



Universidade Federal da Bahia
Instituto de Matemática

Departamento de Ciência da Computação

**VISUALIZAÇÃO MULTIDIMENSIONAL DE
DADOS DE SAÚDE EM DISPOSITIVOS
MÓVEIS**

Fagner Bezerra Batista

TRABALHO DE GRADUAÇÃO

Salvador
21 de Dezembro de 2018

FAGNER BEZERRA BATISTA

**VISUALIZAÇÃO MULTIDIMENSIONAL DE DADOS DE SAÚDE
EM DISPOSITIVOS MÓVEIS**

Este Trabalho de Graduação foi apresentado ao Departamento de Ciência da Computação da Universidade Federal da Bahia, como requisito parcial para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Danilo Barbosa Coimbra

Salvador
21 de Dezembro de 2018

Sistema de Bibliotecas - UFBA

Batista, Fagner Bezerra.

Visualização Multidimensional de Dados de Saúde em Dispositivos Móveis / Fagner Bezerra Batista – Salvador, 2018.

45p.: il.

Orientador: Prof. Dr. Prof. Dr. Danilo Barbosa Coimbra.

Monografia (Graduação) – Universidade Federal da Bahia, Instituto de Matemática, 2018.

1. Visualização de Dados Multidimensionais. 2. Dispositivos Móveis. 3. Dados de Saúde. I. Barbosa Coimbra, Prof. Dr. Danilo. II. Universidade Federal da Bahia. Instituto de Matemática. III Visualização Multidimensional de Dados de Saúde em Dispositivos Móveis.

CDD – XXX.XX

CDU – XXX.XX.XXX

TERMO DE APROVAÇÃO

FAGNER BEZERRA BATISTA

VISUALIZAÇÃO MULTIDIMENSIONAL DE DADOS DE SAÚDE EM DISPOSITIVOS MÓVEIS

Este Trabalho de Graduação foi julgado adequado à obtenção do título de Bacharel em Ciência da Computação e aprovado em sua forma final pelo Departamento de Ciência da Computação da Universidade Federal da Bahia.

Salvador, 21 de Dezembro de 2018

Danilo Barbosa Coimbra
Universidade Federal da Bahia

A Deus, minha amada mãe, meu pai e irmãos.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo seu cuidado e proteção que vem dando a minha família. A ajuda dos meus pais por todo suporte que me ofereceu durante minha vida acadêmica, que não foi curta, tanto financeira como psicológica para suportar os desafios a qual fui submetido. Aos meus amigos da UFBA, do curso de ciências da computação pelos momentos de risadas e de choro. Longas madrugadas foram passadas juntos conectados, conversando e resolvendo questões. Ao meu orientador Danilo Coimbra por sua competência em me ajudar no desenvolvimento deste trabalho, sempre solícito a sanar qualquer dúvida. A todos os professores que tive a oportunidade de conhecer e contribuíram para meu o conhecimento.

Não são os aplausos, nem os prêmios o que define que estamos vencendo, é a certeza de estar vivendo nosso propósito de vida.

—TIAGO BRUNET

RESUMO

A visualização de dados é utilizada amplamente pela comunidade, seja no meio acadêmico ou comercial, seja por seu potencial de sumarizar dados ou pela facilidade em interpretá-los. Contudo, é preciso estar atento para qual tipo de técnica utilizar quando consideramos diferentes características de bases de dados, pois uma escolha equivocada pode ocasionar dificultar o processo de aquisição de informação. No presente trabalho são investigadas técnicas de visualizações multidimensionais empregadas no domínio da saúde, com foco nos dispositivos móveis. Tais visualizações podem auxiliar na captura e extração de informação em ambientes remotos, contudo, é necessário que a responsividade da visualização garanta ao usuário todas as funcionalidades da aplicação, inclusive recursos de interatividade. Dentre as contribuições deste trabalho estão: *i)* uma revisão sistemática acerca do tema supracitado, *ii)* a investigação de técnicas de visualização multidimensional para dados do domínio da saúde em dispositivos móveis, e *iii)* customizações de responsividade e interação em visualizações multidimensionais de modo a garantir a análise exploratória em dispositivos móveis.

Palavras-chave: visualização de dados, dispositivos móveis, visualização de dados multidimensionais, dados de saúde.

ABSTRACT

Data visualization is widely used by the community, whether in the academic or commercial sphere, or for its potential to summarize data or for the ease in interpreting them. However, it is necessary to be aware of which type of technique to use when considering different characteristics of databases, since a wrong choice can cause difficulties in the information acquisition process. In the present work we investigate techniques of multidimensional visualizations used in the health domain, focusing on mobile devices. Such visualizations can aid in the capture and extraction of information in remote environments, however, it is necessary that the responsiveness of the visualization guarantees to the user all the functionalities of the application, including interactivity resources. Among the contributions of this work are: *i)* a systematic review of the above theme, *ii)* the investigation of multidimensional visualization techniques for health domain data on mobile devices, and *iii)* customizations of responsiveness and interaction in multidimensional views to ensure exploratory analysis on mobile devices.

Keywords: data visualization, mobile device, multidimensional data visualization, health data.

SUMÁRIO

Capítulo 1—Introdução	1
1.1 Objetivo	2
1.2 Estrutura	2
Capítulo 2—Revisão Sistemática da Literatura	4
2.1 Considerações Iniciais	4
2.2 Planejamento da Busca	5
2.3 Definição dos Estudos	6
2.4 Uso do StArt para Análise	7
2.5 Análise dos dados - resultados	8
2.6 Análise dos artigos selecionados	10
2.7 Discussão	17
2.8 Considerações Finais	22
Capítulo 3—Fundamentação Teórica	23
3.1 Considerações Iniciais	23
3.2 Visualização de Dados Multidimensionais	23
3.2.1 Star-Plot	25
3.2.2 Coordenadas Paralelas	27
3.2.3 Tree-Map	29
3.3 Considerações Finais	30
Capítulo 4—Aplicando Visualizações Multidimensionais Responsivas em Dados de Saúde	31
4.1 Considerações Iniciais	31
4.2 Base de Dados	31
4.3 Estudos de Casos com Visualizações Multidimensionais	32
4.3.1 STARPLOT	33
4.3.2 COORDENADAS PARALELAS	34
4.3.3 TREEMAP	37
4.3.4 RADAR PLOT	39
4.4 Considerações Finais	41
Capítulo 5—Conclusões	42

LISTA DE FIGURAS

2.1	Exemplo de busca no Web of Science.	6
2.2	Tela inicial do StArt.	7
2.3	Tela do StArt sumarização das bases de dados em um gráfico.	8
2.4	Nuvem de palavras dos autores das duas fontes.	10
2.5	Exemplo dos dados plotados usando o GeoHealth (SA et al., 2016).	18
2.6	Exemplo do hFigure, medição das variáveis separadas por grupos (LED- ESMA; AL-MUSAWI; NIEMINEN, 2016).	18
2.7	Visualização em um smartphone dos alimentos e sua importância de con- sumo (BUTSCHER et al., 2016).	19
2.8	Mapa de calor onde a cor azul forte representa as tags mas citadas no Twitter (HINGLE et al., 2013).	20
2.9	Mapa de calor para visualização do espalhamento da dengue (LATIF et al., 2015).	20
2.10	Evolução do quadro do paciente sendo mostrada num gráfico em círculo (STEVENSON et al., 2014).	21
2.11	O pino mais vermelho representa o rio mais poluído (PINGALI, 2016).	22
3.1	Visualização em tabela da bolsa de valores.	25
3.2	Imagem retirada do artigo Saary (2008) realizando uma comparação do gráfico star-plot, gráfico star-plot figura A e o gráfico de barras tradicional figura B.	26
3.3	Anatomia do Gráfico em radar.	27
3.4	Comparação entre uma pequena tabela e as coordenadas paralelas ((MUNZ- NER, 2015)).	27
3.5	Correlação entre os eixos vizinho das coordenadas paralelas ((MUNZNER, 2015)).	28
3.6	Visualização de centenas de linhas em (a), e milhares de linhas em (b) ((MUNZNER, 2015)).	29
3.7	Regiões da treemap mapeadas por retângulosSaary (2008)	30
4.1	Método do javascript onresize sendo utilizado no redimensionamento de tela no tempo de 100ms.	33
4.2	Figura A - visualização responsiva utilizando o método windows.onresize em um Ggalaxy 5s. Figura B sem o método windows.onresize, apresen- tando alguns eixos fora do espaço visual.	34

4.3	Figura A - visualização responsiva utilizando o método <i>resize</i> em um Iphone 5s. Figura B sem o método <i>resize</i> no Iphone 5s, apresentando muitos eixos fora do espaço visual; Figura C visualização responsiva utilizando o método <i>resize</i> em um LG Optimus L70; Figura D sem o método <i>resize</i> no LG Optimus L70 com eixos fora do espaço visual.	35
4.4	Figura A - visualização responsiva utilizando o método <i>windows.onresize</i> adaptando-se ao tamanho de tela no Ipad Pro. Figura B sem o método <i>windows.onresize</i> de modo a não variar o tamanho do gráfico em uma tela de maior dimensão no Ipad Pro.	35
4.5	Comparação entre dois meses do ano 2018 com polígonos sobrepostos representando meses diferentes, junho em verde e janeiro em azul.	36
4.6	Interação ao passar o mouse, revelando o atributo e o seu valor no gráfico star-plot no ano de 2018 figura A e 2017 figura B.	36
4.7	Coordenadas paralelas visualizada em um Ipad Pro do paciente Veronica H.	37
4.8	Coordenadas paralelas, em que cada ano apresenta cores diferentes.	37
4.9	Interação com a brush, selecionando as linhas do ano 2017, em cor laranja (Figura A), e as linhas do ano 2018 em cor azul (Figura B).	38
4.10	Visão inicial da tree-map (nó raiz) em A, possibilitando a escolha de um dos pacientes para ser visualizada nos nós seguintes; próximo nível da treemap apresenta os anos em B; no nível seguinte apresenta os meses em C; e por fim o último nível os atributos em D.	38
4.11	Visão das telas em um dispositivo móvel - Galaxy S5, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.	39
4.12	Visão das telas em um dispositivo móvel - Iphone 5S, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.	39
4.13	Visão das telas em um dispositivo móvel - LG Optimus L70, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.	39
4.14	Visão das telas em um dispositivo móvel - IPAD PRO, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.	40
4.15	Visão Global da aplicação possibilitando ao usuário escolher quais anos, semestres e pacientes pretende analisar.	40
4.16	Radar visualizado em um computador pessoal A com polígonos sobrepostos, e no dispositivo móvel Galaxy S5.	41
4.17	Visualização de radar em um LG Otimus L70 A, Iphone 5S B e Ipad Pro C.	41

LISTA DE TABELAS

2.1	Sumarização dos artigos avaliados segundo os critérios da RS.	8
2.2	Relação dos países e publicações dos artigos.	9
2.3	Relação das publicações dos autores.	9
2.4	Tabela de referências.	10
3.1	Tabela dos atributos.	26
4.1	Modelos dos dispositivos móveis usados nos testes de responsividade. . .	33

Capítulo

1

INTRODUÇÃO

Grandes quantidades de dados heterogêneos são gerados no sistema de saúde com o passar dos anos (SOMAN; KUMAR; KUMAR, 2016). Segundo a Pesquisa Nacional de Saúde¹ (PNS), e divulgado pelo Instituto Brasileiro de Geografia Estatística² (IBGE) em 2013 142,8 milhões de residentes no Brasil fizeram consultas médicas, as quais geraram inúmeros dados a serem analisados, seja pelos formulários de cadastro ou resultados de exames. À medida que a produção desses dados aumentam, tornam-se cada vez mais complexas a tarefa de tratá-los, analisá-los e extrair informações relevantes para o processo de tomada de decisão.

Visando potencializar o entendimento dos dados analisados e, por consequência, a extração de informações, tornam-se adequados o uso de métodos de análise de dados. Em específico, uma das possibilidades encontra-se na área de visualização de dados, que é definida por Telea (2007) como uma representação visual de um domínio usando para isso, gráficos, imagens, animações (interações) e o comportamento dinâmico de conjuntos de dados grandes e complexos que representam sistemas, eventos, objetos e conceitos. Existem diversas técnicas de visualizações que serão abordadas neste trabalho, para que seja possível extrair as informações presentes nos dados, de modo a permitir que os profissionais do domínio da saúde tomem decisões mais acertadas.

Um grande desafio da visualização é trabalhar com dados multidimensionais, pois segundo Chan (2006) não é uma tarefa fácil mapear dados multidimensionais para um espaço visual bidimensional, como telas e monitores. Este processo envolve associações dos atributos dos dados para entidades gráficas, requerendo sempre um extremo cuidado para evitar sobrecarga a capacidade cognitiva do observador. Ainda, os atributos devem ser selecionados cuidadosamente de modo a permitir serem fáceis de serem analisados. É importante ressaltar a possibilidade de cada dimensão ser analisada pelo usuário separadamente, a fim de tornar a análise mais holística.

¹Disponível em:<https://www.pns.icict.fiocruz.br/>. Acessado em: 05/12/2018

²Disponível em:<https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2015/06/02/ibge-revela-como-anda-a-saude-do-brasil.htm>. Acessado em: 09/12/2018

Os trabalhos que abordam técnicas de visualização de dados no domínio da saúde para dispositivos móveis, em sua maioria, apresentam técnicas tradicionais de visualização. Dentre essas técnicas podemos salientiar os trabalhos com mapa de calor (*heat map*) (HINGLE et al., 2013; LATIF et al., 2015) que geralmente utilizam essa técnica com escala de cores em mapas georreferenciado para mapear concentrações e surtos de doenças. Visualizações espaciais (DASGUPTA, 2010; PINGALI, 2016; KOK et al., 2016) também são comuns, pois tem o intuito de apresentar a posição georreferenciada de eventos dos locais que apresentam pessoas com doenças, como diabete ou hipertensão. Por fim, a visualização de radar (LEDESMA; AL-MUSAWI; NIEMINEN, 2016; LIANG et al., 2016) é adequada para apresentar dados relacionadas a saúde dos pacientes, como o nível de açúcar no sangue, a pressão arterial, qualidade do sono e testes de laboratórios.

Uma lacuna encontrada nesses trabalhos é o pouco uso de visualizações multidimensionais para tratar dados desse domínio, em especial visualizações responsivas, que possam ser apresentadas e permitam interação em dispositivos móveis. Portanto, faz-se necessário explorar e investigar como tais visualizações podem auxiliar na extração de informações em dispositivos móveis, quando o domínio dos dados estão relacionados à saúde.

1.1 OBJETIVO

Desse modo, foi estabelecido o seguinte objetivo para o presente trabalho:

Investigar o uso de técnicas de visualização de dados multidimensionais de modo a auxiliar o usuário a extrair informações relacionadas ao domínio da saúde e que possam ser responsíveis, possibilitando sua utilização em dispositivos móveis.

