

Contributos do GeoGebra para análise gráfica de dados em uma formação inicial estatística para professores

The contributions of GeoGebra to graphical data analysis in an initial statistical education for teachers

José Ronaldo Alves Araújo¹
Celina Aparecida Almeida Pereira Abar²
Zsolt Lavicza³

Resumo: Este artigo é parte de uma pesquisa de doutorado em educação matemática, que investigou contributos de recursos digitais na educação estatística. O recorte buscou identificar e analisar possíveis contribuições de ferramentas do GeoGebra que se mostraram favoráveis em atividades de análise gráfica de dados, sob a perspectiva da transnumeração e da visualização de dados. Com aporte metodológico do ciclo investigativo, o estudo foi implementado com futuros professores em uma disciplina de estatística. Qualitativamente, os resultados evidenciaram que as ferramentas de análises univariadas e bivariadas, presentes no GeoGebra, quando integradas às atividades de análise gráfica de dados, ofereceram oportunidades de compreender os problemas subjacentes a esses dados, com suas representações gráficas, promissoras à visualização e à transnumeração dos dados.

Palavras-chave: Análise gráfica de dados. Transnumeração. Visualização de dados. GeoGebra. Futuros professores.

Abstract: This article is part of a doctoral research in mathematics education, which investigated the contributions of digital resources in statistical education. The study aimed to identify and analyze possible contributions of GeoGebra tools that proved favorable in graphical data analysis activities, from the perspective of transnumeration and data visualization. With the methodological support of the investigative cycle, the study was implemented with future teachers in a statistics course. Qualitatively, the results showed that the univariate and bivariate analysis tools available in GeoGebra, when integrated into graphical data analysis activities, provided opportunities to understand the underlying problems in these data through their graphical representations, promising for data visualization and transnumeration.

Keywords: Graphical data analysis. Transnumeration. Data visualization. GeoGebra. Future teachers.

1 Introdução

Este artigo apresenta um recorte de uma tese de doutorado que investigou contributos de recursos digitais no campo da educação estatística. Como salientam Votto (2018) e Santos, Oliveira Júnior, Santos, Carvalho & Saito (2022), no contexto brasileiro, há carência de pesquisas que se dedicam a investigar a utilização de recursos tecnológicos nesse campo. Além disso, Biehler (2018) e McNamara (2018) enfatizam a necessidade de que os recursos tecnológicos se aproximem das problemáticas latentes à educação estatística, sobretudo porque

¹ Universidade Federal de Ouro Preto • Ouro Preto, MG — Brasil • ✉ jronaldoaraujo@gmail.com • ORCID <https://orcid.org/0000-0002-5352-4137>

² Pontifícia Universidade Católica de São Paulo • São Paulo, SP — Brasil • ✉ abarcaaap@pucsp.br • ORCID <https://orcid.org/0000-0002-6685-9956>

³ Johannes Kepler Universität Linz • Linz, UA — Austria • ✉ zsolt.lavicza@jku.at • ORCID <https://orcid.org/0000-0002-3701-5068>

muitos dos recursos não dispõem de atributos que favoreçam essa aproximação.

Como parte das discussões sobre a viabilidade da incorporação de recursos tecnológicos na promoção desse campo, especificamente, considerou-se delimitar o estudo no âmbito dos possíveis contributos de recursos digitais para a análise gráfica de dados. Como Oliveira (2016) e Sera (2016) indicam, a construção de representações e as análises gráficas ainda são desafios no processo de aprendizagem estatística.

O estudo foi realizado em uma disciplina de estatística ofertada para futuros professores e, com base nas oportunidades de investigação, analisou como ferramentas presentes no *Gapminder* (<https://www.gapminder.org>), no *GeoGebra* (<https://www.geogebra.org/>) e no *Statistical Package for the Social Sciences* (SPSS) (<https://www.ibm.com/products/spss-statistics>) podem contribuir em atividades relativas à análise gráfica de dados. No caso deste artigo, apresenta-se tais oportunidades oferecidas pelo *GeoGebra*, indo na direção de estudos como Gonzales e Lopes (2017), Araújo (2018) e Souza e Calejon (2019), que evidenciam possibilidades de implementações desse recurso digital em aulas de estatística.

O objetivo é identificar e analisar possíveis contribuições de ferramentas presentes no *GeoGebra* que se configuram como favoráveis em atividades de análise gráfica de dados por futuros professores em uma disciplina de estatística.

Ao delimitar tal objetivo, neste artigo, busca-se responder à seguinte questão: *Como a utilização do GeoGebra pode contribuir para análise gráfica de dados na disciplina estatística em uma formação inicial de professores?*

Para compreender o meio emergente dessa questão, consideraram-se discussões, a partir de estudos preliminares circundantes à problemática, de modo a conceber o empreendimento do estudo. Na sequência, são delineados referenciais da transnumeração, conforme Wild e Pfannkuch (1999, p. 227, tradução nossa) “[...] um processo dinâmico de mudanças de representações para gerar compreensão”, mais especificamente nos seus estágios (Pfannkuch, Rubick & Yoon, 2002) e da visualização de dados, que “[...] pode ser visto como incluindo análise exploratória de dados assistida por computador de conjuntos de dados volumosos e complexos” (Prodromou & Dunne, 2017, p. 3, tradução nossa), especificamente nos aspectos *compreensão*, *finalidades* e *verbalização*. Posteriormente, é apresentado o referencial metodológico do ciclo investigativo – PPDAC (Problema, Planejamento, Desenvolvimento, Análise e Conclusão), conforme Wild e Pfannkuch (1999), seguido das apresentações do recurso digital *GeoGebra*, do contexto de implementação do experimento e do material de análise. Em seguida, são apresentados os resultados e as respectivas análises, que permitem constituir considerações para o estudo realizado.

2 Estudos preliminares

Em relação aos adventos dos recursos tecnológicos, Lèvy (1993) destacou que as visualizações por meio de imagens em uma tela (representações de dados), bem como as simulações, desempenhariam um papel de comunicação ou de persuasão importante.

Eckert (2018) parte da premissa de que “à medida que a tecnologia digital se torna mais ou menos uma parte natural da vida cotidiana, a pesquisa em educação continuada fornece *insights* sobre como isso pode ser utilizado em sala de aula para aprimorar o aprendizado dos alunos” (Eckert, 2018, p. 1, tradução nossa). Todavia, Woodard (2018) salienta que há poucas pesquisas avaliando como os recursos computacionais para a estatística afetam as habilidades dos alunos para pensar e raciocinar estatisticamente; “[...] como os alunos usam as ferramentas para resolver problemas estatísticos” (Woodard, 2018, p. 1, tradução nossa).

Como aponta Andre (2020), os pacotes de *software* permitem explorações de conjuntos de dados, muitas vezes, por métodos de visualização, de modo que a conexão entre estatística, as visualizações e tecnologia computacional tem sido indissociável. Por exemplo, Mantilla (2018) se refere aos *softwares* que permitem visualizações no estudo de conceitos estatísticos e nas construções de representações interativas, de forma dinâmica e interativa, como tecnologias que podem favorecer o ensino da estatística. As tecnologias da informação e comunicação desempenham um papel significativo, especialmente, porque “[...] permitem ao aluno explorar, conhecer e analisar dados de forma dinâmica e colaborativa” (Mantilla, 2018, p. 1, tradução nossa).

