

Métodos Quantitativos 101

Métodos Quantitativos
3º quadrimestre de 2025

Prof. Dr. Ricardo Ceneviva
ceneviva@ufabc.edu.br

Monitor: TBA
nonono@aluno.ufabc.edu.br

Objetivos do Curso

Este é um curso de introdução à análise de dados para alunos de pós-graduação em Políticas Públicas. O objetivo principal do curso é fornecer uma introdução básica de métodos estatísticos para a análise de dados sociais, políticos e econômicos. A análise de dados, e a estatística que a fundamenta, podem ser definidas sucintamente como a arte de fazer conjecturas numéricas sobre questões intrigantes (Freedman et al, 2009). Questões que há muito despertam o interesse das Ciências Sociais, tais como: existe alguma relação entre o tipo de regime político e o nível de desenvolvimento econômico dos países? Candidatos que concorrem à reeleição no exercício do cargo político tem mais chances de sucesso eleitoral? Mulheres têm menores rendimentos no mercado e trabalho se comparadas a homens de mesma escolaridade e que ocupam cargos semelhantes? Estas são perguntas que podem ser respondidas com o auxílio de métodos e técnicas estatísticas de pesquisa.

Neste curso serão apresentados os princípios básicos de pesquisa quantitativa em ciências sociais, com ênfase especial em Políticas Públicas e Projetos Sociais. Como definir e mensurar conceitos, tipos de dados e variáveis, construir e interpretar gráficos e tabelas, além de noções usualmente utilizadas nas Ciências Sociais como médias, variação e correlação serão discutidos com cuidado ao longo desse curso.

Ao final do quadrimestre espera-se que os alunos sejam capazes de produzir e analisar seus dados, além de ler e criticar textos que apliquem as técnicas discutidas nas aulas. Para tanto, o curso combina aulas expositivas e sessões práticas no laboratório, onde os alunos serão familiarizados com o manuseio de *softwares* de análise quantitativa de dados. Os comandos desenvolvidos estarão disponíveis no seguinte endereço de *github*: <https://github.com/rceneviva/>. O conteúdo do curso não é exaustivo, contudo, pretende-se encorajar os alunos a continuar estudando métodos mais sofisticados para a análise de dados.

A quem esse curso se destina?

O foco do curso são os alunos do Programa de Pós-graduação em Políticas Públicas da UFABC e os exemplos e casos analisados nas sessões teóricas e práticas serão retirados, principalmente, de artigos e trabalhos do campo de Política Públicas; porém, também serão lidos e discutidos trabalhos de disciplinas afins, tais como: ciência política, sociologia e economia política. Embora o curso tenha um conteúdo bastante técnico, sua abordagem será conceitual e aplicada, em vez de técnica e teórica. Isto é, as aulas expositivas e sessões de laboratório se concentrarão na intuição básica por trás de cada um dos modelos estudados e na sua aplicação prática nas ciências sociais em geral e em Políticas Públicas, em especial. O curso não tem pré-requisitos.

Avaliação

Como se trata de um curso aplicado, a participação em sala de aula terá grande importância. Além das atividades em laboratório e das listas de exercícios, que equivalem a 50% da nota final, os alunos serão encorajados a coletar, analisar e apresentar seus próprios dados ou a replicar algum trabalho ou artigo já publicado. Para tanto, será proposto um projeto final que consistirá em uma apresentação de uma análise de algum banco de dados escolhido pelo aluno, ou no desenvolvimento de um projeto de pesquisa original que envolva análise de dados.

1. Os alunos deverão entregar seis listas de exercícios ao longo do curso. Não serão aceitas listas de exercícios entregues fora do prazo.
2. Projeto final valendo 50% da nota final do aluno.

