



Instituto de  
MATEMÁTICA  
E ESTATÍSTICA

UFRGS



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DEPARTAMENTO DE ESTATÍSTICA

**A EDUCAÇÃO ESTATÍSTICA ATRAVÉS DA VISUALIZAÇÃO DE DADOS DE  
COVID-19 NO ESTADO DO RIO GRANDE DO SUL**

**PEDRO VICTOR DE MEDEIROS FELIPE**

Porto Alegre  
2021

**PEDRO VICTOR DE MEDEIROS FELIPE**

**A EDUCAÇÃO ESTATÍSTICA ATRAVÉS DA VISUALIZAÇÃO DE DADOS DE COVID-19 NO ESTADO DO RIO GRANDE DO SUL**

Trabalho de Conclusão de Curso submetido como requisito parcial para a obtenção do grau de Bacharel em Estatística, em curso de graduação oferecido pela Universidade Federal do Rio Grande do Sul.

Orientador Metodológico: Professora Dra. Luciana Neves Nunes

Banca Examinadora: Professora Dra. Vanessa Bielefeldt Leotti

## **AGRADECIMENTOS**

Agradeço aos meus pais, Mônica Medeiros e Ronaldo Felipe, por todo apoio que me deram e toda influência que exerceram sobre mim para me tornar o homem que sou hoje.

Aos meus amigos e colegas de apartamento, Henrique Moreno, Lucas Moraes e Mauro Mizuta, que nesses anos de graduação me ajudaram a seguir em frente e sempre estiveram comigo em todos os momentos.

À minha orientadora, Luciana Neves Nunes, por toda sua orientação neste projeto, mesmo nos momentos mais delicados, por sua paciência, e por ter abraçado a ideia.

À Professora Vanessa Leotti, por aceitar o convite e fazer parte da banca avaliadora.

Por fim agradeço a todos colegas de sala e amigos que estiveram presentes nesses anos e contribuíram para minha graduação.

## RESUMO

Este trabalho tem como base a dificuldade, tanto de professores como de estudantes, para o ensino e aprendizagem de Estatística e o desenvolvimento do Pensamento Estatístico, principalmente na área de visualização de dados. A partir dessa fundamentação, tem-se como objetivo a apresentação de uma Proposta Didática de Visualização de Dados para auxiliar os professores nas abordagens do conteúdo de visualização de dados de modo que incentive o desenvolvimento do Pensamento e Raciocínio Estatístico nos estudantes, em diferentes níveis de ensino. Para tanto, foram apresentados os conceitos teóricos do Pensamento e Raciocínio Estatístico, e o que compõe um gráfico a partir da Gramática dos Gráficos, e estruturado um Guia composto por três pilares para uma abordagem dinâmica para interpretação de gráficos, buscando o desenvolvimento das habilidades necessárias para formar cidadãos com Pensamento Crítico em Estatística. Ainda, na atividade proposta, há incentivo ao uso da Modelagem Matemática e as ideias de transnumeração, para sua aplicação. O Guia aqui apresentado traz exemplos de gráficos no contexto da pandemia da Covid-19, entretanto os professores que aderirem a estratégia proposta, poderão utilizar diferentes temáticas, oportunizando atividades interdisciplinares para seus estudantes. A partir da ideia destas orientações, espera-se que este trabalho tenha contribuição para a divulgação da Estatística e para a área da Educação Estatística.

**Palavras-chave:** Educação Estatística. Pensamento Estatístico. Raciocínio Estatístico. Visualização de Dados.

## ABSTRACT

This work is based on the difficulty, both of teachers and students, for teaching and learning Statistics and the development of Statistical Thinking, mainly in the data visualization field. For this reason, the objective is to introduce a Didactic Proposal in Data Visualization Practices to assist teachers in approaching the data visualization content in a way that encourages the development of Statistical Thinking and Reasoning in students, of different levels of education. For that, the theoretical concepts of Thought and Statistical Reasoning were described, and what a graph is made of, from the Grammar of Graphs point of view, and a Guide composed of three pillars for a dynamic approach to the interpretation of graphs, seeking the development of the necessary skills to form citizens with Critical Thinking in Statistics. Still, in the proposed activity, there is incentive to the use of Mathematical Modeling and the ideas of transnumeration, for its application. The Guide presented here brings examples of graphs in the context of the Covid-19 pandemic, however, teachers who adhere to the proposed strategy can use different themes, providing interdisciplinary activities for their students. Based on the idea of these guidelines, it is expected that this work will contribute to the dissemination of Statistics and to the area of Statistical Education.

**Keywords:** Statistical Education. Statistical Thinking. Statistical Reasoning. Data Visualization.

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>7</b>
<b>2. EDUCAÇÃO ESTATÍSTICA.....</b>	<b>8</b>
2.1. Letramento e Pensamento Estatístico.....	8
2.2. Transnumeração.....	11
2.3. Letramento em Big Data.....	15
2.4. Representação Semiótica.....	16
<b>3. RECURSOS TECNOLÓGICOS.....</b>	<b>18</b>
3.1. Visualização de Dados.....	18
3.2. A Gramática dos Gráficos.....	19
<b>4. PROPOSTA DIDÁTICA.....</b>	<b>23</b>
4.1. Pandemia COVID-19.....	24
4.2. Visualização de Dados Aplicados.....	25
<b>5. CONSIDERAÇÕES FINAIS.....</b>	<b>35</b>
<b>REFERÊNCIAS.....</b>	<b>37</b>

## 1 INTRODUÇÃO

A Estatística, de uma forma perceptível, está presente a todo momento na rotina das pessoas, podendo aparecer de maneira mais objetiva como em informações apresentadas em noticiários e programas de TV, ou até mesmo de uma maneira mais sutil, como um momento determinante de uma rolagem de dados que ocorre em jogos de tabuleiro. Contudo, apesar de estar presente nas mais diversas áreas, ainda assim é possível observar a dificuldade de muitos sujeitos ao assimilar e lidar, direta ou indiretamente, com essas informações estatísticas. Isso faz com que se perceba que existe a necessidade de uma nova abordagem no Ensino da Estatística, a fim de auxiliar tanto professores quanto alunos a abordar essa disciplina de forma a ir além de fórmulas e cálculos matemáticos, contribuindo para a formação de cidadãos e cidadãs críticas.

A partir desse princípio, a ideia deste trabalho é propor uma abordagem para o Ensino de Estatística diferente do ensino tradicional. Para tanto, focando no tema “Visualização de Dados”, será criada uma Proposta Didática para Educação Estatística, cujo objetivo é auxiliar professores, principalmente da Educação Básica, em uma nova abordagem de ensino, visando principalmente desenvolver o pensamento e raciocínio estatístico no público-alvo de suas aulas.

Com o intuito de desenvolver o tema proposto, este estudo é iniciado com essa introdução que destaca a presença da Estatística e sua importância. Em seguida, o segundo capítulo aborda a Educação Estatística como um objeto de estudo, contextualizando a Base Nacional Curricular Comum (BNCC) e explicando conceitos de Letramento, Raciocínio e Pensamento Estatístico, Transnumeração, Letramento em *Big Data* e Representação Semiótica. Por sua vez, o terceiro capítulo contempla os recursos tecnológicos utilizados na Visualização de Dados e explica a estrutura que compõe um gráfico segundo a “Gramática dos Gráficos”.

A partir dessas premissas, o quarto capítulo traz a consolidação da Proposta Didática para Educação Estatística, trazendo uma abordagem prática para o ensino de Visualização de Dados, cujo tema é a pandemia de COVID-19 no estado do Rio Grande do Sul, com a intenção de trazer uma temática realista e atual para a sala de aula.

## 2. EDUCAÇÃO ESTATÍSTICA

### 2.1 Letramento, Raciocínio e Pensamento Estatístico

O mundo tecnológico no qual vivemos nos permite acesso a um número imensurável de informações e dados, disponíveis muitas vezes em tempo real, e a Estatística é apresentada como uma forma de dar credibilidade para conselhos, argumentos, notícias e até mesmo anúncios de produtos. Contudo, a aptidão de avaliar apropriadamente a evidência (informações) e concluir baseado em ditos dados é uma habilidade que todo estudante deveria aprender como parte de seu programa educacional.

Antes de citarmos as definições que serão os pilares deste estudo, precisamos entender o contexto do Ensino Básico Brasileiro. Segundo o Artigo 21º da Lei nº 9.394 (BRASIL, 1996), a Educação Básica é formada pela educação infantil, ensino fundamental e ensino médio. Com a premissa desse Artigo, “A Base Nacional Comum Curricular (BNCC) é um documento de caráter normativo que define o conjunto orgânico e progressivo de aprendizagens essenciais que todos os alunos devem desenvolver ao longo das etapas e modalidades da Educação Básica” (BRASIL, 2017, p. 7) e está fundamentada em dez competências gerais.

Na BNCC, competência é definida como a mobilização de conhecimentos (conceitos e procedimentos), habilidades (práticas, cognitivas e socioemocionais), atitudes e valores para resolver demandas complexas da vida cotidiana, do pleno exercício da cidadania e do mundo do trabalho. (BRASIL, 2017, p.8).

Dentre essas competências destacamos a de número 2 que sugere “Exercitar a curiosidade intelectual e recorrer à abordagem própria das ciências, incluindo a investigação, a reflexão, a análise crítica...”. A competência 7 aborda “Argumentar com base em fatos, dados e informações confiáveis, para formular, negociar e defender ideias, pontos de vista e decisões...”. (BRASIL, 2017, pp. 9-10).

