gdpR: An R Package for studying differentially private algorithms

by Jordan A. Awan, Kevin Eng, Robin Gong, Nianqiao Phyllis Ju, and Vinayak A. Rao

Abstract This paper serves as a reference and introduction on using the *gdpR* R package. The goal of this package is to provide some tools for exploring the impact of different privacy regimes on a Bayesian analysis. A strength of this framework is the ability to target the exact posterior in settings where the likelihood is too complex to analytically express.

1 Introduction

The ease and pervasiveness of modern data collection technologies has raised concerns about data privacy. (Dwork and Roth 2013) introduced the differential privacy framework as a means to rigorously define privacy. The framework has lead to the development of many "privitized'' versions of existing statistical methods. The process of privitizing usually consist of introducing random noise in someway using a known distribution.

2 overview of the gdpR packge

This section reviews This will show a verbatim inline R expression `` `r 1+1` `` in the output.

3 Background

Some packages on interactive graphics include **plotly** (Sievert 2020) that interfaces with Javascript for web-based interactive graphics, **crosstalk** (Cheng and Sievert 2021) that specializes cross-linking elements across individual graphics. The recent R Journal paper **tsibbletalk** (Wang and Cook 2021) provides a good example of including interactive graphics into an article for the journal. It has both a set of linked plots, and also an animated gif example, illustrating linking between time series plots and feature summaries.

4 Customizing tooltip design with ToOoOlTiPs

ToOoOlTiPs is a packages for customizing tooltips in interactive graphics, it features these possibilities.

5 A gallery of tooltips examples

The **palmerpenguins** data (Horst, Hill, and Gorman 2020) features three penguin species which has a lovely illustration by Alison Horst in Figure 1.

Table 1 prints at the first few rows of the `penguins` data:

Figure 2 shows an plot of the penguins data, made using the **ggplot2** package.

```
penguins %>%
  ggplot(aes(x = bill_depth_mm, y = bill_length_mm,
             color = species)) +
  geom_point()
```

Table 1: A basic table

species	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	sex	year
Adelie	Torgersen	39.1	18.7	181	3750	male	2007
Adelie	Torgersen	39.5	17.4	186	3800	female	2007
Adelie	Torgersen	40.3	18.0	195	3250	female	2007
Adelie	Torgersen	NA	NA	NA	NA	NA	2007
Adelie	Torgersen	36.7	19.3	193	3450	female	2007
Adelie	Torgersen	39.3	20.6	190	3650	male	2007