Nesse sentido, Francisco (2016) considera que os processos fundamentais na análise de dados são a descrição, a organização, a redução, a análise e a representação desses dados, de forma a identificar tendências, fazer inferências ou previsões a partir de representações gráficas. Além do domínio sobre conceitos e aplicações de modelos estatísticos para a resolução de situações-problema, de acordo com Giordano (2017, p. 66), “[...] espera-se que os alunos desenvolvam uma postura crítica e reflexiva diante das informações estatísticas que, a nosso ver, se assemelha à postura característica de pesquisadores científicos”. A autora destaca que “é necessário elaborar conjecturas, buscar regularidades, generalização de padrões, ser capaz de argumentar, dentre outras ações que estão intimamente relacionadas à habilidade de resolver problemas” (Giordano, 2017, p. 97).

Souza (2019), ao investigar a interpretação de representações gráficas de dados por professores de matemática, ressalta que o ensino de estatística deve ser conduzido “a partir da realidade social e da valorização da cultura do estudante, [...] para que haja o desenvolvimento das habilidades necessárias e seja favorecida a tomada de decisões diante das questões que se apresentam no contexto de vida dos sujeitos” (Souza, 2019, p. 15).

Souza (2019) considera que, para a interpretação gráfica, a familiaridade do leitor com determinado tema e com a representação do gráfico são aspectos que influenciam as interpretações em relação aos dados. Bolch e Jacobbe (2018) salientam que, ao ancorar problemas envolvendo dados do mundo real, tem-se uma abordagem propícia aos objetivos de ensino, pois “[...] requer que os alunos apliquem habilidades de pensamento crítico para responder questões de pesquisa usando esses dados” (Bolch & Jacobbe, 2018, p. 2, tradução nossa).

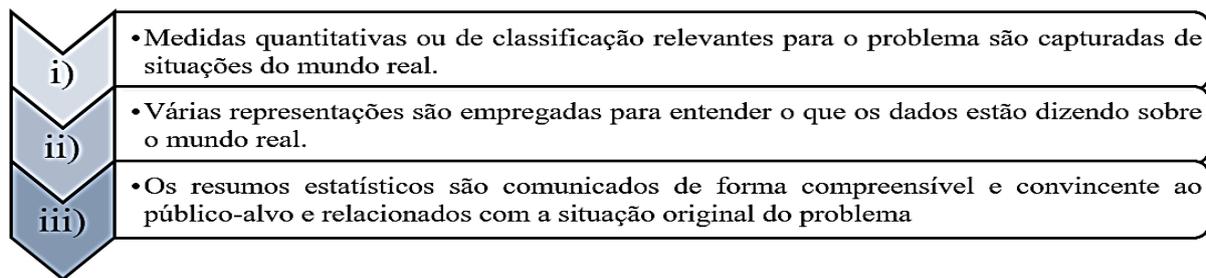
3 Referenciais teóricos

Fazem parte dos referenciais teóricos deste estudo a transnumeração, discutida por Wild e Pfannkuch (1999), delimitada pelos seus estágios, definida por Pfannkuch *et al.* (2002), além da proposição de Prodromou e Dunne (2017) sobre aspectos relativos à *compreensão*, às *finalidades* e à *verbalização* na visualização de dados.

3.1 Transnumeração

A transnumeração se refere às “transformações de numeracia feitas para facilitar a compreensão” (Wild & Pfannkuch, 1999, p. 227, tradução nossa). Consiste em um tipo fundamental de pensamento estatístico que, para Pfannkuch *et al.* (2002), pode ser observada como um processo em três estágios, como mostra a Figura 1:

Figura 1: Estágios de transnumeração



Fonte: Adaptado de Pfannkuch *et al.* (2002).

Conforme Pfannkuch *et al.* (2002), como mostra a Figura 1, no primeiro estágio, o foco está na captura de qualidades ou características da situação real; no segundo estágio, para que sejam apresentadas representações de dados, o foco está na necessidade de os dados serem coletados, transformados de brutos em múltiplas representações gráficas, resumos estatísticos, de modo a permitir a obtenção de significado da situação real a partir deles; no terceiro estágio, dada a obtenção de significados, o foco reside na comunicação a partir dos dados e julgamentos de forma a serem compreendidos, em termos da situação real investigada, pelo público-alvo.

Neste estudo, os estágios de transnumeração permitiram analisar o percurso dos participantes na *Homework* proposta, de modo a elucidar as possíveis contribuições do GeoGebra.

3.2 Visualização de dados

Para Prodromou e Dunne (2017), a visualização de dados, no campo da educação estatística, engloba a análise exploratória de dados assistida por computador, o que possibilita explorar e comunicar efetivamente informações relevantes por meio de representações gráficas. Para esses autores, “[...] a grande importância da disponibilidade de *software* e da popularidade associada, na determinação de quais análises são realizadas e como elas são apresentadas, é um tópico de grande relevância” (Prodromou & Dunne, 2017, p. 2, tradução nossa). Relacionados ao papel da visualização, destacam três aspectos: a *compreensão*, as *finalidades* e a *verbalização*.

A *compreensão*, no contexto de visualização, inclui descrições estáticas e dinâmicas de características multidimensionais. “[...] a visualização envolve a produção de uma representação visual de um ou mais aspectos de um conjunto de dados” (Prodromou & Dunne, 2017, p. 4, tradução nossa). Para a *compreensão* mediada por recursos de visualização, “[...] o usuário acessa processos cognitivos rápidos que permitem o reconhecimento de características estruturais insuspeitas ou latentes de um corpo detalhado de números, cadeias ou símbolos. Em consequência, a visualização faz parte da *compreensão* e integra-se na sua comunicação” (Prodromou & Dunne, 2017, p. 4, tradução nossa).

As *finalidades* da visualização podem oferecer suporte para a produção de significados do contexto e a interpretação por meio de estratégias cognitivas disponíveis (Prodromou & Dunne, 2017). Alguns pontos a serem considerados:

- Descrever e comparar condições ou estados observados em um contexto (em um instante).
- Descrever e avaliar associações entre categorias, contagens e medidas (muitas vezes com aspectos de tempo ignorados).
- Descrever e comparar mudanças ou processos atuais em um contexto (ao longo de um período, às vezes com intervalos de interobservação iguais).

- Descrever e avaliar associações entre mudanças nas variáveis observadas (ao longo de alguns comprimentos de intervalo de tempo implícitos ou especificados) (Prodromou & Dunne, 2017, p. 4, tradução nossa).

Para as condições de análise de dados, “[...] a fluência com a visualização implica uma consciência das possíveis características dos dados que são marginalizadas por qualquer visualização em particular” (Prodromou & Dunne, 2017, p. 4, tradução nossa). “As crenças, inferências e visualizações também podem informar ou permitir *insights* retrospectivos sobre a causalidade associada aos fenômenos que deram origem aos dados e à(s) representação(ões) produzida(s) pelo aluno” (Prodromou & Dunne, 2017, p. 5, tradução nossa).

Prodromou e Dunne (2017) consideram a *verbalização* como um complemento e resultado da visualização. Em um contexto de aprendizagem no qual se considera a visualização, é esperado que emergam comentários, *insights* ou perguntas. Para os autores, as *verbalizações* precipitadas, no contexto de estudo de dados, devem atender aos critérios de explicitações, coerência e precisão de natureza como descrições de variabilidade e pressupostos articulados, necessários para a extrapolação.

