

- > A consolidação vai desde o recebimento dos questionários preenchidos até a compilação das respostas (quantitativas ou qualitativas).
- > A análise consiste na reflexão embasada do pesquisador frente aos dados levantados.
- > A consolidação dos dados pode ser feita por meio da construção de uma matriz. A matriz deve ser construída de forma a atender às especificidades da pesquisa.

Teste seu conhecimento

- > Cite três vantagens e três limitações das pesquisas quantitativas.
- > Explique por que o questionário é o instrumento mais comum para coleta de dados quantitativos.
- > Como se deve organizar o grupo de questões de um questionário?
- > Quais os aspectos mais relevantes a serem considerados na redação das questões de um questionário?
- > Diferencie questões fechadas, questões abertas e questões estratificadas.
- > Conceitue população e amostra.
- > Explique por que, em muitas pesquisas quantitativas em educação, se trabalha com a amostra e não com a população?
- > Estabeleça as principais diferenças entre métodos probabilísticos e não probabilísticos.
- > Por que frequentemente a única solução é utilizar um método não probabilístico?
- > Justifique a importância de se realizar um estudo-piloto antes de executar a pesquisa.
- > Explique as quatro fases de execução de uma pesquisa.

Capítulo 8

Analizando Dados Quantitativos

Estudo de caso

Censo escolar 2008

Em 2009 o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP divulgou o censo escolar. Dentro deste censo é prevista a contagem de professores que atuam na educação básica. Para levantar esta informação, que faz parte de uma pesquisa quantitativa, o INEP utiliza questionários enviados aos estados e municípios, normalmente em meio digital, para que eles identifiquem esses professores. Os dados encontrados foram os seguintes:

Educação Infantil				Ensino Fundamental		
Total	Educação Infantil	Creche	Pré-Escola	Ensino Fundamental	Anos Iniciais	Anos Finais
1.977.978	369.698	127.657	258.225	1.377.483	721.513	783.194

Como foi visto no capítulo anterior, o uso de questionários é uma das formas mais comuns para se levantar dados em uma pesquisa quantitativa. Com base nos dados do censo escolar aqui apresentado, responda:

- > Por qual tipo de organização esses dados já passaram até chegarem à forma com que aparecem nesta página?
- > A que conclusões o pesquisador pode chegar analisando a tabela apresentada?
- > Como apresentar e analisar dados quantitativos em pesquisas na área da educação?

Conceitos para entender a prática

A análise dos dados é um momento extremamente delicado. É muito importante considerar que nem todos os dados poderão ser interpretados seguindo o mesmo modelo. É preciso identificar, dentro da realidade da pesquisa e da pergunta de partida que se formulou, qual a forma adequada de se analisar determinada questão.

Frequentemente os dados coletados em uma pesquisa quantitativa precisam ser convertidos em símbolos, para que se possa fazer sua análise. Esses símbolos são chamados de *escalas de mensuração*: nominal, ordinal, intervalar e de razão.

A escala nominal tem o objetivo de auxiliar na contagem de uma determinada questão. Neste tipo de escala o símbolo representa o evento. Quando se tem a necessidade de conhecer o gênero dos respondentes em uma dada pesquisa, é possível atribuir um valor para cada opção. Por exemplo:

Sexo: Masculino – 1
 Feminino – 2

A utilização de números neste caso simplifica a identificação da frequência dessas respostas para a questão.

A escala ordinal é utilizada para classificar um grupo de respostas atribuindo ordenação. Aqui não se busca somente identificar a frequência, mas a ordem na qual esta é apresentada. Um exemplo seria o uso da escala de Likert, vista no capítulo anterior. Neste caso a ordenação, transposta a números, poderia ser:

- 1 Concordo plenamente
- 2 Concordo parcialmente
- 3 Sem opinião
- 4 Discordo parcialmente
- 5 Discordo plenamente

Perceba que na escala ordinal não se conhece a distância entre uma opção e outra. Não possível “medir” o quanto quem concorda parcialmente se distancia de quem concorda plenamente. Sabemos que os que selecionam o número dois concordam em parte com determinada afirmação, mas não podemos garantir que todos que escolheram esta opção concordam na mesma medida.

A escala intervalar é semelhante a uma escala ordinal, diferenciando-se principalmente pelo fato de se conhecer em detalhes a distância entre as diversas opções. Ao se desejar avaliar as notas de uma turma, supondo que será utilizado um modelo com intervalo de 0,5

ponto e que não serão atribuídos valores quebrados, é possível montar uma escala intervalar. Veja no exemplo:

Nota	Frequência
10	5
9,5	3
9,0	4
8,5	1
8,0	6
97,5	8
7,0	1

A escala de razão é similar à escala intervalar, diferenciando-se por possuir um zero absoluto. É utilizada para medir idade e peso, por exemplo.

Todas as análises de dados quantitativos implicam a escolha de um modelo de escala para apresentação dos dados. A escolha está ligada ao tipo de análise e ao tipo de dados que serão trabalhados. Para facilitar a seleção é possível comparar o uso de cada uma delas.

Escala nominal	São variáveis categóricas. Não permitem que se avalie comparativamente se uma é maior que a outra.
Escala ordinal	Variáveis que permitem ordenar, mas não explicitam a distância que há entre uma opção e a outra.
Escala intervalar	Similar à escala ordinal, apresenta intervalos similares e conhecidos.
Escala de razão	Escala intervalar com a presença do zero (ausência) e que não permite valores abaixo deste.

