very basic)

BUTTERUP YOUR BOSS BONUS (not required): Offer an alternative way to visualize the data they are asking for and make a recommendation for the most effective one to use. (up to +3 points)