A fim de atender este objetivo, foram realizadas as seguintes etapas no decorrer do desenvolvimento da manonografia: condução de uma revisão sistemática da literatura no contexto de visualização de dados no domínio da saúde em dispositivos móveis, estudo dos fundamentos teóricos necessários para a implementação das técnicas elencadas, aplicação e discussão das técnicas de visualização multidimensionais em uma base de dados do domínio da saúde.

1.2 ESTRUTURA

Este trabalho de conclusão de curso está organizado da seguinte forma:

- No Capítulo 2 é apresentado a Revisão Sistemática da literatura, exibindo o mecanismo de busca sistemática, com o intuito de explorar os artigos relacionados ao tema em questão;
- No Capítulo 3 é introduzida a fundamentação teórica, aprofundando nos conceitos das técnicas de visualizações multidimensionais.
- No Capítulo 4 são apresentados e discutidos os resultados obtidos, detalhando a metodologia e escolha das técnicas de visualização multidimensional implementadas, assim como a base de dados escolhida.

- No Capítulo 5 são apresentadas as conclusões do trabalho.

REVISÃO SISTEMÁTICA DA LITERATURA

2.1 CONSIDERAÇÕES INICIAIS

Este capítulo tem como propósito falar sobre a revisão sistemática, que é um método amplamente utilizado no meio científico para obter o máximo de quantidade de informação sobre algum tema, neste caso a construção de um aplicativo móvel para a visualização de dados aplicado à saúde, auxiliando assim no desenvolvimento do respectivo trabalho.

A revisão sistemática da literatura é um processo formal e metuculoso o que significa que o estudo, a revisão, foi feita seguindo um protocolo onde os passos a serem seguidos foram bem definidos. Existem outros tipos de revisões como a tradicional, que não utiliza critérios sistemático claros para a busca e análise da literatura. Nesta técnica de revisão não é aplicada estratégia de busca sofisticada e exaustiva, e a seleção dos estudos e interpretações das informações podem estar sujeitas ao julgamento dos próprios autores, (CORDEIRO¹ et al., 2007). Inicialmente, na revisão sistemática é definida uma pergunta principal para guiar a pesquisa relacionada ao tema, e diversas perguntas secundárias foram criadas para reforçar a pesquisa em torno da pergunta principal. Logo em seguida são utilizadas as bases de dados disponíveis na web para a pesquisa de artigos que tratem do tema em questão. Feito isto, vem o processo mais apurado para a análise dos artigos, com o objetivo principal de responder as perguntas primária e secundária.

O resultado final de uma revisão sistemática apresenta algumas informações estatísticas com gráficos, para ter uma visão mais detalhada dos artigos e sua importância como: o número de publicações anuais dos artigos, países onde mais se publicaram, autores que mais contribuíram, número de citações do respectivo artigo, o nível de relevância dos artigos publicados ao longo de um período, além de outras informações secundárias, porém importantes para o desenvolvimento do trabalho para o qual foi planejado.

Os próximos capítulos detalham todos os processos realizados na Revisão sistemática da literatura com o intuito de chegar ao estado da arte.

2.2 PLANEJAMENTO DA BUSCA

Durante toda a fase da revisão sistemática foi utilizado a ferramenta StArt, para dar o suporte necessário e organização do respectivo trabalho.

Nesta fase inicial da pesquisa é preenchido um formulário, que é o protocolo a ser seguido no StArt, onde todas as características da busca são inseridas como: o objetivo da pesquisa, a pergunta principal e as perguntas secundárias, o repositório de busca a ser usado, as palavras chaves, a língua padrão dos artigos analisados, os critério de inclusão e exclusão para separar os artigos a ser analisados, os artigos também recebem uma pontuação de relevância dependendo do local onde é encontrado a palavra chave no artigo, como no título ou resumo ajudando na sua priorização. Para início dos estudos foram formuladas algumas questões para nortear a pesquisa e seleção dos artigos, as perguntas foram divididas em primária/principal e secundária.

Foi realizada a seguinte pergunta principal:

- Quais as técnicas de visualização são utilizadas para mapear dados da área da saúde em aplicativos móveis?

Foi definido logo em seguida as perguntas secundárias que estão associadas a pesquisa com o intuito de validar a pergunta principal. Este conjunto de perguntas tem como objetivo reforçar a análise dos artigos para melhor avaliá-los.

- A visualização projetada facilita na tomada de decisão?
- Quais são as limitações das técnicas?
- Faz sentido utilizar essa técnica para mapear esse conjunto de dados?
- Quais os mecanismos de avaliação da visualização?
- Quais o(s) grupo(s) de pesquisador(es) que possuem publicações relevantes nessa área?
- Qual a frequência das publicações?

Foi definido como idioma o inglês desse modo artigos escritos em outras línguas não serão consideradas. Em seguida foram definidas as palavras chaves levando em conta a pergunta principal, a pergunta secundária e o objetivo de estudo:

- Visualization
- Health
- Mobile

Definiu-se duas bases de dados como o repositório online de busca dos artigos o que são bem conhecidas entre a comunidade científica:

- Scopus ¹
- Web of Science ²

2.3 DEFINIÇÃO DOS ESTUDOS

Foram criadas inicialmente algumas strings de busca para fazer uma análise mais ampla sobre o estudo em questão, após foi definida a seguinte string de busca: (*health AND visualization AND mobile*) por retornar um valor satisfatório de artigos nas duas bases de dados utilizadas. Esta string compõem as palavras chaves que são inseridas no campo de busca das duas bases de dados, sendo que a base Scopus retornou um número maior de artigos do que a Web of Science. A Figura 2.1 ilustra a pesquisa sendo realizada no Web of Science, para fazer as devidas análises.



Figura 2.1 Exemplo de busca no Web of Science.

Com a string de busca definida e os resultados obtidos, os artigos retornados, foi necessário fazer mais um tipo de refinamento dos artigos, pois eles ainda poderiam não estar relacionados com a pesquisa. Logo, para o próximo passo são definidos os critérios de inclusão e exclusão, parte muito importante da revisão sistemática, para selecionar os documentos mais relevantes e de acordo com o objetivo da pesquisa, incluindo por tanto todos os artigos a partir de 2010 por considerar a tecnologia mobile mais desenvolvida, principalmente com o advento da web 2.0 que popularizou-se a partir de 2004, (ALEXANDER, 2006). Logo foi incluído na análise apenas documentos que atendessem às seguintes condições:

- Lidar com alguma técnica de visualização no contexto da saúde em dispositivos móveis
- Estudos feitos sobre desenvolvimento mobile com visualização
- Documentos apenas em inglês ou português

Os documentos que foram descartados foram baseados nos seguintes critérios de exclusão:

¹Disponível em: <https://www.scopus.com/>. Acessado em: 10/07/2018

²Disponível em: <https://login.webofknowledge.com/>. Acessado em: 10/07/2018

- Screenshot da imagem da tela mobile
- Artigos publicados antes de 2004

O primeiro critério de exclusão foi utilizado pois não foi considerado screenshot das telas como uma técnica de visualização.

2.4 USO DO START PARA ANÁLISE

O StArt foi a ferramenta utilizada para desenvolver a revisão sistemática por apresentar os recursos necessário para tal desenvolvimento. O stArt é uma ferramenta, desenvolvida pela Universidade Federal de São Carlos, para apoio ao processo de revisão sistemática, cujo o objetivo é dar suporte ao pesquisador para a revisão sistemática da literatura, esta ferramenta é organizada em três etapas: o planejamento, execução e sumarização. No planejamento é identificado a necessidade da revisão e preenchido a página inicial, que é o protocolo da RS (revisão sistemática), na execução adiciona e avalia os artigos que compõe a RS e extrai informações daqueles que são relevantes ao tópico da pesquisa em questão, na sumarização são apresentados gráficos, tabelas e dados estatísticos em relação aos artigos que dão uma visão geral sobre a RS, a Figura 2.2 logo abaixo ilustra a tela inicial da ferramenta StArt.



Figura 2.2 Tela inicial do StArt.

O primeiro contato para a classificação dos artigos baseados nos critérios de inclusão e exclusão é realizado pelo título onde clicando-se duas vezes abrir-se-a uma tela que possibilitará o acesso a informação pertinente para a análise como: os autores, o local da publicação e o resumo (*abstract*). Outra funcionalidade importante da ferramenta é a possibilidade de organizar os artigos por ano de publicação e por *score* para facilitar na sua seleção, ela disponibiliza também a classificação desses periódicos como duplicados evitando assim a releitura do mesmo e concentrando-se apenas nas potências publicações para estudo. Na sumarização são mostradas uma série de gráficos apresentando informações que vem a ajudar na revisão sistemática,

como pode ser visto na Figura 2.3 onde é visualizado dados dos dois repositórios de busca, scopus que é representado pela cor vermelha e web of science representado pela cor azul; logo em seguida um gráfico, gerado por padrão pela ferramenta, preenchido pela cor amarela correspondendo os 28 % dos artigos classificados como aceitos para a leitura. Por meio desta ferramenta fica fácil perceber que o scopus possui mais publicações retornadas com a string de busca mencionada do que a web of science.

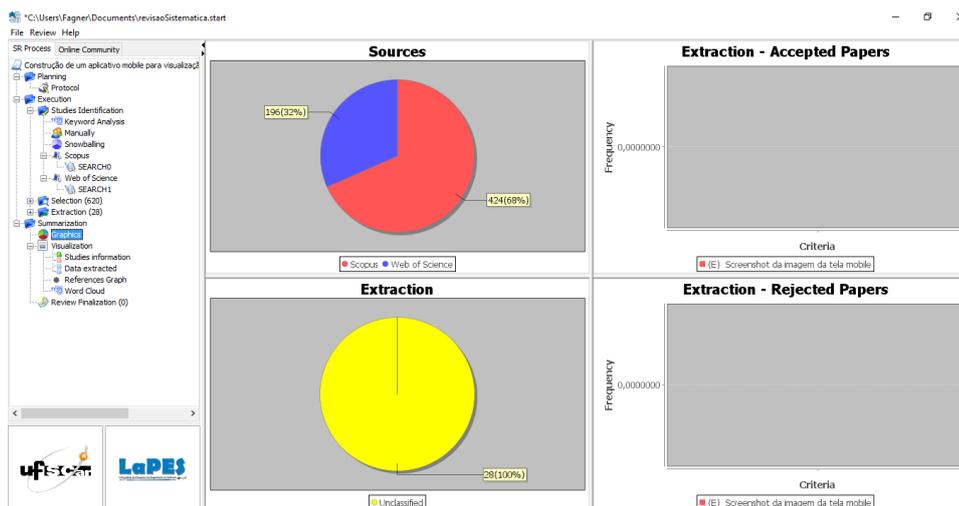


Figura 2.3 Tela do StArt sumarização das bases de dados em um gráfico.

2.5 ANÁLISE DOS DADOS - RESULTADOS

Nesta seção serão apresentadas algumas informações relevantes obtidas com o uso do StArt para fazer as devidas análises dos artigos. A base de dados Scopus retornou um número maior de artigos 424, enquanto que a web of science retornou um número de 196 artigos totalizando desta maneira 620 artigos para a análise, deste total apresentou-se alguns duplicados, que foram identificados pela ferramenta e retirado da pesquisa evitando assim a releitura do mesmo, o número de artigos duplicados foram 30, portanto o total de artigos para as devidas análises foram de 590. Segue abaixo a Tabela 2.1 que sintetiza melhor o que foi falado.

Base de Dados/Critérios	Número de Artigos
Scopus	424
Web of science	196
Total de artigos	620
Artigos duplicados	30
Total para análise	550

Tabela 2.1 Sumarização dos artigos avaliados segundo os critérios da RS.

Logo após feitas algumas análises das publicações foram apresentadas algumas tabelas com informações relevantes para obter uma visão mais ampla do tema em questão, que é relacionado a visualização de dados na saúde. Por meio da base de dados *web of science* foi criada uma tabela, Tabela 2.2 onde apresenta a relação de publicações dos artigos nos países em questão, onde pode ser observado um maior interesse por esse tema nos Estados Unidos, uma vez que apresentou a maior parte das publicações; na base de dados *scopus* a Tabela 2.3 apresenta a relação das publicações feitas pelos autores. Com a ferramenta StArt foi possível gerar uma nuvem de palavras como pode ser visto na Figura 2.4 contendo todos os autores das duas bases de pesquisa, *web of science* e a *scopus*, obtendo como retorno 32% e 68% de artigos respectivamente envolvidos nos artigos aceitos para leitura (*word cloud*) onde pode ser observado que o autor *Krejcar* aparece com maior destaque. A Tabela 2.4 sumariza os artigos selecionados para a leitura na íntegra após todas as etapas seguidas da RS.

País	Número de publicações	Valor em porcentagem %
EUA	53	26,9
GERMANY	15	7,6
PEOPLES R CHINA	12	6,0
CANADA	11	5,5
ENGLAND	10	5,1
AUSTRALIA	9	4,6
CZECH REPUBLIC	9	4,6
SOUTH KOREA	9	4,6
SPAIN	9	4,6
INDIA	7	3,5

Tabela 2.2 Relação dos países e publicações dos artigos.

Autores	Número de publicações	Valor em porcentagem %
JANCKULIK	7	3,5
KREJCAR O	7	3,5
CHITTARO L	3	1,5
KOO B	3	1,5
MOTALOVA L	3	1,5
BATANGAN D	2	1,0
SHON T	3	1,5
ALTEROVITZ G	2	1,0
AMARRA AC	2	1,0
BANOS O	2	1,0

Tabela 2.3 Relação das publicações dos autores.

Depois de implementado a plataforma, outros tipos de dados relacionado a saúde poderá ser colocado para ser visualizado.

- **Técnica:** Utilizada técnica de visualização Radar
 - **Objetivo:** Ajudar os usuários no processo de tomada de decisão e avaliar a situação geral a respeito da saúde e se o programa desenvolvido tem ajudado no fornecimento de benefícios em relação a sua saúde.
 - **Método:** Para o desenvolvimento dessa biblioteca em javascript, foi utilizado a metodologia Extreme programming. O principal requerimento foi fornecer uma visualização que representa a mudança geral na situação da saúde. A implementação desta biblioteca foi possível graças a um grupo com um grande conhecimento em: ciências da saúde, processamento de sinal, design de usuário, engenharia de software e aprendizado de máquina.
2. **Título:** *The Care and Condition Monitor: Designing a Tablet Based Tool for Visualizing Informal Qualitative Healthcare Data.*
- **Dados:** Dados coletados pelos cuidadores em contato com pacientes internados, especialista e não especialista; dados informais , qualitativos e arbitrário.
 - **Técnica:** Utiliza técnica de visualização Radar.
 - **Objetivo:** Usar ferramenta de análise visual para criar um holístico e multidimensional entendimento das condições e saúde dos pacientes internados com problema cognitivo mental e físico, melhorar as interpretações de mudanças sutis no quadro do paciente ao longo do tempo e auxiliar nas decisões médicas usando para isto tais informações.
 - **Método:** Utilizando o Monitoramento de condição e saúde, sigla em inglês(CCM) que é um tablet baseado em ferramentas de visualização analítica que usa uma simples, em inglês, *voting strategy* para coletar e estruturar dados qualitativos e informais em relação a saúde em um formato de visualização circular para representar a informação.
3. **Título:** *Georeferenced and secure mobile health system for large scale data-collection in primary care.*
- **Dados:** Dados obtidos por meio de um *smartphone* (celular inteligente) com um formulário a ser respondido. Tipo de dados: arbitrário.
 - **Técnica:** Visualização de dados georreferenciado.
 - **Objetivo:** Desenvolver um sistema mobile (GeoHealth) que supere e resolva os problemas mencionados no texto, como substituir o processo de obtenção de dados no formato de papel pelo uso do telefone móvel, e implementar esta solução em lugares menos privilegiados (comunidades) na maior cidade do brasil, São Paulo.
 - **Método:** A solução compreende três abordagens: um servidor de aplicação com um banco de dados contendo condições de saúde das famílias; um web browser executando ferramenta de visualização para gerenciar as tarefas e;

um dispositivo de obtenção de dados (celular inteligente) para registrar e georreferenciar os dados relacionado a saúde das famílias. Um framework de segurança dos dados foi implementado para certificar-se da segurança dos dados guardado localmente e transmitido por meio da rede pública.

