



Programação para Cientistas Sociais

2º Semestre de 2017

Professor: Thiago Moreira da Silva

Monitor de pós-graduação: ?

Programa da Disciplina

Este curso tem como objetivo auxiliar as(os) estudantes na construção de bases de dados, visualização de estatísticas descritivas e obtenção de novas informações. Ademais, ao longo do semestre letivo serão introduzidos alguns exames ancorados em redes sociais (*Twitter* e *Facebook*) e em análises de texto. Os conteúdos serão expostos no ambiente do *software R*.

A avaliação das(os) alunas(os) será baseada na aplicação das técnicas ensinadas em bancos de dados e projetos escolhidos pelas(os) próprias(os). Os trabalhos finais serão discutidos em sala de aula com a participação e o apoio das(os) colegas de curso. Não haverá provas. As leituras designadas na seção abaixo visam complementar os conteúdos discutidos em sala de aula e a sugestão é que as(os) alunas(os) as considerem **após** as lições introdutórias. O programa pode ser modificado ao longo do semestre de acordo com as demandas do corpo docente. A responsabilidade pela obtenção dos textos é das(os) estudantes.

Livros:

Bivand, Roger; Pebesma, Edzer; e Gómez-Rubio, Virgilio (2013). *Applied Spatial Data Analysis with R*. Switzerland: Springer.

Arnold, Taylor e Tilton e Lauren (2015). *Humanities Data in R: Exploring Networks, Geospatial Data, Images and Text*. Switzerland: Springer.

Luke, Douglas (2015). *A User's Guide to Network Analysis in R*. Switzerland: Springer.

Boehmke, Bradley (2016). *Data Wrangling with R*. Switzerland: Springer.

Wickham, Hadley (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Switzerland: Springer.

Baumer, Benjamin; Kaplan, Daniel; e Horton, Nicholas (2017). *Modern Data Science with R*. Florida: CRC Press.

Silge, Julia e Robinson, David (2017). *Text Mining with R: a Tidy Approach*. California: O'Reilly.

Wickham, Hadley e Golemund, Garret (2017). *R for Data Science: Import, Tidy, Transform, Visualize and Model Data*. Canada: O'Reilly.

Aula 1: Introdução ao ambiente R

- *Download* e instalação do *software*.
- Apresentação do *Rstudio*.

Aula 2: Análise Exploratória dos Dados (I)

- Importar e exportar dados no R.
- Introdução ao *Tidyverse*.
- Modos de sumarizar os dados.

Leitura base: caps 3, 5, 8 e 10 do livro *R for Data Science*, de Wickham (2017).

Endereço eletrônico: <https://www.rstudio.com/resources/webinars/data-wrangling-with-r-and-rstudio/>.

Aula 3: Análise exploratória dos Dados (II)

- *Datas*.
- *Strings*.
- Formas de combinar bases distintas (*joins* e *merges*)

Leitura base: caps. 9, 11 e 12 do livro *R for Data Science*, de Wickham (2017).

Endereço eletrônico: http://stat545.com/bit001_dplyr-cheatsheet.html.

Aula 4: Visualização dos Dados (I) – Gráficos básicos

- Histogramas.
- Gráficos de linha.
- Gráficos de barra.
- Gráficos de dispersão.

Endereço eletrônico: <http://www.r-graph-gallery.com/> (I). <http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html> (II).

Aula 5: Visualização dos dados (II) – Mapas

- Mapas no R.

Leitura base: cap. 14 do livro *Modern Data Science with R*, de Baumer *et. al.* (2017).

Endereços eletrônicos: <http://eriqande.github.io/rep-res-web/lectures/making-maps-with-R.html>(I). <http://www.milanor.net/blog/maps-in-r-plotting-data-points-on-a-map/>(II).

Aula 6: Apresentações e gráficos dinâmicos

- *Plotly*.
- *Dashboards*.
- Introdução ao *Shiny*.

Endereços eletrônicos: http://rmarkdown.rstudio.com/r_notebooks.html (I). http://rmarkdown.rstudio.com/ioslides_presentation_format.html (II). <https://github.com/ramnathv/rMaps> (III). <https://plot.ly/r/> (IV). <http://rmarkdown.rstudio.com/flexdashboard/> (V). <https://shiny.rstudio.com/> (VI).

Aula 7: Programação funcional

- *Loops*.
- Funções.

Leitura base: caps 15, 16 e 17 de *R for Data Science*, de Wickham (2017)

Aula 8: Importação e raspagem de múltiplas bases

- Importação de múltiplas bases.
- *Download* de múltiplas bases.
- *Web scraping*.

Endereço eletrônico: <https://www.analyticsvidhya.com/blog/2017/03/beginners-guide-on-web-scraping-in-r-using-rvest-with-hands-on-knowledge/>

Aula 9: Análise de redes (I)

- Captação de dados do *Twitter*.
- Captação de dados do *Facebook*.

Leitura base: cap. 15 do livro *Modern Data Science with R*, de Baumer *et. al.* (2017).

Endereços eletrônicos: [https://www.credera.com/blog/business-intelligence/twitter-analytics-using-r-part-1-extract-tweets/\(I\)](https://www.credera.com/blog/business-intelligence/twitter-analytics-using-r-part-1-extract-tweets/(I)). [https://daveatang.org/muse/2013/04/06/using-the-r_twitter-package/\(II\)](https://daveatang.org/muse/2013/04/06/using-the-r_twitter-package/(II)).

Aula 10: Análise de redes (II)

- Visualização de redes (*grapho*).

Leitura base: caps 4, 5 e 6 do livro *A User's Guide to Network Analysis*, de Luke (2013).

Aula 11: Análise de texto

- Frequência dos termos.
- Relações entre as palavras.
- Análise de sentimentos.

Leitura base: caps 1 a 4 do livro *Text Mining with R*, de Silge e Robinson (2017).

Aula 12: Apresentação dos trabalhos finais

Aula 13: Apresentação dos trabalhos finais