

PLANEJANDO ABRIR UM NEGÓCIO? USE ESTATÍSTICA!

PLANNING TO OPEN A BUSINESS? USE STATISTICS!

Lisbeth Kaiserlian Cordani¹

Doris Satie Fontes²

RESUMO

Este artigo é uma síntese de uma atividade de Estatística Descritiva, que foi repetida pelas autoras em diversas ocasiões, com o objetivo de despertar, nos alunos e professores da escola básica, a necessidade de reconhecer padrões de comportamento em dados coletados, que permitam tomar decisões sobre um problema proposto. A abordagem propõe a formulação de projeto para um plano de negócio, que pretende apresentar ideias sobre os tamanhos para confeccionar um produto, no caso, luvas. A sugestão é apresentar os passos de uma coleta de dados reais em sala de aula (medida do palmo da mão), rápida de executar no ambiente escolar. Os conceitos abordados são de fácil compreensão, tratando desde a distribuição (empírica) dos dados, com diferentes representações gráficas, até a apresentação de medidas resumo dos mesmos dados, através de medidas de variabilidade e posição. Verifica-se a convergência dos resultados das diferentes medidas obtidas ao longo de vários anos e é sugerida uma tomada de decisão em termos do negócio proposto, com base nos dados coletados.
Palavras-chave: *Estatística Descritiva. Distribuição. Medidas de Variabilidade. Medidas de posição. Plano de Negócio.*

ABSTRACT

This paper is a synthesis of a Descriptive Statistics activity, which was repeated by the authors in several occasions, with the objective of awakening in the students and teachers of the basic school the need to recognize patterns of behavior in data collected, that allow to make decisions on a proposed problem. The approach proposes designing a project for a business plan, which intends to present ideas about the sizes to make a product, in this case, gloves. The suggestion is to present the step by step of a real data collection (span of the hand), fast to perform in the school environment. The concepts are easy to understand, ranging from the (empirical) distribution of data, with different graphical representations, to the presentation of summary measures of these data set, through measures of variability and position. There is convergence of the results of the different measures obtained over several years and a decision is suggested in terms of the proposed business, based on the data collected.

Keywords: *Descriptive Statistics. Distribution. Variability measures. Position Measures. Business plan.*

¹ Doutora em Educação (USP), Professora. Aposentada do IME USP. E-mail: lisbethkcordani@gmail.com.

² Bacharel em Estatística IME USP, conselheira do Conselho Regional de Estatística da 3ª. Região CONRE-3 (SP, PR, MT e MS). E-mail: dsfontes@gmail.com

Introdução

Se você pretende abrir uma fábrica de sapatos, uma fábrica de bonés ou ainda uma fábrica de luvas e não tem esse negócio já aberto por um membro da família, saiba que a estatística pode vir ao seu encontro. Como?

Bem, vamos nos fixar nas luvas... Alguns pontos básicos devem ser levados em conta:

- o produto é para Mulheres? Homens? Ambos?
- o produto é para Bebês? Crianças? Adultos?
- o produto é de natureza Social? Esportiva? Outras?

Embora a abertura do negócio envolva várias abordagens e linhas de ação, vamos aqui simplificar, fazendo a abordagem de um só ponto de vista, e pensar na luva propriamente dita, destacando como característica principal o *tamanho*, que deve levar em conta diferentes dimensões, mas que para nossos propósitos tomará como referência inicial somente o palmo estendido da mão. Vamos também fixar o interesse no segmento de idade igual ou maior do que 15 anos, para ambos os sexos.

Uma pergunta inicial seria: basta fazer luvas todas iguais que atenderão ao público interessado? Se fizermos essa pergunta para um grupo de amigos, de parentes ou de alunos, virá imediatamente a resposta: NÃO! Ou seja, nossa vivência já demonstrou que nosso palmo da mão não é necessariamente igual ao de um parente ou de uma amiga, mas é uma característica que apresenta variabilidade. Quando perguntarmos: devemos fazer um modelo para homens e um para mulheres ou devemos pensar em um único grupo de tamanhos? Novamente ao fazer essa pergunta aos grupos definidos acima, provavelmente obteremos respostas majoritariamente entre NÃO e TALVEZ NÃO! Novamente nossa experiência pessoal nos ajuda a pensar que mãos masculinas e mãos femininas podem ser diferentes.

Do ponto de vista de negócios, ao escrever um texto sobre a criação de uma empresa, Dolabela (1999) dedica o capítulo IV para Plano de Negócios e sugere: “Faça uma pesquisa junto aos clientes. Mesmo em pequena escala, a pesquisa lhe dará informações preciosas e será melhor do que agir no escuro” (p. 164). Aliás, quase sempre é preciso mais de uma pesquisa para concretizar o empreendimento, mas aqui vamos focar em uma única pesquisa de natureza quantitativa.

Pensando na sala de aula da escola básica, neste caso fundamental II ou ensino médio, podemos lançar essa discussão e pedir sugestões para iniciar o conhecimento sobre tamanho do palmo de mãos. Efetuar as medidas nos próprios alunos poderá aparecer como sugestão de um integrante da própria turma e é importante reforçar, no âmbito escolar, a prática de pesquisas dentro de sala de aula e mesmo entre salas de aula, com possibilidades de ampliação. De fato, uma coleta

de dados sobre os palmos da mão, pesquisa de natureza quantitativa, como já o dissemos, pode ser uma abordagem inicial, a fim de que se comece a pensar na definição e confecção do “produto” luva.

O restante deste artigo se refere à discussão sobre coleta de dados a fim de responder a pergunta “qual o tamanho do palmo da mão das pessoas?”, estudando a *variável* de interesse (*tamanho do palmo da mão*) para efeitos da elaboração de um plano de negócio, com desdobramentos para a aprendizagem de conceitos básicos de Estatística. Detalhes sobre planos de negócio podem ser obtidos em Dolabela (1999).

Motivação

Depois de muitos anos dedicados ao ensino da estatística no âmbito universitário, acompanhando o desconforto de alunos e professores com as disciplinas introdutórias e observando a grande evasão nessas disciplinas, ficou evidente que o caminho era iniciar em etapas de escolaridade básica, para que o raciocínio aleatório substituísse o determinístico com o qual os alunos chegavam ao ensino superior.

