



Universidade Federal do Rio de Janeiro

Luciana Sá Brito

**CARTAS PARA A LITERACIA: DESIGN DE  
UM ARTEFATO PARA A AVALIAÇÃO EM  
LITERACIA DE DADOS**

**DISSERTAÇÃO DE MESTRADO**

Instituto de Matemática



Instituto Tércio Pacitti de Aplicações  
e Pesquisas Computacionais



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
INSTITUTO DE MATEMÁTICA  
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Luciana Sá Brito

CARTAS PARA A LITERACIA: *DESIGN* DE UM ARTEFATO PARA A AVALIAÇÃO EM LITERACIA DE  
DADOS

Dissertação de Mestrado apresentada ao Programa de  
Pós-Graduação em Informática, Instituto de  
Matemática e Instituto Tércio Pacitti, Universidade  
Federal do Rio de Janeiro, como requisito parcial à  
obtenção do título de Mestre em Informática.

Orientadora: Prof. Adriana Santarosa Vivacqua, DSc

Coorientadora: Prof. Juliana Baptista dos Santos França, DSc

Rio de Janeiro  
2020

## CIP - Catalogação na Publicação

B862c Brito, Luciana Sá  
Cartas para a Literacia: design de um artefato  
para a avaliação em literacia de dados / Luciana Sá  
Brito. -- Rio de Janeiro, 2020.  
245 f.

Orientadora: Adriana Santarosa Vivacqua.  
Coorientadora: Juliana Baptista dos Santos França.  
Dissertação (mestrado) - Universidade Federal do  
Rio de Janeiro, Instituto Tércio Pacitti de  
Aplicações e Pesquisas Computacionais, Programa de  
Pós-Graduação em informática, 2020.

1. Educação. 2. Sistemas colaborativos. 3.  
Literacia de dados. 4. Design Science Research. 5.  
Teoria Fundamentada em Dados. I. Vivacqua, Adriana  
Santarosa, orient. II. França, Juliana Baptista dos  
Santos, coorient. III. Titulo.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos  
pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

**Cartas para a Literacia: Design de uma escala para  
a avaliação em Literacia de Dados**

Luciana Sá Brito

**Dissertação de Mestrado** submetida ao  
Programa de Pós-graduação em  
Informática do Instituto de Matemática e  
do Instituto Tércio Pacitti da Universidade  
Federal do Rio de Janeiro - UFRJ, como  
parte dos requisitos necessários à  
obtenção do título de Mestre em  
Informática.

Aprovada em 14/12/2020 por:

*Luciana S. Brito*  
Prof. Adriana Santarosa Vivacqua, DSc, UFRJ (Presidente)

*participou por videoconferência*  
Prof. Juliana Baptista dos Santos França, DSc, UFF

*participou por videoconferência*  
Prof. Angélica Fonseca da Silva Dias, DSc, UFRJ

*participou por videoconferência*  
Prof. Claudia Lage Rebello da Motta, DSc, UFRJ

*participou por videoconferência*  
Prof. Caroline Queiroz Santos, DSc., PUC-RS

Para meu pai, Luiz Ferreira de Brito Filho  
e minha mãe, Luiza Sá Brito.  
Ao lado de minha irmã, Izabela.  
Minha família.

## AGRADECIMENTOS

---

À minha querida orientadora Adriana Santarosa Vivacqua, por ter permitido a realização desta pesquisa não somente através dos seus conhecimentos técnicos, mas também através da sua humanidade e amor pela ciência e pela educação, que se reflete no comprometimento especial com a educação de mulheres na Ciência da Computação. Obrigada pela sugestão do estudo de literacia de dados, pois deu a mim a oportunidade de unir a reflexão acerca do que desejo para o mundo com as minhas próprias buscas e interesses pessoais. Obrigada por ter acreditado que eu poderia me tornar uma cientista e pelos anos de dedicação à minha formação, por ter me doado seu tempo, sua paciência e seu apoio todas as vezes que eu precisei.

À minha coorientadora Juliana Baptista França, por ter permitido a realização desta pesquisa não somente através dos seus conhecimentos técnicos, mas sobretudo pelo trabalho intenso para a sua concretização, seus ensinamentos, sua visão feminista, seu carinho comigo e com todo o grupo de pesquisa. Sua coorientação foi além do âmbito acadêmico e permitiu aprimorar meus senso de responsabilidade e dedicação, além de permitir a ampliação dos meus horizontes no que diz respeito às possibilidades de pesquisa em Ciência da Computação.

À Angélica Fonseca Dias, por ter aceito o nosso convite para a banca e por suas orientações valiosas e correções intensas na versão final da dissertação. Obrigada por sua presença de sabedoria e cuidado na condução desta pesquisa. Os desafios foram tantos e você esteve ao meu lado em cada um deles. Muito obrigada por me ajudar a acreditar em mim e a abraçar a complexidade envolvida na metodologia desta pesquisa. Tenha certeza de que toda a energia positiva e todos os aprendizados serão compartilhados por onde eu for.

À Claudia Mota, professora e pesquisadora pela qual tenho profunda admiração. Obrigada por ter aceito o nosso convite para esta banca feminina incrível, pelas sugestões que ampliaram as possibilidades de aplicações dos nossos artefatos de pesquisa e principalmente pelo seu exemplo como pesquisadora em Informática e nossa representante na Sociedade Brasileira de Computação.

À professora e pesquisadora Caroline Queiroz, minha gratidão por ter aceito o nosso convite para esta banca, pelas sugestões que ampliaram a minha percepção acerca das áreas de Ciência da Computação relacionadas à esta pesquisa e por me inspirar tantas vezes durante esse processo através da sua presença como pesquisadora das áreas de Sistemas Colaborativos e Interação Humano Computador e como pessoa.

Ao professor Cleidson de Souza, pela inspiração que me fez lançar mão da Teoria Fundamentada em Dados como método de análise nesta pesquisa e por ter me emprestado as notas e slides da sua palestra realizada na SBSC 2019.

Ao professor Mariano Pimentel pela inspiração e amizade durante o meu mestrado, por seu estilo questionador da realidade social e educacional que me encanta e por mostrar em suas pesquisas alguns caminhos para tornar teoria e prática aliados do sonho de uma educação de

qualidade, mais crítica e mais plural, através de um *design* educacional científico e participativo.

Meus agradecimentos às professoras Taís Pereira (IFRJ), Carolina Ponce (IFRJ), Sueni Arouca (IFRJ) e Regina Fukuda (IFRJ); ao cientista de dados Luiz Eduardo Gomes (IPEA); e aos professores Felipe Rafael (UNIRIO) e Agnaldo Esquincalha (UFRJ), por terem colaborado gentilmente através dos seus conhecimentos nas das mais de 20 horas de análises atentas na fase de avaliação da escala de literacia de dados.

Meus agradecimentos à Teresa Badu e Marcela Oliveira pela colaboração como pesquisadoras assistentes durante a coleta de dados do estudo de caso.

Minha gratidão ao ilustrador Arthur Torres, pelo apoio técnico na concretização do *design* da versão final das cartas de desafio.

Aos meus professores e professoras do PPGI- UFRJ, por terem me guiado e me dado todo o suporte para o desenvolvimento das habilidades e competências necessárias para esta jornada. Em especial às professoras Mônica Dias, Jonice Oliveira e Carla Delgado. A todos os servidores e servidoras que me acompanharam e me ajudaram durante todo o mestrado, especialmente à Patricia Mamede, secretária do PPGI, Tatiana Ribeiro e Raquel Porto (Bibliotecárias do NCE) por terem me ajudado sempre que precisei (e não foram poucas vezes).

Aos meus amigos e amigas do mestrado, muito obrigada pela troca de experiências. Aos amigos e amigas do laboratório, em especial para Ivison Rubim, Diana Arce e Wallace Ugulino, pela presença acolhedora e por acreditarem em mim. A Felipe Reis e Oscar Firme pela paciência e pelos trabalhos intensos e interessantes que fizemos juntos! Vocês foram demais! Aos pesquisadores e montanhistas Marcelo Goulart, Antônio Revail e Daniel Vieira, pelas horas de conversa, ligações e por terem me permitido escalar durante esse processo.

Agradeço aos meus amigos e amigas, em especial ao Everton Leite, que me acompanhou no difícil caminho desse mestrado. Muito obrigada por tudo.

Aos amigos e amigas do setor de Desenvolvimento de Sistemas da Fundação Cecierj o meu muito obrigada pelas horas de discussões calorosas sobre ciência de dados e análise de dados educacionais, pelas gargalhadas e pela cumplicidade. Ao professor César Bastos, muito obrigada por me inspirar na escolha deste mestrado e por ter me incentivado tantas vezes e em momentos diferentes da minha trajetória profissional na área da educação. À professora Fabrícia Constantino e ao professor Rodrigo Marques, do Consórcio Cederj, muito obrigada pelo apoio incondicional. A todos e todas da fundação, muito obrigada por tudo.

À Universidade Federal do Rio de Janeiro, por me permitir o acesso a uma educação gratuita, de qualidade e comprometida com a sociedade. Ao seu corpo docente, discente e de servidores por lutarem incessantemente pela garantia desses direitos, que viabilizam a realização de pesquisa pública e socialmente referenciada em nosso país.

*“Tornamos o nosso mundo significativo, pela coragem das nossas perguntas e pela profundidade das nossas respostas.”*

(Carl Sagan – O Cosmos)

## RESUMO

---

A partir das duas primeiras décadas do século XXI, grandes volumes de dados passaram a estar amplamente disponíveis para consumo, o que tem criado uma tendência à adoção de narrativas baseadas em dados. Esse estilo de escrita requer conhecimentos relacionados à recuperação de informação e manipulação de dados não somente para criar, mas também para interpretar. A habilidade de ler, trabalhar, analisar e argumentar com base em dados é chamada de Literacia de Dados (D'Ignazio e Bhargava, 2015). Apesar de algumas dessas competências em Literacia de Dados serem mencionadas na Base Nacional Comum Curricular (BNCC), ainda não existem formas de avaliar os conhecimentos de estudantes em Literacia de Dados. É importante que haja apoio para a sistematização dessas competências, uma vez que lacunas na educação básica podem levar a cidadãos incapazes de realizar uma leitura crítica de notícias baseadas em dados, e que poderão ser vítimas de *fake news* ou desinformação. Através da *Design Science Research*, foi desenhado um artefato com o objetivo de tornar explícitos os conhecimentos tácitos de estudantes sobre Literacia de Dados. Aplicando a Teoria Fundamentada em Dados na codificação das narrativas resultantes de dois ciclos de estudo de caso realizados com amostras de estudantes, foi possível criar um mecanismo para avaliação de Literacia de Dados. O artefato foi avaliado por matemáticos, estatísticos e um cientista de dados, que contribuíram com sugestões que viabilizaram a geração de uma versão melhorada da escala. As contribuições desta pesquisa incluem uma escala qualitativa para avaliar a Literacia de Dados de estudantes, considerando 60 competências organizadas em 7 categorias que vão da “coleta de dados” até “tomar decisões informadas por dados”; a atualização da definição de Literacia de Dados e a disponibilização de um *framework* que generaliza o processo de desenvolvimento de uma escala a partir do estudo realizado.

**Palavras-chave:** educação; sistemas colaborativos; literacia de dados; *Design Science Research*.

## ABSTRACT

---

Since the first two decades of the twenty century, large amounts of data have been widely available for consumption. This has created a trend towards data-based narratives. This style of writing requires certain skills related to data retrieval and manipulation not only to create, but also to interpret. The ability to read, work, analyze and argue with data is called data literacy. These skills are not always well developed in schools, even though related aspects are mentioned in Brazil's National Curriculum (BNCC). Holes in basic education will eventually lead to citizens who are not able to critically read a data-based piece, and might more easily fall prey to fake news or misleading news stories. Through Design Science Research we designed an artifact to turn students' tacit knowledge about data literacy into explicit, by using Grounded Theory to code data group storytelling told by two samples of students in two cycles of case study as means to discover data literacy skills to create an early version of a data literacy scale. The artifact was evaluated positively by mathematicians, statisticians, and a data scientist from public institutions, who contributed with suggestions to a final version of the scale. Our contributions include a qualitative data literacy scale, that assesses students on sixty data literacy abilities organized into seven categories that go from "data collection" to "make informed decisions"; actualized data literacy definition, and a framework generalizing the process of scale development.

**Keywords:** education; collaborative systems; data literacy; Design Science Research.

## LISTA DE FIGURAS

Figura 1: Número de artigos retornados através de busca pela expressão “ <i>Data Literacy</i> ” para os anos de 2000 a 28 de novembro de 2020 em bases indexadas Google Scholar, Springer, Willey e SciELO.....	21
Figura 2: Mapa da Estrutura metodológica da pesquisa baseada na DSR.....	26
Figura 3: Estrutura lógica de aplicação da pesquisa.....	28
Figura 4: Na abordagem de Kjelvik e Schultheis (2019), a Literacia de Dados está na interseção entre Ciência de Dados, o raciocínio quantitativo e a compreensão do conteúdo da área estudada. ....	33
Figura 5: Neste mapa mental é apresentada uma proposta de organização visual da definição de D'Ignazio e Bhargava (2015) para <i>Data Literacy</i> . .....	35
Figura 6: Competências destacadas na definição de Literacia de Dados de D'Ignazio e Bhargava (2015).....	37
Figura 7: Diagrama com os domínios de conhecimento em Literacia de Dados para o ensino, de acordo com o construto de Gummer e Mandinach (2015). .....	43
Figura 8: O <i>dataset</i> Cores, como exemplo de um mural de dados genérico, feito por estudantes do curso de Ciência da Computação da UFRJ em uma atividade em sala de aula. As cores representam dados relacionados a características pessoais, preferências e hábitos. ....	44
Figura 9: Modelo de Atividade (adaptado de Engeström, 1987).....	55
Figura 10: Modelo 3C de Colaboração. ....	56
Figura 11: Mapa de elementos esperados de uma pesquisa na abordagem da DSR instanciado para a pesquisa de Santos <i>et al.</i> (2016).....	58
Figura 12: Decomposição da definição de D'Ignazio e Bhargava (2015) segundo competências relacionadas com a Literacia de Dados. ....	65
Figura 13: <i>Pipeline</i> de dados concebido através da definição de literacia em dados de D'Ignazio e Bhargava (2015).....	67
Figura 14: Primeira versão do artefato Cartas para a Literacia criado a partir da definição de Literacia de Dados de D'Ignazio e Bhargava (2015) e utilizada durante a atividade do primeiro ciclo do estudo de caso. ....	69
Figura 15: Cartas do jogo Love Letter.....	70
Figura 16: Etapas de condução do procedimento de estudo de caso. ....	72

Figura 17: As fases da atividade e o tempo para a resolução de cada solicitação representada pelas cartas de desafio. ....	73
Figura 18: Esquema de aplicação da GT, com suas fases.....	77
Figura 19: O artefato “Cartas para a literacia” como uma interface entre os conhecimentos tácitos de cada grupo pesquisado e os conhecimentos explícitos do grupo.....	82
Figura 20: Um dos 45 cartões coloridos do dataset Cores (Figura 8) gerado por estudantes de uma turma de Sistemas de Informação da UFRJ. As cores pintadas à lápis de cor representam as respostas para as 8 perguntas feitas aos estudantes da turma. ....	85
Figura 21: Estudantes da turma de graduação, divididos em grupos, utilizando o artefato Cartas para a Literacia mais o dataset Cores durante a aplicação da atividade em sala de aula. ....	88
Figura 22: Materiais de papelaria disponibilizados para os estudantes durante a atividade proposta no estudo de caso. ....	89
Figura 23: Segunda versão das cartas de desafio, preparadas para serem aplicadas no segundo ciclo do estudo de caso, realizado com a turma de pós-graduação. ....	124
Figura 24: Versão final do conjunto de cartas, obtida após o segundo ciclo de design, realizado com a turma de pós-graduação.....	137
Figura 25: Nuvem de palavras gerada através das respostas dos grupos à quinta carta de desafio. ....	156
Figura 26: Escala de Literacia de Dados apresentada na entrevista piloto, utilizando o software Jamboard. Nas notas autoadesivas azuis constam as definições das categorias, com base na definição de Literacia de Dados de D’Ignazio e Bhargava (2015). ....	158
Figura 27: Escala de Literacia de Dados apresentada nas 7 entrevistas com os especialistas. Nas notas autoadesivas laranjas constam as definições das categorias, com base na definição de Literacia de Dados de D’Ignazio e Bhargava (2015). ....	159
Figura 28: Etapas percorridas durante o desenvolvimento da escala de Literacia de Dados. ....	179
Figura 29: <i>Framework</i> que representa a generalização do processo de desenvolvimento da escala de Literacia de Dados. ....	183

## LISTA DE QUADROS

Quadro 1: Competências específicas e habilidades listadas pela BNCC para o Ensino Médio na área de Matemática e suas Tecnologias. ....	38
Quadro 2: Etapas do <i>pipeline</i> relacionadas às questões pensadas na fase de ideação das cartas. ....	68
Quadro 3: Resultados do grupo 1 para a Fase 4 .....	90
Quadro 4: Resultados do grupo 2 para a Fase 4. ....	91
Quadro 5: Resultados do grupo 3 para a Fase 4. ....	91
Quadro 6: Resultados do grupo 4 para a Fase 4. ....	92
Quadro 7: Resultados do grupo 5 para a Fase 4. ....	93
Quadro 8: Resultados do grupo 6 para a Fase 4. ....	94
Quadro 9: Resultados do grupo 7 para a Fase 4. ....	95
Quadro 10: Resultados do grupo 8 para a Fase 4. ....	95
Quadro 11: Resultados do grupo 1 para a fase 5. ....	96
Quadro 12: Resultados do grupo 2 para a fase 5. ....	96
Quadro 13: Resultados do grupo 3 para a fase 5. ....	96
Quadro 14: Resultados do grupo 4 para a fase 5. ....	96
Quadro 15: Resultados do grupo 5 para a fase 5. ....	97
Quadro 16: Resultados do grupo 6 para a fase 5. ....	97
Quadro 17: Resultados do grupo 7 para a fase 5. ....	97
Quadro 18: Resultados do grupo 8 para a fase 5. ....	98
Quadro 19: Notas de observação do grupo 1 durante a fase 1 do primeiro ciclo do estudo de caso.....	98
Quadro 20: Notas de observação do grupo 2 durante a fase 1 do primeiro ciclo do estudo de caso.....	98
Quadro 21: Notas de observação do grupo 3 durante a fase 1 do primeiro ciclo do estudo de caso.....	99
Quadro 22: Notas de observação do grupo 4 durante a fase 1 do primeiro ciclo do estudo de caso.....	99
Quadro 23: Notas de observação do grupo 5 durante a fase 1 do primeiro ciclo do estudo de caso.....	100

Quadro 24: Notas de observação do grupo 6 durante a fase 1 do primeiro ciclo do estudo de caso.....	100
Quadro 25: Notas de observação do grupo 7 durante a fase 1 do primeiro ciclo do estudo de caso.....	101
Quadro 26: Notas de observação do grupo 8 durante a fase 1 do primeiro ciclo do estudo de caso.....	101
Quadro 27: Notas de observação do grupo 1 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	102
Quadro 28: Notas de observação do grupo 2 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	103
Quadro 29: Notas de observação do grupo 3 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	103
Quadro 30: Notas de observação do grupo 4 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	103
Quadro 31: Notas de observação do grupo 5 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	104
Quadro 32: Notas de observação do grupo 6 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	105
Quadro 33: Notas de observação do grupo 7 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	105
Quadro 34: Notas de observação do grupo 8 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.....	106
Quadro 35: Notas de observação do grupo 1 durante a fase 4 do primeiro ciclo do estudo de caso.....	106
Quadro 36: Notas de observação do grupo 2 durante a fase 4 do primeiro ciclo do estudo de caso.....	106
Quadro 37: Notas de observação do grupo 3 durante a fase 4 do primeiro ciclo do estudo de caso.....	107
Quadro 38: Notas de observação do grupo 4 durante a fase 4 do primeiro ciclo do estudo de caso.....	107

Quadro 39: Notas de observação do grupo 5 durante a fase 4 do primeiro ciclo do estudo de caso.....	108
Quadro 40: Notas de observação do grupo 6 durante a fase 4 do primeiro ciclo do estudo de caso.....	108
Quadro 41: Notas de observação do grupo 7 durante a fase 4 do primeiro ciclo do estudo de caso.....	109
Quadro 42: Notas de observação do grupo 8 durante a fase 4 do primeiro ciclo do estudo de caso.....	109
Quadro 43: Notas de observação do grupo 1 durante a fase 5 do primeiro ciclo do estudo de caso.....	110
Quadro 44: Notas de observação do grupo 2 durante a fase 5 do primeiro ciclo do estudo de caso.....	110
Quadro 45: Notas de observação do grupo 3 durante a fase 5 do primeiro ciclo do estudo de caso.....	110
Quadro 46: Notas de observação do grupo 4 durante a fase 5 do primeiro ciclo do estudo de caso.....	110
Quadro 47: Notas de observação do grupo 5 durante a fase 5 do primeiro ciclo do estudo de caso.....	111
Quadro 48: Notas de observação do grupo 6 durante a fase 5 do primeiro ciclo do estudo de caso.....	111
Quadro 49: Notas de observação do grupo 7 durante a fase 5 do primeiro ciclo do estudo de caso.....	111
Quadro 50: Notas de observação do grupo 8 durante a fase 5 do primeiro ciclo do estudo de caso.....	111
Quadro 51: Categorias encontradas durante a codificação inicial das respostas da fase 4 no primeiro ciclo do estudo de caso. ....	112
Quadro 52: Categorias encontradas na codificação axial relativa aos dados da fase 4 da atividade com a turma de graduação.....	114
Quadro 53: Etapas da codificação dos resultados da fase 4 no primeiro ciclo do estudo de caso. ....	116

Quadro 54: Codificação obtida através das respostas dos estudantes à pergunta realizada na fase 5 do desafio, realizada de acordo com abordagem construtivista da Teoria Fundamentada em Dados.....	117
Quadro 55: Respostas da turma de pós-graduação à fase 4 da atividade.....	126
Quadro 56: Respostas da turma de pós-graduação à fase 5 da atividade.....	128
Quadro 57: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 1 do segundo ciclo do estudo de caso.....	129
Quadro 58: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 2 do segundo ciclo do estudo de caso.....	129
Quadro 59: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 3 do segundo ciclo do estudo de caso.....	131
Quadro 60: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 4 do primeiro ciclo do estudo de caso.....	132
Quadro 61: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 5 do primeiro ciclo do estudo de caso.....	134
Quadro 62: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 6 do primeiro ciclo do estudo de caso.....	135
Quadro 63: Codificação dos resultados obtidos para a fase 4 pelos estudantes da pós-graduação.....	137
Quadro 64: Codificação dos resultados obtidos para a fase 5 pelos estudantes da pós-graduação.....	139
Quadro 65: Resultado consolidado das respostas obtidas pelos estudantes da graduação e da pós-graduação para as respostas da fase 4.....	142
Quadro 66: Versão inicial da escala de Literacia de Dados gerada a partir da codificação dos resultados obtidos pelos estudantes a partir da 4 <sup>a</sup> fase do desafio Cartas para a Literacia.	145
Quadro 67: Resultado consolidado das respostas obtidas pelos estudantes da graduação e da pós-graduação para as respostas da fase 5. Em laranja os códigos gerados através da codificação dos textos da turma de graduação e em cinza os códigos gerados através da codificação dos textos da turma de pós-graduação. ....	149
Quadro 68: Ajustes sugeridos pelos especialistas para a escala de Literacia de Dados.....	159

## **LISTA DE GRÁFICOS**

Gráfico 1: Gráfico de teste obtido para a pergunta de pesquisa “Qual a sua preferência na formação de grupos para trabalhos escolares?”. .....	86
Gráfico 2: Gráfico de teste obtido para a pergunta de pesquisa “Em qual horário o seu estudo rende mais?” .....	86
Gráfico 3: Gráfico de teste obtido para a pergunta de pesquisa “O que é mais importante para você, entre as opções “aprender”, “as pessoas”, “tirar nota boa” e “outro”?.....	87

## **LISTA DE SIGLAS**

DSR	<i>Design Science Research</i>
BNCC	Base Nacional Comum Curricular
CNPQ	Conselho Nacional de Desenvolvimento Científico e Tecnológico
ADR	<i>Action Design Research</i>
FERPA	Family Educational Rights and Privacy Act
VLSS	Vietnamese Living Standard Study
UCLA	<i>University of California Los Angeles</i>
KDD	Knowledge Discovery in Datasets
GT	<i>Grounded Theory</i>
IFRJ	Instituto Federal de Educação, Ciência e Tecnologia
UNIRIO	Universidade Federal do Estado do Rio de Janeiro
FAETEC	Fundação de Apoio à Escola Técnica
UFRJ	Universidade Federal do Rio de Janeiro

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	19
1.1 MOTIVAÇÃO E CARACTERIZAÇÃO DO PROBLEMA.....	19
1.2 PROPOSTA DE SOLUÇÃO .....	22
1.3 OBJETIVOS DA DISSERTAÇÃO .....	24
1.4 METODOLOGIA DA PESQUISA.....	25
1.5 ESTRUTURA DA DISSERTAÇÃO .....	29
<b>2 REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS .....</b>	31
2.1 LITERACIA DE DADOS .....	31
2.2 BASE NACIONAL COMUM CURRICULAR .....	36
2.3 AVALIAÇÃO EM LITERACIA DE DADOS .....	40
2.3.1 Fazendo o bem com dados: desenvolvimento de um modelo de maturidade para alfabetização em dados em organizações não governamentais .....	40
2.3.2 O uso ético e apropriado de dados requer Literacia de Dados .....	42
2.3.3 Percepção gráfica de distribuições de valor: uma avaliação da alfabetização em dados de visualizadores não especialistas.....	45
2.3.4 Incorporando a alfabetização em dados em programas de alfabetização informacional de graduação em ciências sociais: um projeto piloto .....	47
2.3.5 Exploração dos pontos positivos e negativos da literatura revisada .....	49
2.3.6 Diferenças nos <i>designs</i> para a percepção e/ou desenvolvimento de competências em Literacia de Dados .....	52
2.3.7 Considerações sobre os processos de avaliação das propostas de pesquisa.....	53
2.4 COLABORAÇÃO E GROUP STORYTELLING .....	54
2.5 DESIGN SCIENCE RESEARCH .....	57
2.5.1 Trabalhos relacionados à criação de artefatos utilizando a DSR.....	59
<b>3 DISCUSSÃO DA SOLUÇÃO .....</b>	64
3.1 INSTANCIAMENTO DO DSR PARA ESTA PESQUISA.....	71
3.1.1 Etapa 1 (preparação da atividade/estudo de caso): .....	72
3.1.2 Etapa 2 (realização da atividade/estudo de caso): .....	74
3.1.3 Etapa 3 (Pós-atividade/estudo de caso) .....	74
3.2 ANÁLISE FUNDAMENTADA EM DADOS.....	75
3.2.1 A influência do referencial interpretacionista no uso da GT nesta pesquisa .....	79
<b>4 AVALIAÇÃO DO ARTEFATO CARTAS PARA A LITERACIA.....</b>	81
4.1 INTRODUÇÃO SOBRE O ESTUDO DE CASO E SEUS CICLOS .....	81
4.2 PRIMEIRO CICLO: GRADUAÇÃO .....	83

<b>4.2.1 Planejamento do primeiro ciclo.....</b>	83
<b>4.2.2 Atividade do primeiro ciclo .....</b>	87
<b>4.2.3 Resultados do primeiro ciclo.....</b>	90
<b>4.2.4 Discussão do primeiro ciclo.....</b>	112
<b>4.3 SEGUNDO CICLO: PÓS-GRADUAÇÃO.....</b>	122
<b>4.3.1 Planejamento do segundo ciclo .....</b>	123
<b>4.3.2 Atividade do segundo ciclo .....</b>	124
<b>4.3.3 Resultados do segundo ciclo .....</b>	125
<b>4.3.4 Discussão do segundo ciclo .....</b>	136
<b>4.4 SUMARIZAÇÃO DOS DADOS E VERSÃO INICIAL DA ESCALA .....</b>	141
<b>5 AVALIAÇÃO DA ESCALA E APRESENTAÇÃO DOS PRODUTOS DA PESQUISA.....</b>	157
<b>6 CONCLUSÃO .....</b>	185
<b>6.1 CONTRIBUIÇÕES .....</b>	186
<b>6.2 LIMITAÇÕES DA PROPOSTA.....</b>	187
<b>6.3 TRABALHOS FUTUROS.....</b>	188
<b>REFERÊNCIAS.....</b>	190
<b>ANEXO A - TABELAS COM AS UNIDADES TEMÁTICAS DA BNCC RELACIONADAS À LITERACIA DE DADOS .....</b>	198
<b>APÊNDICE A - RESULTADOS DO PRIMEIRO CICLO DO ESTUDO DE CASO .....</b>	209
<b>APÊNDICE B - RESULTADOS DO SEGUNDO CICLO DO ESTUDO DE CASO.....</b>	231

## 1 INTRODUÇÃO

---

Este capítulo apresenta aspectos introdutórios desta dissertação, como o contexto que a motivou, a descrição dos problemas observados, da pergunta de pesquisa levantada e dos objetivos a serem alcançados com a aplicação da metodologia de pesquisa que guiará este estudo.

### 1.1 MOTIVAÇÃO E CARACTERIZAÇÃO DO PROBLEMA

O crescimento sem precedentes históricos da quantidade de dados que são gerados e consumidos tem impacto em vários aspectos da sociedade contemporânea, incluindo negócios, marketing, agências governamentais, desenvolvimento científico e tecnológico, área da saúde, direito, meio ambiente, pesquisas em geral e a educação (Khan *et al.*, 2018). Segundo Tableau, 2019, os dados digitais aumentariam a uma taxa de crescimento composta anual<sup>1</sup> de 42% até 2020.

Por outro lado, vêm surgindo oportunidades de exercício da cidadania a partir do uso e reuso de dados abertos gerados tanto por organizações não governamentais, quanto por agências de governo. Segundo Locke e Heppler (2018), “nota-se que há um problema muito claro no que se refere à capacidade dos cidadãos em acessar e formatar dados para reuso, pois isso requer um grau considerável de conhecimento técnico”. O problema da falta de conhecimento técnico agrava-se quando se leva em conta que os governos muitas vezes disponibilizam dados abertos sem nenhum tratamento para facilitar seu uso pela população. Desta forma, é importante capacitar cidadãos para entender o mundo através de dados e, mais do que isso, construírem narrativas através deles, uma vez que uma parte importante da história deste mundo em que se vive hoje só é possível ser escrita por meio de dados (Knaflic, 2019).

Uma forma de promover a capacitação de cidadãos no entendimento sobre os dados manipulados é desenvolver competências relacionadas à Literacia de Dados (Frank *et al.*,

---

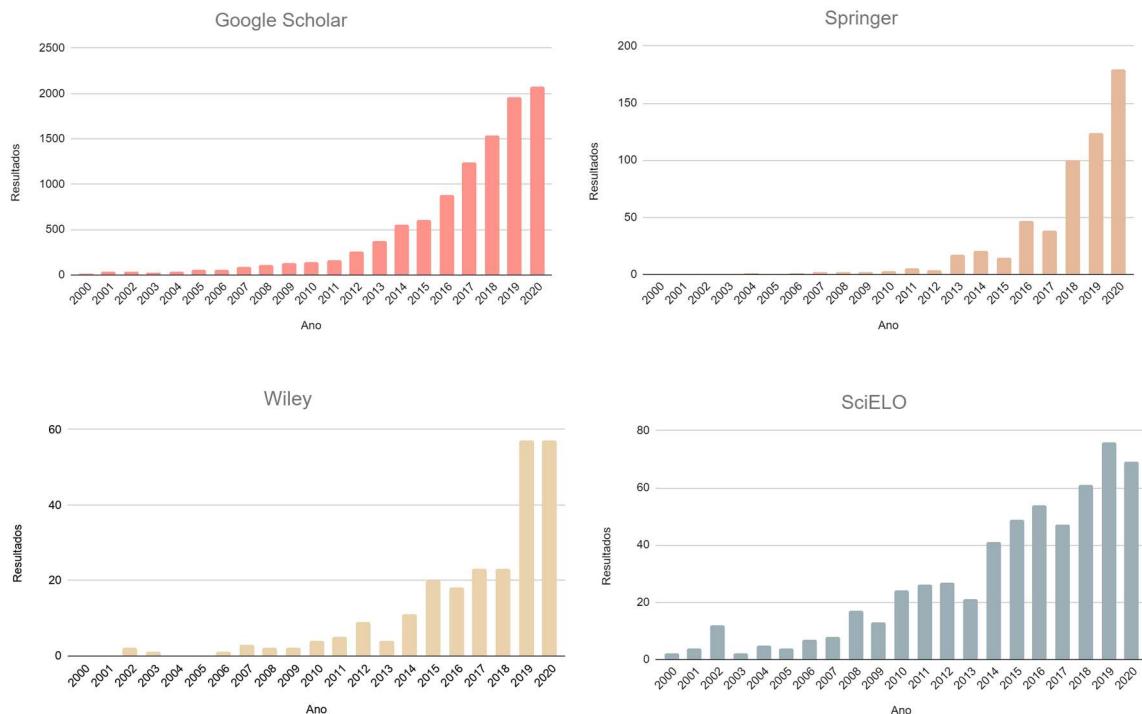
<sup>1</sup> A taxa de crescimento composta anual (Compound Annual Growth Rate) é um número fictício que estimaria à qual taxa um investimento teria crescido, caso os retornos fossem constantes, sendo calculada através da fórmula CAGR =  $(VF/VI)^{1/n}-1$ .

2016). O relatório “*A World that Counts*” apresentado pelo *Independent Expert Advisory Group Secretariat* (2014), motivou a criação do Fórum Mundial de Dados das Nações Unidas, que se reuniu pela 3ª vez no ano de 2020, contando com 6 grupos temáticos, entre eles um inteiramente dedicado à Literacia de Dados, intitulado: “*Understanding the World through Data*” e outro inteiramente dedicado à visibilidade de grupos vulneráveis através de dados: “*Leaving No One Behind*”, não deixando dúvidas de que dialogar sobre um mundo com dados é uma condição urgente e que pressupõe possibilitar que todos os cidadãos estejam aptos a participar do debate de forma plural e em igualdade de condições.

De acordo com Montoya (2018), a literacia é algo que ultrapassa o projeto de aprendizagem delimitado no tempo de curso de um currículo formal, sendo algo que se desenvolve ao longo de uma jornada de vida. Ainda de acordo com Montoya (2018) “literacia é a capacidade de identificar, compreender, interpretar, criar, comunicar e computar, usando materiais impressos e escritos, associados a contextos variados”. Bhargava *et al.* (2015) apenas ampliam o escopo da definição de *Data Literacy* dada por Montoya (2018), afirmando que “[...] é o desejo e a competência para o engajamento construtivo na sociedade através e com dados”.

Segundo Bhargava *et al.* (2015), Literacia de Dados é uma expressão que identifica, de forma mais geral, a capacidade que um indivíduo ou grupo de indivíduos possui para expressar suas ideias, argumentar e compreender informações por meio de dados. Ainda de acordo com os autores, a expressão *Data Literacy* começou a ser notada em relatórios das Nações Unidas publicados entre os anos de 2013 e 2014. Ao realizar buscas pela expressão *Data Literacy* no *Google Scholar*, na *Willey*, na *Scielo* e na *Springer* encontrou-se na base *Google Scholar* o aumento expressivo de 14 resultados recuperados para o ano de 2000 para 2080 resultados recuperados em novembro de 2020 (Figura 1).

Figura 1: Número de artigos retornados através de busca pela expressão “*Data Literacy*” para os anos de 2000 a 28 de novembro de 2020 em bases indexadas Google Scholar, Springer, Wiley e SciELO.



Fonte: reprodução da autora (2020).

*Data Literacy* pode ser traduzido como “literacia em dados”, expressão que foi escolhida para reconhecer o objeto de estudo deste trabalho. Entretanto há publicações que utilizam a expressão “alfabetização em dados” (Tygel e Kirsch, 2016; Yamauti e Tsunoda, 2019) e ainda “inclusão em dados” (Bhargava *et al.*, 2015) para fazer referência ao mesmo campo de estudos da Literacia de Dados. O uso da expressão Literacia de Dados parece razoável para esta pesquisa especificamente, uma vez que é uma expressão recente, mesmo em países anglófonos e que ainda não possuem acordo sobre a sua definição, como será mostrado no Capítulo 2 desta pesquisa. Também nos parece razoável a escolha desta expressão, porque será utilizada uma de suas definições, proveniente da língua inglesa, a fim de criar um artefato para a medição da Literacia de Dados.

Desta forma, principalmente pela atualidade do estudo da Literacia de Dados, e porque pretende-se realizar um diálogo acadêmico e educacional de concentração mundial sobre o tema, optou-se por utilizar a expressão “Literacia de Dados”, para que se possa participar desse diálogo em condições de igualdade na proposição de ações que levem a reflexões conjuntas na comunidade científica em geral a partir da prática educativa nacional.

A partir de pesquisas realizadas em um levantamento bibliográfico inicial, diagnosticou-se os seguintes problemas com relação à Literacia de Dados em contextos estudantis brasileiros: (i) ausência de um debate social aprofundado acerca da necessidade da promoção de Literacia de Dados voltada para uma sociedade de dados plural e inclusiva, (ii) ausência de currículo formal que conte cole as necessidades estudantis básicas de Literacia de Dados, (iii) ausência de programas governamentais voltados para a avaliação de Literacia de Dados, (iv) dificuldade em propor ações educativas e avaliativas que contemplam as necessidades, características e particularidades dos estudantes nacionais.

Com base nos problemas levantados, esta dissertação dará foco à pergunta de pesquisa:

**"De que forma podemos avaliar os conhecimentos de estudantes em  
Literacia de Dados?"**

## 1.2 PROPOSTA DE SOLUÇÃO

Desenhar uma atividade ou um projeto para aprendizagem em qualquer campo do saber passa necessariamente pela fase de exploração contextual de todas as diversas áreas do conhecimento que irão alimentar o processo de organização cognitiva que pretende levar o estudante ao alcance dos objetivos educacionais previstos para cada etapa de seus estudos. Magnusson *et al.* (1999) observaram que as perguntas: “O que devo fazer com meus alunos para ajudá-los a entender esse conceito [...]?”, “Que materiais existem para me ajudar?”, “O que meus alunos provavelmente já sabem e o que será difícil para eles?” e “Qual a melhor forma de avaliar o que meus alunos aprenderam?” distinguem um professor de um especialista que não leciona e transita na mesma área de conhecimento. A essas perguntas ainda podem ser adicionadas: “o que os alunos deverão saber fazer ao final de uma etapa de aprendizagem?” e “qual é o impacto que saber aprender coletivamente pode causar na aprendizagem?”, como norteadoras para a reflexão sobre quais competências os estudantes precisam desenvolver no campo da literacia em dados.

Segundo Guy Le Boterf (1994), competência é “mais um processo do que um conhecimento ou estado possuído”, sendo a expressão da aliança entre conhecimentos, habilidades, atitudes e valores envolvidos nas ações de resolver as solicitações multiformes

da vida, de exercitar irrestritamente a cidadania e de responder aos desafios do contexto laboral.

A expressão “saber fazer” está geralmente relacionada ao conhecimento tácito, que pode ser definido, segundo Chugh (2015), como “habilidades, ideias e experiências que as pessoas têm individualmente, mas não são codificadas e [por esse motivo] podem não ser necessariamente expressas com facilidade”. Já “Saber aprender” pressupõe a interação com outros indivíduos e o acesso ao conhecimento coletivo e à colaboração que, segundo Costa *et al.* (2014) envolve a articulação entre (i) Comunicação: que objetiva a transmissão do conteúdo da informação em linguagem perceptível para os envolvidos, além do uso de ferramentas de comunicação para que ela se processe; (ii) Coordenação: que é o ato de gerenciar as interdependências emergentes entre as atividades básicas para atingir uma meta (Malone e Crowston, 1990) e (iii) Cooperação: que é o trabalho que se desenvolve individualmente, mas com um objetivo comum em um espaço compartilhado.

Esta pesquisa oportunizará investigar quais seriam possibilidades de avaliar o alcance ou não alcance, ou mesmo o alcance em diferentes graduações e formas, dos objetivos de aprendizagem propostos em uma tarefa voltada para a Literacia de Dados. Um motivo para avaliar a Literacia de Dados é identificar indivíduos ou grupos de indivíduos que se beneficiarão com mais instruções sobre dados, quais instruções seriam estas, quais são os formatos apropriados para oferecê-las e avaliar se o que e como se ensina hoje está ajudando a criar condições que facilitem a aprendizagem.

**Como proposta de solução para a pergunta desta pesquisa será sugerido um artefato para avaliação da Literacia de Dados de estudantes a partir da definição de Literacia de Dados de D'Ignazio e Bhargava (2015), utilizando a abordagem metodológica da Design Science Research (DSR) em uma investigação de caráter qualitativo e visão interpretacionista.**

Para realizar a avaliação do artefato criado e das conjecturas teóricas que direcionaram o seu projeto, serão realizadas rodadas de estudo de caso em turmas de diferentes níveis de escolaridade, utilizando técnicas de observação direta do comportamento dos estudantes aliadas aos resultados alcançados por eles em uma atividade proposta em sala de aula. Para analisar os resultados da avaliação, será utilizada a Teoria Fundamentada em Dados aliada à técnica de triangulação metodológica. A triangulação metodológica é uma

técnica em que se utiliza mais de um método de análise, com o objetivo de, ao confrontá-los, obter-se a maximização da validade das conclusões de pesquisa obtidas (Duarte, 2009). A discussão da solução será apresentada no Capítulo 3.

### 1.3 OBJETIVOS DA DISSERTAÇÃO

Com esta pesquisa pretende-se ampliar o conhecimento acerca do tema “Literacia de Dados” e das atividades que têm sido conduzidas para o seu alcance em ambientes educativos, para com isso propor um artefato que permita avaliar características e particularidades da Literacia de Dados de estudantes do ensino médio da educação básica. O problema a ser discutido é a dificuldade de avaliar a Literacia de Dados de estudantes.

Almeja-se então alcançar um maior amadurecimento no que diz respeito às evidências e lacunas de Literacia de Dados dos estudantes, buscando-se com esta pesquisa proporcionar meios para a ampliação da capacidade de profissionais da educação em lidar com a necessidade de abordar o tema Literacia de Dados em diferentes contextos estudantis, com base na análise dos dados de estudos de caso através de atividades que serão propostas em salas de aula.

**Objetivo geral desta pesquisa:** Prover um meio de avaliação da Literacia de Dados dos estudantes.

**Objetivos específicos:** a fim de atender o objetivo geral da pesquisa, espera-se: (i) Identificar e explorar o referencial bibliográfico sobre o tema Literacia de Dados; (ii) identificar pontos de convergência entre o currículo de Matemática nacional e a definição de Literacia de Dados adotada nesta pesquisa; (iii) Propor uma versão inicial de uma escala de Literacia de Dados baseada na literatura da área, a fim de fundamentar a criação de uma escala de Literacia de Dados; (iv) Construir atividades em formato de desafios que favoreçam a investigação do comportamento dos níveis de uma escala de medição da Literacia de Dados.

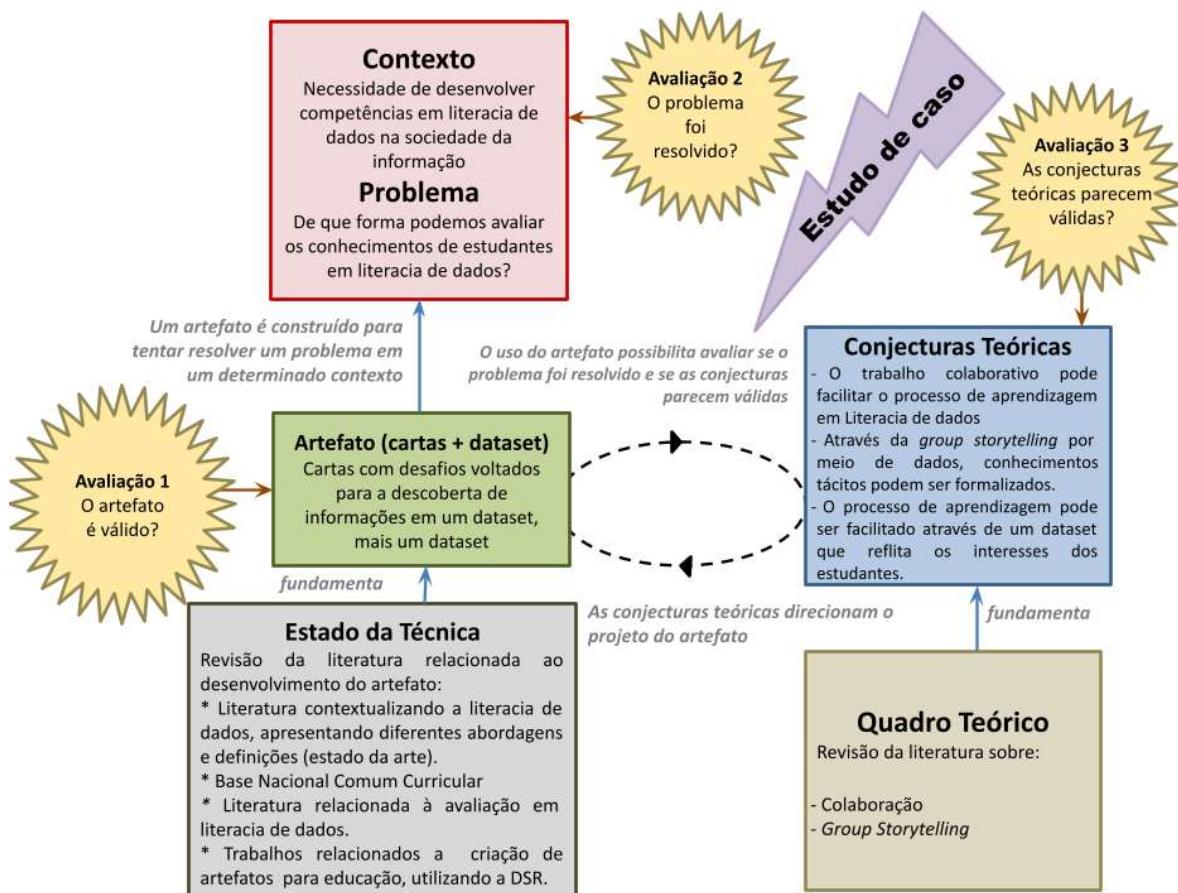
## 1.4 METODOLOGIA DA PESQUISA

Pelo fato de a Literacia de Dados ser um fenômeno que se expressa principalmente por meio da comunicação, entende-se que para iniciar um empreendimento exploratório na sua direção, precisa-se lançar mão de técnicas e métodos que interajam com esta característica inerente ao objeto de estudo.

Ao observar a maioria dos métodos de pesquisa tradicionais de pesquisa científica, não se encontra a previsão de um artefato no contexto da pesquisa, mas somente como produto final da mesma. Escolheu-se a metodologia *Design Science Research* (Ciência do *design* de pesquisa) por apresentar uma abordagem mais adequada a esta pesquisa, uma vez que se pretende desenvolver um artefato e adaptá-lo progressivamente para, a partir da sua aplicação em diferentes contextos estudantis, testar a sua validade, bem como das conjecturas teóricas que embasaram a construção do artefato, a fim de alcançar a resposta para a pergunta desta pesquisa.

Segundo Pimentel *et al.* (2019), a *Design Science Research* (DSR) é uma abordagem epistemológico-metodológica que torna possível a criação de artefatos inovadores através de pesquisas científicas rigorosas. Os mesmos autores apresentam a ideia original de um mapa de elementos esperados de uma pesquisa na abordagem proposta pela DSR com o objetivo de clarificar a correlação entre o desenvolvimento do artefato e o conhecimento científico-teórico envolvidos na busca de soluções para um determinado problema de pesquisa, independentemente dos métodos que serão utilizados para sua realização.

Figura 2: Mapa da Estrutura metodológica da pesquisa baseada na DSR.



Fonte: a autora (2020). Adaptado de Pimentel *et al.* (2019).

A Figura 2 apresenta o mapa metodológico da pesquisa baseada na DSR, de acordo com Pimentel *et al.* (2019). Nela, observa-se que o artefato e as conjecturas teóricas que direcionam o seu projeto encontram-se no centro da pesquisa, e o objetivo da sua criação é resolver um problema de pesquisa que está inserido em um determinado contexto. Para fundamentar a criação do artefato é necessário um estudo aprofundado do estado da arte, que se dá por meio da revisão de literatura relacionada ao desenvolvimento do artefato. Para alimentar as conjecturas teóricas é fundamental traçar um quadro teórico (revisão de literatura) sobre teorias que fundamentam as conjecturas teóricas levantadas. É muito importante perceber que na DSR o método de pesquisa utilizado está diretamente relacionado às três avaliações indispensáveis para a pesquisa: a avaliação da validade do artefato, a avaliação da validade das conjecturas teóricas e a avaliação sobre se o problema foi resolvido através do arcabouço metodológico empenhado (DSR).

Instanciando o mapa da estrutura metodológica para esta pesquisa, foi detalhado o problema de pesquisa. Sua solução está associada à pergunta: “De que forma podemos avaliar os conhecimentos de estudantes em Literacia de Dados?”. O contexto se traduz no público alvo escolhido dentro do cenário sócio-histórico e geográfico em que se encontram estudantes de ensino fundamental, médio, graduação e pós-graduação do ensino brasileiro vivendo na sociedade pós moderna orientada ao dado.

O artefato criado é composto por um conjunto de cartas com desafios voltados para a descoberta de informações em um *dataset* disponibilizado. Os desafios foram criados com base na definição de Literacia de Dados dos autores D’Ignazio e Bhargava (2015).

