Data Viz | Assignment 2

Julian McClellan and Jonathan Camacho

April 23, 2017

Introduction

One of the objectives of data visualization is to facilitate the reader's understanding and extraction of the data. In this sense, Cairo (2016) highlights the importance of correctly selecting the best ways for data coding. He proposes a Hierarchy of Elementary perceptual tasks, developed and tested by William S. Cleveland and Robert McGill, as a guide to determining the best channels for coding data. According to this *Hierarchy of Elementary Perceptual Tasks*, position along identical scales is the most effective mental-visual task; allowing readers to extract information more accurately from plots.

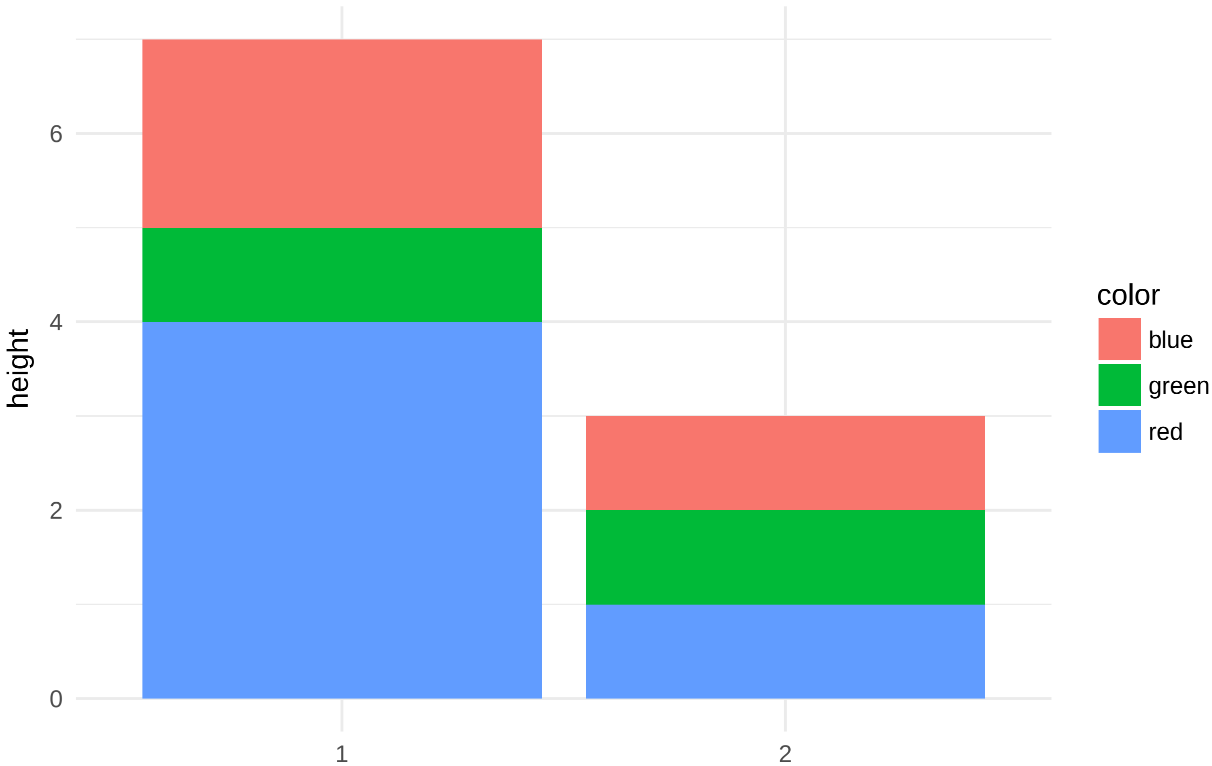
In this assignment, we expand this idea arguing that the context of the position of the data also has an important effect on the effectiveness of the mental-visual task and people's ability to accurately find and extract information from a plot. We pose the following research question: Is data in a stacked bar plot more difficult to estimate and extract than data in a dodge bar plot? We hypothesize that the sum of the bar heights extracted from a stack plot is less accurate than the sum of bar heights extracted from a dodge plot. We assume as a mechanism, the fact that bars in a dodge plot have a baseline based on zero while bars in a stack plot have a variable baseline. The idea is that having a baseline of zero facilitates information estimation and extraction.