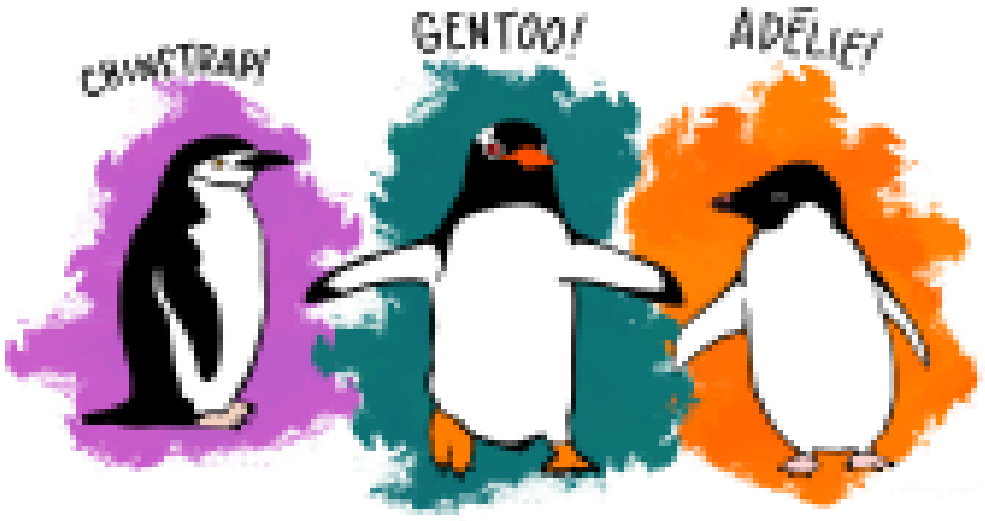

Figure 1: Artwork by allison_horst

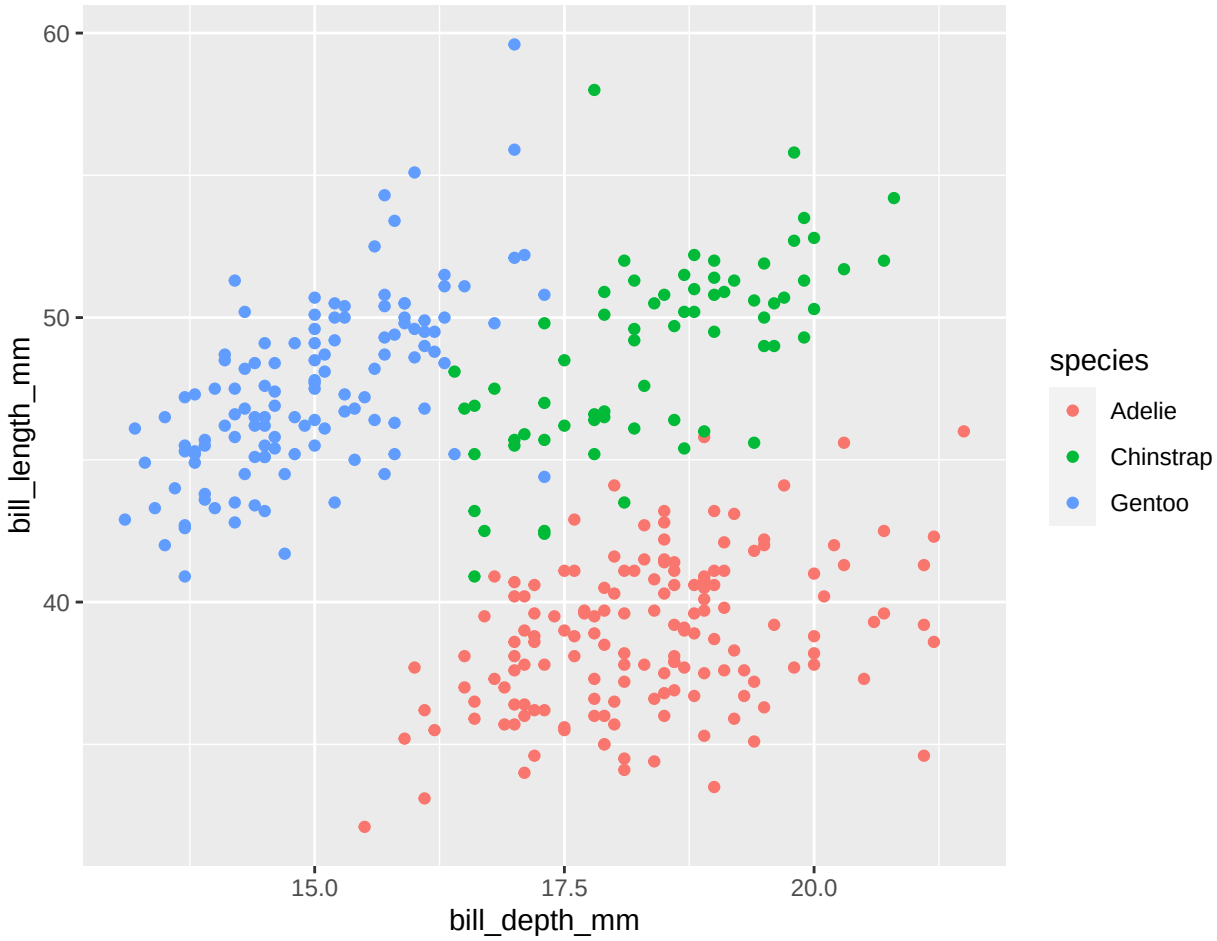

Figure 2: A basic non-interactive plot made with the ggplot2 package on palmer penguin data. Three species of penguins are plotted with bill depth on the x-axis and bill length on the y-axis. Visit the online article to access the interactive version made with the plotly package.

6 Summary

We have displayed various tooltips that are available in the package **ToOoOlTiPs**.

References

Cheng, Joe, and Carson Sievert. 2021. *crosstalk: Inter-Widget Interactivity for HTML Widgets*. https://CRAN.R-project.org/package=crosstalk.

Dwork, Cynthia, and Aaron Roth. 2013. "The Algorithmic Foundations of Differential Privacy." *Foundations and Trends® in Theoretical Computer Science* 9 (3-4): 211–407. https://doi.org/10.1561/0400000042.

Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. https://allisonhorst.github.io/palmerpenguins/.

Sievert, Carson. 2020. *Interactive Web-Based Data Visualization with r, Plotly, and Shiny*. Chapman; Hall/CRC. https://plotly-r.com.

Wang, Earo, and Dianne Cook. 2021. "Conversations in Time: Interactive Visualisation to Explore Structured Temporal Data." *The R Journal*. https://doi.org/10.32614/RJ-2021-050.

Jordan A. Awan
Purdue University
Department of Statistics
West Lafayette, IN 47907
https://www.britannica.com/animal/quokka
jawan@purdue.edu

Kevin Eng
Rutgers University
Department of Statistics
Piscataway, NJ 08854
https://www.britannica.com/animal/quokka
ke157@stat.rutgers.edu

Robin Gong
Rutgers University
Department of Statistics
Piscataway, NJ 08854
https://www.britannica.com/animal/quokka
ruobin.gong@rutgers.edu

Nianqiao Phyllis Ju
Purdue University
Department of Statistics
West Lafayette, IN 47907
https://www.britannica.com/animal/quokka
nianqiao@purdue.edu

Vinayak A. Rao
Purdue University
Department of Statistics
West Lafayette, IN 47907
https://www.britannica.com/animal/quokka
varao@purdue.edu