Para a criação do artefato conjecturou-se que: (i) o trabalho colaborativo pode facilitar o processo de aprendizagem em Literacia de Dados; (ii) através de narrativas por meio de dados, conhecimentos tácitos dos estudantes em Literacia de Dados podem ser formalizados; (iii) a aprendizagem em Literacia de Dados pode ser facilitada através de um dataset que refletia os interesses dos estudantes. O quadro teórico que fundamenta as conjecturas teóricas levantadas compõe-se de conhecimentos científicos sobre: (i) comportamento; (ii) colaboração e (iii) *group storytelling*.

Para avaliar a validade do artefato, das conjecturas teóricas que o fundamentaram e avaliar se o problema foi resolvido, mostrou-se adequada a aplicação de um estudo de caso em 2 interações: a primeira com estudantes recém ingressos no curso de graduação em Ciência da Computação da UFRJ, e a segunda com estudantes de pós-graduação em Informática.

A pesquisa foi conduzida de forma a utilizar um artefato proposto para medir a Literacia de Dados de estudantes. Para tanto, foi realizado um estudo de caso para pesquisar sobre como os estudantes responderiam à aplicação de uma atividade sobre Literacia de Dados utilizando o artefato criado. Espera-se que a partir dessas pesquisas seja possível gerar conhecimento suficiente para estabelecer gradativamente os requisitos para a melhoria do artefato proposto e da proposta de avaliação de competências em Literacia de Dados de estudantes.

Como instrumentos para a análise dos dados coletados durante a etapa de avaliação do artefato, serão utilizadas técnicas de triangulação como as abordadas em *Triangulation* (2014), Flick (2004) e Oliver-Hoyo e Allen (2006). Serão triangulados os resultados das

respostas das atividades propostas durante o estudo de caso com dados da observação direta dos estudantes durante a realização das atividades. Para analisar os dados pretende-se utilizar uma abordagem da Teoria Fundamentada em Dados, que será abordada no capítulo 3, seção 3.3 deste trabalho.

A estrutura lógica de aplicação da pesquisa é formalizada de acordo com os seguintes passos (Figura 3):

Figura 3: Estrutura lógica de aplicação da pesquisa.



Fonte: reprodução da autora (2020).

**Passo 1:** Conduzir pesquisa exploratória da bibliografia sobre os temas:

- Literacia de Dados e suas diferentes definições.
- Base Nacional Comum Curricular.
- Trabalhos relacionados à avaliação em Literacia de Dados.
- Trabalhos relacionados ao uso de artefatos sociotécnicos em contextos educacionais através do paradigma epistemológico-metodológico da DSR.

**Passo 2:** Criar um artefato utilizando uma bricolagem entre cartas de desafio e um *dataset* para análise, que possa auxiliar na aplicação de uma experiência de aprendizagem que auxilie na medição da Literacia de Dados de estudantes em sala de aula.

**Passo 3:** Realizar testes empíricos para avaliação da validade do artefato e das conjecturas teóricas.

- Realizar a primeira interação do estudo de caso com estudantes do primeiro período do curso de graduação em Ciência da Computação da UFRJ. Avaliar a atividade, a validade do artefato e das conjecturas teóricas levantadas para a sua construção e implementar os possíveis redesenhos necessários no artefato e na atividade.
- Realizar a segunda interação do estudo de caso com estudantes do Programa de Pós-Graduação em Informática da UFRJ. Avaliar a atividade, a validade do artefato e das conjecturas teóricas para a sua construção e implementar os possíveis redesenhos necessários no artefato e na atividade.

**Passo 4:** Discutir a análise dos resultados dos testes empíricos implementados para a avaliação da validade do artefato, da validade das conjecturas teóricas e se o artefato ajudou a responder à pergunta de pesquisa e atendeu aos objetivos propostos.

**Passo 5:** Apresentar os resultados obtidos, os produtos prometidos e considerações acerca do processo.

## 1.5 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está estruturada em 6 capítulos, contando com este como Introdução. Os capítulos estão organizados conforme apresentado a seguir:

- **Capítulo 1** – Introdução: expõe a motivação e caracterização do problema, a proposta de solução, os objetivos desta dissertação, a metodologia da pesquisa adotada e a estrutura da dissertação.
- **Capítulo 2** – Referencial teórico: apresenta uma revisão da literatura sobre a definição de Literacia de Dados, levanta pontos de convergência entre a Literacia de Dados e a Base Nacional Comum Curricular e aponta os trabalhos relacionados com avaliação em Literacia de Dados e construção de artefatos utilizando a DSR. Este capítulo também apresenta considerações acerca da colaboração e *group storytelling* para apoiar a educação.

- **Capítulo 3** – Discussão da solução: aponta a proposta de solução e destaca a sua importância. Explica a abordagem metodológica *Design Science Research*. Comenta a instanciação desta pesquisa no Mapa de Elementos Esperados de uma Pesquisa na abordagem da DSR, destacando quais serão os métodos de pesquisa que serão aplicados para a solução do problema.
- **Capítulo 4** – Avaliação do artefato Cartas para a Literacia: detalha a avaliação da implementação do artefato Cartas para a Literacia, incluindo as fases Planejamento da Avaliação da Pesquisa, Execução da Avaliação para cada uma das implementações do estudo de caso e Apresentação dos Resultados para cada uma das implementações do estudo de caso.
- **Capítulo 5** – Avaliação da escala e apresentação dos resultados da pesquisa: avalia a aceitação da escala de Literacia da Dados obtida, da operacionalização das conjecturas teóricas levantadas e se o artefato atendeu os objetivos propostos e ajudou a responder à pergunta de pesquisa.
- **Capítulo 6** – Conclusão: resume os resultados alcançados, aborda as limitações do trabalho e indica os próximos passos desta pesquisa.

## 2 REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS

---

Neste capítulo será apresentado o referencial teórico necessário para a realização da inferência do *design* do artefato que se pretende construir com o objetivo de criar uma escala de medição de Literacia de Dados. Os temas abordados na revisão de literatura foram: definição de Literacia de Dados; Base Nacional Comum Curricular; Avaliação em Literacia de Dados; Colaboração; *Group storytelling* e trabalhos relacionados à criação de artefatos utilizando a DSR.

### 2.1 LITERACIA DE DADOS

Existem diferentes definições para Literacia de dados, e cada uma delas revela as necessidades e aspirações de cada contexto de pesquisa onde têm sido desenvolvidas e empregadas. Para a realização desta etapa de pesquisa foi feita uma busca pelos artigos na base *Google Scholar*, através do emprego da *string* de busca “*Data Literacy*”. Na data em que a busca foi realizada, obteve-se o quantitativo de 1340 artigos. Através da técnica de *snowballing* foi feita a seleção dos artigos que representavam o estado da arte em ensino-aprendizagem de Literacia de Dados em ambientes escolares e acadêmicos.

De acordo com D'Ignazio e Bhargava (2015) “a Literacia de Dados inclui a capacidade de ler, trabalhar, analisar e argumentar com dados. Ler dados envolve a compreensão de quais são os dados necessários para representar um universo e que aspectos desse universo esses dados representam. Trabalhar com dados envolve criação, aquisição, limpeza e gerenciamento. Analisar dados envolve filtrar, classificar, agrregar, comparar e realizar outras operações analíticas neles. Argumentar com dados envolve utilizar dados como suporte para uma narrativa destinada a comunicar alguma mensagem para um público específico”.

Para Gibson e Mourad (2018), a Literacia de Dados é muito mais que numeracia – competência para processar, comunicar e interpretar informações numéricas em uma variedade de contextos (Askew, 1997) – e representação gráfica. Literacia de Dados é um conceito amplo que também inclui o entendimento de quais ferramentas analíticas usar, como aplicar essas ferramentas como um componente integrado de um experimento, como interpretar os dados analisados à luz de uma pergunta específica ou de uma hipótese e como comunicar resultados efetivamente em diferentes plataformas.

O relatório *Strategies and Best Practices for Data Literacy Education*, da Universidade de Dalhousie, define alfabetização de dados como a capacidade de coletar, gerenciar, avaliar e aplicar dados de uma forma crítica (Ridsdale *et al.*, 2015). No Blog da curadora de dados ChartMogul, Shelley (2019) define *Data Literacy* como a habilidade de ler, entender, criar e comunicar dados como informação, sendo ainda a habilidade de trabalhar de forma efetiva com dados.

Rouse (2019) levantou as habilidades relacionadas à Literacia de Dados: (i) Saber quais dados são apropriados para a utilização com um determinado propósito; (ii) Interpretação de dados e visualizações (por exemplo: gráficos e tabelas); (iii) Pensar de forma crítica sobre informações produzidas a partir de análises de dados; (iv) Entendimento sobre ferramentas e métodos de análise de dados e em que contexto (onde e quando) usá-las; (v) Reconhecer quando um dado está deturpado ou iludindo o leitor; (vi) Comunicar-se através de dados com pessoas que possuem ausência ou baixa Literacia de Dados através de data *storytelling* ou outras técnicas.

Sternkopf e Mueller (2018) definem Literacia de Dados como a habilidade de utilizar dados de forma produtiva, pensando sobre eles de uma forma reflexiva. Já na pesquisa de Kjelvik e Schultheis (2019), eles defendem que para entender o que é Literacia de Dados, é preciso voltar o olhar para os diversos campos que a compõem, bem como para as diversas abordagens em que pode se apresentar. Desta forma, apresentam a expressão “Literacia de Dados” como oriunda da interseção entre três campos de estudo distintos: (i) Raciocínio quantitativo, (ii) Ciência de dados e (iii) Conhecimento sobre o contexto de estudo (Figura 4). Sendo o raciocínio quantitativo a “habilidade de aplicar o raciocínio matemático na resolução de problemas cotidianos, através da compreensão de gráficos, tabelas e estatística descritiva para a expressão do raciocínio lógico e coerente sobre informações quantitativas”.

Figura 4: Na abordagem de Kjelvik e Schultheis (2019), a Literacia de Dados está na interseção entre Ciência de Dados, o raciocínio quantitativo e a compreensão do conteúdo da área estudada.



Fonte: a autora (2020). Adaptado de Kjelvik e Schultheis (2019).

Khan *et al.* (2018), partindo da observação de que não há entendimento sobre a definição de literacia em dados, propuseram a si mesmos o desafio de encontrar uma definição holística para o tema, de forma que um conjunto de abordagens diversas e complementares fossem capazes de expressar seu sentido com o máximo de integridade possível. Para isto, empreenderam busca em diversas bases de dados acadêmicas através de termos como “*data*”, “*big data*”, “*research data*”, combinados a outros como “*literacy*”, “*literate*”, “*competency*”, “*knowledge*” e “*skill*”. A partir dos sessenta e nove artigos que retornaram de sua pesquisa, os autores delinearam, de forma preliminar, três abordagens principais para a expressão “*Data Literacy*”: (i) habilidades necessárias para a tomada de decisões orientadas por dados, (ii) Literacia de Dados enquanto atividade de apoio à pesquisa e (iii) Literacia de Dados como um conjunto de práticas que ocorrem nos estágios do ciclo de vida dos dados.

Ghosh (2019) escreveu um artigo para o site *The Customize Windows*, em que define Literacia de Dados como: “a habilidade de entender e utilizar dados efetivamente no contexto de uma coleção de dados, do compartilhamento de dados e da análise de dados”. Para ele a expressão “Literacia de Dados” soa como “familiaridade com dados e seus usos”. Os autores comentam que da mesma forma que a literacia digital —uso seguro, crítico e criativo das Tecnologias da Informação e da Comunicação para o alcance de metas relacionadas com a

inclusão, o lazer, o aprendizado e a empregabilidade na sociedade (Ala-Mutka, 2011) — não tem limites bem definidos, a Literacia de Dados também não tem.

Tygel e Kirsh (2016) em uma parceria do Programa de Pós-Graduação em Informática da UFRJ com a cooperativa EITA<sup>2</sup> propõem, a partir de estudos do método de alfabetização de Freire (1970) o conceito de alfabetização de dados crítica como sendo “o conjunto de habilidades que permite que se use e produza dados de forma crítica. Este conjunto é composto por: Leitura de dados; Compreensão de como os dados foram gerados e de que não são fatos objetivos, mas produtos de um processo social; Manipulação técnica de dados; Uso de ferramentas computacionais e estatísticas para transformar dados em informação, além da interligação entre dados de diferentes fontes e; Comunicação de dados - que compreende encontrar as melhores relações entre os tipos de dados, as ferramentas de comunicação mais eficientes e realizar a avaliação social da mensagem a ser transmitida e do seu público alvo”.

Hu *et al.* (2019) preocupam-se em primeiramente definir literacia da informação como “a capacidade de adquirir, analisar, processar, utilizar, criticar e entender o conteúdo da informação, bem como a atitude e capacidade de integração na sociedade da informação”. Depois ocupam-se em diferenciar “Literacia da Informação” de “Literacia de Dados” e de “Análise de Dados Científicos”. Todavia, em sua pesquisa, abordam a “Literacia de Dados” de forma integrada com a “Literacia da Informação”, mas definindo literacia de dados como as habilidades de posicionar, organizar, entender e criar informação, usando tecnologias digitais para o processamento de dados.

Hoelter (2017) definem literacia estatística — que neste caso específico é tratada como sinônimo de Literacia de Dados — como sendo um conjunto de habilidades tais como: (i) ler e interpretar uma tabela ou gráfico, (ii) calcular porcentagens, (iii) trabalhar usando um modelo científico, (iv) avaliar dados em que argumentos são baseados, utilizando dados para a tomada de suas próprias decisões e argumentações, além de (v) saber quais tipos de dados podem ser úteis para ajudar a responder perguntas distintas.

Para a fundamentação desta pesquisa será utilizada a expressão “alfabetização de dados” como sinônimo para inclusão de dados, letramento em dados e Literacia de Dados. Será adotada a definição de alfabetização de dados publicada por D'Ignazio e Bhargava (2015):

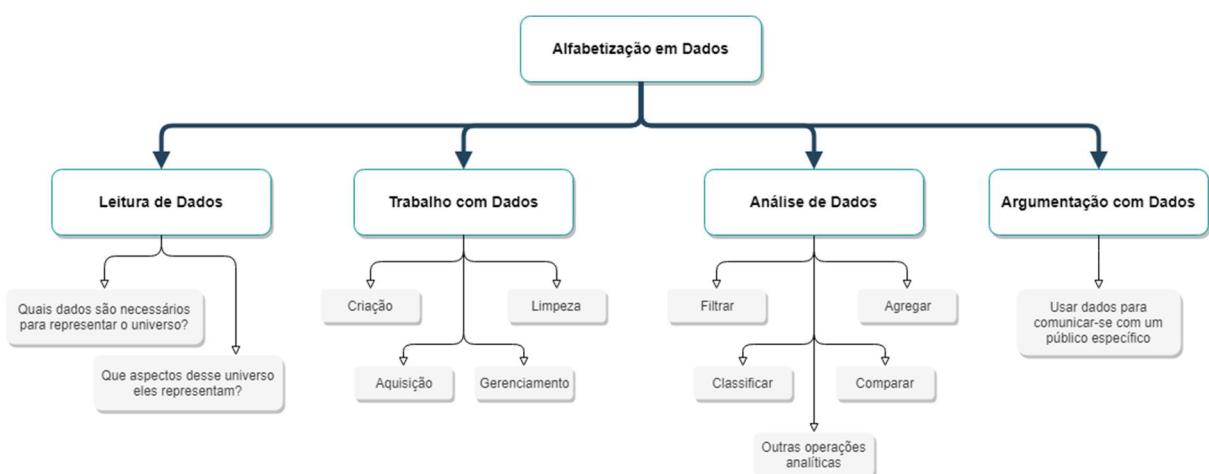
---

<sup>2</sup> <<https://eita.coop.br/>>. Acesso em 09/12/2019.

“A literacia em dados inclui a capacidade de ler, trabalhar, analisar e argumentar com dados. Ler dados envolve a compreensão de quais são os dados necessários para representar um universo e que aspectos desse universo esses dados representam. Trabalhar com dados envolve criação, aquisição, limpeza e gerenciamento. Analisar dados envolve filtrar, classificar, agregar, comparar e realizar outras operações analíticas neles. Argumentar com dados envolve utilizar dados como suporte para uma narrativa destinada a comunicar alguma mensagem para um público específico” (D’Ignazio e Bhargava, 2015).

Na Figura 5, estão relacionadas as quatro capacidades que D’Ignazio e Bhargava (2015) relacionam ao aprendizado de Literacia de Dados, bem como as ações que um indivíduo que possui Literacia de Dados deve ser capaz de realizar. Assim, se alguém possui a capacidade de leitura de dados, essa pessoa deverá ser capaz de identificar quais dados são necessários para representar o universo que está sendo estudado. Para trabalhar com dados, uma pessoa deverá ser capaz de criar dados, adquirir dados, gerenciar dados e limpar dados. Quando uma pessoa analisa dados, ela deverá ser capaz de filtrar dados, agregar dados, classificar dados, comparar dados e ainda realizar outras operações analíticas. Ainda, para argumentar com dados, uma pessoa deverá ser capaz de usar dados para comunicar-se com um público específico.

Figura 5: Neste mapa mental é apresentada uma proposta de organização visual da definição de D’Ignazio e Bhargava (2015) para *Data Literacy*.



Fonte: a autora (2020). Adaptado de D’Ignazio e Bhargava (2015).

## 2.2 BASE NACIONAL COMUM CURRICULAR

No contexto brasileiro, o debate curricular tem sido abordado pelo Ministério da Educação segundo as normas da Base Nacional Comum Curricular (BNCC) (Brasil, 2017), documento que foi homologado no ano de 2018 e que normatiza oficialmente o conjunto de aprendizagens essenciais que todo o aluno deve desenvolver durante a Educação Básica.

Atualmente não há nenhum grupo de pesquisa dedicado ao tema Literacia de Dados cadastrado e certificado no Diretório de Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) (CNPQ, 2019). Como a BNCC trabalha com a ideia de que as habilidades não são desenvolvidas de forma genérica e descontextualizada (Brasil, 2017), procuramos pontos de relacionamento entre as habilidades de outras áreas de estudo e a área da Matemática e que se inserem no campo da Literacia de Dados, apesar de não serem situadas na organização curricular da BNCC como atividades que compõem o campo de estudo da Literacia de Dados.

Apesar disso, é importante notar que a BNCC teve sua construção homologada no ano de 2018 e que, apesar de ter status de norma curricular de abrangência nacional:

i. não menciona o termo literacia em dados e nem suas variações, como letramento de dados, alfabetização em dados ou inclusão em dados, revelando que não foi fruto de um amadurecimento nacional que contivesse um diálogo aprofundado sobre o tema da Literacia de Dados no contexto da sociedade orientada ao dado.

ii. não considera que, para que as habilidades em geral apresentadas em sua documentação sejam desenvolvidas pelos estudantes, levando em conta os seus diversos contextos socioculturais, é necessário que haja um conjunto de ações, que em um primeiro momento podemos dizer que incluem avaliações diagnósticas, de proficiência e de percurso, que indiquem o estado inicial de competência que um estudante possui sobre um determinado tema e acompanhe a sua evolução ao longo do tempo, considerando as situações de aprendizagem a que forem expostos e as particularidades cognitivas de cada estudante. Ao contrário,

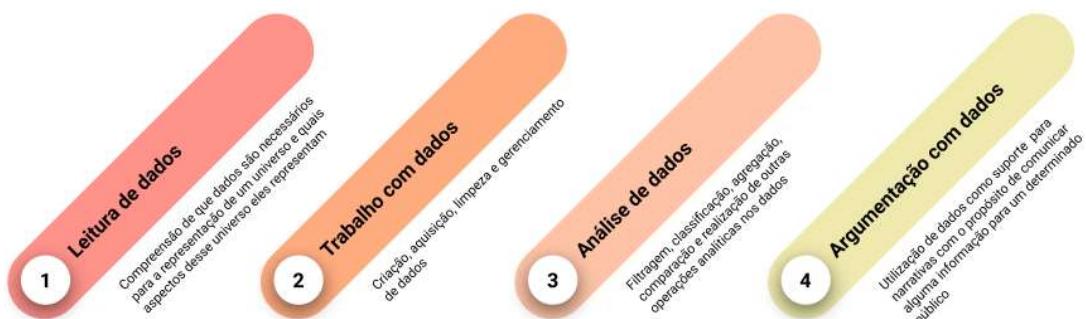
iii. situa as habilidades que um estudante deve desenvolver para alcançar o sucesso escolar em uma lista exaustiva, que, apesar de fornecer aos profissionais da educação um ponto de partida interessante para dialogar sobre os temas adequados à aprendizagem em cada etapa da educação básica, não se ocupa em conduzir à abertura de oportunidades de

progresso na aprendizagem levando em conta diferentes trajetórias comportamentais e socioculturais dos estudantes, professores e comunidades escolares, que percebemos como pedras de toque para alternativas de educação desenhadas para um país com dimensões continentais e tanta diversidade.

Desta forma, a fim de exercitar a dialética para gerar conhecimento em Literacia de Dados a partir das habilidades e competências anotadas na BNCC, compilou-se as habilidades codificadas para a área de Matemática e suas Tecnologias que se percebeu conterem recortes epistemológicos mais evidentes com o estudo da Literacia de Dados. O critério de seleção baseou-se na definição de Literacia de Dados proposta pelos autores D'Ignazio e Bhargava (2015), ou seja, foram escolhidas as habilidades que tivessem o escopo das competências salientes na definição de Literacia de Dados dada pelos autores (Figura 6):

1. Leitura de dados: compreensão de que dados são necessários para a representação de um universo e quais aspectos desse universo eles representam.
2. Trabalho com dados: criação, aquisição, limpeza e gerenciamento de dados.
3. Análise de dados: filtragem, classificação, agregação, comparação e realização de outras operações analíticas nos dados.
4. Argumentação com dados: utilização de dados como suporte para narrativas com o propósito de comunicar alguma informação para um determinado público.

Figura 6: Competências destacadas na definição de Literacia de Dados de D'Ignazio e Bhargava (2015).



Fonte: a autora (2020). Adaptado de D'Ignazio e Bhargava (2015).

Foi proposta uma organização em forma de tabelas que contemplam unidades temáticas, objetos de conhecimento e habilidades da BNCC que percebeu-se conter semelhança epistemológica com o escopo da Literacia de Dados oferecido pela definição de D'Ignazio e Bhargava (2015), organizadas segundo a divisão dos anos escolares vigente para os ensinos fundamental e médio. Essas tabelas se encontram no Anexo 1 desta dissertação (tabelas i-ix). No que diz respeito à Literacia de Dados para o ensino médio, na BNCC há indicações de metas que remetem à alfabetização de dados relacionadas aos campos do saber: Linguagens e suas tecnologias - língua portuguesa; Matemática e suas tecnologias; Ciências da natureza e suas tecnologias; e Ciências humanas e sociais aplicadas. No Quadro 1 estão organizadas competências específicas da área de Matemática e suas Tecnologias relacionadas às suas habilidades que, após análise, apresentaram semelhança epistemológica com a Literacia de Dados.

Quadro 1: Competências específicas e habilidades listadas pela BNCC para o Ensino Médio na área de Matemática e suas Tecnologias.

<b>Competência específica da área de Matemática e suas Tecnologias</b>	<b>Habilidades que apresentam semelhança epistemológica com a literacia em dados</b>
1) utilizar estratégias, conceitos e procedimentos matemáticos para interpretar situações em diversos contextos, sejam atividades cotidianas, sejam fatos das ciências da natureza e humanas, ou ainda questões econômicas e tecnológicas, divulgados por diferentes meios, de modo a consolidar uma formação científica geral.	(EM13MAT102) Analisar gráficos e métodos de amostragem de pesquisas estatísticas apresentadas em relatórios divulgados por diferentes meios de comunicação, identificando, quando for o caso, inadequações que possam induzir a erros de interpretação, como escalas e amostras não apropriadas.
2) articular conhecimentos matemáticos ao propor e/ou participar de ações para investigar desafios do mundo contemporâneo e tomar decisões éticas e socialmente responsáveis, com base na análise de problemas de urgência social, como os voltados a situações de saúde, sustentabilidade, das implicações da tecnologia no mundo do trabalho, entre outros, recorrendo a conceitos, procedimentos e linguagens próprios da matemática.	(EM13MAT202) Planejar e executar pesquisa amostral usando dados coletados ou de diferentes fontes sobre questões relevantes atuais, incluindo ou não, apoio de recursos tecnológicos, e comunicar os resultados por meio de relatório contendo gráficos e interpretação das medidas de tendência central e das de dispersão.

<p>3) Utilizar estratégias, conceitos e procedimentos matemáticos, em seus campos - Aritmética, Álgebra, Grandezas e Medidas, Geometria, Probabilidade e estatística para interpretar, construir modelos e resolver problemas em diversos contextos, analisando a plausibilidade dos resultados e a adequação das soluções propostas, de modo a construir argumentação consistente.</p>	<p>Não se aplica à definição Literacia de Dados de D'Ignazio e Bhargava (2015).</p>
<p>4) Compreender e utilizar, com flexibilidade e fluidez, diferentes registros de representação matemáticos (algébrico, geométrico, estatístico, computacional, etc.), na busca de solução e comunicação de resultados de problemas, de modo a favorecer a construção e o desenvolvimento de raciocínio matemático.</p>	<p>(EM13MAT408) Construir e interpretar tabelas e gráficos de frequências com base em dados obtidos em pesquisas por amostras estatísticas, incluindo ou não o uso de softwares que inter-relacionem estatística, geometria e álgebra.  (EM13MAT409) Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes diagramas e gráficos, como o histograma, o de caixa (box-plot), o de ramos e folhas, reconhecendo os mais eficientes para sua análise.</p>
<p>5) Investigar e estabelecer conjecturas a respeito de diferentes conceitos e propriedades matemáticas, empregando recursos e estratégias como observação de padrões, experimentações e tecnologias digitais, identificando a necessidade, ou não, de uma demonstração cada vez mais formal na validação das referidas conjecturas.</p>	<p>(EM13MAT510) Investigar conjuntos de dados relativos ao comportamento de duas variáveis numéricas, usando tecnologias da informação, e, se apropriado, levar em conta a variação e utilizar uma reta para descrever a relação observada.</p>

Fonte: a autora (2020). Adaptado de Brasil (2017).

Esta iniciativa teve como objetivo identificar pontos de convergência entre o currículo de Matemática nacional e a definição de Literacia de Dados que está sendo adotada nesta pesquisa, para que, a partir dessa identificação seja possível entender melhor o cenário da Literacia de Dados no nosso país e tomar partido das reais necessidades dos nossos estudantes, docentes e sociedade. Ainda, segundo Heinsfeld e Da Silva (2018), é necessário que seja feita a contextualização do seu conteúdo com as realidades diversificadas que se apresentam no cenário escolar nacional brasileiro. Os autores também realizaram um estudo

das três versões da BNCC analisando o papel das tecnologias digitais através da abordagem teórico-metodológica da Análise Crítica do Discurso de Fairclough (1989, 1995, 2008), para além da utilização da base como documento norteador.

Observando ainda os contextos dos cursos de graduação e de pós-graduação *strictu* e *latu sensu* nacionais, não foram encontrados, até então, pesquisas abordando eixos norteadores que remetam à aprendizagem em Literacia de Dados.

## 2.3 AVALIAÇÃO EM LITERACIA DE DADOS

Para a realização desta etapa da revisão de literatura, foi realizada pesquisa na base Google Scholar utilizando as *strings* de busca “*Data Literacy*” e “*Data Literacy*” “*Assessment*”. Foram selecionados os artigos “*Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-governmental Organizations*”; “*Ethical and appropriate data use requires data literacy*”; “*Graphical Perception of Value Distributions: An Evaluation of Non-Expert Viewers’ Data Literacy*” e “*Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project*”.

### 2.3.1 Fazendo o bem com dados: desenvolvimento de um modelo de maturidade para alfabetização em dados em organizações não governamentais

Maturidade, segundo o *Oxford Learner's Dictionaries* (Oxford Learner's Dictionaries, 2019), revela o estado de plenitude e desenvoltura que algo ou alguém apresenta ao realizar uma determinada tarefa em um determinado momento. Modelos de maturidade são capazes de descrever o comportamento predominante em uma organização, considerando os diferentes níveis de maturidade em diversos aspectos do campo que está sendo estudado (Sternkopf e Mueller, 2018), além de possuírem a vantagem de contar previamente com fatores relacionados a mudanças e serem capazes de regular e predizer essas mudanças até um certo ponto (Ross, 2003). Há três tipos principais de modelos de maturidade: descritivo, prescritivo e comparativo (Becker *et al.* 2009 e Maier *et al.* 2012).

Sternkopf e Mueller (2018) desenvolveram um modelo de maturidade descritivo através do método de pesquisa *Action Design Research* (ADR) para responder à pergunta de pesquisa: **Como descrever a Literacia de Dados através de um modelo de maturidade para avaliar a**

**capacidade de manipulação de dados de indivíduos no contexto de entidades não governamentais?** A motivação para que os autores desenvolvessem esse trabalho partiu da percepção de que segundo Chiesa *et al.* (2008), a medição de capacidades, apesar de ser um empreendimento complexo, pode auxiliar as organizações sobremaneira no alcance dos seus objetivos de tornar os esforços educacionais em Literacia de Dados mais escaláveis e mensuráveis.

Sternkopf e Mueller (2018) utilizaram a abordagem metodológica *Action Design Research* (ADR). A ADR, segundo Sein *et al.* (2011) refletem a premissa de que os artefatos de tecnologia da informação são modelados pelo contexto da organização durante o seu desenvolvimento e uso. Para sua estratégia de desenvolvimento da sua escala de maturidade, Sternkopf e Mueller (2018) tomaram como base diferentes aspectos de modelos de maturidade já existentes e levantaram, partindo da definição de Literacia de Dados de Slater (2016), 11 dimensões/áreas da Literacia de Dados que deveriam ser trabalhadas tanto a nível individual quanto organizacional, chegando às dimensões (i) Cultura sobre dados, (ii) Ética sobre dados e segurança, (iii) Competência para fazer perguntas e criar definições, (iv) Competência para procurar dados, (v) Competência para coletar dados, (vi) Competência para verificar dados, (vii) Competência para limpar dados, (viii) Competência para analisar dados, (ix) Competência para visualizar dados, (x) Competência para comunicar-se através de dados, (xi) Competência para acessar e interpretar dados. Também fizeram o levantamento dos níveis de maturidade a partir dos estudos de Maier *et al.* (2012), chegando a 4 níveis: incerteza, iluminação, certeza e fluência em dados. Depois criaram os textos que definem o nível de Literacia de Dados para cada uma das dimensões, também com base nas próprias formulações dos pesquisadores e da definição de Literacia de Dados de Slater (2016). O mecanismo de administração do artefato foi definido como auto avaliativo e foi feito através da ferramenta Excel mais uma grade de maturidade com os resultados apresentados em um diagrama de teia de aranha.

Sternkopf e Mueller (2018) avaliaram os resultados do seu modelo através da aplicação de um questionário auto avaliativo, da representação dos resultados nos gráficos e através do *feedback* de especialistas internos e externos à organização da *Datenschule*<sup>3</sup> (Escola de Dados Alemã), utilizando os critérios: correção, flexibilidade, usabilidade, implementabilidade e

---

<sup>3</sup> < <https://datenschule.de/en/>>. Acesso em 09/12/2019.

eficiência econômica, baseados em Pöppelbuß e Röglinger (2011). O modelo foi avaliado positivamente pelos especialistas e foi apontado que para obter resultados mais significativos, a aplicação do modelo deveria ser realizada em times e em diversas organizações e não somente com alguns indivíduos de duas organizações, como fizeram.

### 2.3.2 O uso ético e apropriado de dados requer Literacia de Dados

Segundo Mandinach *et al.* (2015), Educadores podem potencializar a aprendizagem de estudantes através da tomada de decisão orientada por dados, entretanto, para aproveitar os seus benefícios é necessário que a equipe pedagógica tenha uma maior Literacia de Dados, o que inclui saber o que fazer com os dados, além de utilizar os dados dos estudantes de forma ética e responsável. No contexto de estudo dos autores está a Lei de Privacidade e Direitos Educacionais da Família (*Family Educational Rights and Privacy Act*) - FERPA, sancionada em 1974, com base no artigo I parágrafo 8 da Constituição dos EUA. Essa lei tem como objetivo proteger a privacidade dos dados educacionais dos estudantes de escolas que recebem fundos do Departamento de Educação dos EUA. Entretanto, há um problema: a FERPA está expressa em um documento normativo de difícil entendimento tanto para a população quanto para os educadores que precisam implementá-la como política pública.

A partir dessa motivação os autores Mandinach *et al.* (2015) se fizeram a seguinte pergunta de pesquisa: **Como podemos ajudar educadores a entender o seu papel na proteção dos dados de estudantes e usá-los com ética e responsabilidade?** Para responder esta pergunta, os autores recorreram a um artefato em formato de construto, feito por Gummer e Mandinach (2015) que se chama “Literacia de Dados para a educação” e que pode ser definido como:

“A competência para transformar informação em práticas e conhecimento instrucional acionável através da coleta, análise e interpretação de todos os tipos de dados (avaliação, clima escolar, comportamento, instantâneos, longitudinais, momento-a-momento, etc.) para ajudar a determinar diferentes abordagens e passos instrucionais. Essa competência combina reconhecimento de padrões de dados, conhecimento de práticas disciplinares, conhecimento curricular, de conteúdo pedagógico e compreensão de como crianças aprendem” (Gummer e Mandinach 2015).

Este construto é sustentado por três domínios de conhecimento: 1- o uso de Literacia de Dados para ensinar, 2- o conhecimento do domínio em que o professor é especialista e 3- a competência para aplicar conhecimentos pedagógicos no contexto de sua área de especialidade (Figura 7). No âmbito desses três domínios de conhecimento, os autores formularam seis perguntas: i. como identificar problemas com o uso de dados? ii. Como formular perguntas? iii. Como usar os dados? iv. Como transformar os dados em informação? v. Como transformar informação em decisão e vi. como avaliar os resultados da decisão tomada? (Mandinach *et al.*, 2015).

Figura 7: Diagrama com os domínios de conhecimento em Literacia de Dados para o ensino, de acordo com o construto de Gummer e Mandinach (2015).



Fonte: a autora (2020). Adaptado de Gummer e Mandinach (2015).

A partir do artefato de Gummer e Mandinach (2015) criaram um novo artefato, que se materializou em uma tabela para a avaliação das práticas de uso de dados, que possui uma coluna dedicada às boas práticas de uso de dados em contextos educacionais. Entre elas tem-se exemplos de boas práticas como: utilizar dados para diferenciar instruções e grupos de alunos; usar murais de dados para entender os dados dos estudantes e para conversar com outros professores sobre problemas na prática pedagógica, colocando-os, quando os estudantes estiverem identificados, em um lugar acessível somente para educadores, remover os dados identificadores dos alunos quando os murais de dados forem expostos para

a comunidade escolar; discutir sobre os dados dos estudantes com a equipe de professores e gestores para traçar metas e identificar problemas na prática pedagógica. A outra coluna do artefato é dedicada para as práticas que devem ser evitadas, como por exemplo: falar sobre os fracassos e sucessos de um estudante em público e utilizar dados para rotular os estudantes como incapazes de serem bem-sucedidos. Mandinach *et al.* (2015) não apresentaram no seu artigo o modo como fizeram a avaliação do artefato.

Na Figura 8, há um exemplo de mural de dados realizado pelos estudantes de uma turma de primeiro período de graduação da Universidade Federal do Rio de Janeiro. Os dados expressos são acerca de preferências pessoais dos estudantes com respeito a características pessoais, preferências e hábitos. Para cada resposta dada a uma pergunta, os estudantes realizaram uma ação na moldura quadrada oferecida. Ou pintar de uma cor, ou assinalar um quadradinho, ou colar uma lantejoula. As perguntas realizadas foram quanto às suas atividades preferidas em momentos de lazer, quais os seus animais de estimação e as suas preferências de horário de estudos, entre outras. Cada um dos quadrados representa todas as informações dadas por apenas um estudante a um pequeno questionário estruturado. Note que os estudantes não possuem seus dados pessoais identificados no mural.

Figura 8: O dataset Cores, como exemplo de um mural de dados genérico, feito por estudantes do curso de Ciência da Computação da UFRJ em uma atividade em sala de aula. As cores representam dados relacionados a características pessoais, preferências e hábitos.



Fonte: reprodução da autora (2020).

O dataset da Figura 8 foi construído a partir de uma ideia original de Lupi e Posavec (2016), e foi utilizado para a realização das análises de dados pedidas aos estudantes durante os dois ciclos de estudo de caso relatados mais adiante nesta dissertação.

### **2.3.3 Percepção gráfica de distribuições de valor: uma avaliação da alfabetização em dados de visualizadores não especialistas**

Zubiaga e Namee (2016) descrevem o cenário da contemporaneidade, em que visualizações de dados têm sido cada vez mais exploradas para a finalidade de comunicar, e reforçar informações importantes, de forma ubíqua, em diversos tipos de mídia, principalmente encontradas nas redes sociais, realidade que vem ocorrendo paralelamente ao aumento do poder de processamento computacional e da disponibilização de dados na Web. Diante desse cenário resolveram fazer três investigações: a primeira era se usuários achavam histogramas gráficos mais fáceis para a interpretação de distribuição de dados de uma única variável; a segunda ampliava o escopo da primeira, investigando qual o gráfico comunica mais facilmente distribuições de duas variáveis; a terceira abordou aspectos da literacia de visualização de dados através da interpretação da sobreposição de histogramas tradicionais e de histogramas de rastro.

Zubiaga e Namee (2016) **identificaram como problema de pesquisa a ausência de clareza sobre o nível de Literacia de Dados demonstrada por pessoas leigas em estatística e em análise de dados quando se disponibiliza a elas diferentes tipos de visualização de dados**, então criaram um artefato formado por gráficos de dispersão de 5 tipos diferentes: barras, enxame de abelhas, *boxplots*, barras empilhadas e histogramas – criados na linguagem [R] – mais sentenças sobre os respectivos gráficos. As variáveis utilizadas na criação dos gráficos foram criadas artificialmente e se relacionavam a dados do cotidiano, como por exemplo, notas de estudantes em um exame. Os autores fundamentaram a criação desse artefato com base no construto teórico de que seria necessária uma seleção cuidadosa do tipo de gráfico escolhido quando se planeja oferecer uma informação para uma pessoa, a fim de assegurar que haja uma interpretação mais acurada quanto possível por parte dos usuários, que diferem muito entre si quanto ao nível de literacia em dados (Zubiaga e Namee, 2016).

Para realizar a avaliação da validade do artefato e do construto teórico, Zubiaga e Namee (2016) aplicaram um experimento em que os participantes avaliaram, utilizando 5

pontos na escala *Likert* (concordo fortemente, concordo, sou neutro, discordo, e discordo fortemente), a precisão da correspondência entre uma determinada visualização de dados e uma sentença relacionada à visualização mostrada. Foram coletadas ao todo 50 avaliações das 160 tarefas geradas através da combinação aleatória de 5 tipos de gráficos com 4 diferentes tipos de variáveis e 8 sentenças – umas corretas e outras erradas, mas todas aplicáveis a todos os gráficos – gerando um total de 8.000 avaliações realizadas. As avaliações foram feitas, segundo conselho de Heer e Bostock (2010), por usuários da plataforma *MechanicalTurk*<sup>4</sup> sem nenhuma restrição ao nível de habilidades em análise de dados dos participantes, que deveriam possuir pelo menos 100 tarefas realizadas na plataforma, além de 95% de *HIT acceptance rate*, uma medida interna de confiabilidade das avaliações realizadas pelos usuários no site.

A análise da avaliação foi realizada através de triangulação entre o nível de concordância entre as respostas dos participantes, da precisão das respostas dos participantes com relação à verdade básica (*groundtruth*) e de uma matriz de confusão para a compreensão dos tipos de erros cometidos pelos participantes. O principal resultado desse estudo foi que os usuários acharam que, entre os tipos de gráficos apresentados, o histograma é o mais fácil de ser compreendido para o estudo de dispersões em uma variável. A partir desse resultado os autores empreenderam um segundo estudo de caso, onde **investigaram qual tipo de gráfico se mostra mais eficaz para comparar a distribuição de duas variáveis diferentes** (Zubiaga e Namee, 2016).

Nessa segunda rodada de experimentos Zubiaga e Namee (2016), utilizaram a plataforma *Mechanical Turk* da Amazon para fazer um experimento, mostrando um histograma padrão e um histograma de rastro de densidade para cada participante, fazendo em seguida uma série de perguntas a fim de averiguar a acurácia com que eram capazes de interpretar esses gráficos. Desta vez também utilizaram gráficos feitos em [R], mas utilizando os dados do *Vietnamese Living Standard Study* (VLSS), que contém o valor de gastos familiares *per capita* de 5.999 famílias vietnamitas de áreas rurais e urbanas. Dessa forma criaram as duas distribuições para serem comparadas pelos usuários (Zubiaga e Namee, 2016).

---

<sup>4</sup> A Amazon MechanicalTurk <mturk.com>. Acesso em 02 dez de 2019. É uma plataforma *crowdsourcing* que oferece serviços que necessitam da inteligência humana, de forma terceirizada para empresas e indivíduos.

Como resultado principal dessa segunda etapa da pesquisa, os pesquisadores descobriram que os usuários foram em média mais precisos nas suas respostas quando fizeram interpretações a partir de histogramas tradicionais com relação às interpretações feitas com rastros de densidade, sendo que as taxas de erro para histogramas tradicionais foram levemente menores, o que sugere que histogramas são gráficos mais aderentes para que os usuários pesquisados façam interpretações acuradas. Usuários em média também gastaram 3,4% mais tempo respondendo a questões associadas a histogramas tradicionais, o que foi refletido em 6,4% de melhora em termos de precisão, segundo Zubiaga e Namee (2016).

Então, em uma terceira etapa de pesquisa, Zubiaga e Namee (2016) propuseram a pergunta de pesquisa: **usuários têm um bom desempenho interpretando dados de um histograma tradicional, mas será que dois histogramas tradicionais ou dois histogramas de rastro colocados juntos em um único gráfico otimizam a percepção?** Nesta terceira etapa também utilizaram um experimento para responder à pergunta de pesquisa, perguntando sobre 6 diferentes tipos de gráficos (*overlappedhistogram*, *mirrowedhistogram*, *interpersedhistogram*, *stackedhistogram*, *cumulativehistogram*, *overlappeddensity traces* e *mirroweddensity traces*). Como já esperado pelos autores, os usuários não foram capazes de interpretar histogramas cumulativos de forma precisa, alcançando uma precisão de apenas 20% na maioria dos casos, enquanto que alcançaram em média mais de 70% de precisão na interpretação de outros tipos de gráficos. Os usuários em geral alcançaram precisão de 85% ao interpretar histogramas intercalados (Zubiaga e Namee, 2016). Um resultado ainda interessante é que em mais de uma etapa da pesquisa de Zubiaga e Namee (2016), os usuários tiveram dificuldade de relatar frequências que não tinham *tick mark* no eixo x.

#### **2.3.4 Incorporando a alfabetização em dados em programas de alfabetização informacional de graduação em ciências sociais: um projeto piloto**

Por volta dos anos de 2004 e 2005, Stephenson e Caravello (2007) relataram dificuldade na avaliação e uso da informação, devido a lacunas de conhecimento em Literacia de Dados. Para resolver o problema, os autores então, iniciaram uma pesquisa para a criação de cursos em literacia da informação, que contivessem um módulo de literacia em dados. Primeiro definiram Literacia de Dados como sinônimo de literacia estatística, literacia

quantitativa e numeracia – e, para eles Literacia de Dados teria a ver com o entendimento de tabelas, gráficos, dados e conceitos e terminologias próprias da estatística – e viram a necessidade de integrar a Literacia de Dados como um componente curricular relacionado ao estudo da literacia da informação, integrando o currículo de sociologia do Departamento de Sociologia da *University of California Los Angeles* (UCLA) (Stephenson e Caravello, 2007).

O objetivo da pesquisa empreendida por Stephenson e Caravello (2007) foi analisar a confluência entre os currículos de literacia da informação e Literacia de Dados. A proposta de solução foi desenhar um curso piloto em Literacia de Dados para estudantes do ciclo avançado em ciências sociais da UCLA, avaliando o seu projeto através da metodologia de estudo de caso. Ao final do curso, composto por 2 módulos dedicados a Literacia de Dados, era esperado que os estudantes fossem capazes de: i. encontrar informação que precisam para suas investigações sociológicas, ii. Avaliar a confiabilidade das informações encontradas e iii. Utilizar as informações adquiridas de forma ética.

No primeiro módulo era esperado que os estudantes i. desenvolvessem a habilidade de ler e avaliar criticamente tabelas simples, relacionando duas variáveis ou três variáveis, ii. produzir citações precisas para tabelas de dados, iii. utilizar o *American Factfinder*<sup>5</sup> para criar uma tabela e citá-la corretamente, iv. ler um artigo contendo dados representados de forma gráfica e debater sobre o gráfico com relação ao conteúdo do artigo. O segundo módulo foi dedicado a auxiliar os estudantes a diferenciar fontes acadêmicas de fontes não acadêmicas (Stephenson e Caravello, 2007). Os estudantes foram avaliados nos dois módulos durante as sessões de aprendizagem, através de deveres de casa e formulários de aprendizagem e não foram utilizados pré-testes e pós-testes e nem avaliações padronizadas, devido a quantidade de estudantes participantes do programa piloto: quatro estudantes no primeiro módulo e 7 no segundo módulo (Stephenson e Caravello, 2007).

Segundo Stephenson e Caravello (2007), o resultado foi que a visão crítica dos estudantes foi ampliada e os estudantes obtiveram bons resultados na identificação de componentes de uma tabela, através da introdução de conceitos básicos de Literacia de Dados. Para o futuro, os pesquisadores imaginaram que seria mais interessante ao invés de

---

<sup>5</sup> <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml>. Acesso em 09/12/2019. O *American Factfinder* tem funcionado desde julho de 2019 como um sistema de “arquivo”, sendo substituído pelo <data.census.gov> em definitivo a partir da primavera de 2020.

criar um curso separado das disciplinas próprias da Sociologia, que o ensino da Literacia de Dados deveria estar situado no contexto das disciplinas do curso.

### **2.3.5 Exploração dos pontos positivos e negativos da literatura revisada**

Neste final de capítulo realizou-se uma breve análise da literatura revisada quanto a contemplação dos aspectos “leitura de dados”, “trabalho com dados”, “análise de dados” e “argumentação com dados” com o objetivo de explorar os pontos fortes e fragilidades que possam auxiliar na proposição da solução para a pergunta desta pesquisa.

#### **2.3.5.1 Leitura de dados**

Para D’Ignazio e Bhargava (2015), leitura de dados significa a compreensão de quais dados serão necessários para a compreensão de um universo e quais aspectos desse universo eles representam. Sternkopf e Mueller (2018) citam como sexta dimensão do seu modelo de maturidade em Literacia de Dados a dimensão “verificar dados”, mas não a detalham suficientemente a ponto de poder-se dizer que “verificar dados” tem o mesmo significado de “ler dados”. Ora, parece lógico que para que um sujeito seja capaz de “verificar dados” é necessário que saiba “ler dados”, entretanto, entende-se neste trabalho que, por ser a “leitura de dados” uma ação mais elementar que a “verificação de dados”, conclui-se com base nas poucas informações dadas por Sternkopf e Mueller (2018) que a “verificação de dados” se insere no conceito de “leitura de dados” de D’Ignazio e Bhargava (2015), mas que para que se “verifique dados” é necessário que haja maturidade na “leitura de dados”, sem a qual o sujeito não conseguirá alcançar a etapa de “verificação de dados”.

Mandinach *et al.* (2015), não mencionam a “leitura de dados” como parte de um processo de compreensão de quais dados são importantes para que se entenda o universo dos estudantes (D’Ignazio e Bhargava, 2015), indicando no seu manual de boas práticas diretamente a fase de utilização dos dados para a organização de murais de dados para melhor visualizar os dados dos estudantes, o que está mais próximo da definição de D’Ignazio e Bhargava (2015) no que diz respeito à utilização dos dados em “trabalho com dados” e “análise de dados”.

Zubiaga e Namee (2016), ao escolherem um *dataset* e criarem visualizações de distribuições para representar aspectos desse *dataset*, dão foco ao entendimento de quais aspectos de um universo estudado os dados representam (D'Ignazio e Bhargava, 2015), todavia suprimem, alinhados com o alcance dos seus objetivos de pesquisa, a parte da “leitura de dados” que se relaciona à compreensão de quais dados são necessários para a representação de um universo (D'Ignazio e Bhargava, 2015).

Apesar do pouco detalhamento sobre quais critérios utilizaram para avaliar o alcance dos objetivos de aprendizagem pelos estudantes em seu estudo de caso, Stephenson e Caravello (2007) delimitaram como o primeiro objetivo da sua primeira sessão do curso de Literacia de Dados “ler e avaliar criticamente tabelas simples” e como quarto objetivo “ler um artigo contendo dados representados de forma gráfica [...]”. Entende-se que esse objetivo cumpre os requisitos de “leitura de dados” elencados por D'Ignazio e Bhargava (2015) exceto no que diz respeito à compreensão de “que dados são necessários para a representação de um universo” (D'Ignazio e Bhargava, 2015), que estaria mais relacionado à leitura de dados voltada para a aquisição de dados.

### 2.3.5.2 Trabalho com dados

Sobre o trabalho com dados, Sternkopf e Mueller (2018), em seu modelo de maturidade, elencaram como dimensões do seu modelo de maturidade: competência para procurar dados, competência para coletar dados, competência para limpar dados, competência para acessar dados e ainda trabalhar com os dados de forma ética – com confidencialidade, integridade e disponibilidade de dados e políticas de uso de dados atualizadas constantemente (Sternkopf e Mueller, 2018).