Methods

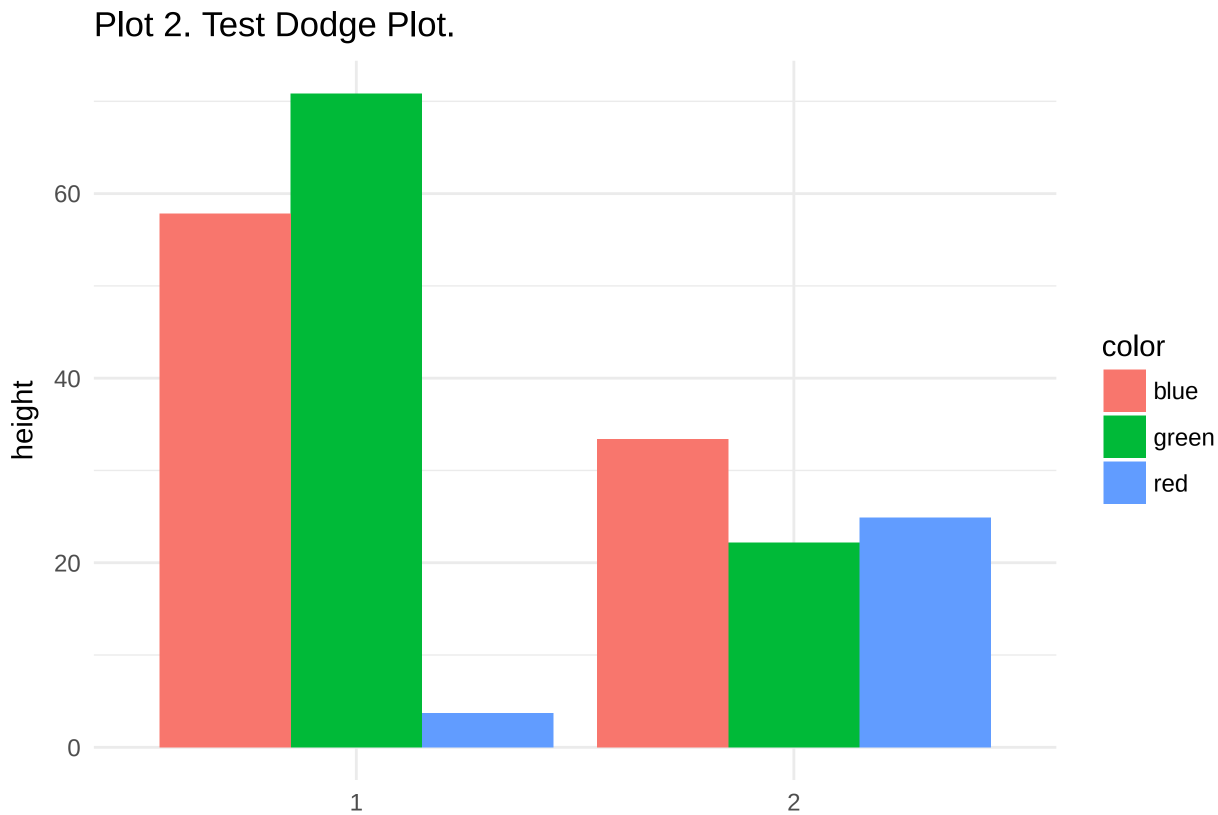
To test this hypothesis, we designed an experiment consisting of one task: estimating the sum of bar heights from a plot. The participants have to read the heights of two bars (the green bars) from data divided into two groups in the same plot and add the heights to create a total height. To compare the effect that stacked vs. dodge positioning of information has on the participants' ability to estimate the bars' sum, the data is presented in two different types of plots: one stacked and the other dodged.

Moreover, to test the accuracy of the experimental procedure, participants are asked to perform the task twice; First, with two *screen* plots (one stacked and one dodged) and with one *test* plot. Both plot sets consist of one stack and one dodge plot. The *screen* set of plots were created with data randomly generated with the purpose of setting a validity benchmark for the data; how accurately participant can extract information with an easy version of the tasks. (Plot 1) The bar heights are more precise and thus easier to sum in the *screen* plots, while the bar heights in the *test* plots are less precise.

Plot 1. Screen Stack Plot.