Software para Análise de Dados

Neste curso, tanto nas sessões de laboratório como para a realização das listas de exercícios, serão usados *softwares* para a análise de dados. Hoje, há vários programas computacionais concebidos especialmente para a análise estatística de dados sociais, políticos e econômicos, além de outros de uso mais geral. Parte importante desse curso destina-se à introdução ao uso de *softwares* para a análise estatística de dados. Espera-se que ao final do semestre o aluno seja capaz de realizar operações básicas para a implementação, análise, diagnóstico e apresentação gráfica dos modelos estatísticos estudados ao longo do curso. Dentre as várias opções de *softwares* para a análise de dados, as sessões práticas em laboratório fornecerão treinamento básico para o manuseio de dois programas computacionais: R e Stata. Caberá ao aluno escolher a opção que mais lhe convém.

Cabe ressaltar que *softwares* e aplicativos para análises estatísticas estão sempre sendo modificados e aprimorados, de modo que qualquer coisa que se diga hoje a respeito deles pode estar incorreta quando uma próxima versão do aplicativo for disponibilizada aos usuários.

R

O R é uma linguagem para computação estatística e gráficos. Uma das grandes vantagens do R é que se trata de um *Software Livre*, que pode ser baixado gratuitamente na página oficial do Projeto R na internet: <http://cran.r-project.org/>. Recentemente um artigo do *New York Times* (“Data Analysts Captivated by R’s Power”, 6 de janeiro de 2009) caracterizou o R como:

a popular programming language used by a growing number of data analysts inside corporations and academia. It is becoming their lingua franca [...] whether being used to set ad prices, find new drugs more quickly or fine-tune financial models. Companies as diverse as Google, Pfizer, Merck, Bank of America, the InterContinental Hotels Group and Shell use it. [...] “The great beauty of R is that you can modify it to do all sorts of things,” said Hal Varian, chief economist at Google. “And you have a lot of prepackaged stuff that’s already available, so you’re standing on the shoulders of giants.”

Como mencionado, o R não é apenas um pacote estatístico; trata-se de uma linguagem para computação estatística e gráfica. Como tal, é muito mais flexível, rápido e poderoso do que os demais *softwares* de análise de dados comerciais disponíveis no mercado, como o SPSS, Stata ou SAS, além de ser gratuito. O R é também mais difícil de ser usado, já que exige alguma noção de programação. Há uma imensa variedade de materiais de apoio, apostilas, e recursos (em português e inglês) destinados à aprendizagem do R disponíveis gratuitamente na Internet. Alguns desses recursos serão comentados e usados nas sessões de laboratório.

Python

Python é uma linguagem de programação de uso geral, livre e de código aberto, com ampla adoção na academia, no setor público e no mercado. Para análise de dados, destaca-se o ecossistema científico composto por NumPy e pandas (manipulação e estruturas de dados), SciPy (métodos numéricos), statsmodels (modelos estatísticos), scikit-learn (aprendizado de máquina), PyMC (Bayes) e bibliotecas de visualização como matplotlib e plotly. O uso combinado com Jupyter Notebooks/JupyterLab favorece a documentação reproduzível de análises, integrando texto, código e resultados em um único arquivo. Materiais didáticos atualizados estão amplamente disponíveis em formato aberto.

Python difere de R por ser uma linguagem de propósito mais amplo, com forte integração a sistemas, APIs e fluxos de produção (ETL, web scraping, automação). Em contrapartida, R oferece um “dialeto” estatístico e de visualização muito coeso. No curso, Python pode ser usado como alternativa (ou complemento) a R e Stata, com foco em boas práticas de programação e reproduzibilidade por notebooks e scripts versionados. A literatura de ensino em estatística recomenda integrar competências computacionais ao currículo desde o início, o que justifica a inclusão de Python ao lado de R.

Para instalação, recomenda-se o uso de ambientes gerenciados (conda/mamba ou venv) e JupyterLab ou VS Code como IDEs. A organização do material em repositórios Git permite

versionamento, revisão por pares e melhor rastreabilidade de resultados. Como em qualquer software, versões e pacotes evoluem continuamente; por isso, procedimentos de instalação e dependências devem ser documentados no início do curso.

