

Dos dados brutos à informação: o papel das técnicas transnumerativas no ensino de Estatística

From raw data to information: the role of transnumerative techniques in the teaching of statistics

De los datos brutos hacia la información: el papel de las técnicas transnumerativas en la enseñanza de la Estadística

Irene Mauricio Cazorla¹
Universidade Estadual de Santa Cruz
<https://orcid.org/0000-0003-3028-5513>

Miriam Cardoso Utsumi²
Universidade Estadual de Campinas
<https://orcid.org/0000-0001-7738-0513>

Carlos Eduardo Ferreira Monteiro³
Universidade Federal de Pernambuco
<https://orcid.org/0000-0003-4355-0793>

Resumo

Um aspecto importante da aprendizagem de estatística refere-se à compreensão de sua essência – extrair informação subjacente aos dados –, o que implica no conhecimento da transformação de grandes massas de dados em estatísticas – tabelas, gráficos e medidas resumo – e na interpretação dos seus significados, permitindo a compreensão do fenômeno em estudo. Esse processo tem sido denominado de transnumeração. Neste artigo, refletimos sobre aspectos teóricos e possíveis implicações pedagógicas da transformação dos dados brutos em estatísticas, e destas em informação, e o papel das técnicas transnumerativas, visando auxiliar o ensino dos conceitos de estatística, na educação básica, normatizados pela Base Nacional Comum Curricular.

¹ icazorla@uol.com.br

² mutsumi@unicamp.br

³ carlos.fmonteiro@ufpe.br

Palavras-chave: Dados estatísticos, Estatísticas, Informação, Transnumeração, Técnicas transnumerativas.

Abstract

An important aspect of learning statistics refers to the understanding of its essence - extracting information underlying the data -, which implies knowledge about transformation of large amount of data into statistics - tables, graphs, and summary measures - and the interpretation of its meanings, allowing the understanding of phenomenon under study. This process has been called transnumeration. In this article we reflect on theoretical aspects and possible pedagogical implications of transformation of raw data into statistics and these into information, as well as on the role of transnumerative techniques, aiming to support the teaching of statistical concepts in basic education level, standardised by the National Common Curricular Base.

Keywords: Statistical data, Statistics, Information, Transnumeration, Transnumerative techniques.

Resumen

Un aspecto importante del aprendizaje de la Estadística se refiere a la comprensión de su esencia -extraer información subyacente a los datos -, lo que implica el conocimiento de la transformación de grandes masas de datos en estadísticas – tablas, gráficos y medidas de resumen - y en la interpretación de sus significados, permitiendo la comprensión del fenómeno en estudio. Este proceso se ha denominado transnumeración. En este artículo reflexionamos sobre los aspectos teóricos y las posibles implicaciones pedagógicas de la transformación de datos brutos en estadísticas y estas en información, y el papel de las técnicas transnumerativas, con el objetivo de ayudar a la enseñanza de los conceptos de estadística, en educación básica, estandarizados por la Base Nacional Común Curricular.

Palabras clave: Datos estadísticos, estadísticas, Información, Transnumeración, Técnicas transnumerativas.

Dos dados brutos à informação: o papel das técnicas transnumerativas no ensino de Estatística

Este artigo foi desenvolvido a partir de reflexões sobre as variáveis estatísticas e as possibilidades de representação em gráficos por ocasião do Seminário Hispano-Brasileiro de Educação Estatística (Cazorla, Utsumi & Monteiro, 2020a), as quais foram expandidas em Cazorla, Utsumi e Monteiro (2020b). Posteriormente, em Cazorla, Utsumi e Monteiro (2021), sistematizamos recomendações para seu ensino na Educação Básica e apresentamos um esquema conceitual inicial visando auxiliar os professores na escolha e construção dos gráficos. Neste artigo, ampliamos o nosso olhar para além da representação gráfica, incluindo a apresentação de tabelas e medidas resumo básicas, normatizadas pela Base Nacional Comum Curricular – BNCC (MEC, 2018).

Assim, este artigo se constitui em um ensaio teórico, no qual discutimos implicações pedagógicas referentes ao percurso dos dados brutos, para o qual adaptamos as técnicas transnumerativas propostas por Chick (2004), até se transformarem em estatísticas, convertendo estas em informação, visando contribuir com o desenvolvimento da transnumeração (Wild & Pfannkuch, 1999) no ensino de Estatística na Educação Básica.

Na sequência a esta introdução, são apresentadas cinco outras seções. Na segunda seção descrevemos as orientações da BNCC (MEC, 2018) sobre as competências em Estatística que devem ser desenvolvidas, bem como os conceitos e procedimentos a serem ensinados na Educação Básica, de forma univariada e bivariada, a partir da sistematização apresentada por Cazorla, Utsumi e Monteiro (2021). Na terceira seção apresentamos o marco teórico, discutindo o conceito de transnumeração proposto por Wild e Pfannkuch (1999). Na quarta seção apresentamos as técnicas transnumerativas propostas por Chick (2004) e propomos uma

adaptação das técnicas transnumerativas; na quinta seção apresentamos dois exemplos de transnumeração e, na sexta seção, tecemos nossas considerações finais e reflexões sobre as implicações pedagógicas. Cabe salientar que este estudo aborda os conceitos estatísticos e não os probabilísticos.

Orientações da BNCC para o percurso dos dados brutos às estatísticas

O ensino de Estatística na Educação Básica brasileira foi introduzido, em 1997, com a publicação dos Parâmetros Curriculares Nacionais (MEC, 1997), sendo ratificado pela BNCC (MEC, 2018). Sua inserção se deve ao reconhecimento da importância do desenvolvimento do pensamento estatístico desde o início da escolarização (Batanero, 2019).

A BNCC aponta que a escolarização deverá ser orientada pelo desenvolvimento de dez competências gerais, das quais destacamos a sétima:

Argumentar com base em fatos, dados e informações confiáveis, para formular, negociar e defender ideias, pontos de vista e decisões comuns que respeitem e promovam os direitos humanos, a consciência socioambiental e o consumo responsável em âmbito local, regional e global, com posicionamento ético em relação ao cuidado de si mesmo, dos outros e do planeta (MEC, 2018, p. 9).

Esta competência enfatiza que os processos educacionais devem estar articulados com aspectos concretos da realidade social. Para o ensino de Matemática, a BNCC (MEC, 2018, p. 267) elenca oito competências, das quais destacamos a quarta, que incentiva a observação sistemática de “aspectos quantitativos e qualitativos presentes nas práticas sociais e culturais, de modo a investigar, organizar, representar e comunicar informações relevantes, para interpretá-las e avaliá-las crítica e eticamente, produzindo argumentos convincentes”, complementada pela sétima competência, que preconiza trabalhar com projetos que abordem “questões de urgência social, com base em princípios éticos, democráticos, sustentáveis e solidários”, e pela oitava, que enfatiza o trabalho coletivo.

Para o Ensino Médio, a BNCC apresenta os Temas Contemporâneos Transversais (MEC, 2019), que buscam explicitar possíveis articulações entre os diferentes componentes

curriculares e estabelecer conexões com as situações cotidianas vividas pelos estudantes, a fim de integrar contextos contemporâneos aos objetos do conhecimento (Cazorla & Giordano, 2021). Nesse sentido, o ensino de Estatística pode contribuir para o desenvolvimento dessas competências, desde que sua abordagem considere a transnumeração.

Segundo a BNCC (MEC, 2018), os objetos de conhecimento de Estatística que devem ser ensinados na Educação Básica são: pesquisa amostral e censitária; variáveis qualitativa e quantitativa; população e amostra; tabelas estatísticas e tabela de distribuição de frequência (TDF), simples e de dupla entrada; gráficos de barras/colunas, linhas, circular, pictograma, diagrama de ramos e folha, da caixa (*box-plot*) e histograma; medidas resumo: frequência absoluta e relativa; medidas de tendência central – MTC (média, mediana e moda); e medidas de dispersão – MD (amplitude e desvio padrão).

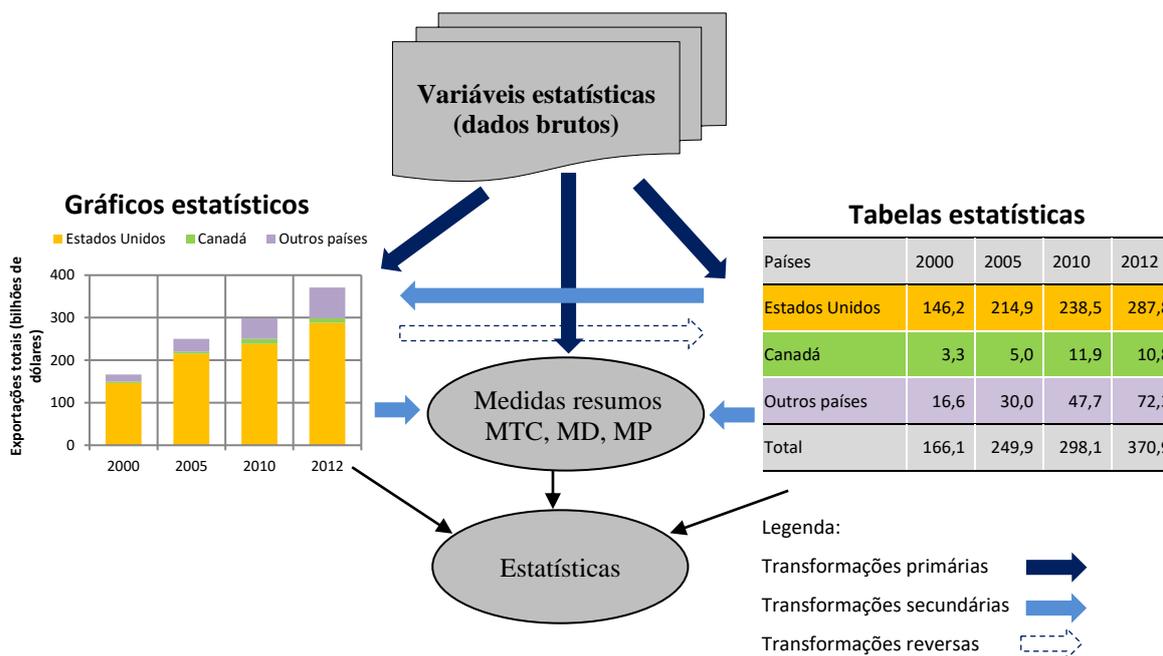
Observa-se, ainda, que os quartis e a amplitude interquartílica não foram elencados na BNCC, mas estão implícitos na construção do *box-plot*. Além disso, alguns objetos não foram contemplados, tais como o diagrama de pontos (*dot-plot*), de hastes ou bastão e o coeficiente de variação (CV), para os quais advogamos seu ensino.

