

MAT02018 - Estatística Descritiva

Organização dos dados

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

Porto Alegre, 2022

Introdução

Introdução

- ▶ Agora que já discutimos alguns **conceitos básicos** de estatística e as etapas gerais de um **levantamento estatístico**, vamos apresentar como é feito o **registro** e a **organização de dados** referentes a uma certa coleta de dados.
- ▶ Começaremos com a **planilha** para o registro dos dados e a **tabela de dados brutos** resultante.
- ▶ Logo em seguida, discutiremos como fazer a **apuração dos dados**.

Coleta de dados

Coleta de dados

Lembrando

A **estatística** é a ciência que tem por objetivo orientar a *coleta*, o *resumo*, a *apresentação*, a *análise* e a *interpretação* de dados.

- ▶ Para **coletar dados**, o pesquisador necessitará armazenar os dados coletados em algum lugar.
- ▶ Assim, se faz necessário organizar uma **planilha**.

Coleta de dados

- ▶ Com o advento da computação, grande parte dos profissionais da área de estatística registram dados em uma **planilha eletrônica**¹.
- ▶ No entanto, os dados também podem ser registrados em meio físico como, por exemplo, fichas, cadernos ou cadernetas, ou seja, a chamada **planilha física**.

¹Softwares como *Calc* (OpenOffice), *Microsoft Excel* (Office) e *Google Sheets* (Google) são exemplos de softwares que trabalham com planilhas eletrônicas.

Coleta de dados

- ▶ As planilhas eletrônicas podem ser construídas a partir de planilhas físicas ou serem alimentadas por algum **instrumento de coleta** em meio eletrônico (formulário ou questionário)².
- ▶ Vamos apresentar como se desenha uma planilha física para registro dos dados.
- ▶ **Se você tiver possibilidade, pode experimentar como organizar os dados em uma planilha eletrônica.**

²O *Google Forms*, por exemplo, cria e alimenta uma planilha eletrônica a partir do formulário de coleta.

Planilha

- ▶ **Planilha** é o documento que armazena os dados coletados, distribuindo-os em linhas e colunas (ou seja, planilhas são “matrizes de dados”).
- ▶ Em planilhas eletrônicas, geralmente, as linhas são numeradas e as colunas são indicadas por letras maiúsculas.

	A	B	C	D
1				
2				D2
3				
4				
5				

Planilha

Exemplo: Um pesquisador está interessado em fazer um levantamento sobre alguns aspectos socioeconômicos dos empregados da seção de orçamentos da Companhia MB, um grupo de 15 pessoas³.

- ▶ Temos a seguinte planilha para registrar os dados do grupo.

³Morettin, P. A. e Bussab, W. O. **Estatística Básica**, Saraiva, 2010.

Planilha

Número	Estado Civil	Grau de instrução	Número de filhos	Salário (x sal. mín.)	Idade	Região de procedência
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

Tabela de dados brutos

- ▶ **Dados brutos** são os dados na forma em que foram coletados, sem qualquer tipo de tratamento.
- ▶ Após a coleta de dados, o pesquisador tem em sua planilha o registro dos dados brutos.

Tabela de dados brutos

Número	Estado Civil	Grau de instrução	Número de filhos	Salário (x sal. mín.)	Idade	Região de procedência
1	solteiro	EF	0	4,00	26	interior
2	casado	EF	1	4,56	32	capital
3	casado	EF	2	5,25	36	capital
4	solteiro	EM	0	5,73	20	litoral
5	solteiro	EF	1	6,26	40	litoral
6	casado	EF	0	6,66	28	interior
7	solteiro	EF	2	6,86	41	interior
8	solteiro	EF	1	7,39	43	capital
9	casado	EM	1	7,59	34	capital
10	solteiro	EM	0	7,44	23	litoral
11	casado	S	2	8,12	33	interior
12	solteiro	EF	0	8,46	27	capital
13	solteiro	EM	2	8,74	37	litoral
14	casado	EF	3	8,95	44	litoral
15	casado	EM	0	9,13	30	interior

Tabela de dados brutos

- ▶ O que podemos falar sobre as variáveis coletadas?
- ▶ Qual a informação podemos apresentar sobre os dados coletados?