Na BNCC, podemos considerar que a Educação Estatística compõe uma abordagem do ensino da Matemática, mais especificamente, faz parte de uma das cinco

“unidades temáticas” que orienta a formulação de habilidades a serem desenvolvidas ao longo do Ensino Fundamental (BRASIL, 2017, p.268), sendo chamada de “Probabilidade e Estatística”.

Dado o objetivo deste estudo e levando em consideração uma abordagem voltada para o Raciocínio Estatístico, definido a seguir, destacamos no 9º ano do Ensino Fundamental o objeto de conhecimento EF09MA21, “Análise de gráficos divulgados pela mídia: elementos que podem induzir a erros de leitura” (BRASIL, 2017, p. 318), que de acordo com a BNCC busca desenvolver a habilidade de:

...identificar elementos que podem induzir, às vezes propositadamente, erros de leitura, como escalas inapropriadas, legendas não explicitadas corretamente, omissão de informações importantes (fontes e datas), entre outros. (BRASIL, 2017, p. 319).

Contudo, segundo Moore (1998), a Estatística é uma disciplina por si só, com seu próprio núcleo ao contrário de ser, conforme abordada na BNCC, um “braço” da matemática. Entretanto, estudantes equiparam a Estatística com a Matemática e esperam o foco em números, computações, fórmulas e uma resposta correta (BEN-ZVI & GARFIELD, 2004). No Brasil, o cenário do ensino tradicional de Matemática, no qual o professor expõe o conteúdo em um primeiro momento, e em seguida os alunos resolvem os exercícios propostos, está presente na maioria das instituições de ensino, e refere a uma abordagem mais técnica do ensino, incluindo os conteúdos de Estatística. Podemos considerar que o currículo brasileiro trata o ensino de Estatística como uma ramificação da Matemática. Entretanto, pesquisas recentes na área de Educação Estatística têm indicado a importância e necessidade de que as competências estatísticas de letramento, raciocínio e pensamento devem ser desenvolvidas nos diferentes níveis de ensino para que se construa uma sociedade com cidadãos e cidadãs críticas (Lopes et al, 2020).

Ben-Zvi e Garfield (2004) nos dizem que, apesar de não haver uma concordância formal a respeito das definições e distinções feitas sobre letramento, raciocínio e pensamento estatístico, a lista a seguir resume as opiniões dos pesquisadores. Garfield, delMas e Chance (2003) propuseram as seguintes definições:

- **Letramento estatístico** inclui habilidades básicas, mas, ainda assim, importantes, que podem ser utilizadas no entendimento de informações estatísticas e resultados de pesquisa. Essas habilidades incluem ser apto a organizar dados, construir e apresentar tabelas, e trabalhar com diferentes representações de dados. Letramento estatístico também inclui a compreensão de conceitos, vocabulário e simbologia, assim como probabilidade como uma medida de incerteza.
- **Raciocínio estatístico** pode ser definido como o modo como pessoas raciocinam com ideias estatísticas e dão sentido às informações estatísticas. Isso envolve interpretações baseadas em conjuntos, representações ou resumos estatísticos de dados. Raciocínio estatístico pode envolver a conexão de um conceito com outro, como, por exemplo, média e amplitude, ou pode combinar ideias referentes a dados e chance. Raciocínio significa entender e conseguir explicar processos estatísticos e interpretar, em sua totalidade, resultados estatísticos.
- **Pensamento estatístico** envolve o entendimento do porquê e do como investigações estatísticas são conduzidas e as “grandes ideias” que constroem a base dessas investigações. Essas ideias incluem a natureza onipresente da variação, além de onde e como usar os métodos apropriados para análise de dados, como resumos numéricos ou visualização de dados. O pensamento estatístico envolve também a compreensão da natureza da amostragem, como fazemos inferência a partir de amostras de uma população e porque designados experimentos são necessários a fim de estabelecer causalidade. Também inclui saber como os modelos podem ser usados para simular fenômenos aleatórios, como dados são produzidos para estimar probabilidades, ser apto a compreender e utilizar o contexto de um problema na formação de investigações e ao chegar em conclusões. Finalmente, pensadores estatísticos estão aptos para criticar e avaliar os resultados de um problema resolvido ou um estudo estatístico.

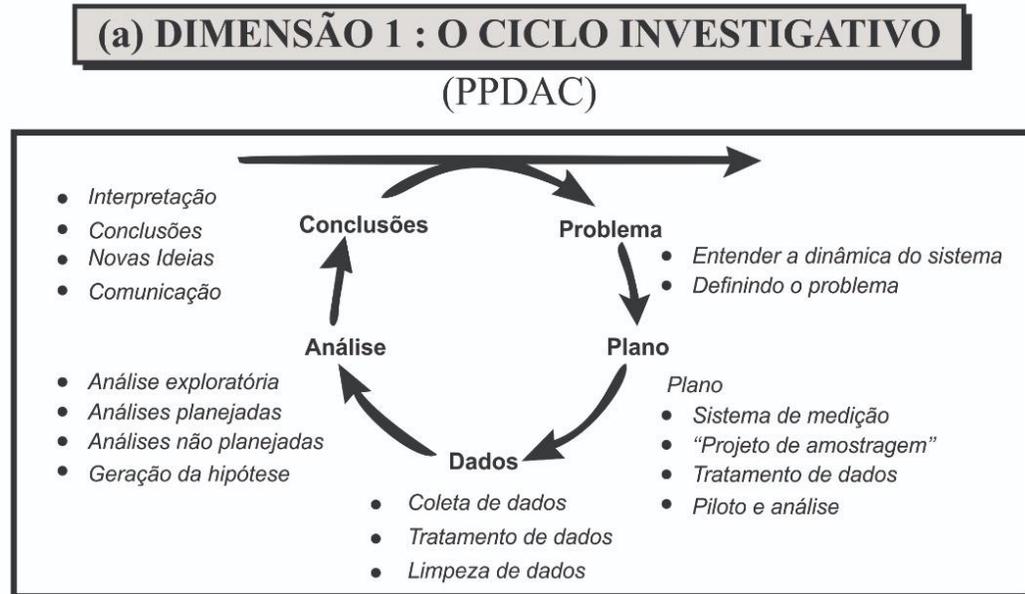
Destarte, o desafio do professor em engajar os alunos a desenvolverem tais habilidades é extenso. Dado o objetivo deste trabalho, a criação da Proposta Didática para Educação Estatística tem como fundamento a proposta da Modelagem

Matemática definida por Barbosa (2001) como “... um ambiente de aprendizagem no qual os alunos são convidados a indagar e/ou investigar situações oriundas de outras áreas da realidade”, trazendo assim situações do dia a dia dos alunos como ferramenta para facilitar a compreensão dos conteúdos de Matemática. Mais especificamente, utilizaremos uma das formas propostas por Barbosa (2001) de se trabalhar com Modelagem Matemática, o Caso 2, onde o professor traz a situação-problema, mas os alunos ficam responsáveis pela coleta das informações necessárias para resolvê-la e pela simplificação.

## **2.2 Fundamentos do Pensamento Estatístico**

Pfannkuch e Wild (1999 apud PFANNKUCH e WILD, 2004) desenvolveram uma estrutura para pensamento estatístico na investigação empírica. Desta, destacamos a Dimensão 1 (Figura 1), centralizando a última etapa do Ciclo Investigativo de MacKay (2000) onde são estruturadas as conclusões, interpretações e comunicação de novas ideias. Para desenvolver a estrutura, foram entrevistados estatísticos e estudantes universitários sobre os projetos estatísticos em que eles estiveram envolvidos. Em seu modelo foram identificados os tipos de pensamentos considerados pelos autores como fundamentais para Estatística, e estão apresentados na Dimensão 2 do modelo, “Tipos de Pensamento” (Figura 2). Dentre os cinco tipos de pensamentos, definimos a seguir aqueles elementares para este estudo.

Figura 1 – **Dimensão 1**: Estrutura quadridimensional para pensamento estatístico na investigação empírica

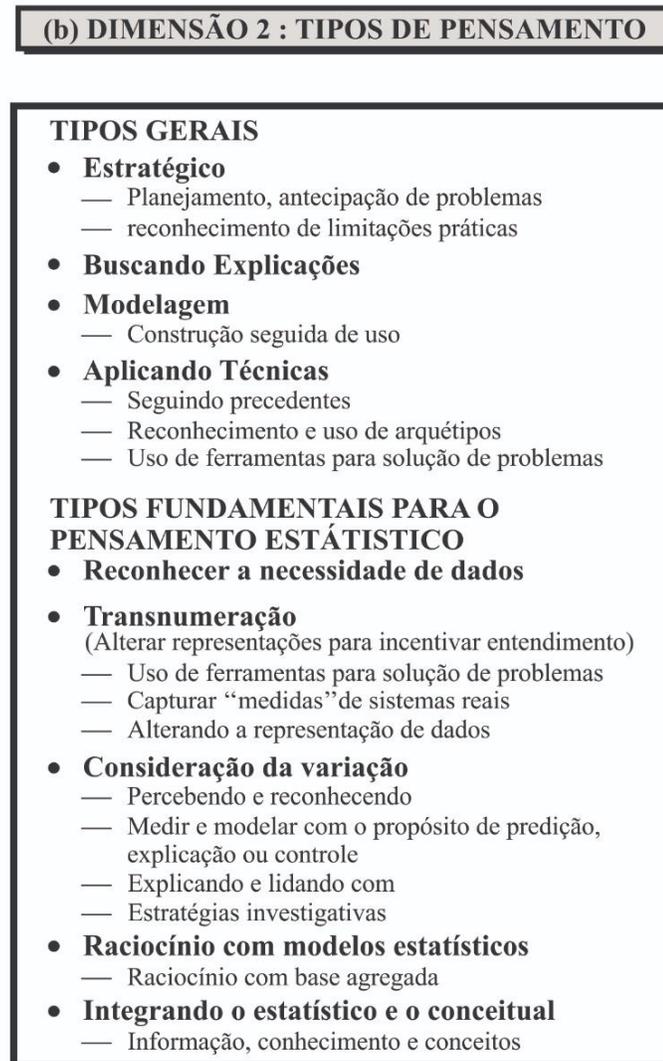


Fonte: Traduzido e adaptado pelo autor a partir de Pfannkuch e Wild (2004)

