UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO MATEMÁTICA E TECNOLÓGICA
CURSO DE MESTRADO

ANDREIKA ASSEKER AMARANTE

O USO DO TINKERPLOTS PARA EXPLORAÇÃO DE DADOS POR PROFESSORES DE ESCOLAS RURAIS

RECIFE
2011

ANDREIKA ASSEKER AMARANTE

O USO DO TINKERPLOTS PARA EXPLORAÇÃO DE DADOS POR PROFESSORES DE ESCOLAS RURAIS

Dissertação apresentada ao Programa de Pós-Graduação em Educação Matemática e Tecnológica, como requisito parcial para obtenção do título de Mestre em Educação Matemática e Tecnológica.

Orientador: Prof. Dr. Carlos Eduardo Ferreira Monteiro

Co-orientadora: Profa. Dra. Irante Maria da Silva Lima

RECIFE
2011

Amarante, Andreika Asseker

O uso do tinkerplots para exploração de dados por professores de escolas rurais / Andreika Asseker Amarante. – Recife: O Autor, 2011.

156 f. : il.; quad., graf.

Orientador: Prof. Dr. Carlos Eduardo Ferreira Monteiro

Dissertação (Mestrado) - Universidade Federal de Pernambuco, CE, Programa de Pós-Graduação em Educação Matemática e Tecnológica, 2011.

Inclui Bibliografia e Anexo.

1. Informática na Educação 2. Software Tinkerplots 3. Educação rural I. Monteiro, Carlos Eduardo Ferreira (Orientador) II. Título

CDD 371.344 UFPE (CE 2011-034)

ANDREIKA ASSEKER AMARANTE

O USO DO TINKERPLOTS PARA EXPLORAÇÃO DE DADOS POR PROFESSORES DE ESCOLAS RURAIS

Comissão Examinadora

Presidente e Orientador
Prof. Dr. Carlos Eduardo Ferreira Monteiro

Co-orientador
Profª. Drª. Iranete Maria da Silva Lima

Examinador Externo
Profª. Drª. Liliane Maria Teixeira Lima de Carvalho

Examinador Interno
Profª. Drª. Gilda Lisbôa Guimarães

Recife, 22 de fevereiro de 2011.

AGRADECIMENTOS

Ultrapassei mais uma etapa da minha vida. Apesar de muitos desafios consegui superá-los e essa dissertação é fruto desse trabalho de superação. Todavia, não conseguiria sem a ajuda de todos que estiveram ao meu lado me incentivando nos momentos de incerteza e me dando força nos momentos de fraqueza.

Gostaria de agradecer inicialmente ao meu companheiro Adenor Almeida que acreditou no meu potencial desde a graduação, incentivou-me no percurso da pós-graduação e ajudou-me nos momentos que mais precisei.

Ao meu amigo e orientador Carlos Eduardo que contribuiu imensamente na construção da pesquisadora e docente que hoje me tornei a partir de seus exemplos diários, dos seus conselhos e de seus puxões de orelhas de vez em quando.

A minha co-orientadora Iranete que interveio nos momentos certos auxiliando a encontrar o caminho metodológico que indicasse respostas aos meus questionamentos de pesquisa.

Obrigada meus pais, Franklin e Edna e minha irmã Erika por torcerem por mim em todas as etapas do meu desenvolvimento.

Obrigada aos participantes do grupo GPEMCE em especial as minhas amigas Iane, Marcela, Michela, Valdenice, Sandra e a professora Liliane Carvalho pela força, pelas discussões em grupo, pela amizade e companheirismo. Foram indispensáveis na construção desse estudo.

Sou grata as professoras e aos alunos da linha de pesquisa Processos de Ensino Aprendizagem em Educação Matemática e Científica que contribuíram significativamente nas discussões das aulas de seminários.

À FACEPE que possibilitou minha inteira dedicação a este estudo através da concessão de uma bolsa durante todo o decorrer da pesquisa.

À Secretaria de Educação, Esporte, Juventude, Ciência e Tecnologia de Caruaru, na pessoa da diretora de ensino Tânia Bazante e a Gestora Zeza que me acompanhou em todas as visitas as escolas rurais. Aos professores participantes do estudo pela disponibilidade em contribuir com a pesquisa.

O meu
Muito obrigada!

RESUMO

Este estudo teve como objetivo investigar o uso do *software TinkerPlots* para exploração de dados por professores de escolas rurais do Agreste de Pernambuco. Uma vez que esse *software* vem sendo apontado como um instrumento que pode auxiliar o ensino da Estatística nas escolas. Buscamos especificamente identificar as representações construídas a partir da manipulação do *software*, analisar o uso das ferramentas *separar*, *empilhar*, *ordenar* e *gradiente* e suas influências na interpretação de dados. Participaram do estudo 12 professores de 8 escolas rurais do Município de Caruaru região Agreste de Pernambuco. Os professores participaram de dois momentos de entrevista semi-estruturada na pesquisa. No primeiro momento foi identificado o perfil dos participantes, no que se refere ao uso do computador e ao ensino de Tratamento da Informação. Em seguida foi realizada uma sessão de familiarização com o *software*, na qual os professores conheceram e manipularam as suas ferramentas. Num segundo momento os professores responderam três questões sobre um banco de dados sobre Violência e três questões sobre um banco de dados sobre Novelas. As questões eram baseadas na perspectiva do contexto bivariado ou cruzamento de duas variáveis. Propomos os tipos de cruzamento entre: variáveis quantitativas, entre as variáveis qualitativas e entre uma variável quantitativa e uma qualitativa. As entrevistas foram videografadas, transcritas e organizadas em protocolos para análise. De uma forma geral os professores apresentaram facilidade de manuseio com as ferramentas do *software* e sua proposta de manipulação de dados. As ferramentas utilizadas proporcionaram uma manipulação dos dados, já que os professores puderam movimentar os *plots* por diversas ações, tais como: organização em categorias em eixos horizontais e verticais; inserção de escalas; empilhamento; ordenação; utilização de recursos de cores e de *gradiente*. Essa manipulação proporcionou em muitas ocasiões uma mudança na reflexão e uma contribuição para a interpretação. Nesse sentido, as ferramentas do *software* foram utilizadas na construção de diferentes representações no momento da exploração dos dados, o que possibilitou uma maior interação com os dados oferecendo diferentes pistas na construção da resposta. Podemos constatar, a partir desse estudo, a importância das ferramentas *separar* e *gradiente*, as quais possibilitaram aos professores a percepção dos valores dispostos na escala, o que auxiliou no processo de interpretação dos dados. No entanto, também apontamos que a ferramenta *gradiente* em algumas situações promoveu uma análise pontual dos dados através da cor mais escura o que indica a necessidade de intervenção do professor. Futuros estudos são necessários para que sejam propostas utilizações desse *software* em um processo de formação de professores, no qual sejam construídos conceitos de Estatística, dando a liberdade de manipulação e construção de diferentes representações. A investigação no âmbito da formação de professores poderá contribuir no seu próprio trabalho com o Tratamento da Informação em sala de aula.

Palavras chave: Ambiente computacional; *Software Tinkerplots*; Interpretação de gráficos; Professores de escolas rurais.

ABSTRACT

Media disseminate daily statistical data using graphs, which has demanded of readers to develop skills in interpreting graphs. Schools should be prepared to teach such skills in interpretation of graphs in order to allow students to participate in social situations of data analysis. The study aimed to investigate the use of software TinkerPlots to handle data by teachers in rural schools in the Agreste of Pernambuco, Brazil, once this software has been appointed as a tool to guide the teaching of statistics in schools. We seek to specifically identify the representations built from the manipulation of software tools to analyse the use of *separate*, *stack*, *order*, and *gradient* and its influence on data interpretation. Study participants were 12 teachers in eight rural schools. Teachers participated in two moments of semi-structured research interview. At first we identified the profile of participants, with regard to computer use and teaching of Information Processing. Then it was performed a familiarization session with the software, in which teachers understand and manipulate his tools. In a second moment, the teachers answered three questions about a database on violence and three questions on a database on Novels. The questions were based on the bivariate context or perspective of the intersection of two variables. We propose the types of crossing: quantitative variables and qualitative variables between variable between quantitative and qualitative. The interviews were recorded in video, transcribed, and organized into protocols for analysis. In general teachers were easy to handle with the software tools and its proposal for data manipulation. The tools used have provided a manipulation of data, since the teachers were able to move the plots for various actions, such as organization into categories in horizontal and vertical axes; inclusion of scales, stacking, sorting, use of resources and gradient colors. During the interview, the manipulation of data generated changes of the interpretations. In this sense, software tools were used in the construction of different representations at the time of use data, which allowed greater interaction with the data providing different insights in the construction of the response. Further studies are needed to propose the use of this software in a process of teacher education, in which statistical concepts are clearly constructed, giving the freedom of manipulation and construction of different representations. The research in teacher education could contribute in teachers' own work with Handling Data in classrooms.

Keywords: Software Tinkerplots; interpretation of graphs; rural schools teachers.

LISTA DE FIGURAS

Figura 1: Demonstração da tela inicial do *TinkerPlots*..... 27
Figura 2: *Card* apresentando um banco de dados 28
Figura 3: Modo de representação *Plots* do banco de dados dos gatos 28
Figura 4: Demonstração da ferramenta *gradiente* representando à variável gênero através dos *Plots* 29
Figura 5: Demonstração da ferramenta *gradiente* representando a variável idade através dos *Plots* 30
Figura 6: Demonstração do uso da ferramenta *separate* em dois agrupamentos da variável idade 31
Figura 7: Demonstração do uso das ferramentas *empilhar* e *ordenar* com a variável idade..... 31
Figura 8: Demonstração da utilização da ferramenta *separate* no banco de dados de peixe..... 32
Figura 9: Barra de ferramentas *Icon Type* aumentada 33
Figura 10: Demonstração de um gráfico pictórico a partir do banco de dados de gatos 33
Figura 11: Barra de ferramentas principal do *software* 34
Figura 12: Demonstração de um cruzamento de variáveis utilizando a ferramenta *gradiente* 34
Figura 13: Mochila de estudantes agrupadas de acordo com o peso. Fonte: Konold, 2006. 36
Figura 14: O peso de mochilas em relação a variável série. Fonte: Konold, 2006 36
Figura 15: Tabela para exemplificar o cruzamento entre duas variáveis qualitativas. (BUSSAB; MORETTIN, 2002, p. 75) 45
Figura 16: Gráfico de dispersão utilizado para exemplificar o cruzamento de variáveis quantitativas (MORETTIN; BUSSAB, 2002, p. 82) 45
Figura 17: Exemplos dos tipos de cruzamentos entre duas variáveis quantitativas. (MORETTIN; BUSSAB, 2002, p. 82) 46
Figura 18: Acesso as escolas rurais e meio de transporte utilizado 49
Figura 19: Foto da Escola Cantinho Feliz situada na região rural 50
Figura 20: Horta e coleta seletiva presentes nas escolas participantes 50
Figura 21: Banco de dados de gasto de água utilizado no estudo piloto 55
Figura 22: Demonstração da ferramenta *gradiente* no banco de dados de gatos no estudo piloto.... 57
Figura 23: Demonstração da ferramenta *separar* no banco de dados do piloto 58
Figura 24: Demonstração de cruzamento de variáveis no Banco de dados gatos no piloto 59
Figura 25: Representação construída pela professora Margarida que cruza tempo e gasto de água 61
Figura 26: Representação construída pela professora Rosa que cruza tempo e gasto de água 62
Figura 27: Banco de dados de violência sexual contra crianças utilizado no 3º momento do estudo principal 66
Figura 28: Banco de dados de Novelas exibidas no horário nobre utilizado no 3º momento do estudo principal 67
Figura 29: Tela inicial apresentada no *TinkerPlots* 75

Figura 30: Tela com o banco de dados dos gatos apresentado na Familiarização 75
Figura 31: Demonstração da ferramenta *gradiente* no banco de dados gatos na Familiarização 76
Figura 32: Demonstração da ferramenta *gradiente* indicando a mudança das cores de acordo com a variável 77
Figura 33: Demonstração da ferramenta *separate* agrupando a idade dos gatos em classes de intervalos..... 78
Figura 34: Demonstração da ferramenta *separate* na construção de classes pontuais para a variável Idade 78
Figura 35: Representação tipo 1 construída na familiarização que cruza a variável genero x tamanho 82
Figura 36: Representação tipo 2 construída na familiarização que cruza a variável gênero x tamanho 83
Figura 37: Representação tipo 3 construída na familiarização que cruza a variável gênero x tamanho 85
Figura 38: Representação tipo 4 construída na familiarização que cruza a variável gênero x tamanho 86
Figura 39: Demonstração da tela do computador no inicio da atividade com primeiro banco de dados de Violência 90
Figura 40: Demonstração da tela do computador no inicio da atividade com segundo banco de dados de Novelas 91
Figura 41: Exemplo 1 de representação dentro da tipificação B utilizada para responder a primeira questão do banco de dados violência 96
Figura 42: Exemplo 1 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados violência 97
Figura 43: Exemplo 2 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados violência 97
Figura 44: Exemplo 1 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados violência 98
Figura 45: Exemplo 1 de representação dentro da tipificação E utilizada para responder a segunda questão do banco de dados violência 99
Figura 46: Exemplo 1 de representação dentro da tipificação D utilizada para responder a segunda questão do banco de dados violência 100
Figura 47: Exemplo 1 de representação dentro da tipificação F utilizada para responder a terceira questão do banco de dados violência 101
Figura 48: Exemplo 1 de representação dentro da tipificação D utilizada para responder a terceira questão do banco de dados violência 101
Figura 49: Exemplo 1 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados Novelas 103
Figura 50: Exemplo 1 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados Novelas 104

Figura 51: Exemplo 2 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados Novelas 104
Figura 52: Exemplo 1 de representação dentro da tipificação A utilizada para responder a primeira questão do banco de dados Novelas 105
Figura 53: Exemplo 1 de representação dentro da tipificação D utilizada para responder a segunda questão do banco de dados Novelas 106
Figura 54: Exemplo 1 de representação dentro da tipificação E utilizada para responder a segunda questão do banco de dados Novelas 107
Figura 55: Exemplo 1 de representação dentro da tipificação F utilizada para responder a terceira questão do banco de dados Novelas 108
Figura 56: Exemplo 1 de representação dentro da tipificação A utilizada para responder a terceira questão do banco de dados Novelas 108
Figura 57: Exemplo 1 de representação dentro da tipificação C utilizada para responder a terceira questão do banco de dados Novelas 109
Figura 58: Exemplo 1 de representação dentro da tipificação D utilizada para responder a terceira questão do banco de dados Novelas 110
Figura 59: Exemplo 1 de uso de a ferramenta *empilhar* para responder a primeira questão do banco de dados Violência 114
Figura 60: Exemplo 2 de uso de a ferramenta *empilhar* associada a *separar* para responder a primeira questão do banco de dados Violência 116
Figura 61: Exemplo 1 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Violência 117
Figura 62: Exemplo 2 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Violência 119
Figura 63: Exemplo 2 de uso de a ferramenta *separar* associada a ferramenta *Gradiente* para responder a primeira questão do banco de dados Violência 120
Figura 64: Exemplo 1 de uso de a ferramenta *separar* associada a *empilhar* para responder a segunda questão do banco de dados Violência 122
Figura 65: Exemplo 2 de uso de a ferramenta *separar* associada a *empilhar* para responder a segunda questão do banco de dados Violência 124
Figura 66: Exemplo 1 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Violência 125
Figura 67: Exemplo 2 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Violência 126
Figura 68: Exemplo 1 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas 129
Figura 69: Exemplo 2 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas 131
Figura 70: Exemplo 3 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas 132

Figura 71: Exemplo 1 de uso de a ferramenta *gradiente* para responder a primeira questão do banco de dados Novelas 134
Figura 72: Exemplo 1 de uso de a ferramenta *separar* (criar categorias) para responder a segunda questão do banco de dados Novelas 136
Figura 73: Exemplo 2 de uso de a ferramenta *separar* (criar categorias) associada a ferramenta *número* para responder a segunda questão do banco de dados Novelas 137
Figura 74: Exemplo 1 de uso de a ferramenta *gradiente* para responder a terceira questão do banco de dados Novelas 139
Figura 75: Exemplo 1 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Novelas 141
Figura 76: Exemplo 2 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Novelas 143

LISTA DE QUADROS

Quadro 1: Etapas propostas para realização da Entrevista semi-estruturada 54
Quadro 2: Questões propostas na atividade realizada no 3º momento da coleta 65
Quadro 3: Representações produzidas pelos professores na familiarização 80
Quadro 4: Exemplos dos diferentes tipos de representações conforme características semelhantes 93
Quadro 5: Ferramentas disponíveis no *Software* e utilizadas pelos professores 111

LISTA DE GRÁFICOS

Gráfico 1: Idade dos Professores participantes da pesquisa 52
Gráfico 2: Formação profissional dos participantes da pesquisa 52
Gráfico 3: Experiência de ensino em salas de aulas de escolas rurais 53
Gráfico 4: Frequência dos tipos de uso de computadores pelos professores 70
Gráfico 5: Frequência de respostas dos participantes no que se refere ao Ensino de Tratamento da Informação 72
Gráfico 6: Representações construídas pelos professores de acordo com cada questão do primeiro banco de dados – Violência 95
Gráfico 7: Representações construídas pelos professores de acordo com cada questão do segundo banco de dados – Novelas 102
Gráfico 8: Percentual de uso de cada ferramenta utilizada para construir as representações 112
Gráfico 9: Ferramentas utilizadas para responder as questões do banco de dados Violência 113
Gráfico 10: Ferramentas utilizadas para responder as questões do banco de dados Novelas 128

SUMÁRIO

Resumo

Abstratc

Introdução 14

Capítulo 1 - Realidade rural e novas perspectivas..... 18

1.1. Estudos em escolas rurais..... 20

1.2. Ações para valorização de escolas rurais..... 21

Capítulo 2 – O uso de *software* para o ensino de estatística: O *TinkerPlots*..... 23

2.1. *Active Graphing*..... 24

2.2. O *software TINKERPLOTS*..... 27

2.3. Pesquisas que utilizam o *TinkerPlots*..... 35

Capítulo 3 – A utilização do contexto bivariado em uma proposta de análise exploratória de dados..... 40

3.1 Analise Exploratória de Dados..... 41

3.2 Contexto Bivariado..... 43

Capítulo 4 – Método e Estudo Piloto 48

4.1. Contexto da pesquisa..... 48

4.2. Os sujeitos da pesquisa 51

4.3. Escolhas metodológicas..... 53

4.4. Estudo Piloto..... 54

4.5. Estudo Principal..... 64

4.6. Desenho da análise..... 67

Capítulo 5 – Primeiro encontro..... 69

5.1. Entrevista semi-estruturada..... 69

5.2. Apresentação do *Software*..... 74

5.2.1. Esse é o *software TinkerPlots*..... 74

5.2.2. Primeiras manipulações com o *software*..... 79

Capítulo 6 – Segundo Encontro: Analisando as representações construídas a partir do *software* 89

6.1. As representações utilizadas 90

6.2. Primeiro banco de dados – violência sexual 94

6.3. Segundo banco de dados – Novelas exibidas no horário nobre 102

Capítulo 7 – Analisando o uso das ferramentas disponíveis no software e sua influência para interpretação de dados 111
7.1. Banco de dados Violência contra crianças – Ferramentas utilizadas 113
7.2. Banco de dados Novelas exibidas – Ferramentas utilizadas 128
Capítulo 8 – Considerações..... 146
Referências 151
Anexo..... 156

INTRODUÇÃO

A construção da cidadania está vinculada a participação do homem e da mulher ativamente na sociedade: nas questões políticas, econômicas e de luta por direitos de melhores condições de vida. Uma das maneiras de garantir a cidadania é oferecer uma educação que habilite as pessoas a lidar com as diversas informações veiculadas nas mídias, para interagir, criticar, entender, construir opiniões e tomar decisões sobre acontecimentos sociais.

Entre os diversos tipos de apresentação de informações utilizadas na sociedade contemporânea podemos citar os gráficos Estatísticos. Lima (1998) e Monteiro (1998) e mais recentemente Cazorla e Castro (2008) têm discutido em seus estudos os gráficos em contextos midiáticos. Cazorla e Castro (2008) apontam as armadilhas contidas em representações estatísticas. As autoras apresentam exemplos de manchetes de jornais de grande circulação que trazem informações tendenciosas e que ganham credibilidade por estarem atreladas a gráficos estatísticos, os quais trazem certas armadilhas, tais como: mudança de escala, valores relativos, não divulgação da amostra utilizada etc., que induzem o leitor a estabelecer determinadas interpretações. Essas autoras discutem a importância de desarmarmos essas armadilhas para compreendermos as informações ali contidas de maneira crítica.

Nesse sentido, podemos apontar a escola como ferramenta importante no processo de reflexão e construção dos conhecimentos estatísticos. No entanto, pude identificar, através dos estudos que venho desenvolvendo em escolas de regiões rurais do interior de Pernambuco, as dificuldades vividas pelos professores no que diz respeito ao trabalho em sala de aula vinculados as conceitos estatísticos (ASSEKER; MONTEIRO, 2008). Por vezes essas dificuldades não estavam apenas atreladas as metodologias de ensino, mas ao desconhecimento dos conceitos básicos de Estatística. Observei que os participantes tentavam melhorar a sua prática com o uso de recursos variados que acreditavam serem motivadores das aulas de Matemática, com os quais os alunos aprenderiam mais facilmente. Essa postura aberta ao novo e em busca por diferentes artifícios que motivassem os alunos a aprenderem foram identificados por nós no decorrer de nossas pesquisas e fizeram com que percebêssemos o quanto ainda poderíamos contribuir a partir de

novas pesquisas nesse campo de investigação. Foi nessa perspectiva que definimos os nossos sujeitos de pesquisa, professores de escolas rurais do Agreste de Pernambuco e a nossa área do conhecimento, a Estatística, enquanto conteúdo curricular.

Ao definirmos nossos sujeitos e objeto de pesquisa outro fator causou-nos inquietação. O aparecimento de *lan-houses* e a implantação de laboratórios de informática em escolas rurais pelo Proinfo Rural (Programa Nacional de Informática na Educação Rural) - que objetivam promover o uso pedagógico da informática na rede pública de Educação Básica em escolas rurais. A possibilidade de acesso aos computadores seja pelas *lan-houses* ou por meio dos laboratórios nas escolas, poderia ser encarada como uma oportunidade de uso para o ensino, no entanto, os professores demonstravam um desconhecimento do uso de computadores na sua própria prática social. Essa situação também nos motivou a investigar a utilização de um software para análise de dados junto a esses professores com o intuito de contribuir com a aproximação entre professores e uso de computadores para construção de conhecimentos. Nesse sentido busquei no grupo de pesquisa CENAGRAF[1] um apoio teórico e de utilização do software *TinkerPlots* (KONOLD; MILLER, 2001) como auxiliar ao ensino do Tratamento da informação.

Esse software vem sendo investigado em realidades fora do Brasil e pela primeira vez o grupo CENAGRAF utiliza-o em escolas brasileiras. Como as informações sobre o uso desse *software* no Brasil são escassas identificamos a necessidade de investigar do uso do *software Tinkerplots* como possível ferramenta para interpretação de dados por professores de escolas rurais do Agreste de Pernambuco.

Para responder esse objetivo buscamos identificar as representações construídas pelos professores a partir da utilização das ferramentas do *software;* analisar o uso das ferramentas *separar*, *empilhar*, *ordenar* e *gradiente* em atividades de cruzamento de variáveis; como também, analisar que influência o uso dessas ferramentas disponíveis no software proporciona no momento da interpretação dos dados em cada questão de pesquisa.

[1] Projeto financiado pela Facepe- Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco – que investiga a contribuição de diferentes recursos para interpretação e construção de gráficos. (2008)

Com o intuito de alcançar os objetivos propostos, selecionamos uma nucleação de escolas rurais em um Município situado na região Agreste de Pernambuco - Caruaru, no qual 12 professores se disponibilizaram em participar do estudo. Os professores, em um primeiro momento, conheceram o *software* e suas ferramentas através de uma sessão de familiarização. Em seguida, em um segundo momento, os professores responderam seis questões de cruzamento de variáveis em dois banco de dados apresentados no *TinkerPlots.* Eles puderam manipular os dados, construir representações e utilizar as ferramentas disponíveis para responder as questões. Todos os dois momentos foram videografados, tanto da fala dos professores como sua imagem e tudo que era construído na tela do computador. Analisamos, em seguida, as representações construídas por eles, seus tipos e especificidades; identificamos as ferramentas utilizadas e sua influência no momento da interpretação dos dados em cada questão proposta.

Delineando o formato desta dissertação, iniciamos o primeiro capítulo descrevendo com mais detalhes o campo de investigação em que foi desenvolvida a pesquisa e discutimos alguns artigos frutos de estudos que foram realizados nessa realidade para que pudéssemos ter clareza das especificidades locais e de como vem se estruturando o ensino da Matemática, campo do conhecimento ao qual está inserido o ensino do Tratamento da Informação (Estatística) para as séries iniciais.

No capítulo 2 apresentamos o *software TinkerPlots* e suas ferramentas para entendermos as possibilidades de representação dos dados oferecidas por ele, também apresentamos exemplos de estudos que utilizaram esse software. No capítulo 3 discutimos o campo da análise exploratória de dados e a utilização da cruzamento de variáveis focando as possibilidades da exploração de dados para construção de conhecimentos Estatísticos.

No capítulo 4 apresentamos o método utilizado, o perfil dos professores participantes, o estudo piloto, quais as atividades realizadas e como essas atividades foram desenvolvidas a partir da utilização do *software.* No capítulo 5 apresentamos o momento da familiarização com o *software* vivido pelos professores ao manipularem o computador. No capítulo 6 apresentamos e discutimos as representações utilizadas pelos professores em cada questão de cruzamento de variável. No capítulo 7 identificamos as ferramentas utilizadas para exploração de dados e suas influências na interpretação de cada questão de cruzamento de

variável. Concluímos apresentando nossas considerações e as perspectivas resultantes do estudo desenvolvido.

CAPÍTULO 1 - REALIDADE RURAL E NOVAS PERSPECTIVAS

Através das experiências de pesquisas vividas no Projeto PREMATER – Projeto Reconceptualizando e Usando Recursos no Ensino de Matemática em Escolas Rurais – pudemos observar e estudar as diversas práticas sociais e culturais vividas nas comunidades rurais e, por vezes, negligenciadas no processo de ensino/aprendizagem. Como identificado por Asseker e Monteiro (2008) os professores de escolas rurais não utilizavam em sua prática pedagógica as vivências culturais dos alunos de comunidades rurais, nem conceptualizavam como recursos para o ensino os materiais sócio-culturias ou a linguagem utilizadas nas comunidades.

Macêdo (2010) reflete que um dos grandes desafios hoje vivenciado pelo ensino, de uma forma geral, e mais evidente em escolas rurais é articular os conhecimentos construídos na escola com as experiências cotidianas. Essa autora discute, especificamente, os conhecimentos Matemáticos e sua utilização no dia-a-dia e como as escolas vêm trabalhando de maneira distanciada das necessidades sociais e culturais diárias.

Essa distância entre os conhecimentos produzidos na escola e os conhecimentos necessários na vida social se torna mais evidente nas comunidades rurais, pois, é fruto do estabelecimento de uma visão dicotômica entre rural e urbano. Trevizan (2003) aponta que o rural esteve ligado, por muito tempo, ao subdesenvolvimento, a parte da população pobre da sociedade, ao que é atrasado. Com isso, as políticas promotoras de desenvolvimento estiveram voltadas para os interesses urbanos e para as indústrias inseridas na cidade. Essa dicotomia era refletida não apenas nas políticas de desenvolvimento, mas, até mesmo, na produção literária que tratavam o rural como exótico, caracterizado pela pobreza intelectual e por uma mentalidade estreita (TREVISAN, 2003). Essa visão dicotômica esteve muito presente nas escolas rurais e, ainda hoje tem resíduos marcantes.

De acordo com Arroyo, Caldart e Molina (2004) a educação nas áreas rurais esteve ligada a uma visão cultural e de necessidades urbanas. Não temos nenhum documento ou política educacional vinculada a construção de uma educação que considerasse a realidade campesina, nesse sentido os conteúdos curriculares, as

metodologias e propostas pedagógicas das escolas da cidade eram transpostos para o campo desconsiderando a atenção ao cidadão enquanto sujeito munido de potencialidades que poderiam ser mais bem desenvolvidas se assegurados mecanismos de promoção enquanto direito humano (MONTEIRO; LEITÃO; ASSEKER, 2010)

Wanderley (2001) e Veiga (2002) em seus trabalhos de investigação vêm discutindo a necessidade de entendermos o "rural" como um universo que possui especificidades sociais, culturais e regionais, mas que está integrado ao conjunto da sociedade brasileira e ao contexto atual das relações internacionais. Essa nova forma de entender os contextos rurais pode garantir o direito as particularidades históricas, sociais, culturais e ecológicas da população campesina atrelada ao desenvolvimento social, tecnológico e de direitos a nível global.

Veiga (2002) aponta que é necessário entender as particularidades do rural como em sua dupla face, simultaneamente natural e social. Dimensão natural pelo predomínio da paisagem composta por elementos da natureza mais do que construções arquitetônicas, e a dimensão social caracterizada por relações sócio-culturais específicas, baseadas nos pequenos grupos humanos nos quais predominariam relações de interconhecimento.

Esses aspectos qualitativos dão identidade aos espaços rurais, também propiciam uma concepção de rural que não se restringe a uma dimensão geográfica, apenas física ou a algo oposto ao que seria urbano. As qualidades da pluralidade rural não interessam apenas aos que vivem nesses espaços, mas constituem-se em valores culturais importantes para a sociedade como um todo. Nessa perspectiva, o rural deixa de ser o não-urbano ou o não lugar, passa a ter suas próprias especificidades, mas ao mesmo tempo adquire uma relação de reciprocidade com os ditos espaços urbanos (WANDERLEY, 2004).

Nesse sentido a escola rural tem como um de seus papeis enxergar a população campesina como indivíduos culturalmente constituídos com vivências específicas, com experiências diferenciadas e meio de produção vinculados a terra e ao mesmo tempo fazem parte de uma realidade global.

No entanto, através de estudos realizados nos últimos anos podemos perceber que essa proposta ainda vem sendo pouco vivenciada nas escolas rurais, como discutimos a seguir.

1.1. Estudos em Escolas Rurais

A partir de estudos realizados nos últimos anos podemos traçar um perfil da realidade das escolas rurais, principalmente, as situadas na região Agreste de Pernambuco. O estudo de Melo, Leitão e Alves (2007) investigou o perfil de 62 professores de escolas rurais e constatou um percentual de 25% de professores que não tinham formação específica para lecionar, 30% dos professores tinham graduação em Pedagogia, 22% estavam cursando Pedagogia e 23% tinham formação a nível de Magistério. Esses dados indicam uma tendência dos professores a buscarem aperfeiçoamento, no entanto, elas identificaram que a concepção de recurso dessas professoras estava vinculada a uma dimensão material e culpabilizavam a falta de recursos para o ensino da Matemática como fator primordial para o não aprendizado dos alunos nessa área. Nesse sentido podemos entender que mesmo a maioria dos professores estando em processo de formação ou tendo concluído a graduação em Pedagogia não é suficiente para perceber os recursos disponíveis e que estão atrelados ao cultural e ao humano. Na verdade eles não conseguem se ver enquanto recurso.

Asseker e Monteiro (2008) discutem o consenso de professores de escolas rurais sobre o ensino de Matemática e constatam em suas falas uma relação difícil em seu histórico de aprendizagem da Matemática, demonstrando certos temores em relação a essa área do conhecimento o que acaba por influenciar sua prática pedagógica. Mesmo demonstrando esses temores e dificuldades conceituais verbalizavam a necessidade de mudar a prática em sala, demonstrando uma predisposição a utilização de jogos para facilitar a aprendizagem da Matemática. No entanto, o que ficou mais explícito nas análises das falas foi a necessidade de formação continuada abordando tanto conceitos teóricos matemáticos como também a abordagem de recursos dentro de um planejamento pedagógico com objetivos claros para os alunos.

Monteiro e Alves (2008) realizaram uma análise na coleção dos guias da Escola Ativa[2] de Matemática que eram utilizados pelas professoras da Rede Municipal de Ensino de Caruaru participantes da pesquisa, livros esses que eram o recurso principal de pesquisa utilizado pelas professoras. Esses livros apresentavam

[2] Projeto do Governo Federal que tem como objetivo a melhoria do ensino nas escolas de áreas rurais brasileiras.

atividades relacionadas ao Eixo[3] do Ensino da Matemática "Números de Operações", as atividades vinculadas ao ensino da Estatística – Eixo Tratamento da Informação - quase que inexistia. Foi observado, ainda, que as atividades continham experiências e vivências baseadas em estereótipos do rural apenas agrícola. Eram atividades iguais as utilizadas na cidade, porém utilizando contextos de realidades rurais que não condiziam com a realidade do rural vivenciado pelos alunos. Por exemplo, os livros mostravam imagens de pessoas com roupas quadriculadas, descalças, usando chapéu de palha, em carroças etc.

Os dados acima apontados permitem uma visão do ensino nas escolas rurais, ainda, muito precário e com necessidade de investimentos, formação para os professores e políticas públicas que garantam a melhoria desse sistema de ensino. Dentre essas propostas de melhoria de ensino nessas escolas podemos identificar o Proinforural e a Escola Ativa que são Programas do Governo Federal vinculados a melhoria da Educação Básica e que abordaremos a seguir.

1.2. Ações para Valorização de Escolas Rurais

Dentre as ações desenvolvidas na atualidade para melhoria da educação básica em áreas rurais podemos citar como exemplos o Proinforural e a Escola Ativa. Esses programas atendem principalmente aos primeiros anos de escolarização de crianças residentes nas áreas rurais.

O Proinforural - Programa Nacional de Informática na Educação Rural - objetiva promover o uso pedagógico de Tecnologias de Informática e Comunicações (TICs) na rede pública de ensino em parceria com Estados e Municípios criado pela Portaria nº 522/MEC, de 9 de abril de 1997.

Sua principal ação é a construção de laboratórios de informática em escolas públicas possibilitando aos alunos o acesso as tecnologias da informação, com o objetivo maior da Inclusão Digital. O programa acontece a partir da compra de computadores pelo MEC que encaminha para as escolas e instala os laboratórios de informática. Em contrapartida, os governos locais (prefeituras e governos estaduais)

[3] De acordo com os PCNs a Matemática foi dividida em quatro grandes eixos do currículo dos conteúdos curriculares, são eles: Números e Operações; Espaço e Forma, Grandezas e Medidas e Tratamento da informação.

providenciam a infra-estrutura das escolas, indispensável para que elas recebam os computadores.

As características das escolas rurais que podem participar do programa são: Escolas de ensino fundamental (1º ao 9º ano), Área Rural, ter mais de 30 alunos, energia elétrica e sem laboratório de informática. Essas escolas ganham cinco terminais de acesso com 1 CPU, monitor LCD, impressora jato de tinta, wirelles, Linux Educacional 3.0, Mobiliário (mesas e cadeiras) – enviados pelo MEC.

O Programa Escola Ativa do Ministério da educação é um Programa do Fundo Nacional de Desenvolvimento da Escola (FNDE - Fundescola - MEC) criado em 1997, especificamente para as escolas rurais, segundo fontes do próprio MEC. O objetivo da Escola Ativa é melhorar a qualidade do ensino fundamental de 1ª a 4ª séries, reduzir a distorção idade/série e aperfeiçoar o nível de aprendizagem nas áreas mais populosas das regiões Norte, Nordeste e Centro-Oeste. Em termos nacionais, o Projeto Escola Ativa foi inicialmente implantado em 384 municípios dessas regiões e se caracteriza por uma metodologia destinada às classes multisseriadas da zona rural que reúne trabalho em grupo, auto-aprendizagem, ensino por meio de livros didáticos específicos, participação da comunidade e acompanhamento de alunos, além de formação continuada para os professores.

Além das políticas governamentais propostas objetivando o desenvolvimento de escolas em áreas rurais, acreditamos que outra forma de levar conhecimento para essas regiões é a realização de pesquisas científicas que tragam em seu arcabouço a concepção de rural com suas especificidades. Ao ser realizado um estudo junto a escolas rurais há uma divulgação de conceitos e troca de experiências entre pesquisador e o pesquisado. Essa troca pode se configurar um processo de valorização do rural e de sensibilização ao estudo pelos envolvidos.

No caso deste estudo temos como dois grandes focos as tecnologias, uso do computador, e a Educação Estatística. Acreditamos que ao levarmos o computador e questões vinculadas à exploração de dados possamos despertar o interesse sobre essas temáticas nos participantes contribuindo, de certa forma, mesmo que indiretamente, para o ensino.

A seguir discutiremos mais aprofundadamente o uso de softwares educacionais em algumas pesquisas e apresentaremos o *software TinkerPlots* investigado neste estudo

CAPÍTULO 2 – O USO DE *SOFTWARE* PARA O ENSINO DE ESTATÍSTICA: O *TINKERPLOTS*

O uso de tecnologias para o ensino vem sendo apontado por alguns autores, como Menezes (2001, 2002), Passos (2003), entre outros com uma ferramenta que pode promover reflexões sobre conhecimentos diversos. Mais especificamente, autores como Menezes (2001, 2002), Bittar (2006) e Bellemain, Bellemain e Gitirana (2006) tem discutido as contribuições da utilização de *softwares* para o ensino, no que refere as possibilidades geradas através de sons, imagens, simulações, fazer e reverter ações, entre outros que possibilitam a construção de novos conhecimentos.

A Estatística, em particular, é um campo do conhecimento que vem proporcionando a criação de diferentes softwares para análise de dados, e mais recentemente, os *softwares* destinados ao ensino de conceitos Estatísticos. Alguns desses *softwares* vêm sendo investigados em estudos como os de Healy, Magina, Costa e Santos (2000) que analisaram o uso do *software Tabletop* para interpretação de gráficos. Eles afirmaram que os objetos dispostos na tela como constituintes da representação gráfica auxiliaram na análise dos dados realizadas pelos participantes de seu estudo.

Santos e Magina (2008) realizaram um estudo de caso com uma professora, de séries iniciais do Ensino Fundamental sobre a utilização do *software Tabletop* para a resolução de problemas. A professora resolveu atividades que envolviam a coleta e organização dos dados, apresentação tabular e gráfica dos mesmos, bem como a análise dos dados para a resolução de problemas através do *software*. Como resultado as autoras apontam para a necessidade do professor vivenciar diferentes situações como organizar, categorizar, coletar dados de uma pesquisa para auxiliar na compreensão de conceitos estatísticos. As autoras identificaram que a manipulação de dados no ambiente computacional contribuiu para a ampliação das possibilidades de representações e interpretações feitas pela professora, contudo elas apresentaram dificuldades em lidar com as representações icônicas (representações que utilizavam figuras para indicar cada caso em particular). Essa dificuldade com a leitura e interpretação gráfica fazia com que a professora utilizasse sempre a representação numérica (com a frequência de ícones), pois lhe parecia

mais familiar, similar a uma tabela. Nesse caso, o uso da tecnologia proporcionou novas possibilidades na prática da interpretação de dados.

Lima (2005) investigou a compreensão de média aritmética por alunos através do uso do ambiente computacional. O estudo foi realizado em 2 turmas de 4ª série de uma escola pública. Todos os alunos participaram de um pré-teste o qual identificava os seus conhecimentos em relação ao conceito e média aritmética. Em uma das turmas foi realizado a intervenção com a utilização do *Tabletop* e ao final as duas turmas realizaram o pós-teste. Como resultado a autora percebeu que os alunos encontraram uma maior facilidade em relacionar os gráficos às tabelas quando participavam da coleta de dados. Os alunos que participaram da intervenção obtiveram resultados superiores aos que não utilizaram o *Tabletop*, nesse sentido, a autora aponta o *software* como uma ferramenta interessante para exploração de dados e a visualização de diferentes representações.

Os estudos apresentados nesta seção sugerem as contribuições da utilização de ferramentas computacionais para o trabalho em sala de aula. Todavia um grande desafio para o professor é colocar todo o potencial da tecnologia a serviço do aperfeiçoamento do processo educacional. É o professor quem vai propor o uso de ferramentas informatizadas capazes de criar as situações favoráveis à aprendizagem dos conceitos e à superação das dificuldades dos alunos.

Além dos dados apontados por essas pesquisas – que o uso de softwares para o ensino da Estatística pode ser ferramenta importante para a compreensão de conceitos dessa área – temos o grupo de pesquisadores Ainley, Nardi e Pratt (1998) que desenvolveram a perspectiva do *active graphing*, no qual afirmam que para promover uma maior reflexão sobre construção e interpretação de gráficos é importante o estudante participar do processo de coleta, organização e construção dos gráficos. A participação nesse processo favorece um maior envolvimento com os dados e por consequência o aparecimento de questões e argumentações interessantes e importantes para a construção de conceitos Estatísticos (AINLEY; NARDI; PRATT; 1998).

2.1. *Active Graphing*

Ainley, Nardi e Pratt (1998) vêm desenvolvendo em seus estudos o potencial do uso de computadores na escola atrelado ao processo de interpretação e

construção de gráficos. Eles desenvolveram pesquisas com crianças das séries iniciais de escolarização com a interpretação de gráficos e o uso do ambiente computacional e identificaram que a aquisição de habilidades de interpretação de gráficos auxilia no desenvolvimento de melhores estratégias para a construção dos mesmos. Assim, crianças que são mais estimuladas a interpretar gráficos podem desenvolver uma maior facilidade em produzir gráficos.

Em suas pesquisas os autores desenvolvem uma atividade prática com os alunos em sala com o auxílio do computador criando um banco de dados a partir de um experimento realizado junto aos alunos que resultou em um gráfico a ser interpretado pelos mesmos. A interpretação de gráficos, nesse sentido, não foi utilizada para responder atividades em outros eixos do ensino da Matemática, mas para entender como se comportou um dado evento, que variação foi percebida, qual a tendência do gráfico e principalmente prever comportamentos futuros da variável em um dado evento.

Ainley (2000) afirma que a prática utilizada nas escolas valoriza a construção de gráficos com um fim em si mesmo, como um conteúdo à parte do próprio processo de interpretação. Pratt (1995) denomina essa prática de *construção passiva*, pois o foco está atrelado à construção, muito tempo é dedicado a medidas das escalas, a produção, a pintura do gráfico ficando a interpretação em segundo plano. Na perspectiva da *construção passiva* é comumente oferecido ao aluno limitadas representações gráficas, como por exemplo, nos livros didáticos e em pesquisas são encontrados com maior frequência os gráficos de barras, com a justificativa de serem “mais atraentes” como mais fáceis de serem construídos (SELVA, 2003).

No entanto, Ainley, Nardi e Pratt (1998) defendem o que eles denominam de *active graphing* que seria a interpretação ativa de gráficos, no qual a ênfase seria na interpretação, no levantamento de hipóteses e na busca por entender os dados e de como eles foram produzidos. O computador, nessa perspectiva, é uma ferramenta que auxilia a construção de diferentes representações sem a demanda de um longo período de tempo, o que liberaria o tempo do aluno para as questões de interpretação. Eles apontam para necessidade que o leitor de gráficos tem em entender os processos, as estratégias relacionadas à própria produção do gráfico (como?, porque? e para que?).

Pratt (1995) sugere três competências essenciais a serem construídas pelos alunos: a primeira se refere ao entendimento de como interpretar e utilizar os gráficos, a segunda seria o conhecimento das convenções e das técnicas envolvidas nas representações gráficas como uso da escala ou, por exemplo, o significado de média e a terceira seria a habilidade prática para produzir gráficos.

Os estudos realizados por esse grupo de pesquisadores (Ainley; Nardi; Pratt; 1998) sugerem que os conceitos estatísticos podem ser menos difíceis do que eles parecem com o uso de ferramentas computacionais e as crianças participantes de sua pesquisa apresentaram níveis interessantes de interpretação mesmo nos gráficos menos utilizados em sala de aula, como os de dispersão e os de linha.

Por exemplo, Pratt (1994, 1995) investigou crianças americanas com idade entre 8 e 11 anos utilizando a perspectiva dos *active graphing*. Ele levou para sala de aula um avião de papel no qual, junto com as crianças, testavam a altura do seu vôo e o comprimento da asa e organizavam em uma planilha em um *software* comercial. Após alguns testes, a partir da planilha já com os dados foi construído um gráfico do tipo dispersão no qual as crianças buscavam entender qual o melhor comprimento para se obter vôos mais altos. As crianças faziam conjecturas, faziam novas experiências e as discussões eram registradas até que eles fossem capazes de decidir sobre o comprimento da asa. Esse é um exemplo de pesquisa em que a interpretação do gráfico e as relações entre variáveis foi o foco principal, os alunos puderam observar a representação, questionar os dados, inserir novos dados para obter uma informação.

No nosso estudo partimos dessa perspectiva por entender as diversas possibilidades que possam surgir a partir da interpretação de gráficos representados de diferentes maneiras. Uma interpretação livre para levantar hipóteses, questionar os dados, entender o porquê das tendências e discordar de certos dados. Nesse sentido, nosso foco foi a interpretação e as possibilidades que as ferramentas do *software TinkerPlots* puderam promover entre professoras com pouco ou nenhum conhecimento em informática e em Estatística.

O *software TinkerPlots* é um programa educacional para o ensino da Estatística a crianças e jovens. O uso desse recurso computacional vem sendo apontado por alguns autores como Konold (2001), Bem-ziv (2006), entre outros, como uma ferramenta que pode contribuir para construção de conceitos Estatísticos.

Para entendermos o seu funcionamento e suas possibilidades discutiremos a seguir seu *layout*, suas ferramentas e as suas possibilidades de uso.

2.2. O *Software TinkerPlots*

TinkerPlots é um programa de computador que foi desenvolvido por Kolnold e Miller (2001) com o objetivo de oferecer ferramentas, para estudantes de 9 anos em diante, que possibilitassem a exploração de dados.

