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 An abstract of less than 150 words.

1 Introduction

Interactive data graphics provides plots that allow users to interact them. One of the most basic types of interaction is through tooltips, where users are provided additional information about elements in the plot by moving the cursor over the plot.

This paper will first review some R packages on interactive graphics and their tooltip implementations. A new package **ToOoOlTiPs** that provides customized tooltips for plot, is introduced. Some example plots will then be given to showcase how these tooltips help users to better read the graphics.

2 Using gdpR

Sampling function

The main function in the **gdpR** is the `gdp_sample` function. The call signature of the function is:

```
gdp_sample(data_model, sdp, nobs, init_par, niter = 2000, warmup = floor(niter / 2),
chains = 1, varnames = NULL)
```

The three required inputs into `gdp_sample` function are the privacy model (`data_model`), the value of the observed privatized statistic (`sdp`), and the total number of observations in the complete data (`nobs`) [MAKE SURE NOTATION IS INTRODUCED]. The **gdpR** package is best suited for problems where the complete data can be represented in tabular form. This is because internally, it is represented as a matrix.

The optional arguments are the number of mcmc draws (`niter`), the burn in period (`warmup`), number of chains (`chains`) and character vector that names the parameters. Running multiple chains can be done in parallel using the **furrr** package. Additionally, progress can be monitored using the **progressr** package.

Creating a `privacy` object

Creating a `privacy` object is done via the `new_privacy` function.

```
new_privacy(post_smpl = NULL, lik_smpl = NULL, ll_priv_mech = NULL,           st_calc
= NULL, add = FALSE, npar = NULL)
```

. The `data_model` input is a `privacy` object that can be constructed using the `new_privacy` constructor. The process of constructing a `privacy` object will be discussed in the next section.

Parallel computing is implemented using the **futures** package.

The `new_privacy` function creates a new data model for input in the `gdp_sample` function.

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 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.

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