- **Transnumeração**, palavra cunhada pelos autores, que significa “mudar representações para gerar compreensão”. O pensamento em transnumeração ocorre em três instâncias específicas, quando (i) são descobertas medidas que capturam qualidades e características de uma situação real; (ii) os dados que foram coletados são transformados de brutos em múltiplas representações gráficas, ou tabelas com o objetivo de obter significado dos dados; e (iii) quando este significado dos dados e o julgamento sobre eles são comunicados de forma que possa ser compreendida em termos da situação real, pelos receptores da informação.
- **Raciocínio com modelos estatísticos**, este tipo de pensamento busca tratar o comportamento geral dos dados através de modelos estatísticos. Em sua forma mais básica, o raciocínio com modelos estatísticos inclui argumentação a partir de gráficos estatísticos, como histogramas e gráficos de dispersão (PFANNKUCH,

2008), mas ainda assim inclui modelos mais avançados, como de séries temporais e modelos de regressão.

Figura 2 – **Dimensão 2**: Estrutura quadridimensional para pensamento estatístico na investigação empírica



Fonte: Traduzido e adaptado pelo autor a partir de Pfannkuch e Wild (2004)

Um exemplo de transnumeração é transformar dados brutos em gráficos, e assim entender que diferentes tipos de representação revelam diferentes aspectos da história

dentro dos dados e que a partir dessas representações, uma história pode ser sintetizada (SHAUGHNESSY e PFANNKUCH, 2002 apud PFANNKUCH, 2008).

A transnumeração também pode estar em uma simples recategorização de dados brutos, quando esta tem por objetivo o melhor aproveitamento e aprendizado a partir dos dados, como por exemplo: um professor pergunta para seus alunos quais são suas séries favoritas, a lista dessas séries é o que consideramos os dados brutos, então essas séries poderiam ser recategorizadas de acordo com o gênero artístico, ou classificação indicativas de idade, ou até mesmo as plataformas nas quais os alunos a assistem. A recategorização dos dados permite que diferentes graus de detalhamento das informações e conclusões sejam obtidos a partir dos dados nesses casos (PFANNKUCH, 2008).

Quanto ao raciocínio com modelos estatísticos, Pfannkuch e Wild (2004) dizem que quando usamos modelos estatísticos para raciocinar, o foco está mais baseado em modos agregados do que individuais de raciocínio. Konold e Higgins (2003), nos trazem o seguinte exemplo para um dos diferentes modos de raciocínio,

...um conceito que é construído através de um esquema quádruplo de visualização de dados inicialmente como um ponteiro (lembança que os estudantes mediram altura), então como um caso individual (a garota mais alta é Poppy), então como um classificador (cinco estudantes tiveram altura entre 150 e 160cm), e por fim como um agregado (estudantes tipicamente tem uma altura entre 150 e 170cm). (pp 193-214).

Ao utilizar esses tipos de pensamentos e trazer o ensino da estatística para situações reais na vida dos alunos, é possível ir além da experiência de ensino tradicional e técnica. Propor situações e contextos reais no ensino pode gerar discussão e debates quanto a natureza dos dados em si, suas transformações e a maneira como foram expostas, como no caso deste trabalho, que propõe a exposição em forma de representações gráficas. Estimulando assim o desenvolvimento do raciocínio estatístico nos estudantes de Educação Básica.

### 2.3 Letramento em Big Data

Ao falarmos de Big Data, são diversos os conceitos para definir um assunto que cresce cada dia mais. Algumas definições são mais simples, como: “refere-se a grandes quantidades de informações que são geradas continuamente por vários meios, a partir de um amplo leque de situações sociais, como, por exemplo, compras *online* e uso de celulares”, como destacado por François, Monteiro e Allo (2020). Outras buscam definir de maneira mais complexa, consolidando uma visão mais técnica, como Laney (2001) ressalta

“O Big Data é comumente caracterizado pelos *três Vs*: volume, grande quantidade de dados; velocidade, rapidez com que os dados vêm e vão; e variiedade, a grande diversidade dos tipos e fontes de dados.”

Ao refletirmos sobre a temática abordada neste estudo, o Painel de COVID-19 do Estado do Rio Grande do Sul, não é possível considerar que este é caracterizado como BigData. Somente pelo conceito dos *três Vs*, o conjunto de dados em questão não se encaixa na definição de “variedade”, e muito menos no segundo V, velocidade, visto que os dados são atualizados diariamente, quando para exemplos em Big Data essas atualizações ocorrem em minutos ou até mesmo segundos.

Ainda assim, abordaremos uma das três lições interdisciplinares propostas por François, Monteiro e Allo (2020). A Lição 1 – Compreender a complexidade de como o conhecimento é criado e decisões são justificadas, nos traz o questionamento de que a Estatística vê a construção do conhecimento de um ponto de vista de como lidamos com a incerteza (KLINE, 1967/1985, apud, FRANÇOIS, MONTEIRO & ALLO, 2020), e assim, adquirimos conhecimento ao reduzir a incerteza e ao quantificá-la. Ao buscarmos desenvolver o letramento e raciocínio estatístico já na Educação Básica é possível consolidar as habilidades críticas necessárias para interpretação através da visualização de dados, desde os especialistas na área, ou qualquer indivíduo assistindo um noticiário de TV, reduzindo assim a incerteza e aos poucos desenvolvendo o conhecimento

estatístico como sustentação de um pensamento crítico para informações divulgadas para o público.

## 2.4 Representação Semiótica

Segundo Duval (1993), um mesmo objeto matemático pode ser representado por formas distintas que causam confundimento de forma geral. Seja por uma notação, ou símbolo a partir de uma forma escrita, ou a partir de traçados e figuras, os objetos matemáticos não devem ser jamais confundidos com a representação que se faz dele.

As representações semióticas são produções constituídas pelo emprego de signos pertencentes a um sistema de representações que têm inconvenientes próprios de significação e de funcionamento (DUVAL, 1993). Podemos destacar alguns exemplos de representações semióticas que possuem diferentes sistemas semióticos, como uma fórmula algébrica, um gráfico e até mesmo um enunciado de um exercício. Duval argumenta que as representações semióticas, como um todo, são críticas para fins de comunicação, mas ainda assim são igualmente essenciais à atividade cognitiva do pensamento, e desempenham um papel primordial no desenvolvimento das representações mentais, na realização de diferentes funções cognitivas e na produção de conhecimentos, ao “permitirem representações radicalmente diferentes de um mesmo objeto, na medida em que elas podem atender sistemas semióticos totalmente diferentes.” (BENVENISTE 1979, BRESSON 1978, apud DUVAL, 1993).

Acerca da construção de conhecimentos, Duval (2009) nos traz que um registro de representação semiótica deve permitir o cumprimento de três atividades cognitivas, sendo elas:

Primeiramente, construir um traço ou um ajuntamento de traços perceptíveis que sejam identificáveis como *uma representação de alguma coisa* em um sistema determinado. Em seguida, transformar as representações apenas pelas regras próprias ao sistema, de modo a obter outras representações que possam constituir uma relação de conhecimento em comparação as representações iniciais. Enfim, converter as representações produzidas em um sistema em representações de um outro sistema, de tal maneira, que estas últimas permitam explicar outras significações relativas ao que é representado. (DUVAL, 2009, p. 36-37).

Ao abordarmos a representação semiótica a partir do que é requerido para o desenvolvimento do Pensamento Estatístico, Coutinho, Silva e Almouloud (2011) destacam que a compreensão da Estatística implica em transitar entre os diversos registros de representação semiótica envolvidos no processo de construção de conceitos estatísticos. A partir de tais premissas surge o questionamento de como as três atividades cognitivas propostas por Duval podem ser analisadas de um ponto de vista estatístico, desta forma Vieira (2008) afirma que:

... em uma análise de um conjunto de dados, deve-se poder passar do banco de dados para uma distribuição em sua forma tabular ou em sua forma gráfica, e ainda para sua representação por meio de medidas-resumo. Assim, a resolução de problemas estatísticos passa não só pela transformação de um registro a outro, como também pelo uso simultâneo desses diversos registros para obtenção do maior número de informações, permitindo a análise crítica dos dados, segundo os princípios da Análise Exploratória de Dados. (p. 24)

Sendo assim, comentamos anteriormente que o desenvolvimento do Pensamento Estatístico, de acordo com Pfannkuch (2008), segue principalmente, do ponto de vista deste trabalho, a etapa de Transnumeração, cuja ideia central é formar e mudar representações de dados para chegar a melhor compreensão destes (Coutinho et al, 2011). Ao relacionarmos os conceitos de Duval quanto à construção de conhecimento aos princípios do Pensamento Estatístico propostos por Pfannkuch e Wild, podemos inferir que a Transnumeração pode ser considerada como um Registro de Representação Semiótica, cujas ideias principais permitem o cumprimento das três atividades cognitivas propostas por Duval.

