

ANÁLISE DIDÁTICA DO USO DOS SOFTWARES R E GEOGEBRA NO DESENVOLVIMENTO DO LETRAMENTO ESTATÍSTICO

Cileda de Queiroz e Silva Coutinho e Fabiano dos Santos Souza

Pontifícia Universidade Católica de São Paulo, Brasil

Universidade Federal Fluminense, Brasil

cileda@pucsp.br

Pesquisas no campo da Educação Estatística (Batanero; Estepa; Godino, 1991) apontam a análise exploratória de dados como uma filosofia importante e eficaz na abordagem dos conteúdos estatísticos, quando objetiva-se o desenvolvimento do letramento estatístico. Abordamos, neste artigo, aspectos e elementos relevantes para o desenvolvimento desse letramento, tais como, a transnumeração (Wild e Pfannkuch, 1999; Pfunckhuk, 2008) e a utilização de múltiplos registros de representação semiótica (Duval, 2003) são estudados e articulados em textos, tais como Coutinho, Silva e Almouloud (2011). Para tanto, faz-se necessário o uso de ambientes computacionais, de forma a potencializar a construção dos conceitos da estatística descritiva. A partir dessa perspectiva, efeturemos a análise didática dos softwares Geogebra e R na construção de gráficos; discutindo critérios, identificando contribuições, fragilidades, ou limitações. Tais critérios serão discutidos à luz da contribuição para a construção e mobilização desses conhecimentos conduzindo o aluno no desenvolvimento e evolução do letramento estatístico (Gal, 2002).

INTRODUÇÃO

Os avanços tecnológicos e computacionais, a partir da metade do século XX, proporcionaram à sociedade viver num mundo moderno e globalizado, onde as informações são apresentadas pelos meios de comunicação basicamente por meio de gráficos e tabelas e, muitas vezes, não são compreendidas pelos cidadãos. Gal (2002) destaca a importância de se proporcionar uma cultura estatística para sociedade, a partir do desenvolvimento da capacidade de interpretar, avaliar criticamente a informação estatística, discutir e argumentar a respeito das informações estatísticas que julgar relevantes. Desse modo, percebemos a importância da reflexão sobre o processo de ensino e de aprendizagem dos conteúdos estatísticos. O exercício profissional e o exercício pleno da cidadania exigem a tomada de decisão e a compreensão dos conceitos estatísticos envolvidos.

É fundamental sublinhar que tais decisões devem ser tomadas a partir da apreensão e da compreensão da variabilidade contida nos dados a serem analisados para subsidiá-las. Daí a necessidade de que o cidadão de hoje seja estatisticamente letrado, para que possa não apenas compreender o noticiário divulgado nas mais diversas mídias, mas também se expressar empregando corretamente termos e noções estatísticas, relacionando-as sempre que necessário para constituir uma análise eficiente e eficaz.

Tomando como referência o conceito de letramento estatístico o uso correto dos conceitos e procedimentos estatísticos pelo sujeito, concordamos com a definição adotada por Garfield, delMas e Chance (2003), citada por BenZvi e Garfield (2004), que reproduzimos abaixo:

Letramento estatístico inclui habilidades básicas e importantes que podem ser usadas para compreender informações estatísticas ou resultados de pesquisa. Estas habilidades incluem estar apto a organizar dados, construir e exibir tabelas, e trabalhar com diferentes representações dos dados. O Letramento Estatístico inclui também uma compreensão de conceitos, vocabulário e símbolos, além de incluir uma compreensão de probabilidade como uma medida da incerteza. (Garfield, delMas e Chance, 2003, apud Ben-Zvi e Garfield, 2004, p.7)

Neste artigo nos propomos a discutir o uso dos softwares Geogebra e R na construção de gráficos, discutindo critérios, identificando contribuições, fragilidades, ou limitações, como ferramentas para o desenvolvimento do letramento estatístico, uma vez que estes permitem a articulação entre os diversos registros de representação semiótica da distribuição de um conjunto de dados. Tais critérios serão discutidos à luz da contribuição para a construção e mobilização

desses conhecimentos conduzindo o aluno no desenvolvimento e evolução do letramento estatístico.