Dessa forma, as autoras se dedicaram à formação de professores da escola básica, através de oficinas e jornadas para mostrar caminhos que forneceriam a eles e a seus alunos um certo traquejo com análise de dados. Parte do material usado nessas oficinas está descrito em Cordani (2012). De fato a coleta de dados e sua análise de natureza estatística é uma área que está entrando cada vez mais na programação da escola básica no Brasil, junto à disciplina de Matemática, e os professores têm demonstrado cada vez mais interesse por atividades que envolvem diretamente os primeiros passos da coleta de dados, o que indiretamente dará ensejo ao desenvolvimento do raciocínio sobre aleatoriedade e variabilidade. Como já falava Moore, em 1992, citado por Ben-Zvi e Garfield (2004), “dados, variação e incerteza são onipresentes na vida moderna”. Esses mesmos autores recomendam a incorporação de dados reais (não meramente realísticos) o que irá mudar a percepção que os alunos têm de que essa disciplina nada mais é do que uma coleção de ferramentas e técnicas, rapidamente esquecidas. A análise de dados reais está na linha de desenvolvimento deste artigo.

O que abordaremos neste artigo é nossa experiência com o desenvolvimento do que chamamos de Estação 1- *o tamanho do palmo da mão direita*, com a coleta de dados para observar a distribuição empírica dessa característica (que chamamos de variável) e construir medidas resumo para entender o seu comportamento, o que encaminhará um possível plano de negócios como o apresentado na introdução (confecção de luvas).

Apresentação dos conceitos teóricos a serem aplicados

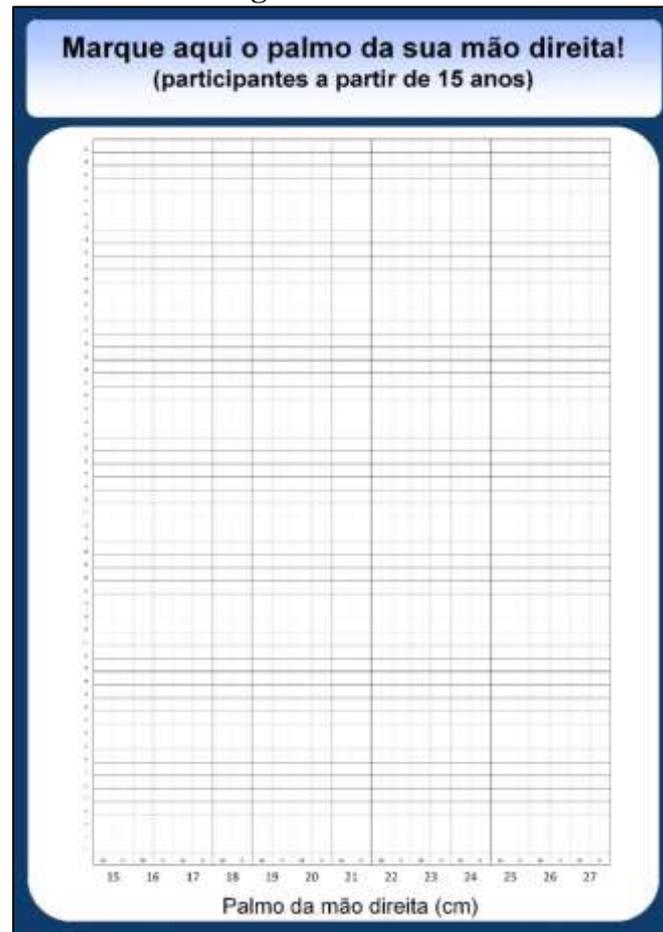
Os conceitos fundamentais a serem abordados serão *distribuição* – comportamento empírico da variável através da construção de gráficos como o gráfico de pontos (*dotplots*), o histograma e o gráfico de caixas (os *boxplots*), *medidas de posição* – serão apresentadas a média, a mediana, a moda, o 1º. Quartil e o 3º. Quartil, e *medidas de variabilidade* – serão discutidas a amplitude e intervalo interquartil.

A apresentação da discussão sobre a distribuição evitará a tendência que os alunos demonstram de olhar os dados como coleção de valores individuais (BAKER; GRAVEMEIJER, 2004), orientando-os a examinar as características agregadas dos dados. A inclusão das medidas de posição, que são muitas vezes referidas como medidas de tendência central, ou compartilhamento justo ou ainda valores típicos (KONOLD; POLLATSEK, 2004) facilita o resumo dos dados, principalmente para comparações entre grupos. No entanto ao resumir dados, não podemos nos contentar somente com essas últimas medidas, pois apesar de elas serem importantes na análise, são insuficientes para descrição dos dados. Daí a necessidade dessas medidas virem acompanhadas das medidas de variabilidade. Também Reading e Shaughnessy (2004) enfatizam que os educadores devem estimular a discussão para além das medidas de tendência central, juntando a essas últimas, sempre que possível, o raciocínio sobre variabilidade. Um dos motivos de não se trabalhar cotidianamente a variabilidade é a ideia de que a única medida a ser trabalhada nesse contexto é o desvio padrão, cujo cálculo é complexo e não facilita o entendimento inicial do verdadeiro conceito que se quer mostrar. Por essa e outras razões a escola básica enfatiza o uso de medidas de posição, fazendo crer que são suficientes para caracterizar um conjunto de dados, quando de fato elas são somente parte da história que os dados podem contar. Assim, completam esses autores, a menos que conheçamos o que os alunos pensam sobre variabilidade perderemos a oportunidade de construir sobre o que eles ou já sabem ou ainda não sabem, e não melhoraremos nosso ferramental pedagógico.

Metodologia de Análise Estatística

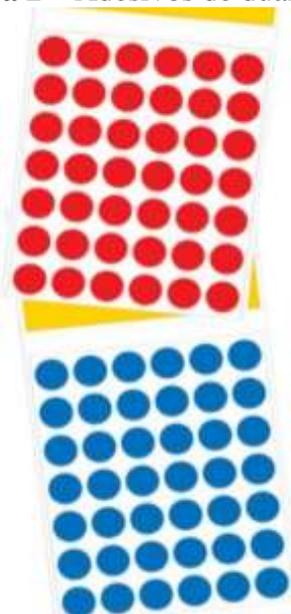
1 – Coleta de Dados

Nosso primeiro objetivo aqui é mostrar algo simples, que pode ser replicado na sala de aula e depois estendido para toda a escola, em diferentes momentos. Como segundo objetivo, vamos mostrar quais conceitos de estatística podem ser ligados a esta atividade de modo a aproveitar diferentes níveis de escolaridade dos alunos.

Figura 1 – Banner

Fonte: Produção nossa.

O material necessário é simples: um banner quadriculado (ou uma folha de cartolina) segundo modelo da Figura 1, adesivos de duas cores (Figura 2), régua de 30 cm e folhas de sulfite.

