Journal of Data Science xx (xx), 1–9 DOI: xx.xxxx/xxxxxxxx
July 2020

Testing Charts: viewer's perceptual accuracy in surveys

Kiegan Rice*[1], Heike Hofmann†[2], Nola du Toit‡[1], Edward Mulrow§[1], and [2]

[1]National Opinion Research Center (NORC)

[2]Department of Statistics, Iowa State University

Abstract

The use of visuals is a key component in scientific communication, and decisions about the design of a data visualization should be informed by what design elements best support the audience's ability to perceive and understand the components of the data visualization. We build on the foundations of Cleveland and McGill's work in graphical perception, employing a large, nationally-representative, probability-based panel of survey respondents to test perception in statistical charts. Our findings provide actionable guidance for data visualization practitioners to employ in their work.

Cleveland and McGill (1984)

This is Heike's color for making changes

Introduction

How do structural design choices in a data visualization impact viewers' ability to identify the larger of two elements?

How do aesthetic design choices in a data visualization impact viewers' ability to identify the larger of two elements?

*Corresponding author. Email: rice-kiegan@norc.org
†Email: hofmann@iastate.edu
‡Email: dutoit-nola@norc.org
§Email: mulrow-edward@norc.org

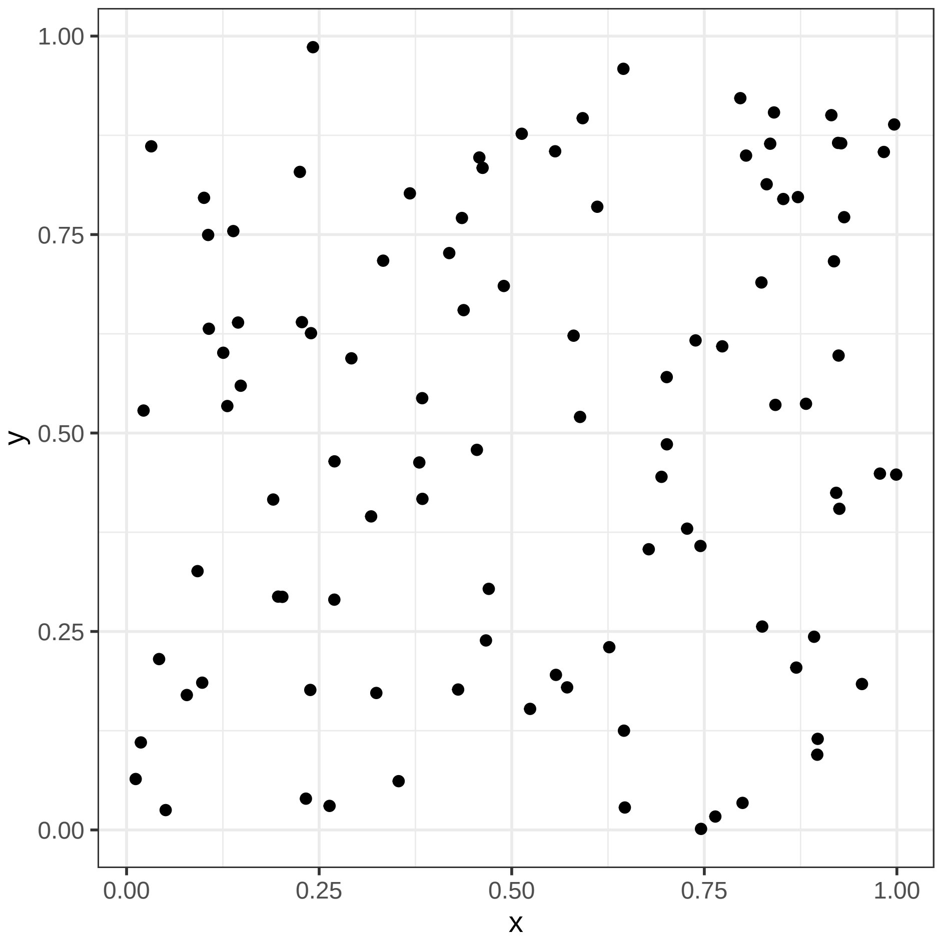

Figure 1: quarto figure caption.

How is viewer behavior (zooming, time spent on question, certainty of response) impacted by structural and aesthetic design choices in a data visualization?