4. **Título:** *Real-Time Health Data Acquisition and Geospatial Monitoring: A Visual Analytics Approach.*

- **Dados:** Os dados são obtidos com as informações prestadas pelos pacientes ao entrar em qualquer serviço médico, foi utilizado o postgre como banco de dados; Tipo de dados arbitrário.
- **Técnica:** Heat Map, height-map, contour-plot.
- **Objetivo:** Solucionar o problema do armazenamento de dados eficientemente e seu devido compartilhamento dos dados. Para isto propuseram o desenvolvimento de um sistema de aquisição de dados (DAS) e um sistema de monitoração de doenças em tempo real (RTDMS) que adquire e armazena os dados do paciente em tempo real rapidamente e eficientemente, além de prever os surtos de qualquer doenças usando técnicas de visualização e monitoração geoespaciais.
- **Método:** Os dados obtidos com a prévia entrevista realizada são armazenadas em uma banco de dados(DAS) postgre onde em conjunto com o RTDMS(um sistema que faz extrações e análise de tendências ocultas de doenças , prover visualização em tempo real dos dados no mapa) usando a técnica de geocondig onde transforma a string do endereço do paciente em coordenadas geográfica (latitude, longitude) somada ao *shapefile* ³, onde contém informações como nome da área; geometria; área; população etc. É gerado um mapa com um marcador da região dos respectivos pacientes, desta maneira torna-se possível o monitoramento em tempo real dos pacientes com determinadas doenças, além de gerar um mapa de calor (heat map) para mapear o espalhamento das doenças.

5. **Título:** *Implementation of u-RPMS using Google Chart in hybrid application for visualization of patient's biometric information.*

- **Dados:**Dados obtidos de diversos sensores instalados no corpo humano, no máximo 12 sensores são usados; Dados discretos e categóricos.
- **Técnica:**Gráficos bidimensionais usando o google chart.
- **Objetivo:**Permitir que qualquer usuário possa visualizar e entender facilmente sua própria informação biométrica (pressão arterial , nível de açúcar no sangue, ECG) usando os dispositivos mobile.
- **Método:** Uso de sensores externos (*wearable*) instalados no corpo humano, alguns internos, usado na medicina para monitorar em tempo real

³Disponível em: <https://doc.arcgis.com/pt-br/arcgis-online/reference/shapefiles.htm>. Acessado em: 27/07/2018.

condições "emergências" dos pacientes, o valor medido do paciente é guardado em um banco de dados no formato XML; o mesmo é gerenciado por meio de um sistema com uma boa interface gráfica.

6. **Título:** *Android-Based Disease Monitoring.*

- **Dados:** Os dados são gerados de diversas organizações governamentais como clínicas municipais, hospitais estaduais e centro de controle de doenças por meio de uma rede do ministério da saúde; Dados contínuos e discretos.
- **Técnica:** Dados espaciais bidimensionais usando Mapa e Modelo de predição.
- **Objetivo:** Notificar o público no momento oportuno por meio de dados confiáveis, o espalhamento de doenças acometidas em determinadas áreas evitando desta maneira um surto repentino de doenças. Provocando uma maior consciência da população em relação às doenças, sendo até mesmo possível o fechamento das escolas para evitar uma maior infecção.
- **Método:** Usa modelo matemático de predição para prever determinadas doenças, como a leptospirose relacionando com os dados obtidos da água usada. Notifica por meio do aplicativo as doenças infecciosas detectadas evitando surtos. Dados visualizados em um mapa, onde a cor dos marcadores são: vermelho (muito sério), amarelo (não é sério) e verde (ainda não é considerado como um surto).

7. **Título:** *A Framework for Mobile Based Geospatial Public Health Management*

- **Dados:** Integração dos dados obtidos de diversas regiões, das organizações públicas e privadas de saúde, disponibilizando os dados conforme a política de privacidade.
- **Técnica:** Dados espaciais bidimensionais utilizando Mapa.
- **Objetivo:** Detectar tendências, padrões de determinados surtos de doenças nas regiões (territórios) baseado em dados espaciais das regiões que foram disponibilizados por pessoas competentes (pessoas envolvidas com a saúde e profissional de TI), deste modo rastreando o espalhamento da doença e possivelmente sua causa e regiões que obteve um maior índice das doenças envolvidas.
- **Método:** Profissionais da saúde uma vez detectado determinada doença (que estar entre os surtos) poderá fazer o *upload* dos dados pelo dispositivo móvel, de modo a permitir as marcações nas respectivas regiões com a cor vermelha; quanto maior a frequência nesta região mais acentuada a cor será. Será realizado uma integração de informação da saúde pública com dados espaciais com o uso de *WebServices* onde será observado a arquitetura: Serviço de dados de saúde espacial, serviço de dados de mapa espacial e Serviço integrado de gerenciamento de dados de saúde pública.

8. **Título:** *Effective Visualization of Long Term Health Data to Support Behavior Change.*

- **Dados:** Os dados foram obtidos por meio dos participantes separados em grupos diferentes e respondendo a perguntas produzidas pelos pesquisadores da universidade de Kaiserslautern; Dados discretos.
- **Técnica:** Gráfico de barra e de pontos multidimensional.
- **Objetivo:** Desenvolver por meio da técnica de visualização uma mudança comportamental dos participantes em relação a sua saúde, bem como o consumo de alimento que pudessem proporcionar um bem-estar maior para eles e conseqüentemente melhorar a qualidade de sua saúde. Os dados precisavam ser visualizados da melhor forma para que as decisões fossem tomadas da melhor maneira possível.
- **Método:** Utilizado o gráfico de barras com a informação de três cores diferentes, onde cada uma estava relacionada o quanto recomendável eram os hábitos medidos dos participantes, como a quantidade de horas de sono, (vermelho: não recomendável, amarelo: ainda pode ser recomendado, verde: recomendável). Em relação ao gráfico de pontos, no eixo x tinha-se o tempo (semanas) medido, no eixo y os dados e o gráfico era preenchido pelas três cores citadas acima.

9. **Título:** *Collection and Visualization of Dietary Behavior and Reasons for Eating Using Twitter.*

- **Dados:** Os dados são obtidos de um aplicativo de mídia social conhecido como *Twitter* onde são analisados hashtags relacionados a comida e bebida dos participantes envolvidos na pesquisa.
- **Técnica:** heat map, Gráfico de linhas onde é mapeado quinze variáveis, mapa de frequência.
- **Objetivo:** Gerar uma consciência nos hábitos alimentares da população dos Estados Unidos, onde uma boa parte dessa população apresenta obesidade devido há uma má alimentação, usando para isso técnica de visualização de dados de maneira simples e de fácil interpretação relacionado a comida e bebidas além de dicas sobre atividades físicas.
- **Método:** Foi utilizado o heat map para mapear a frequência das hashtags reportados pelos participantes da pesquisa, quanto maior a frequência mais concentrada a cor; sobre o mapa de frequência, quanto maior a frequência das hashtags mais concentrado estará no centro do mapa, e a menos frequente localizar-se-a na periferia do mapa; o gráfico de linhas mapeia as oito variáveis do estudo identificadas cada uma por uma cor diferente.

10. **Título:** *Integration and Visualization Public Health Dashboard: The medi+board Pilot Project.*

- **Dados:** Obtidos de diversas fontes, como mídias sociais, pesquisas online, sensores entre outros.
- **Técnica:** Visualização com técnicas bidimensionais e multidimensionais
- **Objetivo:** Simplificar a tarefa de investigação e controle de doenças infecciosas por profissionais da saúde

- **Método:** Para a monitoração e detecção de múltiplas fontes de dados foram requeridos diferentes métodos computacionais de acordo com sua estrutura e confiabilidade, onde os sinais recentemente identificados em cada fluxo de dados são validados pelo sistema, no qual o coeficiente de confiabilidade é ajustado; uma vez que o processamento de uma específica fonte for realizada, os resultados serão cruzados com os outros sinais de dados. O sistema foi desenvolvido para aplicação web html5/javascript permitindo o funcionamento em dispositivos móveis.
11. **Título:** *Lightweight Visual Data Analysis on Mobile Devices - Providing Self-Monitoring Feedback.*
- **Dados:** Obtidos por questionários, pelo sensor; dados arbitrários e contínuo.
 - **Técnica:** gráfico de barras horizontais multidimensionais.
 - **Objetivo:** : Visualizar dados de maneira multidimensional, desenvolvendo e testando empiricamente a eficácia das ferramentas para dispositivos móveis na intervenção em tempo real que visam uma alimentação saudável e atividade física, usando os smartphone e monitoramento corporal.
 - **Método:** Foi utilizado as ferramentas para monitoramento das atividades físicas, informações nutricionais, questionários, notificações, feedback, intervenções e uma ferramenta para armazenamento dos dados. Os dados são mostrados em um gráfico de barras horizontais onde a altura(largura) da barra representa um valor, existem barras mais larga que outras, as cores também são levadas em conta onde a cor verde é o recomendável.
12. **Título:** *mHealth Visual Discovery Dashboard.*
- **Dados:** Dados obtidos de um dispositivo móvel com um sensor; dados contínuos.
 - **Técnica:** Gráfico de linha com múltiplas variáveis.
 - **Objetivo:** Desenvolver um sistema de visualização que oferece uma intuitiva visualização de dados temporais medidos pelo sensor nos dispositivos móveis e oferecer suporte às várias técnicas de interação para um padrão de exploração dos dados.
 - **Método:** Para que fosse possível uma maior interação nos dados com um alta resolução, foi necessário escalar um grande conjunto de dados, para isso foi introduzidas várias camadas de cache. Foi usado um banco de dados relacional (SQLite) para guardar os dados crus e os valores-chaves para armazenar dados remodelados; os dados temporais foram pré-processados e manipulados com o pacote do python Pandas. Os dados calculados foram transmitidos para o cliente por meio de um *WebSockets* e renderizado pelo navegador com o React.js para uma maior performance.
13. **Título:** *Cloud Computing and Crowdsourcing for Monitoring Lakes in Developing Countries Sahithi P, Inventure Academy, Bangalore.*

- **Dados:** Os dados são obtidos de sensores implantados no rio, por meios de fotos e comentários do usuário; Dados são contínuos e arbitrário.
 - **Técnica:** Visualização de dados geográficos
 - **Objetivo:** Realizar a medição/monitoramento dos rios da Índia por meio de dados recebidos por sensores implantados nos rios e por participação dos usuários para que por meio da visualização desses dados em um mapa, governos ou pesquisadores possam realizar estudos e tomar decisões.
 - **Método:** Foi criado um sistema integrado de dados, onde os resultados da medição do nível de poluição do rio e as informações como fotos prestadas pelo usuário fossem enviadas para a nuvem (central cloud platform) no formato XML; também foi construída um aplicação em nuvem.
14. **Título:** *SleepExplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors.*
- **Dados:** Obtidos da API pública do fibit e formulário google; Dados categóricos
 - **Técnica:** Analisar a relação entre dados do sono e fatores contextuais que pode impactar o sono, por meio do aplicativo SleepExplorer.
 - **Objetivo:** Gráfico de barra interativos e gráfico em círculo.
 - **Método:** Foi desenvolvido o SleepExplorer um ferramenta web que importa diversos dados de automonitoramento a partir das ferramentas (aplicativos) comerciais de monitoramento como o "Fibit" e do formulário do google, os dados coletados foram mostrados por meio de gráficos interativos de barras e no formato de círculo com uso de cores correlacionando o sono com fatores contextuais. O sleepExplorer foi implementado na nuvem da Microsoft Azure⁴ baseado na estrutura ASP.NET MODEL VIEW CONTROLLER (MVC). O sleepExplorer obtém os dados automaticamente dos usuários a partir da API "Fibit Public"⁵ via protocolo oauth.
15. **Título:** *Radar plots: a useful way for presenting multivariate health care data*
- **Dados:** Dados obtidos de um cenário fictício; Dados contínuos
 - **Técnica:** Visualização dos dados com o radar plot.
 - **Objetivo:** Comparar três modos de visualização, em tabela, gráfico de barras e por fim visualização em radar.
 - **Método:** Por meio de entrevista os pacientes de três torres hospitalares, torreA, torreB e torreC, foi questionado o nível de satisfação que foi dado por meio de uma escala de 1 (desclassificado) a 5 (totalmente satisfeito) em relação ao: conhecimento médico (MD knowledge), atitude médica (MD attitude), simplicidade do sistema (simplicity), acesso a emergência (Access), tempo de espera ao fazer os testes (Wait Time), valores (Cost) e

⁴Disponível em: <https://azure.microsoft.com/pt-br/>. Acessado em: 10/07/2018

⁵Disponível em: <https://dev.fitbit.com/build/reference/web-api/basics/>. Acessado em 15/08/2018

a disponibilidade do especialista (Availability). Uma vez com esses dados disponíveis, pode ser visualizado nos três tipos de visualizações.

16. **Título:** *Making sense of personal health information: Challenges for information visualization*

- **Dados:** Artigos obtidos dos buscadores, web of science, medline, dados qualitativos e quantitativos.
- **Técnica:** Revisão sistemática; Gráficos bidimensionais
- **Objetivo:** Ajudar as pessoas a melhor entender as suas próprias doenças e obter uma maior consciência sobre suas condições de saúde para que possam controlá-las eficientemente.
- **Método:** Foi usada a técnica da revisão sistemática, onde observou-se três categorias em destaque: visualização de dados, dados obtidos do estilo de vida, objetivos e tarefas do dia a dia. Foi trabalhado cada tema, usando programas de visualização como o Lifelines e o CareView a parti dos dados registrados no banco de dados.