Figura 2 – Adesivos de duas cores

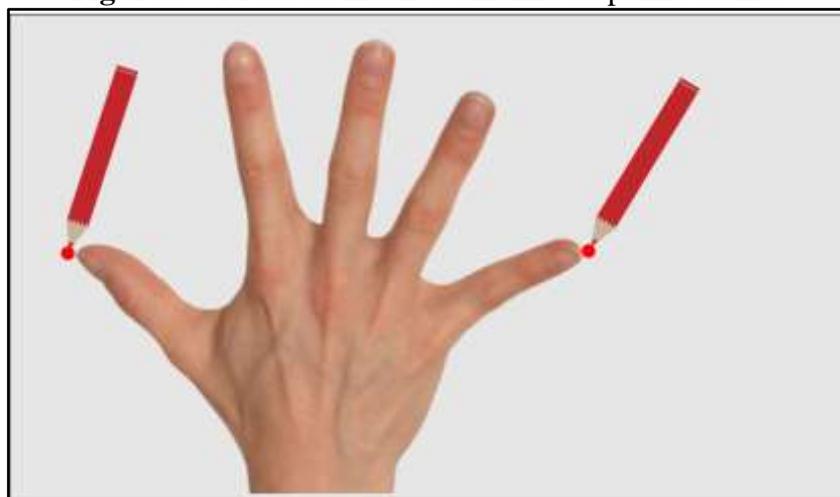
Fonte: Produção nossa.

O procedimento de coleta dos dados será descrito passo a passo.

Passo 1 – Fixar uma idade a partir da qual o gráfico será construído. Uma sugestão poderia ser de 15 anos em diante, mas, se o objetivo for uma escola fundamental, as idades deverão estar num intervalo um pouco menor (com a correspondente redução do tamanho das mãos). Obs.: se a tomada de medidas for numa determinada sala de aula, é interessante estender depois para todas as salas do mesmo ano, por exemplo, fazendo as coletas usando o mesmo banner (ou cartolina).

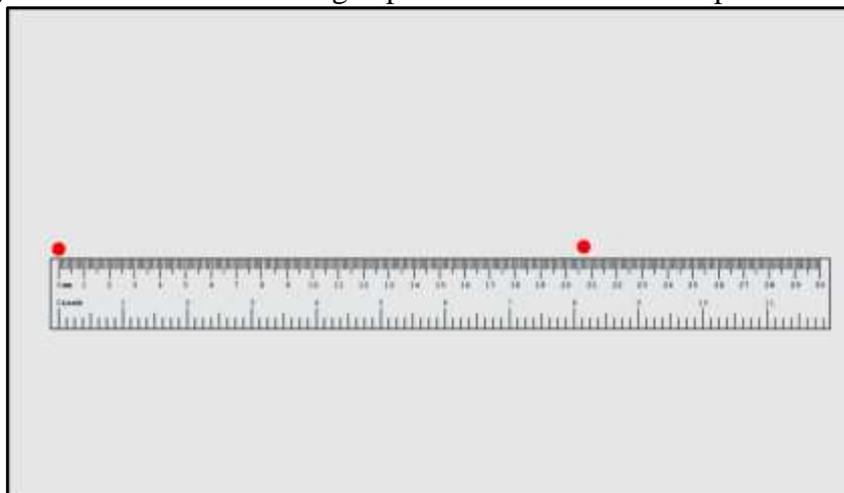
Passo 2 – se for uma feira, todos os passantes farão a medida e se for uma classe, todos os alunos a farão. Pedir para cada pessoa colocar a mão direita bem espalmada em uma folha de papel sulfite (ou numa lousa) fazendo um sinal em dois locais: um na extremidade do polegar e outro na extremidade do dedo mínimo, conforme Figura 3.

Figura 3 – Modo de medir o tamanho da palma da mão



Fonte: Produção nossa.

Figura4 – Como utilizar a régua para medir o tamanho da palma da mão.



Fonte: Produção nossa.

Passo 3 – Retirar a mão e pedir para a pessoa medir com uma régua a distância entre os dois pontos extremos (Figura 4), usando a princípio uma casa decimal. Registrar esse valor e *arredondar para o inteiro mais próximo*. Obs.: aqui vamos fazer uma pausa para discutir este último ponto (arredondamento).

a) Tomemos três exemplos de medidas: 18,2 cm; 19,7cm; 20,5 cm.

Para a primeira, o arredondamento ficaria em 18 cm e para a segunda em 20 cm. Quanto à terceira, valeria pedir sugestão para quem está sendo medido... arredondar para menos? Para mais? Jogar uma moeda para aleatorizar o procedimento, isto é, se cair cara arredondar para cima, caso contrário arredondar para baixo? Escolha uma estratégia e use-a em todos os casos, mas apresente sempre a discussão a todos que estiverem se apresentando para a medida.

b) Esta estratégia de arredondamento não tem alta precisão, o que pode incomodar, em um primeiro momento, tanto os professores quanto as pessoas medidas. Mas nosso interesse não é a precisão em termos de casas decimais, mas sim apresentar as *ideias poderosas* que permeiam o raciocínio estatístico. Assim, manteremos o arredondamento por simplicidade do trabalho em sala de aula, ou mesmo em uma feira de ciências, uma vez que o comportamento da variável é bem descrito através desse procedimento o que veremos por alguns resultados complementares que serão descritos ainda neste artigo.

Passo 4 - Cada pessoa então terá um valor de seu palmo de mão, ou seja, um número inteiro para representar com um adesivo no banner (ou cartolina), seguindo o padrão vertical, com a convenção de colocar uma cor para o sexo feminino, preferencialmente na coluna da esquerda, e uma outra cor para o masculino, na coluna da direita, (claro, poderia ser ao contrário, mas facilitará a leitura do gráfico desta forma). Cada adesivo colocado no gráfico é um ponto e o nome do gráfico assim construído é “gráfico de pontos” (*dotplot*, em inglês). Para efeitos didáticos, o aluno se vê representado explicitamente, o que facilita a construção do gráfico e sua posterior visualização. Além disso, é um gráfico que permite não só ver o comportamento geral da característica estudada, como também observar se há valores discrepantes, ou faltantes, e ainda se há concentração de dados em torno de certos valores. Outra possibilidade de observação é a visibilidade das distâncias entre os valores coletados.