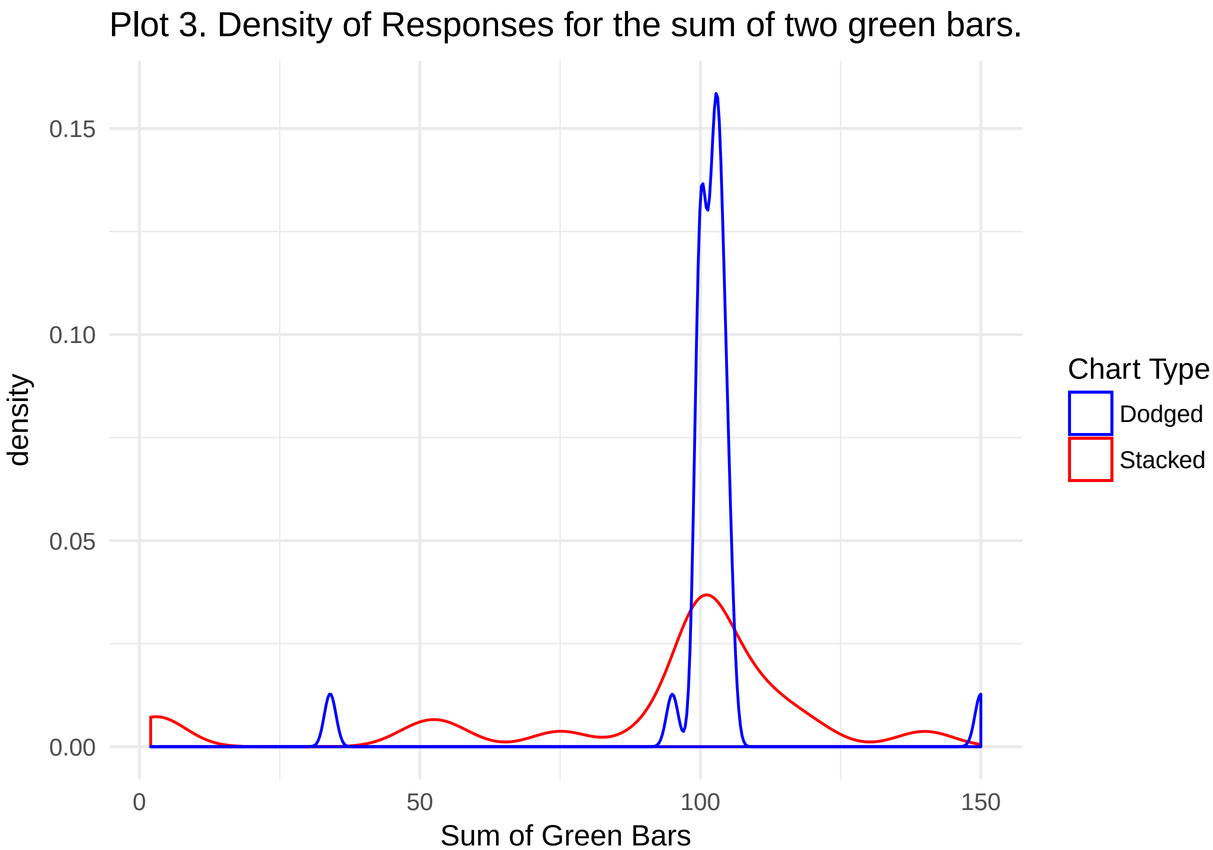


Then, randomly, we assigned each participant to one plot from the *test* set; a stacked or a dodged plot. (Plot 2) The participant has to estimate the total height of the green bars for this task. The data for this experiment was collected through a survey developed in Qualtrics and deployed through the online platform Mechanical Turk from Amazon. Data from this experiment was collected from a 100 participants, 50 to each plot type. The data was cleaned, transformed (the results were separated into a stack, dodge and all results data set), and analyzed. For this experiment, we developed two variables. The independent variable is the type of plot (stack or dodge), and it is operationalized as the type of plot presented to the participants. The dependent variable is the sum of green bar heights estimated by the participant.



Data Summary

Below (Plot 3) shows the density distributions of participants responses: the total sum of the height of the two green bars in the plot. The data is presented according to the randomly assigned plot to the participant: stacked or dodge. The plot suggests that the sums of green bars estimated from the stack plot have less variance than the sums estimated from the dodge plot, which has a flatter distribution of estimated sums, with what seems like a left-skew.

