



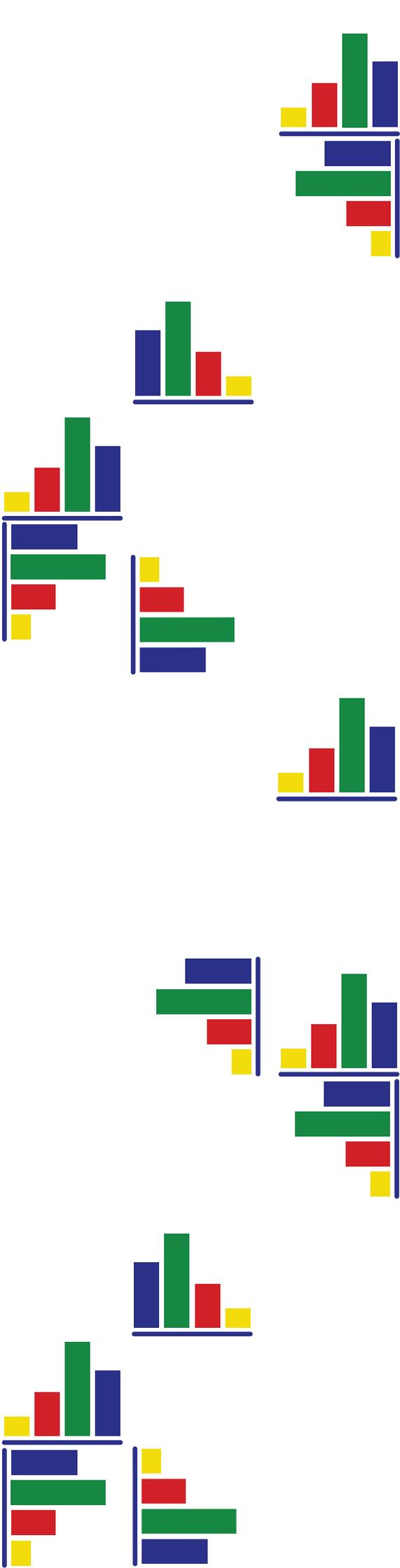
# ESTATÍSTICA PARA OS ANOS INICIAIS DO ENSINO FUNDAMENTAL

Irene Cazorla  
Sandra Magina  
Verônica Gitirana  
Gilda Guimarães

Biblioteca  
do Educador  
Coleção SBEM  
Volume **9**



**Sociedade Brasileira de  
Educação Matemática**



# ESTATÍSTICA PARA OS ANOS INICIAIS DO ENSINO FUNDAMENTAL

Irene Cazorla  
Sandra Magina  
Verônica Gitirana  
Gilda Guimarães



Sociedade Brasileira de Educação Matemática

2017

# ESTATÍSTICA PARA OS ANOS INICIAIS DO ENSINO FUNDAMENTAL

**Organizado:**

Irene Cazorla, Sandra Magina, Verônica Gitiran e Gilda Guimarães

**Editoração:**

Sociedade Brasileira de Educação Matemática

**Ilustrações do miolo:**

(<http://www.freepik.com>)

**Projeto Gráfico e Diagramação:**

Janaína Mendes Pereira da Silva

Dados Internacionais de Catalogação na Publicação (CIP)  
(Câmara Brasileira do Livro, SP, Brasil)

Estatística para os anos iniciais do ensino fundamental [livro eletrônico] / organizado Irene Cazorla ... [et al.]. -- 1. ed. - Brasília : Sociedade Brasileira de Educação Matemática - SBEM, 2017. -- (Biblioteca do Educador - Coleção SBEM ; 9) 6,5 Mb ; PDF.

Outros organizadores: Sandra Magina, Verônica Gitirana, Gilda Guimarães.  
Bibliografia  
ISBN: 978-85-98092-32-4

1. Educação - Finalidade e objetivos 2. Educação matemática 3. Estatística (Ensino fundamental) 4. Matemática - Estudo e ensino 5. Professores Formação  
I. Cazorla, Irene. II. Magina, Sandra. III. Gitirana, Verônica. IV. Guimarães, Gilda V. Série.

17-03296

CDD-519.5

Índices para catálogo sistemático:  
1. Estatística : Formação de professores : Educação matemática 519.5

Esse livro<sup>1</sup> foi escrito por quatro mulheres com formação acadêmicas distintas que viveram em diferentes lugares do Brasil e do mundo e, finalmente, se encontraram para escrever sobre ensino de estatística no Ensino Fundamental.

**Irene Maurício Cazorla** nasceu no Peru. Fez graduação em estatística na Universidade Nacional Mayor de San Marcos (Peru), mestrado em Estatística na Unicamp, doutorado em Educação na Unicamp e pós-doutorado em Educação Matemática na PUC-SP. Atualmente, é professora do mestrado em Educação Matemática da Universidade Estadual de Santa Cruz – UESC.

**Sandra Maria Pinto Magina** é psicóloga pela UFPE com mestrado em Psicologia Cognitiva na UFPE, doutorado em Educação Matemática na Universidade de Londres (Inglaterra) e pós-doutorado na Universidade de Lisboa. Atualmente é professora do mestrado em Educação Matemática da Universidade Estadual de Santa Cruz – UESC.

**Veronica Gitirana Gomes Ferreira** é bacharel e mestre em Matemática pela UFPE, com doutorado em Educação Matemática na Universidade de Londres (Inglaterra). Hoje, atua como professora da Pós-graduação em Educação Matemática e Tecnológica – EDMATEC da Universidade Federal de Pernambuco.

**Gilda Lisbôa Guimarães** é Pedagoga pela PUC-SP, com mestrado e doutorado em Psicologia Cognitiva pela UFPE, pós-doutorado na Universidade de Burgos (Espanha) e na Universidade Laval (Canadá). É professora da Pós-graduação em Educação Matemática e Tecnológica – EDMATEC da Universidade Federal de Pernambuco.

---

<sup>1</sup> Esta obra foi um produto desenvolvido no Projeto “A MATEMÁTICA ESCOLAR: CONSTRUÇÃO E TRANSFORMAÇÃO DO SABER MATEMÁTICO” de Intercâmbio Nacional do PROCAD-CAPES, entre a UNIBAN, UFPE e PUC-SP.

# SUMÁRIO

## **APRESENTAÇÃO 8**

### **INTRODUÇÃO 10**

Por que escrever este livro? 11

Estrutura do livro 12

## **Capítulo 1: A ESTATÍSTICA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL 14**

O que é Estatística? 14

O que é Educação Estatística? 15

O que é Tratamento da Informação? 15

A Estatística desde os anos iniciais do Ensino Fundamental 17

Fases da investigação científica 17

## **Capítulo 2: A IDENTIFICAÇÃO DO PROBLEMA 20**

Da observação e dúvidas à identificação do problema 20

Eu acho que ... do “achismo” às hipóteses 24

Questões didáticas da escolha do problema 25

## **Capítulo 3: A COLETA DE DADOS 27**

De onde se obtêm os dados? 27

A fonte de dados 32

O que é coletar os dados? 33

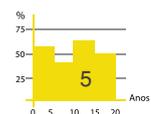
Os instrumentos de coleta de dados 34

Variáveis, seus tipos e sua operacionalização 36

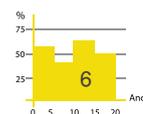
Tipos de variáveis 38

Análise de uma variável qualitativa 38

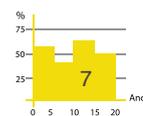
Análise de uma variável quantitativa 40



A necessidade de trabalhar com a classificação	41
<b>Capítulo 4: O TRATAMENTO DOS DADOS</b>	<b>45</b>
Do caos à organização de dados	45
A organização da frequência em pictogramas	45
A tabela de frequência (TDF)	48
A construção da planilha de dados	50
A construção da tabela de distribuição de frequências simples	55
A construção de TDF de dupla entrada	57
Construção de gráficos	58
Gráfico de barras/colunas	58
Gráfico de barras duplas	62
Gráfico de setores	63
Gráfico de linhas	64
Uso e cálculo das medidas estatísticas	67
A moda	67
A média aritmética	69
A Mediana	73
Amplitude total e ponto médio	77
Variáveis contínuas ou discretas que tomam muitos valores: medidas e gráficos	79
O cruzamento dos dados	82
Integração de tabelas, gráficos e medidas estatísticas	85
<b>Capítulo 5: A INTERPRETAÇÃO DOS RESULTADOS</b>	<b>87</b>
A pesquisa sobre o medo das crianças	87
A pesquisa sobre a altura das crianças	89
<b>Capítulo 6: PESQUISA EXPERIMENTAL: O JOGO DA MEMÓRIA</b>	<b>92</b>
<b>REFERÊNCIAS</b>	<b>105</b>
Sugestões de Leitura	106
<b>ANEXOS</b>	<b>109</b>
Anexo A: A variável Idade	109
Anexo B: A classificação com objetos padronizados	111



Anexo C: Construção da TDF sobre o medo	113
Anexo D: Construção do histograma	114
Anexo E: A média aritmética, aprofundamentos	116
Anexo F: Malhas	121



# APRESENTAÇÃO

A experiência na docência e na pesquisa, bem como a formação plural das autoras deste livro, possibilitou a produção de um material de qualidade e de expressiva contribuição para a Educação Estatística. Irene Cazorla, Sandra Magina, Verônica Gitirana e Gilda Guimarães abordam conceitos estatísticos, presentes nos anos iniciais do Ensino Fundamental, de extrema relevância, tanto para a constituição de cidadãos críticos e conscientes, quanto para a construção do pensamento científico.

Ao longo da obra, as autoras adotam uma linguagem simples e acessível. Em algumas partes do texto, estabelecem um diálogo com o leitor, o que torna a leitura agradável e envolvente. As atividades pedagógicas propostas abordam temas que integram o cotidiano de estudantes e professores, o que facilita a compreensão dos conceitos trabalhados. Tal perspectiva também evidencia a presença da Estatística em nossas vidas e deixa o professor mais confortável para abordar os conceitos básicos desta ciência na sala de aula.

As autoras ainda tratam de importantes aspectos inerentes à Estatística, tais como: diferença entre o achismo e a hipótese; fenômenos determinísticos e aleatórios; identificação do problema; coleta e fonte de dados; análise de variáveis qualitativas e quantitativas; e tratamento dos dados. Conforme salientado pelas próprias autoras, uma pesquisa não termina com a organização e o tratamento dos dados. Assim, Irene, Sandra, Verônica e Gilda presenteiam o leitor com um capítulo dedicado à interpretação dos resultados da pesquisa.

Para completar esta obra, apresentam uma atividade experimental, a qual possibilita ir além da proposta de ensinar os conceitos estatísticos através da pesquisa. Desta forma, por meio de um jogo, que pode ser confeccionado pelos próprios estudantes, enaltecem a importância do trabalho coletivo e interdisciplinar. A referida atividade, assim

como as demais apresentadas no livro, pode possibilitar a aprendizagem ativa, na qual a participação e ação do estudante são fundamentais para a construção do seu conhecimento.

Por fim, ressaltamos a relevante contribuição deste livro, o qual, sem dúvida, auxiliará os professores na organização de atividades pedagógicas que promovam a aprendizagem dos conceitos estatísticos nos anos iniciais do Ensino Fundamental. Acreditamos que as atividades aqui propostas também poderão auxiliar outros níveis de ensino, de modo a contribuir para a compreensão dos conceitos estatísticos pelos estudantes em prol da constituição de uma formação científica e cidadã.

Prof<sup>a</sup> Suzi Samá e Prof<sup>a</sup> Mauren Porciúncula  
Coordenadoras do GT12 - SBEM

# INTRODUÇÃO

O livro Estatística para os anos iniciais do Ensino Fundamental é um dos produtos do projeto “A matemática escolar: construção e transformação do saber matemático”. Projeto este de intercâmbio nacional financiado pela CAPES por meio do PROCAD. Este livro foi escrito com a intenção de abrir um canal para discutir temas relativos ao processo de formação e desenvolvimento de conceitos e procedimentos elementares de Estatística para os anos iniciais do Ensino Fundamental, numa perspectiva da formação de professores.

Nosso objetivo com este livro é disponibilizar um material, tanto do ponto de vista teórico quanto prático, que permita ao professor refletir acerca do desenvolvimento de conceitos elementares da Estatística.

Temos o propósito específico de apresentar a Estatística não apenas como uma ferramenta útil e necessária na formação do cidadão, mas também como um instrumento indispensável na construção do pensamento científico. Queremos que este livro incentive os profissionais interessados no tema, especialmente o professor, a experimentar soluções para as inúmeras questões de Estatística em suas salas de aula.

Acreditamos que um professor, para ter condições de coordenar um processo eficiente de ensino e aprendizagem de Estatística, deve desenvolver sua competência em três dimensões de forma integrada: a dos conceitos relacionados à Estatística; a da organização de situações didáticas; e a da compreensão do desenvolvimento da aprendizagem dos conceitos estatísticos pelos alunos.

Nesse sentido, esperamos que este livro auxilie o professor na organização de situações didáticas que, em curto prazo, busquem contribuir para que os alunos compreendam conceitos e procedimentos básicos da Estatística. Conhecimentos esses que, por um lado, podem ser utilizados pelos alunos em tratamentos básicos dos dados e, por outro

lado, formem uma base necessária para o futuro aprofundamento estatístico contribuindo para uma formação científica e cidadã.

Todas essas convicções e desejos sobre o alcance que esta obra pode ter baseiam-se no seu caráter inovador, focando três eixos relacionados e complementares. O primeiro eixo diz respeito à apresentação dos conteúdos de forma sólida, privilegiando a formação conceitual. O segundo eixo vem dos relatos de pesquisas que estarão sempre presentes na discussão desses conceitos. Por fim, o terceiro eixo refere-se à prática da sala de aula, quando propomos situações para serem experimentadas.

Dessa forma, entendemos que este livro pode interessar, além dos professores do ensino básico e dos futuros professores (tanto alunos da Licenciatura em Pedagogia como em Matemática), aos pesquisadores da Educação Matemática e aos pós-graduados das áreas de Educação, Educação Matemática, Ensino de Ciências e Matemática, Psicologia Cognitiva e outras áreas afins.

## Por que escrever este livro?

Vários foram os motivos que nos incentivaram a escrever este livro para o professor, mas gostaríamos de explicitar pelo menos os três mais importantes.

O primeiro deles foi a constatação da pouca disponibilidade de material de apoio para auxiliar os professores no ensino dos conceitos elementares da Estatística, principalmente voltados para os anos iniciais do Ensino Fundamental.

O segundo foi a necessidade de tratarmos a Educação Estatística considerando tanto a formação para a cidadania como para o desenvolvimento do pensamento científico. De fato, saber ler e interpretar informações estatísticas permite ao indivíduo entender, avaliar e se posicionar frente a informações veiculadas costumeiramente na mídia, as quais, muitas vezes, têm influência nos rumos políticos e econômicos da sociedade. Contudo, também é preciso auxiliar o indivíduo no desenvolvimento do pensamento científico. Propiciar situações para que ele saiba, por exemplo, identificar um problema, elaborar questões, levantar hipóteses e testar sua validade, isto é, saber escolher os dados a serem coletados, organizá-los e interpretá-los.

Nesse sentido, é preciso auxiliar e incentivar os indivíduos, especialmente as crianças, no desenvolvimento do pensamento científico. Propiciar situações para que ela saiba, por exemplo, identificar um problema, elaborar questões, levantar hipóteses e testar

sua validade; escolher que dados coletar e como fazê-lo; saber ainda como organizá-los e interpretá-los, estando apta a lançar mão de conceitos e procedimentos estatísticos para interpretar corretamente os dados colhidos em sua pesquisa ou na pesquisa de outrem.

Por fim, o terceiro motivo para termos escrito este livro foi, com certeza, o nosso desejo de oferecer um produto de qualidade, escrito com uma linguagem acessível a todos os educadores e, ainda, disponibilizar para o professor exemplos voltados para a prática docente, por meio dos quais ele possa se sentir mais confortável para abordar as ideias básicas e essenciais da Estatística em sua sala de aula.

## Estrutura do livro

O livro é composto por seis capítulos: 1) A Estatística nos anos iniciais do Ensino Fundamental, 2) A identificação do problema, 3) A coleta de dados, 4) O tratamento dos dados, 5) A interpretação dos resultados e 6) Pesquisa experimental: o jogo da memória. A seguir apresentamos, de maneira resumida, cada um deles.

No Capítulo 1 buscamos posicionar a obra em relação aos conceitos básicos da Educação Estatística, tais como: Estatística; Educação Estatística; a Estatística para a cidadania e nesse o que entendemos por letramento estatístico; a Estatística para o pensamento científico e discutimos o conceito de competência estatística. Também apresentamos as fases da pesquisa estatística.

No Capítulo 2 iniciamos nossa proposta de trabalhar a partir de projetos didáticos. Nesse sentido, escolhemos uma questão e buscamos resolvê-la ao longo do livro, isto é, desde a identificação do problema até o produto final, no caso um relatório, complementando os tópicos com exemplos pontuais. Discutimos a importância dos conhecimentos de outras áreas e a necessidade de se trabalhar de forma interdisciplinar, bem como os fatores que interferem na escolha do problema, tais como genuinidade, motivação, tempo didático, complexidade, adequação à faixa etária, conhecimentos prévios dos alunos, bem como as implicações pedagógicas das escolhas. Destacamos a importância da observação dos fenômenos, o delineamento de experimentos e a necessidade da coleta de dados. Nesse processo, enfatizamos a necessidade de despertar o espírito investigativo inato ao ser humano e que a criança tem por excelência, mas que muitas vezes se perde durante a escolarização.

O Capítulo 3 é dedicado a questões inerentes à coleta de dados, no qual aborda-

mos de onde se obtém os dados, definindo a população a ser investigada, apresentando o que é um censo e uma amostragem, e algumas de suas técnicas. Apresentamos diferentes dados, fontes de dados e procedimentos de coleta dos mesmos. Antes de apresentar a elaboração de instrumento de coleta de dados, trabalhamos o conceito de variável, sua classificação, operacionalização e a relação dessas variáveis com as questões de pesquisa. Discutimos o tipo de dado e a escala de mensuração, bem como as vantagens e desvantagens de se usar categorias definidas a priori e a posteriori e apresentamos vários exemplos de instrumentos de coletas de dados. Neste capítulo ainda fazemos considerações pedagógicas sobre como gerenciar uma turma de crianças em plena atividade de levantamento de dados, fazendo com que elas trabalhem motivadas, busquem bons métodos para coletar seus dados, sejam ativas e, ao mesmo tempo, não percam o foco nem se dispersem ao longo da atividade.

O Capítulo 4 tem por objetivo apresentar os vários tipos de gráficos e tabelas. Nele discutimos sobre sua adequação segundo a natureza do dado coletado. Também apresentamos no capítulo as medidas de tendência central e a amplitude, como a mais rudimentar das medidas de variação. Aqui não nos limitamos apenas à descrição dos algoritmos e a apresentação de fórmulas, mas exploramos de forma intuitiva a lógica subjacente nesses procedimentos, bem como o significado dos resultados de cada medida, de cada representação. Lançamos mão de vários trabalhos publicados nesta área.

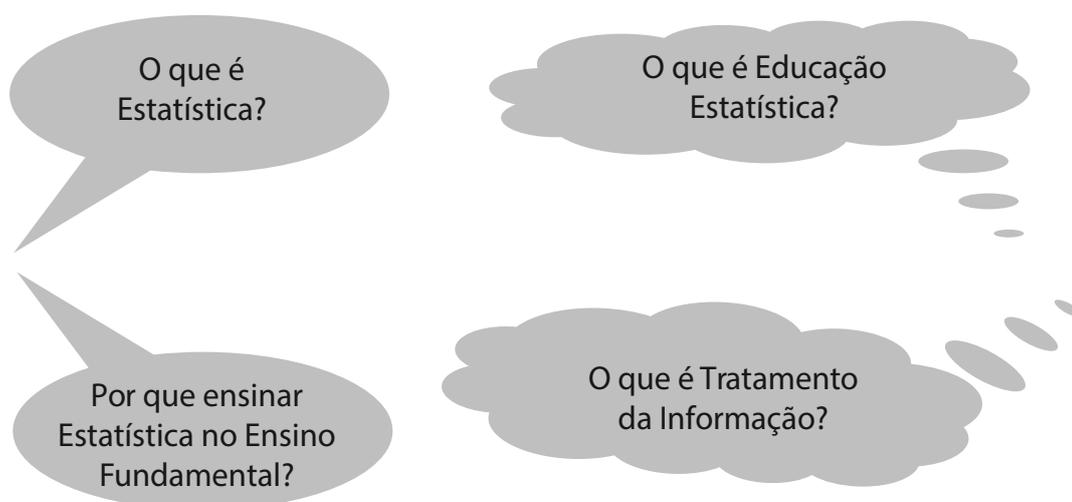
O Capítulo 5 dá ênfase à socialização dos resultados, isto é, à construção da argumentação sobre os resultados a partir dos dados coletados e organizados. Envolve-se aqui a escolha da representação que melhor comunica os argumentos, bem como a elaboração de um relatório. A título de conclusão tecemos algumas reflexões sobre a condução dos trabalhos numa pesquisa como a que propomos.

No Capítulo 6 apresentamos uma atividade voltada inteiramente à prática da sala de aula envolvendo uma pesquisa experimental. Por fim, indicamos referências e sugestões de leituras.

É nosso desejo que este livro proporcione aos educadores reflexões que os mobilizem na direção de olhar e trabalhar com Estatística de forma mais ampla e, ao mesmo tempo, mais focada no planejamento de ambientes e atividades ricas para o desenvolvimento do pensamento estatístico dos alunos do Ensino Fundamental.

# Capítulo 1

## A ESTATÍSTICA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL



Estas são questões que vêm sendo colocadas por educadores matemáticos e professores do Ensino Fundamental quando têm que ensinar conceitos estatísticos. Neste capítulo, situamos o leitor sobre o que entendemos por cada um desses termos.

### O que é Estatística?

O significado da palavra Estatística, enquanto ciência refere-se ao conjunto de ferramentas para obter, resumir e extrair informações relevantes de dados; encontrar e avaliar padrões mostrados pelos mesmos; planejar levantamentos de dados ou delinear experimentos e comunicar resultados de pesquisas quantitativas.

Sua importância reside no auxílio ao processo de pesquisa, que permeia todas as áreas do conhecimento que lidam com observações empíricas. Assim, podemos dizer que a Estatística é a ciência do significado e uso dos dados.

## O que é Educação Estatística?

A Educação Estatística está centrada no estudo da compreensão de como as pessoas aprendem Estatística envolvendo os aspectos cognitivos e afetivos e o desenvolvimento de abordagens didáticas e de materiais de ensino. Para isso, a Educação Estatística precisa da contribuição da Educação Matemática, da Psicologia, da Pedagogia, da Filosofia, da Matemática, além da própria Estatística.

Nesse contexto, o pensamento estatístico pode ser definido como a capacidade de utilizar e/ou interpretar, de forma adequada, as ferramentas estatísticas na solução de problemas. Isto envolve o entendimento da essência dos dados e da possibilidade de fazer inferências, assim como o reconhecimento e a compreensão do valor da Estatística como uma disposição para pensar numa perspectiva da incerteza.

## O que é Tratamento da Informação?

No caso do Ensino Fundamental, os conteúdos de Estatística, Probabilidade e Combinatória são apresentados nos Parâmetros Curriculares Nacionais (PCN) de Matemática (BRASIL, 1997) no bloco denominado Tratamento da Informação.

Segundo os PCN, a finalidade deste bloco é levar o aluno a: construir procedimentos para coletar, organizar e comunicar dados; utilizar tabelas, gráficos e representações que aparecem frequentemente no cotidiano; calcular algumas medidas de tendência central; estabelecer relações entre acontecimentos; fazer previsões e observar a frequência com que ocorre um acontecimento.

Sem dúvida, a proposta dos PCN se constituiu num grande avanço para o ensino de Estatística e Probabilidade na Educação Básica. Suas diretrizes têm como intenção o desenvolvimento do pensamento estatístico que provavelmente dará às novas gerações uma formação básica sólida em Estatística, contribuindo na formação de cidadãos críticos e conscientes.

Esses conceitos e procedimentos vão sendo aprofundados ao longo dos anos escolares a fim de que o aluno aprenda a: formular questões pertinentes para um conjunto de dados; produzir resumos estatísticos; elaborar conjecturas e comunicar informações de modo conveniente; interpretar e construir diagramas e fluxogramas; desenhar experimentos e simulações para fazer previsões.

Para atingir esses objetivos, são sugeridos os seguintes temas/tópicos: Contagem e Combinatória; Estatística (gráficos, tabelas, medidas de tendência central) e Probabilidade. No Quadro 1 encontram-se os tópicos relativos aos dois primeiros ciclos do Ensino Fundamental.

A inserção da Estatística por meio do bloco Tratamento da Informação merece um destaque especial, uma vez que por sua própria natureza a Estatística possibilita trabalhar a Matemática com as outras áreas do conhecimento e com os Temas Transversais. Assim, o ensino de Estatística nesses moldes pode se constituir em um instrumento de base para a formação de uma atitude crítica diante de questões sociais, políticas, culturais e científicas da atualidade.

**Quadro 1:** Conteúdos conceituais e procedimentais de Probabilidade e Estatística (Tratamento da Informação) para os anos iniciais do ensino fundamental

	<b>1º ciclo (1ª e 2ª séries) / (2º e 3º anos)</b>	<b>2º ciclo (3ª e 4ª séries) / (4º e 5º anos)</b>
<b>Estatística</b>	Leitura e interpretação de informações contidas em imagens.	Leitura e interpretação de dados apresentados de maneira organizada (por meio de listas, tabelas, diagramas e gráficos), construção dessas representações.
	Coleta e organização de informações.	Coleta, organização e descrição de dados.
	Exploração da função do número como código na organização de informações (linhas de ônibus, telefones, placas de carros, registros de identidades, roupas, calçados).	
	Interpretação e elaboração de listas, tabelas simples, de dupla entrada e gráficos de barra para comunicar a informação obtida.	Interpretação de dados apresentados por meio de tabelas e gráficos para identificação de características previsíveis ou aleatórias de acontecimentos.
	Criação de registros pessoais para comunicação das informações coletadas.	Produção de textos escritos a partir da interpretação de gráficos e tabelas, construção de gráficos e tabelas com base em informações contidas em textos jornalísticos, científicos ou outros.
	Produção de textos escritos a partir da interpretação de gráficos e tabelas.	Obtenção e interpretação da média aritmética.
<b>Probabilidade</b>		Exploração da ideia de probabilidade em situações-problema simples, identificando sucessos possíveis, sucessos seguros e as situações de “sorte”.
		Utilização de informações dadas para avaliar probabilidades.
		Identificação das possíveis maneiras de combinar elementos de uma coleção e de contabilizá-las usando estratégias pessoais.

Fonte: adaptado dos PCNs (BRASIL, 1997, pp.52 e 61)

Contudo, o termo “Tratamento da Informação” pode induzir as pessoas a terem uma falsa e pobre ideia de que a Estatística se resume apenas a “tratar dados”. Por essa razão, optamos por colocar como título para esta obra “Estatística para os anos iniciais do Ensino Fundamental”.

## **A Estatística desde os anos iniciais do Ensino Fundamental**

O trabalho com Estatística na escola propicia o desenvolvimento do pensamento estatístico, a vivência de um trabalho interdisciplinar e possibilita abordar temas transversais. Como já discutido anteriormente, o pensamento estatístico amplia as formas de pensar valorizando o mundo das incertezas. Muitas vezes o aluno, acostumado a um pensamento determinístico, tende a aceitar como certa a previsão de um resultado a partir da maior frequência de um evento. Por exemplo, ao perceber que todos os seus colegas têm medo do escuro, concluem como certeza que um novo colega terá também medo do escuro. O trabalho com o pensamento estatístico auxiliará o aluno a perceber que sua previsão não necessariamente ocorrerá.

Como a Estatística é parte do método científico, é natural que o trabalho com a mesma parta de problemas de outras áreas do conhecimento e das práticas sociais, viabilizando a interdisciplinaridade e a inserção de temas transversais. Ao trabalhar com projetos em sala de aula, o professor pode partir do levantamento de temas vivenciados pelos alunos, por exemplo, a observação do número de dias ensolarados, o número de alunos que faltam às aulas durante um mês, o maior medo das crianças, a germinação das sementes, dentre outros.

Nesse sentido, sugerimos que, quando realizem projetos escolares coletando dados, não se limitem à coleta de dados, mas os realizem nos moldes da pesquisa científica.

### **Fases da investigação científica**

Na sala de aula podemos ter duas situações em pequena escala: reprodução do conhecimento científico (experimento da refração da luz, a germinação das sementes etc.) ou investigar para a tomada de decisões (investigar o medo das crianças com fins pedagógicos).

Em ambos os casos, o arcabouço metodológico é o mesmo. Conforme Cazorla e

Santana (2010), as fases de uma investigação científica podem ser descritas como segue:

**a) Problematização da pesquisa**

Nesta fase, a escolha do tema é crucial para contextualizar o problema a ser investigado, possibilitar que este faça sentido para o aluno e propiciar o desenvolvimento de uma postura investigativa, incentivando os alunos à observação sistemática dos fenômenos que ocorrem ao seu redor, sejam sociais, culturais ou da natureza, formulando perguntas de pesquisa.

A escolha do tema deve possibilitar um trabalho interdisciplinar, envolvendo aspectos e conteúdos escolares de outras áreas de conhecimento e da Estatística, utilizando seus conceitos e procedimentos que ajudam no planejamento e execução da pesquisa.

Esse tema também deve possibilitar a participação ativa dos alunos, a postura ética, o respeito à opinião do outro, o uso racional dos recursos ambientais etc.

**b) Planejamento da pesquisa**

Escolhido o tema e as perguntas de pesquisa, colocamos em pauta a importância da definição da população a ser investigada, que pode ser por censo (quando se investiga todos os elementos da população) ou por amostragem (quando se investiga uma parte dela).

As perguntas de pesquisa, por sua vez, precisam da escolha adequada das variáveis (características da população) que permitirão sua operacionalização, respondendo à questão levantada. É crucial uma definição clara e precisa dessas variáveis, bem como sua caracterização, o que determina o tipo de tratamento estatístico a ser utilizado.

Após essa etapa, podemos elaborar os instrumentos de coleta de dados, já pensando em responder às perguntas de pesquisa que norteiam o levantamento de dados.

**c) Execução da pesquisa**

Uma vez definida a população a ser investigada e o instrumento para coleta dos dados, o próximo passo é realizar essa coleta. Nesta etapa é preciso uniformizar os procedimentos a fim de que todos os alunos colem os dados da mesma forma. Uma vez coletados os dados, iniciamos o seu tratamento. Nesta fase aproveitamos para apresentar os diversos conceitos e procedimentos que nos ajudam a organizar os dados e a extrair as informações mais relevantes. Isto implica em discutir

como escolher os procedimentos mais adequados para classificar e analisar as variáveis envolvidas.

A interpretação e a comunicação de resultados não se restringem a repetir as informações já contidas nas próprias medidas, mas buscam incentivar a retomada das perguntas de pesquisa que nortearam o levantamento de dados, fechando, assim, o ciclo da investigação científica, como descrevemos na Figura 1.

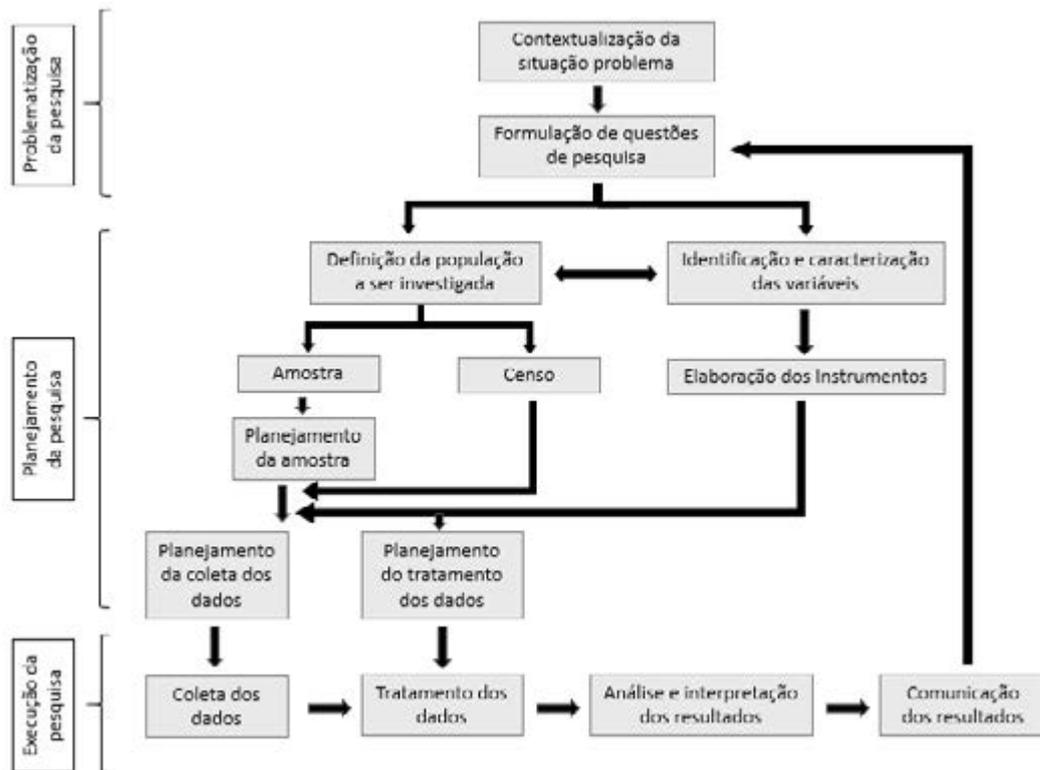


Figura 1: As fases da pesquisa científica.

Fonte: Cazorla e Santana (2010, p. 15).

Nesse livro tomaremos como base essa classificação a qual será detalhada e exemplificada nos capítulos seguintes.

Os alunos devem ter uma participação ativa no processo de construção de seus conhecimentos, ajudando a escolher o tema, as perguntas de pesquisa e as variáveis envolvidas; coletando dados, classificando e analisando os dados (de forma individual ou com a turma); interpretando e comunicando resultados, defendendo suas ideias, desenvolvendo a capacidade de argumentação, aprendendo a ouvir as críticas de seus colegas e, o que é mais importante, aprendendo a respeitar a opinião do outro, dentre outros papéis.

## Capítulo 2

# A IDENTIFICAÇÃO DO PROBLEMA

Neste capítulo discutiremos as fases iniciais da pesquisa estatística. Parte-se da identificação do problema e, então, são levantadas questões a serem respondidas para a solução do mesmo, identificando os fatores envolvidos.