Obs. Essa tomada de dados, se for feita em várias classes de uma escola, facilitará a obtenção de um número grande de medidas de palmos de mão e, à medida que o número de registros cresce, podemos notar algumas características dos grupos medidos, como apresentado no

Passo 4. O mesmo acontece se as medidas forem feitas em uma feira de ciências, por exemplo, local por onde passam centenas de pessoas.

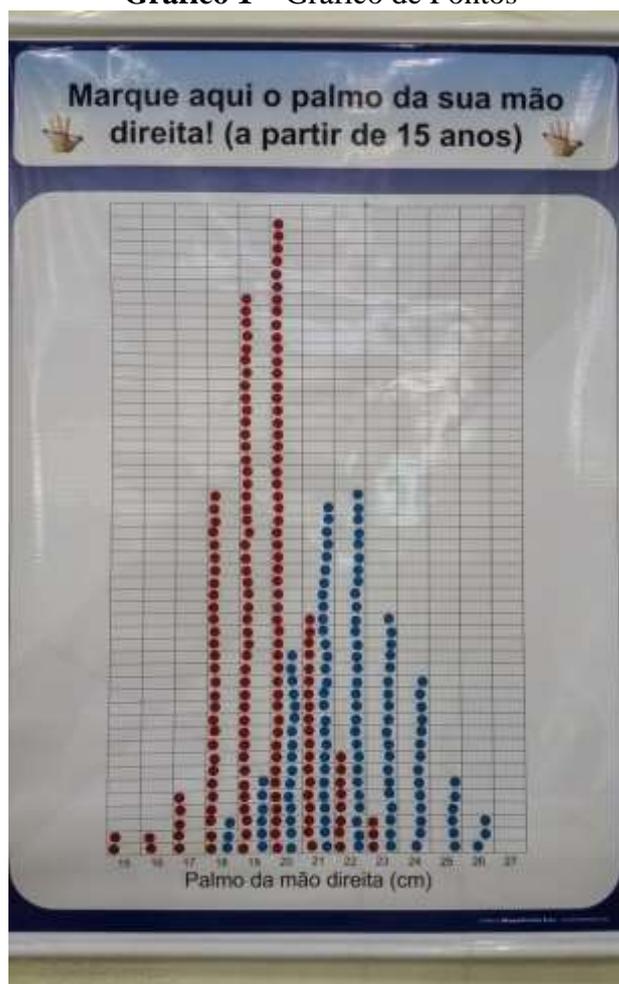
2 - Construção das medidas e dos gráficos

Estamos trabalhando com o que em estatística é denominado de *variável* (definida como uma característica associada ao que estamos medindo): tamanho do palmo da mão direita, em cm. Ao longo da coleta, a superposição de medidas (arredondadas) vai mostrando o comportamento dessa variável, com o aumento do número (n) de palmos medidos. Vai se desenhando um comportamento masculino e um feminino que apresentam uma certa intersecção, mas que, notadamente, separam-se com o palmo das pessoas do sexo masculino à direita e o palmo das pessoas do sexo feminino à esquerda. Pode ser discutido se o aspecto tende a um comportamento simétrico ou não, além de verificar se há valores discrepantes do conjunto geral ou mesmo se a variação entre os valores é análoga para os dois sexos.

O Gráfico 1 mostra o que foi descrito no item a) numa feira de ciências, com 288 passantes ao final do dia, e isto corresponde ao que é chamado em estatística de *distribuição* (empírica) da variável palmo da mão. Dizemos *empírica*, pois foi obtida a partir de dados coletados em um experimento e não a partir de um modelo teórico. O nome *distribuição* é usado em diferentes contextos na área de probabilidade e estatística. Aqui vamos considerar distribuição como o conjunto de valores (ou intervalos de valores) de uma característica (variável) e suas correspondentes frequências observadas (ou probabilidades, se for o caso). Aqui é possível comparar visualmente as duas distribuições e pensar em como seria se não tivéssemos separado por sexo, discutindo com as pessoas medidas.

Além de analisar a distribuição que, como já dissemos, mostra o comportamento gráfico da variável (ou característica) medida, podemos olhar a variável quanto à posição dos valores e comparar inicialmente como se comporta o tamanho do palmo da mão em termos dos sexos masculino e feminino. Como uma medida de posição inicial, podemos definir a *Moda*, que é o valor mais frequente: é fácil num gráfico deste tipo destacar a(s) moda(s) – pode ser uma ou mais de uma (ou até nenhuma). No Gráfico 1, a moda do tamanho do palmo é facilmente identificável para cada sexo, bem como a moda global (sem separação por sexo). Os valores serão apresentados na Tabela 1.

Gráfico 1 – Gráfico de Pontos



Fonte: Produção nossa.

O Gráfico 1 foi obtido ponto a ponto com as medidas de cada passageiro a partir de 15 anos (cada um colocava a sua medida arredondada ao inteiro mais próximo).

Os pontos em vermelho foram marcados por 164 pessoas do sexo feminino e os em azul por 124 pessoas do sexo masculino.

Uma outra medida de posição interessante de calcular, e rápida também, é a *Mediana*, que é o valor que separa metade dos valores (ordenados, ou seja, colocados em ordem geral crescente) de cada lado. Como no Gráfico 1 temos 288 pessoas (número par de medidas), a mediana é o valor médio das medidas das posições 144 e 145 (média dos palmos das posições 144 e 145) obtidas diretamente fazendo a contagem no gráfico, já que os dados se apresentam ordenados. A mediana calculada deixa metade das pessoas à sua esquerda e metade das pessoas à sua direita. O Quadro 5 apresenta o cálculo de medianas tanto para um número par de observações (como é o caso de nosso exemplo) como de um número ímpar. Podemos calcular, também, uma mediana geral e medianas por sexo. Em ambos os casos do Quadro 5, o valor da mediana é 20cm.

Quadro 5 – Esquema para localizar a mediana



Fonte: Produção nossa.

Outra medida de posição é a *Média* (soma de todos os valores dividida pelo número deles), cujo cálculo não é tão imediato a partir do gráfico como o das outras duas medidas – para seu cálculo é necessário registrar todos os dados.