Stata

O Stata é um software comercial muito popular para análise estatística de dados. Este *software* foi criado em 1985 pela StataCorp. O Stata é largamente empregado, tanto por instituições do setor privado como na academia, particularmente, por pesquisadores que lidam com gerenciamento e manipulação de bancos de dados, o que engloba áreas tão diversas como: biologia, epidemiologia , economia, sociologia, demografia e ciência política, entre outras.

Os recursos do Stata incluem não apenas o gerenciamento e manipulação de bancos de dados, mas também, análises estatísticas, gráficos, simulações, além de oferecer ferramentas de programação personalizada para execução de tarefas específicas. O Stata oferece uma interface gráfica o que permite ao usuário a análise via programação ou utilizando os menus (*point-and-click*). O aprendizado no Stata é relativamente fácil, sendo superado apenas pelo SPSS, que não será usado no curso devido a seu elevado custo comercial.

Ms Excel

O Excel, programa de planilhas eletrônicas da Microsoft, embora não seja um *software* concebido especificamente para análise de dados, permite a realização de análises estatísticas diversas. O Excel oferece a opção “Análise de Dados” no menu “Ferramentas” que inclui, entre outras possibilidades, o cálculo de estatísticas descritivas, a construção de histogramas e outros gráficos, a realização de testes de comparações de médias, ANOVA e regressão linear. Usualmente, essa opção não se encontra ativa no programa e, portanto, não aparece no menu de opções. Para ativá-la, é preciso clicar em Ferramentas / Suplementos / Ferramentas de análise.

Uma das principais vantagens do Excel, com relação ao R ou ao Stata, é que se trata de um dos programas de computador mais populares hoje no mercado. Ele é o programa de planilhas eletrônicas dominante para as plataformas Windows e Mac e já vem incluído como parte do Microsoft Office. Assim sendo, ele está instalado na maioria dos computadores pessoais e é, sem dúvida o programa mais popular entre esses três que serão utilizados no curso. Entretanto, o Excel não tem muita flexibilidade e a execução dos comandos não é simples, seja usando as ferramentas nele inclusas ou os programas que a ele podem ser integrados.

Reprodutibilidade de pesquisa e boas práticas de programação

A disciplina adota padrões mínimos de reproduzibilidade para todas as listas e para o projeto final. Entende-se por reproduzibilidade a possibilidade de que terceiros refaçam, do início ao fim, os procedimentos analíticos com os mesmos dados, código e ambiente, obtendo resultados equivalentes aos relatados. A ênfase em práticas reproduzíveis é coerente com os objetivos do curso e com a organização das atividades avaliativas previstas neste programa.

Cada entrega deverá vir acompanhada de documentação suficiente para execução ponta a ponta. Espera-se: um roteiro claro de reprodução (README), a indicação das fontes dos dados e um sumário das transformações realizadas, além do relatório analítico executável que integra texto, código e resultados. A disciplina disponibilizará exemplos e arquivos de apoio nos repositórios indicados pelo professor, preservando a lógica de distribuição de materiais já prevista neste programa.

O ambiente computacional deverá ser fixado e restaurável. Em R, os alunos devem salvar um *lockfile* e registrar as versões utilizadas, garantindo restauração automática do conjunto de pacotes na reprodução. Em Python, devem fornecer um arquivo declarativo de ambiente (por exemplo, `environment.yml` ou `requirements.txt`) e instruções de criação/ativação do ambiente. Em ambos os casos, recomenda-se registrar a versão do interpretador e do sistema operacional ao final do relatório. Essa orientação reconhece que softwares e aplicativos são atualizados continuamente e que a documentação de versões é parte do padrão de qualidade exigido.

