

Geoestatística

Cursos de Ciência de Dados

Escola de Governo do Estado de São Paulo e Fundação SEADE

Lucas Gelape

1 de novembro de 2024

Ementa

A análise espacial procura relacionar dados alfanuméricos com sua dimensão espacial e assim contribuir para melhorar o conhecimento de processos complexos que ocorrem no território. Esse tipo de análise permite potencializar os estudos diagnósticos e prognósticos em diferentes áreas. Este curso tem portanto o objetivo de fornecer noções básicas de geoestatística e prover conhecimento técnico para produção de análises espaciais.

A geoinformação é cada vez mais utilizada em nossa sociedade. O mapa, que é a representação espacial de fenômenos e eventos, ganha relevância no processo de conhecimento da realidade e na comunicação desses saberes. Esse instrumento se mostra valioso em todas as fases de promoção de políticas públicas, desde o diagnóstico até a avaliação de impacto.

O curso oferece noções básicas de geoestatística, com foco na aplicação de diagnósticos socioeconômicos com o uso da linguagem R. Serão abordados: a ideia de vizinhança e matriz de pesos espaciais; análise autocorrelação espacial global e local; modelos de regressão espacial; métodos de interpolação de dados espaciais e krigagem. As aulas deste curso se dividem em abordagens teóricas e práticas, partindo de conceitos sobre linguagem cartográfica e chegando à introdução ao uso de dados espaciais.

São objetivos do curso:

- apresentar noções básicas sobre como o espaço pode ser incorporado em modelagens estatísticas;
- explicar diferentes modelos univariados, bivariados ou multivariados da associação de variáveis no espaço;
- prover conhecimento técnico para o manuseio de dados espaciais em formato *raster*;
- exercitar algumas técnicas de interpolação de dados no espaço e de krigagem.

Conteúdo

Aula 1: Vizinhança e autocorrelação espacial

- A ideia de vizinhança
- Operacionalização do conceito: a matriz de pesos espaciais (W)
- Autocorrelação espacial global: I de Moran e C de Geary
- Autocorrelação espacial local (LISA)

Aula 2: Regressão espacial

- Diagnóstico de dependência espacial em regressões lineares
- Modelos de regressão espacial: panorama
- Modelos de erro espacial (*spatial error*)
- Modelos de defasagem espacial (*spatial lag*)

Aula 3: Dados *raster*: introdução e manipulação

- Importação e manipulação de dados *raster*
- Visualização de dados *raster*
- Operações geográficas
- *Rasterização* de dados vetoriais

Aula 4: Interpolação e krigagem

- Métodos de interpolação de dados espaciais
- Krigagem

Sobre o R e RStudio

O R é tanto uma linguagem de programação quanto um software livre e gratuito, com uso bastante difundido em áreas como ciências sociais, economia e biologia. A sua comunidade de usuários em português é bastante ampla, o que facilita as pesquisas por ajuda com diversos códigos. O RStudio, por sua vez, é um ambiente de desenvolvimento integrado (IDE, em sua sigla em inglês). Isto é, uma espécie de “máscara” para facilitar o uso do R – existem outras IDEs com o mesmo objetivo.

As análises e atividades deste curso serão realizadas por meio de um servidor RStudio disponibilizado pela Fundação SEADE no seguinte endereço: <https://rstudio.seade.gov.br/auth-sign-in>. Para acessá-lo, vocês devem inserir o usuário e senha disponibilizado em documento à parte.

Caso você não esteja confortável com o uso do R/RStudio, não se aflija. Em primeiro lugar, os scripts com as análises realizadas em aula serão disponibilizados a todos os alunos. Além disso, como mencionado, existem materiais disponíveis *online* e que podem te auxiliar a entender melhor como programar nesta linguagem. A seguir, apresentamos uma recomendação de materiais em português que você deveria consultar ao longo do curso, caso considere necessário:

- Phillips, Jonathan. *Análise de Dados para as Ciências Sociais*. Material de curso oferecido no Departamento de Ciência Política da Universidade de São Paulo. 2021.
- Damiani, A. et al. *Ciência de Dados em R*. 2022.

Bibliografia recomendada

- Lovelace, Robin; Nowosad, Jakub; Muenchow, Jannes. *Geocomputation with R*. CRC Press, 2019.
- Moraga, Paula. *Spatial Statistics for Data Science: Theory and Practice with R*. Chapman & Hall/CRC, 2023.
- Darmofal, David. *Spatial Analysis for the Social Sciences*. Cambridge University Press, 2015.
- Barrozo, Ligia Vizeu. *Cartografia Temática em R para estudantes de Geografia*. 2023.