



MINISTÉRIO DA
SAÚDE



Curso

**Análise de Dados de Saúde e Clima:
Estatísticas para Políticas Públicas**

Descritivas

Módulo 2 – 06/08/2024

Priscilla Normando

Conteúdo

1

Definição

2

Pipeline de dados

3

O que fazer

4

Como fazer

<https://bit.ly/rbasico2024>



Descritivas

O que são elas?

Descritivas são relatórios que nos dizem o que está contido em um arquivo ou conjunto de dados.

- Devem dizer as características técnicas e semânticas do arquivo ou dado que estamos descrevendo.
- Os metadados de um arquivo ou dado são extraídos, incluídos ou transformados a partir de sua descrição.
- A extensão de um arquivo, sua data de criação e autor são exemplos de metadados de arquivo e descrevem como e quando ele foi criado.
- Os dados sobre a fonte (provedor, fornecedor, criador) de um dataframe, suas dimensões (tamanho em bytes), formato, licença de uso são metadados do dataframe e descrevem como, quando ele foi criado, além de nos dizer como, quando e para que finalidade pode ser usado. *LGPD / LAI / Outras legislações
- As descrições técnicas e estatísticas de um conjunto de dados nos fornecem informações sobre quantas observações o banco possui, que tipo de variáveis (tipo de dados, qualidade das variáveis [numéricas, categóricas], quantidade de preenchimento, qualidade de preenchimento).

***Por que descrevemos
dados e arquivos?***

- Inspeção → integridade dos arquivos, veracidade, completude
- Conhecimento sobre os dados → tipo de dado, categoria dos dados, utilidade
- Análise (interpretação?) → tipo de variáveis, resposta que as variáveis podem dar ao que desejo medir
- Validação → dados transformados criam novos objetos digitais que precisam ser novamente descritos e representados
OBS: representação de dado não é igual a representação de conhecimentos, assim como arquitetura e eng de dados não é igual a arquitetura e eng de conhecimentos

Vamos explorar algumas coisas?

- Buscar o site do Opendatasus ou INDA e baixar as fichas de notificação e o dicionário de dados do SIM
- Entrar no site da plataforma PCDAS e buscar o material disponível sobre o SIM
- Comparar os metadados e descritivas dos dois conjuntos de dados. Quais são as similaridades e quais as diferenças entre eles? Podemos dizer que se trata do mesmo conjunto de dados? Foram criados pelo mesmo provedor?
- Se você quiser usar os dados do SIM para responder à pergunta do projeto final do curso básico de R, é possível usar os dados fornecidos pela PCDAS? Se sim, de que forma?

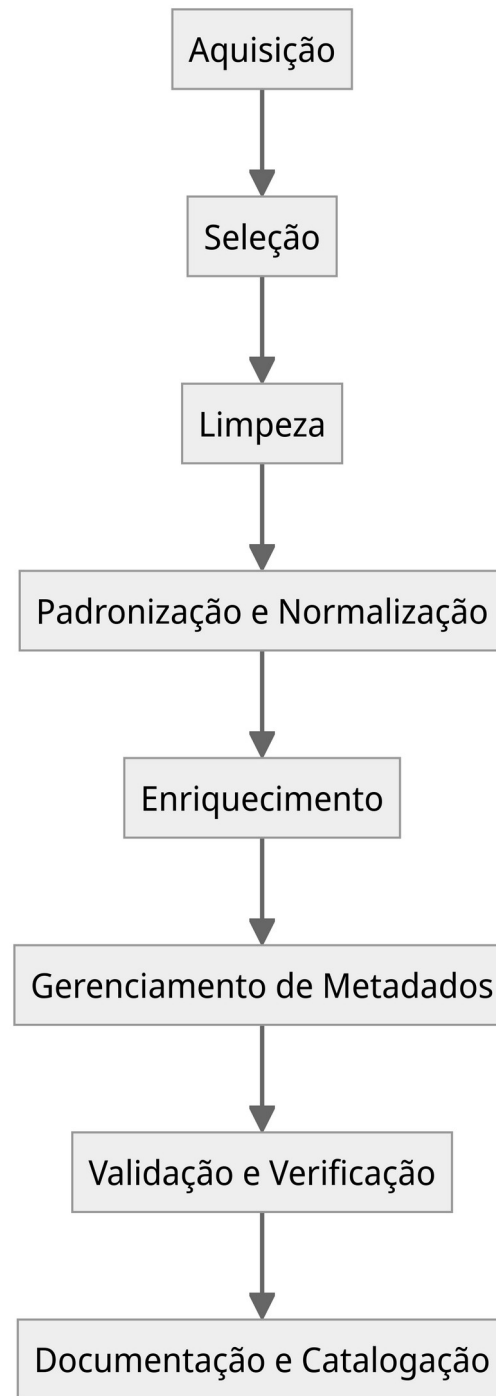
- REGRAS:

- Deve ser realizada em trios
- 20 minutos para realizar a atividade
- 5 minutos para apresentação

Pipeline de dados

O que fazer com os dados?

Pipeline de Dados



Descrevendo os dados

Como fazer?

Como descrever os dados?

- Buscamos e analisamos fontes de dados úteis e confiáveis
- Adquirimos os dados
- Armazenamos os dados com metadados
 - Fonte dos dados
 - Características dos arquivos originais
 - Repositórios
- Carregamos os dados
 - Inspecionamos os dados
 - Descrevemos os dados
 - Salvamos os relatório de inspeção e a descritiva

Funções, Pacotes e Manipulações

Conceitos básicos - Linguagem R

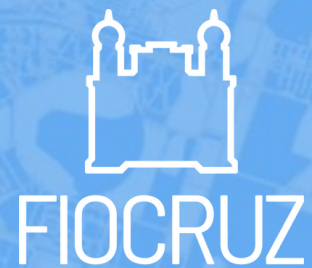
Vamos treinar?

- Vá até o repositório do projeto e baixe os dados do SIM disponibilizados por lá
 - No mesmo repositório baixe os arquivos `descritivaSimples.R` e `descritiva completa.R`
 - Rode os scripts linha a linha
 - Responda às seguintes questões: Quais são as colunas presentes no dataset? Quais as dimensões do dataset? Quais são as colunas com variáveis categóricas e quais são as com variáveis numéricas? Qual a proporção de raça/cor e de sexo biológico no dataset? Qual a idade mínima e máxima de óbito encontrada nos dados? Indique mais três características dos dados presentes na base. O que você acrescentaria na descritiva realizada (gráficos, outras descrições estatísticas, texto, imagem)?
- REGRAS:
 - Deve ser realizada individualmente
 - 20 minutos para realizar a atividade

Obrigada!

priscillanormando@gmail.com

priscilla.normando@fiocruz.br



Sala de
Situação de
Saúde
FS-UnB



MINISTÉRIO DA
SAÚDE



CLIMA
BRONQUIOLITE