

Aula 1 – Apresentação de dados – parte 1

Nesta aula, você aprenderá:

- os conceitos básicos de população e amostra de uma pesquisa estatística;
- a distinção entre variáveis qualitativas e variáveis quantitativas;
- a construir distribuições de frequências para variáveis qualitativas e quantitativas discretas;
- a construir gráficos de colunas e de setores para representar dados qualitativos e quantitativos discretos.

Pesquisa estatística - conceitos básicos

População e amostra

Estatística é a ciência da aprendizagem a partir dos dados. Em geral, fazemos levantamentos de dados para estudar e compreender características de uma população. Por exemplo, um grande banco, querendo lançar um novo produto, precisa conhecer o perfil socioeconômico dos seus clientes e, nesse caso, a população de interesse é formada pelos clientes de todas as agências do banco. A Federação das Indústrias do Estado do Rio de Janeiro – FIRJAN – mede o grau de confiança dos empresários industriais através de uma pesquisa junto às empresas industriais, sendo a população de interesse, aqui, o conjunto das empresas industriais do estado do Rio de Janeiro. Com esses dois exemplos apenas, já podemos ver que o conceito de *população de uma pesquisa estatística* é mais amplo, não se restringindo a seres humanos; ela é definida exatamente a partir dos objetivos da pesquisa. Mais precisamente, *população é o conjunto de elementos para os quais se deseja estudar determinada(s) característica(s)*.

Embora tenham populações bastante distintas, essas duas pesquisas têm em comum o fato de os resultados desejados serem obtidos a partir de dados levantados junto a um subconjunto da população – uma *amostra*. Há várias razões para se trabalhar com *pesquisas por amostragem* – custo e tempo, em geral, são as mais comuns. Mas, além de serem mais baratas e rápidas, as pesquisas por amostragem, se bem planejadas, podem fornecer resultados quase tão precisos quanto aqueles fornecidos por *pesquisas censitárias*, em que todos os elementos da população são investigados. Exemplos clássicos de pesquisa censitária são os Censos Demográficos realizados a cada

POPULAÇÃO

AMOSTRA

CENSO

dez anos no Brasil e em outros países. O objetivo desses censos é levantar informações sobre toda a população do país, de modo a fornecer subsídios para os governantes definirem as políticas públicas.

Variáveis qualitativas e quantitativas

Nas pesquisas estatísticas, as características sobre as quais queremos obter informação são chamadas *variáveis*. Em uma pesquisa domiciliar sobre emprego e renda, algumas variáveis de interesse são sexo, raça, grau de instrução e valor dos rendimentos do morador. Em uma pesquisa sobre o estado nutricional dos brasileiros, o peso e a altura dos moradores de cada domicílio da amostra foram medidos. Para o acompanhamento da atividade industrial no Rio de Janeiro, a FIRJAN obtém informações junto às empresas industriais sobre tipo de atividade econômica, número de empregados, número de horas trabalhadas, valor da folha de pagamento. É importante diferenciar entre variáveis qualitativas e variáveis quantitativas. Sexo, raça, religião e atividade econômica de uma empresa são exemplos de *variáveis qualitativas*. Já valor dos rendimentos, peso, altura, número de empregados, valor da folha de pagamento são exemplos de *variáveis quantitativas*. Podemos ver, então, que as variáveis qualitativas *descrevem* características dos elementos de uma população, enquanto as variáveis quantitativas *mensuram* características desses elementos.

As variáveis quantitativas, por sua vez, podem ser discretas ou contínuas. Quando a variável puder assumir qualquer valor numérico em um determinado intervalo de variação, ela será uma variável *contínua*. Essas variáveis resultam normalmente de medições: peso, altura, dosagem de hemoglobina, renda etc. A interpretação desse tipo de variável leva à noção de valor aproximado, pois não existe instrumento de medição capaz de fornecer precisão absoluta na informação. Assim, quando uma balança mostra o peso de uma pessoa como 65,5 kg, esse valor, na verdade, é uma aproximação para qualquer valor entre, digamos, 65,495 e 65,505 kg. Por outro lado, a variável quantitativa *discreta* só poderá assumir valores pertencentes a um conjunto enumerável; os valores normalmente são obtidos através de algum processo de contagem. Alguns exemplos são: número de filhos de um casal, número de empregados de uma firma de contabilidade etc.

VARIÁVEIS QUALITATIVAS

VARIÁVEIS QUANTITATIVAS

Atividade 1.1

O texto a seguir foi extraído da página do IBOPE na Internet: www.ibope.com.br. Aí temos parte da descrição da pesquisa sociodemográfica realizada por esse instituto. Identifique as variáveis pesquisadas, classificando-as como qualitativas ou quantitativas.

“O Levantamento Socioeconômico (LSE) é a pesquisa do IBOPE Mídia que mapeia as características sociais, demográficas e econômicas das famílias das principais regiões metropolitanas do país. Oferece também outros dados essenciais para traçar a estratégia de marketing para um produto. Com uma base de dados estendida em relação às outras pesquisas do IBOPE Mídia, o LSE serve de base para outros estudos.

