

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO

FILIPE ZEFERINO

VIEW.GOV: VISUALIZAÇÃO DE DADOS DE ENDEREÇOS ELETRÔNICOS DO  
GOVERNO COM A BIBLIOTECA D3

Araranguá, 2017

FILIPPE ZEFERINO

**VIEW.GOV: VISUALIZAÇÃO DE DADOS DE ENDEREÇOS ELETRÔNICOS DO  
GOVERNO COM A BIBLIOTECA D3**

Trabalho de conclusão de curso submetido à Universidade Federal de Santa Catarina como parte dos requisitos necessários para a obtenção do Grau de Bacharel em Tecnologias da Informação e Comunicação no ano de 2017. Sob a orientação do Professor Dr. Robson Rodrigues Lemos.

Araranguá, 2017

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Zeferino, Filipe

VIEW.GOV : visualização de dados de endereços eletrônicos do governo com a biblioteca D3 / Filipe Zeferino ; orientador, Robson Rodrigues Lemos, 2017.  
55 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Campus Araranguá, Graduação em Tecnologias da Informação e Comunicação, Araranguá, 2017.

Inclui referências.

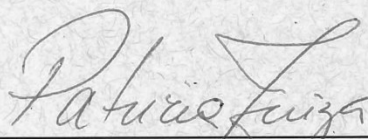
1. Tecnologias da Informação e Comunicação. 2. Tecnologias da Informação e Comunicação. 3. Dados Abertos do Governo Federal. 4. Visualização de Dados. 5. Aplicação WEB. I. Lemos, Robson Rodrigues. II. Universidade Federal de Santa Catarina. Graduação em Tecnologias da Informação e Comunicação. III. Título.

FILIPE ZEFERINO

**VIEW.GOV: VISUALIZAÇÃO DE DADOS DE ENDEREÇOS ELETRÔNICOS DO  
GOVERNO COM A BIBLIOTECA D3**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de  
“Bacharel em Tecnologia da Informação e Comunicação”.

Araranguá, 27 de novembro de 2017



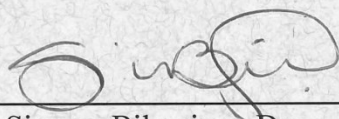
---

Profª. Patricia Jantsch Fiuza, Dra  
Coordenadora do curso  
Universidade Federal de Santa Catarina



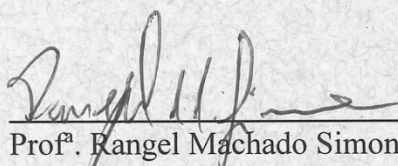
---

Prof. Robson Rodrigues Lemos, Dr  
Orientador  
Universidade Federal de Santa Catarina



---

Profª. Simone Bilessimo, Dra  
Examinador  
Universidade Federal de Santa Catarina



---

Profª. Rangel Machado Simon, MSc  
Examinador  
Universidade Federal de Santa Catarina

## **AGRADECIMENTOS**

Primeiramente quero prestar meu agradecimento ao professor Robson Rodrigues Lemos pela confiança e paciência na elaboração deste trabalho. Agradeço também aos meus pais pelo incentivo a fazer este curso e a minha esposa por estar ao meu lado e me apoiar no decorrer dessa jornada. As minhas duas filhas pelo carinho nas horas difíceis. E a todos os meus colegas de curso que de uma forma ou de outra contribuíram para que juntos pudéssemos concluir essa caminhada.

## RESUMO

A quantidade de dados vem crescendo exponencialmente e cada vez mais cresce a demanda por ferramentas para gerir e visualizar essas informações. O objetivo deste trabalho de conclusão de curso é propiciar ao usuário uma visão geral e uma aplicação que possa exemplificar diferentes formas de visualização de dados e de interações de um conjunto de dados. Para atingir esses objetivos, foi realizada uma pesquisa exploratória a fim de identificar técnicas e algoritmos de visualização de dados. Um conjunto de dados foi obtido para usar como base para a aplicação que foi desenvolvida na linguagem PHP, utilizando banco de dados em MySQL para a consulta dos dados, e a biblioteca D3 (*Data-Driven Documents*) para o processamento e a visualização de dados. Com a apresentação dos resultados, pode-se verificar que a ferramenta desenvolvida intitulada VIEW.GOV contribui para o aprendizado e a tomada de decisão do usuário em relação aos dados de domínios registrados pelo Governo Federal Brasileiro.

**Palavras-chave:** Visualização de Dados, Dados Abertos, Governo Federal, PHP, MySQL, JavaScript, JSON, D3, Tecnologias da Informação e Comunicação.

## **ABSTRACT**

The amount of data has been growing exponentially and the demand for tools to manage and visualize those information is growing. The purpose of this undergraduate project is to provide the user with an overview and an application where is possible to explore data visualization in different ways and interactions of a set of data. To achieve those objectives, an exploratory research was carried out to identify techniques and algorithms for data visualization. A set of data was obtained to use as a base for the application that was developed in the PHP language, using the MySQL database for the data query, and the D3 (Data-Driven Documents) library for processing and data visualization. With the presentation of the results, it can be verified that the developed tool entitled VIEW.GOV contributes to the learning and the decision making of the user in relation to the data of domains registered by the Brazilian Federal Government.

Keywords: Data Visualization, Open Data, Federal Government, PHP, MySQL, JavaScript, JSON, D3, Information and Communication Technologies.

## LISTA DE IMAGENS

Figura 1 - Gráfico Bubble Chart - D3 Examples.....	18
Figura 2 - Force-Directed Graph – D3 Examples .....	19
Figura 3 - 3D Bar Chart.....	20
Figura 4 - Visualização científica e visualização da informação .....	21
Figura 5 - Exemplo de visualização de dados com a NVD3 .....	25
Figura 6 - Tela inicial do AdminLTE.....	26
Figura 7 - Modelo conceitual feito com o programa brModelo .....	28
Figura 8 - Modelo lógico .....	28
Figura 9 - Script DDL ou modelo físico.....	29
Figura 10 - Representação da estrutura da aplicação utilizando o D3.....	30
Figura 11 - Representação da visualização agrupada por ano de cadastro .....	31
Figura 12- Gráfico estilo árvore (Tree) dos domínios registrados por cidade/UF .....	32
Figura 13- Gráfico Bubble Chart com os dados separados por setor de atuação. ....	32
Figura 14 - Exemplo de como visualizar a quantidade de domínios por setor de atuação.....	33
Figura 15 - Visualização de dados representados no mapa.....	34
Figura 16 - Visualização dos dados por <i>timeline</i> do ano de registro.....	35
Figura 17 - Seletor de intervalo de anos.....	35
Figura 18 - Barra de domínios totais com informação da região .....	35
Figura 19 - Dados filtrados por região e por intervalo de tempo. ....	36
Figura 20 - Informações agrupadas em forma de tabela dinâmica com opção de ordenação ..	36
Figura 21 - Página sobre o projeto.....	37
Figura 22 - Página inicial com algumas estatísticas.....	37
Figura 23 - Script para importar as informações para o banco de dados MySQL. ....	40
Figura 24 - Conexão com o banco de dados MySQL.....	41
Figura 25 - Script para extrair a latitude e longitude de cada domínio no banco de dados.....	42
Figura 26 - Script em PHP para gerar uma saída em JSON .....	43
Figura 27 - Inclusão dos arquivos CSS. ....	44
Figura 28 - Inclusão dos scripts em JavaScript no fim das páginas. ....	44
Figura 29 - Carregamento dos dados para o D3 .....	45
Figura 30 - Trecho do código que inicializa o mapa. ....	45
Figura 31 - Trecho do código em que são criados os marcadores e os círculos com o D3.....	46
Figura 32 - Exemplo de marcador criado no mapa .....	46
Figura 33 - Exemplo de valores esperados pelo D3 para iniciar a visualização dos dados .....	47



Figura 34 - Código em PHP para entregar as informações requeridas do banco de dados em um JSON válido. ....	47
Figura 35 - Domínios do Governo Federal registrados por ano .....	49
Figura 36 - Domínios registrados em cidades distintas .....	49
Figura 37 - Total de domínios por setor de atuação .....	50
Figura 38 - Detalhe de visualização no mapa usando o recurso de “zoom” .....	51
Figura 39 - Visualização por linha do tempo, ou <i>timeline</i> .....	51
Figura 40 - Visualização do VIEW.GOV a partir de um Smartphone .....	52

## LISTA DE ABREVIATURAS E SIGLAS

AdminLTE – *Template* para construção de aplicações WEB com várias ferramentas

API – Conjunto de padrões de um software ou plataforma fornecido para uso

CSS – *Cascading Style Sheets*

DDL – *Data definition Language* (Linguagem de definição de dados)

DOM – *Document Object Model* (Modelo de objeto de documento)

HTML – *HyperText Markup Language*

MySQL – *My Structured Query Language*

Open-Source – Software livre

PHP – *HyperText PreProcessor*

SGBD – Sistema de Gerenciamento de Banco de Dados

SQL – *Structured Query Language*

TIC – Tecnologias da Informação e Comunicação

Tooltip – Elemento visual flutuante usado para mostrar alguma informação quando o mouse for passado por cima

Ubuntu – Distribuição Linux *open-source*

UFSC – Universidade Federal de Santa Catarina

WWW – *World Web Wide*

XAMPP – Ambiente de desenvolvimento que serve como servidor local e permite executar PHP e Mysql

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	13
1.1	<b>Justificativa</b>	14
1.2	<b>Objetivos</b>	14
1.2.1	Objetivo Geral	14
1.2.2	Objetivos Específicos	14
2	<b>METODOLOGIA</b>	15
2.1	<b>Organização do Trabalho</b>	15
3	<b>REFERENCIAL TEÓRICO</b>	16
3.1	<b>Representações Visuais</b>	16
3.1.1	Representação dos dados	16
3.2	<b>A visualização de dados</b>	16
3.2.1	Análise visual	17
3.2.2	Técnicas de visualização	17
3.2.3	Visualização da Informação e Visualização Científica	21
4	<b>DESENVOLVIMENTO DA APLICAÇÃO WEB</b>	22
4.1	<b>Infraestrutura de Hardware</b>	22
4.1.1	A linguagem PHP	22
4.1.2	O banco de dados MySQL	22
4.2	<b>Infraestrutura para o <i>front-end</i></b>	23
4.2.1	A linguagem de marcação HTML	23
4.2.2	A Linguagem de estilos CSS	23
4.2.3	A linguagem JavaScript	23
4.2.4	A biblioteca jQuery	24
4.2.5	O padrão de dados JSON	24
4.2.6	A biblioteca D3	24
4.2.7	A biblioteca NVD3	25
4.2.8	AdminLTE	25
5	<b>PROJETO DA APLICAÇÃO WEB</b>	27
5.1	<b>Modelagem de dados</b>	27
5.1.1	Modelo Conceitual	27
5.1.2	Modelo Lógico	28
5.1.3	O Script DDL	29
5.2	<b>Projeto da Interface de Visualização Interativa</b>	29
5.2.1	Visualização por ano	30

