

Universidade Federal da Paraíba Centro de Ciências Sociais Aplicadas Programa de Pós-graduação em Economia do Setor Público

#### PLANO DE DISCIPLINA

# Análise de Dados e Documentos Dinâmicos

DOCENTE: ANTONIO VINÍCIUS BARBOSA

**Objetivo:** Esta disciplina tem como proposta introduzir as principais técnicas computacionais para a leitura, manipulação, visualização e elaboração de documentos dinâmicos de maneira rápida e eficiente. Além disso, este curso tem como objetivo ajudar a criar documentos reprodutíveis, tais como relatórios e apresentações. Ao final do curso, espera-se que o aluno seja capaz de: i) Ler e manipular dados; ii) Computar estatísticas descritivas dos dados; iii) Gerar visualização dos dados; e iv) Criar relatórios, apresentações e páginas web de maneira dinâmica.

**Ementa:** Os principais tópicos do curso incluem:

Introdução e conceitos básicos de análise de dados

Introdução ao R

Leitura e manipulação de dados

Visualização de dados

Elaboração de documentos dinâmicos

Criando artigos científicos

Conteúdo Programático: Os principais tópicos do curso incluem:

# Introdução e conceitos básicos de análise de dados

Conceito de análise de dados; Novas competências e paradigmas; O uso de linguagens de programação; Exemplos de análise de dados.

### Introdução ao R

A linguagem R; Ambientes de desenvolvimento integrado (IDE); RStudio e Posit Cloud; Instalação e gerenciamento de pacotes; Operações básicas; Estrutura de dados

## Leitura e manipulação de dados

Leitura de dados; Criação de variáveis; Ordenação dos dados; Junção (merge) de dados; Manipulação de variáveis contendo datas; Manipulando de variáveis contendo texto; *Tidy data* 

# Visualização de dados

Gráficos básicos: gráfico de barras, gráfico de pizza, histogramas, densidades, box plots, gráfico de dispersão, gráfico de linhas, correlogramas, mosaicos. Introdução ao pacote ggplot2; Gráficos interativos; Visualizando dados em mapas.

### Elaboração de documentos dinâmicos

Introdução ao Markdown; Códigos integrados a documentos; Introdução ao Quarto; Elaboração de relatórios, tabelas, apresentações e páginas web.

## Criando artigos científicos

Artigos acadêmicos; Capítulos e seções; Layout; Referências bibliográficas

■ **Metodologia:** A metodologia adotada é baseada em aulas expositivas do conteúdo previsto, na apresentação e discussão de textos complementares, e na resolução de listas de exercícios. As aulas serão interativas, com a participação dos alunos expressando opiniões e apresentando dúvidas durante sua realização. A principal linguagem utilizada nesta disciplina será o R (r-project.org), juntamente com o ambiente de desenvolvimento integrado RStudio (https://posit.co/). As atividades práticas de laboratório cobrirão os assuntos discutidos durante a exposição do conteúdo e introduzirá tópicos adicionais.

#### Literatura sugerida:

- BRAUN, W. John; MURDOCH, Duncan J. A First Course in Statistical Programming with R. Cambridge University Press, 2016.
- DAYAL, Vikram. An Introduction to R for Quantitative Economics: Graphing, Simulating and Computing, Springer Briefs in Economics, Springer (India), 2015.
- ROCHA, M.; FERREIRA, F. G. Análise e Exploração de Dados com R. FCA, 1ª Edição. 2017.
- TEETOR, Paul. R cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics. O'Reilly Media, Inc., 2011.
- WICKHAM, Hadley; GROLEMUND, Garrett. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media, Inc., 2016.
- CHAMBERS, John. Software for Data Analysis: Programming with R. Springer Science Business Media, 2008.
- KABACOFF, R. I.; ACTION, R. In. Data Analysis and Graphics with R. 2011.
- MATLOFF, Norman. The Art of R programming: A Tour of Statistical Software Design. No Starch Press, 2011.
- XIE, Yihui et al. Tufte: Tufte's styles for R Markdown documents. In: R Markdown: The Definitive Guide. Vienna, Austria: Chapman and Hall/CRC, 2019. p. 1-45.
- RStudio Cheat Sheets (https://www.rstudio.com/resources/cheatsheets/)
- √ Os alunos serão informados ao longo do curso em relação à referências complementares.
- Material: Todo o material utilizado ao longo do curso (slides, scripts, tutoriais e listas de exercícios) serão disponibilizados na página do curso.
- **Comunicação:** Toda a comunicação entre alunos e professor sobre os elementos do curso (atendimento, dúvidas, listas, soluções e notícias) será realizada exclusivamente através da página da disciplina no **Sigaa** (https://sigaa.ufpb.br/sigaa/).
- Avaliação: A avaliação na disciplina será dada através de listas de exercícios e de uma avaliação final, de acordo com a seguinte função:

NF = f(Listas de Exercício, Projeto).

A nota final será uma média ponderada, de acordo com os seguintes pesos:

Descrição	Peso
Listas de Exercício	40%
Projeto	60%

Atendimento Extraclasse: Horário de atendimento às quartas-feiras, 09h-11h, e às sextas-feiras, 14h-16h, na Sala 05 dos professores do CCSA. O agendamento deverá ser feito exclusivamente através da página https://calendly.com/antonio\_barbosa