O USO DOS SOFTWARES GEOGEBRA E R NO ENSINO DE ESTATÍSTICA

Tomamos como referência os termos da filosofia da Análise Exploratória de Dados (Batenero, Estepa e Godino, 1991); e a premissa assumida por Coutinho e Souza (2013), segundo a qual a utilização de ambiente computacional permite ao professor fazer a gestão das atividades de aprendizagem de forma a que os procedimentos para a construção de gráficos não se tornem o foco dessas atividades, permitindo a discussão conceitual sobre a distribuição em jogo.

Assim, assumimos que a Análise Exploratória de Dados, que nos indica a coleta de dados de forma contextualizada, é um elemento basilar da construção do letramento estatístico (Coutinho, Souza, 2013).

Os livros didáticos aprovados pelo Programa Nacional de Livros Didáticos (PNLD), em edições relativas ao Ensino Fundamental e Médio, trazem em sua maioria tarefas cujo objetivo é a construção e leitura de alguns gráficos estatísticos, assim como valorizam a utilização de algoritmos para os cálculos das medidas estatísticas e probabilísticas. Constata-se assim uma maior ênfase no aspecto matemático (aspectos procedimentais), em detrimento dos conceitos e significados, proporcionando uma aprendizagem conceitual.

Coutinho (2014) sublinha que nos livros didáticos há uma forte tendência à utilização de representações gráficas como ponto de partida para o ensino e a aprendizagem dos conteúdos da estatística descritiva, ou seja, são raros os momentos que se utiliza a filosofia da Análise Exploratória de Dados, já que o aluno recebe os dados já coletados e representados em forma tabular. Segundo essa autora, pouco se discute sobre o contexto no qual os dados foram coletados e, principalmente, sobre o questionamento que gerou a necessidade desses dados.

Diante desse cenário, emerge a necessidade sobre a discussão e reflexão sobre o uso de programas computacionais, em nosso caso, o Geogebra e R, para a construção de gráficos de forma a que possam completar o proposto nos livros didáticos, potencializando a aprendizagem e, portanto, o desenvolvimento do letramento estatístico. Essa perspectiva nos permite, também, apreender o fato de que o desafio para os pesquisadores é “comunicar seus achados de forma a impactar o desenvolvimento da prática dos professores e a aprendizagem dos alunos” conforme sublinha Pfannkuch (2008, p.5). Este é o cenário no qual o presente trabalho se insere.

O GeoGebra é um *software* de matemática dinâmica gratuito e multi-plataforma para todos os níveis de ensino, que oferece ferramentas para o estudo de geometria, álgebra, tabelas, gráficos, estatística e cálculo. Quanto ao *software* R, Souza (2013, p. 87) destaca que é uma ferramenta para tratamento estatístico de dados, que oferece gráficos bem desenhados com ótima qualidade para publicação, podendo incluir símbolos matemáticos e fórmulas quando necessário. Esse autor destaca também as vantagens do uso do pacote Rcmdr, mais conhecido com *R Commander*, que foi desenvolvido no ano 2000 visando contornar algumas dificuldades encontradas pelos usuários, principalmente no que se referia à sintaxe dos comandos. A vantagem da utilização desse pacote é a existência de vários *menus* que conduzem à procedimentos e comandos estatísticos, sem a necessidade de programá-los com a sintaxe usual do R. Há também uma interface para saída dos resultados, sendo que os gráficos também são gerados em janela separada.

Destacamos que serão consideradas nesse texto as versões dos *softwares* Geogebra 5.0 e do R 3.1.0. Consideramos a abordagem dos gráficos estatísticos em atividades colaborativas de aprendizagem em sala de aula, desde a construção até a análise, que leva em conta não apenas as representações escolhidas, mas principalmente o conhecimento do contexto no qual os dados foram coletados. Para tanto, utilizamos em nossas análises dois critérios: (i) usabilidade didática; (ii) construção de dois tipos de gráficos em um mesmo sistema de eixos.