### 3. RECURSOS TECNOLÓGICOS

#### 3.1 Visualização de Dados

A evolução da tecnologia deu à sociedade uma das ferramentas mais perfeitas de comunicação visual. Com certas habilidades é possível se construir uma grande variedade de gráficos, dashboards, e apresentações de resultados de maneira visual. Estas são só algumas das ferramentas disponíveis e que são melhoradas com recorrência.

É possível imaginar que as raízes das áreas da Estatística são longas. Em termos de visualização de dados, tais raízes vão de encontro ao princípio histórico da construção de mapas e representação visual, e no mais tardar, dentro da cartografia temática, estatística, e com aplicações e inovações em diversos campos da Medicina e Ciência, as quais são comumente interligadas (FRIENDLY, 2008).

É comum nos noticiários e mídias em geral a necessidade do apelo visual para melhor e mais rapidamente se processar as informações apresentadas, sejam elas em forma de gráficos, tema deste estudo, sejam em imagens referentes a noticiários. Por causa da maneira como o cérebro humano processa informações, usar gráficos e diagramas para visualizar grandes quantidades de dados complexos é mais fácil do que jogar informações em relatórios e planilhas (SAS, 2019). Com gráficos e diagramas bem estruturados e construídos é possível desde identificar padrões em determinadas estruturas de dados.

Contudo, apesar de gráficos serem utilizados extensivamente na Estatística por muito tempo, historicamente, segundo Unwin (2008), não há um corpo substantivo da teoria ou área de conhecimento sobre o tópico. Este tipo de conhecimento é geralmente expresso em fundamentos e princípios a serem seguidos e não em teoria formalizadas. Mesmo assim Cleveland (1993) traz os seguintes estágios da visualização de dados, e a partir deles acredita-se ser possível construir qualquer análise gráfica, tornando possível extrair o máximo de informação de um conjunto de dados.

1. coletar;
2. analisar;
3. filtrar;
4. minerar;
5. representar;
6. refinar;
7. interagir.

Ainda assim, se faz necessário, principalmente no estágio “Representar”, um conjunto de regras práticas para que seja possível entender a partir do que um gráfico é construído, indo além dos dados utilizados. Com isso, na próxima sessão comentaremos uma das abordagens existentes que buscam ressaltar os pilares, ou “partes”, presentes na construção e interpretação de gráficos, tal abordagem é nomeada “A Gramática dos Gráficos” (Grammar of Graphics).

### **3.2 A Gramática dos Gráficos**

Já comentamos a extensão do uso de gráficos na Estatística, contudo há certa dificuldade em responder perguntas como: “O que é um gráfico?” ou “Como posso sucintamente descrever um gráfico?”. A fim de responder tais questões e criar um conjunto de fundamentos que pudessem ser utilizados para descrever e construir uma variedade de gráficos estatísticos, Wilkinson, Anand e Grossan (2005) desenvolveram a Gramática dos Gráficos. O “conjunto de princípios ou regras que regem o funcionamento de uma língua” (DICIO, 2021) forma uma gramática, e ao abordarmos tal conceito do ponto de vista de gráficos, este condiz com a premissa deste trabalho, onde a visualização de dados é dada como forma de comunicação e representação.

Existem, principalmente, duas abordagens para a Gramática dos Gráficos, mas neste trabalho utilizaremos a proposta por Wickham (2010) que nos traz a “Gramática dos Gráficos em Camadas” (*Layered Grammar of Graphics*). Para Wickham (2010), a Gramática dos Gráficos em Camadas define os componentes de um gráfico como (i) um conjunto de dados padrão e um conjunto de mapeamentos de variáveis para estética; (ii)

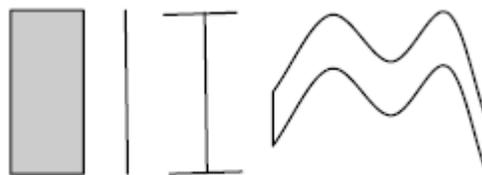
uma ou mais camadas; (iii) uma escala para cada mapeamento estético usado; (iv) um sistema de coordenadas; (v) uma especificação de faceta (característica).

O Conjunto de Dados Padrão é uma parte crítica para criação de um gráfico, mas ainda assim é independente dos outros componentes que formam a gramática. Junto com os dados, precisamos especificar quais variáveis são mapeadas para quais elementos estéticos, como por exemplo mapear “altura” para o eixo vertical.

As Camadas são responsáveis por criar os objetos que serão percebidos no gráfico e é composta de quatro partes:

- **Transformação Estatística:** É responsável por transformar os dados, tipicamente ao resumi-los de alguma maneira. Por exemplo, dividir um intervalo contínuo em caixas e contar o número de pontos em cada caixa, caracterizando um histograma.
- **Objeto Geométrico:** Controla o tipo de gráfico a ser criado e podem ser classificados por sua dimensionalidade:  $0d$ , pontos ou texto;  $1d$ , caminho ou linha;  $2d$ , polígono ou intervalo. A Figura 3 ilustra quatro formas distintas de se representar um intervalo em um gráfico, objeto geométrico em questão.
- **Ajuste de Posicionamento:** Algumas vezes se faz necessário um leve ajuste no posicionamento dos elementos geométricos do gráfico, ou então eles podem obscurecer um ao outro. É mais comum em gráficos de barras, onde podemos posicionar as barras lado a lado, a fim de evitar sobreposição.

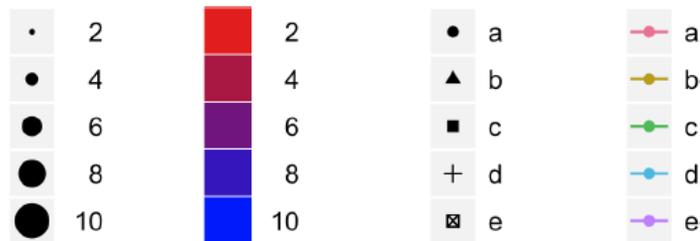
Figura 3: Quatro representações do objeto geométrico intervalo. Da esquerda para direita: como barra, como linha, como barra de erro e como um polígono.



Fonte: Wickham (p. 12, 2010)

A Escala controla o mapeamento a partir dos dados para atributos estéticos, e assim precisamos de uma escala para cada propriedade estética utilizada no gráfico. As Escalas Guias podem ser associadas aos eixos e legendas dos gráficos, e os parâmetros de sua função definem, por exemplo, se o caminho será linear ou curvado, qual paleta de cores utilizar e qual cor começam e terminam tal paleta. Alguns exemplos são destacados na Figura 4.

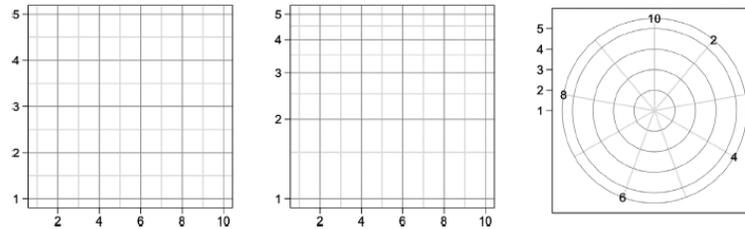
Figura 4: Exemplos de legendas a partir de quatro diferentes escalas. Da esquerda para direita: variável contínua mapeada para tamanho e cor, uma variável categórica mapeada para formato e cor.



Fonte: Wickham (p. 12, 2010)

O Sistema de Coordenadas é responsável por mapear o posicionamento dos objetos no plano do gráfico. A Posição geralmente é especificada por duas coordenadas ( $x$ ,  $y$ ), mas pode ser utilizado qualquer número de coordenadas. O sistema mais comum para duas dimensões é o de Coordenadas Cartesianas. O Sistema de Coordenadas controla como os eixos e a rede de linhas são desenhadas. A Figura 5 ilustra alguns exemplos de Sistemas de Coordenadas.

Figura 5: Exemplos de eixos e rede de linhas para três sistemas de coordenadas: Cartesiano, semi-log e polar.



Fonte: Wickham (p. 14, 2010)

A Faceta (característica) destaca um caso mais geral de gráficos conhecidos como gráficos condicionados ou gráficos de treliça (também conhecido como painel de gráficos), isto é, na etapa de Faceta se torna fácil a criação de pequenos múltiplos de diferentes subconjuntos de uma base de dados. As especificações da faceta descrevem quais variáveis devem ser utilizadas para dividir o banco de dados original e como elas devem ser arranjadas para exibição.

O objetivo de uma gramática é “reunir de forma coerente coisas que antes pareciam não relacionadas, e fundamentar uma base para lidar sistematicamente com novas situações” (COX, 1978 apud WICKHAM, 2010). Ao destrincharmos a construção de gráficos em etapas distintas e com suas próprias regras, podemos compreender melhor os componentes que devem formar um gráfico a ser construído. E, ao melhor entender as formas que um gráfico é elaborado, podemos assim realizar uma melhor interpretação de gráficos e visualizações que nos são expostas no dia a dia.

