

PERÍODO: 1º semestre de 2020
 ELETIVA: Introdução ao R aplicado à ciência de dados
 PROFESSOR(ES): André Arruda Villela
 TUTOR: Diogo Tavares Robaina
 CARGA HORÁRIA: 60 horas

PLANO DE ENSINO

Competências:

- Compreender o funcionamento do pacote estatístico R.
- Descrever bancos de dados por meio de gráficos e tabelas.
- Manipular dados de modo eficiente e rápido.
- Analisar dados usando o R.
- Aprender os rudimentos de programação no R.

<u>Disciplina</u>	<u>Carga horária</u>
2. Introdução ao R aplicado à ciência de Dados	60 h/a

<u>Área</u>	<u>Subárea</u>
Ciência de dados	Ciência de dados/ Estatística

Ementa

Instalando o R e Softwares adjacentes. Gráficos em R. Carregando e manipulando dados em R. Analisando dados com R. Programação e Automação em R.

MÓDULO 1

<u>Conteúdo programático mínimo</u>	<u>Objetivos de aprendizagem</u>
1. Instalando o R e Softwares adjacentes 1.1 O que é o R 1.2 Instalando o R e os softwares auxiliares 1.3 Customizando e entendendo o R	<ul style="list-style-type: none"> • Apresentar o R e softwares adjacentes. • Instalar e configurar softwares adjacentes ao R. • Apresentar os primeiros rudimentos sobre o funcionamento do software

MÓDULO 2

<u>Conteúdo programático mínimo</u>	<u>Objetivos de aprendizagem</u>
2. Gráficos em R. 2.1 O R Básico. 2.2 Gráficos no R: variáveis individuais. 2.3 Gráficos no R: cruzamentos	<ul style="list-style-type: none"> • Entender os comandos básicos do R. • Usar as estruturas básicas do software. • Produzir gráficos utilizando o R. • Identificar quais tipos de gráficos combinam melhor com cada tipo de dados.

MÓDULO 3	
<u>Conteúdo programático mínimo</u>	<u>Objetivos de aprendizagem</u>
3. Carregando e manipulando dados em R. 3.1 Manipulando os dados, parte 1. 3.2 Manipulando os dados, parte 2. 3.3 Arrumando bancos de dados. 3.4 Carregando e descarregando dados no R	<ul style="list-style-type: none"> • Manipular dados usando o R • Compreender o conceito de tidy data • Carregar bancos de dados no R • Salvar bancos de dados para uso em outros softwares.

MÓDULO 4	
<u>Conteúdo programático mínimo</u>	<u>Objetivos de aprendizagem</u>
4. Analisando dados com R. 4.1 Análise univariada. 4.2 Análise Multivariada. 4.3 Análise de Regressão.	<ul style="list-style-type: none"> • Utilizar os comandos principais para análise univariada em R • Utilizar os comandos principais para a análise multivariada no R

MÓDULO 5	
<u>Conteúdo programático mínimo</u>	<u>Objetivos de aprendizagem</u>
5. Programação e Automação em R. 5.1 Programação em R. 5.2 Vetores e Matrizes em R. 5.3 O que mais fazer com o R?.	<ul style="list-style-type: none"> • Aprender sobre como o computador funciona internamente: métodos de calculo e organização da memoria e execução de comandos. • Compreender a automatização de tarefas • Utilizar as principais ferramentas para análise de matrizes em R

Bibliografia Básica

- JAMES, Gareth, et al. An introduction to statistical learning. Vol. 112. New York: springer, 2013.
- MAINDONALD, John, and John Braun. Data analysis and graphics using R: an example-based approach. Vol. 10. Cambridge University Press, 2006.
- REYES, Jose Manuel Magallanes. Introduction to Data Science for Social and Policy Research. Cambridge University Press, 2017.
- TEAM, R. Core. "R language definition." Vienna, Austria: R foundation for statistical computing (2000).

Bibliografia Complementar

- WICKHAM, Hadley, and Garrett Golemund. R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc.", 2016.
- WICKHAM, Hadley, and Romain Francois. "dplyr: a grammar of data manipulation, 2013." URL <https://github.com/hadley/dplyr>. version 0.1.[p 1] (2017).
- WICKHAM, Hadley. "Tidy data." Journal of Statistical Software 59.10 (2014): 1-23.
- WICKHAM, Hadley. Advanced r. Chapman and Hall/CRC, 2014.
- WICKHAM, Hadley. ggplot2: elegant graphics for data analysis. Springer, 2016.