



ESTUDOS AVANÇADOS EM R: A ANÁLISE DE DADOS ESPACIAIS DE FORMA INTEGRADA

Vítor de Souza Silva¹, Gustavo da Silva Ferreira²

Introdução

Nos últimos anos o R (R Core Team, 2016) e em especial o RStudio, têm sido significativamente utilizados para análise de dados provenientes de diversas bases, os quais nem sempre possuem extensões que sejam simples de serem utilizados no R.

Neste sentido, pacotes de visualização, integração e manipulação foram desenvolvidos no intuito de facilitar a interface com os usuários que nem sempre são estatísticos, mas pesquisadores de diversas áreas.

Pacotes como o Leaflet (CHENG, KARAMBELKAR and XIE, 2017) e LeafletR (GRAUL, 2016), permitem a visualização e manipulação de dados espaciais de forma mais simples no R. Além disso, estes pacotes podem ser utilizados na forma de aplicações web, permitindo a disseminação das informações de diversas bases para públicos diversos.

Objetivos

Os principais objetivos deste projeto são a documentação do aprendizado de novos pacotes no R, que permitem apresentar dados espaciais de forma dinâmica, utilizando diversas bases e as integrando; apresentar tais resultados em uma aplicação web de domínio público e explorar os mais diversos tipos de documentos que podem ser gerados utilizando o software R.

Material e Método

Para a realização desse trabalho, foi necessário obter conhecimentos de pacotes que trabalham com dados espaciais, como: Leaflet, LeafletR, GoogleVis (GESMANN and CASTILLO, 2011) e SP (PEBESMA and BIVANT, 2005). Também foi necessário o estudo

¹ Escola Nacional de Ciências Estatísticas (ENCE), vitorssilva1@gmail.com

² Escola Nacional de Ciências Estatísticas (ENCE), gustavo.ferreira@ibge.gov.br

de pacotes de manipulações de dados como por exemplo o Dplyr (WICKHAM, FRANCOIS and MULLER, 2017), e pacotes que permitem a elaboração de apresentações dinâmicas no R, o Markdown (ALLAIRE, HORNER, MARTI and PORTE, 2017) e Knitr (XIE, 2018).

Resultados e Discussão

Resultados preliminares incluem a confecção de documentos de referências para o uso dos pacotes estudados, com a finalidade de serem usados como material didático em disciplinas de graduação em Estatística ou áreas correlatas.

Adicionalmente, foi elaborado um manual de referência para o uso do pacote Leaflet, o qual possui todos os passos necessários a fim de construir mapas no software R. A Figura 1 apresenta um exemplo de resultado fornecido pelo pacote, onde é possível visualizar no R, espacialmente e de forma interativa, informações coletadas de bases virtuais do Instituto Brasileiro de Geografia e Estatística (IBGE), a respeito da população total das Unidades Federativas do Brasil no ano de 2013.

Figura 1 – População Brasileira (mil pessoas) por estados no ano de 2013.



Fonte: Elaboração própria

A elaboração de mapas interativos como o exemplo apresentado na Figura 1 também exige a conversão de arquivos em formatos shapefile, geralmente utilizado em softwares de geoprocessamento, para outros formatos. Com o uso desta conversão, pode-se plotar as informações nos pacotes mencionados e usufruir das ferramentas de visualização que estes oferecem.



Conclusão

O uso de modernos pacotes desenvolvidos para R e RStudio tem se mostrado útil para uma melhor apresentação, análise e integração de dados espaciais. O próximo passo do projeto é a elaboração, com o uso do pacote shiny (CHANG, CHENG, ALLAIRE, XIE and MCPHERSON, 2017), de uma aplicação web de domínio público, onde usuários poderão escolher as variáveis a serem plotadas nos mapas resultantes. Os dados serão obtidos da Pnad, utilizando o pacote sidraR (SIQUEIRA, 2017), que possui uma API do sistema integrado de recuperação automática do IBGE, o SIDRA.

Referências

Graul, Christian (2016): leafletR: Interactive, Web-Maps Based on the Leaflet JavaScript library. R package version 0.4-0, <http://cran.r-project.org/package=leafletR>.

Hadley Wickham, Romain Francois, Lionel Henry and Kirill Müller (2017). dplyr: A Grammar of Data Manipulation. R package version 0.7.4. <https://CRAN.R-project.org/package=dplyr>

JJ Allaire, Jeffrey Horner, Vicent Marti and Natacha Porte (2017). markdown: 'Markdown' Rendering for R. R package version 0.8. <https://CRAN.R-project.org/package=markdown>

Joe Cheng, Bhaskar Karambelkar and Yihui Xie (2017). leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 1.1.0. <https://CRAN.R-project.org/package=leaflet>

Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.5. <https://CRAN.R-project.org/package=shiny>

Markus Gesmann and Diego de Castillo. Using the Google Visualisation API with R. The R Journal, 3(2):40-44, December 2011.

Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), <https://cran.r-project.org/doc/Rnews/>

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.5. <https://CRAN.R-project.org/package=shiny>

Yihui Xie (2018). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.20.