Conceitos (muito) básicos de estatística

Nem sempre, quando se opta por uma pesquisa quantitativa, será necessário aplicar técnicas estatísticas na análise dos dados. Esses dados poderão ser apresentados em tabelas ou gráficos, permitindo uma análise eficiente. Contudo, é muito comum a necessidade de se dominar algumas análises estatísticas para que se chegue a determinadas conclusões. Como nas pesquisas em educação o modelo de análise atrelado às questões estatísticas não é tão utilizado, serão apresentadas as informações básicas.

Medidas de localização central

Quando se tem um conjunto de dados numéricos é preciso identificar o ponto central, por-
será importante compreender como se dá a distribuição das informações e o quanto
is se afastam do centro. Pense em uma pesquisa que visa, por exemplo, comparar as
liações de diversas turmas. Tomemos as notas da turma A:

10	8,5	8,5	7,0	10
7,0	9,5	7,0	9,0	5,0
3,5	7,5	3,0	4,0	9,0
8,5	8,0	8,0	9,5	7,0
9,0	4,5	4,0	5,5	6,5

você quiser saber a localização central dos valores, pode recorrer a três indicadores:
dia, mediana e moda.

1. Média é a soma dos valores apresentados dividida pelo seu número total.
2. Mediana é o ponto que separa 50% dos valores mais baixos dos 50% mais altos.
3. Moda é o valor que aparece com a maior frequência.

omarmos todos os valores apresentados nesse quadro, encontraremos 179. Considerando
foram apresentadas 25 informações, a média seria:

$$\frac{179}{25} = 7,16$$

se caso uma análise inicial dos dados poderia especificar que as notas desta turma, na
dia, são 7,16.

Localização central

Existem dois tipos de média: a aritmética simples e a ponderada. A média aritmética é a
oma das observações dividida pelo número da frequência. A média ponderada é a soma
das observações multiplicada por seus pesos (ou ponderações), dividida pelo total das
observações ponderadas.

Se uma avaliação final de uma dada disciplina considera um trabalho, uma prova e
uma apresentação, cada etapa valendo de zero a dez e o resultado final do aluno foi
trabalho = 5, prova = 6, apresentação = 10, então a média dele será $5 + 6 + 10/3 = 7$.
Perceba que, neste caso, considera-se que todas as etapas da avaliação têm o mesmo

"peso", por isso esta média é chamada de aritmética simples. Contudo, este procedimento
poderia ter feito a seguinte opção: prova vale 5, trabalho vale 3 e apresentação vale 2.
Neste caso, qual seria a nota final? Esta é a situação na qual se usa a média ponderada.
Para chegar ao valor, seria feita a seguinte conta: $(5 * 5) + (6 * 3) + (10 * 2)/(5 + 3 + 2) \Rightarrow$
 $25 + 18 + 20/10 \Rightarrow 6,3$. Agora a nota é menor do que quando feita a média aritmética,
porque a maior nota (10) é a que teve o menor peso (2).

A mediana representa o valor central. Para encontrar a mediana dos dados apresentados
aqui teríamos que encontrar o valor central. Por termos 25 dados, o valor central seria o 13º
(12 ficariam abaixo dele e 12 acima).

3	3,5	4	4	4,5	5	5,5	6,5	7	7	7	7,5	8	8	8,5	8,5	9	9	9,5	9,5	10	10
---	-----	---	---	-----	---	-----	-----	---	---	---	-----	---	---	-----	-----	---	---	-----	-----	----	----

O 7,5, no exemplo, representa a mediana dos dados. Já a moda é o valor que aparece com
a maior frequência.

Valor	Frequência
10	2
9,5	2
9	3
8,5	3
8	2
7,5	1
7	4
6,5	1
5,5	1
5	1
4,5	1
4	2
3,5	1
3	1

Nesta tabela o valor que aparece mais vezes é o 7. Portanto, 7 é a moda nesta distribuição
de frequência.

Afinal, quando utilizar cada uma das medidas de localização central? Esta pergunta não
tem uma resposta pronta. A escolha do método de localização central vai depender profun-
damente do tipo de análise que se quer. Vejamos alguns exemplos:

Suponha que em uma sala de aula com 20 alunos haja 15 alunos com 18 anos de idade e 5 alunos com 8 anos de idade. Ao se desejar expor a distribuição central desta turma, qual seria a melhor medida?

Para facilitar, podemos adiantar que:

Média aritmética: $(15 \cdot 18) + (5 \cdot 8)/20$. Então, $270 + 40/20 \Rightarrow 15,5$.

Moda: 18 anos (que tem a frequência 15, sendo assim a maior).

Mediana: 18 anos (já que na organização dos dados, o aluno que ficasse na posição central teria esta idade).

Ora, se usarmos a média para caracterizar esta turma partiremos do princípio de que todos os elementos que compõem o grupo se aproximam da idade de 15,5 anos. Já ao usar a moda fica claro que a maioria dos alunos tem 18 anos. Então, provavelmente usar a moda como medida central será mais adequado.