Os conceitos e procedimentos da Estatística nos permitem organizar os dados brutos em tabelas, convertê-los em gráficos ou sumariá-los em medidas resumo. Na Figura 1 apresentamos o fluxograma do percurso que os dados brutos podem seguir até se transformarem em estatísticas, conforme sugerido por Silva-Júnior (2018).

A Figura 1 apresenta algumas possibilidades de transformações dos dados. Todavia, essas transformações parecem não ser facilmente compreendidas por professores e estudantes da Educação Básica, em parte pela falta de entendimento de conceitos básicos, tais como variável e dado, conforme mostram os estudos de Matos (2019) com estudantes do 9º ano, e de Oliveira (2020), com licenciandos em Matemática, sobre as dificuldades no reconhecimento do gráfico adequado para uma variável disposta em uma tabela estatística.

Figura 1.

Transformações dos dados brutos em estatísticas (adaptado de Silva-Júnior, 2018, p. 27)



Esses resultados sugerem a importância que devemos dar ao reconhecimento da natureza da variável e dos dados gerados, a fim de fazer escolhas adequadas para seu tratamento estatístico. Isso implica saber definir o que é uma variável estatística e como ela pode ser operacionalizada de modo a gerar certo tipo de dado.

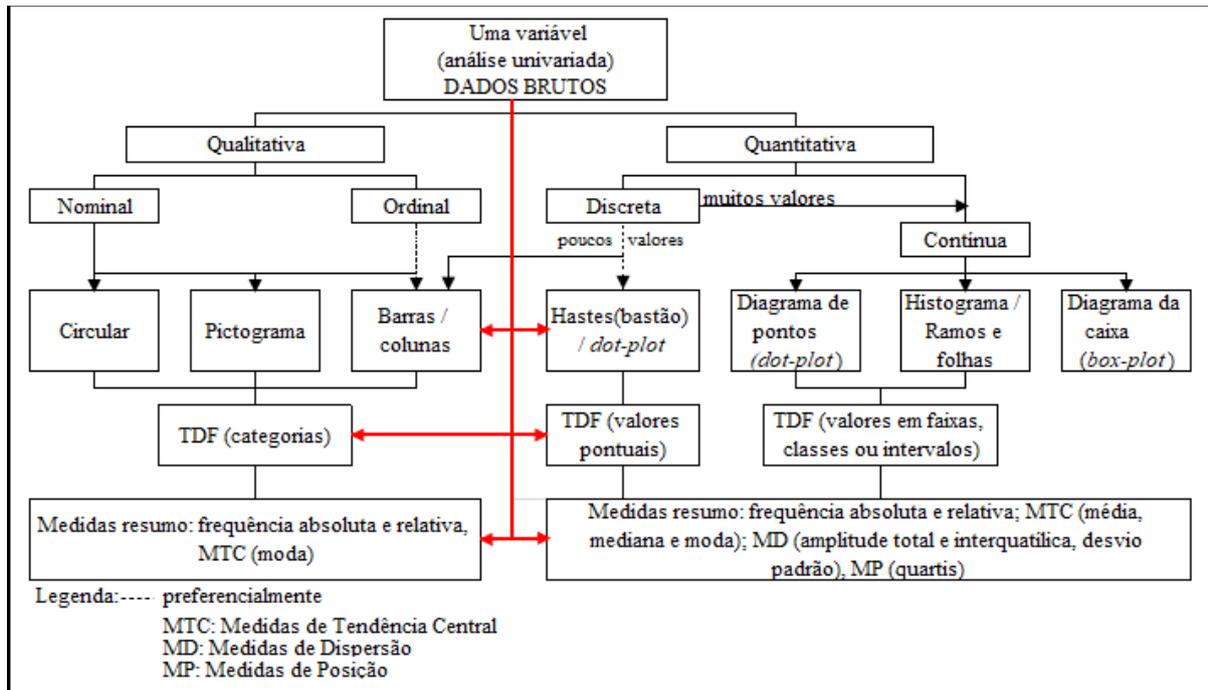
Nesse sentido, Cazorla, Utsumi e Monteiro (2021) apresentam os fundamentos que embasam as variáveis estatísticas e suas representações gráficas: explicitam os conceitos subjacentes aos gráficos, distinguem as variáveis *matemáticas* das variáveis *estatísticas* (fortemente demarcadas pela sua variabilidade), discutem a classificação das variáveis de acordo com a sua natureza e exemplificam como operacionalizar variáveis (empíricas e conceituais). Relativamente aos gráficos, os autores detalham seus componentes, propõem uma classificação de acordo com o arcabouço no qual são construídos, e apresentam um esquema para os tipos de gráficos, tanto em uma análise univariada quanto bivariada.

Tomando como base esse estudo, adicionamos as TDF e as medidas resumo no contexto da análise univariada, conforme Figura 2. Observamos que os fluxos apresentam inicialmente os

gráficos, seguido das tabelas e, por fim, as medidas resumo. Todavia, há diferentes possibilidades de percurso, como ilustrado na Figura 1.

Figura 2.

*O percurso dos dados brutos na transformação em estatísticas, no contexto univariado
(adaptação de Cazorla, Utsumi & Monteiro, 2021, p. 30)*



Quando nos deparamos com variáveis qualitativas (fluxo esquerdo da Figura 2), via de regra, iniciamos o tratamento organizando os dados em uma TDF (categorias), a partir da qual calculamos as medidas resumo (frequências absoluta/relativa e a moda) e, posteriormente, transformamos a TDF em gráficos de barras/colunas, circular ou pictogramas.

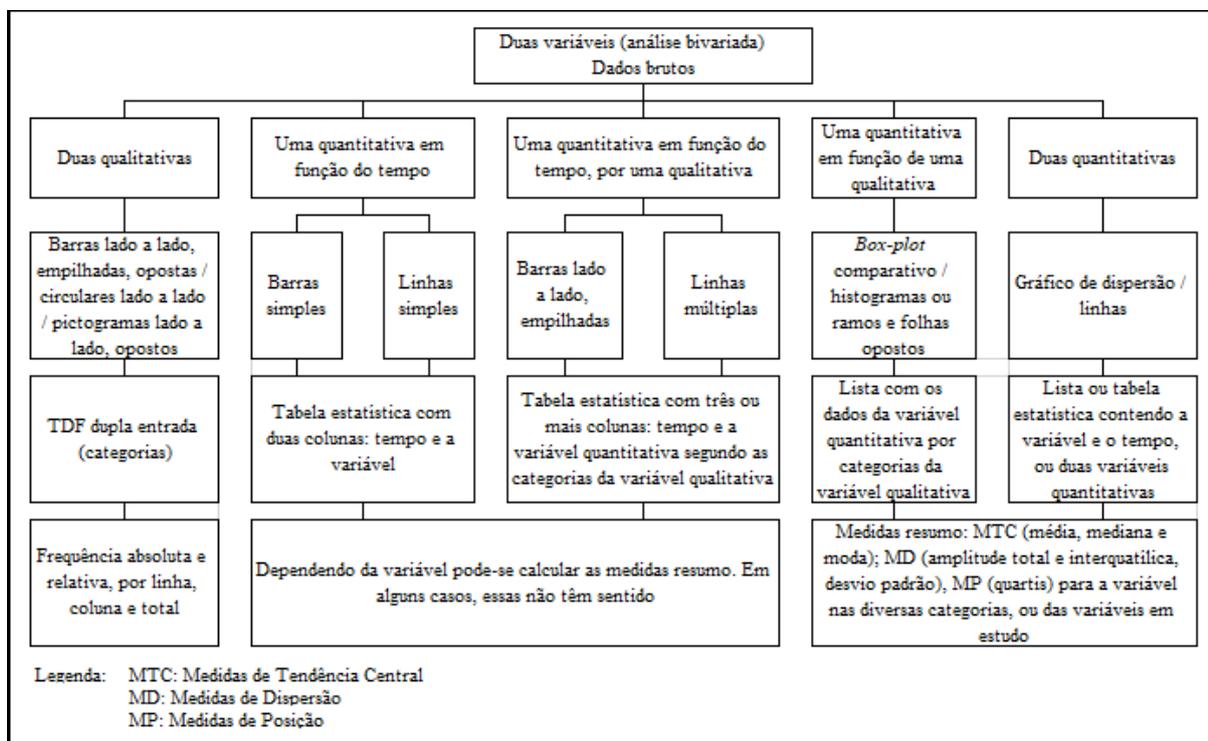
Quando os dados são quantitativos (fluxo direito da Figura 2), no caso das variáveis discretas que tomam poucos valores, como número de irmãos, o tratamento é muito similar ao das variáveis qualitativas. Isto é, construir a TDF (valores pontuais) e o gráfico de barras/colunas (hastes/bastão ou *dot-plot*), e calcular as medidas resumo (média, mediana, moda, desvio padrão etc.). Já quando a variável é discreta que toma muitos valores, ou contínua, como altura dos estudantes, o percurso desejável seria o gráfico de pontos (*dot-plot*), a partição da amplitude total (AT) em classes para construir a TDF (em classes/faixas) e, a partir desta,

construir o histograma e, por fim, calcular as medidas resumo. Essas transformações são denominadas por Chick (2004) de técnicas transnumerativas. Embora possa se calcular as MTC diretamente dos dados brutos.

Na Figura 3 apresentamos o fluxo da análise bivariada, pois o foco é a análise de duas variáveis, embora, em alguns casos, possa haver uma terceira variável.

Figura 3.