Às três medidas de posição apresentadas, *Moda*, *Mediana* e *Média*, podemos agregar mais duas, também chamadas de posição, que podem ampliar a análise descritiva dos dados. São elas o *1º Quartil (Q1)*, medida que separa os primeiros 25% menores valores dos demais, e o *3º Quartil (Q3)*, que é a medida que separa os últimos 25% maiores valores dos demais. Como já vimos que a mediana divide os dados ordenados em 2 metades, pode-se dizer que o *2º Quartil* é a *Mediana*, embora seja este último termo o utilizado no jargão da análise estatística. O cálculo dos *Quartis* é similar ao da *Mediana*. Para os quartis, propomos usar uma das definições possíveis que acreditamos ser a mais direta: *1º Quartil* é a mediana da primeira metade dos dados e *3º Quartil* é a mediana da 2ª metade dos dados. Isto é facilmente calculado colocando os dados em ordem e calculando a *Mediana* global (que é o valor que separa as duas metades). Teremos então duas metades que favorecem o cálculo dos *Quartis* que, como o nome sugere, dividem os dados em quatro partes, cada uma com 25% dos dados. Quando a *Mediana* pertencer ao conjunto coletado (número ímpar de valores) ela fará parte de cada uma das metades para efeito do cálculo dos *Quartis*. Os valores *Máximo* e *Mínimo* são também considerados medidas de posição. Esse termo *posição* deriva do fato de essas medidas poderem ser *posicionadas* no eixo dos valores coletados. A Tabela 1 apresenta um resumo dos dados coletados, com as medidas de posição calculadas.

Nenhuma análise estatística pode prescindir de, pelo menos, uma medida de variabilidade. É preciso evitar a apresentação somente das medidas de posição em uma etapa escolar para somente em outra etapa apresentar a discussão de variabilidade. É importante saber se todos os valores de

palmo são iguais ou se estão mais dispersos ou menos dispersos, quer entre si, quer em relação a alguma referência. Depois de discutir o comportamento gráfico, em termos de simetria, por exemplo, bem como calcular as medidas de posição, é importante colocar em discussão se os valores são todos iguais ou se apresentam variações (dispersão) entre eles. A partir do banner preenchido é possível verificar a existência ou não de dispersão. Tentando caracterizar essa variação por uma medida de variabilidade, apresentamos a mais imediata e simples de calcular, baseada numa variação entre o valor mínimo e o valor máximo coletados. Essa é a chamada *Amplitude* e seu cálculo é feito pela diferença (valor *Máximo* – valor *Mínimo*). Cabe também a indagação: “os palmos da mão do sexo masculino e do feminino variam de modo similar”?

A Tabela 1 apresenta todos os cálculos das medidas definidas de posição e de variabilidade, tanto para o conjunto todo como separadas por sexo. Com este panorama é possível olhar os dados de modo agregado e aqui podemos mais uma vez enfatizar que ter usado os valores inteiros facilita a primeira abordagem do tema.

Tabela 1 – Medidas descritivas (cm)

Distribuição Empírica	<i>GLOBAL</i>	<i>MASCULINO</i>	<i>FEMININO</i>
Tamanho da amostra (n)	288	124	164
Medidas de posição			
Moda	20	22	20
Mediana	20	22	19
1º quartil	19	21	19
3º quartil	22	23	20
Média	20,5	21,9	19,4
Mínimo	15	18	15
Máximo	26	26	23
Medida de variabilidade			
Amplitude	11	8	8

Fonte: Produção nossa.

3 - Complementação de análise gráfica dos dados

O primeiro gráfico que apresentamos foi o Gráfico de Pontos (Gráfico 1), justamente construído durante a tomada das medidas das pessoas. É o gráfico mais intuitivo e deveria ser mais utilizado em coletas de dados de variáveis quantitativas. Em termos de sala de aula, a vantagem do gráfico é que cada elemento da classe se vê representado explicitamente, já faz comparações intuitivas, e passa a ficar interessado sobre o que se fará com o seu valor e com o de seus colegas.

Vamos aqui introduzir mais dois gráficos, usuais em análises estatísticas, deixando claro que não é preciso fazer todos eles em cada análise. Quem elabora a análise decidirá qual o mais apropriado em cada caso.

Assim outro gráfico que poderia ser construído a partir dos dados é o chamado HISTOGRAMA, que não é tão intuitivo como o anterior, mas dá uma ideia agrupada do comportamento da variável estudada (tamanho do palmo da mão direita). Para a sua construção, os dados devem ser colocados em classes com a correspondente construção da tabela de frequências. Um exemplo de tabela deste tipo (classe de 2 em 2 cm aberta à direita e fechada à esquerda) é a Tabela 2, em que as frequências apresentadas são provenientes da quantidade em cada classe (frequências absolutas).

Tabela 2 – Tabela de frequências absolutas

Classes	Feminino	Classes	Masculino
14 -----16	2	18 -----20	9
16 -----18	7	20 -----22	44
18 -----20	74	22 -----24	48
20 -----22	70	24 -----26	20
22 -----24	11	26 -----28	3
Total	164	Total	124

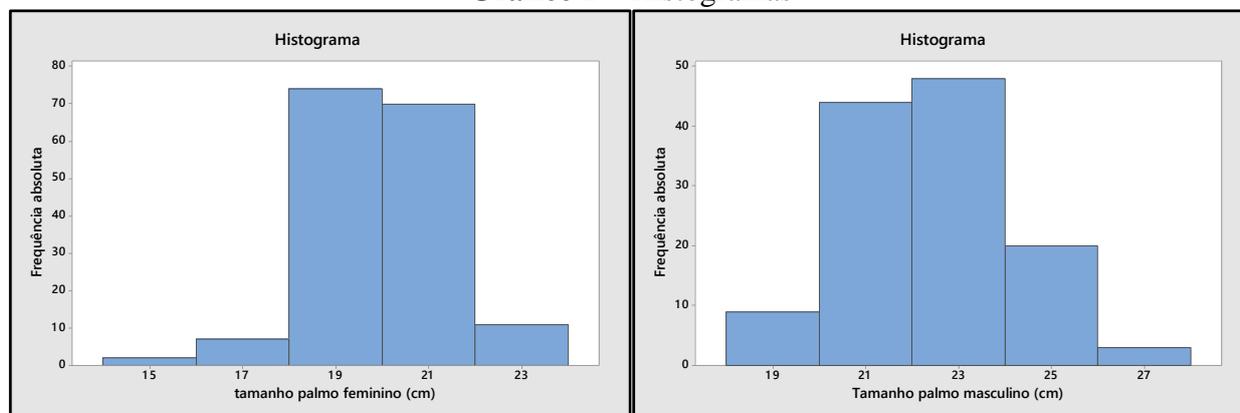
Fonte: Produção nossa.

