



CURSO DE GRADUAÇÃO EM SAÚDE COLETIVA

PLANO DE DISCIPLINA

Disciplina: Tópicos em Análise de Dados II - Software R	Código: NSC048
Carga Horária: 45h	

EMENTA:

Curso básico ao Ambiente R, que é uma linguagem de programação estatística abrangente, além de ser um software livre. Os alunos serão apresentados às principais funções implementadas nos pacotes básicos (core) do programa, importação e exportação de dados, bem como ao gerenciamento de dados. Além do R, os alunos também serão apresentados ao RStudio.

OBJETIVO GERAL:

Apresentar os elementos básicos do software R e do RStudio, incluindo linguagem, operações básicas, criação de objetos, importação e exportação de dados, manipulação de dados da Saúde e construção de tabelas e gráficos.

OBJETIVOS ESPECÍFICOS:

A disciplina tem o objetivo de capacitar o aluno na higienização e manipulação dos dados de Saúde e Populacionais, além da produção e interpretação da Informação, utilizando software estatístico livre e gratuito.

CONTEÚDO PROGRAMÁTICO:

Apresentação do R e RStudio
Instalação do R e RStudio
Apresentação de bases de dados e dicionários
Noções básicas e R como calculadora
Instalação e leitura de pacotes
Mudança de diretório
MANIPULAÇÃO DE BASE DE DADOS

- Importação de base de dados
- Filtrar base de dados
- Selecionar variáveis
- Conferir classificação das variáveis
- (Re)classificar variáveis
- Classificar datas
- Recodificar categorias
- Exportar base de dados
- Renomear variáveis
- Lidando com dados faltantes
- Manipulando a variável Idade e datas
- Criação de nova variável (exemplo: faixa etária)



VISUALIZAÇÃO

- Produção de Tabelas
- Gráficos

FREQUÊNCIA:

Será considerada para aprovação a frequência mínima de 75% (setenta e cinco por cento) da carga horária da disciplina.

METODOLOGIA (PROCEDIMENTOS DE ENSINO):

As aulas ocorrerão no laboratório de informática onde os alunos serão apresentados aos softwares R e RStudio e suas funcionalidades; Além das aulas expositivas, serão corrigidas as listas de exercícios e esclarecimento das dúvidas.

AVALIAÇÃO:

Serão aplicadas 2 avaliações, sendo 1 trabalho individual e um 1 trabalho em grupo.

se Média Final (MF) $\geq 5,0$ o aluno será aprovado

se Média Final (MF) $< 5,0$ o aluno será reprovado

BIBLIOGRAFIA BÁSICA:

MEDRONHO, R. A. *et al.* (ed.). Epidemiologia. 2. ed. São Paulo: Atheneu, 2009.

BIBLIOGRAFIA COMPLEMENTAR:

- VENABLES, William N. *et al.* An introduction to R. Notes on R: a programming environment for data analysis and graphics: version 4.2.1 (2022-06-23). 2022. Disponível em: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>. Acesso em: 11 out. 2022.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. URL <https://www.R-project.org/>.
- Saldanha, Raphael de Freitas, Bastos, Ronaldo Rocha e Barcellos, Christovam. Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS). Cadernos de Saúde Pública [online]. 2019, v. 35, n. 9, e00032419. Disponível em: <https://doi.org/10.1590/0102-311X00032419>.

Observação: a Bibliografia deve estar disponível com pelo menos 3 (três) exemplares nas Bibliotecas da UFRJ. Excepcionalmente, é permitida a utilização de Bibliografia que não atenda a este critério, desde que disponibilizada à turma pelo (a) professor (a).



Data	Conteúdo	Professor
Aula 1 – 05/08	Apresentação da disciplina Introdução ao RStudio e Manipulação de bases de dados Apresentação do R e RStudio Noções básicas e R como calculadora Instalação e leitura de pacotes Mudança de diretório Introdução ao SIVEP-GRIPE, ficha de notificação e dicionário; Importação de base de dados Selecionar variáveis Filtrar base de dados Conferir classificação das variáveis	
Aula 2 – 12/08	Manipulação de Base de Dados (Re)classificar variáveis Classificar datas Recodificar categorias Renomear variáveis Exportar base de dados (csv) Salvar RData	
Aula 3 – 19/08	Exercícios	
Aula 4 – 26/08	Manipulação de Base de Dados Trabalhar com <i>missing</i> Reagrupar categorias Criar faixas de valores Manipular variável Idade Criação de nova variável – categorização	
Aula 5 – 02/09	Exercícios	
Aula 6 – 09/09	Análise Exploratória Tabelas de frequência simples e contingência	
Aula 7 – 16/09	Análise Exploratória Gráficos (pacotes: <i>ggplot2</i>) Agrupar linhas com comando <i>group_by</i> (pacote: <i>dplyr</i>) Criar variáveis mês, ano e semana epidemiológica (pacote: <i>lubridate</i>)	
Aula 8 - 23/09	SIAC	
Aula 9 - 30/09	Revisão/dúvidas Definir: Grupo Tema Pergunta	
Aula 10 – 07/10	Análise Exploratória Gráficos (pacotes: <i>ggplot2</i>)	



	Agrupar linhas com comando <code>group_by</code> (pacote: <i>dplyr</i>) Criar variáveis mês, ano e semana epidemiológica (pacote: <i>lubridate</i>)	
Aula 11 – 14/10	Exercícios	
Aula 12 – 21/10	Manipulação de dados Trabalhar com a variável IDADE no SIM e SINAN Join (similar ao merge) - usando pacote <i>dplyr</i> Diferença entre datas	
Aula 13 – 28/10	Aula de dúvidas	
Aula 14 – 04/11	Aula de dúvidas	
Aula 15 – 11/11	Apresentação do trabalho	