*O percurso dos dados brutos na transformação em estatísticas, no contexto bivariado
(adaptação de Cazorla, Utsumi & Monteiro, 2021, p. 30)*



Sugerimos a inserção dos gráficos/diagramas de hastes/bastão na representação de variáveis discretas que tomam poucos valores (Kataoka & Hernandez, 2010), especialmente quando se deseja representar graficamente o valor da média, com um triângulo (fiel da balança), como discutido por Cazorla, Santana e Utsumi (2019). Também, advogamos a inserção do ensino do *dot-plot*, que é um gráfico que mantém uma relação biunívoca entre o dado e sua representação na reta numérica, auxiliando o estudante no reconhecimento do seu dado. Além disso, permite localizar as diversas estatísticas (Mínimo, Máximo, primeiro e terceiro quartil e a mediana) e construir em cima dele o *box-plot*, denominado por Silva, Magina e Silva

(2010) de *dot-box-plot*. Esse procedimento possibilita aos estudantes compreender o *box-plot*, que é um gráfico poderoso, mas que não mostra mais os dados brutos, o que demanda uma abstração do leitor (Almeida, Sousa & Cazorla, 2021).

Além disso, o *dot-plot*, ao percorrer a reta numérica com todos os valores da variável, permite visualizar a amplitude total (AT) e a distribuição dos dados, permitindo assim uma escolha da partição da AT em classes para construir a TDF e o histograma.

Quanto ao gráfico/diagrama de ramos e folhas, identificamos algumas restrições, uma vez que, para variáveis que tomam valores acima da ordem das dezenas, essas precisam de truncamento, como exemplificado por Morettin e Bussab (2010).

Com esses dois fluxogramas, esperamos auxiliar os professores, a partir dos dados, a identificarem a natureza das variáveis que os geraram e fazerem escolhas adequadas, a fim transformar esses dados em estatísticas e compreender seus significados. A seguir, apresentamos o conceito de transnumeração e as técnicas transnumerativas.

O marco teórico, a transnumeração: das estatísticas à informação

O termo transnumeração foi cunhado por Wild e Pfannkuch (1999) na discussão sobre o pensamento estatístico no contexto de uma investigação empírica (*empirical enquiry*) desenvolvida por produtores de estatísticas, tais como estudantes e profissionais que utilizam a Estatística. Os autores argumentam que transnumeração se refere à ideia mais fundamental em uma abordagem estatística: formar e alterar representações de dados de aspectos de um sistema para chegar a um melhor entendimento desse sistema.

Nós o definimos como ‘transformações de numeramento feitas para facilitar a compreensão’. A transnumeração ocorre quando encontramos formas de obter dados (por meio de medição ou classificação) que capturam elementos significativos do sistema real. Ela permeia todas as análises de dados estatísticos, ocorrendo sempre que mudamos nossa forma de olhar os dados na esperança de que isso nos transmita um novo significado. Podemos examinar muitas representações gráficas para encontrar várias outras realmente informativas. Podemos reexpressar os dados por meio de transformações e reclassificações em busca de novos *insights*. Podemos tentar uma

variedade de modelos estatísticos. E no final do processo, a transnumeração acontece mais uma vez quando descobrimos representações de dados que ajudam a transmitir nossos novos entendimentos sobre o sistema real para outras pessoas. (Wild & Pfannkuch, 1999, p. 227, tradução nossa).

Os pesquisadores usam o termo *sistema real* para se referirem ao fenômeno/situação/problema a ser investigado. A palavra *real*, a nosso ver, refere-se a situações próximas dos estudantes. Neste trabalho, utilizamos a palavra fenômeno como “qualquer evento observável [...] ou [...] inferido a partir de suas manifestações” (Cazorla & Oliveira, 2010, p. 118).

Essa concepção de transnumeração apresentada por Wild e Pfannkuch (1999), no âmbito de um contexto de investigação, também poderia ser explorada em outros contextos de utilização do pensamento estatístico, tais como contextos cotidianos de leitura por adultos (Gal, 2002) ou por professores e estudantes da escola básica (Monteiro & Ainley, 2004). No âmbito escolar, Henriques e Oliveira (2019) enfatizam que as transformações dos dados brutos em estatísticas envolvem o raciocínio sobre os dados que dependem de uma grande diversidade de representações, que por sua vez sustentam a compreensão e favorecem a comunicação de ideias. A identificação de tendências, a formulação de inferências e representações efetivas só são possíveis se os dados forem transformados adequadamente.

Além disso, algumas representações podem destacar ou ocultar informações relevantes contidas nos dados. Por exemplo, utilizar a renda per capita ao invés da renda mediana pode camuflar as desigualdades econômicas de um país. A manipulação de escalas nos gráficos, ou a apresentação dos dados em tabelas ao invés de gráficos, podem induzir o leitor, de maneira inadequada, na tomada de decisões (Cazorla & Castro, 2008).

A transnumeração, junto com o reconhecimento da necessidade de dados, a variabilidade, raciocínio com modelos estatísticos e a integração da Estatística com o contexto fazem parte da segunda dimensão do pensamento estatístico, necessário para a compreensão dos fenômenos em estudo (Wild & Pfannkuch, 1999).

Chick (2003) argumenta que a eficiência das representações para uma efetiva comunicação dos dados depende do conhecimento dos estudantes sobre os tipos de representação que são úteis, bem como de um repertório de operações específicas para transformar dados que conduzam a tais representações ou alterem a sua estrutura, tornando-as convincentes em termos de evidência e significados de relações estatísticas que se obtêm dos dados.

As técnicas transnumerativas e a proposta de adaptação

Segundo Pfannkuch e Rubick (2002):

A transnumeração ocorre em três específicos exemplos. Se o sistema real ou o sistema estatístico é pensado a partir de uma perspectiva de modelagem, então o pensamento de transnumeração ocorre quando: (1) medidas que “capturam” qualidades ou características da situação real são achadas, (2) os dados que têm sido coletados são transformados de dados não tratados em múltiplas representações gráficas, e (3) o significado dos dados, a análise, tem que ser comunicada em uma forma que pode ser entendida em termos da situação real por outros. (p. 5, tradução nossa).

Chick (2004) propõe dez técnicas transnumerativas necessárias à transformação dos dados brutos em estatísticas, a saber: classificação; agrupamento; seleção de subconjunto de dados; mudança do tipo de variável; cálculo de frequência; cálculo de proporção; elaboração gráfica; cálculo de medidas de tendência central; cálculo de medidas de dispersão e outros cálculos.

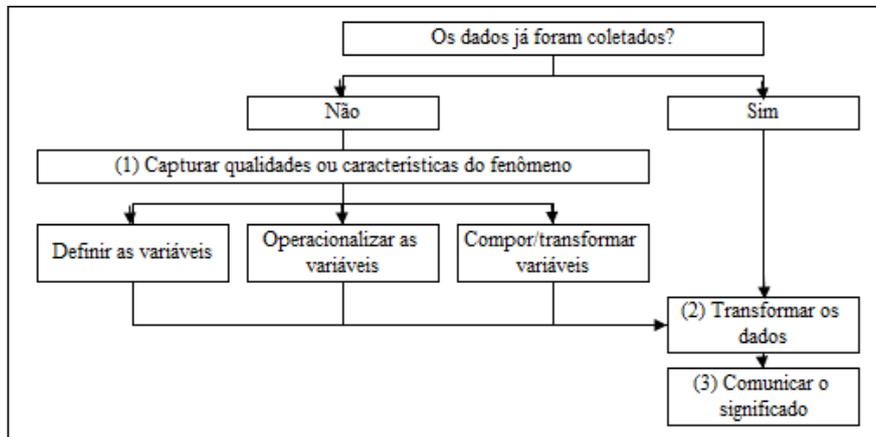
Neste artigo analisamos essas técnicas e propomos uma adaptação, a fim de sistematizar o processo de tratamento de dados brutos, de modo que permita a apreensão do fenômeno em estudo. Para isso, é preciso analisar se os dados já foram coletados ou se eles ainda vão ser coletados, conforme Figura 4.

Desse modo, quando já estamos de posse dos dados, o processo se concentra na sua transformação e posterior comunicação dos significados. Um exemplo deste caso é a situação apresentada por Chick (2003), em que a partir de dados dispostos em uma tabela/gráfico e a mensagem que se queria transmitir, solicitou-se a estudantes universitários procurar outra forma de comunicar a informação. Outra situação é quando não temos os dados e precisamos coletá-los. Neste caso, precisamos nos preocupar com a maneira como vamos “capturar” as

qualidades e características do fenômeno em estudo, isto é, como vamos definir as variáveis, operacionalizá-las e, se preciso for, como vamos compor/transformá-las.

Figura 4.

Percursos para apreensão de um fenômeno de acordo com a disponibilidade dos dados



O processo de transnumeração é potencializado quando utilizamos dados multivariados. Isto porque, ao analisar duas ou mais variáveis, podemos trabalhar as possíveis relações que se estabelecem entre elas, ativando o raciocínio inferencial informal (Makar & Rubin, 2009).

No Quadro 1 apresentamos as dez técnicas transnumerativas propostas por Chick (2004) e, a partir dessas, propomos seis técnicas (Quadro 2), que descrevemos em seguida.

Quadro 1.

Técnicas transnumerativas (Chick, 2004, p. 169, tradução nossa)

Técnica	Descrição*	Exemplo
Classificação	Os dados são classificados segundo algum critério. Nenhuma nova variável surge.	Os dados são classificados segundo quantidade de horas de exercício, do menor para o maior.
Agrupamento	Os dados são agrupados segundo algum critério. Isso cria uma nova variável. Essa técnica pode envolver a mudança do tipo de variável.	Uma nova variável “nível de consumo” é criada usando os dados sobre <i>fast-food</i> , com valores “baixo” (0-1 refeições de <i>fast-food</i> por semana), “médio” (2-3 refeições de <i>fast-food</i> por semana) e “alto” (≥ 4 refeições de <i>fast-food</i> por semana).
Seleção de subconjunto	Um subconjunto dos dados é selecionado para transnumeração posterior.	Dados associados com os níveis “baixo” e “alto” de consumo são considerados (dados de nível “médio” não são).
Mudança do tipo de variável	Uma variável quantitativa (numérica) é pensada em termos qualitativos, ou uma variável qualitativa é pensada em termos quantitativos ou ordinais.	Atividade favorita (variável qualitativa/catórica) pode ser dada <i>status</i> ordinal, ordenando as atividades da mais ativa para menos ativa.
Cálculo de Frequência	As frequências de ocorrência dos valores da variável qualitativa (catórica) são determinadas. Cria-se uma nova variável.	A quantidade de pessoas em cada uma das categorias dos “níveis de consumo” é determinada.
Cálculo de Proporção	Proporções (p.e., frações) são determinadas em relação ao todo. Isso cria uma nova variável.	O percentual de pessoas em cada categoria de atividade é determinado.
Elaboração Gráfica	Algumas ou todas as variáveis nos dados (na sua forma atual) são representadas graficamente ou tabuladas.	Um gráfico de dispersão de horas de exercícios versus o número de refeições de <i>fast-food</i> consumidas é construído.
Cálculo de tendências centrais	Uma medida de tendência central (p.e., média) é calculada para uma variável. Pode criar nova variável.	O número médio de refeições de <i>fast-food</i> consumidas por semana é determinado.
Cálculo da Medida de dispersão	Alguma medida de dispersão de valores associada com uma variável quantitativa (numérica) é determinada. Pode criar nova variável.	O intervalo de valores associados com o número de horas de exercício é reconhecido.
Outros cálculos	Termo genérico, reconhecimento de que outros cálculos estatísticos sobre os dados são possíveis (p.e., somatório, coeficientes de correlação etc.).	Uma reta de melhor ajuste para os dados de horas de exercícios e número de refeições de <i>fast-food</i> consumidas.