4 Referencial e procedimentos metodológicos

Os procedimentos metodológicos para criar um ambiente favorável à proposta da pesquisa tomaram como referencial a dimensão do ciclo investigativo presente no *framework* de Wild e Pfannkuch (1999). Neste recorte, exploram-se possíveis contributos da integração do recurso digital GeoGebra neste ambiente, no qual participaram alunos de diferentes graduações de formação de professores, como matemática, física, geografia etc., matriculados na disciplina de estatística.

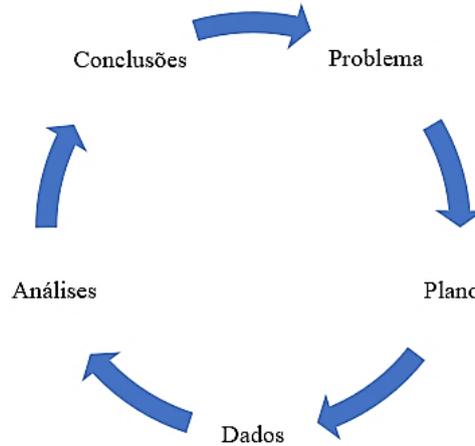
Emergentes da implementação da pesquisa, à luz de pressupostos da abordagem qualitativa, em conformidade com Godoy (1995), os relatórios desenvolvidos pelos futuros professores, ao longo de atividades propostas a disciplina de estatística, foram concebidos como *corpus* de análise. Tais análises têm como fim de elucidar possíveis contributos da integração do recurso digital GeoGebra aos participantes, quanto à visualização de dados e aos processos de transnumeração.

4.1 Ciclo investigativo - PPDAC

Como propõe o *framework* de Wild e Pfannkuch (1999), em atividades estatísticas, o ciclo investigativo permite uma organização para promover a aprendizagem da estatística. O ciclo descreve o processo de investigação estatística, que, eventualmente, pode ser uma possibilidade para questionar, sob alguma medida, algo no mundo. Wild e Pfannkuch (1999) apontam que um ciclo investigativo pode ser definido para atingir objetivos de aprendizagem e o conhecimento adquirido e as necessidades identificadas nesses ciclos podem iniciar novos ciclos investigativos

Como mostra a Figura 2, o ciclo PPDAC é composto por fases. É possível identificar no comportamento das setas que o conhecimento adquirido e as necessidades identificadas em um ciclo podem iniciar novos ciclos investigativos.

Figura 2: Ciclo Investigativo - PPDAC



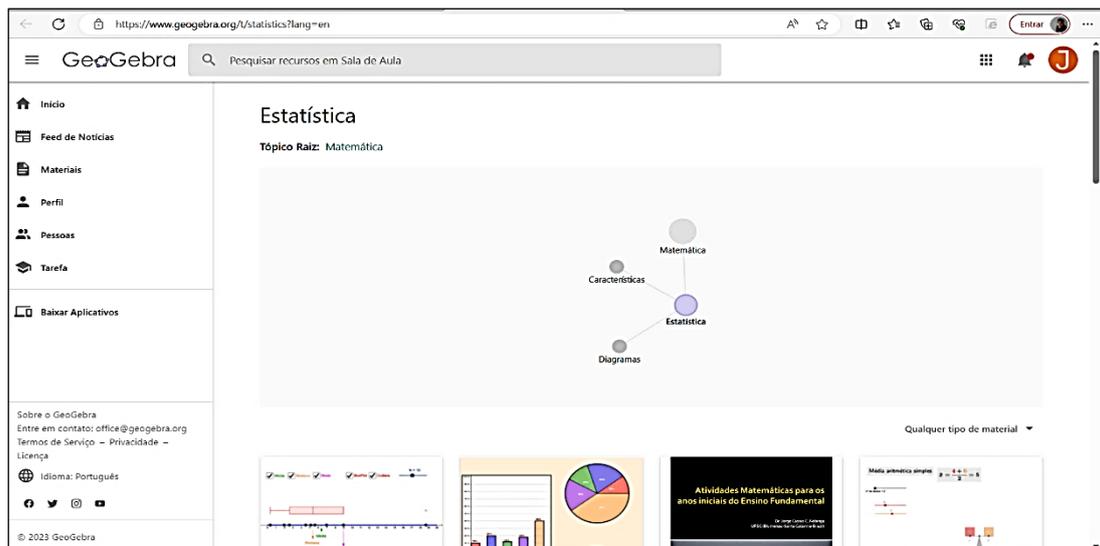
Fonte: Adaptado de Wild e Pfannkuch (1999, p. 226).

Orientada pelo ciclo investigativo proposto por Wild e Pfannkuch (1999), a parte experimental apresentada neste estudo envolveu a integração do GeoGebra, visando elucidar elementos que apontam para o objetivo definido neste recorte. Os resultados relativos à utilização do GeoGebra pelos participantes são, então, explicitados.

4.2 O GeoGebra

O recurso digital GeoGebra consiste em uma ferramenta dinâmica de matemática que abrange tópicos como geometria, álgebra, estatística (Figura 3), probabilidade em uma única plataforma, permitindo trabalhar com planilhas e construir representações gráficas de dados. O *software* é de acesso gratuito e possui versões que podem ser utilizadas *online* ou *offline*, sendo acessível também em dispositivos móveis. Além disso, o GeoGebra oferece uma plataforma *online* com mais de 1 milhão de recursos gratuitos criados pela comunidade em vários idiomas.

Figura 1: Tela do GeoGebra para exploração de ferramentas estatísticas



Fonte: GeoGebra, 2024.

De acordo com as fases do ciclo investigativo – PPDAC, a solicitação de explorar as ferramentas estatísticas do GeoGebra (<https://www.geogebra.org/t/statistics?lang=en>) teve como objetivo investigar, nas atividades desenvolvidas pelos futuros professores, contribuições do recurso digital para análise gráfica de dados.

4.3 A implementação do experimento e o material de análise

A implementação da pesquisa ocorreu durante o estágio doutoral realizado pelo primeiro autor, na Johannes Kepler Universität, Áustria, entre outubro de 2021 e janeiro de 2022. As atividades da pesquisa, denominadas de *Homework*, foram desenvolvidas por 72 alunos (futuros professores) em formação inicial de cursos diferentes de licenciatura, durante uma disciplina de estatística, e contou com a exploração das ferramentas estatísticas disponíveis no *Gapminder*, no GeoGebra e no SPSS. Por estar em um contexto de aulas em inglês, todas as atividades foram desenvolvidas nesse idioma.

Figura 4: Atividade proposta com o GeoGebra

Homework 2

Explorações Estatísticas com GeoGebra

1. Explore algumas ferramentas estatísticas do GeoGebra, semelhantes às apresentadas nas aulas: <https://www.geogebra.org/t/statistics?lang=en>.
Alguns bons exemplos: <https://www.geogebra.org/m/g3vLRuB8>
2. Encontre alguns conjuntos de dados *online* que são interessantes para você, aqui estão alguns sites úteis: <https://www.gapminder.org/tools/>
<https://covid19.who.int/>
<https://www.statista.com/statistics/1109188/coronavirus-confirmed-cases-by-age-group-austria/>
<https://data.humdata.org/event/covid-19?>
<https://www.data.gv.at/covid-19/>
3. Mostre algumas visualizações com o GeoGebra e apresente algumas explicações, do mesmo modo que você fez com o *Gapminder*.