2.7 DISCUSSÃO

A visualização de dados é uma técnica que vem sendo usada largamente em muitas áreas, como educação, economia, segurança, saúde, pois por meio dela é possível enxergar informações que antes estavam escondidas pela simples forma de apresentação dos dados por meio de tabelas ou técnicas ineficientes. A revisão sistemática aqui apresentada aborda os trabalhos que fazem uso da visualização em dispositivos móveis no domínio/área da saúde. Dentre os artigos que relatam dados neste domínio e que utilizam dados públicos brasileiros, encontra-se o trabalho de (SA et al., 2016). Tais dados foram obtidos de formulários em papéis e tinha como desvantagem a longa demora para a inserção desses dados no SIAB (Sistema de Informação de Saúde). Com o intuito de contornar este problema foi desenvolvido o GeoHealth Figura 2.5, um sistema que melhora o processo de análise e coleta desses dados contendo uma versão digital dos formulários. Este sistema possui uma integração com a localização do smartphone por meio de GPS, de modo a permitir uma localização georreferenciada dos dados coletados. A Figura 2.5 ilustra um exemplo, onde cada ponto representa os três tipos de doenças avaliadas, onde a cor vermelha mapeia pessoas com hipertensão e diabetes, a cor verde mapeia pessoas somente com hipertensão e a cor rosa mapeia pessoas somente com diabetes.

Outra técnica utilizada para mapear informações de saúde referente aos pacientes, foi encontrada nos estudos de Ledesma, Al-Musawi e Nieminen (2016) que aborda a visualização em radar. Nesta visualização os pontos representam dados dos pacientes (pressão arterial, consumo de álcool, glicose, exercício, entre outros) e são distribuídos ao redor de uma circunferência. O tamanho da circunferência é ajustada dinamicamente, ou seja, pode aumentar ou diminuir, a depender da quantidade de atributos. Neste estudo foi desenvolvido o hFigure, Figura (2.6) com

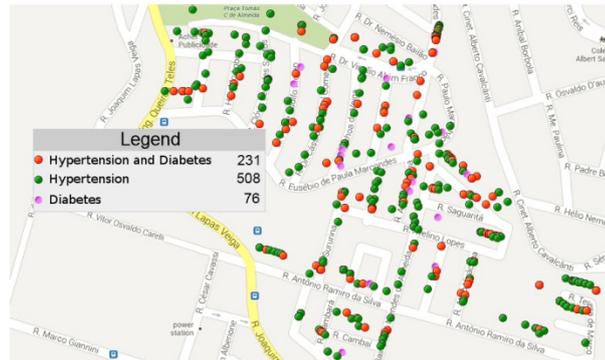


Figura 2.5 Exemplo dos dados plotados usando o GeoHealth (SA et al., 2016).

o objetivo de uma melhor experiência de visualização além de permitir uma comparação entre múltiplas medições relacionandas a saúde dos pacientes modelo ao longo do tempo.

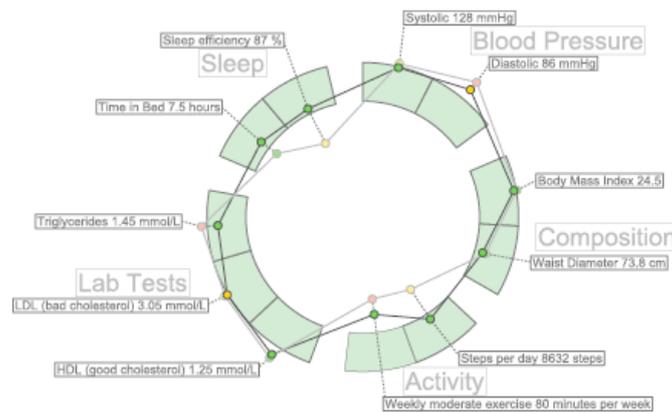


Figura 2.6 Exemplo do hFigure, medição das variáveis separadas por grupos (LEDESMA; AL-MUSAWI; NIEMINEN, 2016).

Butscher et al. (2016) desenvolveu uma visualização multidimensional para dispositivos móveis como smartphones ou tablets, mapeando os seguintes atributos: refeição (café da manhã, almoço, jantar, lanche), tipo de comida no almoço (grãos, vegetais, frutas, carne, leite, óleo e açúcar), quantidade de comida, quantidade ideal, e qual comida foi ingerida (bife, pasta, tomate, manteiga). A Figura 2.7 ilustra essa visualização, na qual a altura das barras de cada tipo de comida (café da manhã, almoço, jantar e refeição leve) corresponde a porcentagem de calorias consumidas durante cada tipo de refeição, a espessura da linha representa a importância da comida e a barra preta o consumo da pessoa (sendo a área verde a quantidade recomendável). Com isso o usuário poderá de maneira mais eficiente controlar a sua alimentação.

Paralelamente a esta idéia, Christmann et al. (2017) além de usar visualização

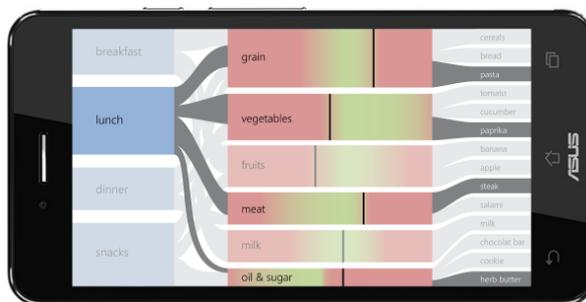


Figura 2.7 Visualização em um smartphone dos alimentos e sua importância de consumo (BUTSCHER et al., 2016).

multidimensional para extrair informações de hábitos alimentares, os autores utilizaram gráficos de barras para comparar o estilo de vida dos usuários da aplicação, como o tempo de sono recomendável, a ingestão diária recomendada de vegetais ou frutas, além de outros valores comparados, permitindo que o próprio usuário fizesse as devidas leituras dessas informações.

Alguns trabalhos abordaram o uso de mapa de calor no contexto da área da saúde (HINGLE et al., 2013; LATIF et al., 2015). O trabalho de Hingle et al. (2013) utiliza a técnica de visualização mapa de calor (*heat map*) para identificar as tags mais citadas no Twitter pelos participantes sendo acompanhados durante duas semanas. Esta técnica propicia uma visualização mais fácil de quais foram as tags do twitter mais citadas por região. Dentre essas tags regionais, as que possuíram um azul mais saturado são as mais citadas, sendo as menos citadas mapeadas pela cor branca, de modo a permitir uma fácil interpretação dos dados; além de apresentar um mapa de frequência onde as tags mais escritas ficaram concentradas no meio do mapa (Figura 2.8).

Em contrapartida, o trabalho de Latif et al. (2015) utiliza a mesma técnica de mapa de calor para monitorar o espalhamento da dengue (cor vermelha mais saturada mapeia maior espalhamento), além de proporcionar o monitoramento de quaisquer outros surtos repentinos de doenças em tempo real, Figura (2.9).

O estudo de Stevens et al. (2014) aborda o uso de visualização em dispositivos móveis para tratar pacientes com problemas cognitivos mentais e físicos. Neste estudo percebeu-se a grande importância da visualização dos dados para que as melhores decisões pudessem ser tomadas, de modo a auxiliar o médico responsável na tomada de decisão sobre os seus pacientes por meio de uma ferramenta, tablet (Figura 2.10). Os dados informais e qualitativos (saúde geral, moral, cognição, mobilidade) são obtidos pelos cuidadores, que conversam com os pacientes, os quais respondem algumas perguntas de um formulário digital. Então, os dados obtidos são dispostos em círculo permitindo uma visão global de suas condições de saúde.

Por meio de uma revisão sistemática os autores Faisal, Blandford e Potts (2013) apresentam dados que foram obtidos dos buscadores web of science e medline, os dados obtidos foram do tipo qualitativos e quantitativos onde o principal objetivo

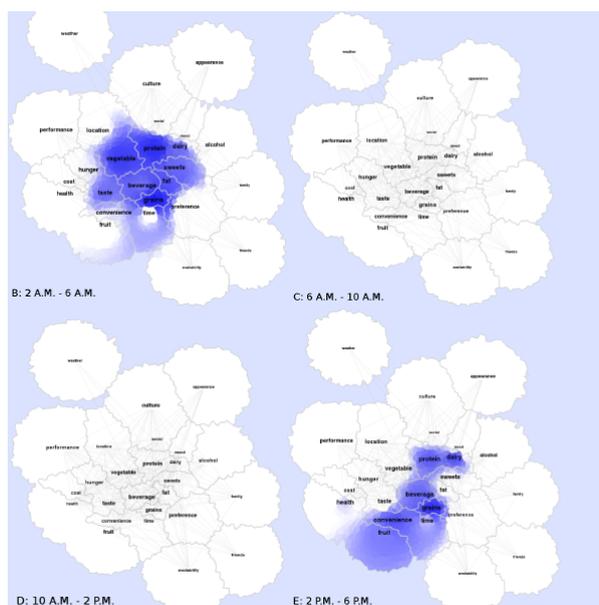


Figura 2.8 Mapa de calor onde a cor azul forte representa as tags mas citadas no Twitter (HINGLE et al., 2013).

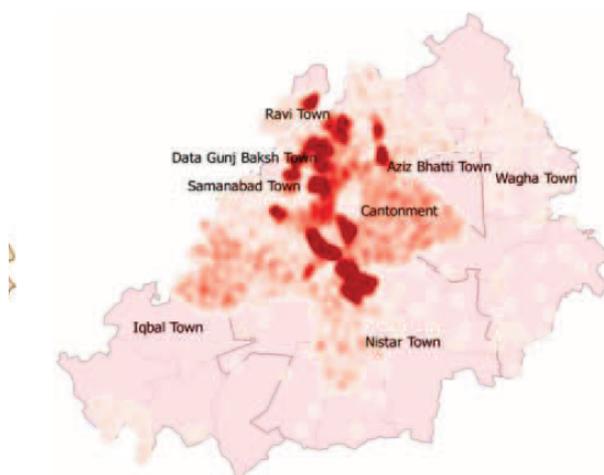


Figura 2.9 Mapa de calor para visualização do espalhamento da dengue (LATIF et al., 2015).

era provocar nas pessoas uma maior consciência sobre a sua própria condição de saúde e obter uma visão mais clara de suas doenças, proporcionando desta maneira uma maior prevenção e possivelmente evitando o avanço delas uma vez que os usuários da aplicação possuirá esta informação facilmente disponível facilitada pela visualização destes dados em um gráfico. Foi selecionado também um estudo relacionando os dados do sono obtido com os fatores contextuais que podia impactar no sono, no qual os autores Liang et al. (2016) obtêm os dados por meio de uma



Figura 2.10 Evolução do quadro do paciente sendo mostrada num gráfico em círculo (STEVENS et al., 2014).

API pública do FIBIT⁶ e de um formulário do Google sendo o tipo dele categórico, foi utilizado gráficos de círculos para sumarizar os dados, mesma técnica utilizada por (LEDESMA; AL-MUSAWI; NIEMINEN, 2016).

Técnicas de visualização geoespacial foram abordadas em vários dos estudos selecionados (DASGUPTA, 2010; PINGALI, 2016; KOK et al., 2016), o autor Dasgupta (2010) utiliza esta técnica apenas em aparelho celular mais simples, não utilizando smartphone para identificar os possíveis surtos de determinadas doenças, onde a região que possui um índice maior de ocorrência torna-se-á mais vermelha. Tais informações serão gerenciadas e enviadas apenas por profissionais qualificados para o ofício. Informações georreferenciadas também foram mapeadas utilizando marcações do tipo pino de Pingali (2016), Figura 2.11 para marcar em um mapa regiões onde possui determinados níveis de contaminação dos rios, quanto mais poluído o rio mais acentuado será a cor da região no mapa, tendo em vista a presença de sensores instalados nos rios realizando a medição do nível de contaminação.

Estes autores Kok et al. (2016) usam a mesma técnica de marcação auxiliada por pinos (marcadores) para marcar áreas no mapa onde existem determinados casos de doenças, evitando desta maneira possíveis surtos além de incorporar modelos de previsões matemáticos para prever o aparecimento de determinadas doenças. Por fim, o trabalho de Lee et al. (2013) utiliza diversos sensores para coleta de informações biométricas (*wearable*, e sensores internos ao corpo humano) para fazer as devidas medições. Tais sensores podem ser invasivos ou não, como os *wearable*, com

⁶Disponível em: <https://www.fitbit.com/home>. Acessado em: 15/08/2018



Figura 2.11 O pino mais vermelho representa o rio mais poluído (PINGALI, 2016).

o objetivo de fornecer ao usuário a visualização de informações como eletrocardiograma (ECG), pressão sanguínea entre outros para fazer os devidos monitoramento, visualizadas em aplicativos móveis híbrido.

2.8 CONSIDERAÇÕES FINAIS

Neste capítulo foi apresentado uma revisão sistemática da literatura para buscar métodos que auxiliem no desenvolvimento das técnicas de visualização de dados aplicada a saúde com o uso dos dispositivos móveis, o qual permitiu enxergar as vantagens e desvantagens dessas técnicas empregadas, além de proporcionar um maior entendimento do tema, possibilitando desta maneira o desenvolvimento do trabalho que será visto nos próximos capítulos.

Por meio das strings de busca empregadas na RS foi possível obter diversos artigos que exploravam tais técnicas de visualizações, sendo a visualização em radar, mapa de calor e a georreferenciada as mais empregadas entre os trabalhos vistos

A visualização de dados é imprescindível para obter uma melhor visão dos dados recebidos e conseqüentemente uma melhor tomada de decisão. A visualização por georreferenciamento é uma técnica que foi muito utilizada nos artigos lidos, principalmente por auxiliar na visualização dos dados em um mapa, permitindo os usuários obter intuitivamente e produtivamente as informações pretendidas; o mapa de calor, por sua vez, também apresentou informações relevantes para a tomada de decisão a exemplo de marcações, que variava a intensidade das regiões em um mapa de acordo com os dados obtidos de uma base de dados; por fim, a visualização em radar apresentou-se de grande utilidade principalmente por permitir que diversos tipos de dados obtidos das pessoas envolvidas nas pesquisas pudessem ser visualizados em formato circular, contribuindo para uma melhor compreensão geral do contexto/domínio dos dados e inferir conhecimento sobre seu estado de saúde.

FUNDAMENTAÇÃO TEÓRICA

3.1 CONSIDERAÇÕES INICIAIS

A visualização é uma ferramenta poderosa para análise de dados e vem sendo utilizada amplamente pela comunidade, seja no âmbito acadêmico ou comercial. Atualmente, existem inúmeras técnicas, sendo elas dependentes de vários fatores, como contexto, tipos de dados, quantidade de atributos e dimensões, entre outros. Neste trabalho, a visualização é aplicada no domínio da saúde, visando especificamente visualizações multidimensionais e focando em responsividade para dispositivos móveis.