5.2.2	Visualização por cidade/UF .....	31
5.2.3	Visualização por setores de atuação .....	32
5.2.4	Visualização por localização no mapa do Brasil .....	33
5.2.5	Visualização por linha do tempo do ano de registro .....	34
5.2.6	Informações Complementares .....	36
6	<b>DESENVOLVIMENTO DA APLICAÇÃO WEB</b> .....	38
6.1	<b>Caracterização dos Dados</b> .....	38
6.2	<b>Técnicas de Visualização</b> .....	39
6.3	<b>Configurações do servidor</b> .....	40
6.3.1	Carga e tratamento dos dados .....	40
6.4	<b>Funcionalidades do <i>front-end</i></b> .....	43
7	<b>RESULTADOS</b> .....	48
8	<b>CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS</b> .....	53
9	<b>REFERÊNCIAS BIBLIOGRAFICAS</b> .....	54

# 1 INTRODUÇÃO

Nos dias de hoje, a quantidade de dados disponíveis de determinados assuntos é extremamente grande. O termo *big-data*, utilizado hoje em dia para indicar um grande volume de dados, demanda por muito poder de processamento e ferramentas cada vez mais robustas para trabalhar, com dados cada vez mais complexos. As Tecnologias da Informação e Comunicação (TIC) tem um grande papel nesse meio, com o objetivo de trazer novas funcionalidades ou aprimoramentos ao contexto atual. Cada dia mais, os dados crescem de forma exponencial. Deste modo, surge a questão de como o usuário comum pode tratar e visualizar esses dados sem uma ferramenta adequada. Os dados, sem um tratamento correto, sem um processamento mais minucioso e sem uma visualização intuitiva, resultam na diminuição do interesse do usuário na busca de informações adicionais. Por exemplo, o estoquista de um supermercado, de posse de uma lista pequena de produtos, poderá vivenciar as seguintes situações: no primeiro olhar, o estoquista saberia onde guardar cada produto. Mas e se fosse uma lista com cinco mil itens? Provavelmente ele perderia muito tempo somente para separar os itens por setor e para organizar os produtos em uma determinada ordem, de uma forma eficiente.

Esse tipo de problema é muito comum e é por isso que cada vez mais a visualização de dados é tratada com atenção por desenvolvedores e usuários avançados envolvidos com as TICs. Os métodos de visualização baseados em gráficos são uma maneira simples e universal de mostrar dados os quais tem como objetivo auxiliar o usuário na compreensão e tomada de decisão.

Uma visualização de dados adequada pode gerar inúmeros benefícios, que vão desde uma comunicação mais efetiva, até a obtenção de *insights* para a tomada de decisão. Por meio da inovação, torna-se possível extrair informações valiosas e novas percepções em conjunto de dados que antes não eram utilizados. Ou seja, seja para aprender, para ensinar ou para compartilhar dados, a visualização pode ser o meio mais eficaz nesses processos (MARQUESONE, 2016, p. 235).

Neste trabalho de conclusão de curso será investigado a importância e os benefícios da visualização de grande volume de dados, mais precisamente a visualização de dados de endereços eletrônicos registrados pelo governo visando proporcionar uma visualização mais clara, objetiva e amigável para o usuário final.

## **1.1 Justificativa**

Os dados abertos do governo são muito úteis para pesquisa de informações no geral. Mas o que torna a experiência não desejável é o método de visualização que a maioria dos conjuntos de dados é apresentada. Os dados não recebem nenhum tipo de tratamento e/ou processamento. Muitas vezes, são apresentados arquivos em formato que o usuário final nem conseguirá ou não terá conhecimento de como abrir ou visualizar os mesmos. Dentro deste contexto, o objetivo deste trabalho de conclusão de curso é desenvolver uma aplicação WEB para ser utilizada por usuários experientes e não experientes no sentido de melhorar o processo de agrupamento e exposição do conjunto de dados abordado. Essa aplicação terá que ser capaz de agrupar os dados vindos de uma base de dados MySQL previamente preenchida com o conjunto de dados escolhido buscando tornar a experiência do usuário mais prática e intuitiva.

## **1.2 Objetivos**

### **1.2.1 Objetivo Geral**

Este trabalho de conclusão de curso tem como objetivo geral desenvolver uma ferramenta WEB para visualização de dados de endereços eletrônicos do Governo Federal para facilitar o processo de análise e visualização de dados.

### **1.2.2 Objetivos Específicos**

Para que seja possível alcançar o objetivo geral, os seguintes objetivos específicos deverão ser obrigatórios:

- Analisar as tecnologias utilizadas em visualização de dados disponíveis no mercado.
- Estudar algoritmos de visualização de dados.
- Investigar as tecnologias que serão usadas para o desenvolvimento da aplicação WEB.
- Escolher um conjunto de dados específico a ser utilizado como base.
- Fornecer o acesso a aplicação WEB e a interface de visualização.

## 2 METODOLOGIA

A metodologia deste trabalho de conclusão de curso foi baseada em uma pesquisa descritiva que de acordo com Gil (2008), possuem como objetivo a descrição das características de uma população, fenômeno ou de uma experiência. Foi realizado um estudo na área de visualização de dados e *big-data*. Para o desenvolvimento da aplicação, foram estudadas tecnologias bastante utilizadas no mercado quando se trata de desenvolvimento WEB e visualização de dados baseados em gráficos vetoriais.

Os dados utilizados na aplicação foram extraídos da base de dados abertos do Governo Federal do Brasil, tratando de todos os endereços eletrônicos registrados no portal do Registro.br e autorizados pelo Ministério do Planejamento de 1995 até outubro de 2017 (Dados Abertos, 2017). A base de dados abertos do governo conta mais de 2900 conjuntos de dados em várias áreas do setor público e é atualizado com frequência.

Para o desenvolvimento da aplicação WEB que servirá de ferramenta para a visualização dos dados, foram utilizados somente softwares gratuitos e com licenças *open-source*. A linguagem de programação escolhida foi a PHP e, como ambiente de desenvolvimento e testes, o servidor local utilizado foi o XAMPP.

No armazenamento dos dados extraídos da base de dados do governo, foi utilizado o MySQL, que já está incluído no pacote do ambiente de desenvolvimento XAMPP e é o Software de Gerenciamento de Base de Dados (SGBD) que o autor tem mais afinidade.

Na disposição e visualização dos dados obtidos, foi escolhido o framework D3 que também utiliza uma licença *open-source*.

### 2.1 Organização do Trabalho

O trabalho segue a seguinte organização:

No segundo terceiro, é apresentada a contextualização da aplicação de visualização de dados.

O quarto capítulo aborda as tecnologias usadas para realização da aplicação.

O quinto capítulo trata das funcionalidades da aplicação WEB.

O sexto capítulo apresenta detalhes do desenvolvimento da aplicação WEB incluindo alguns códigos-fonte importantes no processo de desenvolvimento.

O sexto sétimo visa a relatar os resultados obtidos com a ferramenta de visualização dos domínios GOV.BR.

O oitavo e último capítulo apresenta os resultados obtidos e as conclusões.

### 3 REFERENCIAL TEÓRICO

O referencial teórico tem um papel importante no trabalho de conclusão de curso, pois propicia um universo de pesquisa bibliográfica e garante o estudo correto do tema proposto pelo acadêmico.

Segundo Yin o objetivo do referencial teórico é:

Fornecer um esquema completo o suficiente para desenvolver a pesquisa e isso exige proposições teóricas. É essencial que se desenvolva uma teoria antes que se faça a coleta de dados para qualquer estudo de caso (2001, p. 205).

#### 3.1 Representações Visuais

Para Matthes:

A visualização de dados envolve a sua exploração por meio de representações visuais. Ela está intimamente relacionada ao *data mining* (mineração de dados), que faz uso de algoritmos para explorar padrões e conexões, em um conjunto de dados, o qual pode ser apenas uma pequena lista de números que caiba em uma linha de código ou muitos *gigabytes* de dados (2016, p. 656).

##### 3.1.1 Representação dos dados

Dado, informação, conhecimento e sabedoria constituem a cadeia de evolução de dados e informações. Quando falamos sobre “dados”, na verdade estamos nos referindo à base da matéria-prima necessária para conseguir o que todas as empresas desejam: utilizar o conhecimento das informações para tomar decisões ágeis e corretas (RÊGO, 2013, p. 312).

O repositório de dados abertos do governo conta com diversos conjuntos, dos mais variados assuntos. As técnicas de representação e de visualização dos dados são fortemente influenciadas pela escolha do conjunto.

#### 3.2 A visualização de dados

A visualização de dados tem o objetivo de ser a mais clara possível para permitir ao usuário a fácil compreensão do assunto abordado. Não é somente um conjunto de imagens soltas em um espaço. É um conjunto de fatores visuais que auxiliam o espectador, em um primeiro olhar, para a compreensão dos dados e a tomada de decisão.

O crescimento acelerado da Internet como meio de se obter informações dentro dos mais diversos domínios de aplicações tornam-se necessária a incorporação de técnicas visuais que auxiliem o usuário a interagir com esse vasto universo de informações de forma eficaz e intuitiva. A Visualização de Informação é uma área de pesquisa que estuda o uso de representações visuais e interativas de informações com o objetivo de diminuir a sobrecarga cognitiva dos usuários ao interpretá-las (OSHIRO, 2003, p. 5).



### 3.2.1 Análise visual

Um fator importante a ser levado em consideração para a escolha da técnica de visualização é o conjunto de dados a ser usado. As técnicas de análise visual sempre focam e deixam o objetivo final mais parecido possível com o mundo real. Por outro lado, podem ser levados em consideração o uso de formas geométricas, linhas e pizza, que são mais comuns entre gráficos e vêm sendo usado desde os primeiros passos das técnicas de visualização de dados. O objetivo da análise é montar uma estrutura visual para representar o conjunto de dados.

Segundo Amaral (2016, p. 127), o processo de analisar dados nada mais é do que descrever um fato: contar uma história. A visualização de dados tem um papel fundamental em contar esta história e, por isso, está intimamente relacionada com a ciência de dados.

O processo de visualização de dados envolve várias etapas, desde o estudo do conjunto de dados a ser visualizado até as tecnologias envolvidas para a concepção da aplicação final. Todas as etapas devem ser previamente analisadas para se cumprir o objetivo com efetividade.