Medida de dispersão

Para facilitar a localização central de um dado pode não ser suficiente para sua análise. Veja os conjuntos de dados a seguir:

$A = \{5, 5, 5, 5, 5\}$

$B = \{8, 4, 1, 10, 1\}$

A média aritmética dos dois conjuntos de dados é 5. Mas é fácil perceber que os dados do conjunto A são muito mais uniformes que os do conjunto B. Essa uniformidade é fruto da proximidade de distância entre a média e os dados do conjunto. Veja que a distância entre um dos elementos e a média no conjunto A é zero ($5 - 5 = 0$). Já no conjunto B esta distância varia de 5 ($10 - 5 = 5$) a -4 ($1 - 5 = -4$). Para identificar esta distância a estatística utiliza medidas de dispersão. As principais medidas de dispersão são variância, desvio-padrão e amplitude total. Como o objetivo aqui é oferecer informações básicas, fiquemos com o desvio-padrão, que é o mais utilizado.

O desvio-padrão é a raiz quadrada da média aritmética do quadrado dos desvios, e é representado por σ ou S . Se quiséssemos calcular o desvio-padrão do conjunto B, aqui representado, seria necessário avaliar quanto cada um dos elementos se distancia da média.

Média (a)	Elementos (b)	$[(b) - (a)]^2$
5	8	9
5	4	1
5	1	16
5	10	25
5	1	16

Para os valores de $[(b) - (a)]^2$ é 67. Daí calcula-se a média aritmética ($67/5$) e encontramos o desvio-padrão 3,66. Deste valor é extraída a sua raiz quadrada, que é 3,66. Portanto, o desvio-padrão

deste conjunto é 3,66. Significa dizer que cada um dos elementos do conjunto se afasta na média 3,66 unidades da média.

Como foi dito anteriormente, não é preciso dominar sólidos conceitos de estatística para se fazer uma pesquisa quantitativa, desde que não haja muitos dados envolvidos. É possível apresentar esses dados usando tabelas e gráficos e, ainda assim, se fazer compreender muito bem, além de tirar excelentes conclusões, de forma a ter sucesso em uma investigação científica. Então, vamos ver como apresentar esses dados.

Apresentando dados de questões com apenas uma opção

A apresentação dos dados de um questionário é mais simples quando o respondente tem apenas uma opção para marcar em cada questão. Nesse caso o objetivo é identificar a quantidade de respostas que são escolhidas por questão. Para isso é possível usar medidas simples, como o percentual de respondentes que faz uma determinada opção.

Você sabia?

Percentual é uma medida que descreve quantas vezes uma ocorrência aparece a cada cem. Se, por exemplo, uma ocorrência se repete vinte vezes em um total de cem, temos 20% (vinte por cento).

Veja uma questão de um questionário que utiliza essa lógica:

Em qual das disciplinas você sente mais dificuldade?

A Língua Portuguesa

B Matemática

C História

D Biologia

E Física

Suponha que o questionário que contém essa pergunta tenha sido aplicado para uma amostra de trinta participantes. Na contagem dos resultados observou-se o seguinte:

- > 12 escolheram a opção A;
- > 10 escolheram a opção B;
- > 8 escolheram a opção C;
- > 0 escolheu a opção D.

Se o total de respondentes é trinta, podemos dizer que:

- > 40% escolheram a opção A;
- > 33% escolheram a opção B;

- > 27% escolheram a opção C;
- > 0% escolheu a opção D.

as informações nos permitem realizar uma análise que mostre que, na opinião dos alunos, há mais dificuldade em compreender língua portuguesa do que matemática. Perceba que não é possível dizer, baseado exclusivamente nesta questão, as causas para a realidade indicada. Mas é possível, a partir daí, engendrar uma nova investigação no sentido de esclarecer a razão disso.

Obviamente o exemplo citado foi retirado do contexto, mas em um questionário completo provavelmente a próxima questão serviria para complementar a resposta que se escolheu nesta. Por isso, dando seguimento à compreensão da apresentação de dados de questão com apenas uma opção de resposta, imagine que a questão seguinte fosse a apresentada a seguir:

Por que motivo você julga que a disciplina escolhida na questão anterior é a mais difícil?

- A *Porque o conteúdo é muito difícil.*
- B *Porque não temos ambientes e recursos adequados para aprender.*
- C *Porque o material didático não é bom.*
- D *Porque o professor não ensina direito.*
- E *Porque os conteúdos são apresentados de forma desorganizada.*

A que interessante: a resposta a esta pergunta, além de completar a resposta da pergunta anterior, já permite que o pesquisador faça inferências sobre as causas e as consequências do evento que está pesquisando. Suponha, hipoteticamente, que a maioria dos respondentes optasse pela letra E (conteúdos apresentados de forma desorganizada). Seria possível dizer que há algum problema na apresentação dos conteúdos. Alguns problemas são:

- > Currículo inadequado.
- > Metodologia de ensino pouco atrativa.
- > Professor desorganizado.
- > Recursos insuficientes ou utilizados de forma errada.
- > Outros.

possíveis problemas já poderiam ter sido apresentados no questionário. Mas se não tiver sido, haveria maior facilidade em se deixar uma recomendação para futuras pesquisas.

O pesquisador

não deve "super" o motivo que leva uma pessoa a dar uma resposta. É preciso interpretar somente o que as respostas informam. As especulações não são consideradas científicas, por isso não devem ser parte de uma investigação acadêmica.

Você sabia?

Dificilmente uma pesquisa consegue esgotar um assunto. Por isso é sempre recomendado que o pesquisador declare o que ele não pôde ou não conseguiu abordar, sugerindo que futuras pesquisas busquem dar conta daquele tópico.