Os correspondentes histogramas estão a seguir (Gráfico 2). É fácil ver porque os alunos em geral são mais motivados pelo Gráfico de Pontos, onde cada um se vê representado, o que não ocorre no histograma, onde os dados estão condensados em classes. Mas para ter uma visualização de padrão de comportamento, o Histograma é a continuação didática natural. O Gráfico 2 mostra a representação do Histograma, por sexo, dos dados do palmo da mão direita, tomando como base a Tabela 2 (tabela de frequências absolutas).

O Histograma é construído pela justaposição de colunas, com base em cada classe de tamanho do palmo e com altura igual a cada frequência observada naquela classe. Aqui, como os tamanhos das classes são iguais, a aparência do gráfico representa o comportamento da variável sem distorção. Se os tamanhos dos intervalos fossem diferentes entre si, esta representação não seria adequada (nesse caso, usa-se como ordenada a *densidade*, que é o cociente entre a frequência

relativa e o tamanho da classe correspondente). Os comentários feitos para o Gráfico de Pontos também podem ser percebidos através do Histograma.

Gráfico 2 – Histogramas



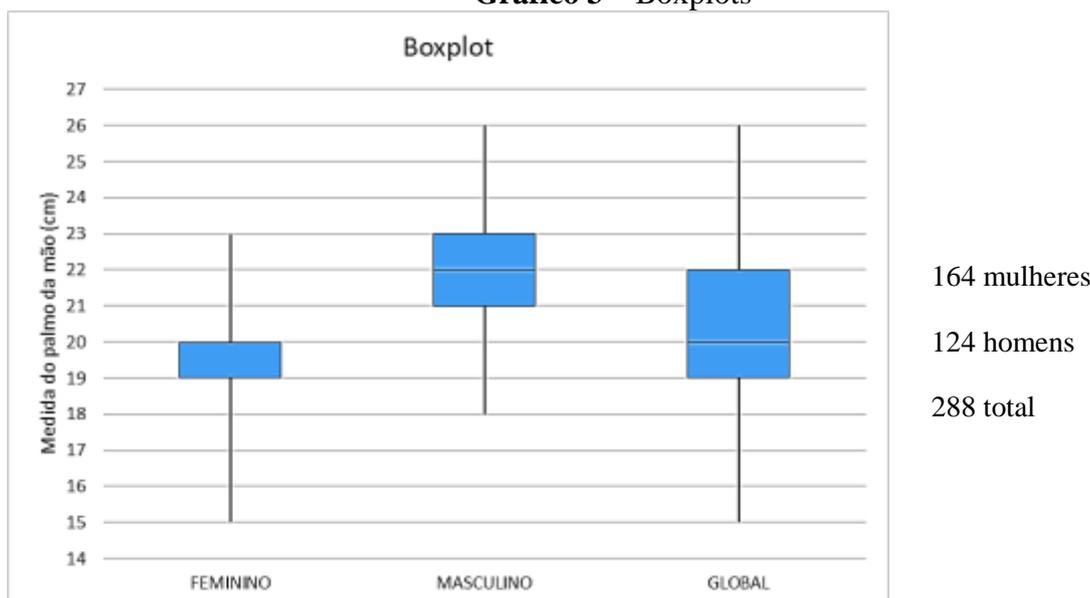
Fonte: Produção nossa.

Um terceiro gráfico que podemos sugerir é o denominado Gráfico de Caixas (em inglês *BoxPlot*) bem fácil de construir, e que utiliza 5 medidas de posição já definidas. Para a sua construção usaremos:

Valor Mínimo, 1º. Quartil (Q_1), Mediana, 3º. Quartil (Q_3) e Valor Máximo.

Construímos uma caixa retangular tendo como base inferior o primeiro quartil (Q_1) e como base superior o terceiro quartil (Q_3). Divide-se a caixa com um traço simbolizando o valor da mediana. Em sua apresentação mais simples, liga-se com uma haste o centro da base inferior ao valor mínimo e o centro da base superior ao valor máximo. O Gráfico 3 apresenta os Gráficos de Caixas (*boxplots*) correspondentes às medidas calculadas na Tabela 1. A colocação de todos juntos facilita a visualização e permite comparações imediatas. Pelas definições anteriores, podemos perceber que os 50% centrais dos dados são delimitados pelos valores da caixa propriamente dita; por exemplo, para o gênero feminino metade dos valores está entre 19 e 20 cm, enquanto que para o masculino, temos a metade entre 21 e 23 cm. Esses valores são justamente os quartis, os quais ensejam a construção de mais uma medida de variabilidade, ou seja, o *Intervalo Interquartil* = $Q_3 - Q_1$. Esta medida é usada para a forma mais elaborada da construção do Gráfico de Caixas, que não será abordada aqui (sugerimos consultar Magalhães e Lima (2016)). Fazê-lo à mão é bem simples também, bastando calcular as 5 medidas de posição já mencionadas.

Gráfico 3 – Boxplots



Fonte: Produção nossa.

4 - Interpretação dos resultados

A partir desses resultados é possível fazer tanto uma análise global quanto uma análise separando por gênero. A seguir apresentamos uma análise rápida do comportamento dos dados.