Structural design choices:

Structural design choices

Mapping, Stacked bar, Vertical, Horizontal, Horizontal wide

Facetted bar

Only have a split sample for this

Pie, Alignment

We have this for all above mappings, but the setup is a little different for facetted bar

Aesthetic design choices (structural choices seems stronger/there could be a lot to talk about there... should we skip aesthetic on this one?)

Colors

Use of gridlines

Outcomes/responses for modeling:

Binary accuracy (correct/incorrect – 'they are the same' is incorrect here)

Ordinal response (a/b/they are the same)

Zooming behavior (zoomed/did not zoom)

Time spent on question (continuous, in seconds)

Certainty

I've noted this below, but: how do model? Ordinal response? Binary (certain or very certain vs everybody else)?

Survey setup - Stimulus description

We decided to show respondents a stacked bar chart (because there is little research on stacked bars) and ask to compare the sizes of two tiles (labelled A vs B or C vs D) to each other.

DATA THAT MAY BE INCLUDED IN ANALYSIS:

ROUNDS 1-2: Color variations on vertical stacked bar, aligned and unaligned

ROUND 3: Horizontal and horizontal wide, aligned and unaligned

ROUND 5: Horizontal wide gridlines (only dark grid split sample)

ROUND 6: Facetted bar (split sample w/o forcing choice)

ROUND 7: Aligned vs unaligned pie (full sample)

Data Analysis and Results

COMBINING SAMPLES AND WEIGHTING NOTES:

We can combine responses across samples into one combined dataset, but we need to adjust weights accordingly so that each sample is weighted equally in the model

Question for Ed: If we compare a full sample to a split sample, do we still want to weight these 'equally'?

Analysis should be done using the 'survey' package and weights should be taken into account.

ANALYSIS PLAN:

Models below structured as:

Response

Covariates to use in each model

STRUCTURAL VARIATION – START HERE

Binary accuracy across structural choices

Model 1:

Alignment only, just vertical stacked bar

Model 2:

Alignment

Bar vs pie (comparable question for pie is A vs B)

Model 3:

Alignment

Vertical x horizontal x horizontal wide

Model 4:

Alignment

Every structure (vertical bar, horizontal bar, horizontal wide bar, facet bar, pie)

Visuals:

% yes across each different structural condition

Facet by aligned/unaligned?

Model estimates + CIs

Ordinal response

Model 1:

Alignment only, just vertical stacked bar

Model 2:

Alignment

Bar vs pie (comparable question for pie is A vs B)

Model 3:

Alignment

Vertical x horizontal x horizontal wide

Model 4:

Alignment

Every structure (vertical bar, horizontal bar, horizontal wide bar, facet bar, pie)

Visuals:

All responses across each different structural condition

Facet by aligned/unaligned?

Model estimates + CIs

Zooming behavior (zoomed/did not zoom)

Model 1:

Device type

Alignment

Vertical x horizontal x horizontal wide

Visuals:

% zoomed by device + alignment (already have this chart)

Model estimates + CIs

Time spent on question (in seconds)

Model 1:

Device type

Zoom

Alignment

Vertical x horizontal x horizontal wide

Model 2 (this may not be feasible for comparison depending on what level the 'TOTALTIME' is captured at):

Device type

Zoom

Alignment

Every structure (vertical bar, horizontal bar, horizontal wide bar, facet bar, pie)

Visuals:

Distribution of time spent variable

Facet by device type, zoom, structural condition, alignment? Play around with it

Average time spent by each of the conditions

Certainty?

Same models as above, but I'm not sure how we want to do the response. Ordinal response? Binary (certain or very certain vs everybody else)?

AESTHETIC VARIATION – ONLY IF TIME

Binary accuracy (correct/incorrect – 'they are the same' is incorrect here) across structural choices

Model 1:

Dark grid vs no grid (only have for horizontal wide)

Response choice (ordinal response)

Model 1:

Dark grid vs no grid (only have for horizontal wide)

Zooming behavior (zoomed/did not zoom)

Model 1:

Device type

Dark grid vs no grid

Time spent on question (in seconds)

Model 1:

Device type

Zoom

Dark grid vs no grid

Certainty?

Same models as above, but I'm not sure how we want to do the response. Ordinal response? Binary (certain or very certain vs everybody else)?

Conclusion

Supplementary Material

- **Participant Data (Linear):** Link to csv file with the data.
- **Data Analysis Code:** Link to an html document with annotated code chunks.

References

Cleveland, William S., and Robert McGill. 1984. "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." *Journal of the American Statistical Association* 79 (387): 531–54. https://doi.org/10.1080/01621459.1984.10478080.