* Chick denomina as variáveis de numéricas e catóricas, mas neste trabalho temos utilizado as nomenclaturas quantitativa e qualitativa.

Quadro 2.

Proposta de Técnicas transnumerativas

Técnica	Descrição	Exemplo
Determinação das variáveis	Definição das variáveis.	Variáveis primárias: idade e altura do estudante, cor favorita.
	Operacionalização das variáveis.	Variáveis quantitativas em faixas. Escalas de Likert para quantificar variáveis qualitativas.
	Composição/transformação de variáveis.	Variáveis secundárias: índices, taxas, razões etc.
Classificação	Criação de categorias.	Variáveis qualitativas (nominais e ordinais).
Ordenação	Ordenação de dados quantitativos.	Encontrar as medidas de posição (Mediana e Quartis).
Construção de tabelas	Representação dos dados em tabelas estatísticas.	TDF simples e de dupla entrada
Construção de gráficos	Representação dos dados em gráficos estatísticos.	Gráficos estatísticos de forma geral.
Cálculo de medidas resumo	Cálculo da frequência absoluta e relativa.	A frequência absoluta/relativa a partir da TDF em categorias; em valores pontuais ou em classes.
	Cálculo das medidas resumo básicas.	MTC (média, mediana e moda); MD (amplitude total e interquartílica, desvio padrão e CV) e MP (Quartis).
	Cálculo de outras medidas resumo.	Os coeficientes de uma função matemática que melhor descreve a nuvem de pontos em um diagrama de dispersão. Coeficiente de correlação.

A seguir, descrevemos as técnicas transnumerativas propostas.

Determinação das variáveis. Nesta técnica agrupamos duas técnicas propostas por Chick (2014): *Mudança do tipo de variável* e *Agrupamento*. Esta técnica tem como objetivo definir como serão as medidas que “capturam” qualidades ou características do fenômeno a ser investigado, isto é, refere-se à fase da coleta de dados. Esta técnica está composta por três subtécnicas:

- a) *Definição das variáveis.* Significa explicitar como aquela característica a ser estudada será apreendida. No caso das variáveis empíricas, que possuem instrumentos de medida ou categorias bem definidas, é bastante simples. Já variáveis conceituais, como “classe social”, depende do quadro teórico no qual esse construto é engendrado.
- b) *Operacionalização das variáveis.* Implica em explicitar como a variável já definida vai ser qualificada/quantificada.
- c) *Composição/transformação de variáveis.* Aqui se encontram as variáveis secundárias, que são geradas a partir de variáveis primárias, que definimos a seguir.

As *variáveis primárias* são aquelas colhidas diretamente da fonte. São muito utilizadas nas pesquisas estatísticas realizadas pelos estudantes em ambiente escolar. As *variáveis secundárias* são criadas a partir das primárias, tais como os índices, razões e taxas, que fazem

parte do currículo de Ciências, Geografia etc. Seu tratamento requer alguns cuidados. Por exemplo, de posse da renda per capita dos estados brasileiros, não se pode aplicar a média (simples) para encontrar a renda per capita do Brasil; seria necessário ponderar pelo tamanho da população de cada estado.

Observamos que, em alguns casos, as variáveis são fáceis de serem definidas, em especial as variáveis empíricas (que têm referente observável no mundo), com instrumentos de medidas consagrados, tais como a altura, massa corpórea, idade etc. Todavia, as variáveis conceituais (não observáveis, mas inferidas pelo comportamento) são complexas, polêmicas e precisam da criação de instrumentos (Cazorla & Oliveira, 2010). Além disso, uma variável pode ser coletada de diferentes maneiras, como exemplificam Cazorla, Utsumi e Monteiro (2021) com a variável “Hábitos de prevenção ao contágio pela Covid-19”.

As variáveis conceituais requerem um trabalho mais sofisticado, como o realizado na pesquisa de Buehring (2006), que trabalhou o medo das crianças (qualitativa nominal); ou na de Souza (2007), que trabalhou o quanto as crianças gostavam da merenda da escola (qualitativa ordinal). Também podem ser utilizadas escalas de Likert, como no trabalho de Silva-Júnior (2018), que investigou quão saudável era a alimentação dos estudantes; no trabalho de Silva e Cazorla (2020), que investigaram a percepção de estudantes sobre o impacto do uso dos agrotóxicos no meio ambiente; e no trabalho de Almeida, Sousa e Cazorla (2021), que trabalharam os hábitos de estudos e o desempenho em Matemática de estudantes da escola básica. Nestes três últimos trabalhos, os autores utilizaram variáveis qualitativas ordinais, que foram transformadas para uma escala numérica e somadas, permitindo que obtivessem a pontuação total das escalas.

Classificação. Esta técnica é própria do manejo das categorias das variáveis qualitativas, que podem ser coletadas de forma “fechada” com as categorias definidas *a priori*, ou de forma

“aberta”, quando as categorias são criadas *a posteriori*. Cazorla, Magina, Gitirana e Guimarães (2017, p. 34) exemplificam como recolher a variável “medo das crianças” de ambas as formas. Para Guimarães, Luz e Ramos (2011, p. 2), “saber categorizar é fundamental para a construção de representações em gráficos e tabelas. Assim, categorizar e representar os dados são atividades imprescindíveis ao cidadão e a escola precisa desenvolver uma atuação didática voltada para tal”. Essas autoras mostram que, em alguns casos, mesmo com variáveis conhecidas pelos estudantes, classificar pode ser bastante complexo tanto para os estudantes quanto para os professores dos anos iniciais.

Ordenação. Esta técnica é própria do tratamento das variáveis quantitativas e é imprescindível quando se calcula as medidas de posição (estatísticas de ordem ou percentis), e, de forma mais específica, quando se calcula a mediana e os quartis para, assim, construir o *box-plot*. Apesar de a ordenação ser uma técnica bem simples, estudantes apresentam dificuldades, como foi verificado por Cobo (2003), Figueiredo (2017) e Cazorla, Utsumi e Santana (2020). Tal dificuldade compromete o cálculo da mediana e dos quartis pelos estudantes. Logo, merece a atenção do professor.

Observamos que existem variáveis que são por natureza ordenadas, como é o caso, por exemplo, dos *rankings* dos países na vacinação contra a Covid-19, do índice de percepção da corrupção, do quadro de medalhas nas olimpíadas etc.

Apenas a título de curiosidade, existe um ramo da Estatística, denominado Estatística não Paramétrica, que trabalha com o *ranking* (postos) das variáveis quantitativas.

Nas variáveis qualitativas ordinais, a recomendação é que, ao se construir as TDF ou gráficos, respeite-se a ordem natural das categorias, pois elas podem trazer informação implícita.

Construção de tabelas estatísticas. As tabelas organizam a informação estatística em forma matricial. Nesta técnica é preciso distinguir a palavra tabela, no sentido coloquial, que diz

respeito a qualquer lista/banco de dados/informações, da tabela estatística, que por definição tem seu conteúdo formado por dados tratados estatisticamente.

As tabelas estatísticas, também conhecidas como séries estatísticas, podem conter informação de diferentes variáveis e podem ser classificadas em *Histórica* (temporal, cronológica), quando são indexadas pelo tempo (5a); *Geográfica* (territorial, regional ou de localização), indexadas pelo espaço ou localização (Figura 5b); e *Específica*, quando o foco são as categorias da variável (5c). Estas tabelas são muito utilizadas nas disciplinas de Geografia, Ciências, Biologia, Sociologia etc.

Figura 5.

Exemplos de tabelas ou séries estatísticas (Crespo, 2002, p. 27-28)

(a) Série Histórica Preço do acém no varejo. São Paulo – 1989-94		(b) Série Geográfica Duração média dos estudos superiores		(c) Série Específica Rebanhos brasileiros	
Ano	Preço médio (US\$)	Países	Nº de anos	Espécies	Quantidade (1.000 cabeças)
1989	2,24	Itália	7,5	Bovinos	154.440,8
1990	2,73	Alemanha	7,0	Bubalinos	1.423,3
1991	2,12	França	7,0	Equinos	549,5
1992	1,89	Holanda	5,9	Asininos	47,1
1993	2,04	Inglaterra	Menos de 4	Muare	208,5
1994	2,62			Suínos	34.532,2
				Ovinos	19.955,9
				Caprinos	12.159,6
				Coelhos	6,1

A *Tabela de Distribuição de Frequência (TDF)* é um caso particular de tabela estatística, em que se tem a contagem do número de casos em cada uma das categorias de uma variável qualitativa (Figura 6a), valores pontuais, como o caso das variáveis discretas que tomam poucos valores (Figura 6b), ou intervalos de classes ou faixas, como o caso das variáveis contínuas ou discretas que tomam muitos valores (Figura 6c). Também existem tabelas conjugadas, em que se pode combinar esses tipos de tabelas.

Figura 6.

Exemplos de Tabelas de Distribuição de Frequência - TDF

(a) categorias		(b) valores pontuais		(c) intervalos de classe	
Mascote	Nº de estudantes	Nº de filhos	Nº de mulheres	Idade	Nº de habitantes
Cachorro	12	0	10	00 -05	10
Gato	15	1	45	05 -10	40
Coelho	2	2	30	10 -15	30
Outro	1	3	15
Total	30	Total	100	Total	100

Construção de gráficos estatísticos. Esta técnica é semelhante a proposta de Chick (2004).