Já sobre Mandinach *et al.* (2015) pode-se dizer que o fato de utilizarem murais de dados para organizar os dados dos estudantes mostra uma importante preocupação com o gerenciamento desses dados, mas os autores ainda vão além, preocupando-se, tal como Sternkopf e Mueller (2018), com a privacidade dos estudantes, apontando como prática a ser evitada o uso de aplicações na nuvem para a guarda e manipulação de dados estudantis e indicando a importância de manter o sigilo da identidade nos murais de dados, a não ser que eles estejam em ambiente de acesso restrito ao corpo docente (Mandinach *et al.* 2015).

Zubiaga e Namee (2016) não abordaram a temática do trabalho com dados, se revisados conforme a definição de literacia de D'Ignazio e Bhargava (2015), uma vez que o foco da sua pesquisa estava relacionado a aderência de diferentes visualizações de gráficos de dispersão com tipos de usuários diversificados. Stephenson e Caravello (2007) incluíram em seu programa de cursos em Literacia de Dados da UCLA o objetivo de aprendizagem “diferenciar fontes acadêmicas de fontes não acadêmicas” (Zubiaga e Namee, 2016). Sendo a diferenciação de fontes uma estratégia utilizada durante a aquisição de dados, interpretou-se nesta pesquisa que Zubiaga e Namee (2016) abordaram com os seus estudantes de sociologia somente esta parte do trabalho com dados, de acordo com suas metas definidas para o curso.

### 2.3.5.3 Análise de dados

Sternkopf e Mueller (2018) formalizam a “análise de dados” como uma das dimensões do seu modelo de maturidade. Mandinach *et al.* (2015) tratam a análise de dados apontando como boas práticas em Literacia de Dados para o ensino: diferenciar tipos de instruções e grupos de estudantes, traçar metas coletivamente e conversar com a equipe sobre possíveis problemas relacionados à prática pedagógica, tudo isto com base em murais de dados dos estudantes. Já Zubiaga e Namee (2016) voltam sua pesquisa inteiramente para a análise de dados, a fim de medir o nível de Literacia de Dados de pessoas leigas utilizando visualizações de dados de dispersão. Stephenson e Caravello (2007) criaram seu curso de literacia em dados orientado por objetivos de aprendizagem que, em seus dois módulos apresentam objetivos relacionados a análise de dados. No primeiro curso apresentam o quarto objetivo de aprendizagem como sendo “ler um artigo contendo dados representados de forma gráfica e debater sobre o gráfico com relação ao conteúdo do artigo” (Stephenson e Caravello, 2007). Já no segundo curso dedicam inteiramente à diferenciação de fontes acadêmicas de fontes não acadêmicas (Stephenson e Caravello, 2007), o que se considerou nesta pesquisa como análise de dados, de acordo com a definição de análise de dados de D'Ignazio e Bhargava (2015).

#### 2.3.5.4 Argumentação com dados

Relembrando que para D'Ignazio e Bhargava (2015), argumentar com dados significa “a utilização de dados como suporte para narrativas com o propósito de comunicar alguma informação para um determinado público”. Sternkopf e Mueller (2018) apontam “comunicar-se através de dados” como penúltima dimensão do seu modelo de maturidade em Literacia de Dados, que em seu pleno domínio deverá apresentar-se como competência para sintetizar e comunicar de maneira adequada a natureza dos dados e a sua finalidade para um determinado público-alvo” (Sternkopf e Mueller, 2018). Apesar de definirem de forma detalhada uma grade de maturidade para a comunicação através de dados, os autores não explicitaram os resultados de performance – a comunicação através de dados, propriamente dita – do seu público avaliado, deixando uma lacuna no entendimento de como a competência da comunicação com dados é formalizada por sujeitos de um processo de aprendizagem.

Mandinach *et al.* (2015) também não formalizam como sujeitos de um processo de aprendizagem comunicam-se através de dados, outrossim, na sua lista de boas práticas e práticas a serem evitadas na Literacia de Dados para o ensino, apontam que os educadores devem usar dados para conversarem uns com os outros sobre problemas na prática pedagógica, para traçarem metas instrucionais e para comunicarem melhor as características dos estudantes com os seus familiares, distinguindo como uma prática a ser evitada abordar assuntos relativos aos sucessos e desafios dos estudantes na frente de outras pessoas (Mandinach *et al.*, 2015). Zubiaga e Namee (2016) não abordaram a comunicação através de dados em seu estudo e Stephenson e Caravello (2017) abordaram o tema através de um objetivo de aprendizado que unia a leitura de um artigo contendo dados representados de forma gráfica com um debate – em sequência – sobre o gráfico com relação ao conteúdo do artigo.

#### 2.3.6 Diferenças nos *designs* para a percepção e/ou desenvolvimento de competências em Literacia de Dados

Stephenson e Caravello (2007), talvez por estarem tratando de um curso com objetivos de aprendizagem bem definidos, apresentam a análise de dados como parte de um processo de aprendizagem que se inicia com a busca de informação para as pesquisas de seus alunos de sociologia, passa pela avaliação da confiabilidade das informações encontradas e se

encerra na utilização das informações adquiridas de forma ética. O modelo de maturidade de Sternkopf e Mueller (2018) apresenta uma grade de maturidade em níveis. Mandinach *et al.* (2015) e Zubiaga e Namee (2016) apresentam um *design* de pesquisa que não se propõe a avaliar processos de aprendizagem multietapas em Literacia de Dados.

### **2.3.7 Considerações sobre os processos de avaliação das propostas de pesquisa**

Sobre a avaliação do modelo de Sternkopf e Mueller (2018), pode-se dizer que indica uma fragilidade porque foi realizada com poucos indivíduos e também porque todos os indivíduos que participaram eram especialistas da *Datenschule* (2 profissionais de dados) ou de uma organização parceira que não foi identificada, mas que também estava envolvida com treinamento em dados. O mesmo ocorreu com a pesquisa de Stephenson e Caravello (2007), que teve em seu programa piloto, cinco estudantes no primeiro módulo e sete estudantes no segundo módulo do curso. Mandinach *et al.* (2015) não mostraram em sua publicação a forma como foi feita a avaliação do seu artefato e Zubiaga e Namee (2016) geraram 8.000 avaliações de visualizações utilizando a *Amazon Mechanical Turk*, processo que apesar de ser generoso em quantidade de avaliações e precisão de métricas, não revela detalhes do processo de aprendizagem dos sujeitos de pesquisa.

Com a finalidade de explorar os fatores positivos e negativos apresentados pela literatura explorada até então, considerando os pontos fortes e fracos dos contextos nacional e local que podem ser aproveitados a fim de gerar conhecimento no tema deste estudo, será apresentada uma pesquisa em que a Literacia de Dados será abordada como um processo derivado da definição de D'Ignazio e Bhargava (2015). Será realizada então, uma pesquisa qualitativa de inclinação interpretacionista com estudantes de diversos níveis de escolaridade, em quantidade razoável para que seja aplicada a triangulação de dados compostos por: atividade em sala de aula e anotações via observação direta. Para que se possa formalizar os conhecimentos em Literacia de Dados dos estudantes pesquisados, pretende-se criar um artefato sociotécnico, utilizando a abordagem epistemológico-metodológica da DSR, que guiará a aprendizagem em dados através de um processo semelhante ao processo de extração de conhecimento em bases de dados. No capítulo 3 será discutida em detalhes a proposta de solução.

## 2.4 COLABORAÇÃO E GROUP STORYTELLING

Aprender com os estudantes é fundamental para que o fenômeno da educação ocorra Freire (1970, 1996). E, para isso é necessário que haja formas de acessar os conhecimentos tácitos dos estudantes a fim de torná-los explícitos. A palavra “tácito”, segundo o dicionário Michaelis<sup>6</sup>, significa “em que há silêncio” ou algo “não expresso por palavras”. Segundo Gargano *et al.* (2019), “o conhecimento tácito não é formal, ele é não mensurável, é subjetivo. É o conhecimento acumulado ao longo da vida, sendo geralmente difícil de ser repassado a outra pessoa”. Existem algumas formas de acessar os conhecimentos tácitos dos estudantes e entre elas, pesquisou-se a técnica de *group storytelling*, que é uma técnica direcionada para a recuperação informações e conhecimentos tácitos em ambientes corporativos (Lage *et al.*, 2017).

De acordo com Lage *et al.* (2017), a *storytelling* ou contação de histórias oportuniza a descoberta do conhecimento tanto para quem procura, quanto para quem deseja construí-lo e, quando realizada coletivamente, a contação de histórias é capaz de se beneficiar de nuances de percepção diversas, dando vida à uma obra de maior complexidade. Segundo o autor, tanto na tarefa de transferência do conhecimento quanto na verificação da aprendizagem, a contação colaborativa de histórias se mostra um instrumento precioso e eficiente para educadores, a fim de obter o conhecimento de como foi a apreensão dos conteúdos de uma disciplina, sobre como foram conduzidas as aulas e também sobre a percepção dos próprios estudantes acerca do método colaborativo de ensino (Lage *et al.*, 2017).

Para Borges *et al.* (2007), o trabalho colaborativo utilizando um instrumento mediador é capaz de influenciar nas interações do grupo, uma vez que os indivíduos tornam explícitas as suas estratégias para a resolução de tarefas. Nesta pesquisa utilizaremos a técnica de *group storytelling* como forma de aprender com os estudantes de que forma podemos medir a sua Literacia de Dados.

Segundo o *Online Etymology Dictionary*<sup>7</sup>, a origem etimológica da palavra colaboração remonta ao século 19, relacionando-se com a palavra *collaborare*, do latim, que significa “trabalhar junto”. A colaboração ocorre através da formação de grupos de trabalho.

---

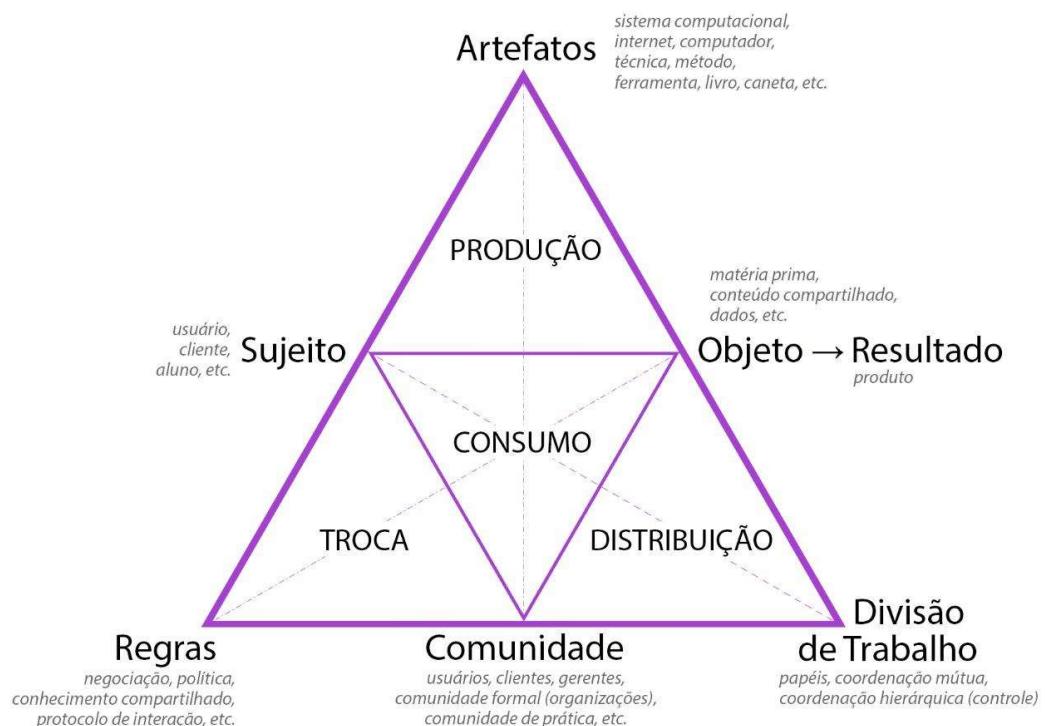
<sup>6</sup> <http://michaelis.uol.com.br/busca?id=3wdk2>. Acesso em 13/09/2020.

<sup>7</sup> <https://www.etymonline.com/word/collaboration>. Acesso em 11/10/2020.

Um grupo ou equipe é um organismo que possui padrões de comportamento próprios e que evoluí, se desenvolvendo ao longo do tempo e se beneficiando da combinação das habilidades de seus integrantes na realização de tarefas ou atividades (Vivacqua *et al.*, 2011). Uma atividade é “uma unidade mínima para compreender as ações de um sujeito”, considerando o sujeito como sendo uma pessoa ou um grupo que realiza ações sobre um objeto a fim de atingir um objetivo (Fuks *et al.*, 2011).

Para realizar atividades de forma colaborativa, os grupos se utilizam de ferramentas que se encontram disponíveis com o objetivo de produzir seu trabalho (Ugulino *et al.*, 2008). De acordo com a Teoria da Atividade, as relações de mediação entre pessoas ou grupos permite a evolução cultural, porque transforma a relação da humanidade com o meio de ecológica — voltada para a adaptação ao meio — para econômica e histórica - voltada para o consumo e caracterizada por produzir algo (cooperação), utilizando artefatos e através de trocas dentro de uma comunidade (comunicação) a partir de uma tarefa que foi distribuída por alguém (coordenação) (Fuks *et al.*, 2011).

Figura 9: Modelo de Atividade (adaptado de Engeström, 1987).

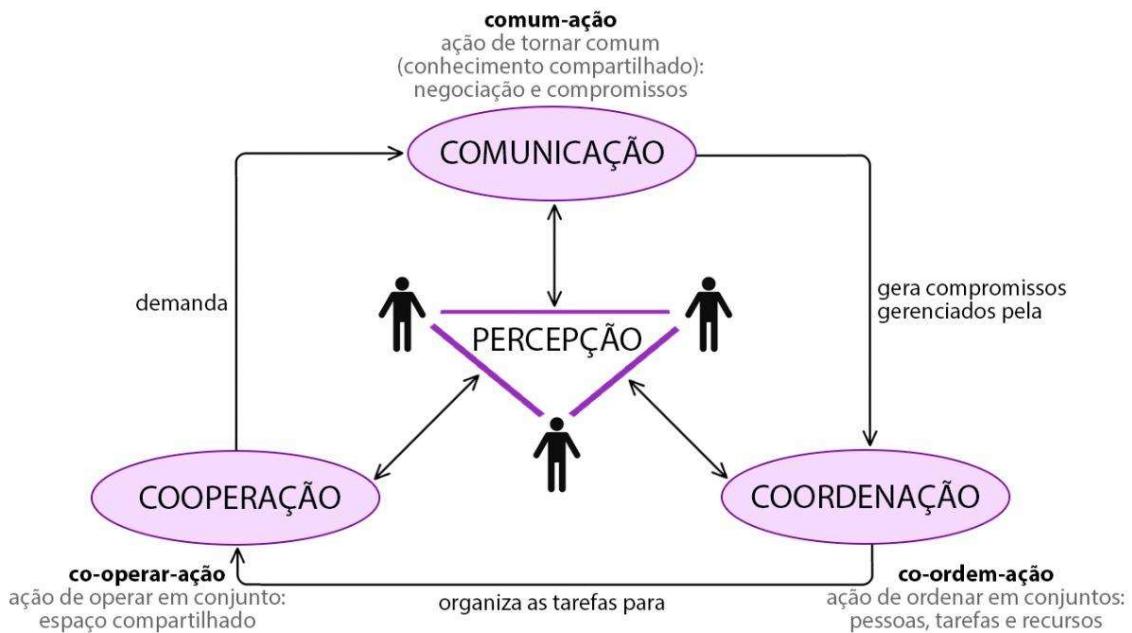


Fonte: a autora. Adaptado de Fuks *et al.* (2011).

Fuks *et al.* (2011) consideram que, entre os Modelos de Colaboração de Tuckman, Padrões de Colaboração e Modelo 3C, este último se destaca por ser mais representativo na literatura e porque relaciona as dimensões Comunicação, Coordenação e Cooperação a partir do entendimento de que a cooperação demanda comunicação que, por sua vez, gera compromissos gerenciados pela coordenação, que organiza as tarefas para a cooperação.

De acordo com Pimentel *et al.* (2006), o Modelo 3C, foi inicialmente proposto por Ellis *et al.* (1991), que levantam a importância de se levar em conta os processos de comunicação, colaboração e coordenação como pilares da relação de colaboração que devem ser levados em conta ao se pensar em *groupware* (Ellis *et al.*, 1991). Pimentel *et al.* (2006) propõem que a colaboração é composta por esses 3 elementos: comunicação, cooperação e coordenação (Figura 10), sendo a troca de mensagens caracterizada pela comunicação, a realização de tarefas em ambiente compartilhado caracterizada pela cooperação e a coordenação caracterizada pelo gerenciamento de pessoas.

Figura 10: Modelo 3C de Colaboração.



Fonte: a autora. Adaptado de Fuks *et al.* (2011).

As expressões “*groupware*” e “*Computer Supported Cooperative work*” na língua inglesa correspondem a expressão “Sistemas Colaborativos” na língua portuguesa, que são sistemas feitos para apoiar a colaboração (Fuks *et al.*, 2011). Segundo Aparicio *et al.* (2012) os

sistemas colaborativos representam ferramentas tecnológicas de apoio ao trabalho de grupo, especializadas o bastante para permitir que haja comunicação, coordenação, controle e colaboração entre os seus usuários, contribuindo assim para o aumento da produtividade. Aparicio *et al.* (2012) ainda levantam 3 critérios que agrupam os sistemas colaborativos: (i) interação e comunicação, (ii) se o espaço de interação é presencial ou à distância e (iii) se a interação é realizada de forma síncrona ou assíncrona. A interação e comunicação são requisitos fundamentais para um sistema colaborativo, enquanto que a relação de tempo na interação pode ser síncrona — como ocorre em bate-papos e videoconferências — ou assíncrona — como nos fóruns e e-mails — e a configuração geográfica da interação pode ser a distância ou presencial.

## 2.5 *Design Science Research*

A *Design Science Research* (DSR) é uma abordagem epistemológico-metodológica utilizada em contextos em que pesquisadores de Informática na Educação lançam mão de um artefato para gerar conhecimento através da resolução de um problema de pesquisa. A DSR, Ciência do Projeto ou Ciência do Artificial é ao mesmo tempo ciência e paradigma epistemológico, porque se debruça sobre a necessidade de produção de conhecimento sobre os artefatos sociotécnicos utilizados nas pesquisas (Pimentel *et al.*, 2019).

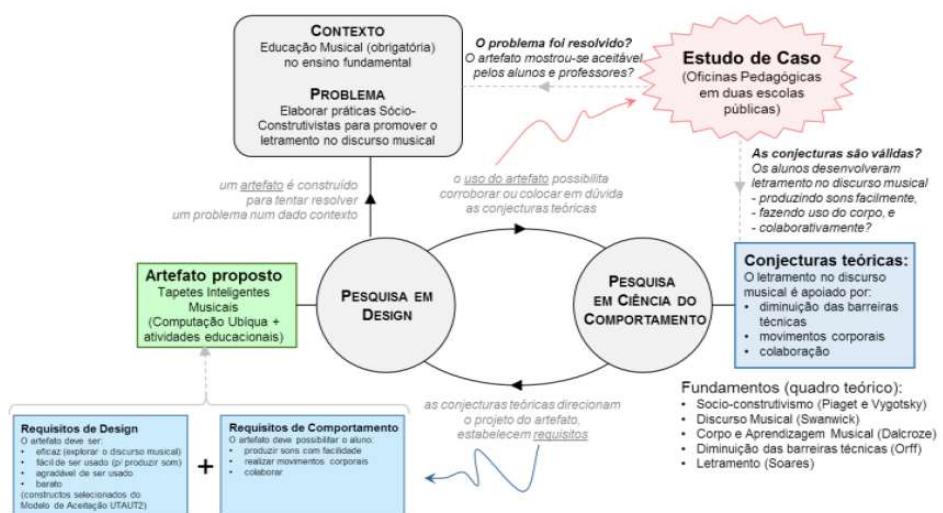
Segundo Pimentel *et al.* (2019) na DSR o pesquisador se compromete com dois objetivos: resolver um problema prático num contexto específico por meio de um artefato e gerar novo conhecimento científico. Escolheu-se a abordagem epistemológico-metodológica da DSR para esta pesquisa, porque havia um problema prático para ser resolvido: a necessidade de avaliar o nível de Literacia de Dados de estudantes. Acreditou-se inicialmente que este problema poderia ser resolvido de forma satisfatória através do *design* de um artefato, tal qual mostramos brevemente no início deste capítulo.

Há três ciclos importantes na DSR, que alimentam uns aos outros durante a pesquisa: o ciclo de engenharia, o ciclo de conhecimento/rigor/empírico e o ciclo de relevância. Esses ciclos conduzem às três respectivas avaliações: (i) se o artefato satisfaz aos requisitos (Expectativa de Performance, Expectativa de esforço, Motivação Hedônica, Intenção de Uso, Condições Facilitadoras, Custo); (ii) se as conjecturas teóricas parecem válidas; (iii) se o problema de pesquisa foi resolvido de forma satisfatória (Pimentel *et al.*, 2019).

Pimentel *et al.* (2019) criaram um mapa dos elementos esperados de uma pesquisa na abordagem DSR. Esse mapa sintetiza os principais elementos de uma pesquisa em DSR propondo um formato de relacionamento entre esses elementos para a organização de uma pesquisa.

O mapa de elementos da Figura 11 auxilia na formalização de elementos epistemológico-metodológicos centrais de uma pesquisa guiada pela DSR: (i) o contexto em que se insere o problema de pesquisa, (ii) o problema de pesquisa, (iii) o artefato que se pretende criar ou lançar mão para a resolução do problema de pesquisa, (iv) o estado da técnica relacionada ao desenvolvimento do artefato, (v) as conjecturas teóricas sobre o comportamento humano-social-organizacional que precisam ser levadas em conta durante o projeto e melhorias do artefato, (vi) o quadro teórico que é relevante para que se formalize as conjecturas teóricas sobre o comportamento humano-social-organizacional na hora de projetar o artefato. (vii) a necessidade da relação de retroalimentação contínua entre o *design* do artefato e o levantamento das conjecturas teóricas sobre o comportamento humano-social-organizacional e (viii) as 3 principais avaliações que precisam ser realizadas para a validação de uma solução baseada em DSR, que, em formato de perguntas são: a) o artefato é válido? b) o problema de pesquisa foi resolvido? E c) as conjecturas teóricas parecem válidas? (Pimentel *et al.*, 2019).

Figura 11: Mapa de elementos esperados de uma pesquisa na abordagem da DSR instanciado para a pesquisa de Santos *et al.* (2016).



Fonte: Santos *et al.* (2016). Adaptado de Pimentel *et al.* (2019).

Pimentel *et al.* (2019), sugerem ainda que seja feita uma revisão de literatura levando-se em conta 3 diferentes eixos: (i) elaboração de um quadro teórico que fundamente as conjecturas teóricas que darão suporte à pesquisa, como foi feita na seção 2.4 e nesta seção 2.5, (ii) caracterização do estado da técnica através do levantamento de outros artefatos, patentes e de princípios de *design* de artefatos, como será mostrado na seção 2.5.1 e (iii) caracterização do problema no contexto em que se encontra, como foi feito na seção 2.2.

### **2.5.1 Trabalhos relacionados à criação de artefatos utilizando a DSR**

Um artefato, segundo Peffers *et al.* (2007) é, “em princípio qualquer coisa projetada para alcançar um objetivo”. Segundo Simon (1969), um artefato pode ser definido pelo objetivo da sua existência, pelas suas características e pelo ambiente em que executa o seu objetivo. Com relação a definição de artefato sociotécnico, podemos então dizer, de acordo com Vermaas *et al.* (2011) que artefatos sociotécnicos são artefatos técnicos – os produtos tangíveis da tecnologia – considerados em seus contextos sociais de uso. Nesse sentido um artefato sociotécnico é diferente de um artefato de tecnologia da informação, porque ele traz em sua definição o ambiente em que é executado. Sistemas sociotécnicos são aqueles em que um amplo número de artefatos técnicos são parte de sistemas mais amplos nos quais o seu correto funcionamento depende não somente de fatores técnicos, mas também de fatores sociais (Vermaas *et al.*, 2011).

Dessa forma, selecionou-se para estudo a monografia Novas Interfaces Tecnológicas: Discutindo Usos Didáticos de Uma Roupa Computacional Para Aulas de Música, de Santos (2013), o artigo Tapetes Musicais Inteligentes: Computação Ubíqua para apoiar a Educação Musical, de Santos *et al.* (2016) e o artigo Desenvolvimento de um Modelo da Participação em Bate-papo seguindo a abordagem de *Design Science Research*, de Rocha *et al.* (2014), por entender que fazem parte de uma constelação de estudos voltados para a geração de conhecimento através da construção de artefatos, utilizando a abordagem da DSR para objetivos diversos e em temáticas diferentes, mas que, por se constituírem como partes de conhecimentos que vêm sendo construídos por uma mesma comunidade de pesquisa, a de Sistemas Colaborativos, possuem uma sinergia importante para a composição de referências para este trabalho.

Para realizar a revisão de literatura levantou-se a seguinte sequência de perguntas guia para o estudo das três publicações selecionadas: Qual é o problema/pergunta de pesquisa que se pretende resolver? Qual é o objetivo principal da pesquisa? Quais são os objetivos secundários da pesquisa? Qual foi o público-alvo da pesquisa? Qual ou quais foram os artefatos criados a fim de apoiar a resolução do problema de pesquisa? Quais foram as conjecturas teóricas levantadas para fundamentar a criação do artefato? Como foi utilizada a abordagem da DSR? Quantas interações foram realizadas com o uso do artefato? Quais foram os métodos de pesquisa utilizados para a avaliação do artefato e das conjecturas teóricas? Conseguiram resolver o problema? Como foi feita a análise dos resultados da pesquisa com o uso do artefato? Quais foram as descobertas e insights proporcionados pela pesquisa?

#### 2.5.1.1 Novas interfaces tecnológicas: discutindo usos didáticos de uma roupa computacional para aulas de música

Santos (2013) projetou uma roupa computacional construída com microprocessadores e componentes eletrônicos, costurada com fio e tecido condutor, para ser utilizada em situações de ensino-aprendizagem voltadas para a exploração do discurso musical e sonoro. O problema de pesquisa que deu origem à investigação através do uso do artefato Camisa-piano foi a constatação de que havia poucas propostas pedagógicas direcionadas para uma possível mudança de paradigma na educação musical, isto é, que apresentassem novas relações com a música para este mundo rápido em inovações. Santos (2013) delineou como objetivo da sua pesquisa investigar ferramentas teóricas e concretas para servirem de suporte à educação musical no contexto do ensino fundamental, tendo como perspectiva a abordagem sociointeracionista, em que o computador é utilizado em contextos educacionais como um meio para a comunicação e a interação. As suposições que deram suporte à criação do artefato diziam que é possível melhorar o aprendizado de música se: (i) os estudantes estiverem no lugar de comunicadores da informação, ao invés de meros receptores; (ii) os estudantes puderem criar conteúdo, estabelecendo uma relação autoral e (iii) se as atividades para a aprendizagem de música forem reconfiguradas para estarem mais aderentes ao ciberespaço (Santos, 2013).

Santos (2013) questiona se a aprendizagem do discurso musical no ensino básico não é em si mais importante do que o ensino de um instrumento em si, apontando a importância

da busca por um conceito de instrumento que seja mais facilitador do que limitador do aprendizado de músicas nas escolas.

#### 2.5.1.2 Tapetes musicais inteligentes: computação ubíqua para apoiar a educação musical

Em Santos *et al.* (2016), o problema de pesquisa que se apresenta é a necessidade da diminuição das barreiras técnicas necessárias para a execução musical. Com a expressão “execução musical”, os autores se referem à execução de ações referentes ao conjunto composto por ritmo, melodia, harmonia e orquestra por humanos com o auxílio de instrumentos musicais tradicionais ou outros artefatos tecnológicos. O objetivo de sua pesquisa foi promover a aprendizagem e a vivência de conceitos musicais fundamentados em Vygotsky (1978), Piaget (2013), Frazee e Kreuter (1987), Jaques-Dalcroze (1921) e Swanwick (2007).

Para isto construíram um artefato sociotécnico que se materializou na forma de Tapetes Musicais Inteligentes construídos a partir de uma combinação de baixo custo entre *software* e *hardware*: um computador ligado a caixas de som e uma placa *MakeyMakey*, fios de ethernet, papel alumínio, fita isolante, tapetes EVA, fita crepe e o *software Soundplant* (Santos *et al.*, 2016). O artefato foi construído para auxiliar a aprendizagem de música por estudantes de 2 turmas de 7º ano de escolas públicas no município do Rio de Janeiro onde o professor-pesquisador atuava. Os Tapetes Musicais emitem sons diversos através do contato corporal entre os estudantes, transformando a sala em um laboratório sonoro (Santos *et al.*, 2016).

As conjecturas teóricas levantadas em Santos *et al.* (2016) focaram em: se a diminuição das barreiras técnicas para a produção de música facilita o aprendizado, se o uso do corpo e de movimentos corporais apoiam a aprendizagem e se a colaboração entre os alunos apoia a aprendizagem. Para avaliar a validade do artefato e das conjecturas teóricas empreenderam um estudo de caso, com abordagem qualitativa de posição epistemológica interpretativista e tendo como base a abordagem de *Design Science Research*. Segundo os autores, o artefato possibilitou aos estudantes “a exploração ativa e autoral do discurso musical”, através do uso do corpo e da colaboração entre pares, tendo sido avaliado positivamente quanto aos aspectos de expectativa de performance, expectativa de esforço, motivação hedônica e intenção de uso.

Um dos insights mais interessantes da pesquisa de Santos *et al.* (2016) foi conquistado através da relação entre o contexto social dos estudantes de uma das turmas, localizada em uma escola na favela da Maré, onde os estudantes tinham receio de entrar em contato físico, foi que a atividade colaborativa com a utilização do artefato, permitiu que houvesse uma expressão mais equilibrada de força e confiança durante o contato físico. Outra descoberta de extrema importância foi que a grande maioria dos estudantes e todos os professores participantes da amostra pesquisada entenderam o artefato Tapetes Musicais como um instrumento musical em si e também capaz de ser um artefato coadjuvante na aprendizagem.

#### 2.5.1.3 Desenvolvimento de um Modelo da Participação em Bate-papo seguindo a abordagem de *Design Science Research*

Rocha *et al.* (2014) utilizaram a DSR para criar um modelo de participação em bate-papo seguindo a abordagem de Design Science *Research*, motivado pelo crescimento da Educação a Distância no Brasil a partir do ano 2000<sup>8</sup>. De acordo com Rocha *et al.* (2014), quando o número de participantes em um bate-papo é excessivo, existe uma tendência a ocorrer sobrecarga de mensagens, quando uma quantidade grande delas é enviada em um dado intervalo de tempo — a chamada “Confusão no bate-papo” (Rocha *et al.* 2014) — que dificulta a interação entre os participantes. A fim de descobrir a quantidade de alunos que pode participar de um bate-papo mantendo o nível de participação desejada pelo professor, Rocha *et al.* (2014) propuseram um modelo matemático da participação de grupos de alunos com base em uma sessão de bate-papo educacional, utilizando a Teoria de Filas (Rocha *et al.* 2014). Na Teoria de Filas, “Cada participante do bate-papo é representado por um servidor, que precisa ler (servir) todas as mensagens (clientes) digitadas pelos demais, que vão sendo enfileiradas para a leitura na ordem em que são recebidas” (ROCHA *et al.* 2014). Para avaliar o artefato, foram comparados os dados coletados em 142 sessões de bate-papo reais com estimativas de logs realizadas no ambiente de modelagem Tangram-II, contendo a duração do

---

<sup>8</sup> A partir do ano 2000, a EAD passou a ser responsável por quase 15% do número de matrículas em cursos de graduação e que em 2010, o uso de sistemas de bate-papo e fórum eram os recursos mais usados na EAD, sendo adotados por 70% das instituições (CENSOEAD.BR, 2010).

bate-papo, a quantidade de participantes e a quantidade de mensagens enviadas (ROCHA *et al.*, 2014).

Os resultados da pesquisa de Rocha *et al.* (2014) indicaram que houve adequação do modelo e a caracterização do comportamento de participação de um estudante em função dos processos elementares “leitura”, “escrita” e “reflexão”, possibilitou: equacionar a quantidade máxima de mensagens em função da quantidade de participantes em um bate-papo; estimar o número de mensagens esperado em um determinado contexto; verificar que quanto mais participantes houverem em um bate-papo, menor é a quantidade de mensagens produzidas, porque um participante levará mais tempo para ler todas as mensagens dos outros participantes (ROCHA *et al.* 2014).

### 3 DISCUSSÃO DA SOLUÇÃO

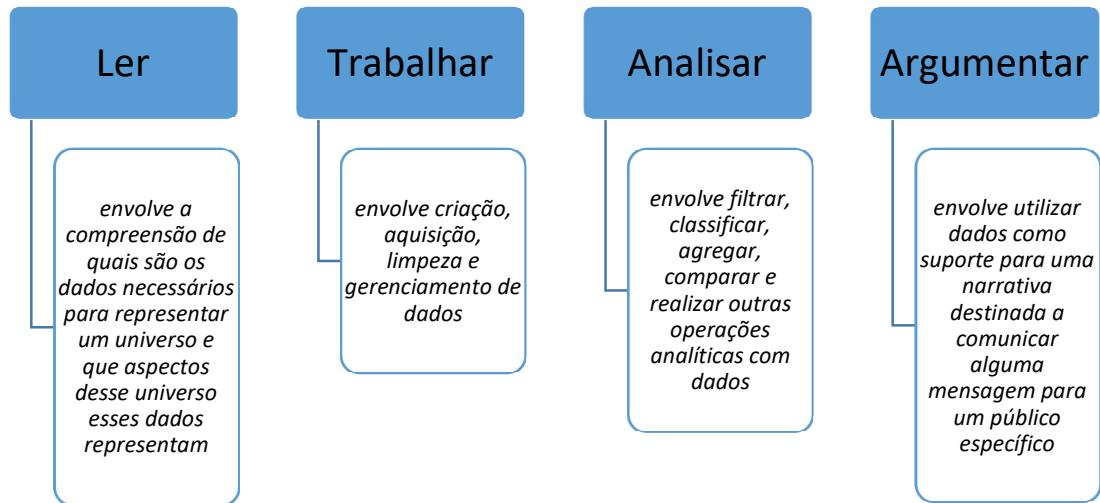
---

Neste capítulo será detalhada a origem da proposta de solução para a pergunta de pesquisa “**De que forma podemos avaliar os conhecimentos em Literacia de Dados de estudantes?**”. Para isto, será apresentado o artefato sociotécnico que foi criado como parte da solução e a metodologia de pesquisa utilizada, bem como um algumas considerações sobre a abordagem epistemológico-metodológica da *Design Science Research* e a sua instanciação para esta pesquisa. Também será detalhado o método de pesquisa aplicado, composto por um estudo de caso mais uma análise qualitativa de orientação interpretacionista e o uso da Teoria Fundamentada em Dados.

A solução proposta para o problema de pesquisa foi desenvolvida a partir da definição de Literacia de Dados de D’Ignazio e Bhargava (2015). Primeiramente, lançou-se mão da definição de Literacia de Dados e identificou-se os verbos que indicavam competências próprias da Literacia de Dados: ler, trabalhar, analisar e argumentar (Figura 12).

Em um segundo passo, separou-se cada um dos verbos, decompondo a definição em suas partes constituintes. Segundo D’Ignazio e Bhargava (2015), ler dados envolve a compreensão de quais são os dados necessários para representar um universo e que aspectos desse universo esses dados representam; trabalhar com dados envolve criação, aquisição, limpeza e gerenciamento de dados; analisar dados envolve filtrar, classificar, agregar, comparar e realizar outras operações analíticas com dados e argumentar com dados envolve utilizar dados como suporte para uma narrativa destinada a comunicar alguma mensagem para um público específico.

Figura 12: Decomposição da definição de D'Ignazio e Bhargava (2015) segundo competências relacionadas com a Literacia de Dados.



Fonte: a autora (2020). Adaptado de D'Ignazio e Bhargava (2015).

Em um terceiro passo, analisou-se qual seria a relação entre as partes constituintes resultantes da decomposição da definição de Literacia de Dados de D'Ignazio e Bhargava (2015) e percebeu-se que havia grande semelhança com o processo de Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Datasets*) - (KDD), dado que suas etapas, apesar de diferentemente nomeadas por D'Ignazio e Bhargava (2015) e pelos diversos autores da área de mineração de dados, transita entre as fases retiradas, como exemplo, de Roiger e Geatz (2003):

1. Identificação de objetivos / Realização de uma pergunta de pesquisa
2. Descoberta de um *dataset* através do qual se poderá buscar as informações desejadas
3. Pré-processamento de dados
4. Transformação de dados
5. Mineração de dados
6. Interpretação dos resultados obtidos e avaliação / Revisão das etapas anteriores
7. Ação através da aplicação da descoberta em seu contexto

Observando que existia uma identidade entre o processo de KDD e as competências em Literacia de Dados listadas na definição de D'Ignazio e Bhargava (2015), quando vistas como um processo que se iniciava com a leitura de dados e terminava com o nascimento da possibilidade de argumentação com dados, compreendeu-se que um processo de KDD simplificado poderia ser abstraído da definição de D'Ignazio e Bhargava (2015) e que isto poderia trazer benefícios em termos pedagógicos para o aprendizado de Literacia de Dados em diversos segmentos de ensino.

A partir da percepção de que se poderia utilizar a definição de D'Ignazio e Bhargava (2015) como uma abstração do conceito de KDD e que essa abstração poderia auxiliar na aprendizagem, buscou-se utilizar esse conceito como base para a criação de um artefato sociotécnico que tivesse como objetivo transformar os conhecimentos tácitos dos estudantes em conhecimentos formalizados (ou explícitos), proporcionando a possibilidade de se avaliar o nível de Literacia de Dados de estudantes. A partir deste ponto foram levantadas três conjecturas (ou suposições teóricas) – com base nas orientações da DSR – que teriam o poder de guiar a criação de um artefato para a avaliação da Literacia de Dados dos estudantes e a consequente resolução do problema de pesquisa. As conjecturas levantadas foram:

i. através da contação de histórias em grupo por meio de dados (*data group storytelling*) os conhecimentos tácitos dos estudantes podem ser formalizados com mais facilidade (Lage et al., 2017).

ii. o processo de aprendizagem e formalização de conhecimentos pode ser facilitado através do trabalho colaborativo, na forma de grupos (ou times) de estudantes, levando-se em conta que cada um traz consigo experiências prévias únicas que somadas podem aumentar a probabilidade de ocorrência de insights durante a pesquisa.

iii. o processo de aprendizagem e formalização de conhecimentos pode ser estimulado através da apresentação de dados que refletem os interesses e contexto dos estudantes envolvidos (Freire, 1970).

Convencionou-se então, que seria criado um *pipeline* de análise de dados composto pelas fases – oriundas da definição de Bhargava e D'Ignazio (2015) - “trabalhar com dados”, “ler dados”, “analisar dados” e “argumentar com dados” (Figura 13), em que a ação de

trabalhar com dados foi deliberadamente posicionada anteriormente às fases de leitura de dados, análise de dados e argumentação com dados, com o objetivo de evitar possível sobrecarga cognitiva durante a utilização do artefato. A sobrecarga cognitiva, segundo Sweller (1988), acontece quando a demanda total de processamento de cognições externas e estímulos internos superam os recursos de atenção disponíveis por uma pessoa. A sobrecarga cognitiva pode impactar grandemente o aprendizado e o desempenho, principalmente em contextos autênticos (Feldon, 2007).

Figura 13: *Pipeline* de dados concebido através da definição de literacia em dados de D'Ignazio e Bhargava (2015).



Fonte: a autora (2020). Adaptado de D'Ignazio e Bhargava (2015).

Mediante as ideias e preocupações levantadas, definiu-se que os pesquisadores “trabalhariam com [os] dados”, isto é, realizariam toda a parte de criação, aquisição, limpeza e gerenciamento – iniciando o *pipeline* (Figura 13) – para que os estudantes pudessem percorrer o restante do pipeline mais confortavelmente no tempo disponível, acreditando que poderiam aprender melhor neste formato minimalista, com foco no melhor aproveitamento das etapas subsequentes do *pipeline*, respeitando a cognição dos estudantes e portanto, auxiliando o fluxo de aprendizagem durante esta proposta específica.

Depois disto, foi pensado o conceito de uma dinâmica que contaria com a criação de cartas de desafio para cada uma das fases do pipeline com questões que os alunos deveriam responder sobre um determinado *dataset*. As cartas, então, foram elaboradas de forma que se assemelhassem com cartas de desafio, para incentivar o engajamento dos alunos na atividade. Desta forma, o artefato construído se materializou em um conjunto de cartas que foi nomeado “Cartas para a Literacia”.

A concepção do artefato surgiu através da reflexão derivada do desenho do *pipeline*, que serviu como estrutura para a percepção de que os estudantes deveriam passar de maneira necessariamente sequencial pelas etapas de leitura de dados, análise de dados e argumentação com dados, para que ao final entregassem como resultado da atividade um texto detalhado que contasse para a audiência de pesquisadores e especialistas envolvidos uma história a partir de um *dataset* oferecido – uma contação de história através de dados.

O modo encontrado para articular essas aspirações foi criando solicitações que estimulassem os estudantes a passar pelas fases do pipeline de forma reflexiva e crítica e que também oferecessem um gancho para a facilitar as respostas das solicitações propostas nas fases subsequentes. As solicitações levantadas foram, para as três etapas do pipeline: (i) Escreva um exemplo de dado contido no *dataset* (Ler dados); (ii) Construa um gráfico de seu interesse com base no *dataset* (Analisar dados); (iii) Explique o que significa o gráfico que você acabou de fazer (iv); Apresente uma descrição textual detalhada sobre as características da amostra representada no *dataset*, com base nas análises que você realizou nas etapas anteriores (Argumentar com dados) (Quadro 2).

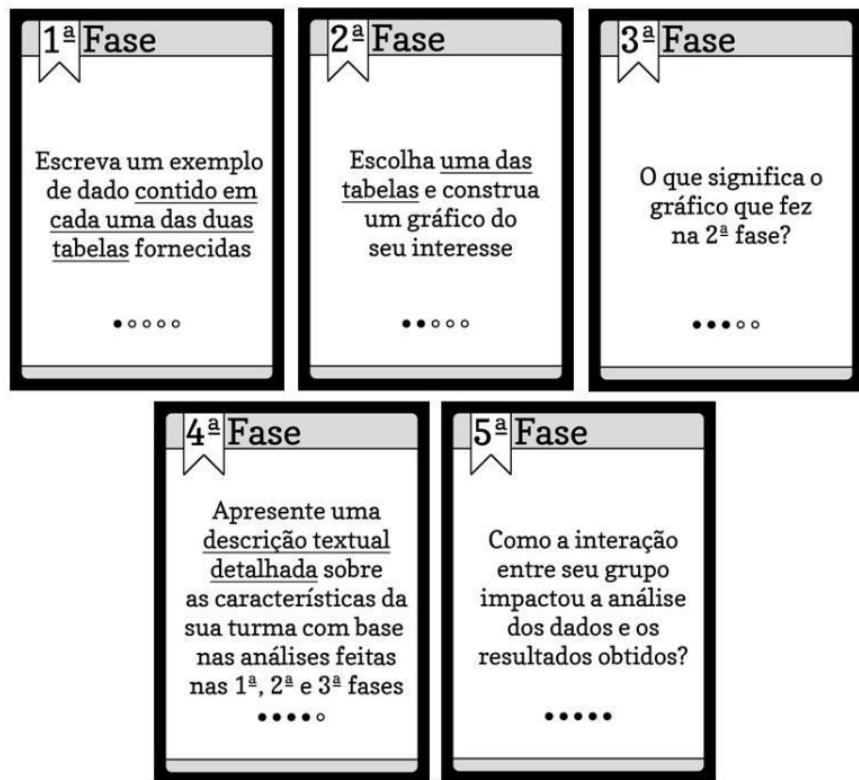
Quadro 2: Etapas do *pipeline* relacionadas às questões pensadas na fase de ideação das cartas.

<b>Etapas do pipeline</b>	<b>Questões/Solicitações</b>
Ler dados	Escreva um exemplo de dado contido no <i>dataset</i>
Analizar dados	Construa um gráfico de seu interesse com base no <i>dataset</i>
	Explique o que significa o gráfico que você acabou de fazer
Argumentar com dados	Apresente uma descrição textual detalhada sobre as características da amostra representada no <i>dataset</i> , com base nas análises que você realizou nas etapas anteriores

Fonte: reprodução da autora (2020).

Pode-se observar a primeira versão do conjunto de cartas na Figura 14 a seguir.

Figura 14: Primeira versão do artefato Cartas para a Literacia criado a partir da definição de Literacia de Dados de D'Ignazio e Bhargava (2015) e utilizada durante a atividade do primeiro ciclo do estudo de caso.



Fonte: reprodução da autora (2020).

A princípio pensou-se que a materialização das questões se daria em forma de cartões nos quais seriam apresentadas as solicitações aos estudantes, respeitando a sequência do *pipeline* (Figura 13). Então iniciou-se uma fase de *brainstorming* que permitiu que a evolução do *design* das cartas passasse por fases de transformação que foram guiadas pela preocupação com a qualidade do material didático que estaria sendo oferecido para os estudantes: qualidade que se desejou que estivesse refletida na sua forma, na sua função, na sua estrutura e no seu processo de aprimoramento. Por forma, entende-se o seu padrão espacial; por função, a tarefa ou objetivo que o artefato se propõe a desempenhar; por estrutura entende-se a maneira com que cada uma das cartas se inter-relacionam; e por processo, quais deveriam ser as modificações implementadas para que o *design* se mantivesse adaptado aos diferentes públicos-alvo que fossem realizar a atividade ao longo das jornadas de estudos de caso.

Sobre a forma, para uma primeira versão dos cartões preferiu-se referências relacionadas a cartas de jogos, porque acreditou-se que um *design* desse tipo poderia fazer com que os estudantes relacionassem a tarefa a um conjunto de desafios e isso impactaria positivamente no estímulo para prosseguirem até o desafio final. Pelo mesmo motivo adicionou-se um elemento identificador de progresso: a inscrição da “Fase” a qual a carta pertenceria. O uso da escala de cinza foi para eliminar estímulos desnecessários no momento da atividade e o uso do sublinhado foi para destacar informações que pareciam mais importantes para a realização da atividade. Utilizou-se como referência visual as cartas do jogo *Love Letter*, publicado pela *Alderac Entertainment Group*, 2012 (Figura 15).

Figura 15: Cartas do jogo Love Letter.



Fonte: © Alderac Entertainment Group, 2012.

Sobre a sua função, cada carta conteve em si uma pergunta que tinha como objetivo tornar explícitos conhecimentos tácitos sobre Literacia de Dados.

Sobre a estrutura, a intenção foi apresentar as cartas com perguntas que conduzissem os estudantes por um caminho de reflexão sobre os dados fornecidos, de forma que esta reflexão, formalizada nas respostas dadas a cada uma das fases, culminasse em uma descrição textual detalhada sobre as características da amostra representada no *dataset* fornecido. Ainda sobre a estrutura, adicionamos uma barra de status estilizada para que o estudante soubesse se estaria mais no início ou mais para o final da atividade.

Sobre o processo, tivemos duas versões iniciais: a primeira, que utilizamos com estudantes da graduação, em um estudo de caso em primeiro ciclo, com 5 cartas, e a segunda,

utilizada com estudantes de pós graduação, em um estudo de caso em segundo ciclo, já contando com 1 carta adicional, revelando alterações importantes no *pipeline*, fruto de reflexões já mais amadurecidas realizadas no âmbito do grupo de pesquisa.

### 3.1 INSTANCIAMENTO DO DSR PARA ESTA PESQUISA

Nesta pesquisa criou-se um artefato que se materializou em um conjunto de cartas de desafio com o objetivo de oferecer uma proposta de solução para a ausência de instrumentos capazes de avaliar a Literacia de Dados de estudantes de diversos níveis de escolaridade. O público-alvo estudado foi composto por estudantes de graduação e pós-graduação. Os estudantes de graduação eram do primeiro período, para que se pudesse contar com indivíduos que tivessem recém concluído o ensino médio. A escolha da amostra composta pelos estudantes da pós graduação deveu-se a uma tentativa de verificar se os resultados obtidos com a turma da graduação, além de verificados também poderiam ser melhor detalhados ou mesmo ampliados.

Para a criação do artefato sociotécnico Cartas para a Literacia foram levantadas três conjecturas teóricas: i. através da contação de histórias por meio de dados é possível externalizar conhecimentos tácitos; ii. o processo de aprendizagem pode ser facilitado através de um *dataset* que reflita os interesses dos estudantes e iii. o trabalho colaborativo pode facilitar o processo de aprendizagem em Literacia de Dados.

Na Figura 2 tem-se a instanciação do mapa de DSR de Pimentel *et al.* (2019) para o contexto desta pesquisa.