Para responder tais perguntadas, precisaremos **resumir** os dados de alguma forma.

Apuração dos dados

Apuração dos dados

- ▶ **Apuração** é o processo de retirar os dados da planilha e organizá-los, para apresentação.

No exemplo apresentado anteriormente, foram coletadas as seguintes variáveis: estado civil, grau de instrução, número de filhos, salário, idade e região de procedência. Note que estas são variáveis de diferentes tipos.

- ▶ **Exercício:** classifique cada uma destas variáveis em **qualitativa nominal**, **qualitativa ordinal**, **quantitativa discreta** e **quantitativa contínua**.

Apuração de dados nominais

- Se quisermos saber quantos solteiros e quantos casados trabalham na seção de orçamentos da Companhia MB devemos escrever os valores possíveis da variável **estado civil**⁴.

solteiro
casado

⁴**Pergunta:** a ordem de escrita dos valores possíveis da variável **estado civil** importa? Por que?

Apuração de dados nominais

- ▶ Logo após, precisamos inspecionar cada registro da tabela de dados brutos e marcar um traço ao lado de **solteiro**, para cada indivíduo solteiro inspecionado, e um traço ao lado de **casado** para cada indivíduo casado inspecionado.
- ▶ A cada quatro traços, corta-se com um traço, e este conjunto representa uma contagem de cinco indivíduos⁵.

<u>solteiro</u>		
<u>casado</u>		

⁵No inglês, *tally marks* (marcas de registro).

Apuração de dados nominais

- ▶ Desta forma, verificamos que na seção de orçamentos da Companhia MB trabalham oito solteiros e sete casados.
- ▶ Duas outras formas alternativas de se fazer a apuração dos dados são apresentadas a seguir.

solteiro
casado

solteiro
casado

Apuração de dados nominais

Comentário

É fácil apurar uma pequena massa de dados, como no caso do exemplo. Já uma grande massa de dados tornará a tarefa difícil e entediante. Além disso, com um grande volume de dados, a **probabilidade** de incorrermos em erros aumenta! Necessitaremos do auxílio de **pacotes estatísticos!**

Apuração de dados ordinais

- ▶ Para apurar dados de grau de instrução (variável qualitativa ordinal), o procedimento é similar ao adotado para apurar dados nominais.
- ▶ A diferença é que, para dados ordinais, **impõe-se uma ordem**.
 - ▶ Contudo, a apuração se faz por contagem.

Ensino Fundamental		= 9
Ensino Médio		= 5
Superior		= 1

Apuração de dados discretos

- ▶ Para apurar o número de filhos (variável quantitativa discreta), também devemos fazer uma contagem.
- ▶ Escrevemos os resultados respeitando a ordem numérica.

0	 	= 6
1	 	= 4
2	 	= 4
3	1	= 1

Apuração de dados contínuos

- ▶ Em geral, os dados contínuos são apresentados na forma como foram coletados, porque assumem valores diferentes, mesmo em amostras pequenas.
- ▶ É o caso da variável idade no exemplo considerado: os empregados da seção de orçamentos da Companhia MB tinham idades diferentes.
- ▶ No entanto, é possível organizar as idades por **faixas**, como veremos nas aulas seguintes.

Para casa

1. Construa a planilha para o registro do levantamento dos dados planejado nas aulas anteriores.
2. Faça uma pequena coleta de dados incluindo pelo menos uma variável de cada tipo (*qualitativa nominal, qualitativa ordinal, quantitativa discreta e quantitativa contínua*).
 - 2.1 Organize uma planilha (física ou eletrônica) para o registro dos dados coletados.
 - 2.2 Faça a coleta e preencha a planilha para obter os dados brutos.
 - 2.3 Faça a apuração dos dados e comente brevemente sobre os resultados encontrados.

Próxima aula

- ▶ Introdução ao R.

Por hoje é só!

Bons estudos!