Existem inúmeros trabalhos abordando este tema, dentre os quais destacamos as teses de Arteaga (2011) e Cazorla (2002). Aqui nos remetemos ao esquema conceitual proposto por Cazorla, Utsumi e Monteiro (2021) para representar variáveis estatísticas em gráficos de acordo com a sua natureza, em contexto univariado e bivariado.

Cálculo das medidas resumo. Nesta técnica englobamos cinco técnicas propostas por Chick (2004): cálculo de frequência; cálculo de proporção, cálculo de medidas de tendência central, cálculo de medida de dispersão e outros cálculos, que agrupamos em três subtécnicas: cálculo de frequências, cálculo das medidas resumo básicas e outros cálculos, que apresentamos a seguir.

- a) *Cálculo de frequências* (absoluta/relativa, simples/acumuladas). Típico das TDF, simples e de dupla entrada. Na TDF simples temos a frequência absoluta e a relativa, que pode ser expressa em probabilidade, em porcentagem e em fração (Cazorla, Henriques & Santana, 2020). Já na TDF de dupla entrada a frequência relativa pode ser calculada em relação ao total da linha, ao da coluna e ao total geral. Cazorla, Utsumi e Oliveira (2020) analisaram a importância dessas frequências na operacionalização das conjecturas (hipóteses) subjacentes à informação, a partir da apresentação de uma TDF.

Neste particular, Chick (2004) argumenta que a partir da frequência podemos criar uma nova variável. Por exemplo, quando analisamos a intenção de voto do eleitor nos candidatos à presidência da república, em um momento determinado, estamos trabalhando uma variável qualitativa nominal; mas quando acompanhamos essa intenção de votos ao longo do tempo, a porcentagem de cada candidato se torna uma variável quantitativa, isto é, uma série temporal. Observamos que, neste caso, o objetivo da análise é capturar as trajetórias, visando ao prognóstico do candidato que será o vencedor.

- b) *Cálculo das medidas resumo básicas.* Engloba o cálculo das medidas de tendência central (MTC), de dispersão (MD) e incluímos as medidas de posição (MP), pois

a BNCC incorporou o *box-plot* no rol de conteúdos de Estatística a serem trabalhados na Educação Básica.

Medidas de tendência central – MTC (média, mediana e moda). Têm sido objeto de investigação de muitos pesquisadores, como Cobo (2003), Figueiredo (2017), Batanero (2000), Cazorla, Santana e Utsumi (2019), entre outros. Os estudos mostram que, apesar de serem conceitos de pouca complexidade, os estudantes não conseguem dar significado a eles e, muitas vezes, não reconhecem as diversas situações em que podem utilizar a média “ponderada”. Nesse sentido, o trabalho de Cazorla, Santana e Utsumi (2019) propõe um esquema conceitual para a média, apresentando diversas situações para seu cálculo.

Medidas de dispersão (MD). Também há várias pesquisas que mostram a dificuldade da compreensão de variabilidade e de sua determinação (Silva, 2007; Moreno, 2010; Moreno 2014). Nesse sentido, é preciso ensinar os estudantes a examinarem a amplitude total e a distribuição dos dados, a presença de valores discrepantes, bem como a possibilidade de fatores que podem explicar uma maior ou menor variabilidade e, para isso, o *dot-plot* é muito útil. Também, consideramos que seria apropriado o ensino do coeficiente de variação (CV), que é uma medida de dispersão relativa ($CV = 100 * \frac{DP}{\bar{x}}$), porque tal medida permite a compreensão da variabilidade em relação à média, retirando a influência da ordem de grandeza da variável.

Medidas de posição (MP). Observamos que, se a estratégia didática de utilizar poucos dados para apresentar as MTC e MD era conveniente, essa perde sentido quando trabalhamos com os quartis, o que nos desafia a trabalhar com um número razoável de dados. Figueiredo (2017) faz uma revisão de como os livros didáticos de Portugal abordam esse conceito. Almeida, Sousa e Cazorla (2021) mostram como trabalhar a mediana e os quartis com estudantes do 1º ano do Ensino Médio, utilizando o *dot-box-plot* para analisar a relação entre os hábitos e frequência de estudo no desempenho em Matemática, mostrando o poder desse gráfico no entendimento da relação entre variáveis.

- c) *Outros cálculos*. Chick (2004) deixa em aberto “outros cálculos” para trabalhar o coeficiente de correlação ou os coeficientes em uma função matemática que melhor descrevem a nuvem de pontos em um diagrama de dispersão, dentre outras possibilidades na Educação Básica.

Na próxima seção exemplificamos o processo de transnumeração em dois contextos, um envolvendo duas variáveis qualitativas e outro envolvendo variáveis quantitativas, para assim tecermos nossas considerações finais.

Exemplificando o processo de transnumeração

Para apresentar as possibilidades de transnumeração entre duas variáveis qualitativas analisamos a tabela de dupla entrada trabalhada por Cazorla, Utsumi e Oliveira (2020, p. 264), que apresenta dados de mil respondentes sobre “Classe Social (CS)” e “Acesso ao Ensino

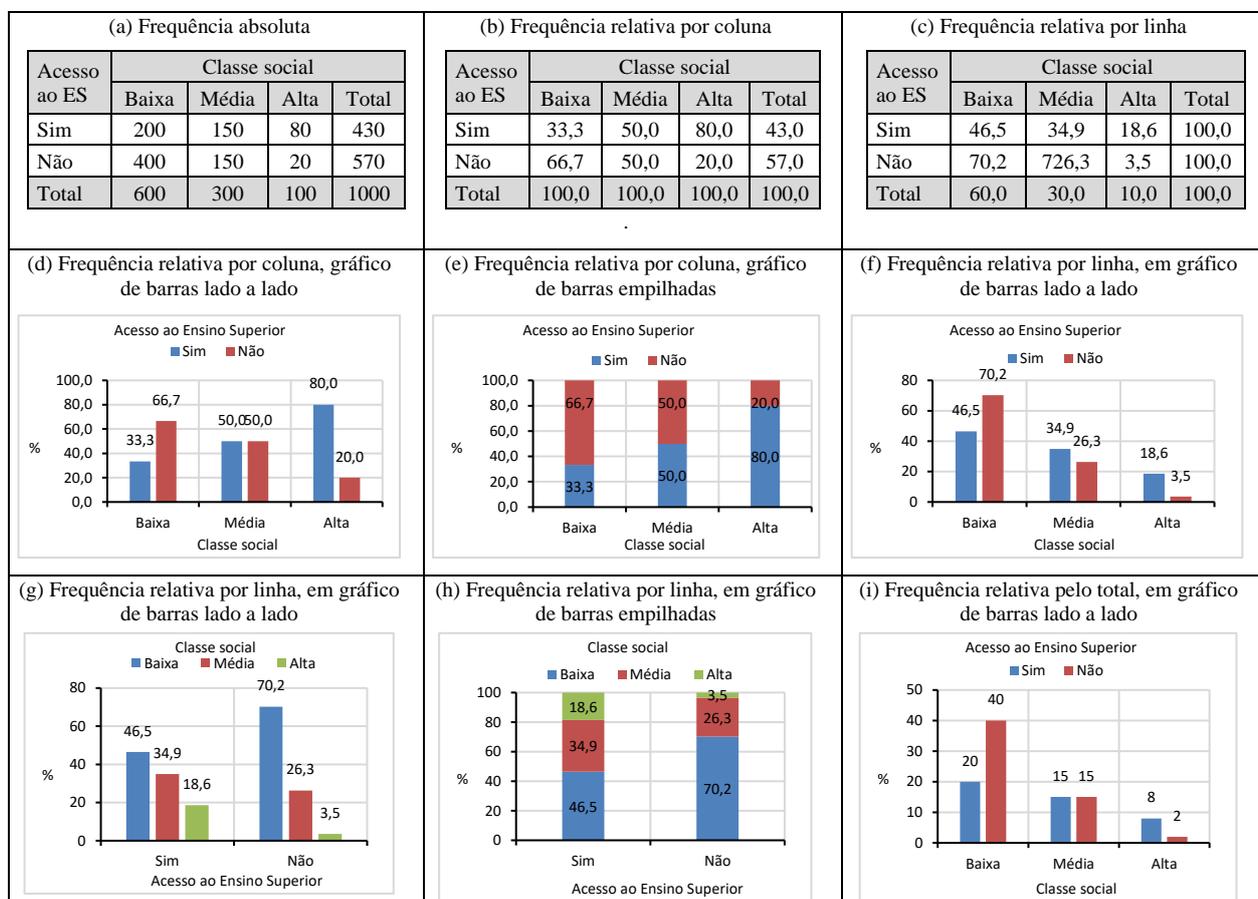
Superior (ES)”. Os participantes deveriam preencher outra tabela com a porcentagem calculada, da forma que lhes parecesse mais conveniente (linha, coluna ou total), a fim de argumentar a existência, ou não, de relação entre essas variáveis, bem como construir um gráfico e escrever uma conclusão acerca da possível relação entre as duas variáveis. Os autores propõem a realização de um *checklist* de perguntas, que o professor pode realizar antes de proceder com o teste de hipótese informal.

Na Figura 7 apresentamos diversas possibilidades das transformações dos dados. Na Figura 7a verificamos que a variável dependente (ES) está na linha, e a variável independente (CS) na coluna. Na Figura 7b apresentamos a frequência relativa por coluna, que foi transformada em gráfico de barra lado a lado (Figura 7d) e gráfico de barras empilhadas (Figura 7e). Essas três representações apresentam a mesma informação, isto é, o percentual de cada classe social que teve acesso ao Ensino Superior.

Na Figura 7c apresentamos a frequência por linha, que foi transformada em gráfico de barras lado a lado (Figura 7f); as Figuras 7g e 7h apresentam a mesma informação, porém, trocando a ordem da variável dependente, uma sendo barras lado a lado e a outra barras empilhadas. Essas quatro representações apresentam a distribuição, por classe social, dos que tiveram acesso ao ensino superior. Na Figura 7i apresentamos a distribuição conjunta, relativa ao total da amostra. Além disso, existem outras possibilidades de representação, sendo que muitas delas podem não ter sentido.

Figura 7.

Possíveis transformações dos dados (construído a partir de Cazorla, Utsumi & Oliveira, 2020, p. 264)



Isto posto, dentre todas essas representações, qual ilustra de forma eficiente a relação entre as duas variáveis consideradas? Como o professor poderia trabalhar essas possibilidades em ambiente papel e lápis/computacional? Isto é desafiador, mas uma estratégia poderia ser dividir a turma em grupos e solicitar que cada grupo elaborasse uma representação e tentasse argumentar com ela.