O primeiro aspecto a destacar diz respeito ao primeiro critério: a usabilidade didática. Coutinho e Souza (2013) concebem usabilidade didática como o custo cognitivo para manipulação do software para a produção de significados a partir de transformações nos registros de representação semiótica mobilizados pelo sujeito.

O primeiro ponto a ser analisado nesse critério é a facilidade de inserir dados no programa. Tal facilidade pode induzir um elemento “distrator”, ou seja, se o aluno julgar difícil a inserção dos dados no *software*, a aprendizagem do seu uso assumirá uma dimensão prioritária em relação ao

conteúdo estatístico visado, conforme Coutinho e Souza (2013). Trata-se aqui de um obstáculo possível para a apreensão da variabilidade pela visualização dos dados em mais de um registro de representação semiótica, tal como sugerido pela filosofia da Análise Exploratória de Dados (BATANERO, ESTEPA e GODINO, 1991). Nesse sentido, destaca-se a importância da familiarização do sujeito com o software, de forma a minimizar a possibilidade de surgimento desse distrator. Ou seja, ressalta-se a necessidade da concepção de atividades didáticas que permitam ao sujeito a apreensão das formas de utilização do software, a partir do resultado de um diagnóstico feito pelo professor.

O segundo critério a ser discutido é a possibilidade de construção de dois tipos de gráficos em um mesmo sistema de eixos, visando potencializar as condições didáticas para que o aluno compare as duas representações sob uma mesma escala, de forma aprofundar a análise dos dados ali representados. Tal possibilidade visa contemplar os princípios da Análise Exploratória de Dados pela utilização simultânea de mais de uma representação dos dados. Coutinho, Almouloud e Silva (2012) assinalam que tal tipo de construção fornece ao aluno condições de análises mais coerentes.

RESULTADOS E DISCUSSÃO

Neste artigo limitaremos nossas análises ao uso simultâneo do gráfico de pontos e do gráfico de caixa (*boxplot*). Souza (2013) ressalta que o *boxplot* permite avaliar a simetria dos dados, sua dispersão e a existência ou não de *outliers* (valores discrepantes) nos mesmos, sendo especialmente adequado para a comparação de dois ou mais conjuntos de dados correspondentes às categorias de uma variável qualitativa. O *boxplot* é um gráfico que possibilita trabalhar com dois tipos de apreensões: a perceptiva e a discursiva. A apreensão perceptiva, no campo da Educação Estatística, refere-se à interpretação das formas da figura no contexto de resolução de problemas que envolvem a análise de dados estatísticos (Vieira, 2008). A apreensão discursiva é imprescindível na explicitação de propriedades matemáticas. Fazendo uma analogia a Estatística pode-se fazer uma apreensão discursiva no gráfico *boxplot* descrevendo e explicitando suas propriedades, como feito anteriormente.

Para tal discussão, utilizamos na análise desse artigo um conjunto de dados obtidos em uma oficina implementada por Coutinho, Souza (2013) no XI Encontro Nacional de Educação Matemática (ENEM) - Educação Estatística na Prática: analisando dados, gráficos e proporcionando saberes. Dentre as variáveis trabalhadas, selecionamos os dados de dois grupos associando a idade (medida em anos), o número de pessoas residentes na casa em moram, e a massa (expressa em kg).

Tabela 1. Grupo A, idade (anos), número de pessoas residentes, e a massa (kg).

Grupo A	1	2	3	4	5	6	7	8	9
Idade	49	51	21	43	31	26	25	31	26
Nº. de pessoas	1	3	3	4	5	3	5	3	3
Massa (kg)	78	51	68	80	70	68	47	71	54

Fonte: os autores

Tabela 2. Grupo B, idade (anos), número de pessoas residentes, e a massa (kg).