A tela inicial do *software* (Figura 1) não oferece maiores atrativos, pois sua utilização está atrelada a construção de bancos de dados pelo próprio usuário ou ao acesso dos bancos de dados dispostos nos seus arquivos.

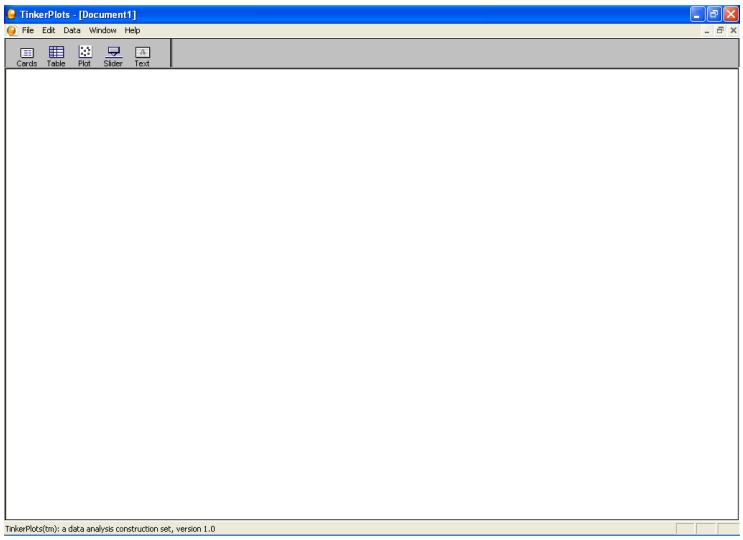


Figura 1: Demonstração da tela inicial do *TinkerPlots*

A partir da tela inicial do *software* temos as possibilidades de apresentação ou construção de um banco de dados. Como proposta de apresentação de banco de dados o *software* oferece a possibilidade de *card* que são os cartões que representam cada caso particularmente (Figura 2). Nesse formato o usuário pode construir e visualizar um banco de dados a partir da determinação do atributo ou variável observada com a respectiva informação de cada sujeito.

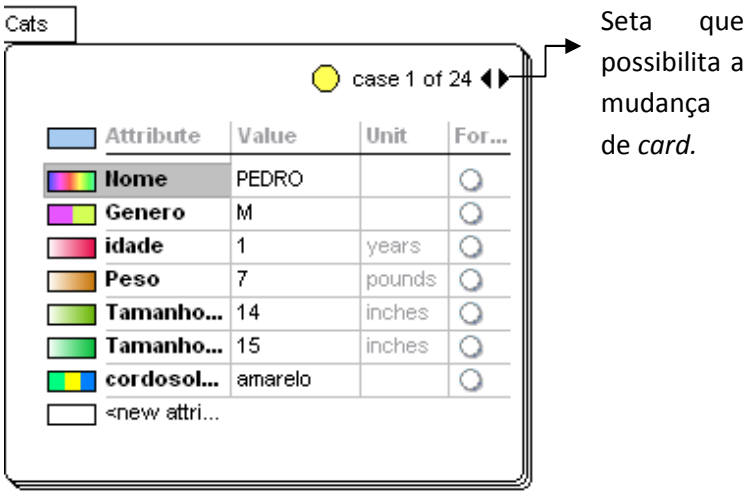


Figura 2: *Card* apresentando um banco de dados

A Figura 2 apresenta um banco de dados sobre “gatos”, no qual identificamos as variáveis: nome, gênero, idade, peso, tamanho etc, como também os dados referentes ao gato Pedro.

Outra forma de apresentação do banco de dados é através dos *Plots*, essa forma oferece possibilidades de manipulação, e de construção de várias representações dos dados (gráficos variados) e de cruzamento de variáveis. A Figura 3 apresenta os *plots* do banco de dados de gatos.

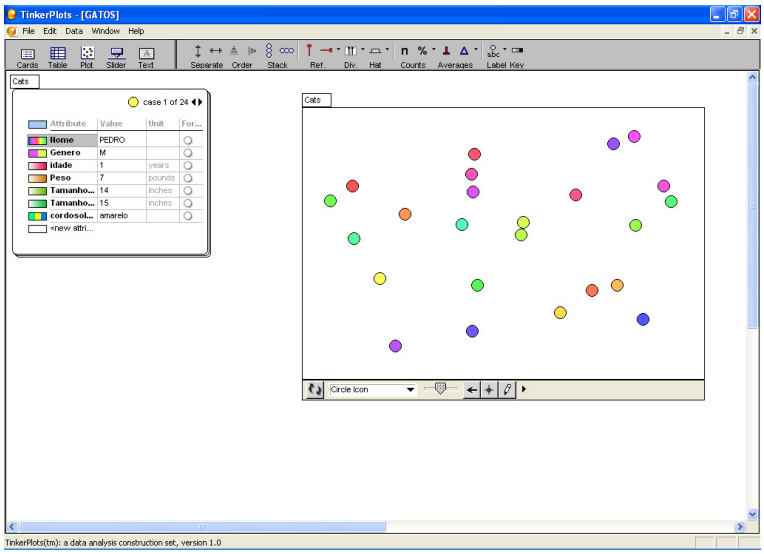


Figura 3: Modo de representação *Plots* do banco de dados dos gatos

As formas de apresentação (*cards* e *plots)* estão interligadas e apresentam o mesmo banco de dados de maneira diferente. No entanto o *cards* e o *plots* são

diretamente relacionados. Por exemplo, cada *card* apresenta as variáveis pesquisadas e as informações de cada sujeito da pesquisa, então quando uma variável está selecionada (achurada) ela automaticamente é representada através da mudança de cor nos *plots.* Os aspectos relacionados as cores dos *plots* são possibilitados pela ferramenta *gradiente*. O exemplo exposto na Figura 4 é a variável gênero (achurada nos *cards*) no qual os *plots* ficam de duas cores diferentes representando as categorias para essa variável (masculino e feminino).

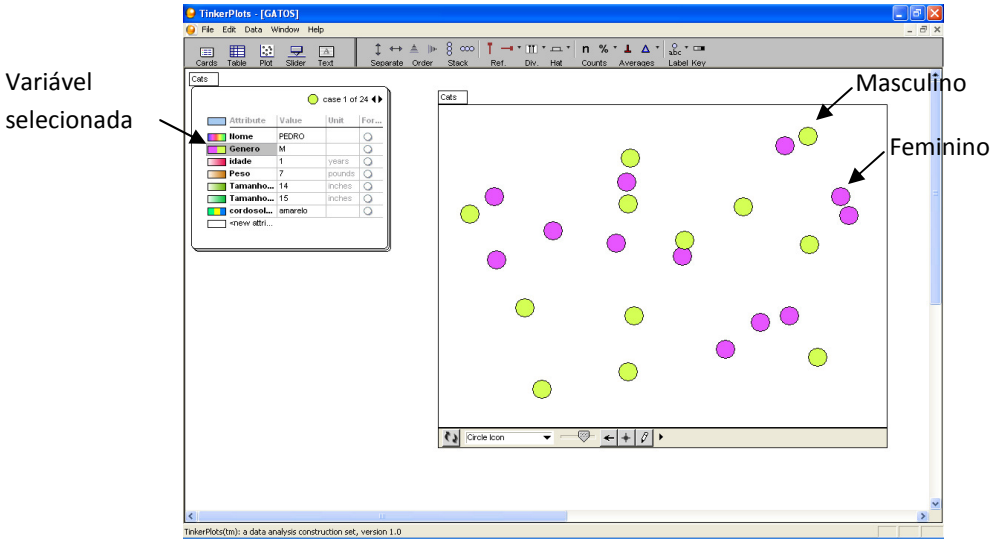


Figura 4: Demonstração da ferramenta *gradiente* representando à variável gênero através dos *Plots*

Nessa perspectiva as cores servem como indicadores das variáveis, quando trata de variável qualitativa, como no exemplo da Figura 4 (gênero). O *software* seleciona cores diferentes que indicam as categorias. Já quando trata de variável quantitativa como no caso da Figura 5 (idade) as cores são dispostas em forma de *gradiente* (dégradé) onde os mais claros representam os gatos mais jovens e a partir do escurecimento da cor apresenta as idades maiores.

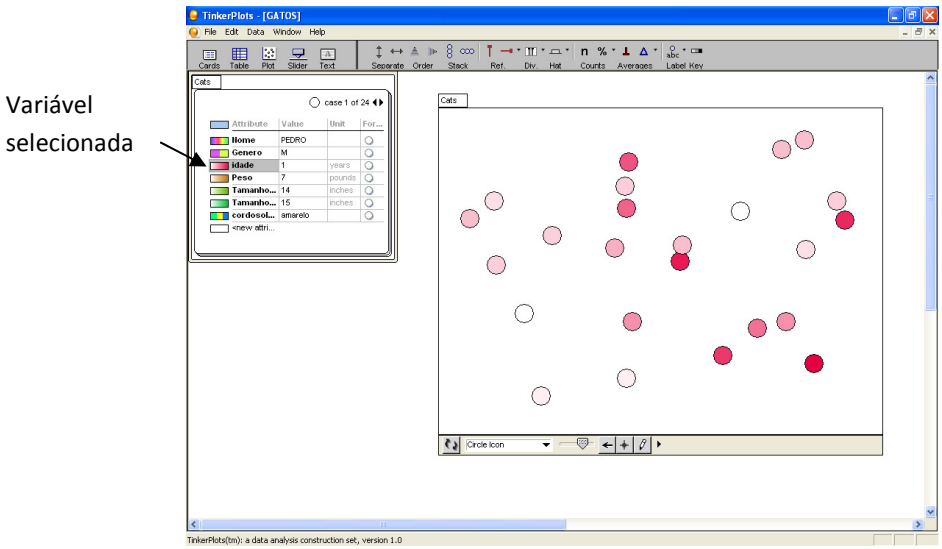


Figura 5: Demonstração da ferramenta *gradiente* representando à variável idade através dos *Plots*

Outras ferramentas disponíveis no software são: *separate*, *order*, *stack* ou *separar*, *ordenar* e *empilhar* respectivamente. Konold (2006) argumenta que essas operações básicas foram observadas em ações de estudantes ao manipular fichas em cartões de papel que representavam dados individuais em um banco de dados. Ele percebeu que os estudantes, para tentar interpretar os dados, tendiam a separar os cartões agrupando e ordenando cada grupo de casos e somente depois analisavam as informações. A partir da observação dessa prática dos estudantes com as fichas em papel foi que ele desenvolveu a idéia de criar um *software* que possibilitasse realizar essas ações, contudo, ampliando as possibilidades.

A ferramenta *separar* é acionada ao clicar-se sobre um *plot* (qualquer que seja), segurar e arrastar. Com o uso dessa ferramenta é possível formar agrupamentos e de acordo com a distância e a direção do arrastar do *plot* podemos chegar a formar escalas (pontuais ou categorias) de acordo com a variável. Na Figura 6 são mostrados dois quadros com tipos diferentes de organização para a variável idade do banco de dados dos gatos (essa possibilidade de agrupar é chamada pelo *software* de *separate*).

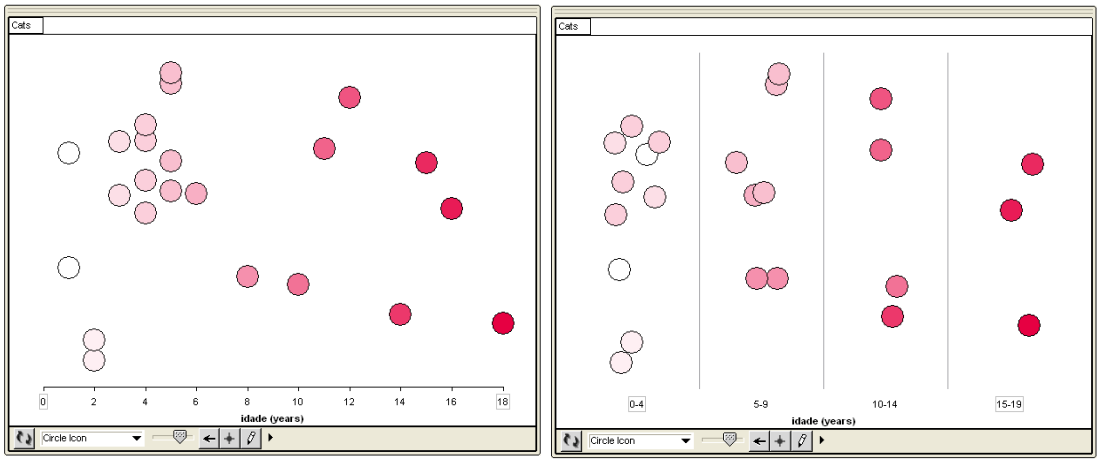


Figura 6: Demonstração do uso da ferramenta *separate* em dois agrupamentos da variável idade

A Imagem 1 da Figura 6 apresenta a idade organizada em uma escala pontual em um eixo horizontal. Essa organização foi possível com a ação de *clicar* em qualquer *plot* segurar e arrastar para o lado direito até o final da tela. Já a Imagem 2 da Figura 6 apresenta a mesma variável idade em uma escala intervalar (intervalos de 5 anos) em um eixo horizontal. Essa organização foi construída com o *clicar* em qualquer *plot*, segurar e arrastar para o lado direito percorrendo uma distância menor que o final da tela.

Outra ferramenta disponível é a *ordenar.* Ela é utilizada para ordenar em ordem crescente os casos de uma variável, essa ferramenta pode ser associada a *separate* para auxiliar na interpretação, também pode ser associada a ferramenta *stack* (empilhar) que concentra os *plots* um em cima do outro, como mostra a Figura 7.

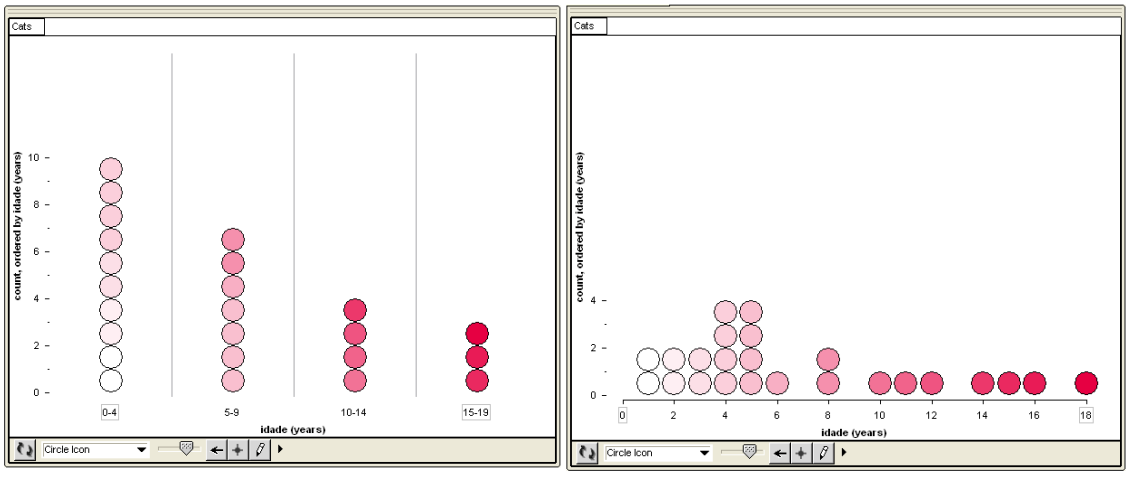


Figura 7: Demonstração do uso das ferramentas *empilhar* e *ordenar* com a variável idade.

A partir da manipulação dos *plots* (clicando e arrastando) é possível construir diferentes formas de organização de dados. A Figura 8, por exemplo, apresenta um banco de dados sobre peixes que foram divididos em peixes normais e geneticamente modificados. Com um simples clicar na categoria tipos no *cards* (indicada pela seta) e clicar e arrastar qualquer *plot* para cima, modifica-se a representação apresentada, a qual visualiza os peixes normais num plano acima dos geneticamente modificados.

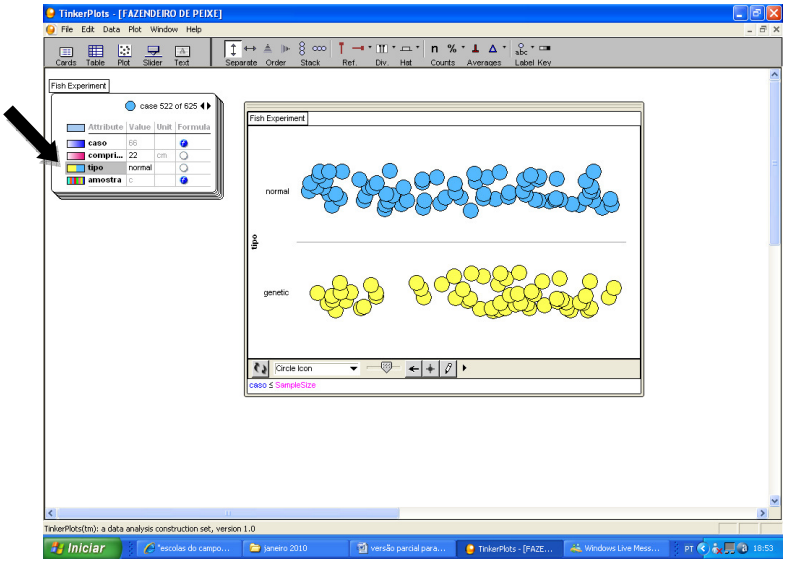


Figura 8: Demonstração da utilização da ferramenta *separate* no banco de dados de peixe

O usuário do *TinkerPlots* ao invés de escolher uma representação padrão para que o *software* apresente uma representação, eles manipulam os *plots* através do *mouse* a partir de ações simples como ordenar, separar e empilhar que em combinações diferentes podem gerar gráficos padrões e outras possibilidades de cruzar atributos, agrupando-os para possibilitar uma maior interação com os dados.

Além dessas ferramentas o *software* dispõe de representações formalizadas na Estatística como gráfico de barras, de setores, dispersão e histogramas a partir da barra de ferramentas abaixo da representação dos *plots* como mostra Figura 9.

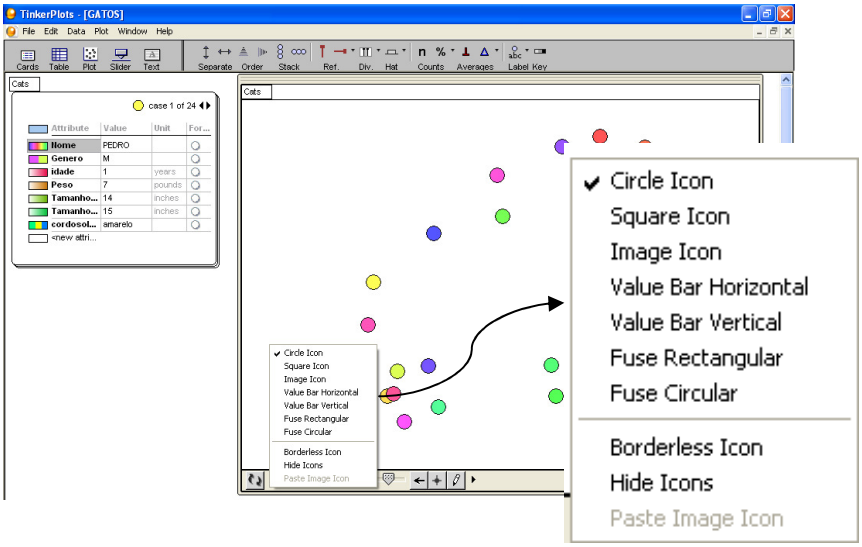


Figura 9: Barra de ferramentas *Icon Type* aumentada

Clicando nas opções indicadas na barra de ferramentas *icon type* é possível construir gráficos pictóricos com a inserção de figuras no lugar dos *plots* como mostra a Figura 10.

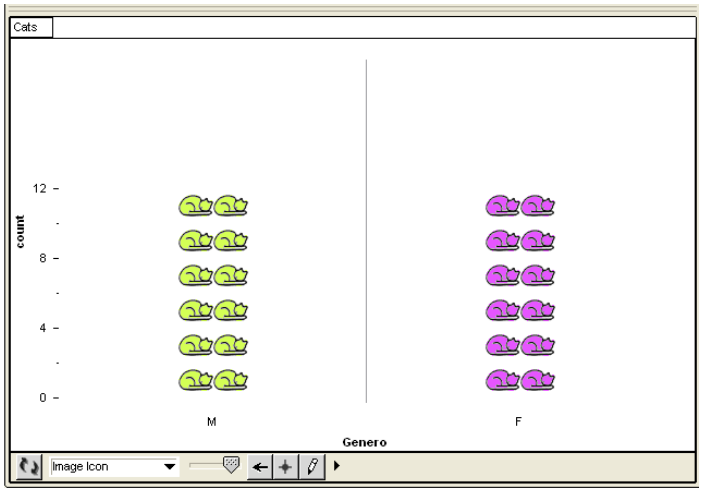


Figura 10: Demonstração de um gráfico pictórico a partir do banco de dados de gatos

A Figura 11 apresenta os ícones dispostos na barra de ferramentas do *software* responsáveis pelas ações, respectivamente, de *separar*, *ordenar*, *empilhar*, *linha de referência*, *média*, *mediana*, *quantidade*, *percentual*, *média*, *nome* e *legenda*.

Figura 11: Barra de ferramentas principal do *software*

Essas ferramentas auxiliam na construção de diferentes formas de visualizar os dados e cruzar os atributos. Em diferentes combinações podemos identificar, por exemplo, a frequência ou variabilidade de um determinado atributo, ou ainda, levantar e testar hipóteses nos cruzamentos de dados.

Além das operações básicas no *TinkerPlots* os estudantes têm a possibilidade de cruzar variáveis. Para isso o usuário pode clicar no nome da variável que deseja arrastar até que forme categorias ou a escala correspondente aos valores da variável e em seguida selecionar a outra variável a ser cruzada que os *plots* mudam de cor representando a variável desejada. Por exemplo, se o aluno separou seus dados de acordo com a idade dos participantes (como quadro um da Figura 12) e quer saber se os mais velhos, na sua maioria, são homens ou mulheres. Para isso, basta clicar na variável gênero que os *plots* mudam de cor apresentando uma cor para homem e outra para mulher (Figura 12).

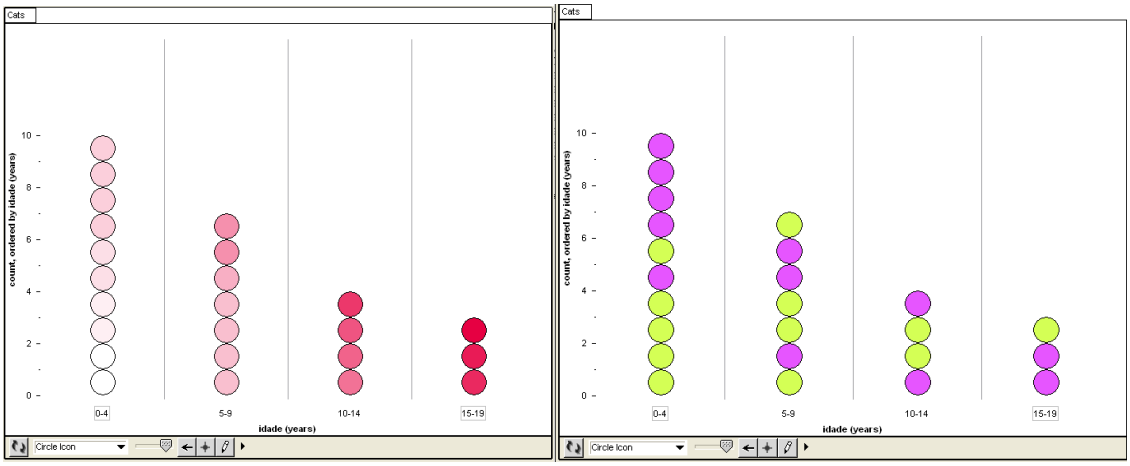


Figura 12: Demonstração de um cruzamento de variáveis utilizando a ferramenta *gradiente*

Na Imagem 1 da Figura 12 observamos o banco de dados apresentando apenas a variável idade, os *plots* estão agrupados por idade, no entanto, na Imagem 2 da Figura 12 a mesma organização é mantida, porém as cores indicam o sexo (lilás indica fêmea e amarelo representa macho). Essa forma de organização possibilita o cruzamento de atributos sendo evidenciado pela ferramenta *gradiente*

(cor). Nesse sentido podemos ao olhar para esse gráfico identificar que existem uma quantidade maior de gatas entre 10 e 19 anos do que gatos machos.

Essas diferentes possibilidades de representação oferecem ao aluno o poder de levantar e testar hipóteses, organizar e reorganizar os dados, fazer conjecturas e inferências. A cada nova ação um campo de possibilidades e questões podem ser levantadas. O aluno é construtor do gráfico através das ferramentas do *software* e não apenas um "escolhedor" de tipos de gráficos em um menu. E nessa atividade de construir e desconstruir os alunos acabam percebendo as diferenças em escalas, em organização dos ícones, um gráfico mais apropriado que represente a informação entre outros aspectos que fazem parte das habilidades de interpretação e construção de gráficos.

Pesquisas realizadas com o *TinkerPlots* já apresentam resultados positivos quando da sua utilização para o ensino. Na seção seguinte apresentaremos algumas dessas pesquisas.

2.3. Pesquisas que Utilizam o *TinkerPlots*

Em diversos países alguns estudos vêm sendo realizados com a utilização do *TinkerPlots* como ferramenta para o ensino da Estatística. Entretanto, no Brasil essas pesquisas são escassas, sendo o grupo de estudos CENAGRAF foi o primeiro a desenvolver pesquisas com esse *Software.* Os estudos realizados fora do país nos dão indicativos para o caminho metodológico utilizado com o uso desse *software*, a seguir iremos apresentar e discutir alguns desses estudos.

Konold (2006) realizou um estudo com crianças americanas de 10 anos, apresentando um banco de dados de uma pesquisa realizada sobre o peso da bolsa de estudantes de escolas americanas. Ele lança uma questão para os estudantes "Será que os estudantes de séries mais avançadas carregam mochilas mais pesadas?". Os alunos prontamente responderam que sim, e justificaram afirmando que eles carregam uma maior quantidade de livros. Então, os alunos recorreram ao banco de dados no *TinkerPlots* para confirmar essa suposição e representar através do *software* e da manipulação dos *plots* uma forma que pudesse responder a questão. Inicialmente os alunos separam os dados de acordo com o peso das mochilas (Figura 13) e em seguida separam de acordo com a série (Figura 17). A hipótese foi confirmada, porém essa representação não foi a única, os alunos

representaram os dados de maneira diversa o que o autor analisou como positivo, pois todos os alunos construíram diferentes gráficos e utilizaram estratégias para responder a questão.

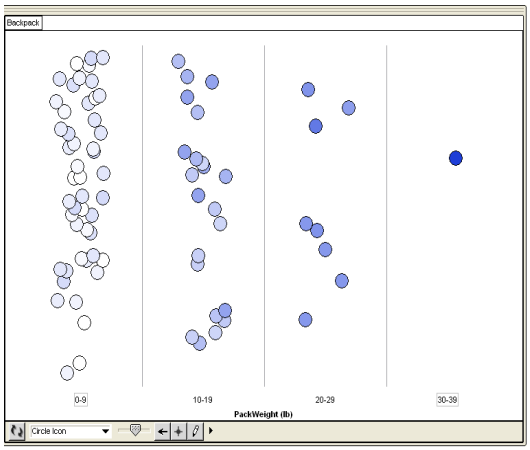


Figura 13: Mochila de estudantes agrupadas de acordo com o peso. Fonte: Konold, 2006.

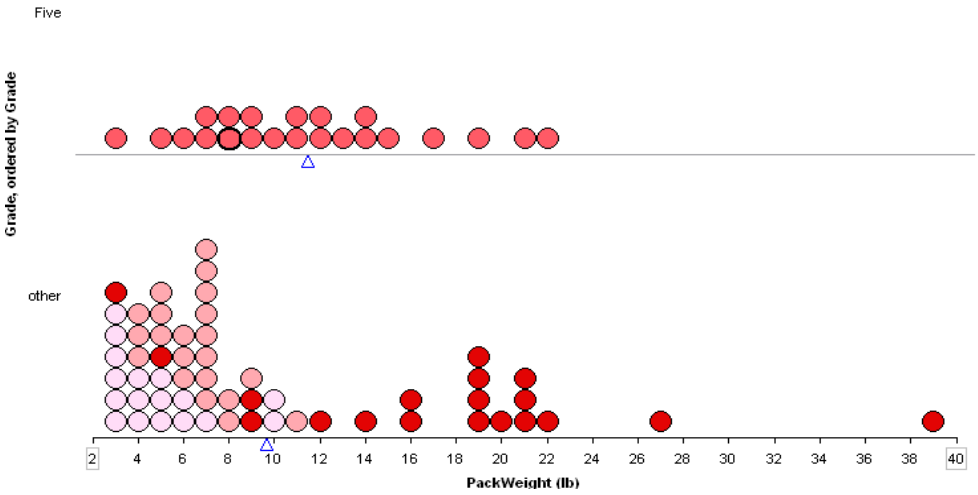


Figura 14: O peso de mochilas em relação a variável série. Fonte: Konold, 2006

O estudo do Konold (2006) sugere que a questão sobre o peso das mochilas oferece um desafio a ser resolvido, nesse caso o *software* tem como função organizar os dados sobre os estudantes em relação ao peso das mochilas e a série. Como os dados são retirados da realidade, eles mobilizam os estudantes a encontrarem uma resposta, Bem-Zvi (2006) também utiliza dados da realidade a serem manipulados no *TinkerPlots.*

Bem-Zvi (2006) foca seu estudo nos processos de inferência informal utilizando o *TinkerPlots* com crianças entre 10 e 12 anos. Ele se utiliza de um

processo de análise de dados que parte de um pequeno banco de dados para um banco de dados maior. Os alunos participantes são convidados a realizar uma coleta de dados com 4 estudantes da escola em que questionam sobre peso, altura, se utiliza computador, distância da residência para escola, entre outros. Após a coleta, os alunos inseriram os dados no *TinkerPlots* e em dupla analisaram os dados coletados (cada dupla analisava os dados de 8 participantes), eles levantavam hipóteses, realizavam inferências e se utilizavam de diversas formas de organização. Após esse primeiro momento, os alunos se uniram em quartetos (analisavam um banco de dados de 16 participantes) e analisavam as questões levantadas e se as tendências permaneciam ou não, essa atividade foi desenvolvida até que os alunos analisaram um banco de dados de 440 participantes. A partir dessa pesquisa foi percebido que a análise dos alunos, sobre os dados, se aproximava da análise realizada por Estatísticos, os alunos foram capazes de construir inferências e argumentações significativas sobre os dados.

A utilização do *TinkerPlots* nessa pesquisa auxiliou imensamente nas formas de organização de dados e nas diferentes formas de representar essas organizações, principalmente, nos cruzamentos de variáveis. O processo de inferência apareceu em vários momentos a partir de suas possibilidades de construção de diferentes formas de representação, com isso, surgiram diversas questões novas a serem analisadas. No entanto, alguns alunos mesmo com essas possibilidades não conseguiram entender os dados de maneira global, de forma crítica, eles permaneciam com o entendimento de casos individuais e pontuais.

Os resultados obtidos por Konold (2006) e Bem-Zvi (2006), com crianças, indicam que o software teve uma boa aceitação e ofereceu instrumentos para que elas pudessem desenvolver um raciocínio sobre os dados analisados, no entanto nas duas experiências os participantes já utilizavam computadores e já manipularam o TinkerPlots anteriormente. As argumentações produzidas por eles pode ser fruto de experiências anteriores com o *software*. Será que a construção dessas argumentações e as analises realizadas pelos participantes aconteceriam se manipulassem o *software* pela primeira vez?

Além de estudos com crianças outros pesquisadores desenvolveram alguns estudos com professores utilizando o *TinkerPlots.* Por exemplo, Rubim e Hammerman (2006) também analisaram os processos de inferência, utilizando os professores de Ensino Médio nos Estados Unidos. Os professores conheceram o

software, anteriormente, para aprender a utilizar suas ferramentas e, somente em seguida, participaram das atividades de análise de dados a partir de um estudo de caso. A atividade proposta pelos pesquisadores foi analisar um caso de um roubo a uma fábrica de escovas de cogumelos, segue a atividade abaixo:

Eles apresentam uma fábrica de escovas que mantinha uma média de produção em torno de 215 escovas em cada 2 minutos. Os autores afirmam que a fábrica foi assaltada, sendo que o ladrão desligou a eletricidade da máquina de produção, o que resultou na queda de produção. O suspeito principal do roubo tinha um álibi no período da meia-noite as três horas da manhã. Os professores tinham que analisar a produção da fábrica no período de toda a noite, através de um banco de dados no *TinkerPlots*, e descobrir qual o horário que a máquina foi interrompida para prender o suspeito ou inocentá-lo.

Os professores da pesquisa de Rubim e Hammerman (2006) cruzaram as variáveis tempo e média de produção e discutiram as possibilidades a partir da variação da produção em cada hora da noite. Os professores analisaram as médias e não pareceram ter nenhuma dificuldade em analisar a variação entre os intervalos de tempo maior e menor, como afirma os autores. Ao mesmo tempo eles realizaram inferências e discutiram a possibilidade da parada da máquina ter sido às duas horas da manhã (o que provocou diversas contestações entre o grupo e novas possibilidades).

As conclusões desse estudo apontam para dois aspectos importantes, primeiro que a possibilidade de representar os dados ora em escala de 2h em 2h, ora de hora em hora ou em escala de 30min e 30min durante o processo de interpretação possibilitou testar hipóteses e interagir mais com as informações. Também foi percebido que a possibilidade de utilizar linhas de demarcação para selecionar variações e médias possibilitou um avanço nas análises de variabilidade.

Em segundo lugar que a utilização de uma situação hipotética, mas que poderia vir a acontecer e que já apresentava variação de dados de forma implícita, poderia ser apontada como fator a potencializar a reflexão dos professores sobre como analisar essas variações observando acréscimo, decréscimo e estabilidade. Por exemplo, a atividade realizada pelas professoras já trazia em seu contexto um decréscimo (queda de energia) que levava as professoras a inferir que a queda na produção ocorreu ao mesmo tempo em que a energia foi interrompida.

A partir das pesquisas citadas pode-se observar as possibilidades trazidas pelo uso do *TinkerPlots* na interpretação e construção de gráficos mesmo com indivíduos com pouco conhecimento em Estatística, como citam os autores. Junto a professores podemos observar o uso do *TinkerPlots* como uma ferramenta que auxilia na formação continuada.

Em termos gerais, observamos que a grande maioria dos estudos utiliza dados da realidade ou de uma realidade hipotética e se utiliza de três eixos principais: o primeiro uma questão instigante a ser resolvida; o segundo tem-se um banco de dados que fornece pistas para resolução da questão proposta e o terceiro os participantes manipulam os dados no *software* emitindo opiniões na tentativa de responder a questão.

A partir dessas experiências, construímos este estudo baseado nos três eixos, mas com participantes que apresentavam pouca experiência com computadores e que conheceram o *software TinkerPlots* no momento da pesquisa. Com isso, buscamos contribuir com as investigações em Estatística entendendo o ambiente do *TinkerPlots* e quais possibilidades o uso de suas ferramentas oferecem para interpretação de dados por professores de regiões rurais. Mais especificamente, trataremos do uso do *software* em questões que utilizam o contexto bivariado, conceito discutido em Estatística que se refere ao cruzamento de duas ou mais variáveis, dentro de uma perspectiva da Análise Exploratória de dados, conceito melhor discutido no capítulo seguinte.

CAPÍTULO 3 – A UTILIZAÇÃO DO CONTEXTO BIVARIADO EM UMA PROPOSTA DE ANÁLISE EXPLORATÓRIA DE DADOS

A Estatística é um componente curricular inserido nos documentos oficiais nos últimos dez anos, por isso é recente e vem causando questionamentos entre os professores ao lidar com essa área do saber sala de aula. Alguns autores como Monteiro e Selva (2001); Asseker e Monteiro (2008); Lima, Lima, Rodrigues e Feitoza (2005) apontam que os professores participantes de suas pesquisas apresentaram dificuldades em interpretar e identificar dados em gráficos, como também, no estudo de Asseker e Monteiro (2008) os professores relataram dificuldades em lidar com conceitos Estatísticos.

Por essas dificuldades relatadas pelos professores o ensino da Estatística nos anos iniciais de escolarização se torna quase ausente. Essa quase ausência acaba por dificultar o desenvolvimento de habilidades necessárias à interpretação de dados, como apresentam alguns estudos (SELVA, 2003; GUIMARÃES, 2002 entre outros). Essas dificuldades estão atreladas aos processos de interpretações de gráficos, principalmente, no que se refere à questões globais (de tendência do gráfico), de interpolação (entre os dados do gráfico) e, também, de questões que se relacionam a conhecimentos de escala, média, entre outros.

Nesse capítulo propomos a perspectiva da Análise Exploratória de Dados como uma possibilidade de aproximação entre o sujeito, no caso desta pesquisa os professores, e o banco de dados a ser manipulado na tentativa de extrair informações percebidas por eles nesses dados. Essa aproximação é uma tentativa de oferecer maior abertura durante a utilização das ferramentas propostas pelo *software* utilizado, para que, possamos analisar que possibilidades o software pode oferecer no momento da interpretação e que possam estar contribuindo para o desenvolvimento do raciocínio para análise de dados

3.1. Analise Exploratória de Dados

A Análise Exploratória de Dados é uma perspectiva defendida por Batanero (1991), entre outros autores, que utiliza diferentes formas de organizar e representar dados para propiciar o desenvolvimento de conhecimentos e novas idéias sobre esse grupo de dados. Godino (1995) aponta, ainda, que a análise exploratória de dados não se configura como um conjunto de métodos, mas sim numa filosofia que consiste em interpretar os dados a partir de diversas perspectivas, sendo que o sujeito levantaria hipóteses e extrairia informações sobre diferentes construções gráficas para apoiar as suas respostas.

De acordo com Batanero, Estepa e Godino (1991) existem especificidades no trabalho com a análise exploratória de dados, tais como: contextualização das atividades; um forte apelo às representações gráficas contribuindo para a percepção da variabilidade nos dados; manuseio de diferentes escalas e diferentes categorizações, como também, a tentativa de evitar ferramentas matemática complexas. A interpretação dos dados se dá a partir da observação das diferentes formas de organizar os dados e das evidências utilizadas pelos sujeitos para justificar suas afirmativas.

Como afirmam Shaughnessy, Garfield e Greer (1996) “trabalhar na Análise Exploratória de Dados é um estado de espírito, um ambiente onde se pode explorar dados e não só um determinado conteúdo estatístico” (p. 205). Nessa perspectiva, o enfoque maior não está nas habilidades de construção de gráficos, mas sim na construção do raciocínio estatístico, visto que, o importante aqui é refletir sobre um grupo de dados. Para isso, o uso de *softwares* para organização de dados configura-se como ferramenta importante nesse processo. Batanero, Estepa e Godino (1991) afirmam que o computador pode facilitar a utilização e mudança de diferentes representações o que faria com que o aluno se envolvesse com maior ênfase na comparação dos dados entre as representações do que na construção de apenas uma representação.

Godino (1995) traz em seu estudo algumas possibilidades de atividades que poderiam ser utilizadas para a interpretação de dados em um *software* a partir dessa perspectiva. Ele cita atividades em contexto univariado e bivariado que permitem observar o comportamento de uma variável em um grupo de dados comparando-os. A partir do ensino de Tratamento da Informação distante da perspectiva da análise

exploratória de dados há um enfoque na resolução de problemas com o uso de dados oferecidos pelo professor. Na perspectiva da analise exploratória de dados Godino (1995) e Biehler (1997) ressaltam que os estudantes podem obter ou completar seus dados a partir de pesquisas de campo ou buscas via internet. Os arquivos contidos na internet, anuários estatísticos, mídia ou dados coletados pelos alunos (por levantamento de medição ou observação) pode ser a base para o levantamento de questões que motivam os alunos e os tornam interessados em aprender.

Nessa perspectiva, além das representações gráficas tradicionais, como gráficos de barras, de colunas, de linhas etc. Surge a necessidade da construção de diferentes organizações de dados em forma de representações que proporcionam, também, a comparação entre as formas de representar um mesmo banco de dados. Nesse caso, o interesse do estudo não incide apenas nas tendências das variáveis, mas também na forma de representar cada informação.

Com a utilização do contexto bivariado a questão a ser investigada, através da Análise Exploratória de Dados, se torna um desafio, no qual, o individuo poderá levantar hipóteses e testá-las mediante a manipulação de dados.

Nos exemplos citados no capítulo anterior, de estudos como o de Kolnold (2006) que propõe um desafio em que os estudantes tiveram que cruzar as variáveis: peso da mochila e série dos alunos temos uma questão dentro de um contexto bivariado, no qual os estudantes se motivaram a resolver essa questão. Outro exemplo de contexto bivariado foi o estudo realizado por Bem-Zvi (2006) no qual os estudantes construíram um banco de dados com variáveis como peso, idade, altura, utilização de computador etc. Estes dados puderam oferecer questões como: os mais pesados são os mais altos? ou quem utiliza mais o computador são os mais velhos ou os mais novos? Todas essas questões são frutos de um cruzamento de variáveis e estão dentro de um contexto bivariado, o que vamos discutir a seguir.

3.2. Contexto Bivariado

Contexto bivariado, Cruzamento de variáveis ou Raciocínio covariacional são denominações utilizadas por diferentes autores como Bisquerra, Sarriera e Martinez (2004) para a teoria que estuda a intensidade da relação ou dependência entre as duas variáveis de uma distribuição bidimensional. Ou seja, identifica o quanto duas variáveis estão relacionadas, se a relação entre elas é fraca ou forte (NOVAES; COUTINHO, 2009).

Para entender o cruzamento entre variáveis é necessário discutir o que vem a ser variável. Existem diferentes formas de conceituar Variável, por exemplo, como Bisquerra, Sarriera e Martinez (2004) que afirmam ser uma “característica”.

> Uma variável é uma característica que pode adotar diferentes valores. Por exemplo, o peso, a idade, a inteligência etc. Opõe-se ao conceito de constante, que se refere àquilo que só pode ter um valor. (BISQUERRA; SARRIERA; MARTINEZ, 2004. P. 20)

Já Novaes e Coutinho (2009) explicam o conceito de variável tomando como base um exemplo de uma pesquisa sobre o estado civil de turistas que escolhem um determinado pacote aéreo. O estado civil de cada um dos turistas pesquisados independe dos demais, e pode variar pra cada um dos pesquisados segundo as possibilidades existentes (casado, solteiro, divorciado, viúvo, solteiro com união estável). Nesse caso, todas as respostas possíveis pertencem a um grupo que pode ser pré-identificado, então, o estado civil é uma variável, pois é uma característica que se quer investigar e que pode variar de um para outro indivíduo dentro de um campo de possibilidades, por ser uma característica do ser também é denominada por alguns autores como atributo. As variáveis são classificadas em qualitativa e quantitativa, Bussab e Morettin (2002) definem essa classificação como:

> Algumas variáveis, como sexo, educação, estado civil, apresentam como possíveis realizações uma qualidade do indivíduo pesquisado, ao passo que outras, como número de filhos, salário, idade, apresentam como possíveis realizações números resultantes de uma contagem ou mensuração. (Morenttin; Bussab, 2002, p. 10)

Assim como a definição desses autores, Novaes e Coutinho (2009) reafirmam essa definição ampliando a classificação. Então para elas variável qualitativa é um certo tipo de característica do grupo pesquisado, que não pode ser mensurado

através de um sistema numérico de medidas, como exemplo: gênero, estado civil, cor de olhos etc. Por sua vez a variável qualitativa pode ser subdividida em ordinal (quando estabelecemos uma ordem. Ex: formato de camisas produzidas em uma fábrica – pequeno, médio e grande) e nominal (quando não é possível estabelecer uma ordem. Ex: cor de olhos, gênero).

A variável quantitativa é aquela que pode ser mensurada como, por exemplo, idade, altura, quantidade de participantes. Ela pode ser discreta, quando entre dois valores consecutivos não podemos inserir nenhum outro valor (ex: número de filhos – contagem de 1 em 1; contagem de números de gols – de 1 em 1). Pode ser contínua quando seus valores podem assumir qualquer valor dentro de um intervalo real (ex: altura; peso).

Os tipos de variáveis podem ser cruzados ou relacionados em um levantamento de hipótese sobre o grupo pesquisado, essa relação é conhecida como Cruzamento ou relação de variáveis, o contexto bivariado, como definido anteriormente. Dentro do contexto bivariado podemos encontrar três situações, de acordo com Bussab e Morettin (2002):

- As duas variáveis são qualitativas;
- As duas variáveis são quantitativas;
- Uma variável é qualitativa e outra é quantitativa.

De acordo com esses autores a análise da relação nas três situações descritas é diferente. Quando as varáveis são qualitativas, os dados são resumidos em tabelas de dupla entrada ou gráficos de barras múltiplas, onde aparecerão as frequências absolutas ou contagem de indivíduos que pertencem simultaneamente a categorias de uma e da outra variável. Como exemplo eles apresentam uma tabela que associa gênero e graduação com um banco de dados hipotéticos, como demonstra a Figura 15.

Tabela 4.5: Distribuição conjunta de alunos segundo o sexo (X) e o curso escolhido (Y).

Y \ X	Masculino	Feminino	Total
Economia	85	35	120
Administração	55	25	80
Total	140	60	200

Fonte: Dados hipotéticos.

Figura 15: Tabela para exemplificar o cruzamento entre duas variáveis qualitativas.
(BUSSAB; MORETTIN, 2002, p. 75)

A Figura 15 apresenta uma tabela de dupla entrada onde encontramos a frequência associada ao gênero e curso de graduação escolhido. Nesse caso temos 85 estudantes homens que escolheram o curso de Economia.