### 3.2.2 Técnicas de visualização

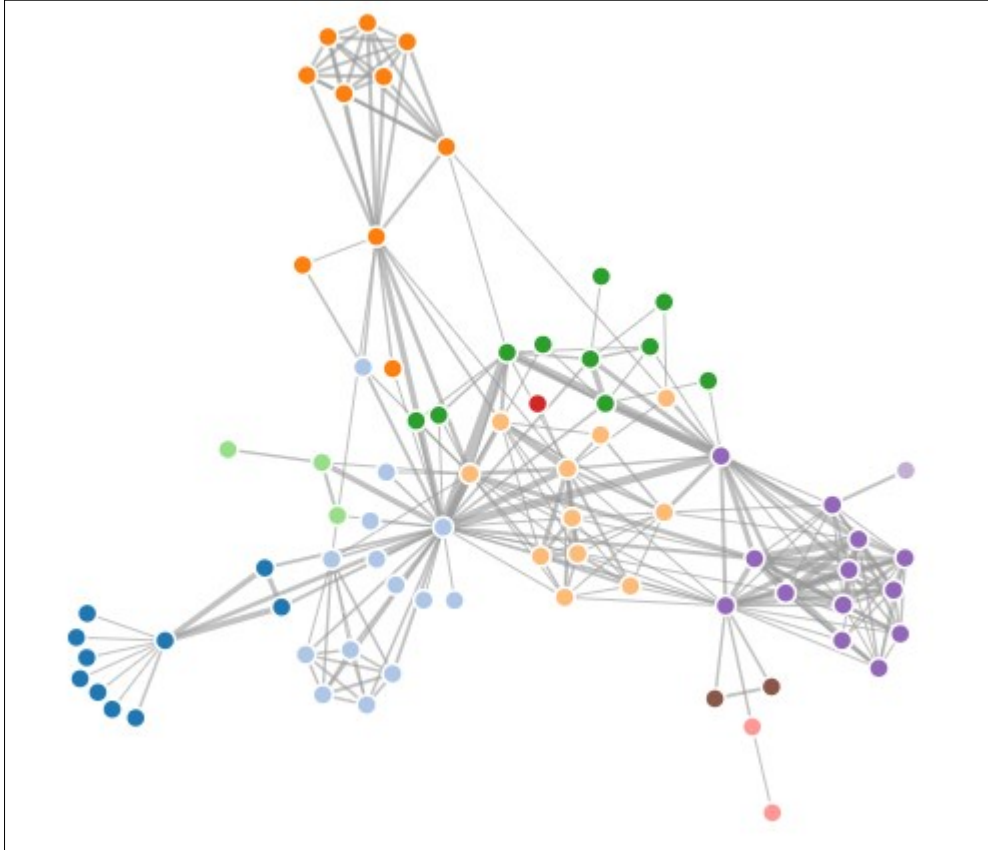
Hoje em dia temos várias técnicas de visualização. As mais usadas são: Técnicas de visualização com pontos, em linhas, em volumes, em áreas e com figuras. Para Marr (1982), o ato de ver é uma tarefa complexa que abrange a aquisição de imagens do ambiente, a representação e o armazenamento adequado das mesmas e a extração de informações relevantes dentro de um contexto. Abaixo descreveremos algumas técnicas mais comuns para a visualização de dados.

Técnicas geométricas são técnicas que apresentam os dados usando transformações geométricas para representação da informação coletada. Um exemplo de gráfico usando forma geométrica é o *Bubble Chart* que usa círculos para representar cada propriedade do conjunto de dados como mostra a Figura 1.



As técnicas baseadas em grafos fornecem uma visualização parecida como uma rede neural. Eles mostram claramente seus vértices, ligados por arestas, chamadas de relações diretamente interligadas. Podem representar, por exemplo, a distância entre cidades.

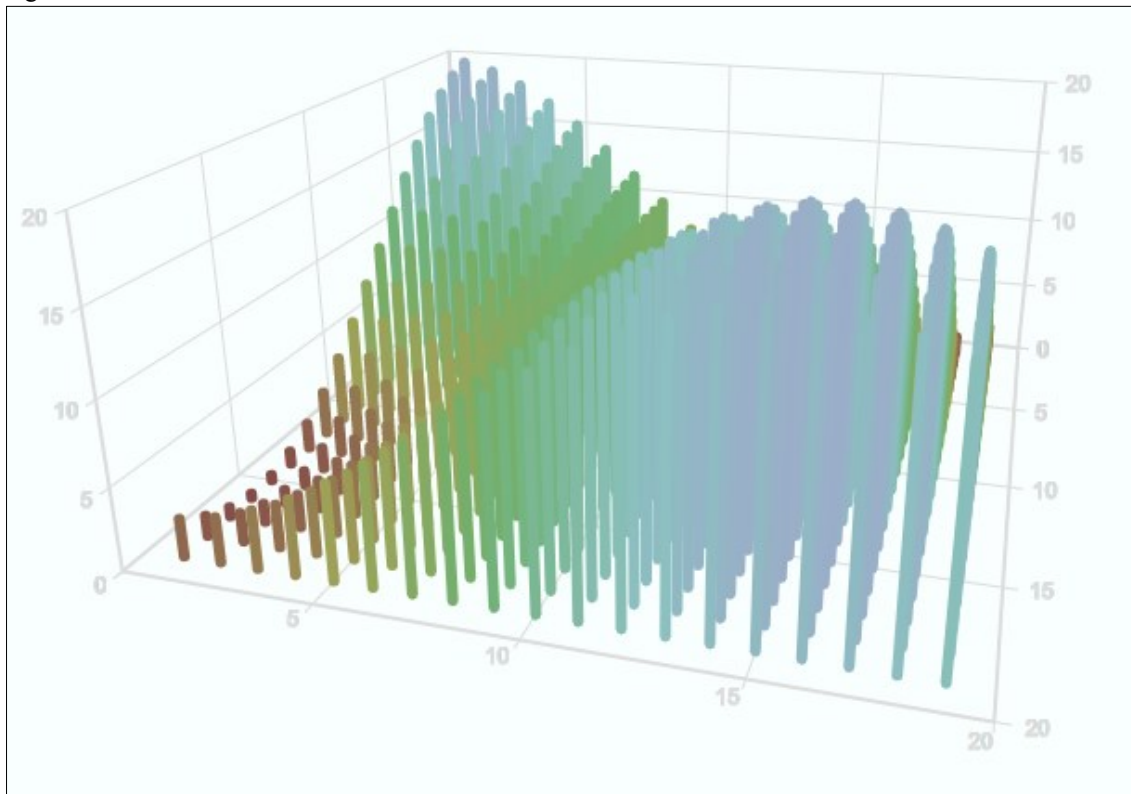
Figura 2 - Force-Directed Graph – D3 Examples.



Fonte: *Force-Directed Graph* (BOSTOCK, 2017).

Técnicas tridimensionais usam gráficos tridimensionais para visualização dos dados e, geralmente, têm uma maior interação com o usuário, o qual atua com filtros, na seleção dos dados. As técnicas tridimensionais necessitam de mais processamento do hardware, visto que existem mais efeitos visuais nessa técnica.

Figura 3 - 3D Bar Chart.



Fonte: 3D Bar Chart (TOUCANO, 2016).

Das técnicas de visualização, a que mais tem abrangência é a de mapas. Podemos representar muitos conjuntos de dados usando mapas. São mundialmente conhecidos e servem de fundo para dados baseados em regiões geográficas. Para a marcação dos pontos os mapas, necessitam somente da latitude e longitude e, para a visualização, várias outras técnicas podem ser empregadas.

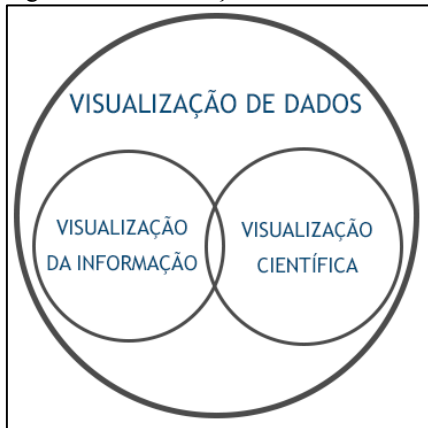
Grinstein e Kein (2015) classificam os tipos de representações visuais em quatro categorias:

- **Dados Multivariados:** Dados que não representam geralmente um atributo espacial explícito.
- **Hierarquias e Árvores:** Repositório de dados em que os itens de dados estão relacionados com subitens de dados.
- **Grafos e Redes:** Grafos permitem a representação da informação através de um conjunto de nodos (os quais podem estar associados a algum tipo de informação) conectados dois a dois por arestas (as quais podem estar associadas a algum tipo de relacionamento). Árvores são um tipo de grafo, conectado, não ponderado e acíclico.
- **Texto e Documento:** A visualização da Informação pode auxiliar na compreensão de uma grande coleção de documentos de uma forma mais rápida.

### 3.2.3 Visualização da Informação e Visualização Científica

Podemos afirmar que a visualização de dados pode ser representada por um conjunto formado pela visualização científica e a visualização da informação (Figura 4).

Figura 4 - Visualização científica e visualização da informação.



Fonte: Elaborado pelo autor (2017).

O conceito de visualização científica é formado pelo uso de recursos computacionais gráficos que ajudam a compreender conceitos científicos geralmente associados a um volume de dados complexo e a visualização da informação utiliza-se de dados abstratos para suas representações com o objetivo de agregar conhecimento e tomadas de decisões.

## 4 DESENVOLVIMENTO DA APLICAÇÃO WEB

Para o desenvolvimento da aplicação, foi usada uma estrutura a fim de fornecer ferramentas necessárias para alcançar o objetivo. A aplicação WEB de visualização de dados pode ser dividida em duas partes principais: o *back-end* e o *front-end*. O *back-end* é onde ficará o hardware (e.g., servidor) e uma parte do software. O software inclui banco de dados MySQL e PHP como linguagem para desenvolvimento ambos *open-source*. Já o *front-end* é o resultado da nossa aplicação. É o que o usuário irá “consumir”.

### 4.1 Infraestrutura de Hardware

O servidor, basicamente é uma unidade computacional com recursos mais aprimorados e específicos. A maioria dos servidores WEB conta com um grande poder de armazenamento e processamento. A existência de softwares da aplicação no servidor o caracteriza como um servidor WEB. Para este trabalho, foi usado a linguagem PHP e o banco de dados MySQL. A junção desses dois fatores, hardware e software, formam o *back-end*.

#### 4.1.1 A linguagem PHP

A principal linguagem usada no trabalho foi a PHP. É uma linguagem muito usada, nos dias atuais, para o desenvolvimento de aplicações e sites para a WEB e suporta aplicações de pequeno, médio e grande porte.

Para Converse e Park (2003, p. 27), o PHP, permite incorporar fragmentos de código em páginas de HTML normais – código esse que é interpretado à medida que suas páginas são oferecidas aos usuários. O PHP também serve com uma linguagem de “cola”, facilitando a conexão de suas páginas WEB com o banco de dados do servidor.

#### 4.1.2 O banco de dados MySQL

O banco de dados MySQL é usado amplamente nos dias de hoje. Ele tem uma comunidade de usuários bem grande e engajada, é o SGBD mais popular atualmente. Sua linguagem é relativamente fácil para o aprendizado e estudo. É compatível com a maioria dos sistemas operacionais. É *open-source*, o que torna a utilização dele viável para este trabalho. Tem desempenho, confiabilidade e facilidade no seu manuseio comprovada e é usada desde sistemas pequenos até aplicações de grande porte.

