cm013

Effective Visualizations

Now that you know how to create graphics and visualizations in R, you are armed with powerful tools for scientific computing and analysis. With this power also comes great responsibility. Effective visualizations is an incredibly important aspect of scientific research and communication. There have been several books (see references) written about these principles. In class today we will be going through several case-studies trying to develop some expertise into making effective visualizations.

Worksheet

The worksheet questions for today are embedded into the class notes.

You can download this Rmd file here

Note, there will be very little coding in-class today, but I've given you plenty of exercises in the form of a supplemental worksheet (linked at the bottom of this page) to practice with after class is over.

Resources

1. Fundamentals of Data Visualization by Claus Wilke.
2. Visualization Analysis and Design by Tamara Munzner.
3. STAT545.com - Effective Graphics by Jenny Bryan.
4. ggplot2 book by Hadley Wickam.
5. Callingbull.org by Carl T. Bergstrom and Jevin West.

Part 1: Warm-up and pre-test [20 mins]

Warmup:

Write some notes here about what "effective visualizations" means to you. Think of elements of good graphics and plots that you have seen - what makes them good or bad? Write 3-5 points.

1. Clear graphs/plots of the data - this means that plots should be displayed in their most paired-down form (simple graphs)
2. Use of colour and data labels helps visualization
3. good graphics are interesting to look at, they link findings to maps or visuals that the target reader will find interesting
4. The creator of the effective visualization understands the audience for the graphics, therefore, they will use targeted graphics.

CQ01: Weekly hours for full-time employees

> Question: Evaluate the strength of the claim based on the data: "German workers are more motivated and work more hours than workers in other EU nations."
>
> Very strong, strong, weak, very week, do not know

- << very weak - this graph shows that romania and the UK have higher hours worked also this graph does not account for motivation - it is only a graph for hours worked and is not a measure of human motivation. Also, the authors did not include how many people were surveyed for this graph. The difference between the countries in the EU is not sufficient enough to make claims in human motivation. >>

- Main takeaway: Summarize the main takeaway from this question/discussion here

CQ02: Average Global Temperature by year

Question: For the years this temperature data is displayed, is there an appreciable increase in temperature?

Yes, No, Do not know

- <<From the data displayed there is no appreciable increase in temperature. The y axis range is not appropriate therefore any claims made from this graph would be hard to make. There seems to be an increase in temperature however it is not displayed in an easily readable format. Years are not labeled. >>
- Main takeaway: Summarize the main takeaway from this question/discussion here

CQ03: Gun deaths in Florida

Question: Evaluate the strength of the claim based on the data: "Soon after this legislation was passed, gun deaths sharply declined."

Very strong, strong, weak, very week, do not know

- <<very weak, it seems to be the opposite, gun deaths increased but the scale is inverted so when first looking at the plot it seems to be the opposite.>>
- Main takeaway: Summarize the main takeaway from this question/discussion here

Part 2: Extracting insight from visualizations [20 mins]

Great resource for selecting the right plot: https://www.data-to-viz.com/ ; encourage you all to consult it when choosing to visualize data.

Case Study 1: Context matters

Causation v. correlation issue. There seems to be a correlation but if causation is not proved then this grpah doesn't mean anything. The axes are completely different, they have a factor of a 100 difference between yaxis right and yaxisleft

Case Study 2: A case for pie charts

Part 3: Principles of effective visualizations [20 mins]

We will be filling these principles in together as a class

1.
2.
3.
4.
5.

Make a great plot worse

Instructions: Below is a code chunk that shows an effective visualization. First, copy this code chunk into a new cell. Then, modify it to purposely make this chart "bad" by breaking the principles of effective visualization above. Your final chart still needs to run/compile and it should still produce a plot.

How many of the principles did you manage to break?

Plotly demo [10 mins]

Did you know that you can make interactive graphs and plots in R using the plotly library? We will show you a demo of what plotly is and why it's useful, and then you can try converting a static ggplot graph into an interactive plotly graph.

This is a preview of what we'll be doing in STAT 547 - making dynamic and interactive dashboards using R!

Supplemental worksheet (Optional)

You are highly encouraged to the cm013 supplemental exercises worksheet. It is a great guide that will take you through Scales, Colours, and Themes in ggplot. There is also a short guided activity showing you how to make a ggplot interactive using plotly.