A organização do projeto deve favorecer a rastreabilidade. Utilizar caminhos relativos; evitar referências a diretórios locais; separar dados brutos e dados derivados; manter funções e rotinas em arquivos próprios; e garantir execução linear do relatório (reiniciar e executar tudo). Notebooks devem servir à comunicação dos resultados e não concentrar toda a lógica de processamento; a lógica principal deve estar em scripts ou módulos importáveis.

A disciplina valoriza a integridade ética e a conformidade legal no uso de dados. Os alunos devem observar licenças e termos de uso, resguardar informações sensíveis e, quando necessário, substituir bases restritas por dados sintéticos ou fornecer instruções de acesso. Todas as decisões de limpeza, seleção de casos e exclusões devem ser descritas, ainda que brevemente, no relatório.

A avaliação levará em conta, além do conteúdo substantivo, a executabilidade do material entregue, a fidelidade do ambiente, a completude documental e a clareza na comunicação do procedimento analítico. Esses critérios se somam à estrutura de listas e projeto final definida nesta ementa.

R e reproduzibilidade

R oferece um ecossistema coeso para relatórios reproduzíveis. O uso de documentos executáveis (por exemplo, via Quarto) permite combinar narrativa, código e resultados em um único arquivo recompilável. A gestão de dependências por *lockfile* garante que versões de

pacotes sejam restauradas com fidelidade em outros computadores. A sintaxe consistente do *tidyverse* facilita a leitura do fluxo analítico e reduz ambiguidade na replicação.

Python e reproduzibilidade

Python dispõe de ambientes declarativos com *Conda* ou *venv* e descreve dependências em arquivos de especificação simples, o que viabiliza clonagem fiel do ambiente. Jupyter e Quarto permitem documentos executáveis com integração de texto, código e gráficos. A linguagem se integra com facilidade a *pipelines*, testes e automação, o que favorece reexecução completa do fluxo analítico em diferentes máquinas.

Por fim, recomenda-se controlar versões do projeto com Git e registrar uma versão arquivada no encerramento do trabalho (por exemplo, por meio de uma *release*). A prática permite vincular a versão submetida às listas ou ao projeto aos materiais efetivamente reproduzíveis, em consonância com o uso de repositórios e materiais complementares previstos neste programa.

Política de colaboração e uso de IA

Esta disciplina incentiva a colaboração acadêmica e o uso responsável de ferramentas de inteligência artificial (IA) como apoio ao aprendizado, desde que o trabalho final reflita a compreensão individual do estudante e observe padrões de transparência e integridade. Salvo orientação em contrário do professor, as listas são individuais. O projeto final pode admitir trabalho em dupla ou equipe, quando explicitamente autorizado, com responsabilidades claramente delimitadas.

Uso aceitável

Considera-se aceitável empregar IA para esclarecer conceitos, revisar texto, sugerir melhorias de estilo, resumir documentação técnica, diagnosticar mensagens de erro e propor caminhos de depuração. É aceitável gerar esqueleto de código e pseudocódigo, desde que acompanhado de documentação completa (comentários e *docstrings*), checagens de integridade de dados, testes mínimos de funcionamento e explicação das escolhas metodológicas. Também é aceitável usar IA para esboçar gráficos e tabelas como protótipo, que devem ser validados e, quando necessário, reescritos pelo estudante. Em todos os casos, o estudante permanece responsável por verificar resultados, reproduzir a análise do início ao fim e justificar cada decisão.

Transparência e registro

Toda entrega deve incluir uma seção “Uso de IA e colaboração” no **README**, descrevendo, de forma sucinta, quais ferramentas e modelos foram utilizados, quais partes do texto ou código foram influenciadas, quais dados (se houver) foram compartilhados com o serviço e quais procedimentos de validação foram realizados. Recomenda-se preservar os principais *prompts*

em anexo e registrar versões de software. Em trabalhos em equipe, delimitar a contribuição individual e a contribuição mediada por IA.