## 4.2 Infraestrutura para o *front-end*

Para entregar uma boa aplicação para os usuários, o *front-end* deverá ser bem estruturado. Ele é quem irá acessar os recursos do servidor e entregar ao usuário. Para que esse objetivo seja cumprido, alguns recursos devem ser usados. Abaixo serão descritos esses recursos que compõe nossa aplicação.

### 4.2.1 A linguagem de marcação HTML

Para a estrutura de páginas que contenham formulários e elementos visuais, por exemplo, usaremos a linguagem HTML. É a linguagem mais usada para a criação de sites e aplicações WEB. Mas o HTML, sozinho, não faz todo o trabalho, por isso a junção do HTML com o PHP.

O HTML é uma linguagem de marcação, utilizada para formatar páginas WEB com texto e informações separadamente. Estas páginas são acessíveis de qualquer parte do mundo, usando qualquer tipo de browser como: Internet Explorer, Firefox, Chrome, Opera ou Safari. (CARRIL,2012, p. 3).

Segundo Quierelli (2012, p. 7), a linguagem HTML se tornou um padrão para o desenvolvimento de web sites na internet. Apesar de ser uma linguagem de marcação, e não programação, a mesma é de fácil aprendizado, não exigindo um computador potente, nem softwares específicos, na criação de sites.

### 4.2.2 A Linguagem de estilos CSS

Para Jobstraibizer (2009, p. 6), o CSS é uma das ferramentas mais amplamente difundidas nos dias atuais. O termo CSS, proveniente de *Cascading Style Sheets*, que significa *Folhas de Estilos em Cascata* é uma linguagem de estilo (ou de formatação) fantástica para construção do layout de suas páginas ou sites.

### 4.2.3 A linguagem JavaScript

O JavaScript é uma linguagem de programação, criada por Brendan Eich e baseada em scripts. Ela foi criada em 1995 para suprir a necessidade de execução de scripts no lado do cliente (usuário). O JavaScript, junto com o HTML, permite dar um dinamismo maior para a aplicação WEB, visto que executa scripts, mesmo que menos complexos, no browser do usuário.

JavaScript é uma linguagem de programação interpretada com recursos orientados a objetos (OO). Sintaticamente, a linguagem core JavaScript se assemelha ao C, C++ e Java. (FLANAGAN, 2006)

#### 4.2.4 A biblioteca jQuery

O jQuery é uma biblioteca baseada em JavaScript que foi criada por John Resig, em 2006 e tem o objetivo de simplificar a programação de efeitos visuais, em páginas HTML. A maior vantagem do jQuery é reduzir as linhas de código escritas, proporcionando ao desenvolvedor maior eficiência e economia de tempo, na execução do projeto. Atualmente, o jQuery está na versão 3.2.1 e opera nos principais navegadores existentes hoje.

Para Rutter (2012), o jQuery é uma biblioteca JavaScript que simplifica a criação de efeitos visuais e de interatividade em sites. Ela pode diminuir drasticamente o tempo necessário para o desenvolvimento WEB, permitindo escrever componentes interativos que parecem impossíveis sem um grande conhecimento de programação.

#### 4.2.5 O padrão de dados JSON

O JSON é conhecido como o padrão para troca de dados, o que implica que ele pode ser usado como formato de dados sempre que houve uma troca de dados. Uma troca de dados pode ocorrer entre o navegador e o servidor, e até mesmo entre um servidor e outro, se for o caso. (SMITH, 2015, p. 66)

#### 4.2.6 A biblioteca D3

Quando se trata de visualização de dados, a biblioteca d3 é bem completa e flexível.

A biblioteca D3 é uma biblioteca baseada em JavaScript para manipular documentos com base em um conjunto de dados. D3 ajuda o desenvolvedor a trazer dados à vida usando HTML, SVG e CSS. A ênfase da D3 nos padrões da WEB oferece-lhe as capacidades completas dos navegadores modernos sem se amarrar a uma estrutura proprietária, combinando componentes de visualização poderosos e uma abordagem baseada em dados para a manipulação de DOM (D3, 2017).



#### 4.2.7 A biblioteca NVD3

Segundo a Partners (2014), o projeto NVD3 é uma biblioteca que tem objetivo de construir gráficos reutilizáveis e componentes gráficos para D3, sem tirar o poder que ela lhe dá. Esta é uma coleção muito nova de componentes, com o objetivo de manter esses componentes muito personalizáveis, mantendo-se longe de suas soluções padrão.

O NVD3 é, atualmente, mantido por uma equipe de engenheiros de software de *front-end* na Novus Partners. A base de códigos da NVD3 é fortemente inspirada pelo trabalho de Mike Bostock.

Figura 5 - Exemplo de visualização de dados com a NVD3.



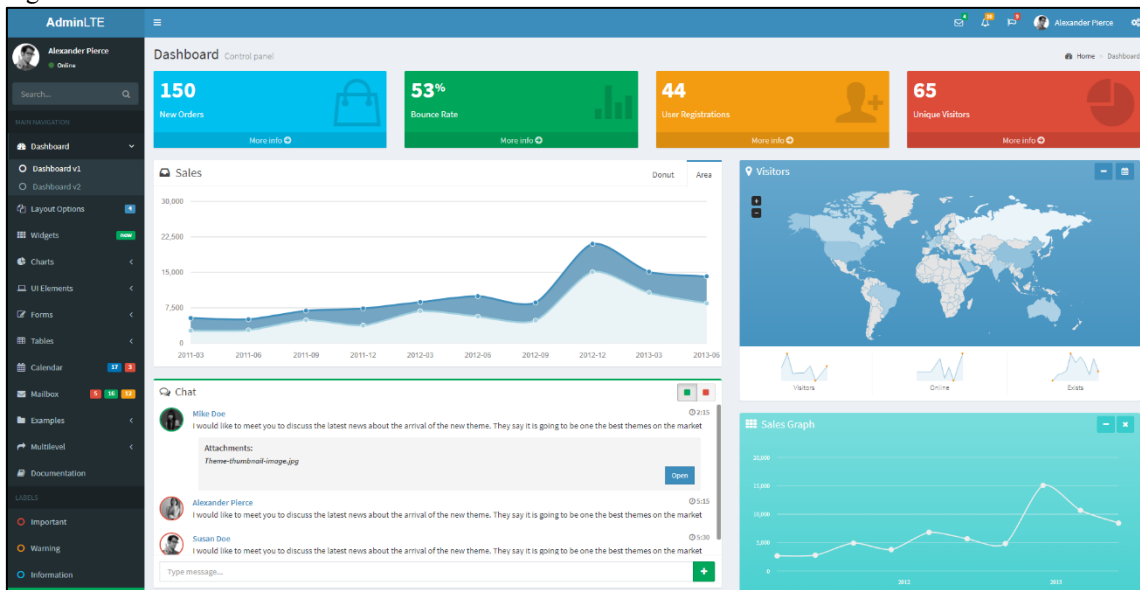
Fonte: NVD3 (2017).

#### 4.2.8 AdminLTE

AdminLTE é um template de administração totalmente responsivo baseado no framework Bootstrap 3. Ele é altamente personalizável e fácil de usar. Adapta-se a muitas resoluções de tela desde pequenos dispositivos móveis até grandes desktops. (ALMSAEED, 2017)

O AdminLTE fornece várias ferramentas junto ao seu pacote, baseadas em jQuery, CSS e HTML, que facilitam na hora do desenvolvimento de funções para a aplicação. Possui também uma documentação completa e bem estruturada, a fim de ajudar o desenvolvedor, na hora da programação dos recursos disponíveis.

Figura 6 - Tela inicial do AdminLTE.



Fonte: AdminLTE (2017).

## 5 PROJETO DA APLICAÇÃO WEB

O projeto da aplicação WEB surgiu da necessidade de se ter uma ferramenta intuitiva e de fácil acesso para a visualização de dados de endereços eletrônicos brasileiros registrados. Dentro deste contexto, o conjunto de dados foi organizado e classificado para melhorar a precisão das informações.

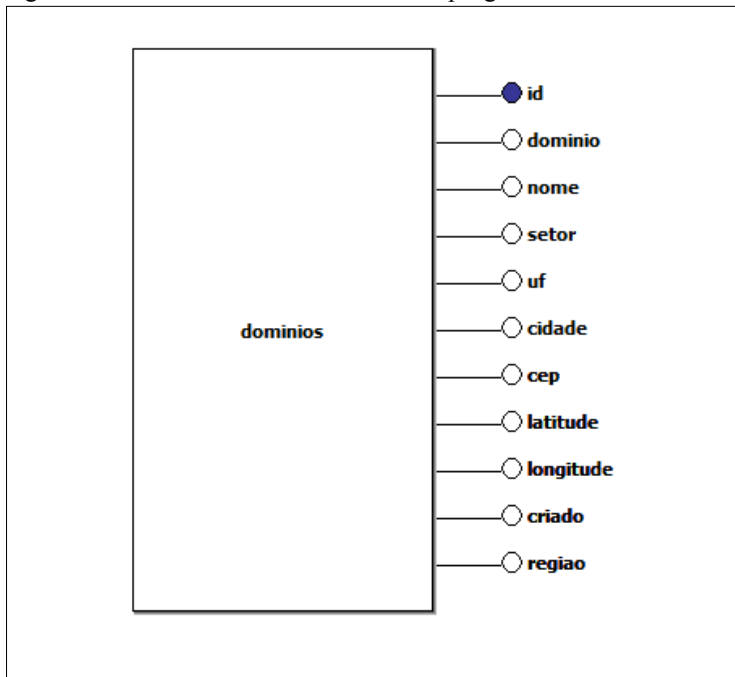
### 5.1 Modelagem de dados

O conjunto de dados não normalizados foi extraído do portal de dados abertos do Governo Federal do Brasil (disponível no site: <<http://dados.gov.br>>) que contém vários outros conjuntos, de diferentes setores públicos. Para os dados abertos de endereços eletrônicos brasileiros registrados, foi utilizado, inicialmente, uma lista bruta, para popular uma tabela, em um banco de dados em MySQL, a qual será a base de dados utilizada para a aplicação. A modelagem dos dados e suas características estão descritas nas seções abaixo.

#### 5.1.1 Modelo Conceitual

O modelo conceitual proposto é apresentado no modelo de diagramas de entidade-relacionamento. Este modelo mostra quais tabelas e relações serão utilizadas. No caso da aplicação em questão, usaremos somente uma tabela chamada “**domínios**” para guardar os dados. Os atributos utilizados para a aplicação, contidos no conjunto de dados são: domínio, nome, cidade, UF, CEP e data de cadastro. Além dessas informações originais, foram criados mais quatro atributos, para complementar as informações e prover formas de visualização mais abrangentes. São eles: setor, região, latitude e longitude. A Figura 7 descreve o modelo proposto.

Figura 7 - Modelo conceitual feito com o programa brModelo.




Fonte: Elaborado pelo autor (2017).

### 5.1.2 Modelo Lógico

O modelo conceitual serve de base para saber como serão as entidades e relacionamentos contidos no banco de dados. Na Figura 7, não temos nenhum relacionamento, visto que nossa aplicação usará somente uma tabela como base. O modelo ainda identifica qual atributo será chave primária no banco. A Figura 8, apresenta o modelo lógico, gerado a partir de um modelo conceitual feito no programa brModelo e apresenta uma visão mais técnica da disposição característica de cada coluna da nossa entidade.