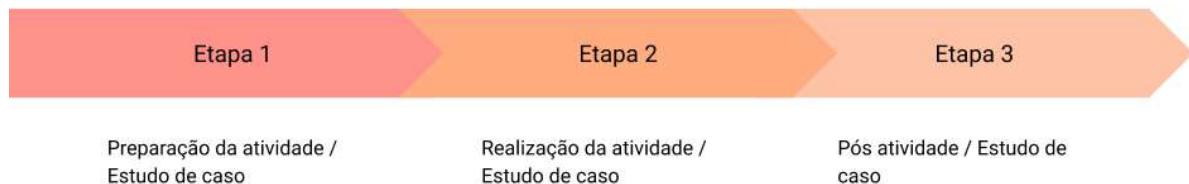
A criação do artefato foi fundamentada pelo estado da técnica, que se refletiu em uma revisão da literatura relacionada ao desenvolvimento do artefato, composta por: i. artigos contextualizando a Literacia de Dados, apresentando diferentes abordagens e definições; ii. Base Nacional Comum Curricular; iii. literatura relacionada à avaliação em Literacia de Dados, representada como referencial teórico; iv. trabalhos relacionados à criação de artefatos para a educação, utilizando a DSR; v. estudo do quadro teórico que fundamenta as conjecturas teóricas utilizadas para direcionar o projeto do artefato. Neste caso, foi realizada uma revisão de literatura com foco em: trabalho colaborativo e *group storytelling*.

Para avaliar a validade do artefato, das suas conjecturas teóricas de base e se o problema de pesquisa foi resolvido satisfatoriamente, foi utilizado como método de pesquisa

um estudo de caso com estudantes em sala de aula, que foram analisados qualitativamente, através da triangulação entre: os dados de observação direta, realizada durante o estudo de caso e os resultados da aplicação de uma tarefa de avaliação de Literacia de Dados, com a utilização de um conjunto de cartas de desafio e um *dataset*. Foi utilizada a análise fundamentada em dados para promover a descoberta de conhecimento e possivelmente geração de teorias através dos dados coletados.

O procedimento de estudo de caso foi orientado pelas seguintes etapas e subetapas (Figura 16):

Figura 16: Etapas de condução do procedimento de estudo de caso.



Fonte: reprodução da autora (2020).

### 3.1.1 **Etapa 1 (preparação da atividade/estudo de caso):**

- criação de uma atividade de literacia em dados em que os estudantes utilizam as cartas de desafio para encontrar informações em um *dataset* com atributos relacionados às características de uma turma de estudantes.
- Delimitação do tempo de resposta para cada fase da atividade, representada por uma carta de desafio: 10 minutos para a Fase 1, 10 minutos para a Fase 2, 15 minutos para a Fase 3, 20 minutos para a Fase 4 e 5 minutos para a Fase 5 (Figura 17).

Figura 17: As fases da atividade e o tempo para a resolução de cada solicitação representada pelas cartas de desafio.



Fonte: reprodução da autora (2020).

- Preparação de uma planilha compartilhada em nuvem, contendo duas versões de um mesmo *dataset* com atributos relacionados a características e preferências de uma turma de estudantes de graduação, devidamente anonimizados e identificados por um número para cada indivíduo cujos dados foram coletados.
- Aquisição de materiais de papelaria diversos (notas autoadesivas, cola colorida, paetês, lantejoulas, carimbos, régulas com moldes variados, papéis coloridos e sulfite, adesivos de formatos variados, tesouras, massa de modelar, entre outros).
- Reserva de laboratório de computadores com acesso à internet disponível, bem com autorização para a instalação de softwares de análise de dados que os estudantes julgassem importantes para sua análise.
- Preparação de um cronômetro para indicar a progressão das fases para os estudantes.
- Preparação de uma breve explicação para os estudantes sobre a dinâmica da atividade.
- Organização da equipe de pesquisa necessária para atuar durante toda a atividade em funções variadas, incluindo a distribuição das cartas para os estudantes nos momentos definidos, a realização de anotações através de observação direta e a documentação digital da atividade.

### **3.1.2 Etapa 2 (realização da atividade/estudo de caso):**

- Recepção dos estudantes pela equipe de pesquisa e explicações gerais sobre a atividade, incluindo a possibilidade e acesso irrestrito à internet, a possibilidade de uso de materiais de papelaria de forma irrestrita, porém compartilhada entre todos os grupos e explicação de que todas as fases terão um tempo cronometrado para a sua realização. Além disso a explicação de que a equipe não poderá fornecer as respostas para os desafios, mas poderá contribuir tirando dúvidas.
- Separação da turma em grupos/times.
- Compartilhamento das planilhas eletrônicas com o *dataset* para a atividade, sendo uma planilha compartilhada com cada grupo.
- Fornecimento de um bloco com 5 folhas de papel sulfite como possibilidade de recurso para a documentação das fases da atividade. Disponibilização de folhas extras para o caso de precisarem.
- Disponibilização e oferecimento explícito do material de papelaria para ser utilizado como recurso opcional durante a atividade.
- Aviso de início da atividade e controle do tempo.
- Realização de observação direta do comportamento dos estudantes durante a prática da atividade.
- Documentação da atividade em forma de relatório de observação e de fotografias.
- Coleta de todo o material impresso e em meio digital resultante das análises realizadas por cada um dos grupos.

### **3.1.3 Etapa 3 (Pós-atividade/estudo de caso)**

- Leitura dos dados produzidos pelos estudantes.
- Análise preliminar dos dados produzidos pelos estudantes.
- Curadoria de dados:
  - Download de todo o material produzido pelos estudantes em formato digital.
  - Digitalização de todo o material impresso produzido pelos estudantes.
  - Organização, análise, entendimento e documentação das anotações obtidas durante a atividade do estudo de caso (nesta etapa se fez necessário conversar com toda a equipe de pesquisa a fim de dirimir dúvidas quanto às anotações e

buscar um entendimento coletivo de momentos importantes, ocorrências interessantes e insights obtidos durante a atividade/estudo de caso.

### 3.2 ANÁLISE FUNDAMENTADA EM DADOS

A seguir será apresentada uma breve descrição da análise fundamentada em dados, bem como a explicação de quais abordagens da análise fundamentada serão utilizadas nesta pesquisa e quais fatores motivaram a sua escolha.

Escolheu-se a Análise Fundamentada em Dados (*Grounded Theory*) - GT como método de análise de dados para esta pesquisa porque ela se propõe a estudar as experiências das pessoas com um processo ou um fenômeno, buscando padrões de comportamento, sendo que o processo do qual este estudo trata é o processo de avaliação de Literacia de Dados de grupos de estudantes de níveis de graduação e pós-graduação da rede pública de ensino. Ele se desenvolve através da utilização de um artefato sociotécnico que se manifesta por meio de um conjunto de cartas que propõem uma sequência de desafios para serem realizados utilizando um conjunto de dados no estilo *small data*. Nesta seção a GT será apresentada, bem como se dará o seu emprego nesta pesquisa.

Para a finalidade desta pesquisa utilizaremos a definição da Teoria Fundamentada em dados segundo Charmaz (1983) em que a GT se assenta em “um conjunto consistente e lógico de dados coletados e processados analiticamente com o objetivo de desenvolver teorias”. A GT é ao mesmo tempo um método e uma teoria. É um método porque apresenta um processo sistemático de pesquisa e é uma teoria porque é um canal para a geração de teorias que permitem estabelecer um fundamento para o entendimento dos conceitos e fenômenos estudados.

A Teoria Fundamentada em Dados foi desenvolvida pelo sociólogos Barney Glaser e o sociólogo médico Anselm Strauss, ambos dedicados ao estudo de campo de doenças crônicas e da consciência da morte na Universidade da Califórnia, São Francisco, que desde sua fundação em 1864, como *Toland Medical College*, é um centro de pesquisa e ensino médico e biológico de grande importância (University Of California San Francisco, 2019). A GT tem sua origem, então, na sociologia médica, com a publicação do livro *Awareness of Dying* em 1964 (Glaser e Strauss, 1964) e posteriormente do livro “*The Discovery of Grounded Theory*” por volta do ano de 1967 (Glaser e Strauss, 1967), legitimando a pesquisa qualitativa em um

contexto histórico complexo, em que a abordagem quantitativa na pesquisa já não dava conta sozinha das demandas em ciências sociais aplicadas à saúde.

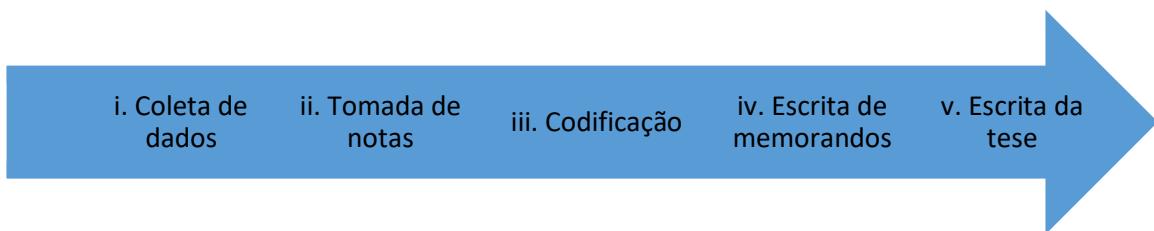
Dentro da conjuntura social de sua época de vida e de pesquisa, Glaser e Strauss foram influenciados pela abordagem sociológica do interacionismo simbólico - que é uma linha de pesquisa sócio psicológica e sociológica que iniciou-se estimulada pelas mazelas sociais consequentes da primeira revolução industrial - em que volta-se o olhar para as pessoas de um ponto de vista humanístico, percebendo-as como sujeitos capazes de atuar de forma ativa e flexível em um determinado contexto, fazendo uso de seu poder de comunicação e interpretação de signos, permitindo sua adaptação ergonômica dependendo de como o seu intelecto comprehende e define a situação em que estão colocados (Carvalho *et al.*, 2010). O interacionismo simbólico lançou-se na criação de uma estrutura de análise social e do comportamento humano que agrupava ideias relacionadas a “grupos humanos ou sociedades, interação social, objetos, o ser humano como ator e a ação humana e as interconexões entre as linhas de ação” (Blumer, 1982), posicionando o pesquisador como observador ativo no ambiente de pesquisa e interativo com os pesquisados, além de livre para interpretações a partir do seu próprio ponto de vista (Jeon, 2004) formando, juntamente com outras abordagens sociológicas, um terreno fértil para o fortalecimento das raízes filosóficas da GT.

A GT em geral foca na construção de teorias através da criação de sentido proveniente da análise qualitativa de dados. Isto é, os dados coletados é que permitem que os pesquisadores descubram a estrutura do conhecimento. Um dos principais benefícios do uso da GT para os pesquisadores é que pode-se utilizar tipos variados de dados - vídeos, entrevistas, grupos focais, imagens, documentos textuais pré-existentes, falas, fotografias, diários e observações - para obter perspectivas em profundidade e, a partir dessas perspectivas, encontrar novas teorias a partir da codificação dos dados em categorias e subcategorias, que terão como objetivo a extração de significados e relações entre ideias (Merriam e Tisdell, 2015; Creswell e Poth, 2016; *Grounded Theory Solutions*, 2019).

A GT pode auxiliar no desenvolvimento de teorias através de fases, que têm como objetivo responder uma pergunta de pesquisa e podem ser resumidas como: (i) Coleta de dados: através de observação, entrevistas e conversas, por exemplo. (ii) Tomada de notas: a cada coleta de dados é importante que se faça a tomada de notas. (iii) Codificação: conforme se codifica os dados, é esperado que surjam teorias. (iv) Escrita de memorandos: nesta fase

há a escrita formal de memorandos, que são relatórios com o objetivo de registrar ocorrências e insights para que se possa recorrer posteriormente. (v) Escrita da tese: espera-se que as informações e insights contidos nos memorandos, bem como sua organização, sirvam como guia principal para a escrita da tese (Merriam e Tisdell, 2015; Creswell e Poth, 2016; *Grounded Theory Solutions*, 2019). Na Figura 18 há um esquema resumido das fases de aplicação da GT.

Figura 18: Esquema de aplicação da GT, com suas fases.



Fonte: a autora (2020). Adaptado de Merriam e Tisdell (2015); Creswell e Poth (2016); *Grounded Theory Solutions* (2019).

A GT evoluiu ao longo do tempo, e para a finalidade desta pesquisa destacam-se 4 abordagens que foram escolhidas para serem comentadas brevemente, dada a sua sequência temporal de surgimento e importância para esta pesquisa: a versão de Strauss e Corbin (Strauss & Corbin, 1990, 1998), a versão de Glaser (Glaser e Strauss, 1967), a Teoria Fundamentada Construtivista (Charmaz, 1983) e a Teoria Fundamentada em Dados Feminista (Wuest, 1995).

A GT teve seu pontapé inicial quando Glaser e Strauss (1964) desenvolvem uma forma de analisar a consciência da morte em pacientes de um hospital. Essa forma é considerada a primeira proposta de GT e consiste na aplicação dos seguintes passos: (i) desenvolver a consciência de quais tipos de dados relacionados ao contexto estão disponíveis; (ii) pesquisar que tipos de dados são relevantes para a pesquisa que está sendo realizada; (iii) com base nas variáveis relevantes, julgar se há outros tipos de variáveis que seriam importantes para analisar os dados.

Strauss e Corbin (1990) sublinham que é muito importante que se siga cânones para a realização da pesquisa em GT e para a avaliação de estudos em GT, são eles: (ii) A coleta de dados e a análise são processos que estão interrelacionados; (ii) Conceitos são as unidades

básicas de análise; (iii) Categorias devem ser desenvolvidas e relacionadas para formar teorias; (iv) Criar uma amostra na GT, que deve obedecer a fundamentos teóricos, sendo necessário ter conhecimento sobre o contexto que será estudado e ter a capacidade de extrair informações detalhadas e conceituais sobre ele, enxergando-o em nuances, intensidades, potenciais e focos diferentes; (v) As análises devem ser feitas por comparação constante; (vi) Padrões e variações devem ser levados em conta, isto é, encontrar padrões e regularidades auxilia na integração e na organização dos dados; (vii) O processo deve estar incorporado na teoria, podendo-se entender o processo como uma sequência de estágios ou ações e interações que se sucedem como respostas a condições estabelecidas; (viii) Escrever memorandos é uma parte integral da prática da GT: os memorandos não devem simplesmente expressar ideias, mas devem estar relacionados com a formulação e revisão da teoria durante o processo de investigação. Devem ser uma ação que se desenvolve durante toda a prática da pesquisa que utiliza a GT; (ix) Hipóteses sobre os relacionamentos entre as categorias devem ser desenvolvidas e verificadas constantemente durante o processo de pesquisa, para que ao final seja apresentada uma teoria com mais solidez; (x) É recomendado que a pesquisa em GT seja realizada em equipe, para diminuir a possibilidade de vieses, aumentar a sensibilidade teórica e proporcionar uma análise colaborativa; (xi) Devem ser levadas em conta as condições relacionadas ao contexto mais amplo onde se localiza a amostra pesquisada, isto é, condições políticas e econômicas, movimentos sociais, valores culturais e etc. (Strauss e Corbin, 1990).

Charmaz (1983) apresenta um olhar construtivista sobre a GT, dialogando com as experiências anteriores de Glaser e Strauss (1964) e Strauss e Corbin (1990) através da corrente de pensamento socioconstrutivista, aderente às atitudes de (i) Tratar o processo de pesquisa em si mesmo como uma construção social, (ii) Analisar minuciosamente as decisões tomadas e indicações oferecidas, (iii) Tomar a liberdade de improvisar no uso de estratégias metodológicas e analíticas durante o processo de pesquisa, (iv) Coletar dados bastantes para distinguir e documentar a forma como os participantes da pesquisa constroem suas próprias experiências de vida (Charmaz, 2006). A autora ainda critica a aplicação da GT como uma receita pronta para a condução de pesquisas qualitativas, chamando os seus leitores para a responsabilidade de responder a questões emergentes relacionadas ao contexto de estudo,

abrir a percepção para novos insights no decorrer da construção do método de análise (GT), assim como da própria análise dos dados que surgirão através da aplicação do método.

### **3.2.1 A influência do referencial interpretacionista no uso da GT nesta pesquisa**

O interpretacionismo e os referenciais críticos e pós-modernos têm se colocado no Brasil, desde a década de 70, aproximadamente, como visões epistemológicas alternativas à doutrinação funcionalista como paradigma de análise organizacional, criticando o seu determinismo obstinado e sua orientação voltada para a manutenção da ordem social, dado que tanto o interpretacionismo quanto os críticos pós-modernos enxergam a realidade como uma teia de representações complexas e subjetivas que são capazes de impulsionar mudanças sociais e inovações (Vergara e caldas, 2005).

A partir dos conceitos e evoluções da GT que foram expostos e por este trabalho adotar uma visão interpretacionista das organizações e sistemas de informações, optou-se por utilizar a teoria fundamentada em dados construtivista de Charmaz (2006), considerada uma GT de segunda geração, devido ao seu surgimento posterior, porém dialógico com as divergências epistemológicas entre as GTs de Glaser e Strauss, juntamente com a perspectiva da GT feminista de Wuest (1995), por representar uma inspiração sobre como dar voz de forma igualitária à pluralidade de sujeitos representados na amostra pesquisada.

A GT de abordagem feminista foi desenvolvida por enfermeiras através reconhecimento do viés androcêntrico e para garantir que as vozes das mulheres fossem ouvidas nas comunidades de pesquisa (Wuest, 1995). A pesquisadora Wuest selecionou elementos das três teorias fundamentadas para apresentar a importância da fusão com a teoria feminista, afirmando que o feminismo não é um método de pesquisa, mas uma perspectiva que deve ser aplicada a qualquer método de pesquisa (Wuest, 1995).

Entende-se que juntas, Charmaz (1983) e Wuest (1995) se adequam melhor à visão interpretacionista, ou seja, ao entendimento de que as organizações são movidas por processos que se dão através da interação humana e que, por isso possuem natureza cinética e espontânea mais parecida com o movimento browniano e processos estocásticos do que com a ideia de determinismo. Segundo Mörters e Peres (2010) movimento browniano é um tipo de processo da teoria probabilística ou estocástica, em que, mesmo se conhecendo as condições iniciais de um sistema, não se pode determinar a forma ou a direção de evolução

do sistema. O movimento browniano é um tipo de movimento aleatório descrito pelas partículas presentes em um fluido, a partir da sua colisão com outras partículas vizinhas (Mörters e Peres, 2010).

Dessa forma, pretende-se entender o processo de aprendizagem de Literacia de Dados através do uso de um artefato sociotécnico a partir do ponto de vista das pessoas envolvidas no processo, sabendo que será imperativo lidar com fatores limitantes próprios da GT, como por exemplo: os vieses das próprias pesquisadoras, o empreendimento de conseguir participantes para serem pesquisados, o tempo para a análise de dados, a necessidade de uma mentalidade analítica sobre dados qualitativos e a ansiedade de trabalhar com um processo de análise em que se detém pouco controle sobre os próximos passos da pesquisa, pois eles serão descobertos durante a trajetória.

## 4 AVALIAÇÃO DO ARTEFATO CARTAS PARA A LITERACIA

---

Neste capítulo será detalhada a avaliação da pesquisa realizada, incluindo as fases Planejamento da Avaliação da Pesquisa, Execução da Avaliação para cada uma das implementações do estudo de caso e Apresentação dos Resultados para cada uma das implementações do estudo de caso.

### 4.1 INTRODUÇÃO SOBRE O ESTUDO DE CASO E SEUS CICLOS

Artificial é, segundo o *Collins Dictionary*<sup>9</sup>, o oposto de natural, sendo algo produzido pela humanidade, feito através da arte ou do trabalho humano e não pela natureza. Ainda segundo o *Collins Dictionary*, objetos, materiais ou processos artificiais não ocorrem naturalmente e são criados por seres humanos usando ciência ou tecnologia, por exemplo. Segundo Dresch *et al.* (2015), “o processo de criação de um artefato adequado e a investigação sobre seu uso num dado contexto se caracterizam como um meio para produzir conhecimento, o que faz do artefato um elemento central nas pesquisas concebidas no paradigma epistemológico das Ciências do Artificial”.

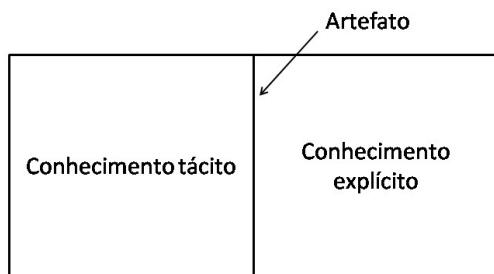
Os artefatos utilizados nesta pesquisa foram um conjunto de cartas de desafio (Figura 14) mais o *dataset*<sup>10</sup> Cores (Figura 8), criado pelos estudantes da turma de graduação que participou do primeiro ciclo do estudo de caso. Para a construção do artefato Cartas para a Literacia, partiu-se da definição de Literacia de Dados criada pelos autores D'Ignazio e Bhargava (2015), em que “a Literacia de Dados inclui a capacidade de ler, trabalhar, analisar e argumentar com dados” para a construção de um pipeline ou conjunto de etapas pelas quais acreditou-se que, se experimentadas por um grupo de estudantes em uma atividade de grupo, poderia oportunizar a avaliação das suas competências em Literacia de Dados, culminando em uma escala de Literacia de Dados. O artefato então, foi materializado através de um artifício composto por um conjunto de 5 cartas com desafios que deveriam ser respondidos em ordenação e intervalos de tempo definidos. O uso das “Cartas para a literacia” teve a intenção de servir como uma interface entre os conhecimentos tácitos dos estudantes sobre Literacia de Dados e os seus conhecimentos explícitos (Figura 19).

---

<sup>9</sup> collinsdictionary.com/pt/dictionary/artificial. Disponível em 14/02/2020.

<sup>10</sup> O *dataset* Cores pode ser acessado em  
[https://docs.google.com/spreadsheets/d/1PlPairMll4rZMZ8\\_WA6uUufs0\\_5VSJjf5-imKxAG-U/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1PlPairMll4rZMZ8_WA6uUufs0_5VSJjf5-imKxAG-U/edit?usp=sharing).

Figura 19: O artefato “Cartas para a literacia” como uma interface entre os conhecimentos tácitos de cada grupo pesquisado e os conhecimentos explícitos do grupo.



Fonte: reprodução da autora (2020).

Para avaliar a validade dos artefatos criados nesta pesquisa realizou-se um estudo de caso. Um estudo de caso é um método de investigação científica que pretende auxiliar na compreensão profunda e rigorosa de fenômenos complexos através de investigação empírica de um fenômeno contemporâneo dentro de seu contexto real (Yin,1989). Em um estudo de caso pode-se traçar comparações entre múltiplos casos, diferentes casos, casos únicos e entre subgrupos de um caso ou subsistemas de um sistema.

Este estudo serviu como um instrumento para a avaliação da validade do artefato em responder à pergunta de pesquisa proposta: “De que forma podemos avaliar os conhecimentos de estudantes em Literacia de Dados?”, feita a partir da observação do contexto de pesquisa, em que estudantes de ensino médio, graduação e pós-graduação estão vivendo em uma sociedade orientada a dados e não possuem meios de avaliar seus conhecimentos em Literacia de Dados. Na intenção de responder à pergunta de pesquisa, optou-se por utilizar o paradigma epistemológico metodológico da *Design Science Research*, aliado à uma visão interpretativista da pesquisa, a fim de se construir um artefato que cumprisse o objetivo de responder o problema desta pesquisa, criando uma solução que tenha a possibilidade de ser empregada em situações reais.

A avaliação do artefato foi realizada através de um estudo de caso com estudantes de instituições públicas de ensino do Rio de Janeiro, considerando os níveis “graduação” e “pós-graduação” como subgrupos para estudo. Nesta pesquisa o estudo de caso será o método utilizado para avaliar a validade do artefato Cartas para a Literacia através de uma averiguação empírica que investigará o fenômeno da Literacia de Dados no contexto estudantil, levando em conta que os limites entre o fenômeno da Literacia de Dados e a possibilidade de medição da Literacia de Dados de estudantes não estão claramente evidentes para a sociedade.

Esta pesquisa foi realizada em ciclos de investigação, sendo que após cada um dos ciclos houve refinamento do artefato. O primeiro ciclo de pesquisa destinou-se à aplicação do artefato em um contexto autêntico, enquanto que no segundo ciclo aplicou-se a atividade com as alterações que se percebeu necessárias com vistas à melhoria de performance do artefato e, após o uso do artefato, ainda foi observada alguma necessidade de aprimoramento, que foi implementada na sequência. Cada ciclo se destinou a responder se o artefato é válido, se as conjecturas teóricas levantadas durante a criação do artefato são válidas e se o artefato criado permite que se responda à questão de pesquisa formulada. A seguir apresentamos o detalhamento de cada ciclo do estudo de caso.

## 4.2 PRIMEIRO CICLO: GRADUAÇÃO

O primeiro ciclo de pesquisa foi realizado com uma turma de estudantes de graduação em Informática da Universidade Federal do Rio de Janeiro. Na seção 4.2.1 é descrita a fase de planejamento do primeiro ciclo do estudo de caso, como foi realizada a criação dos artefatos utilizados na atividade e como ocorreu o planejamento e a condução da atividade. Na seção 4.2.2 é descrito como foi realizada a atividade. Na seção 4.2.3 são apresentados os resultados obtidos para os desafios apresentados nas fases 4 e 5 do artefato e as observações diretas realizadas pela equipe de pesquisa durante a atividade.

### 4.2.1 Planejamento do primeiro ciclo

No primeiro ciclo, projetou-se um conjunto de cartas de desafio voltado para a medição do grau de Literacia de Dados de estudantes e adaptou-se o *dataset* Cores para que os estudantes pudessem nele fazer suas consultas, a fim de responder às perguntas propostas em cada uma das fases propostas nas cartas de desafio. Classificou-se o artefato Cartas para a Literacia como sendo um método para a avaliação da Literacia de Dados de estudantes, enquanto que o *dataset* foi considerado como um instrumento auxiliar para a utilização das cartas.

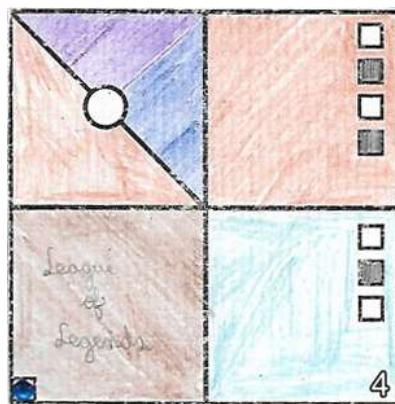
O artefato Cartas para a Literacia apresentou 5 desafios propostos para a turma de graduação (Figura 14). Os desafios propostos para a turma de graduação foram: 1- escrever um exemplo de dado contido em uma das duas tabelas fornecidas; 2- escolher uma das

tabelas e construir um gráfico do seu interesse; 3- explicar o que significa o gráfico que fez na 2ª fase; 4- apresentar uma descrição textual detalhada sobre as características da turma representada nas Tabelas 1 e 2, com base nas análises feitas nas fases 1, 2 e 3; 5- apresentar uma descrição textual contando como a interação do seu grupo impactou a análise dos dados e os resultados obtidos.

A fim de prover meios para que os estudantes pudessem responder os desafios propostos nas cartas, foi feita a adaptação de um *dataset*, criado com base em uma ideia original de Luppi e Posavec (2016), por estudantes de uma turma de Sistemas de Informação do curso de graduação em Ciência da Computação da Universidade Federal do Rio de Janeiro. Cada cartão tinha o formato quadrado e espaços destinados a serem coloridos conforme as respostas dadas por cada um dos estudantes a uma série de perguntas que tinham como temática geral abordar características pessoais dos 45 respondentes (Figura 21). As perguntas que originaram o *dataset* foram: 1- Já conhecia alguém antes da primeira aula?; 2- O que é mais importante para você entre as opções tirar nota boa, aprender, as pessoas e outro?; 3- Qual é o seu animal de estimação entre as opções: cachorro, gato, pássaro ou peixe e outros; 4- O que você faz para esparecer entre as opções: praia, música, dançar, esportes não digitais, maratonar seriados, jogar games e outro?; 5- Vai ter prova, você: já vinha estudando ao longo do período, começa uma semana antes, estuda de véspera, traz cola, vai na sorte; 6- Qual é a sua preferência para grupos de trabalho, entre as opções: prefiro fazer trabalhos sozinho, prefiro fazer trabalhos em dupla, prefiro fazer trabalhos em grupos maiores que duas pessoas; 7- Qual é a frequência com que você joga games, entre as opções: 1 a 8h por semana, 9 a 20 horas por semana, mais de 20h por semana; 8- O estudo rende mais: de manhã, de tarde ou à noite?

Ainda havia uma informação que foi colocada espontaneamente, através do recurso da escrita, por alguns dos respondentes dos cartões coloridos: qual o jogo que mais joga. Na Figura 20, um estudante respondeu a esta pergunta escrevendo “*League of Legends*”.

Figura 20: Um dos 45 cartões coloridos do dataset Cores (Figura 8) gerado por estudantes de uma turma de Sistemas de Informação da UFRJ. As cores pintadas à lápis de cor representam as respostas para as 8 perguntas feitas aos estudantes da turma.



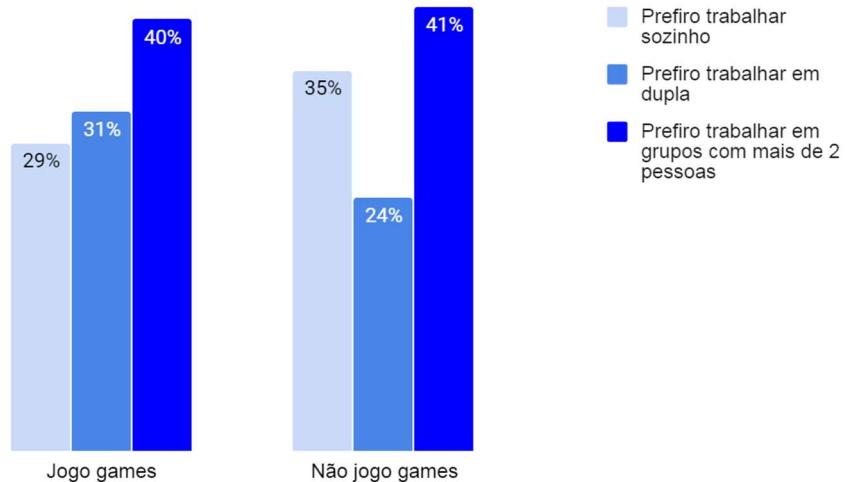
Fonte: reprodução da autora (2020).

Ao preencher os cartões coloridos com as cores que representam as suas respostas para as perguntas realizadas pela professora, os estudantes codificaram suas respostas cada um no seu cartão. Para utilizar o *dataset* formado pelos 45 cartões coloridos, realizou-se a decodificação das respostas contidas nos cartões, com a finalidade de apresentá-las em uma planilha *Google Sheets*.

No momento da decodificação dos dados na planilha eletrônica, houve a percepção de que seria interessante apresentar a tabela em dois formatos diferentes, sendo uma com dados univvalorados e a outra com dados multivvalorados. Os gráficos que foram gerados para testar a tabela do *dataset* estão a seguir (Gráficos 1, 2, e 3) e procuram responder às perguntas: Qual a sua preferência na formação de grupos para trabalhos escolares; Em qual horário o seu estudo rende mais?; Pra você, o que é mais importante entre as opções “aprender”, “as pessoas”, “tirar nota boa”, “outro”?

Gráfico 1: Gráfico de teste obtido para a pergunta de pesquisa “Qual a sua preferência na formação de grupos para trabalhos escolares?”.

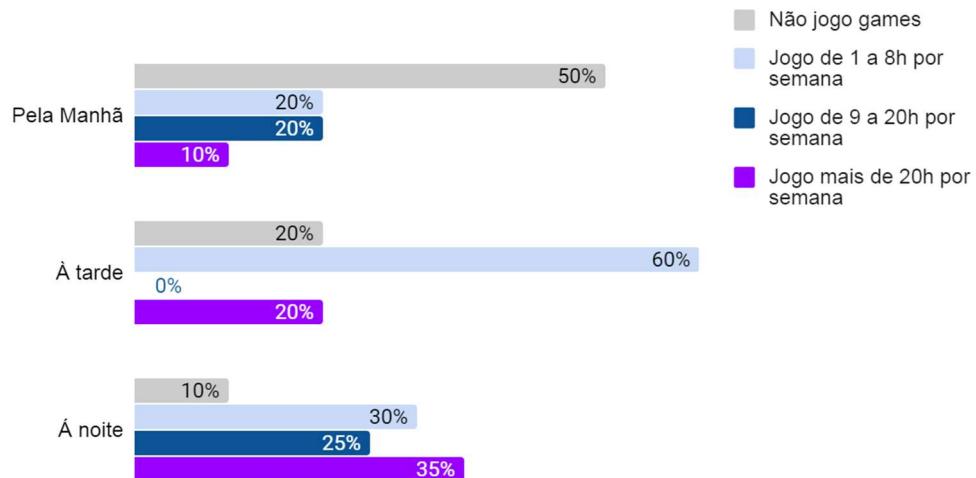
Qual a sua preferência na formação de grupos para trabalhos escolares?



Fonte: reprodução da autora (2020).

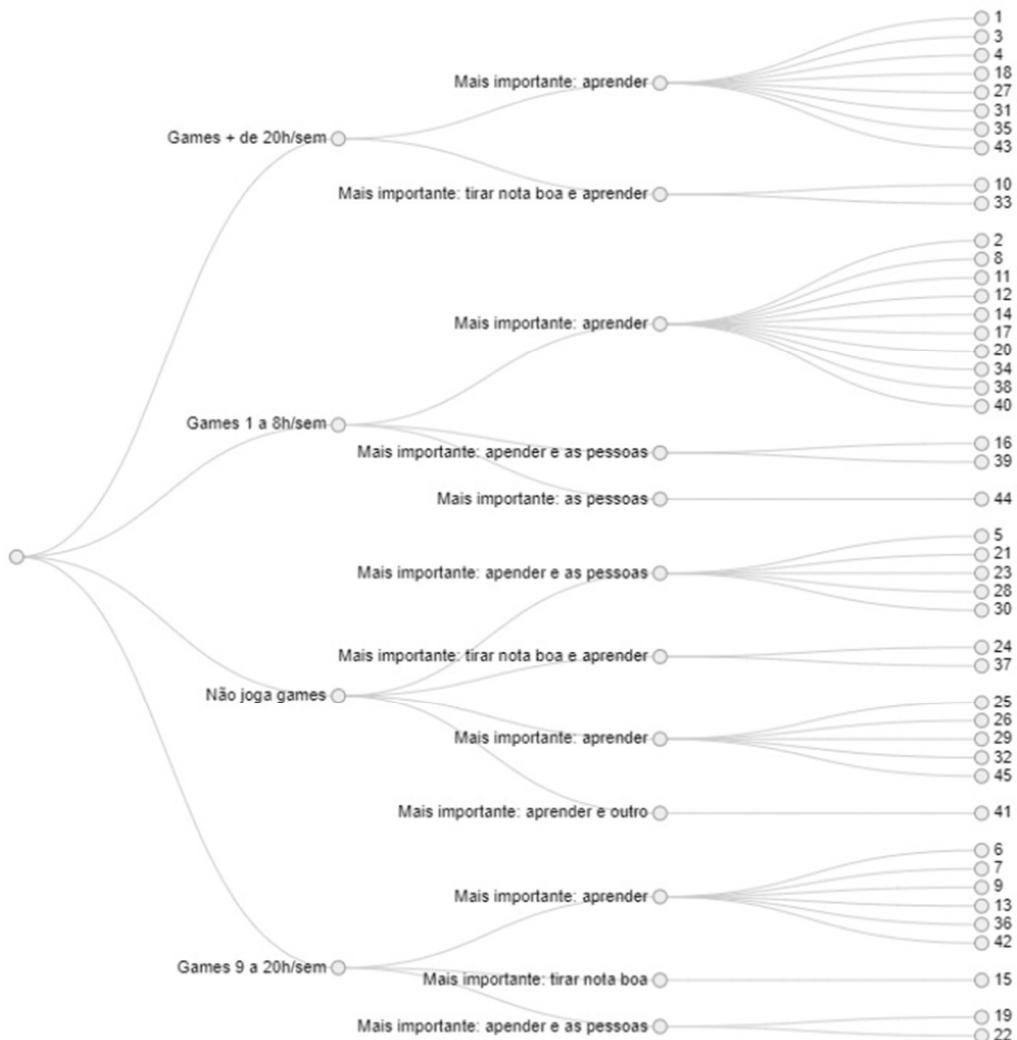
Gráfico 2: Gráfico de teste obtido para a pergunta de pesquisa “Em qual horário o seu estudo rende mais?”.

Em qual horário o seu estudo rende mais?



Fonte: reprodução da autora (2020).

Gráfico 3: Gráfico de teste obtido para a pergunta de pesquisa “O que é mais importante para você, entre as opções “aprender”, “as pessoas”, “tirar nota boa” e “outro”? ”



Fonte: reprodução da autora (2020).

Depois de criado o artefato destinado a solucionar o problema desta pesquisa, passou-se para a sua fase de avaliação, destinada a aplicar o artefato em um contexto real com o objetivo de verificar se ele respondia à pergunta de pesquisa e à realização das demais avaliações prescritas pela DSR.

#### 4.2.2 Atividade do primeiro ciclo

Para avaliar o artefato utilizado na pesquisa, foi realizado um estudo de caso exploratório com estudantes do primeiro período do curso de graduação em Ciência da Computação da Universidade Federal do Rio de Janeiro no dia 10 de outubro de 2019. A

amostra formada foi composta por 55 estudantes, entre eles, 7 mulheres e 48 homens divididos em 8 grupos, três deles com 7 estudantes e os demais com 5, 6, 8 e 9 estudantes cada (Figura 21). Os grupos foram formados por livre escolha dos participantes. O estudo de caso ocorreu em um laboratório de informática, com um computador da universidade disponível para cada estudante.

Figura 21: Estudantes da turma de graduação, divididos em grupos, utilizando o artefato Cartas para a Literacia mais o dataset Cores durante a aplicação da atividade em sala de aula.



Fonte: reprodução da autora (2020).

O estudo de caso foi proposto por meio de uma atividade em sala de aula, utilizando a dupla de artefatos “cartas para a literacia” e *dataset* “Cores”. Durante a atividade foi utilizada a técnica de observação direta com a participação de duas professoras e duas estudantes de graduação como pesquisadoras assistentes. A atividade continha cinco fases, destinadas a saber se os estudantes eram capazes de ler dados, analisar dados, argumentar

com base em dados e oferecer suas impressões sobre como a colaboração intragrupo foi capaz de impactar os resultados obtidos durante a atividade. A atividade foi realizada ao longo de 2 tempos de aula.

Com o objetivo de fornecer insumos para que os estudantes utilizassem a criatividade livremente para expressar sua Literacia de Dados na resolução dos desafios propostos na atividade, foram distribuídas 5 folhas de papel A4 para cada grupo e disponibilizado material de papelaria que incluía: notas autoadesivas e papéis sulfite coloridos, adesivos de diversos formatos e cores, alguns carimbos, massa de modelar, canetas esferográficas coloridas, hidrocor colorida, canetas marca-texto coloridas, lápis coloridos, giz de cera coloridos, cola colorida e lantejoulas coloridas, entre outros itens (Figura 22).

Figura 22: Materiais de papelaria disponibilizados para os estudantes durante a atividade proposta no estudo de caso.



Fonte: reprodução da autora (2020).

A atividade levou o tempo total de 60 minutos para ser concluída, tendo 10 minutos dedicados a resolver o desafio da fase 1, 10 minutos para a fase 2, 15 minutos para a fase 3, 20 minutos para a fase 4 e 5 minutos para a fase 5. A atividade transcorreu tranquilamente, exceto por uma interrupção no acesso à internet, que durou cerca de 15 minutos. Todos os estudantes se mantiveram com seus grupos até o final da atividade.

#### 4.2.3 Resultados do primeiro ciclo

Optou-se, em virtude da necessidade de delimitar o escopo deste trabalho, por apresentar os resultados da atividade obtidos através das cartas correspondentes às fases 4 e 5, por corresponderem aos desafios que propiciaram a entrega dos resultados através do uso de relatos realizados colaborativamente pelos estudantes, que foi a técnica explorada neste estudo com o objetivo de tornar explícitos os conhecimentos tácitos dos estudantes com relação à sua Literacia de Dados. Na seção 4.2.3.1 serão apresentados os resultados dos desafios apresentados nas fases 4 e 5 e na seção 4.2.3.2 serão apresentadas as anotações feitas durante a observação direta do comportamento dos estudantes durante a atividade realizada.

##### 4.2.3.1 Resultados das fases 4 e 5 da atividade com o artefato no primeiro ciclo

Na fase 4 da atividade com o artefato foi apresentada para a turma de graduação uma carta com o desafio: “Apresente uma descrição textual detalhada sobre as características da turma representada nas tabelas 1 e 2 com base nas análises feitas nas fases 1, 2 e 3”. Os resultados da fase 4 apresentados por cada um dos grupos que participaram do estudo de caso com a turma de graduação encontram-se nos quadros 3 a 10 a seguir.

Quadro 3: Resultados do grupo 1 para a Fase 4

##### **Grupo 1**

Aproximadamente metade da turma já conhecia alguém antes da 1<sup>a</sup> aula, a grande maioria tem como mais importante na faculdade o aprendizado, 71% das pessoas têm animais, [...] dentre esses animais 69% das pessoas têm cachorro. Na escolha de trabalhos 43% preferem grupos maiores que 2 pessoas e 25% preferem trabalhos solo. No lado de esparecer 88% preferem músicas, 73% jogos, 59% maratonas seriadas e 89% outros. Mais especificamente em jogos, o jogo mais jogado é *League of Legends* (LOL) e entre frequência há um equilíbrio entre a quantidade de pessoas que jogam muitas horas semanais e as que não jogam. Entre horário de estudo temos uma preferência maior por estudar de manhã com maior rendimento, os dados ficam com 45,5% de manhã, 43,2% de noite e 11,4% de tarde.

Fonte: reprodução da autora (2020).

Quadro 4: Resultados do grupo 2 para a Fase 4.

### **Grupo 2**

Ao analisarmos as tabelas disponibilizadas, podemos perceber que a turma no geral se importa mais na aprendizagem ao invés das notas. Percebe-se também que a maior parte da turma possui animais de estimação de preferência cachorros, além de gostarem de jogar games e escutar música nas horas vagas. Gostam também de realizar tarefas com alguma companhia e de estudar preferencialmente na parte da manhã e da noite, reservando as horas da tarde ao descanso.

Quantidade de alunos = 45

Preferem aprender = 43

Possuem animais = 32

Possuem cachorros = 22

Jogam games = 32

Escutam música = 39

Tarefas em companhia = 29

Estudar de manhã e noite = 40

Fonte: reprodução da autora (2020).

Quadro 5: Resultados do grupo 3 para a Fase 4.

### **Grupo 3**

Com base na análise de gráficos, podemos afirmar que pouco mais da metade da turma não se conhecia antes da primeira aula, tendo a maioria da turma priorizado o aprendizado sobre conhecer outras pessoas e tirar boas notas. Tendo a maioria deles melhor rendimento no período da manhã e noite (44,4% em cada).

Apesar de aproximadamente  $\frac{1}{3}$  da turma não ter animais de estimação, entre aqueles que

possuem, o cão é o mais comum. Podemos afirmar também que entre os hobbies mais comuns é jogar games e ouvir música, sendo o jogo mais jogado *league of legends*, e 88% da turma também têm outros hobbies não informados. Dentro dos que jogam games, a média de horas jogadas por semana encontra-se entre 9 e 20 horas.

Percebe-se também que o perfil dos alunos de CC [Ciência da Computação] é mais cooperativo, preferindo trabalhar em grupos maiores que 2, e a menor parcela dos alunos preferindo trabalho individual.

Fonte: reprodução da autora (2020).

Quadro 6: Resultados do grupo 4 para a Fase 4.

#### **Grupo 4**

Foi observado, entre os alunos que informaram qual game jogam, que a grande maioria dos alunos da turma jogam LOL, e com base na tabela 1, notamos por observação e contagem que essas conseguem render nos estudos mais a noite, além de pela tabela 2, boa parte deles não estudam na véspera da prova e que não gostam de ir à praia.

Em geral, uma boa parcela da turma já se conhecia antes da primeira aula, assim como, o mais importante para a maioria das pessoas da turma é aprender. É possível tirar das planilhas que a maioria das pessoas tem pelo menos um animal de estimação e que uma boa parte da turma prefere esparecer com praia, música, seriados e games bem como estudar em duplas e grupos.

Obs: Para a análise de que o pessoal que joga LOL não gosta de ir à praia (entre os que informaram) e que não estudam na véspera da prova, usamos o Libre Office, onde para analisar bem todas as informações dos que informaram que jogam LOL, na tabela 2 do dataset, copiamos as linhas referentes à estes alunos e colamos mais abaixo, da tabela original

Além de que foi utilizado o *Jupyter Notebook* para ajudar também nas análises da tabela 1.

Fonte: reprodução da autora (2020).

Quadro 7: Resultados do grupo 5 para a Fase 4.

### **Grupo 5**

A priori, para fornecer uma descrição textual detalhada sobre as características da turma, analisamos os gráficos correspondentes a cada atributo da tabela.

10% da turma considera importante tirar nota boa, em contraste com a maioria que discorda. Em comparação, 95% acha essencial aprender durante seu percurso na faculdade. Cerca de 22% dos alunos considera importante conhecer pessoas, podendo assim firmar conexões. Uma minoria tem outros objetivos em mente.

29% da turma não tem um animal de estimação, e dos que possuem, metade tem cachorro, 29% tem gato, 13% tem peixe ou pássaro e 13% tem outros, podendo ter mais de um tipo de animal.

Para descansar, um pouco menos da metade da turma gosta de ir a praia, cerca de 90% gosta de ouvir música, 20% gosta de dançar, cerca de metade gosta de praticar esportes, 60% gosta de maratonar séries, 75% gosta de jogar e a maioria também tem outras formas de relaxar.

Um pouco mais da metade vem estudando ao longo do período, em comparação aos que preferem estudar a noite, tendo uma minoria que prefere estudar a tarde. Ninguém cola e menos de 10% vão na sorte, de acordo com os dados.

Um pouco menos de  $\frac{1}{3}$  prefere trabalhar sozinho e na mesma proporção alguns preferem trabalhar em dupla. O resto prefere trabalhar em grupos maiores do que 4 alunos.

Apenas  $\frac{1}{3}$  da turma tem algum jogo favorito, com a maioria preferindo jogar LOL. Os outros  $\frac{2}{3}$  ou não possuem um jogo favorito ou não informaram. 25% da turma não joga, 25% joga de 1 a 8 horas por semana, 25% de 9 a 20 horas e 25% joga mais de 20 horas semanais.

Fonte: reprodução da autora (2020).

Quadro 8: Resultados do grupo 6 para a Fase 4.

### **Grupo 6**

- A maior parte da turma não possuía amizades antes do primeiro dia de aula.
- Menos da metade da turma já conhecia pelo menos uma pessoa antes do primeiro dia de aula.
- A maior parte da turma considera o aprendizado como o mais importante no curso. Logo, o conhecimento é algo primordial para os estudantes da mesma.
- Poucas pessoas consideram como mais importante conseguir conhecer pessoas novas. Logo, as relações interpessoais são subestimadas por boa parte dos estudantes da turma.
- A parte considerável da turma não possui animais de estimação.
- Os que possuem apenas um animal, em maioria possui um cachorro.
- Os que possuem mais de um, a maior parte possui cães e gatos.
- A turma se divide entre pessoas que adoram estudar em grupo e pessoas que estudam solo. Logo, a diferença acerca das preferências da metodologia de estudo são bastante opostas na turma.
- As exceções preferem as interseções entre as metodologias.
- 3\4 da turma se interessa por games. A maior parte dos que se interessam por games optaram por não informar suas preferências. O game LOL foi maioria nas preferências dos estudantes que decidiram optar por informar.
- Os dados referentes aos modos de esparrecer nos mostram que a maior parte da turma tem preferência por ir à praia ou escutar música.
- A maioria da turma não se interessa por dançar ou praticar esportes não digitais.
- Apenas 1\4 não se interessa por games.
- A turma ficou dividida entre pessoas que estudam melhor a noite e pessoas que estudam melhor de manhã.

Fonte: reprodução da autora (2020).

Quadro 9: Resultados do grupo 7 para a Fase 4.

### **Grupo 7**

Estabelecendo um perfil geral do grupo analisado, isto é, concluindo os dados por meio de um indivíduo hipotético que melhor representa a turma, é possível apontar esse “aluno(a) médio” como alguém mais produtivo em seus estudos pelas manhãs e noites. Ele gosta de vídeo game, tem 10 horas semanais diante de um aparelho para tal finalidade e prefere não revelar o seu título preferido. Em questão de trabalhos, lhe convêm mais em grupos acima de dois integrantes. Sobre animais de estimação, esse aluno possui um cão.

Fonte: reprodução da autora (2020).

Quadro 10: Resultados do grupo 8 para a Fase 4.

### **Grupo 8**

Analizando os dados da turma, percebemos que cerca de 55% dos alunos não conhecia ninguém do curso até entrar na universidade, a grande maioria não considera tirar uma boa nota como principal fator e 64,4% acha que o mais importante é aprender. Além disso 32% dos alunos possuem animais domésticos. desses 32, 22 tem cachorros, 13 tem gatos, 6 possuem pássaros ou peixes e 6 tem outros. 13 alunos não possuem animais domésticos.

Cerca de 25 alunos começam a estudar para uma prova durante o período e 20 começam uma semana antes. Uma porcentagem de 55,56% e 44,44% respectivamente. Cerca de 44% dos alunos consideram que o estudo rende mais pela manhã, essa mesma quantidade considera o estudo melhor a noite e somente 11% estuda melhor pela tarde.

Fonte: reprodução da autora (2020).

Na Fase 5 foi apresentada para a turma de graduação uma carta com o desafio: “Como a interação entre seu grupo impactou a análise dos dados e resultados obtidos?”. Nos quadros 11 a 18 a seguir estão transcritos os resultados para cada um dos grupos estudados no primeiro ciclo.