### **Da observação e dúvidas à identificação do problema**

É comum a criança, por meio de suas observações, buscar entender o mundo que vê com questões ou afirmações do tipo:

- Por que o céu é azul?
- Quando vai ser amanhã?
- Rosa é cor de menina.
- Menino é mais forte que menina?
- As meninas sentem medos diferentes dos meninos?
- O Brasil vai ser campeão da copa do mundo?

Por meio de sua curiosidade a criança é levada a questionar, investigar e descobrir coisas novas. Ela age de forma similar à investigação científica ao levantar questionamentos a partir de suas observações. Cabe a nós, professores da escola, aproveitar a curiosidade infantil como um primeiro elemento na condução de uma pesquisa estatística, a qual pode ajudar na compreensão de aspectos do mundo que a cerca. Aguçar a identificação das dúvidas tem, portanto, um papel fundamental no desenvolvimento do pensamento estatístico das crianças.

Uma investigação estatística parte da observação dos fenômenos e da identificação de um problema. Portanto, este é o primeiro elemento a ser construído numa pes-

quisa. É a partir dele que identificamos as perguntas que queremos responder. Assim, o problema, também chamado de questão de pesquisa, é o motivo pelo qual resolvemos fazer uma investigação, é o ponto inicial e motivador.

O problema, do ponto de vista formal, é um enunciado e, do ponto de vista semântico, uma dificuldade ainda não resolvida, uma pergunta ainda não respondida. Ser ou não respondida, precisa ser considerada em relação ao contexto da investigação. O fato de uma questão já ter resposta científica não implica em sua inviabilidade de uso em sala de aula. Essas investigações são feitas para que o aluno observe ou reconstrua o conhecimento, ou parte dele, a partir de experimentos ou de observação dos fenômenos.

Fenômenos são todos os acontecimentos observáveis, algo que pode ser visto. Estes podem ser observados em condições naturais ou experimentais. Os experimentos são réplicas dos fenômenos naturais em condições controladas pelo experimentador.

Por exemplo, o fenômeno da refração da luz, o arco-íris, é um fenômeno natural que pode ser observado na natureza num dia de sol após uma chuva ou reproduzido (de forma experimental) utilizando o prisma de Newton ou, ainda, direcionando um jato de água contra o sol.

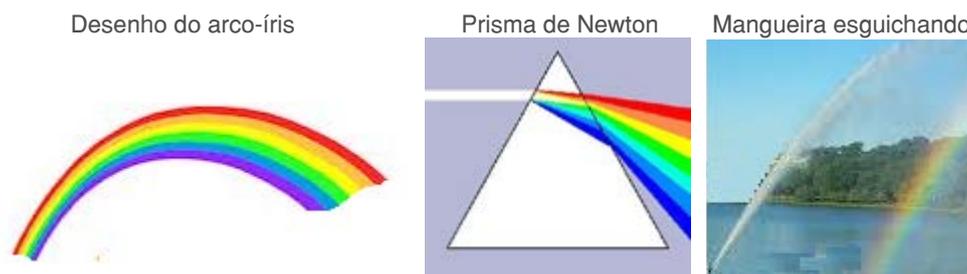


Figura 2: Diversas aparições da refração da luz

Outro exemplo, bastante intuitivo, é a queda dos corpos, como conta a história da maçã que caiu na cabeça de Issac Newton. Observar a queda das maçãs ou qualquer outra fruta da região diretamente na natureza levaria muito tempo e poderia ser inviável.

Apesar dessa dificuldade, os fatores que interferem na queda dos corpos podem ser observados em condições experimentais, controlando volume, formato, peso (massa) etc.

Estes dois exemplos são de fenômenos determinísticos, pois há uma garantia de

certeza do arco-íris ter sempre sete cores numa mesma ordem, assim como podemos afirmar que todos os corpos (mais pesados que o ar) ao serem soltos eles cairão. Esses fenômenos são denominados de determinísticos, pois conhecemos os resultados a priori.

Contudo, existem fenômenos que não são determinísticos, pois não sabemos qual será o resultado de sua realização. Um exemplo é a germinação de uma semente, que pode ou não germinar e só saberemos após plantá-la. Outro exemplo é o clima de nossa cidade no dia seguinte, esse pode ser ensolarado, nublado ou chuvoso. Dependendo da região e da estação do ano a chance<sup>1</sup> de haver uma mudança de clima de um dia para o outro pode ser grande ou muito pequena. Por exemplo, no verão, no período do fenômeno “El niño”, a chance de chuva em São Paulo será altíssima; já no sertão nordestino será pequeníssima. Aliás, este é um tema interessante para ser trabalhado na sala de aula, pois desenvolve a capacidade de prever o resultado de eventos aleatórios.

Esses fenômenos são denominados de aleatórios e alguns deles podem ser replicados via experimentação. No caso da germinação das sementes, ao invés de esperar as sementes caírem na natureza, podemos reproduzir o fenômeno em sala de aula. Neste caso, plantamos as sementes em vasos e se quisermos, ainda, podemos controlar fatores que interferem no resultado tais como: luz (com ou sem luz), adubação (com ou sem adubo), irrigação (controlando a quantidade de água por dia), dentre outras possibilidades.

Em geral, utilizamos os experimentos para conhecer melhor os fenômenos e, muitas vezes, para controlá-los, otimizando os seus resultados. Por exemplo, em condições naturais a chance de uma semente germinar pode ser muito baixa, pois pode cair em um terreno infértil, os animais podem comê-la, pode não chover e morrer por falta de água etc. Já em uma situação experimental essa chance poderá ser bastante alta.

Professor,

Aproveite este momento para discutir com as crianças projetos como o TAMAR, que visa salvar as tartarugas marinhas, cuidando e controlando os locais de reprodução para o aumento da taxa de natalidade das tartarugas.

No caso do clima será impossível de ser recriado de forma experimental. O máximo que podemos fazer é estudar seu comportamento ao longo do tempo, bem como utilizar

<sup>1</sup> Chance é a possibilidade de ocorrer aquele evento, probabilidade é a medida dessa possibilidade.

aparelhos cada vez mais sofisticados para a compreensão deste fenômeno.

No próximo quadro apresentamos exemplos de fenômenos determinísticos ou aleatórios e formas de investigações naturais ou experimentais.

**Quadro 2:** Os tipos de fenômenos e as formas de investigá-los

TIPO	FENÔMENO/ QUESTÃO	FORMA DE INVESTIGAÇÃO	
		OBSERVAÇÃO NATURAL	EXPERIMENTAÇÃO
Determinístico	<i>Refração da luz: o arco-íris tem sempre as mesmas cores na mesma ordem?</i>	Observar o arco-íris em diferentes dias, locais etc.	Observar a formação do arco-íris utilizando diferentes instrumentos como o prisma de Newton e um jato de água em um dia ensolarado.
Aleatória	<i>Germinação de sementes: todas as sementes germinam?</i>	Observar se todas as sementes que caem de uma árvore germinam.	Plantar sementes em diferentes vasos e verificar se todas elas germinam.

Outro aspecto ligado aos fenômenos é sua qualidade de ser observável. Os fenômenos provenientes das ciências exatas, naturais e biológicas, em geral, são de natureza observável e envolvem grandezas que podem ser medidas sem muitas controvérsias.

Por exemplo, a quantidade de sementes que germinam pode ser contada; a altura de uma criança, o espectro da luz e a intensidade de um terremoto precisam de instrumentos, mas podem ser medidos. Em geral, os instrumentos de medida são padronizados, assim como as unidades de medida, que respeitam convenções internacionais.

Ao contrário, os fenômenos ligados às ciências humanas não são diretamente observáveis, são inferidos pela manifestação das pessoas envolvidas, como, por exemplo, o medo que uma pessoa sente, a capacidade de memória, o conhecimento aprendido, o gosto pela Matemática, dentre outros.

Nestes casos enfrentamos dois problemas cruciais: como definir e como medir o fenômeno em estudo.

Por exemplo, o que é “medo” e como medi-lo? Para os adultos pode ter um significado, já para as crianças outro. Para saber do que as crianças têm mais medo é preciso decidir a partir do que serão inferidos os medos das crianças. Não podemos criar situações experimentais, por exemplo, situações que levassem as mesmas a sentirem medos, pois seríamos antiéticos. Neste caso, podemos: perguntar diretamente à criança do que ela tem mais medo; mostrar um rol de situações e pedir que ela marque de qual ela tem mais medo; ou perguntar ao pai do que seu filho(a) tem mais medo.

Já para investigar quem tem mais memória, as crianças ou os adultos, podemos

criar uma situação experimental a partir de uma investigação interessante e fácil de ser realizada pelas próprias crianças. Esse é o caso, por exemplo, de medir a memória das pessoas por meio do “Jogo da Memória”, que será apresentado no Capítulo 6.

No momento interessa-nos apenas levantar problemas possíveis de serem investigados na sala de aula. Esses problemas poderiam ser entendidos a partir das questões:

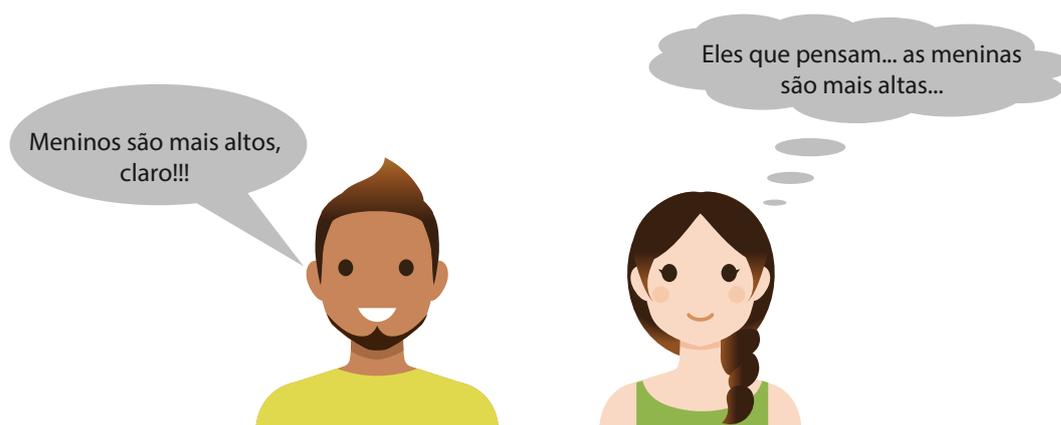
Do que as crianças têm mais medo?

Será que os adultos têm melhor memória do que as crianças?

Todas as sementes germinam quando plantadas?

Os meninos são sempre mais altos do que as meninas?

## Eu acho que ... do “achismo” às hipóteses



É natural que toda criança tenha uma resposta para cada um dos problemas que foram levantados. Um(a) acha que os meninos sempre são mais altos, enquanto outras acham o contrário. Essas respostas das crianças podem ser aproveitadas pelo professor para estimular a explicitação de suas afirmações, as quais, quando acompanhadas de uma explicação, são denominadas de hipóteses.

**Hipótese** é uma afirmativa elaborada e que será colocada à prova, de maneira que poderá ser rejeitada ou não.

As hipóteses, quando testadas, podem se transformar nas conclusões da pesquisa. Nesse sentido, a geração de hipóteses com os alunos é uma etapa fundamental para a Educação Estatística. A criança pode afirmar que os meninos são mais altos que as

meninas, por observar que os homens adultos são mais altos do que as mulheres adultas. Isso é uma hipótese, porque ela afirma e justifica a afirmação. Entretanto, a hipótese pode não ser verdadeira e é por isso que se realiza a pesquisa.

Observamos que existem pesquisas de cunho descritivo ou exploratório as quais não partem de hipóteses. O mesmo exemplo dos medos das crianças pode ser um estudo exploratório se o objetivo for fazer um mapeamento dos principais medos dos alunos. Neste caso, não faz sentido levantar que certos medos são mais frequentes do que outros.

A hipótese, em geral, relaciona pelo menos duas variáveis. No exemplo da altura dos meninos relacionamos gênero, altura e idade. A altura é chamada de variável dependente, pois é ela que sofre a interferência das variáveis gênero e idade, sendo que estas duas últimas são denominadas de variáveis independentes por serem os fatores que modificam a variável dependente.

O diagrama a seguir apresenta uma síntese das variáveis envolvidas nessa hipótese:

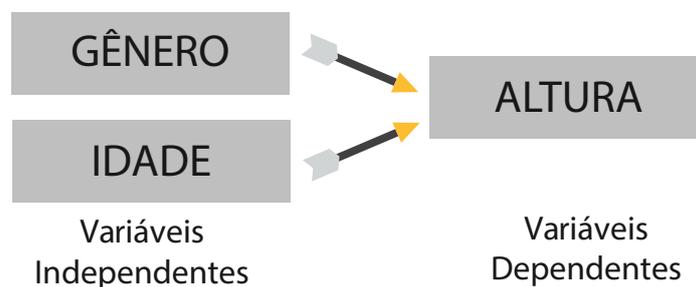


Figura 3: Variáveis dependente e independente

## Questões didáticas da escolha do problema

A escolha do problema ou da questão a ser investigada pode ser uma proposição do professor, de um aluno ou de um grupo de alunos. O que importa é que todos estejam motivados em pesquisar sobre o mesmo. Um trabalho de pesquisa em sala de aula pode ser realizado em uma aula ou em várias, perpassando todo um bimestre letivo. Para que os alunos não desistam da pesquisa no meio do caminho é fundamental que o problema seja, de fato, interessante e desafiador para todos.

Quando falamos “desafiador” estamos enfatizando que o professor precisa refletir se a pesquisa que será realizada permitirá a produção de um conhecimento novo para

esses alunos, sobre o qual eles devem estar efetivamente interessados em saber.

Uma pesquisa científica requer a produção de um conhecimento novo, mas na escola também é realizada a replicação de uma pesquisa, a qual permite que os alunos compreendam um determinado fenômeno e suas variações. Para o psicólogo e pesquisador francês Vergnaud (1994), para que haja aprendizagem é preciso que a criança, por meio de suas ações, construa, mesmo que apenas em parte, esse conhecimento. Só dessa maneira ela se apropria dele.

É importante ressaltar que algumas vezes se confunde pesquisa com estudo. Entretanto, a diferença entre os dois está exatamente na produção de um conhecimento e não na apropriação por alguém de um conhecimento já produzido. Uma pessoa que não conheça a teoria de Piaget poderá estudá-la por meio de seus livros ou de autores que escrevem sobre ela para aprender. Entretanto, para saber se um aspecto da teoria de Piaget é válido, essa pessoa terá que elaborar uma pesquisa para confirmar ou não o que Piaget está argumentando. Da mesma forma, um aluno pode não saber sobre o comportamento da germinação de uma semente e estudar nos livros sobre isso, ou planejar um experimento que lhe permita compreender esse fenômeno.

Adequar uma pesquisa à faixa etária dos alunos é um dos fatores determinantes para o sucesso da mesma. Assim, o tema, os instrumentos, as grandezas envolvidas (força, velocidade, comprimento), a precisão das medidas, devem estar coerentes ao nível de escolaridade. Salientamos que a quantidade de variáveis, que vão ser relacionadas, também deve ser cuidadosamente estipulada.

Ao escolhermos o problema precisamos também considerar o tempo que temos para solucioná-lo. Se um professor, por exemplo, quiser implementar o experimento da germinação das sementes, bastará organizar a classe de tal forma que todos os alunos plantem as sementes e, após um dia ou dois, proceder à contagem daquelas que germi-naram. Mas se esse professor quiser também acompanhar o crescimento das plantinhas, então isso levará mais tempo e envolverá outros procedimentos.

Dessa forma, desde o início é preciso saber quanto tempo se tem para a realização da pesquisa, bem como a adequação das tarefas à idade e aos conhecimentos prévios das crianças sobre o tema a ser investigado.

## Capítulo 3

# A COLETA DE DADOS

No capítulo anterior refletimos sobre a definição do problema de uma pesquisa, o levantamento de hipóteses e as questões de pesquisa. Agora é preciso determinar a população que será investigada, como os dados serão coletados, as variáveis que importam a questão de pesquisa e como serão medidas ou classificadas.

### **De onde se obtêm os dados?**

Ao se questionar do que as crianças têm mais medo, primeiro devemos delimitar o que consideraremos por “crianças”. Podemos considerar “crianças” pelo critério idade, por exemplo, “todas as pessoas de 6 a 11 anos” (ou qualquer outra faixa etária similar). Também, podemos delimitar “crianças” como “todos os alunos matriculados do 1º ao 5º ano do Ensino Fundamental”.

Analisemos cada uma dessas delimitações. A primeira leva em consideração um critério bastante claro, a idade. Contudo, ao se operar a coleta de dados com essa delimitação, seria muito trabalhoso, pois implicaria localizar crianças nos diversos ambientes em que elas se encontram (residências, escolas etc.). Já a segunda delimitação facilita a coleta, já que as crianças estão nas escolas. Contudo, devemos lembrar que esta delimitação se refere a um subconjunto da população de crianças, pois refere-se apenas a crianças brasileiras, que estão na escola. Ao considerar que a Ensino Fundamental é obrigatório no país, estaríamos pegando todas as crianças nessa faixa de escolaridade. Portanto o estudo poderá estar refletindo apenas os medos das crianças brasileiras e não o medo das crianças.

Podemos pensar em todas as crianças do mundo, nas crianças brasileiras, nas

crianças de nossa cidade, nas crianças de uma escola ou nas crianças de nossa sala de aula. Portanto, precisamos definir a abrangência em termos espaciais de nossa investigação. Além disso, as crianças de 2010 não são as mesmas de 2017. Assim, é preciso definir também a abrangência temporal.

Cada um desses grupos se constitui em um tipo de população. Assim, temos diferentes populações para responder a uma mesma questão de pesquisa, o que muda é a abrangência espacial e temporal da investigação.

**População**, em Estatística, é um conjunto de elementos (pessoas, objetos, aulas, etc.), agrupados a partir de pelo menos um critério.

Suponhamos que decidimos fazer a pesquisa na nossa sala de aula e definimos como nossa população todos os alunos matriculados, no 3º ano, turma A, do período matutino, da Escola ABC, da cidade de Recife, em 2016.

Além de definir a população, devemos definir também como vamos obter os dados: perguntar as próprias crianças sobre seus medos; pedir ao seu responsável para apontar em uma lista de medos quais o que a criança tem, etc. Tudo vai depender do objetivo da pesquisa.

Se investigarmos os medos a partir do depoimento das próprias crianças, então estaremos respondendo à questão “quais são os maiores medos das crianças”. Se investigarmos esses medos a partir do depoimento do responsável pela criança, então estaremos respondendo a questão “*quais são os maiores medos das crianças a partir da opinião de seus responsáveis*”. Obviamente, são duas pesquisas diferentes, pois os adultos sentem outros medos que as crianças ainda podem nem conhecer. Portanto, esses resultados poderão nos dar uma visão não fidedigna do medo das crianças.

Assim, a definição dos sujeitos da pesquisa é fundamental para que a investigação atinja o objetivo desejado.

Da mesma forma, no caso da pesquisa sobre a germinação de sementes, não é possível que cada aluno traga um tipo de semente, pois nesse caso, a germinação também dependerá do tipo de semente utilizada. Assim, a semente precisa ser de uma variedade de planta e de preferência de uma mesma procedência, a fim de evitar que fatores alheios interfiram na germinação delas.

Outra questão para a qual devemos chamar a atenção é que a Estatística é a ciên-

cia do *significado e uso dos dados*. Sua grande missão é a compreensão dos fenômenos a partir da análise dos dados, desvendando os padrões subjacentes aos mesmos. Portanto, precisa-se de uma quantidade de dados que possam representar o comportamento do fenômeno em estudo.

Por exemplo, não podemos inferir o comportamento da germinação das sementes a partir da observação de apenas uma ou duas sementes. É preciso ter uma quantidade maior. Entretanto, uma quantidade muito grande pode deixar as crianças perdidas entre os dados. Em geral, sugere-se que para conduzir uma pesquisa em sala de aula cada aluno seja responsável por uma quantidade pequena e fixa de sementes e que os dados de todos os alunos formem o conjunto de dados necessários à pesquisa estatística.

Além da delimitação da população, é preciso definir se coletamos os dados com todos os sujeitos que compõem a população (**censo**) ou escolhemos uma parte representativa da população (**amostra**).

**Censo:** quando investigamos todos os elementos da população.

**Amostra:** quando investigamos uma parte da população.

**Amostragem:** métodos e processos para definir a amostra.

No caso da pesquisa sobre o medo das crianças, podemos escolher pesquisar o medo das crianças de nossa sala de aula. Neste caso a realização do censo é viável, mas a abrangência dos resultados é limitada àquela turma.

No entanto, podemos querer investigar o medo de todas as crianças da escola. Se a escola for de pequeno porte, ainda podemos pensar em realizar um censo; mas se a escola for maior, o censo pode se tornar inviável ou muito trabalhoso. Neste caso, é mais viável utilizar uma amostra dos alunos da escola.

Contudo, definir a amostra não é tão simples, ela precisa levar em consideração as hipóteses. Se o gênero, por exemplo, for importante, não adianta ter uma amostra só de meninas. Da mesma forma, se a idade for importante, não adianta selecionar só crianças de seis anos ou só as de onze anos. Por essa razão a seleção da amostra deverá levar em consideração as características essenciais da população.

Suponha que nossa escola, no turno matutino, tenha doze turmas, sendo três do 1º ano, três do 2º ano; três do 3º ano; duas do 4º ano e uma do 5º ano, sendo que em média cada turma tem 30 alunos, totalizando 360 alunos. Agora devemos responder as

seguintes perguntas: Agora devemos responder as seguintes perguntas:

- Quantos alunos devemos entrevistar? Isto é, definir o tamanho da amostra.
- Como vamos selecionar a amostra?

Ano	A	B	C	Total
1º	30	30	30	90
2º	30	30	30	90
3º	30	30	30	90
4º	30	30		60
5º	30			30
Total				360

- Como levar em consideração as duas variáveis: gênero e idade?

Por uma questão de viabilidade decidimos que cada aluno deve entrevistar dois colegas. Como temos 30 alunos, então o tamanho da amostra será de 60 entrevistados.

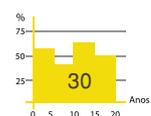
Definido o tamanho da amostra, agora devemos decidir como selecionar os alunos que comporão a mesma. Vejamos algumas estratégias:

- Cada aluno entrevista dois alunos, um menino e uma menina, durante o recreio, sem que uma mesma criança seja entrevistada duas vezes.
- Sorteamos duas turmas e entrevistamos todos os alunos das duas turmas.
- Como são doze turmas, sorteamos cinco alunos de cada turma, utilizando os números da ordem de chamada.
- Listamos todos os alunos, enumeramos cada um de 1 a 360, sorteamos 60 números e entrevistamos os alunos correspondentes.

Analisemos cada um destes procedimentos para verificar a qualidade da amostra como representativa da população. Lembrar que acreditamos que o medo das crianças depende do gênero e da idade e queremos conhecer qual é o medo das crianças da escola, do 1º ao 5º ano, do período matutino, população da qual nossa classe faz parte.

- Se cada aluno entrevista dois alunos, um menino e uma menina, durante o recreio, teremos a garantia de que amostra estará formada metade de meninos e metade de meninas. Porém, não teremos nenhuma garantia de haver crianças de todas as idades e, além disso, corremos o risco de que nossos alunos escolham seus amigos e assim a amostra não representará a escola. Logo poderá não ser representativa. Este tipo de amostragem é **não-probabilística**.

- Selecionamos duas turmas aleatoriamente e entrevistamos todos os alunos das duas turmas. Este tipo de amostragem é denominada **probabilística por conglomerados**, pois os alunos estão agrupados em turmas e, ao invés de selecionarmos os alunos, selecionamos aleatoriamente duas turmas e entrevistamos todos os alunos dessas duas turmas. No entanto, a distribuição por idades não



será garantida, pois no melhor dos casos teremos apenas dois anos diferentes.

c) Como são doze turmas, sorteamos cinco alunos de cada turma, utilizando os números da ordem de chamada. Este tipo de amostragem pode ser considerada **estratificada**, pois utilizamos a idade como critério de estratificação, podemos supor que os alunos do 1º ano tem, na sua maioria, seis anos; os do 2º ano, sete anos e assim por diante. Contudo, não está garantida a representatividade por gênero.

d) Listamos todos os alunos, os enumeramos de 1 a 360, sorteamos 60 números e entrevistamos os alunos correspondentes. Este tipo de amostragem probabilística é denominada de **amostra aleatória simples**. Contudo, a representatividade por gênero e idade não está garantida.

e) Porém, para garantir a representatividade por gênero podemos ter uma **amostra sistemática**. Com uma lista de meninos e outra de meninas, se as duas tiverem a mesma quantidade, sorteia-se 30 de cada uma. Entretanto, se tiver 60 meninos e 300 meninas, será necessário buscar a razão entre o número de meninos e o número de meninas e mantê-la no sorteio da amostra. Nesse caso, para cada 5 meninas, 1 menino será sorteado. Uma sistematização também pode ser feita para a idade.

É preciso considerar também a proporcionalidade da amostra a partir das variáveis importantes para a pesquisa. No exemplo acima, temos muito mais alunos do 1º ano do que do 5º ano. Se pegarmos 5 alunos de cada turma teremos 15 alunos do 1º e 5 do 5º. Assim, se os alunos mais jovens têm medos diferentes dos mais velhos podemos ter distorções sobre os medos dos alunos como um todo.

Professor,

Observe que a ideia é discutir as diversas formas de selecionar a amostra, refletindo junto com as crianças o que pode acontecer com um ou outro procedimento ou, pelo menos, fazer ver a elas que há diferentes formas de selecionar as amostras.

Essa discussão é fundamental, pois de um lado não podemos deixar que os alunos acreditem que qualquer amostra serve para generalizar para toda a população e, de outro, é esse tipo de indagação que nos ajudará a desenvolver o pensamento estatístico.

## A fonte de dados

A fonte de dados é composta pelos sujeitos da pesquisa ou elementos da população que fornecem os dados, que pode ser uma pessoa, como no caso da pesquisa sobre o medo; a semente, no caso da pesquisa da germinação, etc.

Dependendo da pesquisa, a fonte de dados pode ser o próprio aluno, seus colegas, os professores, a semente, os livros da biblioteca, as pedras do pátio, a conta de água etc. No entanto, as fontes poderão ser classificadas em função de quem coleta os dados.

### a) **Fonte primária**

Quando o pesquisador coleta os dados diretamente da fonte são denominados de dados primários. Por exemplo, na pesquisa sobre o medo, os dados obtidos pelo próprio pesquisador a partir das respostas dos alunos ou de seus responsáveis são dados primários. O mesmo ocorre quando registramos os dados da observação da germinação das sementes.

### b) **Fonte secundária**

Quando os dados foram coletados por outras pessoas e nós trabalhamos em cima destes dados, esses são denominados de dados secundários. Por exemplo, se quisermos investigar o padrão do consumo de água ou energia elétrica das famílias de nossos alunos, a fonte de dados será a conta de água ou de energia deles.

Vejamos um exemplo. Suponhamos que queremos investigar qual é a disciplina que mais reprova na escola. Esta pesquisa pode ser realizada de duas formas. A primeira entrevistando diretamente os alunos sobre as disciplinas que reprovaram no ano anterior ao da pesquisa. Outra forma é recorrer às atas finais da escola e fazer esse levantamento a partir desses dados. Assim, o aluno é a fonte primária e a planilha final de notas é uma fonte secundária.



Figura 4: Diferentes fontes de dados

## O que é coletar os dados?

Após a definição dos sujeitos e a definição da fonte de dados é preciso decidir como os dados serão coletados, ou seja, buscar as informações que respondam à questão da pesquisa, que são denominadas de variáveis. Essa é uma oportunidade de solicitar que a turma levante ideias de como a coleta pode ser realizada.

No caso da pesquisa sobre o medo das crianças, como dissemos anteriormente, precisamos definir o que será inferido como medo da criança. A partir do pressuposto que vamos coletar os dados das próprias crianças, a pergunta agora é: como vamos coletar esses dados? Isto é, qual será o procedimento para coletar os dados?

Analisemos alguns procedimentos, elencando suas vantagens e desvantagens:

- a) Realizar uma entrevista: os alunos fazem as perguntas e anotam ou gravam as respostas dos colegas.
- b) Aplicar um questionário, perguntando de que ele tem mais medo. O registro pode ser por escrito ou por desenho. Pode-se solicitar à criança que faça um desenho, como fez, por exemplo, a professora Roberta Buehring (2006) com seus alunos de 2º ano. Esta pesquisadora optou por este procedimento, pois seus alunos ainda não estavam completamente alfabetizados para registrar de forma escrita seus sentimentos.
- c) Caso a população escolhida seja apenas uma turma, podemos, ainda:
  - a. Chamar aluno por aluno, solicitar que eles digam em voz alta qual é seu maior medo e vamos anotando no quadro. Este procedimento pode induzir as crianças a imitar os colegas, ou algumas por constrangimento, não manifestarem sua opinião.
  - b. Fazer uma lista dos medos no quadro, perguntar aos alunos se existe algum outro tipo de medo para ser listado e depois solicitar a eles que levantem a mão à medida que vamos fazendo a leitura dos tipos de medos. Este tipo de procedimento também pode induzir as crianças a optarem por um tipo de medo que a maioria opta.

Esses questionários, desenhos, áudio ou outras formas de registro dos dados são denominados de instrumentos, que coletam os dados das variáveis que vamos estudar.

## Os instrumentos de coleta de dados

Como já mencionamos, há várias formas de coletar os dados. Podemos realizar entrevistas, criar questionários e fichas de observação. Podemos utilizar materiais concretos, fotografias, adesivos, desenhos etc. Podemos, ainda, utilizar instrumentos de medida como réguas e balanças.

### a) Questionário

No caso da pesquisa sobre o medo das crianças, podemos elaborar um questionário que deve contemplar as variáveis: idade, gênero e tipo de medo (Figura 5).

Questionário da pesquisa: “Do que você tem mais medo?”	
Nome do aluno:	_____
Gênero:	( ) Masculino    ( ) Feminino
Idade:	_____ anos completos
Do que é que você tem mais medo?	_____
	_____

Figura 5: Exemplo questionário com categorização *a posteriori*.

Caso a pesquisa seja feita na escola, isto é, incluindo os alunos do 1º ao 5º ano, será preciso incluir a variável ano escolar.

Observamos que como a pergunta é aberta, a criança tratará de exprimir seu sentimento com palavras que ela conhece e, certamente, vamos precisar criar categorias (a posteriori) a partir do registro delas. Também poderíamos deixar as opções prontas (Figura 6), trabalhando com as categorias elaboradas a priori.

Questionário da pesquisa: “Do que você tem mais medo?”	
Nome do aluno:	_____
Gênero:	( ) Masculino    ( ) Feminino
Idade:	_____ anos completos
Do que é que você tem mais medo? Marque com X a opção que expresse seu sentimento:	
( ) Bandido e ladrão	
( ) Fantasmas, espíritos, mula sem cabeça, lobisomem	
( ) Animais (jacaré, hipopótamo, leão, cachorro, barata, aranha etc.)	
( ) Altura, escuro,	
( ) Ficar de castigo, reprovar de ano	

Figura 6: Exemplo de um questionário com categorização a priori.

Na pesquisa que sugerimos que utiliza o jogo da memória com os alunos da clas-

se e um adulto (Figura 7), as variáveis envolvidas são: gênero, idade, autopercepção de capacidade de memória, bem como o tempo gasto no jogo.

A ficha pode ser conforme modelo a seguir:

<b>Questionário da pesquisa: “Do que você tem mais medo?”</b>	
Nome da criança: _____	Nome do adulto: _____
Idade: _____ anos completos	Idade: _____ anos completos
Gênero: ( ) Masculino ( ) Feminino	Gênero: ( ) Masculino ( ) Feminino
Você acha que a sua memória é:	Você acha que a sua memória é:
( ) Muito Boa	( ) Muito Boa
( ) Boa	( ) Boa
( ) Regular	( ) Regular
( ) Fraca	( ) Fraca
( ) Muito Fraca	( ) Muito Fraca
Tempo que gastou no jogo:	Tempo que gastou no jogo:
minutos: _____	minutos: _____
segundos: _____	segundos: _____
Conversão em segundos: _____	Conversão em segundos: _____

Figura 7: Exemplo de ficha para coletar dados

#### b) **Fichas de observação**

No caso da germinação de sementes, o dado será coletado pela observação direta na natureza ou no experimento, portanto, precisaremos de uma ficha para anotar os dados.

A observação na natureza é muito mais complexa. Nesse tipo de observação muitas variáveis podem interferir sem termos como controlá-las, como umidade do ar e da terra, fertilidade do solo, época de germinação etc. Já em uma situação experimental podemos controlar vários desses fatores e escolher os que deixaremos variar.

Para registrar a germinação em uma situação experimental, em que estejam garantidos a mesma quantidade de água, o mesmo tipo de terra, a mesma semente e a mesma exposição ao sol, entre outros, podemos elaborar uma ficha (Figura 8) na qual cada criança irá registrar a germinação das sementes sob sua responsabilidade do 1º ao 5º dia, após plantada.

Ficha de observação: “A germinação das sementes”	
Nome do aluno: _____	
Dias após plantadas	Número de sementes que germinaram
1 (24 h)	
2 (48 h)	
3 (72 h)	
4 (96 h)	
5 (120 h)	

Figura 8: Ficha de observação do fenômeno “a germinação das sementes”.

A depender da idade das crianças, podemos coletar o dado em um único dia, simplificando a coleta, por exemplo, depois de 7 dias. Contudo, o acompanhamento ao longo dos dias possibilitará à criança perceber que existe uma variação natural no tempo de germinação, pois nem todas as sementes germinam ao mesmo tempo.

## Variáveis, seus tipos e sua operacionalização

A todo o momento, estamos falando de variáveis e esse é um conceito-chave na Estatística. Por essa razão vamos nos deter um pouco na sua definição, características e, o que é mais importante, de que forma vamos coletá-la, o que chamamos de operacionalização.

**Variável**, em Estatística, é uma característica da população que assume diferentes valores ou categorias.

Vejamos alguns exemplos. Na pesquisa sobre o medo, caso a população seja formada pelos alunos de nossa classe, os sujeitos da pesquisa são nossos alunos. Que características importantes dos sujeitos vão coletar para poder responder nossa questão de pesquisa? Neste exemplo, vamos considerar as variáveis: idade, gênero e medo.

Contudo, se estivéssemos investigando o desenvolvimento físico dessas crianças, as variáveis poderiam ser: idade, sexo, altura, peso, etc. Se estivéssemos investigando o desempenho escolar, as variáveis poderiam ser: notas nas disciplinas, disciplina favorita, etc. Ainda podemos ter outras variáveis a depender do tema de investigação.

