

Disciplina CE003 – Estatística II

Dados

Prof. Me. Lineu Alberto Cavazani de Freitas

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.

Conjunto de dados

- ▶ Em Estatística, em geral, lidamos com dados estruturados em um formato tabular.
- ▶ Variáveis são as características observadas em cada um dos elementos que pertencem à amostra.
- ▶ Na prática precisamos de uma forma para representar indivíduos e variáveis.

Conjunto de dados

- ▶ 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.
- ▶ O conjunto de dados completo e sem tratamentos é denominado conjunto de dados brutos.

Conjunto de dados

FIGURA TIDY DATA



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.

Variáveis quantitativas

- ▶ Variáveis Quantitativas: assumem valores numéricos.
 - ▶ Discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros.
 - ▶ Contínuas: características mensuráveis que assumem valores em uma escala contínua, isto é, na reta real.

Variáveis Qualitativas

- ▶ Variáveis Qualitativas: são as características definidas 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.

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:

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

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.

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?