Compreensão demonstrável

Submissões sem compreensão demonstrável serão recusadas. A critério do professor, o estudante poderá ser convocado para uma defesa técnica breve, replicação ao vivo, explicação linha a linha do código, modificação pontual do fluxo analítico ou interpretação de resultados alternativos. A incapacidade de executar essas tarefas caracteriza insuficiência de domínio e poderá implicar refação, perda de pontos ou nota zero, conforme gravidade.

Práticas proibidas

É vedado entregar soluções geradas por IA sem validação independente, copiar trechos extensos de código ou texto sem atribuição e fazer-se passar por autor de conteúdo que não domina. É proibido fabricar dados, inventar referências ou resultados, burlar procedimentos de avaliação, contornar limitações técnicas do software por meios opacos e compartilhar dados sensíveis ou restritos com serviços de IA sem autorização explícita e anonimização adequada. Também é proibido submeter material cuja execução dependa de serviços externos indisponíveis ao avaliador.

Proteção de dados e conformidade

O uso de IA deve observar a legislação aplicável, as licenças dos dados e a política de reproduzibilidade desta ementa. Informações sensíveis devem ser anonimizadas. Sempre que houver restrição legal ou contratual, substituir a base por dados sintéticos e documentar o procedimento de acesso.

Avaliação

A avaliação considerará a qualidade substantiva da análise, a clareza da comunicação, a reproduzibilidade integral e a transparência sobre colaboração e IA. O histórico de versões (*commits* e *releases*) poderá ser solicitado para verificar autoria e evolução do trabalho. A política aqui definida complementa as demais normas do programa e não exime o estudante de cumprir os prazos e formatos de entrega estabelecidos.

Programa

Aula 1: “Introdução: Teoria e Dados”

Apresentação do curso: o modelo científico nas ciências sociais, estudos observacionais e estudos experimentais.

Leitura obrigatória: não há leitura obrigatória para essa aula

Leitura complementar:

Kellstedt & Whitten (2009) capítulo 1: “*The Scientific Study of Politics*” e capítulo 2: “*The Art of Theory Building*”

Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier. “Political science methodology.” *The Oxford handbook of political methodology*. 2008.

Aula 2: “Conceitos, Medidas e Hipóteses”

Medidas: Os conceitos das ciências sociais podem ser medidos? Quais as principais dificuldades da mensuração? Mensuração, tipos de dados e variáveis.

Leitura obrigatória:

Goertz, Gary. “Concepts, Theories, and Numbers: A Checklist for Constructing, Evaluating, and Using Concepts or Quantitative Measures.” *The Oxford Handbook of Political Methodology*.

Pedhazur & Schmelkin (1991). capítulo 1: “*Measurement and Scientific Inquiry*”

Leitura complementar:

Pereira (2004) capítulo 1: “O Dado Qualitativo” (pp.29-42) e capítulo 2: “Definição de Medidas”

Agresti, Franklin & Klingenberg (2017) capítulo 4: “*Gathering Data*”

Pollock (2012) capítulo 1: “The Definition and Measurement of Concepts” (pp.6-27).

OBSERVAÇÃO: Distribuição da “Lista de Exercícios 1”

Aula 3: “Definindo e Medindo suas Variáveis”

Validade e confiabilidade na operacionalização e mensuração de conceitos nas ciências sociais e seus problemas.

Leitura obrigatória:

Kellstedt & Whitten (2009) capítulo 5: “*Getting to Know your Data: Evaluating Measurement and Variations*”

Leitura complementar:

Pedhazur & Schmelkin (1991). capítulos 2 - 5 : “*Criterion-Related Validation*”, “*Construct Validation*”, e “*Reliability*”

Freedman, Pisani & Purves (2007) capítulo 6: “*Measurement Error*”

Aula 4: “Estatísticas Descritivas: Medidas de Tendência Central e Medidas de Dispersão”