Fonte: Acervo da pesquisa.

Neste trabalho, a *Homework* (Figura 4) e seus resultados são apresentados de forma traduzida para o idioma português. Constam os resultados observados a partir do relatório produzido por uma das duplas de futuros professores, Sarah e Ian (nomes fictícios).

A escolha de bancos de dados e suas explorações, por meio da utilização do GeoGebra, faz sentido se houver um problema a ser abordado com base nesses dados. Sob o ponto de vista metodológico, o experimento colocou os participantes em uma situação de investigação, seguindo o ciclo investigativo – PPDAC. Era esperado que os grupos explorassem diferentes fontes de dados sobre diversas problemáticas e escolhessem algumas delas, o que correspondia à primeira fase do ciclo investigativo.

A fase de planejamento, a segunda do ciclo investigativo - PPDAC, os participantes foram responsáveis por delimitar amostras de dados, gerenciar essas informações, definir unidades de medida e realizar análises para compreender o problema. O processo de gerenciamento e limpeza dos dados e sua importação para o GeoGebra caracterizava a terceira fase do ciclo investigativo, o trabalho com os dados. Esse trabalho, já com os dados importados na planilha de cálculo, envolvia a escolha de ferramentas do GeoGebra que possibilitassem a construção de representações gráficas e a extração de estatísticas desses dados.

O objetivo desses procedimentos era gerar uma interface do GeoGebra com janelas que possibilitassem a realização de análises detalhadas dos dados, correspondendo à quarta fase do ciclo investigativo. A partir dessas análises, as conclusões, características da quinta e última fase do ciclo investigativo, seriam formuladas com base nas representações e estatísticas extraídas através das funcionalidades do GeoGebra.

Com os delineamentos metodológicos estabelecidos, que permitiram o desenvolvimento da atividade mediada pelo GeoGebra, chamada *Homework 2*, seguem-se os resultados observados nos relatórios produzidos pela dupla de futuros professores, Sarah e Ian. As análises subsequentes são fundamentadas na teoria da transnumeração proposta por Wild e Pfannkuch (1999), Pfannkuch *et al.* (2002), e na visualização de dados segundo Prodromou e Dunne (2017).

5 Resultados e análises

O relatório da dupla não apresentou detalhes sobre as primeiras explorações realizadas no GeoGebra ou sobre a experiência com esse recurso tecnológico digital. No relatório, Sarah e Ian delimitaram dois tópicos abordados: o primeiro relativo à altura média e ao peso médio da população mundial por sexo; e o segundo relacionado ao peso médio da população mundial por sexo. Para este estudo, destaca-se, ao longo desta seção, recortes das análises dos dados relativos aos dois tópicos escolhidos pela dupla de futuros professores.

5.1 Altura média e peso médio da população mundial por sexo

O problema a ser abordado, com o estabelecimento das variáveis, evidenciou uma delimitação característica da primeira fase do ciclo investigativo. A dupla destaca:

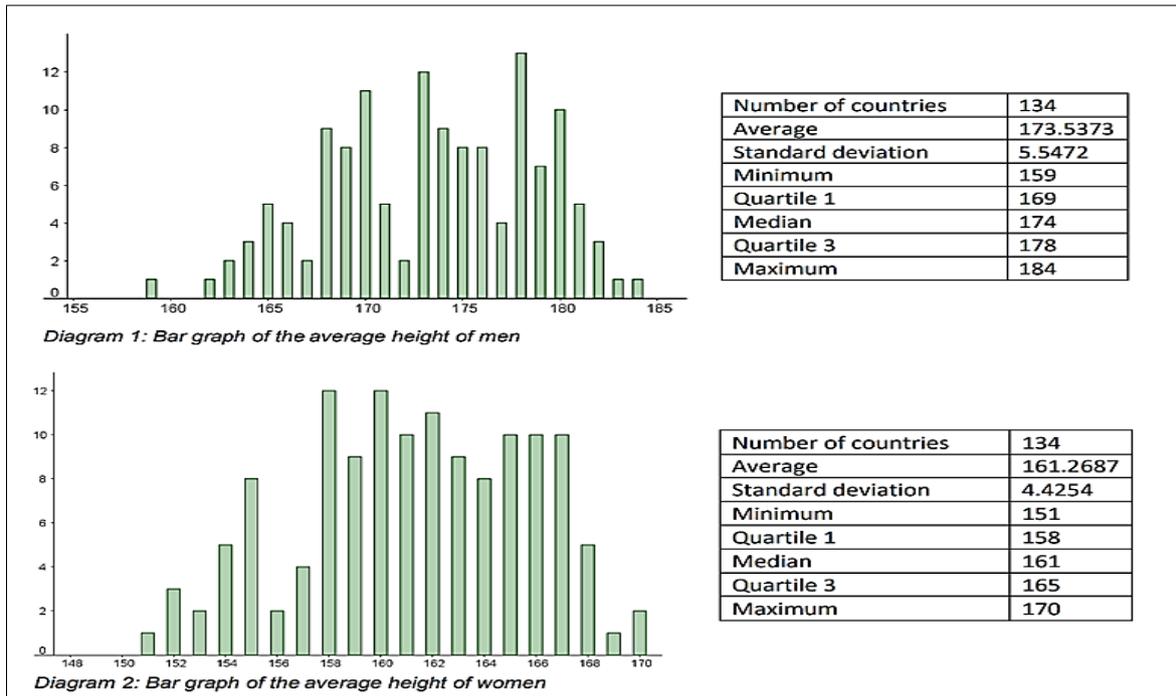
Para o primeiro tópico, escolhemos tratar sobre a altura média por sexo da população humana. As datas utilizadas para nossas visualizações foram encontradas neste link "<https://www.laenderdaten.info/durchschnittliche-koerpergroessen.php>" Estes são os valores médios dos países. (Recorte 1)

Sarah e Ian apresentaram os resultados das manipulações realizadas com os dados, os quais envolveram procedimentos de escolha de amostra, observação das unidades de medida e análises, que são parte da gestão dos dados, segunda fase do ciclo investigativo e necessárias no primeiro estágio de transnumeração. Embora esses procedimentos não estejam explicitamente evidenciados no relatório, é importante reconhecer que foram realizados.

Os procedimentos de importação dos dados para a planilha de cálculo do GeoGebra, limpeza e organização dos dados são essenciais para o trabalho com os dados, terceira fase do ciclo investigativo. Como resultado desses procedimentos, as possibilidades de representações gráficas foram utilizadas para compreender os dados e suas relações com o problema em questão, o que corresponde ao segundo estágio de transnumeração, de acordo com Pfannkuch *et al.* (2002). Essas representações gráficas, sob o ponto de vista da visualização de dados, no sentido de Prodromou e Dunne (2017), podem ser consideradas contribuições do GeoGebra para a análise gráfica dos dados sobre a altura média da população mundial por sexo.

As primeiras representações gráficas que Sarah e Ian exploraram para o estudo da altura média da população mundial por sexo foram os gráficos de colunas e as estatísticas representados na Figura 5.

Figura 2: Altura média mundial por sexo de 134 países



Fonte: Acervo da pesquisa.