O uso de recursos computacionais para a construção de questionários é muito rico, porque permite que cada uma das questões se desdobre em diversas opções, o que é inviável quando se aplica um questionário em papel. Na situação em que o pesquisador possui hipóteses diferenciadas para cada uma das questões essa estratégia é interessante. Retomando o exemplo citado aqui, imagine que a escolha de cada uma das disciplinas como a mais difícil suscitasse um conjunto diferenciado de hipóteses. Nesse caso teríamos uma questão seguinte diferente, dependendo do que foi escolhido nesta primeira pergunta. Este recurso é interessante na medida em que permite uma análise mais aprofundada das considerações que são trazidas pelo público convidado a participar da etapa de coleta de dados.

Atenção!

Quando se calcula o percentual em uma pesquisa só se deve considerar o total de questionários respondidos. Por exemplo, se foram enviados 120 questionários, mas apenas 100 foram respondidos e uma opção de uma questão recebeu 80 respostas, então se tem 80% de concordância (80/100), não 66% (80/120).

Apresentando dados de questões com múltiplas opções

Vimos anteriormente que é possível construir um questionário no qual uma pergunta busque sua explicação na resposta da pergunta seguinte. Contudo, não é incomum encontrar questionários que estabeleçam um paralelo entre a causa e a consequência de um evento. Por exemplo, o pesquisador pode querer compreender a relação que existe entre a evasão escolar e a distorção idade/série. Para isso ele formula a seguinte questão:

Com relação à evasão e à distorção idade/série, marque a opção que melhor apresente esta relação:

- A *Quanto maior a evasão, maior a distorção idade/série.*
- B *Quanto maior a evasão, menor a distorção idade/série.*
- C *Quanto menor a evasão, maior a distorção idade/série.*
- D *Quanto menor a evasão, menor a distorção idade/série.*

É possível notar que neste caso se convida o respondente a trazer uma resposta que congrega duas vertentes: causa e consequência. A análise desse tipo de questão pressupõe um olhar atento do investigador. Caso se opte pelas letras A e D com alta frequência, será possível infe-

ir que a evasão justifica a distorção idade/série de maneira linear (quanto maior a primeira, maior a segunda. Quanto menor a primeira, menor a segunda). Contudo, se houver alta escolha das opções B e C, não se pode fazer esse tipo de análise de forma direta, já que pelo raciocínio lógico não faz sentido que ao reduzir a evasão se aumente a distorção idade-série.

Por isso, esse tipo de questão sempre exigirá validação nas questões a seguir. Ou seja, será necessário apresentar outras perguntas que levem o respondente a confirmar o que se apresentou justificando um resultado que, à primeira vista, não teria sentido.

Esse tipo de questionário costuma enriquecer as pesquisas, mas exige muito mais trabalho do investigador para que se possa chegar a alguma conclusão. Isso acontece porque, na comparação das respostas, muitas vezes perde-se a linha de raciocínio e não é possível inferir uma conclusão. Tomemos a questão citada como exemplo. Imagine que a maioria dos que receberam o questionário o devolveram marcando a letra A (quanto maior a evasão, maior a distorção idade/série) e que, a seguir, tenha sido apresentada uma nova questão para confirmá-la.

Os alunos com distorção idade/série, na sua percepção, tendem a:

- A Se esforçar para terminar o curso.
- B Esmorecer e aumentar a distorção.
- C Abandonar o curso.
- D Buscar caminhos diferenciados para lidar com a distorção (cursos noturnos, etc.), conseguindo não ser reprovados.

e nesta questão que visava confirmar a anterior a maioria responder letra C, terá confirmado a resposta anterior. Mas se a escolha recair na letra D, haverá uma contradição interna na pesquisa. Se a evasão leva à distorção idade, como os alunos não evadem e ainda conseguem evitar novas reprovações?

Este tipo de conflito interno de pesquisa não é incomum e exige habilidade para efetuar análises. Por isso, inclusive, as pesquisas quantitativas que usam questionário para coletar dados costumam dedicar um espaço para o uso de questões abertas. Talvez a explicação para o caso supracitado esteja na diferença do perfil dos alunos do passado e dos alunos do presente.

Seguindo o exemplo anterior, no qual o respondente marcou A (maior evasão, maior distorção) e em seguida D (aluno se esforça e não é reprovado), pode ser que nos comentários certos você encontre:

Os alunos que cursavam o ensino fundamental até o ano de 2009 eram adultos, empregados, chefes de família em sua maioria, o que explica a alta evasão e a alta distorção, pois já chegavam muito fora da série na qual deveriam estar. A nova turma tem um perfil diferente: alunos jovens, sem compromissos familiares, trabalhando em meio período, por isso têm se esforçado mais e evitado a reprovação.

Agora sim, é possível analisar os resultados explicitando que as respostas trouxeram uma confusão quanto à questão do tempo. E isso possibilita o investigador a ter conclusões mais lógicas e coerentes com a realidade.

Apresentando dados de questões por escala

Existem dois tipos de questão por escala e, consequentemente, pelo menos duas formas de se fazer esta análise. O primeiro tipo utiliza o modelo da escala de Likert. Nesse caso, as opções de resposta para a questão costumam ser:

- 1 Concordo totalmente.
- 2 Concordo parcialmente.
- 3 Não concordo nem discordo.
- 4 Discordo parcialmente.
- 5 Discordo totalmente.