Medidas de tendência central: média, moda e mediana. Medidas de dispersão: variância, desvio padrão e outras medidas de variabilidade dos dados em torno da média. Como usar gráficos para examinar e entender seus dados.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 2: “*Exploring Data with Graphs and Numerical Summaries*”

Leitura complementar:

Kellstedt & Whitten (2009) capítulo 6: “*Descriptive Statistics and Graphs*”

Monogan, James (2015), capítulo 4.1. Political Analysis using R. Springer

Aula 5: “Gráficos: analisando os dados de maneira descritiva”

Como usar gráficos para examinar e entender os dados. Nesse tópico cobriremos os principais quadros de representação usados nas ciências sociais: histogramas, *scatter plots*, gráficos de barra, gráficos de linhas, *boxplots* e mapas.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 2: “*Exploring Data with Graphs and Numerical Summaries*”

Leitura complementar:

Kellstedt & Whitten (2009) capítulo 6: “*Descriptive Statistics and Graphs*”

Freedman, Pisani & Purves (2007) capítulos 3 e 4: “*The Histogram*” e “*The Average and the Standard Deviation*”

Wickham, Hadley (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

OBSERVAÇÃO: Devolução da “Lista de Exercícios 1”

Aula 6: “Associação: Contingência, Correlação e Regressão.”

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 3: “*Association: Contingency, Correlation, and Regression*”

Leitura complementar:

Freedman, Pisani & Purves (2007) capítulo 8 e 9: “*Correlation*” e “*More about Correlation*”

Aula 7: “Distribuições de Probabilidade”

Como a probabilidade quantifica a aleatoriedade? As principais distribuições de probabilidade.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 5: “*Probability in Our Daily Lives*” e capítulo 6 “*Probability Distributions*”

Leitura complementar:

Kaplan (2011) capítulo 11 “*Modeling Randomness*”

Freedman, Pisani & Purves (2007) capítulo 13: “*What are the Chances?*”

Aula 8: “Os Fundamentos da Inferência Estatística”

Inferência e Causalidade nas Ciências Sociais. O que é inferência estatística? O teste de significância, amostras e populações.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 8: “*Statistical Inference: Confidence Intervals*”

Leitura complementar:

Kellstedt & Whitten (2009) capítulo 6: “*Probability and Statistical Inference*” (120-132)

Pollock (2012) capítulo 5: “*Foundations of Statistical Inference*” (pp.122-152).

Aula 9: “Teste de Hipóteses”

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 9: “*Statistical Inference: Significance Tests About Hypotheses*”

Leitura complementar:

Kaplan (2011) capítulo 13 “*The Logic of Hypothesis Testing*”

Kellstedt & Whitten (2009) capítulo 7: “*Bivariate Hypothesis Testing*” (120-132)

Aula 10: “Comparações Envolvendo Médias e Proporções”

Como comparar dois grupos? O teste *t* para duas amostras. Testando hipóteses. O teste de diferenças entre médias e o teste de diferenças entre proporções.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 10: “*Comparing Two Groups*”.

Leitura complementar:

Kaplan (2011) capítulo 4: “*Group-wise Models*”.

Aula 11: “Analizando a Associação entre Dados Categóricos”

Associação e independência. Frequência observada e frequência esperada. Correção de Yates. Teste exato de Fisher.

Leitura obrigatória:

Agresti, Franklin & Klingenberg (2017) capítulo 11: “*Analyzing the Association Between Categorical Variables*”.

Leitura complementar:

Sirkin (2006) capítulo 12: “The Qui-Square Test” (pp.397-437).

Aula 12: “Analizando a Associação entre Dados Quantitativos: Análise de Regressão”

O modelo de regressão linear simples. Estimação via MQO. Interpretação da reta de regressão. Correlação de Pearson.

Leitura obrigatória:

Kennedy (2009) capítulo 3: “O Modelo Clássico de Regressão Linear” (pp.38-49).