As representações gráficas, presentes na Figura 5, mostraram as alturas médias de acordo com o sexo e a frequência de países, visualizadas nos eixos horizontal e vertical, respectivamente. Como a variável altura média é contínua, considera-se que uma representação gráfica adequada no estudo realizado pela dupla seria o histograma. A partir da representação do gráfico de colunas, a dupla introduziu algumas análises:

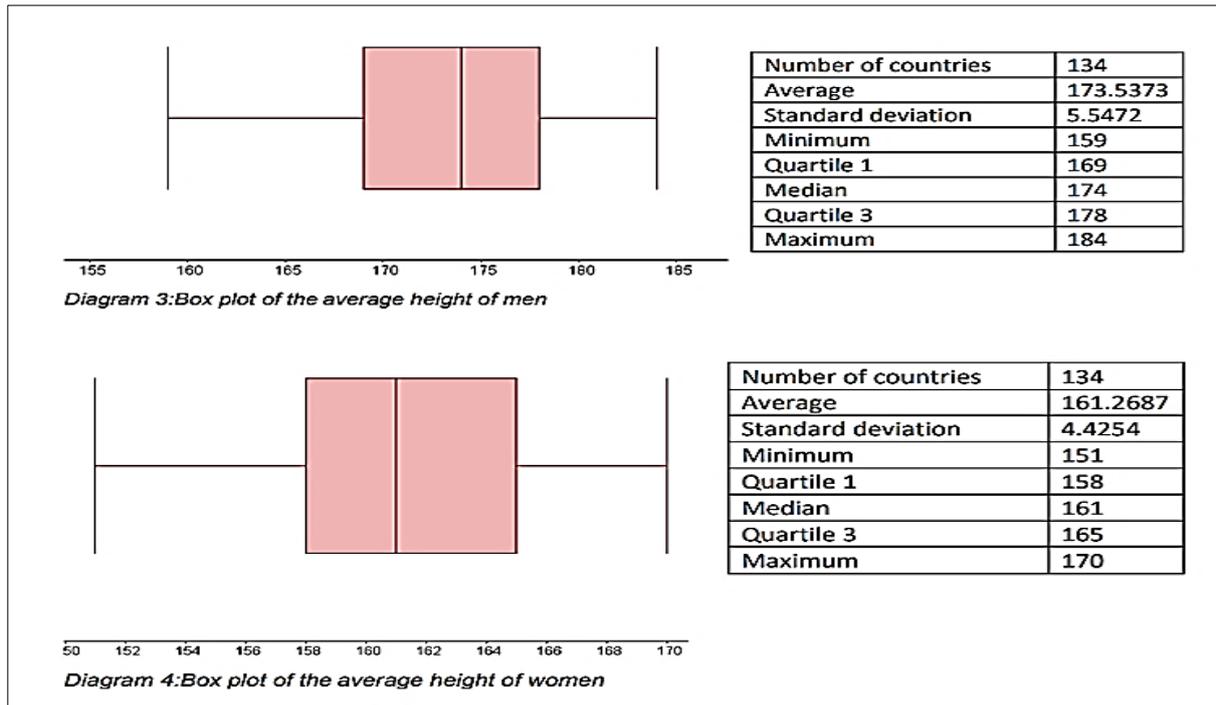
Estes dois diagramas mostram a altura média de mulheres e homens de 134 países. Para os homens, o menor valor é de 159 cm, apenas alguns países têm essa altura média. No caso das mulheres, o menor valor é de 151 cm. A altura média dos homens é de 173,54 cm, isso é mostrado no diagrama 1. Este valor é maior do que a altura média das mulheres, que é de 161,27 cm e pode ser visto no diagrama 2. A altura média da população feminina e masculina foi determinada calculando-se uma média dos valores médios dos homens e mulheres de 134 países. A mediana da altura dos homens é de 174 cm e das mulheres é de 161 cm. (Recorte 2)

É relevante apontar que, embora seja possível observar a frequência de países relativa a cada valor de média das alturas, não é possível, pela representação gráfica dos dados, visualizar quais países estão relacionados em cada coluna de frequência. Todavia, a representação pode ser relevante para compreensão de como essas alturas médias da população humana varia a depender do sexo e de cada país.

Sarah e Ian, ao destacarem valores como mínimos, médias e medianas, permitiram identificar, nas suas análises, a apresentação de comparações entre esses valores relativos a cada sexo. A descrição é resultado do que é possível visualizar na janela gráfica e nas estatísticas apresentadas pelo GeoGebra, que caracterizam contribuições da tecnologia digital para a realização dessa análise inicial da dupla de futuros professores, a quarta fase do ciclo investigativo.

Outra representação gráfica que Sarah e Ian exploraram foi o *box-plot* (Figura 6). A dupla utilizou os dados dos 134 países, representados em quartis e, por meio dessa representação, observam a dispersão das médias de altura.

Figura 6: *Box-plot* das alturas médias da população, por sexo, de 134 países



Fonte: Acervo da pesquisa.

A visualização oferecida pela representação do *box-plot* para os dados dos 134 países permitiu que a dupla ampliasse a análise descritiva apresentada a partir da representação do gráfico de colunas.

No box-plot podemos ver que a mediana para os homens é de 174 cm e para as mulheres de 161 cm. O valor do primeiro quartil é de 158 cm e do terceiro quartil de 165 cm. (Recorte 3)

É possível identificar, por meio da representação gráfica do *box-plot*, as diferenças entre as médias de altura entre os sexos. A representação do diagrama *box-plot* contribuiu para que a dupla de futuros professores verificasse como os dados, divididos em quartis, estão distribuídos. Além da representação gráfica, a exposição das estatísticas das alturas médias foi relevante para possibilitar a compreensão dos dados explorados sobre essa variável.

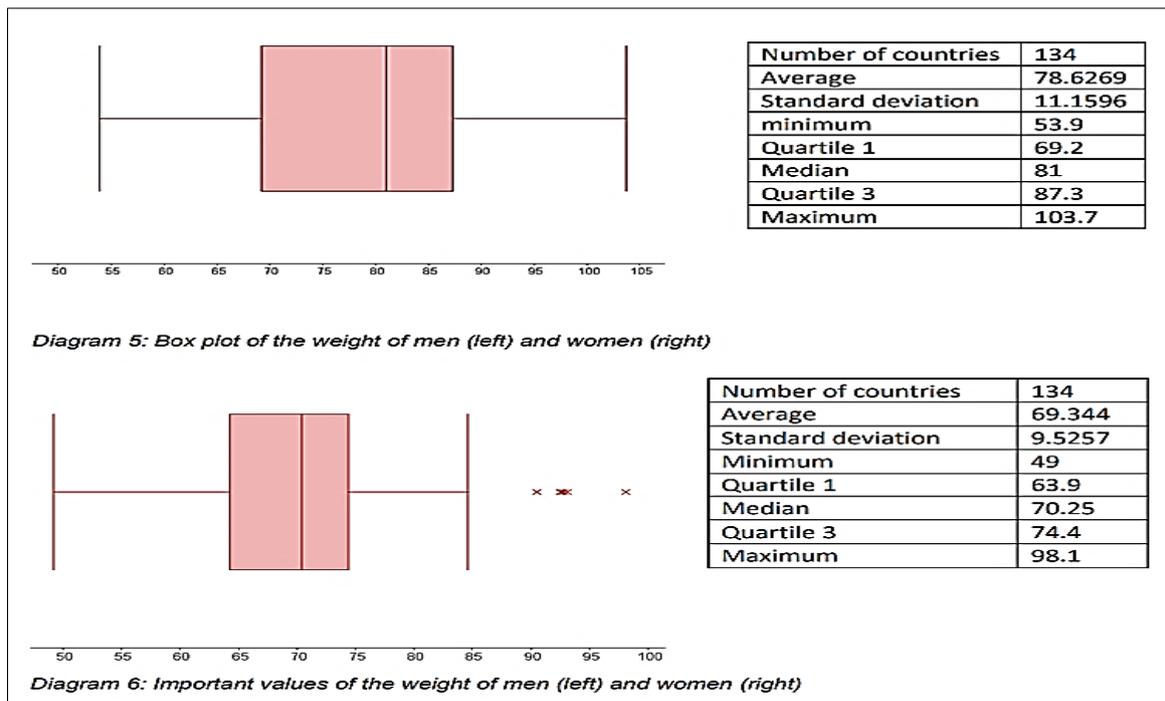
Sob o ponto de vista da transnumeração, as estatísticas capturadas inicialmente colaboraram para um processo transnumerativo em um primeiro estágio, enquanto a admissão de algumas representações, gráfico de colunas e o diagrama *box-plot*, disponibilizadas pela utilização do GeoGebra, favoreceram o estabelecimento de processos de transnumeração em um segundo estágio. Como se observa, os resultados do trabalho com os dados permitiram gerar uma compreensão de como a altura média da população humana pode variar de acordo com o sexo e o país.