É comum encontrar estudos nos quais um valor é atribuído a cada uma dessas opções de resposta e, daí, o investigador consegue calcular indicadores estatísticos como desvio-padrão, média, variância... Essa abordagem para este tipo de escala não procede! Perceba que não há nenhuma relação de grandeza nas opções oferecidas ao respondente. Portanto, concordar parcialmente não significa concordar três ou quatro vezes menos do que concordar totalmente.

Por isso, a escala de Likert deve ter sua análise realizada por meio da contagem do número de respondentes por opção e, a partir de então, o pesquisador deve chegar às suas conclusões com base nas hipóteses que havia formulado ou com base no referencial que sustenta seu trabalho.

Por exemplo, foi feita a seguinte afirmação a um grupo de vinte professores em um questionário de pesquisa, solicitando que preenchessem como encaravam tal afirmação:

A participação dos pais nas reuniões escolares é fundamental para o sucesso da aprendizagem dos alunos.

As respostas encontradas foram:

Opções	Quantidade de respostas
Concordo totalmente	10
Concordo parcialmente	5
Não concordo nem discordo	3
Discordo parcialmente	1
Discordo totalmente	1

Para se fazer a análise desta situação será preciso explicitar de forma descritiva como a distribuição da frequência se deu.

Você sabia?

A distribuição da frequência em estatística representa o agrupamento dos dados de forma que se facilite a compreensão das classes. Ou seja, os dados são agrupados utilizando-se uma tabela, por exemplo, para facilitar a visualização do resultado.

este exemplo, a frequência mostra que 75% dos professores concordam de alguma forma com a afirmação, 15% não têm opinião formada e 10% de alguma forma discordam. Com base nos dados apresentados só se pode chegar até aqui. Lembre-se de que não é possível ter informações que não foram dadas pelos respondentes. Então, o aprofundamento da interpretação dessas informações será feito pelo pesquisador esclarecendo que tais análises são dadas com base no referencial teórico dado.

Um outro tipo de questão é aquele que atribui um valor numérico à sensibilidade do respondente. Se no exemplo dessa questão a escala de Likert fosse substituída pela escala de se segue, a análise seria outra.

0	9	8	7	6	5	4	3	2	1	0
---	---	---	---	---	---	---	---	---	---	---

no caso anterior não era possível estabelecer uma relação de valor entre as opções; neste caso é diferente. A estrutura da resposta já é um valor, o que permite que sejam feitas análises que identifiquem indicadores centrais de distribuição, bem como a estrutura da distribuição dos dados. Aqui podem ser utilizadas as fórmulas apresentadas anteriormente, pois as afirmações que considerem média, moda e variância fazem sentido.

Nesse caso, suponha que a distribuição das respostas ficou da seguinte forma:

Valor	Frequência
10	5
9	2
8	3
7	2
6	2
5	0
4	0
3	1
2	2
1	1
0	2

Aqui há algumas opções para realizar a análise. Podemos buscar os pontos centrais, que seriam:

- Moda: 10
- Mediana: 8
- Média: 6,3

Observação: lembre-se de que, nesse caso, para se calcular a média foi multiplicado o valor pela frequência, e o resultado foi somado e dividido pela frequência total, que é 20.

Essas informações são suficientes para aprofundar a análise da distribuição das respostas e se buscar novas conclusões. Além disso, é possível representar esses resultados por meio de gráficos dos mais diversos tipos, o que será apresentado posteriormente.

Apresentando dados de questões que envolvem ordenação

Diversas questões costumam solicitar que o respondente chegue a uma ordem frente a um problema. Por exemplo, se é perguntada qual a maior dificuldade com relação aos alunos recém-chegados à universidade, é possível pedir que o respondente marque de 1 a 5, sendo 1 o que mais se relaciona à dificuldade e 5 o que menos se relaciona.

- () Adaptação com novo local.
- () Dificuldade com aumento significativo na quantidade de conteúdos.
- () Ausência da sólida responsabilização da família.
- () Sentimento de liberdade.

Em situações como esta não existe uma resposta certa e uma resposta errada. Apenas uma ordem de concordância. Neste caso também é preciso lembrar que os valores 1, 2, 3, 4 e 5 não permitem análises estatísticas porque não se conhece com exatidão o intervalo. São valores referenciais que facilitam a compreensão no momento da análise, mas não se prestam a desenvolver relações estatísticas. Portanto, falar em média nesta situação, nem pensar!

A análise desses dados pode ser feita de diversas maneiras. A descritiva, como foi visto anteriormente, na qual os dados são apresentados para posteriormente o investigador chegar às suas conclusões, principalmente com base no referencial que sustentou a pesquisa. A categórica, na qual as respostas são agrupadas por categorias, e a relacional, na qual as escalas são comparadas com outras questões.

A análise descritiva tem o objetivo principal de descrever detalhadamente os dados e se apoiar em pesquisas anteriores que conduzam a uma conclusão.