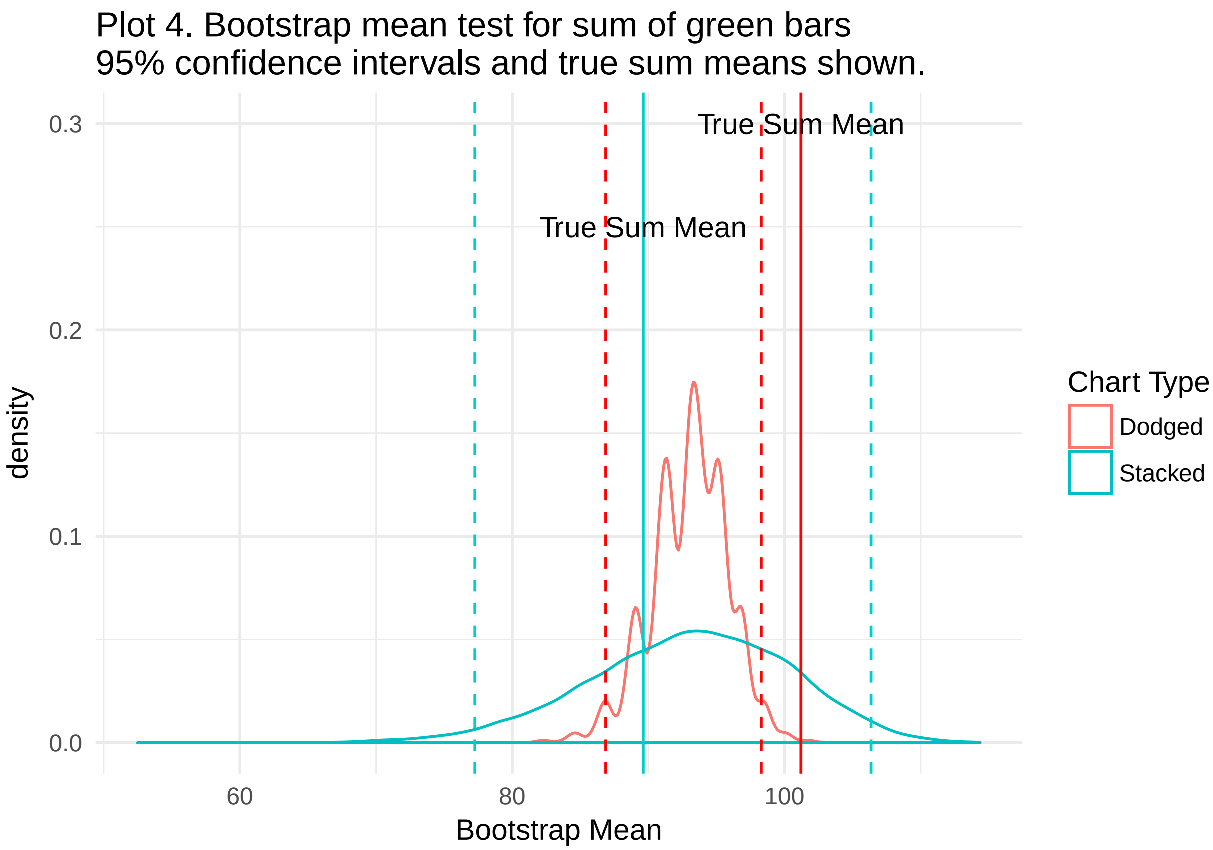


Analysis

To determine whether the distributions of dodged and stack bar heights captured the true sum of the green bars, we conduct a bootstrap mean test was conducted with a 95% confidence interval. Finally, an F-test was also conducted to compare two variances. However, the F-test might not be the most powerful variance test available, given that both distributions are not perfectly normal.

In Plot 4 below, we see that the confidence interval for the stacked means is much wider compared to that of the dodged means. Still, the stacked confidence interval actually captures the original mean, while the dodge confidence interval does not.

Finally, to compare the two variances (means from stack and dodge plots), an F-test was conducted. It is important to note that the p-value is significantly low, indicating that the variances are probably unequal.



Conclusion.

For this experiment, we hypothesized that information presented in a stack plot it is more difficult to estimate accurately than information presented in a dodge plot. The results from the statistical analysis seem to contradict this hypothesis, but in an surprising way.

While the estimations of dodge plot readers was more tightly distributed (more precise), a bootstrap means test shows that the true value of the green bar heights was not captured by the distribution, showing that the dodged chart readers were more precise but less accurate. On the other hand, the distribution of estimations from the stacked plot readers was more spread (less precise), but that distribution was able to capture the true sum of the green bars, and was thus more accurate. These results seem to indicate that presenting data in dodged bars format facilitates more precise, if more inaccurate, information estimation compared to presenting data in a stacked bar format.

We do note however that the distribution of the dodged bar readers appear clustered around `100`, indicating that there might be some "hundreds" round up bias. The true value: 93.0186174 might have been too closely convenient to `100` for readers of the dodged chart to more precisely refine their estimates.

Citations.