**Quadro 3:** Exemplos de variáveis para Sujeitos – pessoas

Tipo de Sujeito	Exemplo de variáveis
 Sujeito (pessoas)	Gênero: Feminino, Masculino. Idade: 9, 10, 11, 12 (anos completos). Medo: barata, mula sem cabeça.
	Altura (em centímetros, de 120 a 160, por exemplo). Peso (em quilogramas, de 30 a 60, por exemplo). Perímetro cefálico (em centímetros, de 30 a 45, por exemplo).
	Nota em Matemática (escala de zero a dez). Nota em Português, Nota em Matemática, ... Disciplina favorita (Matemática, Português, ...). Disciplina que não gosta (Matemática, Português, ...).
	Número de irmãos (0, 1, 2,...).
	Número de letras de seu nome (3, 4, 5, ...).

Esse questionário pode ser, ainda, a apresentação de uma lista de medos na qual a criança marca a sua resposta. É importante ressaltar que quando a população é compreendida por objetos (como as sementes, no caso da pesquisa sobre germinação) ou por uma coleção de objetos (cada aluno trabalha com um grupo de sementes) ou, ainda, de um coletivo (família), os tipos de variáveis podem ser como as expostas no quadro a seguir:

**Quadro 4:** Exemplos de variáveis para outros tipos de sujeitos

Tipo de Sujeito	Exemplo de variáveis
 Objeto (semente)	Germina: sim ou não. Tempo de germinação: 1 dia, 2 dias, ... O tempo é uma variável que não é da semente, mas que interfere na ocorrência da germinação, pois algumas sementes demoram mais do que outras para germinar, assim esta variável deve ser coletada.
 Coleção de objetos (várias sementes)	Nº de sementes que germinaram (0, 1, 2, ...)  7 das 11 plantadas germinaram
Coletivo (família do aluno)	Número de pessoas da família que mora com o aluno Quantidade de metros cúbicos de água que consomem por mês Renda familiar (em reais R\$) Religião predominante (Católica, Evangélica, ...)

## Tipos de variáveis

As variáveis se classificam em qualitativas e quantitativas (Figura 9). Uma variável qualitativa é aquela cujos resultados se enquadram em categorias. Se as categorias assumem algum tipo de ordenação, elas são denominadas de ordinais, por exemplo, gosto pela Matemática (pouco, regular e muito). Caso contrário, são denominadas de nominais, como, por exemplo, gênero, tipos de medo, entre outros.

Uma variável quantitativa (também denominada de numérica) é aquela cujos resultados assumem valores numéricos. Se essa for passível de contagem é chamada de discreta, como número de irmãos ou número de sementes que germinam. Se a variável é resultante de mensuração, tomando qualquer valor, então são chamadas de contínuas, como peso (kg), altura dos alunos (cm), renda familiar (R\$), entre outras. Assim, podemos classificar as variáveis de diferentes formas, como apresentado no esquema a seguir.

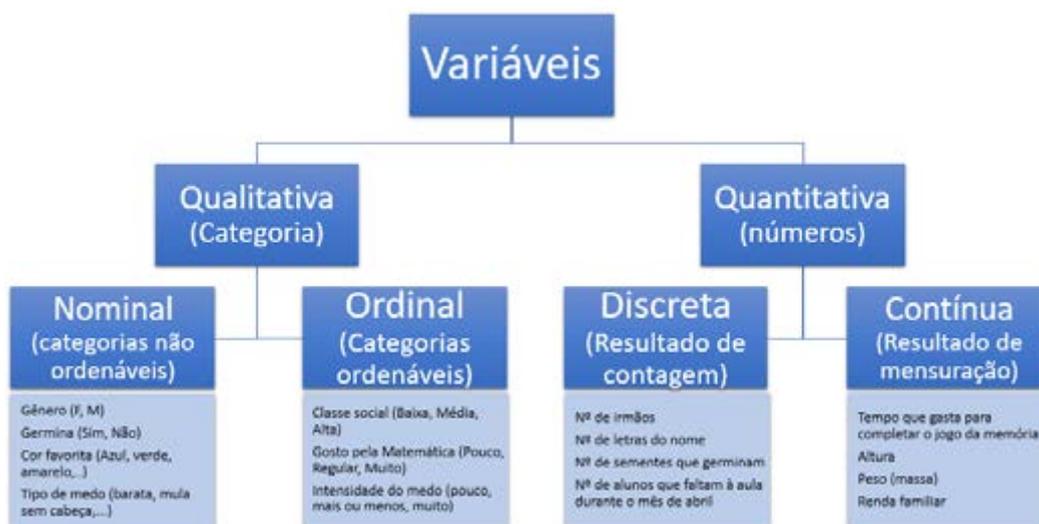


Figura 9: Classificação das variáveis estatísticas de acordo com sua natureza.

### a) Operacionalização das variáveis

Além de aprendermos a reconhecer os tipos de variáveis é importante saber que as variáveis podem ser coletadas (operacionalizadas) de diferentes maneiras.

## Análise de uma variável qualitativa

Quando as categorias das variáveis já estão definidas a priori, por exemplo, gênero (masculino, feminino), gosto pela Matemática (pouco, regular, muito), a coleta de dados é

simples.

Já quando a variável não possui naturalmente as categorias predefinidas, sua coleta se torna mais complexa e é preciso trabalhar a “classificação” da variável, que pode ser *a priori* ou *a posteriori*. Vejamos um exemplo com a pesquisa sobre o medo.

- i. Pergunta aberta. Neste caso formulamos a pergunta de tal maneira que damos completa liberdade ao respondente para expressar seu sentimento:

“Do que você tem mais medo?” \_\_\_\_\_  
\_\_\_\_\_

Esta pergunta permite qualquer tipo de resposta, inclusive aquelas que não têm nada a ver com a pergunta, ou vir em branco, o que é pior. Contudo, podemos ter respostas mais fidedignas, isto é, mais próximas ao sentimento dos alunos.

Essa forma de coletar os dados implica em criar categorias *a posteriori*, isto é, a partir das respostas dos alunos criamos as categorias.

- ii. Pergunta fechada. Neste caso formulamos a pergunta e damos opções para o aluno responder, isto é, o respondente não tem tanta liberdade, pois as categorias já estão criadas. Entretanto, como somos nós os que criamos as categorias, podemos estar induzindo as respostas ou distorcendo completamente a natureza da pesquisa.

- iii. Pergunta Semi-aberta. Quando temos uma pergunta fechada, o sujeito muitas vezes não se encaixa nas respostas apresentadas. Por isso, várias pesquisas incluem a categoria outras. Nessas, o sujeito pode explicitar outra resposta de forma a ampliar o leque de opções e, assim, necessitar de uma recategorização.

- Exemplo de pergunta com uma única escolha (categorias mutuamente excludentes)

“Marque com X a alternativa que você considera lhe dar mais medo:  
( ) animais (cachorros, jacarés, insetos,...)  
( ) fantasmas, espíritos, alma penada, ...  
( ) altura, escuro, não saber as respostas, falar em público,..  
( ) pessoas más, bandido, ladrão, ...  
( ) outro, explique: \_\_\_\_\_

- Exemplo de pergunta com múltipla escolha (categorias complementares)  
Suponha que estamos fazendo uma pesquisa com os professores da escola para identificar os principais problemas da escola.

Na sua opinião, quais são os maiores problemas que impedem sua escola atingir as metas traçadas pelo governo em relação ao IDEB?

- baixos salários dos professores
- professores com formação inadequada para o ensino
- a omissão dos pais dos alunos no processo educativo
- a política educacional
- a infra-estrutura da escola
- outro, explicita: \_\_\_\_\_

Neste caso o respondente pode marcar com X todas aquelas alternativas que ele acredita serem os maiores problemas, que podem ser uma, duas, até todas. Com este tipo de opção corremos o risco de não podermos discriminar quais são os maiores problemas. Uma forma de evitar isso é solicitar ao respondente que de todas as alternativas, escolha apenas três e que coloque 1º (ao maior problema), 2º (ao segundo maior) e 3º (ao terceiro maior).

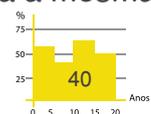
Ainda há uma terceira opção, podemos solicitar ao respondente que coloque V (verdadeiro) ou F (Falso) a cada uma das alternativas. Neste caso teremos a indicação da gravidade do problema a partir da maior frequência de Verdadeiro de cada alternativa.

## Análise de uma variável quantitativa

Quando trabalhamos com variáveis quantitativas precisamos definir a unidade, estar atento ao grau de precisão da mesma, e definir o instrumento pelo qual vai ser medido, dentre outras. Vejamos um exemplo de uma variável quantitativa: idade

- a) Data de nascimento \_\_\_/\_\_\_/\_\_\_
- b) Quantos anos você tem? \_\_\_\_\_
- c) Idade \_\_\_\_\_ em anos completos
- d) Faixa etária (marque com x)
  - De 0 a 20 anos
  - De 21 a 40 anos
  - De 41 a 60 anos
  - De 61 a 80 anos
  - 81 anos ou mais
  - outro, explicita: \_\_\_\_\_

Cada forma de coletar o dado nos fornecerá informações diferentes para a mesma



variável. Analisemos os prós e contras de cada forma de coleta:

- a) A data de nascimento nos permite calcular com precisão de dias a idade da pessoa, que poderá ser crucial se a pesquisa for sobre desnutrição infantil. Porém, será uma informação inútil e no caso de uma pesquisa eleitoral, por exemplo, pois neste caso só interessa a faixa etária.
- b) Quando perguntamos “quantos anos você tem?”, estamos dando liberdade ao respondente para fornecer dados arredondados ou mais detalhados, assim poderemos ter respostas tais como: 9 anos, 9 anos e 5 meses, 9 anos e 8 dias, ou 9 anos e meio. Neste caso, não sabemos se o aluno que respondeu 9 anos é porque ele tem exatamente nove anos ou se ele arredondou para anos completos. Novamente, questionamos se vale a pena ter esses dados nesse nível de precisão ou se não seria melhor solicitar apenas a idade em anos completos, como no item c. Uma discussão mais detalhada sobre o uso da idade em anos e meses para realização de estatísticas encontra-se no anexo A.
- c) Em alguns estudos não temos interesse na idade específica, apenas em faixas etárias. Por exemplo, nas pesquisas eleitorais, podemos querer saber se os mais jovens pensam e votam de forma diferenciada dos que os mais velhos. Também podemos querer saber se as pessoas que não são obrigadas a votar têm interesse em votar. Assim podemos criar diferentes faixas etárias. Isto facilita o tratamento dos dados. Ao fazermos isso, estamos convertendo uma variável intrinsecamente numérica em uma variável qualitativa ordinal, pois uma faixa é mais velha que a anterior. Contudo, se coletamos os dados em faixas etárias, não teremos o dado original (idade), assim não dá refazer as faixas.

## A necessidade de trabalhar com a classificação

O trabalho com classificação precisa de uma atenção especial. Infelizmente, o que se tem observado é que o ensino tem se preocupado muito mais com que os alunos memorizem formas de classificar do que no desenvolvimento do pensamento lógico que permite o classificar.

Classificar significa verificar em um conjunto de elementos os que têm a mesma propriedade.

Uma categorização deve atender a duas condições: ser *exaustiva* (todos os elementos precisam estar em alguma categoria) e *exclusiva* (nenhum elemento pode estar em mais de uma categoria).

Na Figura 10, por exemplo, observamos a tentativa de uma criança de 8 anos em classificar os brinquedos. Ela agrupa os elementos a partir de diferentes critérios, o que termina por não atender a exclusividade.



Figura 10: Classificação de brinquedos realizada por uma criança de 8 anos.

Fonte: Luz e Guimarães (2009).

Já na Figura 11, outra criança consegue realizar uma classificação utilizando como critério o tempo cronológico:

*“Esse que tem o pião é do grupo de antigamente, esse que tem a bicicleta é do grupo que estão brincando muito no presente e esse que tem o jogo da velha é do grupo dos de futuramente, que ainda vão brincar.”*



Figura 11: Classificação dos brinquedos realizada por uma criança de 8 anos

Fonte: Luz e Guimarães (2009).

No primeiro exemplo, a criança utiliza os critérios: tipo de material (plástico e madeira) e envolver pensamento lógico (sim ou não), além de alocar outros como “diferentes”. Assim, era preciso que ela tivesse optado por um desses critérios. Outra possibilidade, bem mais complexa, seria combinar esses critérios gerando quatro categorias: lógicos de madeira, lógicos de plástico, não lógicos de madeira e não lógicos de plásticos, como podemos ver na Figura 12.

	Material	
	Madeira	Plástico
Lógicos		 brinquedos lógicos
Não Lógicos	 brinquedos de madeira	 brinquedos de plástico brinquedos diferentes

Figura 12: Exemplo de classificação com dois critérios.

Fonte: Luz e Guimarães (2009).

Outra forma de representar essa combinação de categorias é o diagrama da árvore (Figura 13), muito utilizado para representar todas as possibilidades no raciocínio combinatório:

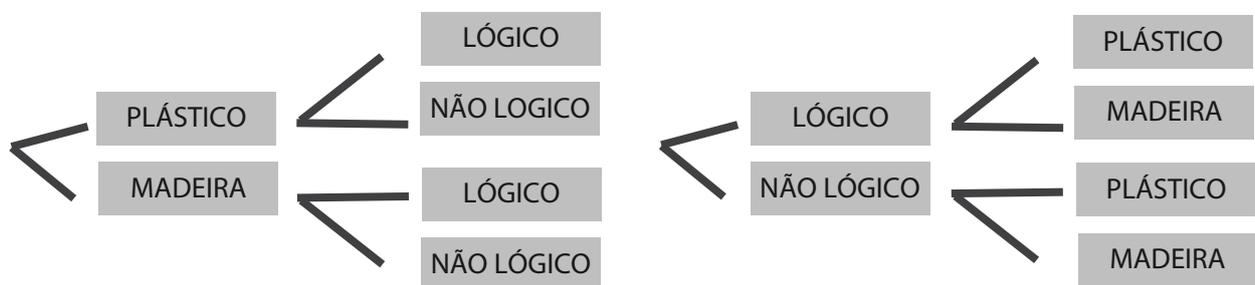


Figura 13: Representação das possibilidades utilizando o diagrama da árvore.

Obviamente, a tarefa de classificação se torna mais complexa quando trabalhamos com variáveis que estão relacionadas aos sentimentos ou ao comportamento das pes-

soas.

Como vimos no caso da variável “tipo de medo”, qualitativa nominal, podemos deixá-la em aberto e provavelmente teremos respostas muito variadas, ou podemos criar categorias a priori para assim coletar os dados. De qualquer forma, teremos a tarefa de classificação a partir das respostas dadas pelos alunos.

Suponhamos que deixamos a pergunta em aberto “De que você tem mais medo?” e que as respostas foram: barata, mula sem cabeça, bandido, altura, rato, escuro, lobisomem.

Percebe-se que as respostas apresentam diferentes tipos de medo. Podemos dizer que esses alunos têm medo de coisas reais (barata, rato, altura e escuro) e coisas imaginárias (mula sem cabeça e lobisomem). Podemos, também, dizer que esses alunos têm medo de bichos (barata e rato), de assombrações (mula sem cabeça e lobisomem) e de situações (escuro e altura). Assim, os mesmos elementos podem ser classificados de diferentes formas que dependem do objetivo de quem classifica.

Além disso, é importante que determinemos como vamos registrar. Por exemplo, se classificarmos os medos em duas classes podemos registrar, apenas, sim e não, ou podemos registrar real e imaginário (Figura 14).

Medo	Ser real
Barata	Sim
Mula sem cabeça	Não
Lobisomem	Não
Rato	sim
Escuro	sim
Altura	sim

Medo	Tipo
Barata	real
Mula sem cabeça	imaginário
Lobisomem	imaginário
Rato	real
Escuro	real
Altura	real

Figura 14: Diversas formas de registrar uma classificação.

Não há, de fato, uma maneira melhor do que a outra. Apenas é preciso evitar a mistura das duas numa mesma anotação.

## Capítulo 4

# O TRATAMENTO DOS DADOS

Iniciamos esse capítulo lembrando que a Estatística tem como objetivo organizar e resumir os dados brutos em poucas medidas ou representações que mostrem de forma sintética o perfil dos dados, as tendências e as relações entre as variáveis.

Para realizar essas tarefas podemos contar com representações em tabelas e gráficos e com as medidas estatísticas tais como: frequências (absoluta e relativa), amplitude, média, mediana, moda, entre outras.

### **Do caos à organização de dados**

Uma vez que coletamos os dados precisamos sistematizá-los. Para isso contamos com várias estratégias didáticas que devem ser avaliadas para verificar qual delas se adéqua à natureza dos dados e à faixa etária dos alunos. Uma das primeiras medidas estatísticas exploradas com as crianças é a frequência.

### **A organização da frequência em pictogramas**

Este tipo de configuração só é possível para dados coletados do próprio aluno e para poucos dados, adequado para trabalhar com os dados de nossos alunos de uma classe que, em geral, tem menos do que 30 alunos. Isto se deve ao fato de que trabalharemos com a escala unitária. A ideia é focar a relação biunívoca entre o dado e sua representação.

Pode-se iniciar a compreensão sobre a função de uma representação em gráficos solicitando que os alunos registrem suas opiniões utilizando caixa de fósforo como utilizadas pela pesquisadora Buering (2006).

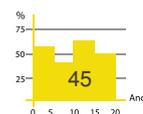




Figura 15: Gráfico de "barras" utilizando material concreto.

Fonte: Buehring (2006, p.64)

Nesse exemplo, cada aluno escreveu seu nome na caixa de fósforos e o empilhou segundo o gênero.

Também podemos utilizar ícones, escudos ou camisetas dos times confeccionados com cartolina, emborrachados ou com adesivos, de tamanhos padronizados. Na Figura 16 utilizamos camisetas dos times confeccionados com emborrachado.

Neste caso, cada aluno pega a camiseta de seu time e o cola no quadro ou na cartolina.

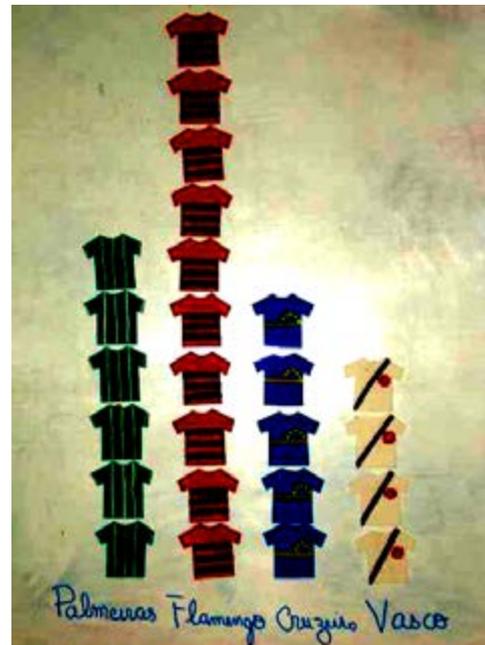


Figura 16: Pictograma construído com material emborrachado para o time favorito no quadro.

Fonte: arquivo das autoras

Na Figura 17 são utilizados ícones para frutas e sorvetes com bolas de sabores coloridos para cada tipo.

Essa abordagem conduz os alunos a entender os gráficos chamados de pictogramas, pois utilizamos ícones para representar os dados.

**Pictograma** é uma representação gráfica que utiliza ícones para representar os dados

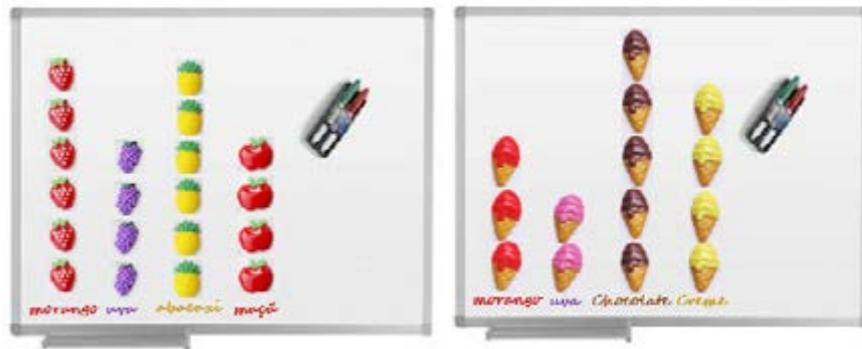


Figura 17: Exemplo de pictograma para a fruta e sabor de sorvete favorito

Na Figura 17 utilizamos ímã de geladeira nos ícones e um quadro magnético (que pode ser substituído por uma placa de metal). Observe que as representações do morango, da uva, do abacaxi e da maçã têm o mesmo tamanho, apesar de, na realidade, essas frutas terem diferentes tamanhos, sendo o abacaxi, por exemplo, bem maior que a morango. Esse problema já não será enfrentado com a representação do sabor de sorvete favorito, pois todos podem ter o mesmo tamanho, variando apenas a cor do sabor.

O tamanho e o espaçamento entre os ícones precisam ser iguais para que visualmente possamos comparar as alturas. Se tivermos 3 abacaxis grandes e 10 maçãs pequenas poderemos ter a mesma altura, o que é um equívoco, porque nesse caso a quantidade de maçãs seria bem maior do que a de abacaxis.

Se solicitarmos às crianças que desenhem é preciso ter atenção na padronização do tamanho dos papezinhos dos desenhos. Todos devem ter as mesmas dimensões. Este tipo de recurso foi utilizado por Buehring (2006), quando solicitou às crianças pequenas que desenhassem do que tinham mais medo. Depois, a pesquisadora solicitou que cada criança colasse o seu desenho no papel madeira a fim de construir o pictograma. A Figura 18 apresenta uma estrutura similar ao que foi realizado pela autor.

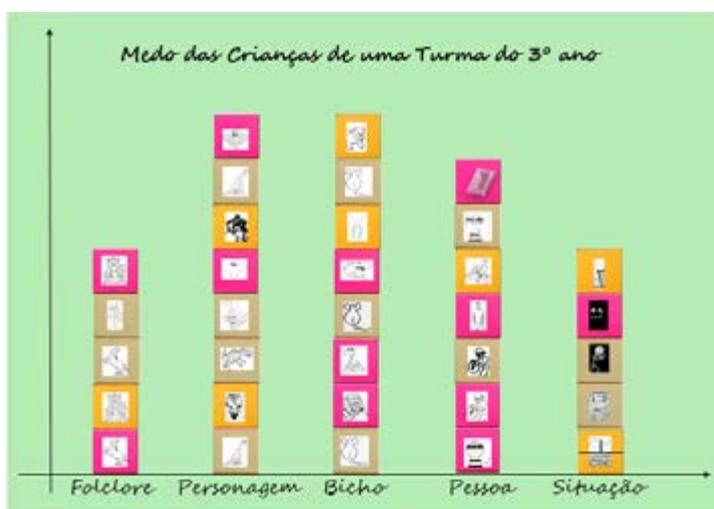


Figura 18: Esquema de pictograma construído colando papéis com desenho em cartolina  
Fonte: Dados fictícios, estratégia similar à de Buehring (2006).

A percepção da correspondência biunívoca entre aluno e o desenho no papel é facilitada pelo fato de cada aluno ter a sua própria representação. Como veremos mais adiante, quando tratarmos os dados diretamente de uma planilha, essa percepção se perde.

## A tabela de frequência (TDF)

Antes de iniciar o tratamento se faz necessário esclarecer o que é uma tabela do ponto de vista estatístico. Atualmente, utilizamos o termo “tabela” para nomear várias coisas, tais como, uma lista de compras, um rol de dados, um quadro, uma planilha, um banco de dados. A rigor uma tabela é qualquer organização matricial composta por linhas, colunas, cujas interseções são denominadas de células, onde se encontram os dados, que podem ser números, categorias, palavras, frases, etc.

Uma tabela de frequência é quando ela apresenta os dados de forma resumida por meio de frequências. As Tabelas de distribuição de Frequência (TDF) (Figura 19) são utilizadas para verificar como se distribuem os dados nas categorias das variáveis qualitativas (a), nos valores pontuais da variável discreta, que toma poucos valores (b) ou nas faixas ou classes, para o caso de variáveis contínuas e discretas que tomam muitos valores (c).

**Tabela de Distribuição de Frequência (TDF)** é um tipo de tabela estatística formada pelas categorias (variável qualitativa), valores pontuais (variável discreta) ou intervalos (variável contínua) e sua frequência absoluta ou relativa.

Distribuição de frequência por categorias (a)

Mascote em casa	Nº de alunos
Cachorro	3
Pássaro	2
Gato	2
Outro	3
Nenhum	15
Total	25

Distribuição de frequência por pontos (b)

Nº de filhos	Nº de famílias
0	40
1	100
2	60
3	40
4	10
Total	250

Distribuição de frequência por intervalos (c)

Altura (em cm)	Nº de alunos
125 – 129	2
130 – 134	3
135 – 139	11
140 – 144	8
145 - 149	1
Total	25

Figura 19: Exemplo de tabelas de distribuição de frequência.

Fonte: dados hipotéticos

**Frequência absoluta**, chamada apenas de frequência, é o número de vezes em que ocorre cada uma das categorias, valores ou faixas da variável.

**Frequência relativa** é a porcentagem da população que se enquadra na categoria (valores ou faixas).

Podemos sistematizar os dados em uma tabela fazendo uma contagem direta. Suponha que estamos investigando o time de futebol favorito dos alunos de nossa classe. Neste caso podemos solicitar às crianças para levantar a mão à medida que vamos enunciando o time favorito. Por exemplo, “Levantem a mão todos os alunos cujo time favorito é o Palmeiras”, fazemos a contagem e anotamos em uma tabela, no quadro, conforme mostra a Figura 20.

Também podemos utilizar essa estratégia para sistematizar uma variável quantitativa que toma poucos valores, como por exemplo, número de irmãos dos alunos da classe, número do calçado etc. Assim, podemos solicitar às crianças para levantar a mão a medida em que vamos enunciando o número de irmãos. Por exemplo, “Levantem a mão todos os alunos que não têm irmãos”, fazemos a contagem e anotamos em uma TDF, no quadro negro, conforme mostra Figura 21.

Time de futebol favorito	Nº de alunos
Palmeiras	6
Flamengo	10
Cruzeiro	5
Vasco	4
<b>Total</b>	<b>25</b>

Figura 20: A TDF a partir da contagem direta de uma variável qualitativa.

Fonte: dados hipotéticos

Número de irmãos	Nº de alunos
0	4
1	10
2	5
3	4
4	0
5	2
<b>Total</b>	<b>25</b>

Figura 21: TDF construída a partir da contagem direta de uma variável quantitativa discreta.

Fonte: dados hipotéticos

Contudo, nesse tipo de sistematização podemos esquecer de contar um aluno ou algum aluno distraído pode levantar a mão duas vezes. Nesse caso, a quantidade de alunos da sala não será igual à soma das frequências representadas na tabela. Perderemos, assim, o controle da sistematização.

Para evitar essa confusão, podemos pegar a lista de chamada, chamar aluno por

aluno e solicitar para que cada um diga o seu time favorito ou número de irmãos e a cada indicação vamos fazendo um risco numa tabela. Ao final, realizamos a contagem e estamos prontos para gerar a TDF, como mostramos na Figura 22.

Time de futebol favorito	Contagem	Time de futebol favorito	Nº de alunos
Palmeiras		Palmeiras	6
Flamengo		Flamengo	10
Cruzeiro		Cruzeiro	5
Vasco		Vasco	4
<b>Total</b>		<b>Total</b>	<b>25</b>

Figura 22: Forma de registro utilizando a lista de chamada e gerando a TDF.

Fonte: dados hipotéticos

## A construção da planilha de dados

Uma forma de garantirmos a fidelidade dos dados é registrando-os em uma planilha ou banco de dados. Essas planilhas nos auxiliam na organização dos dados a partir das variáveis observadas em cada um dos sujeitos da pesquisa.

Suponha que estamos investigando a quantidade de produtos que as famílias de nossos alunos compram para um mês, a partir da lista de compras mensal. Essa lista de compras é uma planilha ou banco de dados, pois os dados são brutos, não receberam nenhum tratamento, como podemos observar na Figura 23.

Lista de compras do mês da Família de Ana			Lista de compras do mês da Família de Ana		
Item	Unidade	Quantidade	Item	Unidade	Quantidade
Açúcar	Sacos de 1 kg	2	Açúcar	Sacos de 1 kg	3
Arroz	Sacos de 1 kg	3	Arroz	Sacos de 1 kg	4
Óleo	Garrafas de 1 litro	1	Óleo	Garrafas de 1 litro	2
Pasta de dente	Unidade de 300 g	1	Pasta de dente	Unidade de 300 g	1

Figura 23: Exemplos de listas de compras das famílias de dois alunos

Se sistematizássemos os dados dessas listas em uma planilha, fazendo apenas a listagem da quantidade dos itens consumidos pelas famílias dos alunos (Figura 24), essa planilha ainda não seria uma tabela estatística, pois ela é composta apenas pelos dados brutos das famílias de nossos alunos. Alguns livros chamam este arranjo de tabela. Neste livro, optamos por chamar este tipo de tabela de planilha de dados, como veremos logo

a seguir, pois apenas transcrevemos os dados brutos para uma lista conjunta, que não receberam nenhum tratamento estatístico.

Família de	Açúcar (Sacos de 1 kg)	Arroz (Sacos de 1 kg)	Óleo (Garrafas de 1 litro)	Pasta de dente (Unidade de 300 g)
Ana	2	3	1	1
Bruna	3	4	2	1
...				
Vitor	1	2	1	1

Figura 24: Exemplo de uma planilha de dados.

Fonte: dados hipotéticos

A planilha de dados deverá ser construída de acordo com o instrumento de pesquisa. Por exemplo, na pesquisa sobre os medos, cada aluno preenche sua ficha e construímos uma planilha que contém os dados brutos como ilustramos a seguir.

a) **A pesquisa sobre o maior medo das crianças**

Os dados de cada aluno são colocados nas linhas conforme as variáveis descritas nas colunas. Assim, a nossa planilha de dados tem na primeira coluna o nome do aluno, na segunda o gênero, na terceira a idade e, na quarta o tipo de medo (Figura 25).

Lembramos que quando construímos a planilha de dados utilizamos códigos que nos ajudam a simplificar o trabalho. Por exemplo, ao invés de escrever por extenso “Feminino”, escrevemos apenas a letra “F” maiúscula e “M” maiúscula para “Masculino”.

Professor,

Reforce a observação da natureza das variáveis. A idade é uma variável quantitativa, já o gênero e os tipos de medo são variáveis qualitativas, e neste caso, ambas nominais, já que suas categorias não podem ser ordenadas, pois ter medo de fantasma não é maior, igual ou menor que ter medo de rato.

Exemplo da ficha em branco:

Ficha da pesquisa: "Do que você tem mais medo?"

Nome do aluno: \_\_\_\_\_

Gênero: ( ) Masculino ( ) Feminino

Idade: \_\_\_\_\_ anos completados

Do que é que você tem mais medo? \_\_\_\_\_

Arcabouço da planilha de dados:

Nome	Gênero	Idade	Tipo de Medo
Ana			
Artur			
Bianca			
Bruna			
Beto			
...			

Exemplo da ficha preenchida:

Ficha da pesquisa: "Do que você tem mais medo?"

Nome do aluno: Ana

Gênero: ( ) Masculino (X) Feminino

Idade: 8 anos completados

Do que é que você tem mais medo? Ladrão

Planilha de dados preenchida:

Nome	Gênero	Idade	Tipo de Medo
Ana	F	8	Ladrão
Artur	M	8	Rato
Bianca	F	8	Altura
Bruna	F	9	Leão
Beto	M	9	Mula sem cabeça
...			

Figura 25: Exemplo de ficha de coleta de dados e arcabouço da planilha de dados em branco e preenchidas.

Para desenvolver essa atividade em sala de aula, é importante que o professor discuta com os alunos que trabalhar com pesquisa, em geral, não é finalizado em uma aula. Assim, é fundamental que os dados sejam registrados em papel ou em um software no computador para que os mesmos possam ser recuperados em outro momento.

Outro fator fundamental de ser refletido é sobre a uniformidade do banco de dados. Se cada aluno fizer um banco de dados é preciso checar se todos estão procedendo da mesma forma, senão, as análises terão resultados diferentes. Se, por exemplo, forem feitos registros como "M" para "masculino" ou para "menina". Outro tipo de problema é quando utilizamos as planilhas eletrônicas, nas quais uma letra diferente é compreendida pelo sistema computacional como uma outra categoria.

Por essa razão, é recomendável que o professor esteja atento e conduza o preenchimento da planilha de forma coletiva.

A partir dessa planilha com os dados brutos (Quadro 5) podemos iniciar o tratamento dos mesmos categorizando os tipos de medo em classes.

Dados primários (brutos)				Dados secundários		
Nome	Gênero	Idade	Tipo de medo	Classe de Medo	Subclasse de Medo	Nº de letras do nome
Ana	F	8	Ladrão	Real	Real Pessoa	3
Artur	M	8	Rato	Real	Real Bicho	5
Bianca	F	8	Altura	Real	Real Situação	6
Bruna	F	9	Leão	Real	Real Bicho	5
Beto	M	9	Mula sem cabeça	Imaginário	Imaginário Folclore	4
Breno	M	8	Bandido	Real	Real Pessoa	5
Carla	F	7	Mula sem cabeça	Imaginário	Imaginário Folclore	5
Camila	F	8	Bicho papão	Imaginário	Imaginário Folclore	6
Daniel	M	9	Tiranossauro	Imaginário	Imaginário Personagem	6
Denise	F	9	Lugar alto	Real	Real Situação	6
Deise	F	8	Escuro	Real	Real Situação	5
Emilio	M	8	Escuro	Real	Real Situação	6
Fabio	M	7	Cobra	Real	Real Bicho	5
Felipe	M	9	Marginal	Real	Real Pessoa	6
Gilda	F	8	Rato	Real	Real Bicho	5
Gabriela	F	8	Coringa	Imaginário	Imaginário Personagem	8
Irene	F	8	Altura	Real	Real Situação	5
José	M	9	Homem mascarado	Real	Real Pessoa	4
Juliana	F	9	Palhaço	Real	Real Pessoa	7
Luiz	M	8	Dinossauro	Imaginário	Imaginário Personagem	4
Luciana	F	7	Bandido	Real	Real Pessoa	7
Mariana	F	8	Bruxa	Imaginário	Imaginário Personagem	7
Marcelo	M	8	Saci	Imaginário	Imaginário Folclore	7
Milton	M	8	Bicho papão	Imaginário	Imaginário Folclore	6
Paulo	M	8	Tubarão	Real	Real Bicho	5
Pâmela	F	9	Fantasma	Imaginário	Imaginário Personagem	6
Pedro	M	7	Feiticeira	Imaginário	Imaginário Personagem	5
Rui	M	7	Barata	Real	Real Bicho	3
Renata	F	8	Rato	Real	Real Bicho	6
Sandra	F	8	Tiranossauro	Imaginário	Imaginário Personagem	6
Saulo	M	8	Ladrão	Real	Real Pessoa	5
Vera	F	8	Cachorro	Real	Real Bicho	4
Vanessa	F	9	Fantasma	Imaginário	Imaginário Personagem	7

Fonte: Dados fictícios de uma turma do 3º ano.