A pesquisa categórica tem o objetivo de incluir os dados em categorias. Por exemplo, se as opções são:

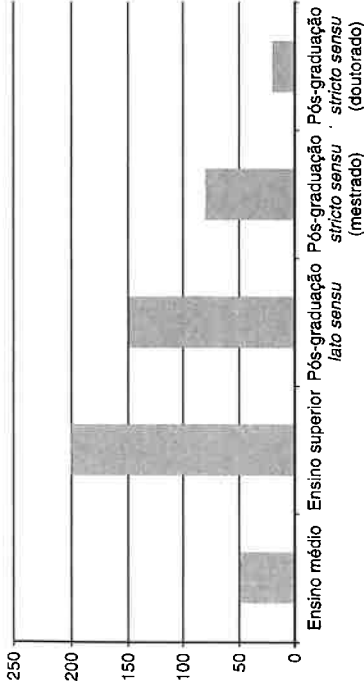
possível considerar apenas três categorias: bom, regular e ruim. Ou, em alguns casos, apenas duas (bom e ruim), dependendo do interesse de quem investiga. Nesse caso, no entanto, a migração de categorias exige habilidade e conhecimento do pesquisador. Se essas ações fossem transformadas em apenas duas categorias, para onde iria a opção “regular”? a teria que ir para o lado que melhor representa regular, e isso teria que ser identificado em estudo-piloto, quando o questionário ainda estava em fase de teste. A simples escolha não retrata a realidade e coloca em cheque a validade dos resultados da pesquisa. A análise relacional das questões que buscam ordenação se sustenta na comparação das respostas que são atribuídas a uma questão com as respostas de outras questões. Nesses casos, o que se busca é a lógica nas respostas ou a relação que se estabelece entre elas.

belas e gráficos como suporte à apresentação de resultados

pesquisas quantitativas têm a vantagem de permitir que a realidade se torne quantificável. Se possível transformar um determinado evento em um número, desde que exista uma relação de aneaza, é possível tornar essas informações mais visuais. Suponha que uma pesquisa se prestou a identificar o nível de formação dos professores do ensino fundamental em uma dada cidade; 10 professores responderam à pesquisa e os dados foram apresentados na seguinte tabela:

Formação	Quantidade
Ensino médio	50
Ensino superior	200
Pós-graduação lato sensu	150
Pós-graduação stricto sensu (mestrado)	80
Pós-graduação stricto sensu (doutorado)	20

ta mesma informação pode aparecer de outras formas.



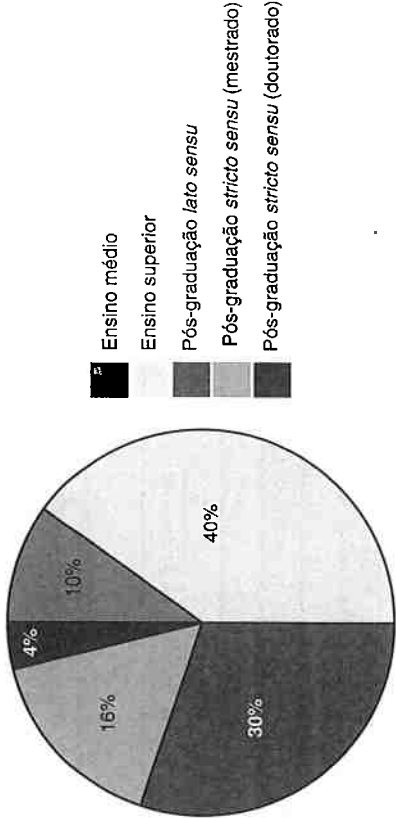
Esta apresentação em gráfico de barras mostra a distribuição da frequência, o que permite fazer diversas inferências mesmo sem aprofundar a análise. Por exemplo, é fácil perceber que a maioria dos professores se encontra no grupo que possui curso superior ou pós-graduação lato sensu. Ao mesmo tempo, é simples notar que há um número reduzido de docentes com doutorado e, em seguida, com ensino médio.

Atenção!

É preciso ter cuidado na escolha do tipo de gráfico que comporá a redação da análise de dados. Nem todos os tipos servem para todas as questões. Os principais tipos de gráfico e seus usos são:

- Gráfico de colunas – mostra alteração de frequência em um período de tempo ou compara itens.
- Gráfico de linhas – mostra uma evolução ao longo do tempo. Ajuda a inferir tendências.
- Gráfico de pizza – mostra o tamanho da participação de um dado em relação ao todo.
- Gráfico de área – mostra a mudança no decorrer de um período.
- Gráfico de dispersão – mostra as relações entre valores numéricos em dada sequência.

Outra forma de apresentar as informações neste exemplo é utilizando o gráfico de pizza.



Neste caso, a relação quantitativa é reforçada em relação ao todo. É possível perceber com muita facilidade que “ensino superior” é o item que mais aparece, seguido pela pós-graduação lato sensu.

vale lembrar:

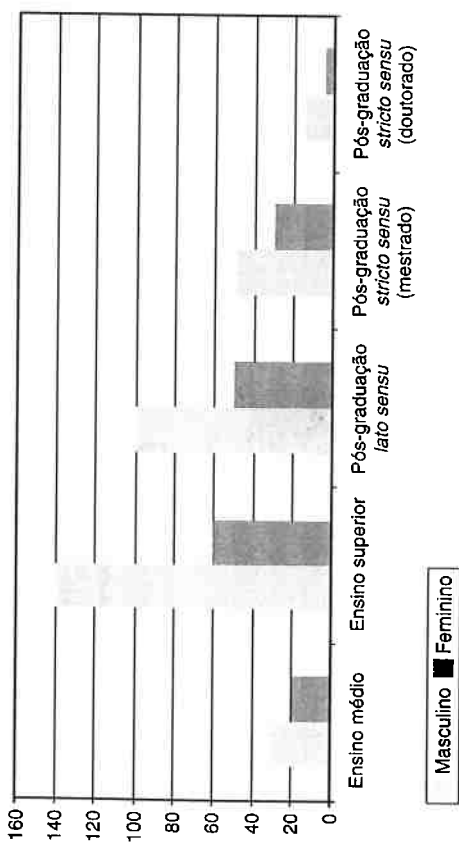
As planilhas eletrônicas, que fazem parte dos principais softwares, facilitam a construção de gráficos e tabelas.