Figura 8 - Modelo lógico.

<b>dominios</b>	
	<b>id: Int(11)</b>
	<b>dominio: Varchar(100)</b>
	<b>nome: Varchar(255)</b>
	<b>setor: Varchar(100)</b>
	<b>uf: Varchar(2)</b>
	<b>cidade: Varchar(100)</b>
	<b>cep: Varchar(10)</b>
	<b>latitude: Varchar(15)</b>
	<b>longitude: Varchar(15)</b>
	<b>criado: Varchar(20)</b>
	<b>regiao: Varchar(20)</b>

Fonte: Elaborado pelo autor (2017).

### 5.1.3 O Script DDL

A partir do modelo lógico, foi gerado o script DDL também chamado de modelo físico. De posse desse modelo, poderemos criar no banco de dados as colunas da entidade a ser usada na aplicação. A Figura 9 apresenta o resultado final do script gerado.

Figura 9 - Script DDL ou modelo físico.

```
-- Geração de Modelo físico
-- Sql ANSI 2003 - brModelo.

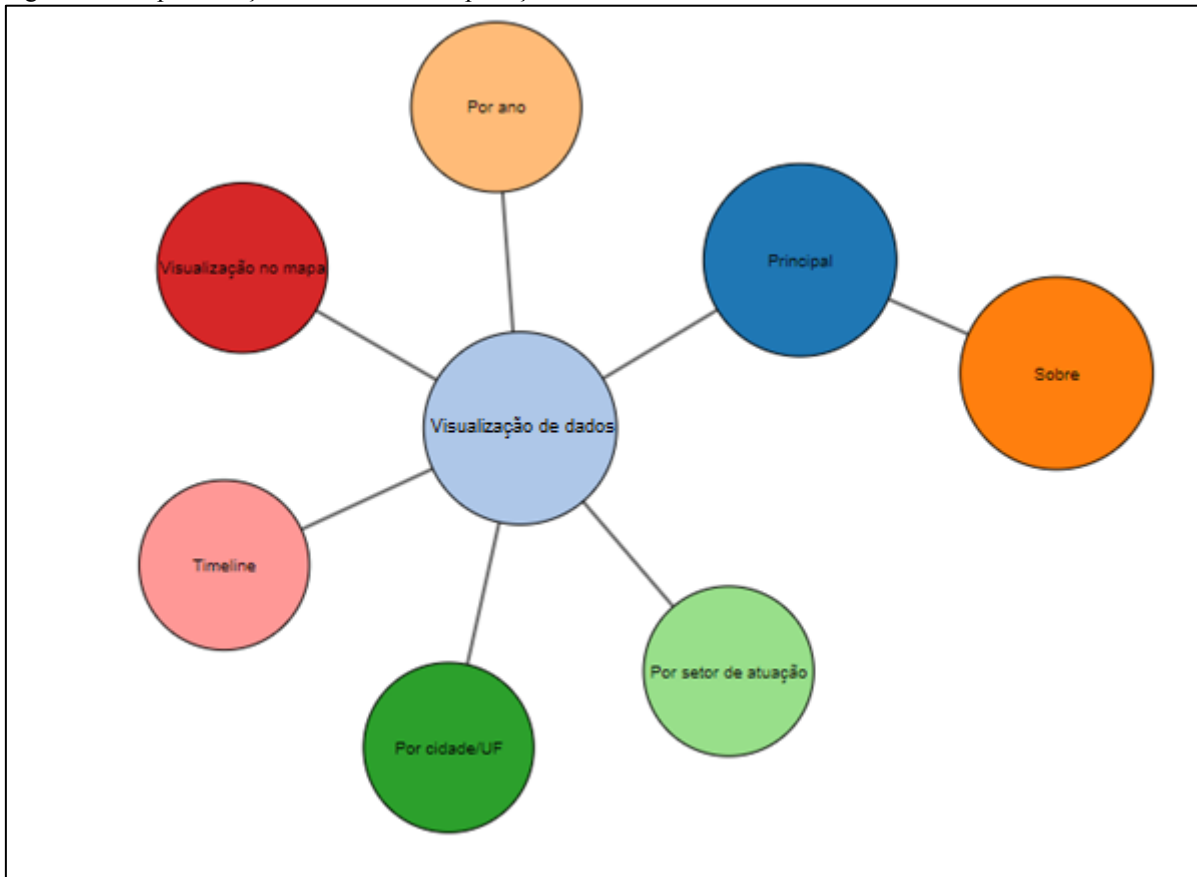
CREATE TABLE dominios (
  id Int(11) PRIMARY KEY,
  dominio Varchar(100),
  nome Varchar(255),
  setor Varchar(100),
  uf Varchar(2),
  cidade Varchar(100),
  cep Varchar(10),
  latitude Varchar(15),
  longitude Varchar(15),
  criado Varchar(20),
  regioao Varchar(20)
)
```

Fonte: Elaborado pelo autor (2017).

## 5.2 Projeto da Interface de Visualização Interativa

A interface de visualização dos dados tem um design simples e objetivo, na qual o usuário pode facilmente visualizar a informações de uma forma interativa. Dentre os gráficos, o usuário pode contar com formas de interação dos mesmos, selecionando por região, por área de atuação ou manuseando o mapa com a base vindo diretamente da API do Google Maps. Para exemplificar, foi feito um grafo com a própria biblioteca D3, como mostra a Figura 10 abaixo.

Figura 10 - Representação da estrutura da aplicação utilizando o D3.

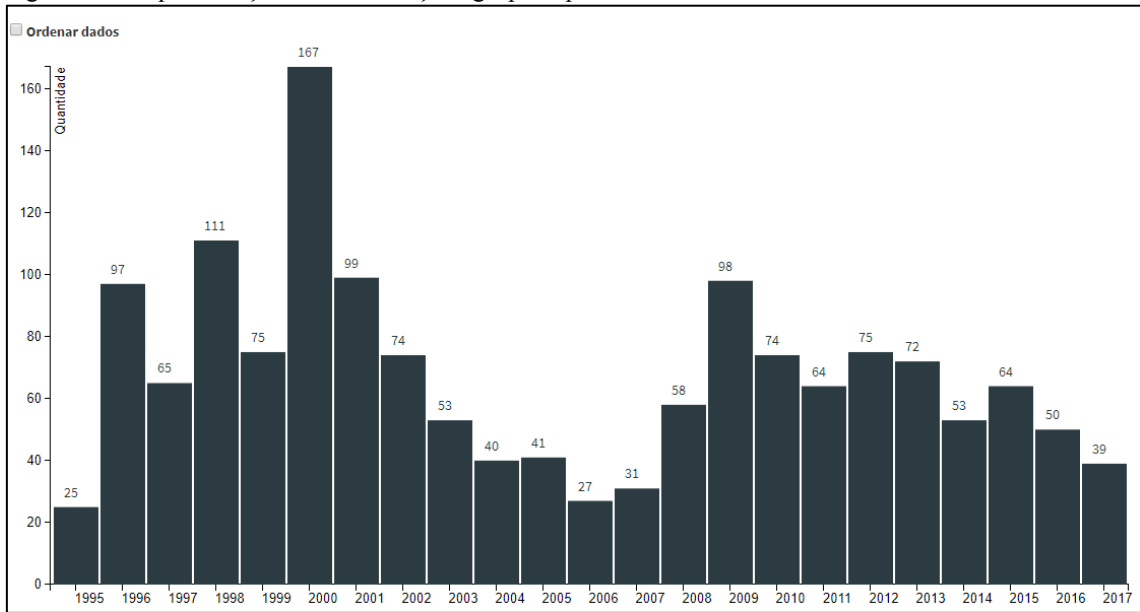


Fonte: Elaborado pelo autor (2017).

### 5.2.1 Visualização por ano

Conforme mostra a Figura 11, a visualização dos dados é obtida pelos domínios registrados e agrupados de acordo com o ano em que foram criados. Neste primeiro gráfico, podemos comparar em quais anos temos mais ou menos registros de domínios do governo. Cada agrupamento de cada ano apresenta também a quantidade de domínios.

Figura 11 - Representação da visualização agrupada por ano de cadastro.



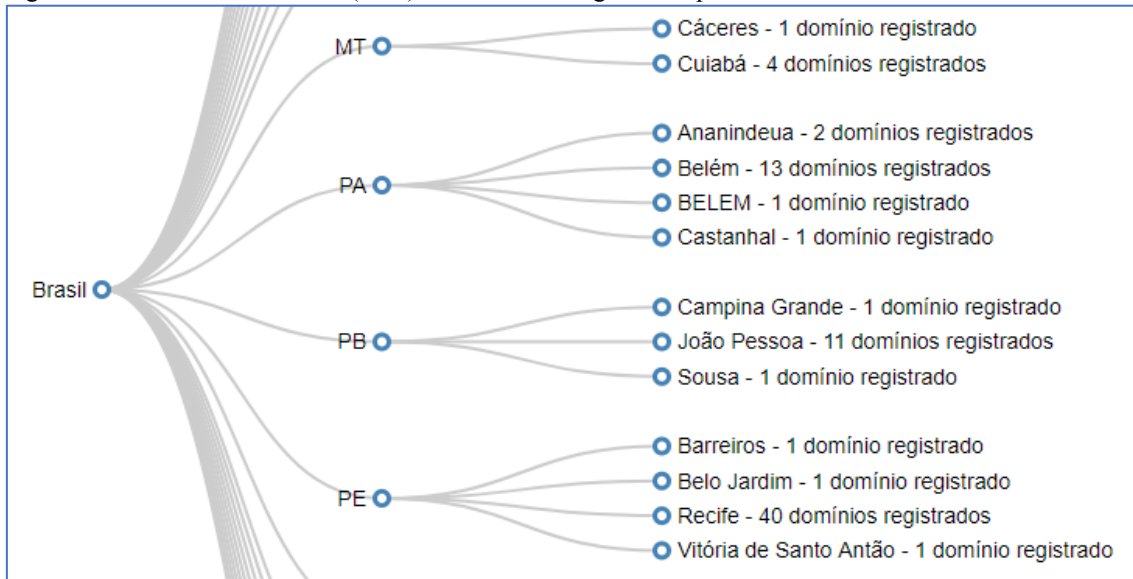
Fonte: VIEW.GOV (2017).

Este gráfico em forma de barras é muito utilizado quando se trata de visualização de dados. Em um primeiro olhar, o usuário pode saber com clareza em qual ano teve mais registros de domínios e em qual ano tivemos menos registros juntamente com a quantidade em cada ano. O usuário também marcar o *checkbox* “Ordenar dados” e, desta forma, obter a ordenação da maior para a menor quantidade de endereços eletrônicos criados por ano.