Neste capítulo serão explicados os conceitos relacionados à visualização, apresentando e discutindo algumas técnicas de visualização. Em específico, serão apresentadas visualizações com características multidimensionais, tais como: gráfico em radar, gráfico estrela (star-plot), coordenadas paralelas, e a tree-map representando também características hierárquicas. A técnica de visualização radar ou starplot foi escolhida para estudo e implementação por apresentar estudos realizados e retornados na revisão sistemática da literatura; as coordenadas paralelas foi escolhida por ser eficiente em correlacionar os atributos da base de dados; e por fim a treemap foi escolhida por apresentar características hierárquicas, uma vez que a base de dados foi encontrado no formato *json*¹ e apresenta tais características.

Tais conceitos são de fundamental importância para compreender o escopo e resultados desta monografia.

3.2 VISUALIZAÇÃO DE DADOS MULTIDIMENSIONAIS

Telea (2007) define visualização de dados multidimensional como sendo N um conjunto de dados $D = \{p_i\}$, $1 \leq i \leq N$, onde cada ponto p_i tem k -dimensional vetor

¹Disponível em: <https://www.json.org/> Acessado em: 10/05/2018

de atributo $(\hat{a}_{il} . . . , \hat{a}_{ik}) \in A^k$, cada atributo é definido sobre um domínio A . Este conjunto de dados é chamado multivariado, pois contém diversos atributos (K neste exemplo). A visualização de dados em seu escopo global é definido por Ward, Grinstein e Keim (2015) como uma forma de comunicar informação por meio de representações gráficas, seu principal objetivo é potencializar a apropriação da informação pelo usuário por meios de tais representações. Ward, Grinstein e Keim (2015) expressa ainda que uma simples figura pode conter enormes quantidades de informações, e que podem ser processadas mais rapidamente do que uma página de palavras. Pois a interpretação da imagem é executada paralelamente dentro de um sistema perceptual humano, enquanto que a análise textual é limitada por um sequencial processo de leitura.

Sob a perspectiva da computação pode-se classificar visualização da informação, conhecida também como *infovis*, como um subcampo da visualização de dados em que Freitas et al. (2001) define a visualização da informação em uma área de aplicação de técnicas de computação gráfica interativa, que objetivam auxiliar a análise e a compreensão de um conjunto de dados. Algumas áreas que a *infovis* abrange segundo Telea (2007) são: raciocínio visual, modelagem de dados visuais, programação visual, visualização de execução de programas, interface visual, linguagem visual entre outras que são áreas mas específicas da *infovis*.

Muitos ainda estão familiarizados com o uso das tabelas para obter algum tipo de informação referente ao domínio em questão, porém existirão alguns dados cujo os significados não estariam necessariamente claros. Outro problema que é observado neste método de visualização é a escalabilidade, ou seja a medida que os atributos medidos crescem, maior será a dificuldade em descrever ou analisar seus significados como mostra a Figura 3.1 estudada por Telea (2007), onde são apresentados nove colunas da base de dados de uma bolsa de valores, além da grande dificuldade em se obter uma visão global dos dados em questão, gerando problemas na interpretação das informações. Segundo Telea (2007), a visualização em gráficos seria a melhor solução para os problemas citados, pois por meio deste método torna-se possível a visualização global dos dados (*overview*) em uma imagem, uma melhor relação dos atributos medidos, além de potencializar a interpretação das informações. Entretanto, para que esses benefícios trazidos pelo uso dos gráficos possam ser aproveitados, será preciso o uso adequado (dependendo do domínio) dos diversos tipos de gráficos existentes, será necessário avaliar o cenário em um escopo mais amplo para indicar qual a melhor visualização será abordada.

Na visualização de dados comumente são encontrados bases de dados com uma grande quantidades de atributos, dados multivariados, e portanto será preciso adotar técnicas eficientes de visualização de dados multivariados, a fim de que as correlações, *outliers*, grupos e tendências tornem-se visíveis (TELEA, 2007). Tecnicamente, dados multivariados, é definido pela maior dimensionalidade, ou seja a partir de 3 dimensões. Contudo segundo Chan (2006) visualizar dados em um espaço tridimensional(3D) não é uma tarefa fácil, por exemplo comparar dois pontos no mesmo eixo. Conseqüentemente o uso de técnicas de visualização multidimensional se faz

id	category	name	date	time	open	high	low	close
636	sif	SIF1	2004-11-29	13:00	0.800000	0.800000	0.800000	0.800000
635	sif	SIF1	2004-11-29	14:00	0.800000	0.800000	0.800000	0.800000
633	sif	SIF1	2004-11-29	16:00	0.795000	0.795000	0.795000	0.795000
630	sif	SIF1	2004-11-30	14:00	0.795000	0.795000	0.795000	0.795000
632	sif	SIF1	2004-11-30	12:00	0.800000	0.800000	0.795000	0.795000
631	sif	SIF1	2004-11-30	13:00	0.795000	0.795000	0.795000	0.795000
628	sif	SIF1	2004-11-30	16:00	0.795000	0.795000	0.795000	0.795000
629	sif	SIF1	2004-11-30	15:00	0.795000	0.795000	0.795000	0.795000
627	sif	SIF1	2005-00-02	12:00	0.785000	0.790000	0.785000	0.790000
626	sif	SIF1	2005-00-02	13:00	0.790000	0.795000	0.790000	0.795000
625	sif	SIF1	2005-00-02	14:00	0.795000	0.795000	0.795000	0.795000
624	sif	SIF1	2005-00-02	15:00	0.800000	0.800000	0.800000	0.800000
620	sif	SIF1	2005-00-03	15:00	0.795000	0.795000	0.795000	0.795000
623	sif	SIF1	2005-00-03	12:00	0.795000	0.795000	0.795000	0.795000
622	sif	SIF1	2005-00-03	13:00	0.795000	0.795000	0.795000	0.795000
621	sif	SIF1	2005-00-03	14:00	0.795000	0.795000	0.795000	0.795000
619	sif	SIF1	2005-00-03	16:00	0.795000	0.795000	0.795000	0.795000
618	sif	SIF1	2005-00-06	11:00	0.790000	0.790000	0.790000	0.790000
614	sif	SIF1	2005-00-06	15:00	0.795000	0.795000	0.795000	0.795000
617	sif	SIF1	2005-00-06	12:00	0.795000	0.795000	0.795000	0.795000
616	sif	SIF1	2005-00-06	13:00	0.795000	0.795000	0.795000	0.795000
615	sif	SIF1	2005-00-06	14:00	0.795000	0.795000	0.795000	0.795000
613	sif	SIF1	2005-00-06	16:00	0.795000	0.795000	0.795000	0.795000
609	sif	SIF1	2005-00-07	14:00	0.790000	0.795000	0.790000	0.795000
612	sif	SIF1	2005-00-07	11:00	0.795000	0.795000	0.795000	0.795000
611	sif	SIF1	2005-00-07	12:00	0.795000	0.795000	0.795000	0.795000
610	sif	SIF1	2005-00-07	13:00	0.790000	0.790000	0.790000	0.790000
608	sif	SIF1	2005-00-07	15:00	0.790000	0.790000	0.790000	0.790000
606	sif	SIF1	2005-00-08	13:00	0.795000	0.795000	0.795000	0.795000
607	sif	SIF1	2005-00-08	12:00	0.790000	0.790000	0.790000	0.790000
605	sif	SIF1	2005-00-08	14:00	0.795000	0.795000	0.795000	0.795000

Figura 3.1 Visualização em tabela da bolsa de valores.

necessária, aumentando desta maneira a fidelidade na interpretação dos dados. O autor Chan (2006) ainda cita que estritamente falando, multidimensional refere-se a dimensionalidade das dimensões independentes, enquanto que o termo multivariado refere-se aos atributos dependentes. Entretanto a propriedade multidimensional está implícita no uso comum.

Nas subseções seguintes serão apresentadas algumas técnicas de visualização de informação, notadamente multi-dimensionais.

3.2.1 Star-Plot

No trabalho apresentado por Saary (2008) o Star-plot, conhecido também como radar charts ou web charts, é um método de visualização utilizado para mostrar dados multivariados, ou seja, múltiplas características a serem observadas. Neste tipo de visualização são frequentemente usadas para mostrar diferentes observações do mesmo tipo de dados. Segundo Saary (2008) a percepção das informações em dados com muitos atributos tornam-se mais claras, se comparado aos gráficos convencionais como o gráfico de barras.

Uma outra alternativa de visualizar as informações no Star plot, pois permitiria uma melhor comparação entre os dados, seria em dispor-los no mesmo gráfico usando para isto diferentes cores e estilos de linhas, facilitando desta maneira a tomada de decisão. Entretanto, essa visualização seria de difícil percepção se fossem dispostos grandes quantidades de observações, que são os polígonos instanciados. Foi realizado um comparativo no trabalho de Saary (2008) com dados simulados de três torres/prédios hospitalares, com três modos de visualização, a visualização tabular (Tabela 3.1), o Star-plot (Figura 3.2a) e o gráfico de barras (Figura 3.2b). Analisando esta comparação nos estudos de Saary (2008) comprovou-se uma maior facilidade na interpretação dos atributos no star plot, uma vez que eles estão dis-

postos de uma maneira mais intuitiva. A exemplo da percepção do atributo MD-Atitude (simplicidade do sistema, menos burocracia) que apresentou um maior valor para as três torres/prédios, o que difere na visualização por meio da tabela, exigindo um esforço maior para notar tal informação. Ainda neste estudo, no gráfico de barras observou-se que ele não é eficiente para mostrar grandes quantidades de atributos, dificultando deste modo a percepção do leitor, uma vez que a imagem cresce na tela e não apresenta uma boa escalabilidade, pois possui uma maior dificuldade em comportar mais atributos. Isto não ocorre com o Star-plot, que pode comportar o acréscimo dos atributos medidos sem necessariamente alterar o tamanho da imagem na tela.

Variáveis	MD Knowledge	MD Attitude	Simplicity	Access	Wait Time	Cost
Town A	5	4	4	5	3	3
Town B	5	4	5	5	5	4
Town C	5	3	2	3	2	1

Tabela 3.1 Tabela dos atributos.

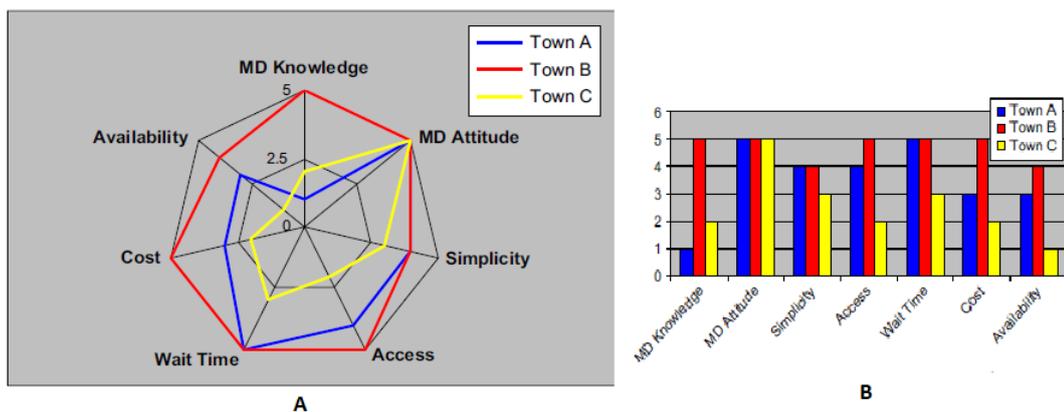


Figura 3.2 Imagem retirada do artigo Saary (2008) realizando uma comparação do gráfico star-plot, gráfico star-plot figura A e o gráfico de barras tradicional figura B.

Em consonância, Munzner (2015) compara diferentes tipos de modelos de visualizações radiais.

Um exemplo de um gráfico de radar pode ser observado na Figura 3.3 ², no qual é formado por atributos onde cada uma deles possuem um eixo, que começa a partir do centro. Todos eles são formados radialmente, com igual distância entre cada atributo, enquanto mantém a mesma escala entre todos os eixos. As linhas verticais que conecta eixo a eixo são frequentemente usadas como guia. Cada valor do atributo é plotado ao longo de seu eixo individual e todos as linhas, representadas pelas instâncias, são conectadas formando um polígono.

²Disponível em: <https://datavizcatalogue.com/methods/radar-chart.html>. Acessado em: 15/10/2018

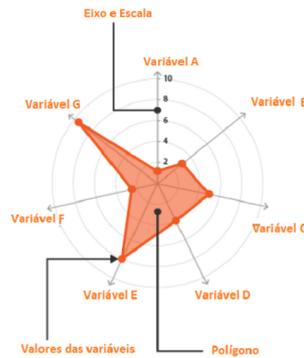


Figura 3.3 Anatomia do Gráfico em radar.
3

3.2.2 Coordenadas Paralelas

Um modelo de visualização que é potencialmente indicado para uso de maior quantidade de dados multivariados são as coordenadas paralelas. Como o nome sugere, todos os eixos são dispostos paralelamente numa posição vertical na tela, onde cada atributo possui seu próprio eixo. Cada eixo pode possuir sua própria escala ou podem trabalhar com unidades diferentes de medição, podendo ser normalizadas para manter uma escala uniforme. As instâncias são plotados por uma série de linhas retas que são conectadas entre todos os eixos exatamente uma vez, formando uma polilinha. A Figura 3.4 ilustra um exemplo deste modelo associado com uma visualização de tabela para o mesmo conjunto de dados.

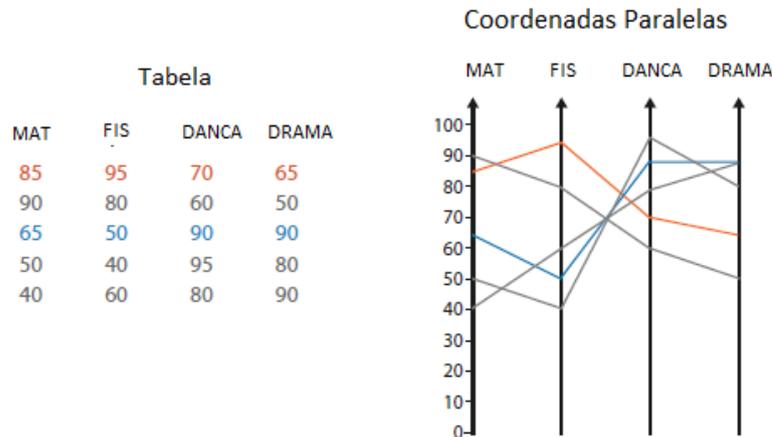


Figura 3.4 Comparação entre uma pequena tabela e as coordenadas paralelas ((MUNZNER, 2015)).

Uma real motivação para uso das coordenadas paralelas está na sua eficiência de correlação entre os atributos, uma vez que eles estão dispostos paralelamente. Segundo Munzner (2015) sua correlação é dada por três padrões visuais. A primeira

é a correlação positiva onde as linhas que cruzam os dois eixos são em grande parte paralelas, o segundo padrão apresenta uma correlação negativa pois as linhas na maior parte cruzam umas com as outras entre os eixos, e por último, o não correlacionamento entre os eixos uma vez que existem uma mistura de ângulos de cruzamento, como pode ser visto na Figura 3.5. Este modelo de visualização são frequentemente usados também para se obter uma visão global de todas os atributos em questão, encontrar intervalos entre eles, selecionar um intervalo específico e detecção de *outlier*. No exemplo da Figura 3.5 o terceiro eixo apresenta um intervalo maior que os outros, assim como no item superior do quarto eixo que possivelmente possa ser um *outlier*.