Um ponto que se deve ter bem claro é que gráficos e tabelas não falam por si sós. Eles utilizam os recursos visuais para dar credibilidade às interpretações feitas pelo pesquisador, além de se apoiarem no que está disponível para concluir. Mas é papel do pesquisador chegar a uma análise que, algumas vezes, pode contradizer o gráfico. A análise sempre é sustentada por duas vertentes: os dados trazidos no momento da coleta e o referencial teórico utilizado. Para pesquisadores extremamente experientes a vivência na coleta e nas análises será um diferencial.

Ainda sobre o uso de gráficos e tabelas, suponha que a mesma informação anterior seja agora apresentada por separação entre os gêneros masculino e feminino e que tenhamos a seguinte tabela:

Formação	Masculino	Feminino
Ensino médio	30	20
Ensino superior	140	60
Pós-graduação lato sensu	100	50
Pós-graduação stricto sensu (mestrado)	50	30
Pós-graduação stricto sensu (doutorado)	15	5

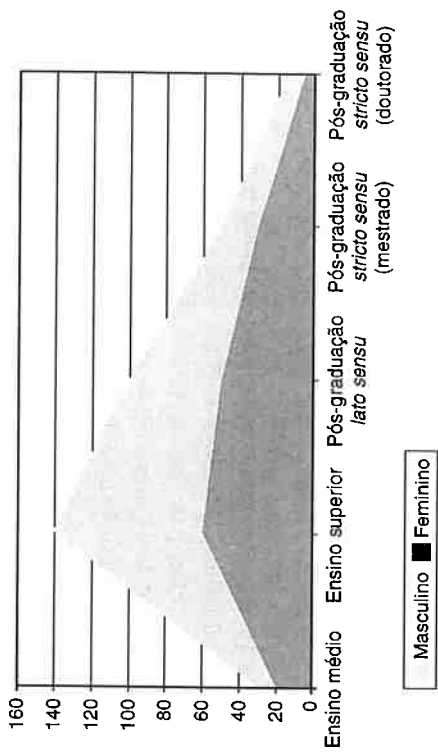
Nesta nova tabela é gerado o seguinte gráfico de barras:



Este gráfico mais uma vez permite diversas novas análises. A começar, é possível perceber que há um número bem maior de professores do gênero masculino do que do gênero feminino. No entanto, a maior quantidade não parece se refletir em uma formação melhor, já que as professoras têm a curva de tendência de formação similar à dos professores.

Note que por ter sido acrescentada mais uma variável, novas possibilidades de análise aparecem. E essas possibilidades aumentarão sempre que uma nova informação for inserida. Na situação em foco há diversas outras possibilidades não abordadas. Por exemplo, por que há menos professoras do que professores? Há relação com o projeto pedagógico da escola? Os dados apresentados não nos permitem fazer tais inferências, mas estas podem ser levantadas pelo pesquisador no sentido de enriquecer sua análise.

Para concluir, outra forma eficaz de se apresentar os dados da tal pesquisa, separando professores do sexo masculino dos professores do sexo feminino, seria o gráfico de área.



Perceba que a visualização deste gráfico reforça a informação de que há muito menos sujeitos do sexo feminino do que do masculino (são 335 contra 165).

As pesquisas que usam uma abordagem quantitativa parecem, inicialmente, conseguir traduzir a verdade da forma mais objetiva. Contudo, é frequente que se termine uma investigação com a sensação de que ainda há muito para se conhecer do fenômeno que foi investigado. Por isso é preciso sempre lembrar que não se deve ler nos dados informações que não existem. As conclusões que são feitas devem ser sempre baseadas nos dados que foram levantados e no referencial teórico que foi escolhido para dar suporte ao trabalho. Qualquer conclusão que não considere essas duas vertentes será vista com receio pelo meio acadêmico e pelos interessados nos resultados.

exercício de aplicação

Um aluno de graduação resolveu analisar os resultados das notas obtidas por todas as séries iniciais do ensino fundamental de uma dada escola. Seu objetivo maior era analisar em qual disciplina os alunos desse segmento encontravam maior dificuldade.

O resultado obtido, apresentando as médias aritméticas por disciplina e por série, foi o seguinte:

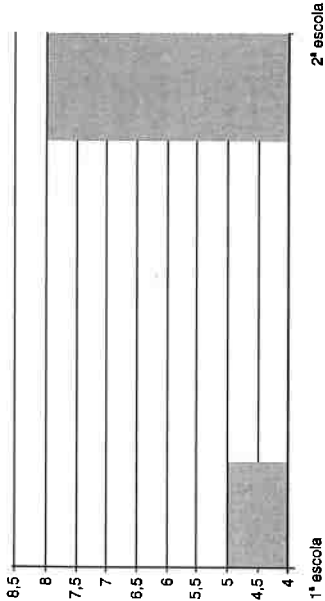
	Matemática	Português	Estudos Sociais	Ciências	Artes
1ª série	6,0	8,5	6,5	5,5	9,0
2ª série	5,0	8,0	7,0	9,5	9,5
3ª série	7,0	9,0	9,5	8,5	8,5
4ª série	8,5	7,5	9,0	6,5	8,5
5ª série	7,5	7,0	8,5	8,0	8,0