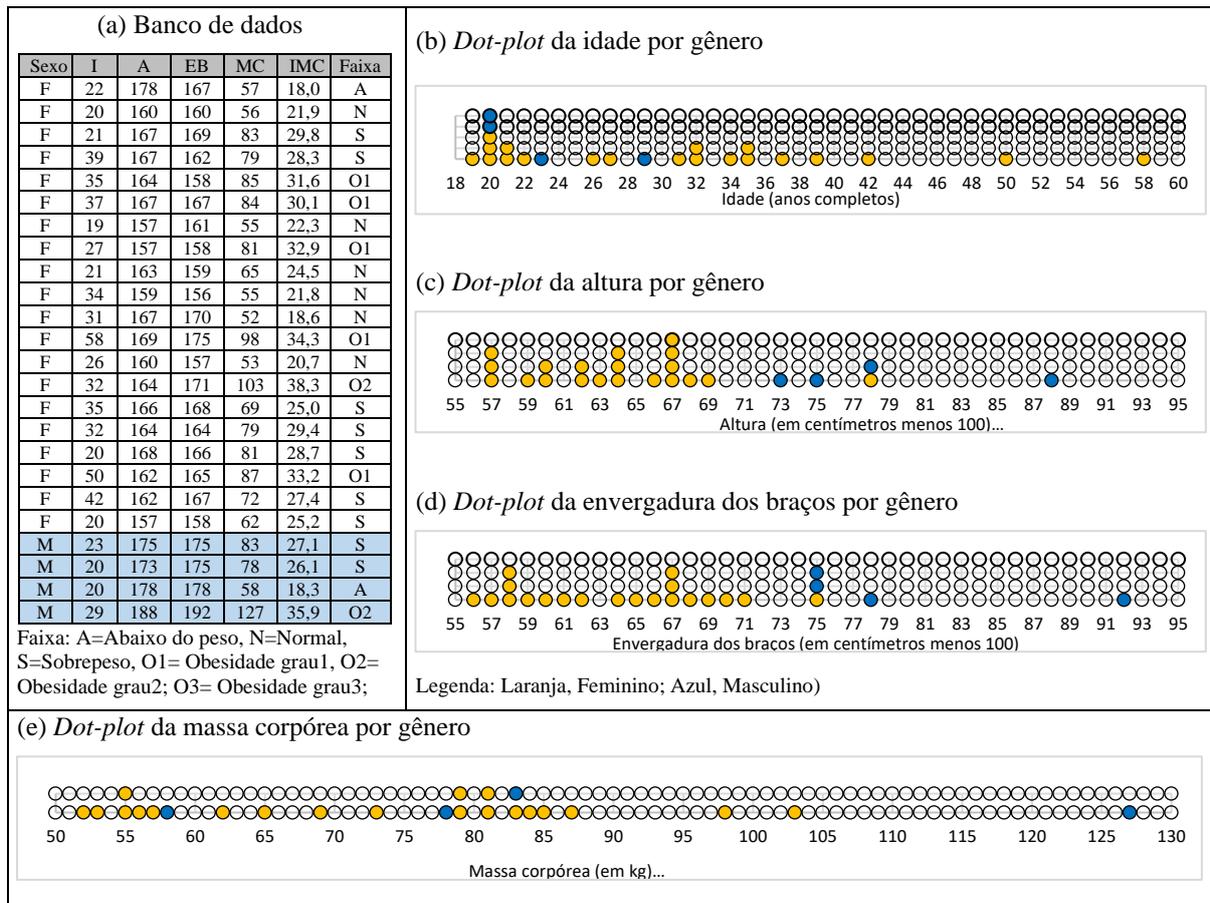
A nosso ver, situações como estas propiciam o desenvolvimento da transnumeração, do pensamento estatístico e do letramento estatístico, entendidos como a capacidade de compreensão de como o fenômeno funciona e de apresentar argumentos com base nas estatísticas e no contexto.

A seguir, apresentamos as possibilidades de transnumeração analisando a relação entre variáveis quantitativas com base nos dados de Santana (2020), que estudou a contribuição das variáveis estatísticas na contextualização da função afim com estudantes do 1º ano de um curso de Técnico em Enfermagem, com o tema obesidade, percorrendo as fases do ciclo investigativo. Os dados coletados foram sobre Gênero (G), Idade (I), Altura (A), Envergadura dos braços (EB), massa corpórea (MC), o IMC e as faixas do IMC. Os estudantes trabalharam com o diagrama de dispersão e foram desafiados a escolher dois pontos e encontrar a função afim que descrevesse a possível relação entre as variáveis. Na Figura 8 apresentamos o banco de dados (8a), os *dot-plots* da Idade (8b), da altura (8c), da envergadura dos braços (8d) e da massa corpórea (8e) por gênero.

O *dot-plot* da idade nos mostra que as mulheres apresentam maior variabilidade do que os homens. Além disso, parece que existem dois grupos, aqueles que saíram diretamente do Ensino Fundamental e aqueles que deixaram a escola e retornaram, buscando um curso profissionalizante. Analisando a distribuição da altura e envergadura dos braços, ambas as variáveis na mesma unidade de medida, verificamos que os homens são mais altos e têm os braços mais compridos, uma vez que os pontos azuis estão localizados mais à direita. Já a massa corpórea apresenta maior variabilidade e não se verifica que os homens sejam mais pesados do que as mulheres. Observamos que não podemos comparar a variabilidade entre estas três variáveis, pois as duas primeiras estão em centímetros, e a terceira em quilogramas. Neste caso, é preciso calcular o coeficiente de variação.

Figura 8.

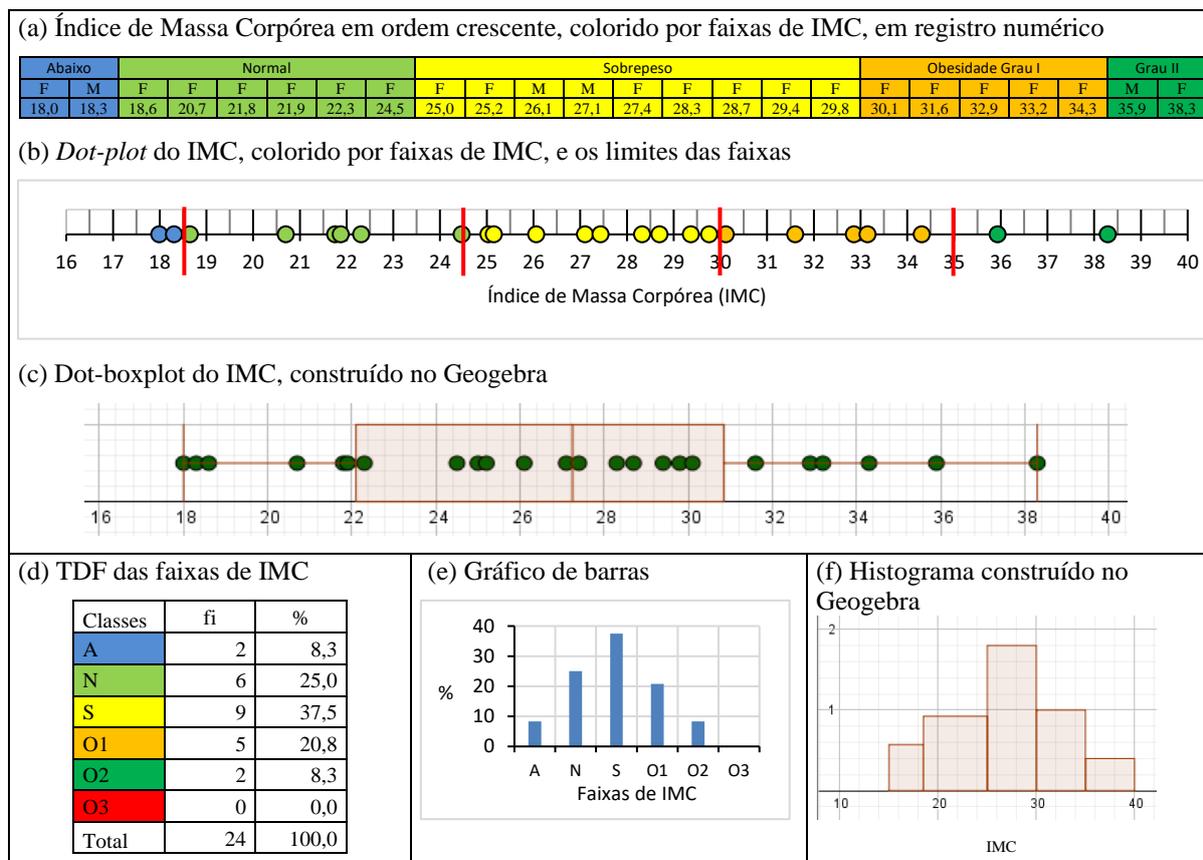
O processo de transnumeração envolvendo variáveis quantitativas (Santana, 2020)



Na Figura 9 apresentamos algumas possibilidades de transformação dos dados do IMC. Aqui, salientamos que, embora altura, envergadura e massa corpórea sejam variáveis contínuas, via de regra, são tratadas como discretas. Já o IMC é uma variável contínua genuína e o arredondamento pode comprometer a sua classificação nas faixas. Devemos ter cuidado ao construir o histograma: se as classes tiverem a mesma amplitude, a altura dos retângulos é proporcional à frequência absoluta/relativa, de tal forma que a soma das áreas seja igual ao total de dados (100% ou 1). Se as classes não tiverem o mesmo tamanho, então, deve-se trabalhar com a densidade, a fim que a soma das áreas dos retângulos recomponha o total de dados (Morettin & Bussab, 2010).

Figura 9.

O processo de transnumeração envolvendo uma variável contínua genuína (Santana, 2020)



Observamos que é muito comum representar os dados da TDF em faixas com um gráfico de barras, o que ocasiona uma distorção dos dados.