Quadro 11: Resultados do grupo 1 para a fase 5.

### **Grupo 1**

A interação entre o grupo impactou de forma positiva a análise e os resultados, com cada membro tendo ideias pertinentes e distintas. A troca de informações se tornou mais eficiente pela divisão de tarefas e pela utilização sinérgica das ferramentas digitais.

Fonte: reprodução da autora (2020).

Quadro 12: Resultados do grupo 2 para a fase 5.

### **Grupo 2**

A interação entre o grupo proporcionou uma análise mais efetiva e ampla dos dados analisados. Assim como incentivou o grupo a ser criativo, contribuindo na realização das tarefas de forma proativa.

Fonte: reprodução da autora (2020).

Quadro 13: Resultados do grupo 3 para a fase 5.

### **Grupo 3**

Cada um teve as suas interpretações para os dados analisados, promovendo uma discussão que contribuiu para uma melhor qualidade e precisão do trabalho. Além disso, a divisão de tarefas contribuiu muito para a eficiência do trabalho como um todo.

Fonte: reprodução da autora (2020).

Quadro 14: Resultados do grupo 4 para a fase 5.

### **Grupo 4**

A interação do grupo de 7 pessoas para a realização dessa atividade foi feita da seguinte forma. Nos separamos em 2 grupos (um de 3 e um de 4) que realizamos pesquisas meio-independentes, onde o resultado de ambas as pesquisas eram compartilhados de forma dinâmica por meio de um integrante que intermediava os grupos. Acreditamos que a liberdade e a independência de cada grupo foram essenciais para um bom desenvolvimento

do trabalho, pois permitiu que diversos pontos de vista para a resolução do problema fosse[m] explorado[s].

Fonte: reprodução da autora (2020).

Quadro 15: Resultados do grupo 5 para a fase 5.

### **Grupo 5**

Um trabalho que exige análise de dados possui várias fases e, por ser feito em grupo, necessita que haja uma interação interpessoal. O nosso grupo respondeu bem a essa necessidade de interação, obtendo os resultados esperados durante as etapas deste trabalho.

Fonte: reprodução da autora (2020).

Quadro 16: Resultados do grupo 6 para a fase 5.

### **Grupo 6**

Nós analisamos conjuntamente os dados tabelados e discutimos a melhor forma de representá-los criativamente. Nos dividimos em grupos (desenhistas, analisadores, etc.) para conseguirmos concluir os exercícios no tempo estimado.

Em resumo, a criação só foi possível por conta da sinergia dos membros e do compromisso em realizar a tarefa.

Fonte: reprodução da autora (2020).

Quadro 17: Resultados do grupo 7 para a fase 5.

### **Grupo 7**

O fluxo de ideias foi a principal moeda de troca entre os integrantes. Novos pensamentos foram lançados ao longo do debate. Assim, foi possível a produção de todos os dados apresentados nas fases anteriores. A grande maioria do conteúdo produzido foi realizada fora do meio digital, apesar da coleta ter sido quase inteiramente por meio do computador.

Fonte: reprodução da autora (2020).

Quadro 18: Resultados do grupo 8 para a fase 5.

**Grupo 8**

Todos os integrantes do grupo interagiram e participaram na análise, tornando fácil a obtenção de dados e agilizando as etapas da atividade. A interação impactou a análise das informações de forma positiva, pois a divisão das tarefas entre os integrantes do grupo tornou possível a execução dos desafios sem ultrapassar o tempo previsto.

Fonte: reprodução da autora (2020).

#### 4.2.3.2 Anotações realizadas durante a atividade com o artefato, durante o primeiro ciclo

No primeiro ciclo, a observação foi conduzida com a ajuda de duas pesquisadoras assistentes simultaneamente, cada uma fazendo suas anotações de forma individual, transitando entre os 8 grupos pesquisados. As anotações realizadas durante a fase 1 do estudo de caso são mostradas nos quadros 19, 20, 21, 22, 23, 24, 25 e 26.

Quadro 19: Notas de observação do grupo 1 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Concentrados, mas trabalhando separados</li> <li>- Focados, mas desunidos, atrasados</li> <li>- Quietos, atrasados, desunidos e sem criatividade</li> <li>- Quietos, atrasados</li> </ul>	<ul style="list-style-type: none"> <li>- Dividiram-se em duplas para analisar as tabelas (concentrados)</li> </ul>	

Quadro 20: Notas de observação do grupo 2 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Perdidos, dialogando</li> </ul>	<ul style="list-style-type: none"> <li>- Trabalham em</li> </ul>	<ul style="list-style-type: none"> <li>- Um integrante:</li> </ul>

<ul style="list-style-type: none"> <li>- Focados, mas desunidos</li> <li>- Atrasados</li> <li>- Usando a criatividade</li> <li>- Unidos</li> <li>- Quietos</li> </ul>	<p>conjunto dividindo suas informações (concentrados)</p>	<p>“[Dado] Pode ser qualquer coisa”</p> <p>- Outro integrante: “Mas o que a gente vai dar de exemplo de dado?”</p>
---	---	--

Fonte: reprodução da autora (2020).

Quadro 21: Notas de observação do grupo 3 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Desenvolvendo</li> <li>- Focados</li> <li>- Adiantados, já fazendo o gráfico</li> <li>- Usando pouca criatividade</li> <li>- Focados, dialogando</li> </ul>	<ul style="list-style-type: none"> <li>- Uniram-se em todas as etapas</li> <li>- Cooperação</li> </ul>	<ul style="list-style-type: none"> <li>- Não entenderam o pedido da fase a princípio.</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 22: Notas de observação do grupo 4 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Trocando informações e tirando dúvidas</li> <li>- Dispersos, olharam fotos no computador</li> <li>- Atrasados, demorando para começar</li> <li>- Discutindo entre si como fazer e como começar</li> <li>- Com dúvidas</li> </ul>	<ul style="list-style-type: none"> <li>- Por etapas</li> <li>- Cooperando entre as tabelas</li> <li>- Dados anotados, mas se dispersaram</li> </ul>	<ul style="list-style-type: none"> <li>- Não entenderam o pedido da fase a princípio</li> </ul>

<ul style="list-style-type: none"> <li>- Finalizando nos últimos segundos</li> <li>- Uma parte do grupo continua interagindo com material sobre outra matéria</li> </ul>		
--	--	--

Fonte: reprodução da autora (2020).

Quadro 23: Notas de observação do grupo 5 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Trabalhando em conjunto</li> <li>- Há liderança</li> <li>- Algumas pessoas estão perdidas</li> <li>- Focados</li> <li>- Unidos</li> <li>- Gráfico já sendo montado</li> <li>- Usando pouca criatividade</li> <li>- Algumas pessoas não estavam fazendo nada</li> <li>- Há liderança</li> <li>- Conversando entre si</li> </ul>	<ul style="list-style-type: none"> <li>- Dividiram-se por gênero</li> <li>- Cooperação</li> <li>- Concentração</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 24: Notas de observação do grupo 6 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Trabalhando em conjunto</li> <li>- Quase todo o grupo</li> </ul>	<ul style="list-style-type: none"> <li>- Dispersos</li> <li>- Alguns estavam fora do grupo</li> </ul>	<ul style="list-style-type: none"> <li>- Um integrante ao tentar tirar dúvida: "A gente pegou uma coluna e viu se a</li> </ul>

<ul style="list-style-type: none"> <li>- tentando desenhar</li> <li>- Dúvidas sobre como começar</li> <li>- Unidos</li> <li>- Adiantados e usando a criatividade</li> <li>- Gráfico virtual já sendo montado (adiantados)</li> <li>- Unidos</li> <li>- Conversando entre si</li> </ul>	<ul style="list-style-type: none"> <li>- Envolvidos</li> <li>- Concentrados!</li> <li>- Dados para o papel</li> </ul>	<ul style="list-style-type: none"> <li>- resposta era sim ou não”.</li> <li>- Outro integrante respondendo ao primeiro: “Pega um quadrado só!”</li> </ul>
--	---	---

Fonte: reprodução da autora (2020).

Quadro 25: Notas de observação do grupo 7 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Trabalhando em conjunto</li> <li>- Debatendo sobre o tema</li> <li>- Dispersos</li> <li>- Unidos</li> <li>- Uma pessoa desfocada, conversando sobre outras coisas</li> </ul>	<ul style="list-style-type: none"> <li>- Múltiplas opiniões</li> <li>- Trabalho em conjunto</li> <li>- Passaram os dados para o papel</li> </ul>	<ul style="list-style-type: none"> <li>- Não entenderam o pedido da fase a princípio</li> <li>- Na hora que faltou internet tentaram usar o LibreOffice, mas logo desistiram.</li> <li>- Um integrante falou: “O Sheets é muito melhor!”</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 26: Notas de observação do grupo 8 durante a fase 1 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Não estão discutindo tanto</li> <li>- Focados</li> </ul>	<ul style="list-style-type: none"> <li>- Organização e cooperação</li> <li>- Há liderança</li> </ul>	<ul style="list-style-type: none"> <li>- Não entenderam o pedido da fase a princípio e começaram a utilizar</li> </ul>

<ul style="list-style-type: none"> <li>- Unidos</li> <li>- Algumas pessoas não estão fazendo parte do debate</li> <li>- Focados, não estão usando a criatividade e não estão debatendo entre si</li> </ul>		a tabela dinâmica
--	--	-------------------

Fonte: reprodução da autora (2020).

Foram realizadas Anotações gerais para todos os grupos. A observadora 1 relatou que: de início, alguns estudantes estavam perdidos observando os outros fazendo o trabalho; Todos os grupos utilizaram o tempo dado para responder os desafios propostos; Em alguns grupos observou-se a predileção por gráficos de barras, enquanto que em outros observou-se a predileção por gráficos de setores; Nem todos os grupos utilizaram o material de papelaria para a criatividade e os grupos começaram a focar e acelerar suas ações de acordo com o desenvolvimento da atividade. De forma complementar, a observadora 3 relatou que durante a fase 1 a internet caiu, então alguns grupos prontamente abriram o Libre Office, enquanto que outros tentaram rotear a internet do celular e outros ainda aguardaram o retorno da internet. Alguns estudantes chegaram a verbalizar que a usabilidade do Google Sheets era muito melhor que a do Libre Office.

As anotações realizadas durante as fases 2 e 3 do estudo de caso foram tomadas conjuntamente e estão mostradas nos quadros 27, 28, 29, 30, 31, 32 e 34 a seguir.

Quadro 27: Notas de observação do grupo 1 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Pouca interação entre os membros</li> <li>- Pouco compartilhamento</li> </ul>	<ul style="list-style-type: none"> <li>- Um integrante: "Vai ter que contar mesmo, porque se usar...[o recurso de gráficos do sheets], vai dar mais que o</li> </ul>

	<ul style="list-style-type: none"> <li>- Gráfico de barras</li> <li>- Utilizaram pouco os materiais de papelaria</li> </ul>	total de alunos que existem"
--	---	------------------------------

Fonte: reprodução da autora (2020).

Quadro 28: Notas de observação do grupo 2 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Dividiram as funções</li> <li>- 1 pessoa isolada</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 29: Notas de observação do grupo 3 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Agrupados em conjunto</li> <li>- Agitados</li> <li>- Redondo</li> <li>- Distraídos</li> <li>- Pouco criativos</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 30: Notas de observação do grupo 4 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Atrasados, desunidos</li> <li>- Uma parte não fazendo nada e pesquisando sobre</li> </ul>	<ul style="list-style-type: none"> <li>- Compartilhando dados de produção</li> <li>- Anotando</li> </ul>	<ul style="list-style-type: none"> <li>- Os integrantes bateram palmas quando conseguiram gerar um gráfico de</li> </ul>

<ul style="list-style-type: none"> <li>- outras coisas</li> <li>- Devagar</li> </ul>	<ul style="list-style-type: none"> <li>- Calmos</li> <li>- Barra</li> <li>- Fazendo gráficos no papel colorido</li> </ul>	<ul style="list-style-type: none"> <li>- frequências</li> <li>- Um integrante comentou com outro: “cara, é a quantidade de períodos [nos quais as pessoas jogam] que é a frequência!”</li> <li>- Mudaram a cor do gráfico</li> <li>- Utilizaram um protótipo em papel (feito com um post it) para guiar a geração de gráficos virtuais</li> </ul>
--	---	---

Fonte: reprodução da autora (2020).

Quadro 31: Notas de observação do grupo 5 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Com dúvidas sobre o que fazer</li> <li>- Desunidos. Algumas pessoas não estavam fazendo nada</li> <li>- Uma pessoa só montando a resposta</li> <li>- Usando a criatividade</li> </ul>	<ul style="list-style-type: none"> <li>- Nervosos</li> <li>- Divididos</li> <li>- Compartilhando etapas</li> <li>- Anotando dados</li> <li>- Redondo</li> <li>- Trabalho pouco criativo</li> <li>- Utilizaram poucos recursos oferecidos</li> </ul>	<ul style="list-style-type: none"> <li>- “A gente não está conseguindo utilizar a planilha do LibreOffice”</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 32: Notas de observação do grupo 6 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Ainda desenhando</li> <li>- Usando a criatividade</li> <li>- Desenhando sobre</li> </ul>	<ul style="list-style-type: none"> <li>- Dividiram-se em dois grupos</li> <li>- Calmos e em equipe</li> <li>- Rápidos</li> <li>- Redondo</li> <li>- Não externalizaram a atividade</li> </ul>	<ul style="list-style-type: none"> <li>- Um integrante comentou com a observadora: “a gente vai fazer um gráfico com o desenho do animal e preencher”</li> <li>- Utilizaram a pesquisa por imagens do Google como referência para o esboço dos animais em post its.</li> <li>- Estavam trabalhando também com gráficos de barras</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 33: Notas de observação do grupo 7 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Agrupados</li> <li>- Agitados</li> <li>- Produtivos</li> <li>- Concentrados</li> <li>- Redondo</li> <li>- Desenhos de animais coloridos</li> <li>- Organização das ideias através das cores</li> <li>- Inclusivos</li> </ul>	<ul style="list-style-type: none"> <li>- Utilizaram Python para fazer cálculos simples. Estavam com o <i>shell</i> do python aberto.</li> <li>- Um integrante, ao perceber que estavam sendo observados, comentou com os outros: “mas não é melhor usar a calculadora”?</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 34: Notas de observação do grupo 8 durante as fases 2 e 3 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Usando os próprios meios tecnológicos para achar inferências em equipe</li> <li>- Redondo</li> <li>- Anotações com destaque em rosa</li> <li>- Utilizaram pouco os recursos de papelaria</li> </ul>	<ul style="list-style-type: none"> <li>- Estavam usando tabelas dinâmicas</li> <li>- Na fase 3 utilizaram fórmulas no LibreOffice para calcular “qual game joga”</li> </ul>

Fonte: reprodução da autora (2020).

Os quadros 35, 36, 37, 38, 39, 40, 41 e 42 a seguir mostram as anotações realizadas durante a fase 4 do primeiro ciclo do estudo de caso.

Quadro 35: Notas de observação do grupo 1 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Lentos! divididos em etapas anteriores</li> <li>- Texto feito, metade do grupo integrada</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 36: Notas de observação do grupo 2 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Montando o texto em grupo</li> <li>- Rápidos</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 37: Notas de observação do grupo 3 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Adiantados</li> <li>- Focados</li> <li>- Conversando entre si</li> <li>- Fazendo</li> </ul>	<ul style="list-style-type: none"> <li>- Montando o texto em grupos, alguns dispersos</li> <li>- Texto sendo montado a partir da análise dos gráficos, simultaneamente</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 38: Notas de observação do grupo 4 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Desunidos</li> <li>- Apenas 3 pessoas do grupo estavam conversando sobre a atividade</li> <li>- Fazendo outras coisas de disciplinas diferentes</li> <li>- Conversando sobre outras coisas</li> <li>- Uma parte do grupo fazendo a atividade e outra continua pesquisando sobre outra coisa</li> </ul>	<ul style="list-style-type: none"> <li>- Texto no papel</li> <li>- Rápidos</li> <li>- Detalhistas</li> <li>- Objetivos</li> <li>- PC</li> </ul>	<ul style="list-style-type: none"> <li>- Usaram um <i>Jupyter Notebook</i> para tentar realizar análises</li> <li>- Ao serem observados, um integrante comenta com a observadora 3: “Não deu certo isso não, professora, esquece isso!”</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 39: Notas de observação do grupo 5 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Quietos</li> <li>- Usando a criatividade</li> <li>- Unidos</li> <li>- Conversando entre si sobre liderança</li> <li>- Uma parte não fazendo nada (2 pessoas)</li> </ul>	<ul style="list-style-type: none"> <li>- Texto no papel, compartilhamento de dados</li> <li>- Dividiram-se em 2 grupos</li> <li>- Tensos</li> <li>- Pouco criativos</li> <li>- Objetivos</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 40: Notas de observação do grupo 6 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
	<ul style="list-style-type: none"> <li>- Em equipe</li> <li>- Decisões em conjunto</li> <li>- proposta</li> <li>- Ainda no gráfico</li> </ul>	<ul style="list-style-type: none"> <li>- O estudante que possui baixa visão olha informações muito pertinho do celular e está totalmente integrado ao grupo. Orienta as ações junto ao colega que escreve no computador. Está com uma régua na mão em postura de liderança.</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 41: Notas de observação do grupo 7 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Unidos</li> <li>- Alguns conversando sobre o trabalho</li> <li>- Alguns sem falar nada</li> <li>- Dispersos</li> <li>- Unidos</li> <li>- Perdidos</li> </ul>	<ul style="list-style-type: none"> <li>- Trabalhando a externalização ainda</li> <li>- 1 pessoa escrevendo</li> <li>- Externalizando com números e imagens coloridas</li> <li>- Preocupados com a apresentação dos dados coletados</li> </ul>	

Fonte: reprodução da autora (2020).

Quadro 42: Notas de observação do grupo 8 durante a fase 4 do primeiro ciclo do estudo de caso.

Observadora 1	Observadora 2	Observadora 3
<ul style="list-style-type: none"> <li>- Desunidos</li> <li>- Focados</li> <li>- Sem criatividade</li> <li>- Unidos</li> <li>- Quietos</li> <li>- Perdidos</li> </ul>	<ul style="list-style-type: none"> <li>- Liderança</li> <li>- Cooperação</li> <li>- Análise por etapa</li> <li>- 1 pessoa escrevendo</li> <li>- questionadores, mas não muito produtivos em criação</li> </ul>	

Fonte: reprodução da autora (2020).

Durante a fase 5 do estudo de caso foram realizadas as anotações que constam nos quadros 43, 44, 45, 46, 47, 48, 49 e 50 a seguir. As anotações foram tomadas, nesta fase, por apenas duas das 3 observadoras.

Quadro 43: Notas de observação do grupo 1 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
<ul style="list-style-type: none"> <li>- Quietos</li> <li>- Desunidos</li> <li>- Devagar</li> </ul>	<ul style="list-style-type: none"> <li>- Só 2 pessoas fazendo (2/6)</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 44: Notas de observação do grupo 2 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
<ul style="list-style-type: none"> <li>- Unidos</li> <li>- Focados</li> <li>- Atrasados</li> </ul>	<ul style="list-style-type: none"> <li>- Todos colaborando (7/7)</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 45: Notas de observação do grupo 3 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
	<ul style="list-style-type: none"> <li>- Contribuição de 3/6 na análise</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 46: Notas de observação do grupo 4 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
	<ul style="list-style-type: none"> <li>- Análise feita por 4/7 do grupo</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 47: Notas de observação do grupo 5 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
	<ul style="list-style-type: none"> <li>- 2 fazendo a análise em grupo</li> <li>- 4 dispersos</li> <li>- (2/4)</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 48: Notas de observação do grupo 6 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
<ul style="list-style-type: none"> <li>- Conversando entre si</li> <li>- Criativos</li> </ul>	<ul style="list-style-type: none"> <li>- 5 colaborando para análise do grupo</li> <li>- 2 salvando os arquivos</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 49: Notas de observação do grupo 7 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
	<ul style="list-style-type: none"> <li>- 1 escrevendo</li> <li>- 1 colaborador</li> <li>- 5 analisando os dados antes de enviar</li> </ul>

Fonte: reprodução da autora (2020).

Quadro 50: Notas de observação do grupo 8 durante a fase 5 do primeiro ciclo do estudo de caso.

<b>Observadora 1</b>	<b>Observadora 2</b>
	<ul style="list-style-type: none"> <li>- 1 pessoa escrevendo</li> <li>- O líder colaborando (lago) (2/9)</li> </ul>

Fonte: reprodução da autora (2020).

As anotações obtidas no primeiro ciclo do estudo de caso serão utilizadas em conjunto com os resultados do questionário e as respostas dos desafios da atividade proposta durante a análise de dados, na discussão dos resultados.

#### **4.2.4 Discussão do primeiro ciclo**

Neste tópico será apresentada a análise e discussão das fases 4 e 5 da atividade proposta em sala de aula no primeiro ciclo do estudo de caso. Na seção 4.2.4.1 será apresentada a discussão relativa à análise dos dados obtidos através das respostas dos estudantes à pergunta da carta da fase 4 da atividade do primeiro ciclo do estudo de caso. Na seção 4.2.4.2 será apresentada a discussão relativa à análise dos dados obtidos através das respostas dos estudantes à pergunta da carta da fase 5 da atividade do primeiro ciclo do estudo de caso.

A análise das fases 4 e 5 da atividade com as cartas seguiu as seguintes etapas: 1) Codificação inicial, 2) Codificação axial e 3) Saturação. A codificação inicial foi feita incidente por incidente, através da leitura e interpretação do texto escrito pelos grupos participantes da atividade, buscando identificar as situações vivenciadas pelos grupos durante a atividade de análise de dados. Os incidentes foram identificados principalmente através da percepção do emprego de formas verbais regulares e verbos de ligação. A Codificação axial foi realizada buscando identificar as ações mais globais que representavam as situações narradas pelos estudantes e sistematizadas na codificação inicial. Na saturação foram codificadas as categorias que se revelaram salientes na codificação axial.

##### **4.2.4.1 Discussão da fase 4 do primeiro ciclo**

Na análise dos resultados da fase 4 para a turma de graduação, durante a codificação inicial, foram obtidos os seguintes códigos listados no Quadro 51, a seguir:

Quadro 51: Categorias encontradas durante a codificação inicial das respostas da fase 4 no primeiro ciclo do estudo de caso.

Lendo dados
Fazendo comparação por meio de porcentagem

Calculando frequências
Analizando tabelas
Analizando gráficos
Identificando interseções
Utilizando linguagem de programação para a análise de dados
Utilizando planilha eletrônica para a análise de dados
Representando resultados por meio de porcentagem
Representando resultados por meio de expressões partitivas
Representando resultados por meio de frequências
Representando resultados por meio de frações
Representando resultados com expressões indefinidas
Representando resultados por meio de subconjuntos com expressões partitivas
Representando resultados por meio de subconjuntos com porcentagens
Representando resultados por meio de medidas aproximadas
Representando todos os subconjuntos de um conjunto através de porcentagens
Representando resultados por meio de subconjuntos com frequências
Representando resultados por meio de subconjuntos de tamanho indeterminado
Escrevendo em forma de tópicos

Escrevendo em forma de texto
Argumentando com proposições condicionais
Relacionando teoria com dados

Fonte: reprodução da autora (2020).

Na realização da etapa de codificação axial, foram identificadas novas categorias a partir dos códigos obtidos na codificação inicial. As categorias identificadas na codificação axial podem ser lidas no Quadro 52, a seguir:

Quadro 52: Categorias encontradas na codificação axial relativa aos dados da fase 4 da atividade com a turma de graduação.

Codificação inicial	Codificação axial
Lendo dados	Lendo dados
Lendo dados	
Fazendo comparação por meio de porcentagem	Comparando dados
Calculando frequências	agregando dados
Analizando tabelas	
Analizando gráficos	
Identificando interseções	Analizando dados
Utilizando linguagem de programação para a análise de dados	
Utilizando planilha eletrônica para a análise de dados	Analizando dados com ajuda de software
Representando resultados por meio de porcentagem	Representando resultados

Representando resultados por meio de expressões partitivas	
Representando resultados por meio de frequências	
Representando resultados por meio de frações	
Representando resultados com expressões indefinidas	
Representando resultados por meio de subconjuntos com expressões partitivas	
Representando resultados por meio de subconjuntos com porcentagens	
Representando resultados por meio de medidas aproximadas	
Representando todos os subconjuntos de um conjunto através de porcentagens	
Representando resultados por meio de subconjuntos com frequências	
Representando resultados por meio de subconjuntos de tamanho indeterminado	
Escrevendo em forma de tópicos	
Escrevendo em forma de texto	
Argumentando com proposições condicionais	Argumentando com dados
Relacionando teoria com dados	

Fonte: reprodução da autora (2020).

Na fase de saturação, as categorias encontradas durante a codificação axial foram abstraídas em categorias mais gerais o quanto possível. Dessa forma, foram obtidos os

seguintes códigos finais: Lendo dados, Analisando dados, Informando resultados e Argumentando com dados. Um panorama das etapas de codificação pode ser obtido através do Quadro 53, a seguir.

Quadro 53: Etapas da codificação dos resultados da fase 4 no primeiro ciclo do estudo de caso.

Codificação inicial	Codificação axial	Saturação
Lendo dados em planilhas	Lendo dados	Lendo dados
Fazendo comparação por meio de porcentagem	Comparando dados	
Calculando frequências	agregando dados	
Analizando tabelas		
Analizando gráficos		
Identificando interseções	Analizando dados	Analizando dados
Utilizando linguagem de programação para a análise de dados		
Utilizando planilha eletrônica para a análise de dados	Analizando dados com ajuda de software	
Representando resultados por meio de porcentagem		
Representando resultados por meio de expressões partitivas		
Representando resultados por meio de frequências		
Representando resultados por meio de frações		
Representando resultados com expressões indefinidas	Representando resultados	Informando resultados
Representando resultados por meio de subconjuntos com expressões partitivas		
Representando resultados por meio de subconjuntos com porcentagens		
Representando resultados por meio de medidas		

aproximadas		
Representando todos os subconjuntos de um conjunto através de porcentagens		
Representando resultados por meio de subconjuntos com frequências		
Representando resultados por meio de subconjuntos de tamanho indeterminado		
Escrevendo em forma de tópicos		
Escrevendo em forma de texto		
Argumentando com proposições condicionais	Argumentando com dados	Argumentando com dados
Relacionando teoria com dados	Argumentando com dados	Argumentando com dados

Fonte: reprodução da autora (2020).

#### 4.2.4.2 Discussão da fase 5 do primeiro ciclo

A análise das respostas dos estudantes da turma de graduação à pergunta da fase 5 – “Como a interação entre seu grupo impactou a análise dos dados e os resultados obtidos?” – foi realizada com base na abordagem construtivista da teoria fundamentada em dados, assim como a análise das respostas dos estudantes para a pergunta realizada na fase 4. Os códigos iniciais, as categorias axiais e a saturação obtidos através das respostas dos estudantes podem ser lidos no Quadro 54, a seguir:

Quadro 54: Codificação obtida através das respostas dos estudantes à pergunta realizada na fase 5 do desafio, realizada de acordo com abordagem construtivista da Teoria Fundamentada em Dados.

Inicial	Axial	Saturação
Trocando informações de forma eficiente	Comunicando com o grupo	Comunicando
Promovendo discussão	Comunicando com o grupo	Comunicando
Interação agilizando as etapas da atividade	Comunicando com o grupo	Comunicando

Interagindo em grupo	Interagindo em grupo	Comunicando
Interagindo em grupo	Interagindo em grupo	Comunicando
Interagindo em grupo	Interagindo em grupo	Comunicando
Interagindo	Interagindo em grupo	Comunicando
Interagindo em grupo	Interagindo em grupo	Comunicando
Respondendo à necessidade de interação	Interagindo em grupo	Comunicando
Ideias fluindo	Interagindo em grupo	Comunicando
Utilizando o fluxo de ideias como moeda de troca entre os integrantes	Interagindo em grupo	Comunicando
Lançando novos pensamentos ao longo do debate	Interagindo em grupo	Comunicando
Pluralidade de ideias	Pluralidade de ideias	Comunicando
Explorando diversos pontos de vista	Pluralidade de ideias	Comunicando
Explorando diversos pontos de vista através da liberdade e independência de pensamento	Pluralidade de ideias	Comunicando
Tendo ideias pertinentes e distintas individualmente	Tendo ideias pertinentes e distintas individualmente	Comunicando
Compartilhando resultados internamente ao grupo	Compartilhando resultados com o grupo	Comunicando
Compartilhando resultados internamente ao grupo de forma dinâmica	Compartilhando resultados com o grupo	Comunicando
Compartilhando resultados entre os	Compartilhando resultados	Comunicando

subgrupos por meio de um integrante	entre subgrupos	
Produzindo conteúdo fora do meio digital	Produzindo conteúdo em grupo fora do meio digital	Comunicando
Apresentando dados para cada uma das fases	Apresentando resultados para cada fase	Comunicando
Interagindo em grupo de forma a impactar positivamente a análise e os resultados	Interação impactando no resultado do trabalho realizado	Comunicando
Interagindo com o grupo de forma a proporcionar uma análise mais efetiva e ampla dos dados analisados	Interação impactando no resultado do trabalho realizado	Comunicando
Interagindo em grupo de forma a incentivar a ser criativo	Interação impactando na criatividade do grupo	Comunicando
Interagindo em grupo de forma a contribuir para a proatividade	Interação impactando na proatividade do grupo	Comunicando
Alcançando uma melhor qualidade e precisão do trabalho através da discussão	Interação impactando na qualidade e precisão do trabalho	Comunicando
Interação facilitando a obtenção de dados	Interação facilitando a obtenção de dados	Comunicando
Interação facilitando a divisão de tarefas	Interação facilitando a divisão de tarefas	Comunicando
Dividindo tarefas entre os integrantes do grupo	Dividindo tarefas	Cooperando
Realizando pesquisas de forma meio	Pesquisando em subgrupos	Cooperando

independente por meio de subgrupos		
Desenvolvendo um bom trabalho através da liberdade e independência dos subgrupos	Desenvolvendo trabalho de forma livre e independente em subgrupos	Cooperando
Realizando tarefas	Realizando tarefas em grupo	Cooperando
Resolvendo problemas	Resolvendo problemas em grupo	Cooperando
Trabalhando em grupo	Trabalhando em grupo	Cooperando
Facilitando a criação através do compromisso do grupo	Atuando de forma comprometida com o grupo	Cooperando
Realizando a tarefa	Realizando tarefas em grupo	Cooperando
Coletando dados por meio do computador	Coletando dados em grupo	Cooperando
Obtendo dados	Coletando dados em grupo	Cooperando
Participação facilitando a obtenção de dados	Obtendo dados através da participação	Cooperando
Interpretando dados de forma independente	Interpretando dados	Cooperando
Analizando dados	Analizando dados em grupo	Cooperando
Analizando os dados tabelados conjuntamente	Analizando dados tabelados em grupo	Cooperando
Analizando conjuntamente	Analizando dados em grupo	cooperando
Participando da análise em grupo	Analizando dados em grupo	Cooperando

Utilizando ferramentas digitais de forma sinérgica	Utilizando ferramentas digitais sinericamente	Cooperando
Obtendo os resultados esperados	Obtendo resultados em grupo	Cooperando
Facilitando a criação através da sinergia do grupo	Criando em sinergia com o grupo	Cooperando
Produzindo informações	Produzindo conteúdo em grupo	Cooperando
Produzindo conteúdo	Produzindo conteúdo em grupo	Cooperando
Dividindo tarefas	Dividindo tarefas	Coordenando
Dividindo tarefas	Dividindo tarefas	Coordenando
Alcançando eficiência no trabalho através da divisão de tarefas	Alcançando eficiência no trabalho através da divisão de tarefas	Coordenando
Separando o grupo em subgrupos	Dividindo-se em subgrupos	Coordenando
Intermediando os subgrupos por meio de um dos integrantes do grupo	Mediando os subgrupos	Coordenando
Trabalhando em diversas etapas	Trabalhando em etapas	Coordenando
Trabalhando em etapas	Trabalhando em etapas	Coordenando
Discutindo a melhor forma de representar os dados	Planejando a representação de resultados	Coordenando
Dividindo-se em subgrupos por competências	Dividindo-se em subgrupos por competências	Coordenando
Dividindo-se em subgrupos para otimizar o tempo	Dividindo-se em subgrupos para otimizar o tempo	Coordenando

Participação agilizando as etapas da atividade	Otimizando o tempo	Coordenando
Executando os desafios no tempo previsto devido a divisão de tarefas	Otimizando o tempo através da divisão de tarefas	Coordenando
Executando os desafios sem ultrapassar o tempo previsto	Trabalhando dentro do tempo previsto	Coordenando

Fonte: reprodução da autora (2020).

Durante a codificação axial foi percebido que alguns códigos indicavam uma relação com conceitos relacionados ao modelo de colaboração 3C de Fuks *et al.* (2011), que analisa a colaboração com base em 3 dimensões: comunicação, coordenação e cooperação.

A partir dos resultados obtidos no primeiro ciclo do estudo de caso, em que foi criado o artefato para a resolução do problema de pesquisa, determinou-se que seria necessário um segundo ciclo de estudo de caso com o objetivo de observar se as conjecturas teóricas levantadas durante a concepção do artefato são válidas e se é possível responder à pergunta desta pesquisa: “Como podemos avaliar os conhecimentos em Literacia de Dados de estudantes?”.

#### 4.3 SEGUNDO CICLO: PÓS-GRADUAÇÃO

O segundo ciclo de pesquisa foi realizado com uma turma de estudantes de pós-graduação em Informática da Universidade Federal do Rio de Janeiro. O método de pesquisa utilizado foi um estudo de caso, tal qual empregou-se no primeiro ciclo. O objetivo deste segundo ciclo foi verificar se as conjecturas levantadas na fase de *design* do artefato permaneceriam válidas para um outro grupo de estudantes, levantar possibilidades de melhorias a serem realizadas no artefato e possibilitar uma reflexão mais aprofundada acerca da resolução do problema de pesquisa. Escolheu-se a nova amostra de estudantes entre alunos de uma turma de pós-graduação com a intenção de oportunizar maior detalhamento nas respostas aos desafios, além de um maior alcance na amplitude do conhecimento de Literacia de Dados externalizado. Na seção 4.3.1 é descrita a fase de planejamento do segundo

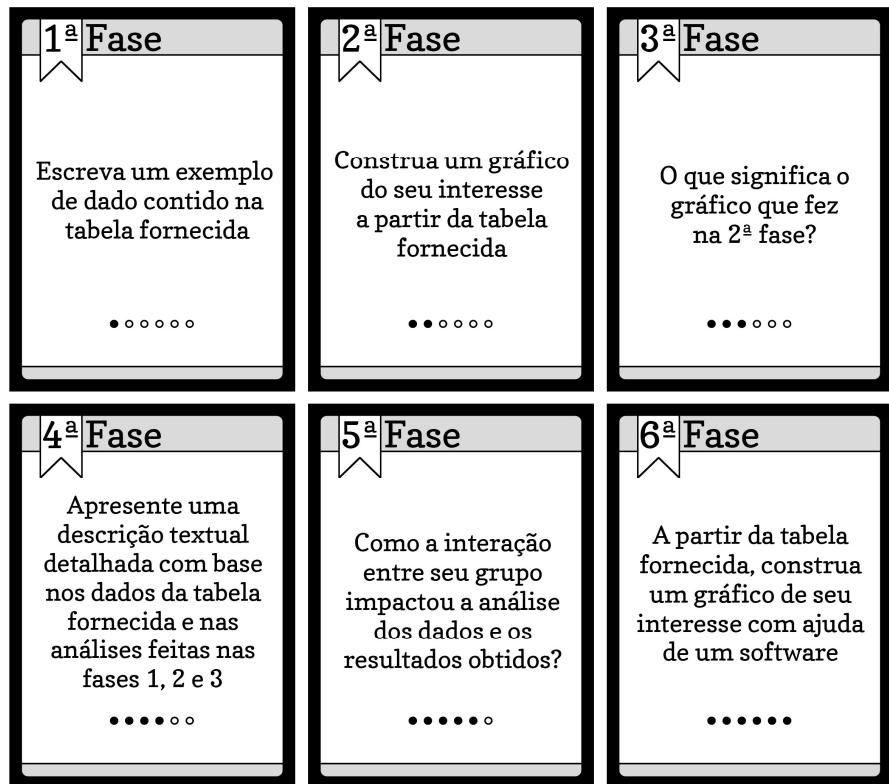
ciclo do estudo de caso, quais e como foram implementadas as adaptações com objetivo de melhorar a performance do artefato. Na seção 4.3.2 são apresentados os resultados obtidos para os desafios das fases 4 e 5 do artefato e as observações diretas realizadas pela equipe de pesquisa durante a atividade.

#### 4.3.1 Planejamento do segundo ciclo

No segundo ciclo foram realizadas melhorias no artefato, mediante os resultados do primeiro ciclo. Para a aplicação do artefato no segundo ciclo do estudo de caso, modificou-se a escrita da carta da fase 2 do conjunto a fim de que sua solicitação apresentasse uma escrita que direcionasse qualquer usuário para o uso do *dataset* – e não somente os estudantes que criaram o *dataset* utilizado na atividade. Também foi adicionada uma sexta carta às cartas de desafio projetadas no primeiro ciclo, com o objetivo de fazer com que os estudantes respondessem à solicitação feita no desafio da fase 2 obrigatoriamente com o uso de um software de livre escolha – que no primeiro ciclo poderia ser solucionado com o auxílio do material de papelaria e/ou de softwares quaisquer.

As cartas de desafio apresentaram então 6 desafios propostos para a turma de pós-graduação, tal como mostra a Figura 23. Os desafios propostos para a turma de pós-graduação foram: 1- escrever um exemplo de dado contido em uma das duas tabelas fornecidas; 2- escolher uma das tabelas e construir um gráfico do seu interesse; 3- explicar o significado do gráfico que fez na 2<sup>a</sup> fase; 4- apresentar uma descrição textual detalhada sobre as características da turma representada nas Tabelas 1 e 2, com base nas análises feitas nas fases 1, 2 e 3; 5- apresentar uma descrição textual sobre como a interação entre seu grupo impactou a análise dos dados e os resultados obtidos; 6- com ajuda de um software, escolher uma das tabelas e construir pelo menos 1 gráfico de seu interesse (Figura 23). A fim de prover meios para que os estudantes pudessem responder os desafios propostos nas cartas, foi utilizado o mesmo *dataset* do primeiro ciclo.

Figura 23: Segunda versão das cartas de desafio, preparadas para serem aplicadas no segundo ciclo do estudo de caso, realizado com a turma de pós-graduação.



Fonte: reprodução da autora (2020).

#### 4.3.2 Atividade do segundo ciclo

Para avaliar o artefato no segundo ciclo, foi realizado um estudo de caso exploratório com estudantes do programa de pós-graduação em Informática da Universidade Federal do Rio de Janeiro no dia 5 de novembro de 2019. A amostra formada foi composta por 13 estudantes, entre eles, 3 mulheres e 10 homens divididos em 4 grupos, três deles com 3 estudantes e o restante com 4 estudantes cada. Os grupos foram formados por livre escolha dos participantes, influenciados pela localização que ocupavam na sala de aula no início da atividade. O estudo de caso ocorreu em uma sala de aula convencional e os estudantes disponibilizaram 4 notebooks de sua propriedade para uso pela turma, tendo cada grupo utilizado um desses computadores para a realização da proposta apresentada.

O estudo de caso foi proposto por meio de uma atividade em sala de aula, utilizando o artefato “Cartas para a Literacia” e o *dataset* “Cores”. Durante a atividade foi utilizada a técnica de observação direta com a participação de duas professoras e duas estudantes de graduação como pesquisadoras assistentes. A atividade continha seis fases, destinadas a saber

se os estudantes eram capazes de ler dados, analisar dados, analisar dados com ajuda de um software de livre escolha, argumentar com base em dados e oferecer suas impressões sobre como a colaboração intragru po foi capaz de impactar os resultados do grupo formado pelos estudantes para a atividade. A atividade foi realizada ao longo de 2 tempos de aula.

Com o objetivo de fornecer insumos para que os estudantes utilizassem a criatividade livremente para expressar sua literacia na resolução dos desafios propostos na atividade, foram distribuídas 5 folhas de papel A4 para cada grupo e disponibilizado material de papelaria que incluía: notas autoadesivas coloridas, papéis sulfite coloridos, adesivos de diversos formatos e cores, alguns carimbos, massa de modelar, canetas esferográficas coloridas, hidrocor colorida, canetas marca-texto coloridas, lápis coloridos, giz de cera coloridos, cola colorida e lantejoulas coloridas — mesmo material que foi disponibilizado durante a atividade aplicada no primeiro ciclo de avaliação, com a turma de graduação.

A atividade levou o tempo total de 70 minutos para ser concluída, tendo 10 minutos dedicados a resolver o desafio da fase 1, 10 minutos para a fase 2, 15 minutos para a fase 3, 20 minutos para a fase 4, 5 minutos para a fase 5 mais 10 minutos para a sexta fase. A atividade transcorreu tranquilamente e todos os estudantes se mantiveram com seus grupos até o final.

#### **4.3.3 Resultados do segundo ciclo**

Assim como no primeiro ciclo, optou-se, em virtude da necessidade de delimitar o escopo deste trabalho, por apresentar os resultados da atividade do segundo ciclo obtidos através das cartas correspondentes às fases 4 e 5, por corresponderem aos desafios que propiciaram a entrega dos resultados através do uso de *group storytelling*, que foi a técnica utilizada neste estudo com o objetivo de tornar explícitos os conhecimentos tácitos dos estudantes com relação à sua Literacia de Dados. Na seção 4.3.3.1 serão apresentados os resultados dos desafios apresentados nas fases 4 e 5 e na seção 4.3.3.2 serão apresentadas as anotações feitas durante a observação direta do comportamento dos estudantes durante a atividade realizada no segundo ciclo.

#### 4.3.3.1 Resultados das fases 4 e 5 da atividade no segundo ciclo

Na fase 4 da atividade com o artefato foi apresentada para a turma de pós-graduação uma carta com o desafio: “Apresente uma descrição textual detalhada sobre as características da turma representada nas Tabelas 1 e 2 com base nas análises feitas nas fases 1, 2 e 3”. Os resultados da fase 4 apresentados por cada um dos grupos que participaram do estudo de caso com a turma de pós-graduação encontram-se no Quadro 55 e os resultados da fase 5 encontram-se no Quadro 56.

Quadro 55: Respostas da turma de pós-graduação à fase 4 da atividade.

##### Grupo 1

Se escolhêssemos um aluno dessa turma, o mais provável que encontrássemos:

- Alguém que não conhecia outro aluno antes da 1<sup>a</sup> aula (56%)
- Citou que considera aprender como o mais importante (95%) entre notas, aprendizado e pessoas. (Poderiam escolher mais de uma alternativa)
- Já vinha estudando ao longo do período (56%)
- Tem um animal de estimação (71%) e ele é um cão (69%)
- Alguém prefere espairecer ouvindo música (87%)
- Não vai colar na prova (Ninguém respondeu que iria!)
- Prefere fazer trabalhos em grupo com mais de 2 integrantes (47%)
- Joga games de 1 a 8 horas por semana (29%)
- Prefere estudar de manhã ou à noite (empate de 44% cada)

##### Grupo 2

Analizando os dados que observamos que são dados de alunos

Nesses dados observamos que:

- Quem não conhecia ninguém antes da 1<sup>a</sup> aula prefere trabalhar em grupos > 2 do que quem conhecia alguém antes da 1<sup>a</sup> aula.

- Em ambos os grupos (quem conhecia alguém antes da 1<sup>a</sup> aula e quem não conhecia) os alunos que preferem trabalhar sozinho preferem jogos individualistas (como Zelda)

#### Grupo 3

A Tabela 1 trata-se de um questionário estruturado acerca do comportamento dos indivíduos que os responderam. São feitas perguntas de cunho psicográfico, como gosto e momentos de lazer/estudo. As respostas presentes nesta tabela são de difícil tratamento, visto que precisam de uma excessiva limpeza dos dados. Esta tabela, devido a sua estrutura visual, permite uma fácil leitura humana, sendo rápida na retirada de insights.

A Tabela 2 trata-se de um detalhamento mais aprofundado com questões bem semelhantes que a Tabela 1. Pela sua estrutura booleana, permite um fácil tratamento.

A partir da estrutura das tabelas, cogitou-se a possibilidade de predizer o horário de aulas da turma, e de acordo com os horários declarados de maior rendimento foi possível supor que, devido à menor incidência de bom rendimento nos estudos durante a tarde, a turma em que foi aplicado o questionário deve ter aulas neste turno.

Também se cogitou a possibilidade de entender a dinâmica de trabalhos em grupo considerando o fato dos alunos já conhecerem algum de seus colegas ou não antes do início das aulas, mas não foi possível ter algum indício considerável de tendência comportamental.

Uma análise feita sobre os horários preferidos para estudo e a carga horária semanal despendida em jogos demonstra a tendência dos jogadores a terem maior eficiência no estudo à noite, mas não foi possível concluir, por exemplo, qual é o horário que os entrevistados costumam jogar.

Ainda dentro dos jogos, os que são considerados *e-sports* costumam demandar mais tempo dentre os entrevistados, ao mesmo tempo que os entrevistados que consideram as pessoas importantes da escola costumam gastar poucas horas de jogos.

#### Grupo 4

Como coletamos fragmentos de dados de forma aleatória não existe uma conclusão precisa. Especulamos que os alunos que possuem um cachorro como animal de estimação apresentam um menor desempenho nos estudos por dividir seu tempo com seu animal de estimação.

Fonte: reprodução da autora (2020).

Quadro 56: Respostas da turma de pós-graduação à fase 5 da atividade.

Grupo 1

Os perfis diferentes (um mais técnico, outro mais artístico e outro mais analítico) dos integrantes do grupo permitiu a análise das questões através de diferentes perspectivas.

Isso foi considerado bom, porém o tempo certo para a realização das tarefas limitou um pouco a exploração dessas habilidades.

Grupo 2

Impactou na forma de ver os dados. Observamos critérios diferentes para observação de alguns dados. Cada um dos membros do grupo [teve] perspectivas diferentes sobre os dados chegando ao mesmo objetivo como nas fases anteriores.

Grupo 3

Cada integrante do grupo possuía uma skill que complementava na criação de uma análise mais coesa, entretanto as diferenças foram fator dificultoso para chegar a conclusões em comum acordo. A falta de um computador para cada integrante do grupo dificultou análises individuais.

**Grupo 4**

A interação com o grupo proporcionou debates sobre a escolha dos dados a serem analisados. Apesar de divergências, a interação foi importante para concluir a análise.

Fonte: reprodução da autora (2020).

#### 4.3.3.2 Anotações realizadas durante a atividade no segundo ciclo

No segundo ciclo, a atividade foi conduzida com a ajuda de duas pesquisadoras. A pesquisadora assistente realizando a contagem do tempo e a pesquisadora principal distribuindo as cartas de desafio e realizando as anotações fruto da observação direta. As anotações realizadas durante a fase 2 do estudo de caso estão mostradas nos Quadros 57, 58, 59, 60, 61, 62.

Quadro 57: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 1 do segundo ciclo do estudo de caso.

<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
Teve dificuldade de saber se era o que estava em uma célula ou em uma linha inteira	Teve a mesma dúvida que o grupo 1 sobre o que seria um dado	x	Teve a mesma dúvida sobre se um dado estaria dentro de 1 célula ou seria outra coisa

Fonte: reprodução da autora (2020).

Quadro 58: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 2 do segundo ciclo do estudo de caso.

<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
“Talvez a gente tenha que	Passaram a fazer gráficos de barras no	Nenhum grupo quis utilizar o material de	Nenhum grupo quis utilizar o material de

“normalizar”	<i>sheets</i>	papelaria	papelaria
Nenhum grupo quis utilizar o material de papelaria	Nenhum grupo quis utilizar o material de papelaria	Demoraram mais que os demais para terminar a fase 2	Conversando sobre outros assuntos. Muito dispersos
Depois da metade do tempo escolhe utilizar o material de papelaria (todos participam). Há discussão interna	Mudou de barras verticais para horizontais e depois passaram para a folha de papel	Estão usando uma tela preta. Já fizeram um gráfico de pizza com uma contagem e estão conferindo a contagem para ver se está conferindo com a tabela  “Põe os títulos nos eixos”  “Se você for fazer uma barra então é só selecionar tudo e trocar para a barra”  “Não é melhor colocar 1 para cada tipo de animal?”  “Vai ser tipo	Aluna x: “De que formas a gente pode fazer esse gráfico?”  Aluna x: “Precisamos escolher os dados de uma tabela e pegar o que nos interessa. Não precisa ser todos os dados”  Uma aluna (Aluna y) opera o computador e outra (Aluna x) conversa sobre outros assuntos, mas direciona o grupo. Um outro aluno, mais no canto da sala olha o telefone celular

		<p>Diagrama de Venn quase”</p> <p>“A não ser que a gente faça na mão”</p> <p>“Ela mostrou vários softwares que fazem Diagrama de Venn”</p> <p>Abriram um software de visualização e estão tentando fazer um gráfico preenchendo espaços destinados às variáveis</p>	
--	--	---	--

Fonte: reprodução da autora (2020).

Quadro 59: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 3 do segundo ciclo do estudo de caso.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
“A gente fez um investimento, fez um gráfico bom e agora resolveu todos os problemas”	Motivação média	Muito motivados  Falando   Todos conversando sobre	Estão escrevendo (Aluna y) o resultado da terceira fase no docs. Optaram por não usar o papel

"Ele tirou até o relógio para digitar"		outros assuntos Riem alto, dando gargalhadas todos   Alto	"Uma das alunas (Aluna x) olha o celular"  Motivação média  Um dos alunos que estava com o celular trocou de lugar e passou a ajudar
Muito motivados			

Fonte: reprodução da autora (2020).