Com base nessas informações, tente responder às seguintes indagações:

- > Que medidas poderiam ser utilizadas para se apresentar esses dados?
- > Como esses dados poderiam ser apresentados sem se utilizar conceitos estatísticos?
- > É possível, com base nessas informações, responder à pergunta de pesquisa proposta pelo aluno (não se esqueça de considerar somente os dados apresentados)?
- > Quais as principais vantagens de se utilizar a abordagem quantitativa nessa pesquisa?
- > Quais as principais limitações de se utilizar essa abordagem nessa pesquisa?

para debater

Os quantitativos podem ser facilmente manipulados, ainda que se atenham a retratar mente a verdade. Isso é mais comum quando se apresentam os resultados em forma de gráfico. Pense, por exemplo, na comparação da média de avaliação de um dado indicador educacional de duas escolas e suponha que a primeira atingiu a nota 5, enquanto a segunda atingiu a nota 8. Vejamos como essa informação foi apresentada no gráfico:



Apesar de o valor real ter sido mantido, a mudança na unidade mínima da escala leva o leitor a concluir que a segunda escola está quase quatro vezes acima da primeira.

Você já viu notícias em jornais fazerem isso? Discuta as implicações éticas dessa distorção com seu professor e seus colegas de classe.

Resumo executivo

- > A análise de dados quantitativos está atrelada às necessidades específicas da pesquisa.
- > Os dados quantitativos utilizam frequentemente escalas para sua apresentação: nominal, ordinal, intervalar ou de razão.
- > É possível realizar análise de dados quantitativos em pesquisas na área educacional mesmo sem se dominar conhecimentos estatísticos com profundidade.
- > Quando se opta por utilizar a estatística para a análise de dados, dois conceitos são fundamentais: distribuição central e dispersão dos dados.
- > A distribuição central busca identificar o ponto central de um conjunto de dados. As mais comuns são a média, a moda e a mediana.
- > A distribuição de dados busca identificar o quanto cada elemento de um conjunto se afasta da medida central. Os mais comuns são o desvio-padrão, a variância e a amplitude total.
- > Para apresentar dados com apenas uma opção é possível utilizar somente o percentual, destacando a participação de uma opção no todo.
- > Questões com múltiplas opções exigem esforço do pesquisador para se realizar a análise, além da validação da resposta de uma questão em outras questões.
- > Questões por escala podem utilizar a escala Likert ou a escala numérica.
- > A escala Likert não permite a interpretação por meio de análises estatísticas, já que não é possível identificar o intervalo entre uma opção e outra.
- > A escala numérica pode ser interpretada por meio de análises estatísticas.
- > As formas mais simples de se apresentar os dados de uma pesquisa quantitativa são o uso de tabelas e o de gráficos.
- > Existem diversos tipos de gráfico. Cada um é utilizado para trabalhar um determinado tipo de análise.

Teste seu conhecimento

- > Defina escala nominal, ordinal, intervalar e de razão.
- > Defina com suas palavras o que são medidas de dispersão central e cite as três principais medidas.
- > Diferencie medidas de dispersão de desvio-padrão.
- > Explique como apresentar dados:
 - em questões com apenas uma opção;
 - em questões de múltiplas opções;

~ em questões com opções por escala;

~ em questões que envolvem ordenação.

> Diferencie tabelas de gráficos e diga, com suas palavras, quando usar cada um na apresentação dos dados.

Coletando Dados Qualitativos

Estudo de caso

Memórias Póstumas de Brás Cubas (Machado de Assis)

Nunca Virgília me parecera mais expansiva, mais sem reservas, menos preocupada dos outros e do marido. Não eram remorsos. Imaginei também que a concepção seria um puro invento, um modo de prender-me a ela, recurso sem longa eficácia, que talvez começava de oprimi-la. Não era absurda esta hipótese; a minha doce Virgília mentia às vezes com tanta graça!

Naquela noite descobri a causa verdadeira. Era medo do parto e vexame da gravidez. Padecera muito quando lhe nasceu o primeiro filho; e essa hora, feita de minutos de vida e minutos de morte, dava-lhe já imaginariamente os calafrios do patíbulo. Quanto ao vexame, complicava-se ainda da forçada privação de certos hábitos da vida elegante. Com certeza, era isso mesmo; dei-lho a entender, repreendendo-a, um pouco em nome dos meus direitos de pai. Virgília fitou-me; em seguida desviou os olhos e sorriu de um jeito incrédulo.

No trecho acima, extraído do livro de Machado de Assis, tenta-se compreender o que se passa na cabeça de uma mulher grávida. Virgília não parece feliz, e a causa não está, a princípio, aparente. Para descobri-la foi preciso ouvir o que ela tinha a dizer e comparar com uma atenta observação.

Em muitas situações na vida não é possível compreender a realidade por meio da mensuração dos eventos. Isso ocorre porque diversos fenômenos têm sua compreensão somente pelo ponto de vista do outro. Por isso não basta contar, é preciso observar, compreender, analisar o outro.

E na educação? Você acha que existem questões que não podem ser mensuradas, mas só podem ser compreendidas pela imersão na realidade e na forma de compreender o mundo de outras pessoas? Certamente sim. A essa forma de compreender a realidade chamamos de abordagem qualitativa ou, simplesmente, pesquisa qualitativa.