Neste exemplo utilizamos duas classes: real e imaginário. Podemos, ainda, criar subclasses. Para a classe real podemos ter: “real-pessoa”, “real-bicho” e “real-situação” e para imaginários é possível ter: “imaginário-folclore” e “imaginário-personagem”. Assim, os tipos de medo: ladrão, bandido e marginal, são dados primários e foram classificados como “real-pessoa”. Dessa forma, os tipos de medo são dados primários, ou seja, dados brutos, enquanto que a classe “real-pessoa” é um dado secundário, deriva-se da junção de três tipos de medo.

Da mesma forma a variável “número de letras do nome” é um dado secundário, pois as crianças apenas escreveram seus nomes (dado bruto) e, a partir do nome, conta-

mos e registramos a quantidade de letras do nome na coluna correspondente. Ressaltamos que os dados são fictícios.

Observe que nem todo mundo pode concordar com que Tiranossauro ou Dinosauru sejam categorizados como “imaginário personagem”, talvez alguém queira sugerir a criação de uma nova categoria “imaginário bicho”. Já com a resposta “fantasma”, não sabemos se a criança se referiu ao personagem fantasma do filme ou ao fantasma de assombração, neste caso, seria mais adequado classificá-lo em “imaginário folclore”.

#### b) **A pesquisa sobre a altura das crianças**

No caso da pesquisa sobre a altura das crianças, havia sido questionado se os meninos são mais altos do que as meninas, e levantou-se a hipótese que sim, pois em geral os homens são mais altos que as mulheres.

Assim, consideramos as variáveis idade, gênero e altura. Para garantirmos uma variação nas idades é importante ter na amostra alunos de diferentes anos escolares. Vamos fazer, então, uma escolha aleatória de 3 meninas e 3 meninos das duas turmas (A e B) dos cinco anos (1º ao 5º), ver Quadro 6.

Ficha de pesquisa: “Os meninos são mais altos do que as meninas?”	Ficha de pesquisa: “Os meninos são mais altos do que as meninas?”
Nome do aluno: _____	Nome do aluno: Alberto
Turma: _____ Ano: _____	Turma: A Ano: 1º
Gênero: ( )M ( )F	Gênero: ( X )M ( )F
Idade: _____	Idade: 5 anos e 7 meses
Altura: _____	Altura: 1,10 m

Figura 26: Exemplo de ficha de coleta de dados em branco e preenchida.

Observe que os dados foram coletados em anos. Entretanto, alguma criança pode coletar os dados considerando anos e meses, por exemplo 5 anos e 3 meses, 8 anos e 1 mês. Como essa conversão é trabalhosa para este nível escolar aconselhamos simplificar a coleta de dados da variável idade, solicitando apenas os anos completos.

Da mesma forma, no caso da altura, devemos decidir em qual unidade vamos registrar os dados. Neste exemplo decidimos trabalhar com metros. Como estamos preocupados com uma maior precisão, utilizaremos duas casas decimais, ou seja, o centésimo do metro. Assim, devemos registrar todos os dados com duas casas decimais, mesmo que um aluno meça um metro (1,00). Isto é muito importante, pois estudos mostram que os alunos enfrentam problemas ao registrar, ler e operar com números decimais.

**Quadro 6:** Extrato da planilha de dados da pesquisa sobre a altura dos alunos

Dados originais				
Nome	Turma	Gênero	Idade	Altura (m)
Alberto	1º ano A	M	5 anos	1,10
João	1º ano A	M	6 anos	1,00
Tereza	1º ano A	F	5 anos	1,15
Pedro	1º ano B	M	6 anos	1,04
...	...	...	...	...
Soma				...

Fonte: dados fictícios

Uma vez preenchida a planilha (veja planilha completa no Anexo A) com os dados brutos podemos iniciar o tratamento estatístico dos dados: construir tabelas e gráficos e calcular as medidas estatísticas.

## A construção da tabela de distribuição de frequências simples

Para construir a TDF, a partir da planilha de dados basta contar o número de vezes que se repete uma categoria (variável qualitativa) ou o valor da variável em estudo (discreta).

Na Figura 27 apresentamos as fases da construção da TDF para a variável qualitativa “gênero”. Este tipo de tabela contém a frequência absoluta com que aparece cada categoria e sua frequência relativa (expressa em porcentagem)

Nome	Gênero
Ana	
Artur	
Bianca	
Bruna	
Beto	
...	

Gênero	Contagem
Feminino	
Masculino	
Total	

Gênero	Nº de alunos	%
Feminino	18	54,5
Masculino	15	45,5
Total	33	100,0

**Tabela 1:** Distribuição dos alunos por gênero

Figura 27: Construindo a TDF

No caso dessa tabela a frequência absoluta é o número de meninas (18) e a frequência relativa é percentual de crianças que são meninas (54,5%)

### Como se obtém a porcentagem?

Note que 18 do total de 33 alunos são meninas (do gênero feminino), portanto, teríamos uma fração de  $18/33$  do total de alunos como meninas. Quando trabalhamos com porcentagem, precisamos transformar para uma fração de base 100. A pergunta seria então: se o total fosse 100 crianças quantas seriam meninas?

Estaríamos diante do problema multiplicativo de proporcionalidade simples.

Meninas	Meninos
?	100
18	33

Meninas	Meninos
?	100
18	33

São várias as formas de calcular o valor desejado. Uma delas é calcular a fração  $18/33$  da quantidade 100. Portanto, basta calcular  $18/33 \times 100 = 54,5$ . Então teríamos que 54,5% dos alunos são meninas, ou seja, do gênero feminino. Outra maneira é calcular a razão entre 33 e 100. Essa razão deve ser a mesma entre 18 meninas e a quantidade de meninas procurada. Portanto,  $?/18 = 100/33$ , nesse caso também temos que:  $? = 18 \times 100 / 33 = 54,5$ .

É importante salientar que o professor precisa estar atento para verificar se seus alunos já são capazes de trabalhar com porcentagem.

Na Tabela 2 apresentamos a distribuição do medo por classes e na Tabela 3 por subclasses.

**Tabela 2:** Distribuição das classes de medo dos alunos (nominal)

Classe de medo	Nº de alunos	%
Imaginário	13	39,4
Real	20	60,6
Total	33	100,0

**Tabela 3:** Distribuição das subclasses de medos alunos (nominal)

Classe de medo	Subclasse de medo	Nº de alunos	%
Imaginário	Imaginário folclore	5	15,2
	Imaginário personagem	8	24,2
Real	Real bicho	8	24,2
	Real pessoa	7	21,2
	Real situação	5	15,2
Total		33	100,0

Do mesmo modo podemos proceder com as variáveis quantitativas discretas, que tomam poucos valores, como mostramos nas Tabelas 4 e 5.

**Tabela 4:** Distribuição da idade dos alunos (discreta)

Idade	Nº de alunos	%
7	5	15,1
8	19	57,6
9	9	27,3
Total	33	100,0

**Tabela 5:** Distribuição do número de letras do nome (discreta)

Nº de letras do nome	Nº de alunos	%
3	2	6,1
4	4	12,1
5	11	33,3
6	10	30,3
7	5	15,2
8	1	3,0
Total	33	100,0

## A construção de TDF de dupla entrada

Apresentamos a seguir as fases da construção de uma tabela de dupla entrada a qual relaciona a variável gênero e com a variável classe de medo (Figura 28).

Nome	Gênero	Idade	Tipo de medo	Classe de Medo
Ana	F	8	Ladrão	Real
Artur	M	8	Rato	Real
Bianca	F	8	Altura	Real
Bruna	F	9	Leão	Real
Beto	M	9	Mula sem cabeça	Imaginário
Breno	M	8	Bandido	Real

Medo	Feminino	Masculino
Imaginário		
Real		
Total		

Figura 28: Fases da construção de uma TDF de dupla entrada.

A partir dessa contagem podemos construir a tabela de distribuição de medos segundo o gênero (Tabela 6).

**Tabela 6:** Distribuição das classes de medos segundo gênero<sup>1</sup>

Classe de Medos	Feminino		Masculino		Total	
	N	%	N	%	N	%
Imaginário	7	38,9	6	40,0	13	39,4
Real	11	61,1	9	60,0	20	60,6
Total	18	100,0	15	100,0	33	100,0

Observe que, aqui, estamos interessados em saber se as meninas têm medos diferentes dos meninos e, para isso, só podemos realizar uma comparação a partir do percentual, uma vez que temos mais meninas do que meninos. A Tabela 6 mostra que a

<sup>1</sup> Usa-se a letra N para denotar a quantidade de sujeitos, ou seja, a frequência absoluta de sujeitos.

maioria dos meninos e das meninas tem mais medo de coisas reais, do que imaginárias. A tabela também mostra que a diferença entre as classes de medos de meninas e meninos é tão pequena (1,1%), que nos permite dizer que não há diferença.

Essa análise que fizemos mostra que a leitura de uma tabela não é habilidade intuitiva e simples. É preciso um trabalho sistematizado sobre a representação em si, focando na relação entre as variáveis, nas frequências relativas, entre outros.

Além das tabelas, a Estatística disponibiliza diferentes formas gráficas para representar os dados.

## Construção de gráficos

Os gráficos são representações poderosas, pois em um golpe de vista podem proporcionar a compreensão do padrão dos dados. Existe uma grande diversidade de gráficos e softwares que os constroem de forma rápida e esteticamente bonitos. Contudo, recomendamos que os alunos também sejam solicitados a construir gráficos com lápis e papel quadriculado para que possam se apropriar melhor dos conceitos e representações envolvidos.

### Gráfico de barras/colunas

O gráfico de barras é apropriado para representar as variáveis qualitativas, assim para cada categoria é levantada uma barra vertical (coluna) ou barra horizontal.

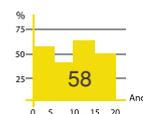
Esse tipo de gráfico pode utilizar uma escala unitária ou não-unitária. Ao se introduzir na escola a construção de um gráfico de barras é recomendável iniciar com uma escala unitária para que os alunos possam perceber a relação de cada sujeito e sua representação como discutido no item relativo ao pictograma (página 44).

#### a) Gráfico de colunas ou de barras verticais

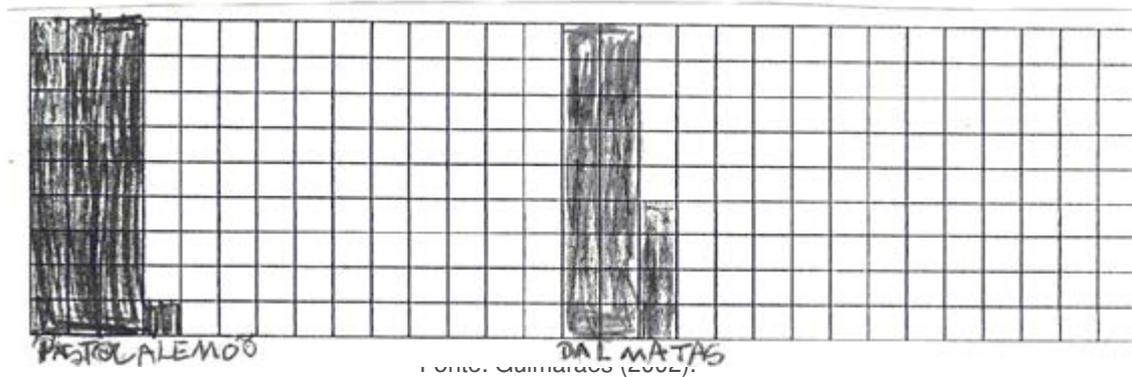
No gráfico de colunas ou de barras construído no papel quadriculado, cada quadradinho equivale a um sujeito ou ícone do pictograma.

Essa relação um quadrado para cada dado (aluno) precisa ser bem compreendida pelos alunos. Veja na Figura 29 o que um menino de 9 anos, que cursava o 4º ano, fez para representar 28 cachorros da raça pastor alemão e 22 dálmatas.

Observamos que o menino estabeleceu a relação um quadrado para cada unidade.

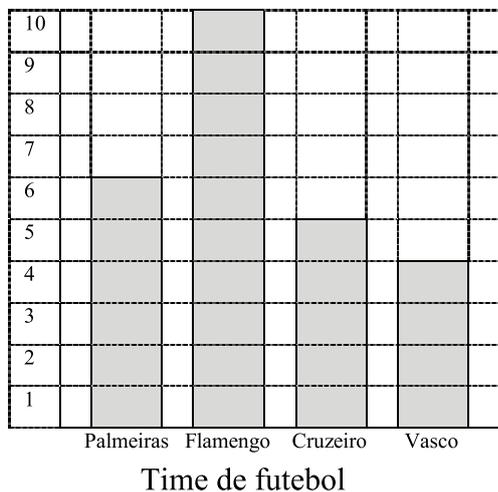


Porém, achou que se os quadrados estivessem pintados, estava construindo um gráfico de barras vertical, independente de se constituírem, de fato, como uma coluna ou barra vertical.



No gráfico de barras verticais, a altura da barra vertical indica o número de alunos. Neste caso, para saber a frequência devemos contar quantos quadradinhos tem cada barra vertical, o que é fácil se tivermos uma malha por trás do gráfico (a), ou se tivermos o número (rótulo) em cima da coluna (b). A rigor o gráfico de barras vertical é formado por colunas contínuas (Figura 30).

a) Desenhado em uma malha



b) Gerado por uma planilha eletrônica

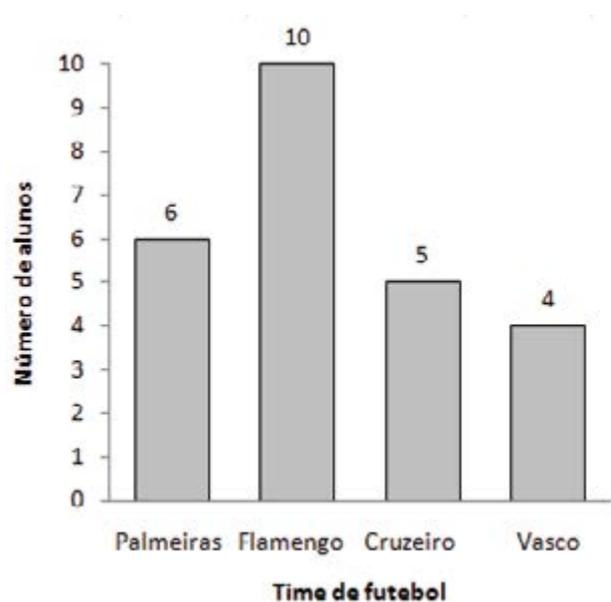


Figura 30: Exemplo de construção de gráfico de colunas com escala unitária.

Professor,

Na construção do gráfico com o papel e lápis lembre-se da necessidade das barras verticais terem a mesma largura. O mesmo vale para o espaçamento entre elas. Lembre ainda de não colar as barras verticais umas às outras.

Os gráficos que aparecem com barras verticais coladas são de outro tipo, são os histogramas, próprio de variáveis contínuas ou discretas que tomam muitos valores, como mostraremos mais adiante.

Imagine se tivéssemos que representar a preferência dos torcedores que estavam presente na final do campeonato carioca no jogo do Flamengo com o Fluminense no Maracanã. Neste caso seria impossível utilizar a escala unitária.

Precisamos sair de uma escala unitária para uma escala proporcional, que precisará ser calibrada, sendo que para isso o aluno precisará conhecer proporcionalidade, para entender quanto vale cada valor da reta numérica, em geral, apresentada em intervalos explícitos. Outra possibilidade é na malha quadriculada um quadradinho representar 5 unidades, 10 unidades, ou qualquer outro número.

Se na escola tivéssemos 37 alunos que torcem pelo Santos, 50 pelo Fluminense, 40 pelo Santa Cruz e 7 pelo Palmeiras, a melhor opção será utilizar um gráfico de colunas com escala proporcional, como apresentamos na Figura 31.

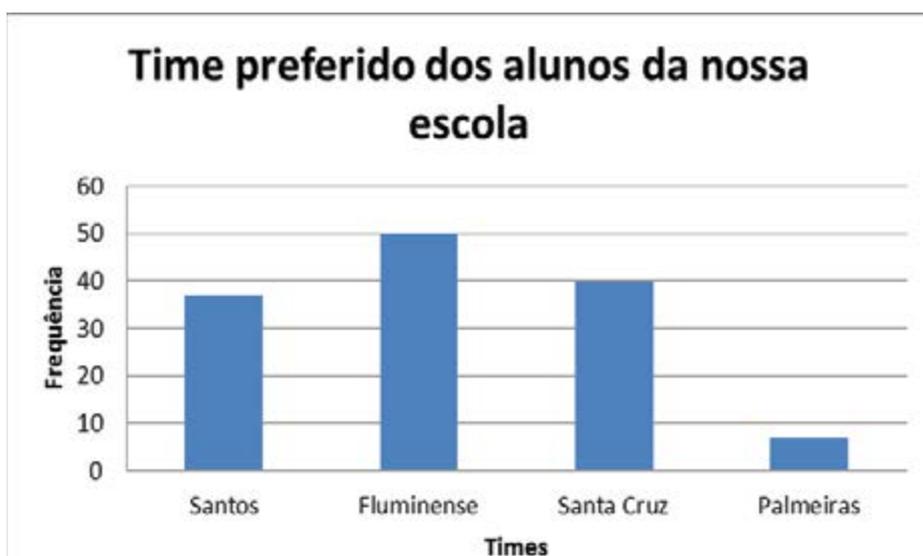


Figura 31: Gráfico com escala não unitária

Os alunos terão facilidade em perceber que na escola tem 50 torcedores do Fluminense e 40 do Santa Cruz, mas apresentarão dificuldades em identificar a quantidade de

torcedores do Santos e do Palmeiras. Por que será?

Na escala que utilizamos de 10 em 10, facilmente podemos conferir que a altura da coluna dos torcedores do Fluminense termina em 50 marcado na escala. Entretanto, a quantidade de santistas fica entre 30 e 40. Alguns alunos podem achar que é 30 e meio, outros 31 e outros 35. De fato a partir desse gráfico saber que são 37 não é uma tarefa fácil.

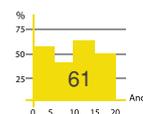
Por outro lado, é possível ver que 30 e meio ou 31 são quantidades pequenas quando consideramos a proporcionalidade do intervalo entre 30 e 40 e a altura da coluna que expressa a quantidade de santistas. De fato, compreender o valor não expresso na escala dificulta bastante a compreensão dos alunos. Entretanto, essa é uma habilidade possível de ser compreendida pelos alunos como evidenciam Evangelista e Guimarães (2015).

A dificuldade de calibrar as escalas pode ser superada com o uso de planilhas eletrônicas, que disponibilizam diversos tipos de gráficos, e calibram as escalas automaticamente, mesmo quando o grau de precisão é maior, como no uso de uma casa decimal para uma escala de 5 em 5. Entretanto, é fundamental propor atividades que levem os alunos a refletirem sobre a construção de escalas representadas em gráficos.

Precisamos ter muita atenção a esse ponto. Muitas vezes os gráficos divulgados, inclusive na mídia, apresentam distorções como foi investigado por Cavalcanti, Natrielli e Guimarães (2010). Essas autoras verificaram que 39% dos gráficos apresentados na mídia impressa analisados por elas apresentavam erros de proporcionalidade na escala. Assim, é fundamental que os alunos compreendam uma escala para serem leitores críticos das informações, como é o desejado.

Observe que esse trabalho não é trivial, pois exige o domínio de proporcionalidade, porém algumas distorções são facilmente identificadas. Avalie a possibilidade de seus alunos compreenderem essa discussão. Por outro lado, o trabalho com a proporcionalidade é fundamental desde os primeiros anos. A literatura infantil apresenta algumas histórias que podem ser utilizadas para deflagrar a discussão como a história dos ursos e a menina de cachinhos de ouro.

É importante também mostrar a relação entre os valores da variável e sua frequência. Via de regra, os valores da variável vão no eixo horizontal (abscissa) e sua frequência (número de alunos) no eixo vertical (ordenada). Porém, podemos trocar os eixos, e obter um gráfico de barras horizontais.



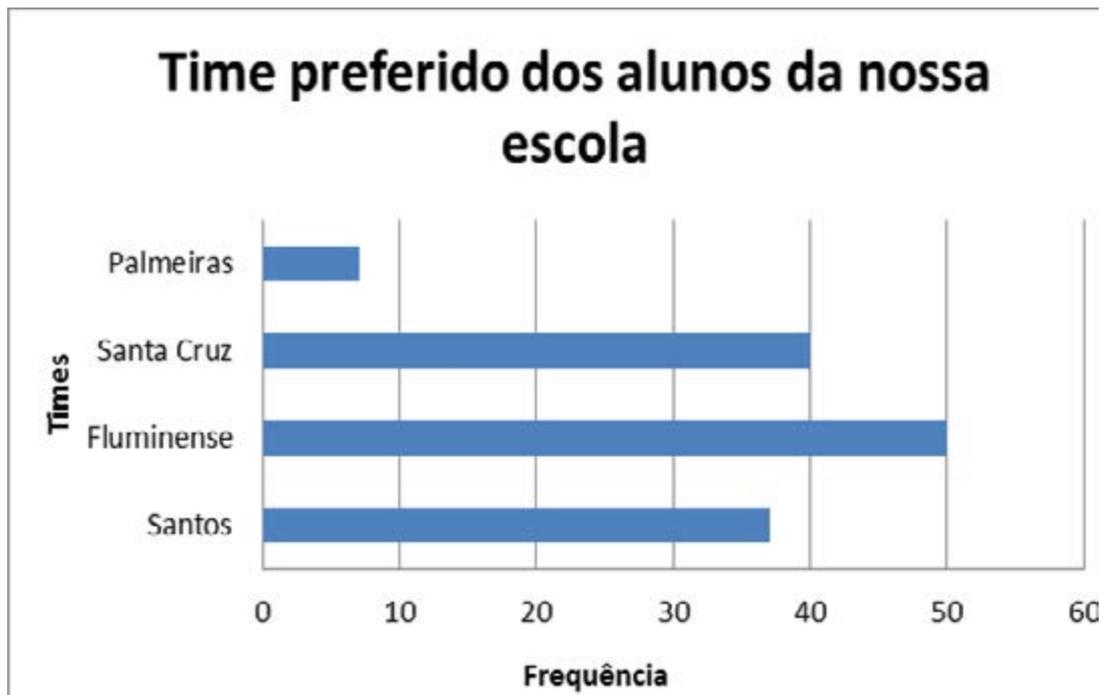


Figura 32: Exemplo de gráfico de barras horizontal

Em algumas situações prefere-se o uso do gráfico de barras horizontais, como é o caso de se ter uma variável com muitas categorias, por exemplo com 10 categorias. Nesse caso é preferível o uso das barras horizontais.

A rigor o gráfico de barras de uma variável deve utilizar apenas uma cor, não deveríamos utilizar duas cores ou mais. Devemos ter cuidado, pois em gráficos de barra ou de colunas as cores são utilizadas para diferenciar variáveis e não categorias em uma variável. O uso de cores diferentes em gráficos de barras/colunas vai, portanto, surgir nos gráficos de barras múltiplas.

### Gráfico de barras duplas

A tabela sobre os medos (Tabela 6) mostra que a maioria dos meninos e das meninas tem mais medo de coisas reais, do que imaginárias. A tabela também mostra que a diferença entre os medos de meninas e meninos é pequena (menos de 2%), observe como isto fica mais evidente com o gráfico de barras duplas, como podemos ver na figura a seguir. Isto é ainda mais evidente quando se utiliza a frequência relativa.

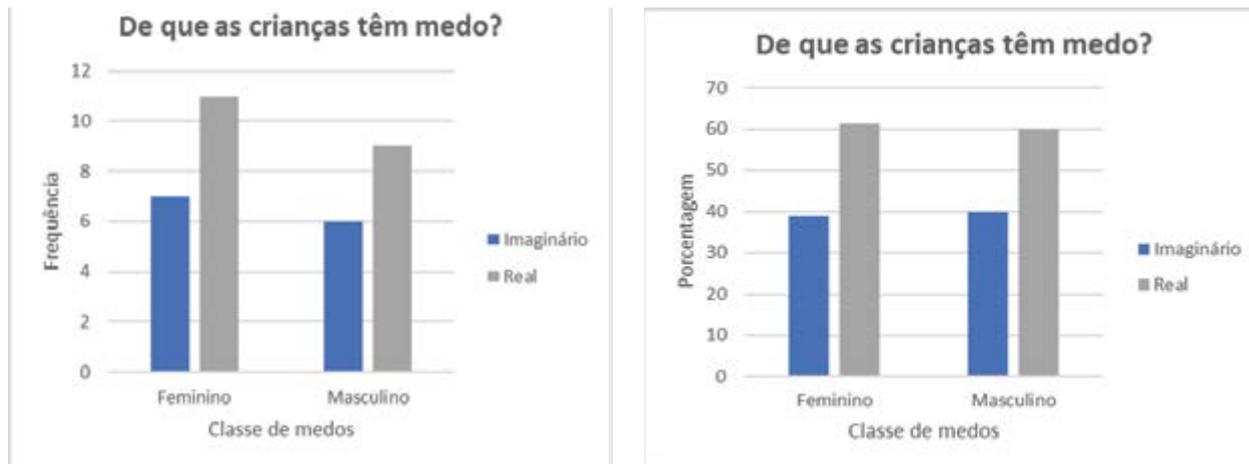


Figura 33: Exemplo de gráfico de barras duplas.

Fonte: Tabela 6.

Professor,

Observe que quando os tamanhos dos grupos forem muito diferentes, o valor absoluto pode nos induzir ao erro. Por essa razão, nestes casos, aconselhamos trabalhar com a estrutura percentual que elimina esse problema.

## Gráfico de setores

Popularmente, conhecido como gráfico de pizza, esse tipo é utilizado para representar variáveis qualitativas quando estamos interessados em observar a relação parte-todo, em especial, as variáveis nominais. No caso das variáveis ordinais, pode ser que exista algum padrão relacionado a ordem das classes e, nesses casos, é mais indicado o uso do gráfico de barras.

A interpretação desse tipo de gráfico pode ser trabalhada com crianças pequenas, entretanto, sua construção não é muito simples. Para construirmos um gráfico de setor é preciso compreender a relação parte-todo expressa nas frações, a divisão dos ângulos de uma circunferência e a proporcionalidade entre frequência e ângulo das partes (categorias) em relação ao todo.

Há, porém, opções para se abordar essa representação. A primeira opção é iniciar um trabalho com frequências ou percentuais mais facilmente desenhados como  $\frac{1}{2}$  e  $\frac{1}{4}$ . Assim, metade equivale a meio círculo, um quarto a quarta parte do círculo e, assim por diante.

## Exemplo de equivalência

Categoria	Nº de alunos	Frações	%
Não	5	1/8	12,5
Pouco	5	1/8	12,5
Regular	10	1/4	25,0
Muito	20	1/2	50,0
Total	40	1	100,0

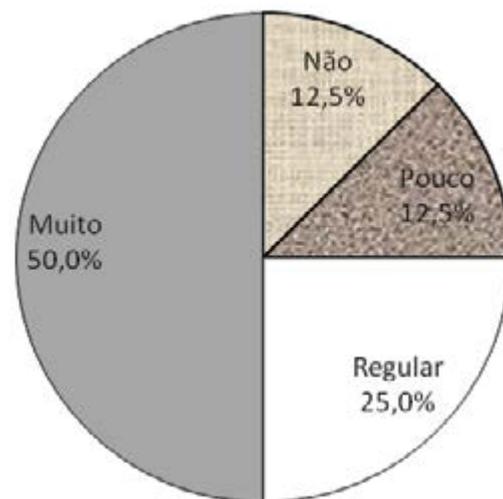


Figura 34: Exemplo de equivalência entre frações, porcentagem e setores do círculo.

Fonte: dados fictícios

A segunda opção é disponibilizar uma malha circular (Anexo F), como mostra a Figura 35, onde cada setor (fatia) corresponde a 5%. Outra opção é construirmos o gráfico em uma planilha eletrônica como o do CALC do Open Office, que é gratuito. Essas planilhas realizam todos os cálculos e calibram as escalas automaticamente, como mostra a Figura 36.

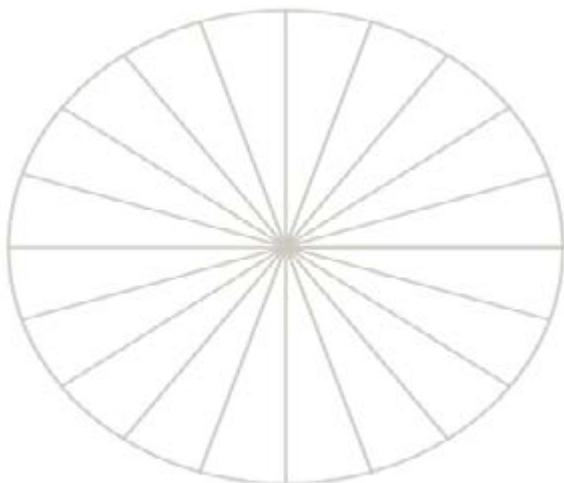


Figura 35: Malha circular

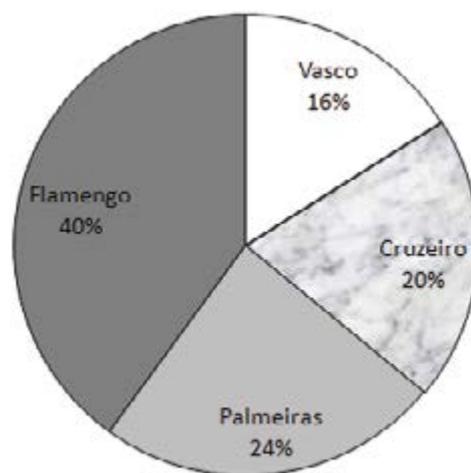


Figura 36: Gráfico construído no Excel.

Fonte: Planilha dos times (Figura 20, p. 47)

## Gráfico de linhas

O gráfico de linhas normalmente é utilizado quando queremos mostrar uma tendência nos nossos dados ao longo do tempo. Se quisermos, por exemplo, saber como as sementes vão germinando e prevermos qual se desenvolverá melhor, podemos ir anotando

do a sua altura em um período de tempo. Na Figura 37 mostramos vários copinhos com cinco sementes no dia em que plantamos (dia zero) e à direita os copos no 5º dia após as sementes terem sido plantadas. Observe as diferentes alturas.



Figura 37: Experimento com a germinação das sementes, no dia zero e 5º dia.

Para medir a altura das plantinhas algumas alternativas podem ser realizadas. Colocar uma régua dentro do copo ou fazer uma escala, encostar cada copo e verificar a altura como mostra a Figura 38.



Figura 38: Alternativas para medir a altura da planta

Podíamos, então, fazer um gráfico no qual os alunos fossem registrando a altura de cada plantinha a cada dia (Figura 39).

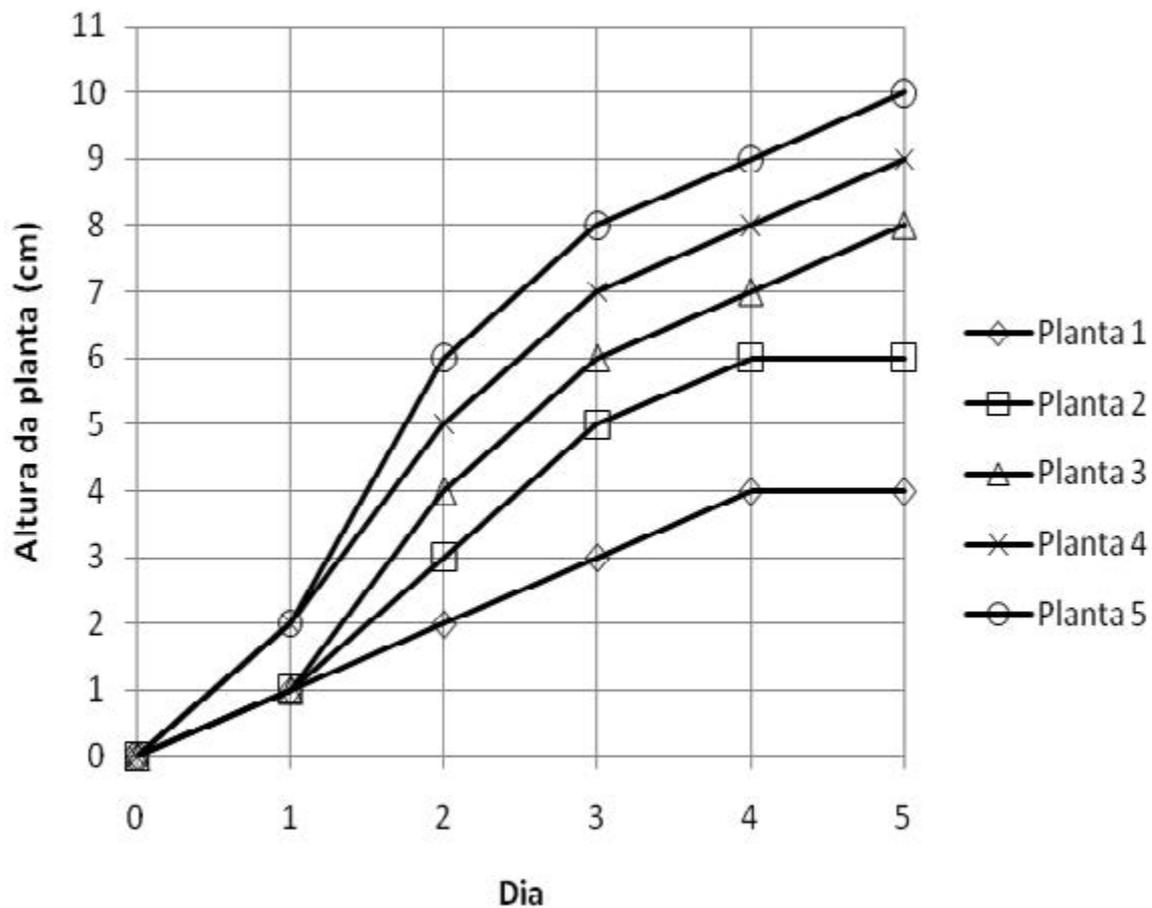


Figura 39: Altura das plantinhas durante os cinco primeiros dias.

A partir de um gráfico como esse é possível refletir sobre o desenvolvimento das plantinhas e fazer previsões sobre a continuidade do desenvolvimento das mesmas ou relacionar o desenvolvimento delas com os tratamentos dados as mesmas como: quantidade de água, luz etc.

Ainda trabalhando com a compreensão das escalas, é fundamental que os alunos compreendam a importância dos intervalos da escala. A Figura 40 apresenta os mesmos dados da Figura 39. Entretanto, a escala vai de 5 em 5 cm até 50 cm e não de 1 em 1, com o valor máximo da de 11, como na figura anterior. Visualmente os gráficos parecem não mostrar os mesmos dados. Mais do que isso, o gráfico que segue dificulta a percepção das diferenças do crescimento das plantinhas.