Quadro 60: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 4 do primeiro ciclo do estudo de caso.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
"A gente está fazendo a descrição toda de como a gente fez o gráfico"  Muito empenhados mesmo  "Vocês estão filosofando"	Depois de bastante tempo (mais de 10 min) de entregue a 4ª fase, ainda possuem somente a 1ª linha de texto  A aluna do grupo olha outras coisas no celular enquanto os outros membros do grupo operam o computador,	"Ai, esse é o que tem que fazer um textão"  Riem muito  Há 2 polos. 2 computadores. 2 subgrupos que separados analisam e	"Para de olhar o dos outros!"  "É complicado"  Estão distraídos. 2/4 do grupo interage  A aluna (aluna y) que está operando o

<p>“Tá difícil agregar os dados”</p> <p>Conversam entre si sobre como fazer uma análise de dados geral e conversam sobre se podem criar uma analogia dessa situação geral com os dados que possuem</p> <p>“vamos dizer que a maioria prefere cachorro?”</p>	<p>mas parecem confusos. Um deles põe as mãos na cabeça com postura de dificuldade extrema</p> <p>Depois de 20 minutos, ainda apenas 1 linha está escrita no papel.</p> <p>Debatem muito e parecem não chegar a um consenso</p> <p>Começam a encontrar algumas respostas após meia hora. Já têm 4 ou 5 linhas escritas</p> <p>Agora começam a produzir mais texto a partir de resultados de uma tabela dinâmica   motivados</p>	<p>interagem com suas análises</p> <p>Escrevem muito após meia hora na fase. 1 escreve e outro trabalha com a planilha, o terceiro foi pegar café. Usam filtro em 1 das máquinas e contam no olho.</p> <p>Um deles sai do computador onde está escrevendo o texto e vai ajudar o outro que está analisando no <i>sheets</i>. Eles trabalham de forma coordenada.</p> <p>O tempo dado não foi suficiente</p>	<p>computador relembra todas as fases do experimento. A outra aluna (aluna x) do grupo interrompe o grupo para falar de assuntos particulares e suas emoções.</p>
---	---	---	---

		para que o grupo conclúisse a fase	
--	--	------------------------------------	--

Fonte: reprodução da autora (2020).

Quadro 61: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 5 do primeiro ciclo do estudo de caso.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
Estão distraídos. Apenas 1 dos integrantes responde a pergunta	Rápida   Unidos	<p>“Ih... a maioria venceu... essa é uma boa resposta”</p> <p>“Eu tive que ameaçar dar uma chinelada no outro para que ele pudesse participar” (neste momento este outro integrante não está na sala)</p> <p>Uma integrante (Aluna x), que não participa porque está no celular, pede desculpas pela dificuldade de participar. “São muitas modernidades para eu lidar [no celular]”</p>	x

		Procuram na internet e resmungam por não conseguir	
--	--	--	--

Fonte: reprodução da autora (2020).

Quadro 62: Notas de observação dos grupos 1, 2, 3 e 4 durante a fase 6 do primeiro ciclo do estudo de caso.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
Respondem a 6ª questão colaborativamente. Um que está no computador dita a resposta, um escreve e o outro controla, observando o que os demais fazem e eventualmente opina	Desunidos (o estudante que sabe sql trabalha sozinho)  Há um estudante estrangeiro no grupo. Este sugere utilizarem sql para fazer a análise dos dados  O estudante que faz a análise em sql tentou procurar o que era Literacia de Dados na internet. Estava preocupado com os resultados que entregariam	x	“Mas é um software? Qualquer software?”  “Ai, isso é muito difícil (e continuou a olhar o celular)  Riem muito  “Não tem correspondência numérica!”  Ouvem uma notícia com som alto no celular. Demonstram dificuldade para iniciar a fase

	Resolveram empreender		
--	--------------------------	--	--

Fonte: reprodução da autora (2020).

#### 4.3.4 Discussão do segundo ciclo

Neste tópico será apresentada a discussão das fases 4 e 5 da atividade proposta em sala de aula no segundo ciclo do estudo de caso. Na seção 4.3.4.1 será apresentada a discussão relativa à análise dos dados obtidos através das respostas dos estudantes à pergunta da carta da fase 4 da atividade do segundo ciclo do estudo de caso. Na seção 4.3.4.2 será apresentada a discussão relativa à análise dos dados obtidos através das respostas dos estudantes à pergunta da carta da fase 5 da atividade do segundo ciclo do estudo de caso.

##### 4.3.4.1 Discussão da fase 4 do segundo ciclo

A codificação dos resultados obtidos para a fase 4 pelos estudantes da pós-graduação foi realizada, tal como a codificação dos resultados obtidos para as fases 4 e 5 pela turma de graduação, através da Teoria Fundamentada em Dados, utilizando a abordagem construtivista de Charmaz (1983).

Como primeiro passo, foi feita uma leitura das contações de histórias realizadas pelos grupos de estudantes que participaram do segundo ciclo do estudo de caso. O objetivo dessa primeira leitura foi conhecer o texto que foi elaborado pelos grupos de estudantes a fim de descrever detalhadamente o processo de análise que experimentaram durante as fases 1, 2 e 3 da atividade com as cartas.

Como segundo passo, com intenção analítica, foram feitas diversas leituras críticas, a fim de aprofundar a compreensão sobre o texto, dando partida à codificação inicial do *dataset* formado pelas histórias contadas pelos grupos de estudantes. Foram encontrados 13 códigos iniciais no segundo passo: “lendo dados em planilhas”, “explorando os dados”, “planejando o tipo de análise que será realizada”, “comparando dados”, “levantando hipóteses”, “analizando tendências”, “inferindo/concluindo”, “representando resultados por meio de porcentagem”, “representando todos os subconjuntos de um conjunto através de porcentagem”, “representando resultados por meio de expressões partitivas”,

“representando resultados por meio de medidas aproximadas”, “escrevendo a análise por meio de tópicos”, “argumentando com proposições condicionais”.

Como terceiro passo, através dos códigos obtidos durante a codificação inicial, foi realizada a codificação axial, quando através dos códigos da codificação inicial, foram abstraídas categorias mais gerais. Foram encontradas 9 categorias durante a codificação axial: “lendo dados em diferentes visualizações”, “explorando dados”, “planejando a análise”, “comparando dados”, “levantando hipóteses”, “analizando tendências”, “inferindo”, “representando resultados”, “argumentando com dados”.

Como quarto passo, buscou-se a saturação das categorias encontradas durante a codificação axial, através da abstração dessas categorias em outras, dessa vez tão gerais quanto possível. Dessa forma, obteve-se os códigos: “lendo dados”, “analizando dados”, “informando resultados” e “argumentando com dados”. O Quadro 63, a seguir, mostra toda a codificação realizada para o conjunto de dados gerado pelos estudantes no segundo ciclo da análise de dados ao responderem a fase 4.

Quadro 63: Codificação dos resultados obtidos para a fase 4 pelos estudantes da pós-graduação.

Codificação inicial	Codificação axial	Saturação
Lendo dados em planilhas	Lendo dados em diferentes visualizações	Lendo dados
Explorando os dados	Explorando dados	
Planejando o tipo de análise que será realizada	Planejando a análise	
Comparando dados	Comparando dados	Analizando dados
Levantando hipóteses	Levantando hipóteses	

Analisando tendências	Analizando tendências	
Inferindo / Concluindo	Inferindo	
Representando resultados por meio de porcentagem		
Representando todos os subconjuntos de um conjunto através de porcentagem		
Representando resultados por meio de expressões partitivas	Representando resultados	Informando resultados
Representando resultados por meio de medidas aproximadas		
Escrevendo a análise por meio de tópicos		
Argumentando com proposições condicionais	Argumentando com dados	Argumentando com dados

Fonte: reprodução da autora (2020).

#### 4.3.4.2 Discussão da fase 5 do segundo ciclo

A análise das respostas dos estudantes da turma de pós-graduação à pergunta da fase 5 – “Como a interação entre seu grupo impactou a análise dos dados e os resultados obtidos?” – foi realizada com base na abordagem construtivista da teoria fundamentada em dados, assim como na análise das respostas dos estudantes para a pergunta realizada na fase 4. Os códigos iniciais, as categorias axiais e a saturação obtidos através das respostas dos estudantes podem ser lidos no Quadro 64, a seguir:

Quadro 64: Codificação dos resultados obtidos para a fase 5 pelos estudantes da pós-graduação.

Inicial	Axial	Saturação
Interagindo		
Debatendo sobre a escolha dos dados a serem analisados		
Interagindo com o grupo	Interagindo	
Proporcionando debates através da interação		Comunicando
Interagindo para concluir a análise		
Divergindo		
Chegando a um objetivo comum através de diferentes perspectivas	Contribuindo com ideias pertinentes e distintas	
Trabalho em grupo impactando a forma de ver os dados		
Apresentando habilidades complementares na criação de uma análise mais coesa		
Analizando através de diferentes perspectivas oriundas de diferentes perfis técnicos		
Analizando questões em grupo	Analizando dados	Cooperando
Analizando de forma coesa		
Proporcionando debates sobre a análise de dados através da interação		
Analizando de forma coesa a partir de diferentes perfis de competência		

Cada membro tendo diferentes perspectivas sobre os dados		
Utilizando critérios diferentes para a observação de alguns dados		
Enfrentando dificuldades devido a falta de um computador para cada integrante		
Enfrentando dificuldades devido às diferenças	Enfrentando dificuldades	
Enfrentando dificuldades para chegar em comum acordo devido às diferenças		
trabalhando com menos de um computador por integrante		
Limitando a exploração das atividades através do "tempo certo"	Otimizando o tempo	Coordenando

Fonte: reprodução da autora (2020).

No final do segundo ciclo do estudo de caso, o conjunto de cartas que compunha o artefato Cartas para a Literacia foi validado com alterações nas cartas da 1<sup>a</sup> fase, da 2<sup>a</sup> fase, da 4<sup>a</sup> fase e da 6<sup>a</sup> fase. Os ajustes realizados tiveram como objetivo tornar as solicitações das cartas mais aderentes ao formato do *dataset* utilizado (uma planilha com duas tabelas), além de fornecer um aspecto mais lúdico por meio do uso de ilustrações e criar destaque para as informações escritas julgadas importantes através do uso de cores.

A versão final do conjunto de cartas, com todas as alterações realizadas após os dois ciclos de *design* representados pelos estudos de caso com as turmas de graduação e pós graduação, pode ser visualizada na Figura 24.

Figura 24: Versão final do conjunto de cartas, obtida após o segundo ciclo de design, realizado com a turma de pós-graduação.



Fonte: reprodução da autora (2020).

#### 4.4 SUMARIZAÇÃO DOS DADOS E VERSÃO INICIAL DA ESCALA

Depois de concluída a codificação dos resultados obtidos pelos estudantes da pós-graduação para a fase 4, passou-se para a fase de consolidação dos resultados da codificação das respostas obtidas na fase 4 no primeiro e no segundo ciclos. Primeiro foram listados os códigos iniciais, axiais e de saturação obtidos através da codificação da fase 4 no primeiro ciclo, depois uniu-se a eles os resultados obtidos através da codificação da fase 4 no segundo ciclo, de forma que os códigos iniciais iguais para a codificação do primeiro e do segundo ciclos, fossem superpostos, ou seja, surgissem ocupando uma única linha da codificação inicial e os códigos iniciais diferentes oriundos da codificação do primeiro e do segundo ciclos fossem acumulados. O mesmo procedimento foi repetido na codificação axial e para a saturação. Dessa forma, obtivemos os 29 códigos iniciais, 19 códigos axiais e 4 códigos de saturação constantes no Quadro 65, a seguir.

Quadro 65: Resultado consolidado das respostas obtidas pelos estudantes da graduação e da pós-graduação para as respostas da fase 4.

Codificação inicial	Codificação axial	Saturação
Lendo dados em planilhas	Lendo dados em diferentes visualizações	Lendo dados
Explorando os dados	Explorando dados	
Planejando o tipo de análise que será realizada	Planejando a análise	
Comparando dados		
Fazendo comparação por meio de porcentagem	Comparando dados	
Identificando interseções		
Calculando frequências	agregando dados	
Levantando hipóteses	Levantando hipóteses	Analizando dados
Analizando tendências	Analizando tendências	
Analizando tabelas	Analizando dados em diferentes tipos de visualizações	
Analizando gráficos		
Utilizando linguagem de programação para a análise de dados	Analizando dados com ajuda de linguagem de programação	
Utilizando planilha eletrônica para a análise de dados	Analizando dados com ajuda de software	
Inferindo / Concluindo	Inferindo	
Representando resultados por meio de	Representando resultados por	Informando

porcentagem	meio de porcentagens	resultados
Representando resultados por meio de subconjuntos com porcentagens		
Representando todos os subconjuntos de um conjunto através de porcentagens		
Representando resultados por meio de subconjuntos com expressões partitivas	Representando resultados por meio de expressões partitivas	
Representando resultados por meio de expressões partitivas		
Representando resultados por meio de frequências	Representando resultados por meio de frequências	
Representando resultados por meio de subconjuntos com frequências		
Representando resultados por meio de frações	Representando resultados por meio de frações	
Representando resultados por meio de medidas aproximadas		
Representando resultados com expressões indefinidas	Representando resultados por meio de medidas aproximadas	
Representando resultados por meio de subconjuntos de tamanho indeterminado		
Escrevendo em forma de tópicos	Representando resultados por meio da escrita	
Escrevendo em forma de texto		
Argumentando com proposições condicionais	Argumentando com dados	Argumentan

Relacionando teoria com dados	Relacionando teoria com dados	do com dados
-------------------------------	-------------------------------	--------------

Fonte: reprodução da autora (2020).

A partir da consolidação dos resultados da codificação da fase 4 para o primeiro e segundo ciclos do estudo de caso, notou-se que os códigos resultantes para a saturação — Lendo dados, analisando dados, informando resultados e argumentando com dados — apresentavam etapas semelhantes às descritas na definição de Literacia de Dados de D'Ignazio e Bhargava (2015) — Ler dados, Trabalhar com dados, Analisar dados e Argumentar com base em dados.

Desta forma, a partir da percepção do potencial dos códigos consolidados para as respostas à fase 4 dos desafios propostos para responder à pergunta desta pesquisa, foi proposta a versão preliminar de uma escala de Literacia de Dados, instrumento através do qual seria possível, a princípio, avaliar a Literacia de Dados de grupos de estudantes.

A escala foi construída de forma que os códigos encontrados na etapa de saturação da codificação originassem os níveis da escala, os códigos encontrados na codificação axial originassem as categorias da escala e os códigos encontrados na codificação inicial às suas subcategorias.

Os níveis da versão preliminar da escala de Literacia de Dados, segundo o resultado das três etapas de codificação foram: ler dados, analisar dados, informar resultados e argumentar com base em dados. Houve a adição do nível “Gerenciar dados” com as categorias “Armazenar dados” e “Limpar dados” na escala, devido à percepção de que os estudantes tanto no primeiro ciclo quanto no segundo ciclo realizaram ações voltadas para o armazenamento de dados e que, se outro *dataset* que não o Cores fosse oferecido para a realização da atividade, poderia haver a necessidade de realização de tarefas relacionadas à limpeza dos dados. Sendo assim, a escala de Literacia de Dados se materializou em sua versão inicial no Quadro 66 a seguir.

Quadro 66: Versão inicial da escala de Literacia de Dados gerada a partir da codificação dos resultados obtidos pelos estudantes a partir da 4<sup>a</sup> fase do desafio Cartas para a Literacia.

<b>1- Ler dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
	Ler Dados em diferentes formatos	Ler dados em tabelas Ler dados em gráficos Ler dados em textos Ler dados em infográficos
<b>2- Gerenciar dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
	Armazenar dados	Salvar tabela (ou dados em outros formatos) em um diretório no computador ou em nuvem
	Limpar dados	Identificar dados incompletos e incorretos
<b>3- Analisar dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
Quanto à tecnologia utilizada	Analizar dados manualmente	Analizar dados utilizando material de papelaria

	Analizar dados com ajuda de software	Analizar dados utilizando planilhas eletrônicas
		Analizar dados utilizando outros softwares (Tableau, RapidMiner, Microsoft BI Studio, Google Data Studio, etc.)
	Analizar dados com ajuda de linguagem de programação	Analizar dados com ajuda de linguagem de programação
		Analizar dados com ajuda de linguagem de consulta
Quanto aos aspectos metodológicos	Planejar a análise	Identificar previamente a pergunta de pesquisa principal e as perguntas de pesquisa secundárias que guiarão a análise
		Levantar hipóteses
		Desenhar os passos metodológicos que serão percorridos na análise de dados
		Fazer o levantamento prévio das ferramentas que serão utilizadas na análise de dados (softwares,

		linguagens de programação, etc.)
	Explorar dados	<p>Examinar medidas de tendência central</p> <p>Examinar desvio padrão</p> <p>Examinar dados de forma gráfica</p> <p>Examinar correlações</p> <p>Examinar tendências</p> <p>Examinar pontos fora da curva</p>
	Comparar dados	Procurar semelhanças e disparidades entre os dados
	Analizar tendências	Analizar a possibilidade de um evento futuro, com base em eventos passados ou presentes
<b>4- Informar resultados</b>		
	Categorias	Subcategorias
	Representar resultados em formato textual	Representar resultados utilizando expressões partitivas (a maioria, parte

		de, metade de, a maior parte de) na composição do sujeito de uma oração
		Representar resultados utilizando expressões que remetem a quantidades aproximadas (aproximadamente, aproximadamente metade, quase todos, quase 50%)
		Representar resultados por meio de frações
		Representar resultados por meio de porcentagens
		Representar resultados por meio de frequências
	Representar resultados em forma de tabela	Representar resultados em forma de tabela
	Representar resultados utilizando gráficos	Representar resultados utilizando gráficos
	Representar resultados utilizando infográficos	Representar resultados utilizando infográficos
<b>5- Argumentar por meio de dados</b>		

	Categorias	Subcategorias
	Argumentar por meio de dados	Argumentar utilizando dados com o objetivo de apresentar fatos, ideias, razões lógicas, provas que comprovem uma afirmação
		Argumentar por meio de dados, entremeando dados e teorias para apresentar os fatos, as ideias, as razões lógicas, as provas, que compõem uma afirmação, uma tese

Fonte: reprodução da autora (2020).

Foi gerada também uma tabela consolidando os resultados obtidos a partir das respostas dos estudantes no primeiro e no segundo ciclos do estudo de caso para a fase 5, que pedia para descreverem como o trabalho em grupo impactou a análise dos dados e os resultados obtidos (Quadro 67).

Quadro 67: Resultado consolidado das respostas obtidas pelos estudantes da graduação e da pós-graduação para as respostas da fase 5. Em laranja os códigos gerados através da codificação dos textos da turma de graduação e em cinza os códigos gerados através da codificação dos textos da turma de pós-graduação.

Codificação inicial	Codificação axial	Saturação
Interagindo em grupo		
Interagindo em grupo	Interagindo	Comunicando
Interagindo em grupo		

Interagindo		
Interagindo em grupo		
Respondendo à necessidade de interação		
Ideias fluindo		
Utilizando o fluxo de ideias como moeda de troca entre os integrantes		
Lançando novos pensamentos ao longo do debate		
Interagindo		
Debatendo sobre a escolha dos dados a serem analisados		
Interagindo com o grupo		
Proporcionando debates através da interação		
Interagindo para concluir a análise		
Divergindo		
Trocando informações de forma eficiente	Trocando	
Promovendo discussão	informações	
Pluralidade de ideias		
Explorando diversos pontos de vista	Contribuindo	
Explorando diversos pontos de vista através da liberdade e independência de pensamento	com ideias pertinentes e distintas	
Tendo ideias pertinentes e distintas individualmente		
Chegando a um objetivo comum através de diferentes		

perspectivas		
Compartilhando resultados internamente ao grupo		
Compartilhando resultados internamente ao grupo de forma dinâmica	Compartilhando resultados	
Compartilhando resultados entre os subgrupos por meio de um integrante		
Apresentando dados para cada uma das fases		
Interagindo em grupo de forma a impactar positivamente a análise e os resultados		
Interagindo com o grupo de forma a proporcionar uma análise mais efetiva e ampla dos dados analisados		
Interagindo em grupo de forma a incentivar-lo a ser criativo	Facilitando a	
Interagindo em grupo de forma a contribuir para a proatividade	cooperação	
Alcançando uma melhor qualidade e precisão do trabalho através da discussão	através da	
Interação facilitando a obtenção de dados	interação	
Interação facilitando a divisão de tarefas		
Interação agilizando as etapas da atividade		
Dividindo tarefas entre os integrantes do grupo	Dividindo tarefas	
Realizando pesquisas de forma meio independente por meio de subgrupos	Cooperando	
	Pesquisando	

Desenvolvendo um bom trabalho através da liberdade e independência dos subgrupos	Desenvolvendo um trabalho	
Realizando tarefas		
Resolvendo problemas		
Trabalhando em grupo	Resolvendo problemas	
Facilitando a criação através do compromisso do grupo		
Realizando a tarefa		
Coletando dados por meio do computador		
Obtendo dados	Coletando dados	
Participação facilitando a obtenção de dados		
Interpretando dados de forma independente	Interpretando dados	
Analizando dados		
Analizando os dados tabelados conjuntamente		
Analizando conjuntamente		
Participando da análise em grupo		
Trabalho em grupo impactando a forma de ver os dados	Analizando dados	
Apresentando skills complementares na criação de uma análise mais coesa		
Analizando através de diferentes perspectivas oriundas de diferentes perfis técnicos		
Analizando questões em grupo		

Analizando de forma coesa		
Proporcionando debates sobre a análise de dados através da interação		
Analizando de forma coesa a partir de diferentes perfis de competência		
Cada membro tendo diferentes perspectivas sobre os dados		
Utilizando critérios diferentes para a observação de alguns dados		
Utilizando ferramentas digitais de forma sinérgica	Utilizando ferramentas digitais	
Enfrentando dificuldades devido a falta de um computador para cada integrante		
Enfrentando dificuldades devido às diferenças	Enfrentando dificuldades	
Enfrentando dificuldades para chegar em comum acordo devido às diferenças		
trabalhando com menos de um computador por integrante		
Obtendo os resultados esperados	Obtendo resultados	
Facilitando a criação através da sinergia do grupo		
Produzindo informações	Produzindo conteúdo	
Produzindo conteúdo		

Produzindo conteúdo fora do meio digital		
Dividindo tarefas		
Dividindo tarefas	Dividindo tarefas	
Alcançando eficiência no trabalho através da divisão de tarefas		
Separando o grupo em subgrupos		
Dividindo-se em subgrupos por competências	Dividindo-se em subgrupos	
Dividindo-se em subgrupos para otimizar o tempo		
Intermediando os subgrupos por meio de um dos integrantes do grupo	Mediando os subgrupos	
Trabalhando em diversas etapas		Coordenando
Trabalhando em etapas	Trabalhando em etapas	
Participação agilizando as etapas da atividade		
Executando os desafios no tempo previsto devido a divisão de tarefas		
Executando os desafios sem ultrapassar o tempo previsto	Otimizando o tempo	
Limitando a exploração das atividades através do "tempo certo"		
Discutindo a melhor forma de representar os dados	Planejando a representação de resultados	

Fonte: reprodução da autora (2020).

Após a sumarização dos códigos encontrados através da Teoria Fundamentada em Dados para a fase 5, encontramos os códigos: “comunicando”, “cooperando” e “coordenando”, na fase de saturação, representando um reflexo da utilização do Modelo 3C de Colaboração nesta pesquisa.

As ações listadas na codificação axial que deram origem ao código “comunicando” revelam que os grupos experimentaram situações de interação para a troca de informações e compartilhamento de resultados, com o objetivo de contribuir com ideias pertinentes e distintas, facilitando a cooperação. As ações que culminaram no código “cooperando” revelam que os grupos desenvolveram um trabalho de produção de conteúdo por meio de pesquisa, que envolveu coleta, interpretação, análise de dados e utilização de ferramentas digitais para a resolução de problemas, algumas vezes enfrentando barreiras. As ações que permitiram encontrar o código “coordenando” na saturação, revelam que os grupos se dividiram em subgrupos que trabalharam em etapas e com a ajuda de mediadores para a realização das tarefas e planejamento da representação dos resultados, otimizando o tempo.

Na Figura 25 a nuvem de palavras gerada através dos itens da codificação inicial revela os termos mais frequentes na contação de histórias relativa à fase 5 da atividade com o artefato, contribuindo para a identificação de características e processos relativos à colaboração que os grupos experimentaram durante os ciclos de estudo de caso.

Figura 25: Nuvem de palavras gerada através das respostas dos grupos à quinta carta de desafio.



Fonte: reprodução da autora (2020).

## 5 AVALIAÇÃO DA ESCALA E APRESENTAÇÃO DOS PRODUTOS DA PESQUISA

---

Seguindo as orientações do Mapa de DSR, neste capítulo foi realizada a avaliação do artefato Cartas para a Literacia a partir das novas perguntas: 1) O artefato funciona? 2) O artefato operacionaliza todas as conjecturas teóricas? 3) É possível gerar uma escala de Literacia de Dados a partir do uso do artefato? 4) Que fenômenos emergiram do uso do artefato? 5) O artefato tem potencial para ser aceito pelos alunos, professores e escola?

Para realizar a análise, foi planejada uma entrevista semiestruturada com professores e especialistas em Matemática e Estatística. Primeiro foi realizada uma entrevista piloto com uma estudante de graduação, com o objetivo de validar os instrumentos e as técnicas utilizadas para coletar os dados, além de melhorar a organização dos elementos que comporiam a entrevista semiestruturada. O estudo piloto foi realizado através de entrevista remota no Google Meet com uma estudante de Iniciação Científica do curso de Bacharelado em Matemática de uma universidade pública, utilizando um quadro colaborativo *Jamboard* (Figura 26) e um roteiro de entrevista elaborado para esta finalidade (Apêndice 3). Foi solicitada a autorização de imagem da estudante e logo depois iniciou-se a entrevista com a explicação dos seus objetivos e com a visualização de um vídeo que explicava brevemente o estudo de caso realizado na pesquisa. Depois foi apresentada a escala de Literacia de Dados e dada à estudante a oportunidade de realização de perguntas gerais sobre a pesquisa. Na sequência, foi solicitado para a estudante que realizasse uma análise crítica da escala apresentada. Depois da realização da entrevista, levantou-se que houve limitações quanto às possibilidades de colaboração utilizando o software *Jamboard* no computador da entrevistada, quanto ao *design* visual da escala apresentada, além da percepção de aspectos da escala e do seu desenvolvimento que mereceriam uma explicação mais detalhada durante a entrevista.

Figura 26: Escala de Literacia de Dados apresentada na entrevista piloto, utilizando o software Jamboard. Nas notas autoadesivas azuis constam as definições das categorias, com base na definição de Literacia de Dados de D'Ignazio e Bhargava (2015).

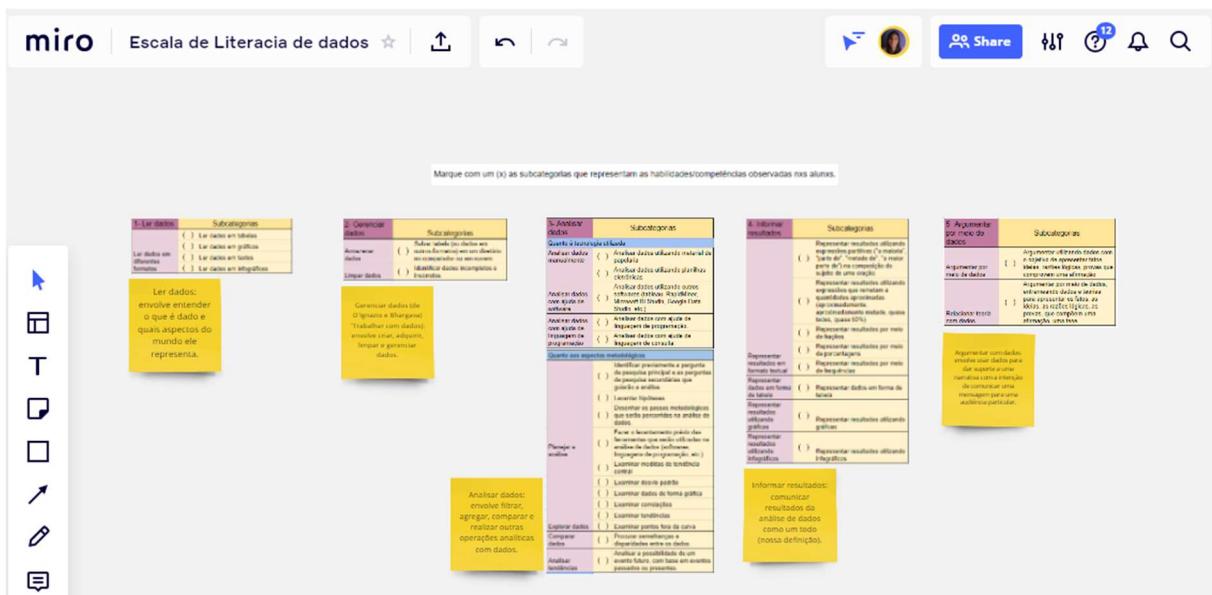


Fonte: reprodução da autora (2020).

Depois da entrevista piloto, seguiram-se as entrevistas com 7 especialistas docentes das instituições Instituto Federal do Rio de Janeiro (IFRJ), Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Fundação de Apoio à Escola Técnica (FAETEC) e Universidade Federal do Rio de Janeiro (UFRJ). O especialista 1 (UNIRIO), com 10 anos de docência em estatística; a especialista 2 (IFRJ), com 11 anos de docência em matemática; o especialista 3 (FAETEC), com 8 anos de experiência em Ciência de Dados; a especialista 4 (IFRJ), com 10 anos de docência em estatística; a especialista 5 (IFRJ) com 15 anos de docência em estatística; a especialista 6 (IFRJ), com 5 anos de docência em estatística e o especialista 7 (UFRJ), com 16 anos de docência em matemática.

As entrevistas foram realizadas remotamente através do Google Meet, utilizando o quadro colaborativo Miro (Figura 27) como ferramenta de suporte, além do mesmo vídeo explicativo da pesquisa que foi utilizado na entrevista piloto, o roteiro de entrevista que se encontra no apêndice 3 e uma versão da escala de Literacia de Dados com *design* visual melhorado.

Figura 27: Escala de Literacia de Dados apresentada nas 7 entrevistas com os especialistas. Nas notas autoadesivas laranjas constam as definições das categorias, com base na definição de Literacia de Dados de D'Ignazio e Bhargava (2015).



Fonte: reprodução da autora (2020).

As sugestões dos especialistas para a versão da escala de Literacia de Dados mostrada na Figura 26 foram organizadas no Quadro 68 a seguir.

Quadro 68: Ajustes sugeridos pelos especialistas para a escala de Literacia de Dados.

<b>Especialista 1</b>	
Nível 5 Informar resultados	Adição de uma categoria sobre “saber escolher o melhor formato para a representação de dados”

Nível 6 Argumentar por meio de dados	Reorganização das categorias em 1) “argumentar por meio dos resultados informados, 2) “argumentar por meio dos resultados relacionando teoria com dados, 3) “argumentar por meio de dados gerais informados antes da análise” e 4) “argumentar relacionando teoria com dados informados antes da análise”
<b>Especialista 2</b>	
Nível 3 Gerenciar dados	Na categoria “Limpar dados”: Trocar “dados incompletos e incoerentes” para “dados incompletos, incorretos, incoerentes”
Nível 4 Analizar dados	Houve uma dúvida sobre se a categoria “planejar a análise” deveria vir antes de “explorar dados” ou se estas deveriam ser categorias paralelas, ou ainda, se “explorar dados” deveria vir antes de “planejar a análise”
Nível 6 Argumentar por meio de dados	Adição da subcategoria “argumentar criticamente acerca de possíveis tendências e previsões” (categoria mais complexa que “Argumentar por meio de dados” e diferente da categoria “Argumentar relacionando teoria com dados”)
<b>Especialista 3</b>	
Nível 1 Coletar dados	Adicionar o nível “Coleta de dados”, com a categoria “Planejar a coleta de dados” e as subcategorias 1) identificar a pergunta de pesquisa e 2) levantar hipóteses.

	Adicionar também a categoria “coletar dados”
Nível 4 Analizar dados	Na categoria “Planejar a análise”, suprimir as subcategorias “Indicar previamente a pergunta de pesquisa principal e as perguntas de pesquisa secundárias que guiarão a análise” e “Levantar hipóteses”
Nível 5 Informar resultados	<p>“Subdividir este nível nos eixos: “Quanto aos aspectos da Língua Portuguesa” e “Quanto à forma de representação dos resultados”</p> <p>No eixo “Quanto aos aspectos da Língua Portuguesa”, adicionar a categoria “Representar resultados utilizando coesão e coerência na informação dos resultados”. Nesta categoria, adicionar as subcategorias “Informar resultados utilizando coesão na Língua Portuguesa” e “Informar resultados utilizando coerência na linguagem”</p>
Nível 6 Argumentar por meio de dados	Adicionar a categoria “Relacionar os resultados obtidos nesta pesquisa com dados e informações de outras pesquisas” e a categoria “Relacionar resultados obtidos nesta pesquisa com diferentes contextos da realidade social”, com as subcategorias “Dentro da sua comunidade” e “Fora da sua comunidade”
<b>Especialista 4</b>	
Nível 1 Coletar dados	Adicionar o nível “Coletar dados” antes do nível “Ler dados”

	<p>Adicionar o eixo: “Quanto aos passos metodológicos da coleta de dados”, com a categoria: “Planejar a coleta e análise de dados”, com as subcategorias: “Identificar previamente a pergunta de pesquisa principal e as perguntas de pesquisa secundárias que guiarão a análise”, “Levantar hipóteses”, “Desenhar os passos metodológicos que serão percorridos na coleta e análise dos dados” e “Fazer o levantamento prévio das ferramentas que serão utilizadas na coleta e na análise de dados (pesquisa na internet, entrevistas, pesquisa bibliográfica, pesquisa em arquivos)”</p> <p>Adicionar eixo: “Quanto aos métodos de coleta”, com as categorias “Coletar dados em amostragem probabilística (amostragem aleatória, por exemplo)” e “Coletar dados em amostragem não probabilística (amostragem accidental, por exemplo)”</p> <p>Adicionar eixo: “Quanto aos tipos de mídias coletadas”, com as categorias “Perguntado (boca-a-boca)”, “Pesquisando na internet”, “Em arquivos de papel”</p>
Nível 4 Analizar dados	<p>Se o professor não utilizar o nível zero (coleta de dados), então eu posso trazer a categoria “Planejar a análise” para antes ou dentro do nível “Análise de dados”</p> <p>Dentro da categoria “Explorar dados”, substituir a</p>

	subcategory “Examinar desvio padrão” por “Examinar medidas de dispersão absolutas e relativas”
Nível 6 Argumentar por meio de dados	Ajustar a escrita da segunda categoria para “Argumentar relacionando teoria com dados”
<b>Especialista 5</b>	
Nível 3 Gerenciar dados	<p>Substituir a categoria “Limpar dados” pela categoria “Tratar dados”</p> <p>Na nova categoria “Tratar dados”, incluir as subcategorias: “Limpeza de dados” e “Transformação de dados (identificar os dados que estão em um formato e transformá-lo para outro mais desejável)”</p>
Nível 4 Analisaar dados	<p>Na categoria “Explorar dados”, levar a subcategoria “Examinar dados de forma gráfica” para o primeiro lugar da lista de subcategorias</p> <p>Ainda na categoria “Explorar dados”, colocar como segunda subcategoria “Examinar medidas de tendência central”</p> <p>Ainda na categoria “Explorar dados”, substituir a subcategoria “Examinar desvio padrão” por “examinar medidas de variabilidade (desvio padrão, variância, desvio médio, variância relativa, coeficiente de</p>

	<p>variabilidade, etc.)” e colocá-la como terceira subcategoria</p> <p>Como quarta subcategoria da categoria “Explorar dados”, colocar “Examinar pontos fora da curva”</p> <p>Como quinta subcategoria da categoria “Explorar dados”, colocar: “Examinar correlações”</p> <p>Como sexta subcategoria da categoria “Explorar dados”, colocar: “Examinar possibilidades”</p> <p>Substituir a categoria “Analisar tendências” por “Analisa possibilidades ou estimativas”</p>
Nível 5 Informar resultados	<p>Na categoria “Representar resultados em formato textual”, manter a primeira e a segunda subcategorias e reordenar as restantes para: 3- Representar resultados por meio de frequências/contagem/ocorrências; 4- Representar resultados por meio de frações ou comparações (por ex.: uma em cada 10 pessoas possuem carro de passeio); 5- Representar resultados por meio de porcentagem</p>
<b>Especialista 6</b>	
Nível 1 Coletar dados	<p>Localizar o nível “Organizar dados” aqui opcionalmente, no caso de os estudantes serem solicitados a realizar coleta de dados antes da leitura de dados</p>

	<p>Substituir o nível “Gerenciar dados” por “Organizar dados”, com os subníveis “Compilar ou coletar dados”; “Tratar dados (limpeza e transformação de dados)”; “Armazenar dados (pontualmente + durante todo o processo de análise de dados)”</p>
Nível 2 Ler dados	<p>Adicionar o eixo: “Quanto ao formato dos dados”, para abrigar as subcategorias: “Ler dados em tabelas”, “Ler dados em gráficos”, “Ler dados em textos” e “Ler dados em Infográficos”</p> <p>Adicionar a subcategoria “Ler dados em planilhas” logo depois da subcategoria “Ler dados em Tabelas”</p> <p>Adicionar o eixo “Quanto à tecnologia utilizada”, com as categorias e subcategorias a seguir:</p> <ul style="list-style-type: none"> <li>1) “Ler dados visualmente”: Ler dados utilizando recursos visuais e material de papelaria</li> <li>2) “Ler dados com ajuda de software”: “Ler dados utilizando planilhas eletrônicas”, “Ler dados utilizando softwares especializados”</li> <li>3) “Ler dados utilizando codificação”: “Ler dados com ajuda de linguagem de programação”; “Ler dados com ajuda de linguagem de consulta”</li> </ul>
Nível 3	Localizar o nível “Organizar dados” aqui opcionalmente,

Gerenciar dados	<p>no caso de os estudantes não serem solicitados a realizar coleta de dados.</p> <p>Substituir o nível “Gerenciar dados” por “Organizar dados”, com os subníveis “Compilar ou coletar dados”; “Tratar dados (limpeza e transformação de dados)”; “Armazenar dados (pontualmente + durante todo o processo de análise de dados)”</p>
Nível 5 Informar resultados	<p>Adicionar a categoria “Representar os resultados da análise exploratória”, com as subcategorias: “Representar resultados em forma textual”, “Representar resultados em forma de tabelas” e “Representar resultados utilizando gráficos”</p>
Nível 7 Tomar: decisões	<p>Adicionar o nível 6: “Tomar decisões por meio de dados”</p> <p>No novo nível 6, adicionar os subníveis: “Tomar decisões a partir dos dados analisados até então” e “Tomar decisões por meio de dados, utilizando ferramentas e técnicas estatísticas que permitam a tomada de decisões mais precisas e assertivas”</p>
<b>Especialista 7</b>	
Nível 2 Ler dados	<p>Adicionar a categoria “Ler dados comparativamente”, por ser importante ler o dado dentro do seu contexto, problematizando os dados lidos.</p> <p>Subdividir a categoria “Ler dados em gráficos” em</p>

	subcategorias de acordo com os tipos de gráficos em que o professor pretender avaliar a literacia dos estudantes
Nível 3 Gerenciar dados	<p>Adicionar uma fase anterior a “armazenar dados”, em que os estudantes sejam solicitados a “criar dados”, como os que os estudantes criaram quando escreveram a <i>storytelling</i> com base em dados</p> <p>Exemplo: categoria: criar/adquirir dados. Subcategoria 1: propor situações que levem à criação de dados. Subcategoria 2: Obter dados em diferentes situações.</p> <p>Adicionar a categoria “tratamento de dados” e adicionar a essa categoria a subcategoria “escolher qual é a melhor forma de representação para os dados (se em tabelas, gráficos, infográficos, frases, etc.)</p>
Nível 4 Analizar dados	<p>Adicionar a categoria “Analisar dados apenas visualmente” no campo “Quanto à tecnologia utilizada”.</p> <p>Substituir a subcategoria “Analisar dados utilizando outros softwares (Tableau, RapidMiner, Microsoft BI Studio, Google Data Studio, etc.)” por “Analisar dados utilizando software especializado em análise estatística”</p> <p>Dependendo da atividade e/ou pesquisa que será realizada pelo estudante, a etapa de exploração de dados (categoria “Explorar dados”) poderá surgir antes do planejamento da análise (categoria “Planejar a análise”)</p>

Nível 5 Informar resultados	Transferir a subcategoria “Representar resultados por meio de frequências” para o primeiro lugar na lista de subcategorias da categoria “Representar resultados em formato textual”
Nível 6 Argumentar por meio de dados	Suprimir a palavra “lógica” da subcategoria “Argumentar por meio de dados entremeando dados e teorias para apresentar os fatos, as ideias, as razões lógicas, as provas que compõem uma afirmação, uma tese

Fonte: reprodução da autora (2020).

A conclusão pós entrevista com os especialistas foi que o artefato Cartas para a Literacia foi capaz de tornar explícitos os conhecimentos tácitos sobre Literacia de Dados de estudantes, proporcionando a geração de uma escala de Literacia de Dados validada pelos professores especialistas das áreas de matemática, estatística e ciência de dados entrevistados. Depois de realizadas as alterações recomendadas pelos especialistas, obteve-se a versão da escala apresentada na Tabela 69.

Tabela 69: Escala de Literacia de Dados gerada a partir da codificação dos resultados obtidos pelos estudantes a partir da 4ª fase do desafio Cartas para a Literacia.

<b>1- Coletar dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
	Planejar a coleta de dados	Identificar previamente a pergunta principal e as perguntas secundárias que guiarão a pesquisa
		Levantar as hipóteses que

		guiarão a pesquisa
		Desenhar os passos metodológicos que serão percorridos na pesquisa
		Levantar as ferramentas que serão utilizadas na pesquisa
		Estimar o tempo necessário para a realização dos passos metodológicos
	Colher dados	Colher dados em amostragem probabilística
		Colher dados em amostragem não probabilística
<b>2- Ler dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
Quanto ao formato utilizado	Ler dados em diferentes formatos	Ler dados em tabelas Ler dados em planilhas Ler dados em gráficos Ler dados em infográficos
Quanto ao recurso utilizado	Ler dados utilizando os	Ler dados utilizando os

	sentidos	órgãos dos sentidos e material de papelaria
	Ler dados com ajuda de software	Ler dados com ajuda de softwares especializados
	Ler dados utilizando comandos em linha de código	Ler dados com ajuda de linguagem de programação
		Ler dados com ajuda de linguagem de consulta
Quanto ao método	Ler dados analiticamente	Ler dados com o objetivo de compreender de forma lógica e global um conjunto de dados
	Ler dados comparativamente	Ler os dados de forma crítica dentro do seu contexto, problematizando os dados e confrontando-os com a realidade
<b>3- Gerenciar dados</b>		
	Categorias	Subcategorias
	Armazenar dados	Salvar dados pontualmente em um diretório no computador ou em nuvem
		Salvar dados continuamente

		em um diretório no computador ou em nuvem
	Tratar dados	Limpar dados (identificar dados incompletos, incorretos e incoerentes)
		Transformar dados (identificar dados que estão em um formato e transformá-lo para outro mais desejável)
		Escolher qual é a melhor representação para os dados (tabelas, frases, planilhas, infográficos, etc.)
	Proteger dados	Proteger dados sensíveis quando necessário
	Diferenciar dados confiáveis de dados não confiáveis	Diferenciar dados confiáveis de dados não confiáveis, através da verificação da veracidade dos dados
<b>4- Analisar dados</b>		
	Categorias	Subcategorias
Quanto à tecnologia utilizada	Analisisar dados sem auxílio de computador	Analisisar dados utilizando somente os órgãos dos

		<p>sentidos</p>
		<p>Analizar dados utilizando os órgãos dos sentidos e material de papelaria</p>
	<p>Analizar dados com ajuda de software</p>	<p>Analizar dados utilizando planilhas eletrônicas</p>
		<p>Analizar dados utilizando software especializado em análise estatística</p>
	<p>Analizar dados utilizando codificação</p>	<p>Analizar dados com ajuda de linguagem de programação</p>
		<p>Analizar dados com ajuda de linguagem de consulta</p>
<p>Quanto aos aspectos metodológicos</p>	<p>Planejar a análise</p>	<p>Desenhar os passos metodológicos que serão percorridos na análise de dados</p>
		<p>Fazer o levantamento prévio das ferramentas que serão utilizadas na análise de dados (softwares, linguagens de programação, etc.)</p>

	Explorar dados	<p>Examinar os dados de forma gráfica</p> <p>Examinar medidas de tendência central</p>
		<p>Examinar medidas de variabilidade (desvio padrão, variância, desvio médio, variância relativa, coeficiente de variabilidade)</p>
		<p>Examinar pontos fora da curva</p>
		<p>Examinar correlações</p>
		<p>Examinar possibilidades</p>
	Comparar dados	<p>Procurar semelhanças e disparidades entre os dados</p>
	Analizar possibilidades ou estimativas	<p>Analizar a possibilidade de um evento futuro, com base em eventos passados ou presentes</p>
<b>5- Informar resultados</b>		
Quanto aos resultados da análise exploratória		
	Categorias	Subcategorias

	Representar resultados em diversos formatos	Representar resultados em formato textual  Representar resultados em forma de tabelas  Representar resultados utilizando gráficos
Quanto aos resultados da análise dos dados		
Categorias	Subcategorias	
Quanto aos aspectos da língua portuguesa	Representar resultados utilizando coesão e coerência na informação dos resultados	Informar resultados utilizando coesão na língua portuguesa  Informar resultados utilizando coerência na linguagem
Quanto à forma de representação dos resultados	Representar resultados em formato textual	Representar resultados utilizando expressões partitivas (“a maioria”, “parte de”, “a maior parte de”) na composição do sujeito de uma oração  Representar resultados utilizando expressões que remetem a quantidades

		aproximadas (aproximadamente, aproximadamente metade, quase todos, quase 50%)
		Representar resultados por meio de frequências, contagem, ocorrências
		Representar resultados por meio de frações ou comparações (exemplo: uma pessoa em cada 10 não conseguiram se proteger do vírus)
		Representar resultados por meio de porcentagens
	Representar resultados em forma de tabela	Representar resultados em forma de tabela
	Representar resultados utilizando gráficos	Representar resultados utilizando gráficos
	Representar resultados utilizando infográficos	Representar resultados utilizando infográficos
Quanto à escolha da forma de representação dos resultados		
	Saber escolher o melhor	Saber escolher o melhor

	formato para a representação dos resultados	formato para a representação dos resultados
<b>6- Argumentar por meio de dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
	Argumentar por meio de dados	Argumentar utilizando dados com o objetivo de apresentar fatos, ideias, lógicas, provas que comprovem uma afirmação
	Argumentar acerca de previsões	Argumentar criticamente acerca de previsões e tendências
	Argumentar relacionando teoria com dados	Argumentar por meio de dados, entremeando dados e teorias para apresentar os fatos, as ideias, as razões, as provas que compõem uma afirmação, uma tese
	Argumentar relacionando dados de uma pesquisa com os dados de outras pesquisas	Argumentar relacionando os dados obtidos na pesquisa principal com dados e informações de outras pesquisas
	Argumentar relacionando	Argumentar relacionando

	<p>resultados obtidos em uma pesquisa com diferentes contextos da realidade social</p>	<p>resultados obtidos na pesquisa com diferentes contextos da realidade social, dentro de sua própria comunidade</p>
		<p>Argumentar relacionando resultados obtidos nesta pesquisa com diferentes contextos da realidade social fora da sua comunidade</p>
<b>7- Tomar decisões por meio de dados</b>		
	<b>Categorias</b>	<b>Subcategorias</b>
	<p>Tomar decisões a partir dos dados analisados até então</p>	<p>Tomar decisões a partir dos dados analisados, sem o uso de ferramentas e técnicas estatísticas que permitam a tomada de decisões mais precisas e assertivas</p>
	<p>Tomar decisões por meio de dados utilizando ferramentas e técnicas para tomada de decisões</p>	<p>Tomar decisões por meio de dados, utilizando ferramentas e técnicas estatísticas que permitam a tomada de decisões mais precisas e assertivas</p>

Fonte: reprodução da autora (2020).

O artefato operacionalizou duas das três suposições levantadas durante a sua fase de *design*. O artefato mostrou que é válida a primeira conjectura teórica levantada: **o trabalho colaborativo facilitou o processo de aprendizagem em Literacia de Dados**, uma vez que através de atitudes e valores relacionados ao uso das dimensões comunicação, cooperação e coordenação do Modelo 3C os estudantes conseguiram resolver os desafios propostos pelo artefato Cartas para Literacia. O artefato também validou a segunda conjectura teórica levantada: **através do uso de group storytelling os conhecimentos tácitos dos estudantes puderam ser formalizados**, permitindo a criação de uma escala de Literacia de Dados através da codificação das histórias contadas pelos grupos de trabalho formados durante o estudo de caso.