### 5.2.2 Visualização por cidade/UF

A visualização por cidade/UF agrupa em forma de árvore (*Tree*) três níveis de informação. Brasil que é o topo da árvore, as unidades federativas, e o terceiro nível que são todas as cidades que tem domínio registrado. A consulta apresenta a cidade somente uma vez. Seguindo as linhas, é possível identificar com clareza quais cidades que registraram domínios, ligadas a cada estado. A Figura 12 apresenta o gráfico mencionado. O usuário pode ainda recolher os nós das árvores clicando no círculo.

Figura 12- Gráfico estilo árvore (Tree) dos domínios registrados por cidade/UF.

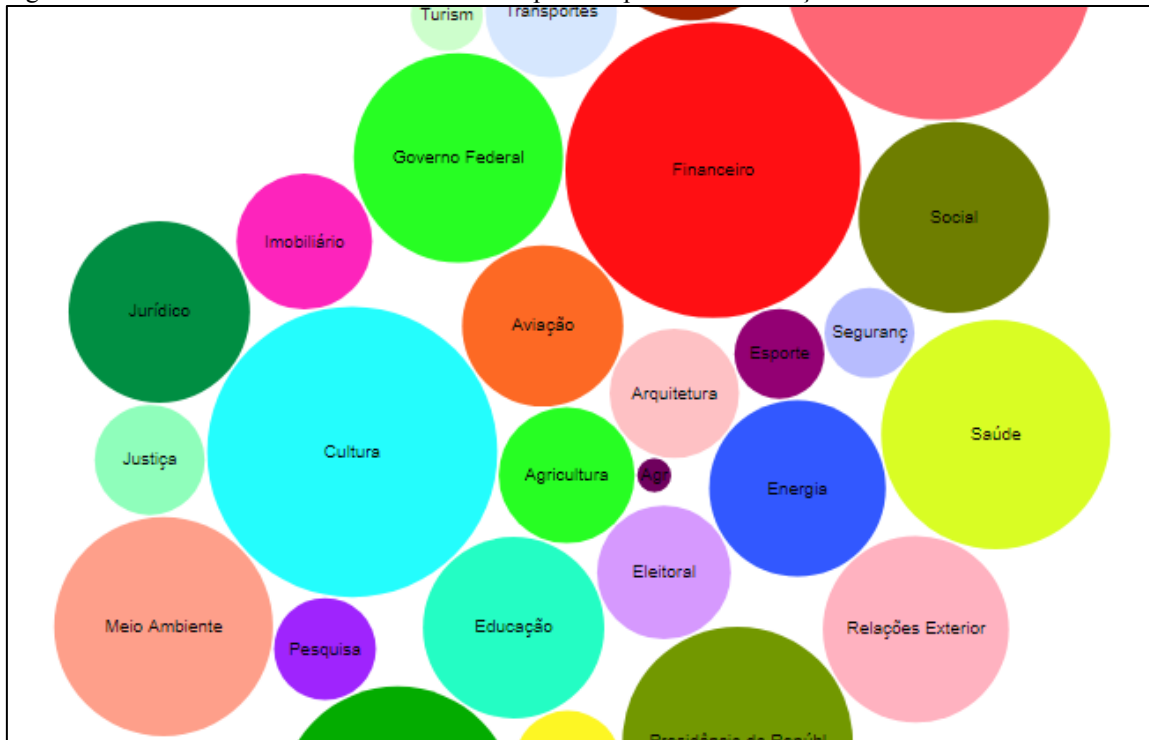


Fonte: VIEW.GOV (2017).

### 5.2.3 Visualização por setores de atuação

Neste gráfico os dados são agrupados por setores de atuação onde se pode observar que quanto maior o tamanho do círculo, mais domínios aquele setor possui. A representação das cores é aleatória. A Figura 13 abaixo mostra o resultado final obtido.

Figura 13- Gráfico Bubble Chart com os dados separados por setor de atuação.

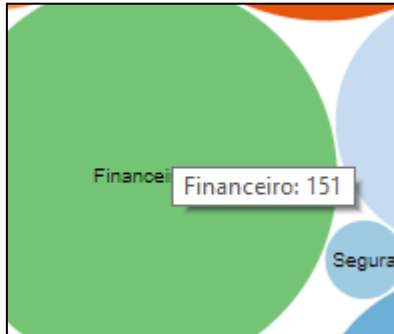


Fonte: VIEW.GOV (2017).



Ainda é possível nesse estilo de gráfico, saber a quantidade de domínios registrados por setor, simplesmente passando o mouse por cima de cada círculo, como exemplifica a Figura 14 mostrada abaixo.

Figura 14 - Exemplo de como visualizar a quantidade de domínios por setor de atuação.

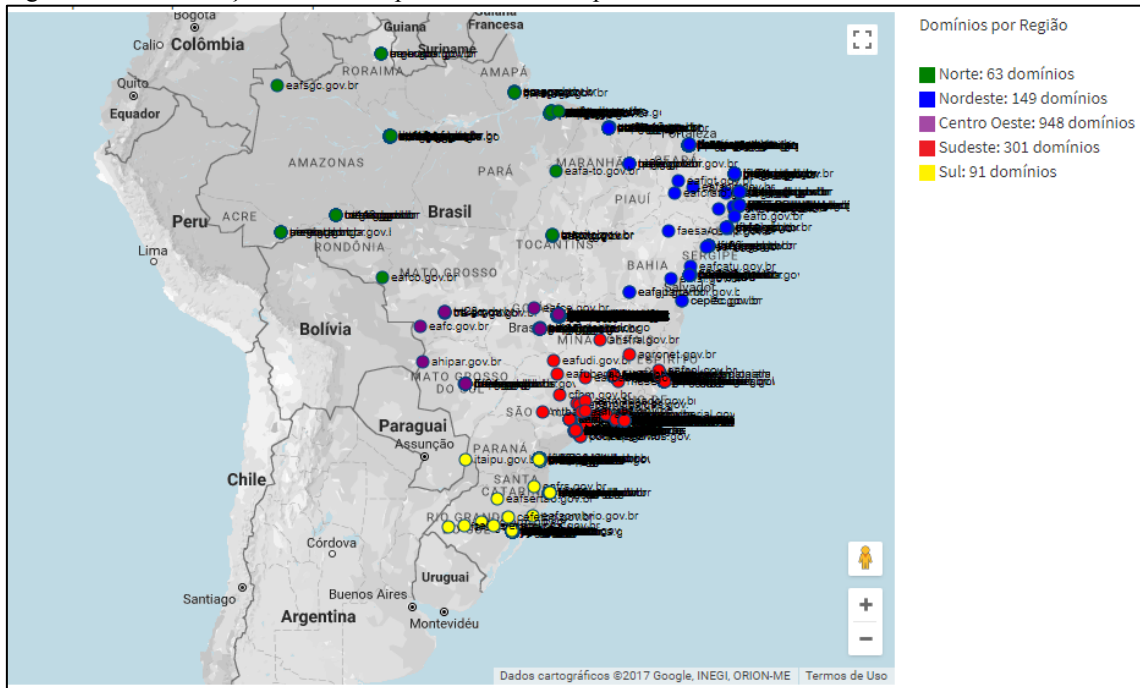


Fonte: VIEW.GOV (2017).

#### 5.2.4 Visualização por localização no mapa do Brasil

A visualização aplicada no mapa o mostra centralizado na latitude e longitude do Brasil e marca, com pequenos pontos, a cidade onde o domínio foi registrado. Esta representação do mapa não é uma imagem fixa, mas sim um mapa real, utilizando a API do Google Maps. O usuário pode interagir com o mapa, quando necessário, aplicando um “zoom”, por exemplo. Os pontos são marcados no mapa conforme sua latitude e longitude. A obtenção dessa informação será descrita na seção de caracterização dos dados. Abaixo, como mostra a Figura 15, podemos observar a disposição dos domínios, dentro do território brasileiro.

Figura 15 - Visualização de dados representados no mapa.



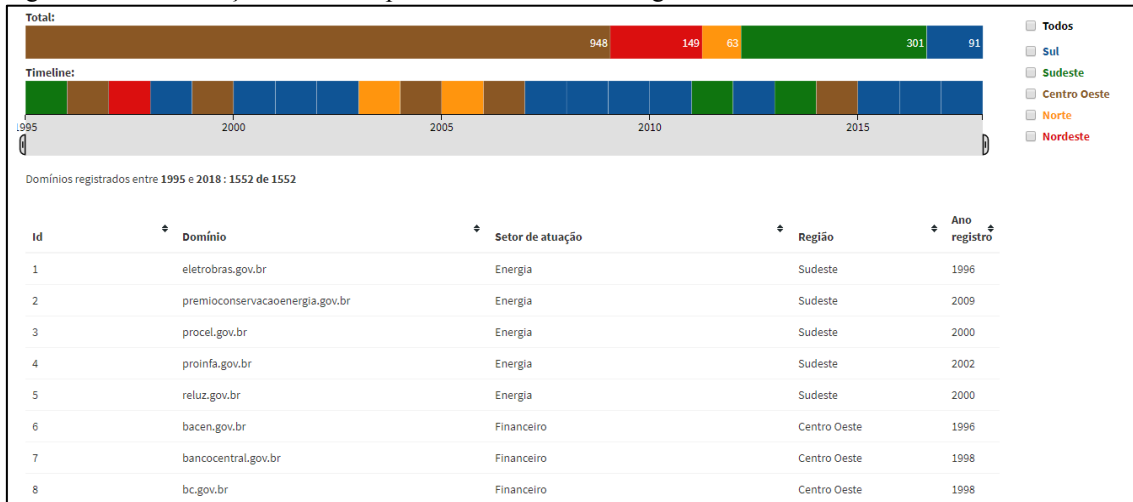
Fonte: VIEW.GOV (2017).

Ainda neste tipo de visualização, o usuário pode conferir no canto superior direito as informações de quantos domínios estão registrados por região. Podemos observar que temos domínios registrados em todas as regiões do Brasil. As diferentes regiões são identificadas por cores distintas.

### 5.2.5 Visualização por linha do tempo do ano de registro

A visualização por linha do tempo ou *timeline* traz os domínios, as regiões e o ano de registro. Essas informações serão usadas para preencher uma linha do tempo que inicia no ano em que o primeiro domínio foi registrado e termina no último ano em que teve um domínio registrado. A Figura 16 exemplifica uma visão geral desse tipo de gráfico.

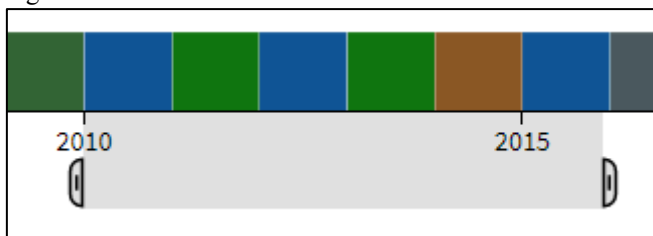
Figura 16 - Visualização dos dados por *timeline* do ano de registro.



Fonte: VIEW.GOV (2017).