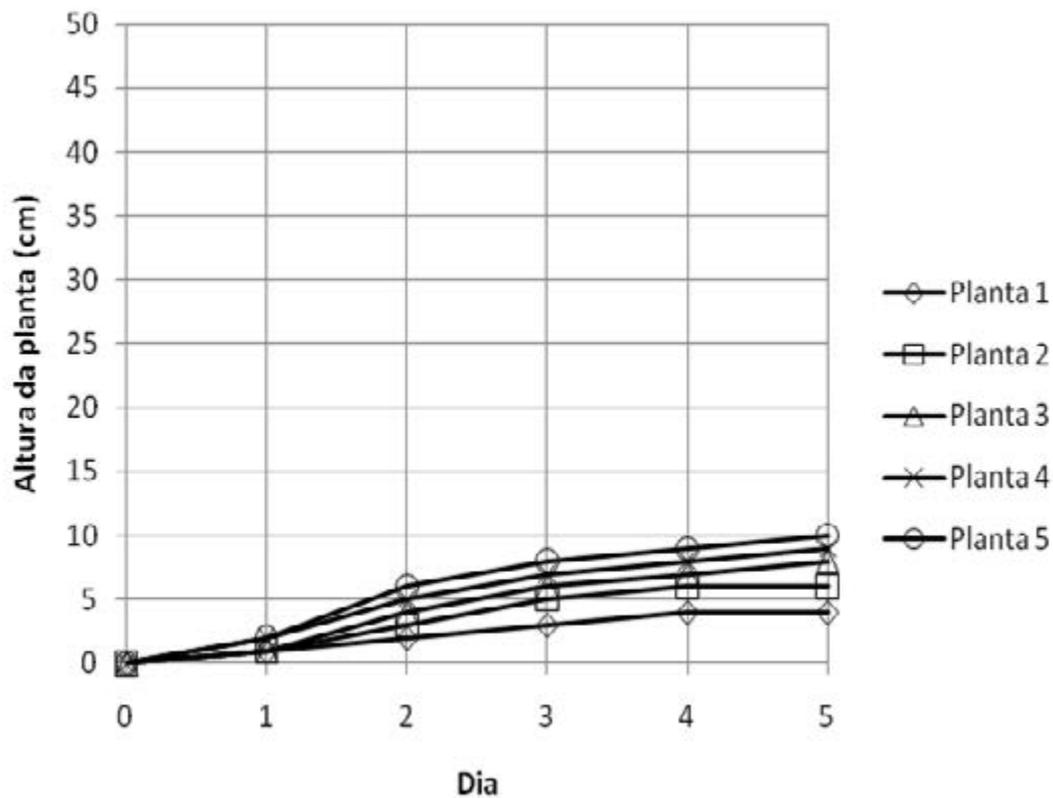


Figura 40: Altura das plantinhas em escala de 0 a 50, em intervalos de 5 cm.

## Uso e cálculo das medidas estatísticas

Assim como organizamos as variáveis em tabelas e gráficos, também podemos sintetizar os dados em um único valor.

Neste nível escolar trabalhamos as medidas de tendência central, sendo as mais conhecidas a moda, a média e a mediana. Essas medidas são chamadas assim, pois indicam em que lugar a massa de dados tende a se concentrar. Também apresentamos a amplitude, que é uma medida de dispersão.

### A moda

A moda é a medida mais intuitiva de todas as medidas, pois se refere a categoria da variável qualitativa ou ao valor da variável quantitativa que se repete com maior frequência.

No caso do exemplo do medo, a moda é a classe de medo real, pois 20 alunos têm

medo de coisas reais (Tabela 6). Já se quisermos saber qual é a idade mais frequente, basta examinar a idade que se repete mais vezes. Neste caso, a idade mais frequente foi 8 anos.

Observamos que para as variáveis discretas que tomam muito valores ou as variáveis contínuas a moda pode não ter nenhum sentido, por isso, em geral não é calculada.

A moda também é a categoria em que se encontra o ponto máximo de um gráfico de barras, de uma variável qualitativa ou discreta que toma poucos valores, resultantes de uma Tabela de Distribuição de Frequência, como por exemplo, o mascote que as crianças têm em casa ou o número de filhos por família (Figura 41).

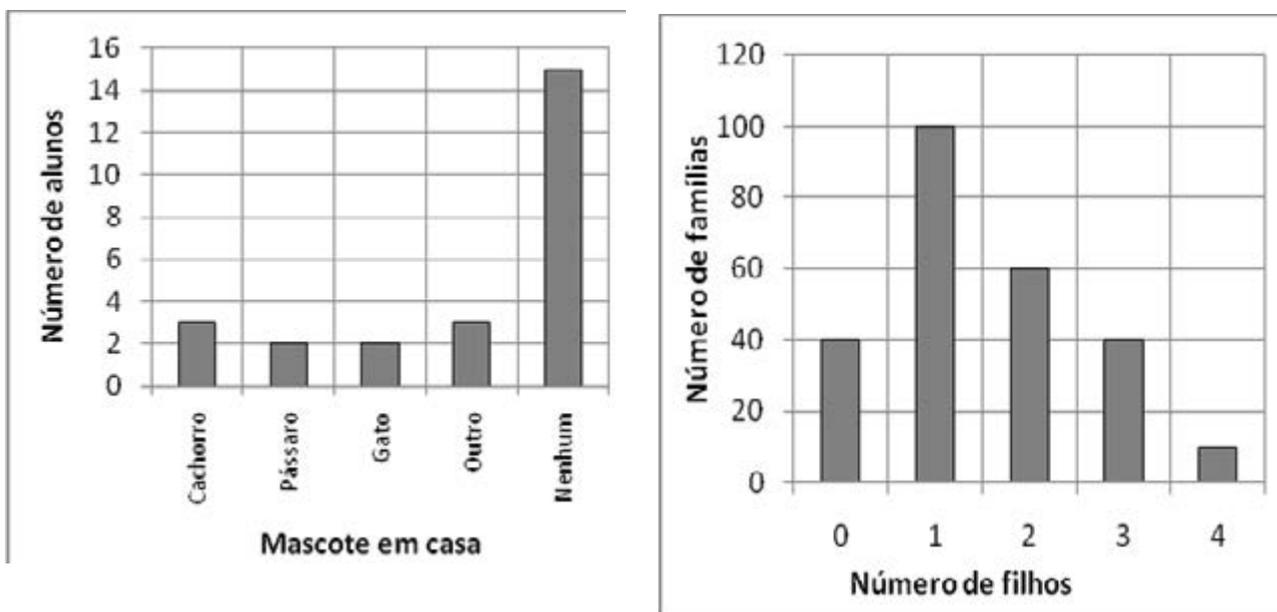


Figura 41: Extraindo a moda de um gráfico de barras.

Assim lendo o gráfico podemos concluir em relação às mascotes que a maioria dos alunos não tem mascote. Já em relação ao número de filhos, a maioria das famílias tem um único filho.

Aqui devemos alertar aos alunos que não podemos generalizar de que a categoria com o ponto máximo de um gráfico representa a moda. Por exemplo, nos gráficos de barras que representam uma série temporal (que estão em função do tempo), o ponto máximo do gráfico raramente poderá ser a moda.

## A média aritmética

Quem não conhece a média aritmética, ou simplesmente média? A maioria das pessoas lida com este conceito de forma bastante familiar e intuitiva, mesmo aquelas que nunca tiveram acesso à escola. As pessoas estão acostumadas a estimar o tempo médio que demoram para chegar ao trabalho, que ficam na fila de banco, dentre outras estimativas.

Esse processo faz parte do cotidiano e está tão arraigado que, às vezes, as pessoas nem percebem o grau apurado de suas estimativas. Para chegar a essas estimativas, ninguém anotou sistematicamente o tempo gasto em cada viagem, somou e dividiu pelo número de viagens; aliás, muitas dessas pessoas nem conhecem o algoritmo da média, mas continuam a utilizar seu conhecimento intuitivo no planejamento de suas atividades rotineiras.

Além desse conhecimento intuitivo, a média também faz parte do cotidiano dos cidadãos, pois cada vez mais a mídia a utiliza junto com outras informações estatísticas. É comum ler nos jornais ou ouvir nas reportagens frases do tipo: “a expectativa de vida da mulher é maior que a do homem”, ou informações referentes à chuva média mensal, à escolaridade média, ao número médio de filhos por casal etc.

Todavia, apesar do aparente conhecimento desse conceito em uso, quando sistematizado, o conceito escolar de média refere-se, em geral, ao domínio do algoritmo: “soma dos valores da variável dividida pelo número de dados envolvidos na soma”:

$$\text{Média} = \frac{\text{Soma dos valores da variável}}{\text{Número de parcelas que compõem a soma}} = \frac{\text{Soma}}{n}$$

na qual “n” é o número de elementos

Podemos observar que a média é a razão entre duas variáveis. No numerador temos a soma dos valores da variável em estudo e no denominador o número de parcelas que compõem essa soma.

Também podemos pensar a média desta forma: a reunião de todos os valores em um único valor (somatório) e depois a distribuição em partes iguais, dentre os elementos que compõem esse todo. Portanto, a compreensão da média implica na compreensão da divisão, dos números decimais e das propriedades da média.

Vejamos alguns exemplos. Suponha que as notas em Matemática de Ana foram: 8 no primeiro bimestre; 6 no segundo; 7 no terceiro e 7 no quarto. Para calcular a nota média, primeiro temos que saber qual foi o total de pontos obtidos nos quatro semestres. Para isso basta somar as quatro notas, o que resulta em 28 pontos  $(8 + 6 + 7 + 7)$  e, depois dividir por quatro, resultando em 7  $(28/4 = 7)$ :

$$\text{Nota média} = \frac{\text{soma\_de\_pontos}}{\text{n.de\_bimestres}} = \frac{8+6+7+7}{4} = \frac{28}{4} = 7 \text{ pontos por bimestre}$$

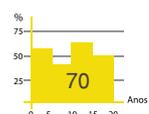
Assim, a nota média de Ana ao longo do ano escolar foi de 7 pontos por bimestre. Porém, o que teria acontecido se ao invés de 8 tivesse obtido 6 no primeiro bimestre. Nesse caso o total de pontos seria de 26  $(6 + 6 + 7 + 7)$ , que dividido por 4 resulta 6,5  $(26/4 = 6,5)$  ou seis e meio.

$$\text{Nota média} = \frac{\text{soma\_de\_pontos}}{\text{n.de\_bimestres}} = \frac{6+6+7+7}{4} = \frac{26}{4} = 6,5 \text{ pontos por bimestre}$$

Neste caso, as crianças precisam conhecer os números decimais. Voltaremos a este ponto logo mais.

Agora analisemos o que acontece quando trabalhamos com variáveis discretas, resultante de contagens, isto é, seus valores tomam números inteiros.

Suponha que temos quatro alunos: Alex, Ana, João e Taís. Alex tem três (3) irmãos, Ana nenhum (0), João um (1) e Taís dois irmãos (2). Para encontrar o número médio de irmãos por aluno somamos o número de irmãos dos quatro alunos, o que resulta seis  $(3 + 0 + 1 + 2 = 6)$ . A seguir, dividimos por quatro, que é número de alunos, resultando 1,5  $(6/4 = 1,5)$ , ver figura que segue. Assim, em média, esses quatro alunos têm 1,5 irmãos. Logo a média é a razão entre o número de irmãos e o número de alunos.



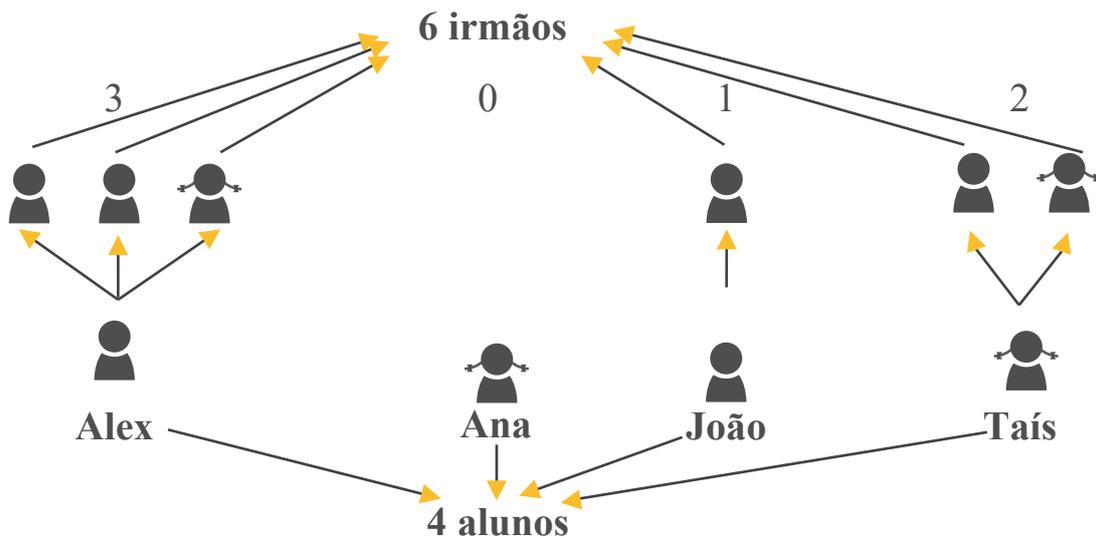


Figura 42: A média de variável discreta.

$$\text{Média de irmãos} = \frac{\text{soma\_n.de\_irmãos\_dos\_alunos}}{\text{n.de\_alunos}} = \frac{6\_irmãos}{4\_alunos} = 1,5$$

irmãos por aluno

Muitas crianças não conseguem compreender que a média pode ter como resultado um valor não inteiro, em um caso como este, que se refere a pessoas. Muitos tendem a arredondar o número ou para um (1) ou para dois (2). Vejamos o que acontece em cada um desses casos. Se cada aluno tem em média um irmão, então os quatro alunos terão 4 irmãos, dois a menos que o verdadeiro número. Se arredondarmos para dois, então os quatro alunos terão oito irmãos, dois a mais do que o verdadeiro número.

Por essa razão é importante que as crianças compreendam a relação inversa que estabelece entre o todo (soma dos valores), a média e o número de elementos que compõem a média. Isto é, se conhecemos a média e o número de elementos que a compõem, podemos conhecer a soma de todos os valores da variável:

$$\text{Média} \times \text{N}^\circ \text{ de elementos} = \text{soma dos valores}$$

Assim, não é possível arredondar os dados de forma indiscriminada, pois em alguns casos comprometemos os mesmos.

Neste nível escolar, as crianças já estão familiarizadas com a divisão e elas têm várias estratégias de distribuição, como mostra o trabalho de Selva e Borba (2005). O problema aparece quando a divisão é inexata ou quando o resultado não tem um referente

na vida real.

Essa dificuldade foi investigada Watson (1996) numa pesquisa que envolveu estudantes da 4<sup>o</sup> ao 9<sup>o</sup> ano do ensino fundamental e 1<sup>o</sup> ano do ensino médio. A pesquisadora solicitou a interpretação da seguinte afirmação: “Em média, os casais jovens têm 2,3 filhos”. Respostas típicas incluíam as seguintes: “existem duas crianças mais velhas e uma mais jovem”, “virgula três significa que, mais tarde, a criança menor contará como três”, “a mãe tem dois filhos e está grávida do terceiro”, dentre outras do gênero. Para a autora, esses estudantes estavam sofrendo conflito cognitivo, pois para eles (0,3) estaria indicando que o filho ainda não nasceu ou que é pequeno para ser contado como um número inteiro. Outro tipo de resposta também apareceu: “a maioria das famílias tem dois filhos (conceito de moda), “poucas famílias têm três filhos”, “mais famílias têm dois filhos do que três”, “a maioria têm dois filhos, mas algumas têm três ou cinco”, “somar todas as crianças e dividir pelo número de famílias, porém o resultado não dá um número inteiro”. Essas expressões denotam uma aproximação do conceito de média, mas, segundo a autora, esse tipo de resposta não foi da maioria. Finalmente, poucos estudantes se referiram ao resultado como um resultado estatístico: “2,3 é apenas uma estatística, uma forma resumida dos dados”.

Outra dificuldade encontrada é que muitos alunos acreditam que porque um elemento não apresenta a característica esse não deve fazer parte da média. Voltemos ao exemplo do número de irmãos. Observe que Ana por não ter irmãos não deve fazer parte do cálculo da média e assim dividem por três ao invés de dividir por quatro resultando uma média de dois ( $6/3 = 2$ ).

Uma forma de evitar esse conflito é colocar os dados em uma lista, como a que segue. Assim, os alunos podem ver que existem 4 alunos e que esses quatro alunos tem ao todo 6 irmãos. Portanto, a média será o resultado de dividir 6 (irmãos) por 4 (alunos), que resulta 1,5 irmãos por aluno.

Aluno	Nº de irmãos
Alex	3
Ana	0
João	1
Táís	2
<b>Total</b>	<b>6</b>

Figura 43: Planilha de dados da quantidade de irmãos.

### a) **Propriedades da média**

De acordo com Strauss e Bichler (1988), a média aritmética possui sete propriedades cujo conhecimento pelo sujeito denota o domínio do conceito:

1. a média está localizada entre os valores extremos  
(mínimo média máximo);
2. a soma dos desvios a partir da média é zero ( $\sum (X_i - \text{média}) = 0$ );
3. a média é influenciada por cada um e por todos os valores  
(média =  $\sum X_i / n$ );
4. a média não necessariamente coincide com um dos valores;
5. a média pode ser uma fração que não tem uma contrapartida na realidade física (por exemplo, o número médio de mascotes por criança);
6. o cálculo da média leva em consideração todos os valores, inclusive os nulos (zeros) e os negativos;
7. a média é um valor representativo dos dados a partir dos quais ela foi calculada. Em termos espaciais, a média é aquela que está mais próxima de todos os valores.

Ainda segundo os autores, as três primeiras propriedades se referem ao aspecto estatístico da média e são básicas enquanto função matemática; as três seguintes referem-se ao aspecto abstrato, permitindo valores não observados; já a última refere-se ao aspecto representativo de um grupo de valores individuais e é o aspecto central da média. Voltaremos a esta discussão quando analisarmos as três medidas juntas: moda, média e mediana.

## **A Mediana**

Para apresentar a mediana vamos nos reportar ao trabalho de Cazorla e Oliveira (2010, pp. 131-134). Essas autoras apresentam e discutem este conceito de forma mais detalhada. Aqui vamos nos concentrar apenas na ideia intuitiva de mediana.

A mediana divide em duas partes iguais um conjunto de dados ordenado. Para encontrar seu valor, primeiro devemos ordenar os dados, a seguir localizamos o lugar que ela ocupa, para assim encontrar seu valor.

Um exemplo fácil e prático é encontrar a altura mediana dos alunos de nossa sala, mas antes vamos fazer um exercício com um número pequeno de dados. Para isso so-

licite aos alunos que anotem em um papel de tamanho padronizado sua altura em centímetros. Em seguida, solicite que cinco alunos da classe se coloquem em pé na frente da sala, mostrando numa folha de papel sua altura (Figura 44):



Figura 44: Exemplo utilizado para calcular a mediana da altura de cinco alunos.

Para encontrar a mediana, o primeiro passo é ordenar os dados. Para isso, solicite aos cinco alunos que se posicionem em ordem crescente de estatura (Figura 45): do mais baixo ao mais alto (pode ser de forma decrescente também).

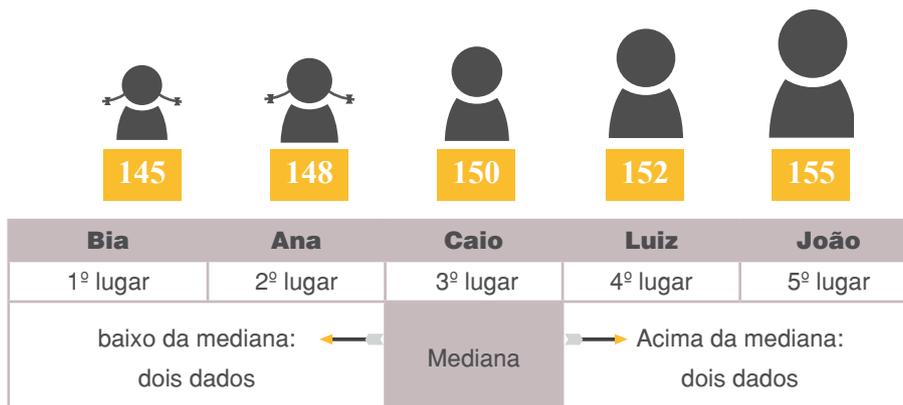


Figura 45: Exemplo do cálculo da mediana da estatura de cinco alunos

Como existem cinco dados ( $n$  é ímpar), a posição que a mediana ocupa será o terceiro lugar (3º), pois abaixo do terceiro lugar temos dois dados e acima também temos dois dados; conseqüentemente, a estatura mediana será a estatura de Caio, que é 150 cm.

Logo, o valor da mediana será a estatura do aluno que ocupa o 3º lugar:

**Mediana = 150 cm**

A interpretação da mediana é bastante intuitiva: no mínimo 50% dos alunos têm estatura menor ou igual a 150 cm e os outros 50% assumem valores maior ou igual a 150 cm.

Se o número de alunos for par, então precisaremos calcular a média dos valores que ocupam as posições centrais. Suponhamos que Francisco, que mede 180 cm, se unisse ao grupo. Agora, o número de alunos seria seis, um número par. Logo, a mediana deveria estar entre o terceiro e o quarto lugares, pois teríamos três dados abaixo e três acima desse valor (Figura 46).

Bia	Ana	Caio	Luiz	João	Francisco	
145 cm	148 cm	150 cm	152 cm	155 cm	180 cm	
1º lugar	2º lugar	3º lugar	4º lugar	5º lugar	6º lugar	
Abaixo da mediana: três alturas			← — — — — — →	Acima da mediana: três alturas		

Figura 46: Exemplo do cálculo da mediana da estatura de cinco alunos.

Logo, o valor da mediana seria o ponto médio desses valores centrais:

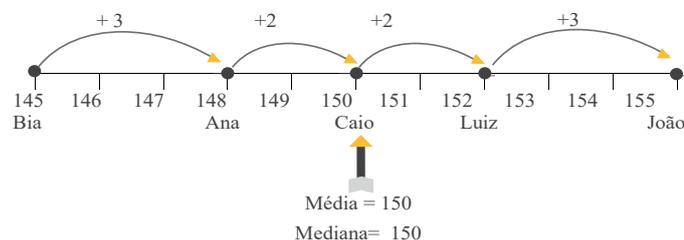
$$\text{Mediana} = \frac{150 + 152}{2} = 151 \text{ cm}$$

A tarefa de ordenar é bastante trabalhada nos anos iniciais, então podemos aproveitar para ensinar o conceito de mediana, quando estivermos trabalhando a ordenação.

Assim, só precisaremos aprender a encontrar a posição da mediana:

- Se o número de dados for ímpar, adicionar um e dividir por dois. Por exemplo 33. Nesse caso  $33 + 1 = 34$ , que dividido por 2 resulta 17. Logo a mediana de um conjunto que tiver 33 elementos será o décimo sétimo lugar.
- Se o número de dados for par, dividir por dois e calcular a média desse valor e de seu sucessor (dados ordenados). Por exemplo, se o número de dados for 34, então  $34/2 = 17$ , assim a mediana será a média dos valores que ocupam o 17º e 18º lugares.

Neste ponto, quando os alunos já aprendem a trabalhar com a reta numérica é importante utilizá-la, pois a ordenação nos traz problemas conceituais, uma vez que todos os alunos estão um ao lado do outro, dando a falsa ideia de que a diferença entre dois alunos é a mesma, o que não é verdade. Vejamos no nosso exemplo, colocando os dados na reta numérica:



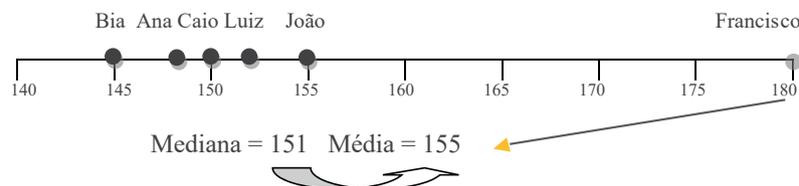
Como podemos ver, a distância entre os alunos não é a mesma e isso piora quando colocamos altura de Francisco no conjunto.

Aproveitamos esta representação para sinalizar onde se localiza a média e a mediana. Ambas tomam o valor de 150:

$$\text{Média} = \frac{145 + 148 + 150 + 152 + 155}{5} = \frac{750}{5} = 150$$

A diferença da mediana em relação à média é que a mediana não é afetada pelos valores fora do padrão (outliers ou discrepantes). No exemplo com os cinco alunos, a mediana e a média coincidem; logo, qualquer uma delas representa bem o conjunto de dados, pois as estaturas não são muito diferentes.

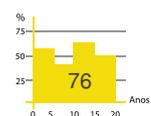
Todavia, quando Francisco, com 30 cm acima da média, junta-se ao grupo, eleva o valor da média em cinco centímetros pois aos 750 cm dos cinco alunos, devemos adicionar 180 cm da altura de Francisco, totalizando 930 cm, que dividido por seis resulta 155 cm. Isto é, cinco cm a mais, enquanto que o impacto na mediana foi de apenas um centímetro.



Neste exemplo, vemos como um valor “fora do padrão” eleva a média e já a mediana não sofre tanto. Assim, a pergunta é: qual das duas medidas devemos usar? É o que discutiremos logo mais.

A mediana também pode ser encontrada a partir dos dados agrupados numa TDF. Calculemos a mediana das idades dos 33 alunos. Nos dados simples (sem agrupar), a Mediana ocupará a 17ª posição  $(33+1)/2$ , que neste caso será 8 anos, como podemos ver no esquema a seguir:

1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°
7	7	7	7	7	8	8	8	8	8	8	8	8	8	8	8	8
16 dados abaixo																←



continuação

continua

17°	18°	19°	20°	21°	22°	23°	24°	25°	26°	27°	28°	29°	30°	31°	32°	33°
8	8	8	8	8	8	8	8	9	9	9	9	9	9	9	9	9
(*)	 16 dados acima															

(\*) Estamos repetindo o 17º lugar apenas com fins didáticos.

Uma forma prática para encontrar a mediana, quando os dados estão agrupados em uma TDF, é calcular a frequência acumulada. Para isso basta perguntar quantos alunos temos até aquela idade e somar todos os valores anteriores incluindo o valor da idade em questão. Por exemplo:

- Quantos alunos temos com idades até 7 anos? 5 alunos
  - Quantos alunos temos com idades até 8 anos?  $5 + 19 = 24$  alunos
  - Quantos alunos temos com idades até 9 anos?  $5 + 19 + 9 = 33$  alunos
- Sendo que o último valor deve coincidir com o número total de alunos

Para encontrar a mediana basta examinar esses dados e encontrar o número que contenha pela primeira vez 17. Assim, constatamos que o número 25 contém o número 17 pela primeira vez, logo a idade mediana será 8 anos.

Idade	Nº de alunos	Nº de alunos acumulados
7	5	5
8	19	24
9	9	33
Total	33	

## Amplitude total e ponto médio

Como vimos, a Estatística trata de dados que variam, alguns mais do que outros. Assim a variabilidade dos dados é o coração da Estatística e a justificativa da sua existência. Observando os fenômenos que permeiam nossas vidas, vemos que esses são fortemente marcados pela existência da variabilidade.

Existem várias estatísticas que medem essa variabilidade, chamadas de medidas de dispersão, como amplitude total, variância, desvio padrão, coeficiente de variação. Quanto maior a variabilidade dos dados, maiores serão os valores dessas estatísticas. Neste livro apresentaremos apenas a amplitude total, por ser bastante intuitiva.

A Amplitude Total (AT) é a medida do comprimento do intervalo de possíveis valores da variável, calculado pela distância entre o valor máximo e o mínimo da variável. É uma medida de dispersão absoluta, a mais simples e mais intuitiva de todas:

**Amplitude** = valor máximo – valor mínimo

Calculemos a amplitude no caso das alturas dos cinco alunos a amplitude, como o menino mais baixo mede 145 cm e o mais alto 155 cm, então a amplitude será de 10 centímetros ( $155 - 145 = 10$ ). Já quando Francisco entra essa amplitude vai para 325 centímetros ( $180 - 145 = 35$ ).

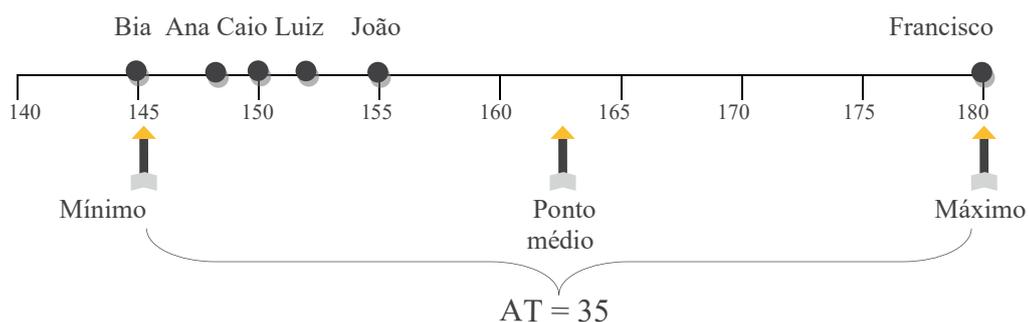
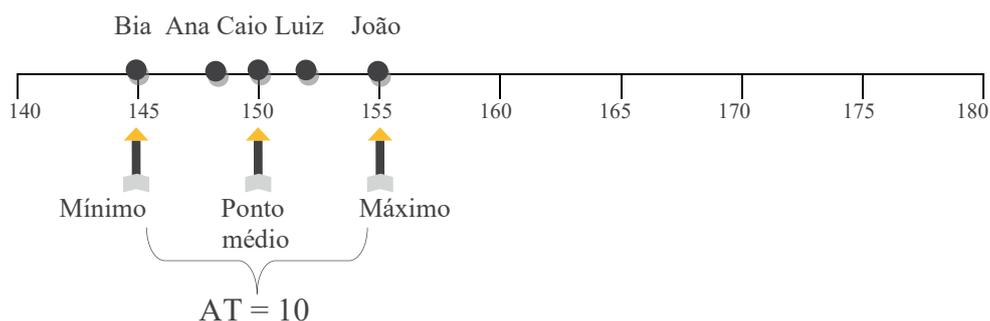
Assim, podemos concluir que quando temos os cinco alunos temos um conjunto de alunos cujas alturas são similares. Mas quando o Francisco entra no grupo ele é tão alto que “discrepa” do grupo, logo a variação é maior.

Um ponto de referência importante é o ponto médio, definido como a média do valor mínimo e do valor máximo. O ponto médio divide exatamente ao meio a amplitude da variável.

**Ponto Médio** = (valor máximo + valor mínimo) / 2

O ponto médio serve para analisar se o conjunto de dados se concentra para a esquerda (a maioria dos dados assumem valores menores) ou à direita (a maioria assume os maiores valores) ou se distribui em torno desse ponto.

Calculemos o ponto médio da altura dos cinco alunos:  $(145 + 155)/2 = 150$ , já quando Francisco entra, esse valor aumenta para  $(145 + 180)/2 = 162,5$ .



## Variáveis contínuas ou discretas que tomam muitos valores: medidas e gráficos

Para trabalhar este tipo de dado vamos utilizar os dados da pesquisa da altura das crianças, neste exemplo coletados em metros, constantes no Quadro 6. Examinando a planilha de dados, os alunos devem decidir quem é mais alto: os meninos ou as meninas. Os alunos podem observar na planilha e notar que a criança mais alta é um menino. Mas será que isto significa que os meninos são mais altos? Caso isto ocorra mostre aos alunos que o mais baixo também é um menino. Portanto, continuamos a mesma questão. Como decidir quem é mais alto, os meninos ou as meninas?

Para discutir melhor vamos calcular a média e a mediana, assim como vamos usar outros tipos de gráficos: o dotplot e o histograma.

Construir gráficos para variáveis contínuas ou para variáveis discretas que tomam muitos valores pode ser complexo para este nível escolar. Contudo o dotplot (Figura 47) é bastante intuitivo, uma vez que a localização dos números na reta numérica é cada vez mais utilizada pelos livros didáticos.

Para construir o dotplot, primeiro ordenamos os dados.

1,00	1,01	1,04	1,08	1,08	1,10	1,10	1,10	1,10	1,13	1,15	1,15	1,20	1,20	1,20
1,20	1,30	1,30	1,34	1,35	1,35	1,35	1,35	1,36	1,40	1,43	1,45	1,46	1,48	1,52

Checamos o valor mínimo e o valor máximo, desenhamos a reta numérica e representamos os valores com um ponto.

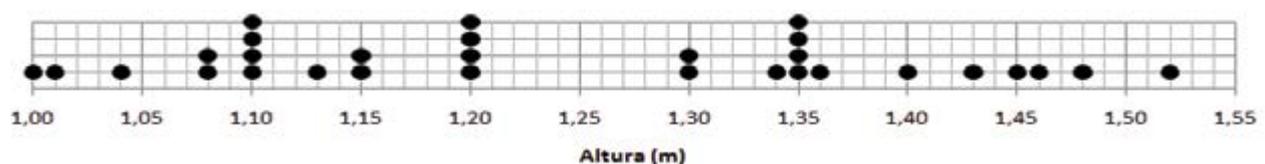


Figura 47: Dotplot da altura das crianças.

Agora precisamos decidir em quantas classes ou faixas vamos trabalhar. Se optarmos por classes de 10 centímetros de amplitude, teremos seis classes, como apresentamos na Tabela 7.

O histograma (Figura 48) é construído com retângulos (barras/colunas) contíguos,

colados uns aos outros. A base dos retângulos é formada pelas classes e a altura pela frequência absoluta (relativa) respectiva. Só haverá espaços caso a frequência absoluta da classe for zero.

Altura (m)	Nº de alunos	%
1,00 a 1,09	5	16,7
1,10 a 1,19	7	23,3
1,20 a 1,29	4	13,3
1,30 a 1,39	8	26,7
1,40 a 1,49	5	16,7
1,50 a 1,59	1	3,3
<b>Total</b>	<b>30</b>	<b>100,0</b>

Fonte: Quadro 13, Anexo A.

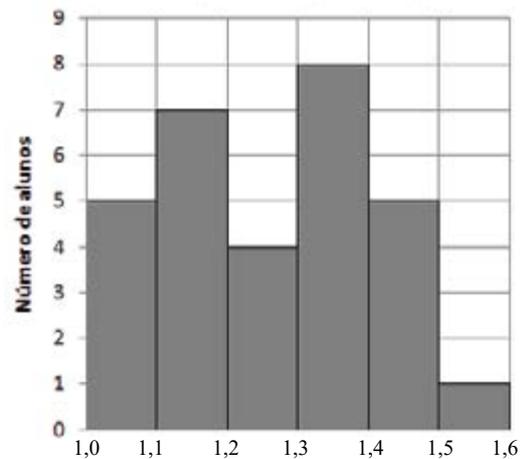


Figura 48: Exemplo da construção do histograma.