Já quando as variáveis são quantitativas as observações são provenientes de mensuração e técnicas como gráficos de dispersão. Um gráfico de dispersão é representado através de um eixo cartesiano (x e y), no qual cada eixo apresenta a escala relativa a variável que está sendo analisada. Como exemplo, apresentamos um gráfico de dispersão utilizado pelos autores que cruza a variável número de clientes e anos de serviço (Figura 16)

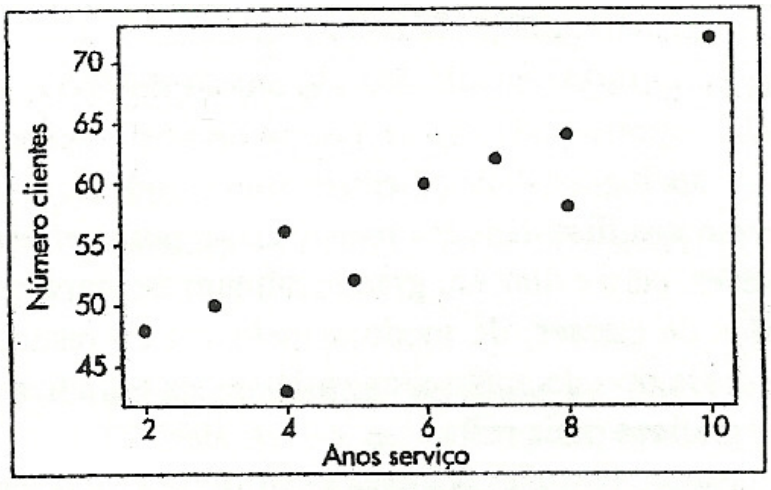


Figura 16: Gráfico de dispersão utilizado para exemplificar o cruzamento de variáveis quantitativas
(BUSSAB; MORETTIN, 2002, p. 82)

Como os autores afirmam, a partir desse gráfico podemos perceber que existe um tipo de relação entre as variáveis. “Para o exemplo, vemos que parece haver uma associação entre as variáveis, por que no conjunto, à medida que

aumenta o tempo de serviço, aumenta o número de clientes" (MORETTIN; BUSSAB, 2002, p. 82).

Eles demonstram três tipos de cruzamentos de variáveis que podemos encontrar nos gráficos de dispersão, como mostra Figura 17.

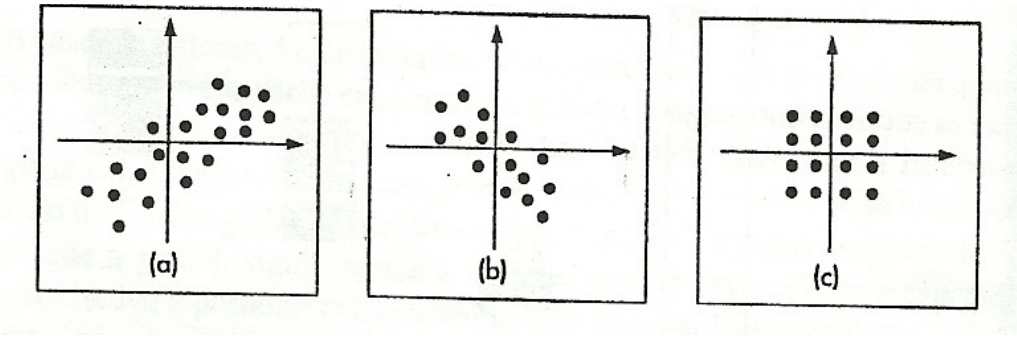


Figura 17: Exemplos dos tipos de relação entre duas variáveis quantitativas.(BUSSAB; MORETTIN, 2002, p. 82)

A Figura 17 apresenta três exemplos de tendências de cruzamentos de variáveis que podem acontecer em um gráfico de dispersão. No exemplo (a) temos um cruzamento que possui uma associação linear direta ou positiva, os pontos se concentram no primeiro e terceiro quadrante. No exemplo (b) observamos uma dependência linear inversa ou negativa, os pontos se concentram no segundo e quarto quadrante. No exemplo (c) não há associação entre as variáveis, pois os pontos se concentram no centro não apresentando nenhuma tendência evidente.

Nos cruzamentos entre uma variável quantitativa e outra qualitativa é comum analisar o comportamento da variável quantitativa dentro de cada categoria da variável qualitativa. Essa análise pode ser realizada através de histogramas, gráficos de barras, *Box plot*, entre outros.

Questões que utilizam o contexto bivariado tendem a requerer do sujeito uma reflexão sobre as tendências das variáveis, propondo, ainda, uma maior observação das informações gerais dos dados não apenas pontos específicos. O desafio proposto por uma questão dentro de um contexto bivariado oferece ao sujeito diferentes possibilidades de resolução, já que partem de hipóteses defendidas pelos sujeitos e que necessitam de evidências para sustentar suas argumentações. Esse processo de análise de dados requer a observação da tendência da variável e sua relação com uma outra variável o que não se configura atividade fácil. Para isso, o sujeito precisa interagir com os dados, tendo a possibilidade de organizá-los de

diferentes maneiras objetivando a busca por todo tipo de informação que possa contribuir com a resolução da questão.

A proposta da utilização do contexto bivariado em uma perspectiva da Análise Exploratória de dados é uma aproximação entre o sujeito e a Estatística. Os bancos de dados utilizados nesta pesquisa foram retirados de dados reais e as questões utilizadas foram elaboradas a partir de um contexto bivariado o que acreditamos contribuir para uma interpretação mais livre, com mais levantamento de hipóteses e com mais de uma resposta correta. O nosso interesse é com o processo de reflexão sobre os dados.

Para este estudo organizamos as questões envolvidas nas duas atividades de análise de dados a partir dos tipos diferentes de cruzamento em variáveis apontados por Bussab e Morettin (2002). Também utilizamos a nomenclatura e os conceitos para variáveis defendidos por Novaes e Coutinho (2009) como variável qualitativa e variável quantitativa.

A utilização do software TinkerPlots para análise de dados está apoiada na Filosofia da Análise Exploratória de Dados em um contexto bivariado, no qual os professores tentam responder as questões a partir da manipulação dos dados e argumentar sobre suas afirmações.

CAPÍTULO 4 – MÉTODO E ESTUDO PILOTO

Neste capítulo explicitamos as escolhas metodológicas da pesquisa. Como ponto de partida, iniciamos relembrando que o objetivo deste estudo foi investigar o uso do *TinkerPlots* para exploração de dados por professores de escolas rurais. Para responder a esse objetivo buscamos identificar as representações construídas pelos professores a partir da utilização das ferramentas do *software;* analisar o uso das ferramentas *separar*, *empilhar*, *ordenar* e *gradiente* em atividades de cruzamento de variáveis; como também, analisar que influência o uso dessas ferramentas disponíveis no software proporciona no momento da interpretação dos dados em cada questão. Então, apresentamos na seção seguinte o contexto da pesquisa e os sujeitos envolvidos e os resultados do estudo piloto que nortearam as questões do estudo principal.

4.1. Contexto da Pesquisa

A pesquisa foi realizada com professores de escolas rurais do Município de Caruaru situado no Agreste de Pernambuco nas quais funcionam 121 salas de aula em regime seriado, bisseriado e multisseriado. O regime seriado é caracterizado pela existência de alunos de uma mesma série em cada sala de aula com um professor; o regime bisseriado é a existência de alunos de duas séries em uma mesma sala de aula com um professor e o regime multisseriado é a existência de alunos de todas as séries dos primeiros anos de escolarização juntos em uma sala de aula com um professor apenas. A rede de ensino municipal de Caruaru possui, de acordo com o educacenso (BRASIL, 2009), cerca de 9 mil alunos matriculados no Ensino Fundamental.

As escolas situadas nas áreas rurais desse município são classificadas em dois tipos, as independentes e as nucleadas. De acordo com Farias (2010) as escolas Independentes possuem mais de 100 alunos matriculados, nesse caso elas possuem salas de aulas seriadas, uma gestora, uma supervisora e uma merendeira. Já as escolas nucleadas são as que possuem menos de 100 alunos matriculados,

nesse caso funcionam em regime bisseriado e multisseriado. Por serem consideradas de pequeno porte elas acabam funcionando em um regime de nucleação, na qual, uma gestora e uma supervisora ficam responsáveis pelo acompanhamento pedagógico de até 10 escolas, realizando o acompanhamento pedagógico através de um carro disponibilizado pela Secretaria de Educação.

Em relação à infra-estrutura, as escolas são bastante simples, funcionando com uma sala de aula ou duas no máximo, uma cozinha onde são feitas as merendas, um banheiro e uma área livre onde os alunos brincam nos intervalos. A secretaria e a diretoria funcionam na própria sala de aula.

Fez parte deste estudo 8 escolas de uma mesma nucleação por disponibilidade da gestora e do acesso a essas escolas. Foram 4 escolas que funcionavam com uma sala de aula cada em regime multisseriado nos horários da manhã e tarde e 4 escolas que funcionavam com 2 salas de aulas cada em regime bisserado, também nos dois horários manhã e tarde. Essa nucleação é considerada de difícil acesso, com muitos trechos sem estrada asfaltada (Figura 18). O transporte da gestora e supervisora era realizado através de Toyotas Bandeirantes um tipo de jipe que chegou ao Brasil nos anos 50 e que tem como características principais a potência do motor, durabilidade do veículo e possibilidade de travessia em lugares de difícil acesso como mostra a Figura 18.

Figura 18: Acesso as escolas rurais e meio de transporte utilizado

Nosso acesso as escolas foi realizado junto com a gestora e a supervisora pelo transporte fornecido pela Secretaria Municipal de Educação que apoiou nosso estudo desde o início. A Figura 19 apresenta a foto de uma das escolas onde funciona uma sala de aula em regime bisseriado no qual a professora desta sala foi

participante da pesquisa. Vamos utilizar nomes fictícios para nomear as escolas e resguardar o sigilo e anonimato aos participantes.

Figura 19: Foto da Escola Cantinho Feliz situada na região rural

Além dos aspectos estruturais específicos das escolas rurais, também foi observado em todas as escolas participantes do estudo a presença da horta e de um sistema de coleta seletiva produzidos pelos alunos como mostra a Figura 20. Essas informações delineiam os aspectos específicos dessas escolas e permitem que possamos conhecer melhor essa realidade.

Figura 20: Horta e coleta seletiva presentes nas escolas participantes

A merenda das escolas era produzida utilizando alguns legumes e vegetais produzidos na horta e os próprios alunos eram responsáveis por cuidarem do plantio, de regar e de colher o que era produzido. Além da horta eles eram responsáveis por separarem o lixo através da coleta seletiva, como também coletarem em casa o lixo reciclável e encaminharem à coleta seletiva da escola.

A região onde as escolas estão situadas apresenta em seu entorno muitas paisagens naturais, uma efetiva atividade agrícola (plantação de tomate, milho, macaxeira, feijão etc.) e de criação de animais (porcos, bois, galinhas etc.). Como meio de sobrevivência, a população também trabalha no comércio local e nas proximidades do centro da cidade, em fábricas e empresas na área urbana do município, isso devido à implementação de uma linha de ônibus (2 viagens por dia do centro urbano à comunidade rural – uma pela manhã, outra no final da tarde). Todos os serviços básicos de saúde, previdência, educação dos anos finais de escolarização, dentre outros ficam localizados no centro do município – área urbana.

Essas informações são importantes para entendermos a nova dinâmica que se constitui nas áreas rurais, a população tem acesso a centros urbanos com mais facilidade, o sistema de informação está presente através do sinal de televisão e rádios mantendo alguns aspectos agrícolas (pequenas plantações) e de criação de animais (em pequena quantidade) e os vínculos estreitos de afetividade e culturais.

4.2. Os sujeitos da Pesquisa

Os sujeitos da pesquisa foram 12 professores que lecionavam nas 8 escolas da nucleação Nossa Senhora Aparecida. Apresentamos a seguir o perfil desses professores que foi identificado a partir de uma entrevista estruturada (Anexo 1) realizada na primeira sessão da pesquisa.

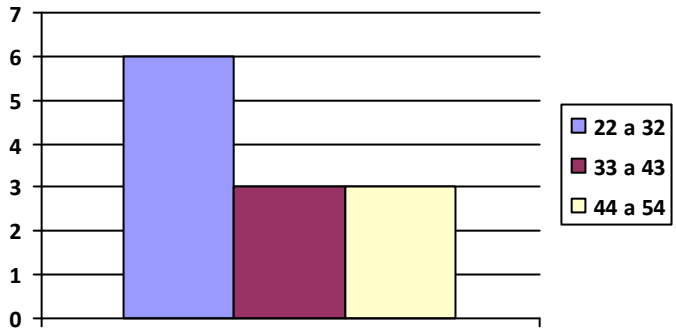


Gráfico 1: Idade dos Professores participantes da pesquisa

Como mostra o gráfico 1 a grande maioria dos professores possui de 22 a 43 anos de idade, sendo que quase a metade dos professores tem até 32 anos.

No que se refere a formação profissional dos professores percebemos que 5 dos 12 participantes estão cursando Pedagogia e 4 já cursaram Pedagogia, sendo que 2 desses já estão cursando Pós-graduação a nível de especialização. Apenas 2 professores cursam apenas o magistério. Percebemos que todos os professores possuem formação mínima exigida para lecionarem, como demonstra o Gráfico 2.

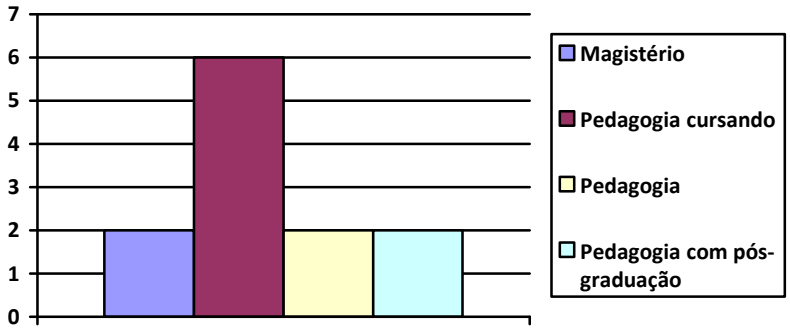


Gráfico 2: Formação profissional dos participantes da pesquisa

No Gráfico 3 podemos perceber que a maioria dos professores tem mais de 8 anos de experiência de ensino em salas de aulas de escolas rurais.

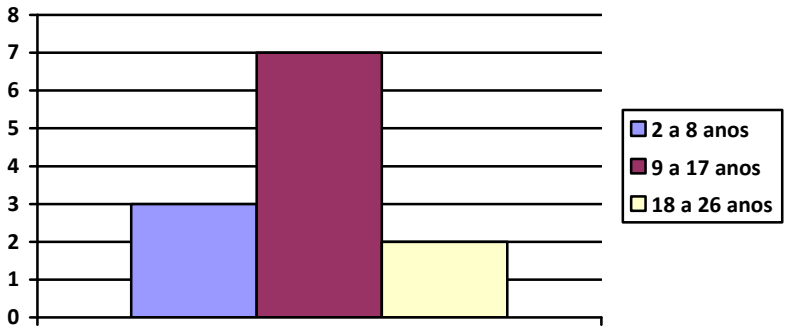


Gráfico 3: Experiência de ensino em salas de aulas de escolas rurais

Podemos refletir que os professores participantes da pesquisa, de uma forma geral, possuem o mínimo de formação para atuarem em salas de aulas e que quase a metade deles ainda está em processo de formação (os professores que estão cursando Pedagogia e Especialização) e que já possuem experiências anteriores com o ensino em áreas rurais.

4.3. Escolhas Metodológicas

Para o alcance dos nossos objetivos de investigação do uso do software pelos professores de escolas rurais realizamos entrevistas semi-estruturadas baseadas em atividades que foram respondidas por meio da exploração de dados no *TinkerPlots*. Através da sua utilização pode-se obter informações que vão além de dados objetivos, podendo o entrevistador ter acesso as atitudes e valores atrelados às opiniões dos entrevistados. (HAGUETTE, 1997)

Para garantirmos o registro de todos os aspectos que pudessem ser relevantes para análise, videografamos todas as entrevistas, captando as falas e imagens dos professores, assim como, todas as ações que eles realizaram no computador usando o *software* e que apareciam na tela do computador. Essa videografia foi possível através do *software Camtasia Studio 4* que permitiu a captura da tela do computador através de vídeo de alta resolução.

As entrevistas foram divididas em 3 etapas, como mostra o Quadro 1, e foram realizadas em dois dias diferentes, com um intervalo de uma semana.

Quadro 1: Etapas propostas para a realização da Entrevista semi-estruturada

	Etapas	Encaminhamento
1º ENCONTRO	**1ª. Etapa – Perfil**	Questões que versavam sobre o perfil dos professores, suas concepções e praticas do uso do computador e da Estatística.
	2ª. Etapa – Familiarização	Primeiros contatos com o *TinkerPlots* a partir de um banco de dados simples.
2º ENCONTRO	**3ª. Etapa – Atividades**	Realização de duas atividades de exploração de bancos de dados diferentes. Para cada banco foram propostas três questões.

A estrutura proposta no estudo principal teve como base os dados obtidos no estudo piloto realizado com duas professoras de escolas rurais dessa mesma região. Iremos apresentar, a seguir, o estudo piloto e alguns resultados encontrados a partir da manipulação do *software* pelas professoras. Os resultados encontrados indicaram o caminho a ser percorrido no estudo principal, o contexto dos bancos de dados utilizados e a definição dos dois encontros de atividades.

4.4. Estudo Piloto

Foram entrevistadas duas professoras pertencentes as escolas da mesma região rural onde iria ser realizado o estudo principal. No estudo piloto planejamos a realização de duas sessões para cada participante. No primeiro momento foi realizada a entrevista semi-estruturada para conhecermos o perfil das professoras a partir do roteiro do anexo 1. O momento da familiarização aconteceu logo após a entrevista inicial com o banco de dados dos gatos que apresentamos no capítulo do método (Figura 21).

A atividade proposta no segundo momento foi baseada em um banco de dados que supostamente acreditávamos ser de interesse das professoras, Gasto de Água. Nesse banco de dados estavam dispostas as variáveis tempo (a medida em minutos de água); tipo de atividade (em que atividade se utilizava água) e litros de água (a quantidade em litros de água gasta nas atividades). Esse banco de dados

foi extraído da Fundação Nacional da Saúde (2009) e as participantes manipularam os dados a partir dos *plots* para responder as seguintes questões:

1. Se olharmos esse banco de dados que informações poderemos obter?
2. E se cruzarmos as variáveis tempo com gasto de água podemos observar alguma tendência no gráfico?

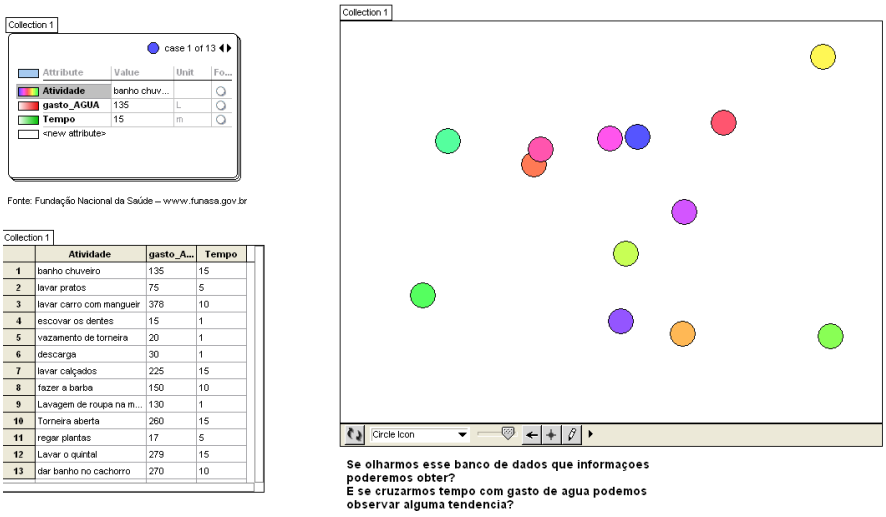


Figura 21: Banco de dados de gasto de água utilizado no estudo piloto

As duas sessões foram videografadas e realizadas em dois dias com intervalo de oito dias entre cada encontro. A seguir relataremos alguns resultados obtidos através do estudo piloto.

a) Analisando a entrevista inicial

Toda videografia foi transcrita em protocolos para posterior análise das falas, das ações e realizadas no *software* e das representações construídas por eles. Inicialmente a entrevista buscava conhecer as duas professoras, então utilizaremos nomes fictícios para resguardar o anonimato das participantes. A primeira professora a participar foi Rosa que leciona em uma turma bisseriada (3ª e 4ª séries). Rosa tem nove anos de experiência em sala de aula de escola rural, está cursando pedagogia, possui pouca experiência com computador e atualmente o utiliza com uma frequência de três vezes na semana devido as atividades da faculdade. Em relação a Estatística ela afirma trabalhar em sala de aula e cita, ainda, questões sobre interpretação de gráficos presentes na Provinha Brasil.

A outra professora a participar do estudo foi Margarida que lecionava em uma turma multisseriada (1ª, 2ª, 3ª e 4ª séries). Ela tem 25 anos de experiência em sala de aula de escolas rurais, também está cursando Pedagogia. Em relação a experiência com o uso do computador ela afirma não ter muito acesso, já que ela não possui computador e sempre que precisa pede a sua filha para fazer na *lan-house*. Sobre Estatística ela também afirma trabalhar com gráficos em sala de aula, mas não com muita frequência.

Em seguida apresentaremos trechos das falas das professoras e imagens produzidas por elas no momento da exploração do software na interpretação dos bancos de dados.

b) Analisando a familiarização

As professoras tiveram facilidade em entender a proposta do *software*, bem como em manusear o teclado e o *mouse*, conforme o extrato da entrevista com a professora Rosa:

> Pesquisadora: então você vai clicar no card em atributo e colocar o nome do gato.
> Rosa: aqui eu coloco nome [...] [ela inclui a categoria nome sem dificuldade]
> P: aqui é uma categoria, um atributo que quero do gato, depois colocamos o nome dele. E agora qual o outro atributo que queremos ver dos gatos...
> R: a idade... a idade né?! [ela realiza o comando mais rapidamente e sem intervenção da pesquisadora]
> P: aquilo que você tinha dito que tinha mais gatos mais novos do que mais velhos continua?
> R: tem mais gatos novos [...].

No extrato da professora Rosa podemos perceber que mesmo afirmando na entrevista ter pouco conhecimento e acesso a computadores demonstrou facilidade em compreender o uso das ferramentas e de realizar os comandos sem ajuda e de construir diferentes representações. Essa facilidade também foi percebida com a professora Margarida, que proferiu frases otimistas e de valorização do seu desempenho com o *software*, como é exemplificado a seguir.

> [A professora sorri por ter conseguido realizar o comando rapidamente]
> Margarida: vixe que chique... como eu to chique [sorri muito] eu dividi... é, é. Masculino e feminino [apontando para os grupos]
> Pesquisadora: não foi legal?
> M: foi ótimo.

Essa fala da professora demonstra seu contentamento por estar realizando uma atividade não usual e por ter conseguido desempenhar uma atividade que, por vezes, é considerada difícil pela população que não tem acesso ao computador.

A seguir apresentamos alguns extratos da fala das professoras ao se depararem com representações que utilizam a ferramenta *gradiente* do Tinkerplots, no qual a pesquisadora apresenta a representação do banco de dados dos gatos, levando em conta a variável idade e pergunta as professoras o que elas percebem nos dados. Na Figura 22 podemos observar a representação do atributo idade dos gatos, o tom vermelho claro indica idades menores e o tom vermelho escuro o gatos com idade mais velha.

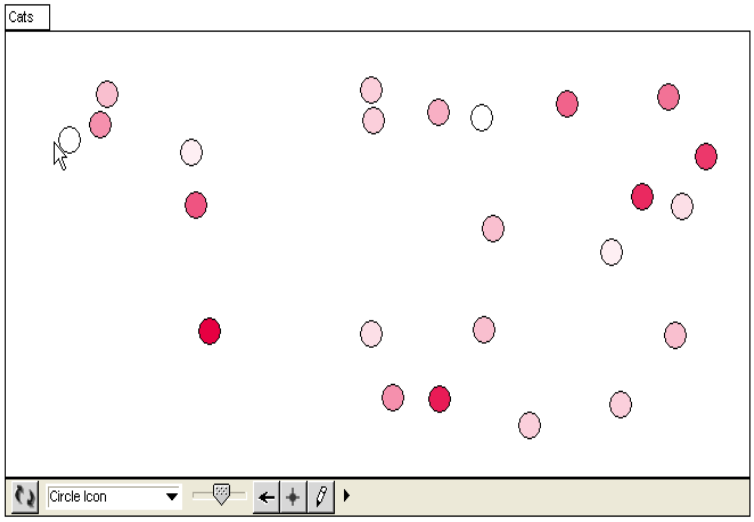


Figura 22: Demonstração da ferramenta *gradiente* no banco de dados de gatos no estudo piloto

Segue extrato das entrevistas das professoras Rosa e Margarida

> Rosa: A cor mais clara são os mais novos e os escuros os mais velhos...
> Pesquisadora: Dá pra dizer se eu tenho a maioria mais velho ou mais novo?
> R: Mais novo! Por que tem mais bolinhas clarinhas do que escuras.
>
> Pesquisadora: Deu pra perceber alguma coisa? [ela mexe a cabeça confirmando] o que você percebeu?
> Margarida: Que no caso... assim... o branquinho é mais novo e o escuro é o mais velho.
> P: Se eu olhar assim já dá pra eu dizer que tem mais gatos novos ou mais gatos mais velhos?
> M: Mais novos.
> P: Por que?
> M: Por que os tons tão mais claros do que escuros

Essas falas das professoras indicam que a ferramenta *gradiente* favoreceu a discriminação entre os gatos novos e mais velhos, como também a quantidade

maior entre gatos novos e velhos. Dessa maneira podemos perceber que a representação gráfica gerada a partir do *software* pode enfatizar tanto aspectos pontuais dos dados, como também aspectos globais.

Ressaltamos que essas são hipóteses a serem exploradas no estudo principal. No entanto, os resultados obtidos podem ser considerados indícios da possibilidade de percepção das informações globais, como pontuais no banco de dados.

Em um momento posterior da entrevista utilizamos a associação das ferramentas separar e *gradiente*. A função separar permite que os dados sejam organizados em categorias de acordo com a necessidade do usuário, assim, cada vez que clicamos em um *plot* e arrastamos os agrupamentos vão se estabelecendo. A Figura 23 mostra as idades dos gatos agrupadas em intervalos.

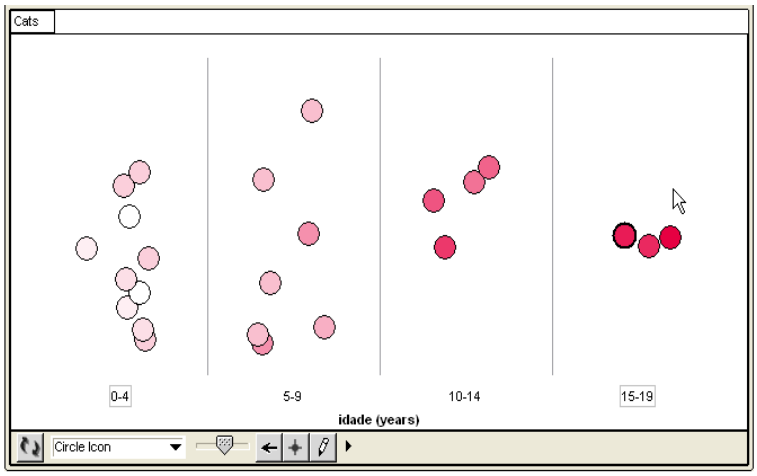


Figura 23: Demonstração da ferramenta *separar* no banco de dados do piloto

As professoras interpretaram essa imagem conforme extrato a seguir:

> Pesquisadora: E agora o que aconteceu?
> Rosa: Aqui veja... tá... aqui... tá separado... assim mais novos dos mais velhos, agora em uma forma diferente aqui... acho que tá pela sequência de idade... digamos quem tem 15 aí tá... ele agrupou...
>
> Pesquisadora: Isso... o que aconteceu aí?
> Margarida: Eles foram divididos. Os mais novos... [aponta para os plots brancos] e os... os.. que vão chegando aos mais velhos. E aqui praticamente o gráfico já está formado. Não é? Quantos têm de dois anos... vamos dizer assim... quantos têm de três anos, de seis anos. Mais ou menos assim.

Nesses extratos observa-se que as professoras mudam as representações que relatam a nova forma de agrupar e organizar os dados de acordo com intervalos

de idades. A professora Margarida relata que essa forma de representação se assemelha a um gráfico, se referindo ao gráfico de barras e suas categorias. Além das funções sugeridas pela pesquisadora, as professoras construíram outras representações com no intuito de auxiliar na interpretação.

A professora Rosa opta por separar o grupo de gatos em macho e fêmea e construir categorias que representam o peso dos gatos. A Figura 24 indica o cruzamento das variáveis gênero, representada no eixo vertical (macho e fêmea), e a variável peso, representada em duas categorias intervalares no eixo horizontal.

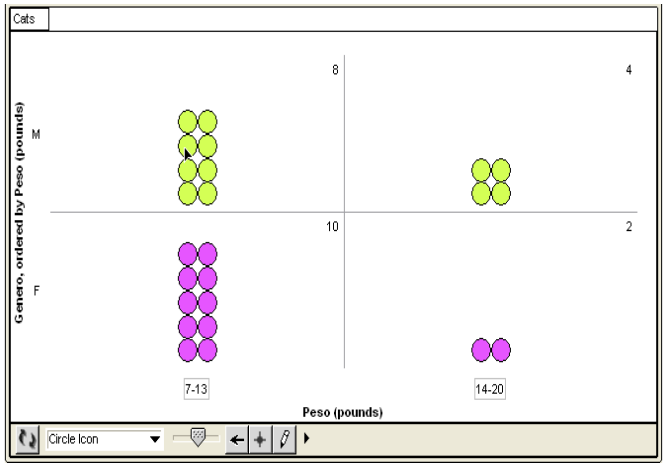


Figura 24: Demonstração de cruzamento de variáveis no Banco de dados gatos no piloto

Segue um extrato da fala da professora Rosa com relação a esse cruzamento de atributos:

> [A professora afirma que os gatos são mais pesados que as gatas e procura outra representação para confirmar sua afirmação]
> Pesquisadora: E ai posso ter certeza mesmo??? Da sua resposta...
> *Rosa*: Acho que a gente ainda pode facilitar mais... agora deixa eu ver como.. pode dividir em dois grupos...
> [A professora clica em um plot e arrasta com o objetivo de dividir os gatos em dois grupos em relação a idade – como mostra figura 18]
> P: E agora? eu tenho ai dois grupos... ai já dá pra afirmar?
> *R*: Agora já da pra afirmar. Ai aqui ... tem... os que ta entre 14 né isso... e entre 20 quilos..
> os que estão entre 14 e 20 quilos e aqui eles estão entre 7 e 13. Eu acho que é fêmea.
> P: Por que?
> *R*: Por que tem mais gatos entre 14 e 20 quilos.

A Rosa precisou mudar a representação para confirmar sua hipótese, que estava relacionada ao cruzamento de duas variáveis, peso e gênero. Na verdade a representação utilizada pela professora requer uma comparação ente grupos, não

se trata de localizar pontos em gráficos, mas de analisar a relação peso-idade o que requer uma maior compreensão da relação dos dados no gráfico. Nesse sentido, a proposta do *software* de mudar de representação conforme necessidade do sujeito auxiliou em uma maior interação com os dados e maior percepção de que representação poderia ser melhor para responder alguma hipótese, como foi o caso da Rosa que afirma “acho que a gente ainda pode facilitar mais...” se referindo utilizar uma representação que possibilitasse uma confirmação da hipótese defendida por ela.

As mudanças na representação e as possibilidades de ferramentas do *software* parecem ter favorecido uma maior reflexão por parte das professoras. Isso indica uma possível contribuição do software com o uso das diferentes possibilidades de representação, nesse sentido a atividade desenvolvida com elas sobre o banco de dados gatos foi interessante para que elas se apropriassem do uso das ferramentas.

c) Analisando a atividade realizada após a familiarização no estudo piloto

Na atividade 2 as professoras demonstraram uma independência ainda maior, por vezes sugerindo representações e manipulando os dados sem intervenção prévia do pesquisador. A professora Margarida, por exemplo, também não encontra dificuldades com o uso das ferramentas na atividade 2, como segue no trecho abaixo a pesquisadora pede que a professora mude a forma de representação, porém não indica a representação que a professora deveria construir.

> Pesquisadora: certo... vamos mudar um pouquinho.
> [Margarida clica em um *plot* e arrasta o *mouse* sem intervenção da pesquisadora]

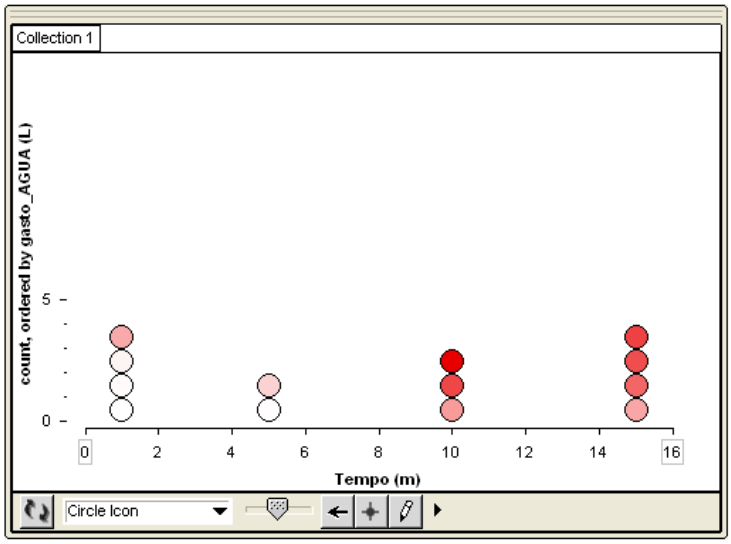


Figura 25: Representação construída pela professora Margarida que cruza tempo e gasto de água

> Margarida: [olhando atentamente] Tempo médio... eu continuo achando que eu gasto muito... muita água
> Pesquisadora: O que é que me indica ai que eu gasto muita água?
> E: A cor mais escura eu acho que tem mais

Margarida realiza o cruzamento das variáveis tempo e gasto de água, no qual tempo está representado na escala de pontos e gasto de água pela cor (quando mais escuro o vermelho maior é o gasto de água). Ela não requer nenhuma ajuda para realização desse comando, também não demonstra dificuldade em entender a imagem construída na tela. Ela relata que gasta muita água baseada nas informações contidas no gráfico quando afirma que esta observando *plots* mais escuros do que claros.

Esse trecho, dentre outros, indicam que as professoras não encontraram dificuldades em manusear o *software*. Como também, demonstraram entender a proposta das diferentes formas de representação dos dados, já que elas ao realizarem as representações emitiam falas de entendimento e interpretação dos dados.

Na representação da Figura 26 temos o cruzamento das variáveis: tempo de torneira aberta e gasto de água indicado pelo dégradé de cores (os *plots* mais claro gasto menor de água os mais escuros maior gasto de água). A pesquisadora perguntou se ela consegue identificar alguma tendência no gráfico.

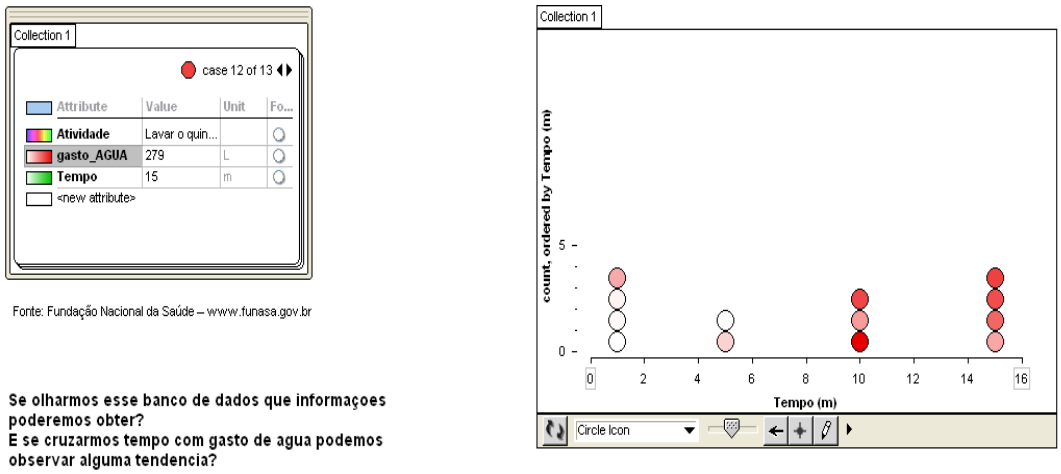


Figura 26: Representação construída pela professora Rosa que cruza tempo e gasto de água

Na Figura 26 observamos o cruzamento da variável tempo de torneira aberta no eixo horizontal com a variável gasto de água que está representado pelas cores (o vermelho claro até o vermelho escuro indica o aumento do gasto de água). Na questão proposta a professora tinha que observar e analisar se existe alguma tendência nessa relação. A resposta esperada seria que quanto mais tempo utilizamos a torneira aberta maior é o gasto de água. No entanto, a ferramenta *gradiente* não contribui para essa reflexão como nas outras atividades.

> Rosa: Minutos. 16 minutos.. foi La... que a gente gasta...ai eu acho que é assim... aqui gastou 1 minutos pra lavar 10 pratos... pra lavar um... 16 minutos para serem lavados... por que essa cor aqui representa que lavou mais, né isso?
> Pesquisadora: É! Lavou mais! Na verdade que gastou mais água... e aqui ta o tempo.
> R: Por que a atividade... aqui é o que? lavar o quintal...
> P: É mas a gente não sabe onde ele ta... não ta clicado em nada.
> R: Por que eu acho que gastou mais tempo lavando prato.
> P: Por que tu diz isso?
> R: Eu penso assim... é que .. pela quantidade de bolinhas.. por que essa bolinha aqui ta representando essa quantidade aqui... a quantidade que lavou mais, a quantidade de prato, né isso! Eu acho que tem que deixar mais... pra deixar mais claro aqui... a divisão.

A professora atribui nome aos *plots* para justificar sua afirmação de que lavar pratos gasta muita água. Ela não conseguiu perceber que a cor esta representando uma variável e a escala outra, ela quer justificar sua experiência de que lavar pratos gasta muita água. Essa foi uma limitação encontrada no uso do *software*, pois mesmo a professora já estando familiarizada com o *software* é muito complexo

entender varáveis quantitativa se relacionando através de cores. Esse dado também foi percebido na mesma atividade com a Margarida que não consegue entender o cruzamento de variáveis quantitativas com a utilização de cores, como segue.

> Pesquisadora: E essa linha de baixo?
> Margarida : Essa linha de baixo é... a... é... a ba... é... a... essa aqui que você ta falando? [aponta para a escala]
> P: Ah ram.
> M: É a base de quan... de.. quanto tempo. Se eu for somar... a base da somatória... eu iria somar no gráfico de quanto tempo eu tenho.... quanto mais eu gasto mais eu economizo.
> P: Então pra gente fechar o que eu posso afirmar daí?
> M: Que eu... eu.. gasto muita água...
> P: Posso dizer mais alguma coisa?
> M: Não se for em relação a água não. Só somar ela pra saber se eu gasto ou não, se há desperdício ou não pra trabalhar em ciências, se for em matemática saber dividir esses dados.

Ela também recorre a experiência do cotidiano para entender os dados do gráfico, pois não consegue compreender esse cruzamento de variáveis quantitativas em relação as cores. As duas professoras apresentaram dificuldade com o uso dessa mesma ferramenta *gradiente* nesse tipo de situação.

d) Possibilidades para o estudo principal

A partir do estudo piloto podemos observar que o primeiro momento, no qual realizamos as etapas entrevista, inicialmente as etapas de coleta do perfil e de familiarização se mostraram como um importante momento de pesquisa. Conseguimos observar que as professoras utilizaram as ferramentas e realizaram os comandos demonstrados pela pesquisadora e foram além, com autonomia buscaram novas representações a partir de diferentes usos das ferramentas disponíveis no *software* demonstrando um nível alto de aceitação e disponibilidade para o trabalho em ambientes computacionais.

Já no segundo momento as professoras demonstraram dificuldade na atividade de cruzamento de variáveis e no entendimento das cores como representação do *gradiente*, no entanto, não podemos afirmar se essa dificuldade é inerente ao uso do *software* ou a natureza da questão de cruzamento de variável (quantitativa x quantitativa). Nesse sentido decidimos mudar o banco de dados e incluir mais uma atividade como demonstraremos no subitem a seguir do método,

como também sistematizar as questões de cruzamento de variáveis utilizando os tipos diferentes de cruzamento nos dois bancos de dados.

4.5. Estudo Principal

A 1ª etapa foi realizada no primeiro contato com os professores, no qual a pesquisadora se apresentava e iniciava uma conversa de maneira informal. Todos os professores estavam cientes e consentiram a realização da pesquisa e na utilização da videografia como parte do processo da entrevista. Após essa primeira conversa, os professores recebiam o roteiro (Anexo 1) da entrevista semi-estruturada para que pudessem acompanhar as questões que iriam ser realizadas. Essa prática trouxe uma maior confiança na relação entre o pesquisador e os professores durante a realização da entrevista fazendo com que os professores se sentissem mais a vontade para falar sobre o que acreditavam. Durante a realização da entrevista outras questões surgiram e foram também respondidas pelos professores.

Na 2ª etapa acontecia a familiarização com o *software*, no qual a pesquisadora apresentava desde a tela inicial do *software* até as ferramentas e representações mais usadas, utilizando um banco de dados sobre gatos e suas características, o mesmo utilizado no estudo piloto.

No momento da familiarização os professores manipulavam o *mouse* e a realização de todos os comandos sugeridos pela pesquisadora, como também eles puderam realizar o comando de forma autônoma conforme iam aprendendo.

Foi realizada uma pergunta a partir desse banco de dados:

1. Quem tende a ser maior os gatos ou as gatas? Porquê?

Essa questão foi parte de uma atividade que teve como função principal o primeiro contato com o software e com um banco de dados disposto em um *software* para análise de dados. As atitudes, os gestos, as falas, o olhar, as reações frente a cada comando foram capturados a partir do vídeo e foram indícios importantes desse momento de iniciação ao *TinkerPlots*.

A 3ª etapa foi realizada em um momento posterior, pois, a partir dos indícios apresentados no estudo piloto, realizado inicialmente, os professores utilizaram em torno de uma hora para cada sessão, então continuamos realizando a atividade em dias diferentes para contribuir para minimizar o cansaço nos professores.

Na 3ª etapa, realizada uma semana após a familiarização, os professores eram questionados sobre as lembranças que tinham do uso do *software* e convidados a manipularem o *TinkerPlots* de maneira autônoma. Foram apresentados dois bancos de dados e realizadas 3 questões a partir de cada banco de dados conforme apresentado no Quadro 2.

Quadro 2: Questões propostas na atividade realizada no 3º momento da coleta

Banco de dados	Tipo de cruzamento de variável	Questões
Casos de Estupros contra crianças	Qualitativa X Quantitativa	1. A partir de que idade os meninos apresentam menor possibilidade de serem estuprados? Por quê?
	Qualitativa X Qualitativa	2. O estupro a meninas é mais frequentemente realizado por qual agressor?Por quê?
	Quantitativa X Quantitativa	3. Existe alguma relação entre a idade do agressor e a idade da vitima? Por quê?
Novelas exibidas no horário nobre da Rede Globo	Quantitativa X Quantitativa	1. Se relacionarmos ano de exibição das Novelas e os índices de audiência o que podemos observar? Por quê?
	Qualitativa X Qualitativa	2. Que autor escreveu mais Novelas com protagonistas formados por casais?
	Qualitativa X Quantitativa	3. A partir do ano de 2000 qual tipo de protagonista foi mais frequente nas Novelas?

As questões referentes aos dois bancos de dados vincularam-se a proposta do contexto bivariado, no qual cruzamos variáveis de diferentes tipos de maneira sistemática. Essa proposta sugere a realização de questões que possam desafiar os participantes, para que eles se motivem na descoberta das respostas.

Como podemos observar no Quadro 2 para os dois banco de dados cruzamos variáveis quantitativas x qualitativas; quantitativas x quantitativas e qualitativas x qualitativas. Com essa sistematização tínhamos como intenção perceber se existia alguma diferença na utilização das ferramentas e na construção de representações conforme natureza do cruzamento de variáveis.

Os bancos de dados foram escolhidos a partir da realidade de vida dos sujeitos pesquisados. No momento do Estudo Piloto os professores relataram a ocorrência de casos de violência sexual contra crianças nessas comunidades, como também do trabalho doméstico exercido por crianças. Esse dado foi importante para pesquisarmos as informações contidas no banco de dados sobre caso de estupros contra crianças, que foi retirado da – DPCA – Delegacia de Proteção a Criança e ao Adolescente (2000). Os casos presentes no banco de dados foram registrados na DPCA e ocorreram em Recife no ano de 2000. Esperou-se que informações polêmicas e próximas a realidade dos pesquisados pudessem contribuir para a exploração de dados, mostra a Figura 27.

