

FLS 6183 - Métodos Quantitativos de Pesquisa II

Manoel Galdino

2025-20-06

Objetivos

O objetivo do curso é introduzir os métodos de regressão múltipla de maneira aplicada, por meio da utilização da linguagem R, para a análise dos fenômenos sócio-políticos. Ao final do curso, o aluno será capaz de entender e interpretar pesquisas quantitativas que utilizem métodos de regressão, aplicar esses métodos em suas pesquisas e análise de dados em geral.

O aluno também deverá ter adquirido familiaridade suficiente com a linguagem R para importar banco de dados, processar dados e realizar análises descritivas e de regressão. O curso possui também um componente crítico, de forma a permitir aos alunos entender os limites das técnicas de pesquisa na ciência política. Objetiva também estimular a capacidade crítica necessária para a seleção de técnicas apropriadas para coleta de dados e pesquisa.

Justificativa

A análise empírica quantitativa em ciências sociais envolve, na maior parte do tempo, utilização de técnicas de regressão. Embora cada vez mais outras técnicas estejam sendo empregadas, regressão continua sendo a principal ferramenta de modelagem estatística nas ciências sociais e é fundamental que os alunos de mestrado e doutorado possuam conhecimento sobre essa técnica.

Ao utilizarmos a linguagem R, estaremos também introduzindo o estudante de pós-graduação em uma das principais linguagens de programação voltada a análise de dados que existe hoje, com amplas possibilidades de aplicação no setor acadêmico, governamental e do mercado privado. Por ser uma linguagem open source, existe uma ampla comunidade de praticantes que se ajudam, criando assim espaço para que o aluno possa prosseguir seus estudos da linguagem posteriormente.

Professores

O Curso será ministrado pelos professores Manoel Galdino, professor do DCP-USP e Rodrigo Martins, Pós-Doc no DCP-USP.

Pré-requisitos

O pré-requisito para o curso é Métodos I ou equivalente. Pressupomos apenas familiaridade básica com o R.

Formato das aulas

No geral, as aulas serão divididas em duas partes: a primeira parte da aula será expositiva, com base no tema da aula e supondo-se que os alunos leram a bibliografia indicada previamente. Na segunda parte da aula, iremos utilizar o R para implementar e testar o que aprendemos na parte teórica.

Bibliografia

O curso irá seguir o livro Theory of Agnostic Statistics by Aronow and Miller (AM) e o Livro do Hansen, Econometrics, ambos disponíveis no Moodle.

Leitura adicional

Wooldridge, Jeffrey. Introductory Econometrics. New York: South-Western. 4th edition.

Moore, William H. & David A. Siegel . 2013. A Mathematics Course for Political & Social Research. Princeton University Press.

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

<https://github.com/iamericfletcher/awesome-r-learning-resources>

Plataforma Digital

A Plataforma Moodle será o principal canal de comunicação entre professor, monitor e alunos e de entrega das atividades. Especificamente, o Moodle será utilizado para as seguintes funções:

Os avisos ordinários e extraordinários do curso serão postados no Moodle; As atividades semanais dos laboratórios em sala de aula serão entregues pelo Moodle em arquivos PDF e arquivo em Rmarkdown. O trabalho final será entregue pelo Moodle em um arquivo PDF e com o script do software R (em Rmarkdown).

Avaliação

1. Participação em aula 10%;
2. Listas de exercícios 40%;
3. Trabalho final 50%: Os alunos da pós-graduação deverão entregar um trabalho final utilizando uma base de dados de interesse individual do discente. Espera-se que o texto cubra a maior parte dos tópicos trabalhados ao longo do curso. As especificações a respeito do trabalho serão combinadas com a turma durante o semestre.

Listas de Exercício

Serão passadas listas de exercício (aprox. 8), quase todas envolvendo implementações práticas no R, que devem ser entregues pelo Moodle, em RMarkdown (ou Quarto, se alguém preferir). Algumas das listas serão corrigidas por vocês mesmos (cada uma corrige a lista de uma outra colega), outras pelo professor. Isso será combinado ao longo do semestre

Agenda preliminar

14 de agosto: Aula 1 - Apresentação da disciplina e álgebra linear

21 de agosto: Aula 2 - Revisão de stat e probabilidade (1.1 e 1.2; 2.1 e 2.2.1 e 2.2.2 de AM)

28 de agosto: Aula 3 - R, RMarkdown, Github

04 de setembro: Semana da Pátria. Não haverá aula.

11 de setembro: Aula 4 - Causalidade - Potential Outcomes (chap 7 de AM) -

18 de setembro: Aula 5 - CEF (2.2.3-2.2.5 de AM)

25 de setembro: Aula 6 - Regressão Linear - plug-in e OLS (4.1 de AM; 1 W)

02 de outubro: Aula 7 - Pressupostos e Propriedades de OLS (4 H)

09 de outubro: Aula 8 - Estimação de CEF não-linear e preditores binários (4.3 AM; 2.16 e 2.17 H)

16 de outubro: Aula 9 -Termos de interação (a definir)

23 de outubro: ANPOCS. Não haverá Aula

30 de outubro: Aula 10 - Viés de variável omitida (2.21 - 2.25 H)
06 de novembro: Aula 11 - Erros padrão e Bootstrap (2.10-2.13 H; 4.2 AM)
13 de novembro: Aula 12 - Inferência (5 H)
20 de novembro: Feriado. Não haverá aula
27 de novembro: Aula 13 - Modelos Lineares Generalizados (5.1 AM)
04 de dezembro: Aula 14 - Missing Data (6 AM)
18 de dezembro: data limite para entrega do trabalho final.