São levantados dados sobre a condição do domicílio entrevistado (condição da rua, tipo de imóvel) e sobre a condição socioeconômica do domicílio (informações sobre renda e classificação econômica). Também são pesquisados o número de pessoas no domicílio, a presença e a quantidade de crianças e adolescentes, a idade, grau de instrução e condição de atividade do chefe da casa e da dona-de-casa. A pesquisa levanta também dados sobre a posse de bens, como geladeira, máquina de lavar, automóvel, rádio, computador, telefone, entre outros, e acesso a serviços de mídia, como TV por Assinatura, Internet, etc.”

Apresentação de dados qualitativos

Vamos considerar o seguinte exemplo fictício, mas verossímil. A direção de uma empresa está estudando a possibilidade de fazer um seguro saúde para seus funcionários e respectivos familiares. Para isso, ela faz um levantamento junto a seus 500 funcionários, obtendo informação sobre sexo, estado civil, idade, número de dependentes e salário. Como são 500 funcionários, temos que achar uma forma de resumir os dados. Nesta aula você irá aprender a resumir dados qualitativos em forma de uma distribuição (ou tabela) de frequência e também em forma gráfica. Você verá que os gráficos complementam a apresentação tabular.

Distribuições de frequência

Consideremos inicialmente a variável qualitativa sexo. O que interessa saber sobre essa variável não é que João é do sexo masculino e Maria é do sexo feminino, mas, sim, quantos funcionários e quantas funcionárias há na empresa. Esse resultado pode ser resumido em uma tabela ou distribuição de frequências da seguinte forma:

Sexo	Número de Funcionários
Masculino	270
Feminino	230
Total	500

Os números 270 e 230 resultaram da contagem das frequências de ocorrência de cada uma das categorias da variável sexo. Essa contagem é também chamada de *frequência simples absoluta* ou simplesmente *frequência*. O total de 500 é obtido somando-se o número de homens e de mulheres.

É interessante também expressar esses resultados em forma relativa, ou seja, considerar a frequência de cada categoria em relação ao total:

$$\frac{270}{500} = 0,54$$

ou seja, 54% dos funcionários da empresa são do sexo masculino e

$$\frac{230}{500} = 0,46$$

ou seja, 46% dos funcionários são mulheres. A **Tabela 1.1** apresenta a versão completa.

Tabela 1.1: Distribuição do número de funcionários por sexo

Sexo	Frequência Simples	
	Absoluta	Relativa
Masculino	270	0,54
Feminino	230	0,46
Total	500	1,00

Note que a soma das frequências relativas é sempre 1, enquanto a soma das frequências absolutas deve ser igual ao número total de elementos sendo investigados.

De maneira análoga, obteríamos a **Tabela 1.2** para a variável estado civil. Note que, aí, a frequência relativa está apresentada em forma percentual, ou seja, multiplicada por 100. Por exemplo, para os casados temos:

$$\frac{280}{500} \times 100 = 0,56 \times 100 = 56\%$$

Em geral, essa é a forma mais usual de se apresentarem as frequências relativas e neste caso, a soma deve dar 100%.

Tabela 1.2: Distribuição do número de funcionários por estado civil

Estado Civil	Frequência Simples	
	Absoluta	Relativa %
Solteiro	125	25,0
Casado	280	56,0
Divorciado	85	17,0
Viúvo	10	2,0
Total	500	100,0

Exemplo 1.1

Consideremos que, na situação descrita anteriormente, os dados tenham sido levantados por departamento, para depois serem totalizados. Para o Departamento de Recursos Humanos, foram obtidas as seguintes informações:

Nome	Sexo	Estado civil	Número de dependentes
João da Silva	M	Casado	3
Pedro Fernandes	M	Viúvo	1
Maria Freitas	F	Casada	0
Paula Gonçalves	F	Solteira	0
Ana Freitas	F	Solteira	1
Luiz Costa	M	Casado	3
André Souza	M	Casado	4
Patrícia Silva	F	Divorciada	2
Regina Lima	F	Casada	2
Alfredo Souza	M	Casado	3
Margarete Cunha	F	Solteira	0
Pedro Barbosa	M	Divorciado	2
Ricardo Alves	M	Solteiro	0
Márcio Rezende	M	Solteiro	1
Ana Carolina Chaves	F	Solteira	0

Para pequenos conjuntos de dados, podemos construir a tabela à mão e para isso precisamos contar o número de ocorrências de cada categoria de cada uma das variáveis. Varrendo o conjunto de dados a partir da primeira linha, podemos ir marcando as ocorrências da seguinte forma:

Masculino		Solteiro	
Feminino		Casado	
		Divorciado	
		Viúvo	

Obtemos, então, as seguintes tabelas:

Sexo	Frequência Simples		Estado Civil	Frequência Simples	
	Absoluta	Relativa %		Absoluta	Relativa %
Masculino	8	53,33	Solteiro	6	40,00
Feminino	7	46,67	Casado	6	40,00
Total	15	100,0	Divorciado	2	13,33
			Viúvo	1	6,67
			Total	15	100,00