Neste tipo de visualização podemos filtrar os domínios registrados por ano dentro de um determinado período. Para isso o usuário pode utilizar controles correspondentes aos anos para determinar o intervalo é mostrado na Figura 17. Quando a visualização é carregada, o período padrão selecionado é de 2010 a 2016.

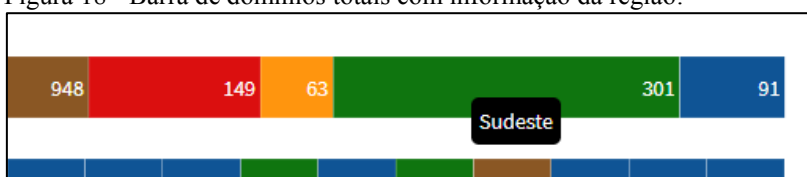
Figura 17 - Seletor de intervalo de anos.



Fonte: VIEW.GOV (2017).

Acima, pode-se visualizar uma barra horizontal com o total de domínios, que se atualiza de acordo com a seleção do intervalo e é separada pelas regiões brasileiras. Cada cor representa uma região e, ao passar o mouse em cima da cor desejada, um alerta *tooltip* é apresentado, identificando a região como exemplifica a Figura 18.

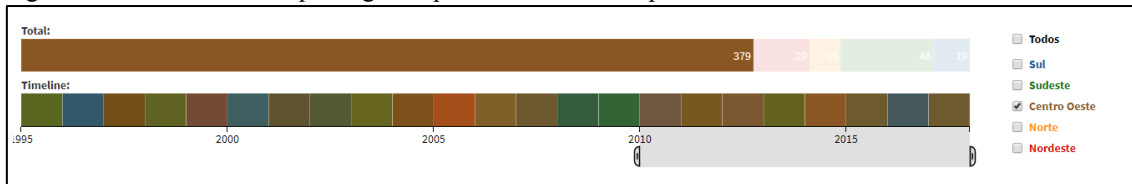
Figura 18 - Barra de domínios totais com informação da região.



Fonte: VIEW.GOV (2017).

No canto superior direito, podemos, ainda, filtrar por região ou selecionar todas as regiões. A seleção de todas as regiões é o padrão utilizado e é carregado quando o gráfico é inicializado. A Figura 19 apresenta o gráfico filtrado pela região Centro Oeste no intervalo de 2010 a 2017.

Figura 19 - Dados filtrados por região e por intervalo de tempo.



Fonte: VIEW.GOV (2017).

Abaixo da visualização gráfica, temos uma tabela com os dados que foram agrupados acima. Essa tabela mostra todos os domínios com o seu setor de atuação, região e ano de registro. Esses dados podem ser pelos critérios apresentados, alternando por ordem crescente ou decrescente, como mostra a Figura 20 a seguir.

Figura 20 - Informações agrupadas em forma de tabela dinâmica com opção de ordenação.

Id	Domínio	Setor de atuação	Região	Ano registro
21	redegeochronos.gov.br	Pesquisa	Sudeste	2006
61	caixacultural.gov.br	Financeiro	Centro Oeste	2005
141	dpu.gov.br	Governo Federal	Centro Oeste	2007
148	pormin.gov.br	Energia	Centro Oeste	2008
152	bnfsp.gov.br	Financeiro	Sudeste	2006
162	confiancabrasil.gov.br	Presidência da República	Centro Oeste	2008
163	confiancanobrasil.gov.br	Presidência da República	Centro Oeste	2008
171	juventude.gov.br	Presidência da República	Centro Oeste	2007
172	maisbrasil.gov.br	Presidência da República	Centro Oeste	2007
174	nae.gov.br	Presidência da República	Centro Oeste	2005
179	portosdobrasil.gov.br	Presidência da República	Centro Oeste	2007
189	splp.gov.br	Presidência da República	Centro Oeste	2007
191	cbnb.gov.br	Aviação	Sudeste	2005
193	brmcc.gov.br	Aviação	Sudeste	2005

Fonte: VIEW.GOV (2017).

## 5.2.6 Informações Complementares

A última seção intitulada “Sobre”, traz uma breve apresentação do autor, um resumo dos objetivos pretendidos com essa aplicação, um pequeno resumo sobre o que é o Laboratório de Ciência e Visualização de Dados (LabData), que foi de grande ajuda com a execução do projeto.

Figura 21 - Página sobre o projeto.

**Sobre o Projeto**

**O Autor**

Trabalho de Conclusão do Curso de Bacharelado em Tecnologias na Informação e Comunicação (TIC) com o nome de VIEW.GOV: Visualização dos dados de endereços eletrônicos do governo do Brasil do acadêmico Filipe Zeferino sob orientação do Professor Robson Rodrigues Lemos, defendido em novembro de 2017 nas dependências da UFSC no campus de Araranguá em Santa Catarina.

**Resumo**

O objetivo final do trabalho é o desenvolvimento de uma aplicação Web para visualização de dados de endereços eletrônicos registrados para o governo do Brasil com uma biblioteca D3 (Data-Driven Documents), e disponibilizar na internet para o livre acesso de todos, facilitando o entendimento do conjunto de dados em um primeiro olhar. Ao término do desenvolvimento da aplicação, comprovou-se que o seu uso, ou o uso de similares, auxilia muito na compreensão de um conjunto de dados grande ou pequeno. O estudo das Tecnologias da Informação e Comunicação (TIC) deu o suporte necessário para o desenvolvimento da aplicação de visualização de dados VIEW.GOV. O projeto e a aplicação foram desenvolvidos com um conjunto de dados do portal brasileiro de dados abertos (<http://dados.gov.br>) que agrupa informações sobre os domínios Gov.br registrados no Registro.Br e autorizados pelo ministério do planejamento. A real clareza da aplicação comprova a eficácia no seu uso. Sem uma ferramenta semelhante, seria muito difícil o usuário tirar alguma conclusão a respeito dos dados.

**LabData**

O Laboratório de Ciência e Visualização de Dados (LabData), do Campus Araranguá da Universidade Federal de Santa Catarina, tem como objetivo a análise e visualização de dados para identificação de padrões, tomada de conhecimento, e tomada de decisões em diversas áreas do conhecimento. Uma boa análise de dados a partir de técnicas de extração e análise de dados e a partir de técnicas de representação visual e análise visual de dados permite um melhor gerenciamento do que está acontecendo atualmente e meios mais adequados de prever o que poderá acontecer sobre condições particulares em um futuro próximo, de tal forma que se possam criar oportunidades e prevenir problemas.

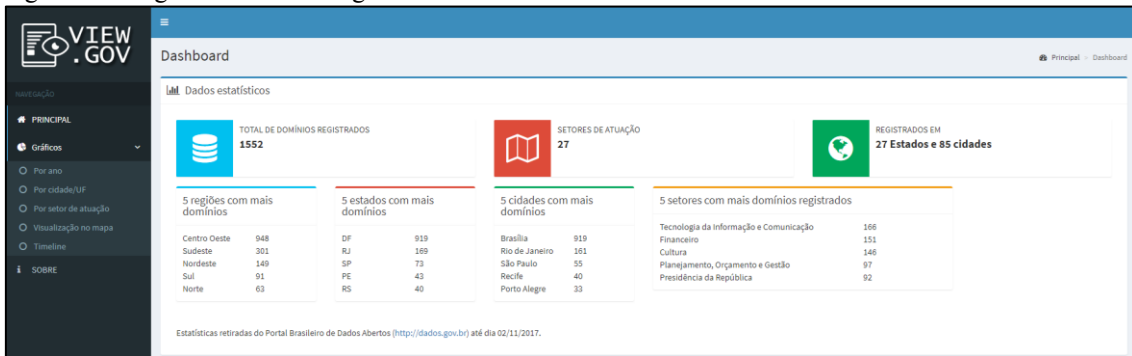
**Orientador**

Dr. Robson Rodrigues Lemos ([robson.lemos@ufsc.br](mailto:robson.lemos@ufsc.br))

Fonte: VIEW.GOV (2017).

A página inicial traz algumas estatísticas relevantes do conjunto de dados escolhido, como mostra a Figura 22.

Figura 22 - Página inicial com algumas estatísticas.



Fonte: VIEW.GOV (2017).

## **6 DESENVOLVIMENTO DA APLICAÇÃO WEB**

O objetivo deste capítulo é detalhar o processo do desenvolvimento da aplicação WEB VIEW.GOV e mostrar como os dados foram extraídos, tratados no banco de dados e como eles são recuperados na aplicação. Também são apresentadas algumas técnicas de obtenção de dados e o uso de ferramentas de modo a complementar o desenvolvimento.

### **6.1 Caracterização dos Dados**

Os dados utilizados para a construção da aplicação foram baixados do Portal Brasileiro de Dados Abertos (Dados Abertos, 2017) em um único arquivo no formato JSON. O portal online oferece vários outros conjuntos de dados de forma aberta para acesso a qualquer usuário.

A escolha de um conjunto de dados relevante serve de base para uma análise mais detalhada e intuitiva para o usuário que irá acessar a aplicação proposta pelo trabalho de conclusão de curso.

Olhando para o conjunto de dados a ser utilizado, pode-se identificar várias informações que se mostram desordenadas e com pouca ou quase nenhuma forma de interação entre o usuário e os dados.

A disposição dos dados segue como base uma planilha simples do Microsoft Excel onde cada linha representa um registro de domínio e cada coluna contém um tipo de informação referente a este domínio.

De posse dos dados, foram acrescentadas mais algumas informações relevantes para o complemento do registro. São elas: O setor de atuação, a região em que o domínio foi criado, a latitude e a longitude da origem do domínio.

## 6.2 Técnicas de Visualização

As técnicas de visualização utilizadas foram escolhidas com base nas informações disponíveis e com o intuito de obter uma melhor forma de representação e exploração da informação. Foram escolhidas cinco técnicas diferentes de visualização para o mesmo conjunto de dados e cada gráfico não usa todos os dados disponíveis, um gráfico complementa o subsequente.

### 1. Técnica de visualização por ano

Para a visualização de dados multivariados organizados por ano do registro do domínio, foi utilizada a técnica de visualização em barras (*Bar Chart em inglês*). Uma técnica simples e de fácil entendimento, dá ao usuário informações precisas em um olhar rápido.

Nesse tipo de gráfico, temos dois eixos principais. Um deles representa a quantidade de domínios registrados e o outro representa cada ano, desde o registro do primeiro domínio até o último domínio. No canto superior direito, se o *checkbox* for acionado, teremos a ordenação dos dados por ordem decrescente.

### 2. Técnica de visualização por cidade/UF

Essa técnica conhecida como hierarquia, agrupa os dados em uma árvore hierárquica horizontal relacionando de forma distinta a cidade ligada ao estado e ligada ao país. A lista das cidades é formada por todas as cidades que tem pelo menos um domínio registrado.

### 3. Técnica de visualização por setor de atuação

A técnica usada para computar os domínios por setor de atuação é a de visualização de texto, que agrupa o total de domínios registrados, em forma de bolhas ou *Bubble Chart em inglês*. Os setores de atuação do domínio registrado são diferenciados por cores distintas e o seu tamanho é determinado pela quantidade de domínios registrados.

