



---

## **Plano de Ensino**

EST032 - PACOTES ESTATÍSTICOS

Ano: 2018 Semestre: 1

### EMENTA

Prática computacional, motivada por problemas reais e usando linguagem livre para análise de dados, com atividades de laboratório para introdução de técnicas e conceitos básicos que permitam a iniciação com a prática da Estatística. Compreenderá: fundamentos da linguagem R; introdução à programação em R; obtenção e limpeza de dados; manipulação de grandes bases de dados.

### CONTEÚDO

1. Fundamentos da linguagem R: Introdução aos comandos básicos; manipulação e exploração de diferentes tipos de objetos de dados (vetores, matrizes, data frames); operações básicas; importação e exportação de dados; uso de scripts para organizar comandos.
2. Princípios de programação em R: Estruturas de controle: if-else, loops em for, loops em while, repeat-next-break; funções em R; padrões de codificação; debugging: diagnóstico dos problemas, uso de ferramentas básicas para debugging.
3. Extração de conjuntos de dados: Leitura de diferentes tipos de arquivos; apresentação dos sistemas de armazenamento de dados mais comuns; ferramentas para extração de dados da web ou de bancos de dados.
4. Limpeza de dados: Recodificação e conversão de dados; manipulação de caracteres: normalização de strings, problemas de codificação de caracteres; detecção, localização e correção de erros; imputação de dados: modelos básicos de imputação numérica, imputação hot deck, imputação kNN, ajuste de valor mínimo.
5. Tópicos especiais: Manipulação de grandes conjuntos de dados no R; visualização de dados; comandos de otimização; pacotes especiais: R Shiny e outros; geração de relatório com R Markdown, knitr ou similar; outros tópicos.

### BIBLIOGRAFIA

- FERREIRA, D. F. Estatística computacional utilizando R. Notas de aula. Lavras: UFLA, 2014.  
DALGAARD, P. Introductory statistics with R. New York: Springer, 2002.

### BIBLIOGRAFIA COMPLEMENTAR

- DE JONGE, E.; VAN DER LOO, M. An introduction to data cleaning with R. Statistics Netherlands, The Hague, 2013.  
PRAJAPATI, V. Big data analytics with R and hadoop. Packt Publishing, 2013.  
VAN DEN BROECK, J. et al. Data cleaning: detecting, diagnosing, and editing data abnormalities. PLoS Med, v. 2, n. 10, p. e267, 2005.