## 4. PROPOSTA DIDÁTICA

A Proposta Didática para Visualização de Dados proposta neste trabalho envolve a realização de uma dinâmica em sala de aula onde os envolvidos possam explorar temáticas atuais e a partir destas debater em conjunto, somando opiniões, a fim de desenvolver criticidade ao interpretar gráficos, mesmo fora da sala de aula, explorando assim o desenvolvimento das habilidades centrais presentes no Letramento e Raciocínio Estatístico, definidas anteriormente. Este Guia propõe uma abordagem de ensino para ser utilizada em sala de aula, separada em três pilares:

- i) **Tema de Estudo**, onde os envolvidos devem selecionar o tema no qual os dados a serem observados se originaram, seguindo o conceito de Modelagem Matemática de Barbosa (2001), no qual uma temática real, ou seja, atual e presente na vida dos alunos envolvidos, auxilia no processo de aprendizagem. O tema proposto para este Guia é a Pandemia de Covid-19, definida a seguir;
- ii) **Visualização de Dados Aplicados**, onde os envolvidos devem interpretar os gráficos escolhidos destacando pontos positivos e negativos, quando existirem, em cada um deles. Dentre os pontos observados, podemos destacar a adequabilidade do gráfico escolhido, contraste de cores e falta ou excesso de informação.
- iii) **Debate**, onde os envolvidos reúnem os pontos destacados e suas interpretações e os debatem em conjunto. Nesta etapa não existe certo ou errado, é nela que os alunos expõem suas opiniões a respeito dos gráficos interpretados, assim como é nesta etapa onde o professor abre o questionamento dos “por quês?”, indagando sobre os objetivos de quem construiu o gráfico, o motivo do uso de certas cores e gráficos em si, e o que poderia ser feito para corrigir ou melhorar as visualizações estudadas.

## 4.1 Pandemia COVID-19

Em sua definição, uma Pandemia é definida como “uma epidemia ocorrendo ao redor do mundo, ou sobre uma larga área territorial, atravessando fronteiras internacionais e geralmente afetando numerosos indivíduos” (OMS, 2011). A Covid-19 é uma infecção respiratória aguda causada pelo coronavírus SARS-CoV-2, potencialmente grave, de elevada transmissibilidade (BRASIL, 2021) e, sendo corretamente classificada na definição de Pandemia, de distribuição global. Este vírus foi descoberto a partir da análise de amostras de pacientes com pneumonia por causa desconhecida, na cidade de Wuhan, província de Hubei, China, em dezembro de 2019.

Dentre os principais sinais e sintomas clínicos da Covid-19 estão: febre, que não cessa com remédio; tosse seca ou com catarro e; dificuldade para respirar. De forma geral, segundo o Ministério da Saúde (BRASIL, 2021) os casos podem ser classificados em:

- **Caso assintomático:** Ocorre quando o teste laboratorial para Covid-19 retorna positivo, porém o paciente não apresenta sintomas. Podendo evoluir para um caso leve, onde ocorre presença de sintomas não específicos, como tosse, dor de garganta ou coriza, fadiga, entre outros.
- **Caso moderado:** São os casos cujos sintomas mais frequentes podem incluir desde os sinais leves da doença até sinais de piora de outro sintoma relacionado à Covid-19 (adinamia, prostração, hiporexia ou diarreia), além da presença de pneumonia sem sinais ou sintomas graves.
- **Caso grave:** É considerado como Síndrome Respiratória Aguda Grave, um caso de Síndrome Gripal que apresente dispneia/desconforto respiratório, ou pressão persistente no tórax, ou saturação de oxigênio menor que 95% em ar ambiente, ou coloração azulada de lábios ou rosto.
- **Caso crítico:** São os casos cujos principais sintomas são sepse, síndrome de desconforto respiratório agudo, insuficiência respiratória grave, disfunção de múltiplos órgãos, pneumonia grave, necessidade de suporte respiratório e internações em unidades de terapia intensiva.

Em resposta à emergência da situação e o reconhecimento da pandemia pela Organização Mundial da Saúde (OMS), o Ministério da Saúde do Brasil estabeleceu medidas de resposta e enfrentamento da Covid-19. Dentre as indicações destacamos as não farmacológicas, como:

“... lavar as mãos frequentemente; usar máscara obrigatoriamente ao sair de casa; evitar tocar nos olhos, nariz e boca com as mãos não lavadas; evitar contato próximo com pessoas doentes; ficar em casa quando estiver doente; cobrir a boca e nariz ao tossir e espirrar com um lenço de papel ou antebraço; limpar e desinfetar objetos e superfícies tocados com frequência; evitar contato físico como, abraços, beijos e apertos de mãos; manter distanciamento das pessoas em público de no mínimo 1,5 metros.” (BRASIL, 2021).

## 4.2 Visualização de dados aplicados

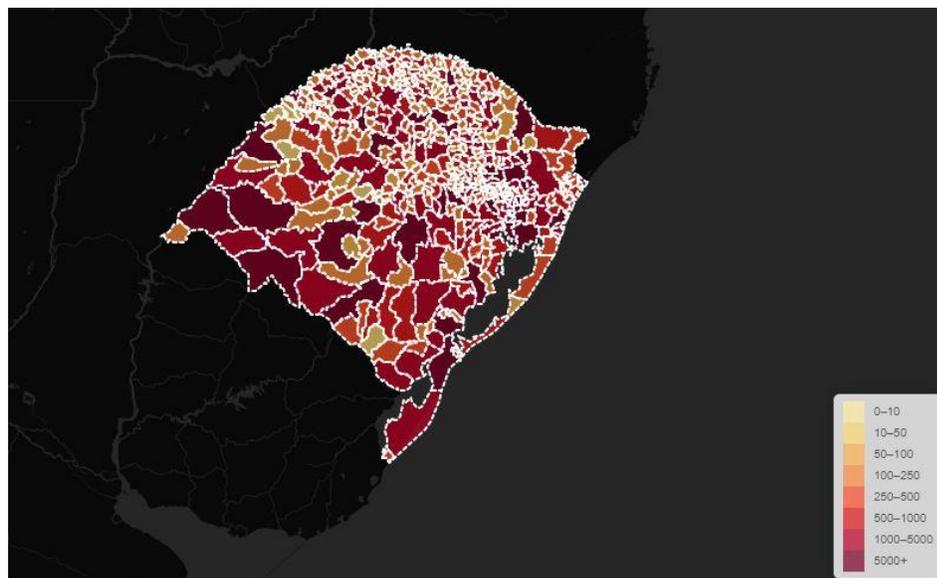
Nesta seção, vamos interpretar alguns gráficos da temática proposta: Pandemia Covid-19, que foram divulgados em diferentes mídias. Abordaremos visualizações de dois tipos, sob o ponto de vista do autor deste trabalho:

- As que estão adequadas e possuem pontos positivos a serem destacados;
- As que podem ser melhoradas para auxiliar a interpretação do leitor, destacando também os pontos nos quais tal melhoria é necessária.

Este Guia Prático foi construído a partir da premissa de que seus leitores já possuem certo Letramento Estatístico para exploração de dados, e já estão familiarizados com temas como: tipos de variáveis, adequabilidade de gráficos para cada tipo de variável e os tipos mais comuns de gráficos utilizados. Para mais referências neste tema Barbetta (2007) destaca a descrição e exploração de dados categorizados (qualitativos) e quantitativos, assim como algumas sugestões de quais representações gráficas utilizar para determinados tipos de variáveis. Caso haja preferência por uma abordagem prática do assunto, ou seja, com utilização de recurso computacional para produção de gráficos, Aquino (2014) traz diretamente no software estatístico R diversas aplicações para visualização de dados, especialmente no capítulo 8.

A Figura 6 traz um exemplo de gráfico que está adequado a seu propósito, portanto é um exemplo de gráfico do primeiro tipo. Nesta situação, retratado um Mapa de Calor onde a variável em questão é o número de casos confirmados de Covid-19, por município, no Rio Grande do Sul. Quanto a adequabilidade do gráfico selecionado aos dados a serem expostos, tendo em vista o objetivo do autor em retratar uma visão geral do estado em questão, percebemos que essa escolha está adequada. Podemos destacar também a paleta de cores utilizada, onde ao aumentar o número de casos, aumenta também o contraste e presença da cor Vermelha, destacando assim que municípios com mais casos (mais “vermelhos”) estão em situação de maior alerta, significado comumente associado ao vermelho. Quanto a clareza e objetividade, podemos destacar a presença da legenda no gráfico, de modo que toda a informação necessária para interpretação desta visualização está presente de forma destacada. Ainda assim, alguns questionamentos se originam neste caso, como: “o número de casos é proporcional à população do município?” avaliando a possibilidade de trasnumerar a informação original, e “é necessário mais grupos para diferenciar municípios com mais de cinco mil casos?” compondo uma recategorização da informação e alterando sua representação semiótica.

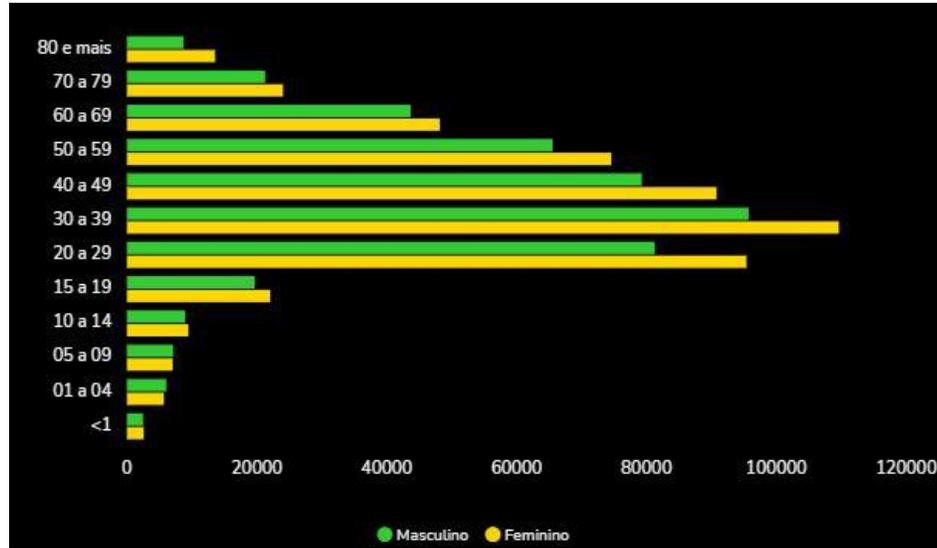
Figura 6: Mapa da quantidade casos confirmados de Covid-19 no Rio Grande do Sul.



