autoplotly: An R package for automatic generation of interactive visualizations for statistical results

Yuan Tang[1]

1 H2O.ai

Summary

Often times users only want to quickly iterate the process of exploring data, building statistical models, and visualizing the model results, especially the models that focus on common tasks such as clustering and time series analysis. Some of these packages provide default visualizations using base R's `plot()` for the data and models they generate. However, they look out-of-fashion and these components require additional transformation and clean-up before using them in `ggplot2` (Wickham 2009) and each of those transformation steps must be replicated by others when they wish to produce similar charts in their analyses. Creating a central repository for common/popular transformations and default plotting idioms would reduce the amount of effort needed by all to create compelling, consistent and informative charts. The `ggfortify` (Tang, Horikoshi, and Li 2016; Horikoshi and Tang 2018) package provides a unified `ggplot2` plotting interface `autoplot()` to many statistics and machine learning packages and functions in order to help these users achieve reproducibility goals with minimal effort. `ggfortify` package has a very easy-to-use and uniform programming interface that enables users to use one line of code to visualize statistical results of many popular R packages using `ggplot2` as building blocks. This helps statisticians, data scientists, and researchers avoid both repetitive work and the need to identify the correct `ggplot2` syntax to achieve what they need. Users are able to generate beautiful visualizations of their statistical results produced by popular packages with minimal effort. `plotly` (Sievert et al. 2018) is an R package built on top of Javascript visualization framework `plotly.js` that become popular building blocks for creating interactive visualizations in R. Its `ggplotly()` function can convert graphics created by `ggplot2` to interactive plots with `plotly` style.

The `autoplotly` (Tang 2018b, 2018a) package is an extension built on top of R packages `ggplot2`, `plotly`, and `ggfortify` to provide a unified `autoplotly()` function to automatically generate interactive visualizations for many popular statistical results supported by `ggfortify` package, such as time series, PCA, clustering and survival analysis, with `plotly` and `ggplot2` style. The generated visualizations can also be easily extended using `ggplot2` and `plotly` syntax while staying interactive.

References

Horikoshi, Masaaki, and Yuan Tang. 2018. *Ggfortify: Data Visualization Tools for Statistical Analysis Results*. http://CRAN.R-project.org/package=ggfortify.

Sievert, Carson, Chris Parmer, Toby Hocking, Scott Chamberlain, Karthik Ram, Marianne Corvellec, and Pedro Despouy. 2018. *Plotly: Create Interactive Web Graphics via 'Plotly.js'*.

Tang, Yuan. 2018a. "Autoplotly - Automatic Generation of Interactive Visualizations for Popular Statistical Results." *arXiv Preprint arXiv:1803.08424*.

—————. 2018b. *Autoplotly: Automatic Generation of Interactive Visualizations for Popular Statistical Results*. https://github.com/terrytangyuan/autoplotly.

Tang, Yuan, Masaaki Horikoshi, and Wenxuan Li. 2016. "Ggfortify: Unified Interface to Visualize Statistical Result of Popular R Packages." *The R Journal* 8 (2):478–89. https://journal.r-project.org/.

Wickham, Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer Science & Business Media. https://doi.org/10.1007/978-0-387-98141-3.