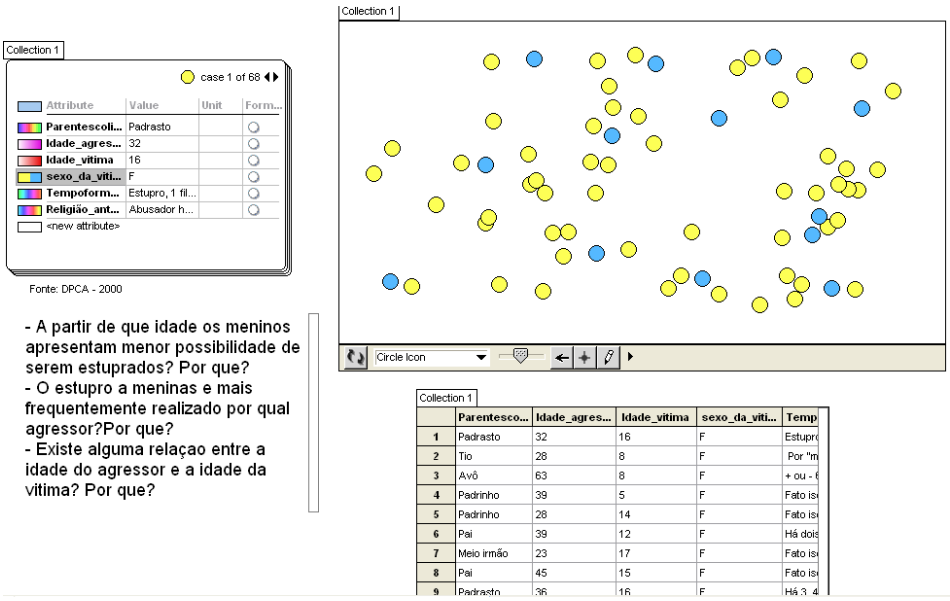


Figura 27: Banco de dados de violência sexual contra crianças utilizado no 3º momento do estudo principal

Já o Banco de dados das Novelas (Figura 28) partiu das diversas conversas que as professoras que participaram do estudo piloto tiveram sobre a novela que estava assistindo. Nesse caso pesquisamos no site da Emissora Globo as Novelas do horário nobre que haviam sido exibidos desde 81 até o ano atual e suas características. Então, os dois bancos de dados foram construídos a partir do pressuposto que estariam ligados a realidade dos sujeitos pesquisados.

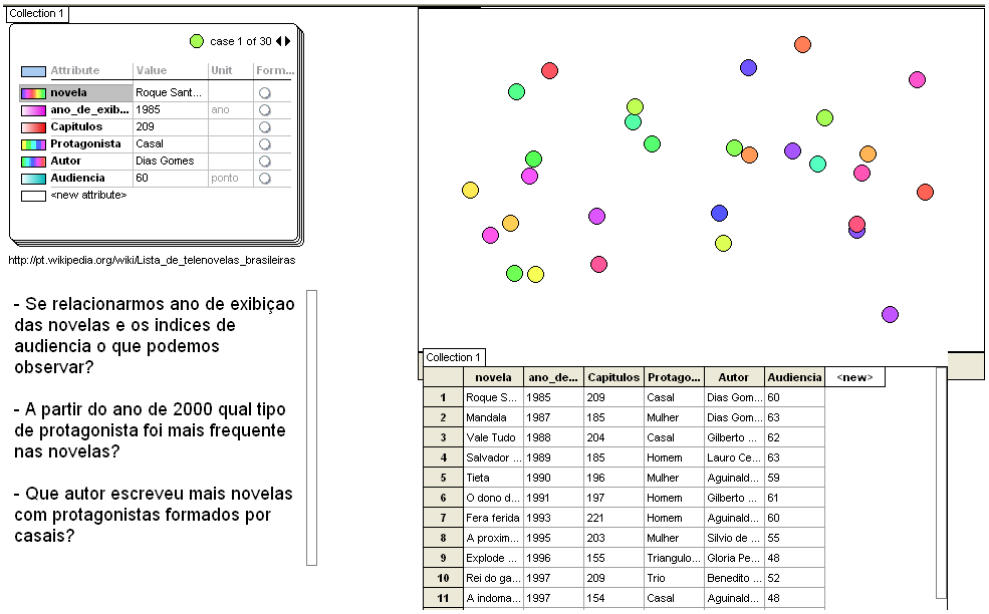


Figura 28: Banco de dados de Novelas exibidas no horário nobre utilizado no 3º momento do estudo principal

Os professores construíram representações de maneira autônoma através da manipulação do software, no entanto, quando o pesquisador intervinha sempre que era requerido ou questionado sobre alguma ferramenta que auxiliasse o professor. Ao término da realização das atividades os professores eram convidados a falar sobre suas impressões a respeito das atividades e do *software*.

4.6. Desenho da análise

Para realizarmos a análise das videografias das entrevistas semi-estruturadas foi necessário de início a transcrição de todos os vídeos. Eram transcritas as falas e os gestos que os professores realizavam durante cada momento da entrevista. Para melhor analisar os dados desta pesquisa eles foram sistematizados em protocolos escritos. Inicialmente, realizamos uma leitura dos protocolos como um todo para em seguida estabelecer passagens significativas que evidenciavam os objetivos desse estudo, sendo assim, vamos discutir os dados de pesquisa a partir de nossas categorias.

Durante a análise focamos na utilização das ferramentas do software e quantificamos de acordo com o seu uso na diferentes questões, como também analisamos a influência do uso dessas ferramentas na interpretação. Outro fator

analisado foram as representações construídas para cada questão de cruzamento de variável e suas especificidades.

A seguir iniciamos as análises apresentando os dados da primeira etapa da entrevista com informações sobre os participantes e o processo de familiarização com o *software*, buscando apontar os pontos relevantes que possam contribuir para a análise da atividade realizada na segunda etapa.

CAPÍTULO 5 – PRIMEIRO ENCONTRO

O primeiro encontro foi dividido em dois momentos, a entrevista que visava o conhecimento do perfil dos professores e em seguida a familiarização o momento em que os professores conheciam o *software*. O primeiro momento da coleta foi a realização da entrevista que teve como objetivo central conhecer os 12 professores envolvidos na pesquisa e seus perfis. Assim, o roteiro das questões objetiva identificar se os professores tinham experiências com o uso de computadores e o que eles falavam sobre o trabalho em sala com o Tratamento da informação.

Em seguida, no segundo momento apresentamos o *software TinkerPlots* e suas ferramentas propondo uma atividade que motivasse os professores a utilizá-lo em uma atividade de interpretação. Desse modo, não tínhamos como intuito apresentar uma sequência de ensino ou propor uma intervenção para a aprendizagem de conceitos Estatísticos. Focamos a experiência do uso do computador e, especificamente, o *software TinkerPlots* e suas ferramentas.

Os encontros foram realizados nas escolas em que cada professor lecionava. Em cada escola procuramos um lugar ameno e distante de qualquer interrupção, tentando promover um momento tranqüilo para que os professores se sentissem à vontade no momento da entrevista e em manipular o computador. Todos os momentos foram videografados através do *software Camtasia Studio 4* que capturava a imagem da tela do computador, a imagem do professor e suas falas. Pudemos observar tanto a dimensão da fala dos professores como também seus gestos no momento da resposta para cada pergunta e as etapas do uso das ferramentas do *software* no momento da manipulação.

5.1. Entrevista semi-estruturada

No primeiro momento do encontro propomos uma entrevista semi-estruturada, como segue roteiro anexo (anexo 1), no qual os professores participantes tiveram acesso ao roteiro durante a realização da entrevista. Acreditamos que acompanhar o roteiro da entrevista ofereceu uma maior segurança aos professores os quais demonstraram uma disponibilidade maior em participar da pesquisa.

As perguntas da entrevista tiveram dois enfoques, o uso do computador e o Tratamento da Informação. Referente ao primeiro foco, o uso do computador, tínhamos como questões o acesso aos computadores, a frequência desse uso e as atividades que eram realizadas. A maioria dos professores afirmou utilizar computadores, dez professores afirmaram ter contato com computares e dois relataram nunca ter utilizado computadores na vida. No entanto, dos dez professores que afirmaram utilizar computadores 5 afirmaram utilizar todo dia e 5 utilizavam esporadicamente. Dentre os motivos citados para o uso dos computadores os professores indicaram: pesquisas em geral, trabalhos da faculdade, correio eletrônico e redes sociais. Conforme é apresentado no Gráfico 4.

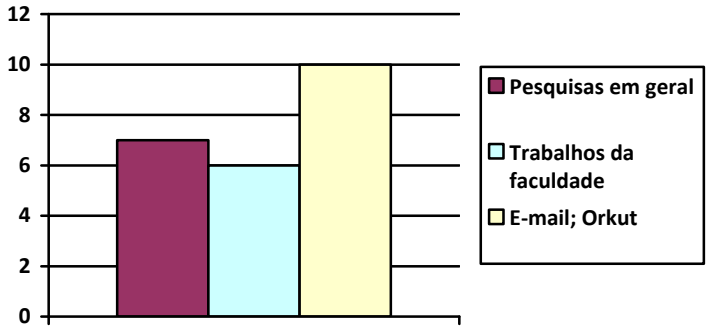


Gráfico 4: Frequência dos tipos de uso de computadores pelos professores

A partir dessas justificativas de usos podemos inferir que a maioria dos professores são motivados a usar os computadores pelo processo de formação inicial em que eles estão envolvidos. Seis dos doze professores utilizam computadores para realização de atividades da faculdade os outros são motivados pelos sites de relacionamento (Orkut), correio de mensagem e pesquisas em geral, já que todos que afirmam utilizar computador apontam o E-mail e Orkut como algo que geralmente acessam.

Na fala dos professores perpassa a existência de diferentes utilizações para o computador, como segue nos extratos abaixo.

> Pesquisadora: Tu utiliza o computador na tua vida? No dia-a-dia...
> Poliana: Utilizo.... masssss... não com muita frequência... assim de vez em quando.
> Pesquisadora: de vez em quando... e quando tu utiliza tu utiliza mais o que?
> Poliana: Só mais Orkut... e-mail essas coisas assim... e... de vez em quando pesquisa.
> Pesquisadora: Na tua área profissional tu usa pra alguma coisa? Tu como professora utiliza o computador pra fazer alguma coisa?
> Poliana: Não... não uso não

> “tenho uns trabalhos que eu faço da faculdade, é muita coisa pra fazer. quase todo dia tenho que ir lá”. (Sandra)
>
> “ultimamente eu to usando para terminar minha monografia, ou então pra pesquisar na internet. Eu gosto muito de pesquisar... é... atividades da Nova Escola. Pra trabalhar junto com eles, em sala de aula. Mas, assim... eu gosto muito de pesquisar texto, poesia. Eu trabalho muito com texto e poesia. Eu gosto de pesquisar na internet pra eles”. (Fábia)

Podemos identificar, a partir da análise das falas dos professores, que a maioria não identifica o computador como uma ferramenta de auxilio no trabalho docente, seja como uma fonte de pesquisa, ou seja, na construção de novos conhecimentos. Seu uso está vinculado, muito fortemente, a necessidades individuais ou de formação profissional.

Quando perguntamos sobre computadores nas escolas ou nas comunidades em que as escolas estavam situadas. Nenhuma das escolas possuía computador e apenas 2 escolas ficavam próximas a *lan-houses* em que os alunos pagavam para acessar os computadores.

No segundo enfoque das questões sobre o tema Tratamento da Informação, buscávamos identificar pelas falas dos professores se eles abordavam esse assunto, que exemplos eles davam para essa temática e qual a importância desses conhecimentos para os alunos.

Quando os professores foram questionados sobre o trabalho com leitura de gráficos, apenas dois afirmaram nunca ter trabalhado com Tratamento da Informação em sala de aula e dez afirmaram já ter trabalhado. As duas professoras que afirmaram nunca ter trabalhado com esse assunto são as que possuem formação apenas no nível médio com o Magistério. Abaixo um extrato de fala da professora Marcela comprovando o pouco conhecimento nessa área de atuação.

> Marcela: Eu trabalhei assim... mais de primeiro e segundo ano e também os meninos da tarde... eu recorto sempre gravuras... e peço para eles fazerem... histórias.
> Pesquisadora: Os gráficos que você trabalha mais... Que gráficos você tá falando?..
> M: Do livro mesmo... eu recorto e... antes tinha jornal, mas agora é só do livro mesmo.
> P: Essa aula com gráficos e em que disciplina?
> M: Envolve todas né! Porque assim... dependendo do texto, porque para fazer interpretação né, e ô. Mando eles terminar a história, observa a gravura... dos pequenos têm os animais... e faço porque assim tem vários tipos, da construção do texto no caso. se for do primeiro ano... aí cada um eu faço é um tipo diferente.

> P: E na Matemática, na aula de Matemática trabalha com.... aí eu vou falar especificamente em gráficos estatísticos... você já trabalhou?
> M: Hummm... não pouco. meu trabalho é mais com material dourado, é o que eu mais utilizo, com fichas também, para somar ou dividir, é nesse estilo ai.

A professora Marcela tem 26 anos de ensino e o Magistério como formação. Durante essa fase da entrevista, pode-se identificar que ela confunde “gravuras” com “gráficos”, no entanto essa confusão não se dá pela semelhança entre as palavras, mas, sobretudo, pelo desconhecimento do que seria o trabalho com gráficos na área da Estatística, como sugere o extrato de fala. Uma grande parte dos professores afirmou abordar gráficos em sala, no entanto, percebemos que não há uma sistematização, uma regularidade como demonstra o Gráfico 5.

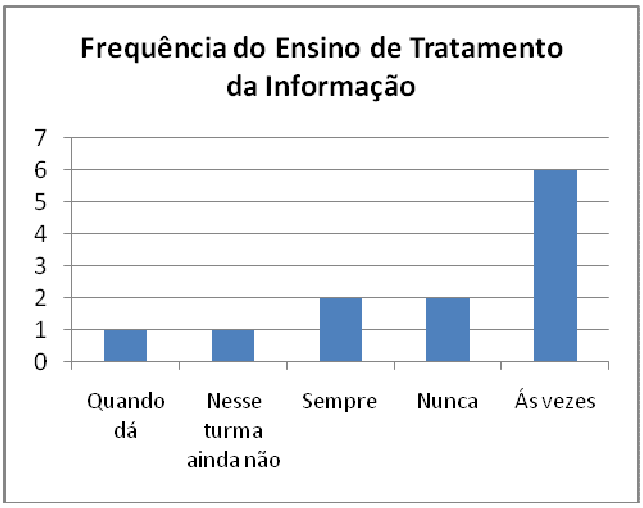


Gráfico 5: Frequência de respostas dos participantes no que se referes ao Ensino de Tratamento da Informação

A grande maioria dos professores afirmou abordar o conteúdo, no entanto, não existe uma regularidade, eles afirmam abordar esse assunto ás vezes. De uma maneira geral, as falas dos professores indicam um certo desconhecimento sobre a temática e uma despreocupação na sua sistematização. Então, perguntamos qual a importância em se trabalhar com Tratamento da Informação em sala. A grande maioria apontou como uma importante ferramenta que auxilia no entendimento de outros conhecimentos. Cinco dos 12 professores relataram que os gráficos auxiliam os alunos a resolverem problemas e entenderem as quantidades, como exemplo a fala da professora Keila.

> Pesquisadora: E você acha que é importante aprender o Tratamento da Informação?
> Keila: Sem dúvida... porque é assim.... com a aprendizagem de gráficos... é porque é assim, veja bem . é tema de dificuldade.... é que eu sempre me analiso no final do que ensinei a um aluno e se ele aprendeu, na aula seguinte em que eu vejo que ele não teve a condição de abssorver o que eu falei... nas quantidades eles têm dificuldades então se trabalho com gráficos eles entendem melhor. aí quando é um trabalho com gráficos eles se saem melhor nas quantidades.

> Pesquisadora: E você acha importante trabalhar com gráficos?
> Mercia: Acho
> P: Porque?
> M: Assim... acho mais fácil para eles aprender. eles entendem mais.
> P: Em que sentido?
> M: A matemática, a Geografia... quando eu coloco o gráfico eles entendem melhor os outros conteúdos.

A partir das falas das professoras Mércia e Keila podemos inferir que o trabalho com gráficos é entendido como uma ferramenta que auxilia na compreensão de outros conteúdos da Matemática ou de outras áreas do ensino. Não identificamos uma preocupação em abordar os conhecimentos Estatísticos necessários à compreensão das informações transmitidas em gráficos e tabelas, por exemplo.

Além dessas justificativas apresentadas nos extratos de falas acima, quatro professores relataram ser importante para a vida dos alunos, eles desenvolverem conhecimentos e habilidades Estatísticas. Como segue na fala abaixo.

> Pesquisadora: E você acha importante... que eles aprendam a trabalhar com gráficos?
> Carmem: Acho importante. Para saber interpretar né... e ler né.
> P: Por que?
> C: Porque assim... facilita mais a vida deles, assim.... a vida do aluno... e é importante... assim participar de coisas na vida deles. Para ler nos jornais e revistas as informações.

A fala da professora Carmem é um exemplo de fala que demonstra a preocupação em trabalhar com o Tratamento da Informação como habilidade necessária para a vida do cidadão. Podemos apontar diferenças nas falas desses dois grupos de professores, os que relatam o uso da Estatística para auxiliar o entendimento de conceitos em outras áreas e os que relatam o uso da Estatística necessários a vida social.

Podemos a partir da entrevista perceber que a maioria dos professores utiliza pouco o computador, sendo esse uso direcionado para os trabalhos da formação a

que estão vinculados. Nenhum professor utiliza o computador para o ensino, como também, nenhuma escola possui computador. A vivência com a tecnologia está mais associada às necessidades da vida de cada um e não as necessidades sociais vinculadas ao processo de mudança da própria escola em relação ao uso de tecnologias pelos alunos. Já no que se refere ao Tratamento da informação também percebemos que é pouco abordado e pelas justificativas apontadas pelos professores tem como objetivo maior o ensino de outros conteúdos ou assuntos de outras áreas. O foco na maioria das vezes não são os conceitos Estatísticos.

Após essas considerações sobre os momentos inicias das entrevistas discutiremos o momento de familiarização com o *software*.

5.2. Apresentação do *Software*

Após os momentos iniciais a pesquisadora iniciou a apresentação do *TinkerPlots* para cada participante. A proposição para esse momento foi uma apresentação do *software* e em seguida o participante iria manipular as ferramentas demonstradas e outras que possivelmente puderam ser descobertas para responder uma questão sobre um banco de dados sobre as características de 24 gatos, como descrito no método.

Essa apresentação inicial foi pensada a partir de uma sequência pré-definida realizada com todos os professores utilizando a mesma ordem até a realização da questão realizada nesse momento. Após a realização da questão problema o encaminhamento diferiu a partir da manipulação de cada participante.

5.2.1. Esse é o *software TinkerPlots*

De início o pesquisador apresentou o *software* e seu *layout* como segue.

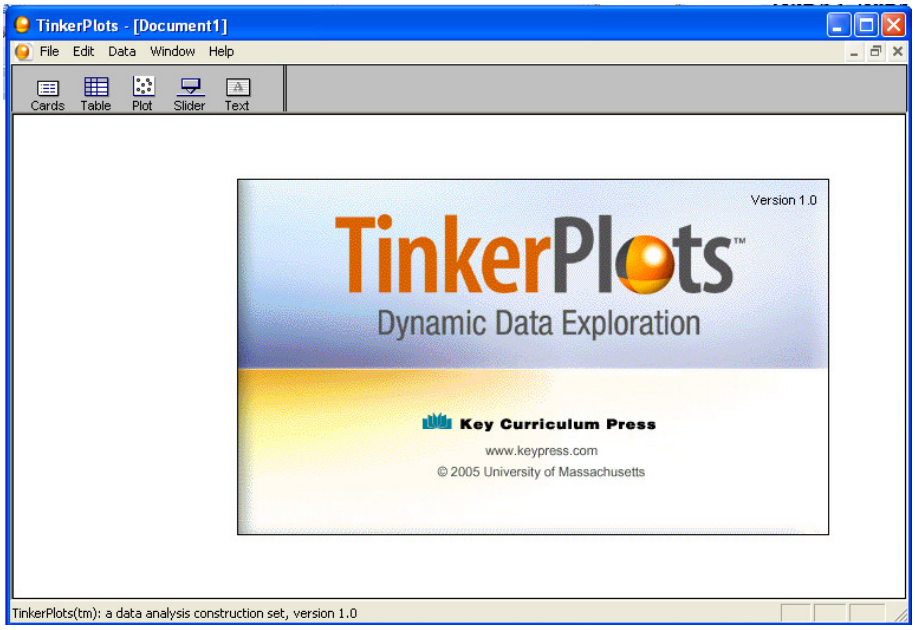


Figura 29: Tela inicial apresentada no *TinkerPlots*

Foi explicado que esse era um software desenvolvido para o ensino de Estatística para crianças. Nele poderíamos construir um banco de dados e/ou utilizar os já existentes no programa. Também poderíamos manipular as variáreis buscando entender o comportamento de cada uma. Após essa explicação o pesquisador apresenta o banco de dados de gatos como mostra a Figura 30.

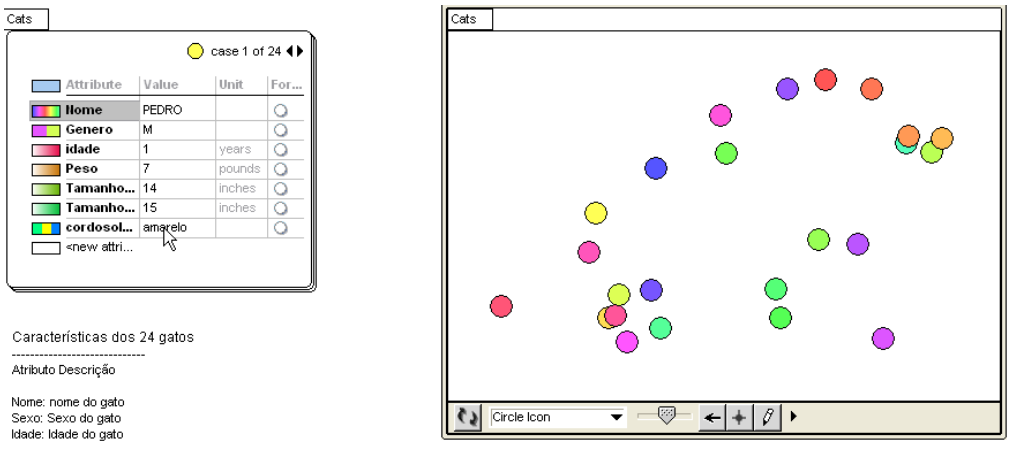


Figura 30: Tela com o banco de dados dos gatos apresentado na Familiarização

A pesquisadora apresenta cada forma de representação do software focando no *cards* e *plots.*

> Pesquisador: Então para começar vou te mostrar o que ele faz. Esse quadrado aqui [aponta com o *mouse* para os *cards*] é o que ele chama de *cards*, que são o lugar onde colocamos nossos dados. É como se fosse fichas individuais para cada gato. Então, por exemplo, esse banco de dados é de gatos, a gente tem características dos gatos como nome do gato, idade do gato, peso, tamanho... [aponta com o *mouse* para cada variável falada no *card*] eu tenho vinte e quatro gatos como mostra aqui no *card* se eu clicar aqui [aponta para o número 24 no *cards*] vão aparecer os outros gatos. Charles, Bichano, Penélope... e as características de cada gato. Então aqui temos um banco de dados e cada gato desse é representado por uma bolinha ou *plot* [apontando com o *mouse* para cada *plot*]. Então se eu clicar no *plot* [clica no *plot*] automaticamente aparece quem é o meu gato no *card* ou se eu clicar no *card* aparece quem é o meu gato no *plot*. Então essa bolinha verde [clica em um *plot* verde] é a gata [aponta para os *cards*] Raquel, ela tem 5 anos, pesa 10 quilos, mede 12 cm... Esse *software* trabalha diretamente com cores, veja não tem nenhuma cor igual, por que todos os gatos tem nomes diferentes. Já se eu colocar em gênero [aperta no atributo gênero] aparecem duas cores...

A fala da pesquisadora descrita acima se preocupa em apresentar cada parte importante no *software*. No começo, apresenta o *card*, uma possibilidade de apresentar o banco de dados através de fichas individuais, depois apresenta as variáveis do banco de dados e os seus diferentes casos, ou seja, gatos. Em seguida, demonstra a outra possibilidade de representação os *plots*, forma de bolas diversas que representam cada variável de cada gato do banco de dados e, também, demonstra a relação entre os *cards* e os *plots*. Também demonstra a utilização das cores para indicar o comportamento das variáveis, como no caso do gênero que é representado por duas cores.

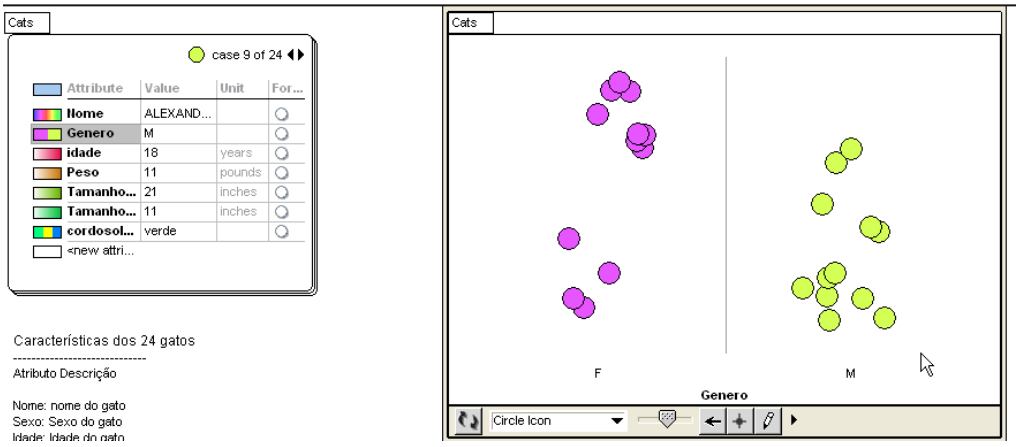


Figura 31: Demonstração da ferramenta *gradiente* no banco de dados gatos na Familiarização

Na Figura 31 a representação do gênero dividido em agrupamentos foi a primeira representação que a pesquisadora realizou para demonstrar uma segunda

ferramenta, *separate*, que pode ser utilizada ao clicar em qualquer *plot*, segurar o clique e arrastar para o lado direito. Em seguida a pesquisadora apresenta a ferramenta *gradiente* na variável idade, como segue.

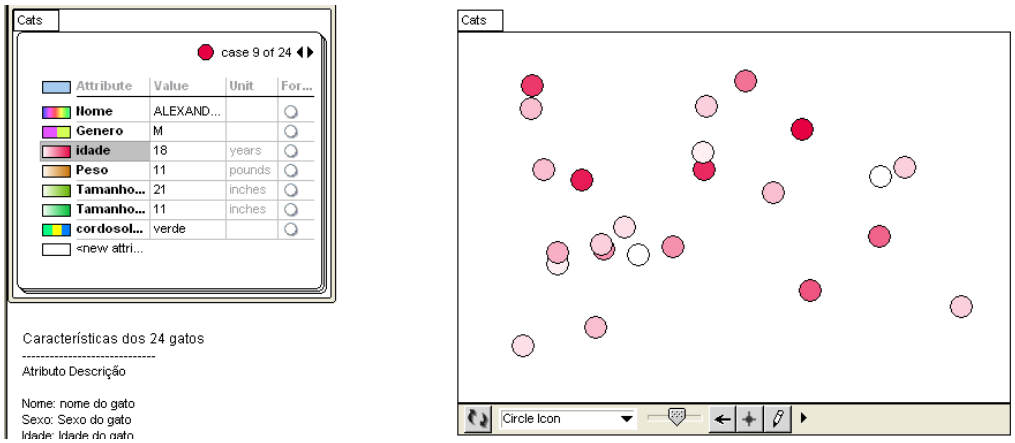


Figura 32: Demonstração da ferramenta *gradiente* indicando a mudança das cores de acordo com a variável

Essa representação é utilizada pela pesquisadora para demonstrar que as cores dos *plots* modificam-se de acordo com a variável. Por exemplo, na Figura 32 está representada uma situação na qual, a variável que esta sendo selecionada é idade, então, as bolinhas mais escuras são os gatos mais velhos e as mais claras são os gatos mais novos. O trecho a seguir apresenta um extrato de fala do protocolo da professora Carmem no momento da apresentação dessa ferramenta.

> Pesquisadora: E se eu clicar em idade... [clica no *cards* no nome da variável idade] o que aconteceu?...
> Carmem: A idade também... tá mais claro e mais escuro...
> P: Por quê?
> C: Por conta que cada um tem sua idade...não são todos iguais..
> P: Você viu que tem bolinha mais claras e mais escuras.. o que você acha que quer dizer isso?
> C: Posso dizer que o mais... o mais escuro é o de maior idade... né!?.. e os mais claros diminuiu a idade. É isso...?
> P: Perfeito. Isso mesmo. Ele usa cor em forma de gradação... quanto mais escuro mais velho, quanto mais claro mais novo. A mesma coisa ele faz com peso, quanto mais escuro mais pesado, quanto mais claro.

Após essa apresentação o pesquisador clica em um *plot* e arrasta, como segue na Figura 33.

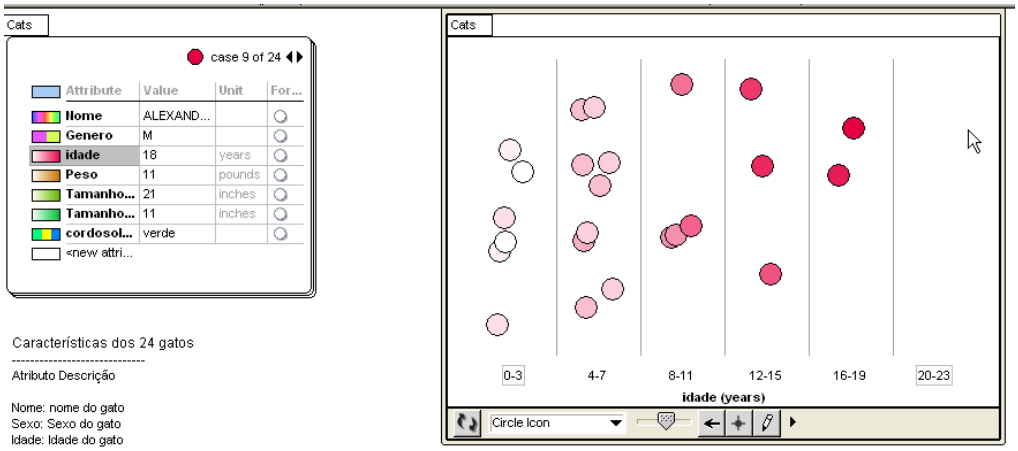


Figura 33: Demonstração da ferramenta *separate* agrupando a idade dos gatos em classes de intervalos

Na situação representada pela Figura 33, o professor observa à variável idade agrupada em classes de intervalos, essa ação é possível através da ferramenta separar. A pesquisadora novamente clica e arrasta para demonstrar a variável em outra representação.

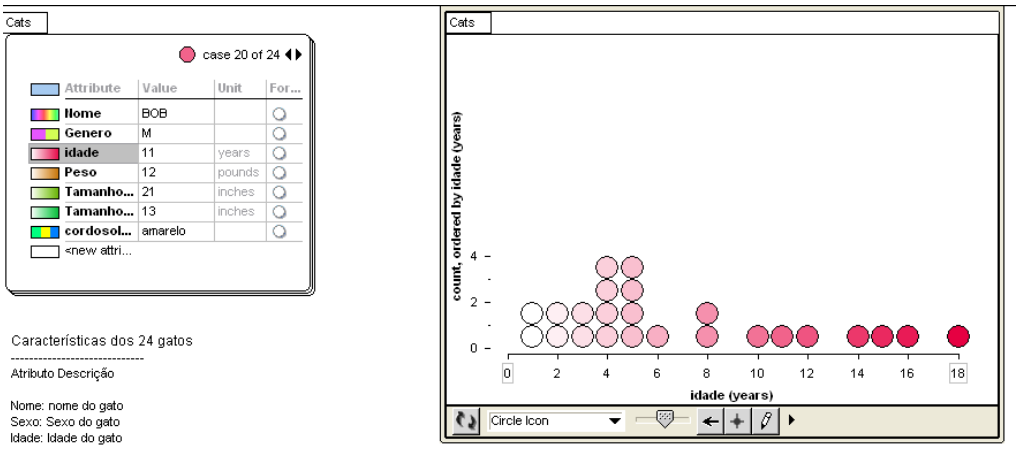


Figura 34: Demonstração da ferramenta *separate* na construção de classes pontuais para a variável Idade

Quando a pesquisadora mostra a representação referente a Figura 34 apresenta uma representação das idades dos gatos dispostas em classes pontuais, no entanto a pesquisadora, ainda, seleciona as ferramentas ordenar e empilhar para que os *plots* ficassem organizados em fileiras ordenadas. A pesquisadora realiza a construção de todas essas representações demonstradas até o momento e vai realizando perguntas sobre os dados, como segue.

> Pesquisadora: Qual é o gato mais novo que eu tenho aí?
> Mercia: Esses dois ai [aponta para os *plots* que estão no intervalo entre 0 e 2 na escala]
> P: Tem quantos anos esses aí?
> M: Um ano
> P: E o mais velho?
> M: Tem dezoito

A professora Mércia interage com a pesquisadora e responde sem dificuldade o que foi questionado, mesmo quando aponta para o valor de 1 ano que esta oculto na escala. Isso demonstra que a professora está compreendendo os dados e buscando através das representações construídas as respostas às questões.

Na sequência é apresentada a possibilidade do cruzamento de variáveis, ou o contexto bivariado, como é denominado por alguns autores, como segue trecho seguinte.

> Pesquisadora: Então... você viu que é bem tranquilo né! Agora esse *software* também ajuda a gente a cruzar variáveis. Que danado é isso? Por exemplo, se eu perguntasse a você quem tem a tendência de ser maior são os gatos ou as gatas? Quem tende a ser maior gato ou a gata?
> Poliana: O gato
> Pesquisadora: Será que é? Vamos ver nesse grupo de dados?

A pesquisadora levanta uma hipótese sobre o banco de dados. A partir da manipulação dos *plots* o participante irá tentar chegar a resposta. Nesse momento os professores pegam o *mouse* e começam a utilizar o computador de maneira autônoma. O pesquisador, ainda, indica o que pode ser feito, no entanto, o professor pode ir testando as ferramentas do modo que preferir.

A partir desse momento da entrevista cada professor dá um encaminhamento diferenciado, já que as ferramentas disponíveis no software oferecem diferentes possibilidades de representação. Essas diferenças entre a manipulação realizada pelos professores puderam ser identificadas nas transcrições durante o processo de análise e categorizadas em grupos semelhantes como iremos apresentar nos subitens a seguir.

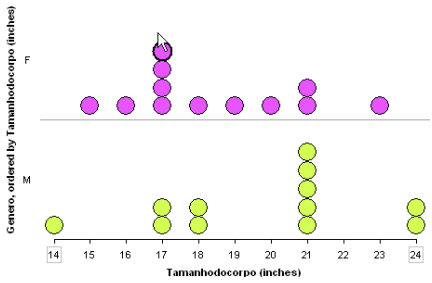
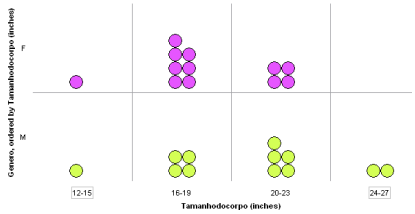
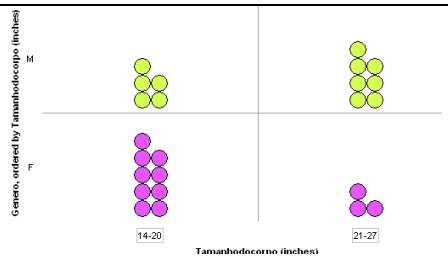
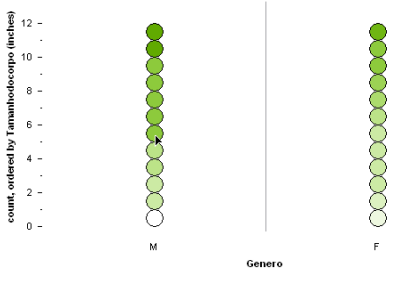
5.2.2. Primeiras manipulações com o *software*

Após observarem as ferramentas do software os professores foram motivados a utilizá-lo na busca pela resolução de um questionamento levantado pela pesquisadora: Quem tende a ser maior o gato ou a gata? Quando essa questão era feita eles tinham liberdade para testar as ferramentas e construir diferentes

representações. A pesquisadora servia de mediadora auxiliando nas construções das representações e intervindo através de questionamentos que auxiliassem a reflexão no uso do *software*.

Categorizamos em quatro as diferentes representações construídas pelos professores para responder a questão proposta e a resumimos no quadro 3 a seguir.

Quadro 3: Representações produzidas pelos professores na familiarização

Tipo de Representação	**Representação**	**Quantidade de professores que utilizaram/Tipo de representação**
Tipo 1		12
Tipo 2		5
Tipo 3		3
Tipo 4		12

O Quadro 3 apresenta os quatro tipos diferentes de representações utilizadas pelos professores para responder a questão proposta pelo pesquisador. Nesse momento esperávamos que os professores conseguissem de forma autônoma construir diferentes representações na tentativa de responder a questão proposta.

Então em uma análise global podemos perceber as representações mais utilizadas foram a Tipo 1 e 4. A representação Tipo 1 apresenta as variáveis de gênero separadas com atribuição de cores: rosa para o gênero feminino e verde para o gênero masculino. São dispostos dois eixos, um para a escala dos *plots* rosas e um eixo para os *plots* verdes. A representação do Tipo 4 apresenta os *plots* separados em duas fieiras cada uma correspondendo a um gênero. Na representação a cor verde é atribuída tanto ao gênero masculino como ao gênero feminino. Todavia, existe uma variação da tonalidade de verde para cada *plot* de acordo com o tamanho dos gatos. A seguir analisaremos cada representação construída.

a) Tipo 1 de representação

A representação do Tipo 1 foi construída por todos os professores como primeira alternativa de representação. Essa escolha pode estar atrelada ao processo de apresentação do *software*, já que a última representação construída pela pesquisadora antes de realizar o questionamento foi uma variável (idade) representada em escala de classes pontuais, conforme apresentada na Figura 35. No entanto, tivemos diferentes reflexões sobre essa representação, como exemplo temos a fala do professor José que representa a maioria dos professores que tendiam a responder, utilizando essa representação, o ponto máximo das barras.

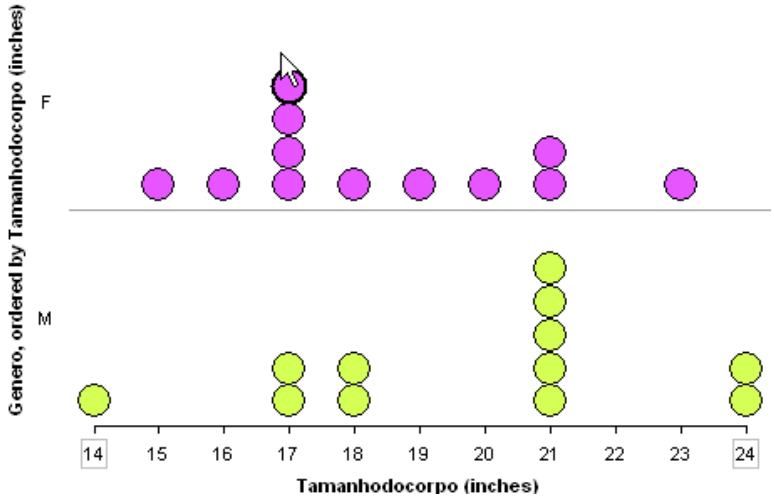


Figura 35: Representação tipo 1 construída na familiarização que cruza a variável gênero x tamanho

Pesquisadora: E agora quem é maior gato ou gata?
José: Gatos!!!
P: Os gatos são maiores? Por quê?
J: Pela quantidade....
P: Como assim? Explica melhor...
J: Pela quantidade de bolinhas.... que... representa, né?.... [fica pensando durante 2 min]
P: A pergunta é quem é maior em tamanho gato ou gata?
J: É os gato. Por que aqui em 21 tem cinco e lá em 17 tem quatro. Tem mais em gato. [se refere a frequência nos valores 21 e 17 da escala]
P: Ah! Então a resposta é...
J: Os gato são maiores.

A fala do professor José é um exemplo da maioria das respostas baseadas no ponto máximo ou pilha maior de bolinhas (barra mais alta) que representa maior frequência. Quando ele compara as pilhas (barras) do grupo gato e gata a partir da frequência apenas, ele não considera a variável tamanho que está sendo representada através da escala horizontal, o que indicaria que os maiores gatos estariam situados a direita da escala a partir de 21 cm. Essa resposta foi dada por 8 dos 12 professores se baseando no ponto máximo a partir dessa representação.

Os 4 professores que não basearam-se no ponto máximo responderam considerando a escala e afirmaram que os gatos eram maiores por chegarem a 24 cm e que no grupo das gatas chegavam até 23cm. Essa afirmativa também está apoiada em um ponto extremo da escala, já que os professores não consideram o comportamento da variável, ou a quantidade de *plots* (casos) em cada valor da escala, de uma forma global, mas apenas o último valor da escala para cada grupo. O extrato de fala da professora Poliana exemplifica essa consideração.

> Poliana: É gato!
> Pesquisadora: Por quê? Onde você tá vendo isso?
> Poliana: Aqui. [aponta para escala] os gatos tem 24 centímetros, né! E as gatas têm 23. O gato é maior.

Essa fala sugere o que havíamos discutido anteriormente, que a professora se apoiou com maior ênfase no ponto extremo da escala. A resposta a essa questão precisa levar em consideração o ponto extremo da escala, como realziado por 4 dos professores, no entanto não apenas nesse ponto. A análise precisa considerar o comportamento da variável, tendendo a uma interpretação global. Com isso, teríamos que observar que os *plots* no grupo dos gatos estão em sua grande maioria dispostos nos valores maiores da escala, enquanto que nas gatas há uma tendência de regularidade o que pode indicar que os gatos tendem a ser maiores.

b) Tipo 2 de representação

Outra forma de representação utilizada pelos professores foi o Tipo 2 que apresenta intervalos de tamanhos dos casos. Nesse tipo 2 os *plots* para cada gênero estão agrupados em cada faixa de valor e dispostos em eixos horizontais. Nessa representação os *plots* verdes representam os gatos e os rosas representam as gatas. A seguir um exemplo de extrato de entrevista.

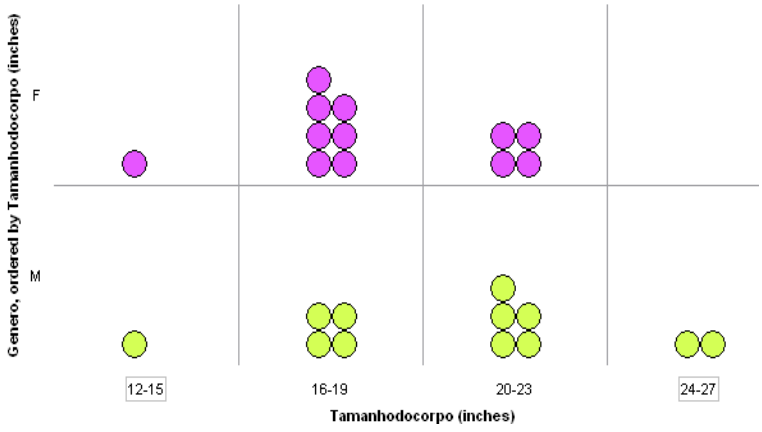


Figura 36: Representação tipo 2 construída na familiarização que cruza a variável gênero x tamanho

> Pesquisadora: E então assim mudou alguma coisa?
> Keila: Sim.
> P: E quem são maiores?
> K: Gato! Por que... [fica olhando para o tela por 1 min] Pelo valor daqui oh. Vê! Do tamanho dele... [aponta para escala] Né tamanho?
> P: É. E mudou?
> K: Então, 24..., mas se fosse por quantidade seria elas.
> P: Não entendi.
> K: Por quantidade de bolinhas assim... seria elas [aponta com o *mouse* para a barra mais alta – gatas intervalo de 16-19]. Mas, se for pelo número... é os gatos [aponta a categoria 24 – 27].. pela sequência numérica... o centímetros do gatos... o maior seria o gato.

A professora Keila, que havia respondido através do ponto máximo nas barras diante da representação anterior, muda de opinião a partir da nova representação. Ela afirma que pode obter a resposta tanto pela escala, quando aponta para o maior número no intervalo, como também, pela quantidade de bolinhas que representam a frequência representada pelo tamanho das barras. Nas duas respostas ela enfoca nos pontos extremos, tanto da escala como da barra mais alta. Um ponto relevante a ser observado é que essa professora aparentemente demonstra entender o que representam as classes de intervalos, isso fica evidente quando relata que os números contidos nas categorias representam o tamanho dos gatos em centímetros.

Dos cinco professores que utilizaram essa representação 3 apresentaram entender as classes de intervalos e identificaram a resposta a partir delas e não do ponto máximo da barra. No entanto, a dúvida entre responder pelo tamanho da barra ou pelo valor do intervalo pode ser um indício de que o *software* proporcionou, através da ferramenta separar, uma reflexão sobre as possibilidades de diferentes evidências a serem consideradas no momento de uma interpretação de gráficos.

c) Tipo 3 de representação

Outro formato de representação bem semelhante a esse foi o Tipo 3, porém com apenas duas categorias, o qual foi utilizado por 3 professores. Nessa representação o agrupamento dos *plots* estão dispostos em dois intervalos de valores. Assim como a representação tipo 2, os *plots* estão dispostos em eixos horizontais paralelos e as cores verde e rosa são atribuídos respectivamente aos casos de gatos e gatas. Um exemplo de fala a partir dessa representação é o extrato da professora Sandra.