No Quadro 3 apresentamos as medidas resumo de três variáveis (altura, massa corpórea e IMC), por gênero. Com relação às MTC, verificamos que a média e a mediana traduzem a informação do *dot-plot*, isto é, as mulheres são mais velhas, mais baixas e menos pesadas do que os homens. Não calculamos a moda, pois, a nosso ver, quando trabalhamos variáveis que tomam muitos valores diferentes, essa medida perde o sentido, a menos que seja calculada a partir da TDF ou do histograma. Ainda, acreditamos que o esforço não compensa o pouco ganho que esta medida traz para a compreensão da tendência dos dados.

Quadro 3.

Medidas resumo da altura, massa corpórea e IMC, por gênero

Medidas resumo		Altura			Massa corpórea			IMC		
		Fem	Mas	Geral	Fem	Mas	Geral	Fem	Mas	Geral
MTC	Média	163,9	178,5	166,3	72,8	86,5	75,1	27,1	26,9	27,1
	Mediana	164,0	176,5	165,0	75,5	80,5	78,5	27,9	26,6	27,3
MD	Mínimo	157,0	173,0	157,0	52,0	58,0	52,0	18,0	18,3	18,0
	Máximo	178,0	188,0	188,0	103,0	127,0	127,0	38,3	35,9	38,3
	Amplitude total (AT)	21,0	15,0	31,0	51,0	69,0	75,0	20,3	17,6	20,3
	Desvio padrão (DP)	4,95	5,77	7,45	15,04	25,18	17,89	5,40	6,25	5,55
	Coef. de variação (CV%)	3,0	3,2	4,5	20,7	29,1	23,8	19,9	23,3	20,5
	Amplitude interquartilica (AI)	7,0	6,0	6,8	26,5	21,0	25,5	8,3	5,2	8,3
MP	Quartil 1	160,0	174,5	161,5	56,8	73,0	57,8	22,2	24,1	22,2
	Quartil 3	167,0	180,5	168,3	83,3	94,0	83,3	30,5	29,3	30,5

Quanto à variabilidade, verificamos que a massa corpórea e o IMC apresentam maior variabilidade do que a altura, e que os homens apresentam maior variabilidade absoluta (DP) e relativa (CV) do que as mulheres. Neste ponto, o professor poderia explorar a aderência entre a distribuição dos dados e as medidas resumo, conduzindo os estudantes a partir de perguntas do tipo: quem são mais altos, as mulheres ou os homens; qual dos dois grupos apresenta maior variabilidade, dentre outras, como realizado por Santana (2020), que solicitou aos estudantes que “chutassem” o valor da média, observando a distribuição dos dados no *dot-plot*.

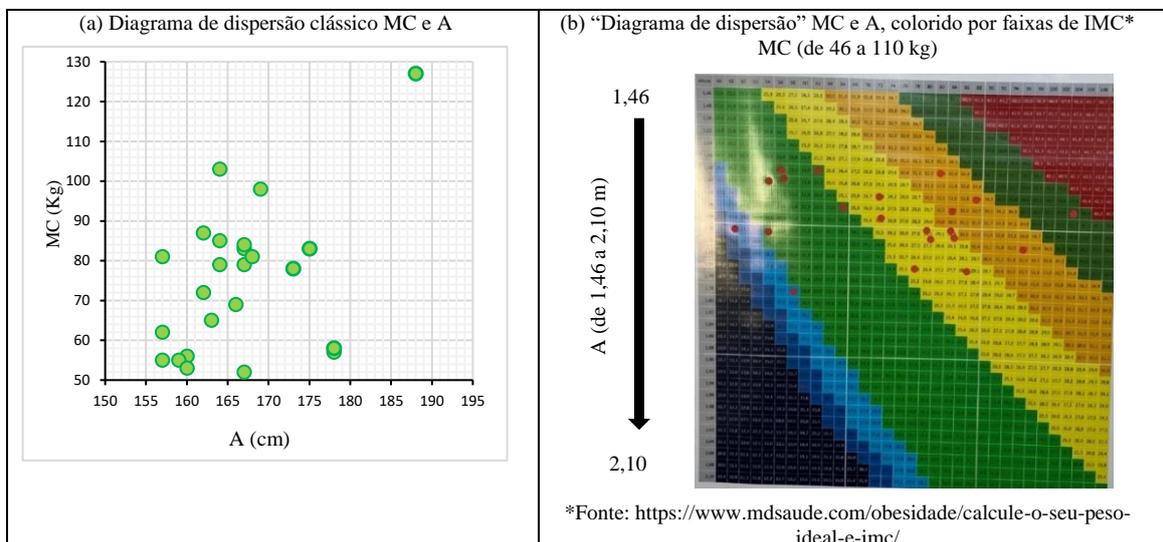
Para analisar a relação entre a massa corpórea e a altura (análise bivariada), na busca de padrões sobre o IMC, Santana (2020) trabalhou com dois diagramas de dispersão. O primeiro no padrão clássico (Figura 10a), que tinha a altura (em cm) na abscissa (v. independente), variando de 150 até 200 cm, e a massa corpórea (em kg) na ordenada (v. dependente), variando de 50 a 130 kg, iniciando no ponto (150, 50), no canto inferior esquerdo. A leitura deste gráfico é iniciada no canto inferior esquerdo, da esquerda para a direita, e de baixo para cima. Nesse diagrama, podemos ver que as pessoas mais baixas e mais magras se encontram no canto inferior esquerdo, e as mais altas e mais pesadas no canto superior direito.

Já o diagrama colorido, próprio para analisar o IMC (Figura 10b), tinha a altura (em metros) na ordenada (v. independente), variando de 1,46 até 2,10 m, e a massa corpórea (em kg) na

abscissa (v. dependente), variando de 46 a 110 kg, iniciando no ponto em (1,46, 46), no canto superior esquerdo. Na interseção das células (unitárias) está o valor do IMC. O diagrama é colorido por faixas de IMC. A leitura deste gráfico é iniciada no canto superior esquerdo, da esquerda para a direita, e de cima para baixo. Nesse diagrama, podemos ver que as pessoas mais baixas e mais magras se encontram no canto superior esquerdo, as baixas e obesas no canto superior direito, e as altas e magras no canto inferior esquerdo. As pessoas dentro do padrão de normalidade se situam na diagonal.

Figura 10.

O processo de transnumeração envolvendo o nível de obesidade (Santana, 2020)



Analisando os diagramas, verificamos que ambos mostram os mesmos dados, mas, enquanto o diagrama de dispersão clássico não permite capturar a mensagem, pois a nuvem de pontos é totalmente dispersa sem nenhum padrão, o diagrama colorido mostra de forma contundente a tendência à obesidade dos estudantes. Santana (2020) relata o susto que os estudantes tomaram quando constatarem essa situação.

Os exemplos ilustram a importância de dois aspectos a serem levados em consideração quando ensinamos Estatística na Educação Básica. De um lado, a necessária compreensão de como os dados brutos podem ser transformados em estatísticas e, portanto, a relevância do conhecimento e domínio das técnicas transnumerativas. De outro lado, a importância do

objetivo da investigação e das perguntas orientadoras, que se tornarão guias do processo de geração de estatísticas e informação, isto é, alcançar a transnumeração. Ambas as competências devem ser trabalhadas de forma simultânea, imbricadas, de maneira a acionar o raciocínio inferencial e o uso ético da informação estatística.

Ensinar Estatística visando a transnumeração ultrapassa os cálculos matemáticos, devendo oportunizar aos estudantes a tomada de consciência de que serão eles os que tomarão decisões que podem impactar a sua vida e a vida de sua comunidade.

Ensinar Estatística nesta perspectiva tem implicações pedagógicas, pois, para conseguir colocar o estudante no centro da análise de dados, o professor necessitaria assumir o papel de condutor do processo de transnumeração. Isto é, mostrar aos estudantes as vantagens e desvantagens das estatísticas (tabelas, gráficos, medidas resumo e suas propriedades matemáticas); as implicações da escolha dessas estatísticas na compreensão do fenômeno/problema/situação em estudo e suas implicações na tomada de decisões.

Observando esses resultados, podemos concluir que um trabalho desta envergadura em sala de aula pode auxiliar os estudantes a compreenderem o real significado de transnumeração, de modo que seu desenvolvimento requer um planejamento minucioso.

Considerações Finais

Há bastantes evidências de experiências de trabalhos com projetos, nos quais os estudantes coletam dados multivariados. Recomendamos a leitura dos trabalhos de Santos (2015); Oliveira Junior e Vieira (2015); Schreiber e Porciúncula (2019); e de Oliveira e Paim (2019), que realizaram mapeamentos/estado da arte da produção científica no campo da Educação Estatística na Educação Básica, onde podem ser encontradas boa parte das pesquisas brasileiras que envolvem os conceitos estatísticos aqui trabalhados.

Apesar de ser desafiador ensinar Estatística nesta perspectiva, verificamos na BNCC uma nova tendência de trabalho no Ensino Médio com os temas transversais e com os itinerários

formativos, como ressaltam Cazorla e Giordano (2021), ou seja, existem condições para se trabalhar em uma perspectiva transnumerativa.

Neste artigo, apresentamos as técnicas transnumerativas e sua importância e refletimos sobre possíveis implicações pedagógicas da transformação dos dados brutos em estatísticas, e destas em informação. Nossas reflexões sugerem que é necessário apontar possibilidades de transnumeração dos dados para os professores, a qual pode ser feita com dados já coletados ou percorrendo todas as fases da investigação estatística e, assim, contribuir com o desenvolvimento das habilidades e competências normatizadas pela BNCC.

Agradecimentos

Os autores agradecem a Cláudio Vitor Santana pela cessão do banco de dados de sua dissertação de mestrado (Santana, 2020).

Referências

- Almeida, C., Sousa, H. & Cazorla, I. (2021). Letramento estatístico na Educação Básica: os desafios de ensinar o diagrama da caixa (*box-plot*) em contexto. *EMP*, 23(1), p. 499-529. <http://dx.doi.org/10.23925/1983-3156.2021v23i1p499-529>
- Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores* [Tese de doutorado em Didática da Matemática, Universidade de Granada]. <http://www.ugr.es/~batanero/pages/ARTICULOS/artega.pdf>.
- Batanero, C. (2000). Significado y comprensión de las medidas de tendencia central. *Uno*, 25, 41-58.
- Batanero, C. (2019), Treinta años de investigación en educación estadística: Reflexiones y desafíos. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. www.ugr.es/local/fqm126/civeest.html.
- Buehring, R. (2006). *Uma análise de dados no início da escolaridade: uma realização de ensino por meio dos registros de representação semiótica* [Dissertação de mestrado em Educação Científica e Tecnológica, Universidade Federal de Santa Catarina].
- Cazorla, I. (2002). *A relação entre as habilidades viso-pictóricas e o domínio de conceitos estatísticos na leitura de gráficos* [Tese de Doutorado em Educação, Universidade Estadual de Campinas].