A seguir vamos calcular a média e a mediana. Para calcular a média devemos somar a altura dos 30 alunos que resulta 37,28 metros, que dividido por 30 resulta 1,24. O valor da mediana é bem mais simples, pois como temos 30 dados, a mediana será a média dos valores que ocupam o 15º e 16º lugares, que são 1,20 e 1,20. Assim, vemos que a média é ligeiramente maior do que a mediana.

Agora, vamos examinar se o gênero é responsável por essa configuração. Para facilitar os cálculos, colocamos os dados ordenados e como temos 15 dados para cada grupo, a mediana ocupará o 8º lugar, que está sombreado na Figura 49. Na Figura 50 apresentamos o dotplot, que nos dá uma ideia da distribuição dos dados.

Gênero	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º	11º	12º	13º	14º	15º
Meninas	1,08	1,08	1,10	1,15	1,20	1,20	1,20	1,30	1,35	1,35	1,36	1,40	1,45	1,46	1,48
Meninos	1,00	1,01	1,04	1,10	1,10	1,10	1,13	1,15	1,20	1,30	1,34	1,35	1,35	1,43	1,52

Figura 49: Apresentação dos dados ordenados e da mediana por gênero.

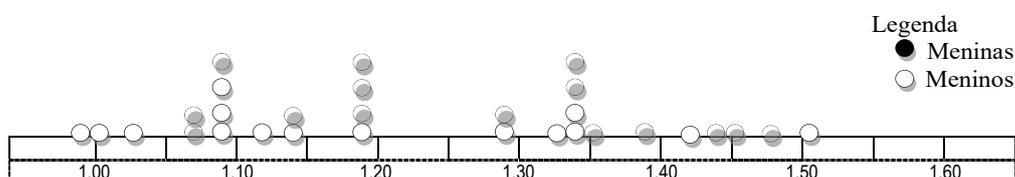


Figura 50: Diagrama de pontos (dotplot) da altura dos alunos.

O dotplot nos permite visualizar que a distribuição da altura dos meninos é mais dispersa, pois são eles que têm o menor e o maior valor da altura ( $A = 0,52$ ), enquanto que a variação das alturas das meninas é menor ( $A = 0,40$ ). Contudo, talvez o histograma seja mais eficaz na comunicação do padrão dos dados. Para construí-lo, primeiro construímos o TDF.

**Tabela 8:** Distribuição das alturas dos alunos por gênero

Altura (m)	Meninas		Meninos		Total	
	N	%	n	%	N	%
1,00 a 1,09	2	13,3	3	20,0	5	16,7
1,10 a 1,19	2	13,3	5	33,3	7	23,3
1,20 a 1,29	3	20,0	1	6,7	4	13,3
1,30 a 1,39	4	26,7	4	26,7	8	26,7
1,40 a 1,49	4	26,7	1	6,7	5	16,7
1,50 a 1,59	0	0,0	1	6,7	1	3,3
Total	15	100,0	15	100,0	30	100,0

A disposição visual dos dados permitida pelo histograma facilita que identificamos que, apesar de ser menino a criança mais alta, as alturas dos meninos se concentram mais nos intervalos de menores alturas, enquanto que as meninas parecem estar com alturas melhor distribuídas (Figura 51).

Cada representação vai facilitar a leitura e identificação das informações melhor. No caso, passamos a desconfiar que o fato de a maior criança ser um menino não implica que os meninos não são os mais altos.

Outra possibilidade seria lançar mão das médias da altura dos meninos e das meninas. Para isso somamos as alturas das meninas (19,16 m) e dividimos por 15, resultando uma média de 1,28 metros por menina. Igualmente somamos as alturas dos meninos (18,12 m) e dividimos por 15, resultando uma média de 1,21 metros por menino.

$$\text{Meninas} \quad \text{Soma} = (1,08 + 1,08 + 1,10 + 1,15 + 1,20 + 1,20 + 1,20 + 1,30 + 1,35 + 1,35 + 1,36 + 1,40 + 1,45 + 1,46 + 1,48) = 19,16$$

$$\text{Média} = 19,16 / 15 = 1,28 \text{ m/menina}$$

$$\text{Meninos} \quad \text{Soma} = (1,00 + 1,01 + 1,04 + 1,10 + 1,10 + 1,10 + 1,13 + 1,15 + 1,20 + 1,30 + 1,34 + 1,35 + 1,35 + 1,43 + 1,52) = 18,12$$

$$\text{Média} = 18,12 / 15 = 1,21 \text{ m/menino}$$

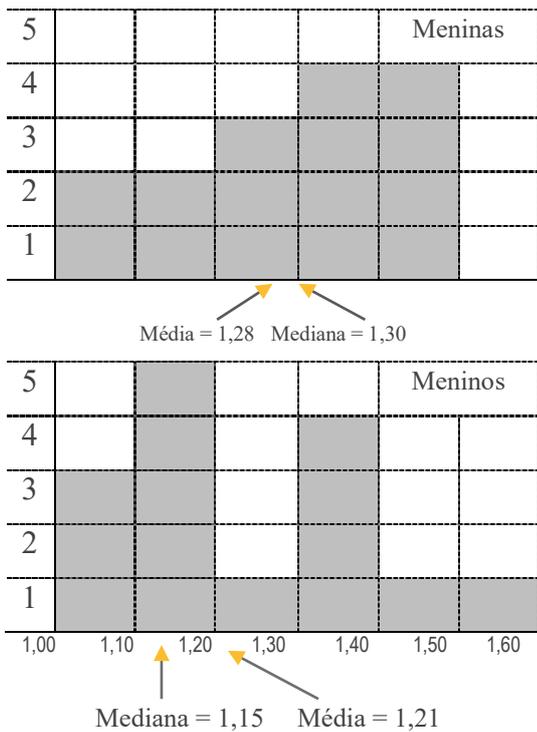


Figura 51: Distribuição da altura dos alunos por gênero

outras questões podem surgir.

## O cruzamento dos dados

Uma investigação é feita de forma que, durante o processo de levantamento das variáveis, observam-se diversos fatores que terminam por influenciar a variável que se está buscando investigar, como é o fato da altura aqui estudada. Nesse caso, já observamos desde o Capítulo 2 que a idade é importante, é um fator a se controlar. Portanto, várias outras perguntas ou afirmações podem ser feitas pelas crianças, como:

- Será que, em média, as meninas de cada idade são sempre mais altas?
- Os meninos de 10 e 11 anos são mais altos, mas antes disso não?

A cada pergunta feita teremos a boa escolha de uma representação e de medidas estatísticas diferentes.

Para responder essas perguntas sugerimos dividir a classe em duplas e, a cada dupla, calcula a média de apenas uma idade e um gênero, registrando o número de alunos e a média. Por exemplo, a Dupla 1 calcula a média das meninas de 5 anos, formada apenas por Ana e Tereza, logo  $(1,15 + 1,20)/2$  resultando uma média de 1,175 que será arredondado para 1,18. No Quadro 7 apresentamos o resultado do cálculo das médias por

Analisando visualmente o histograma podemos ver que as alturas das meninas se concentram ao lado direito, assim a média fica a esquerda da Mediana.

Já o histograma da altura dos meninos, vemos que há dois grupos, um concentrado nos menores valores e o segundo em torno do centro, em geral à esquerda. Isto é confirmado pela mediana que fica à esquerda da média, indicando que a altura dos meninos se tende aos menores valores.

Nesse caso, podemos verificar que em média as meninas são mais altas do que os meninos. No entanto, algumas ou-

idade e gênero.

Neste ponto é preciso discutir com os alunos o significado da média de um grupo formado por um único aluno. Qual é a representatividade?

Outra questão que vamos aproveitar para discutir aqui é a adequação do gráfico de barras para representar a relação da altura média com a idade, assim como a “manipulação” da escala.

**Quadro 7:** Resultados do cálculo das médias por idade e gênero

Nome	Turma	Gênero	Idade	Altura	Nº de alunos	Média	Nº de alunos	Média
Tereza	1º ano A	F	5	1,15	2	1,18	3	1,15
Ana	1º ano B	F	5	1,20				
Alberto	1º ano A	M	5	1,10				
Geisa	2º ano B	F	6	1,08	2	1,08	5	1,04
Telma	1º ano B	F	6	1,08				
João	1º ano A	M	6	1,00	3	1,02		
Valter	2º ano A	M	6	1,01				
Pedro	1º ano B	M	6	1,04				
Maria de Fátima	3º ano B	F	7	1,20	3	1,22	4	1,19
Gertrudes	2º ano B	F	7	1,35				
Marta	3º ano B	F	7	1,10				
Marcos	2º ano A	M	7	1,10	1	1,10		
Telma	2º ano A	F	8	1,20	3	1,29	7	1,23
Aída	4º ano A	F	8	1,36				
Maria	3º ano A	F	8	1,30				
Manoel	3º ano A	M	8	1,10	4	1,18		
José	3º ano A	M	8	1,20				
Miguel	3º ano B	M	8	1,13				
Maurício	2º ano B	M	8	1,30				
Michelle	4º ano A	F	9	1,40	1	1,40	3	1,39
Michael	4º ano B	M	9	1,35	2	1,39		
Arquimedes	4º ano B	M	9	1,43				
Severina	4º ano A	F	10	1,45	3	1,46	6	1,37
Maria	5º ano A	F	10	1,46				
Tereza	5º ano B	F	10	1,48				
João	5º ano A	M	10	1,15	3	1,28		
Severo	4º ano B	M	10	1,35				
Josué	5º ano A	M	10	1,34				
Miriam	5º ano B	F	11	1,35	1	1,35	2	1,44
Norberto	5º ano B	M	11	1,52	1	1,52		

Na Tabela 9 e na Figura 52 apresentamos as médias das alturas por idade e gênero. Nessas representações podemos ver que, em geral, a altura média das meninas é maior do que os meninos em todas as idades com exceção dos 11 anos, onde se inverte essa tendência.

Observamos na que o gráfico da esquerda utiliza uma escala que inicia em zero, já o gráfico da direita utiliza uma escala que inicia em 0,9 m. A escala iniciada em 0,9 acen-tua as diferenças. Esse recurso tem como fim realçar as diferenças entre os valores re-presentados. Este procedimento pode ter impactos positivos ou negativos, de acordo com o que queremos comunicar. Entretanto, muitas vezes é utilizado para mascarar dados.

**Tabela 9:** Média da altura por gênero e idade

Gênero	Idade em anos completos						
	5	6	7	8	9	10	11
Feminino	1,18	1,08	1,22	1,29	1,40	1,46	1,35
Masculino	1,10	1,02	1,10	1,18	1,39	1,28	1,52
Total	1,15	1,04	1,19	1,23	1,39	1,37	1,44

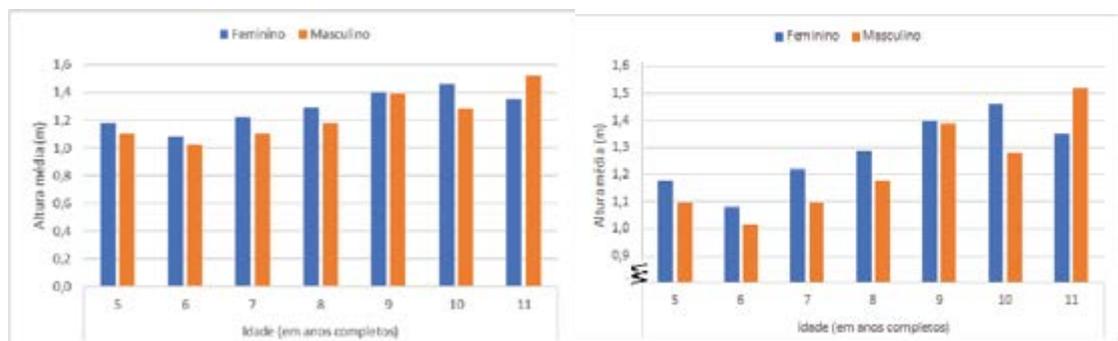


Figura 52: Média da altura por gênero e idade

Na Figura 53 apresentamos os mesmos dados, mas ao invés de utilizarmos o gráfico de barras, utilizamos o gráfico de linhas. Observe como o gráfico de linhas é mais eficaz para transmitir a informação da tendência dos dados, do que o gráfico de barras. Entretanto, isso só é possível porque estamos utilizando uma variável numérica contínua e não uma variável discreta.

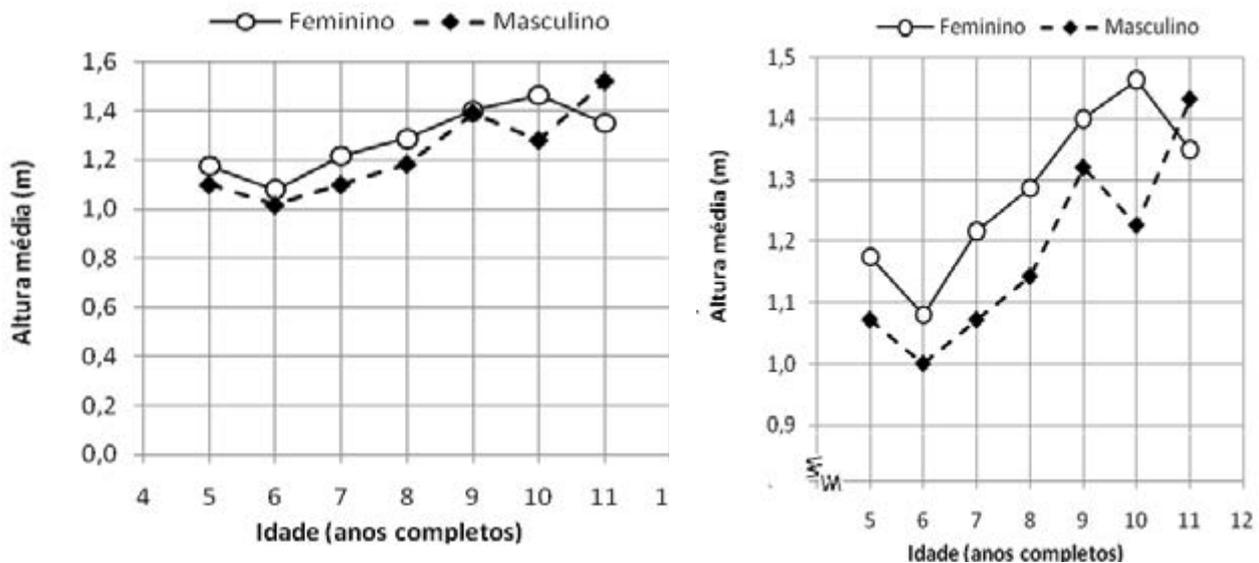


Figura 53: Média da altura por gênero e idade

## Integração de tabelas, gráficos e medidas estatísticas

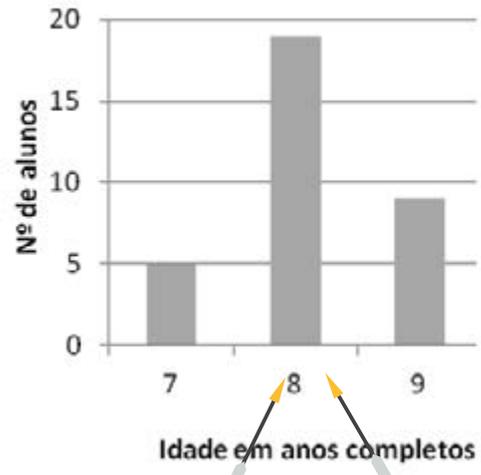
A compreensão das propriedades das medidas, média, mediana, moda e amplitude, assim como a configuração em tabelas e gráficos, permite ao aluno aprimorar seu conhecimento sobre a natureza dos dados.

Para analisar a variável “idade” da pesquisa sobre o medo, muitas vezes, podemos intuir o valor da média analisando visualmente os dados. Por exemplo, suponha que desejamos, sem fazer cálculos, apenas analisando a tabela ou o gráfico, conhecer o valor da média de idade dos alunos.

O nosso primeiro palpite, neste exemplo, deveria ser a moda, 8 anos, pois a maioria dos alunos tem essa idade. Agora, analisemos a idade 7 anos, que tem cinco alunos e, na idade de 9 anos, temos 9 alunos, isto é, 4 alunos a mais. Assim a média será um pouco maior do que 8 anos, provavelmente estará ente 8,1 e 8,2, que será nosso palpite definitivo. Na seção anterior calculamos a média da idade que resultou 8,1 anos por aluno. Vejamos, a média foi 8,1, a mediana e a moda igual a 8 e a amplitude total dois anos. Aparentemente estamos diante de um conjunto de dados bastante homogêneo (Figura 54). Porém, temos um número significativo de crianças com 9 anos (27,2%), que possuem uma defasagem de um ano. Lembramos que se trata de uma turma do 3º ano, cuja idade recomendada é de 8 anos.

Idade	Nº de alunos	%
7	5	15,2
8	19	57,6
9	9	27,3
Total	33	100,0

Média = 8,12



Mediana = Moda = 8      Média = 8,1

Figura 54: Distribuição das idades dos alunos da pesquisa sobre o medo.

Fonte: Tabela 4.

## Capítulo 5

# A INTERPRETAÇÃO DOS RESULTADOS

Uma pesquisa não termina com a organização e tratamento dos dados. Ao final precisamos voltar às questões que deram origem à pesquisa e buscar responder essas questões. E, ainda, se tivessem sido geradas novas hipóteses, verificar sua validade.

Até o capítulo anterior aprendemos a traduzir em números nossas questões de pesquisa. Escolhemos um tema para investigar, formulamos questões de investigação, aprendemos a construir instrumentos para coletar os dados, coletamos e tratamos os dados.

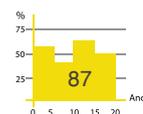
Agora já temos em mãos tabelas, gráficos e estatísticas, chegou a hora de responder as questões de pesquisa e socialização dos resultados, isto é, a construção da argumentação sobre os resultados, o que implica na escolha da representação que melhor comunica os argumentos, bem como a elaboração de um relatório.

### **A pesquisa sobre o medo das crianças**

Na pesquisa sobre o maior medo das crianças, queríamos saber quais eram os maiores medos das crianças de nossa sala e conjecturamos que a idade e o gênero poderiam ser dois fatores importantes. Então, agora retomemos os dados e a partir da leitura conjunta vamos tentar responder às perguntas.

Foi realizada uma pesquisa com 33 alunos do 3º ano A, do Colégio ABC, da cidade de Recife, dos quais 18 eram meninas (54,5%) e 15 eram meninos (45,5%), conforme Tabela 1. Suas idades variaram de 7 a 9 anos sendo que a maioria tinha 8 anos e a idade média foi de 8,12 anos.

Embora a idade média estivesse próxima da idade recomendada para este ano escolar, 27,3% dos alunos tinha 9 anos, isto é, estavam defasados em um ano relativamente ao esperado (Tabela 5).



A leitura geral da planilha de dados (Quadro 5) indica que o medo mais frequente foi de rato (3 alunos) e de altura (3 alunos, considerando “lugar alto” como altura), os outros medos só se repetiram duas vezes ou apenas uma única vez.

Em relação à idade, também não parece haver interferência dessa variável em relação à classe de medo, como podemos observar na Tabela 10.

**Tabela 10:** Distribuição dos tipos de medo em relação à idade dos alunos

Classe de medo	7 anos		8 anos		9 anos		Total	
	N	%	n	%	N	%	N	%
Imaginário	2	40	7	36,8	4	44,4	13	39,4
Real	3	60	12	63,2	5	55,6	20	60,6
Total	5	100,0	19	100,0	9	100,0	33	100,0

Na Tabela 11 observamos que 60,6% dos alunos têm medo de coisas reais e 39,4% de coisas imaginárias e, que isso não difere por gênero, pois a diferença é de apenas 1,1%.

**Tabela 11:** Distribuição dos tipos de medo por subclasse em relação ao gênero e a idade.

Subclasses de medo	Gênero		Idade em anos			Total
	Feminino	Masculino	7	8	9	
Imaginário Folclore	11,1	20,0	20,0	15,8	11,1	15,2
Imaginário Personagem	27,8	20,0	20,0	21,1	33,3	24,2
Real Bicho	22,2	26,7	40,0	26,3	11,1	24,2
Real Pessoa	16,7	26,7	20,0	15,8	33,3	21,2
Real Situação	22,2	6,7	0,0	21,1	11,1	15,2
Total	100,0	100,0	100,0	100,0	100,0	100,0

Contudo, quando analisamos dentro das subclasses de medo, percebemos sutis diferenças. Por exemplo, as meninas tendem a ter mais medo de personagens imaginários e os meninos de bichos e pessoas reais. Quanto à idade, parece que quanto mais velho, vai diminuindo o medo de bichos, mas aumenta o medo em relação aos personagens imaginários e pessoas, como pode ser visto na Tabela 11. Essas diferenças são mais visíveis na Figura 55.

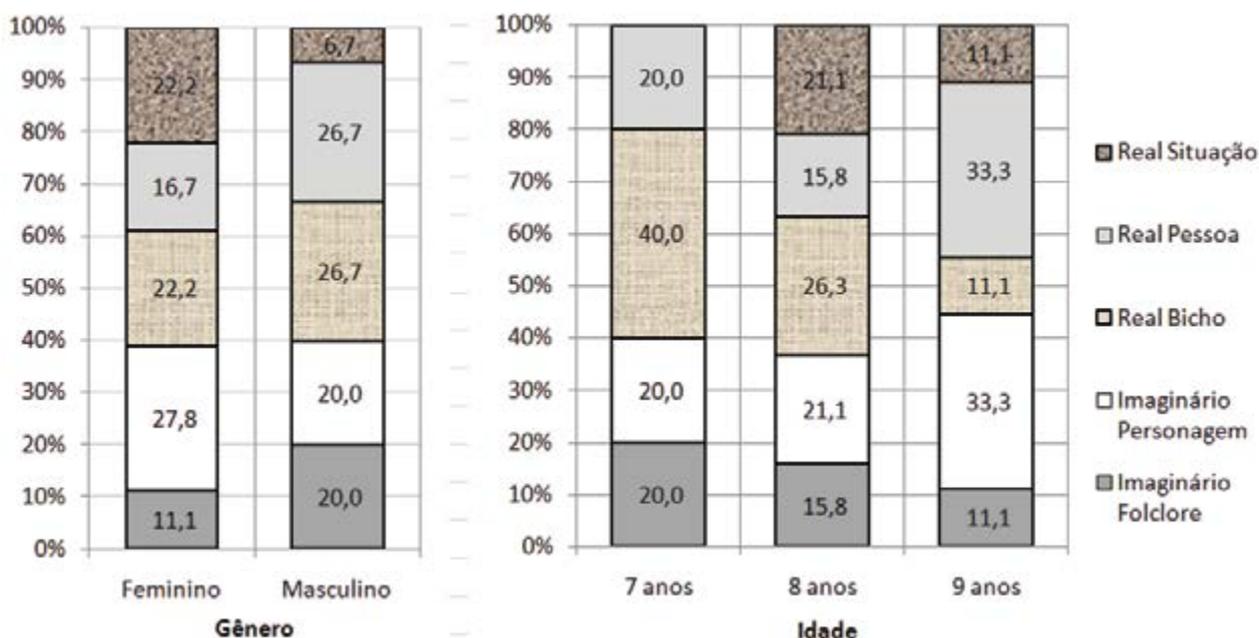


Figura 55: Distribuição do medo por subclasses segundo gênero e idade.

## A pesquisa sobre a altura das crianças

Retomando a questão de investigação sobre a altura das meninas e dos meninos podemos fazer nosso relatório:

Foram investigados 30 alunos, do 1º ao 5º ano, sendo seis de cada ano, três meninas e três meninos. A altura média dos 30 alunos foi de 1,24 metros, sendo que a altura média das meninas foi de 1,28 metros e a altura média dos meninos foi de 1,21 metros. Isto é, em média, as meninas são mais altas do que os meninos.

Ao analisar a altura média dos alunos por gênero segundo a idade, observamos que em todas as idades as meninas apresentam médias maiores, com exceção do último ano (11).

Como este resultado pode entrar em conflito com as concepções prévias dos alunos (eles acreditam que os meninos são sempre mais altos do que as meninas) é preciso que o professor esteja preparado para dar uma explicação científica para este resultado.

Para isso, o professor pode solicitar ajuda ao professor de Educação Física e ao professor de Ciências e juntos trabalhar o tema transversal “Saúde da Criança e do Adolescente”.

Nas Figura 56 e Figura 57 apresentamos as curvas de crescimento para meninos e meninas. Nesses gráficos podemos ver que o ritmo de crescimento é diferenciado. Em

geral as meninas crescem mais rápido de 7 a 11 anos, enquanto que os meninos crescem mais a partir dos 13 anos, fato melhor visualizado no gráfico da que segue.

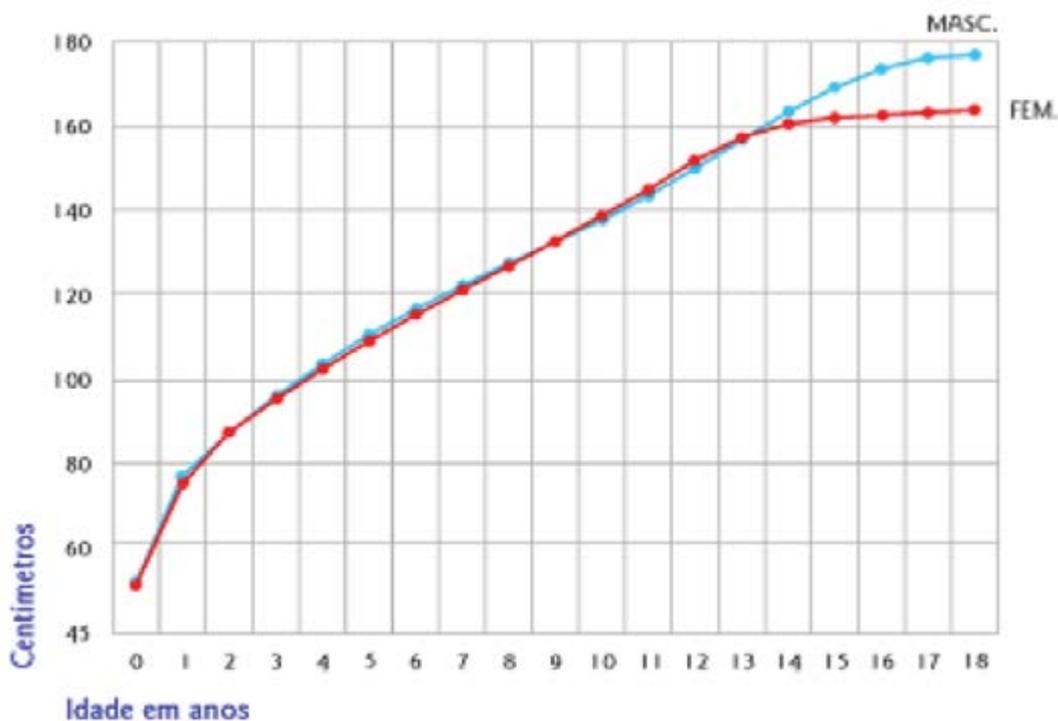


Figura 56: Alturas médias por idade da população de referência do NCHS – Meninos e meninas de 0 a 18 anos.  
Fonte: Gráfico 2, Brasil (2002, p. 21).

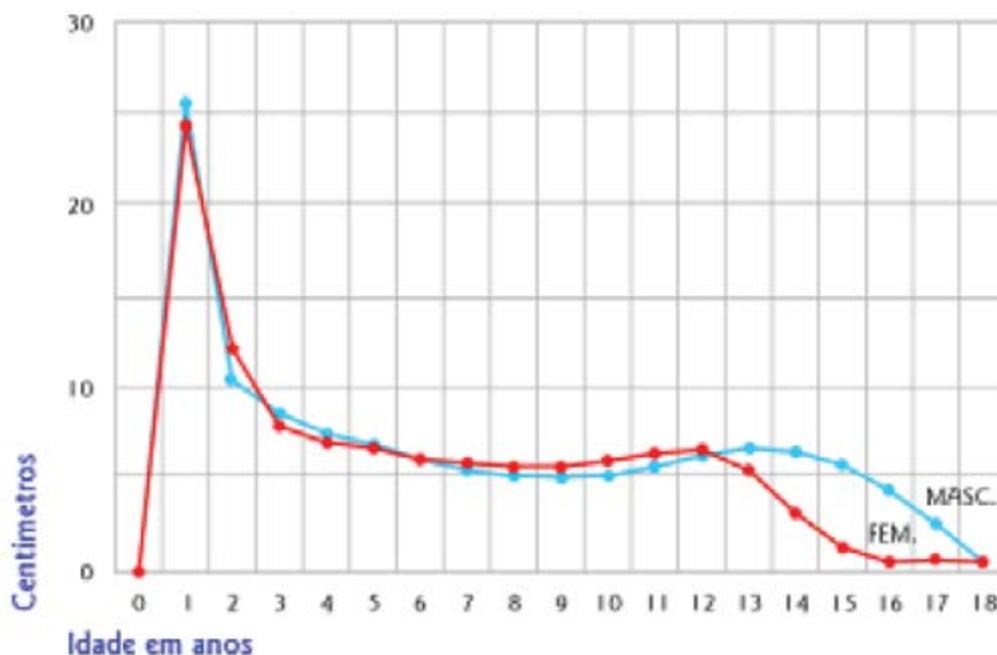


Figura 57: Velocidade média de crescimento pós-natal da população de referência do NCHS – meninos e meninas de 0 a 18 anos.  
Fonte: Gráfico 3, Brasil (2002, p. 21).



## Capítulo 6

# PESQUISA EXPERIMENTAL: O JOGO DA MEMÓRIA

O Jogo da Memória pode ser trabalhado mesmo com crianças pequenas, do 1º ou 2º ano. Investigar sobre a memória em uma situação experimental é bem interessante de ser realizada, pois podemos pensar em uma atividade em que a memória possa ser medida cronometrando o tempo que uma pessoa gasta para identificar um número X de pares, em um Jogo da Memória.

Nesse caso um questionário, como o apresentado na Figura 5, pode ser construído coletivamente com a classe. Para a coleta dos dados, é preciso ter vários jogos da memória iguais. Os jogos podem ser adquiridos com verba da escola, mas uma boa possibilidade é a confecção do jogo pelas próprias crianças, em sala de aula ou na aula de artes, o qual poderia ter, por exemplo, oito pares de figuras iguais<sup>1</sup>.

Uma decisão essencial, que cabe ao professor tomar, é sobre o registro dos dados. Como medir o tempo gasto no jogo? Que unidade de medida do tempo será utilizada: minutos ou segundos? Mas, isto nos traz mais um desafio: como tal medida será anotada na planilha ou banco de dados?

Se a opção for a de desconsiderar os segundos, as diferenças entre o tempo gasto pelas pessoas podem levar a uma grande distorção dos dados. Esse seria o caso de uma pessoa ter gasto 1 minuto e 4 segundos para identificar todos os pares do jogo, enquanto outra gastou 1 minuto e 58 segundos realizando a mesma tarefa. Note que a primeira pessoa gastou praticamente 1 minuto, enquanto o tempo da segunda foi quase 2 minutos.

---

1 O número de pares do jogo é decidido pelo professor. Esse é um fator importante para o sucesso da atividade, pois um número muito grande de pares pode tornar o jogo longo, conseqüentemente, cansativo e pode dificultar a coleta de dados. Não percamos de vista que o foco não é o jogo nem a investigação da memória em si, mas sim permitir que os estudantes tenham a vivência de coletar dados experimentalmente. Se o jogo for muito longo, isto pode desviar a atenção da criança para marcação do tempo, levando a obtenção de uma coleta imprecisa. Para crianças da Educação Infantil, talvez fosse mais interessante não ultrapassar os seis pares; já para crianças de 10 ou 11 anos, o número de pares pode chegar a 12.

E então, ficarão essas duas pessoas com o tempo gasto de 1 minuto?

Se a opção for por considerar o tempo em segundos, então será necessário fazer a conversão de minutos para segundos, de tal forma que possamos registrar apenas os segundos na planilha de dados. Nesse caso, há a necessidade de trabalhar com a classe a maneira de como fazer essa conversão da unidade minutos para segundos. Esse pode ser um bom momento para discutir as unidades utilizadas para a medição do tempo. Aqui se estabeleceria o trabalho interdisciplinar, pois o professor estaria trabalhando tanto dentro do bloco Tratamento da Informação, quanto no de Grandezas e Medidas.

Contudo, temos a clareza que tal atividade não será muito fácil de trabalhar com crianças muito pequenas. Para essas talvez seja o caso de registrar no banco de dados Quem Ganhou, o dado “A”, se foi o adulto, “C” se foi “criança” e “E” se foi empate.

O quadro que segue apresenta uma planilha de dados fictícia, a partir da qual tentaremos responder à questão de pesquisa levantada: quem tem mais memória: adultos ou crianças? Além disso, colocamos no questionário uma pergunta sobre a percepção sobre sua memória.

**Quadro 8:** Planilha de dados da pesquisa “Quem tem mais memória?”

Dados relativos à criança					Dados relativos ao pai ou responsável					Dife- rença	Quem ganhou?
Nome	Idade	Gênero	Percep- ção	Tempo	Nome	Idade	Gênero	Percep- ção	Tempo		
Ana	8	F	MB	38	Aldo	40	M	B	51	13	C
Artur	8	M	MB	38	Felipe	50	M	R	55	17	C
Bianca	8	F	MB	35	Rafaela	36	F	R	56	21	C
Bruna	9	F	R	45	Clara	32	F	R	53	8	C
Beto	9	M	B	36	Clara	43	F	R	48	12	C
Breno	8	M	B	38	Diego	28	M	MB	45	7	C
Carla	7	F	R	38	Amanda	30	F	B	41	3	C
Camila	8	F	MB	47	Suzane	54	F	R	61	14	C
Daniel	9	M	B	35	João	30	M	B	46	11	C
Denise	9	F	B	43	André	42	M	B	57	14	C
Deise	8	F	MB	33	Cláudio	28	M	MB	49	16	C
Emilio	8	M	MB	40	Karla	39	F	B	47	7	C
Fabio	7	M	B	44	Aline	36	F	B	47	3	C
Felipe	9	M	MB	37	Flora	41	F	R	56	19	C
Gilda	8	F	MB	45	Pedro	27	M	B	40	-5	A
Gabriela	8	F	B	41	Jorge	35	M	B	58	17	C
Irene	8	F	R	37	Artur	29	M	MB	45	8	C
José	9	M	B	38	Vitor	47	M	R	62	24	C
Juliana	9	F	MB	39	Tiago	43	M	R	56	17	C
Luiz	8	M	MB	47	Safira	40	F	B	45	-2	A
Luciana	7	F	B	36	Sara	39	F	B	57	21	C
Mariana	8	F	B	40	Hélio	41	M	R	57	17	C
Marcelo	8	M	R	44	Melissa	24	F	MB	43	-1	A
Milton	8	M	MB	41	Maria	50	F	R	59	18	C
Paulo	8	M	B	32	Inês	32	F	MB	50	18	C
Pâmela	9	F	R	43	Hugo	37	M	MB	45	2	C
Pedro	7	M	MB	45	Márcia	35	F	B	47	2	C
Rui	7	M	MB	39	Nilza	28	F	B	45	6	C
Renata	8	F	MB	46	Júlia	42	F	MB	52	6	C
Sandra	8	F	B	40	Max	36	M	B	45	5	C
Saulo	8	M	B	43	Lucas	34	M	R	43	0	E
Vanessa	8	F	R	36	Manoel	27	M	MB	39	3	C
Vera	9	F	B	45	Igor	28	M	MB	41	-4	A
Soma	268			1324		1203			1641		

Dependendo do ano escolar e dos conhecimentos prévios dos alunos, podemos trabalhar simplesmente quem ganhou e registrar apenas as variáveis: nome e tempo da criança e do pai (responsável) e a coluna quem ganhou. Observe que mesmo nessa tarefa, os alunos já têm que saber trabalhar com números na casa dois dígitos e saber comparar esses números.