Ao considerar a visualização de dados, as representações do gráfico de colunas sobre a frequência de países, com diferentes médias de altura em ambos os sexos, e da representação do *box-plot*, em quartis, é possível compreender como essas médias de altura da população humana se distribuem. Tais representações se mostraram contributos para a *compreensão* dos dados. Nesse contexto, o GeoGebra desempenhou um papel importante (*finalidade*), pois não apenas facilitou a construção dessas representações, mas também destacou suas funções analíticas. Como explicitado, as representações gráficas e as estatísticas obtidas foram relevantes na *verbalização* do discurso de Sarah e Ian.

5.2 Peso médio da população mundial por sexo

Em relação ao segundo tópico de interesse de Sarah e Ian, como no primeiro tópico, as fases de planejamento e de trabalho com os dados, respectivamente, segunda e terceira fases do ciclo investigativo – PPDAC ocorreram. Sob a perspectiva dos estágios de transnumeração, a seleção de amostras e a definição das unidades de medida são etapas relevantes, que certamente foram alcançadas pela dupla de futuros professores. A figura 7 mostra a representação gráfica visualizada pela dupla, resultado da manipulação dos dados no GeoGebra.

Figura 3: Peso médio da população mundial por sexo



Fonte: Acervo da pesquisa.

A partir da visualização gráfica, apresentada na Figura 7, Sarah e Ian realizaram uma análise que descreve os resultados obtidos em relação à média de peso da população por sexo nos 134 países, remetendo à quarta fase do ciclo investigativo.

Graças aos dois box-plot pode-se ver o peso médio mundial das mulheres e dos homens. O peso médio dos homens é de 78,63 kg e das mulheres é de 69,34 kg. O desvio padrão do box-plot mostrando a distribuição do peso médio dos homens é de 11,16 e das mulheres é de 9,53. Pode-se também perceber que existem alguns valores atípicos no diagrama 6). (Recorte 4)

A análise realizada por Sarah e Ian sobre o peso médio da população dos 134 países é descritiva e se enquadra na análise univariada, assim como no tópico anterior. Ao atribuir um valor médio para o peso de cada sexo, a dupla considerou valores representativos para toda a população desses 134 países. Os valores do desvio padrão para os dados de cada sexo permitiram a eles explicitarem a variação do peso médio observada nos diferentes países em relação aos valores médios calculados na atividade. A representação do *box-plot* possibilitou aos futuros professores identificarem possíveis valores atípicos presentes nos dados. A representação gráfica do diagrama de *box-plot* viabilizou a compreensão sobre o peso médio da população humana apontado pelos dados.

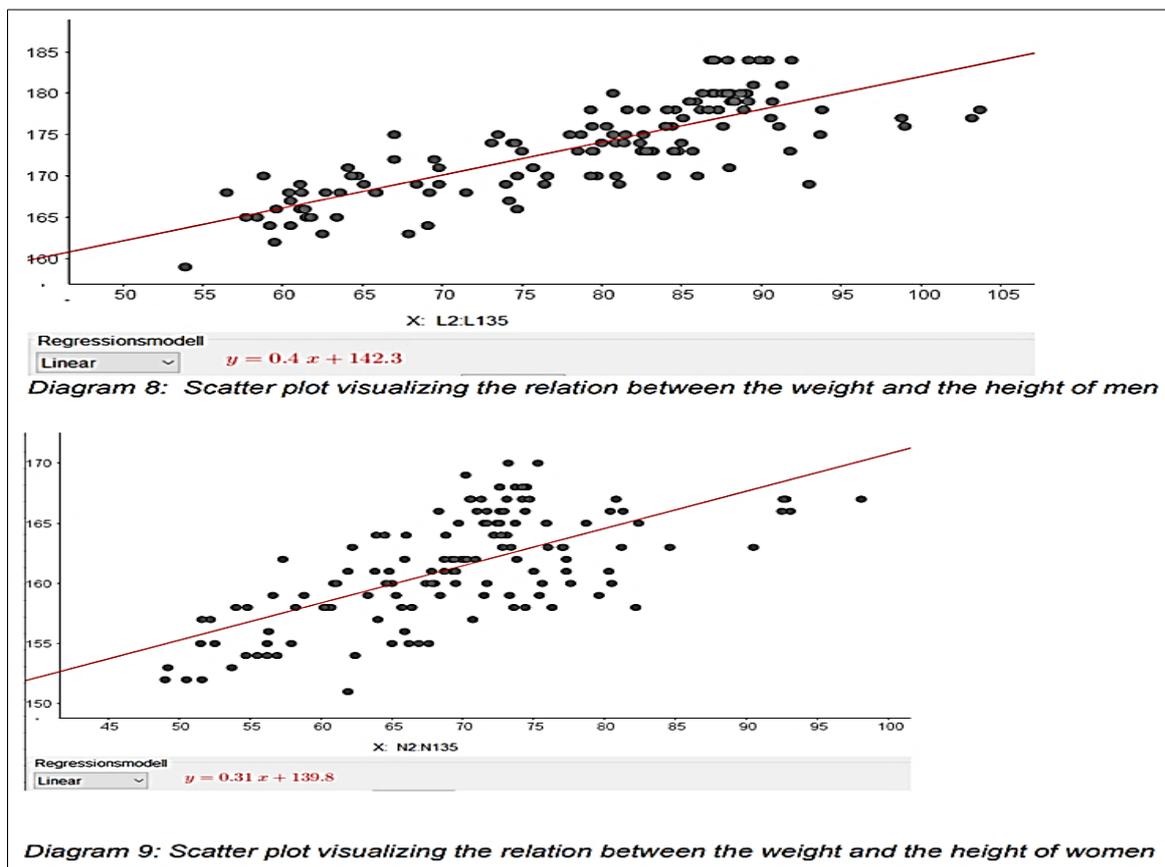
As estatísticas apresentadas e as representações gráficas do *box-plot*, referenciadas no relatório, indicam a compreensão da dupla de futuros professores acerca de como os dados

sobre o peso médio se distribuem e como os valores encontrados resumem a situação estudada, cumprindo a finalidade da representação gráfica.