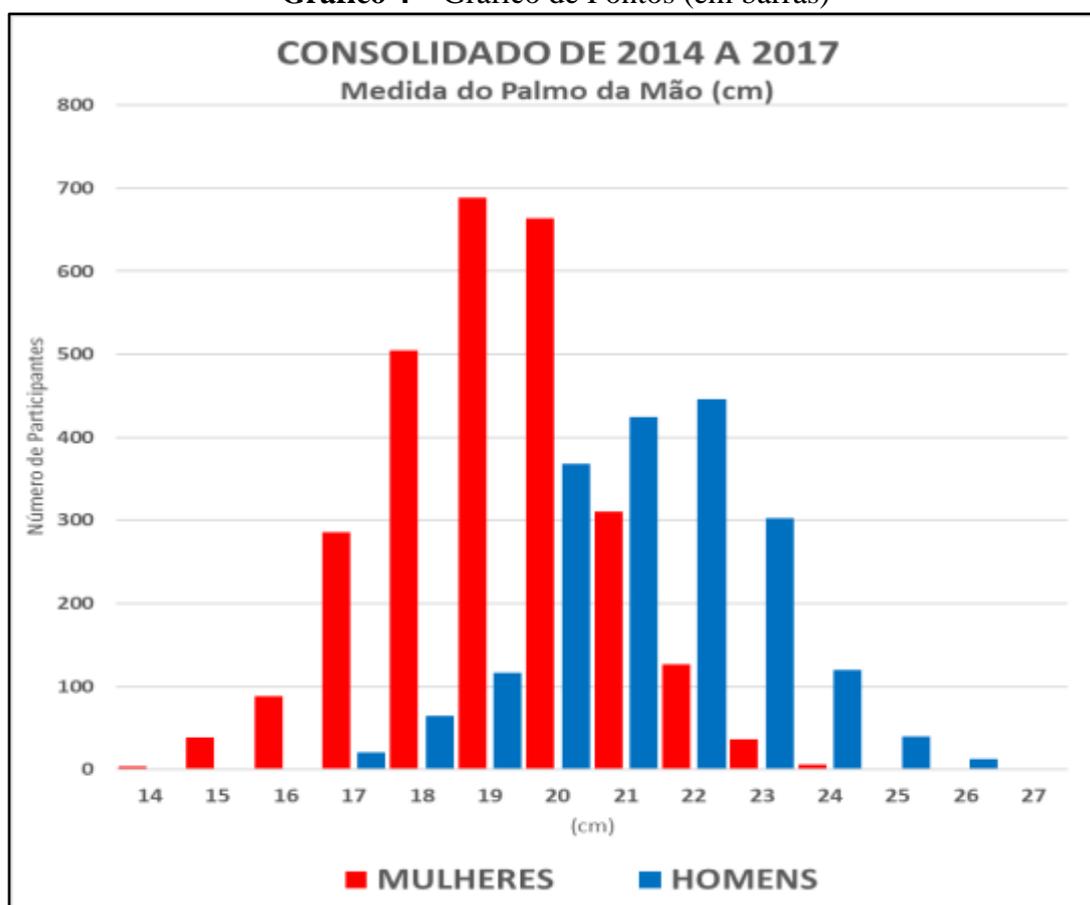
- 1 Verifica-se que 50% dos valores centrais dos palmos do sexo feminino estão entre 19 e 20 cm, enquanto que a mesma porcentagem do sexo masculino está entre 21 e 23 cm, sugerindo que valores masculinos são geralmente maiores e variam mais do que os valores femininos. A existência de variabilidade dentro de cada gênero (vista pela *Amplitude* ou *pelo Intervalo Inter-Quartil*) sugere que não façamos um tamanho único por gênero, e poderia ser adotada a classificação P (pequena), M (média) e G (grande). Os gráficos apresentados corroboram com essas sugestões.
- 2 A *Moda*, a *Mediana* e a *Média* são relativamente próximas dentro de cada grupo (global, masculino, feminino), mas dão indicações interessantes sobre as diferenças de gênero, incluindo possíveis limites por tamanho. Todas as medidas de posição indicam superioridade de tamanho para o gênero masculino o que sugere separar o tamanho (P, M, G) das luvas por GÊNERO (é claro que o plano de negócios pode decidir por mais categorias, como por exemplo, tamanhos PP e GG).
 - Luvas femininas de tamanho P podem considerar medidas de 15 cm a 18 cm; M podem ser de 19 cm a 20 cm, enquanto que G podem ser de 21 cm a 23 cm.

- Luvas masculinas de tamanho P podem considerar medidas de 18 cm a 20 cm; M de 21 cm a 23 cm e G de 24 cm a 27 cm.
 - Obs. 1 - Se o tamanho vai ser pelo valor máximo do intervalo ou pelo valor médio do mesmo é para se decidir durante a elaboração do plano de negócio.
 - Obs. 2 – Essa é uma linha de ação. Podem ser sugeridas outras.
- 3 Com relação às quantidades de fabricação, isso vai depender de outras variáveis, como, por exemplo, a escala de produção. Mas, de acordo com o Gráfico 1, o encaminhamento seria para fabricar maior quantidade de luvas M e menor quantidade de P ou G.

5 - Resultados “mais além”

Em diversas oportunidades, como já o dissemos, realizamos esta atividade da Estação 1 e apresentaremos o gráfico de pontos consolidados obtidos em oito reuniões em diversos pontos do país e as estatísticas descritivas correspondentes.

Gráfico 4 – Gráfico de Pontos (em barras)



Fonte: Produção nossa.

A Tabela 3 mostra as estatísticas descritivas associadas ao conjunto maior de medidas (n=4662 observações), as mesmas consideradas no Quadro 4.

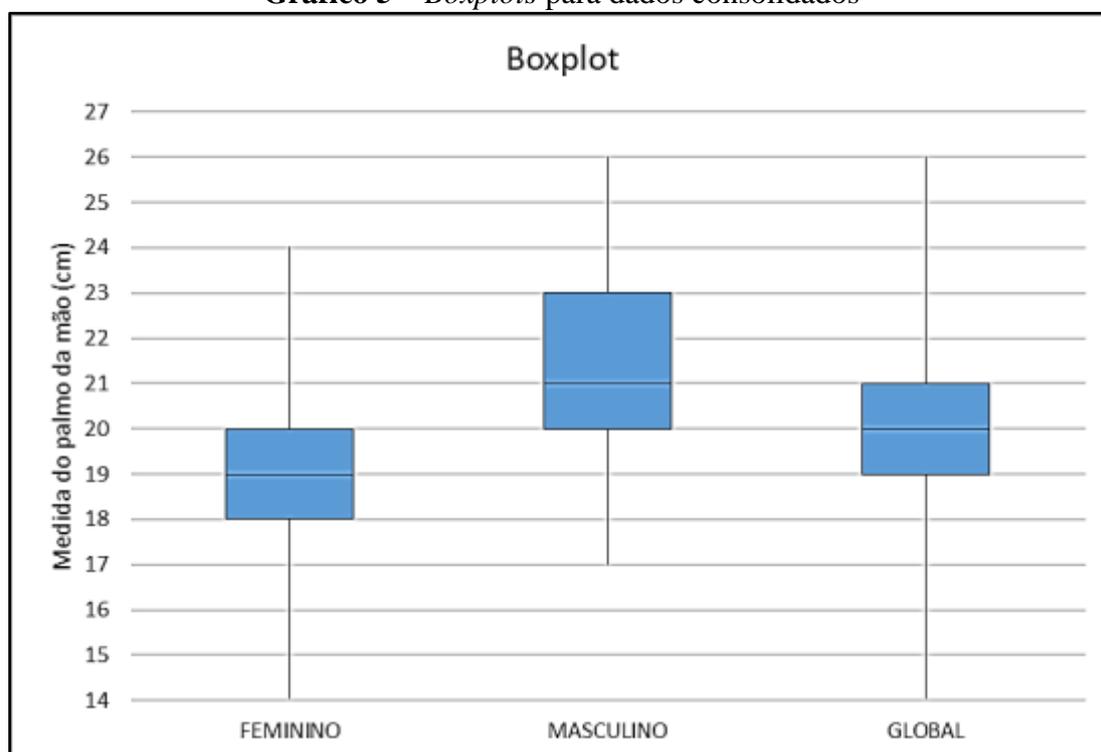
Tabela 3 – Medidas descritivas (cm)

<i>Distribuição Empírica</i>		GLOBAL	MASCULINO	FEMININO
<i>GLOBAL (n)</i>		4662	1912	2750
<i>Medidas de posição</i>	Moda	20	22	19
	Mediana	20	21	19
	1º quartil	19	20	18
	3º quartil	21	23	20
	Média	20,1	21,4	19,1
	Mínimo	14	17	14
	Máximo	26	26	24
<i>Medida de variabilidade</i>	Amplitude	12	9	10

Fonte: Produção nossa.

