Modern Data Science with R 1st edition Baumer Solutions Manual

SKU: 9781498724487-SOLUTIONS

Chapter 2

Data visualization I

2.1 Exercises

Exercise 2.1

What would a Cartesian plot that used colors to convey categorical values look like?

Answer of exercise 2.1

A heatmap.

Exercise 2.2

Consider the two graphics related to *The New York Times* "Taxmageddon" article at `http://www.nytimes.com/2012/04/15/sunday-review/coming-soon-taxmageddon.html`. The first is "Whose Tax Rates Rose or Fell" and the second is "Who Gains Most From Tax Breaks."

1. Examine the two graphics carefully. Discuss what you think they convey. What story do the graphics tell?

2. Evaluate both graphics in terms of the taxonomy described in this chapter. Are the scales appropriate? Consistent? Clearly labelled? Do variable dimensions exceed data dimensions?

3. What, if anything, is misleading about these graphics?

Answer of exercise 2.2

1. Answers will vary. The main take-aways are that tax rates on the rich have fallen, while tax rates on the poor have not. Moreover, in terms of the "cost" to the US Treasury, most tax breaks go to those in the top 20% in income.

2. Answers will vary. In the first graphic, the y-scale is the tax rate, but there is no axis. The labels about the segment of the income distribution are not reflected in any quantity. The color scheme is not immediately helpful. The change in real pre-tax income figures to not have a corresponding quantitative representation.

   In the second graphic, the y-scale is very confusing, and possibly misleading. The scale is a percentage, and yet the income amounts on the left are *not* on a linear scale.

Modern Data Science with R 1st edition Baumer Solutions Manual

3. Answers will vary. Most obviously, the y-scales in either data graphic.

Exercise 2.3

Choose *one* of the data graphics listed at `http://mdsr-book.github.io/exercises.html#exercise_23` and answer the following questions. Be sure to indicate which graphical display you picked.

1. Identify the visual cues, coordinate system, and scale(s).
2. How many variables are depicted in the graphic? Explicitly link each variable to a visual cue that you listed above.
3. Critique this data graphic using the taxonomy described in this chapter.

Answer of exercise 2.3

Answers will vary.

Exercise 2.4

Answer the following questions for each of the following collections of data graphics listed at `http://mdsr-book.github.io/exercises.html#exercise_24`.

Briefly (one paragraph) critique the designer's choices. Would you have made different choices? Why or why not?

Note: Each link contains a collection of many data graphics, and we don't expect (or want) you to write a dissertation on each individual graphic. But each collection shares some common stylistic elements. You should comment on a few things that you notice about the design of the collection.

Answer of exercise 2.4

Answers will vary.

Exercise 2.5

Consider one of the more complicated data graphics listed at `http://mdsr-book.github.io/exercises.html#exercise_25`.

1. What story does the data graphic tell? What is the main message that you take away from it?
2. Can the data graphic be described in terms of the taxonomy presented in this chapter? If so, list the visual cues, coordinate system, and scales(s) as you did in Problem 2(a). If not, describe the feature of this data graphic that lies outside of that taxonomy.
3. Critique and/or praise the visualization choices made by the designer. Do they work? Are they misleading? Thought-provoking? Brilliant? Are there things that you would have done differently? Justify your response.

Answer of exercise 2.5

Answers will vary.

Exercise 2.6

Consider the data graphic (`http://tinyurl.com/nytimes-unplanned`) about birth control methods.

Modern Data Science with R 1st edition Baumer Solutions Manual

1. What quantity is being shown on the y-axis of each plot?
2. List the variables displayed in the data graphic, along with the units and a few typical values for each.
3. List the visual cues used in the data graphic and explain how each visual cue is linked to each variable.
4. Examine the graphic carefully. Describe, in words, what *information* you think the data graphic conveys. Do not just summarize the *data*—interpret the data in the context of the problem and tell us what it means.

Answer of exercise 2.6

Modern Data Science with R 1st edition Baumer Solutions Manual