Ao visualizar como os dados se distribuem em quartis e ao identificar pontos que representam valores atípicos no conjunto de dados estudado, assim como ao considerar a descrição presente na análise desses valores, é possível reconhecer outra contribuição da representação gráfica do *box-plot*, resultante da utilização do GeoGebra na atividade. Além disso, as medidas de média e desvio padrão, consideradas pela dupla de futuros professores, são relevantes para questionar seus significados em relação à complexidade subjacente, mesmo antes do cálculo das médias dos pesos das populações por sexo de cada país. Ou seja, houve uma reflexão sobre o conhecimento do contexto de origem dos dados.

Ao analisar separadamente as variáveis altura média e peso médio por sexo da população dos 134 países, Sarah e Ian realizaram análises descritivas sobre cada uma delas, obtendo suas estatísticas e representações gráficas possibilitadas pela utilização do GeoGebra. Esse processo de análise, característico da quarta fase do ciclo investigativo, foi ampliado pela percepção da dupla de futuros professores, que considerou a possibilidade de existir alguma relação entre a altura média e o peso médio por sexo observados nos diferentes países.

Figura 4: Relação entre a altura média mundial e o peso por sexo



Fonte: Acervo da pesquisa.

Com a utilização da ferramenta de análise bivariada do GeoGebra, representada na Figura 8, a dupla obteve o resultado que relaciona as duas variáveis: altura média e peso médio por sexo. É importante ressaltar que a representação gráfica pode ser interpretada como uma representação da correlação entre as variáveis, e a reta exibida, obtida a partir do modelo linear disponível no GeoGebra, representa um estudo de regressão linear simples para essas duas variáveis. Diante dos resultados, a dupla de futuros professores destacou:

O diagrama 8 mostra a relação entre o peso e a altura dos homens e o diagrama 9 das mulheres. Em ambos os diagramas, os valores dos 134 países foram aproximados por uma equação linear. Ambos os diagramas confirmam que geralmente pode ser afirmado, que quanto mais alta uma pessoa é, mais ela pesa. A variância da reta de regressão mostrada no diagrama 8 é de 0,64 e no diagrama 9 é de 0,44, o que significa que a linha de regressão que se aproxima dos valores dos 134 países dos homens é melhor que a das mulheres, pois apresenta uma variância maior. (Recorte 5)

Os apontamentos de Sarah e Ian, baseados nos resultados da análise bivariada representada na janela gráfica do GeoGebra, evidenciaram considerações gerais sobre a relação entre altura média e peso médio da população dos 134 países. A dupla de futuros professores identificou que a reta de regressão linear é um modelo que aproxima a altura média do peso médio, por sexo, da população dos 134 países. Como é possível observar na representação da reta, em ambos os sexos, faz sentido afirmar que, geralmente, quanto mais alta uma pessoa é, mais ela pesa.

Os valores 0,64 para o sexo masculino e 0,44 para o sexo feminino, concebidos pela dupla e apresentados como variância nas estatísticas do GeoGebra, deveriam ser considerados como coeficientes de determinação da reta de regressão sobre o peso médio e a altura média. A interpretação desses resultados foi evidente quando Sara e Ian afirmaram que “a reta de regressão que se aproxima dos valores dos 134 países dos homens é melhor que a das mulheres, pois apresenta uma variância maior”.

A representação gráfica resultante da análise bivariada realizada pela dupla de futuros professores, utilizando os dados dos 134 países em relação às variáveis altura média e peso médio por sexo, explicitou outra possibilidade de contribuição do GeoGebra para a análise gráfica de dados.

Como analisado, a partir do relatório da dupla, a transnumeração, conforme Wild e Pfannkuch (1999), ocorreu a partir dos resultados obtidos pela utilização do GeoGebra na análise bivariada, que permitiu compreender a existência de diferenças na variabilidade dos dados, dependendo do sexo, e que, embora possam ser modelados por uma regressão linear simples, um modelo se mostrou mais adequado do que o outro, devido a essa diferença na variabilidade dos dados.

Ao considerar o referencial de Pfannkuch *et al.* (2002), a escolha de representar a relação entre peso médio e altura média da população dos 134 países como uma regressão linear simples significou ações correspondentes ao segundo estágio de transnumeração, no qual se observou a contribuição da utilização do GeoGebra para que a dupla de professores apontasse considerações sobre como os dados das variáveis estão relacionados.

Ao tecerem considerações sobre o modelo de regressão linear simples visualizado, se observa que, para a *compreensão* da dupla, quando afirma que “quanto mais alta uma pessoa é, mais ela pesa”, evidencia-se uma *finalidade* atendida da representação gráfica da linha de regressão visualizada. A afirmativa sobre a representação do modelo de regressão linear simples em relação aos dados é, explicitamente, a *verbalização* daquilo que Sarah e Ian visualizaram diante do resultado da análise bivariada mediada pelo GeoGebra.

Em relação à fase de conclusão do ciclo investigativo – PPDAC, no *Homework 2*, as análises apresentadas nos dois tópicos colaboraram para elucidar como há uma variação em relação à altura média e o peso médio da população em 134 países. Em ambos os tópicos, a percepção sobre essas diferenças foi mediada pela utilização do GeoGebra, que possibilitou a construção de representações gráficas e o acesso às estatísticas dos dados relativos às variáveis estudadas.

Para concluir a quinta fase do ciclo investigativo, além dos resultados e como esses emergiram em suas análises interpretativas das representações gráficas dos dados, novas ideias surgiram, de modo que a dupla de futuros professores buscou encontrar explicações para as diferentes alturas médias e pesos médios.

Os desvios podem resultar de má nutrição. A assistência médica é, portanto, também importante. O tamanho do corpo pode ser influenciado por doenças e alergias. Nos países pobres, a altura é menor do que nos países mais ricos. Os genes determinam a altura que você pode alcançar. As famílias em que as pessoas são mais altas transmitirão estes genes à próxima geração. As pessoas estão ficando cada vez maiores. Nos últimos 100 anos, as pessoas cresceram cerca de 15 cm. (Recorte 6)

Na tentativa de explicar as diferenças observadas nos dados, a dupla apresentou novas ideias que levaram em consideração aspectos do contexto de origem dos dados. Considerando o terceiro estágio de transnumeração proposto por Pfannkuch *et al.* (2002), a comunicação sobre os resultados observados pela dupla de futuros professores ocorreu de forma que, em cada tópico, ao analisar os dados, as afirmações se mostraram como resultados convincentes sobre os dados estudados.

Um apontamento final do estudo realizado pelos futuros professores Sarah e Ian, ao dissertarem sobre possíveis causas de haver diferenças relevantes em relação aos dados de cada país, emerge como uma tentativa de, para além de apresentar resultados estatísticos, problematizar o porquê de eles serem como são. O sentido do problema abordado pela dupla revela, possivelmente, problemas sociais e econômicos ao redor do mundo, razões para o empreendimento de novos ciclos investigativos.