Grupo B	1	2	3	4	5	6	7	8	9
Idade	59	32	31	34	27	35	24	31	31
Nº. de pessoas	3	3	2	4	2	3	2	3	2
Massa (kg)	67	42	60	80	58	75	73	63	55

Fonte: os autores

Nos programas aqui discutidos, R e Geogebra, a inserção dos dados é feita pela digitação direta na tabela do programa; no R, a tabela é elemento do pacote RComander (Rcmdr) e no Geogebra é elemento do menu “exibir”. No software R, a digitação dos dados na tabela do pacote Rcmdr é simples, mas deve-se atentar ao tipo de variável (numérica ou categórica) para assinalar antes da inserção dos dados. No Geogebra, o tratamento que permite a construção de gráficos a partir de uma tabela só é possível para variáveis quantitativas.

No software Geogebra, a construção dos gráficos é feita de acordo com o tipo de análise que se deseja realizar (univariada, bivariada, multivariada). Na análise univariada, há seis tipos de

gráficos disponíveis: histogramas, boxplot, diagramas (de pontos, ramos e folhas, de barras, quantil normal). No software R há mais opções de gráficos, que vão além dos explorados na Escola Básica. Neste trabalho, em particular, analisaremos apenas a construção de dois tipos de gráficos: o diagrama de pontos e o diagrama de caixa (*dotplot* e *boxplot*, respectivamente).

A construção destes gráficos depende do trabalho sobre um conjunto ordenado de valores. Nesse sentido, indicamos que uma primeira introdução seja feita em ambiente “papel e lápis” (ou seja, fora de ambiente computacional), e para isso o conjunto de dados seja adequado, de forma a que o tratamento para a construção dos gráficos não se torne um obstáculo a essa construção. Versões anteriores do Geogebra tinham a manipulação para ordenação dos dados manual, o que muitas vezes tornava vantajosa a ordenação em uma planilha eletrônica, como por exemplo, o Excel, para posterior importação dos dados para a planilha Geogebra. No entanto, versões mais atuais, particularmente a que apresentamos neste texto, já faz a ordenação no próprio comando para construção do gráfico. Logo, no que se refere à construção do gráfico de pontos, a consideração da ordenação dos dados é feita automaticamente nos dois softwares, R e Geogebra.

É necessário assinalar que na construção do gráfico *boxplot* da variável idade de grupo A, no Geogebra e no R não apresentaram diferenças, porém na construção do *boxplot* da variável idade de grupo B, os softwares apresentaram resultados distintos (figura 1) em sua construção. Essas diferenças devem ser analisadas, discutidas e refletidas, em função da possibilidade do uso de formas distintas para determinação dos quartis, conforme aponta Langford (2006).

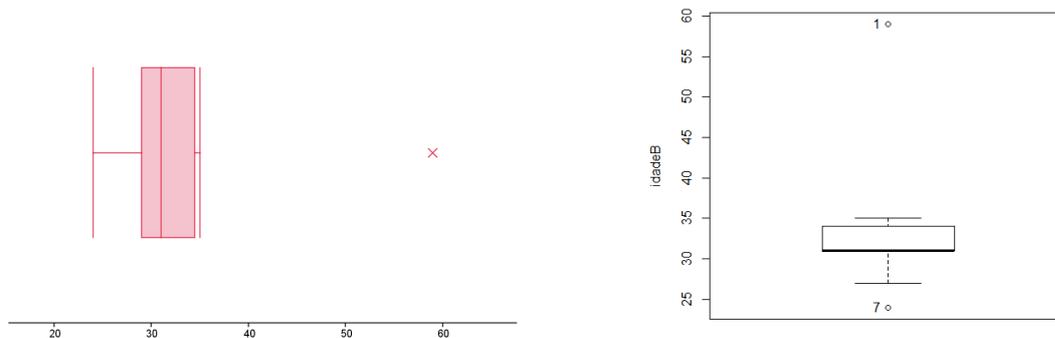


Figura 1. *boxplot* da idade de grupo B, gráfico da esquerda (Geogebra) e da direita (R).

Na análise se evidencia, também, o fato de que no software Geogebra, existe a possibilidade para uma análise univariada de dois tipos de registros na tela, como por exemplo, um *dotplot* e um *boxplot* de uma mesma variável, em um mesmo sistema de eixos, conforme figura 2.

