CHAPMAN & HALL/CRC STATISTICS BOOK PROPOSAL

TITLE AND AUTHOR/EDITOR(S)

1. Provisional title of your book.

Interactive web-based data visualization with R and plotly

2. Names, titles, affiliations, and email addresses for all authors/editors

Dr. Carson Sievert

CEO, Sievert Consulting LLC

cpsievert1@gmail.com

CONTENTS

3. Please include a complete table of contents including chapter and section headings.

Why interactive data visualization?

- Data science workflow
- Interactivity augments exploration
  - Some historical context and examples (https://talks.cpsievert.me/20180202/#7)
  - "Interactive graphics enable the analyst to pursue follow-up questions" – Cook & Swayne, GGobi book, p15
  - A case study of US election data
- GUI and command-line - conflict or synergy?
  - https://talks.cpsievert.me/20180305/#12
- Web-based visualization
  - New capabilites (e.g. easy sharing, portability) brings new set of concerns (e.g., client-server, security, etc)
  - Great tools for expository vis (d3.js, vega, plotly.js)
  - Lack of tools for exploratory vis (i.e. tools for iteration)

The plotly for R ecosystem

- plotly is a FOSS tool built with plotly.js and htmlwidgets
- Plotly's cloud services
  - collaboration, point-and-click editing & collaboration
- Getting started (installation, etc)
- Two approaches, one object

The `plot_ly()` toolbox

- Scatter traces
- Maps
- Bars & histograms
- 2D frequencies
- Raster images
- 3D graphics

The `ggplotly()` toolbox

- Makes some plots easier to acheive (e.g., `geom_hex()`, `geom_smooth()`, etc)

- Caveats and limitations
- Learning from and extending `ggplotly()` ()

Arranging multiple views

- Arranging htmlwidgets
- Merging plotly objects
- Navigating many views

Multiple linked views

- Linking views with shiny
- Linking views without shiny

Animating views

- Key frame animations
- Linking animated views

Other miscellaneous topics

- Manipulating widgets with https://github.com/rte-antares-rpackage/manipulateWidget
- Layering on JavaScript
- Saving, exporting, and embedding local images

Advanced applications

- storm paths – https://vimeo.com/257149623
- https://github.com/cpsievert/pedestrians
- https://github.com/cpsievert/zikar
- https://github.com/cpsievert/eechidna
- https://github.com/cpsievert/accidents

plotly extension packages

- Hmisc, heatmaply, TSstudio, autoplotly

SUBJECT/AUDIENCE

4. Please describe in detail the subject of your book. Why will this book be important, who will find it useful, and what is new? What background will you assume and can you name any books that will suffice?

This book is about interactive data visualization using the R package plotly. It focuses on tools and techniques that data analysts should find useful for asking follow-up questions from their data using interactive web graphics. A basic understanding of R will be assumed – the book *R for Data Science* (Wickham & Grolemund, 2017) should be sufficient for understanding the majority of the material.

5. Will your book be primarily a reference/monograph or textbook? If a textbook, for which courses will it be the primary text and at what level is the course taught? Will you include exercises sets and supply a solutions manual?

Monograph

6. What related books are available and what advantages will your book have?

This book focuses *exploring* data with interactive *web* graphics. Most books on interactive web graphics are written for web developers, but this book is written for data analysts that want to quickly leverage interactive techniques to understand their data better and easily share those findings.

- Interactive and Dynamic Graphics for Data Analysis: With R and GGobi
- Graphical Data Analysis with R
- R Graphics Cookbook: Practical Recipes for Visualizing Data
- ggplot2: Elegant Graphics for Data Analysis
- Interactive Data Visualization for the Web: An Introduction to Designing with D3

PRODUCTION

7. Approximately how many printed pages will your book contain? Are color figures essential to your book? If so, about how many would have to be in color? Color printing is still very expensive and color figures will increase the price so black and white should be used unless color is essential.

About 150-200 printed pages. Color printing throughout the book will be essential.

8. When would you hope to be able to submit the final draft of the book to us? Will you use Latex or Word? We will supply a style file for LaTeX authors and request an unformatted file from Word authors.

I assume a final draft means after reviews have been made? In that case, I would be happy if a final draft was ready by the end of the year. I'll be using LaTeX via rmarkdown.

REVIEWS

9. Please give the names and e-mail addresses of four people who would be qualified to give an opinion on your proposed book.

Dr. Dianne Cook visnut@gmail.com

Dr. Heike Hofmann heike.hofmann@gmail.com

Dr. Hadley Wickham hadley@rstudio.com

Dr. Antony Unwin au50au@me.com

Dr. Barret Schloerke schloerke@gmail.com

MARKETING

10. If your book is aimed at a professional/research market, are there key societies outside of statistics to which it should be marketed? (All major statistical societikes will be covered.)

IEEE, IEEE VIS, vgtc

11. Please list up to six key features of your proposed book that we can use in bulleted form.

- Convert static **ggplot2** graphics to an interactive web-based form
- Link, animate, and arrange multiple plots in standalone HTML from R

- Embed, modify, and respond to plotly graphics in a shiny app
- Learn best practices for visualizing continuous, discrete, and multivariate data
- Learn numerous ways to visualize geo-spatial data
- Learn how to further customize interactive behavior with JavaScript

12. Please list up to six key words or phrases that people interested in this topic may use to search Amazon or the web. Do not repeat words in the title as these will already be found.

Interactive data visualization, web graphics, data science, statistical graphics, information visualization

13. Please select the three most important markets for your book. Other categories are available including education, psychology, and economics so please mention other important disciplines.

Is there a section on statistical graphics?

STA07J-Statistics-Statistical Learning & Data Mining STA08A-Statistics-Computational Statistics STA09A-Statistics-Statistics for Business, Finance & Economics