

Amostragem e o SIPD

Curso de Programação - 2023

João Pedro de Freitas Gomes

15 de Agosto, 2023

FEA - USP



Made
centro de
pesquisa em
macroeconomia
das desigualdades

1. Resolução do Exercício Proposto na Aula Anterior
2. Pesquisa Amostral e Amostragem Complexa

Resolução do Exercício Proposto na Aula Anterior

Resolução

```
library(tidyverse)
library(nycflights13)
dados_resumidos <- flights %>% group_by(carrier)
%>% summarize(
  arr_delay = mean(arr_delay, na.rm = TRUE),
  dep_delay = mean(dep_delay, na.rm = TRUE))

dados_resumidos %>% ggplot(aes(x=carrier, y=arr_delay))
+ geom_bar(stat="identity", width=0.5)

dados_resumidos %>% ggplot(aes(x=carrier, y=dep_delay))
+ geom_bar(stat="identity", width=0.5)
```

Pesquisa Amostral e Amostragem Complexa

- No mundo ideal, teríamos informações sobre a população completa, ou seja, saberíamos as informações demográficas de todos, assim, toda estimativa que fizéssemos seria uma estimativa verdadeira.
- Em geral, é bem difícil obter dados desse tipo.
- Para contornar isso, utilizamos pesquisas amostrais, ou seja, selecionamos uma amostra da população e, a partir dela, iremos inferir informações sobre todo o nosso universo. Esta amostra não deve conter viéses, ou seja, deve ser uma amostra aleatória, justamente para garantir que realmente estamos representando a população.

Qual é o possível problema de se fazer uma amostra aleatória na população brasileira com o objetivo de se produzir uma pesquisa com informações representativas do Brasil?

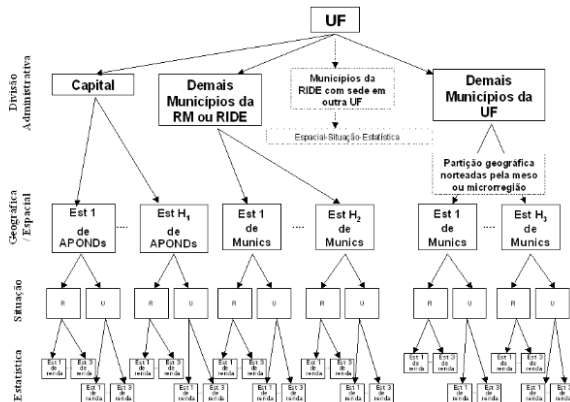
Qual é o possível problema de se fazer uma amostra aleatória na população brasileira com o objetivo de se produzir uma pesquisa com informações representativas do Brasil?

A população brasileira não é homogênea, então, ao realizar uma amostragem plenamente aleatória, existem grandes chances de que você tenha uma amostra composta somente de pessoas de uma determinada característica, a qual, apesar de ser representativa na população, não é única.

Conceitos Relevantes para a Amostragem do IBGE

- Setor Censitário: Unidade territorial estabelecida para fins de controle cadastral. Ela possui tamanho e um número de domicílios que permite o levantamento por parte de um recenseador.
- Cadastro Mestre: Base com todos os setores censitários elegíveis para composição da amostra mestra.
- Amostra Mestra: Conjunto de setores censitários selecionados com probabilidade proporcional a uma medida de tamanho de um Cadastro Mestre. A partir dela, seleciona-se as subamostras que são utilizadas nas pesquisas do SIPD.
- UPA: São setores censitários ou conjunto de setores censitários com pelo menos 60 domicílios.

- PNADc: Características econômicas e sociais da população.
- POF: Informações sobre padrões de consumo e gastos das famílias.
- PNS: Caracterizar a situação de saúde e estilo de vida da população.



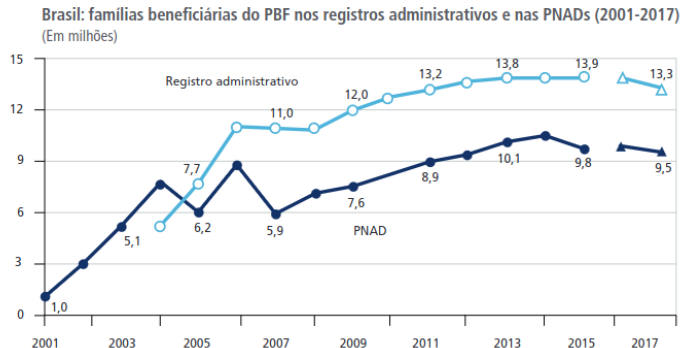
Municípios da Amostra Mestra



Alguns Problemas de Pesquisas Amostrais

- Quem responde a pesquisa muitas vezes pode não responder corretamente perguntas relacionadas à um tema por uma série de razões.
- Podem fornecer estimativas viesadas para subconjuntos muito pequenos da população.

Alguns Problemas de Pesquisas Amostrais



Fonte: PNADs (2001-2015), PNADs Contínuas (2016-2017) e dados da Matriz de Informações Sociais da Secretaria de Avaliação e Gestão da Informação (SAGI/MCidadania).
Elaboração dos autores.

Obs.: Informações da PNAD incluem os programas predecessores do PBF e excluem as áreas rurais dos estados da região Norte (exceto Tocantins) até 2003. Informações dos registros administrativos incluem somente o PBF e são referentes a setembro (2001-2015) e junho (2016-2017).