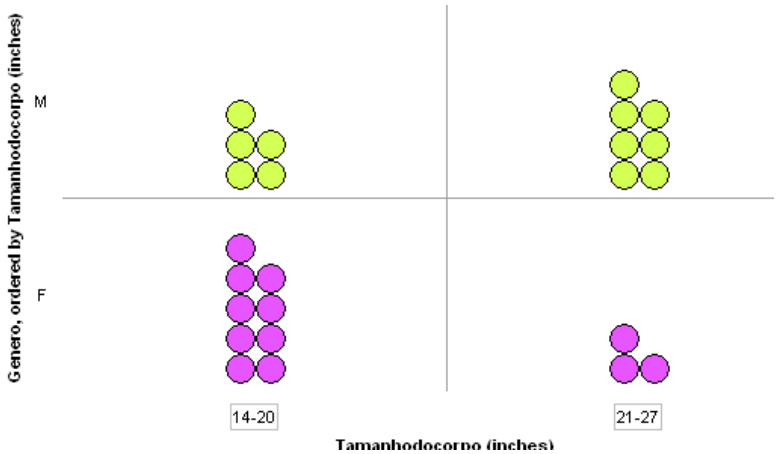


Figura 37: Representação tipo 3 construída na familiarização que cruza a variável gênero x tamanho

> Pesquisadora: E aí mudou um pouquinho né?
> Sandra: É...
> P: E então dá pra dizer quem é maior?
> S: Acho que é fêmea...
> P: Por quê?
> S: [pensa durante 2 min] pelo tamanho... [passa o *mouse* pela barra maior das gatas e depois dos gatos] tem mais gata aqui nessa... nesse lado do que gato aqui.

A professora Sandra que também utilizou o Tipo 1 de representação não muda sua resposta a partir dessa nova representação. Ainda utiliza o ponto máximo para responder a questão. Como essa representação foi bem semelhante a do tipo 2, já era esperado que os professores, ainda, se baseassem no ponto máximo para apoiar suas respostas.

d) Tipo 4 de representação

O Tipo de representação utilizado por todos os professores foi o 4. Essa representação utilizava as cores para identificar a variável tamanho do corpo. Então os *plots* foram divididos por categoria de gênero em forma de barras, sendo empilhados e ordenados de acordo com o dégradé de cores (as cores representavam o tamanho – quanto menor o valor mais claro os *plots*, quanto maior o valor mais escuro os *plots*).

Nessa representação visualizada na Figura 38 os professores responderam que os gatos eram maiores que as gatas. A seguir trecho da fala de um professor.

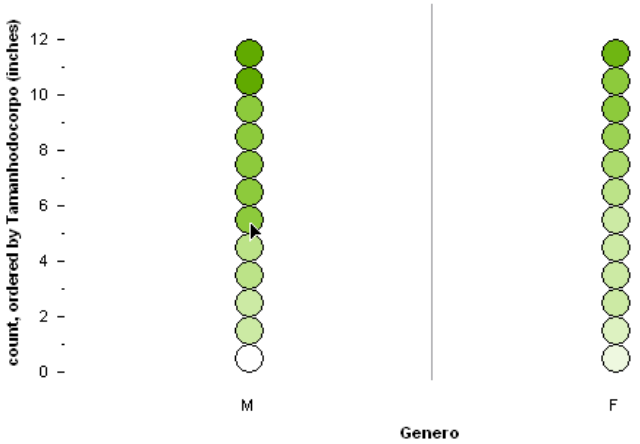


Figura 38: Representação tipo 4 construída na familiarização que cruza a variável gênero x tamanho

> Pesquisadora: Agora eu tenho gato e gata... e a cor indica tamanho. Melhora assim?
> Mércia: Masculino!
> P: Por quê?
> M: Por que aqui tem mais escuro.... mais escuro né maior...né!
> P: Ahhh.. Mas peraí. Finalmente é gato ou gata?
> M: Agora é gato. Por esse gráfico aqui é macho... por que os mais pesados são os mais escuros. Aqui tem mais escuro [aponta para masculino] do que aqui [aponta para feminino]

Essa representação propõe um olhar não para o tamanho das barras, mas, principalmente, para a intensidade da cor, que representa o tamanho dos gatos. Pela fala da Professora Mércia, podemos inferir, que ela compreendeu que as cores representavam o tamanho do corpo dos gatos e estabelece uma comparação entre as bolinhas mais escuras dos gatos e das gatas chegando a conclusão de que havia mais bolinhas escuras nos gatos do que nas gatas. Esse tipo de ferramenta se configurou como um importante mecanismo de reflexão para os professores, já que 5 dos professores ao se deparar com as barras iguais (mesmo tamanho) questionaram inicialmente que os gatos e gatas eram iguais, como segue o trecho da professora Carmem.

> Carmem: Os dois são iguais, oxê... [risos]
> Pesquisadora: Por quê?
> C: O mesmo tamanho.
> P: Ah.... Mas a gente não tinha visto que tínhamos a mesma quantidade de gatos e gatas?...
> [a professora balança a cabeça]
> P: Lembra ai quem tá dizendo o tamanho?
> C: Hummm [dá um sorriso] então é a cor!!!!... Então são os gatos.
> P: Por quê?

> C: A quantidade de verdes mais escuros aqui é maior [aponta para os masculinos].
> P: Tem certeza.
> C: Agora tenho sim.

As barras, do mesmo tamanho, causou estranheza a professora Carmem, o que podemos inferir através dos risos no momento da reflexão. Após a intervenção do pesquisador enfocando que as quantidades eram iguais para gatos e gatas, ela relembra a função das cores na representação. Naquele momento de sua fala Carmem afirma com convicção que os gatos são maiores por que tem mais bolinhas escuras no grupo dos gatos. Esse momento foi extremamente rico, pois fez com que ela repensasse as afirmações anteriores e forneceu novas evidências, através das cores, para contribuir com a resposta. Todavia, o uso dessa representação tipo 4 em um momento de interpretação de dados pode encobrir algumas informações importantes e que precisam ser consideradas no momento da interpretação.

Como exemplo de dados encobertos através da representação tipo 4, podemos citar esse mesmo caso do tamanho dos gatos. Por exemplo, se os *plots* (bolinhas) escuras, presentes no grupo das gatas, representasse tamanhos extremamente maiores do que os tamanhos contidos no grupo dos gatos esses valores poderiam interferir no todo da análise da variável, o que, poderia ocasionar uma diferença significativa e dar como resposta as gatas sendo maiores, mesmo os gatos possuindo uma maior quantidade de bolinhas verdes escuras.

Então, podemos identificar que as ferramentas foram utilizadas de maneira diversa pelos professores, como forma de auxiliar a reflexão sobre os dados estatísticos. Os professores não tiveram dificuldades ao realizar as representações e o pesquisador teve uma função importante nesse momento indicando o uso das ferramentas e levantando reflexões sobre as respostas dos professores.

De uma maneira geral, esse momento de familiarização com o *software* não demonstrou dificuldades no uso com o computador pelos professores. Essa etapa da pesquisa teve como objetivo apresentar o software e suas ferramentas. Os professores manipularam o *mouse* e construíram diferentes representações de maneira autônoma, como também pôde propiciar um momento de interação para que os professores se sentissem seguros ao utilizá-lo no momento da realização das atividades que foram propostas no segundo encontro.

No Capítulo a seguir abordaremos as representações construídas pelos professores ao responder as atividades, que ferramentas foram mais utilizadas e que contribuições o uso dessas ferramentas proporcionaram para a interpretação dos dados de dois bancos de dados utilizados nas atividades.

CAPÍTULO 6 – SEGUNDO ENCONTRO: ANALISANDO AS REPRESENTAÇÕES CONSTRUÍDAS A PARTIR DO *SOFTWARE*

Para a segunda sessão de coleta de dados foram realizadas 6 questões de cruzamento de variáveis, sendo que 3 eram sobre um banco de dados de violência sexual contra crianças e 3 eram sobre Novelas exibidas no horário nobre de uma emissora de televisão, como exposto no Quadro 2 do capítulo do Método.

Para iniciarmos a análise iremos inicialmente procurar evidenciar os aspectos de análise global e depois de maneira específica. Então, começaremos analisando as representações construídas pelos professores sem nos atermos inicialmente aos diferentes tipos de questões e em seguida analisaremos essas representações e suas características de acordo com os tipos de cruzamento de variáveis.

O *software TinkerPlots*, como descrito no capítulo específico, é interativo e de difícil explicação a partir de palavras, já que ele é extremamente manipulável, nesse sentido as ferramentas disponíveis nele estão ligadas as representações construídas, já que para construir qualquer representação são utilizadas as ferramentas separar, empilhar, ordenar e *gradiente*. Nesse sentido iremos analisar as ferramentas associando-as as representações construídas e a suas mudanças a partir das intervenções da pesquisadora. Chamamos aqui intervenção toda observação, ou indicação de uso de ferramenta realizada pelo pesquisador no momento da interpretação e que possa de alguma forma ter contribuído para interpretação. Nesse sentido, apresentaremos um levantamento das ferramentas utilizadas de uma maneira geral e na sequência identificaremos o uso dessas ferramentas em cada questão de cruzamento de variáveis e sua possível influência na interpretação.

6.1. As representações utilizadas

No segundo encontro os professores apresentaram uma maior disponibilidade em participar do estudo, já que eles conheceram o *software* através do encontro da familiarização. Eles demonstraram entusiasmo e ansiedade para participar do momento de realização da pesquisa. Então, inicialmente era mostrado o banco de dados da Violência sexual, nas quais as 3 questões ficavam expostas na tela do *software*. Os professores manipulavam o *software* para responder as questões de cruzamento de variáveis, em seguida, apresentávamos o segundo banco de dados das Novelas, no qual, também eram expostas na tela as 3 questões de cruzamento de variáveis. Eles poderiam manipular as ferramentas do software buscando construir representações que os auxiliassem na busca pela resposta às questões.

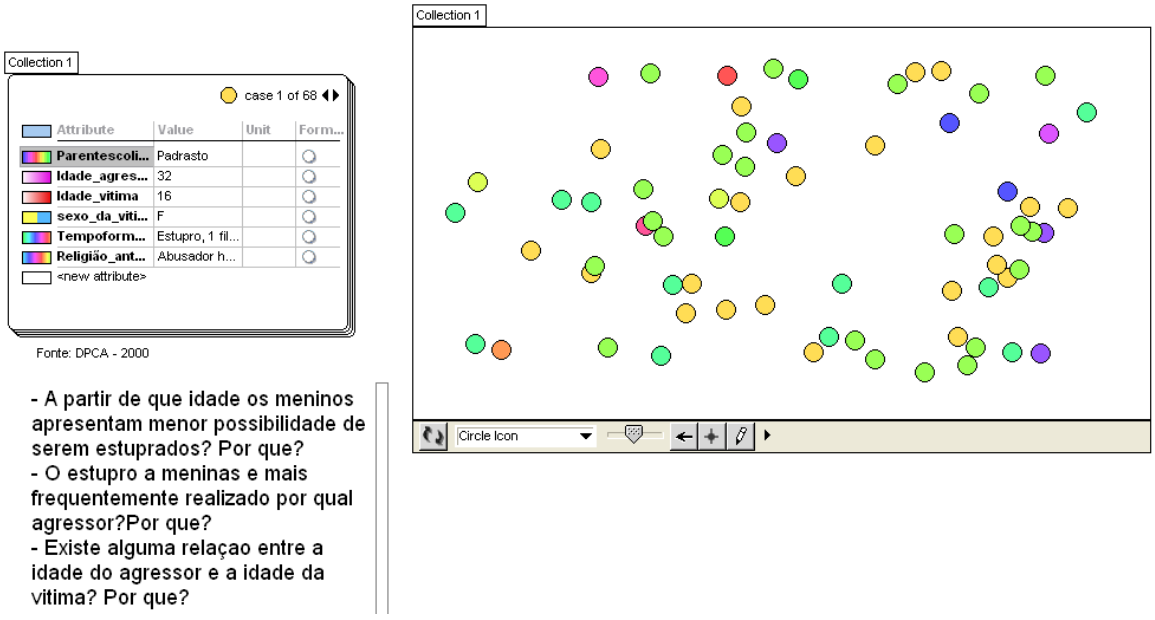


Figura 39: Demonstração da tela do computador no inicio da atividade com o primeiro banco de dados de Violência

Na Figura 39 podemos observar as questões propostas para os professores. Ao todo foram três, que cruzavam dados nas suas diferentes possibilidades (quantitativoXqualitativo; qualitativoXqualitativo; quantitativoXquantitativo). Também podemos identificar a tela inicial do *software* que era apresentado aos professores, todos receberam o computador com essa tela inicial do *software* e puderam utilizar as duas formas de visualização e manipulação de dados, os *cards* e os *plots*.

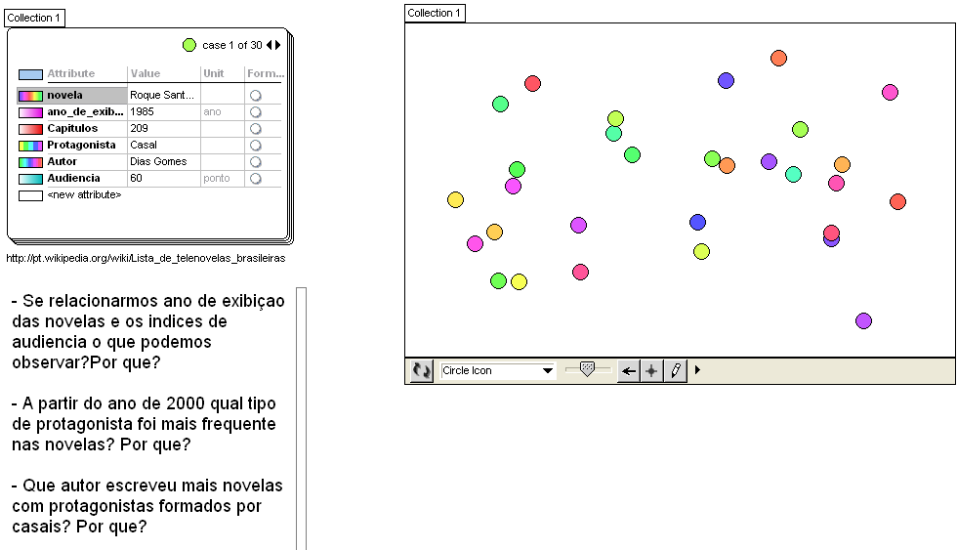


Figura 40: Demonstração da tela do computador no inicio da atividade com segundo banco de dados de Novelas

Na Figura 40 observamos as questões propostas para os professores no segundo banco de dados, também foram propostas três questões de cruzamento de variáveis como no primeiro banco de dados (quantitativoXqualitativo; qualitativoXqualitativo; quantitativoXquantitativo). Os professores poderiam ler as questões a qualquer momento e manipular os *plots* com a utilização do *mouse*.

Inicialmente iremos analisar, de uma forma geral, as representações construídas a partir do *software*. Diferentemente da Familiarização, na atividade principal proposta nesse segundo encontro tínhamos, ao todo, seis questões. Pela maior quantidade de questões e diferentes possibilidades do uso do software, obtivemos muitas representações diferentes. Por esse motivo não tornou-se viável a tipificação das representações uma a uma, como foi proposta na análise da questão da familiarização.

A partir da busca pela resposta para as seis questões os professores construíram ao todo 144 representações durante o processo de análise de dados. Para iniciarmos uma categorização, buscamos as características similares entre essas representações e percebemos que essas características similares estavam presentes no tipo de escala escolhida pelos professores para responder cada questão de cruzamento de variáveis.

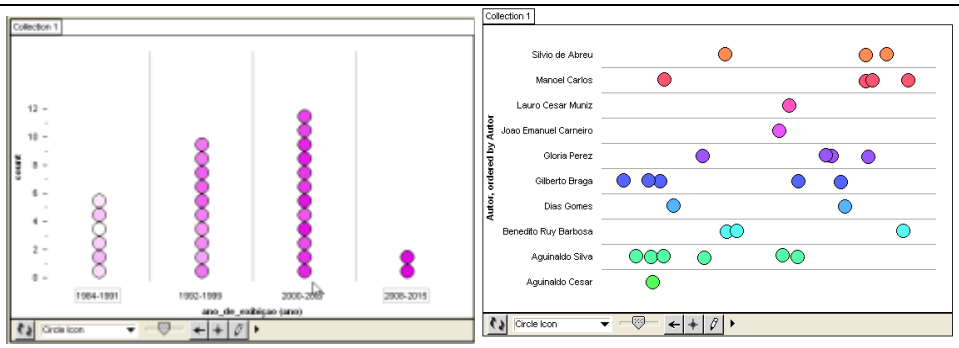
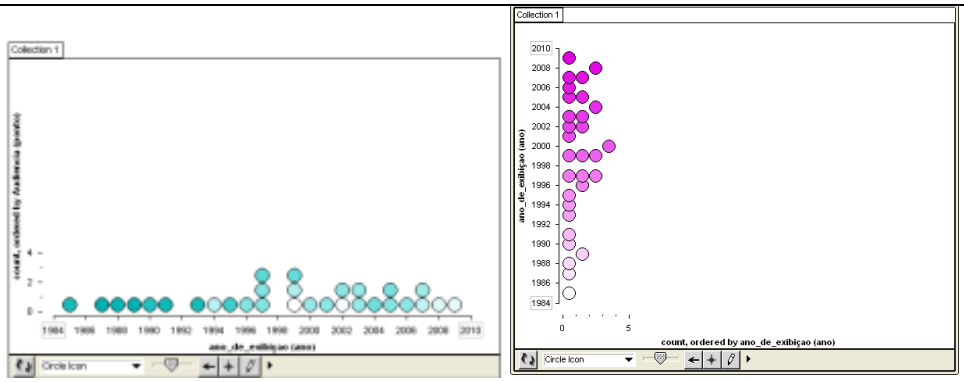
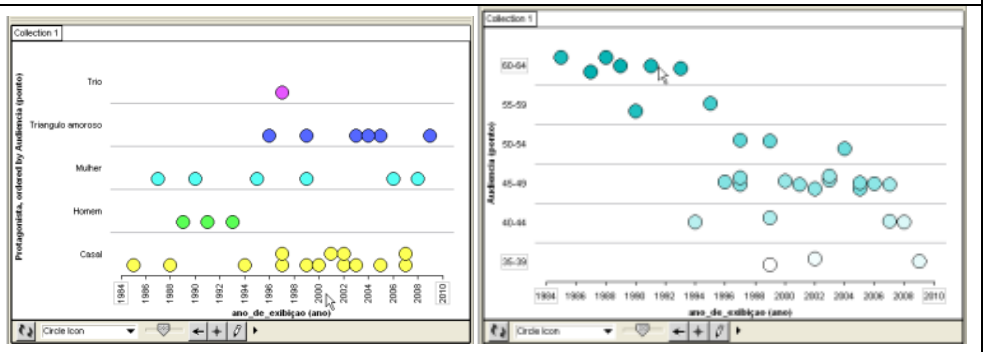
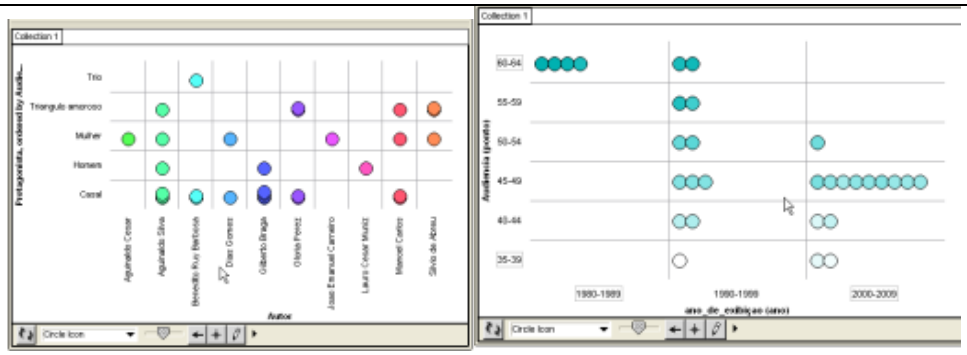
Nesse estudo, como forma de facilitar a percepção entre as diferenças entre os agrupamentos de tipos de representações construídas pelos professores, iremos denominar a organização de variáveis qualitativas de categoria e a organização de

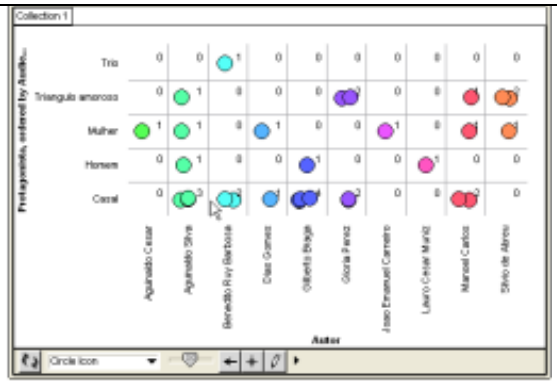
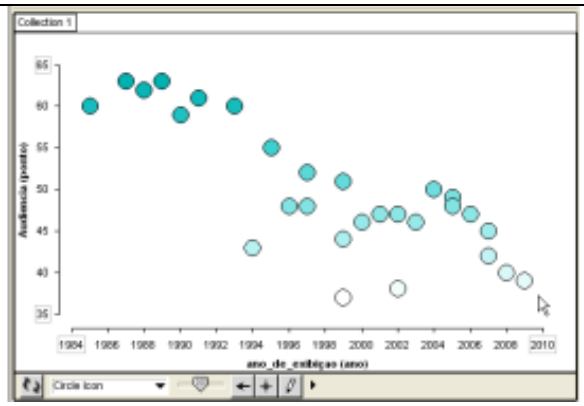
variáveis quantitativas de classes, sendo que classes pontuais, para a disposição de valores pontuais em uma reta e classes itervalares (ou classes de intervalos) as classes que apresentam intervalos iguais entre os valores. Então:

- *Classes pontuais* – organização de variáveis quantitativas em valores pontuais dispostos em um eixo;
- *Classes intervalares* – organização de variáveis quantitativas em intervalos de igual valor;
- *Categoria* – organização de variáveis qualitativas no qual são classificadas em agrupamentos com características semelhantes.

Seguindo essas denominações classificamos em seis características os tipos de representações construídas pelos professores. O Quadro 4 apresenta exemplos de representações dentro de cada tipificação. Podemos identificar as características comuns ao agrupamento de representações e exemplos de representações.

Quadro 4: Exemplos dos diferentes tipos de representações conforme características semelhantes

Tipificação	Exemplos de representações
A Uma variável distribuída em categorias ou classes de intervalo.	
B Uma variável distribuída em classes pontuais	
C Cruzamento de duas variáveis organizadas: uma em classe pontual e outra em classe de intervalo ou categoria	
D Apresenta o cruzamento de duas variáveis organizadas: em categorias ou classes de intervalos	

E Apresenta o cruzamento de duas variáveis organizadas: em categorias ou classes de intervalos com o auxilio do número	
F Apresenta um gráfico de dispersão	

As tipificações A e B apresentam a escala de apenas uma variável, mesmo se tratando do cruzamento entre variáveis. Isso foi possível pelo uso da ferramenta *gradiente*, que utiliza a cor para representar uma variável. As tipificações C, D, E e F apresentam as escalas em classes pontuais, classes de intervalos ou categorias no cruzamento de variáveis. Iremos analisar e discutir mais aprofundadamente nas seções seguintes essas características e os exemplos expostos no Quadro 4.

6.2. Primeiro banco de dados – violência sexual

Para responder as questões sobre o primeiro banco de dados os professores construíram 61 representações. Como forma de análise das representações, categorizamos conforme Quadro 4 e observamos a frequência do uso das representações em cada questão de cruzamento proposta no banco de dados de violência. A seguir o Gráfico 6 apresenta o percentual de frequência do uso das diferentes representações em cada questão proposta.

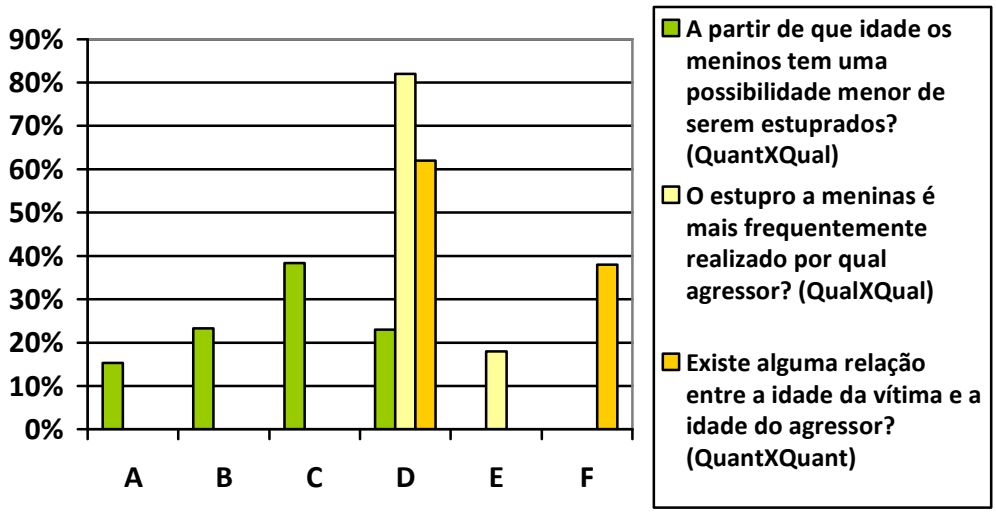


Gráfico 6: Representações construídas pelos professores de acordo com cada questão do primeiro banco de dados - Violência

De uma forma geral as representações mais utilizadas nas três questões foram as agrupadas no tipo D que se utiliza da construção de classes intervalares e categorias (as duas formas de organização de dados se assemelham no formato visual da representação) , tanto com as variáveis qualitativas, como também com as variáveis quantitativas, iremos analisar as representações utilizadas em cada questão.

a) Primeira questão do Banco de dados Violência

Era esperado que os professores cruzassem uma variável quantitativa (idade) e uma variável qualitativa (gênero) para responder “A partir de que idade os meninos tinham menor possibilidade de serem estuprados?”.

Foram utilizadas as representações com características das tipificações A, B, C e D para responder essa questão, sendo que as representações mais utilizadas estavam dentro da tipificação C que apresenta os dados em forma de classes pontuais e categoria. No entanto, tivemos um uso evidente de representações com características das tipificações A e B, no qual apresenta, respectivamente, apenas uma variável em classe pontual ou classes de intervalo e a outra variável foi representada através de cores. Nesse caso os professores utilizaram a cor

(ferramenta *gradiente*) para representar uma variável, como no exemplo da Figura 41.

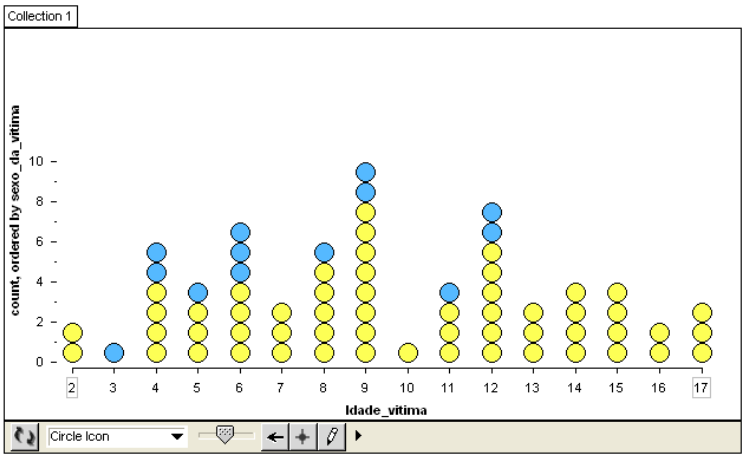


Figura 41: Exemplo 1 de representação dentro da tipificação B utilizada para responder a primeira questão do banco de dados violência

Esse tipo de representação, apresentada na Figura 41, foi a única construída dentro da tipificação B para essa questão de cruzamento. Isso porque a variável quantitativa foi representada em classes pontuais e a outra variável qualitativa através das cores. Dos 12 professores 7 utilizaram esse tipo de representação como primeira forma de manipular os dados.

As representações dentro da tipificação C também foram utilizadas pelos professores, um exemplo de representação dentro dessa tipificação é o apresentado na Figura 42, no qual se assemelha a representação da Figura 41. A diferença entre elas se dá nas cores que identificam o gênero, que na Figura 41 aparece empilhada e na Figura 42 as cores aparecem separadas em dois eixos horizontais, cada vinculado a uma categoria em particular, como exemplo a seguir.

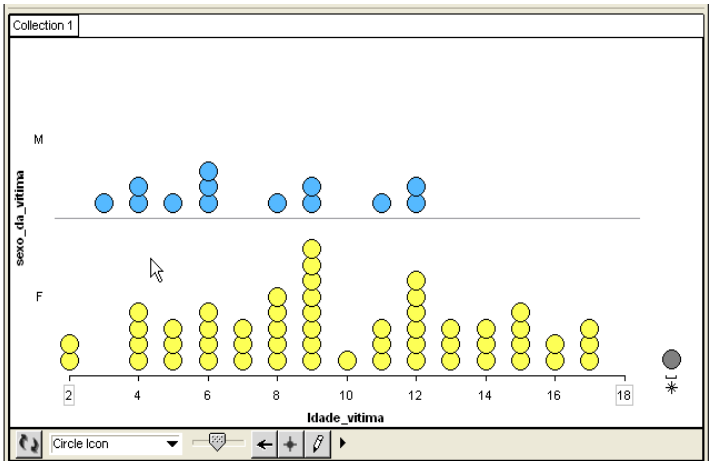


Figura 42: Exemplo 1 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados violência

Os 7 professores que iniciaram com a representação da Figura 41 em seguida separaram as variáveis, ou por iniciativa pessoal ou por indicação da pesquisadora. Outra possibilidade para representação C, bem semelhante a demonstrada na Figura 42 foi a disponível na Figura 43 e utilizada por 5 dos professores, no qual, a idade da vítima era apresentada através das cores. Quando mais escuros os *plots* mais velhos eram as crianças.

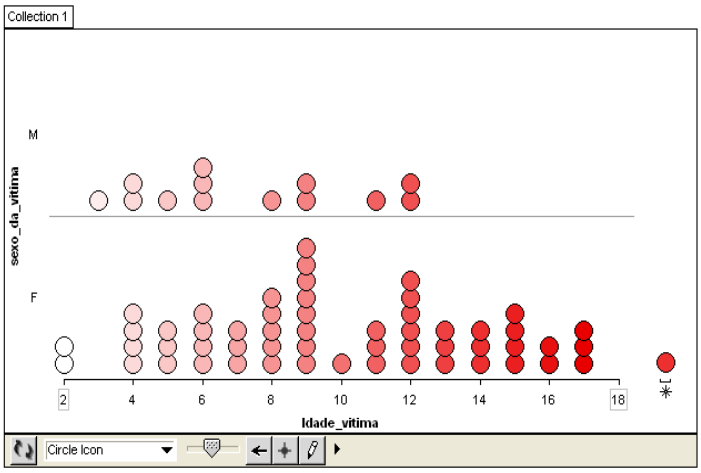


Figura 43: Exemplo 2 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados violência

As duas representações mostradas nas Figuras 41 e 42 tiveram a mesma estrutura, escala para idade e cores para a categoria gênero, já a representação da Figura 43 a cor que representava o gênero passou a representar a idade. Essa mudança faz com que o sujeito ao observar o gráfico não olhe apenas para a barra

mais alta, ponto máximo, mas também chama atenção para as cores, nesse caso os *plots* mais escuros.

Outro grupo de representações utilizadas para responder a questão um do banco de dados de Violência foi a tipificação D. Esse grupo de representações teve como característica o cruzamento a partir da utilização de categorias e classes intervalares. Esse resultado é bem interessante, pois pode indicar que os professores se utilizaram, de uma forma significativa, de intervalos para organizar as variáveis quantitativas ao invés da utilização das classes pontuais, como demonstra o exemplo a seguir.

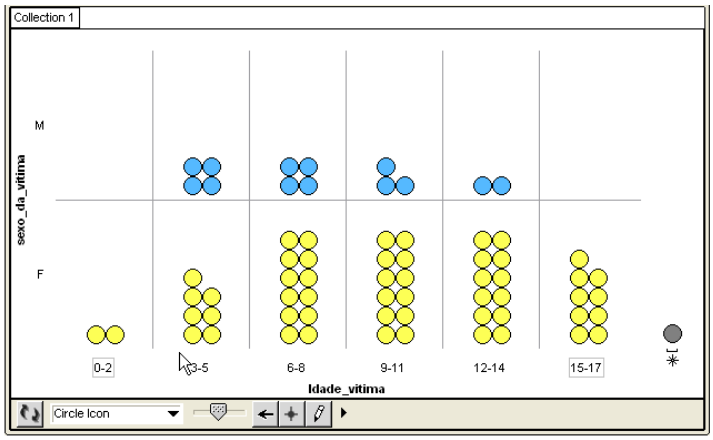


Figura 44: Exemplo 1 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados violência

Tivemos 8 professores que utilizaram representações dentro dessa tipificação para responder a primeira questão do banco de dados Violência. Nessa representação a variável idade aparece organizada em intervalos de três anos e separada em dois eixos verticais que representam o gênero masculino (azul) e feminino (amarelo). Além do intervalo de três anos, como visto na Figura 43, a variável idade também é organizada em intervalos de 4 e 5 anos. Essa forma de organizar uma variável quantitativa pode ter auxiliado os professores na interpretação dos dados, iremos discutir e analisar mais aprofundadamente essa questão no capítulo das *ferramentas* do *software*.

b) Segunda questão do Banco de dados Violência

A segunda questão desse banco de dados cruzava uma variável qualitativaXqualitativa. Os professores tiveram que responder “O estupro a meninas

é mais frequentemente realizado por qual agressor?", como se tratava de duas variáveis qualitativas eles teriam que cruzar os dados em categorias, o que foi bem utilizado por eles como vemos no Gráfico 6 da página 96. As representações dentro da tipificação D e E foram as mais utilizadas. Essas eram as possibilidades, mais evidentes, que o *software* oferecia para representar esse cruzamento.

Nessa questão que tratava do cruzamento de variáveis qualitativas as representações mais utilizadas foram as que continham características da tipificação D que cruza variáveis em categoria. Essa frequência já era esperada pela própria natureza das variáveis que oferecia a organização em categorias, no entanto, um artifício utilizado pelos professores foi o recurso *número*, que acabava fazendo com que a representação ficasse muito semelhante a uma tabela, como segue.

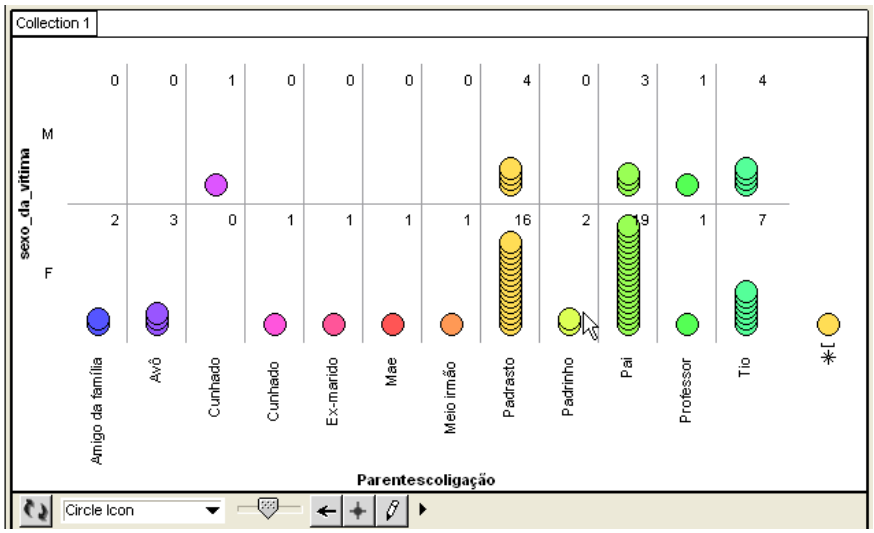


Figura 45: Exemplo 1 de representação dentro da tipificação E utilizada para responder a segunda questão do banco de dados violência

Nessa representação temos a associação das linhas e colunas e o valor numérico de cada intersecção das categorias, essa organização se assemelha a uma tabela. E por vezes, os professores apresentavam dúvidas na contagem, já que os *plots* estavam sobrepostos e a disposição deles não contribuía para uma possibilidade de contagem. Dessa maneira a apresentação dos valores numéricos pelo uso da ferramenta *número* foi importante.

Esse tipo de representação foi pouco utilizado e só requerido nessa questão específica. No entanto, quando a variável *Parentesco* estava no eixo vertical a tela

do *software* tinha mais espaço para empilhar os *plots* sem que existisse a sobreposição, como mostra Figura 46.

Figura 46: Exemplo 1 de representação dentro da tipificação D utilizada para responder a segunda questão do banco de dados violência

Esse tipo de representação da Figura 46 foi utilizado pela maioria dos professores. Como podemos observar temos o mesmo cruzamento de variáveis (sexo e parentesco) apresentado na Figura 45, entretanto, houve a inversão dos eixos, auxiliando, assim, na visualização dos *plots*. Esse resultado indica que o tamanho da tela do software pode se tornar limitado para uma quantidade grande de casos. Nesse estudo, trabalhamos com uma amostra de até 68 casos e já observamos algumas dificuldades de visualização para certas representações.

c) Terceira questão do Banco de dados Violência

A terceira questão tratava de um cruzamento de variáveis quantitativasXquantitativas, para responder a questão “Existe alguma relação entre a idade da vítima e a idade do agressor?”. Para observar a relação entre essas duas variáveis quantitativas, o gráfico mais utilizado é o de dispersão, que seria construído com o uso do *software* de maneira simples, arrastando o *plot* até o final da escala nos dois sentidos (vertical e horizontal). Com essa representação seria evidente perceber que a relação existente é muito fraca entre a idade da vítima e a idade do agressor, já que os *plots* se apresentam de forma dispersa, sem apresentar nenhuma tendência crescente ou decrescente.

Para essa questão os professores utilizaram com muita frequência o gráfico de dispersão (tipos F), pois esse tipo de gráfico é específico para os cruzamentos de

variáveis quantitativas (ver Figura 47). A construção desse tipo de representação foi desenvolvida pela maioria dos participantes a partir da interação com a pesquisadora no momento da interpretação. Então, 8 professores utilizaram esse tipo de gráfico, sendo que 3 deles de forma autônoma e 5 pela indicação da pesquisadora.

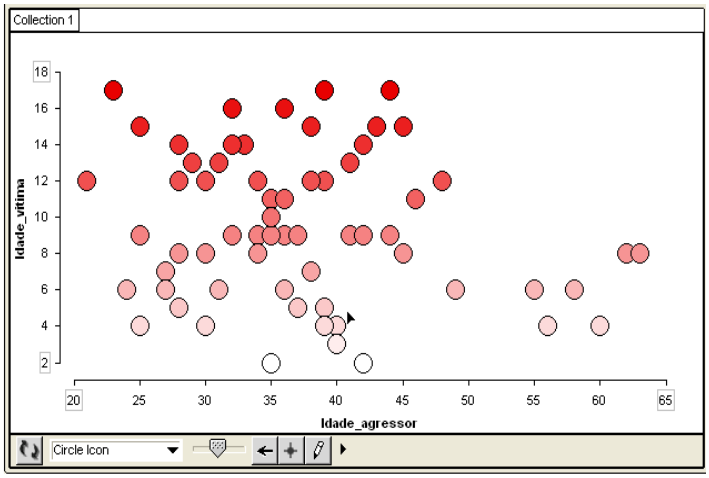


Figura 47: Exemplo 1 de representação dentro da tipificação F utilizada para responder a terceira questão do banco de dados violência

Outra forma de representar esse cruzamento para analisar os dados e chegar a resposta foi a partir da utilização de classes intervalares para as duas variáveis quantitativas, como segue na Figura 48.

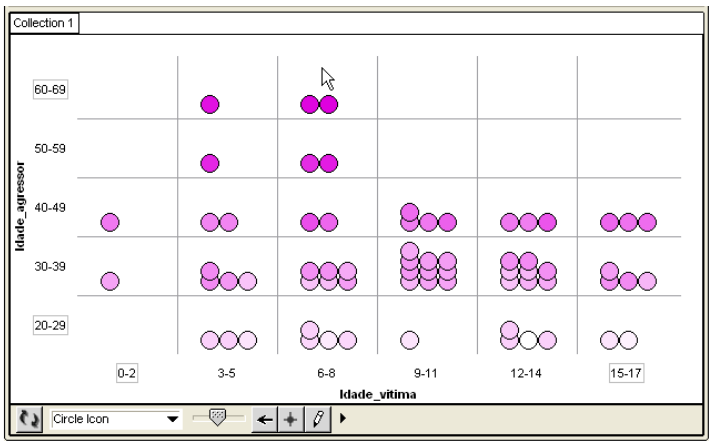


Figura 48: Exemplo 1 de representação dentro da tipificação D utilizada para responder a terceira questão do banco de dados violência

Todos os professores utilizaram a representação indicada na Figura 48, mesmo os que iniciaram utilizando o gráfico de dispersão, foi sugerida pela pesquisadora a mudança de classes pontuais para classes intervalares, como veremos com mais detalhes no capítulo sobre as ferramentas utilizadas. De uma maneira geral, as representações mais utilizadas pelos professores relacionavam as variáveis através de intervalos.

A seguir apresentaremos os dados produzidos a partir do segundo banco de dados que tratava de Novelas.

6.3. Segundo banco de dados – Novelas exibidas no horário nobre

Com o banco de dados sobre as Novelas exibidas no horário nobre de uma emissora de televisão os professores tiveram que responder, também, três questões de cruzamento de variáveis. A seguir o Gráfico 7 apresenta as 83 representações agrupadas nas 6 tipificações (apresentadas no Quadro 4) para cada questão associada a esse banco de dados.

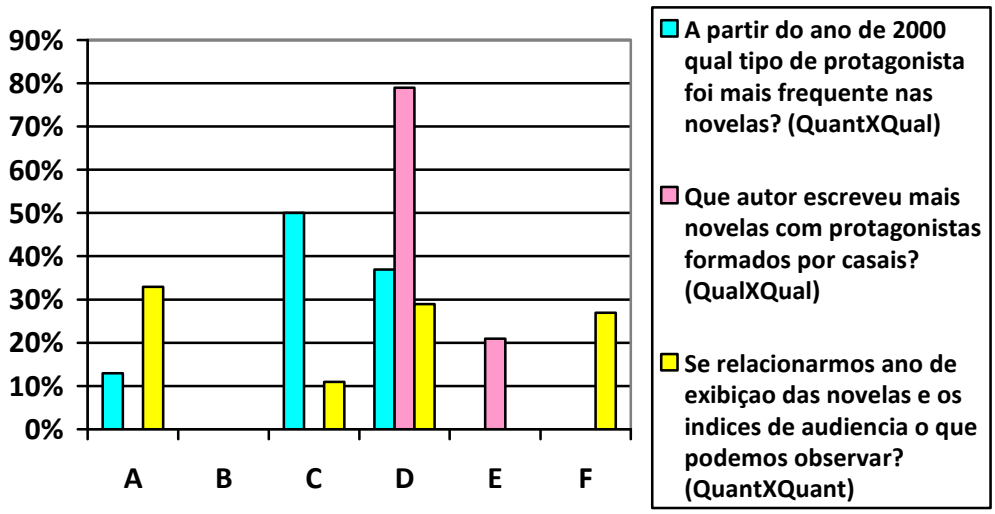


Gráfico 7: Representações construídas pelos professores de acordo com cada questão do segundo banco de dados - Novelas

No gráfico 7 podemos observar que as representações da tipificação C e D foram mais utilizadas nesse banco de dados, o que repete a mesma tendência com relação ao primeiro banco de dados. Essas representações estão ligadas ao uso de categorias e classes intervalares para cruzar as variáveis. As representações com

características dentro da tipificação B não foram utilizadas nesse banco de dados e segue a mesma tendência do primeiro banco de dados, no qual teve um pequeno percentual de uso. Iremos a seguir analisar mais aprofundadamente essas representações de acordo com as questões propostas.

a) Primeira questão do Banco de dados Novelas

Na primeira questão foi necessário realizar o cruzamento de variáveis quantitativas e qualitativas para responder a questão “A partir do ano de 2000 qual tipo de protagonista foi mais frequente nas Novelas?”. Para responder essa questão uma das possibilidades seria representar um eixo com classes pontuais para variável ano e, o outro eixo, com as categorias dos tipos de protagonistas das Novelas. Com isso, poderíamos identificar que o tipo de protagonista casal foi o mais utilizado do ano de 2000 até o ano 2010. As representações construídas por 6 professores relacionavam-se a tipificação C, conforme exemplo da Figura 49.

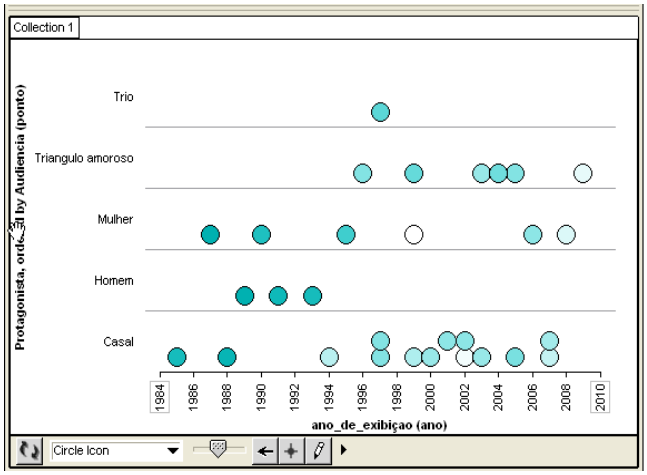


Figura 49: Exemplo 1 de representação dentro da tipificação C utilizada para responder a primeira questão do banco de dados Novelas

Nessa representação temos no eixo horizontal a variável ano organizada em classes pontuais e no eixo vertical as categorias relacionadas aos tipos de protagonista. Para responder a questão proposta seria necessário observar a freqüência maior entre os tipos de protagonistas a partir do ano de 2000.

Já na representação disposta na Figura 50 observamos o mesmo cruzamento de variáveis anos e protagonista, no entanto com uma organização diferenciada.