- Cairo, Alberto. The Truthful Art: Data, charts, and maps for communication. New Riders, 2016.


```
##  setting  value
##  version  R version 3.3.3 (2017-03-06)
##  system   x86_64, linux-gnu
##  ui       X11
##  language en_US
```

```
##  collate  en_US.UTF-8
##  tz       America/Chicago
##  date     2017-04-28
##
##  package    * version  date       source
##  assertthat   0.2.0    2017-04-11 CRAN (R 3.3.3)
##  backports    1.0.5    2017-01-18 CRAN (R 3.3.3)
##  bootstrap  * 2017.2   2017-02-27 CRAN (R 3.3.3)
##  broom        0.4.2    2017-02-13 CRAN (R 3.3.3)
##  cellranger   1.1.0    2016-07-27 CRAN (R 3.3.3)
##  colorspace   1.3-2    2016-12-14 CRAN (R 3.3.3)
##  DBI          0.6-1    2017-04-01 CRAN (R 3.3.3)
##  devtools   * 1.12.0   2016-12-05 CRAN (R 3.3.3)
##  digest       0.6.12   2017-01-27 CRAN (R 3.3.3)
##  dplyr      * 0.5.0    2016-06-24 CRAN (R 3.3.3)
##  evaluate     0.10     2016-10-11 CRAN (R 3.3.3)
##  forcats      0.2.0    2017-01-23 CRAN (R 3.3.3)
##  foreign      0.8-68   2017-04-24 CRAN (R 3.3.3)
##  ggplot2    * 2.2.1    2016-12-30 CRAN (R 3.3.3)
##  gtable       0.2.0    2016-02-26 CRAN (R 3.3.3)
##  haven        1.0.0    2016-09-23 CRAN (R 3.3.3)
##  hms          0.3      2016-11-22 CRAN (R 3.3.3)
##  htmltools    0.3.5    2016-03-21 CRAN (R 3.3.3)
##  httr         1.2.1    2016-07-03 CRAN (R 3.3.3)
##  jsonlite     1.4      2017-04-08 CRAN (R 3.3.3)
##  knitr        1.15.1   2016-11-22 CRAN (R 3.3.3)
##  labeling     0.3      2014-08-23 CRAN (R 3.3.3)
##  lattice      0.20-35  2017-03-25 CRAN (R 3.3.3)
##  lazyeval     0.2.0    2016-06-12 CRAN (R 3.3.3)
##  lubridate    1.6.0    2016-09-13 CRAN (R 3.3.3)
##  magrittr     1.5      2014-11-22 CRAN (R 3.3.3)
##  memoise      1.1.0    2017-04-21 CRAN (R 3.3.3)
##  mnormt       1.5-5    2016-10-15 CRAN (R 3.3.3)
##  modelr       0.1.0    2016-08-31 CRAN (R 3.3.3)
##  munsell      0.4.3    2016-02-13 CRAN (R 3.3.3)
##  nlme         3.1-131  2017-02-06 CRAN (R 3.3.3)
##  plyr         1.8.4    2016-06-08 CRAN (R 3.3.3)
##  psych        1.7.3.21 2017-03-22 CRAN (R 3.3.3)
##  purrr      * 0.2.2    2016-06-18 CRAN (R 3.3.3)
##  R6           2.2.0    2016-10-05 CRAN (R 3.3.3)
##  Rcpp         0.12.10  2017-03-19 CRAN (R 3.3.3)
##  readr      * 1.1.0    2017-03-22 CRAN (R 3.3.3)
##  readxl       1.0.0    2017-04-18 CRAN (R 3.3.3)
##  reshape2     1.4.2    2016-10-22 CRAN (R 3.3.3)
##  rmarkdown    1.4      2017-03-24 CRAN (R 3.3.3)
##  rprojroot    1.2      2017-01-16 CRAN (R 3.3.3)
##  rvest        0.3.2    2016-06-17 CRAN (R 3.3.3)
##  scales       0.4.1    2016-11-09 CRAN (R 3.3.3)
##  stringi      1.1.5    2017-04-07 CRAN (R 3.3.3)
##  stringr      1.2.0    2017-02-18 CRAN (R 3.3.3)
##  tibble     * 1.3.0    2017-04-01 CRAN (R 3.3.3)
##  tidyr      * 0.6.1    2017-01-10 CRAN (R 3.3.3)
##  tidyverse  * 1.1.1    2017-01-27 CRAN (R 3.3.3)
##  withr        1.0.2    2016-06-20 CRAN (R 3.3.3)
```

```
## xml2          1.1.1    2017-01-24 CRAN (R 3.3.3)
## yaml          2.1.14   2016-11-12 CRAN (R 3.3.3)
```