Fonte: SESRS (2021). Disponível em: < <https://ti.saude.rs.gov.br/covid19/> >

Na Figura 7, a visualização se dá por um gráfico de barras para retratar o número total de casos confirmados de Covid-19 no Rio Grande do Sul sendo distinguido por Faixa Etária e Sexo. Imagina-se que o objetivo do autor na construção deste gráfico fosse avaliar como era a distribuição do número de casos confirmados para cada uma das faixas etárias. Quanto a adequabilidade dos dados para o gráfico final, podemos afirmar que em termos de “contagem” o gráfico de barras é o mais indicado, principalmente quando as variáveis de eixo e legenda são categóricas. Vale destacar a ausência da Camada de Nomes para Eixos, tanto no eixo vertical quanto no horizontal o que significa a informação exposta não está descrito. Percebemos também que, a escolha da paleta de cores neste gráfico traz uma seleção mais neutra, diferente dos tons de vermelho utilizados na Figura 6, as cores escolhidas aqui quando utilizadas em conjunto não destacam uma informação além da separação de categorias. Tanto o objetivo quanto as informações apresentadas pelo gráfico estão claros, de modo que a informação necessária para aqueles indivíduos que vão olhar e interpretar este gráfico estão expostas de forma coerente e concisa. Ainda assim, podemos questionar algumas melhorias e gerar um debate em conjunto sobre: “Que tipo de transformações na variável ‘contagem de casos’ poderíamos fazer e qual informação a mais essa mudança faria?”, ou até mesmo, “Imagine que você seja o Administrador da Saúde na Região que gerou este gráfico, quais conclusões você poderia chegar ao interpretar este gráfico e qual decisão tomaria?”.

Figura 7: Número de casos confirmados de Covid-19 no Rio Grande do Sul por Faixa Etária e Sexo desde o início da contabilização.

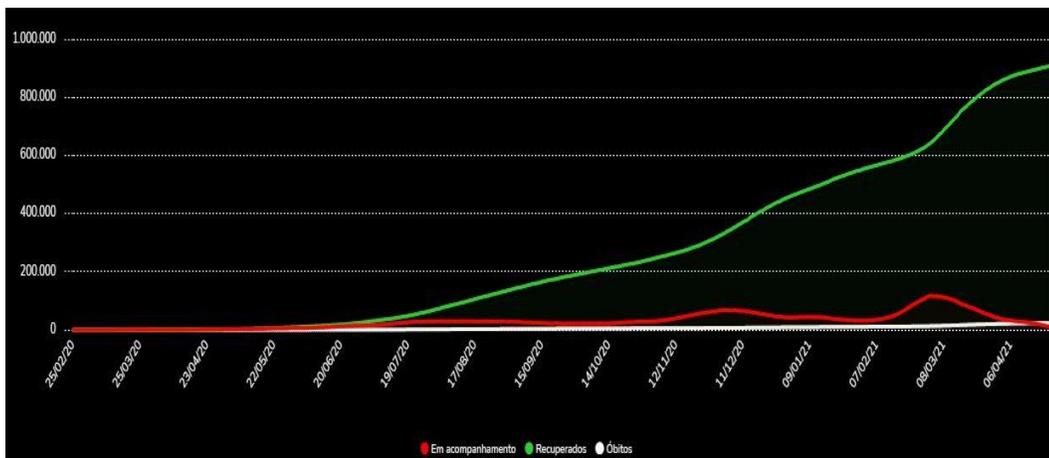


Fonte: SESRS (2021). Disponível em: < <https://ti.saude.rs.gov.br/covid19/>>

Ainda considerando o primeiro tipo, contudo com traços do segundo, na Figura 8, é apresentada a evolução dos casos confirmados de Covid-19 no Rio Grande do Sul através do tempo, destacando o Status do Paciente. Quanto a adequabilidade do gráfico, e as variáveis utilizadas para compô-lo, fica claro pela simplicidade e clareza da informação exposta que um gráfico de linhas, neste caso, é o melhor a ser utilizado, principalmente quando há a presença da variável “tempo”. Basta imaginar a quão poluída ficaria esta visualização, caso fosse utilizado um gráfico de barras com a mesma abertura por Status. Podemos destacar também a ótima seleção de cores utilizadas neste gráfico, principalmente para os status de “Em acompanhamento” e “Recuperados”, pelo contraste entre as cores vermelho e verde, instigando o leitor a concluir, corretamente, que a evolução dos casos recuperados deve ser considerada positivamente, e negativamente para os casos em acompanhamento, quando comparados entre si. A clareza do gráfico para estes dois status é facilmente notável, contudo, quando vamos observar a evolução de óbitos por conta de Covid-19, alguns fatores tornam difícil a interpretação desta informação: a cor escolhida para definir óbitos é a mesma para os eixos e escalas em si, podendo causar confundimento; além disso, a escala para contagem dos casos é muito

desproporcional entre os três status de paciente, dificultando a visualização da evolução de óbitos. Vale ressaltar que o comportamento das três variáveis apresentadas não é o mesmo, visto que a quantidade de óbitos e casos recuperados são cumulativos, ou seja, nunca diminuirão, já o número de casos em acompanhamento varia de acordo com o tempo e pode sofrer quedas ou aumentos. Mesmo que o objetivo do autor do gráfico seja dispor as informações dos três status em conjunto, ainda assim alguns questionamentos podem surgir para serem debatidos em conjunto, como: “De qual forma poderíamos transformar a variável ‘contagem de casos’ a fim de melhorar a escala do gráfico em si?”, “Qual cor ou paleta de cores você escolheria para compor o gráfico?”, tratando principalmente a camada estética de acordo com a Gramática dos Gráficos, ou até mesmo “De que outra forma poderíamos construir essa visualização de forma que a mesma informação seja exposta?” a partir da Transnumeração, ou seja transformação para melhor compreender, dos dados utilizados para criação desse gráfico.

Figura 8: Evolução dos casos de Covid-19 no Rio Grande do Sul por Status do Paciente.

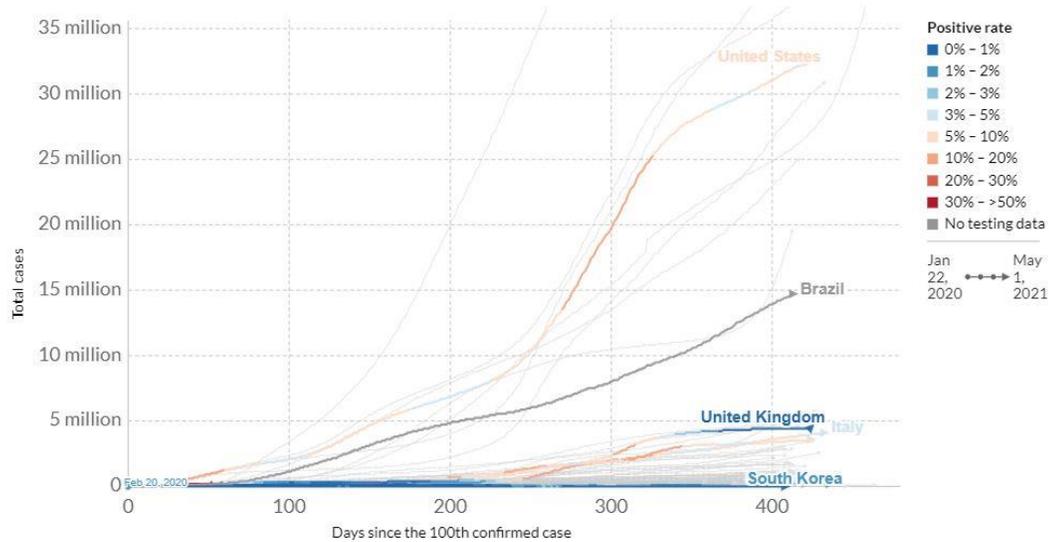


Fonte: SESRS (2021). Disponível em: < <https://ti.saude.rs.gov.br/covid19/>>

Acima, interpretamos e discutimos alguns exemplos onde o objetivo do autor, a adequabilidade dos gráficos aos dados e a clareza do conteúdo estão de acordo. As próximas visualizações destacam casos em que há confundimento e a quantidade de questionamento e dúvidas pode vir a ser maior do que as informações claramente extraídas e suas interpretações.