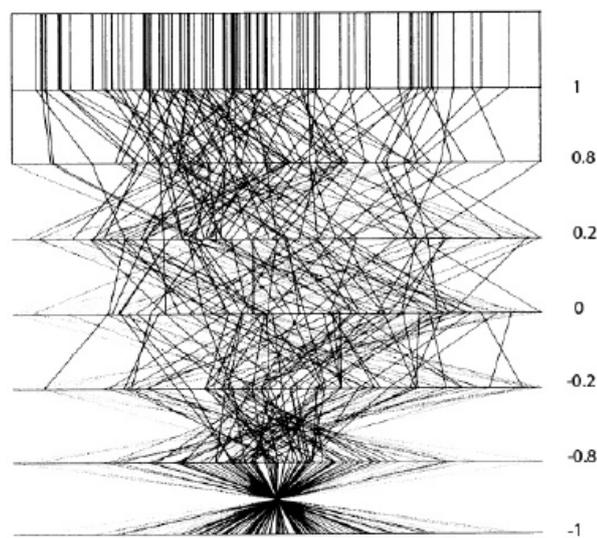


Figura 3.5 Correlação entre os eixos vizinhos das coordenadas paralelas ((MUNZNER, 2015)).

Ainda nos estudos de Munzner (2015) as coordenadas paralelas trabalham com os dados multivariados usando duas dimensões, onde cada atributo requer apenas uma dimensão espacial, e a segunda dimensão é usada para dispor os múltiplos eixos presentes, observando o limite do espaço de projeção dos eixos. Dezenas de atributos são comuns neste tipo de visualização, porém a medida que a quantidade de eixos cresce, maior será a necessidade de espaço horizontal. Contudo, é preciso estar atento aos limites dos eixos na tela, onde será preciso haver um espaço suficiente entre eles, para permitir discernir os padrões de interseção ou paralelismo dos segmentos de linhas que passam entre os eixos. Outra questão a ser considerada nas coordenadas paralelas são as quantidades de linhas que podem ser usadas sem provocar uma má interpretação dos dados, pois neste modelo o recomendável é o uso de centenas de itens a serem visualizados e não milhares de itens Munzner (2015) como mostra a Figura 3.6.

Um recurso disponível nesta técnica de visualização, é a opção de determinar a ordem dos eixos no gráfico, permitindo ao leitor movimentá-los para esquerda ou direita para um melhor correlacionamento e interpretação dos atributos, além

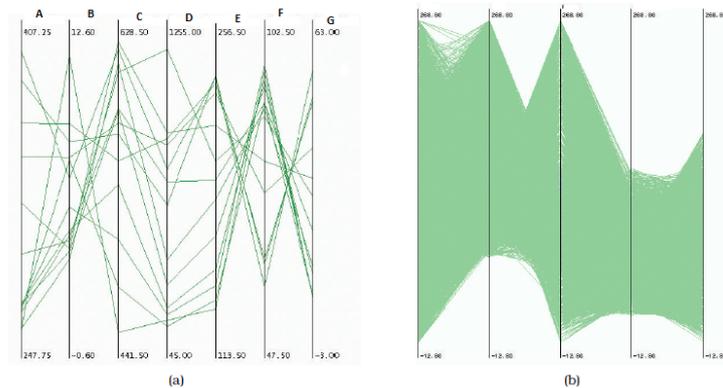


Figura 3.6 Visualização de centenas de linhas em (a), e milhares de linhas em (b) ((MUNZNER, 2015)).

de dispor de outro recurso chamado brush, cujo o objetivo é a seleção dos eixos e linhas desejadas para ser visualizada, potencializando o estudo das relações entre os atributos e filtrando as análises.

3.2.3 Tree-Map

Continuando o estudo sobre a visualização de dados multivariados tem-se a Treemap, que foi originalmente desenvolvido por Shneiderman como uma maneira de visualizar grandes diretórios de arquivos em um computador, sem ocupar muito espaço na tela. Outra característica é o fato desta técnica ter a capacidade de representar estrutura hierárquica, o que permite uma visão global das características dos dados em questão (MUNZNER, 2015). A Treemap, representa um determinado conjunto de dados usando retângulos aninhados, é uma maneira alternativa de visualizar a estrutura hierárquica de um diagrama de árvore (*Tree Diagram*) enquanto mostra as quantidades de cada categoria por meio do tamanho da área do retângulo. A Figura 3.7 desenvolvida por Saary (2008) apresenta a visualização estática, sem recursos de interação, o que é um ponto negativo quando se trata de visualização de dados, pois o mesmo permitiria uma melhor análise dos dados visualizados.

Cada categoria da Treemap é atribuída a uma área retangular com suas subcategorias retangulares aninhadas internamente. Quando uma quantidade é atribuída a uma categoria, a área do retângulo é formada proporcionalmente conforme esta quantidade. O tamanho da área do nó (retângulo) pai é o resultado da soma de todas as subcategorias (nó filho), contudo se nenhuma quantidade é atribuída às subcategorias, então esta área será dividida igualmente entre as outras categorias dentro do nó pai.

A maneira que o retângulo é dividido e ordenado dentro dos sub-retângulo (nó filho) é dependente dos algoritmos de ladrilho (*tiling*) usado, muitos algoritmos de

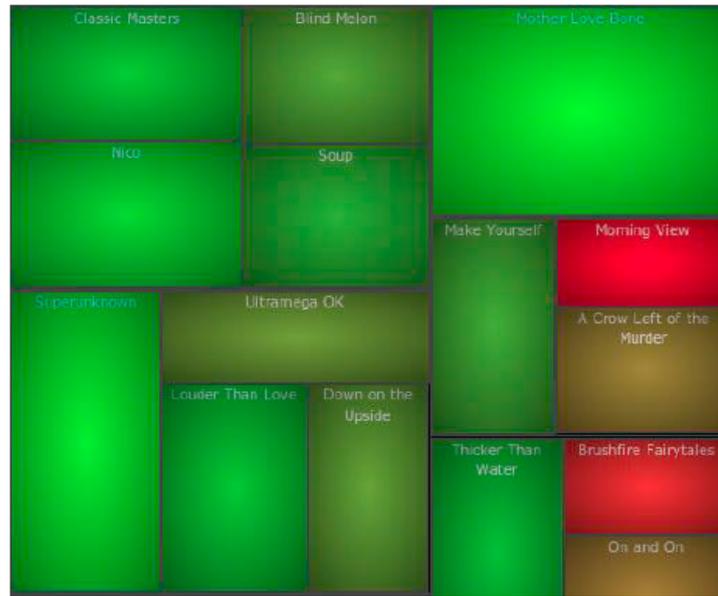


Figura 3.7 Regiões da treemap mapeadas por retângulos Saary (2008)

ladrilho são usado na construção da treemap, mas o mais comum é o “*squarified algorithms*” que mantém cada retângulo o mais quadrado possível.

3.3 CONSIDERAÇÕES FINAIS

As técnicas apresentadas neste capítulo fazem parte do escopo de visualizações multidimensionais. As técnicas de visualização radiais e coordenadas paralelas são baseadas em eixos, enquanto a treemap de característica hierárquica possui uma representação retangular.

As visualizações radiais permitem ter uma visão global por meio dos polígonos instanciados, eles também possuem limitações por exemplo, se muitos atributos forem visualizados. O gráfico que se mostrou de grande utilidade na interpretação dos atributos por correlaciona-los de uma melhor maneira, foi as coordenadas paralelas, por meio do recurso de interação fornecido por esta técnica, se implementada, de modo a permitir ao leitor selecionar por meio da brush qual eixo ou linha analisar, potencializando as análise. Contudo esta técnica apresenta problema de escalabilidade, pois não suportam muitas dimensões; por fim outro modelo de visualização conceituado foi a treemap, que apresentou uma característica peculiar de representar estruturas hierárquicas, disponibilizando os dados em forma retangular, seguindo uma lógica hierárquica (pai e filho) onde o mesmo varia dependendo da quantidade atribuída a cada categoria.

O próximo capítulo utiliza as técnicas supracitadas no domínio de dados da saúde considerando aspectos de responsividade.

APLICANDO VISUALIZAÇÕES MULTIDIMENSIONAIS RESPONSIVAS EM DADOS DE SAÚDE

4.1 CONSIDERAÇÕES INICIAIS

Este capítulo apresentará os resultados obtidos por meio da investigação de técnicas de visualização multidimensional no domínio da saúde, considerando a responsividade em dispositivos móveis. A avaliação realizada sobre a responsividade dos gráficos nos dispositivos móveis foi embasada no conceito sobre responsividade Frain (2012) e validada visualmente.

4.2 BASE DE DADOS

A base de dados utilizada para o desenvolvimento deste trabalho foi encontrada por meio dos estudos de Ledesma, Al-Musawi e Nieminen (2016). Os dados ¹ encontram-se no formato *json* e descreve um grupo de cinco pacientes, dentre os quais foram considerados para esta monografia dois pacientes, em específico Diego.C e Veronica.H. Esses dados foram escolhidos por apresentarem características quantitativas contínuas e possuem relação hierárquica.

Os atributos encontrados na base de dados foram: *Access*, *Environment*, *Behavior*, *Social Circumstances*, *Genetics/Biology*, *Medical Care* e *Health Literacy*. O atributo *access* refere-se ao acesso ao serviço de saúde, importante para promover e manter a saúde; *environment* preocupa-se com todos os aspectos relacionados ao ambiente natural que afeta a saúde humana; *social circumstances* refere-se às

¹Disponível em: <https://github.com/goinvo/VisualizationComponents/blob/master/src/data/data.json>. Acessado em: 10/12/2018

condições do local onde as pessoas vivem, aprendem, trabalham ou brincam que afetam a saúde; *medical care* são os planos de saúde; *genetics/biology* refere-se a doenças advindas de um histórico familiar; *health literacy* é a capacidade do paciente em entender, ler e usar informações de saúde para tomar decisões adequadas.

Os valores dos atributos variaram em um intervalo entre 0 e 1. Esses valores referem-se aos atributos medidos, representando a sua classificação (positivo ou negativo) de cada um desses atributos, sendo que os valores próximo a 0 ou a 1 foram avaliados negativamente. Como exemplo, considere os atributos da base de dados, *Access* e *Social circumstances* no ano 2018 para o usuário Diego.C. Para serem avaliadas positivamente neste estudo, tais valores tinham que estar no intervalo de 0,29 e 0,75. Porém esses atributos apresentaram valores 0.15 e 0.99 respectivamente, sendo desta maneira avaliados negativamente por Ledesma, Al-Musawi e Nieminen (2016), justamente por estarem fora da faixa recomendada.

4.3 ESTUDOS DE CASOS COM VISUALIZAÇÕES MULTIDIMENSIONAIS

Os gráficos implementados nesta monografia deu-se por meio da linguagem de marcação HTML, da folha de estilo em cascatas CSS, da linguagem de programação javascript e sua biblioteca D3² que possui inúmeros recursos para a manipulação e visualização de dados. Juntamente com o D3 foi utilizado o padrão de imagem SVG³, que compreendem gráficos vetoriais escaláveis, uma linguagem XML para descrever de forma vetorial desenhos e gráficos bidimensionais estáticos, dinâmicos ou animados. Em conjunto, eles garantem a responsividade das visualizações.

Todas as implementações foram desenvolvidas com o uso da linguagem javascript bem como o desenvolvimento da responsividade dos gráficos que é definido por Frain (2012) como a representação de conteúdo de maneira acessível para qualquer tamanho tela, permitindo adaptações a diversos tipos de telas como os smartphone e tablets. Os códigos das visualizações estão abertos e disponíveis no Github Pages⁴, um serviço de hospedagem oferecido pelo Github⁵ para hospedar páginas Web.

Durante a fase de pesquisa alguns gráficos encontrados não apresentaram responsabilidades, oferecendo portanto uma maior dificuldade na interação ao ser visualizado em dispositivos móveis. O uso da biblioteca D3 resolveu esta questão em dois modelos de gráficos implementados, a tree-map e o gráfico radar, porém o star-plot e as coordenadas paralelas continuavam sem este importante recurso. Para solucionar este problema, foi necessário desenvolver a responsividade desta técnica, utilizando para isto um recurso do javascript que é o atributo **window.onresize**. Este evento é acionado quando ocorre o redimensionamento do documento *view* na

²Disponível em: <https://d3js.org>. Acessado em: 01/11/2018

³Disponível em: <https://developer.mozilla.org/pt-BR/docs/Web/SVG>. Acessado em: 10/10/2018

⁴Os gráficos, star-plot, gráfico radar, coordenadas paralelas e a treemap, podem ser encontradas nesses links, respectivamente: [fafa1.github.io/d3-res](https://github.com/fafa1/d3-res), [fafa1.github.io/coo-paral/](https://github.com/fafa1/coo-paral/) e [fafa1.github.io/tree-map](https://github.com/fafa1/tree-map).

⁵Disponível em <https://github.com/>. Acessado em 15/12/2018.

janela de visualização, permitindo que qualquer objeto na tela se adapte ao tamanho atual. A Figura 4.1 apresenta o método **window.onresize** onde no seu escopo existe a chamada de duas funções implementadas, *tam_Padrao* e *plotarGrafico*, que são responsáveis pelo desenho do gráfico na tela.

```
let timer;
window.onresize = function(event) {
  clearTimeout(resizeTimer);
  timer = setTimeout(function()
  {
    tam_Padrao();
    plotarGrafico();
  }, 100);
}
```

Figura 4.1 Método do javascript onresize sendo utilizado no redimensionamento de tela no tempo de 100ms.

A seguir, serão apresentadas e discutidas as técnicas de visualizações multidimensionais em diferentes configurações de dispositivos móveis. Os testes foram efetuados utilizando um recurso de responsividade do navegador Chrome que simula as configurações e resoluções de diferentes dispositivos móveis (Tabela 4.1). Dentre os quais foram escolhidos:

Fabricante	Modelo	Tamanho da Tela	Resolução da Tela(px)
Apple	Iphone 5S	320 X 560	1136 X 640
LG	LG Optimus L70	384 X 640	800 X 480
Samsung	Galasx S5	360 X 640	1920 X 1080
Apple	Ipad pro	1024 X 1366	2388 x 1668

Tabela 4.1 Modelos dos dispositivos móveis usados nos testes de responsividade.

4.3.1 STARPLOT

A primeira visualização multidimensional investigada foi o Starplot, onde as dimensões/atributos são representados por eixos. Esta técnica de visualização foi implementada com o propósito de apresentar os valores de cada atributo de cada mês separadamente, de modo a permitir uma análise mais apurada dos dados em questão. Foi utilizado para o redimensionamento do gráfico star-plot um tempo de cem milissegundos (100ms) permitindo que as interações ocorram em tempo no uso da visualização. A seguir, serão apresentados os resultados desta técnica nos quatro dispositivos móveis supracitados.