As tabelas do *dataset* Cores utilizadas durante o estudo de caso, refletiam as respostas da turma de graduação para perguntas referentes às suas preferências com relação à animais de estimação, jogos eletrônicos e opções de lazer, entre outros. Para a turma de pós-graduação não foi dada a oportunidade de trabalhar com um *dataset* gerado a partir de seus próprios dados, o que poderia permitir observar se houve algum impacto pelo uso de um *dataset* cuja intenção era refletir os interesses dos estudantes. Entretanto, **não se pode ter certeza de que o artefato operacionalizou a terceira conjectura teórica levantada**, apesar de a turma de graduação ter apresentado alguns resultados bem mais expressivos que a de pós-graduação aos desafios propostos durante a atividade com as cartas — resultados relacionados à atitudes de maior familiaridade e interesse sobre as informações contidas no *dataset* Cores, bem como de maior criatividade na expressão da sua Literacia de Dados — pois houveram algumas diferenças na aplicação da atividade com asas amostras que podem ter impactado os resultados. Uma diferença importante foi que os estudantes da pós-graduação não tiveram o acesso ao laboratório de informática, tendo que utilizar notebooks de sua propriedade na proporção de 1 notebook para cada 3 estudantes. Outra diferença se deve ao número de estudantes em cada um dos grupos, que era bem maior na amostra retirada da turma de graduação com relação à amostra retirada entre os estudantes da pós-graduação.

**A pergunta de pesquisa, que questionava de que forma podemos medir a Literacia de Dados de estudantes de ensino médio, foi respondida satisfatoriamente através da obtenção de uma escala de Literacia de Dados**, que, segundo a avaliação dos professores

especialistas entrevistados, permite avaliar qualitativamente competências relacionadas à Literacia de Dados de estudantes de ensino médio em 7 níveis de Literacia de Dados: 1- Coletar dados, 2- Ler dados, 3- Gerenciar dados, 4- Analisar dados, 5- Informar resultados, 6- Argumentar por meio de dados e 7- tomar decisões com base em dados (Quadro 77).

A escala de Literacia de Dados teve como ponto de partida para o seu desenvolvimento, a criação do artefato Cartas para a Literacia, que foi idealizado tendo como base a definição de Literacia de Dados fornecida por D'Ignazio e Bhargava (2015) abstraída para um processo de KDD simplificado. A Figura 28 mostra as etapas percorridas durante o desenvolvimento da escala de Literacia de Dados.

Figura 28: Etapas percorridas durante o desenvolvimento da escala de Literacia de Dados.



Fonte: reprodução da autora (2020).

Apesar de o artefato ter sido planejado com base nas etapas constantes na definição de D'Ignazio e Bhargava (2015), que previa as etapas de leitura de dados, gerenciamento de dados, análise de dados e argumentação com base em dados, a análise realizada pelos

especialistas indicaram a importância de considerar também na escala de literacia as etapas de coleta de dados, informação de resultados — que vem após a análise de dados — e de tomada de decisões baseada em dados. Desta forma foi proposta uma nova definição de Literacia de Dados, que amplia o escopo da de D'Ignazio e Bhargava (2015), adicionando as categorias “coletar dados”, “informar resultados” e “tomar decisões com base em dados”, além de definições um pouco mais aprimoradas de cada uma das etapas contidas no processo de análise descrito na definição. As definições atualizadas segundo esta pesquisa são reveladas a seguir.

**Literacia de Dados** é a capacidade de coletar dados, ler dados, gerenciar dados, analisar dados, informar resultados, argumentar com base em dados e tomar decisões com base em dados.

**Coletar dados** envolve planejar a Coleta de Dados e colher dados. Planejar a coleta de dados passa por identificar previamente a pergunta principal e as perguntas secundárias que guiarão a pesquisa, desenhar os passos metodológicos que serão percorridos na pesquisa, levantar as ferramentas que serão utilizadas e estimar o tempo necessário para a realização dos passos metodológicos. Colher os dados passa por colher dados em amostragem probabilística e amostragem não probabilística.

**Ler dados** envolve ler dados em diferentes formatos, ler dados utilizando os sentidos, ler dados com ajuda de software, ler dados utilizando comandos em linhas de código, ler dados analiticamente e ler dados comparativamente. Ler dados em diferentes formatos passa por: ler dados em tabelas, ler dados em planilhas, ler dados em gráficos, ler dados em textos e ler dados em infográficos. A leitura de dados pode ser realizada através dos órgãos dos sentidos apenas, sem ajuda de software, ou com ajuda de software, utilizando planilhas eletrônicas ou softwares especializados em análise de dados. Ler dados analiticamente é ler dados com o objetivo de compreender de forma lógica e global um conjunto de dados. Ler dados comparativamente é ler dados de forma crítica dentro do seu contexto, problematizando os dados e confrontando-os com a realidade.

**Gerenciar dados** envolve armazenar dados e tratar dados. Armazenar dados passa por salvar dados pontualmente e continuamente em um diretório no computador ou em nuvem. Tratar dados passa por limpar dados, identificando dados incompletos, incorretos e

incoerentes; transformar dados, identificando dados que estão em um formato e transformando-os para outro mais desejável e escolher qual é a melhor representação para os dados (tabelas, frases, planilhas, infográficos, etc.).

**Analizar dados** envolve analisar dados sem auxílio de computador e com ajuda de software, analisar dados utilizando codificação, planejar a análise, explorar os dados, comparar os dados e analisar possibilidades ou estimativas. Analisar dados sem auxílio do computador passa por: analisar dados utilizando os órgãos dos sentidos somente ou com ajuda de material de papelaria. Analisar dados com ajuda de software passa por analisar dados utilizando planilhas eletrônicas e softwares especializados em análise estatística. Analisar dados utilizando codificação passa por analisar dados com ajuda de linguagem de programação e linguagem de consulta. Planejar a análise passa por: desenhar os passos metodológicos que serão percorridos na análise de dados e fazer o levantamento prévio das ferramentas que serão utilizadas na análise de dados (softwares, linguagens de programação, etc.). Explorar dados passa por: examinar dados de forma gráfica, examinar medidas de tendência central, examinar medidas de variabilidade (desvio padrão, variância, desvio médio variância relativa, coeficiente de variabilidade), examinar pontos fora da curva, correlações e possibilidades. Comparar dados passa por procurar semelhanças e disparidades entre os dados. Analisar possibilidades e estimativas passa por analisar a possibilidade de um evento futuro, com base em eventos passados ou presentes.

**Informar resultados** envolve representar resultados em diversos formatos; representar resultados utilizando coesão e coerência na informação dos resultados; representar dados em formato textual, em forma de tabela, utilizando gráficos e infográficos; saber escolher o melhor formato para a representação dos resultados.

**Argumentar por meio de dados** envolve: argumentar por meio de dados com o objetivo de apresentar fatos, ideias, lógicas e provas que comprovem uma afirmação; argumentar criticamente acerca de previsões e tendências; argumentar por meio de dados, entremeando dados e teorias para apresentar os fatos, as ideias, as razões, as provas que compõem uma afirmação, uma tese; argumentar relacionando resultados obtidos na pesquisa com diferentes contextos da realidade social, dentro e fora da sua própria comunidade;

**Tomar decisões por meio de dados** envolve tomar decisões a partir dos dados analisados utilizando ou não técnicas estatísticas que permitam tomadas de decisões mais precisas e assertivas.

Foi observada também a emergência de alguns fenômenos através do uso do artefato. Observou-se os fenômenos quanto à colaboração, quanto ao aprendizado de Literacia de Dados e quanto a grupos de alunos com especificidades.

**Quanto à colaboração:** os processos de trabalho seguiram a lógica do Modelo 3C, sendo que a coordenação dos grupos de trabalho, realizada nas turmas de graduação por 1 ou 2 membros e na turma de pós-graduação por um membro do grupo, em sua maioria. A maioria dos grupos teve dificuldade em realizar as tarefas no tempo proposto, demonstrando que as tarefas não eram fáceis de serem cumpridas.

**Quanto à Literacia de Dados:** Tanto na turma de graduação quanto na turma de pós-graduação houve dúvida quanto à definição de “dado”, quando os estudantes responderam o desafio proposto na carta da Fase 1, que pedia para dar um exemplo de dado contido na tabela fornecida. Um integrante chegou a verbalizar que “[dado] poderia ser qualquer coisa”. O tipo de gráfico mais utilizado para a representação de informações quando os estudantes utilizaram software como recurso, foi o gráfico de barras, enquanto que durante o uso de material de papelaria, emergiram pictogramas, infográficos e diagramas de Venn, demonstrando que a criatividade teve terreno mais fértil para aflorar quando os estudantes se viam diante de materiais de papelaria, que permitiam uma maior liberdade de criação, com relação aos softwares que escolheram para trabalhar.

**Quanto à apresentação de necessidades educacionais especiais:** um dos estudantes, que possuía baixa visão, utilizou como recurso tecnológico o seu telefone celular para visualizar mais de perto as tabelas fornecidas e os gráficos que estavam sendo feitos pelo seu grupo. Ele olhava na tela do computador e depois aproximava o celular do rosto. Esse estudante assumiu papel de liderança no seu grupo, em alguns momentos figurando com uma régua grande na mão. Ele, de pé, debatia intensamente sobre a atividade com seus colegas e tomava as decisões do grupo junto com um de seus colegas que, sentado, ocupava um computador à sua frente.

No final de todo o processo de obtenção de resultados narrado, obteve-se uma

generalização do processo de desenvolvimento da escala de Literacia de Dados. A generalização possui oito passos (Figura 29): 1- Identificação da definição; 2- Abstração para desafios; 3- *Design* de atividade; 4- Codificação; 5- Desenvolvimento da escala; 6- Avaliação; 7- Ajustes; 8- Apresentação dos produtos. Através dos passos listados, em teoria, é possível construir uma escala qualitativa de habilidades, competências e valores relacionada a um tema de interesse escolhido, partindo da definição desse campo de interesse.

Figura 29: Framework que representa a generalização do processo de desenvolvimento da escala de Literacia de Dados.



Fonte: reprodução da autora (2020).

Isto é, para criar uma escala qualitativa para a avaliação de competências a partir da definição de uma determinada área de aprendizagem, deve-se realizar as etapas da Figura 29, descritas a seguir.

- 1. Identificação da definição:** Identificação de uma definição, escrita através de uma construção linguística em forma de processo, ou seja, uma definição de uma área de estudos que seja descrita em etapas.
- 2. Abstração para desafios:** Abstração das etapas contidas na definição em desafios que correspondam a cada uma das etapas e mais um desafio, que corresponda a uma tarefa de relato das ações realizadas pelos estudantes para a resolução de cada um dos desafios propostos.
- 3. Design de atividade:** Criação de uma atividade com os desafios e a tarefa definidas na etapa anterior e a aplicação dessa atividade com estudantes organizados em grupos, realizando ciclos de estudo de caso.
- 4. Codificação:** Codificação dos relatos dos estudantes através da Teoria Fundamentada em Dados e descoberta das categorias representativas das competências, habilidades e valores que comporão a escala de avaliação que se deseja criar.
- 5. Desenvolvimento da escala:** Desenvolvimento de uma escala representando os níveis, subníveis, categorias e eixos descritivos dos conhecimentos, habilidades, competências e valores relacionados à área de estudo.
- 6. Avaliação da escala:** Avaliação da escala através de entrevistas semiestruturadas com especialistas e professores na área de estudo.
- 7. Realização de ajustes:** Implementação das alterações e ajustes apontados pelos professores e especialistas na etapa de avaliação da escala.
- 8. Apresentação do produto / Escala:** Apresentação da escala desenvolvida e das novas definições geradas para cada nível do processo descrito na definição original identificada na etapa 1. Apresentação das considerações dos pesquisadores acerca do processo de colaboração experimentado no desenvolvimento da escala.

## 6 CONCLUSÃO

---

Neste capítulo resumimos os resultados alcançados, as limitações do trabalho e indicamos os próximos passos desta pesquisa.

A partir do entendimento de que a cidadania no contexto da sociedade de dados depende da promoção de esforços educacionais para ampliar a capacidade dos indivíduos em lidar com dados, foi conduzida uma investigação inicial que identificou algumas instituições dentro e fora do país que têm conduzido tarefas no sentido de avaliar a Literacia de Dados de estudantes e colaboradores. Apesar disso, ainda existem lacunas com relação à existência de escalas que permitam a identificação de competências relacionadas à Literacia de Dados em estudantes de ensino médio, que se manifestam, por exemplo, na ausência de iniciativas curriculares voltadas para a Literacia de Dados na Base Nacional Comum Curricular (Brasil, 2017).

Com a identificação do problema de pesquisa foi desenvolvido o artefato Cartas para a Literacia, tendo como suporte o paradigma metodológico-epistemológico da *Design Science Research* e do mapa de elementos da DSR de Pimentel *et al.* (2019). O artefato foi construído com base na definição de Literacia de Dados dos autores D'Ignazio e Bhargava (2015) e em 3 suposições teóricas. A definição dos autores foi escolhida por apresentar um formato em etapas, que poderia ser abstraído em um processo de Descoberta de Conhecimento em Bases de Dados. As suposições teóricas que deram suporte à construção do artefato foram: (i) conhecimentos de um grupo podem ser explicitados através do trabalho colaborativo (ii) utilizando a técnica de *group storytelling* (iii) associada a um *dataset* que desperte o interesse dos estudantes.

Uma vez desenhado o artefato, foram realizados dois ciclos de estudo de caso em que se aplicou uma atividade cronometrada contando com a participação de amostras de estudantes de um curso de graduação e outro de pós-graduação divididas em grupos. Desta forma foi possível obter dados resultantes da observação direta realizada durante a atividade proposta, bem como dados relativos às respostas dos estudantes aos desafios, de forma suficiente para realizar uma análise qualitativa apoiada na abordagem construtivista da Teoria Fundamentada em Dados (Charmaz, 1983).

Através da extração de conhecimento por meio da codificação dos resultados das atividades, obteve-se a versão inicial de uma escala qualitativa para avaliar a presença de

competências em Literacia de Dados no contexto do ensino médio, mais uma nuvem de palavras revelando os termos mais frequentes relativos ao processo de colaboração experimentado pelos grupos durante o estudo de caso.

Como resultado da avaliação do artefato Cartas para a Literacia, considerou-se que colaboração aliada à *group storytelling* podem fornecer meios para a explicitação dos conhecimentos tácitos dos estudantes em Literacia de Dados, considerando a manifestação das dimensões do modelo 3C, de Fuks, *et al.* (2011) durante a atividade de análise de dados em grupos, com o uso das cartas de desafio. Especialistas das áreas de matemática e estatística atuantes em instituições de ensino superior realizaram considerações acerca da escala de Literacia de Dados gerada através da atividade realizada com o artefato Cartas para a Literacia, apontando limitações e oportunidades de melhoria. Ao serem questionados sobre a validade da escala, todos responderam que a proposta é pertinente diante da ausência de um instrumento formal para a medição da Literacia de Dados de seus estudantes, que utilizariam a escala em suas aulas e que o instrumento propõe a avaliação de competências de Literacia de Dados adequadas para serem trabalhadas no ensino médio.

## 6.1 CONTRIBUIÇÕES

Como contribuições desta pesquisa pode-se destacar que:

- A pergunta de pesquisa foi respondida de forma satisfatória, isto é, pode-se avaliar em estudantes de ensino médio a presença de 60 competências em Literacia de Dados, distribuídas entre os 7 níveis através da escala qualitativa desenvolvida através do estudo conduzido.
- O artefato Cartas para a Literacia se mostrou um instrumento válido para responder à pergunta de pesquisa, uma vez que, através dele, os conhecimentos tácitos em Literacia de Dados dos grupos de estudantes pesquisados tornaram-se explícitos, podendo ser utilizado em sala de aula com diversas finalidades pedagógicas.
- A escala desenvolvida permitiu a atualização das definições de Literacia de Dados, Leitura de Dados, Gerenciamento de Dados (ou Trabalho com Dados), Análise de Dados, e Argumentação com base em dados. Além disso, a escala permitiu escrever

definições para Coleta de Dados, Informação de Resultados e Tomada de Decisão por Meio de Dados.

- Foi possível generalizar o processo de desenvolvimento da escala de Literacia de Dados em um *framework* em 8 passos para a criação de escalas a partir de definições de temas de interesse que estejam escritos em forma de processos.
- Quanto à colaboração entre os estudantes, foi possível identificar características e processos relativos à colaboração experimentados pelos grupos durante os ciclos de estudo de caso.

## 6.2 LIMITAÇÕES DA PROPOSTA

- Não foi possível afirmar que o artefato operacionalizou a terceira conjectura teórica levantada — “O processo de aprendizagem pode ser facilitado através do uso de um *dataset* que reflete os interesses dos estudantes” — uma vez que não foram realizados testes utilizando outros *datasets* que não o Cores com a turma de graduação, com o objetivo de estabelecer um quadro comparativo de performance. Também não foi possível utilizar um *dataset* a partir dos dados dos estudantes da turma de pós-graduação.
- O *design* do artefato e o formato em que a atividade foi apresentada não garantia meios para a plena participação de estudantes portadores de necessidades especiais educacionais.
- O trabalho colaborativo experimentado pelos grupos de estudantes poderia ser melhor investigado, em busca de um entendimento mais aprofundado sobre as dimensões do modelo 3C e das situações, características e processos envolvidos na colaboração.
- A turma de graduação foi mais favorecida com relação a algumas condições de aplicação da atividade, como o acesso a um computador por estudante em um laboratório de informática e quantidade de integrantes por volta de 8 nos grupos de trabalho formado, enquanto que na turma de pós-graduação os grupos eram de 3 estudantes, com 1 notebook para cada grupo.

### 6.3 TRABALHOS FUTUROS

Apesar de haver na literatura alguns esforços institucionais direcionados para a avaliação da Literacia de Dados de estudantes, esse tema é motivo de preocupação mundial tanto de gestores como de educadores, devido à falta de políticas públicas e estratégias de gestão educacional que possam dar suporte à promoção e garantia da cidadania através da Literacia de Dados. Esta pesquisa apresentou o desenvolvimento de uma escala que permite avaliar a existência de competências em Literacia de Dados em estudantes do ensino médio, conforme foi apresentado no Capítulo 5 deste trabalho. Todavia, esta pesquisa representa uma primeira etapa de investigações sobre gestão e representação de habilidades e competências e atitudes em Literacia de Dados, uma vez que o seu foco esteve em, através da externalização de conhecimentos de estudantes, identificar conceitos, competências, habilidades e atitudes que poderiam compor uma escala qualitativa de Literacia de Dados.

Dessa forma, é possível recomendar:

- a produção de novos ciclos de estudo de caso, a fim de avaliar a Literacia de Dados de estudantes de ensino médio utilizando a escala desenvolvida neste estudo, a fim de cumprir etapa de avaliação do artefato que não pôde ser realizada com os estudantes devido ao contexto da pandemia de COVID-19.
- a realização de teste de validade da proposta de *framework* apresentada na Figura 29, que representa a generalização do processo de desenvolvimento da construção de escalas para a construção de outras escalas, em campos de conhecimento distintos.
- A adaptação do artefato Cartas para a Literacia para a aplicação e estudo em contextos em que haja estudantes com necessidades educacionais especiais.
- a criação de um dicionário de termos de Literacia de Dados com as definições envolvidas na área de estudo.
- a realização de estudos para avaliar a profundidade em que um estudante possui domínio de cada uma das competências listadas na escala de Literacia de Dados que foi apresentada. Acredita-se que este

tipo de avaliação pode servir como instrumento pedagógico para promover um entendimento mais preciso de quais ações são necessárias para o atingimento de um nível de Literacia de Dados a ser estabelecido como meta para estudantes e grupos de estudantes em determinada etapa escolar ou acadêmica e também para profissionais que desempenham suas atividades no mercado de trabalho.

- o aprofundamento do estudo da codificação realizada durante a fase de análise com a Teoria Fundamentada em Dados com o objetivo de compreender os padrões de linguagem capazes de revelar a presença de determinadas competências em Literacia de Dados no discurso dos estudantes, obtido através da *group storytelling* de dados.
- Criação de um sistema de gestão de conhecimento que realize avaliação da Literacia de Dados de indivíduos de forma automatizada, considerando o alcance de níveis de atendimento às distintas competências listadas na escala de Literacia de Dados.

## REFERÊNCIAS

---

- ALA-MUTKA, Kirsti. **Mapping digital competence**: towards a conceptual understanding. Sevilla: Institute for Prospective Technological Studies, 2011.
- APARICIO, Manuela; COSTA, Carlos J. Collaborative systems: characteristics and features. In: ACM INTERNATIONAL CONFERENCE ON DESIGN OF COMMUNICATION, 30., 2012, Seattle. **Proceedings [...]**. New York: ACM, 2012. p. 141-146.
- ASKEW, Mike. **Effective teachers of numeracy**. London: King's College London, 1997.
- BECKER, Jörg; KNACKSTEDT, Ralf; PÖPPELBUß, Jens. Developing maturity models for IT management. **Business & Information Systems Engineering**, [S. l.], v. 1, n. 3, p. 213-222, 2009.
- BHARGAVA, Rahul et al. **Beyond data literacy**: reinventing community engagement and empowerment in the age of data. 2015. Disponível em:  
<https://hdl.handle.net/1721.1/123471>. Acesso em: 02 dez. 2019.
- BLUMER, Herbert et al. **El interaccionismo simbólico**: perspectiva y método. Barcelona: HORA, 1982.
- BORGES, Robson M. et al. Usando o modelo 3C de colaboração e Vygotsky no ensino de programação distribuída em pares. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 28., 2007, São Paulo. **Proceedings [...]**. [S. l.: s.n.], 2007. p. 51-60.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO. (MEC/SEED). **Base Nacional Comum Curricular**. 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: 02 dez. 2019.
- CARTER, Nancy. The use of triangulation in qualitative research. **Oncology Nursing Forum**, Pittsburgh, v. 41, n. 5, p. 545-547, 2014.
- CARVALHO, Virgínia D. de; BORGES, Livia de O.; RÉGO, Denise Pereira do. Interacionismo simbólico: origens, pressupostos e contribuições aos estudos em Psicologia Social. **Psicologia: ciência e profissão**, Brasília, DF, v. 30, n. 1, p. 146-161, 2010.
- CENSO EAD.BR: relatório analítico da aprendizagem a distância no Brasil. São Paulo: Pearson Education do Brasil. Disponível em: [http://abed.org.br/censoEADbr0809\\_portugues.pdf](http://abed.org.br/censoEADbr0809_portugues.pdf). Acesso em: 22 set. 2021.
- CHARMAZ, Kathy. Loss of self: a fundamental form of suffering in the chronically ill. **Sociology of Health & Illness**, Boston, v. 5, n. 2, p. 168-195, 1983.
- CHARMAZ, Kathy. **Constructing grounded theory**: a practical guide through qualitative analysis. [S. l.]. Los Angeles: Sage, 2006.

CHIESA, V. et al. Designing a performance measurement system for the research activities: A reference framework and an empirical study. **Journal of Engineering and Technology Management**, Amsterdam, v. 25, n. 3, p. 213-226, 2008.

CHUGH, Ritesh. Do Australian Universities Encourage Tacit Knowledge Transfer? In: INTERNATIONAL JOINT CONFERENCE ON KNOWLEDGE DISCOVERY, KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT, 7., 2015, Lisboa. **Proceedings [...]**. [S, I.: s.n.], 2015. p. 128-135.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Consulta Parametrizada**. 2019. Disponível em:  
[http://dgp.cnpq.br/dgp/faces/consulta/consulta\\_parametrizada.jsf](http://dgp.cnpq.br/dgp/faces/consulta/consulta_parametrizada.jsf). Acesso em: 02 dez. 2019.

COCKING, Rodney R.; MESTRE, Jose P. **Cognitive Science**. 1989. Disponível em:  
<https://files.eric.ed.gov/fulltext/ED307104.pdf>. Acesso em: 02 dez. 2019.

COSTA, António Pedro; LOUREIRO, Maria João; REIS, Luís Paulo. Do Modelo 3C de Colaboração ao Modelo 4C: Modelo de Análise de Processos de Desenvolvimento de Software Educativo. **Revista Lusófona de Educação**, Lisboa, n. 27, p. 181-200, 2014.

CRESWELL, John W.; POTTH, Cheryl N. **Qualitative inquiry and research design: choosing among five approaches**. Los Angeles: Sage, 2016.

D'IGNAZIO, Catherine; BHARGAVA, Rahul. Approaches to building big data literacy. In: BLOOMBERG DATA FOR GOOD EXCHANGE CONFERENCE, 2015, New York. **Proceedings [...]**. [S, I.: s.n.], 2015.

DRESCH, Aline; LACERDA, Daniel Pacheco; JÚNIOR, José Antonio Valle Antunes. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. Porto Alegre: Bookman, 2015.

DUARTE, Teresa. **A possibilidade da investigação a 3: reflexões sobre triangulação (metodológica)**. Lisboa: CIES, 2009.

ELLIS, Clarence A.; GIBBS, Simon J.; REIN, Gail. Groupware: some issues and experiences. **Communications of the ACM**, New York, v. 34, n. 1, p. 39-58, 1991.

FAIRCLOUGH, N. **Language and Power**. London: Longman, 1989.

\_\_\_\_\_. **Critical Discourse Analysis: The Critical Study of Language**. London: Longman, 1995.

\_\_\_\_\_. **Discurso de mudança social**. Brasília, DF: Editora Universidade de Brasília, 2008.

FELDON, David F. Cognitive load and classroom teaching: the double-edged sword of automaticity. **Educational Psychologist**, Hillsdale, v. 42, n. 3, p. 123-137, 2007.

FLICK, Uwe. Triangulation in qualitative research. In: FLICK, Uwe; KARDORFF, Ernst von; STEINKE, Ines. **A companion to qualitative research**. London: Sage Publications, 2004. p. 178-183.

FRANK, Mark et al. Data Literacy-What is it and how can we make it happen? **The Journal of Community Informatics**, [S. l.], v. 12, n. 3, p. 4-8, 2016.

FRAZEE, J.; KREUTER, K. **Discovering Orff**: A Curriculum for Music Teachers. London: Schott, 1987.

FREIRE, Paulo. **Pedagogia da Autonomia**: saberes necessários à prática educativa. 25. ed. São Paulo: Paz e Terra, 1996. 76 p.

\_\_\_\_\_. **Pedagogia do Oprimido**. 37. ed. São Paulo: Paz e Terra, 1970.

FUKS, Hugo et al. Teorias e modelos de colaboração. In: PIMENTE, Mariano (Org.). **Sistemas colaborativos**. Rio de Janeiro: Elsevier, 2011. p. 16-33.

GARGANO, Leonardo; VIVACQUA, Adriana; MONTEIRO, Rodrigo. Colaboração Através do Reuso de Ontologias. In: SIMPÓSIO BRASILEIRO DE SISTEMAS COLABORATIVOS, 15., 2019, Rio de Janeiro. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2019. p. 29-33.

GLASER, Barney G.; STRAUSS, Anselm L. Awareness contexts and social interaction. **American Sociological Review**, Menasha, v. 29, n. 5, p. 669-679, 1964.

\_\_\_\_\_. **The Discovery of Grounded Theory**: strategies for qualitative research. New Jersey: Aldine Transaction, 1967. 262 p.

GHOSH, Abhishek. **What is Data Literacy And Why Data Literacy Matters**. 2018. Disponível em: <https://thecustomizewindows.com/2018/09/what-is-data-literacy-and-why-data-literacy-matters/>. Acesso em: 01 jun. 2019.

GIBSON, J. Phil; MOURAD, Teresa. The growing importance of data literacy in life science education. **American Journal of Botany**, Lancaster, v. 105, n. 12, p. 1953-1956, 2018.

GROUNDED THEORY SOLUTIONS. **Grounded Theory Online**. 2019. Disponível em: <http://www.groundedtheoryonline.com/what-is-grounded-theory>. Acesso em: 07 dez. 2019.

GUMMER, Edith; MANDINACH, Ellen. Building a Conceptual Framework for Data Literacy. **Teachers College Record**, New York, v. 117, n. 4, 2015.

HEER, Jeffrey; BOSTOCK, Michael. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2010, Atlanta. **Proceedings [...]**. [S. l.: s. n.], 2010. p. 203-212.

HEINSFELD, Bruna Damiana; SILVA, Maria Paula Rossi Nascentes da. As versões da Base Nacional Comum Curricular (BNCC) e o papel das tecnologias digitais: conhecimento da

técnica versus compreensão dos sentidos. **Curriculum sem Fronteiras**, v. 18, n. 2, p. 668-690, 2018.

HOELTER, Lynette. Introduction to Statistical Literacy. In: FONTICHIARO, Kristin; OEHRLI, Jo Angela; LENNEX Amy (Ed.). **Creating Data Literate Students**. Ann Arbor: Maize Books, 2017.

HU, Jianfei; SHAO, Hua; CAO, Xin. Evaluation and Analysis of College Students' Information Literacy in the New Media Age. In: INTERNATIONAL CONFERENCE ON ECONOMICS, MANAGEMENT AND HUMANITIES SCIENCE, 5., 2019, Bangkok. **Proceedings [...]**. [S. I.: s.n.], 2019. p. 1139-1142.

INDEPENDENT EXPERT ADVISORY GROUP. **A World That Counts**: mobilizing the data revolution for sustainable development. 2014. Disponível em: <https://www.undatarevolution.org/report/>. Acesso em: 18 set. 2021.

JAQUES-DALCROZE, Émile. **Rhythm, Music and Education**. New York: G. P. Putnam's Sons, 1921.

JEON, Yun-Hee. The application of grounded theory and symbolic interactionism. **Scandinavian Journal of Caring Sciences**, [S. I.], v. 18, n. 3, p. 249-256, 2004.

KHAN, Hammad R.; KIM, Jeonghyun; CHANG, Hsia-Ching. **Toward an Understanding of Data Literacy**. 2018. Disponível em: [https://www.ideals.illinois.edu/bitstream/handle/2142/100243/Khan-Hammad\\_20180417\\_V01.pdf?isAllowed=y&sequence=1](https://www.ideals.illinois.edu/bitstream/handle/2142/100243/Khan-Hammad_20180417_V01.pdf?isAllowed=y&sequence=1). Acesso em: 18 set. 2020.

KJELVIK, Melissa K.; SCHULTHEIS, Elizabeth H. Getting Messy with Authentic Data: Exploring the Potential of Using Data from Scientific Research to Support Student Data Literacy. **CBE—Life Sciences Education**, [S. I.], v. 18, n. 2, p. es2, 2019.

KNAFLIC, Cole Nussbaumer. **Storytelling com dados**: um guia sobre visualização de dados para profissionais de negócios. Rio de Janeiro: Alta Books, 2019.

LAGE, Bruno et al. Aplicando o group storytelling no compartilhamento de experiências e na avaliação em sala de aula. In: WORKSHOP DE INFORMÁTICA NA ESCOLA, 23., 2017, Recife. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2017. p. 165-174.

LE BOTERF, Guy. **De la compétence**: essai sur un attracteur étrange. [S. I.]: Editions d'Organisation, 1994.

LOCKE, Brandon T.; HEPPLER, Jason A. Teaching Data Literacy for Civic Engagement: Resources for Data Capture and Organization. **KULA: knowledge creation, dissemination, and preservation studies**, [S. I.], v. 2, n. 1, p. 23, 2018.

LUPI, Giorgia; POSAVEC, Stefanie. **Dear data**. London: Particular Books, 2016.

MAGNUSSON, Shirley; KRAJCIK, Joseph; BORKO, Hilda. Nature, sources, and development of pedagogical content knowledge for science teaching. In: GEES-NEWSOME, J.; LEDERMAN, N. G. (Ed.). **Examining pedagogical content knowledge**. Dordrecht: Springer, 1999. p. 95-132.

MAIER, A. M.; MOULTRIE, J.; CLARKSON, P. J. Developing maturity grids for assessing organisational capabilities: practitioner guidance. **IEEE Transactions on Engineering Management**, [S. l.], v. 59, n. 1, p. 138-159, 2012.

MALONE, T. W.; CROWSTON, K. What is coordination theory and how can it help design cooperative work systems? In: ACM CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK, 1990, Los Angeles. **Proceedings [...]**. [S. l.: s.n.], 1990. p. 357-370.

MANDINACH, Ellen B. et al. Ethical and appropriate data use requires data literacy. **Phi Delta Kappan**, Bloomington, v. 96, n. 5, p. 25-28, 2015.

MERRIAM, Sharan B.; TISDELL, Elizabeth J. **Qualitative research: a guide to design and implementation**. San Francisco: John Wiley & Sons, 2015.

MONTOYA, Silvia. **Defining Literacy**. 2018. Disponível em: [http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/12/4.6.1\\_07\\_4.6-defining-literacy.pdf](http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/12/4.6.1_07_4.6-defining-literacy.pdf). Acesso em: 17 out. 2018.

MÖRTERS, Peter; PERES, Yuval. **Brownian motion**. Cambridge: Cambridge University Press, 2010.

OLIVER-HOYO, Maria; ALLEN, DeeDee. The Use of Triangulation Methods in Qualitative Educational Research. **Journal of College Science Teaching**, [S. l.], v. 35, n. 4, p. 42-47, 2006.

OXFORD LEARNER DICTIONARIES. Maturidade. 2019. Disponível em: [https://www.oxfordlearnersdictionaries.com/definition/american\\_english/maturity#:~:targetText=1the%20quality%20of%20thinking,has%20maturity%20beyond%20his%20years.&targetText=\(of%20a%20person%2C%20an%20animal,100%20years%20to%20reach%20maturity.](https://www.oxfordlearnersdictionaries.com/definition/american_english/maturity#:~:targetText=1the%20quality%20of%20thinking,has%20maturity%20beyond%20his%20years.&targetText=(of%20a%20person%2C%20an%20animal,100%20years%20to%20reach%20maturity.) Acesso em: 02 dez. 2019.

PEFFERS, Ken et al. A design science research methodology for information systems research. **Journal of Management Information Systems**, Armonk, v. 24, n. 3, p. 45-77, 2007.

PIAGET, Jean. **A psicologia da inteligência**. Rio de Janeiro: Vozes, 2013.

PIMENTEL, Mariano et al. Modelo 3C de Colaboração para o desenvolvimento de Sistemas Colaborativos. In: SIMPÓSIO BRASILEIRO DE SISTEMAS COLABORATIVOS, 3., 2006, Natal. **Anais [...]**. [S. l.: s.n.], 2006. p. 58-67.

PIMENTEL, Mariano; FILIPPO, Denise; SANTORO, Flávia Maria. Design Science Research: fazendo pesquisas científicas rigorosas atreladas ao desenvolvimento de artefatos computacionais projetados para a educação. In: JAQUES, Patrícia Augustin et al. (Org.)

**Metodologia de pesquisa em informática na educação:** concepção da pesquisa. Porto Alegre: SBC, 2019.

PÖPPELBUß, Jens; RÖGLINGER, Maximilian. What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management. In: EUROPEAN CONFERENCE ON INFORMATION SYSTEMS, 2011, Helsinki. **Proceedings** [...]. [S. l.: s.n.], 2011.

RIDSDALE, Chantel et al. **Strategies and best practices for data literacy education:** knowledge synthesis report. [S. l.]: Dalhousie University, 2015.

ROCHA, Edmilson Barcelos; PIMENTEL, Mariano; DINIZ, Morgana Carmen. Desenvolvimento de um Modelo da Participação em Bate papo seguindo a abordagem Design Science Research. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO, 10., 2014, Londrina. **Anais** [...]. Porto Alegre: Sociedade Brasileira de Computação, 2014. p. 32-43.

ROIGER, Richard J.; GEATZ, M. W. **Data mining:** a tutorial-based primer. Boston: Addison Wesley, 2003.

ROSS, Jeanne W. Creating a strategic IT architecture competency: learning in stages. **MIS Quarterly Executive**, [S. l.], v. 2 n. 1, p. 31-43, 2003.

ROUSE, Margaret. What is data literacy? 2019. Disponível em:  
<https://whatis.techtarget.com/definition/data-literacy>. Acesso em: 01 jun. 2019.

SANTOS, Thiago Marcondes et al. Uma roupa computacional para apoiar a educação musical no contexto da sala de aula. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO MUSICAL, 14., 2013, Brasília, DF. **Anais** [...]. Porto Alegre: Sociedade Brasileira de Computação, 2013. p. 201-212.

SANTOS, Thiago Marcondes; PIMENTEL, Mariano; FILIPPO, Denise. Tapetes Musicais Inteligentes: Computação Ubíqua para apoiar a Educação Musical. **Revista Brasileira de Informática na Educação**, Florianópolis, v. 24, n. 2, p. 123-142, 2016.

SHELLEY, Ed. **SaaS Metrics Refresher #8:** Data Literacy. 2019. Disponível em:  
<https://blog.chartmogul.com/saas-metrics-refresher-data-literacy/>. Acesso em: 01 jun. 2019.

SIMON, Hebert A. **The Sciences of the Artificial**. Cambridge: MIT Press, 1969.

SLATER, D. **Research results part 1:** Defining data literacy. 2016. Disponível em:  
<https://schoolofdata.org/2016/01/08/research-results-part-1-defining-data-literacy/>. Acesso em: 01 jun. 2019.

STEPHENSON, Elizabeth; CARAVELLO, Patti Schifter. Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project. **Reference Services Review**, Ann Arbor, v. 35, n. 4, p. 525-540, 2007.

STERNKOPF, Helena; MUELLER, Roland M. Doing Good with Data: Development of a Maturity Model for Data Literacy in Non-governmental Organizations. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 51., 2018, Waikoloa Village. **Proceedings [...]**. [S. l.: s.n.], 2018.

STRAUSS, Anselm; CORBIN, Juliet. **Basics of qualitative research**. Newbury Park: Sage, 1990.

\_\_\_\_\_. **Basics of qualitative research techniques**. Thousand Oaks: Sage, 1998.

SWANWICK, Keith. **Teaching Music Musically**. New York: Routledge, 2007.

SWELLER, John. Cognitive load during problem solving: effects on learning. **Cognitive Science**, Norwood, v. 12, n. 2, p. 257-285, 1988.

TABLEAU. **2019 Tendências de Business Intelligence**. 2019. Disponível em: <https://www.tableau.com/pt-br/reports/business-intelligence-trends>. Acesso em: 18 set. 2021.

TYGEL, Alan Freihof; KIRSCH, Rosana. Contributions of Paulo Freire for a Critical Data Literacy: a Popular Education Approach. **The Journal of Community Informatics**, [S. l.], v. 12, n. 3, p. 108-121, 2016.

UGULINO, Wallace et al. Dos processos de colaboração para as ferramentas: a abordagem de desenvolvimento do projeto CommunicaTEC. In: BRAZILIAN SYMPOSIUM ON MULTIMEDIA AND THE WEB, 14., 2008, Vila Velha. **Proceedings [...]**. [S.I.: s.n.], 2008. p. 233-240.

VERGARA, Sylvia Constant; CALDAS, Miguel Pinto. Paradigma interpretacionista: a busca da superação do objetivismo funcionalista nos anos 1980 e 1990. **RAE-Revista de Administração de Empresas**, Rio de Janeiro, v. 45, n. 4, p. 66-72, 2005.

VERMAAS, Pieter et al. **A philosophy of technology**: from technical artefacts to sociotechnical systems. [S. l.]: Morgan & Claypool, 2011.

VIVACQUA, Adriana Santarosa et al. Ontologia de colaboração. **Sistemas Colaborativos**, v. 1, 2011.

VYGOTSKY, Lev Semenovich. **Mind in Society**: the development of higher psychological processes. London: Harvard University Press, 1978.

WUEST, Judith. Feminist grounded theory: an exploration of the congruency and tensions between two traditions in knowledge discovery. **Qualitative Health Research**, Thousand Oaks, v. 5, n. 1, p. 125-137, 1995.

YAMAUTI, Rogerio Koji; TSUNODA, Denise Fukumi. Análise bibliométrica das publicações sobre Alfabetização em Dados. In: SEMINÁRIO HISPANO-BRASILEIRO DE PESQUISA EM INFORMAÇÃO, DOCUMENTAÇÃO E SOCIEDADE, 8., 2019, São Paulo. **Anais [...]**. [S. l.], 2019.

YIN, Robert K. **Case study research**: design and methods. Los Angeles: Sage Publications, 2009.

ZUBIAGA, Arkaitz; MAC NAMEE, Brian. Graphical Perception of Value Distributions: An Evaluation of Non-Expert Viewers' Data Literacy. **The Journal of Community Informatics**, [S. l.], v. 12, n. 3, p. 138-159, 2016.

## Anexos

---

No Anexo A são apresentadas as tabelas resultantes da compilação dos objetos de conhecimento e habilidades relacionadas à literacia de dados para o ensino fundamental, presentes na Base Nacional Comum curricular (Brasil, 2017).

### ANEXO A - TABELAS COM AS UNIDADES TEMÁTICAS DA BNCC RELACIONADAS À LITERACIA DE DADOS

Tabela i: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 1º ano do Ensino Fundamental.

<b>Unidade temática</b>	<b>Objetos de conhecimento</b>	<b>Habilidades</b>
Probabilidade e estatística	Noção de acaso	(EF01MA20) Classificar eventos envolvendo o acaso, tais como “acontecerá com certeza”, “talvez aconteça” e “é impossível acontecer”, em situações do cotidiano.
	Leitura de tabelas e de gráficos de colunas simples	(EF01MA21) Ler dados expressos em tabelas e em gráficos de colunas simples.
	Coleta e organização de informações	(EF01MA22) Realizar pesquisa, envolvendo até duas variáveis categóricas de seu interesse e universo de até 30 elementos, e organizar dados por meio de representações pessoais.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela ii: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 2º ano do Ensino Fundamental.

<b>Unidade temática</b>	<b>Objetos de conhecimento</b>	<b>Habilidades</b>
Probabilidade e estatística	Análise da ideia de aleatório em situações do cotidiano	(EF02MA21) Classificar resultados de eventos cotidianos aleatórios como “pouco prováveis”, “muito prováveis”, “improváveis” e “impossíveis”.
	Coleta, classificação e representação de dados em tabelas simples e de dupla entrada e em gráficos de colunas	(EF02MA22) Comparar informações de pesquisas apresentadas por meio de tabelas de dupla entrada e em gráficos de colunas simples ou barras, para melhor compreender aspectos da realidade próxima.  (EF02MA23) Realizar pesquisa em universo de até 30 elementos, escolhendo até três variáveis categóricas de seu interesse, organizando os dados coletados em listas, tabelas e gráficos de colunas simples.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela iii: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 3º ano do Ensino Fundamental.

<b>Unidade temática</b>	<b>Objetos de conhecimento</b>	<b>Habilidades</b>
Probabilidade e estatística	Análise da ideia de acaso em situações do	(EF03MA25) Identificar, em eventos familiares aleatórios, todos os resultados possíveis,

	cotidiano: espaço amostral	estimando os que têm maiores ou menores chances de ocorrência.
	Leitura, interpretação e representação de dados em tabelas de dupla entrada e gráficos de barras	(EF03MA26) Resolver problemas cujos dados estão apresentados em tabelas de dupla entrada, gráficos de barras ou de colunas. (EF03MA27) Ler, interpretar e comparar dados apresentados em tabelas de dupla entrada, gráficos de barras ou de colunas, envolvendo resultados de pesquisas significativas, utilizando termos como maior e menor frequência, apropriando-se desse tipo de linguagem para compreender aspectos da realidade sociocultural significativos.
	Coleta, classificação e representação de dados referentes a variáveis categóricas, por meio de tabelas e gráficos	(EF03MA28) Realizar pesquisa envolvendo variáveis categóricas em um universo de até 50 elementos, organizar os dados coletados utilizando listas, tabelas simples ou de dupla entrada e representá-los em gráficos de colunas simples, com e sem uso de tecnologias digitais.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela iv: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 4º ano do Ensino Fundamental.

Unidade temática	Objetos de conhecimento	Habilidades
Probabilidade e estatística	Análise de chances de eventos aleatórios	(EF04MA26) Identificar, entre eventos aleatórios cotidianos, aqueles que têm

		maior chance de ocorrência, reconhecendo características de resultados mais prováveis, sem utilizar frações.
	Leitura, interpretação e representação de dados em tabelas de dupla entrada, gráficos de colunas simples e agrupadas, gráficos de barras e colunas e gráficos pictóricos	(EF04MA27) Analisar dados apresentados em tabelas simples ou de dupla entrada e em gráficos de colunas ou pictóricos, com base em informações das diferentes áreas do conhecimento, e produzir texto com a síntese de sua análise.
	Diferenciação entre variáveis categóricas e numéricas  Coleta, classificação e representação de dados de pesquisa realizada	Realizar pesquisa envolvendo variáveis categóricas e numéricas e organizar dados coletados por meio de tabelas e gráficos de colunas simples ou agrupadas, com e sem uso de tecnologias digitais.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela v: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 5º ano do Ensino Fundamental.

Unidade temática	Objetos de conhecimento	Habilidades
Probabilidade e estatística	Espaço amostral: análise de eventos aleatórios	(EF05MA22) Apresentar todos os possíveis resultados de um experimento aleatório, estimando se esses resultados são igualmente prováveis ou não

	Cálculo de probabilidade de eventos equiprováveis	(EF05MA23) Determinar a probabilidade de ocorrência de um resultado em eventos aleatórios, quando todos os resultados possíveis têm a mesma chance de ocorrer (equiprováveis).
	Leitura, coleta, classificação e interpretação de dados em tabelas de dupla entrada, gráfico de colunas agrupadas, gráficos pictóricos e gráficos de linhas	(EF05MA24) Interpretar dados estatísticos apresentados em textos, tabelas e gráficos (colunas ou linhas), referentes a outras áreas do conhecimento ou a outros contextos, como saúde e trânsito, e produzir textos com o objetivo de sintetizar conclusões.  (EF05MA25) Realizar pesquisa envolvendo variáveis categóricas e numéricas, organizar dados coletados por meio de tabelas, gráficos de colunas, pictóricos e de linhas, com e sem uso de tecnologias digitais, e apresentar texto escrito sobre a finalidade da pesquisa e a síntese dos resultados.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela vi: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 6º ano do Ensino Fundamental.

Unidade temática	Objetos de conhecimento	Habilidades
Probabilidade estatística	Cálculo de probabilidade como a razão entre o número de resultados	(EF06MA30) Calcular a probabilidade de um evento aleatório, expressando-a por número racional (forma fracionária,

	<p>favoráveis e o total de resultados possíveis em um espaço amostral equiprovável</p> <p>Cálculo de probabilidade por meio de muitas repetições de um experimento (frequências de ocorrências e probabilidade frequentista)</p>	<p>decimal e percentual) e comparar esse número com a probabilidade obtida por meio de experimentos sucessivos.</p>
	<p>Leitura e interpretação de tabelas e gráficos (de colunas ou barras simples ou múltiplas) referentes a variáveis categóricas e variáveis numéricas</p>	<p>(EF06MA31) Identificar as variáveis e suas frequências e os elementos constitutivos (título, eixos, legendas, fontes e datas) em diferentes tipos de gráfico</p> <p>(EF06MA32) Interpretar e resolver situações que envolvam dados de pesquisas sobre contextos ambientais, sustentabilidade, trânsito, consumo responsável, entre outros, apresentadas pela mídia em tabelas e em diferentes tipos de gráficos e redigir textos escritos com o objetivo de sintetizar conclusões.</p>
	<p>Coleta de dados, organização e registro</p> <p>Construção de diferentes tipos de gráficos para representá-los e</p>	<p>(EF06MA33) Planejar e coletar dados de pesquisa referente a práticas sociais escolhidas pelos alunos e fazer uso de planilhas eletrônicas para registro, representação e interpretação das</p>

	interpretação das informações	informações, em tabelas, vários tipos de gráficos e texto.
	Diferentes tipos de representação de informações: gráficos e fluxogramas	(EF06MA34) Interpretar e desenvolver fluxogramas simples, identificando as relações entre os objetos representados (por exemplo, posição de cidades considerando as estradas que as unem, hierarquia dos funcionários de uma empresa etc.).

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela vii: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 7º ano do Ensino Fundamental.

Unidade temática	Objetos de conhecimento	Habilidades
Probabilidade e estatística	Experimentos aleatórios: espaço amostral e estimativa de probabilidade por meio de frequência de ocorrências	(EF07MA34) Planejar e realizar experimentos aleatórios ou simulações que envolvem cálculo de probabilidades ou estimativas por meio de frequência de ocorrências.
	Estatística: média e amplitude de um conjunto de dados	(EF07MA35) Compreender, em contextos significativos, o significado de média estatística como indicador da tendência de uma pesquisa, calcular seu valor e relacioná-lo, intuitivamente, com a amplitude do conjunto de dados.
	Pesquisa amostral e pesquisa censitária.	(EF07MA36) Planejar e realizar pesquisa envolvendo tema da realidade social,

	Planejamento de pesquisa, coleta e organização dos dados, construção de tabelas e gráficos e interpretação das informações	identificando a necessidade de ser censitária ou de usar amostra, e interpretar os dados para comunicá-los por meio de relatório escrito, tabelas e gráficos, com o apoio de planilhas eletrônicas.
	Gráficos de setores: interpretação, pertinência e construção para representar conjunto de dados	(EF07MA37) Interpretar e analisar dados apresentados em gráfico de setores divulgados pela mídia e compreender quando é possível ou conveniente sua utilização.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela viii: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 8º ano do Ensino Fundamental.

Unidade temática	Objetos de conhecimento	Habilidades
Probabilidade e estatística	Princípio multiplicativo da contagem  Soma das probabilidades de todos os elementos de um espaço amostral	(EF08MA22) Calcular a probabilidade de eventos, com base na construção do espaço amostral, utilizando o princípio multiplicativo, e reconhecer que a soma das probabilidades de todos os elementos do espaço amostral é igual a 1.
	Gráficos de barras, colunas, linhas ou setores e seus elementos constitutivos e adequação para determinado conjunto de dados	(EF08MA23) Avaliar a adequação de diferentes tipos de gráficos para representar um conjunto de dados de uma pesquisa.