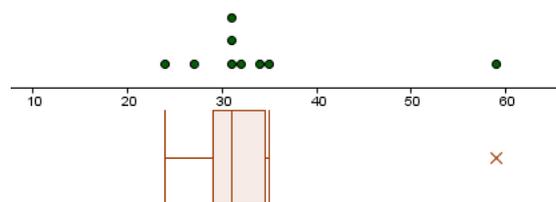


Figura 2. *boxplot* e *dotplot* referentes à variável “idade”, grupo B, Geogebra

Um ponto que não pode deixar de ser destacado refere-se a utilização desses gráficos no mesmo sistema de eixos facilita a apreensão da variação, assim como a análise da forma da distribuição ao modificarmos os valores dos dados: o professor pode modificar tais valores substituindo alguns deles por valores discrepantes (*outliers*), por exemplo. A mudança na tabela do Geogebra gera, automaticamente, uma mudança nos dois gráficos construídos graças ao aspecto dinâmico do programa, valorizando a abordagem dos gráficos estatísticos como início da construção do letramento estatístico. Referimo-nos aqui à construção de habilidades relativas à leitura de dados representados graficamente, relacionando as informações contidas nos gráficos

escolhidos – *dotplot* e *boxplot*, assim como, à descrição desses mesmos dados relacionando-os ao contexto no qual foram coletados.

Quanto à construção desses dois gráficos com o programa R, particularmente utilizando-se o pacote Rcmdr, construímos um gráfico de dispersão, tendo cada variável a associação de um *boxplot* marginal, conforme ilustrado na Figura 3.

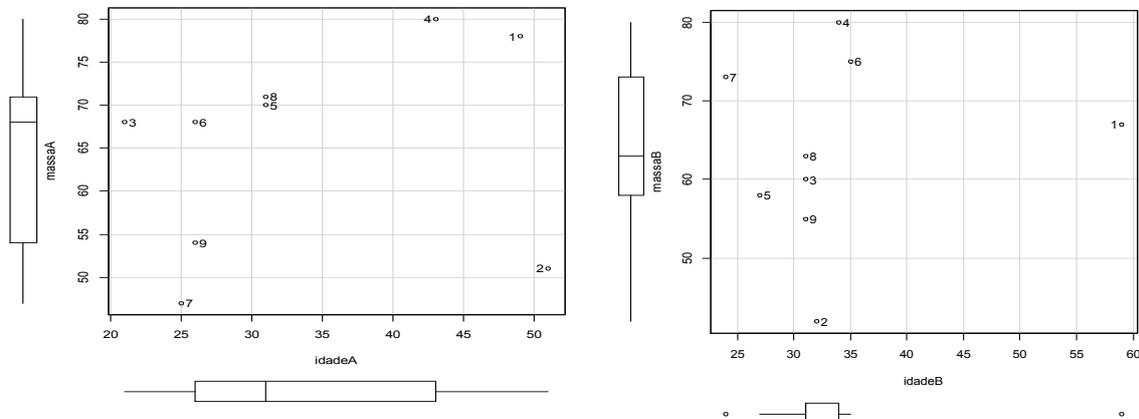


Figura 3. Gráfico de dispersão idade do Grupo A e a massa (kg) do Grupo A com os respectivos *boxplot* marginais construídos no R por meio do pacote Rcmdr.

Verifica-se assim o aumento da complexidade do registro de representação semiótica, o que ocorre sem um correspondente aumento na complexidade da construção. A análise da figura (apreensão a partir da visualização), no entanto, exige do professor maior atenção, já que envolve a apreensão da variabilidade para cada variável e para a distribuição conjunta. É necessário que o aluno mobilize habilidades mais complexas do que aquelas mobilizadas para a análise de gráficos como apresentado nas Figuras 1 e 2, mas tal construção em ambiente computacional é favorecida. O professor pode, dessa forma, oferecer ao aluno uma situação de aprendizagem com complexidade crescente: uma análise unidimensional seguida por uma análise bidimensional.