Esse tipo de representação apresenta características da tipificação D, no qual temos o cruzamento entre categorias e ou classes intervalares

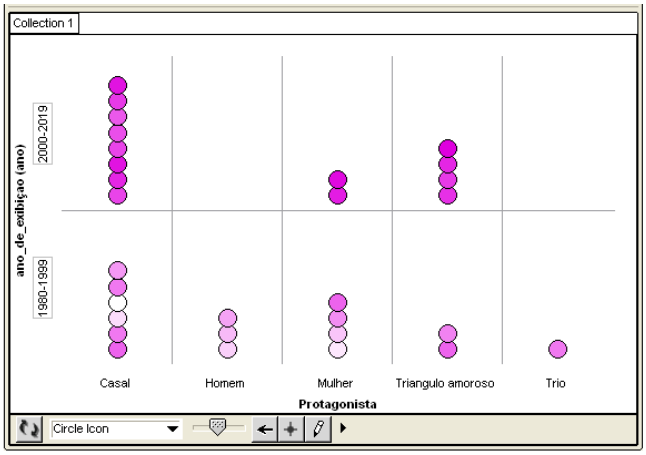


Figura 50: Exemplo 1 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados Novelas

Essa representação do tipo D obteve uma freqüência significativa, 4 dos 12 professores optaram por ela. Nessa representação os professores utilizam intervalos de 20 anos para organizar os anos no eixo vertical e categorias que representam os tipos de protagonistas. A resposta para essa representação está vinculada a frequência das categorias a partir do ano de 2000, o ponto máximo.

Mas, nem todos os professores utilizaram essa representação com a variável protagonista na horizontal e o ano de exibição na vertical. Assim como a professora Fabiola mais 3 professores utilizaram de forma invertida, como segue.

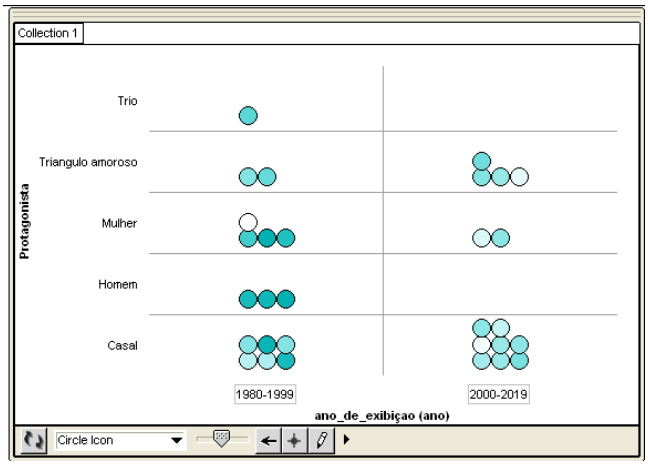


Figura 51: Exemplo 2 de representação dentro da tipificação D utilizada para responder a primeira questão do banco de dados Novelas

Nesse tipo de representação o ano é representado em intervalos de 10 anos no eixo horizontal e os tipos de protagonista na vertical.

Outra representação utilizada pelos professores para responder essa questão foi a apresentada na Figura 52 dentro da tipificação A na qual é representada uma variável em forma de cores e a outra em forma de classes pontuais, intervalares ou categorias.

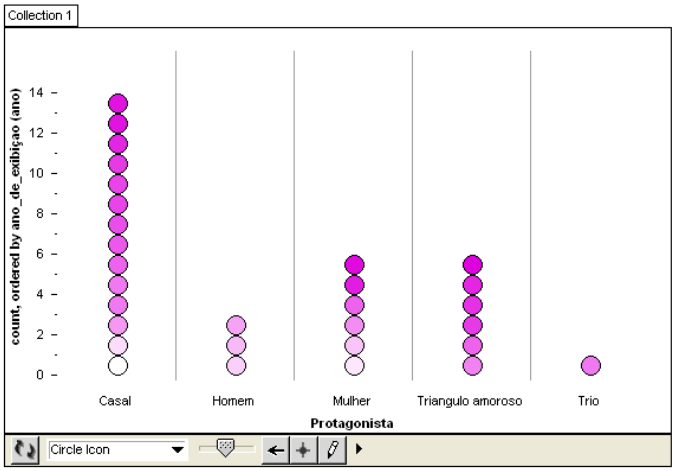


Figura 52: Exemplo 1 de representação dentro da tipificação A utilizada para responder a primeira questão do banco de dados Novelas

A representação demonstrada na Figura 52 foi construída por 2 professores, essa representação é semelhante a um gráfico de barras. Ela só apresenta uma variável (protagonista), a outra variável foi representada pela cor (os mais escuros anos maiores os mais claros anos menores). Para analisar esse gráfico temos que observar os tons de rosa escuro presentes nos *plots* em maior quantidade em relação aos protagonistas. Essa representação pode trazer equívocos no momento da interpretação, já que as cores escuras podem representar na sequência dos *plots* tanto anos próximos, como 2000, 2001, 2002; como também podem representar anos que se distanciam como 2000, 2008, 2010. Essa diferença nos valores numéricos não fica evidente através dos tons escuros da cor rosa dificultando a interpretação. A ferramenta *gradiente* que é responsável por essa função de atribuir cores para representar variáveis será melhor discutida e analisada no capítulo do uso das ferramentas.

b) Segunda questão do Banco de dados Novelas

A segunda questão requeria o cruzamento entre duas variáveis qualitativas para responder a questão “Que autor escreveu mais Novelas com protagonistas formados por casais?”. Como no banco de dados da violência sexual contra crianças, as possibilidades de representação para esse cruzamento estão dentro das tipificações D e E, que cruzam categorias. E como podemos perceber no Gráfico 7 foram os tipos mais utilizados.

As representações construídas que possuem características da tipificação D cruza duas categorias sem a utilização da quantidade em cada cruzamento, como segue no exemplo abaixo.

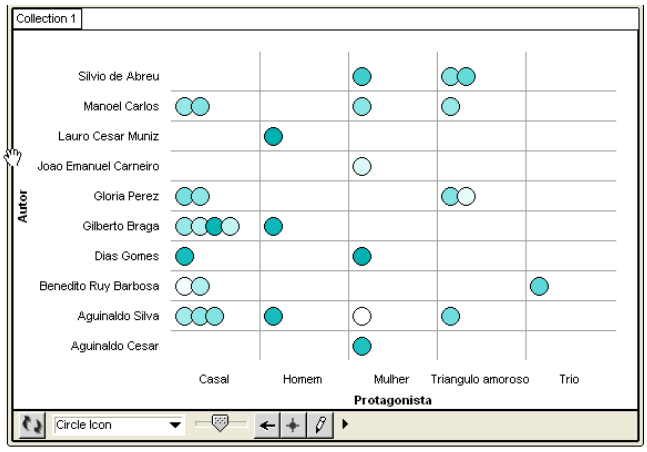


Figura 53: Exemplo 1 de representação dentro da tipificação D utilizada para responder a segunda questão do banco de dados Novelas

A representação exposta na Figura 53 foi utilizada por 9 professores para responder a questão que cruzava variáveis qualitativas. A utilização dessa representação segue a tendência do primeiro banco de dados, no qual os professores também construíram uma representação semelhante à Figura 53.

Também foi utilizada a ferramenta *número* para auxiliar na contagem dos *pots* nas categorias, como segue exemplo abaixo.

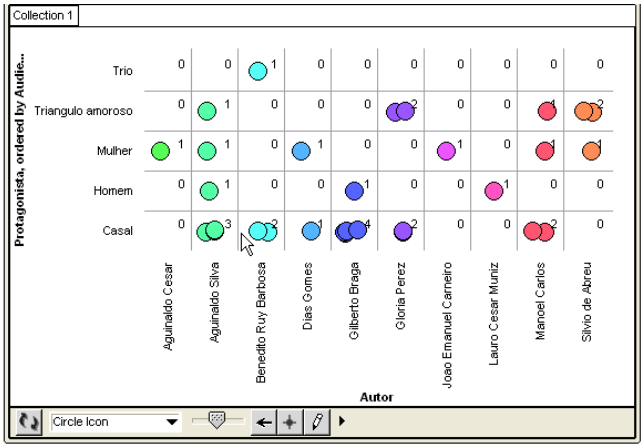


Figura 54: Exemplo 1 de representação dentro da tipificação E utilizada para responder a segunda questão do banco de dados Novelas

Na Figura 54 encontramos o mesmo cruzamento de variáveis apresentados na Figura 53, entretanto os eixos foram invertidos. Temos na Figura 54 os *autores* no eixo horizontal e protagonista no eixo vertical, com essa inversão o espaço da tela tornou-se pequeno para apresentar de forma horizontal uma variedade grande de autores, então alguns *plots* ficaram sobrepostos dificultando a contagem e foi necessário utilizar a ferramenta *número*. Essa mesma dificuldade de contagem dos *plots* foi identificado no primeiro banco de dados, nesse sentido o tamanho da tela do software pode limitar a quantidade de casos para esse tipo de representação.

c) Terceira questão do Banco de dados Novelas

A Terceira questão relacionava-se a um cruzamento de variáveis quantitativas, no qual, os professores tiveram que responder “Se relacionarmos ano de exibição das Novelas e os índices de audiência o que podemos observar?”. Nessa questão era esperado que os professores construíssem representações que cruzassem as variáveis ano de exibição e índice de audiência e observassem que a audiência diminui ao passar dos anos, nesse caso há uma relação inversa forte entre as variáveis, enquanto o ano aumenta o índice de audiência diminui. Essa relação é diferente da relação proposta no primeiro banco de dados, no qual a relação entre idade da vítima e idade do agressor era fraca.

Para essa questão os professores utilizaram representações dentro da tipificação A, D e F e, particularmente, eles utilizaram uma quantidade maior e uma diversidade de representações para responder a essa questão. Então, a

representação esperada para responder essa questão seria dentro da tipificação F, um gráfico de dispersão, no qual poderíamos observar a tendência dessa relação. O gráfico de dispersão foi construído por 10 dos professores, no entanto apenas 3 construíram esse gráfico como primeira opção de representação, enquanto os outros iniciaram construindo outras representação e finalizaram com a construção desse gráfico, como apresenta a Figura 55.

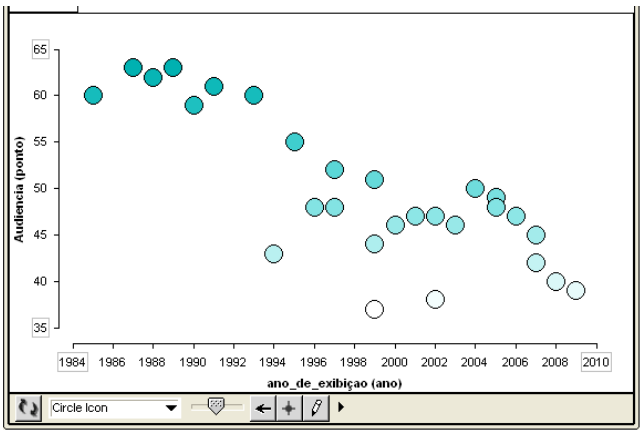


Figura 55: Exemplo 1 de representação dentro da tipificação F utilizada para responder a terceira questão do banco de dados Novelas

Uma grande maioria dos professores construiu esse gráfico e podemos observar que a disposição dos *plots* estão evidentes, com um formato de linearidade decrescente. No qual quanto maior o tempo menor a audiência.

Outra representação utilizada pelos professores para auxiliar suas interpretações foi a representação na tipificação A.

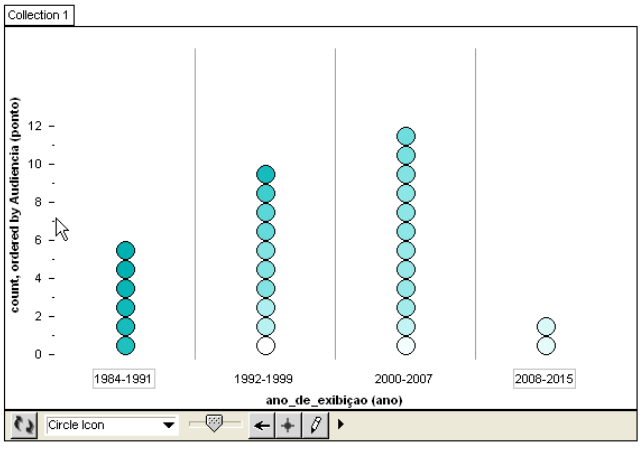


Figura 56: Exemplo 1 de representação dentro da tipificação A utilizada para responder a terceira questão do banco de dados Novelas

Nessa representação temos o ano em forma de classes intervalares, em intervalos de 8 anos, e as cores que indicam a audiência nesses anos, então quanto mais escuro maior é a audiência. No caso específico dessa representação demonstrada na Figura 56 podemos perceber claramente a tendência das cores que inicia no sentido da esquerda para direita azul escuro e termina azul mais claro. Porém as cores não oferecem valores específicos do índice de audiência em cada categoria, elas servem de indicativo apenas.

A representação dentro da tipologia C também foi utilizada. A característica dessa representação é cruzar uma variável quantitativa em forma de classes pontuais e a outra em forma de categoriais ou intervalos, como segue exemplo abaixo.

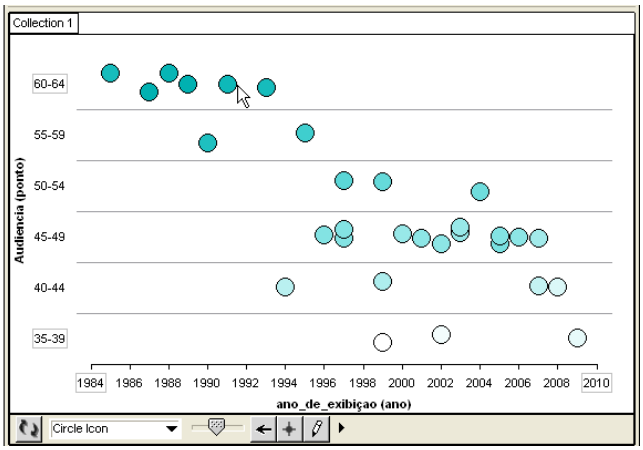


Figura 57: Exemplo 1 de representação dentro da tipificação C utilizada para responder a terceira questão do banco de dados Novelas

Na representação da Figura 57 os anos de exibição das Novelas estão em escala e os índices de audiência categorizados em intervalos de 4 pontos de audiência. Essa representação foi utilizada por 5 professores.

A representação do tipo D também foi utilizada, ela tem como característica o cruzamento de variáveis a partir de intervalos. Esse tipo foi bem utilizado no primeiro banco de dados na tentativa de melhorar o entendimento da relação entre as idades da vítima e do agressor. Como exemplo a Figura 58.

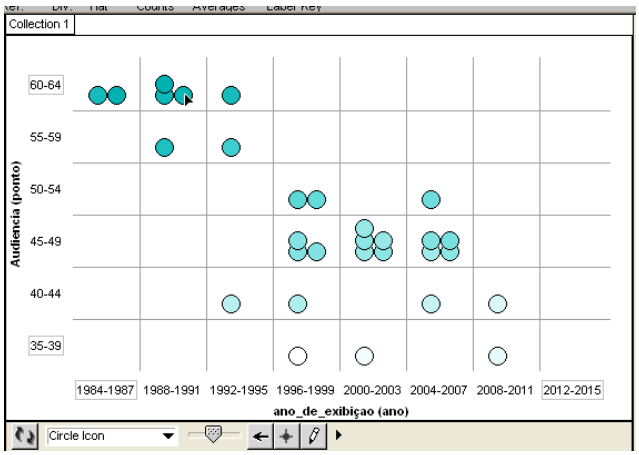


Figura 58: Exemplo 1 de representação dentro da tipificação D utilizada para responder a terceira questão do banco de dados Novelas

Nessa representação os anos e os índices de audiência são apresentados em classes intervalares, no qual podemos observar tanto a cor que indica o índice de audiência, como as intersecções entre as categorias dos anos e dos índices de Audiência.

De uma maneira geral podemos perceber que os professores não encontraram dificuldades em construir diferentes representações, mesmo em se tratando de um contexto bivariado, o que é mais complexo do que construir representações com apenas uma variável. Até mesmo utilizando gráficos de dispersão que não é muito frequente nos livros didáticos, como indica Pratt (2008).

Nesse caso, podemos atribuir ao *software TinkerPlots* um *layout* simples e suas ferramentas acessíveis para professores com pouca experiência com o uso de computadores. Dentre as representações construídas podemos apontar como as maiores utilizações as que apresentavam as variáveis em categorias e ou intervalos. O uso das cores também despertou interesse nos professores, já que tivemos uma grande utilização de representações de cruzavam duas variáveis, sendo uma delas através das cores.

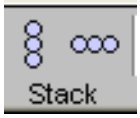
No entanto, ainda não podemos apontar a influência que cada representação ofereceu a interpretação, esta questão será analisada a partir do uso das ferramentas, já que o seu uso está ligado a construção das representações, o que iremos tratar a seguir.

CAPÍTULO 7 – ANALISANDO O USO DAS FERRAMENTAS DISPONÍVEIS NO SOFTWARE E SUA INFLUÊNCIA PARA INTERPRETAÇÃO DE DADOS

Os professores ao construírem as representações utilizaram as ferramentas *separar*, *empilhar*, *ordenar* e *gradiente* para suas análises. A pesquisadora no decorrer da realização das atividades propostas sugeria/intervinha com o uso de algumas ferramentas de forma intencional buscando promover a reflexão no momento da interpretação. Nos momentos em que os professores apresentavam dúvidas ou equívocos na interpretação dos dados, a pesquisadora intervinha sugerindo alguma ferramenta (*separar*, *empilhar*, *ordenar* e *gradiente*) que pudesse contribuir para o entendimento dos dados ou para a percepção de informações contidas nos dados, antes não vistas. Desse modo, procuramos identificar quais ferramentas foram utilizadas para construção das representações e o que elas influenciaram na análise de dados para o alcance das respostas das questões propostas.

No quadro abaixo destacamos as quatro ferramentas escolhidas e utilizadas na pesquisa, já que, o *software* oferece outras ferramentas.

Quadro 5: Ferramentas disponíveis no *Software* e utilizadas pelos professores

Ferramenta	Funções
Separar	- Sua função é organizar os dados em categorias ou escala. - Ao clicar direto no ícone o *software* constrói categorias utilizando critérios próprios. - Ao clicar em um *plot* e arrastar para o lado direito podemos escolher quantas categorias ou intervalos queremos formar. Essa ação também se dá de forma inversa quando queremos diminuir a quantidade de categorias ou intervalos basta arrastar para o lado esquerdo.
Ordenar	- Sua função é ordenar os *plots* de forma crescente de acordo com a categoria e a variável. - A ordenação acontece tanto na vertical como também na horizontal.
 *Empilhar*	- Sua função é empilhar os *plots* em forma de barras; - O empilhar pode ocorrer em forma de barras na vertical ou na horizontal.
Gradiente	- Sua função é estabelecer cores para indicar o comportamento da variável; - Quando a variável é quantitativa as cores são definidas em forma de dégradé, no qual, os *plots* mais claros representam os valores menores da variável assim, quanto mais o valor aumenta mais o tom do *plot* escurece; - Quando a variável é qualitativa as cores aparecem de acordo com as categorias, para cada categoria diferente uma cor diferente.

O Quadro 5 fornece na primeira coluna o nome da ferramenta utilizada e seu respectivo ícone exposto na tela do software. Na segunda coluna temos a descrição das funções correspondentes a cada ferramenta.

Como forma de analisar o uso das ferramentas nas atividades propostas dos dois bancos de dados, iniciamos destacando a quantidade de vezes que cada ferramenta foi requerida/utilizada. O Gráfico 8 apresenta o percentual de uso de cada ferramenta.

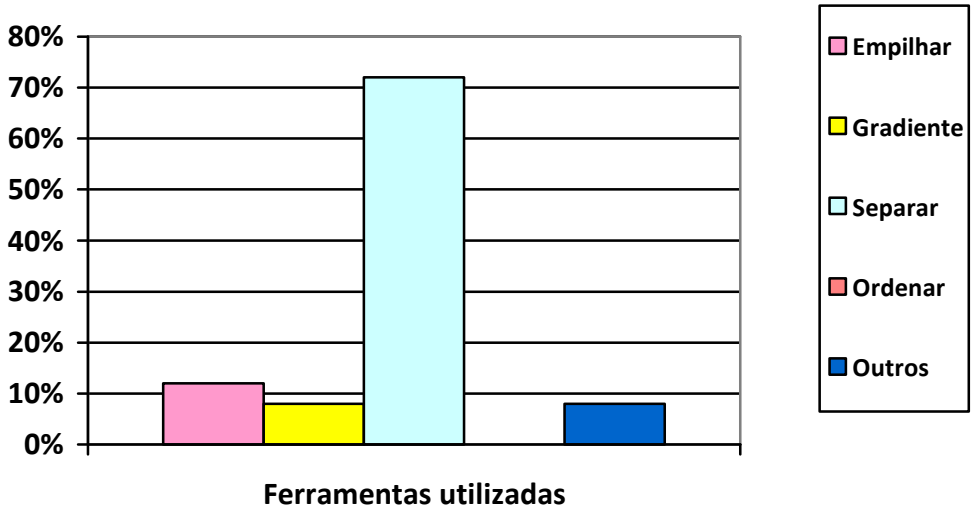


Gráfico 8: Percentual de uso de cada ferramenta utilizada para construir as representações

Foram ao todo 59 utilizações/sugestões de ferramentas no decorrer de cada questão de cruzamento de variáveis. De acordo com o Gráfico 8 a ferramenta mais utilizada foi a *separar*. 70% de utilização de ferramentas estava relacionada a manipulação de escalas (função da ferramenta *separar*). No entanto, essa ação pode estar associada a organizar os dados em intervalos ou categorias que estão dispersos na tela, como também de diminuir os intervalos utilizados para as variáveis; acreditamos que essas duas maneiras de uso precisam ser analisadas de maneira dissociada, pois elas podem promover diferentes influências na interpretação. Com isso verificamos que dos 70% de uso da ferramenta *separar* 52% foi para organizar os dados em classes pontuais, intervalos ou categorias e 18% esteve associado a diminuir os intervalos utilizados para representar uma variável. As ferramentas *empilhar* e *gradiente* foram utilizados com menor frequência no momento da interpretação. E a ferramenta *ordenar* não foi utilizada no momento da interpretação.

A categoria outros está incluindo o uso de duas ferramentas *número* e *linha de referência* que foram utilizadas de forma casual, pela necessidade dos professores no momento da visualização de algumas representações. Para entendermos o uso dessas ferramentas vamos, a partir de agora, identificar o uso dessas ferramentas em cada questão de cruzamento e sua influência na resposta dos professores.

7.1. Banco de dados Violência contra crianças – Ferramentas utilizadas

Nas três questões propostas para esse banco de dados os professores manipulavam os *plots* na tentativa de construir uma representação de auxiliasse na obtenção da resposta. Para construção essas representações os professores utilizavam as ferramentas disponíveis no *Software* ou a pesquisadora por vezes sugeria a utilização de alguma ferramenta que contribuísse na interpretação.

Para o primeiro banco de dados tivemos utilização as ferramentas expostas no Gráfico 9. Nas categorias utilizadas pelo Gráfico 9 dissociamos a ferramenta *separar* nas suas duas funções, construir escalas e reduzir intervalos, como discutido anteriormente.

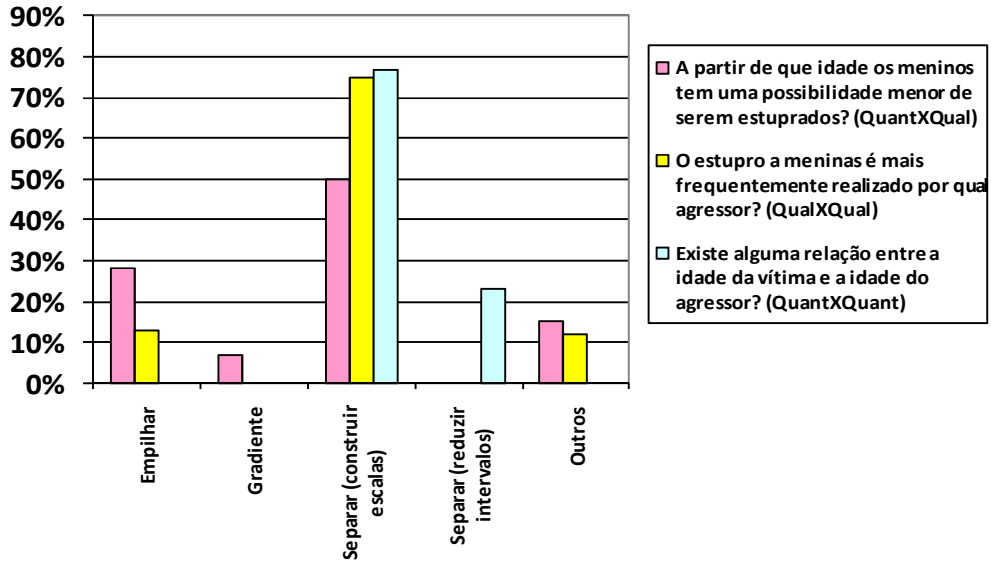


Gráfico 9: Ferramentas utilizadas para responder as questões do banco de dados Violência

O Gráfico 9 apresenta na legenda as três questões de cruzamento de variáveis, no qual estão indicadas pelas cores. No eixo horizontal encontramos as

categorias que se referem ao uso das ferramentas empilhar, *gradiente*, separar (criar categorias ou escala), separar (reduzir intervalos) e outros. Como explicado anteriormente a ferramenta *separar* foi dissociada a partir de suas funções. Já a ferramenta *ordenar* não foi utilizada como intervenção no momento da interpretação, nesse sentido não aparece no gráfico.

A seguir iremos analisar a partir de cada questão proposta.

a) *Primeira questão do Banco de dados Violência*

Para a primeira questão encontramos a ferramenta *empilhar* e *separar* (criar categorias ou escalas) como as mais utilizadas, apresentadas no Gráfico 9. Em seguida temos as *gradiente* e outros. A categoria outros está englobando algumas ferramentas disponíveis no software, porém que não foram foco deste estudo. A categoria *separar* (reduzir intervalos) não foi utilizada para responder a essa questão.

Para entendermos melhor a utilização dessas ferramentas apresentaremos exemplos de trechos de falas de professores analisando a passagem de uma representação para outra com o uso da ferramenta especificada. Por exemplo, como estava o estado inicial da tela do software antes da utilização da ferramenta e após a utilização da ferramenta e o que influenciou na interpretação do professor.

Em relação a ferramenta *empilhar* tivemos em torno de 30% de uso. O trecho a seguir mostra fala da professora Poliana com o uso dessa ferramenta.

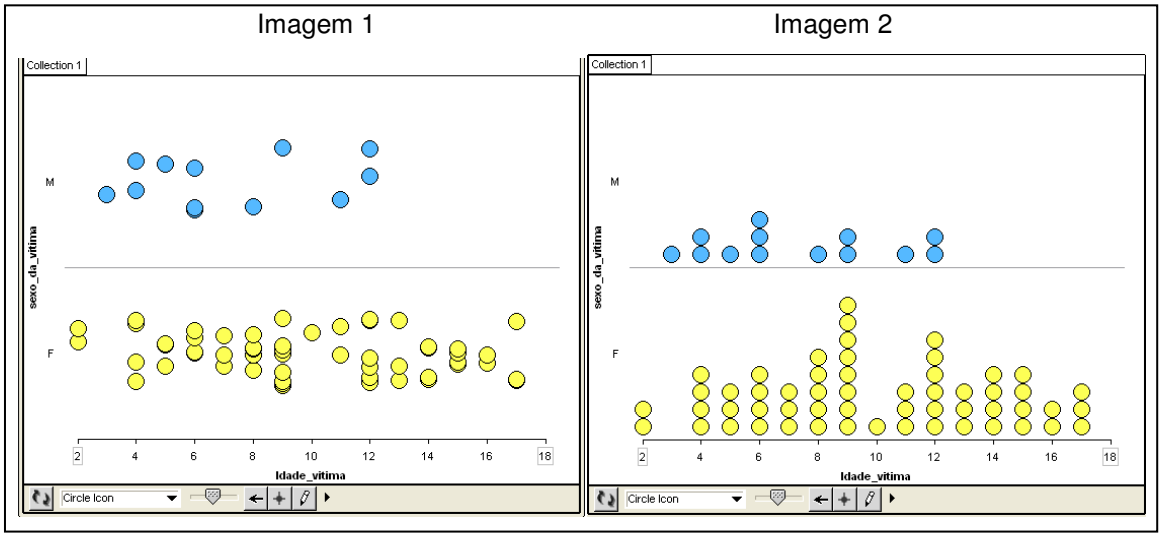


Figura 59: Exemplo 1 de uso de a ferramenta *empilhar* para responder a primeira questão do banco de dados Violência

Na Imagem 1 da Figura 59 podemos observar a representação construída pela professora Poliana no cruzamento entre idade e gênero da vítima. Na Imagem 2 temos a mesma representação construída anteriormente, no entanto, com a utilização da ferramenta empilhar. Abaixo segue a fala da professora.

> Pesquisadora: Os azuis são os meninos e os amarelos as meninas. E agora? A partir de que idade os meninos deixam de serem estuprados?
> Poliana: Hummm [aponta com o *mouse* para a escala (idade) e relaciona com o gênero masculino] deixa eu ver.... [pensa por 3 min]
> Pesquisadora: Se você quiser trazer as bolinhas mais pra perto da linha é só apertar nos empilhadinhos. [refere-se a ferramenta empilhar]
> [Professora faz da forma que a pesquisadora orienta]
> Pesquisadora: Pronto. A partir de que idade os meninos deixam de ser estuprados?
> Poliana: Ah! não sei.... porque se repete quatro vezes né! A idade que é mais é com 6 anos.
> Pesquisadora: Certo.
> Poliana: Aí, aqui poderia ser.... Entre 3 anos, 5, 8 e 11 são as idades que se iguala pela quantidade.
> Pesquisadora: Sim. Mas, assim, a partir de que idade os meninos tendem a deixar de ser estuprados?
> Poliana: A partir dos 13 anos, porque não tem mais.
> Pesquisadora: Não tem mais? Por quê?
> Poliana: Porque não tem mais nenhuma quantidade de vitima aqui! A partir dos 13 anos. Só tem até o 12.

A professora Poliana através da sua fala demonstra entender as variáveis e a correspondência entre elas. Também observa a escala que representam as idades das vítimas e mesmo os valores ocultos da escala ela consegue identificar as idades de 3, 5 e 11 anos. Antes da utilização da ferramenta empilhar ela demonstra dificuldade em relacionar os *plots* com a escala, talvez por estarem distantes do eixo e alguns *plots* sobrepostos. A pesquisadora sugere a utilização do empilhar e logo em seguida a professora consegue emitir suas primeiras análises sobre os dados.

A fala da professora, ainda pode nos sugerir que ela identifica a idade, entre os meninos, que mais ocorre estupro, no caso 6 anos e as idades que possuem a mesma quantidade de vítimas. Essa análise deixa de ser pontual, passando a entender o comportamento da variável e a identificar a escala a partir do cruzamento entre a idade e o gênero.

Em diversos momentos a ferramenta *empilhar* foi utilizada com o intuito de empilhar os *plots* para que deixassem de estar sobrepostos oferecendo uma melhor

visualização dos dados. Essa ferramenta foi frequentemente utilizada associada a ferramenta separar, como segue na figura abaixo.

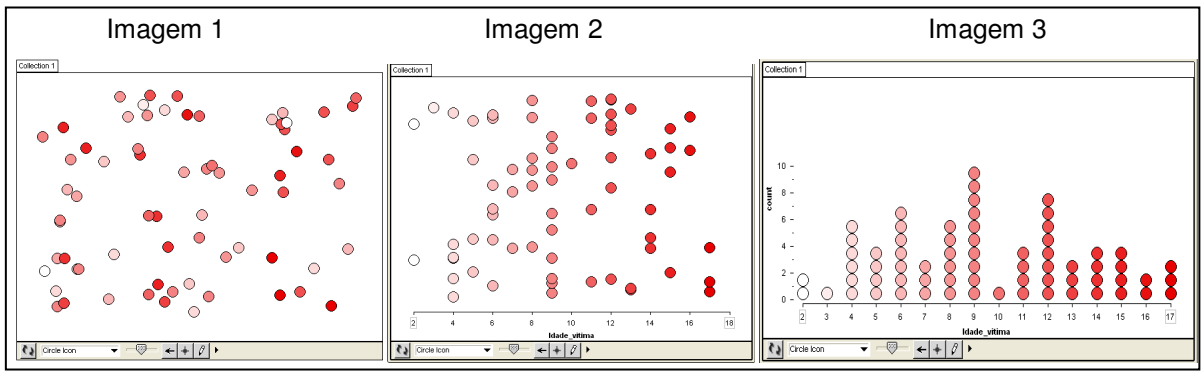


Figura 60: Exemplo 2 de uso de a ferramenta *empilhar* associada a *separar* para responder a primeira questão do banco de dados Violência

Na Figura 60 podemos observar na Imagem 1 os *plots* representam a variável idade da vítima a través das cores o vermelho mais claro representa as idades menores quanto mais escuro o vermelho se torna representa as idade maiores. Na imagem 2 apresenta a ação realizada pela professora de clicar em um *plot* e arrastar utilizando, assim, a ferramenta *separar* (construir escala), e, por último na Imagem 3 a professora utiliza a ferramenta *empilhar* para organizar os *plots* próximos ao eixo e dispô-las em forma de barras. Para essas passagens a professora Nara emite alguns comentários sobre os dados, a partir da Figura 60 como segue.

[olhando a Imagem 1]
Pesquisadora: Hum... ai temos idade da vitima, certo?
Nara: Sim.
P: Você pode dizer alguma coisa sobre a idade deles?
N: Os mais escuros são maiores, né?
P: São.
N: Tem mais claros que escuros.
P: Como assim?
N: Acho que tem crianças mais novas né!
P: Vamos colocar em escala, então?
[utiliza a ferramenta separar] [olhando a Imagem 2]
P: E agora. Que idade as crianças são mais estupradas?
N: [risos] não dá pra ver.
P: Hummm, será que se apertar em empilhar melhora?
[professora clica em empilhar] [olhando a Imagem 3]
N: Muito! 9 anos
P: E a quantidade de criança com maior idade que são estupradas?
N: 17. Oh! 3 crianças com 17 anos é isso?

A fala da professora Nara demonstra a importância do uso dessas duas ferramentas, *separar* e *empilhar,* isso pode ser constatado pela certa facilidade em responder os questionamentos realizados pela pesquisadora. Principalmente quando utilizada as duas ferramentas juntas, como na Imagem 3 da Figura 60. Essa

facilidade pode estar associada a semelhança dessa representação ao gráfico de barras, no entanto, ela consegue não só identificar os pontos máximos, mas consegue também perceber a escala utilizada para apresentar a variável idade e utilizá-la como resposta.

A ferramenta *separar* foi dissociada de acordo com suas funções, como apontado anteriormente. Ela pode ser utilizada como no exemplo da Figura 60, construindo escalas para a variável, ou pode ser utilizada reduzindo intervalos já existentes. A utilização dessa ferramenta para construção de escalas foi muito presente, como em outro exemplo na Figura 61.

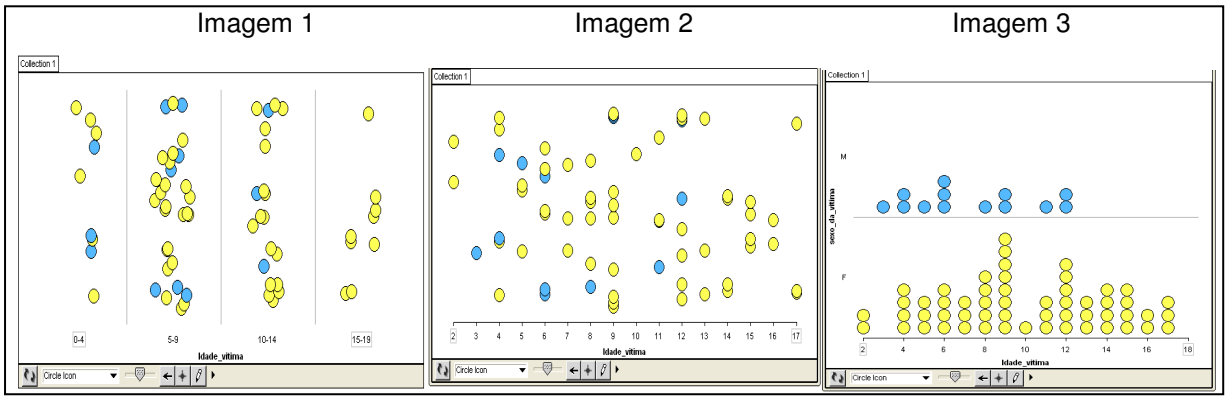


Figura 61: Exemplo 1 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Violência

Para esse caminho utilizado pela Professora Fabíola foi necessário inicialmente utilizar a ferramenta *separar* para criar intervalos (clicar em um *plot* e arrastar com o *mouse*) e em seguida clicar em gênero para que a cor pudesse representar meninos e meninas, presente na Imagem 1 da Figura 61. Em seguida ela clicou em um *plot* e arrastou com uma distância maior, então conseguiu dispor a idade em uma escala organizada em classes pontuais (ainda utilizando a ferramenta *separar*), como mostra a Imagem 2. E por último ela utiliza novamente a ferramenta *separar*, agora para variável gênero, dividindo em categorias menino e menina e, ainda, empilha os *plots,* como mostra Imagem 3. Nesse percurso ela emite algumas reflexões como mostra trecho abaixo.

> [olhando a Imagem 1]
> Pesquisadora: Aí já dar para dizer alguma coisa?.
> Fabíola: Já.
> P: Será que dá pra dizer qual a idade do grupo mais velho de crianças são estuprados?
> F: De 15 à 19.
> P: São os mais velhos?

F: Os mais velhos.
P: E a quantidade de criança maior?
F: A quantidade maior, é de 5 à 9 anos.
P: Mas o que nós queremos saber é com que idade os meninos tem menor possibilidade de serem estuprados?
[a professora clica e arrasta]
[olhando a Imagem 2]
P: O que você fez?
F: Hum... [risos]. Acho que deixei tudo esticadinho.
P: E já dá pra responder assim?
F: Não!
P: E agora?
F: Aí vem pro sexo né?
P: Isso ai você puxa agora, clica em qualquer um e puxa para cima, pra cima.
[olhando a Imagem 3]
P: E agora já dá pra dizer. Olha... se você quiser organizar as bolinhas é só apertar nos empilhar.
F: É... não sei se tá certo. Acho que é 12 mais ou menos.
P: Por quê?
F: Aqui pela quantidade que tem.
P: Hammm. Quantidade?
F: É. Aqui não tem mais nenhum, já as meninas coitadas dá pena mulé.
P: Por quê?
F: Oxe,vai até os 17.

A professora Fabíola inicia uma análise apenas da variável idade, que se encontra disposta em intervalos como na Imagem 1 da Figura 61. Ela identifica a maior idade das crianças que foram violentadas, como também identifica a idade com maior incidência de casos de estupro. Na sequência, com o auxílio da pesquisadora, ela constrói a representação cruzando as variáveis com o uso da ferramenta *separar*. Então ela consegue identificar a idade em que os meninos tendem a deixar de serem estuprados pela ausência de *plots* na idade de 13 anos e ainda estabelece uma comparação entre o grupo das meninas identificando até que idade as meninas são estupradas. Esse percurso permitiu a professora acompanhar o caminho em que o uso da ferramenta auxiliou no entendimento dos dados. Ela acompanhou a construção das escalas para a idade e separou o gênero em categorias, essas ações podem ter contribuído para a análise desses dados pela professora.

Além da utilização de classes pontuais para responder a essa questão os professores utilizaram os intervalos para representar a variável idade, como segue.

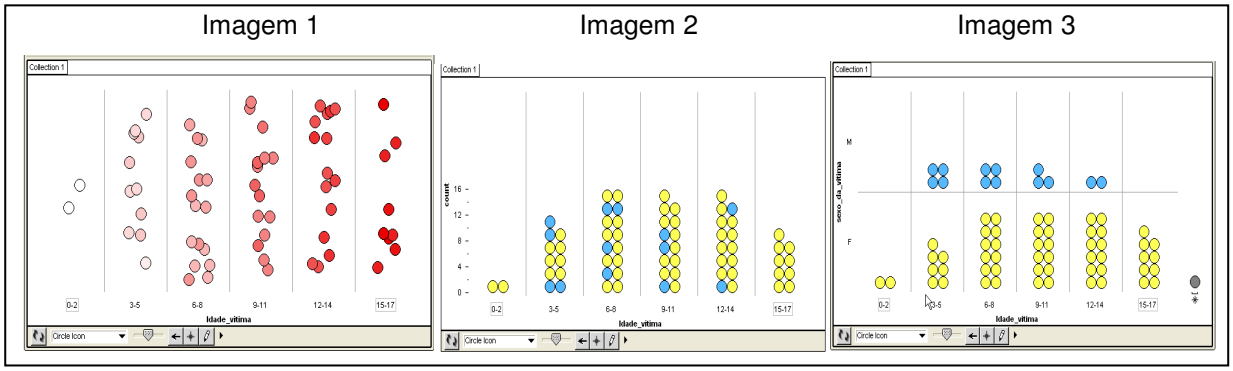


Figura 62: Exemplo 2 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Violência

No percurso demonstrado na Figura 62 foi necessário selecionar a variável idade e utilizar a ferramenta separar para criar intervalos, como na Imagem 1. Em seguida, na Imagem 2, foi selecionada a variável gênero e por último a professora utilizou a ferramenta *separar* para construir categorias da variável gênero na vertical. Com esse percurso a professora pôde refletir sobre a variável idade e buscou entender a relação existente entre as variáveis a partir da imagem 3 da Figura 62.

> [olhando a Imagem 1]
> Pesquisadora: A partir de que idade os meninos deixam de ser mais estuprados?
> Nara: Ainda não dá pra ver...
> P: Por que?
> N: Tá misturado.
> P: Então clica em gênero e em empilhar.
> [olhando a Imagem 2]
> N: Vamo ver... dá pra separar né?!
> P: dá.
> [olhando a Imagem 3]
> N: Hum... Menor de 12 à 14.
> P: O quê? Não entendi.
> N: Porque o índice é pequeno né! No caso, o número de vítimas é menor de 12 a 14. Dois e... vai diminuindo. E vai diminuindo.

A professora Nara utiliza a ferramenta *separar*, mas nesse caso constrói intervalos e categorias. As representações construídas na Figura 62 assemelham-se a o gráfico de barras. Isso se deve ao empilhamento dos *plots* dispostos em categorias ou intervalos separados por uma reta. Com a representação disposta na imagem 3 a professora pode responder que a partir dos 12 e aos 14 anos os casos de violência foram diminuindo.

É possível que a participação do professor nas etapas da formação dos intervalos das idades pelo *software* possa ter contribuído para que 8 dos 12 professores apresentassem resultados semelhantes aos de Nara. O uso dessa ferramenta pode ter despertado o olhar dos professores para as idades e com isso obtivemos mais respostas baseadas nos intervalos e não no ponto máximo.

Nem todos os professores acompanharam a formação dos intervalos. A ferramenta *separar* do *software* é acionada quando clicamos em um *plot* e arrastamos com o *mouse*, a partir desse arrastar o sujeito pode acompanhar a formação dos intervalos ou categorias(quando se trata de variável qualitativa), quanto mais arrastamos mais separamos (formamos mais intervalos). No entanto, alguns professores não apresentaram habilidades na utilização do *mouse*, nesses casos os *plots* eram arrastados de forma rápida, sem a preocupação de observar a formação dos intervalos essa ação resultava na construção automática da escala. Era como se a escala surgisse sem que eles dessem conta, já que a preocupação era no arrastar do *mouse*. Essa estratégia de resolução foi observada entre 4 professores, como exemplo apresentamos trecho da Professora Fábia.

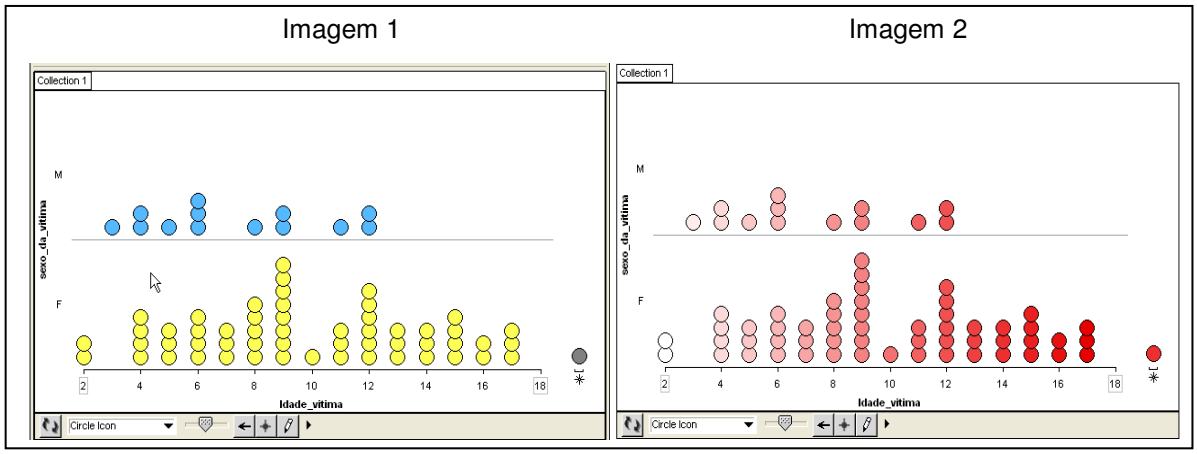


Figura 63: Exemplo 2 de uso de a ferramenta *separar* associada a ferramenta *Gradiente* para responder a primeira questão do banco de dados Violência

Para construir essa representação a professora inicialmente seleciona a variável idade e utiliza *separar*, no entanto como tem pouca familiaridade em manipular o *mouse* ela clica em um *plot* e arrasta de forma rápida sem observar a tela do computador. Nesse momento surge a escala, que causa espanto a professora.

