

Coleta e análise de dados secundários

Professor



Frederico Bertholini



Atendimento: por e-mail



<https://www.lapcipp.org/>



frederico.bertholini@unb.br

Info do Curso



Material: goo.gl/5xVnvn



Segundas e Quintas



19h-22h(seg) 9h-12h(qui)



Laboratório 104

Objetivos

Avançar os conhecimentos dos participantes na linguagem R de programação estatística priorizando a coleta e análise de dados secundários. Desenvolver a habilidade dos alunos em coletar dados estruturados e não estruturados e organizá-los de forma inteligível para extrair intuições acerca de fenômenos sociais relevantes, bem como apresentá-los de forma visualmente atraente.

Ementa

O Universo Tidyverse; Fontes de dados secundários e estratégias de coleta; Estruturas de dados individuais e agregados; Limpeza de dados; Trabalhando com dados amostrais; Análises descritivas multivariadas; Versionamento e publicação no GitHub; Visualizações interativas e relatórios no RMarkdown.

Metodologia

Aulas expositivas/práticas em laboratório de informática. Em cada aula novos conceitos são apresentados aos alunos que devem aplicar esses conceitos em exercícios envolvendo coleta e análise de dados.

Avaliação

A nota final da disciplina será calculada como uma média de uma Prova (50%) e um Trabalho Final (50%).

- Prova: Avaliação prática presencial sobre pontos relacionados ao programa.
- Relatório Final: Apresentação interativa de dados em HTML utilizando a ferramenta Rmarkdown e contendo código R utilizado. Os temas serão escolhidos pelos alunos e a rubrica com os itens de avaliação do relatório será disponibilizada na etapa de definição dos temas.

Bibliografia básica

Baquero, O. (2017). *Manipulação e visualização de dados no R*

Wickham, H., Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc."

Coletânea de material do curso-r

Docente (mini-currículo)

Frederico Bertholini é professor do Instituto de Ciência Política da Universidade de Brasília (IPOL/UnB). Tem doutorado em Administração (Instituições Políticas e Governo) pela Escola Brasileira de Administração Pública e Governo da Fundação Getúlio Vargas (FGV/EBAPE), mestrado em Demografia e Planejamento Urbano pela Escola Nacional de Ciências Estatísticas (ENCE / IBGE), Especialização em Análise de Políticas Públicas na UFRJ, é Bacharel em Relações Internacionais e cursou Economia. Foi Gerente de Estudos Regional e Metropolitano na Companhia de Planejamento do Distrito Federal (Codeplan-DF) entre 2015 e 2018, pesquisador associado do IPGLab/FGV de 2011 a 2018 e esteve de 2013 a 2014 como Pesquisador Visitante no Departamento de Política da Universidade de Nova Iorque. Entre seus interesses de pesquisa estão o efeito das instituições políticas sobre os resultados das políticas públicas; relações Executivo-Legislativo; gerência de coalizões políticas em presidencialismos multipartidários; incluindo a aplicação de métodos quantitativos e qualitativos sobre a análise política e de políticas públicas.

Programação

Lista de itens das aulas e dos tópicos correspondentes que serão abordados durante o semestre. A programação poderá sofrer alguns ajustes ao longo do curso, consulte sempre o professor sobre mudanças na programação.

Semana 1	20.set	O Universo tidyverse
Semana 2	24.set	Fontes de dados secundários e estratégias de coleta
	27.set	Estruturas de dados individuais e agregados
Semana 3	01.out	Limpeza de dados
	04.out	Trabalhando com dados amostrais
Semana 4	08.out	Análises descritivas multivariadas
	11.out	Versionamento e publicação no GitHub
Semana 5	15.out	Visualizações interativas e relatórios no RMarkdown
	18.out	Prova
Semana 6	22.out	Apresentação de trabalhos finais