

Dados Digitais

Plano de Aula: Summer Institutes in Computational Social Science 2023

Preparado por SICSS-FGV ECMI, e adaptado do SICSS-Rutgers e SICSS-Princeton.

Resumo

Uma atividade em grupo com objetivo de dar mais práticas em processos de coleta de dados da internet, a formular perguntas de pesquisa, e pensar limites, benefícios e extensões com base nestes dados.

Objetivos de Aprendizado

- Dar mais experiências aos participantes com processos de raspagem de dados e uso de APIs. Muitos participantes possuem experiência neste tipo de atividade, portanto, a ideia é que vocês trabalhem em conjunto e compartilhem conhecimento.
- Gerar potenciais ideias para a segunda semana do SICSS-FGV ECMI.

Atividade

- Dividir em grupos menores de forma aleatória.
- Fase Um: Desenho de Pesquisa e Coleta de Dados.
- Discutir as perguntas e dados com todos-as no Zoom.
- Fase Dois: Problemas, benefícios e extensões.
- Discutir sobre os projetos e idéias com todos-as no Zoom.

Programação

- Fase Um: 10:15am - 11:30pm
- Discussão das Perguntas e Dados: 11:30am - 12:00pm
- Fase Dois: 1:15-2:30pm
- Discussão das Extensões, Problemas e Vantagens: 2:30-3:30pm

Atividade

Fase um:

- Por 15 minutos, trabalhem juntos para identificar um tópico que você acredita que os dados de rastreamento digital (por exemplo, dados coletados por meio de captura de tela ou trabalhando com APIs) podem ajudar a descrever ou ser usados para abordar uma questão de pesquisa específica.
- Identifique uma base de amostragem relevante. Por exemplo, se o seu tópico de pesquisa é sobre política, sua base pode ser uma amostragem de PLs propostas em um determinado período.
- Comece a coletar os dados!

Fase Dois:

Aprimore os dados de rastreamento digital. Escolha uma das seguintes opções:

- Continue coletando os dados.
- Discuta os pontos fortes / fracos dos dados para abordar seu tópico de pesquisa.
- Desenvolva um projeto de pesquisa híbrido que permite combinar os dados de rastreamento digital com algum outro tipo de dados que irá gerar mais informações e abordar alguns dos pontos fracos que você identificou com os dados.

Recursos

Lembrem-se destes códigos das aulas on-line sobre raspagem de dados. Eles serão úteis para suas atividades:

[Screen-Scraping in R](#)

[Application Programming Interfaces in R](#)

[Building Apps & Bots for Social Science Research](#)

Além destes códigos, o Tiago Ventura possui alguns tutoriais em português sobre raspagem de dados:

[Tutorial Raspagem](#)

[Tutorial APIs](#)

Por fim, segue uma lista de recursos sobre APIs que podem ser úteis para a criação de idéias.

 APIs for CSS

Se você tiver sugestões de APIs em português, adicione na lista abaixo:

 APIs Brasil