

Dados

Conjunto de dados, tipos de variáveis, fontes de dados

Prof. Me. Lineu Alberto Cavazani de Freitas

CE003 – Estatística II

Departamento de Estatística
Laboratório de Estatística e Geoinformação



O que são dados?

- ▶ Dados são **conjuntos de valores**.
- ▶ Podem ser de diferentes fontes, tais como **estudos** e **experimentos**.
- ▶ Podem conter variáveis de diferentes tipos.
- ▶ Podem surgir em formatos **estruturados** e **não estruturados**.



Figura 1. Extraído de pixabay.com.

Conjunto de dados

- ▶ Em Estatística, em geral, lidamos com **dados estruturados em um formato tabular**.
- ▶ O conjunto de dados completo e sem tratamentos é denominado conjunto de **dados brutos**.
- ▶ Um conjunto de dados considerado **arrumado** é aquele em que:
 - ▶ Cada **coluna** representa uma **variável**.
 - ▶ Cada **linha** representa uma **observação**.
 - ▶ Cada **célula** representa o **valor** observado.

Conjunto de dados

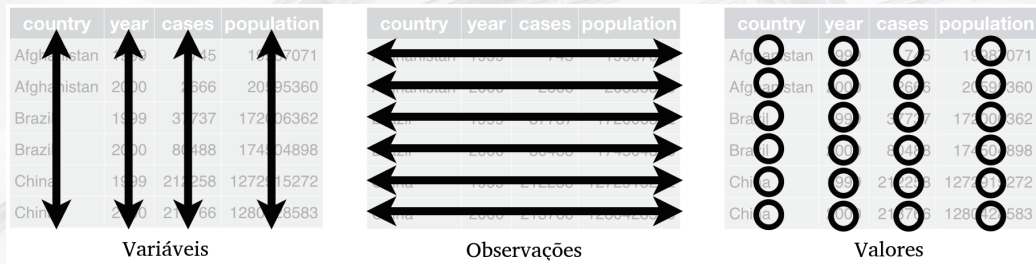


Figura 2. Extraído de <https://r4ds.had.co.nz>.

Conjunto de dados

Tabela 1. Exemplo de conjunto de dados

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa

Tipos de variáveis

- ▶ Na prática, podemos coletar variáveis de diferentes tipos e naturezas.
- ▶ Existem dois tipos (básicos) de variáveis:
 - ▶ Numéricas (**quantitativas**).
 - ▶ Não numéricas (**qualitativas**).
- ▶ Cada uma das classes tem suas ramificações.



Figura 3. Tipos básicos de variáveis.

Variáveis quantitativas

Exemplos

- ▶ **Variáveis Quantitativas:** assumem valores numéricos.
 - ▶ **Discretas:** características mensuráveis que podem assumir apenas um número finito ou infinito **contável** de valores.
 - ▶ **Contínuas:** características mensuráveis que assumem valores em uma **escala contínua**, isto é, na reta real.
- ▶ Altura.
- ▶ Peso.
- ▶ Idade.
- ▶ Percentual de gordura corporal.
- ▶ Número de filhos.
- ▶ Número de fraturas.
- ▶ Número de faltas.
- ▶ Número de peças defeituosas em um lote.

Variáveis qualitativas

Exemplos

- ▶ **Variáveis Qualitativas:** são as características definidas por categorias, ou seja, representam uma classificação dos indivíduos e não uma característica numérica.
 - ▶ **Nominais:** não existe ordenação nem peso entre as categorias.
 - ▶ **Ordinais:** existe uma ordenação entre as categorias.
- ▶ Estado civil.
- ▶ Orientação sexual.
- ▶ Turma.
- ▶ Posição em que joga em um time.
- ▶ Severidade de uma lesão.
- ▶ Escolaridade.
- ▶ Grau de proficiência em língua inglesa.
- ▶ Risco de infarto.

Cuidados com variáveis

- ▶ Existem particularidades na classificação de variáveis devido a situações como:
 - ▶ Discretização de variáveis contínuas.
 - ▶ Limitações em instrumentos de mensuração.
 - ▶ Utilização de quantidades numéricas para representação de variáveis categóricas
 - ▶ Dentre outras.
- ▶ Deve-se sempre estar atento a este tipo de situação pois podem levar a implicações nas análises e consequentemente nos resultados.
- ▶ Existem outros tipos de variáveis que ocorrem em situações particulares que requerem técnicas específicas de análise.

De onde vêm os dados?

Alguns exemplos:

- ▶ Estudos de caso.
- ▶ Experimentos.
- ▶ Pesquisas.
- ▶ Registros administrativos.
- ▶ Dados em repositórios online.
- ▶ Bancos de dados corporativos.



Figura 4. Extraído de pixabay.com.

Dados observacionais x dados experimentais

Dados observacionais

- ▶ **Observação passiva** da realidade.
- ▶ Sem modificação das condições.

Dados experimentais

- ▶ **Intervenção** na realidade.
- ▶ Condições controladas.
- ▶ Observação dos efeitos das intervenções.



Figura 5. Extraído de pixabay.com.

No que devemos pensar antes de analisar nossos dados?

- ▶ O que estamos interessados em avaliar?
- ▶ Quais são as variáveis de interesse?
- ▶ Quais são as variáveis que queremos avaliar se influenciam a variável de interesse?
- ▶ Quais são os métodos disponíveis para análise de variáveis deste tipo?
- ▶ Quais os métodos disponíveis que permitem responder nossa pergunta de pesquisa?
- ▶ Como coletar os dados?
- ▶ Os dados são válidos?

O que foi visto:

- ▶ Dados.
- ▶ O que são dados.
- ▶ Conjunto de dados.
- ▶ Tipos de variáveis.
- ▶ Fontes de dados.
- ▶ Estudos observacionais e experimentais.

Próximos assuntos:

- ▶ Amostras.
- ▶ Métodos de amostragem.
- ▶ Amostragem probabilística.
- ▶ Amostragem não probabilística.