



Cruzeiro do Sul Virtual
Educação a distância

Apresentação da Disciplina



Programação Estatística

Objetivo Geral

Estudo de linguagem de programação estatística com suporte à visualização de dados, no âmbito da área de Ciência de Dados.





Conteúdo Programático das Unidades

Unidade 1: Conceitos Gerais

Conceitos básicos sobre como trabalhar com a linguagem R.

Unidade 2: Variáveis, tipos de Dados e Estruturas de Controle

Criação e atribuição de valores para variáveis; tipos de dados “primitivos” ou nativos | Estruturas de controle, *if*, *if/else*, laços de repetição.

Unidade 3: Vetores, Matrizes, Fatores e Funções

Utilização de estruturas de dados convencionais unidimensionais e bidimensionais | Apresentação de fatores | Funções: explicação da importância, regras de criação.

Unidade 4: Manipulação de Dados

Permitir que o aluno importe dados de arquivos e manipule os mesmos conforme sua necessidade.

Unidade 5: Visualização de Dados

Aprender a elaborar gráficos em R com o comando *plot* e o pacote *ggplot2*.

Unidade 6: Estudo de Caso

Permitir que o aluno tenha uma experiência simulada com problemas reais, utilizando a linguagem R.





Contextualização à disciplina

Contexto Acadêmico:

O ambiente de programação R é uma opção gratuita aos softwares estatísticos, pois estes muitas vezes possuem alto valor de licença. O R está sob licença GNU, ou seja, é código fonte aberto, possui uma comunidade ativa que suporta à linguagem.

É por meio de tal comunidade que o R possui vários pacotes para se trabalhar com grandes volumes de dados desde a sua coleta, armazenamento, processamento e visualização, além de poder realizar técnicas estatísticas clássicas e modernas.

Contexto Profissional:

Em um mundo onde as pessoas estão cada vez mais conectadas, gerando grande quantidade de dados diariamente ao consumir, compartilhar e criar conteúdos digitais, faz-se necessário tratar tais dados de maneira a gerar informações precisas. Contudo, quanto maior a quantidade de dados, mais difícil é processá-los.

Nesse contexto, a utilização de linhagens de programações específicas para atender tal demanda se faz necessária como, por exemplo, a linguagem R, criada para ser uma opção à ferramenta de apoio ao trabalho com estatística; todavia, atualmente possui vários recursos para compilar grandes volumes de dados e gerar informações coerentes, portanto, a uma linguagem de programação para Ciência de Dados.



Bibliografia Básica e Complementar

Básica

THE R PROJECT. *What is R?* [20--]. Disponível em: <<https://www.r-project.org/about.html>>. Acesso em: maio 2019.

WICKHAM, H. *R for Data Science*. [S.l.]: O'Reilly, 2017. Disponível em: <<https://r4ds.had.co.nz>>. Acesso em: ago. 2019.

YIN, R. K. *Estudo de caso: planejamento e métodos*. [S.l.]: Bookman, 2015.