6 Considerações

A utilização do GeoGebra contribuiu para que a dupla de futuros professores investigasse a altura média e peso médio da população humana por sexo e a relação entre essas variáveis, tendo como referência valores médios observados em 134 países. Retomando o objetivo de pesquisa delimitado para o recorte apresentado neste artigo, a partir do ciclo investigativo trilhado pela dupla, a contribuição do GeoGebra pode ser identificada nas possibilidades de construção de representações do gráfico de colunas e do *box-plot*, resultantes das análises univariadas e do acesso às estatísticas descritivas relativas aos dados estudados. Da mesma forma, ao investigar possíveis relações entre as variáveis, a representação do gráfico de regressão linear simples, possibilitada pela manipulação dos dados no GeoGebra, permitiu que a dupla de futuros professores chegasse a uma conclusão sobre a possível relação existente entre os dados. As representações do gráfico de colunas, do diagrama *box-plot* e da linha de regressão linear simples possibilitadas pelo GeoGebra favoreceram que a dupla de futuros professores visualizasse os dados e transnumerasse sobre eles.

Ao considerar a questão delimitada neste recorte da pesquisa: “*Como a utilização do GeoGebra pode contribuir para análise gráfica de dados na disciplina estatística em uma formação inicial de professores?*”, as possibilidades de análises univariadas e bivariadas utilizadas pela dupla de futuros professores podem ser identificadas como ferramentas presentes na tecnologia digital GeoGebra, que, quando integradas às atividades de análise gráfica de dados, oferecem oportunidades de compreender os problemas subjacentes a esses dados.

À luz dos referenciais teóricos, como analisado a partir da *Homework* desenvolvida pela dupla de futuros professores, foram as ferramentas de representação gráfica, aliadas às estatísticas, advindas das análises univariadas e bivariadas, que se mostraram favoráveis nos estágios de transnumeração dos dados e que permitiram aprimoramentos nas análises

apresentadas. No contexto da visualização de dados, os resultados emergentes do trabalho com os dados, nomeadamente, as representações gráficas e as estatísticas resultantes das análises univariadas e bivariadas, se mostraram relevantes para a *compreensão* dos dados visualizados, assim como as representações gráficas forneceram elementos *fins* para posteriores *verbalizações* sobre os dados, inclusive, sem desconectá-los do contexto de origem.

Como este trabalho é um recorte de uma tese, outras produções poderão considerar a integração do *Gapminder* e do SPSS.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo financiamento da pesquisa origem deste estudo.

Referências

- Andre, M. (2020) *Implementing the statistical investigative process in secondary school education*. 2020. 131f. Thesis (Doctorate in Science). Johannes Kepler University. Linz, Austria.
- Araújo, J. R. A. (2018) *Atividades para o estudo das medidas de tendência central: uma proposta com o apoio do GeoGebra*. 2018. 145 f. Dissertação (Mestrado em Educação Matemática). Pontifícia Universidade Católica de São Paulo. São Paulo, SP.
- Biehler, R. (2018). Design principles, realizations and uses of software supporting the learning and the doing of statistics—a reflection on developments since the late 1990s. In: Sorto, M. A.; White, A. & Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (pp. 1-6). Kyoto, Japan.
- Bolch, C. & Jacobbe, T. (2018). Students’ Understanding of Data Visualizations. In: Sorto, M. A.; White, A. & Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (pp. 1-3) Kyoto, Japan.
- Eckert, A. (2018). Implementing research results in technology induced teaching of statistics – a literature review. In: Sorto, M. A.; White, A. & Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (p. 1-4). Kyoto, Japan.
- Francisco, V. R. (2016). *Interpretação de dados estatísticos: um estudo com alunos do ensino médio da Educação de Jovens e Adultos*. 2016. 131f. Dissertação (Mestrado em Educação Matemática e Tecnológica). Universidade Federal de Pernambuco. Recife, PE.
- Giordano, L. V. O. (2017). *Panorama da competência estatística no ensino médio brasileiro: das ideias e práticas dos professores ao desempenho dos alunos no ENEM*. 2017. 274f. Tese (Doutorado em Ensino de Ciências e Matemática). Universidade Estadual de Campinas. Campinas, SP.
- Godoy, A. S. (1995). Pesquisa qualitativa: Tipos fundamentais. *Revista de Administração de empresas*, 35(3), 20–29.
- Gonzales, F. & Lopes, C. E. (2017). Estudo de viabilidade do uso do software GeoGebra, no ensino estatístico. *FESPM, Federación Española de Sociedades de Profesores de Matemáticas (Ed.), VIII Congreso Ibero-Americano de Educación Matemática* (pp. 14-22). Madri, España.
- Lévy, P. (1993). *As tecnologias da inteligência: o futuro do pensamento na era da informática*.

Rio de Janeiro, RJ: Editora 34.

- Mcnamara, A. (2018). Imagining the Future of Statistical Education Software. In: Sorto, M. A.; White, A.; Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (pp. 1-3). Kyoto, Japão.
- Mantilla, C. (2018). Using ICT's for to Introduce to Statistic. In: Sorto, M. A.; White, A. & Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (pp. 1-4). Kyoto, Japan.
- Oliveira, S. A. P. (2016). *Educação Estatística em escolas do povo Xukuru do Ororubá*. 2016. 152f. Dissertação (Mestrado em Educação Matemática e Tecnológica). Universidade Federal de Pernambuco. Recife, PE.
- Pfannkuch, M.; Rubick, A. & Yoon, C. (2002). Statistical thinking and transnumeration. In: B. Barton, B.; Irwin, K. C.; Pfannkuch, M. & Thomas, M. O. J. (Eds.), *Mathematics Education in the South Pacific. Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia* (pp. 567-574). Auckland, New Zealand.
- Prodromou, T. & Dunne, T. (2017). Data Visualisation and Statistics Education in the Future. In T. Prodromou (Ed.), *Data Visualization and Statistical Literacy for Open and Big Data* (pp. 1-28). IGI Global.
- Santos, P. G.; Oliveira Júnior, A. P.; Santos, L. R. S.; Carvalho, D. M. & Saito, S. (2022). The use of technological resources in teacher education as a contribution to the teaching of statistics and probability in Brazil. *Bridging the Gap: Empowering and Educating Today's Learners in Statistics. Proceedings of the Eleventh International Conference on Teaching Statistics* (pp. 1-6). Rosario, Argentina.
- Sera, E. K. (2016). *Conhecimento de professores para o ensino da leitura e construção de gráficos estatísticos na educação básica*. 2016. 203f. Dissertação (Mestrado em Educação Matemática). Universidade Anhanguera de São Paulo. São Paulo, SP.
- Souza, J. M. G. (2019). *Interpretação de gráficos: explorando o Letramento estatístico dos professores de escolas públicas no campo nos espaços de oficinas de formação continuada*. 2019. 144 f. Dissertação (Mestrado em Educação Matemática e Tecnológica). Universidade Federal de Pernambuco. Recife, PE.
- Souza, R. & Calejon, L. M. (2019). Uso da tecnologia da informação e comunicação em uma sequência didática incluindo *software* Geogebra no ensino da estatística descritiva. *REnCiMa. Revista de Ensino de Ciências e Matemática*, 10(4), 227-244.
- Votto, T. R. (2018). *As potencialidades lúdicas nas estratégias para o ensino e a aprendizagem estatística nos anos iniciais do ensino fundamental*. 2018. 175 f. Dissertação (Mestrado em Educação em Ciências Química da Vida e Saúde). Universidade Federal do Rio Grande. Porto Alegre, RS.
- Wild, C. J. & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *Internacional Statistical Review*, 67(3), 223-265.
- Woodard, V. W. (2018). Defining the relationship between statistical thinking and statistical computing. In: Sorto, M. A.; White, A. & Guyot, L. (Eds.) *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)* (pp. 1-6). Kyoto, Japan.