Fabia: Ui! [risos] é assim mesmo é?
Pesquisadora: Como assim?
F: Fica assim mesmo? [se refere a escala formada com a variável idade]
P: Sim. É uma escala.

Após a construção da escala ela seleciona gênero e utiliza separar novamente para construir a escala no eixo vertical. Nesse momento temos idade da vítima na horizontal e gênero na vertical, mas ainda temos a cor que ressalta o gênero (amarelo meninas e azul meninos) como podemos observar na Imagem 2 da Figura 63. A partir dessa representação a professora inicia suas análises.

[olhando a Imagem 1]
Fabia: Ah... olha! Tem mais meninas que meninos estuprados né! Os meninos né o azul?
Pesquisador: É sim.
F: Então.... deixam de ser a partir dos 6 anos.
P: O quê?
F: Deixam de ser estuprados?
P: Humm.
F: É a partir dos 6 anos, assim, eu tô tendo a base pelo número de bolinhas azuis.
P: Certo. E se você clicasse em idade, será que muda alguma coisa.
[a professora clica em idade] [olhando a Imagem 2]
P: Assim ajuda alguma coisa?
F: Bem... mudou a cor....e dá... pela cor... a gente ... é o número maior de idade é de 12 anos de meninos.
P: Certo.
F: Então os mais escuros. São os mais velhos...
P: E para responder a pergunta, é só para bater o martelo. Em que idade os meninos apresentam menor possibilidades de serem estuprados?
F: Eu acho que é 12 anos.
P: 12?
F: 12 anos.
P: Por quê?
F: Pela informação que eles obtém.

Através da fala da professora Fábia podemos perceber que ela entendeu que existiam duas variáveis sendo cruzadas e que uma delas estava sendo representada pela cor, no caso amarelo meninas e azul meninos. No entanto, sua resposta foi baseada no maior número de *plots* de cor azul, que representava os meninos, no caso o ponto máximo. Essa primeira reflexão pode ser ligada a semelhança dessa representação ao gráfico de barras e a tendência em responder a partir do ponto máximo. Também entendemos que a formação da escala não foi uma opção da professora e sim uma construção aleatória sem intenção, já que o arrastar do *mouse* se tornou um desafio para ela.

No entanto, a pesquisadora percebendo que o olhar da professora estava baseado no ponto máximo, indicou a utilização da ferramenta *gradiente*, nesse momento, a professora teve o seu olhar atraído pelo *gradiente* de cores representando a idade dos meninos. Ela sabia que os mais claros eram os mais novos e os mais velhos os mais escuros, então, por conseqüência ela afirmou que a partir dos 12 anos os meninos tendiam a ser menos estuprados, no entanto, não ficou clara sua compreensão do gráfico de maneira global, já que sua justificativa foi baseada em conhecimentos da sua realidade de vida. Mas podemos inferir que a ferramenta *gradiente* atraiu o olhar da professora para a escala, que anteriormente havia sido desconsiderada.

b) Segunda questão do Banco de dados Violência

Na segunda questão do banco de dados sobre violência os professores tinham que responder "O estupro a meninas é mais frequentemente realizado por qual agressor?". De acordo com o Gráfico 9 as ferramentas mais utilizadas para essa questão foram a *empilhar* e *separar*. Por se tratar de um cruzamento de variáveis qualitativas o uso de categorias esteve muito presente. De uma forma geral essa questão apresentou um pequeno índice de intervenção por parte da pesquisadora e os professores tenderam a responder satisfatoriamente a questão proposta, como segue o exemplo da professora Sandra.

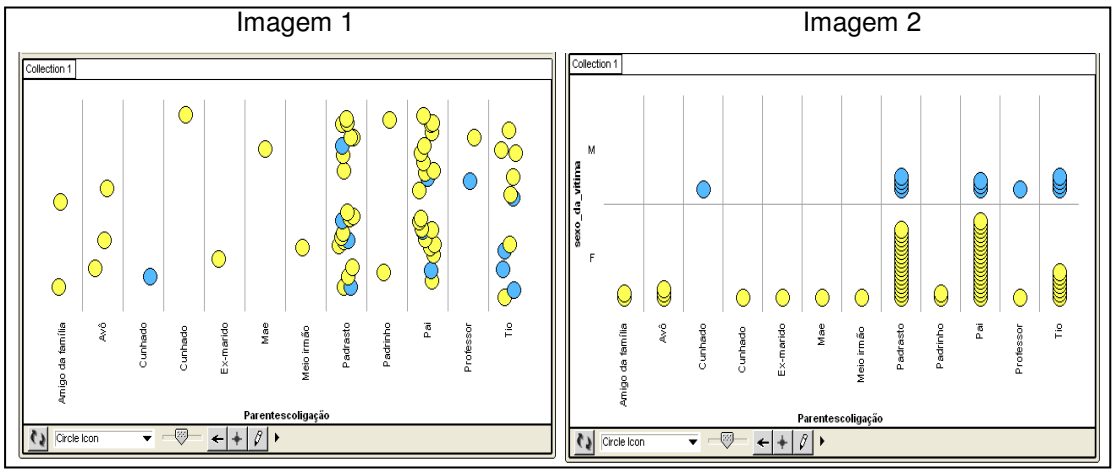


Figura 64: Exemplo 1 de uso de a ferramenta *separar* associada a *empilhar* para responder a segunda questão do banco de dados Violência

Para essa representação a professora inicialmente selecionou a variável parentesco da vítima e separou em categorias, como na Imagem 1. Em seguida, utilizou a ferramenta separar para a categoria gênero, empilhando por último, como mostra a Imagem 2 da Figura 64.

> Pesquisadora: Segunda questão. O estupro a meninas é mais frequentemente realizado por qual agressor?
> [a professora clica em parentesco e separa as categorias] [Imagem 1]
> Sandra: Pai ou padrasto... eu acho que é padrasto.
> P: Já dá pra dizer?
> S: Posso separar?
> P: Sim.
> [utiliza a ferramenta *separar* para a variável gênero] [Imagem 2]
> S: O pai, meu Deus do céu. Por quê?
> P: Hummm.
> S: Isso.... [demonstra surpresa na expressão do olhar] é.... que pelo amor de Deus. O estupro de meninas é mais frequentemente realizado por quem? [lê novamente a questão na tela do *software*] Pelo pai!
> P: Por quê?
> S: Eu acho que é por causa da formação viu.
> P: Como assim?
> S: Olhe! Você... é... eu penso assim, a pessoa tem que ver muitas vezes com quem vai casar, porque muitas vezes você tá casando com uma pessoa que você não conhece. E é difícil entendeu? E muitas vezes a pessoa tem um filho simplesmente porque ficou e isso aí vai prejudicar em que? Ele não tem amor, não tem respeito e se ele não possui isso é, principalmente eu penso assim se você não tiver o temor, você pode ter qualquer religião mas, se você não tiver temor acontece isso.

A fala da professora Sandra ao olhar a Imagem 1 da Figura 64 sugere que ela identifica os grupos de categorias e apresenta dúvida, enquanto resposta a questão, entre duas categorias, pai e padrasto, isso pode se dar pela sobreposição dos *plots*.

No entanto, ao separar o gênero e empilhar os *plots* ela afirma no ato que o pai é o agressor principal das meninas. Esse tema é tão motivador para essa professora que ela justifica sua resposta com argumentos do seu cotidiano e das suas crenças. Apesar da justificativa ser pessoal, fica claro que no início da fala em que se estabelece a dúvida entre pai e padrasto ela acredita ser o padrasto o maior agressor das meninas, entretanto ao separar as categorias e empilhar os *plots* ela afirma com segurança ser o pai e demonstra insatisfação com a informação que ela identificou a partir da representação e com uma reflexão baseada em opiniões pessoais sobre os dados apresentados.

Outra forma de utilizar a ferramenta *separar* foi invertendo os eixos. Um exemplo dessa utilização está no caso da professora Janice. Como representação final obteve o que mostra a Figura 65.

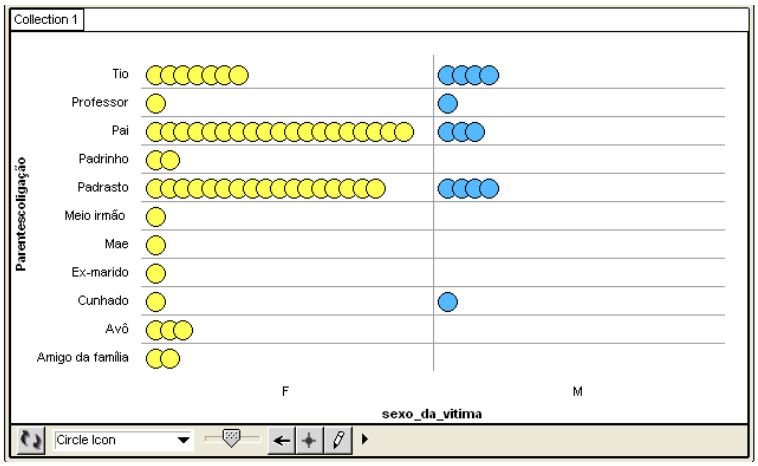


Figura 65: Exemplo 2 de uso de a ferramenta *separar* associada a *empilhar* para responder a segunda questão do banco de dados Violência

Janice: Legal, mais fácil. O pai e o padrasto são os concorrentes né?
Pesquisadora: Hummm.
J: Avô, amigo da família, tio, cunhado. Ai só quer meninas não é?
P: É meninas e meninos.
J: Ainda bem que professor teve pouco, por que senão eu confiava mais não.
P: Então, vamos lá. O estupro a meninas é mais frequentemente realizado por qual agressor?
J: O pai.
P: Por quê?
J: Porque é o que tem mais casos, já os meninos é mais pelo padrasto....
P: Padrasto né? interessante isso.
J: Não é.

Essa questão demonstrou ser de fácil entendimento e identificação da resposta pelos professores. Todos responderam satisfatoriamente e utilizaram essas duas formas de representação e essas duas ferramentas *separar* e *empilhar* para responder. Aparentemente acreditamos que o cruzamento entre variáveis qualitativas oferece uma maior facilidade de compreensão, já que, para analisá-lo seria necessário a quantificação dos casos em relação às categorias, uma análise baseada em frequências. No caso específico da questão dois o ponto máximo foi a resposta a ser encontrada.

c) Terceira questão do Banco de dados Violência

A terceira questão tratava de um cruzamento de variáveis quantitativasXquantitativas, para responder a questão “Existe alguma relação entre a idade da vítima e a idade do agressor?”. Como analisado no Capítulo 6 sobre representações, essa relação entre idade do agressor e idade da vítima é fraca. Isso pode ser observado em um gráfico de dispersão pela falta de tendência na disposição dos pontos ou no nosso caso *plots*.

A ferramenta mais utilizada, de acordo com o gráfico 9 foi a *separar*, esse resultado está muito associado ao tipo de representação construída por eles, já que em grande parte foram construídos os gráficos de dispersão e as representações que utilizaram intervalos para representar as variáveis. Como exemplo de gráfico de dispersão, temos a representação construída pela Professora Keila.

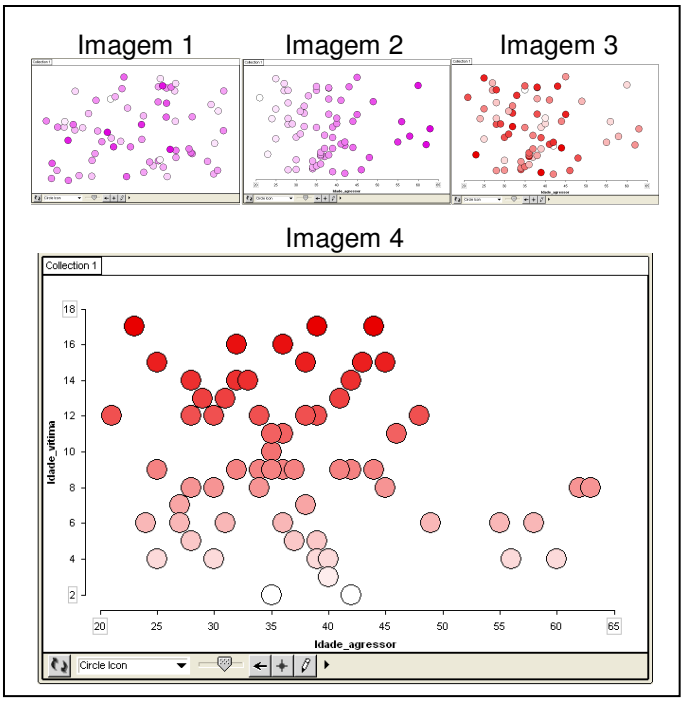


Figura 66: Exemplo 1 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Violência

Para construir o gráfico de dispersão da Figura 66 foi necessário que a professora selecionassem primeiro a variável idade do agressor, como Imagem 1; clicasse e arrastasse até o final da tela um *plot* qualquer para construção da escala horizontal, Imagem 2; selecionasse a variável idade da vitima, Imagem 3 e por último clicasse e arrastasse para cima um *plot* qualquer até o final da tela. Com isso,

teremos o que mostra a Imagem 4 da Figura 66, em evidência, o gráfico de dispersão da relação idade da vítima e idade do agressor.

Como reflexões a professora Keila afirma.

> Pesquisadora: Embaixo é a idade da vítima e de lado é a idade do agressor, e aí tem relação?
> Keila: Essa daqui é a idade da vítima esses aqui ? [aponta com o *mouse* para escala vertical]
> P: É.
> K: E aqui é idade do agressor. [aponta com o *mouse* para escala horizontal]. Hum... [pensa por alguns instantes] Existe alguma relação entre a idade do agressor e a idade da vítima? [ela lê a questão novamente] É se eles escolhem mais novos ou mais velhos não é?....
> P: Humrum.
> K: Eu achava que eles escolhiam os mais novos, eu tô um pouco confusa com isso aqui... Não tô entendendo direito...

A professora Keila encontra muita dificuldade em responder essa questão através desse tipo de gráfico. Essa mesma dificuldade foi percebida nos 8 professores que utilizaram esse gráfico, este resultado pode estar atrelado ao tipo de gráfico utilizado, como também podemos inferir que o tipo de relação apresentada nesse gráfico não contribuiu para interpretação, já que a relação fraca não apresenta nenhuma tendência de linearidade evidente.

Percebendo essa dificuldade a pesquisadora interveio sugerindo a utilização da ferramenta separar novamente, só que de dessa vez transformando as classes pontuais em classes de intervalos, como é exemplificado nas Imagens da Figura 67.

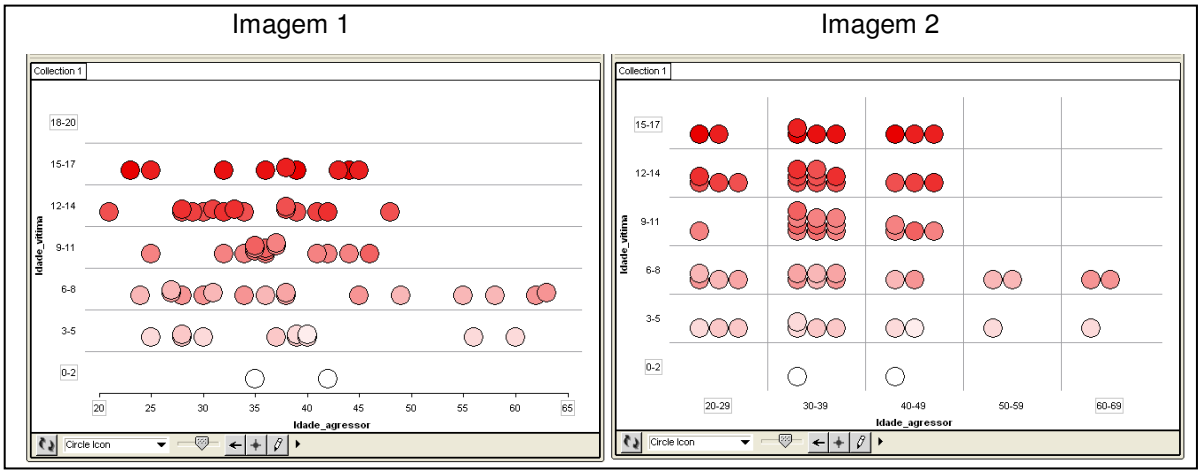


Figura 67: Exemplo 2 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Violência

Para construir essa representação a professora clicou em um dos *plots* presentes no gráfico de dispersão e arrastou ao contrário da formação das classes pontuais, utilizando a ferramenta separar, isso quer dizer que se a escala foi horizontal a professora clicou e arrastou para esquerda, se a escala for vertical a professora clicou e arrastou para baixo. Aos poucos a professora vai observado a quantidade de intervalos que deseja formar.

A professora Keila organizou a idade da vitima em sete intervalos (eixo vertical), como mostra Imagem 1 da Figura 67. Em seguida clica mais uma vez em qualquer plot e arrasta para o lado esquerdo formando cinco intervalos para idade do agressor como mostra Imagem 2 da Figura 67. Keila continua sua reflexão sobre os dados após essa mudança.

> Keila: Idade do agressor né, 69 Ave Maria! Isso é um véio. Existe alguma relação entre idade do agressor e idade da vitima? Não! [lê a questão na tela do computador]
> Pesquisador: Por quê?
> K: Eles gostam de pegar os mais novos, os mais novos sempre é o que acontece muito ou criança ou adolescente, quando pegam uma velhinha coitado é porque tá no último caso.
> P: Ham.
> K: Não tem mais o que pegar vai ser essa mesmo, mais são os mais novos.
> P: Onde você tá vendo isso?
> K: Aqui! Olha... [aponta com o *mouse* para as categorias de 30 a 39] tanto faz esses pegarem os mais novos de 3 anos como os mais velhos de 15 anos... É um crime isso né.... já viu!

Nesse trecho do protocolo da sessão com a professora Keila consegue perceber a fraca relação entre as duas variáveis. Evidente que ela não utiliza a nomenclatura específica da Estatística, no entanto, já percebemos o avanço realizado por ela ao perceber essa relação e perceber que não existe tendência entre as variáveis. Essa mesma professora que no extrato da Figura 66 apresentou dificuldades em entender o gráfico de dispersão. Então, no início de sua fala há um distanciamento dos dados e uma aproximação com a sua realidade quando ela emite frases de surpresa e de opiniões pessoais. Porém, ao ser questionada pela pesquisadora sobre os indícios em que ela se baseava para afirmar sua resposta, acaba demonstrando ter compreendido o cruzamento das variáveis, principalmente, quando cruza o intervalo 30 a 39 com a idade das crianças, e estabelece uma resposta baseada na ausência de tendência na relação.

Todos os professores utilizaram essa representação, mesmo os que iniciaram utilizando o gráfico de dispersão foi sugerido pela pesquisadora a mudança de escala para categoria, com essa mudança podemos perceber que 9 dos professores conseguiram perceber o cruzamento entre as variáveis, o que apenas no gráfico de dispersão não foi possível.

De uma forma geral, as ferramentas mais utilizadas para o banco de dados sobre Violência foram a *separar* e *gradiente* auxiliando na mudança de representação para que os professores pudessem observar indícios, antes não observados, e contribuir com a interpretação.

7.2. Banco de dados Novelas exibidas – Ferramentas utilizadas

Como discutido no Capítulo 6 sobre as representações, o banco de dados sobre Novelas vem abordando características de Novelas exibidas no horário nobre da Televisão. Para ele são feitas três questão de cruzamento de variáveis, as quais serão apresentadas a seguir no Gráfico 10 que apresenta o percentual do uso das ferramentas no decorrer da atividade.

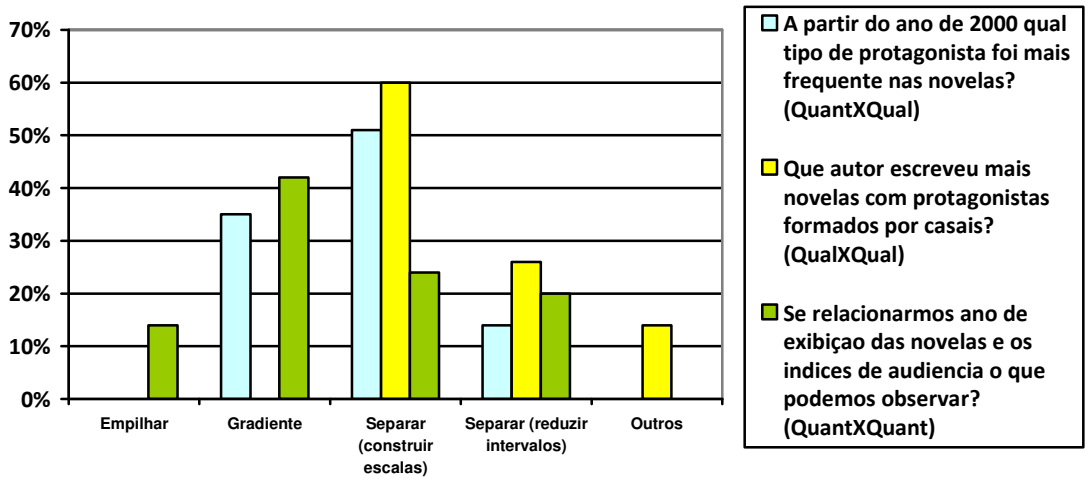


Gráfico 10: Ferramentas utilizadas para responder as questões do banco de dados Novelas

A partir do Gráfico 10 podemos observar o percentual do uso das ferramentas *empilhar*, *gradiente*, *separar* (com as suas duas funções: criar categorias e reduzir intervalos) e outros. Para esse banco de dados as ferramentas mais utilizadas foram *separar* e *gradiente*, assim como no banco de dados anterior. Também podemos

observar o percentual do uso das ferramentas em cada questão proposta, o que iremos analisar a seguir.

a) Primeira questão do Banco de dados Novelas

A primeira questão relaciona-se ao cruzamento de uma variável quantitativa com uma qualitativa, os professores responderam “A partir do ano de 2000 qual tipo de protagonista foi mais frequente nas Novelas? Porquê?”. Como discutido anteriormente os professores teriam que analisar a frequência de Novelas exibidas a partir do ano de 2000 em relação ao tipo de protagonista. As ferramentas utilizadas foram a *gradiente* e *separar*, para demonstrar a representação que utilizou a ferramenta *separar* apresentamos o exemplo da professora Keila.

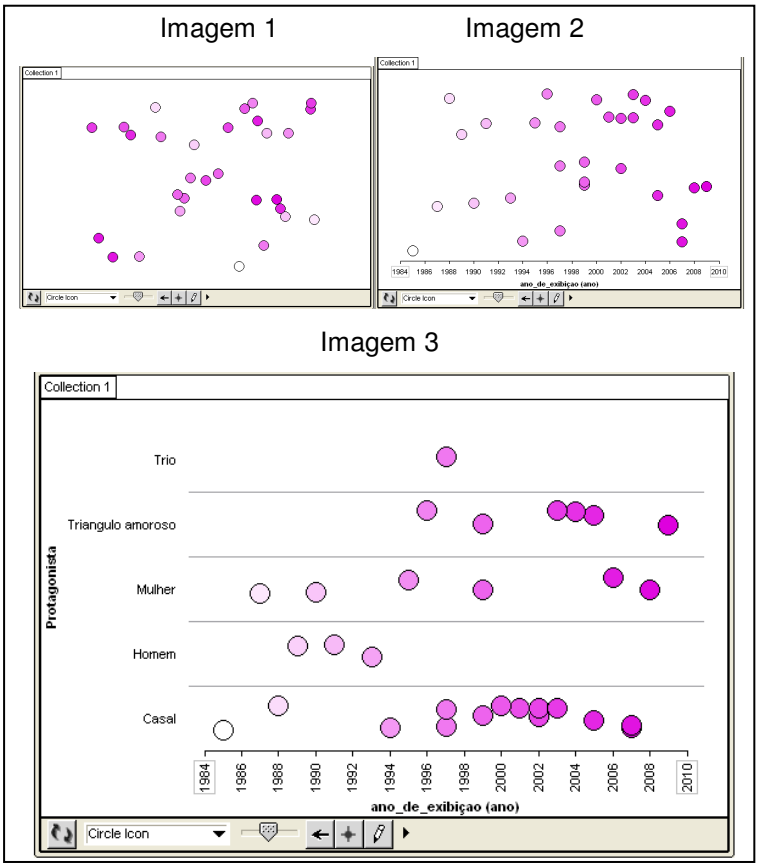


Figura 68: Exemplo 1 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas

Para essa construção foi necessário que a professora selecionasse ano de exibição, como na Imagem 1 da Figura 10. Em seguida clicou em um dos *plots*, para

utilizar a ferramenta *separar*, e arrastou para o lado esquerdo até o final da tela formando a escala, como na Imagem 2. E por último, utilizou *separar* novamente, clicando na variável protagonista e arrastando para cima até que as categorias se formassem, como na Imagem 3 da Figura 68. A professora Keila refletiu sobre essa representação conforme extrato do protocolo de sua sessão.

> Keila: Os mais claros é homem é?...Ah! Não tem cor né!?
> Pesquisadora: Não. A cor representa os anos.
> K: Triângulo amoroso.... O trio.
> P: Trio, por quê? A partir do ano 2000 que tipo de protagonista foi mais frequente nas Novelas? [Lê a questão na tela do *Software*]
> K: Aqui tá dizendo trio né.
> P: Por quê? Onde é que tu tá vendo isso?
> K: Ah!... é a partir de 2000 não é?
> P: É.
> K: A partir de 2000 então, Triângulo amoroso.
> P: É o mais frequente? Tu tá dizendo isso baseado em que?
> K: No triângulo amoroso olha aqui 2000.
> P: Tem mais é isso?
> K: Tem mais.

Na fala da professora Keila podemos observar que ela se confunde entre os protagonistas: *Trio* e *Triângulo amoroso*. Ela parece estar observando o ponto mais alto no gráfico, nesse caso, o ponto mais alto estava no *Trio*. Ponto mais alto aqui está relacionado a linha mais alta, o *plot* que está mais alto. Então, quando o pesquisador pergunta novamente qual protagonista foi mais frequente a partir de 2000 ela percebe que é a partir de 2000, então refaz sua resposta afirmando ser *Triângulo amoroso*. Ela continua com o ponto mais alto, no entanto, ela associa o ponto mais alto no ano de 2000.

Dos 6 professores que construíram esse tipo de representação três deles deram a mesma resposta que a professora Keila. Os outros analisaram afirmando que casal foi o tipo de protagonista mais utilizado a partir de 2000 devido a quantidade de *plots* existentes após o ano de 2000 na categoria casal. Quando o pesquisador percebe o equívoco da professora Keila sugere a inversão dos eixos, como mostra no trecho a seguir e a utilização da ferramenta separar para organizar os anos em intervalos.

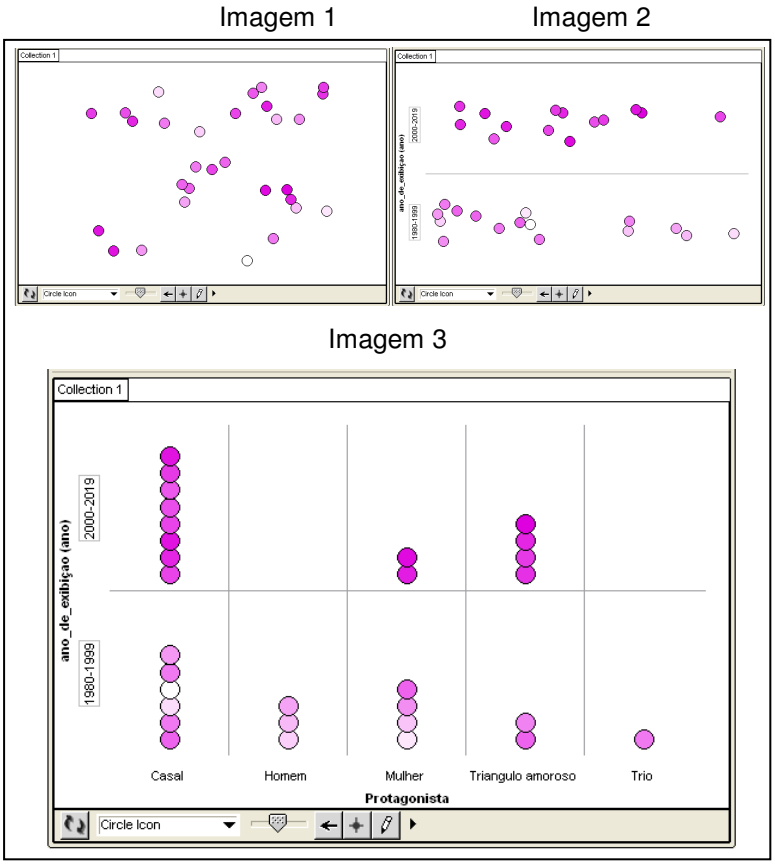


Figura 69: Exemplo 2 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas

Para essa representação a pesquisadora sugeriu que a professora retornasse ao ponto inicial, como na Imagem 1, com apenas a seleção da variável ano. Em seguida, separasse em dois intervalos o ano, como na Imagem 2 e finalmente selecionasse a variável protagonista e separa-se até formar as categorias, como na Imagem 3 da Figura 69. Com isso a professora muda de opinião, como segue trecho abaixo.

> Pesquisador: E agora melhorou pra ver assim?
> Keila: Casal.
> P: O quê?
> K: Casal foi mais.
> P: Por quê?
> K: Porque as cores...., tem mais bolinhas.... e as cores são mais fortes.
> P: Mais fortes onde?
> K: No ano de 2000.
> P: Hummm e ficou melhor assim?
> K: Ficou, ficou sem dúvida ficou. Ficou mais organizado, bem melhor.

Nessa representação encontramos uma grande semelhança com um gráfico de barras e a resposta está vinculada a maior barra, ponto máximo. Então, a fala da professora Keila demonstra que nessa representação ela não teve nenhuma

dificuldade em identificar que Casal foi o mais produzido, porém ela se apóia em um único ponto o mais alto ou o mais escuro como ela mesma relata. Não conseguimos observar na fala da professora o entendimento de que a barra está representando um intervalo de anos, então apenas podemos afirmar que ela se baseou na barra mais alta, ponto máximo.

A professora Nara também utilizou o mesmo tipo representação de Keila para iniciar respondendo a questão (Figura 69) e produziu uma reflexão semelhante a da professora Keila, em relação ao ponto máximo no gráfico. No entanto a intervenção da pesquisadora apenas promoveu a transformação da escala do ano em categorias, sem inverter os eixos, como exemplificado na Figura 70.

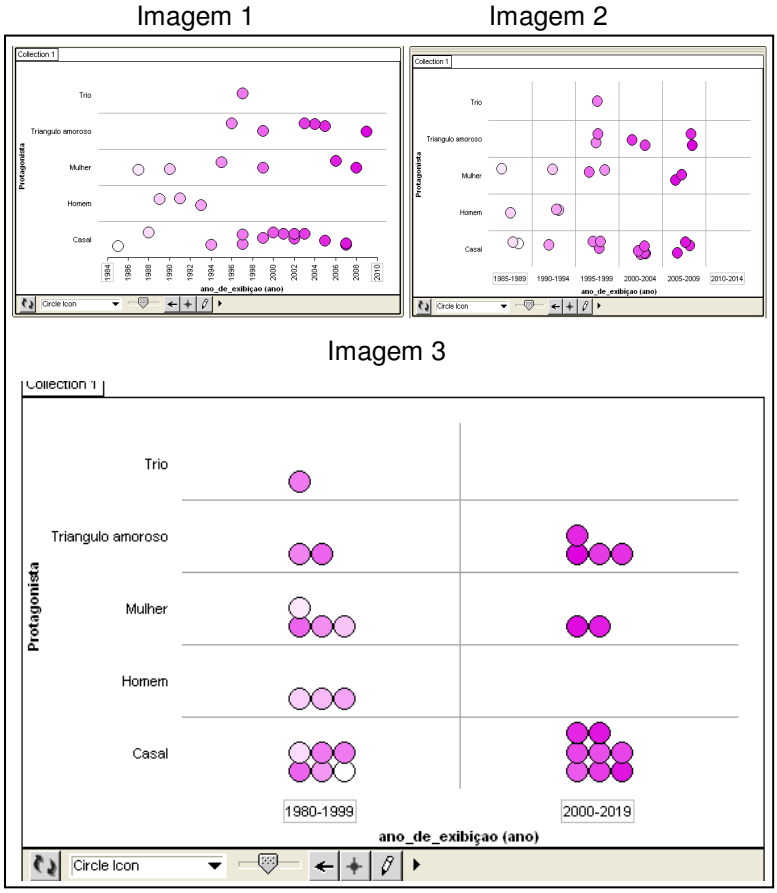


Figura 70: Exemplo 3 de uso de a ferramenta *separar* para responder a primeira questão do banco de dados Novelas

Para essa representação a pesquisadora sugeriu, a partir da representação disposta na Imagem 1 da Figura 70 em que a professora Nara deu como resposta *Triângulo amoroso*, a transformação da escala em categorias. Para isso ela utilizou-se da ferramenta *separar* clicando em um *plot* e arrastando para o lado esquerdo, de

forma inversa de maneira bem lenta, observando a organização dos intervalos. A professora construiu a Imagem 2 e continuou arrastando até que parou construindo a representação disposta na Imagem 3 da Figura 70.

> Nara: Casal no caso, não é?
> Pesquisador: Por quê? Mudou?
> N: [risos] É... Vê... né por aqui não? [aponta com o *mouse* para a categoria de 2000-2019] Por causa desse dado aqui.
> P: Sim mais o que tem aí?
> N: [risos] tem mais né!
> P: Mais comparado com quem?
> N: Mais... mais do que todos.

Com a intervenção da pesquisadora e a formação de intervalos a professora mudou sua resposta, mesmo a representação sendo semelhante a da apresentada na Figura 68 em que ela afirmou ser *Triângulo amoroso* a resposta. Acreditamos que observar a transformação da escala possa ser algo que contribua na observação dos dados de forma global e não apenas vinculado a observação dos pontos extremos. No entanto não podemos afirmar que a professora compreendeu por completo as informações dos dados, já que ela nesse momento afirma que sua resposta estava baseada no grupo de tem mais, não apenas na comparação da quantidade de Novelas exibidas na categoria 2000-2019 em relação ao tipo de protagonista.

Outra ferramenta utilizada para essa questão foi a *gradiente*. Conforme exemplifica a Figura 71.

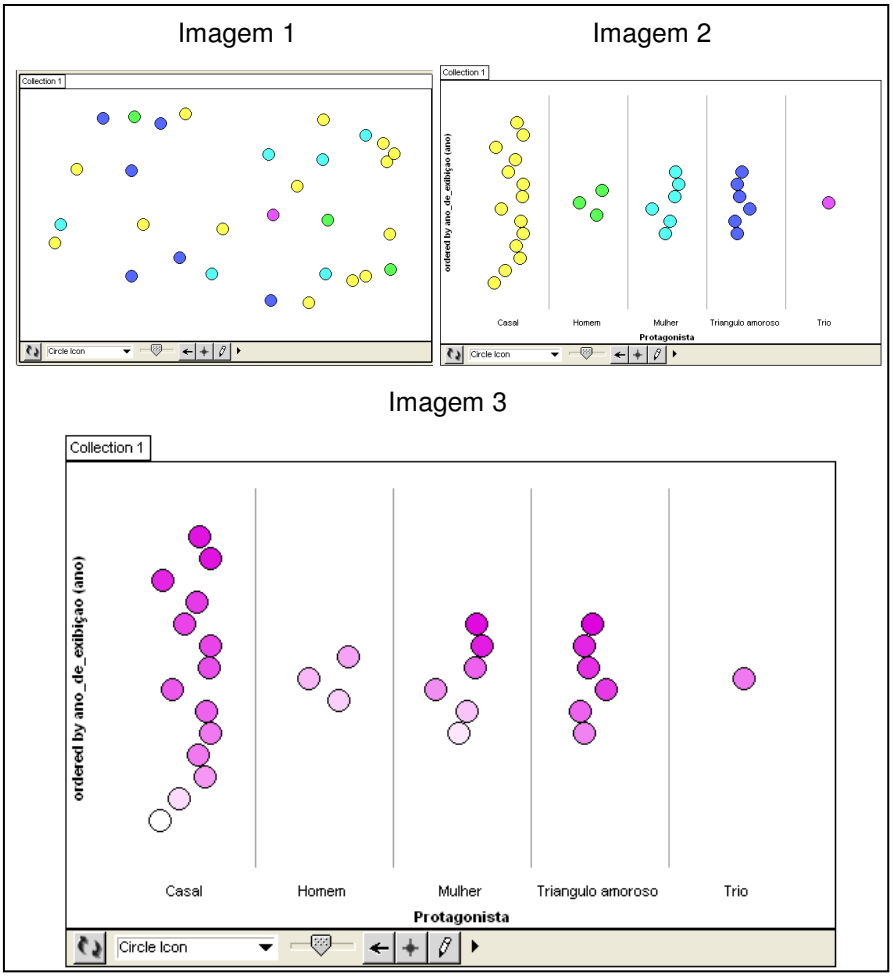


Figura 71: Exemplo 1 de uso de a ferramenta *gradiente* para responder a primeira questão do banco de dados Novelas

A professora inicialmente selecionou a variável protagonista, como mostra na Imagem 1; em seguida, utilizou a ferramenta *separar* para construir as categorias dos tipos de protagonistas, como na Imagem 2 da Figura 73. Por último selecionou a variável ano, com isso ela deixou evidente a gradação de tonalidades de cor rosa referente a variável ano, ferramenta *gradiente*. Como isso o rosa mais escuro representa os anos mais recentes (ou valores maiores) e quanto mais claro os anos mais antigos (ou valores menores). O extrato da sessão com a professora Sandra exemplifica o uso da ferramenta *gradiente.*

> Pesquisadora: Será que não dava pra ver se você apertar em ano?
> [se refere a seleção da variável ano para que o *gradiente* possa aparecer nos *plots*]
> Sandra: Eu não consigo [tenta e não consegue da primeira vez, até que consegue selecionar] Casal.
> P: Por quê?
> S: Porque tem mais.

> P: Sim mas a pergunta é a partir de 2000 que o protagonista esteve mais presente nas Novelas?
> S: Sim... [olha por alguns minutos] Casal!
> P: Por quê?
> S: Pela quantidade e pela cor.
> P: Como assim?
> S: A cor né o ano maior né? [a pesquisadora balança a cabeça confirmando]. Então, em casal temos mais escuro, mais Novelas nos anos maiores de 2000 pra cá.

A fala da professora Sandra indica que inicialmente ela observou apenas a tamanho das barras (quantidade maior nos agrupamentos). No entanto, quando a ferramenta *gradiente* é utilizada provoca um momento de reflexão, a professora olha por alguns minutos e inicia uma análise considerando a relação entre as variáveis. Então o uso dessa ferramenta sensibilizou o olhar da professora para outros aspectos além, apenas, do ponto máximo.

Entretanto, a utilização dessa ferramenta não garante uma análise aprofundada dos dados, já que os valores são representados por cores. Ela pode ser utilizada como auxiliar, mas nunca como uma única forma de análise. Assim não podemos garantir que os *plots* rosas escuros são os anos acima de 2000, então é uma análise superficial, porém não podemos negar que a cor atraiu o olhar da professora para variável ano, e com isso, ela pode refletir mais um pouco sobre os dados.

b) Segunda questão do Banco de dados Novelas

A segunda questão requeria o cruzamento entre duas variáveis qualitativas para responder a questão “Que autor escreveu mais Novelas com protagonistas formados por casais?”. Como discutido no Capítulo das representações os professores precisavam cruzar protagonista e autor das Novelas, com isso, iriam observar que Gilberto Braga escreveu quatro Novelas com casais, comparando com os outros autores ele seria escreveu mais Novelas com esse tipo de protagonista. E os professores construíram representações bem semelhantes e responderam sem dificuldade a essa questão.

Nesse caso a quantidade na utilização de ferramentas foi pequena nessa questão e as intervenções realizadas pela pesquisadora foram poucas e estavam mais relacionadas a formação das categorias iniciais, como exemplo da Figura 72.

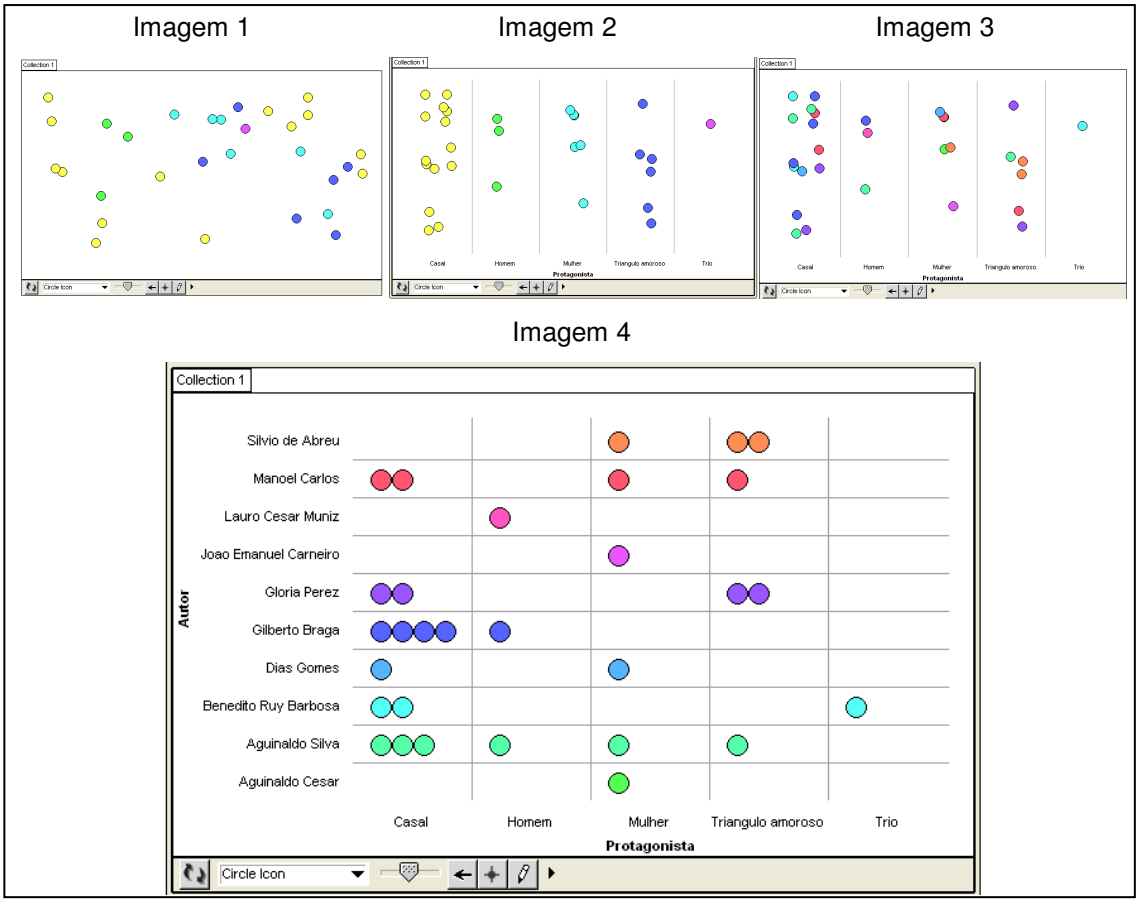


Figura 72: Exemplo 1 de uso de a ferramenta *separar* (criar categorias) para responder a segunda questão do banco de dados Novelas

Para chegar a essa representação o professor João selecionou a variável protagonista, como na Imagem 1; em seguida utiliza a ferramenta *separar* para formar as categorias de protagonistas; então ela seleciona a variável autor, como na Imagem 3 e por último separa em categorias com a ferramenta *separar*, como mostra a Imagem 4 da Figura 72. Esse caminho foi utilizado por todos os professores e as respostas foram semelhantes as de Raquel e Mércia como segue.

> Pesquisador: Que autor escreveu mais Novelas com protagonistas formados por casais?
> Raquel: Gilberto Braga.
> P: Gilberto Braga, Por quê?
> R: Por quê? [Professora sorri] Por causa das bolas de novo, porque a quantidade de bolas estão maior em Gilberto Braga... mais Novelas né?....
>
> Mercia: É pra dizer os nomes dos autores, qual dos autores escreveu mais Novelas... [a professora observa a representação por 3 segundos e coloca o *mouse* em cima dos *plots* do autor Gilberto Braga] é Gilberto Braga.
> Pesquisador: Por quê?

> M: [a professora observa a representação e coloca o *mouse* em cima dos plots de Agnaldo Silva] Agnaldo Silva... não mas só casal, né? É assim mesmo no caso... Gilberto Braga.
> P:Tem certeza?
> M: [a Professora conta com o cursor do *mouse* o número de plots que está no espaço do autor Gilberto Braga e depois balança a cabeça afirmando]

Podemos observar nas falas das Professoras Raquel e Mércia que elas identificaram a relação entre autor e protagonista e conseguiram responder a questão satisfatoriamente. Identificaram a frequencia de Novelas com protagonistas casais escrita por Gilberto Braga. No caso de Mércia, ela ainda, comparou com outro autor Agnaldo Silva que também escreveu uma quantidade expressiva de Novelas com casais. Assim como na questão dois do primeiro banco de dados sobre Violência, essa questão necessitava de um cruzamento simples de variáveis para os professores entenderem e identificarem a resposta. Todos os professores responderam a essa questão, os que utilizaram uma organização invertida nos eixos das variáveis, precisaram da ferramenta *número* para identificar a resposta, uma vez que espaço da tela do software não possibilitava a visualização total de todos os *plots*, como é exemplificado pela Figura 75 e pelo extrato de fala da entrevista com o professor José.

Imagem 1 Imagem 2

Figura 73: Exemplo 2 de uso de a ferramenta *separar* (criar categorias) associada a ferramenta *número* para responder a segunda questão do banco de dados Novelas

> [olhando para Imagem 1]
> José: Deixa ver aqui, o autor né?
> Pesquisador: Que autor escreveu mais Novelas com protagonista formados por casais?
> J: Pera aí, deixa eu ver se. [com o *mouse* conta os *plots* no autor Gilberto Braga e Aguinaldo Silva] Não consigo ver direito.