Na Figura 9, estão expostas as informações para o acumulado de casos confirmados de Covid-19 por País dentro do período de 22 de janeiro de 2020 até 1 de maio de 2021. As cores escolhidas para compor a legenda estão adequadas, visto o que já foi discutido na Figura 6: uma taxa de resultados positivos dentre os testados baixa, de paleta azul, indica uma situação boa para o país em questão, já as taxas de resultados positivos dentre o total de testados retratadas com a paleta vermelha indicam uma alta proporção de testes positivos, um indicativo de um volume baixo de testes sendo aplicados, que se traduz em uma situação ruim para o país. Contudo, apesar da boa escolha e atribuição para paleta de cores, podemos perceber que assim como na Figura 8, este gráfico apresenta um problema de escala, causando poluição visual para os países com um número de casos confirmados acumulado menor do que 5 milhões. Tomando a Coreia do Sul (South Korea) como exemplo, não é possível tirar conclusões ou interpretar as informações deste país em específico por conta da escala gráfica. É possível interpretar que a Coreia do Sul possui um número de casos acumulados abaixo de 2 milhões e 500 mil, entretanto, informações como a taxa de crescimento dos casos fica perdida. Com isso, muitos questionamentos podem surgir para serem debatidos em conjunto. Podemos destacar indagações do tipo: “O valor bruto de casos confirmados é a melhor forma de se relatar esta informação?” aplicando a transnumeração na camada de transformação estatística destacada na Gramática dos Gráficos, “Quais transformações poderiam ser aplicadas a fim de melhorar a visualização final?”, e “Como podemos levar em consideração a população dos países, com o objetivo de compará-los visualmente?”.

Figura 9: Cumulativo de casos confirmados de Covid-19 por País.



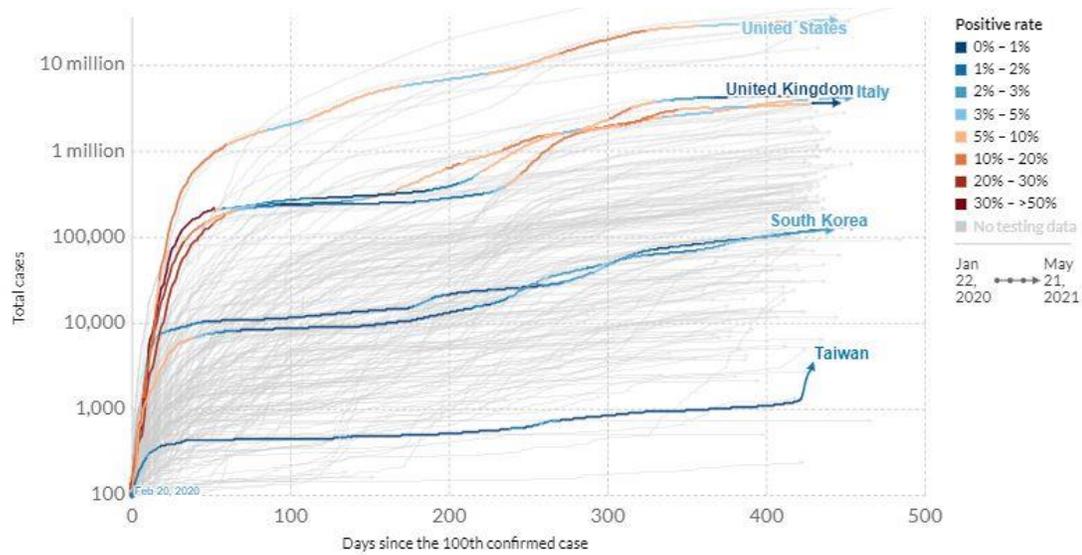
Source: Johns Hopkins University CSSE COVID-19 Data, Official data collated by Our World in Data

CC BY

Fonte: Global Change Data Lab (2021) Disponível em: <<https://ourworldindata.org/how-to-embed-charts>>

Na Figura 10, realizamos uma transformação da variável presente na Figura 9, transnumerando-a de “cumulativo de casos positivos” para “cumulativo de casos positivos em escala logarítmica”. Ao transnumerarmos essa variável dessa forma, podemos perceber o impacto direto no gráfico, principalmente em países com uma população pequena. Como o caso da Coreia do Sul (South Korea), onde na Figura 9, apesar de estar destacado, era difícil tirarmos quaisquer conclusões ou comparações levando em consideração a Coreia do Sul, visto que ao possuir uma população muito menor em escala do que outros países, o número de casos positivos também fica proporcionalmente menor. Ao alterarmos a escala da variável, uma das camadas que compõe a Gramática dos Gráficos, podemos melhor observar e comparar países de baixa população, como o caso da Coreia do Sul.

Figura 10: Cumulativo de casos confirmados de Covid-19 por País em escala logarítmica.

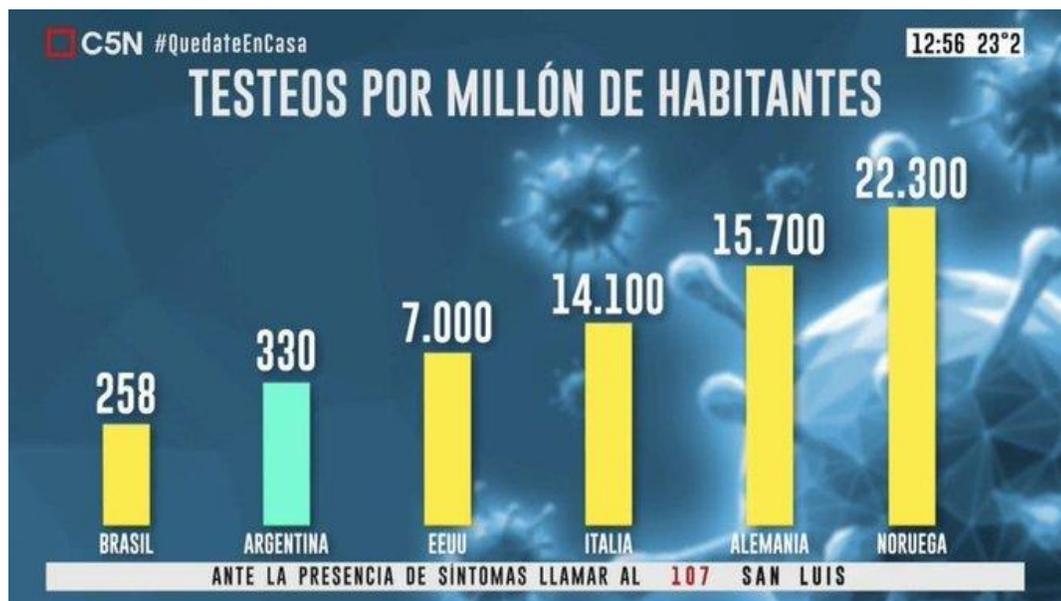


Fonte: Global Change Data Lab (2021) Disponível em: <<https://ourworldindata.org/how-to-embed-charts>>

Na Figura 11, observamos as informações do número de testes aplicados, na população de cada país em destaque, por milhão de habitantes reportadas em um dos canais de televisão Argentinos C5N (Canal 5 Notícias). Para as variáveis em questão - quantidade de testes por milhão e país – o gráfico de barras escolhido pelo autor está adequado, assim como as camadas estéticas do gráfico, como a escolha de cores, onde é destacado o país do qual o canal de notícias pertence, e os tamanhos de fonte para facilitar a leitura dos valores numéricos, vale destacar além da ausência total do eixo vertical a falta de títulos neste gráfico, por se tratar de uma visualização apresentada em televisão ainda assim é necessário que o próprio telespectador possa saber quais informações estão sendo expostas, através de seus títulos e legendas. Contudo a manipulação na escala do eixo de contagem realizada pelo autor é preocupante e induz o telespectador a uma interpretação enviesada das informações. Podemos perceber, ao comparar o número de testes feitos na Noruega (22.300) que este é aproximadamente 67 vezes maior quando comparado ao número de testes realizados na Argentina (330), porém observando somente a altura das barras, que deveriam ser proporcionais aos

números expostos, é possível observar que a diferença de altura entre as barras da Argentina e Brasil, assim como entre as barras da Argentina e Itália são semelhantes, todavia os valores numéricos dessas mesmas comparações em termos de diferença são de, 72 e 13.770 respectivamente. Tal manipulação do eixo através de sua ausência, pode levar o telespectador a achar que a situação de testes para Covid-19 na Argentina está indo bem, principalmente por se assemelhar e estar próxima à situação na Itália, por exemplo. Neste caso, podemos nos perguntar: “Qual o objetivo real do autor ao escolher essa apresentação?” destacando principalmente os conceitos na Gramática dos Gráficos como: a ausência de um sistema de coordenadas para melhor referenciar as comparações dos países expostos, ou a alteração e ocultação da escala na qual as barras foram baseadas. Ou ainda “De que maneira poderíamos alterar a representação semiótica destes dados de forma a apresentar uma informação mais fidedigna?”, contemplando principalmente a definição de Transnumeração e sua relação com as representações semióticas.

Figura 11: Número de testes aplicados para Covid-19 por milhão de habitantes.



Fonte: Towards Data Science (2020). Disponível em: <<https://towardsdatascience.com/stopping-covid-19-with-misleading-graphs-6812a61a57c9>>. Fonte Original: Rede de Televisão Argentina C5N.

Como etapa final de uma atividade proposta a partir deste guia, o ideal é que todos os questionamentos e sugestões de melhorias sejam debatidos por todo o grupo no qual esta abordagem para visualização de dados está sendo aplicada. É nesta etapa onde o grupo poderá, em conjunto, expor e discutir suas opiniões com o objetivo de melhorar a interpretação dos gráficos trazidos, e cabe comentar que não há resposta certa ou errada, e, todo comentário e proposta deve ser analisada e debatida entre os envolvidos. É neste momento que se espera desenvolver as habilidades necessárias para construção do Pensamento e Raciocínio Estatístico nos envolvidos.