A Figura 4.2 apresenta os resultados obtidos desta técnica no Galaxy S5 utilizando o método responsivo. Nota-se que o StarPlot da Figura 4.2A, se adapta

totalmente ao tamanho da tela, ao passo que quando não aplicado o método, fica evidente que alguns eixos ficaram fora do espaço visual, Figura 4.2B. Os mesmos resultados foram obtidos no modelo LG Optimus L70, na Figura 4.3C, nota-se todos os eixos apresentam-se dentro do espaço visual, diferente do resultado obtido na Figura 4.3D com mais eixos fora do espaço visual. No Iphone 5S Figura 4.3A/B por apresentar um menor tamanho de tela comparado aos outros modelos, obteve um pior resultado quando não utilizado o método *resize* Figura 4.3B, verificando desta maneira a eficiência da contribuição realizada empregando este método Figura 4.3A. No Ipad Pro Figura 4.4 fica evidente o uso deste método verificando a adaptação do gráfico na tela, de modo que o gráfico desenhado apresenta dimensões maiores 4.4A, diferente do que pode ser visto na Figura 4.4B, sem o uso do método, onde o tamanho do gráfico não se altera, de modo a permanecer indiferente ao tamanho de tela o qual é submetido.

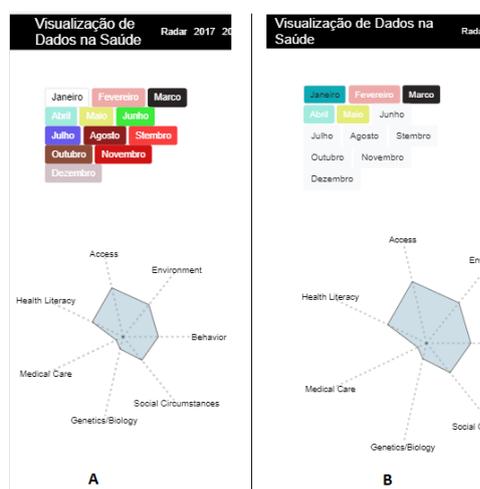


Figura 4.2 Figura A - visualização responsiva utilizando o método `windows.onresize` em um Galaxy S5. Figura B sem o método `windows.onresize`, apresentando alguns eixos fora do espaço visual.

A Figura 4.5 faz uma comparação entre dois meses, janeiro e junho, do ano 2018 com as sete variáveis disponíveis nos eixos. A Figura seguinte 4.6 mostra a interação com o uso do mouse ou do dedo, se for em dispositivos móveis, nos eixos onde ao passar o mouse ou tocar com o dedo em cima de cada eixo, é mostrado o nome da variável e seu respectivo valor.

4.3.2 COORDENADAS PARALELAS

Foi implementado no gráfico das coordenadas paralelas com o objetivo de contornar o problema da responsividade o mesmo método implementado no Star-plot, *resize*. Entretanto, com o uso deste método um recurso de interatividade deste modelo de visualização ficou comprometido, que é a *brush*. Como explicado no capítulo de

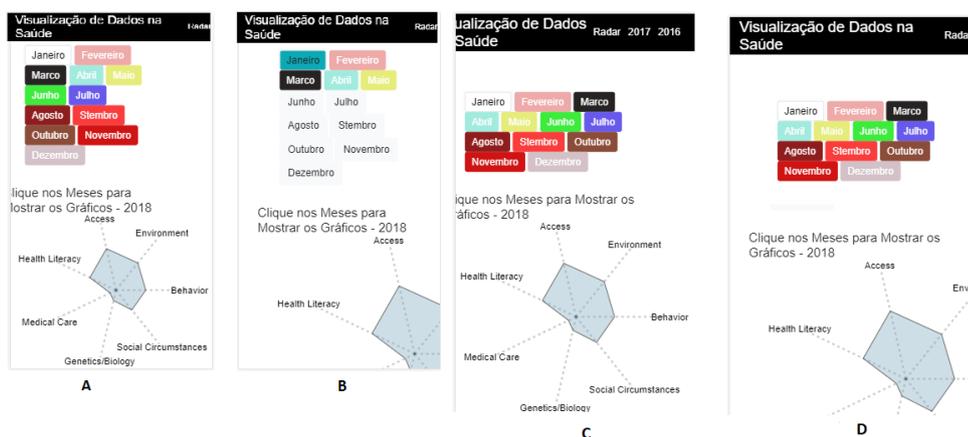


Figura 4.3 Figura A - visualização responsiva utilizando o método *resize* em um Iphone 5s. Figura B sem o método *resize* no Iphone 5s, apresentando muitos eixos fora do espaço visual; Figura C visualização responsiva utilizando o método *resize* em um LG Optimus L70; Figura D sem o método *resize* no LG Optimus L70 com eixos fora do espaço visual.



Figura 4.4 Figura A - visualização responsiva utilizando o método *windows.onresize* adaptando-se ao tamanho de tela no Ipad Pro. Figura B sem o método *windows.onresize* de modo a não variar o tamanho do gráfico em uma tela de maior dimensão no Ipad Pro.

fundamentação teórica, este recurso permite a seleção dos eixos e linhas das coordenadas paralelas, de modo a possibilitar um filtro dos dados resultando em linhas selecionadas com a cor azul e as linhas de fundo com a cor cinza; por este motivo optou-se pelo não uso do método de redimensionamento do javascript, preservando desta maneira a brush das coordenadas paralelas.

Porém, sem o uso do método *resize* e tendo-se em vista que este modelo de visualização foi implementado com o uso da biblioteca D3, uma boa experiência



Figura 4.5 Comparação entre dois meses do ano 2018 com polígonos sobrepostos representando meses diferentes, junho em verde e janeiro em azul.

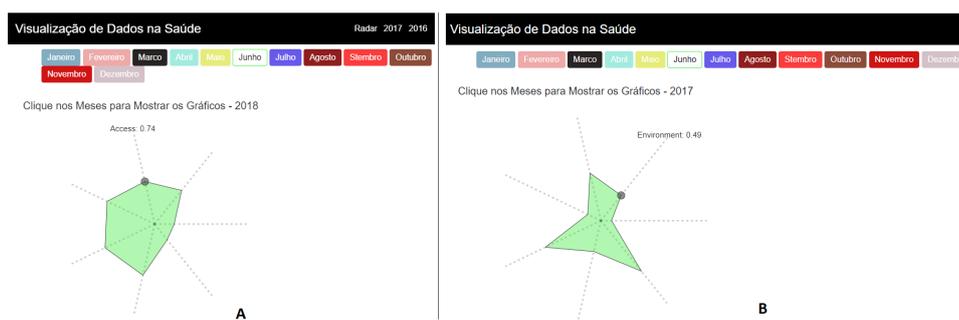


Figura 4.6 Interação ao passar o mouse, revelando o atributo e o seu valor no gráfico star-plot no ano de 2018 figura A e 2017 figura B.

do usuário é garantida quando utilizado o tablet ou ipad, com resolução a partir de 1024 X 1136, Figura 4.7 e todos os recursos de interação estão disponíveis, a exemplo da brush e o movimento horizontal de cada eixo. Nos dispositivos móveis supracitado, Samsung Galaxy S5, Iphone 5S e LG Optimus L70, não foi verificado a responsividade para esta técnica. As figuras a seguir mostram os resultados adquiridos com as implementações realizadas, onde os anos 2018, 2017 e 2016 (Figura 4.8) são mapeados com cores diferentes (Azul 2018, Laranja 2017 e Verde 2016). A Figura 4.9A e Figura 4.9B exemplifica o uso da brush para selecionar um conjunto de linhas de anos diferentes bem como as possíveis interações no gráfico; a Figura 4.9C apresenta a interação realizada sobre o eixo com o atributo *SocialCircumstances* arrastando para a esquerda, de modo a trocar de posição com o eixo com atributo *Behavior*.

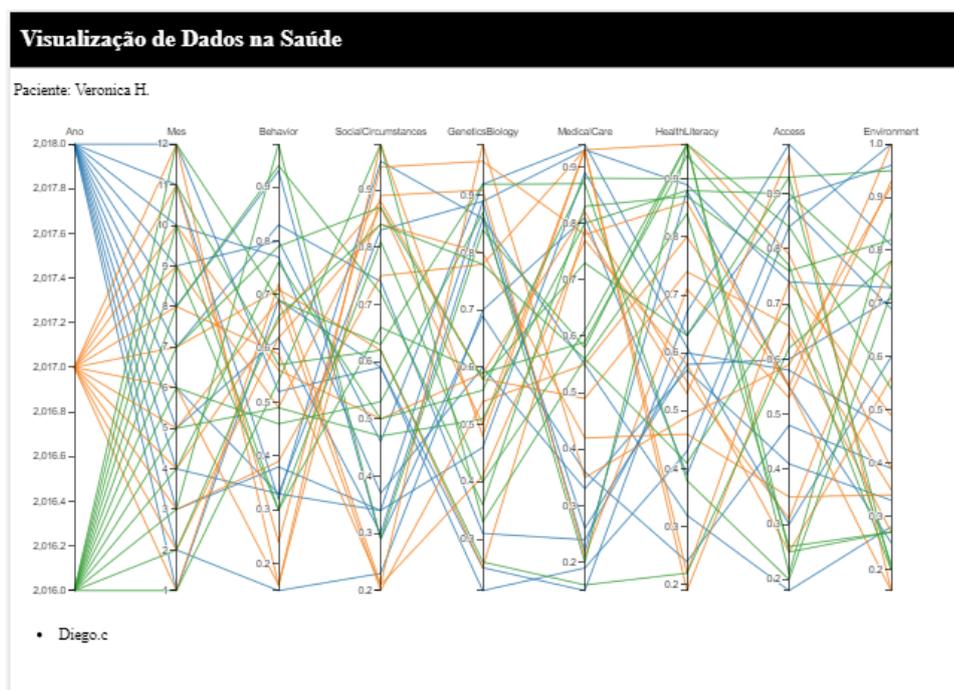


Figura 4.7 Coordenadas paralelas visualizada em um Ipad Pro do paciente Veronica H.

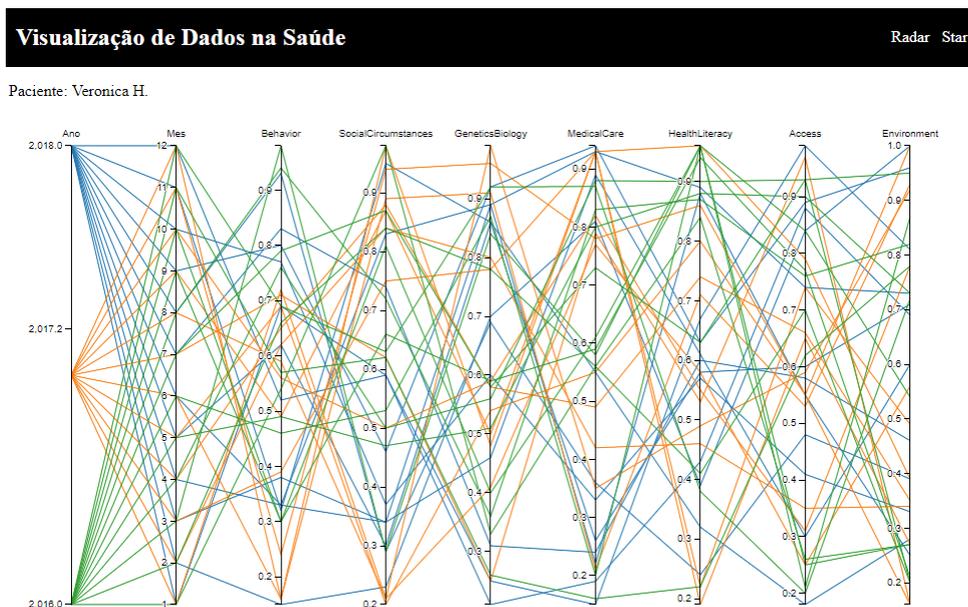


Figura 4.8 Coordenadas paralelas, em que cada ano apresenta cores diferentes.

4.3.3 TREEMAP

Foi desenvolvido, também com o uso do D3 e uma biblioteca auxiliar Highcharts ⁶, o modelo de visualização da Tree-map, o qual por meio desta biblioteca a responsi-

⁶Disponível em: <https://www.highcharts.com/docs>. Acessado em: 16/07/2018

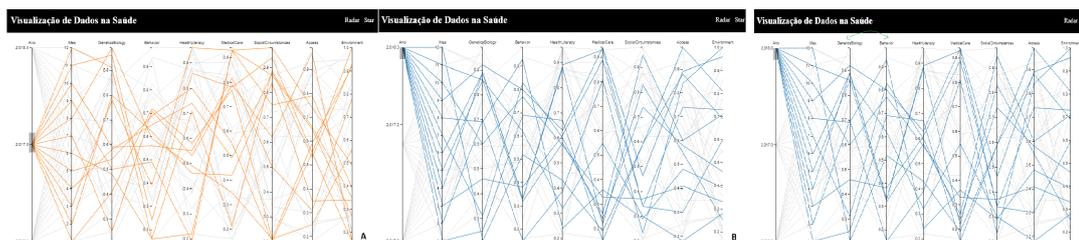


Figura 4.9 Interação com a brush, selecionando as linhas do ano 2017, em cor laranja (Figura A), e as linhas do ano 2018 em cor azul (Figura B).

vidade foi garantida, sendo assim adaptada aos diversos tamanhos de telas seja na posição vertical ou horizontal dos dispositivos móveis.

Um recurso importante desta técnica é a interatividade, porém este recurso foi encontrado durante a fase de pesquisa de maneira estática, de modo a permitir ao usuário apenas visualizar os nomes dos atributos e seus respectivos valores ao passar o mouse por cima das áreas retangulares. Neste trabalho foi desenvolvido uma nova interação, que permite ao usuário interagir com cada nível da treemap, começando do primeiro nível (nó raiz) até o último nível (nó folha).

As figuras a seguir elucidam bem esta nova abordagem desenvolvida inteiramente em javascript. A Figura 4.10A apresenta o primeiro nível da treemap contendo os dois pacientes da base de dados, ao clicar ou pressionar com o dedo em um deles é levado para o próximo nível, como demonstra a Figura 4.10B, onde apresenta os anos (2016, 2017 e 2018). Ao escolher um dos anos para análise é levado ao próximo nível da Treemap, que são os doze meses dividido igualmente no espaço Figura 4.10C por fim ao escolher um dos meses presentes é direcionado ao último nível da Treemap, onde encontra-se todos os atributos da base de dados, cujo seus respectivos valores são apresentados ao passar o mouse por cima em cada um deles (Figura 4.10D). É possível notar nesta imagem que cada dimensão é mapeada com cores diferentes, facilitando deste modo a identificação de cada um dos atributos. Para voltar aos níveis anteriores, basta clicar ou pressionar, no botão localizado na parte superior à direita da Treemap.