- Cazorla, I. & Castro, F. (2008). Papel da Estatística na leitura do mundo: o letramento estatístico. *Publicatio UEPG: Ciências Humanas, Ciências Sociais Aplicadas, Linguística, Letras e Artes*, 16 (1), p. 45-53.
- Cazorla, I. & Oliveira, M. (2010). Para saber mais. In I. Cazorla e E. Santana (Orgs.). *Do tratamento da Informação ao Letramento Estatístico* (pp. 113-144). Itabuna: Via Litterarum.
- Cazorla, I., Magina, S., Gitirana, V. & Guimarães, G. (2017). *Estatística para os anos iniciais do ensino fundamental*. Brasília: Sociedade Brasileira de Educação Matemática – SBEM. http://www.sbem.com.br/files/ebook_sbem.pdf.
- Cazorla, I., Santana, E. & Utsumi, M. (2019). O campo conceitual da média aritmética: uma primeira aproximação conceitual. *Revemat*, 14, *Edição Especial Educação Estatística*, p. 1-21.
- Cazorla, I., Henriques, A. & Santana, C. (2020). O papel dos ostensivos na representação de variáveis estatísticas qualitativas. *Bolema*, 34, 1243-1263.
- Cazorla, I., Utsumi, M. & Oliveira, T. (2020). Reflexões sobre o ensino de tabelas de dupla entrada a partir do raciocínio inferencial informal. *Hipátia*, 5(2), 246-270.
- Cazorla, I., Utsumi, M. & Santana, E. (2020). Desempenho em Estatística de estudantes do Ensino Fundamental, no contexto do D-Estat. *Zetekiké*, 28, p. 1-25. DOI: 10.20396/zet.v28i0.8656917.
- Cazorla, I., Utsumi, M. & Monteiro, C. (2020a). Variáveis estatísticas e seus gráficos. In M. M. Gea. R. Álvarez-Arroyo y J. A. Garzón (Eds.), *Seminario Hispano Brasileño de Educación Estadística*. (pp. 69-72). Granada: Grupo PAI FQM-126. <https://digibug.ugr.es/bitstream/handle/10481/61731/Actas%20definitivas.pdf?sequence=1&isAllowed=y>.
- Cazorla, I., Utsumi, M. & Monteiro, C. (2020b). Reflexões sobre as variáveis estatísticas e suas representações em gráficos. In C. Campos & A. Perin (Orgs.). *Investigações Hispano-Brasileiras em Educação Estatística*. (pp. 189-195). Taubaté: Editora Akademy. <https://www.akademy.com.br/wp-content/uploads/2020/10/Livro-GT12.pdf>.
- Cazorla, I., Utsumi, M. & Monteiro, C. (2021). Variáveis estatísticas e suas representações em gráficos: reflexões para seu ensino. *Números*, 106, 23-32.
- Cazorla, I. & Giordano, C. (2021). O papel do letramento estatístico na implementação dos Temas Contemporâneos Transversais da BNCC. In C. Monteiro & L. Carvalho (Orgs.), *Temas emergentes em Letramento Estatístico*. Ebook, UFPE. (no prelo).
- Chick, H. (2003). Transnumeration and the art of data representation. In L. Bragg, C. Campbell, G. Herbert, e J. Mousley (Eds.), *Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia* (pp. 207-214). Sydney: MERGA.
- Chick, H. (2004). Tools for transnumeration: Early stages in the art of data representation. In I. Putt, R. Faragher, e M. McLean (Eds.), *Proceedings of the 27th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 167-174). Sydney: MERGA.
- Cobo, B. (2003). *Significados de las medidas de posición central para los estudiantes de secundaria* [Tese de doutorado em Didática da Matemática, Universidade de Granada]. <https://www.ugr.es/~batanero/pages/ARTICULOS/tesiscobo.pdf>.

- Crespo, A. (2002). *Estatística fácil*. São Paulo: Saraiva.
- Figueiredo, T. (2017). *Mediana e quartis: um caso de estudo das dificuldades de aprendizagem de alunos do 8.º ano de escolaridade* [Dissertação de mestrado em Matemática para professores, Universidade de Aveiro]. https://ria.ua.pt/bitstream/10773/21920/1/Disserta%C3%A7%C3%A3o_Teresa_Figueiredo.pdf.
- Gal, I. (2002). Adult statistical literacy: Meanings, components, responsibilities, *International Statistical Review*, 70(1), 1-25.
- Guimarães, G., Luz, P. & Ramos, M. (2011). Classificar: uma atividade difícil para alunos e professores dos anos iniciais do Ensino Fundamental? *Anais do XIII CIAEM*, UFPE: Recife.
- Henriques, A. & Oliveira, H. (2019). Transnumeração em contexto STEM: Um cenário de aprendizagem desenvolvido por futuros professores com alunos do 8.º ano. Em J. M. Contreras, M. M. Gea, M. M. López-Martín & E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. www.ugr.es/local/fqm126/civeest.html.
- Kataoka, V. & Hernandez, H. (2010). Perfil da turma. In I. Cazorla e E. Santana (Orgs.). *Do Tratamento da Informação ao Letramento Estatístico*. (pp. 23-44). Itabuna: Via Litterarum.
- Makar, K. & Rubin, A. (2009). A framework for thinking about informal Statistical inference. *Statistics Education Research Journal*, 8(1), 82-105. [https://www.stat.auckland.ac.nz/~iase/serj/SERJ8\(1\).pdf#page=85](https://www.stat.auckland.ac.nz/~iase/serj/SERJ8(1).pdf#page=85).
- Matos, L. (2019). *Leitura e construção de gráficos estatísticos por estudantes do ensino fundamental* [Dissertação de mestrado em Educação Matemática, Universidade Estadual de Santa Cruz].
- Ministério da Educação - MEC. (1997). *Parâmetros Curriculares Nacionais*. Brasília: MEC.
- Ministério da Educação - MEC. (2018). *Base Nacional Comum Curricular – BNCC. Educação é a base*. Brasília: MEC.
- Ministério da Educação - MEC. (2019). *Temas Contemporâneos Transversais na BNCC: Contexto Histórico e Pressupostos Pedagógicos*. Brasília: MEC. http://basenacionalcomum.mec.gov.br/images/implementacao/contextualizacao_temas_contemporaneos.pdf.
- Monteiro, C. & Ainley, J. (2004). Exploring the complexity of the interpretation of media graphs. In O. McNamara, & R. Barwell (Eds.), *Research in Mathematics Education*, 6(1), 115-128. Taylor & Francis. <https://doi.org/10.1080/14794800008520133>.
- Moreno, M. (2010). *Ensino e aprendizagem de estatística com ênfase na variabilidade: um estudo com alunos de um curso de licenciatura em Matemática* [Dissertação de mestrado em Educação Matemática, Pontifícia Universidade Católica de São Paulo].
- Moreno, M. M. (2014). *Utilização do dotplot e do boxplot na aprendizagem da variabilidade estatística no Ensino Médio* [Dissertação de mestrado em Educação Matemática, Universidade Estadual de Santa Cruz]. <http://www.biblioteca.uesc.br/biblioteca/bdtd/201270066D.pdf>.
- Morettin, P. & Bussab, W. (2010). *Estatística Básica*. São Paulo: Saraiva.

- Oliveira Junior, A. & Vieira, M. (2015). A Educação Estatística nos anos iniciais do Ensino Fundamental: estado da arte. *Triângulo*, 8(2), 33-52.
- Oliveira, P. & Paim, S. (2019). O mapeamento de pesquisas brasileiras sobre o letramento estatístico de 2006 a 2018. *Revista Brasileira de Educação em Ciências e Educação Matemática*, 3(2), 669-699. <https://doi.org/10.33238/ReBECCEM.2019.v.3.n.2.22631>.
- Oliveira, T. (2020). *Contribuições das disciplinas de estatística na formação do futuro professor de Matemática para a Educação Básica* [Dissertação de mestrado em Educação Matemática, Universidade Estadual de Santa Cruz].
- Pfannkuch, M. & Rubick, A. (2002). An exploration of students' statistical thinking with given data. *Statistics Education Research Journal*, 1(2), 4-21.
- Santana, C. (2020). *Relações entre variáveis estatísticas na contextualização e apropriação da função afim* [Dissertação de Mestrado em Educação Matemática, Universidade Estadual de Santa Cruz].
- Santos, R. (2015). *Estados da arte e história da pesquisa em Educação Estatística em programas brasileiros de pós-graduação* [Tese de doutorado em Educação, Universidade Estadual de Campinas].
- Schreiber, K. & Porciúncula, M. (2019). Mapeamento das pesquisas sobre Educação Estatística na Biblioteca Digital Brasileira de Teses e Dissertações: um olhar para a formação do professor de matemática. *Revemat*, 14, 1-17. <https://doi.org/105007/1981-1322.2019e.62799>.
- Silva, C. (2007). *Pensamento estatístico e raciocínio sobre variação: um estudo com professores de matemática* [Tese de doutorado em Educação Matemática, Pontifícia Universidade Católica de São Paulo].
- Silva, C., Magina, S. & Silva, E. (2010). Homem Vitruviano. In I. Cazorla e E. Santana (Orgs.). *Do Tratamento da Informação ao Letramento Estatístico*. (pp. 80-93) Itabuna: Via Litterarum.
- Silva, S. & Cazorla, I. (2020). O ensino de Estatística e a percepção de estudantes sobre o impacto do uso dos agrotóxicos. *ReBECCEM*, 4(4), 694-720.
- Silva-Junior, A. (2018). *Efeitos do Ciclo Investigativo PPDAC e das transformações de representações semióticas no desenvolvimento de conceitos estatísticos no Ensino Fundamental* [Dissertação de mestrado em Educação Matemática, Universidade Estadual de Santa Cruz].
- Souza, A. (2007). *A Educação Estatística na Infância* [Dissertação de mestrado em Educação em Ciências e Matemática, Universidade do Cruzeiro do Sul].
- Wild, C. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67, 223-265. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>.

Recebido: 17/03/2021

Aceito: 15/05/2021