Os Gráficos de Caixas (*boxplots*) correspondentes aos dados consolidados estão apresentados no Gráfico 5.

Gráfico 5 – *Boxplots* para dados consolidados



Fonte: Produção nossa.

Analisando os resultados da compilação de todas as medidas obtidas nas diversas oficinas e feiras, podemos concluir que os tamanhos de luvas escolhidos para o primeiro caso são semelhantes nesta análise global. Para luvas femininas, valeria pensar se seria realmente necessário fabricar um

PP (tamanho da medida 14 cm), já que a incidência foi mínima. Assim como um PP masculino para o tamanho 17 cm, cuja incidência também foi muito pequena. O mesmo raciocínio valeria para luvas femininas de tamanho GG (equivalente à medida 24 cm). A Estatística fornece os indícios quantitativos, mas a decisão da empresa deve se basear também em aspectos que otimizam o plano de negócios.

Comentários Finais

Já comentamos que os estudantes, ao olharem o comportamento gráfico dos dados, não o veem como agregado de valores e sim como valores individualizados, o que sugere que é preciso construir passo a passo o conceito de distribuição. Nossa sequência de apresentação, começando com o Gráfico de Pontos para depois introduzir o Histograma e o Gráfico de Caixas, usa a estratégia mais compatível ao raciocínio dos alunos, ou seja, começar a ver o problema através dos dados individualizados para depois pensar nos dados agregados, como um conjunto de observações. Assim, aprenderão a olhar o conjunto e não o indivíduo. Também a importância da variabilidade é reforçada bem como o uso de dados reais e possibilidade de elaborar as análises com ferramentas simples, como lápis e papel.

Este artigo descreve os passos de uma análise estatística, contemplada numa atividade oferecida aos alunos e professores da escola básica. Coleta, cálculo de medidas-resumo (posição e variabilidade) e comportamento da distribuição empírica através de gráficos para descrever o comportamento dos dados foram apresentados e discutidos, tendo como objetivo dar respaldo a um plano de negócios.

Pela importância do tema, apresentamos uma última discussão sobre variabilidade, para a qual apresentamos até o momento duas medidas: a *Amplitude* e o *Intervalo Interquartil*. Com relação à *Amplitude*, é importante a discussão com os alunos sobre pontos fortes e fracos desta medida, para que eles adquiram traquejo com análise de dados. Como *ponto forte*, podemos citar a simplicidade de cálculo, pois basta ordenar os dados e tomar o valor máximo e o valor mínimo e fazer a diferença entre eles. No caso extremo de todos os valores serem iguais entre si, a *Amplitude* será igual a zero o que nos dá uma indicação de variabilidade nula. Isso então mostra que a medida é adequada para o início da discussão sobre o tema. Como *ponto fraco*, podemos citar o uso de somente dois valores para caracterizar o comportamento da dispersão dos dados, uma vez que essa medida ignora o corpo dos dados, ou seja, o comportamento de todas as demais medidas. Mas mesmo assim, o início da discussão de variabilidade deve passar pela *Amplitude*, pode prosseguir pelo Intervalo Interquartil, o qual nada mais é do que a *Amplitude* de variação dos 50% valores centrais, e que possibilita a comparação entre os “miolos” dos dois grupos quanto à variação, para somente depois definir outras medidas mais abrangentes. O leitor interessado pode consultar

Magalhães e Lima (2016) para definir a conhecida medida de variabilidade, o *desvio padrão*, ao qual voltaremos em outra oportunidade.

Normalmente nessas atividades temos contado com a colaboração de monitores, em grande parte alunos provenientes ou da Licenciatura em Matemática ou do Bacharelado em Estatística. Os depoimentos desses alunos são sempre entusiasmados e encorajam a repetição dessas atividades. Terminamos com um deles, sem edição, escrito no final da atividade por uma aluna da Licenciatura em Matemática de uma universidade federal:

“A tenda da estatística foi ótima. Teve um grande envolvimento das pessoas e ao mesmo tempo possibilitou o acesso à Estatística para o público de qualquer faixa etária. As atividades desenvolvidas foram bem interessantes e atrativas e isso mostrou que Estatística pode ser aprendida por qualquer pessoa de forma simples e divertida.”

Recebido em: 30/ 04/ 2018

Aprovado em: 30/ 05/ 2018

Referências

BAKKER, A.; GRAVEMEIJER, K. P. E.. Learning to reason about distribution. In: BEN-ZVI, D.; GARFIELD, J. (eds). **The Challenge of Developing Statistical Literacy, Reasoning and Thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p. 147–168.

BEN-ZVI, D.; GARFIELD, J. Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In: BEN-ZVI, D.; GARFIELD, J. (Eds.). **The Challenge of Developing Statistical Literacy, Reasoning and Thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p. 3–15.

CORDANI, L. K.. **Estatística para Todos: atividades para sala de aula**. 1ª. ed. São Paulo: CAEM/IME-USP, 2012.

DOLABELA, F.. **O Segredo de Luísa**. São Paulo: Cultura, 1999.

KONOLD, C.; POLLATSEK, A.. Conceptualizing an Average as a Stable Feature of a Noisy Process. In BEN-ZVI, D.; GARFIELD, J. (eds). **The Challenge of Developing Statistical Literacy, Reasoning and Thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p. 169–199.

MAGALHÃES, M. N.; LIMA, A. C. P.. **Noções de Probabilidade e Estatística**. São Paulo: EDUSP, 7ª. ed., 2016.

READING, C.; SHAUGHNESSY, J. M.. **Reasoning about variation**. In: BEN-ZVI, D.; GARFIELD, J. (Eds.). **The Challenge of Developing Statistical Literacy, Reasoning and Thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, p. 201–226.