Se coletarmos todos os dados da ficha, o que podemos fazer é distribuir tarefas,

isto é, dividir a classe em grupos e cada um deles trabalha com uma variável. A seguir apresentamos as tabelas e os passos intermediários que nos ajudam na construção das tabelas. Para calcular a média geral podemos somar o tempo diretamente da planilha de dados.

Idade	Nº de alunos	%	Gênero	Nº de alunos	%	Memória	Nº de alunos	%
7	5	15,2	Feminino	18	54,5	Regular	6	18,2
8	19	57,6	Masculino	15	45,5	Boa	13	39,4
9	9	27,3	Total	33	100,0	Muito Boa	14	42,4
Total	33	100,0				Total	33	100,0

Figura 59: Distribuição dos dados dos alunos.

Idade (*)	Nº de alunos	%	Gênero	Nº de alunos	%	Memória	Nº de alunos	%
Até 29	5	15,2	Feminino	16	48,5	Regular	11	33,3
De 30 a 39	19	57,6	Masculino	17	51,5	Boa	13	39,4
40 ou +	9	27,3	Total	33	100,0	Muito Boa	9	27,3
Total	33	100,0				Total	33	100,0

Figura 60: Distribuição dos dados dos pais (responsáveis).  
 (\*) Ver Quadro 8 para compreender o agrupamento dos dados

Para facilitar o trabalho podemos ordenar os dados por idade. O Quadro 9 apresenta os dados das crianças ordenados por idade e tempo gasto no jogo, assim como a média do tempo gasto por idade.

**Quadro 9:** Planilha para calcular o tempo médio gasto por idade das crianças

Nº	Nome	Idade	Tempo	Nº	Soma	Média
1	Fábio	7	44	5	229	45,8
2	Pedro	7	45			
3	Rui	7	46			
4	Carla	7	47			
5	Luciana	7	47			
6	Breno	8	36	19	772	40,6
7	Luiz	8	36			
8	Irene	8	37			
9	Ana	8	38			
10	Artur	8	38			
11	Camila	8	38			
12	Renata	8	39			
13	Emílio	8	40			
14	Marcelo	8	40			
15	Sandra	8	40			
16	Gilda	8	41			
17	Milton	8	41			
18	Deise	8	43			
19	Paulo	8	43			
20	Saulo	8	43			
21	Mariana	8	44			
22	Bianca	8	45			
23	Gabriela	8	45			
24	Vera	8	45			
25	Pâmela	9	32	9	323	35,9
26	Denise	9	33			
27	Beto	9	35			
28	Daniel	9	35			
29	Vanessa	9	36			
30	Felipe	9	37			
31	Bruna	9	38			
32	José	9	38			
33	Juliana	9	39			
	Soma	268	1324	33	1324	40,1
	Média	8,1	40,1			

Para apresentar a idade dos adultos podemos criar faixas etárias. Por exemplo, podemos pensar que os pais mais jovens, até 29 anos, são pessoas da “geração” internet, dos vídeo games, da tecnologia, que podem ter ajudado a desenvolver algumas estratégias de memória, portanto mais rápidos no jogo; já os pais com 40 anos ou mais sem esse acesso, mais resistentes a esse tipo de jogos, portanto menos rápidos. Observamos que essas faixas são arbitrárias. O Quadro 10 apresenta os dados dos adultos, ordenados por idade e tempo gasto no jogo, assim como a média do tempo gasto por idade.

**Quadro 10:** Planilha para calcular o tempo médio gasto por faixa etária dos adultos

Nº	Nome	Idade	Tempo	Nº de adultos	Soma	Média
1	Melissa	30	41	8	347	43,4
2	Pedro	30	46			
3	Manoel	32	53			
4	Diego	32	50			
5	Cláudio	34	43			
6	Nilza	35	58	13	635	54,9
7	Igor	35	47			
8	Artur	36	56			
9	Amanda	36	47			
10	João	36	45			
11	Clara	37	45			
12	Inês	39	47			
13	Lucas	39	57			
14	Jorge	40	51			
15	Márcia	40	45			
16	Rafaela	41	56			
17	Aline	41	57			
18	Max	42	57			
19	Hugo	42	52			
20	Karla	43	48			
21	Sara	43	56			
22	Aldo	47	62			
23	Safira	50	55			
24	Flora	50	59			
25	Hélio	54	61	12	659	54,9
26	André	1203	1641			
27	Julia	9	35			
28	Clara	9	35			
29	Tiago	9	36			
30	Vitor	9	37			
31	Felipe	9	38			
32	Maria	9	38			
33	Suzane	9	39			
	Soma	1203	1641	33	1641	49,7
	Média	36,5	49,7			

Também podemos querer saber se a auto percepção de memória está relacionada ao desempenho no jogo. Isto nos obriga a trabalhar os dados não por idade e sim pela auto percepção. O Quadro 11 apresenta os dados das crianças e o Quadro 12 dos adultos, ordenados por auto percepção de memória e tempo gasto no jogo, assim como a média do tempo gasto pelas categorias de auto percepção.

**Quadro 11:** Planilha para calcular o tempo médio gasto por auto percepção das crianças

Nº	Nome	Percepção	Tempo	Nº	Soma	Média
1	Vanessa	R	36	6	243	40,5
2	Irene	R	37			
3	Camila	R	38			
4	Paulo	R	43			
5	Mariana	R	44			
6	Bianca	R	45			
7	Pâmela	B	32	13	511	39,3
8	Daniel	B	35			
9	Breno	B	36			
10	Luiz	B	36			
11	Bruna	B	38			
12	José	B	38			
13	Marcelo	B	40			
14	Sandra	B	40			
15	Gilda	B	41			
16	Deise	B	43			
17	Saulo	B	43			
18	Fábio	B	44			
19	Vera	B	45			
20	Denise	MB	33	14	570	40,7
21	Beto	MB	35			
22	Felipe	MB	37			
23	Ana	MB	38			
24	Artur	MB	38			
25	Juliana	MB	39			
26	Renata	MB	39			
27	Emílio	MB	40			
28	Milton	MB	41			
29	Gabriela	MB	45			
30	Pedro	MB	45			
31	Rui	MB	46			
32	Carla	MB	47			
33	Luciana	MB	47			
	Soma		1324	33	1324	40,1
	Média		40,1			

**Quadro 12:** Planilha para calcular o tempo médio gasto por auto percepção dos adultos

Nº	Nome	Peercepção	Tempo	Nº de Adultos	Soma	Média
1	Lucas	R	43	11	606	55,1
2	Clara	R	48			
3	Clara	R	53			
4	Felipe	R	55			
5	Rafaela	R	56			
6	Flora	R	56	13	626	48,2
7	Tiago	R	56			
8	Hélio	R	57			
9	Maria	R	59			
10	Suzane	R	61			
11	Vitor	R	62			
12	Pedro	B	40			
13	Amanda	B	41			
14	Safira	B	45			
15	Nilza	B	45			
16	Max	B	45			
17	João	B	46			
18	Karla	B	47			
19	Aline	B	47			
20	Márcia	B	47			
21	Aldo	B	51			
22	André	B	57			
23	Sara	B	57			
24	Jorge	B	58			
25	Manoel	MB	39	9	409	45,4
26	Igor	MB	41			
27	Melissa	MB	43			
28	Diego	MB	45			
29	Artur	MB	45			
30	Hugo	MB	45			
31	Cláudio	MB	49			
32	Inês	MB	50			
33	Júlia	MB	52			
	Soma		1641	33	1641	49,7
	Média		49,7			

Esses quadros são criados como suporte para compreender de onde resultam as estatísticas, pois o que nos interessa é saber como a idade e a auto percepção de memória se relacionam com o desempenho no jogo.

Assim, a partir dos resultados dos quadros criamos tabelas e gráficos que nos ajudem a compreender a natureza das variáveis e suas relações.

Na Tabela 12 podemos ver que o tempo gasto pelas crianças melhora com a idade, já para os adultos essa tendência se inverte, os mais jovens tendem a ser mais rápidos

do que os mais velhos, conforme Tabela 13. Lembrar que quanto menor o tempo, maior a capacidade de memória.

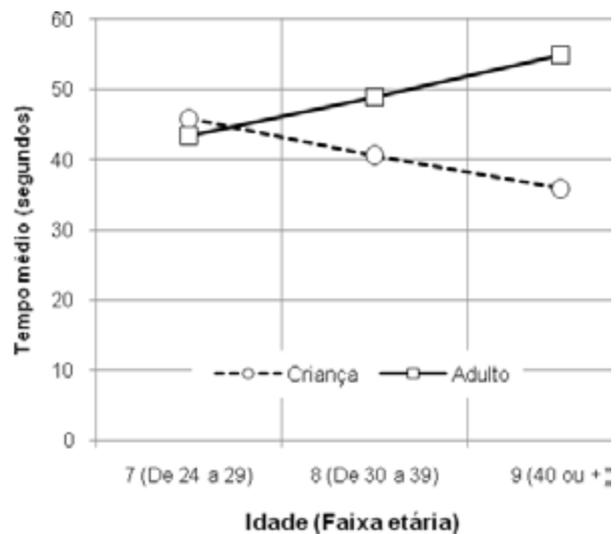
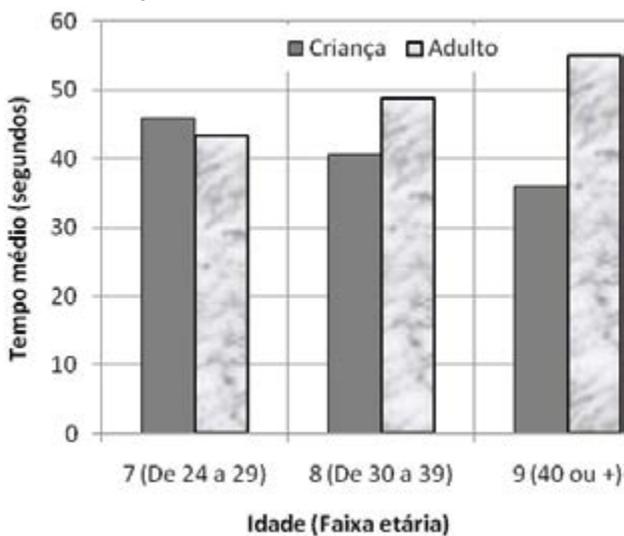
**Tabela 12:** Tempo médio gasto pelas crianças por idade

Idade (anos)	Nº de alunos	Soma	Média
7	5	229	45,8
8	19	772	40,6
9	9	323	35,9
Total	33	1324	40,1

**Tabela 13:** Tempo médio gasto pelos adultos por faixa etária

Faixa etária	Nº de adultos	Soma	Média
De 24 a 29	8	347	43,4
De 30 a 39	13	635	48,8
De 40 ou mais	12	659	54,9
Total	33	1641	49,7

Na Figura 61(a) apresentamos em um gráfico de barras o tempo médio gasto no jogo da memória, tanto pelas crianças, quanto pelos adultos e os mesmos dados são apresentados em um gráfico de linhas na Figura 61(b). Observamos como os gráficos são mais eficientes para transmitir os dados do que as tabelas, e que o gráfico de linhas é mais eficiente para transmitir a tendência dos dados.



(a) Gráfico de barras

(b) Gráfico de Linhas

Figura 61: Tempo médio gasto no jogo da memória das crianças e dos adultos segundo a idade.

A partir das tabelas de auto percepção constantes nas Tabela 14 e Tabela 15 construímos um gráfico de barras para comparar a distribuição da percepção das crianças e dos adultos, conforme Figura 62. Observamos que as crianças são muito mais autoconfiantes na sua memória do que os adultos. Contudo, ser autoconfiante não necessariamente significa ter o melhor desempenho, por essa razão precisamos analisar essa relação.

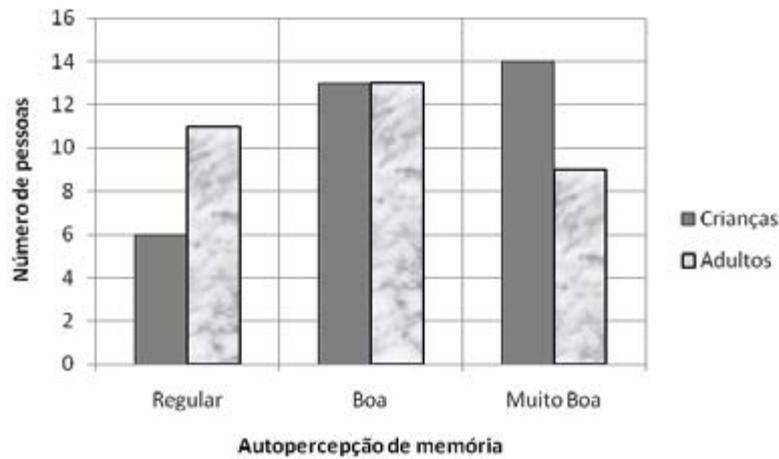


Figura 62: Tempo médio gasto no jogo da memória das crianças e dos adultos segundo a idade.

Em relação à auto percepção das crianças parece não haver relação com o tempo gasto, isto é, crianças que são autoconfiantes na sua memória tiveram desempenho similar aos que não confiam tanto na sua capacidade de memorizar (Tabela 14). Já no caso dos adultos, Tabela 15, parece que quanto mais autoconfiante melhor o desempenho, isto é, gasta menos tempo. A Figura 63 apresenta esses dados em gráficos de barras e linhas.

**Tabela 14:** Tempo médio gasto pelas crianças por autopercepção

Autopercepção	Nº de alunos	Soma	Média
Regular	6	243	40,5
Bom	13	511	39,3
Muito Bom	14	570	40,7
<b>Total</b>	<b>33</b>	<b>1324</b>	<b>40,1</b>

**Tabela 15:** Tempo médio gasto pelos adultos por autopercepção

Autopercepção	Nº de alunos	Soma	Média
Regular	11	606	55,1
Bom	13	626	48,2
Muito Bom	9	409	45,4
<b>Total</b>	<b>33</b>	<b>1641</b>	<b>49,7</b>

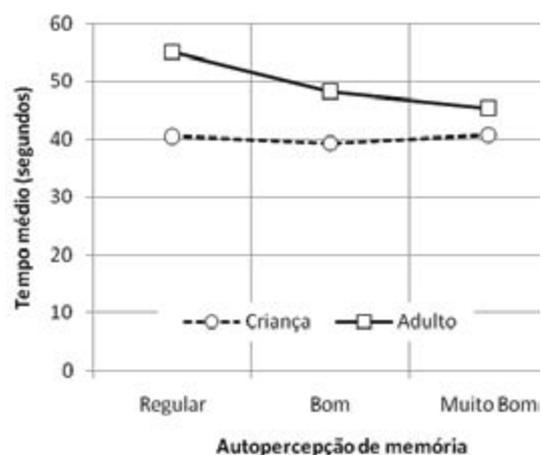
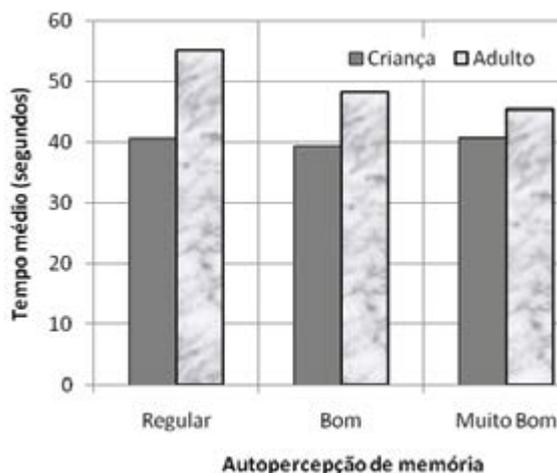


Figura 63: Tempo médio gasto no jogo da memória das crianças e dos adultos segundo a idade.

Professor, você poderá estar questionando a validade de termos construído um gráfico de linhas para uma variável qualitativa. Podemos fazer isso, pois a variável é ordinal, assim podemos analisar se há padrões de tendência segundo a intensidade da variável ordinal, neste caso, o grau de autopercepção da memória.

Mas, afinal, quem tem mais memória: os adultos ou as crianças? Para responder essa pergunta precisamos de um único dado que nos permita comparar de forma global o desempenho de ambos os grupos e para isso temos a Média e a Mediana.

Observamos que o melhor gráfico para analisar o desempenho de forma global é o dotplot (Figura 66) ou o histograma (Figura 67).

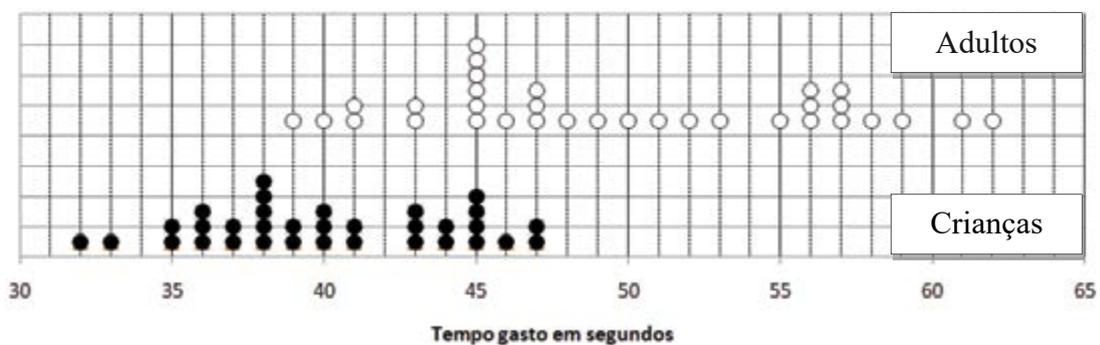


Figura 64: Gráfico de pontos do tempo gasto no jogo da memória das crianças e adultos.

Para construir o histograma, vamos ter que construir a TDF, para isso devemos decidir a amplitude das classes ou intervalos. Podemos escolher intervalos de 5 segundos, conforme mostramos na Tabela 16. A partir dessa tabela construímos o histograma, como mostramos na Figura 65.

**Tabela 16:** Distribuição do tempo gasto no jogo da memória por grupo

Tempo	Crianças		Adultos	
	Nº	%	Nº	%
De 30 a 34	2	6,1	0	0,0
De 35 a 39	14	42,4	1	3,0
De 40 a 44	10	30,3	5	15,2
De 45 a 49	7	21,2	12	36,4
De 50 a 54	0	0,0	4	12,1
De 55 a 59	0	0,0	9	27,3
De 60 a 64	0	0,0	2	6,1
Total	33	1641	49,7	

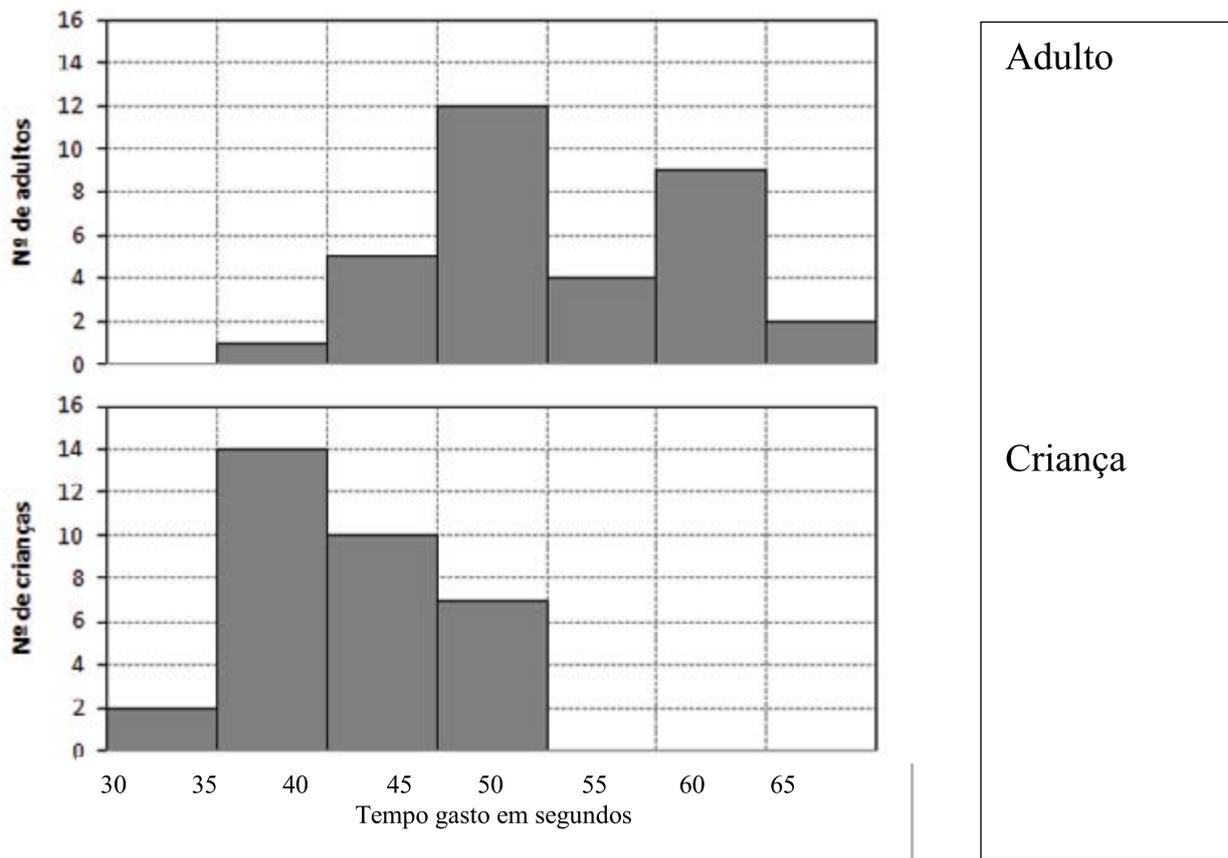


Figura 65: Distribuição do tempo gasto no jogo da memória por grupo.

Analisando a configuração dos dados, tanto no dotplot quanto no histograma podemos concluir que as crianças são mais rápidas e tem um desempenho mais homogêneo; já os adultos, além de apresentar um desempenho mais variável, tem uma quantidade deles, quase a metade, que demora mais do que qualquer criança.

Na Tabela 17 apresentamos as estatísticas do tempo gasto pelas crianças e pelos adultos. Com todas essas informações podemos concluir que as crianças são mais rápidas do que seus pais (ou responsáveis), para executar a tarefa, em quase 10 segundos. Além disso, as crianças têm um desempenho mais homogêneo, pois os adultos, além de demorar mais, alguns executam a tarefa rapidamente, enquanto que outros demoram mais.

**Tabela 17:** Estatísticas do tempo gasto pelas crianças e pelos adultos

Grupo	Mínimo	Máximo	Média	Mediana	Amplitude
Crianças	32	47	40,1	40	15
Adultos	39	62	49,7	48	23
Diferença	-5	24	9,6	8	29

Na Figura 66 apresentamos o dotplot da diferença do tempo gasto pelos adultos em relação ao tempo gasto pelas crianças. Observe que apenas quatro crianças gastaram mais tempo do que os seus responsáveis (valores negativos), uma criança empatou

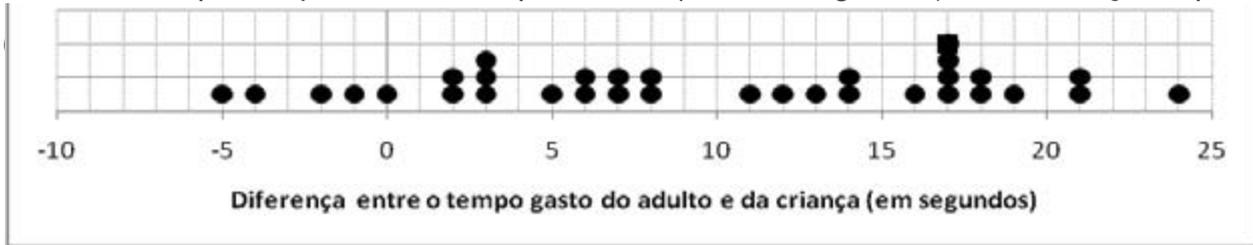


Figura 66: Gráfico de pontos da diferença do tempo gasto no jogo da memória do adulto em relação às crianças sob seus cuidados.

Por fim, apresentamos a variável “quem ganhou”, calculada a partir da comparação do tempo gasto pelo adulto e pela criança. Como dissemos, anteriormente, esta é a variável mais simples, pois mesmo crianças pequenas, que ainda não sabem o algoritmo da subtração com reserva, conseguem comparar e saber quem ganhou.

A Tabela 18 mostra claramente que foram as crianças que ganharam. É interessante notar que muitas vezes maiores pontuações são sinônimo de ganhar, mas neste caso, não é, pois ganha quem consegue encerrar o jogo no menor tempo possível.

**Tabela 18:** Distribuição de “Quem ganhou?”

Quem ganhou?	Nº	%
Adultos	4	12,1
Crianças	28	84,8
Empate	1	3,0
Total	33	100

Chegamos ao fim. Esperamos ter contribuído com uma reflexão acerca do desenvolvimento de conceitos elementares da Estatística.

Acreditamos que a Estatística não é apenas como uma ferramenta útil e necessária na formação do cidadão, mas também um instrumento indispensável na construção do pensamento científico.

Esperamos que este livro tenha auxiliado você na organização de situações didáticas que, em curto prazo, busquem contribuir para que os alunos compreendam conceitos

## REFERÊNCIAS

ANDRADE, A. M. dos S. Matemática e Cidadania. Monografia de conclusão de curso. Universidade Estadual de Santa Cruz - UESC, Ilhéus, Bahia, 2006.

BRASIL. Ministério da Educação. Secretaria de Educação Fundamental. Parâmetros Curriculares Nacionais: Matemática. Brasília: Ministério da Educação/Secretaria de Educação Fundamental, 1997.

BRASIL. Ministério da Saúde. Secretaria de Políticas de Saúde. Departamento de Atenção Básica. Saúde da criança: acompanhamento do crescimento e desenvolvimento infantil. Brasília: Ministério da Saúde, 2002. Disponível em [http://bvsms.saude.gov.br/bvs/publicacoes/crescimento\\_desenvolvimento.pdf](http://bvsms.saude.gov.br/bvs/publicacoes/crescimento_desenvolvimento.pdf).

BUEHRING, R. S. Uma análise de dados no início da escolaridade: uma realização de ensino por meio dos registros de representação semiótica. Dissertação de mestrado da UFSC, 2006. disponível em [www.ppgect.ufsc.br/dis/27/Dissert.pdf](http://www.ppgect.ufsc.br/dis/27/Dissert.pdf).

CABRERA, P. C.; SOSA, A. B. Matemática con sentido: una propuesta que replantea el modo de enseñar la matemática en nivel inicial y EGB 1. Córdoba: Comunic-Arte, 2006.

CAVALCANTI, M.; NATRIELLI, R. e GUIMARÃES, G. Gráficos na mídia impressa. Bolema. Boletim de Educação Matemática, vol. 23 - UNESP. Rio Claro, 2010.

CAZORLA, I. M.; SANTANA, E. R. dos S. (Org.). Do Tratamento da Informação ao Letramento Estatístico. Itabuna-BA: Via Litterarum, 2010.

EVANGELISTA, B.; GUIMARÃES, G. Escalas representadas em gráficos: um estudo de intervenção com alunos do 5º ano. Revista Portuguesa de Educação, v.28, p.117 - 138, 2015.

GUIMARÃES, G. L. Interpretando e construindo gráficos de barras. Tese de doutorado

na Pós Graduação em Psicologia Cognitiva da UFPE. Recife, 2002. Disponível em [www.ufpe.br](http://www.ufpe.br).

LUZ, P.; GUIMARÃES, G. L.. Classificações nos anos iniciais do ensino fundamental: o papel das representações. In: Anais do X ENEM. SBEM, Salvador., 2010.

SELVA, A. C. V.; BORBA, R. E. de S. R. O uso de diferentes representações na resolução de problemas de divisão inexata: analisando a contribuição da calculadora. In Anais da 28 Reunião Anual da ANPED, 2005. Disponível: [www.anped.org.br/reunioes/28/textos/gt19/gt19635int.rtf](http://www.anped.org.br/reunioes/28/textos/gt19/gt19635int.rtf)

STRAUSS, S.; BICHLER, E. The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19(1), 64-80, 1988.

VERGNAUD, G. Epistemology and Psychology of Mathematics Education. In Pearlman Neshet e Jeremy Kilpatrick (Eds.), *Mathematics and Cognition*. Cambridge, Cambridge University Press, 1994.

WATSON, J. What's the point? *The Australian Mathematics Teacher*. 52(2): 40-43, 1996.

## Sugestões de Leitura

BATANERO, C. Dificultades de los estudiantes en los conceptos estadísticos elementales: el caso de las medidas de posición central. In Loureiro, C.; Oliveira F.; Brunheira, L. (Eds.) *Ensino e Aprendizagem da Estatística*. Sociedade Portuguesa de Estatística, 2000.

BATANERO, C.; GODINO, J.; GREEN, D. ; HOLMES P.; VALLECILLOS, A. Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547, 1994.

BORBA, R.; MONTEIRO, C. ; GUIMARÃES, G.; COUTINHO, C.; KATAOKA, V. Educação Estatística no Ensino Básico: Currículo, pesquisa e prática em sala de aula. *EM TEIA: Revista de Educação Matemática e Tecnológica Iberoamericana*, v.2, 1 - 18, 2011.

CARVALHO, C. Olhares sobre a Educação Estatística em Portugal. Anais do SIPEMAT. Recife, Programa de Pós-Graduação em Educação-Centro de Educação – Universidade Federal de Pernambuco, 2006.

CAZORLA, I. M. Média aritmética: um conceito prosaico e complexo. Anais do IX Seminário

rio de Estatística Aplicada, Rio de Janeiro, 2003.

EVANGELISTA, B. ; GUIMARÃES, G. Como alunos do 5º ano compreendem o conceito de escala em graficos? EM TEIA: Revista de Educação Matemática e Tecnológica Iberoamericana, v.6, 1-17, 2015.

GAL, I. Adult's Statistical Literacy: meanings, components, responsibilities. International Statistical Review. 70(1), 1-25, 2002.

GUIMARÃES, G. Formação de professores: trabalhando com gráficos e tabelas na Educação Infantil. Revista Iberoamericana de Educación Matemática - UNION, n.21, 87-101. Disponível em: <http://www.fisem.org> 2010.

GUIMARÃES, G.; GITIRANA, V.; MARQUES, M. C.; CAVALCANTI, M. A Educação estatística na educação infantil e nos anos iniciais. Zetetike (UNICAMP), v.17, 11-28, 2009.

GUIMARÃES, G. Refletindo sobre a Educação Estatística na sala de aula. In GUIMARÃES, G. e BORBA, R. (Org.). Reflexões sobre o ensino de Matemática nos anos iniciais de escolarização. Sociedade Brasileira de Educação Matemática (SBEM) – Grupo de Trabalho 1 (GT1), 2009.

GUIMARÃES, G.; GITIRANA, V. Classificações: o que sabem os alunos de 3ª série do ensino fundamental? Anais do XVI Encontro de Pesquisa Educacional Norte-Nordeste - EPENN. Aracaju, 2003.

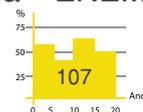
GUIMARÃES, G.; GITIRANA, V.; CAVALCANTI, M.; MARQUES, M. Análise das atividades sobre representações gráficas nos livros didáticos de matemática. Anais do 2º SIPEMAT, Recife, 2008.

GUIMARÃES, G.; CAVALCANTI, M.; EVANGELISTA, B. Compreensão de Escala nos Anos Iniciais. In: Pesquisa e atividades para o aprendizado matemático na educação infantil e nos anos iniciais do Ensino Fundamental. Ed. Brasília: SBEM, 47-69, 2015.

GUIMARÃES, G.; GITIRANA, V.; OLIVEIRA, I. ; PESSOA, C.; CARVALHO, I.; LIMA, C.; ROCHA, C. Educação Estatística. Brasília: Secretaria de Educação Básica, Diretoria de Apoio à Gestão Educacional, v.1. p.80, 2014.

LAJOIE, S.; JACOBS, V.; LAVIGNE, N. Empowering Children in the Use of Statistics. Journal of Mathematical Behavior, 14, 401-425, 1993.

LIMA, I.; SELVA, A. Investigando o desempenho de jovens e adultos na construção e interpretação de gráficos. Anais do X Encontro Nacional de Educação Matemática – ENEM.



Salvador, 2010.

LIMA, I.; SELVA, A. Investigando o desempenho de jovens e adultos na construção e interpretação de gráficos. Anais do X Encontro Nacional de Educação Matemática – ENEM. Salvador, 2010.

LIMA, R.; MAGINA, S. O uso de diferentes escalas na leitura de gráficos por crianças das séries iniciais do ensino fundamental. Anais do X Encontro Nacional de Educação Matemática – ENEM. Salvador, 2010.

LINS, W. Procedimentos lógicos de classificação através de um banco de dados: um estudo de caso. Anais do II Simpósio Internacional de Pesquisas em Educação Matemática – SIPEM. Santos, 2003.

LOPES, C.; CARVALHO, C. Literacia Estatística na educação básica. In: A. Nacarato e C. Lopes (Eds), *Escritas e Leituras na educação matemática* (pp 77- 92). Belo Horizonte: Autêntica, 2005.

MAGINA, S. M. P.; CAZORLA, I. M.; GITIRANA, V.; GUIMARÃES, G. L. Concepções e concepções alternativas de média: um estudo comparativo entre professores e alunos do Ensino Fundamental. *Educar em Revista (Impresso)*, v.1, 2011.

MARQUES, M.; GUIMARÃES, G.; GITIRANA, V. Compreensões de Alunos e Professores sobre média aritmética. *Bolema. Boletim de Educação Matemática*, v.24, p.725 - 746, 2011.