Leitura complementar:

Agresti, Franklin & Klingenberg (2017) capítulo 12: “*Analyzing the Association Between Quantitative Variables: Regression Analysis*”.

Aula 13: “Entendendo e Interpretando o Modelo de Regressão”

O modelo de regressão linear multipla. Interpretação dos coeficientes da Regressão. Resíduos. ANOVA.

Leitura obrigatória:

Kennedy (2009) capítulo 4: “Estimação de Intervalo e Teste de Hipótese” (pp.50-65) e capítulo 5: “Especificação” (pp.70-88) .

Leitura complementar:

Agresti, Franklin & Klingenberg (2017) capítulo 13: “*Multiple Regression*”.

Kellstedt & Whitten (2009) capítulo 10: “*Multiple Regression Models I: The Basics*” (183-200).

Bibliografia

Os bancos de dados e os *scripts* dos programas computacionais de análise estatística utilizados nas sessões de laboratório e nas listas de exercícios serão distribuídos por e-mail e, posteriormente, postados na pasta compartilhada da disciplina no site: www.dropbox.com Dropbox.com. As leituras exigidas para as aulas dividem-se em obrigatórias e complementares. Espera-se que os alunos já tenham feito a leitura dos textos indicados como obrigatórios antes das respectivas aulas. Cabe observar que o conteúdo dessa disciplina tem uma caráter fortemente cumulativo. Ou seja, a compreensão dos tópicos abordados numa determinada aula dependem do domínio dos conteúdos abordados nas aulas passadas. Uma bibliografia complementar será apresentada ao final de cada aula como sugestão de leitura para aqueles que desejem se aprofundar no tema.

Livros textos

1. AGRESTI, Alan, Christine FRANKLIN and Bernhard KLINGENBERG. *The art and science of learning from data*. Upper Saddle River, NJ: Prentice Hall, 2007.
2. BUSSAB, Wilton e MORETTIN, Pedro. *A Estatística Básica*. 6^a edição. São Paulo: Saraiva, (2009).
3. FREEDMAN, David, Robert PISANI, and Roger PURVES. "Statistics (International Student Edition)." Pisani, R. Purves.-4th edition.-NY, USA: WW Norton & Company 720 (2007).
4. GELMAN, Andrew, and Jennifer HILL. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.(2007).
5. KENNEDY, Peter. *Manual de Econometria*. Rio de Janeiro: Elsevier, 2010.
6. LEVIN, Jack e FOX, James A. *Estatística para Ciências Humanas*. 9^a edição. São Paulo: Prentice Hall, 2004.
7. MONOGAN, James III. 2015. *Political Analysis Using R*. Springer.
8. PEREIRA, Julio Ignacio Piovani. *Análise de Dados Qualitativos: Estratégias Metodológicas para as Ciências de Saúde, Humanas e Sociais*. 3^a edição. São Paulo: EDUSP, 2004.
9. PEDHAZUR, Elazar J.; SCHMELKIN , Liora Pedhazur. *Measurement, design, and analysis: An integrated approach*. Psychology Press, New York, 1991.
10. WOOLDRIDGE, Jeffrey M. 2010. *Introdução à Econometria: Uma Abordagem Moderna*. Tradução da 4 ed. norte-americana. São Paulo: Cengage Learning.

Leitura Complementar

1. KELLSTEDT, Paul M. & WHITTEN, Guy D. *The Fundamentals of Political Research*. Cambridge: Cambridge University Press, 2009.

2. POLLOCK, Philip H. *The Essentials of Political Analysis*. 4th ed. Washington: CQ Press, 2012.
3. SIRKIN, R. Mark. *Statistics for the Social Sciences*. 3rd ed. Thousand Oaks: Sage, 2006.

Versão preliminar e sujeita a (pequenas) alterações, atualizada em: 17 de setembro de 2025.

Documento produzido em L^AT_EX

<https://github.com/rceneviva/metodosquanti101/>