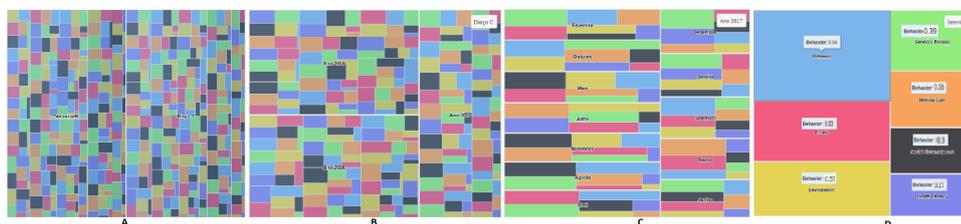


Figura 4.10 Visão inicial da tree-map (nó raiz) em A, possibilitando a escolha de um dos pacientes para ser visualizada nos nós seguintes; próximo nível da treemap apresenta os anos em B; no nível seguinte apresenta os meses em C; e por fim o último nível os atributos em D.

As Figuras a seguir apresentam os resultados obtidos dos dispositivos móveis in-

investigados. A Figura 4.11, Figura 4.12, Figura 4.13 e Figura 4.14, Samsung Galaxy s5, Iphone 5S, LG Optimus L70 e Ipad Pro respectivamente, obtiveram os mesmos resultados de modo a apresentar a treemap responsivamente (todos os atributos dentro do espaço visual) contendo os quatro níveis de interações desenvolvidos e todos os atributos com suas respectivas cores disponíveis para o usuário.



Figura 4.11 Visão das telas em um dispositivo móvel - Galaxy S5, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.



Figura 4.12 Visão das telas em um dispositivo móvel - Iphone 5S, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.



Figura 4.13 Visão das telas em um dispositivo móvel - LG Optimus L70, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.

4.3.4 RADAR PLOT

Outra técnica de visualização desenvolvida por Ledesma, Al-Musawi e Nieminen (2016) e que se assemelha ao star-plot é o gráfico radar, sendo a diferença entre

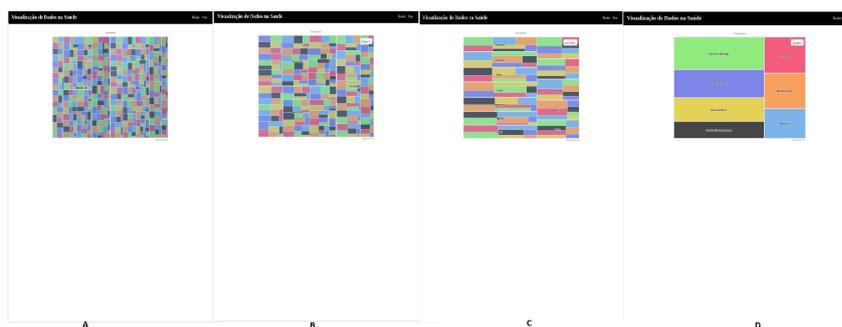


Figura 4.14 Visão das telas em um dispositivo móvel - IPAD PRO, pacientes disponíveis em A, anos disponíveis em B, meses disponíveis em C e atributos disponíveis em D.

eles as linhas radiais existentes. No presente trabalho o radar plot foi utilizado com o propósito de visualizar as médias de cada atributo dos meses dos anos (2016, 2017 e 2018). Este modelo foi também implementado com a biblioteca D3 de modo a permitir a responsividade nos dispositivos supracitados, estando ele na posição horizontal ou vertical.



Figura 4.15 Visão Global da aplicação possibilitando ao usuário escolher quais anos, semestres e pacientes pretende analisar.

Na Figura 4.15A apresenta a visão global do gráfico de modo a permitir escolher um ou os dois pacientes, localizado a direita, bem como as porcentagens de cada atributo ao clicar ou pressionar com o dedo no gráfico Figura 4.15B. É possível sobrepor os dois polígonos, Figura 4.16A, representado pelos dados dos dois pacientes, e fazer uma comparação entre eles. Na mesma Figura 4.16A é apresentado o gráfico sendo visualizado em um computador pessoal e em um dispositivo móvel da marca Samsung modelo Galaxy S5 Figura 4.16B.

A Figura 4.17 apresenta o gráfico em outros três dispositivos móveis, o LG Optimus L70, Iphone 5S e no Ipad Pro, onde é possível notar a responsividade nos três modelos testados. A contribuição implementada para este gráfico, foi oferecer ao usuário um recurso de filtrar quais semestres dos anos (2018, 2017, 2016), como pode ser visto na Figura 4.17C ele pretende analisar, potencializando desta maneira suas percepções sobre os dados de saúde analisados.

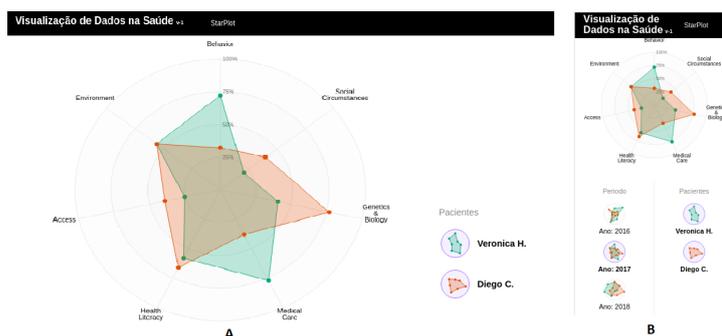


Figura 4.16 Radar visualizado em um computador pessoal A com polígonos sobrepostos, e no dispositivo móvel Galaxy S5.

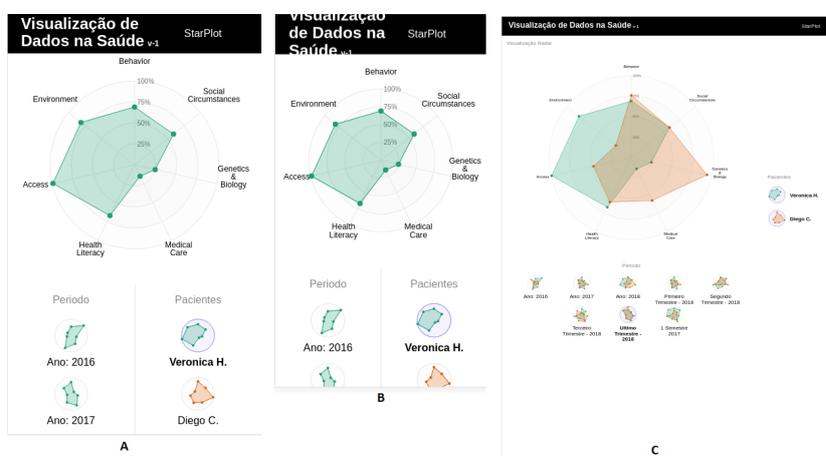


Figura 4.17 Visualização de radar em um LG Optimus L70 A, Iphone 5S B e Ipad Pro C.

4.4 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os resultados obtidos com a implementação das visualizações multidimensionais inteiramente escritos em javascript e sua biblioteca D3. Os gráficos desenvolvidos foram responsivos, de modo a permitir sua utilização em dispositivos móveis. Foi desenvolvido também as interatividades de todas as técnicas de visualizações, de modo a oferecer para o utilizador da aplicação uma melhor análise dos dados de saúde.

Neste trabalho as visualizações radar plot e o star plot, foram utilizadas com finalidades diferentes; enquanto o star plot averiguava os valores dos atributos de cada mês separadamente, o radar plot apresentava como propósito averiguar a média geral de cada atributos ao longo de cada ano.

De modo geral, telas com resoluções maiores suportam bem as visualizações, enquanto telas menores possui limitações quanto ao tamanho da tela e respectivas resoluções. Os melhores resultados foram encontrados com o dispositivo LG Optimus L70, enquanto os piores foram com o Iphone 5s.

CONCLUSÕES

Este trabalho de conclusão de curso apresentou um estudo da utilização de técnicas de visualização multidimensional para dados da saúde em dispositivos móveis. Inicialmente, foi realizada uma revisão sistemática da literatura sobre técnicas de visualização multidimensional, abordando a responsividade de cada gráfico apresentado, sendo portanto adaptadas aos tamanho de telas existentes, a exemplo dos celulares inteligentes (smartphone) ou tablets. Foi inicialmente apresentado todo o processo de seleção de artigos, classificação e estudo.

Uma das contribuições desenvolvidas neste trabalho foi a revisão sistemática da literatura que proporcionou uma larga base de conhecimento, contribuindo para o estudo das técnicas abordadas: gráfico radar, star plot, coordenadas paralelas e a visualização hierárquica, treemap. Durante a fase da revisão algumas técnicas de visualização multidimensional, como o gráfico star plot, não apresentou responsividade dificultando sua análise em dispositivos móveis; conseqüentemente foi proposto e implementado uma solução para este problema, sendo desenvolvida inteiramente em javascript. Outro desafio superado foi a responsividade no gráfico de coordenadas paralelas, pois o mesmo apresenta nove atributos para serem analisados, o que dificulta o processo de visualização e interação nos dispositivos móveis. Uma abordagem adotada para contornar este problema, foi tratar sua visualização para dispositivos com telas maiores, como tablets ou ipads garantindo todos os recursos de interações fornecidas por esta técnica. A responsividade desta técnica para os dispositivos móveis com tela de tamanho a partir de 360X640 px será considerada como trabalhos futuros.

A base de dados usada foi encontrada por meio do estudo de Ledesma, Al-Musawi e Nieminen (2016), a mesma apresentou características quantitativos contínuos, abordando atributos para medir a qualidade da saúde do paciente, de modo a permitir um conhecimento maior refinado sobre seu estado de saúde. A mesma base de dados foi utilizada em todas as técnicas abordadas neste trabalho, promovendo

perspectivas diferentes de percepções ao analisar as visualizações multidimensionais investigadas. As interações implementadas contribuíram para uma melhor análise dos resultados, uma vez que permitiu ao utilizador da aplicação fazer os devidos filtros de todas as visualizações apresentadas, como nos semestres dos anos (2016, 1017e 2018) facilitando desta maneira a tomada de decisão.

Todos as visualizações implementadas, apresentadas na seção de resultados, foram inteiramente desenvolvidas em javascript e sua biblioteca D3. Sua responsividade foram testadas por meio do navegador Chrome.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALEXANDER, B. Web 2.0. *A New Wave of Innovation for Teaching and learning*, p. 32–44, 2006.
- BUTSCHER, S. et al. Lightweight visual data analysis on mobile devices: Providing self-monitoring feedback. In: *VVH 2016-1st International Workshop on "Valuable visualization of healthcare information"*. [S.l.: s.n.], 2016. p. 28–34.
- CHAN, W. W.-Y. A survey on multivariate data visualization. *Department of Computer Science and Engineering. Hong Kong University of Science and Technology*, v. 8, n. 6, p. 1–29, 2006.
- CHITTARO, L.; SIONI, R. Evaluating mobile apps for breathing training: The effectiveness of visualization. *Computers in Human Behavior*, Elsevier, v. 40, p. 56–63, 2014.
- CHRISTMANN, C. A. et al. Effective visualization of long term health data to support behavior change. In: SPRINGER. *International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management*. [S.l.], 2017. p. 237–247.
- CORDEIRO¹, A. M. et al. Revisão sistemática: uma revisão narrativa. 2007.
- DASGUPTA, A. A framework for mobile based geospatial public health management. In: IEEE. *Students' Technology Symposium (TechSym), 2010 IEEE*. [S.l.], 2010. p. 82–87.
- FAISAL, S.; BLANDFORD, A.; POTTS, H. W. Making sense of personal health information: challenges for information visualization. *Health informatics journal*, Sage Publications Sage UK: London, England, v. 19, n. 3, p. 198–217, 2013.
- FRAIN, B. *Responsive web design with HTML5 and CSS3*. [S.l.]: Packt Publishing Ltd, 2012.
- FREITAS, C. M. D. S. et al. Introdução à visualização de informações. *Revista de informática teórica e aplicada. Porto Alegre. Vol. 8, n. 2 (out. 2001), p. 143-158*, 2001.
- HINGLE, M. et al. Collection and visualization of dietary behavior and reasons for eating using twitter. *Journal of medical Internet research*, JMIR Publications Inc., v. 15, n. 6, 2013.
- HUANG, J. et al. Seamless data visualization for frost detection. In: IEEE. *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*. [S.l.], 2011. p. 1970–1975.

- KOK, W. C. et al. Android-based disease monitoring. In: IEEE. *Information and Communication Technology (ICICTM), International Conference on*. [S.l.], 2016. p. 97–103.
- LATIF, S. et al. Real-time health data acquisition and geospatial monitoring: A visual analytics approach. In: IEEE. *Open Source Systems & Technologies (ICOSST), 2015 International Conference on*. [S.l.], 2015. p. 146–150.
- LEDESMA, A.; AL-MUSAWI, M.; NIEMINEN, H. Health figures: an open source javascript library for health data visualization. *BMC medical informatics and decision making*, BioMed Central, v. 16, n. 1, p. 38, 2016.
- LEE, J.-G. et al. Implementation of u-rpms using google chart in hybrid application for visualization of patient's biometric information. In: IEEE. *Information Science and Applications (ICISA), 2013 International Conference on*. [S.l.], 2013. p. 1–4.
- LIANG, Z. et al. Sleepexplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. *Personal and Ubiquitous Computing*, Springer, v. 20, n. 6, p. 985–1000, 2016.
- MEYER, J. et al. Visualization of complex health data on mobile devices. In: ACM. *Proceedings of the 2016 ACM Workshop on Multimedia for Personal Health and Health Care*. [S.l.], 2016. p. 31–34.
- MUNZNER, T. *Visualization analysis and design*. [S.l.]: AK Peters/CRC Press, 2015.
- PINGALI, S. *Cloud Computing and Crowdsourcing for Monitoring Lakes in Developing Countries*. 2016. 161–163 p.
- SA, J. H. et al. Georeferenced and secure mobile health system for large scale data collection in primary care. *International journal of medical informatics*, Elsevier, v. 94, p. 91–99, 2016.
- SAARY, M. J. Radar plots: a useful way for presenting multivariate health care data. *Journal of clinical epidemiology*, Elsevier, v. 61, n. 4, p. 311–317, 2008.
- SOMAN, S.; KUMAR, S.; KUMAR, S. Healthanalytic: A concept application for customizable visualization and analysis of health informatics datasets. In: IEEE. *Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on*. [S.l.], 2016. p. 1033–1038.
- STEVENS, A. et al. The care and condition monitor: Designing a tablet based tool for visualizing informal qualitative healthcare data. In: IEEE. *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*. [S.l.], 2014. p. 279–280.
- TELEA, A. C. *Data visualization: principles and practice*. [S.l.]: AK Peters/CRC Press, 2007.
- WARD, M. O.; GRINSTEIN, G.; KEIM, D. *Interactive data visualization: foundations, techniques, and applications*. [S.l.]: AK Peters/CRC Press, 2015.