REFERÊNCIAS

- Batanero, C., Estepa, A., & Godino, J. D. (1991). *Análisis Exploratorio de Datos: sus Posibilidades en la Enseñanza Secundaria*. Suma, n. 9, p. 25-31. Recuperado de <http://www.ugr.es/~batanero>. Consultado em 05/06/2014.
- Ben-Zvi, D., & Garfield, J (eds.) (2004). *The challenge of developing statistical literacy, reasoning and thinking*, p. 6-7.
- Coutinho, C. Q. S., Almouloud, S. Ag., & Silva, M. J. F. (2012) O desenvolvimento do letramento estatístico a partir do uso do Geogebra: um estudo com professores de matemática. In: *Revemat: R. Eletr. de Edu. Matem.* eISSN 1981-1322. Florianópolis, 07(2), 246-265.
- Coutinho, C. Q. S., Silva, M. J. F., & Almouloud, S. Ag. (2011). Desenvolvimento do Pensamento Estatístico e sua Articulação com a Mobilização de Registros de Representação Semiótica. In *Bolema*, Rio Claro (SP), 24(39), 495-514.
- Coutinho, C. Q. S.; Souza, F. S. (2013). Aprendizagem da Estatística e o uso de ambientes computacionais: uma análise didática de programas para construção de gráficos estatísticos. In: *VII Congreso Iberoamericano de Educación Matemática*, Montevideo, Uruguai. p.6221-6228.
- Coutinho, C. Q. S., & Souza, F. S. (2013). Desenvolvimento do letramento estatístico e a leitura e análise de gráficos: uma análise didática sobre a contribuição de ambientes computacionais como R e Geogebra. In *Boletim Labem*, Rio de Janeiro, ano 5, n. 8, p. 8-15, jan/jun.

- Coutinho, C. Q. S.; Souza, F. S. (2013). Educação Estatística na Prática: analisando dados, gráficos e proporcionando saberes. In *Anais do XI Encontro Nacional de Educação Matemática* – ISSN 2178-034X
- Duval, R. (2003). Registros de representações semióticas e funcionamento cognitivo da compreensão da matemática. In Machado (org.), *Aprendizagem em matemática, registros de representação semiótica*. Campinas: Papirus, p.11-33.
- Gal, I. (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25. Recuperado de <http://www.stat.auckland.ac.nz/~iase/cblumberg/gal.pdf>. Consultado 05/06/2014.
- Langford, E. Quartiles in elementary statistics. *Journal of Statistics Education*. v.3, n14, 2006. Recuperado de: http://www.amstat.org/publications/jse/jse_archive.html. Acesso em: 2 março 2015.
- Pfannkuch, M. (2008). Training teachers to develop statistical thinking. In: the *ICMI STUDY 18 and 2008 IASE Round Table Conference. Proceedings*. Recuperado de <http://www.ugr.es/~icmi/iase_study/Files/Topic4/T4P2_Pfannkuch.pdf>. Consultado: 05/06/2014.
- Ponte, P. J.; Brocado, J.; Oliveira, H. (2006). *Investigações Matemáticas na Sala de Aula*. Belo Horizonte: Autêntica.
- Souza, F. S. (2013). A Importância das Atividades para o Ensino de Estatística na Educação Básica com o Uso do Software R e seu Pacote de Dados Rcmdr no III SIPEMAT. In *Perspectivas da Educação Matemática*, Campo Grande, MS, v.6, n. 11, p. 85-92, jan/jun.
- VIEIRA, M. (2008) *Análise Exploratória de Dados: uma abordagem com alunos do Ensino Médio*. Dissertação (Mestrado em Educação Matemática) - Programa de Estudos Pós-graduados em Educação Matemática. Pontifícia Universidade Católica de São Paulo, São Paulo. Disponível em: <<http://www.pucsp.br/pos/edmat>>. Acesso em: 3 mar. 2015.
- Wild, C.; Pfannkuch, M. (1999). *Statistical thinking in empirical enquiry*. *International Statistical Review*, Auckland, v. 6, p. 223–265, 1999.