

Coleta e Análise de Dados Secundários

Aula 01

Frederico Bertholini

Introdução

O que são dados secundários?

- ▶ “No sentido mais amplo, análise de dados coletados por outra pessoa” (Boslaugh, 2007)
- ▶ Uso de dados para responder a uma pergunta diversa da qual originou sua coleta (Vartanian, 2010)
- ▶ Em contraste com a análise de dados primários em que o mesmo indivíduo/equipe de pesquisadores desenha, coleta e analisa os dados

O que são dados secundários?

Muitas fontes

- ▶ Grandes conjuntos de dados financiados pelo governo
- ▶ Registros administrativo
- ▶ Suplementos de periódicos
- ▶ websites dos autores
- ▶ Etc.

O que são dados secundários?

- ▶ Disponível para um número aparentemente ilimitado de temas
- ▶ Quantitativo ou qualitativo
- ▶ Uso restrito ou público
- ▶ Direto ou observação indireta

Fontes essenciais de dados secundários

- ▶ Portal de dados do Governo Federal
<http://dados.gov.br/dataset?groups=governo-politica>
- ▶ IBGE <https://www.ibge.gov.br/estatisticas-novoportal/downloads-estatisticas.html>
- ▶ IPEADATA <http://www.ipeadata.gov.br/Default.aspx>
- ▶ DATASUS <http://www2.datasus.gov.br/DATASUS/index.php?area=0205&id=6936>
- ▶ INEP <http://portal.inep.gov.br/web/guest/dados>

Internacionais (US)

- ▶ Inter-University Consortium for Political and Social Research
http:
[//www.icpsr.umich.edu/icpsrweb/ICPSR/access/index.jsp](http://www.icpsr.umich.edu/icpsrweb/ICPSR/access/index.jsp)
- ▶ Data.gov <http://www.data.gov>
- ▶ National Center for Education Statistics <http://nces.ed.gov>
- ▶ U.S. Census Bureau <http://www.census.gov>
- ▶ Simple Online Data Archive for Population Studies (SodaPop)
<http://sodapop.pop.psu.edu/data-collections>

Vantagens de dados secundários

- ▶ Desenho do estudo e coleta de dados já concluídos
- ▶ Economiza tempo e dinheiro
- ▶ Acesso a dados internacionais e históricos que caso contrário, levariam vários anos e milhões de reais para coletar

Vantagens de dados secundários

- ▶ Ideal para uso em exemplos de sala de aula, projetos semestrais, mestrados teses, dissertações
- ▶ Normalmente os dados têm qualidade superior
- ▶ Estudos financiados pelo governo geralmente envolvem amostras maiores que são mais representativos da população-alvo (maior validade externa)

Vantagens de dados secundários

- ▶ A sobreamostra de grupos/comportamentos de baixa prevalência permite maior precisão estatística
- ▶ Os conjuntos de dados geralmente têm amplitude considerável (milhares de variáveis)

Desvantagens de dados secundários

Desenho do estudo e coleta de dados já concluídos

- ▶ Os dados podem não facilitar uma questão de pesquisa específica
- ▶ Informações sobre desenho do estudo e procedimentos de coleta de dados pode ser escassas

Desvantagens de dados secundários

- ▶ Os dados podem ter falta de profundidade (quanto maior a largura, mais difícil para medir qualquer construção em profundidade)
- ▶ Certos campos ou departamentos (por exemplo, programas experimentais) podem ver menor valor na análise de dados secundários
- ▶ Pode exigir conhecimento de estatística/métodos de pesquisa que não são geralmente fornecidos por cursos de graduação ou pós-graduação

Entenda seus dados

Familiarize-se com o estudo e os dados originais!

- ▶ Leia todos os manuais
- ▶ Para quem os resultados são generalizáveis?

Entenda seus dados

- ▶ Como os dados faltantes (perdidos) são tratados?
- ▶ Quais são os pesos de análise apropriados?
- ▶ Quais variáveis compostas estão disponíveis e como elas são construídas?

Entenda seus dados

- ▶ Protocolos de coleta
- ▶ Questionários
- ▶ Atualizações

Preparo de dados

- ▶ Documente TUDO!

1. Transfira ou leia diretamente
2. Lide com missing data
3. Recodifique variáveis
4. Crie novas variáveis

Análise de dados

- ▶ Com base na sua questão de pesquisa, identificar análise estatística apropriada
- ▶ Selecione o pacote de software que implementará a análise e viabilizará a amostragem complexa
- ▶ Examine estatísticas descritivas não ponderadas para identificar erros de codificação e determinar a adequação do tamanho da amostra

Análise de dados

- ▶ Identifique pesos
- ▶ Identifique método de estimação de variância (e variáveis correspondentes)
- ▶ Realize análises de diagnóstico (identificar outliers, não normalidade, etc.)
- ▶ Realize análises preliminares e interprete os resultados!