	Organização dos dados de uma variável contínua em classes	(EF08MA24) Classificar as frequências de uma variável contínua de uma pesquisa em classes, de modo que resumam os dados de maneira adequada para a tomada de decisões.
	Medidas de tendência central e de dispersão	(EF08MA25) Obter os valores de medidas de tendência central de uma pesquisa estatística (média, moda e mediana) com a compreensão de seus significados e relacioná-los com a dispersão de dados, indicada pela amplitude.
	Pesquisas censitária ou amostral.  Planejamento e execução de pesquisa amostral	(EF08MA26) Selecionar razões, de diferentes naturezas (física, ética ou econômica), que justificam a realização de pesquisas amostrais e não censitárias, e reconhecer que a seleção da amostra pode ser feita de diferentes maneiras (amostra casual simples, sistemática e estratificada).  (EF08MA27) Planejar e executar pesquisa amostral, selecionando uma técnica de amostragem adequada, e escrever relatório que contenha os gráficos apropriados para representar os conjuntos de dados, destacando aspectos como as medidas de tendência central, a amplitude e as conclusões.

Fonte: a autora (2020). Adaptado de Brasil (2017).

Tabela ix: Unidades temáticas, objetos de conhecimento e habilidades da BNCC para o 9º ano do Ensino Fundamental.

<b>Unidade temática</b>	<b>Objetos de conhecimento</b>	<b>Habilidades</b>
Probabilidade estatística	Análise de probabilidade de eventos aleatórios: eventos dependentes e independentes	(EF09MA20) Reconhecer, em experimentos aleatórios, eventos independentes e dependentes e calcular a probabilidade de sua ocorrência, nos dois casos.
	Análise de gráficos divulgados pela mídia: elementos que podem induzir a erros de leitura ou de interpretação	(EF09MA21) Analisar e identificar, em gráficos divulgados pela mídia, os elementos que podem induzir, às vezes propositalmente, erros de leitura, como escalas inapropriadas, legendas não explicitadas corretamente, omissão de informações importantes (fontes e datas), entre outros.
	Leitura, interpretação e representação de dados de pesquisa expressos em tabelas de dupla entrada, gráficos de colunas simples e agrupadas, gráficos de barras e de setores e gráficos pictóricos	(EF09MA22) Escolher e construir o gráfico mais adequado (colunas, setores, linhas), com ou sem uso de planilhas eletrônicas, para apresentar um determinado conjunto de dados, destacando aspectos como as medidas de tendência central.

	<p>Planejamento e execução de pesquisa amostral e apresentação de relatório</p>	<p>(EF09MA23) Planejar e executar pesquisa amostral envolvendo tema da realidade social e comunicar os resultados por meio de relatório contendo avaliação de medidas de tendência central e da amplitude, tabelas e gráficos adequados, construídos com o apoio de planilhas eletrônicas.</p>
--	---	--

Fonte: a autora (2020). Adaptado de Brasil (2017).

## Apêndices

---

Apresentação dos resultados dos dois ciclos de estudo de caso realizados na fase de avaliação do artefato “cartas para a literacia”. Os resultados obtidos se materializaram em relatos realizados manualmente pelos participantes divididos em grupos, através da utilização de material de papelaria.

### **APÊNDICE A – RESULTADOS DO PRIMEIRO CICLO DO ESTUDO DE CASO**

No Apêndice A são apresentados os resultados obtidos no primeiro ciclo do estudo de caso, realizado com 8 grupos de estudantes do primeiro período de um curso de graduação em Ciência da Computação de uma universidade Federal brasileira.

Geupot

Page 1

Tabelle 1

Coluna B - 20 pessoas já conheciam alguém / 25 não conheciam ninguém

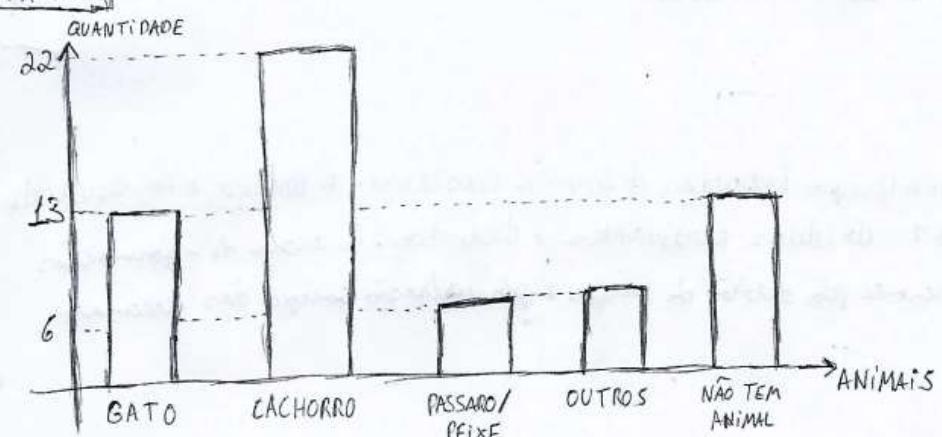
Tabela 2 - Celunc C - 43 pessoas consideraram mais importante superender.

2 pessoas consideram mais importante ~~é~~ aprender  
mão

Dentre esses deus, uma escolhe tirar nota boa e  
outro escolher os pessoas como mais importante.

Chegamos nesse resultado ordenando a tabela e contendo o incidência desses dado

FAISE 2



FASE 3

**FASE 3** Relação entre quantas pessoas na turma tem um animal ou não tem animais.

## FASE 4

• Aproximadamente metade da turma já conhecia alguém antes da 1<sup>a</sup> aula, a grande maioria tem como mais importante na faculdade o aprendizado, 71% das pessoas têm animais, tendo dentre esses animais 69% das pessoas tem cãozinho. Na Motta de trabalhos 43% preferem grupos maiores que 2 pessoas e 25% preferem trabalhos solo. No lado de esportes 88% preferem nadar, 73% jogos, 59% maratonas semanais e 89% outros. Mais especificamente em jogos, o jogo mais jogado é League of Legend (LOL) e entre frequência há um equilíbrio entre a quantidade de pessoas que jogam muitas horas semanais e as que não jogam. Entre horário de estudo temos uma preferência maior por estudar de manhã com maior rendimento, os dados ficam com 45,5% de manhã, 43,2% de noite e 11,4% de tarde.

## FASE 5

O interação entre o grupo impactou de forma positiva o trabalho e os resultados, com cada membro tendo ideias pertinentes e distintas. A troca de informações se tornou mais eficiente pela divisão de tarefas e pela utilização sinérgica das ferramentas digitais.

## GRUPO 2

1<sup>a</sup> Fase → Quantas pessoas não tem animais?

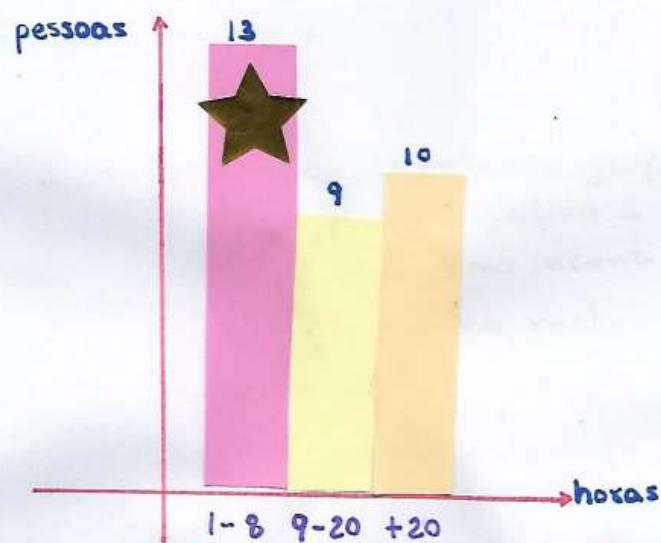
Resposta: 13

Quantas pessoas não jogam games?

Resposta: 13

Como chegamos? Usamos o filtro de tabela.

2<sup>a</sup> Fase ~ Entre as pessoas que jogam, quantas horas jogam por semana?



3<sup>a</sup> Fase ~

O gráfico representa a relação de duas variáveis (pessoas e horas jogadas) a quantidade de horas jogadas é o número de pessoas que utilizam tal tempo nessa atividade.

**4<sup>a</sup> Fase** ~> Ao analisarmos as tabelas disponibilizadas, podemos perceber que a turma no geral se importa mais na aprendizagem ao invés das notas. Percebe-se, também que a maior parte da turma possui animais de estimação de preferência cachorros, além de gostarem de jogar games e escutar música nas horas vagas. Gostam também de realizar tarefas com alguma companhia e de estudar preferencialmente na parte da manhã e da noite, reservando as horas da tarde ao descanso.

- Quantidade de alunos → 45
- Preferem aprender → 43
- Possuem animais → 32
- Possuem cachorros → 22
- Jogam games → 32
- Escutam música → 39
- Tarefas em companhia → 29
- Estudar de manhã e noite → 40

**5<sup>a</sup> Fase** ~> A interação entre o grupo proporcionou uma análise mais efetiva e ampla dos dados analisados. Assim como incentivou o grupo a ser criativo, contribuindo na realização das tarefas de forma proativa.



### Grupo 3

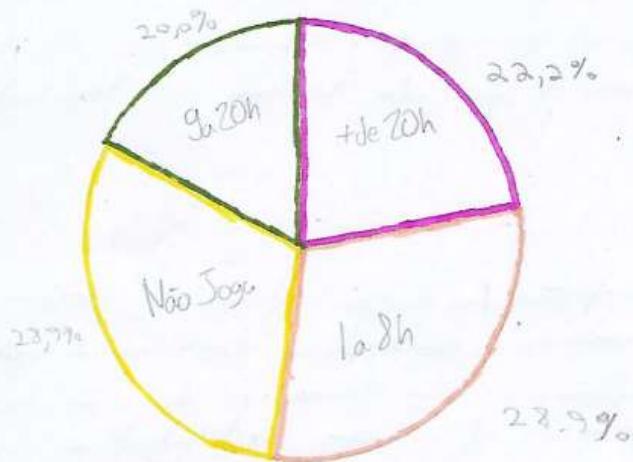
1<sup>a</sup> Fase:

Tarefa 1 - Análise de Estimativas: "Tendo cão"

Tarefa 2 - Experiência-Prática: "Sair"

Excluiremos um campo adicional em cada folha e pegaremos os resultados de uma única questão para cada um destes campos.

2<sup>a</sup> Fase: construirmos um gráfico de pizza com base na Coluna G (Ganhar - Frequência).



3<sup>a</sup> Fase: O gráfico é uma representação visual de um grupo de dados (Coluna G: Ganhar - Frequência) expresso em porcentagens.

Vou a folha ?

4º fisi: Com base na análise de gráficos, podemos afirmar que poucos mais da metade da turma não se consideram satisfeitos com a primeira aula, tendo a maioria da turma desejado o aprendizado sobre colher outros pessoas e terem suas opiniões. Tendo a maioria dizer melhor rendimento no período da manhã a noite (99,9% em cada).

Aproveitar de abandonando um tempo da turma não tem animais de estimação, entre aqueles que possuem, o cão é o mais comum. Podemos afirmar também que entre os hobbies mais comuns é jogar games e ouvir música, sendo o jogo mais jogado logo de ler livros, a 88% da turma também têm outros hobbies não informados. Dentro dos que jogam games, a média de horas jogadas por semana encontra-se entre 9 e 20 horas.

Possivelmente também que é perfil dos alunos de CC é mais cooperativo, preferindo trabalhar em grupos maiores que 2, e a menor ponta dos alunos preferindo trabalho individual.

5º fisi:

Cada um tem suas suas interpretações para os dados analisados, promovendo uma discussão que contribui para uma melhor qualidade e prazeres do trabalho. Além disso, a discussão de temas contribui muito para a eficiência do trabalho como um todo.

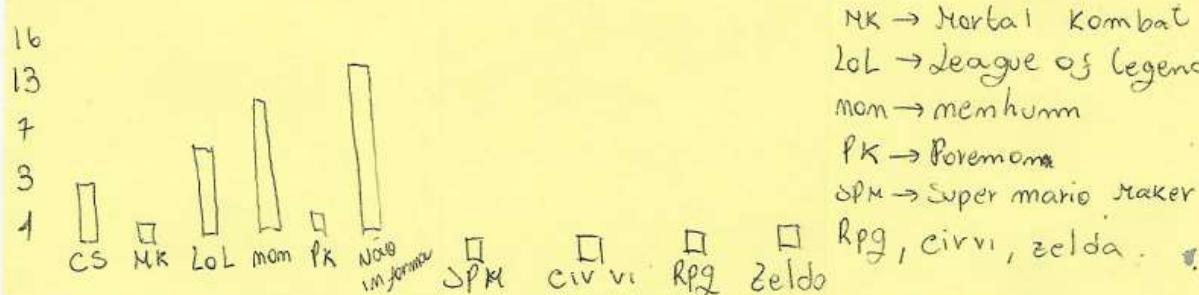
## Grupos

1º fase: Escreva um exemplo de dado contido em cada uma das duas tabelas fornecidas.

Um exemplo de dado é LOL, visto na coluna "Qual game joga", presente na Tabela 1 e 2. É além disso usamos o recurso filtro, para verificar quantos indivíduos jogam que é f. simetricamente, foi visto que pelo mesmo método verificamos que 22 pessoas possuem cais, "Tenho cais".

2º: Escolha uma das tabelas e construa um gráfico do seu interesse na tabela 1, na coluna de "Qual game joga", pegamos os dados que estão contidos, fizemos um gráfico formato barras, e descobrimos quantas pessoas jogam cada jogo e quantas não jogam. Além do gráfico feito virtualmente, fizemos um paralelo dos dados obtidos em um post-it (colado abaixo).

Qual game jogam os alunos da turma? Legenda:



3º O que significa o gráfico que fiz na 2º fase:

- Com a junção dos dados da coluna "Qual game joga", obtemos um gráfico do tipo "barra", que mostra o jogo mais jogado pela turma. Além de verificar que há pessoas que não jogam nenhum jogo. Porém para fazer uma análise mais robusta, falta dados visto que um número considerável de alunos não informaram qual game joga, onde essas podem:

- Jugar, mas não informaram;
- Não jogam nenhum;
- já jogaram, mas não jogam com frequência

4º fase: Apresenta uma descrição textual detalhada sobre as características da sua turma com base nas análises feitas nas 1º, 2º e 3º fases.

Foi observado, entre os alunos que informaram qual game joga, que a grande maioria dos alunos da turma jogam LOL, e com base na tabela 4, notamos por observação e contagem que essas conseguem render nos estudos mal à noite, além de pela tabela 2, boa parte deles não estudam na véspera da prova e que não gostam de ir à praia.

Em geral, uma boa parcela da turma já se conhecia antes da primeira aula, assim como, o mais importante para a maioria das pessoas da turma é nordestina. É possível tirar das plomílias que a maioria das pessoas tem pelo menos um animal de estimação e que uma boa parte da turma prefere esparrecer com praia, música, seriados e games bem como estudar em duplas e grupos.

Obs: Para o análise de que o pessoal que joga LOL não gosta de ir à praia (dentre os que informaram) e que não estudam na véspera de prova, usamos o libre office, onde para analisar bem todas as informações dos que informaram que jogam LOL, uma tabela é do batallt, copiamos as linhas referentes a estes alunos e colamos mais abaixo, da tabela original.

Além de que foi utilizado o Jupyter notebook para ajudar também na análise da tabela 1.

Fase final = 5º fase

Como a interação entre seu grupo impactou?

A interação do grupo de 7 pessoas para a realização dessa atividade foi feita da seguinte forma: nos separamos em 2 grupos (um de 3 e um de 4) que realizaram pesquisas meio-independentes, onde o resultado de ambas as pesquisas eram compartilhadas de forma dinâmica por meio de um integrante que intermediava os grupos. Acreditamos que a liberdade e independência de cada grupo foi essencial para um bom desenvolvimento do trabalho, pois permitiu que diversos pontos de vista para solução do problema fosse explorado.

GRUPO 5

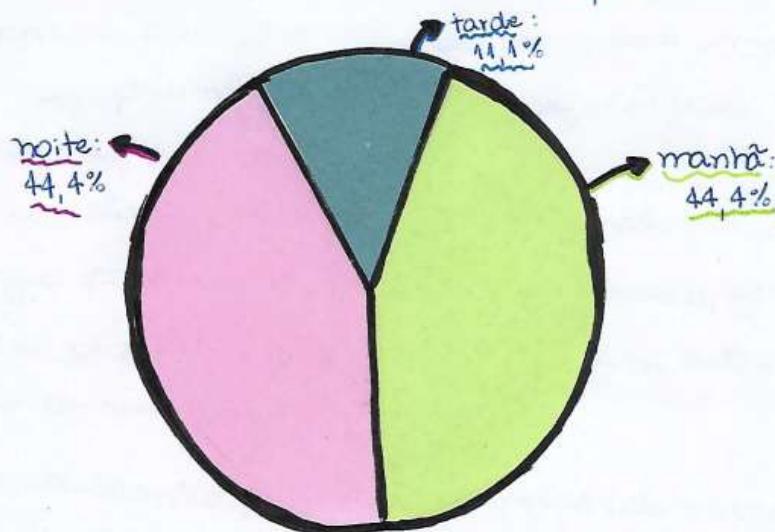
# atividade

## análise de dados

**1<sup>a</sup> fase:** Analisando juntas as tabelas, percebemos que o atributo "O estudo rende mais..." contém os mesmos tipos de dados nas idades. Então o valor "O estudo rende mais de noite" está em ambas, por exemplo.

**2<sup>a</sup> fase:** Tabela escolhida: Tabela-2  
Atributo: O estudo rende mais...

Dados	Total de alunos	Porcentagem
manhã	20	44,4%
tarde	5	11,1%
Noite	20	44,4%



**3<sup>a</sup> fase:** O gráfico construído em (2) significa que dentro desse universo de 45 alunos, 20 alunos, isto é, 44,4% dos alunos acham que o estudo vende mais à noite. E, o mesmo acontece para o período da manhã. No entanto, apenas 5 alunos, isto é, 11,1% dos alunos acham que o estudo vende mais no período da tarde.

**4<sup>a</sup> fase:** Aqui, para fornecer uma descrição textual detalhada sobre as características da turma, analisaremos os gráficos correspondentes a cada atributo da tabela.

10% da turma considera importante tirar nota boa, em contraste com a maioria que discorda. Em comparação, 95%acha essencial aprender durante seu percurso na faculdade. Cerca de 22% dos alunos considera importante conhecer pessoas, pedindo assim formar conexões. Uma minoria tem outros objetivos em mente.

29% da turma não tem um animal de estimação, e todos que possuem, metade tem coelhos, 29% tem gatos, 13% têm peixe ou pássaro e 13% têm outros, pedindo ter mais de um tipo de animal.

Para descontrair, um pouco menos da metade da turma gosta de ir à praia, cerca de 90% gosta de ouvir música, 20% gosta de dançar, cerca de metade gosta de praticar esportes, 60% gosta de maratonar séries, 45% gosta de jogar e a maioria também tem outras formas de relaxar.

Um pouco mais da metade tem estudos nos longos  
do período, em comparação aos que estudam na  
semana anterior. Tendo 20% que também estudam de  
sextas. Em questões de horário de estudo, as turmas se  
igualmente divididas ao meio, entre os que preferem estudar  
de manhã e os que preferem estudar à noite, tendo uma  
minoria que prefere estudar à tarde. Ninguém coloca  
menos de 10% nessa na sorte, de acordo com os dados.

Um pouco menos de 1/3 prefere trabalhar sozinho e na  
mesma proporção alguns preferem trabalhar em  
dupla. O resto prefere trabalhar em grupos maiores do que  
2 alunos.

Apenas 1/3 das turmas têm algum jogo favorito, com  
a maioria preferindo jogar só. Os outros 2/3 ou não  
possuem um jogo favorito ou não informaram. 25% das  
turmas não jogam, 25% jogam de 1 a 8 horas por semana, 25%  
de 9 a 20 horas e 25% jogam mais de 20 horas semanais.

**5ª fase:** Um trabalho que exige análise de dados possui  
várias fases e, por ser feito em grupo, necessita que haja  
uma interação interpessoal. O nosso grupo responde bem  
à essa necessidade de interação, obtendo os resultados  
esperados durante as etapas deste trabalho.

## GRUPO 6

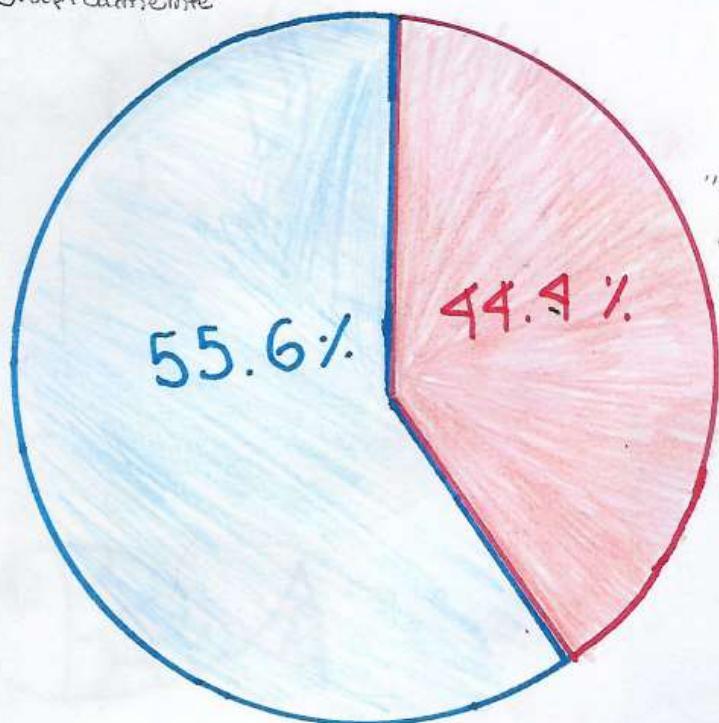
Datalio 1:

- = Já conhecia alguém antes da primeira aula?

SIM - 25

NÃO - 20

Graficamente

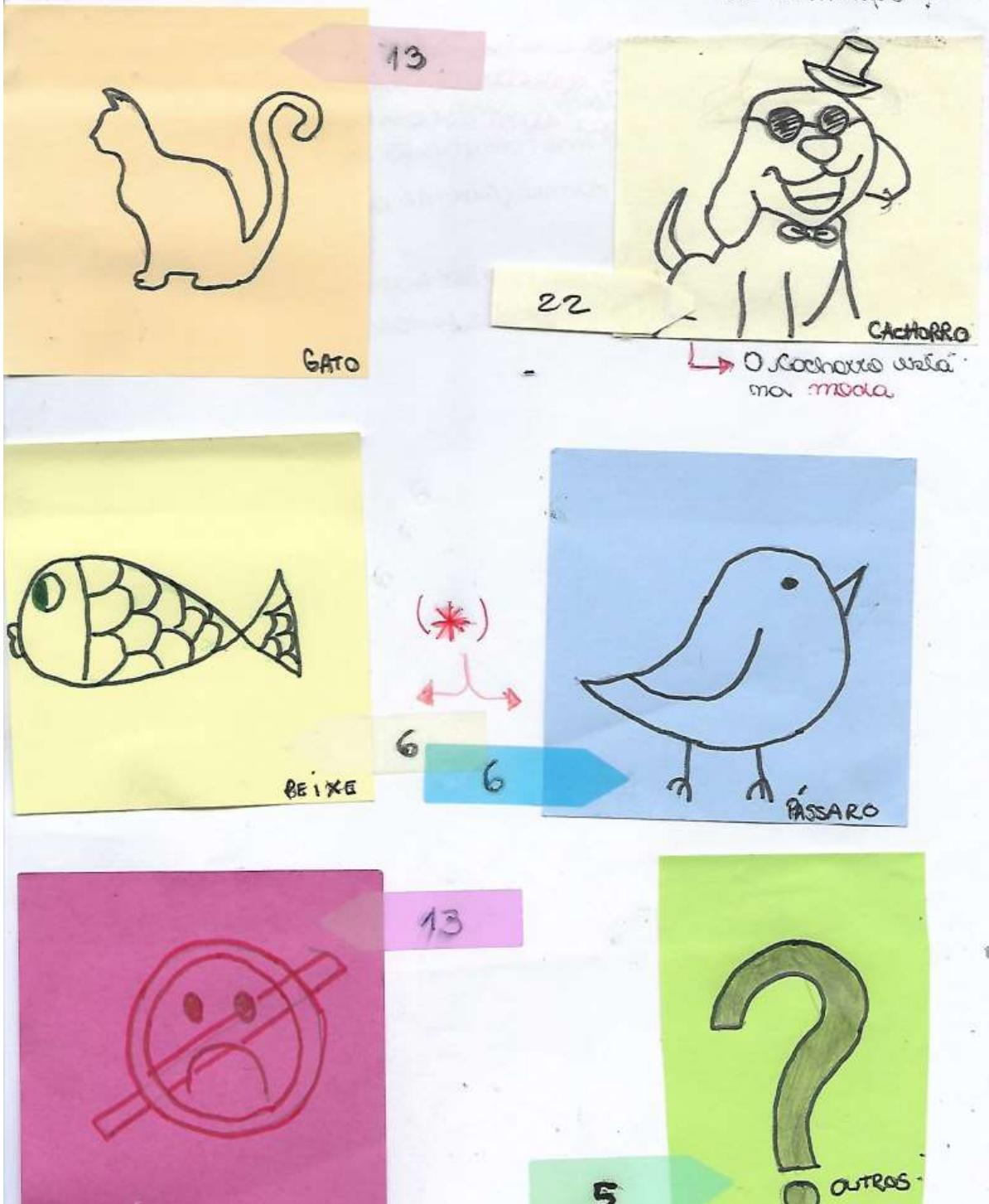


Nós contarmos quantos "Sim" e "Não" tem e separamos em duas cores distintas e realizamos um gráfico.

→ A maioria dos alunos já conhecia alguém antes de entrar no hotel.

## Desafio 2:

Fizemos um gráfico que contabiliza, por meio de desenhos, a quantidade de pessoas que possui um animal de estimação:



Desafio 3: O objetivo do gráfico é representar sobre  
mais dos elementos a velocidade da turma com os ami-  
mãos da estimocão.

As cores são dada de acordo com o número de  
verdades relacionadas a cada animal.

\* No gráfico a turma é dada por "gatozinho ou urso",  
Isso, podemos que é possível quantificar & dos 2, mas  
é mais possível determinar com certeza.

Desafio 4 : Relatórios  
Enviado digitalmente

Desafio 5: Nós analisaremos conjuntamente os dados tabulares e diremos a melhor forma de representá-los visualmente. Nós dividimos em grupos (eletricistas, encarregados, etc) para conseguirmos concluir os exercícios no tempo estipulado.

Com certeza, a execução não será possível dentro da visão da maioria dos membros e da impossibilidade em realizar a tarefa.

## Grupo 6 (Desafio 4)

Relatório de dados da turma:

- A maior parte da turma não possuía amizades antes do primeiro dia de aula.
  - Menos da metade da turma já conhecia pelo menos uma pessoa antes do primeiro dia de aula.
  - A maior parte da turma considera o aprendizado como o mais importante no curso.
- Logo, o conhecimento é algo primordial para os estudantes da mesma.
- Poucas pessoas consideram como mais importante conseguir conhecer pessoas novas. Logo, as relações interpessoais são subestimadas por boa parte dos estudantes da turma.
  - A parte considerável da turma não possui animais de estimação.
  - Os que possuem apenas um animal, em maioria possui um cachorro.
  - Os que possuem mais de um, a maior parte possui cães e gatos.
  - A turma se divide entre pessoas que adoram estudar em grupo e pessoas que estudam solo. Logo, a diferença acerca das preferências da metodologia de estudo são bastante opostas na turma.
  - As exceções preferem as interseções entre as metodologias.
  - $\frac{3}{4}$  da turma se interessa por games. A maior parte dos que se interessam por games optaram por não informar suas preferências. O game LOL foi maioria nas preferências dos estudantes que decidiram optar por informar.
  - Os dados referentes aos modos de esparrecer nos mostram que a maior parte da turma tem preferência por ir à praia ou escutar música.
  - A maioria da turma não se interessa por dançar ou praticar esportes não digitais.
  - Apenas  $\frac{1}{4}$  não se interessa por games.
  - A turma ficou dúvida entre pessoas que estudam melhor a noite e pessoas que estudam melhor de manhã.

/ **Grupo 4**1<sup>a</sup> FASE

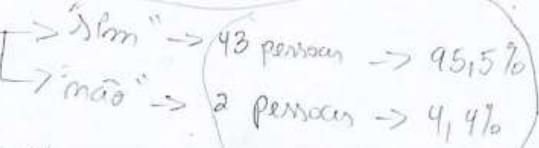
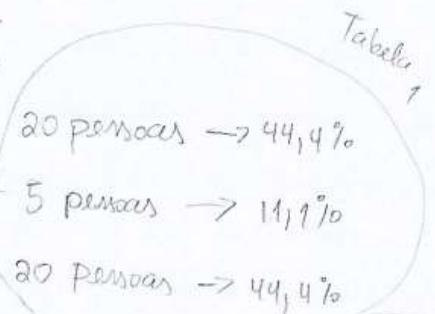
"Estudo geralmente mais de muito" - 20 pessoas  $\rightarrow 44,4\%$

"Estudo geralmente mais de tarde"

"Estudo geralmente mais de muito"

"mais importante aprender"

$\rightarrow$  Contam-se diretamente da planilha

2<sup>a</sup> FASE

44,4 %

11,1 %

44,4 %

"Estudo geralmente mais de muito"

"Estudo geralmente mais de tarde"

"Estudo geralmente mais de manhã"

$\rightarrow$  relacionamos a coluna "i" da Tabela\_1, em seguida, clicamos em "inserir"  $\rightarrow$  "gráfico", e o gráfico foi construído.

### 3<sup>ª</sup> FASE

o O gráfico evidencia, por meio de informações acerca do horário preferido para estudo, que 44,4% das pessoas preferem de manhã e a mesma porcentagem correspondeu aos indivíduos que preferem estudar à noite. Em contrapartida, apenas 11,1% preferem o estudo de tarde. Os dados foram produzidos a partir de 45 pessoas.

### 4<sup>ª</sup> Fase

o Estabelecendo um perfil geral do grupo analisado, isto é, concluindo os dados por meio de um indivíduo hipotético que melhor representa a turma, é possível apontar esse "aluno médio" como alguém mais produtivo em seus estudos pelas manhãs e noites. Ele gosta de videogames, tem 30 horas semanais diante de um aparelho para tal finalidade e pretende não revelar o seu título profissional. Em questão de trabalhos, lhe convém mais em grupos acima de dois integrantes. Sobre animais de estimação, esse aluno possui um cão.

### 5<sup>ª</sup> FASE

o O fluxo de ideias foi a principal molde de troca entre os integrantes. Novos pensamentos foram sendo lançados ao longo do debate. Assim, foi possível a produção de todos os dados apresentados nas fases anteriores. A grande maioria dos conteúdos produzidos foi realizada fora do meio digital, apesar da coleta ter sido quase integralmente por meio do computador.

## /

### GRUPO 8

1) Tabela 1 → Já conhecia alguém antes da primeira aula?

Tabela 2 → Mais importante tirar nota boa?

T<sub>1</sub> ~ cerca de 225 alunos não conheciam ninguém antes da primeira aula e 20 conheciam.

T<sub>2</sub> ~ 40 alunos não consideram tirar nota boa o fator mais importante em uma disciplina, 5 consideram.

Utilizamos a tabela dinâmica para obter os dados.

2) Utilizamos a Tabela dinâmica para calcular a quantidade de pessoas que responderam: Tabela 1 - Coluna C

1 - mais importante é aprender: 64.9%.

2 - mais importante é aprender e as pessoas: 20.0%.

3 - mais importante é aprender e outros: 2.2%.

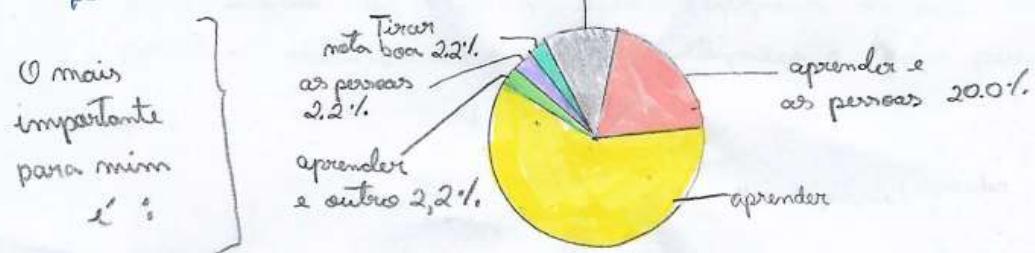
4 - mais importante é só as pessoas: 2.2%.

5 - mais importante é tirar nota boa: 2.2%.

6 - mais importante é tirar nota boa e aprender: 8.9%.

Utilizamos a tabela dinâmica para formular o gráfico.

3) O gráfico mostra que as pessoas consideram mais importante no curso.



4) Analisando os dados da turma percebemos que cerca de 55% dos alunos não conhecia ninguém do curso até entrar na universidade, a grande maioria não considera trair uma nota boa como principal fator e 64,4% acha que o mais importante é aprender. Além disso, 32 alunos possuem animais domésticos. desses 32, 22 tem cachorros, 13 tem gatos, 6 possuem passaros e ou peixes e 6 tem outros. 13 alunos não possuem animais domésticos.

Cerca de 25 alunos começam a estudar para uma prova durante o período e 20 conseguem睡iria se o antes. uma porcentagem de 55,56% e 44,44% respectivamente. Cerca de 44% dos alunos consideram que o estudo rende mais pela manhã, essa mesma qualidat considera o estudo melhor a noite e somente 11% estudar melhor pela tarde.

5) Todos os integrantes do grupo interagiram e participaram na análise, tornando fácil a obtenção de dados e agilizando as etapas da atividade. A interação impactou a análise das informações de forma positiva, pois a divisão das tarefas entre os integrantes do grupo permitiu a execução dos desafios sem ultrapassar o tempo previsto.

**APÊNDICE B – RESULTADOS DO SEGUNDO CICLO DO ESTUDO DE CASO**

Apêndice B são apresentados os resultados obtidos no segundo ciclo do estudo de caso, realizado com 4 grupos de estudantes de um curso de pós-graduação em Informática de uma universidade Federal brasileira.

## 1º FASE

## GRUPO 1

TABELA 1: "tenho cão"

TABELA 2: "Mortal Kombat"

- Inicialmente o grupo ficou em dúvida sobre a definição de "Dado" as dúvidas foram:

- Dado seria uma célula? Uma linha? Uma coluna?

Foi feita uma busca na Wikipédia em português e Inglês.

Por fim o grupo concordou que "Dado" é apenas uma representação de um estado do mundo sem, necessariamente, isoladamente ser intérprete pelo leitor. Logo foi selecionada uma célula aleatória de cada tabela.

## 2º FASE

O QUE É MAIS IMPORTANTE PARA VOCÊ ?

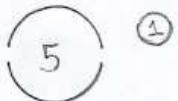
TER UMA BOA NOTA  5

APRENDER  43

AS PESSOAS  10

OUTROS  1

LEGENDA



### 3<sup>º</sup> FASE

O grupo ficou em dúvida entre um gráfico feito à mão ou por software. Decidiu-se pela 1<sup>º</sup> abordagem como uma forma de exercer uma habilidade que não é mais utilizada atualmente de forma a entender suas possibilidades e dificuldades.

O grupo pensou em apresentar no gráfico uma correlação entre os dados da tabela 2. Exemplos: Será que um aluno que considera a nota como mais importante, estuda ao longo do período? No entanto, essa abordagem parece difícil de ser realizada no tempo proposto sem a ajuda de um software.

Por fim o grupo decidiu por representar a que fator os estudantes dão mais importância. Foi escolhida uma representação utilizando "smiles" em dois tamanhos para agregar a resposta de cada item, utilizando uma legenda.

Os "smiles" usados foram diferentes para representar a característica de cada resposta.

## 4º FASE

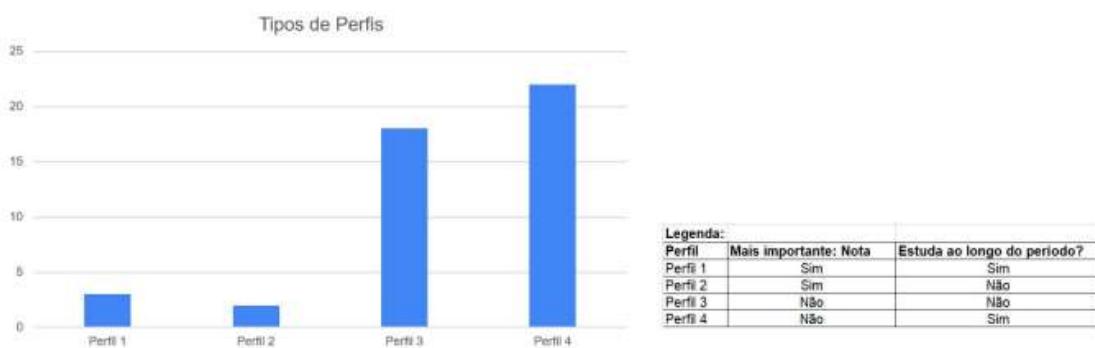
Se escolhessemos um aluno dessa turma, o mais provável que encontrasssemos:

- Alguém que não conhecia outro aluno antes da 1º aula (56%)
- Alguém que considera aprender como o mais importante (93%) entre notícias, aprendizado e pessoas.  
(Podriam escolher mais de uma alternativa)
- Já vinha estudando ao longo do período (56%)
- Tem um animal de estimação (41%) e ele é um cão (69%)
- Alguém que prefere escutar surtindo música (87%)
- Não vai a clôr na praia (Ninguém respondeu que iria!)
- Prefere fazer trabalhos em grupo com mais de 2 integrantes (47%)
- Joga Games de 1 a 8 horas por semana (29%)
- Prefere estudar de manhã ou à noite (Empate de 44% cada).

### 5º FASE

Os perfis diferentes (um mais técnico, outro mais artístico e outro mais analítico) dos integrantes do grupo permitem a análise das questões através de diferentes perspectivas.

Isto foi considerado bom, porém o tempo curto para a realização das tarefas limitou um pouco a exploração dessas habilidades.



— 1<sup>º</sup> fase

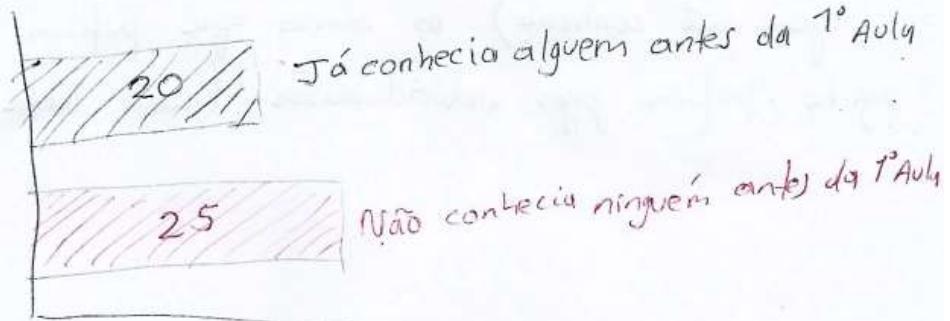
Tabela 1: Praia e outros

Grupos 2:

Tabela 2: Não tenho

2<sup>º</sup> fase

de acordo com o gráfico criado foi identificado que a maioria das pessoas  
gráficó: ~~não~~ ~~gráficó~~ ~~conhecia~~  
não se conheciam antes da 1<sup>º</sup> aula



3<sup>º</sup> Fase

Descrição do gráfico. da Tabela 1.

O gráfico representa a distribuição de dois grupos complementares onde é possível visualizar a proporção de pessoas que conheciam algum ou não conheciam antes da 1<sup>º</sup> aula.

4<sup>o</sup> fase

Analisando os dados observamos que não dados de alunos  
Nesses dados observamos que:

- Quem não conhecia ninguém antes de 1º aula preferiu trabalhar em grupos >2 do que quem conhecia alguém de 1º aula.
- Em ambos os grupos ( quem conhecia alguém antes de 1º aula e quem não conhecia) os alunos que preferiram trabalhar sozinho preferiram jogos individualistas (como zelda)

### 5º fase

Impactou na forma de ver os dados. Observamos critérios diferentes para observação de alguns dados. Cada um dos membros do grupo tinham perspectivas diferentes sobre os dados chegando ao mesmo objetivo como nas fases anteriores.

## Grupo 3

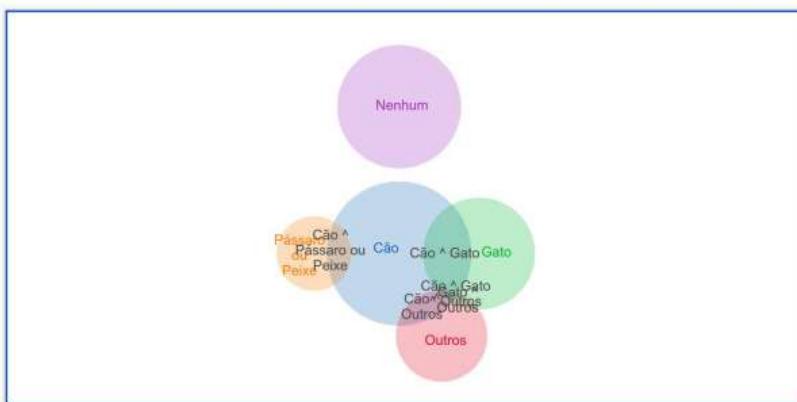
1 -> Escreva um exemplo de dado contido em uma das duas tabelas fornecidas

Dados são as respostas brutas das pessoas entrevistadas, sem fazer interpretações sobre o que foi coletado. Com isso, um exemplo de dado contido na tabela 1 é o tipo de jogo mais consumido por cada indivíduo entrevistado.

2 -> Escolha uma das tabelas e construa um gráfico de seu interesse

Pensamos em fazer um gráfico que relate o tipo de animal de estimação dos entrevistados.

<https://www.meta-chart.com/venn#/display>



3 -> O que significa o gráfico que você fez na 2<sup>a</sup> fase?

O gráfico de Venn apresentado acima descreve a intersecção entre os animais possuídos pelos entrevistados. Há intersecções entre entrevistados que possuem mais de um animal de estimação. Pessoas sem animais são colocados aparte do gráfico.

4 - Apresente uma descrição textual detalhada sobre as características representada nas tabelas 1 e 2, com base nas análises feitas nas fases 1,2 e 3:

A Tabela 1 trata-se de um questionário estruturado acerca do comportamento dos indivíduos que os responderam. São feitas perguntas de cunho psicográfico, como gosto e momentos de lazer/estudo. As respostas presentes nesta tabela são de difícil tratamento, visto que precisam de uma excessiva limpeza dos dados. Esta tabela, devido a sua estrutura visual, permite uma fácil leitura humana, sendo rápida na retirada de insights.

A Tabela 2 trata-se de um detalhamento mais aprofundado com questões bem semelhantes que a Tabela 1. Pela sua estrutura booleana, permite um fácil tratamento.

A partir da estrutura das tabelas, cogitou-se a possibilidade de predizer o horário de aulas da turma, e de acordo com os horários declarados de maior rendimento foi possível supor que, devido à menor incidência de bom rendimento nos estudos durante a tarde, a turma em que foi aplicado o questionário deve ter aulas neste turno.

Também cogitou-se a possibilidade de entender a dinâmica de trabalhos em grupo considerando o fato dos alunos já conhecerem algum de seus colegas ou não antes do início das aulas, mas não foi possível ter algum indício considerável de tendência comportamental.

Uma análise feita sobre os horários preferidos para estudo e a carga horária semanal despendida em jogos demonstra a tendência dos jogadores a terem maior eficiência no estudo à noite, mas não foi possível concluir, por exemplo, qual é o horário que os entrevistados costumam jogar.

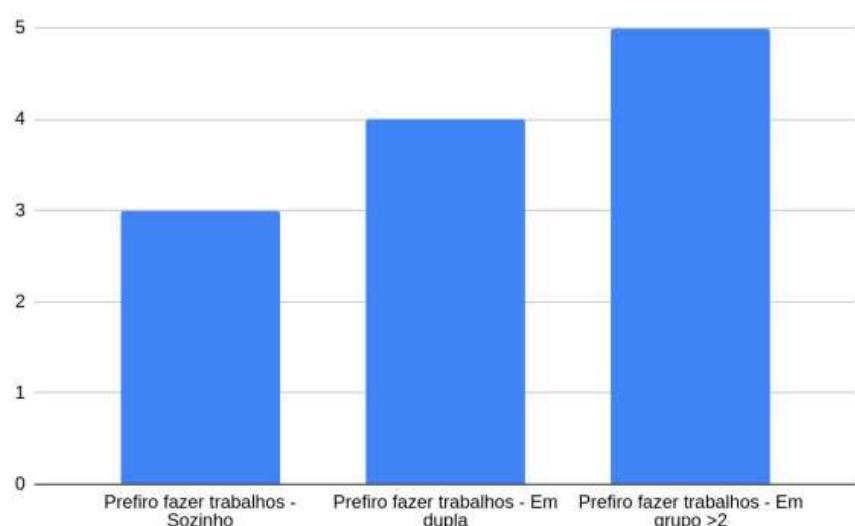
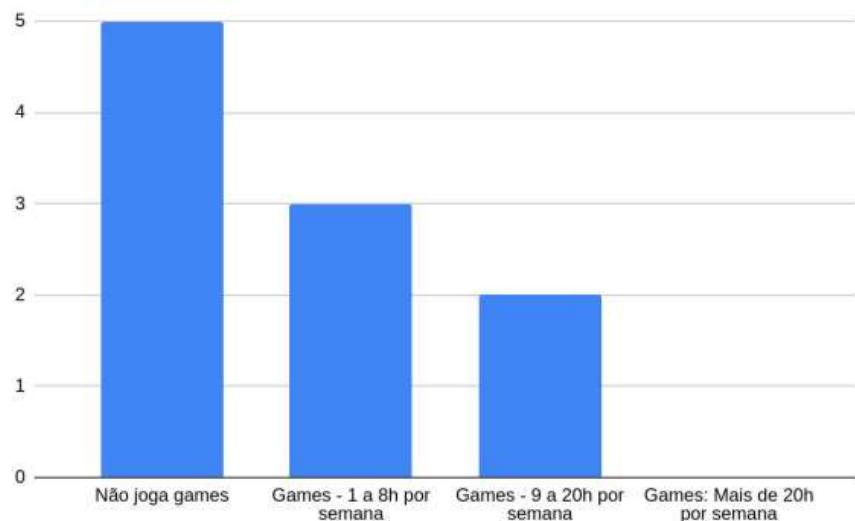
Ainda dentro dos jogos, os que são considerados e-sports costumam demandar mais tempo dentre os entrevistados, ao mesmo tempo que os entrevistados que consideram as pessoas importantes da escola costumam gastar poucas horas de jogos.

5 - Como a interação entre seu grupo impactou a análise dos dados e os resultados obtidos?

Cada integrante do grupo possuia uma skill que complementava na criação de uma análise mais coesa, entretanto as diferenças foram fator dificultoso para chegar a conclusões em comum acordo. A falta de uma computador para cada integrante do grupo dificultou análises individuais.

O gráfico abaixo mostra a quantidade de horas de tempo dispendida jogando games para pessoas que consideram dentro os assuntos mais importantes conhecer outras pessoas na escola. Nesta mesma população, o segundo gráfico apresenta preferência por trabalhar em

grupos maiores.



Grupo 4: .

1<sup>a</sup> Fase: Escreva um exemplo de dado contido em cada uma das duas tabelas fornecidas:

**Tabela 1:**

*Exemplo:* Tenho cão

**Tabela 2:**

*Exemplo:* Não

2<sup>a</sup> Fase: Escolha uma das tabelas e construa um gráfico do seu interesse.

Gráfico gerado a partir da tabela 2:

Campo que informa o período do dia em que o estudo rende mais:



3<sup>a</sup> Fase: O que significa o gráfico que fizemos na 2<sup>a</sup> fase?

O gráfico demonstra o resultado da coleta de dados que determina o período do dia em que o estudo rende mais comparado com o fato de ter um cachorro como animal de estimação.

4<sup>a</sup> Fase: Apresente uma descrição textual detalhada sobre as características da turma representada nas tabelas 1 e 2, com base nas análises feitas nas fases 1, 2, 3.

Como coletamos fragmentos de dados de forma aleatória não existe uma conclusão precisa. Especulamos que os alunos que possuem um cachorro como animal de estimação apresentam um menor desempenho nos estudos por dividir seu tempo com seu animal de estimação.

5<sup>a</sup> Fase: Como a interação entre seu grupo impactou a análise dos dados e os resultados obtidos?

A interação com o grupo proporcionou debates sobre a escolha dos dados a serem analisados. Apesar de divergências, a interação foi importante para concluir a análise.

6ª Fase: Com ajuda de um software, escolha uma das tabelas e construa pelo menos 1 gráfico de seu interesse.

<Relação do rendimento nos estudos com os games>

Qual game joga (Tabela11)	O estudo rende mais... (Tabela11) / Games (frequência)		
	Estudo rende mais de manhã Games 1 a 9 a Games + Bh/sem 20h/sem de 20h/s. games	Estudo rende mais de noite Games 1 a 9 a Games + Bh/sem 20h/sem de 20h/s. games	Estudo rende mais de tarde Games 1 a 9 a Games + Bh/sem 20h/sem de 20h/s. games
CIV VI	*	*	*
CS		*	*
LOL	*	*	*
Mortal Kombat		*	*
Não informou	*	*	*
Nenhum	*	*	*
Pokemon Soul Silver		*	*
RPG			*
Super Mario Maker			*
Zelda			*