## 5. CONSIDERAÇÕES FINAIS

Com o crescimento do volume de dados sendo utilizados para comunicar informações gerais é perceptível a presença da Estatística nas mais diversas áreas, o que induz a uma maior necessidade de que as habilidades que formam o Pensamento e Raciocínio Estatístico em um cidadão estejam desenvolvidas. Contudo, é notável a dificuldade demonstrada pelos estudantes em assimilar e “aprender” os conteúdos e conceitos de Estatística, geralmente apresentada na disciplina de Matemática na Educação Básica. Com isso, os pesquisadores da área da Educação Estatística têm enfrentado nas duas últimas décadas o desafio de contribuir na disseminação do conhecimento estatístico com o intuito de contribuir para a formação crítica dos cidadãos e cidadãs.

Através deste trabalho, propusemos uma abordagem de ensino não tradicional, cujo objetivo é auxiliar professores e estudantes envolvidos no ensino e aprendizado da Visualização de Dados a desenvolver as habilidades necessárias para a formação do Pensamento e Raciocínio Estatístico. Tal abordagem se fez a partir da proposta de uma Proposta Didática para Visualização de Dados. A proposta é voltada para professores em geral, desde a Educação Básica até Ensino Superior, e envolve a ideia da realização de uma dinâmica em sala de aula onde os envolvidos possam explorar temáticas atuais e a partir destas debater em conjunto, somando opiniões, a fim de desenvolver criticidade ao interpretar gráficos, mesmo fora da sala de aula. A Proposta Didática é adequada para projetos multidisciplinares que envolvam leitura e interpretação de informações expostas por meio de gráficos, visto que a Estatística, como comentado diversas vezes neste trabalho, está presente nas variadas áreas de conhecimento abordadas, principalmente na Educação Básica.

É válido ressaltar a dificuldade pessoal na escrita deste trabalho, por ser um TCC do curso de Estatística, e principalmente por ter como público-alvo os professores, destacando aqui os de Educação Básica. Contudo, a importância da Educação Estatística em si e do desenvolvimento do Letramento, Pensamento e Raciocínio Estatístico cresce proporcionalmente ao volume de dados e informações apresentadas

diariamente nos diversos meios de comunicação acessados, cada vez com maior frequência, por indivíduos de todas as idades. Além disso, destaco a relevância que essas habilidades, Pensamento e Raciocínio Estatístico, tem, ou seja, devem ser abordadas e desenvolvidas a fim de produzir conhecimento mínimo e básico para qualquer aluno. Com isso, podemos evitar que gráficos e visualizações como, principalmente, o discutido na Figura 11, manipulem informações e tragam viés e desinformação para a vida de todas e todos.

Esta Proposta Didática não é um produto finalizado e ainda há o que evoluir e melhorar. Entretanto, sua base está construída e com ela já acredito ser possível dar um passo a mais para que alunos, independente de qual grau de ensino, estejam desenvolvendo o Pensamento Estatístico e melhor entendam e critiquem o mundo no qual vivem.

## REFERÊNCIAS

- BEN-ZVI, D. & GARFIELD, J. **Statistical literacy, reasoning, and thinking: goals, definitions and challenges** In: The Challenge of Developing Statistical Literacy, Reasoning and Thinking, p. 3-16. Dordrecht, Holanda: Kluwer Academic Publishers, 2004.
- PFANNKUCH, M. & WILD, C. **Towards an Understanding of Statistical Thinking**. In: The Challenge of Developing Statistical Literacy, Reasoning and Thinking, p. 17-46. Dordrecht, Holanda: Kluwer Academic Publishers, 2004.
- BRASIL. Lei nº 9.394, de 20 de dezembro de 1996. **Estabelece as diretrizes e bases da educação nacional**. Diário Oficial da União, Brasília, 23 de dezembro de 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/l9394.htm](http://www.planalto.gov.br/ccivil_03/leis/l9394.htm)>. Acesso em: 17 mar. 2021
- BRASIL. Ministério da Educação. **Base Nacional Comum Curricular - Educação é a Base: Introdução**. Brasília: MEC, 2017. Disponível em: <[http://basenacionalcomum.mec.gov.br/images/BNCC\\_EI\\_EF\\_110518\\_versaofinal\\_site.pdf](http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf)>. Acesso em: 17 mar. 2021.
- Moore, D. (1998). **Statistics among the liberal arts**. *Journal of the American Statistical Association*, 93, 1253–1259.
- PFANNKUCH, M. **Training Teachers to Develop Statistical Thinking**. Estudo conjunto: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. ICMI Study 18. 2008 IASE Round Table Conference. 2008.
- SHAUGHNESSY, M. & PFANNKUCH, M. (2002). **How faithful is Old Faithful? Statistical Thinking: A story variation and prediction**. *Mathematics Teacher*, 95(4), p. 252-259.
- MACKAY, R. J. & OLDFORD, R. W.. **Scientific Method: Statistical Method and the Speed of Light**. *Statistical Science*, v15 n 3 pp 254-278. Agosto, 2000. Publicado por Institute of Mathematical Statistics
- KONOLD, C., & HIGGINS, T.. **Reasoning about data**. A research companion to principles and standards for school mathematics. National Council of Teachers in Mathematics, 2003.
- FRANÇOIS, K., MONTEIRO, C., ALLO, P., **Big-Data Literacy as a new vocation for statistical literacy**. *Statistics Education Research Journal*, v19, n 1 pp 195-205. 2020

LANEY, D. **3D data management: Controlling data volume, velocity and variety.** META Group Research Note. 2001.

KLING, M. **Mathematics for the Nonmathematician.** New York: Dover Publications. 1967/1985.

DUVAL, R. **Registres de représentation sémiotique et fonctionnement cognitif de la pensée.** *Annales de Didactique et de Sciences Cognitives.* p. 37- 64. Strasbourg: IREM - ULP, 1993. Traduzido em: Revemat: R. Eletr. de Edu. Matem. eISSN 1981-1322. Florianópolis, v. 07, n. 2, p.266-297, 2012.

DUVAL, R. **Semiósis e pensamento humano** – registros semióticos e aprendizagens intelectuais. Tradução de L.F. Levy; M. R. Silveira. São Paulo: Editora Livraria da Física, 2009. (fascículo 1).

COUTINHO, C. Q. S., SILVA, M. J. F., ALMOULOU, S. A.. Desenvolvimento do Pensamento Estatístico e sua Articulação com a Mobilização de Registros de Representação Semiótica. UNESP. v 24 n 39. 2011.

UNWIN, A., CHEN, C., HÄRDLE, W. K., **Introduction.** In: Handbook of Data Visualization. p. 4. Springer Handbooks of Computational Statistics. Editora Springer. ISBN: 978-3-540-33036-3. 2008.

FRIENDLY, M. **A Brief History of Data Visualization.** In: Handbook of Data Visualization. p. 15-56. Springer Handbooks of Computational Statistics. Editora Springer. ISBN: 978-3-540-33036-3. 2008.

LOPES, C. E., BARBOSA, G. C., SANTOS, S. S., **Trajetórias e Perspectivas da Educação Estatística a partir dos trabalhos apresentados no SIPEM.** Educação Matemática Pesquisa., São Paulo, v. 22, n 1 pp. 584-609, 2020.

WICKHAM, H., **A Layered Grammar of Graphics.** Journal of Computational and Graphical Statistics, v 19, n 1 pp. 3-28, 2010.

DICIO. **Dicionário Online de Português,** 2021. Disponível em <<https://www.dicio.com.br/gramatica/>>. Acesso em 10 abr. 2021.

OMS, Organização Mundial da Saúde. **Boletim da Organização Mundial da Saúde.** v 89, n 7 pp. 540-541, 2011. Doi:10.2471/BLT. 11.088815.

COX, D. R., **Some Remarks on the Role in Statistics of Graphical Methods,** Applied Statistics, v 27, n 1 pp 4-9, 1978.

BRASIL. **Sobre a doença**, Ministério da Saúde, Governo Federal. Disponível em <<https://coronavirus.saude.gov.br/sobre-a-doenca>>. Acesso em 17 abr. 2021.

BRASIL. **Como se proteger?**. Confira medidas não farmacológicas de prevenção e controle da pandemia do novo coronavírus. Ministério da Saúde, Governo Federal. Disponível em: < <https://www.gov.br/saude/pt-br/coronavirus/como-se-proteger> >. Acesso em 22 mai. 2021.

BARBETTA, P. A., **Estatística Aplicada às Ciências Sociais**. Editora da USFC. 7 ed. cap 4 e 5. 2007. ISBN 9788532803962.

AQUINO, J. A.. **R para Cientistas Sociais**. EDITUS: Editora da UESC. 2014. ISBN 9788574553696

SESRS. **Painel Coronavírus RS**. Secretaria Estadual de Saúde do Rio Grande do Sul. Disponível em: < <https://ti.saude.rs.gov.br/covid19/>>. Acesso em 24 abr. 2021.

GLOBAL CHANGE DATA LAB. **How to embed interactive COVID-19 charts in your articles**. Disponível em: <<https://ourworldindata.org/how-to-embed-charts>>. Acesso em 01 mai. 2021.

TOWARDS DATA SCIENCE. **Stopping COVID-19 with Misleading Graphs**. Disponível em:<<https://towardsdatascience.com/stopping-covid-19-with-misleading-graphs-6812a61a57c9>>. Acesso em 06 mai. 2021.

BARBOSA, J. C.. **Modelagem na Educação Matemática: contribuições para o debate teórico**. In: REUNIÃO ANUAL DA ANPED, 24., 2001, Caxambu. Anais... Rio de Janeiro: ANPED, 2001. 1 CD-ROM.