

Graduação & Pós-Graduação em Ciência Política

Disciplina - Introdução à captura de dados com a linguagem de programação em R.

Carga Horária: 30

Coordenador: Felipe Borba

Professor: Steven Dutt-Ross

Semestre: 2º Ano: 2022

PLANO DE ENSINO

Ementa

Disciplina síncrona mediada por tecnologia que aborda a Linguagem de Programação R para captura de dados. Objeto: Variáveis e Banco de dados; Banco de dados de dados não estruturado (texto). Acesso a API, Visualização de dados, e raspagem de dados.

Objetivos

O objetivo central do curso é promover a capacidade de criar scripts para extrair dados da internet utilizando a linguagem de programação R.

Avaliação

A avaliação será sobre o trabalho final: estudantes deverão realizar uma pesquisa empírica a partir da coleta de dados de um site ou do twitter. A nota final será o resultado do trabalho final com o código R em anexo.

Programação

1. **Introdução ao R.** Introdução ao R & ao Tidyverse.
2. **Extração de dados do twitter.** Definição de API. Como acessar a API do twitter.
3. **Introdução ao ggplot.** Introdução à gramática de gráficos.
4. **Introdução ao ggplot - parte 2.** Construindo visualizações de dados.
5. **Nuvens de palavras, redes e outras análises de dados.**

6. Nuvens de palavras, redes e outras análises de dados - parte 2.

7. Introdução à raspagem de dados com R Introdução ao método para raspar dados de sites. Inspeccionando elementos.

8. Introdução à raspagem de dados com R - parte 2. Exercício aplicado: Aqui vamos raspar as informações do site do GIEL (<http://giel.uniriotec.br/>) e do site da Escola de Ciência Política da UNIRIO (<http://www.unirio.br/ccjp/escola-de-ciencia-politica>)

Bibliografia básica

BRUCE, Andrew; BRUCE, Peter. Estatística Prática para Cientistas de Dados. Alta Books, 2019.

WICKHAM, H. and GROLEMUND, G. *R for data science: import, tidy, transform, visualize, and model data.* O'Reilly Media, Inc.", 2016.

WICKHAM, H. ggplot2: Elegant graphics for data analysis. Springer International Publishing, 2016.

Bibliografia complementar

AQUINO, J. A. de *R para cientistas sociais* – Ilhéus, BA: EDITUS, 2014. 157 p.: il

KEARNEY, M. W. *rtweet: Collecting and analyzing Twitter data.* Journal of open source software, v. 4, n. 42, p. 1829, 2019.

PUSCHMANN, C.; BURGESS, J. *The politics of Twitter data.* 2013.

DUTT-ROSS, S. *Como comprar um carro usado com o R.* Estatística é com R! UFF. Disponível em <http://www.estadisticacomr.uff.br/?p=869> Acesso em 12/07/2022.

DUTT-ROSS, S.; CRUZ, B. de P. A. *Análise Quantitativa de Textos: Apresentação e Operacionalização da Técnica via Twitter.* Administração: Ensino e Pesquisa, v. 22, n. 1, 2021.

FILGUEIRA, L. *WordCloud no R* – Estatística é com R! UFF. Disponível em <http://www.estadisticacomr.uff.br/?p=322> Publicado em julho 2, 2016. Acesso em 12/07/2022.

Para saber mais, visite os sites:

<http://giel.uniriotec.br/> &

<https://dataunirio.github.io/>