RUMSEY, D. J. Statistical Literacy as a Goal for Introductory Statistic Courses, *Journal of Statistics Education*. V. 10, n. 3, 2002.

SELVA, A C. V. Gráficos de barras e materiais manipulativos: analisando dificuldades e contribuições de diferentes representações no desenvolvimento da conceitualização matemática em crianças de 6 a 8 anos. Tese de doutorado na Pós Graduação em Psicologia Cognitiva da UFPE. Recife, 2003. Disponível em [www.ufpe.br](http://www.ufpe.br).

WATSON, J.; KELLY, B. Can grade 3 students learn about variation? In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*, Cape Town, South Africa [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute, 2002.

WATSON, J.M.; CALLINGHAM, R. A. Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3-46, 2003.

## ANEXO A: A VARIÁVEL IDADE

Em geral, a idade é trabalhada em anos completos o que simplifica bastante o trabalho.

Mas podemos querer maior precisão. Nesse caso, se os alunos já conhecem os números decimais e o arredondamento, podemos utilizar o critério até 5 meses considerar apenas o ano, 6 meses ou mais adicionar um ano. Por exemplo, 9 anos e 4 meses será registrado com 9 anos e 9 anos e 6 meses como 10 anos.

Um outro ponto importante é que, quando trabalhamos com a idade, precisamos que nossos alunos compreendam a equivalência entre meses e a base decimal, como podemos ver na Figura 67.

### Variável idade

Meses	0	1	2	3	4	5	6	7	8	9	10	11	12
Base 10	0,00	0,08	0,17	0,25	0,33	0,42	0,50	0,58	0,67	0,75	0,83	0,92	1,0

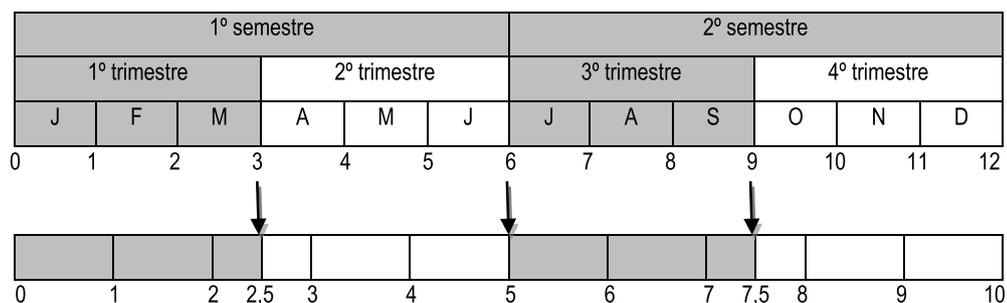


Figura 67: Equivalência entre os sistemas base 12 e base 10 para a coleta de dados da variável tempo.

A variável idade pode ser trabalhada não como os anos vividos por uma pessoa,

mas pela percepção que esta tem em relação a algum empreendimento na sua vida. Por exemplo, podemos investigar como adultos analfabetos se sentem em relação à sua idade para aprender a ler e escrever, ou seja, serem alfabetizados.

Com relação à sua idade, como o(a) senhor(a) se sente diante a possibilidade de aprender a ler e escrever:

- ( ) Muito jovem  
 ( ) Jovem  
 ( ) Velho  
 ( ) Muito velho

Neste caso, podemos até coletar o dado real da idade e podemos estudar se a idade cronológica é determinante na percepção de idade para ser alfabetizado. Estamos trabalhando com uma variável conceitual, que não há como medi-la, a não ser pelo depoimento do respondente, diferente da idade cronológica que é uma variável empírica, pois podemos “observá-la”. Maiores detalhes podem ser encontrados em Cazorla e Oliveira (2010).

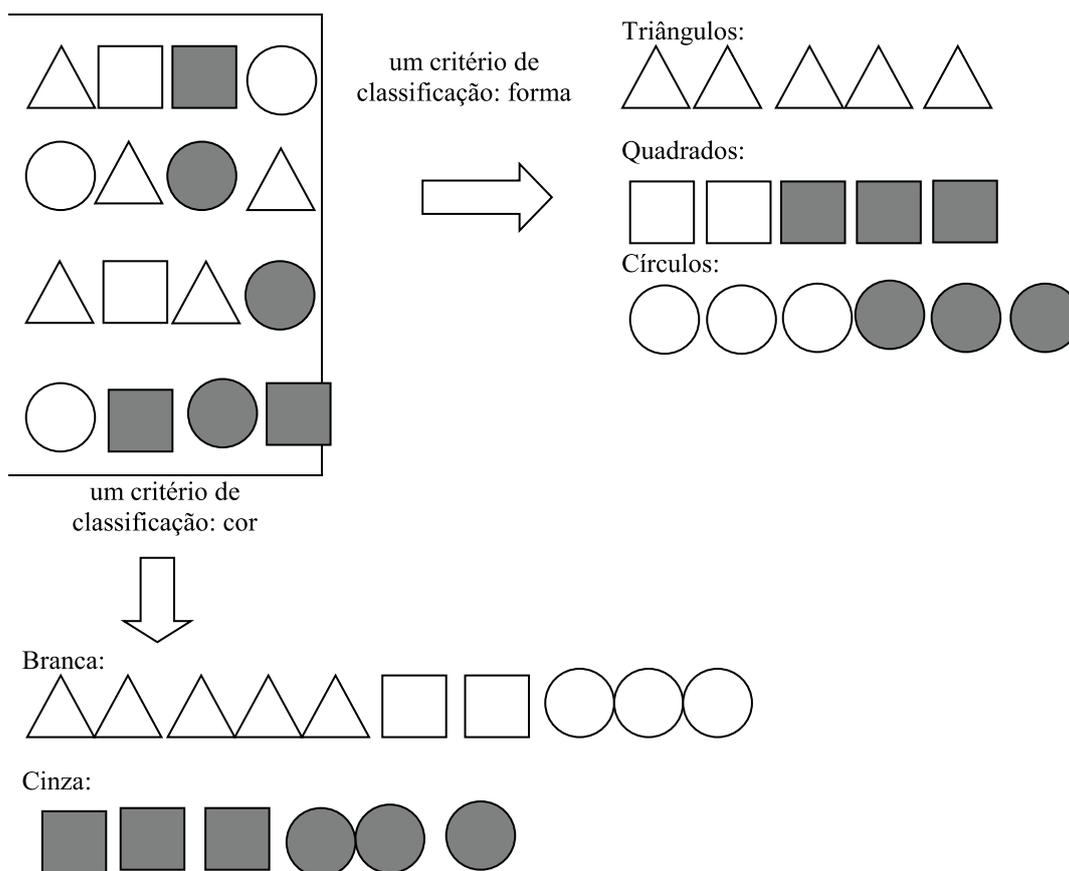
**Quadro 1:** Planilha completa de dados da pesquisa sobre a altura dos alunos

Dados originais					Idade (anos)	
Nome	Turma	Gênero	Idade	Altura	Completo	Base
Alberto	1º ano A	M	5 anos e 7 meses	1,10	5	5,58
João	1º ano A	M	6 anos	1,00	6	6,00
Tereza	1º ano A	F	5 anos e 10 meses	1,15	5	5,83
Pedro	1º ano B	M	6 anos e 1 mês	1,04	6	6,08
Ana	1º ano B	F	5 anos e 2 meses	1,20	5	5,17
Telma	1º ano B	F	6 anos e 5 meses	1,08	6	6,42
Valter	2º ano A	M	6 anos	1,01	6	6,00
Marcos	2º ano A	M	7 anos e 3 meses	1,10	7	7,25
Telma	2º ano A	F	8 anos	1,20	8	8,00
Geisa	2º ano B	F	6 anos e 11 meses	1,08	6	6,92
Gertrudes	2º ano B	F	7 anos e 5 meses	1,35	7	7,42
Maurício	2º ano B	M	8 anos e 6 meses	1,30	8	8,50
Manoel	3º ano A	M	8 anos	1,10	8	8,00
José	3º ano A	M	8 anos e 10 meses	1,20	8	8,83
Maria	3º ano A	F	8 anos e 5 meses	1,30	8	8,42
Marta	3º ano B	F	7 anos e 5 meses	1,10	7	7,42
Maria de Fátima	3º ano B	F	7 anos e 10 meses	1,20	7	7,83
Miguel	3º ano B	M	8 anos e 2 meses	1,13	8	8,17
Aída	4º ano A	F	8 anos e 11 meses	1,36	8	8,92
Michelle	4º ano A	F	9 anos e 10 meses	1,40	9	9,83
Severina	4º ano A	F	10 anos	1,45	10	10,00
Severo	4º ano B	M	10 anos e 2 meses	1,35	10	10,17
Michael	4º ano B	M	9 anos e 2 meses	1,35	9	9,17
Arquimedes	4º ano B	M	9 anos e 4 meses	1,43	9	9,33
João	5º ano A	M	10 anos	1,15	10	10,00
Josué	5º ano A	M	10 anos e 3 meses	1,34	10	10,25
Maria	5º ano A	F	10 anos e 7 meses	1,46	10	10,58
Miriam	5º ano B	F	11 anos	1,35	11	11,00
Tereza	5º ano B	F	10 anos e 9 meses	1,48	10	10,75
Norberto	5º ano B	M	11 anos e 3 meses	1,52	11	11,25
<b>Soma</b>				<b>758</b>		

Fonte: dados fictícios

## ANEXO B: A CLASSIFICAÇÃO COM OBJETOS PADRONIZADOS

A classificação é um tópico que deve ser trabalhado nos anos iniciais. Analisando os livros didáticos observamos que os exemplos utilizam formas geométricas, cores, tamanhos, grossura etc. Por exemplo, suponha que temos 16 objetos, de três formas e duas cores diferentes. Vejamos como fica quando classificamos por um único critério (Figura 68) ou forma ou cor:



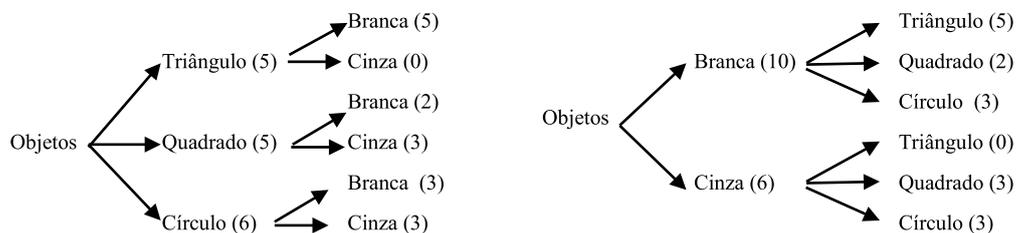
**Figura 1:** Exemplo de classificação de acordo com um único critério

Agora vejamos como fica a classificação quando usamos dois critérios de classificação: cor e forma simultaneamente:

Forma	Cor	
	Branca (10)	Cinza (6)
Triângulos (5)		
Quadrados (5)		
Círculos (6)		

**Figura 1:** Representação da classificação com dois critérios: a tabela de dupla entrada

Outra representação que podemos utilizar é o diagrama da árvore ou a árvore de possibilidades, conforme a seguinte figura:



**Figura 1:** Representação da classificação com dois critérios: o diagrama da árvore.

A criança deve perceber que a despeito de não ter nenhum triângulo de cor cinza, essa combinação existe apenas neste conjunto, não houve nenhum objeto dessa forma com essa cor. Esse tipo de classificação pode ser facilmente utilizado por crianças a partir dos cinco ou seis anos em tarefas de classificação.

É importante que tenhamos clareza que o trabalho com representações de dados implica também o conhecimento da simbologia específica desse tipo de representação. Pode ser que crianças muito pequenas não entendam o diagrama da árvore, mas certamente estão acostumadas a trabalhar com tabelas de dupla entrada.

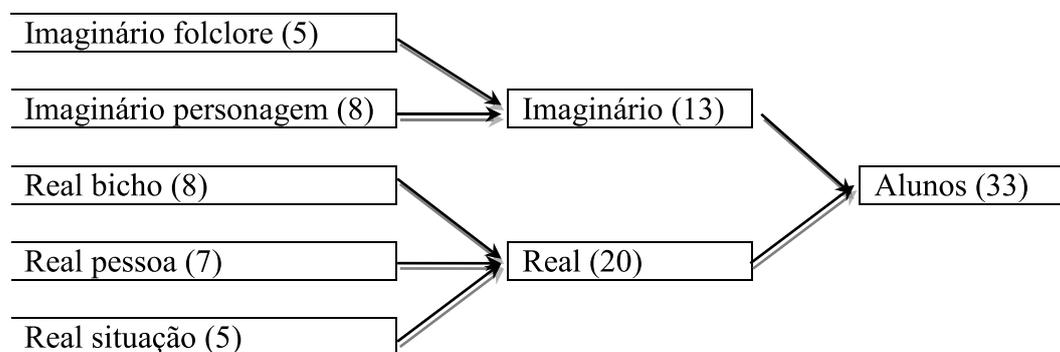
## ANEXO C: CONSTRUÇÃO DA TDF SOBRE O MEDO

Para compreender de onde resultaram os dados da Tabela 19 apresentamos a distribuição do medo segundo as subclasses, junto com as respostas dos alunos, a fim de que o aluno compreenda de onde resultaram os números 5, 8, e assim por diante.

**Tabela 1:** Distribuição do medo das crianças

Subclasse de medo	Nº de alunos	%
Imaginário folclore (Bicho papão=2, Mula sem cabeça=2, Saci=1)	5	15,2
Imaginário personagem (Bruxa=1, Coringa=1, Dinossauro=2, Tiranossauro=1, Fantasma=2, Feiticeira=1)	8	24,2
Real bicho (cobra=1, leão=1, rato=3, barata=1, cachorro=1, tubarão=1)	8	24,2
Real pessoa (bandido=2, homem mascarado=1, ladrão=2, marginal=1, palhaço=1)	7	21,2
Real situação (altura=3, escuro=2)	5	15,2
<b>Total</b>	<b>33</b>	<b>100,0</b>

Para apresentação final desta tabela deixamos apenas os nomes das subclasses como mostra a Tabela 3 e, a partir dela podemos construir a Tabela 4, pois é só agrupar as subclasses em classes.



## ANEXO D: CONSTRUÇÃO DO HISTOGRAMA

Para construir a TDF e o histograma, primeiro devemos decidir como vamos particionar (dividir) a amplitude total (valor máximo – valor mínimo) em classes de igual comprimento. No nosso caso, podemos optar por classes de dez centímetros de comprimento.

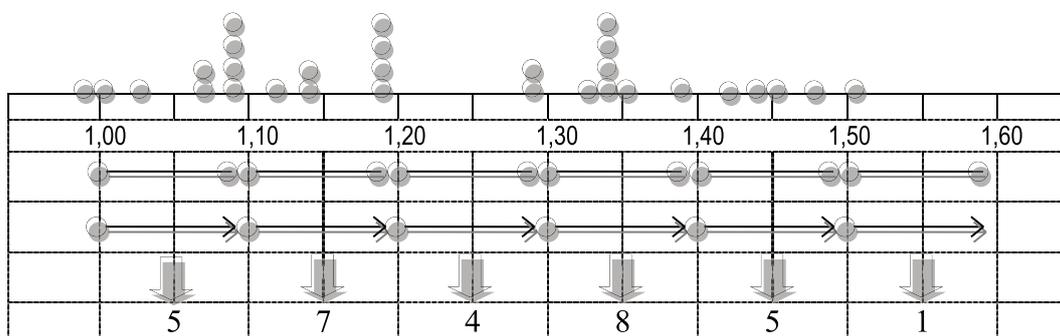
Neste ponto devemos compreender que, como os dados foram coletados em metros e centímetros, não existem dados entre 1,09 e 1,10; entre 1,19 e 1,20 etc., logo os intervalos de 1,00 a 1,09; de 1,10 a 1,19 etc. são chamados de intervalos fechados, pois os seus limites são contados no intervalo e a sua união recompõe todos os valores da variável.

Neste exemplo, o intervalo fechado 1,00 a 1,09 terá o mesmo efeito do semi-aberto, fechado à esquerda e aberto à direita (1,00 | 1,10) e assim poderemos colar os retângulos.

Intervalo fechado	Intervalo semi-aberto fechado a esquerda, aberto à direita	
Altura (m)	Altura (m)	Nº de alunos
De 1,00 até 1,09	1,00  — 1,10	5
De 1,10 até 1,19	1,10  — 1,20	7
De 1,20 até 1,29	1,20  — 1,30	4
De 1,30 até 1,39	1,30  — 1,40	8
De 1,40 até 1,49	1,40  — 1,50	5
De 1,50 até 1,59	1,50  — 1,60	1
	1,10 é contado no segundo intervalo 1,20 é contado no terceiro intervalo	

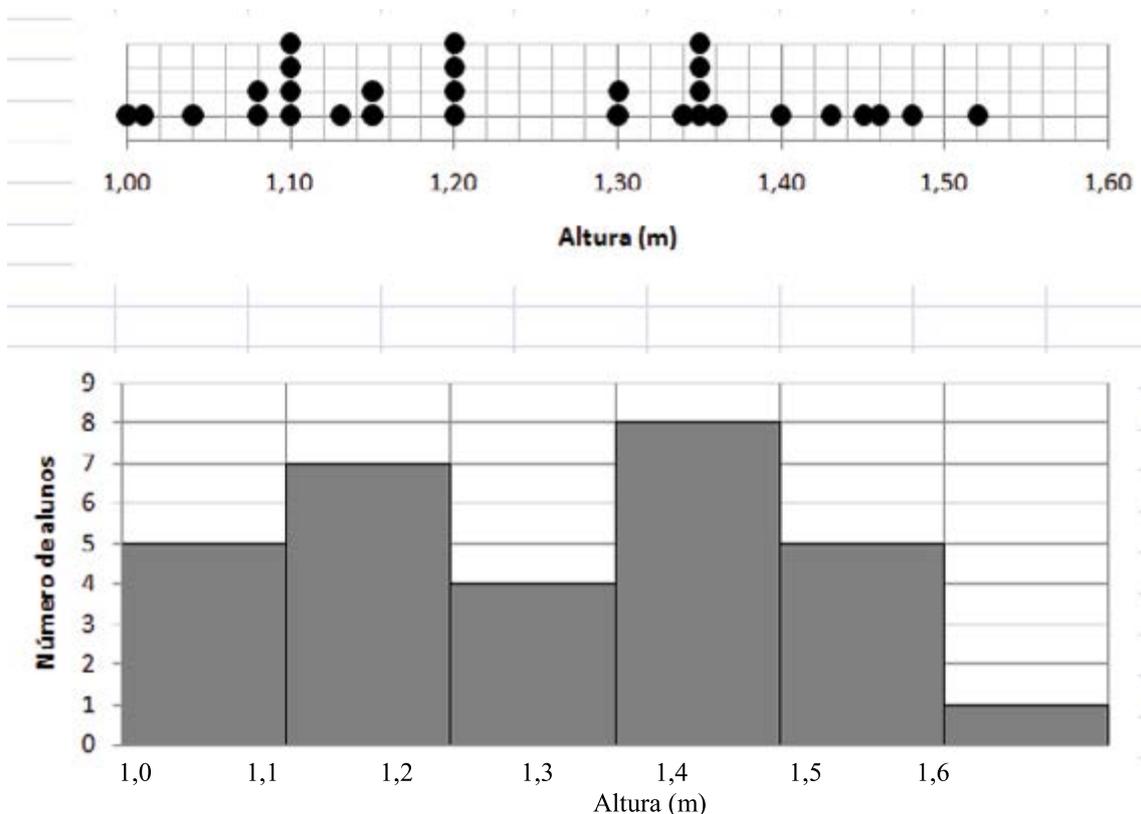
Acreditamos que esta sofisticação não será necessária neste nível escolar, o importante é que os alunos compreendam, que colamos os retângulos (barras) porque, em tese, a variável pode tomar qualquer valor entre o valor mínimo e máximo.

1,00	1,01	1,04	1,08	1,08	1,10	1,10	1,10	1,10	1,13	1,15	1,15	1,20	1,20	1,20
1,20	1,30	1,30	1,34	1,35	1,35	1,35	1,35	1,36	1,40	1,43	1,45	1,46	1,48	1,52



**Figura 1:** Exemplo da construção do diagrama de pontos (dotplot).

O histograma é construído com retângulos (barras/colunas) contíguos, colados uns aos outros. A base dos retângulos é formada pelas classes e a altura pela frequência absoluta (relativa) respectiva. Só haverá espaços caso a frequência absoluta da classe for zero. A Figura a seguir mostra o histograma correspondente.



**Figura 1:** Exemplo de construção do histograma da altura dos alunos.

## ANEXO E: A MÉDIA ARITMÉTICA, APROFUNDAMENTOS

Além das dificuldades já mencionadas no item sobre média aritmética. Outras dificuldades são apontadas na literatura.

Se ao invés de nota, os números tomados na sessão de média se referissem à idade de 4 crianças, então no numerador teríamos a soma das idades das quatro crianças e no denominador o número de crianças:

$$\text{Idade Média} = \frac{\text{soma das idades}}{\text{n.de crianças}} = \frac{6+6+7+7}{4} = \frac{26}{4} = 6,5 \text{ anos por criança}$$

Logo, a média das idades das quatro crianças seria seis anos e meio. Tanto neste problema como no de notas, não há problema do resultado ser um número decimal, pois esses valores existem por se tratar de variáveis contínuas (podem tomar qualquer valor).

Vejamos outro exemplo, agora, envolvendo dinheiro. Suponhamos que temos quatro alunos, com a seguinte distribuição de dinheiro no bolso: Ana tem 10 reais, Maria 7, João 3 e Pedro 4 reais. Solicitamos aos quatro alunos que coloquem o dinheiro em cima da mesa e juntamos tudo, a soma dará 24 reais (10 + 7 + 3 + 4). A seguir distribuimos esse montante em quatro partes iguais (24/4 = 6) resultando seis reais. Esse valor representa o valor médio de dinheiro por aluno.

Agora vamos devolver o dinheiro aos alunos, mas ao invés de devolver o valor original, vamos devolver o valor da média. Nesse caso, Ana perderá 4 reais e Maria 1 real, juntos perderão 5 reais. Em compensação, João ganhará 3 reais e Pedro dois reais, juntos ganharão cinco reais. Observe que o ganho recompensa a perda, zerando a diferença. Esta é uma propriedade da média. A planilha a seguir ilustra a distribuição

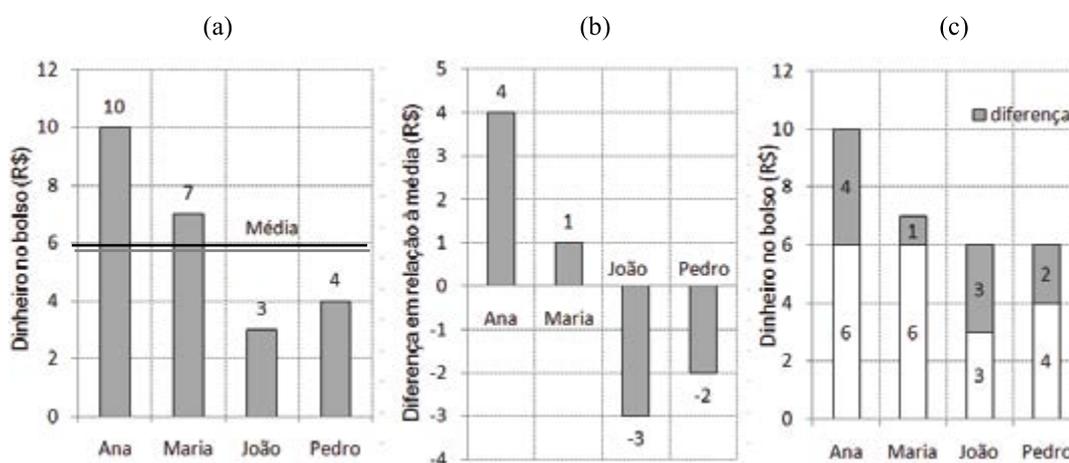
Aluno	Dinheiro que tem no bolso (R\$)	Devolução pela média	Diferença	Diferença
Ana	10	6	Perde 4 reais	- 4
Maria	7	6	Perde 1 real	-1
João	3	6	Ganha 3 reais	+3
Pedro	4	6	Ganha 2 reais	+2
Total	24	24	Zero	0

**Figura 1:** Exemplo de perdas e ganhos devido à devolução do dinheiro pela média.

Observamos que neste nível de ensino ainda não se trabalha com os números

negativos, mas as crianças compreendem o conceito como “estar devendo”. Logo, sua inserção no 5º ano já é possível. Veja como os gráficos nos ajudam a compreender o significado das diferenças dos valores da variável em relação à média.

Na Figura 74, o item (a), o gráfico de barras com a uma linha horizontal na média (6), nos indica em quanto os valores excederam a média ou quanto falta para atingi-la. O item (b) apresentamos os mesmos dados, só que com a origem no valor zero, isto é apenas as diferenças dos valores em relação à média. Já o item (c) sombreamos a porção que excede e a porção que falta para chegar à média. Assim, os alunos saberão que valores acima da média indicam excesso em relação à média, já valores abaixo da média indicam falta.



**Figura 1:** Diferença dos valores em relação à média.

Para melhor compreender a média, analisemos algumas situações. Suponhamos que nossos quatro alunos: Alex, Ana, João e Taís têm ao todo quatro balas. Mantendo fixo o todo (número de balas) e o número de elementos (número de crianças), analisemos algumas situações:

a) Cada criança tem uma bala, logo, em média, uma criança tem uma bala:

a)

Alex	Ana	João	Taís	Soma	Média
				4 balas	1 bala por criança
				4 crianças	

b) Alex e João têm duas balas cada um e as meninas não têm balas. Neste caso, a média continua a ser uma bala por criança:

a)

Alex	Ana	João	Taís	Soma	Média
				4 balas	1 bala por criança
				4 crianças	

c) Alex tem quatro balas e as outras três crianças não têm balas. Neste caso a média continua a ser uma bala por criança:

a)

Alex	Ana	João	Taís	Soma	Média
				4 balas	1 bala por criança
				4 crianças	

Discuta com os alunos esta situação. Pergunte a eles o que achariam se essa situação fosse verdadeira. Se juntássemos todas as balas das quatro crianças e as redistribuíssemos entre elas pela média.

No primeiro caso todas as crianças concordarão que está tudo bem, pois cada uma tinha uma bala e receberá uma bala. No segundo caso, algumas crianças acharão a distribuição pela média mais justa, pois as meninas que não tinham balas ganharão uma, sendo que todos terão balas de forma igualitária. Mas outras crianças poderão achar essa distribuição injusta, pois os meninos que tinham duas balas vão perder uma. Essa situação ficará mais tensa no terceiro caso, pois apenas o Alex tem 4 balas e as outras crianças não tem balas, logo Alex seria fortemente prejudicado.

			
			
Perde 3 - 3	Ganha 1 +1	Ganha 1 +1	Ganha 1 +1

Observe que a composição do todo é a mesma, o que varia é a distribuição dos dados entre os elementos que o compõem. Esta é uma característica importante da média. Quanto mais homogênea for a distribuição dos dados (caso a), a média representará melhor esse conjunto de dados, porém quanto mais dispersa for a distribuição, a média não será uma medida adequada para representar os dados (caso c). Compare agora a moda com a média. Enquanto a média é uma bala por criança, a moda é nenhuma bala.

Outra característica da média é que o que os desvios (diferença entre o valor da variável e a média) se anulam. Alex perdeu 3 balas, mas Ana, João e Taís ganharam uma bala cada, isto é a perda de um foi compensada pelo ganho dos outros.

Analisemos outra situação, mantendo fixa a soma de quatro (4) mascotes, variando agora o número de elementos (crianças):

a) Quatro crianças, cada uma tem uma mascote, logo o número médio de mascotes por criança é um.

a)

Alex	Ana	João	Taís	Soma	Média
				4 mascotes	1 mascote por criança
				4 crianças	

b) Alex tem três mascotes e Ana tem uma. Neste caso, a média é duas mascotes por criança:

a)

Alex	Ana	Soma	Média
  		4 mascotes	2 mascotes por criança
		2 crianças	

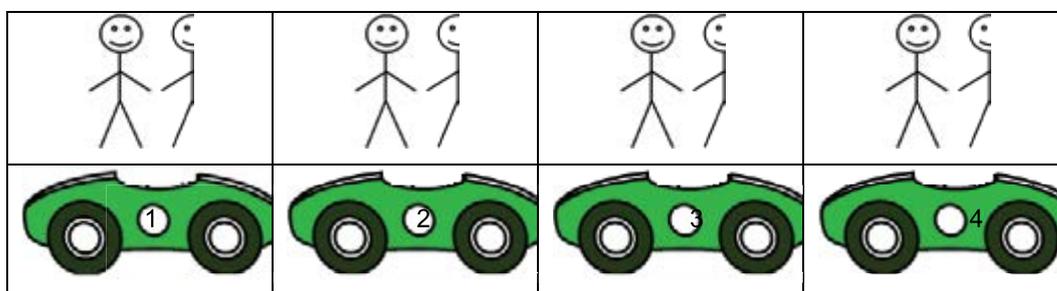
c) Alex tem três mascotes, Ana tem um e João não tem mascote. Neste caso, o número médio de mascotes por criança é 1,33 mascotes por criança:

a)

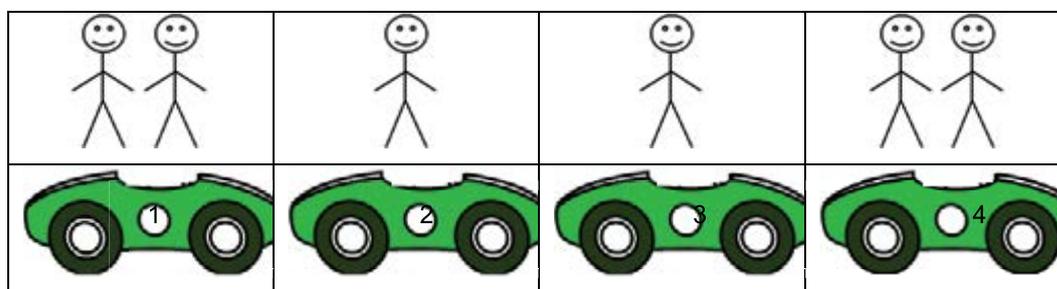
Alex	Ana	João	Soma	Média
  			4 mascotes	1,33 mascotes por criança
			3 crianças	

Veja que terrível seria pensar em distribuir as quatro mascotes pelas 3 crianças. Isso implicaria em cortar uma mascote em três pedaços iguais. Isto é um dos obstáculos que enfrentamos quando vamos ensinar a média para crianças que ainda não compreendem que a média é um número que representa um conjunto de dados.

A pesquisadora Watson (1996) também relata que após observar o número de pessoas em carros de passeio que passaram por uma ponte, num período determinado, solicitou às crianças que calculassem a média e que a representassem com desenhos (a média foi 1,5 pessoas por carro). A pesquisadora observou que crianças pequenas desenharam a ponte com carros e em cima dos carros um boneco e metade de outro (Figura 47), apenas as crianças maiores conseguiram desenhar de forma adequada (Figura 48).



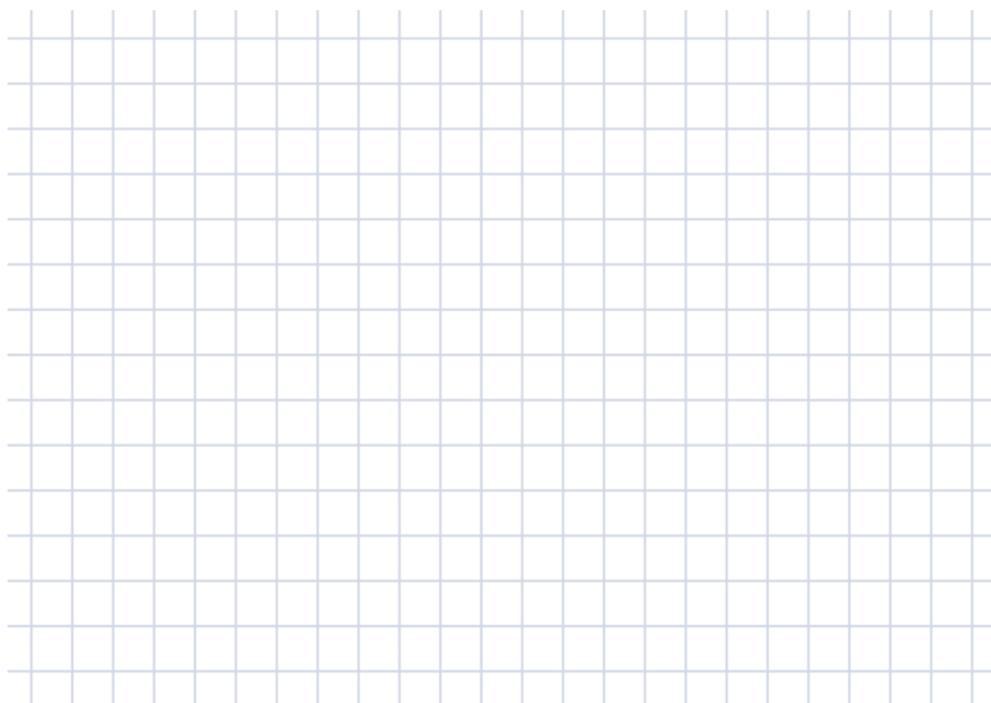
**Figura 1:** Exemplo fictício de representação inadequada baseado na pesquisa de Watson (1996).



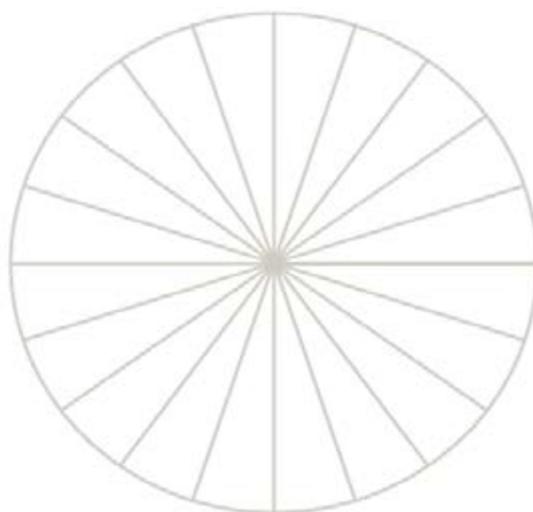
**Figura 1:** Exemplo fictício de representação adequada baseado na pesquisa de Watson (1996).

No Brasil, encontramos os trabalhos das pesquisadoras Selva e Borba (2005) que apresentam resultados relevantes relativos à compreensão da divisão inexata. Elas observaram como crianças comparam os resultados de um mesmo problema de divisão com resto resolvido por meio de diferentes representações (papel e lápis, calculadora versus papel e material manipulativo). Essa análise nos permite compreender melhor as razões pelas quais as crianças apresentam dificuldades no cálculo da média.

## ANEXO F: MALHAS



Malha quadriculada.



Malha circular com divisões relativas a 5%.