> P: Tenho uma ferramenta que pode ajudar. Clica em número. [se refere a ferramenta *number* que possibilita a visualização dos valores numéricos]
> [A pesquisadora clica na ferramenta *número*][olhando para Imagem 2]
> José: Foi Gilberto Braga
> Pesquisador: Por quê?
> J: Porque sim. Pela quantidade. Aqui você não diria que tem três, nem aqui que tem quatro, não é?
> P: É porque tá uma por cima do outro.

Essa passagem do professor José retrata o uso da ferramenta *número* para indicar a quantidade de cada intersecção das variáveis. Na primeira representação, Imagem 1, os *plots* estavam sobrepostos e para que ele confirmasse a resposta foi necessário a utilização do recurso número do *software,* Imagem 2.

c) Terceira questão do Banco de dados Novelas

A Terceira questão foi um cruzamento de variáveis quantitativas, no qual, os professores tiveram que responder “se relacionarmos ano de exibição das Novelas e os índices de audiência o que podemos observar?”. No entanto, como discutido no Capítulo das representações, o cruzamento das variáveis *ano de exibição* e *audiência* das Novelas possui uma relação forte com uma tendência decrescente.

Para essa questão todos os professores obtiveram um bom desempenho logo na primeira representação construída. Os professores de uma maneira geral não encontraram dificuldade em responder a essa questão diferente da questão três do banco de dados de violência que possuía a mesma natureza de cruzamento (quantitativaXquantitativa). Para construção das representações os professores utilizaram as ferramentas *empilhar*, *separar* e *gradiente. C*omo apresenta o Gráfico 10.

Utilizando a ferramenta *gradiente* os professores puderam observar a tendência da relação entre as variáveis através do *gradiente* na tonalidade da cor azul, como segue trecho da professora Poliana e a Figura 73 da representação construída.

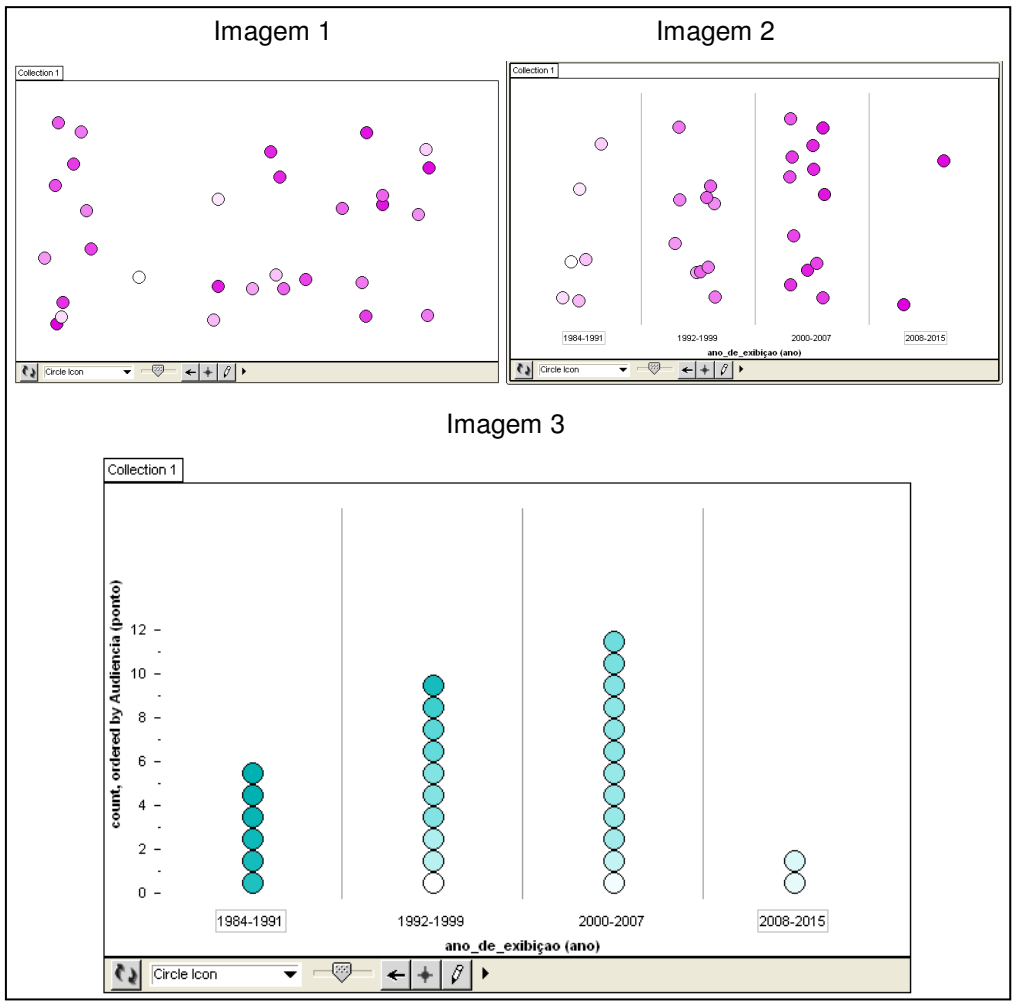


Figura 74: Exemplo 1 de uso de a ferramenta *gradiente* para responder a terceira questão do banco de dados Novelas

Para a representação apresentada na Figura 74 a professora iniciou selecionando a variável ano, Imagem 1, clicando em um *plot* e arrastando para o lado direito, com isso, ela formou intervalados de 8 anos, como mostra Imagem 2 quadro da Figura 74. Por último selecionou a variável audiência que deixou aparente a cor azul em diferentes tonalidades que representam os pontos de audiência das Novelas, Imagem 3. Então podemos observar na Imagem 3 que os tons de azul escuro representam os índices de audiência maiores e os mais claros os menores índices de audiência.

[olhando para Imagem 2]
Poliana: Então teve mais audiência entre o ano de 2000 à 2007.
Pesquisadora: Não, ai é só o ano de exibição. Não tem audiência ainda, da um clique em audiência.
Poliana: Pode?
Pesquisadora: Pode.

> [olhando a Imagem 3]
> Poliana: Ano de exibição e as audiências. Aí aqui, continua do mesmo jeito dos gráficos, de tudo? Quer dizer que as cores mais fortes teve mais audiência entre 84 e 91 e menos audiência entre 2008 e 2015.
> Pesquisadora: Por quê? Onde tu tá vendo isso?
> Poliana: Tem mais bolinha escura nos anos anteriores, olha aqui! [aponta com o *mouse* para a categoria de 1984-1991]. O pessoal assistia mais novela entre 84 e 91. Tinha mais tempo de assistir novela.
> Pesquisadora: Antigamente...
> Poliana: Hoje em dia não tem tempo de assistir Novelas mais não.

A fala da professora Poliana indica que ela percebeu os intervalos como uma variável e relacionou-a com as cores que indicavam a audiência das Novelas. Nesse caso ela pode tecer uma análise sobre essa relação afirmando que antigamente se assistia mais Novelas do que hoje em dia, essa fala esta ligada as cores dos *plots* (audiência), mas também está muito relacionada a sua vivência pessoal, já que ela conclui que nos dias de hoje ninguém tem tempo de assistir Novelas.

5 professores utilizaram esse tipo de representação e obtiveram resultados similares aos da professora Poliana em associar a resposta às gradações de cor dos *plots*. Isso indica que o *gradiente* pode ter contribuído para que os professores não olhassem apenas para a maior barra. No entanto, responder questões de análise de dados apenas pela cor pode levar a interpretações limitadas dos dados, pois a gradação relaciona-se a valores, mas não possuem a mesma precisão da representação dos dígitos numéricos. Nesse caso fica evidente que nos anos de 84 a 91 tínhamos índices de audiência maiores, pela tonalidade da cor azul estar bem mais escura nesse período, porém, não podemos identificar os valores numéricos dos índices desse período, já que só temos as cores como indícios. Nesse sentido, o uso da gradação de cor contribuiu como uma pista para auxiliar na interpretação, mas é necessário utilizar essa ferramenta associada a outras representações para que possamos ter uma análise mais consistente.

Outra ferramenta que auxiliou a análise de dados para essa questão foi a *separar*. Com essa ferramenta houve a construção de um gráfico de dispersão, como no caso do banco de dados de violência. Entretanto, os professores não tiveram dificuldades em responder por esse gráfico nessa questão proposta. Temos o exemplo das professoras Nara e Keila.

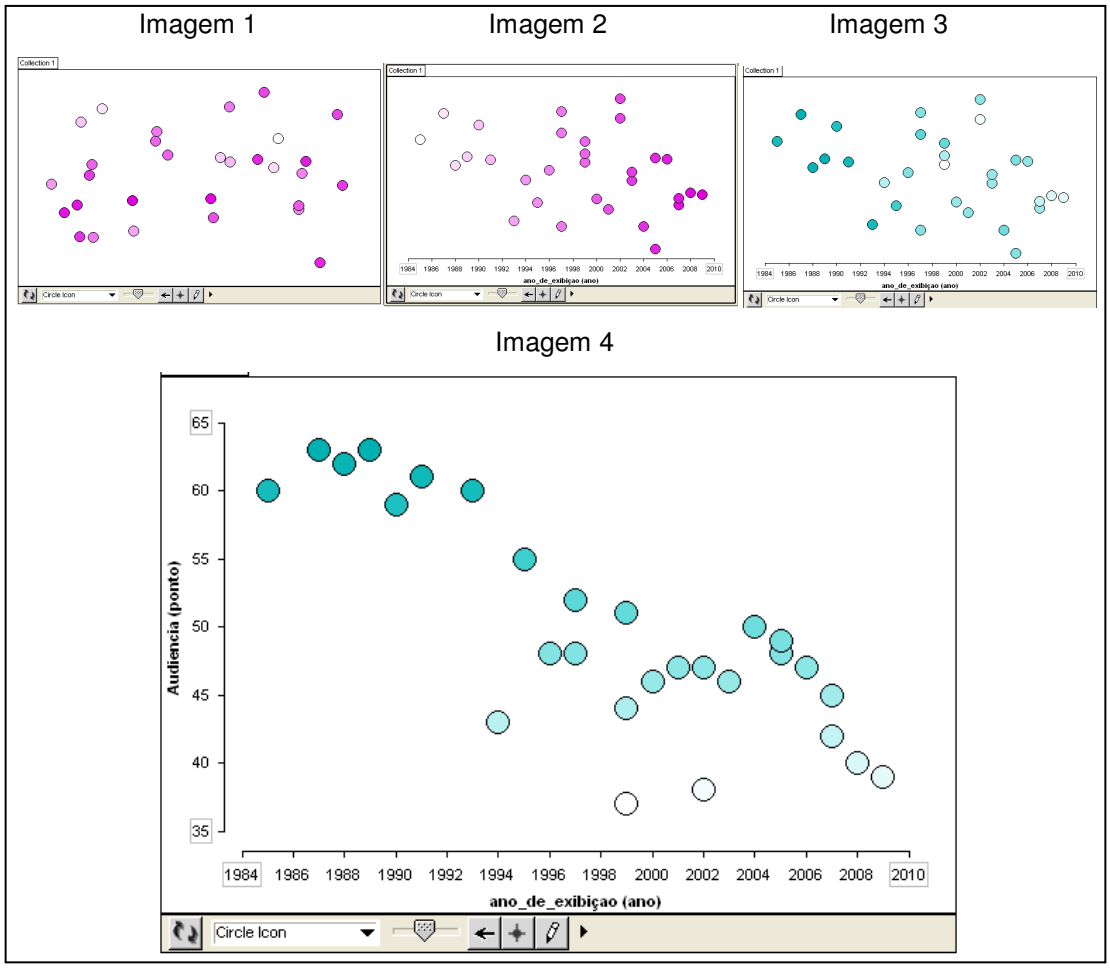


Figura 75: Exemplo 1 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Novelas

Para a construção desse gráfico as professoras inicialmente selecionaram a variável *ano*, como na Imagem 1 da Figura 75. Depois clicaram em um *plot* e arrastaram para o lado direito formando a escala da variável *ano*, como na Imagem 2, em seguida selecionaram a variável *audiência*, como na Imagem 3 e por último clicaram e arrastaram um *plot* pra cima para formação da escala da *audiência* na vertical, como mostra Imagem 4 da Figura 75.

Nara: É,... Antigamente tinha mais.
Pesquisador: Tinha mais, o quê?
N: Audiência
P: Por quê?
N: Porque aqui dá para perceber que antes o índice era maior de 50... 45... o pessoal assistia mais.
P: E hoje em dia?
N: Hoje em dia, pouquíssimo... de 50 pra baixo.

> Keila: Se relacionarmos o ano de exibição das Novelas e o índice de audiência o que podemos observar?... [a professora lê a pergunta disposta no *software*]... Que... as Novelas antigas foram mais assistidas. Foi..? Pelo menos eu acho.
> Pesquisador: Hum. Onde tais vendo?
> K: Eu tô vendo... vou dizer, tô vendo assim.... pela escala que as bolinhas ficaram mais altas.
> P: Como assim?
> K: Aqui elas estão mais altas [aponta para os *plots* que apresentam Audiência 60] e depois vai diminuindo.
> P: O que aconteceu?
> K: Que diminuiu né, 2010 diminuiu o índice de audiência. Já pensou?!

As falas das Professoras indicam que elas conseguiram observar a tendência do gráfico, como também perceber a relação entre as duas variáveis. Elas tecem uma análise baseadas na disposição dos *plots* que iniciam mais altos, em relação a audiência, e decrescem ao passar dos anos. Podemos identificar, ainda, que a professora Keila apresentou um desempenho muito interessante nessa questão diferente do apresentado na questão 3 do banco de dados da Violência quando ela afirmou não ter entendido esse mesmo tipo de representação, como podemos ver no trecho da Figura 66. 10 Professores construíram esse tipo de representação, 5 dos professores construíram por indicação da pesquisadora e todos que construíram essa representação obtiveram resultados similares com as das professores Keila e Nara.

Outro uso da ferramenta separar foi na criação de intervalos, como mostra o trecho da professora Janice e Sandra.

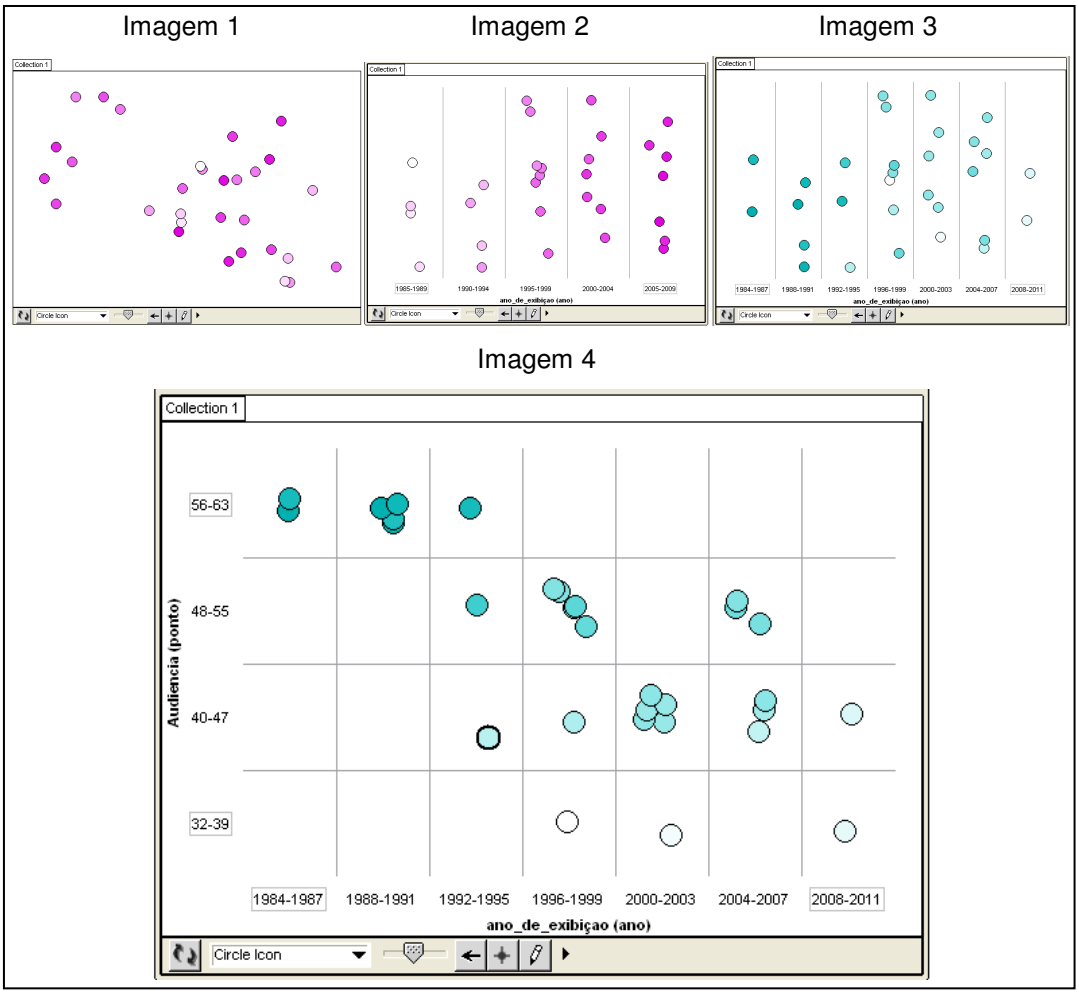


Figura 76: Exemplo 2 de uso de a ferramenta *separar* para responder a terceira questão do banco de dados Novelas

Para a construção desse gráfico as professoras inicialmente selecionaram a variável *ano*, como na Imagem 1 da Figura 76, depois clicaram em um *plot* e arrastaram para direita formando os intervalos de 4 anos, como na Imagem 2. Em seguida selecionaram a variável *audiência*, como na Imagem 3 e por último clicaram e arrastaram um *plot* pra cima para formação dos intervalos de audiência em 8 pontos no eixo vertical, como mostra Imagem 4 da Figura 76. 4 professores construíram esse tipo de representação, como exemplo apresentamos extratos de falas das professoras Janice e Sandra.

> Janice: Audiência é aqui é? [aponta para as categorias no eixo vertical]
> Pesquisadora: Aqui embaixo é o ano [aponta para as categorias na horizontal] e a audiência é aqui [aponta para as categorias na vertical]
> J: No caso a novela tem...., mais... o ano foi 84 a 87 que teve mais audiência né isso?
> P: Certo.
> J: De 2011 à 92, tá ao contrário né? Tá subindo.

P: O que isso quer dizer?
J: Que tinha mais audiência antes. Também o pessoal tinha mais tempo que logo quando todo mundo começou a ver novela e tudo. O que tinha para fazer era isso.
P: Hoje em dia a gente tem uma audiência?
J: Bem menos. Porque eu acredito que seja, hoje em dia ninguém mais tem tempo.

Sandra: Bem aqui a audiência.... aqui é de 84 até 91.
Pesquisador: Hummm.
S: Aí as pessoas tinha mais tempo de assistir televisão, hoje as pessoas estudam, fazem um curso a noite, um cursinho ai não tem tempo de assistir que nem eu.
P: Aí assim nesse tempo a audiência era o quê?
S: Mais alta. Por que as pessoas tinham mais tempo, ficavam mais em casa.
P: Você diz isso a partir de quê?
S: Nos anos de 84 a audiência era maior que 56... já hoje é bem poquinha 32....
P: Aí... assim, olhando aqui no gráfico o que você pode dizer mais?
S: Que vem declinando. E outra coisa hoje a gente... Você pegou a globo não foi?
P: Foi.
S: Hoje têm outras é... [risos], tem outras rede de TV que tem mais Novelas agora aí as pessoas tem opções, programas. E principalmente hoje tem os DVDs as pessoas preferem assistir.
P: Do que ver novela?
S: É.

Com o uso dessa ferramenta foi possível observar a relação entre as variáveis e a professora Janice conseguiu perceber a tendência do gráfico. Ela identificou que existia uma relação inversa, mesmo que sem a utilização dessa nomenclatura quando afirma que de “2011 a 1992 está ao contrário”, o que para a análise do cruzamento seria uma relação inversa, decrescente.

A professora Sandra também percebe a relação entre as variáveis identificando a tendência do gráfico a diminuir de acordo com o passar dos anos. No entanto, ela traz em sua fala muitas informações do cotidiano para justificar ou entender melhor essa diminuição dos índices de audiência. Por exemplo, quando ela fala no aparecimento de outras emissoras de TV e a influência no índice de audiência.

Todos os professores conseguiram responder a essa questão estabelecendo relação entre as variáveis e identificando a relação forte e inversa de cruzamento. Diferente da questão três do banco de dados de violência, no qual os professores

tiveram muita dificuldade em entender a relação entre os eixos e identificar uma relação fraca.

De uma maneira geral podemos perceber que os professores não encontraram dificuldade em responder a essa questão, mesmo em se tratando de um cruzamento de variáveis quantitativas, até mesmo quando utilizaram gráficos de dispersão. Nesse caso podemos atribuir esse êxito a dois fatores, o primeiro é a própria tendência do gráfico, já que havia uma relação forte existente entre as variáveis no banco de dados sobre Novelas, o que não acontecia no banco de dados sobre Violência. E o segundo fator pode estar relacionado ao próprio processo de aprendizagem ocorrido durante a resolução das atividades, já que essa foi a última questão dentre as 6 sugeridas para os dois bancos de dados e em todas havia a relação de variáveis, no qual os professores analisavam. No entanto, qualquer que seja as situações o que podemos afirmar é que a possibilidade de observar a construção da escala ou sua organização em intervalos e a utilização das cores e gradação de uma cor, foi muito utilizada pelos professores e em certos momentos auxiliou na análise dos dados.

As ferramentas utilizadas nos dois bancos de dados possibilitou aos professores um repensar em suas afirmações oferecendo indícios que serviram de pistas para que os professores observassem ora a escala, ora o ponto máximo, ora as cores mais fortes dando uma visão mais ampla e aprofundada da interpretação de dados.

CAPÍTULO 8 – CONSIDERAÇÕES

A partir da realização das atividades propostas pudemos observar que os professores apresentaram facilidade de manuseio com as ferramentas do *software*. Apesar de o software utilizar a língua inglesa na sua *interface* isso não foi um obstáculo para o seu uso. Nesse sentido, os *ícones* foram elementos importantes que auxiliaram na identificação das ferramentas. Outro aspecto que pode ter facilitado o uso do *software* foi o fato de que as ferramentas utilizadas nesse estudo podiam ser acionadas através do clicar na variável ou clicar e arrastar o *plot*.

Essa "facilidade" em manusear o software foi constatada também em estudos internacionais (por exemplo, KONOLD; MILLER, 2001; BEN-ZVI, 2006; BAKKER; DERRY, 2006) bem como nos estudos do CENAGRAF (por exemplo, LIRA, 2010; ALVES, 2011; PEREIRA; MONTEIRO, 2010; AZEVEDO; MONTEIRO, 2010).

Os estudos vinculados ao projeto CENAGRAF investigaram o uso do *TinkerPlots* entre diversos atores e em diversos contextos escolares, tais como estudantes de escolas particulares com experiências com computador (LIRA, 2010); estudantes de escolas públicas rurais sem experiência com computador (ALVES, 2011); estudantes de Educação de Jovens e Adultos sem experiência com computador (ALMEIDA; SOUZA; MONTEIRO, 2010); professores em formação inicial (PEREIRA; MONTEIRO, 2010); e professores de escolas públicas rurais (AZEVEDO; MONTEIRO, 2010). Esses estudos indicam que crianças, jovens, adultos e professores ao utilizarem o *TinkerPlots* entenderam com certa facilidade as suas ferramentas; *icones* e a forma de manipulação do banco de dados.

As questões utilizadas neste estudo requeriam a percepção da relação entre variáveis. Essa tarefa de pesquisa foi configurada de maneira a atribuir um sentido desafiador, pois para interpretar os dados era necessário testar hipóteses e observar os indícios que podiam ser utilizados para justificar a resposta construída. Os professores envolvidos nesse estudo construíram diferentes tipos de representações através das ferramentas do *software* que puderam auxiliar na interpretação.

Tomando como base o quantitativo total de representações construídas pelos professores participantes, ficou evidenciado que para as questões relacionadas ao primeiro banco de dados (Violência contra crianças) foram construídas um total de

61 representações, enquanto que para a resolução das questões vinculadas ao segundo banco de dados (Novelas) foram construídas 83 representações. Podemos inferir que esse aumento no número de representações construídas quando os professores refletiam sobre o segundo banco de dados, pode estar associado ao fato de que, naquele momento da sessão de pesquisa, eles estivessem mais familiarizados com o *software* e suas possibilidades de representar os dados. Nesse sentido, os participantes puderam ousar mais, inclusive, requerendo ajuda a pesquisadora para que ela pudesse esclarecer como poderia mudar as representações.

Observamos que para responder à cada uma das primeiras questões de cruzamento de variáveis quantitativas e qualitativas dos dois bancos de dados, os participantes construíram representações semelhantes. Tanto para aquelas representações que utilizavam classes pontuais ou classes de intervalos como aquelas com gradação de cor. No entanto, o que foi observado é que quando os professores utilizavam o *gradiente* de cor, sua resposta estava atrelada a cor mais escura, e quando era utilizada escala os professores tendiam a responder através da identificação do ponto máximo das barras. Quando os professores construíram categorias ou intervalos ou transformavam as classes pontuais em intervalos eles tendiam a perceber os valores contidos na escala, o que contribuía na análise de dados.

Na questão dois para cada banco de dados, com o cruzamento de variáveis qualitativas, pudemos observar que os professores não tiveram dificuldade em responder ou afirmar sua hipótese em ambos os bancos de dados. A representação utilizada pelos professores cruzava categorias e se assemelhava a uma tabela. Entranto, o *software* apresentou uma limitação na visualização dos *plots* nessa segunda questão. Como tínhamos uma quantidade de 68 casos para o primeiro banco de dados e de 31 casos para o segundo banco de dados, os *plots* não ficaram completamente visíveis nas representações, pois alguns plots apareceram sobrepostos, o que dificultou a contagem pelos professores. Então, foi requerida a explicitação dos valores numéricos pelo próprio *software*, o que deixou a representação, ainda mais parecida a uma tabela.

Na terceira questão para cada banco de dados, de uma maneira geral os professores utilizaram um gráfico de dispersão para representar a relação das variáveis. No primeiro banco de dados (violência) as variáveis quantitativas

estabeleciam uma relação fraca, os professores encontraram dificuldade em perceber e entender o gráfico de dispersão construído por eles. Houve a necessidade de organizar a variável em intervalos para contribuir com a reflexão sobre os dados. Já no segundo banco de dados (Novelas), a relação era forte e decrescente, com o uso do gráfico de dispersão os professores não encontraram dificuldade em responder essa questão.

A relação forte e a tendência dos *plots* dispostos no gráfico de dispersão para os dados do segundo banco de dados podem explicar a facilidade de entendimento dessa relação pelos professores, no entanto também podemos considerar o processo de aprendizagem ocorrido durante a realização das questões, pois esta foi a última questão das seis propostas nesse estudo.

Nas duas situações, no primeiro banco de dados e no segundo banco de dados, a possibilidade de mexer com a escala e utilizar o *gradiente* proporcionou aos professores um repensar de suas respostas e a identificação de indícios antes não percebidos que possibilitaram um novo olhar sobre os dados.

De uma forma geral, as ferramentas utilizadas proporcionaram uma manipulação dos dados, já que os professores puderam movimentar os *plots* por diversas ações, tais como: organização em intervalos ou categorias em eixos horizontais e verticais; empilhamento; ordenação; utilização de recursos de cores e de *gradiente*. Essa manipulação proporcionou em muitas ocasiões uma mudança na reflexão e uma contribuição para a interpretação. Uma dessas contribuições pode ser exemplificada através das falas de professores que conseguem perceber os números implícitos na escala das questões propostas. Resultados encontrados em estudos como de Cavalcanti (2010), Lemos e Gitirana (2007) e Monteiro e Selva (2001), apontam que alunos e professores encontram dificuldades em identificar valores implícitos em escalas. Tendo em vista que outros fatores podem interferir na compreensão de escalas e que o presente estudo não teve como foco investigar esse aspecto, parece oportuno que outros estudos que possam abordar a compreensão de escalas através do uso da ferramenta *separar* do *TinkerPlots*.

A ferramenta *gradiente* também pode ter sido outra contribuição do *software*, pois em diversos momentos das sessões de pesquisa, principalmente quando os participantes respondiam às primeiras questões dos dois bancos de dados, o recurso das cores promoveu o repensar sobre os dados. Nesse sentido, as cores puderam chamar atenção para outro ponto do gráfico que não fosse o ponto

máximo. Não estamos afirmando que ao responder a questão pela cor mais escura seja uma interpretação global dos dados, ao contrário, tal tipo de interpretação podia estar vinculada a um enfoque pontual. No entanto, se essa ferramenta for utilizada junto a uma reflexão sobre os aspectos gerais dos dados e como motivação de mudança de reflexão, então pode servir de instrumento para que os professores percebam outros aspectos no gráfico, considerando o comportamento das variáveis. Como foi o caso de alguns professores desse estudo que iniciam dando a resposta pelo ponto máximo do gráfico e ao utilizar o *gradiente* mudam de opinião.

No entanto, não podemos supervalorizar a ferramenta *gradiente*, já que através das discussões promovidas neste estudo, podemos perceber que seu uso, sem a representação numérica da variável quantitativa, pode promover erros na interpretação. A cor e os tons de gradação representam os valores quantitativos, mas para compreendermos o comportamento de uma variável e comparar entre diferentes categorias necessitamos dos valores numéricos (em escalas ou intervalos), com isso podemos fazer uso das medidas de tendência central utilizada em Estatística. Nesse caso, a cor serve de indicativo para chamar atenção, principalmente, da tendência de variáveis. Um exemplo disso foi no caso da questão três do banco de dados das Novelas, na qual o *gradiente* evidenciou a relação decrescente através das cores que iniciam mais fortes e ficavam mais claras.

Nesse sentido, as ferramentas do *software* foram utilizadas na construção de diferentes representações no momento da exploração dos dados, o que possibilitou uma maior interação com as informações oferecendo diferentes pistas na construção da resposta. Ainley, Nardi e Pratt (1998) discutem a importância do uso do computador como auxiliar para interpretação de gráficos, principalmente, no que se refere a levantar e testar hipóteses através do uso de diferentes possibilidades de construção de gráficos. Este estudo aponta que o *TinkerPlots* pode ser uma ferramenta possível para o trabalho com a análise de dados por professores com pouco conhecimento em Estatística e pouco uso de computadores apontando como uma possibilidade para a formação Inicial e Continuada de Professores.

Futuros estudos são necessários para que sejam propostas utilizações desse software em um processo de formação de professores, no qual sejam construídos conceitos de Estatística, dando a liberdade de manipulação e construção de diferentes representações. A investigação no âmbito da formação de professores

poderá contribuir no seu próprio trabalho com Tratamento da Informação em sala de aula.

REFERÊNCIAS

AINLEY, J.; NARDI, E.; PRATT, D. Constructing Meaning for Formal Notation in Active Graphing. In: *Conference of the European society for Research in Mathematics Education*, 1., 1998, Osnabrueck. **Proceedings**... Osnabrueck : Germany, 1998.

_____. The construction of meaning for trend in active graphing. **International Journal of Computers for Mathematical Learning**, Norwell, v. 5, n. 2, 2000. p. 85-114.

ARROYO, M. ; CALDART, R. ; MOLINA, M. (Org). **Por uma Educação do Campo**. Petrópolis:Vozes, 2004.

ASSEKER, A.; MONTEIRO, C. E. Entre os consensos sociais e a prática pedagógica do ensino de matemática: explorando as falas de professoras de escolas rurais. In*:* Simpósio Internacional de Pesquisa em Educação Matemática. 2., Recife, 2008, **Anais...,** Recife: UFRPE, 2008. CD-ROM.

_____. Explorando a reconceptualização de recursos no ensino de matemática na formação continuada. In*:* Simpósio Internacional de Pesquisa em Educação Matemática. 2., Recife, 2008, **Anais...,** Recife: UFRPE, 2008. CD-ROM.

ALVES, I,; MONTEIRO, C. E. Analisando os guias de aprendizagem de matemática de 1ª a 4ª série da escola ativa. In*:* Simpósio Internacional de Pesquisa em Educação Matemática. 2., Recife, 2008, **Anais...,** Recife: UFRPE, 2008. CD-ROM.

ALBUQUERQUE, M. R. **Como adultos e crianças compreendem a escala.** 2010. 122 f. Dissertação (Mestrado) – Pós-Graduação em Educação Matemática e Tecnológica, UFPE, Recife, 2010.

BAKKER, A.; DERRY, J.; KONOLD, C. Technology to support diagrammatic reasoning about center and variation. In: International Conference on Teaching Statistics, 7., Salvador, 2006, **Proceedings**... Salvador: Brazil. 2006. CD -ROM.

BELLEMAIN, F.; BELLEMAIN, P. M. B.; GITIRANA, V. Simulação no ensino da matemática: um exemplo com cabri-géomètre para abordar os conceitos de área e perímetro. In: Seminário Internacional de Pesquisa em Educação Matemática, 3., 2006, **Anais...**, Águas de Lindóia: São Paulo. G06 - Educação Matemática novas tecnologias e educação a distância. CD-ROM.

BEN-ZVI, D. Scaffolding students informal inference and argumentation. In: International Conference on Teaching Statistics, 7., Salvador, 2006, **Proceedings...** Salvador: Brazil. 2006. CD -ROM.

BITTAR, M. Possibilidade e dificuldades da incorporação do uso de *softwares* na aprendizagem da matemática. In: Seminário Internacional de Pesquisa em Educação Matemática,3., Águas de Lindóia, 2006, **Anais...**, Águas de Lindóia: São Paulo. G06 - Educação Matemática novas tecnologias e educação a distância. CD-ROM.

BISQUERRA, R.; SARRIERA, J. C.; MARTINEZ, F. **Introdução à estatística: enfoque informático com o pacote estatístico SPSS**. Porto Alegre: Artes Médica; 2004.

BONILLA, M. H. O Brasil e a alfabetização digital. **Jornal da Ciência**, Rio de Janeiro, Abr, 2001. Disponível em: < http:// www.faced.ufba.br/~bonilla/artigojc.htm>. Acesso em: 12 mar 2008.

BRASIL. **Parâmetros Curriculares Nacionais: Matemática**. Brasilia: Ministério da Educação e Desporto - Secretaria do Ensino Fundamental, 1997.

Baterno, C., Estepa, A., & Godino, J. D. Evolution of students' understanding of statistical association in a computer based teaching environment. In: IASE Round Table Conferenc. Voorburg, 1991. **Proceedings…** Voorburg, The Netherlands: International Statistical Institute, 1991.

BATANERO, C.; ESTEPA, A.; GODINO, J. D.Análisis exploratório de datos: sua posibilidades em La enseñanza secundaria. **Suma**, 9, p. 25 - 31.1991.

CALDART, R. S. Elementos para a Construção de um Projeto Político Pedagógico da Educação no Campo. In: MOLINA, M. **Por uma Educação no Campo:** Contribuições para a Construção de um Projeto de Educação no Campo. Brasília, 2004.

CAMPOS, T. M. M.; CARVALHO, L. M. T. L. Interpretação de gráficos: Analisando o efeito de diferentes informações. In: Encontro de Pesquisa Educacional do Norte e Nordeste – EPENN.19.João Pessoa, 2009, **Anais...**, João Pessoa:UFPB, 2009. CD-ROM.

CAZORLA, I., CASTRO, F. O papel da estatística na leitura do mundo: O letramento estatístico. **Publ. UEPG Humanit. Sci., Appl. Soc. Sci., Linguist., Lett. Arts**, Ponta Grossa, V.16 (1). P. 45-53. Jun, 2008.

FARIAS, M.; MONTEIRO, C.; Recursos Materiais em aulas de Matemática: explorando concepções de professoras. In*:* Simpósio Internacional de Pesquisa em Educação Matemática. 2., Recife, 2008, **Anais...,** Recife: UFRPE, 2008. CD-ROM.

GARFIELD, J.; BEN-ZVI, D. **Developing students' statistical reasoning: Connecting research and teaching**. New York: Springer, 2008.

GUIMARÃES, G. L. **Interpretando e Construindo Gráficos de Barras**. 2002. 273 f. Tese (Doutorado em PSICOLOGIA COGNITIVA) – Programa de Pós-Graduação. Centro de Filosofia e Ciências Humanas. Universidade Federal de Pernambuco. Recife:, 2002.

GUIMARÃES, G. L.; GITIRANA, V.; ROAZZI, A. Interpretando e construindo gráficos. In: Reunião Anual da ANPED, 24., Caxambu, 2001, **Anais...**, Caxambu,2001. CD-ROM.

GODINO, Juan D. **Relaciones Dialécticas entre Teoría, Desarrollo y Práctica en Educación Matemática: un mera-análisis de tres investigaciones.** Granada: Universidad Granada,1995.Disponível em: http://www.ugr.es/~jgodino/dialectes.html.

KONOLD, C.; MILLER, C. (2001). **TinkerPlots**, version 0.42. Data analysis software for the middle school. Amherst, MA: University of Massachusetts.

KONOLD, C. Handling complexity in the design of educational *software* tools. In: International Conference on Teaching Statistics – ICOTS, 7.,Salvador, 2006, **Proceedings...**, Salvador, 2006. CD-ROM.

LIRA, O. **O uso de ferramentas do software TinkerPlots para a interpretação de gráficos**. 2010. 195f. Dissertação (Mestrado) Programa de Pós-Graduação em Educação. Centro de Educação. UFPE, 2010.

LIMA, R. **Introduzindo o conceito de média aritmética na 4ª série o Ensino Fundamental usando o ambiente computacional**. 2005. 156f. Dissertação (Mestrado) Programa de Pós-Graduação em Educação Matemática. PUC. SP, 2005.

LEMOS, M. P.; GITIRANA, V. Interpretação de gráficos de barras: analise a priori enquanto recurso na formação de professores. In: Encontro Nacional de Educação Matemática- ENEM. 10., Belo Horizonte, 2007. **Anais...**, Belo Horizonte, UniBH, 2007. CD-ROM.

MACÊDO, M.; MONTEIRO, C. Discursos de estudantes de uma escola rural sobre recursos no ensino e aprendizagem de matemática. In: Encontro de Pesquisa Educacional do Norte e Nordeste – EPENN.19.João Pessoa, 2009, **Anais...**, João Pessoa:UFPB, 2009. CD-ROM.

BRASIL. Ministério da Educação. **Diretrizes Operacionais para a Educação Básica nas Escolas do Campo**. Brasília, 2003

MELO S.; SILVA, V. L.; LEITÃO, V.; ALVES, I. Conceptualizando e Categorizando Recursos no Ensino de Matemática em Escolas do Campo. In: Encontro de Pesquisa educacional do Norte e Nordeste- EPENN, 18.Maceió, 2007., **Anais...**, Maceió:UFAL, 2007.CD-ROM.

MENEZES. Informática e *softwares* na Educação Matemática: impressões e inserções. In: Encontro de Pesquisa Educacional do Norte e Nordeste – EPENN.15.São Luiz , 2001, **Anais...**, São Luiz: MA, 2001. CD-ROM.

MONTEIRO, C. E. F. **Projeto conceptualizando e usando recursos no ensino de matemática em escolas rurais** – PREMATER. Relatório CNPq, 2007.

MONTEIRO, C.E.F.; SELVA, A.C.V. Investigando a atividade de interpretação de gráficos entre professores do Ensino Fundamental. In: Reunião Anual da Associação Nacional de Pós-graduação - ANPED. 24., Caxambu, 2001. **Anais...**, Caxambu: MG, 2001. CD-ROM.

MONTEIRO, C. E. F. ; ASSEKER, A. e FARIAS, M. R. Investigando as concepções e usos de recursos no ensino da matemática por professores de escolas rurais. In: Encontro de Pesquisa educacional do Norte e Nordeste- EPENN, 18.Maceió, 2007., **Anais...**, Maceió:UFAL, 2007.CD-ROM.

MONTEIRO, C.; ALVES, I. Explorando a concepção de pais sobre recursos no ensino de Matemática numa escola rural: explorando concepções de professoras. In*:* Simpósio Internacional de Pesquisa em Educação Matemática. 2., Recife, 2008, **Anais...,** Recife: UFRPE, 2008. CD-ROM.

MONTEIRO, C. E. F. ; AINLEY, J. . Investigating the interpretation of media graphs among student teachers. **International Electronic Journal of Mathematics Education**. v. 2, p. 187-207, 2007.

MORAN, J. M. **Novas tecnologias e o re-encantamento do mundo Tecnologia Educacional**. Rio de Janeiro, vol. 23, n.126, setembro-outubro 1995, p. 24-26*.*

NOVAES, D.V.; COUTINHO, C.Q.S. **Estatística para Educação Profissional**. São Paulo: Atlas, 2009.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**, 5ª edição, Editora Saraiva: São Paulo, 2002.

PASSOS, M. S. C. Navegar é impreciso: considerações sobre a formação de professores e as Tics. **Educação e Contemporaneidade**. Revista da FAEEBA. UNEB, vol 12, número 20, jul/dez, 2003.

PRATT, D. Active Graphing in a Computer-Rich Environment. In: Annual Conference of the International Group for the Psychology of Mathematics Education. 18., Lisbon, 1994., **Proceedings...**, Lisbon: University of Lisbon, 1994. CD-ROM.

PRATT, D. Young Children's Active and Passive Graphing. ***Journal of Computer Assisted Learning***, V.11, pp. 157-169, 1995.

PRENSKY, M. **Don't Bother me, Mom, I'm Learning!**:How computer and video games are preparing your kids for 21 st century success and how you can help! St. Paul, Minnesota: Paragon House, 2006.

RUBIN, A., HAMMERMAN, J., & KONOLD, C. Exploring informal inference with interactive visualization *software*. In: International Conference on Teaching Statistics – ICOTS, 7.,Salvador, 2006, **Proceedings...**, Salvador, 2006. CD-ROM.

SHAUGHNESSY, J.; GARFIELD, J.; GREER, B. Data Handling. In A. Bishop et al. (eds.), **International handbook on mathematics education***,* 205-237. Netherlands: Kluwer, 1996.

SANTOS, M.S.; GITIRANA,V. A interpretação de gráficos de barra, com variáveis numéricas, em um ambiente computacional de manipulação de dados. In: Encontro de Pesquisa Educacional do Nordeste – EPEN. 15.,Salvador, 1999. **Anais...** Salvador: Bahia, 1999. CD-ROM.

SANTOS, S; MAGINA, S. Estratégias de Interpretação Gráfica de uma Professora Polivalente ao Manipular Dados no Ambiente Computacional. **Bolema**, Rio Claro: SP, n. 29, p. 157-174, 2008.

SELVA, A. C. V. **Gráficos de barras e materiais manipulativos: analisando dificuldades e contribuições de diferentes representações no desenvolvimento da conceitualização Matemática em crianças de seis a oito anos**. Tese (Doutorado em PSICOLOGIA COGNITIVA) – Programa de Pós-Graduação. Centro de Filosofia e Ciências Humanas. UFPE. Recife,2003.

SILVA, H. ; JAMBEIRO O. ; LIMA, J. e BARNADÃO, M. A. Inclusão digital e educação para a competência informacional: uma questão de ética e cidadania. **Ciência da Informação**. Vol 34. n. 1, Jan/Abr, P. 55-62. 2005.

SILVA, J. B. **As representações sociais dos professores em classes multisseriadas sobre a formação continuada**. Dissertação (Mestrado) - Pós-graduação em Educação, UFPE, Recife, 2002.

TANNER, C.; ESTEPA, A.; GODINO, JD. A análise exploratória de dados: o seu potencial no ensino secundário. **Revista Brasileira de Educação Matemática em Ciência e Tecnologia**. p. 25-31, 1992.

TREVIZAN, S. D. P.. O que é rural? O que é urbano? E a educação*?.* Ilhéus, BA: **Forum Estadual de Educação do Campo**, 2003 (Texto de Palestra). Disponível em: < http://www.forumeja.org.br/ec/files/Texto%20Salvador%20Trevisan.pdf> Acesso em: 13 abr. 2008.

VEIGA, E. **Cidades Imaginárias : O Brasil é menos urbano do que se calcula**. São Paulo: Autores Associados, 2002.

WANDERLEY, M. (org). **Globalização e desenvolvimento Sustentável: dinâmicas sociais rurais no Nordeste brasileiro**. São Paulo: Polis; Campinas: Ceres - Centro de Educação de estudos rurais do IFCH - Unicamp, 2004.

_____, M. A emergência de uma nova ruralidade nas sociedades modernas avançadas: o rural como espaço singular e ator coletivo. **Estudos sociedades e agricultura**. v.15, out. p. 87-144. 2001.

WATSON, J. **Exploring beginning inference with novice grade7 studesnts**. Statistics Education Research Journal 7(2), 59-82, http:// ouckland.ac.nz/serj@International Association for Statistical Education (IASE/ISI), november, 2008.

ANEXO

	Universidade Federal de Pernambuco **EDUMATEC – Pós-graduação em Educação Matemática e Tecnológica**

ENTREVISTA SEMI-ESTRUTURADA

1. Qual o seu nome?
2. Qual sua idade?
3. Quanto tempo você tem de sala de aula, ensino?
4. Atualmente ensina a que ano?
5. Utiliza computador na sua vida pessoal? E na profissional? Com que frequência?
6. Você possui computador em casa?
7. A sua escola possui laboratório de informática? Você utiliza? Como? Em que momentos?
8. Quais experiências você tem com leitura e interpretação de gráficos no seu dia-a-dia? Por exemplo você costuma ver em jornais, revistas, livros didáticos etc, que contenham gráficos?
9. E você aborda a interpretação de gráficos nas suas aulas? Como, cite exemplo? Qual importância do ensino de gráficos pra você?