### 4. Técnica de visualização por localização no mapa do Brasil

A técnica de visualização aplicada no mapa usa a API do Google Maps para dispor a visualização de dados multivariados e o mapa base. Quando é carregado, o mapa é centralizado de forma que a visualização apresente todo o território brasileiro. Os pontos são marcados no mapa, de acordo com a sua latitude e longitude obtida. Eles ainda são identificados por cores diferentes, de acordo com a sua região. Ao lado, uma legenda de cores para cada região é exposta, junto com a quantidade de domínios por região.

## 5. Técnica de visualização estilo por linha do tempo do ano de registro

Essa técnica de visualização de dados multivariados, no estilo *timeline* ou visualização por linha do tempo do ano de registro propicia ao usuário selecionar um intervalo de datas para a visualização dos dados. Uma barra identifica o total de domínios registrados, de acordo com o intervalo escolhido. Os dados ainda podem ser filtrados por região, utilizando o menu, no canto superior direito, como mostrou a Figura 19.

## 6.3 Configurações do servidor

Para o desenvolvimento da aplicação, foi utilizada a linguagem de programação PHP que é a linguagem responsável por fazer a consulta no banco de dados, fazer o devido processamento destes dados e entregar para o *front-end* as informações aninhadas e prontas para produzir a visualização escolhida. O servidor do LabData está localizado nas dependências da UFSC (Universidade Federal de Santa Catarina). O sistema operacional utilizado neste servidor é o ubuntu 4.22, a versão do PHP é a 5.5.9-1, com Apache 2.4.7.

### 6.3.1 Carga e tratamento dos dados

Para carregar os dados para o servidor de banco de dados MySQL, primeiramente foi necessário baixar o arquivo em formato JSON para o computador local e enviar o mesmo para o servidor. Junto com o arquivo do conjunto de dados ainda não tratado, foi desenvolvido um script em PHP que carrega esse arquivo JSON e extrai os seus dados inserindo diretamente no banco de dados. A Figura 23 abaixo mostra o script.

Figura 23 - Script para importar as informações para o banco de dados MySQL.

```
<?php
$jsonurl = "dominios.json";
$json = file_get_contents($jsonurl,0,null,null);
$json = json_decode($json, true);
$total = $json['result']['total'];
include('conecta.php');

for($i = 0; $i < $total; $i++){
    $dominio = $json['result']['records'][$i]['dominio'];
    $nome = $json['result']['records'][$i]['nome'];
    $uf = $json['result']['records'][$i]['uf'];
    $cidade = $json['result']['records'][$i]['cidade'];
    $cep = $json['result']['records'][$i]['cep'];
    $criado = substr($json['result']['records'][$i]['data_cadastro'], 0, -15);
    $result = mysql_query("INSERT INTO dominios (dominio, nome, uf, cidade, cep, criado)
        VALUES ('{$dominio}', '{$nome}', '{$uf}', '{$cidade}', '{$cep}', '{$criado}')");
}
echo "Dados de dominios.json importados para o banco com sucesso.";
?>
```

Fonte: Elaborado pelo autor (2017).



No arquivo JSON, o campo da data de cadastro do domínio apresenta a data exata com dia, mês e ano. Para a nossa aplicação, não foi usado o dia e o mês, então, neste script, antes de persistir a informação no banco de dados, é realizado um tratamento para deixar somente o ano.

O script de conexão chamado “conecta.php”, como mostra a Figura 24 abaixo, faz a conexão da aplicação com o banco de dados e deve ser incluído em todas as páginas que necessitam de alguma interação com o banco, como consultas, atualizações ou remoção de informação.

Figura 24 - Conexão com o banco de dados MySQL.

```
1  <?
2  if (!$link = mysql_connect('host', 'usuario', 'senha')) {
3      echo 'Não foi possível conectar ao mysql';
4      exit;
5  }
6
7  if (!mysql_select_db('banco', $link)) {
8      echo 'Não foi possível selecionar o banco de dados';
9      exit;
10 }
11 ?>
```

Fonte: Elaborado pelo Autor (2017).

Como informação adicional para popular a nossa base de dados, foram incluídas a latitude e a longitude. Esses dois parâmetros são essenciais quando precisamos apresentar informações baseadas em mapas. Para conseguir essas informações com facilidade, foi desenvolvido um script em PHP, que automatiza a obtenção delas quando executado. A Figura 25 mostra o script completo abaixo.

Figura 25 - Script para extrair a latitude e longitude de cada domínio no banco de dados.

```

lat_long.php
1 <?
2 include('../json/conecta.php');
3
4 $result = mysql_query("SELECT id, cep, from dominios");
5
6 while ($row = mysql_fetch_array($result)) {
7     $cep = $row['cep'];
8     $id = $row['id'];
9
10    $maps = file_get_contents("https://maps.googleapis.com/maps/api/geocode/
        json?key=AIzaSyBRLLc9KgVupv1Lqi0Pj1U886B_gj_wg9A&components=postal_code:". $cep."&sensor=false
        ");
11
12    $result_json = json_decode($maps, true);
13    if($result_json[status] != "ZERO_RESULTS"){
14        $lat = $result_json[results][0][geometry][location][lat];
15        $long = $result_json[results][0][geometry][location][lng];
16
17        $execmysql = mysql_query("UPDATE dominios SET latitude = '$lat', longitude = '$long' WHERE
            id = '$id'");
18    }
19
20 }
21 echo "Execução finalizada.";

```

Fonte: Elaborado pelo autor (2017).

O script faz uma consulta no banco de dados e resgata o id e o CEP dos domínios. Com essa informação, ele faz uma consulta na API do Google Maps, junto com o CEP desejado e recebe de retorno um arquivo JSON com uma série de dados. Um destes dados é latitude e o outro é a longitude. Com elas, é feito uma atualização daquele dado com o ID resgatado, adicionando as para aquele domínio.

Para que a aplicação no *front-end* possa carregados os dados, esse novo conjunto com as novas informações deve ser passados como um arquivo JSON. Para esse objetivo, foi criado mais um script para gerar o JSON tratado para a aplicação.

Um exemplo usado para entregar os dados para a visualização por setor pode ser visto abaixo, na Figura 26.

Figura 26 - Script em PHP para gerar uma saída em JSON.

```

1  <?php
2
3  include('conecta.php');
4
5  $query = "SELECT setor, COUNT(*) as total FROM dominios group by setor";
6  $exec = mysql_query($query);
7
8  if ( ! $exec ) {
9      echo mysql_error();
10     die;
11 }
12
13 mysql_close($link);
14
15 $dados = array();
16 while ($rw = mysql_fetch_assoc($exec)) {
17     if($rw["setor"] == ""){$rw["setor"] = "Sem Categoria";}
18     array_push($dados, array("name" => $rw["setor"], "size" => $rw["total"]));
19 }
20
21 $areas = array(
22     "name" => "Setores",
23     'children' => $dados
24 );
25
26 header('Content-Type: application/json');
27 echo json_encode($areas, JSON_PRETTY_PRINT);
28
29 ?>

```

Fonte: Elaborado pelo autor (2017).

#### 6.4 Funcionalidades do *front-end*

Para o desenvolvimento do *front-end*, que é onde o realmente o usuário vai interagir, foi utilizado o “*template*” AdminLTE, descrito na seção 3.2.8 deste trabalho. Além do *framework* Bootstrap, o *template* usa seu próprio CSS para estilizar as páginas, conforme Figura 27.

Figura 27 - Inclusão dos arquivos CSS.

```

1 <!DOCTYPE html>
2 <html>
3 <head>
4 <meta charset="utf-8">
5 <meta http-equiv="X-UA-Compatible" content="IE=edge">
6 <title>VIEW.GOV: Visualização de dados de endereços eletrônicos do governo com a biblioteca D3</title>
7 <!-- Tell the browser to be responsive to screen width -->
8 <meta content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no" name="viewport">
9 <!-- Bootstrap 3.3.7 -->
10 <link rel="stylesheet" href="bower_components/bootstrap/dist/css/bootstrap.min.css">
11 <!-- Font Awesome -->
12 <link rel="stylesheet" href="bower_components/font-awesome/css/font-awesome.min.css">
13 <!-- Ionicons -->
14 <link rel="stylesheet" href="bower_components/ionicons/css/ionicons.min.css">
15 <!-- Theme style -->
16 <link rel="stylesheet" href="dist/css/AdminLTE.min.css">
17
18 <link rel="stylesheet" href="dist/css/skins/_all-skins.min.css">
19
20 <link rel="icon" href="dist/img/favicon.ico">
21 <!-- Google Font -->
22 <link rel="stylesheet" href="dist/css/
css.css?family=Source+Sans+Pro:300,400,600,700,300italic,400italic,600italic">
23 </head>

```

Fonte: Elaborado pelo autor (2017).

Para complementar, a execução dos gráficos da biblioteca D3 e outros efeitos dependem do *framework jQuery*, que são incluídos no fim da página junto com a biblioteca D3, como apresentado na Figura 28 abaixo. A inclusão desses arquivos em CSS e JS deverá ser feito em todas as páginas.

Figura 28 - Inclusão dos scripts em JavaScript no fim das páginas.

```

<!-- jQuery 3 -->
<script src="bower_components/jquery/dist/jquery.min.js"></script>
<!-- jQuery UI 1.11.4 -->
<script src="bower_components/jquery-ui/jquery-ui.min.js"></script>
<!-- Bootstrap 3.3.7 -->
<script src="bower_components/bootstrap/dist/js/bootstrap.min.js"></script>
<!-- AdminLTE App -->
<script src="dist/js/adminlte.min.js"></script>
<script src="http://d3js.org/d3.v3.min.js"></script>

```

Fonte: Elaborado pelo autor (2017).

Com a biblioteca D3 carregada, podemos criar os gráficos e seus atributos e métodos. Na página onde serão exibidas as visualizações, precisamos pegar os dados vindos do banco de dados dispostos em forma de arquivo JSON e carregá-los para a aplicação via D3 (Figura 29).

Figura 29 - Carregamento dos dados para o D3.

```
d3.json("json/json_mapa.php", function(error, data) {
  if (error) throw error;
```

Fonte: Elaborado pelo Autor (2017).

Com os dados carregados, é iniciado o script para criação da visualização escolhida. Para exemplificar como o D3 apresenta uma visualização aplicada no mapa, algumas linhas de código são necessárias. A Figura 30 apresenta o trecho do código onde o mapa é criado e alguns atributos são previamente configurados.

Figura 30 - Trecho do código que inicializa o mapa.

```
var map = new google.maps.Map(d3.select("#map").node(), {
  zoom: 4,
  center: new google.maps.LatLng(-17.9364178, -54.9087281),
  styles: [{stylers:[{saturation:-100},{gamma:1}],{elementType:"labels.text.stroke",stylers:[{visibility:"off"}]},
  {featureType:"poi.business",elementType:"labels.text",stylers:[{visibility:"off"}]},
  {featureType:"poi.business",elementType:"labels.icon",stylers:[{visibility:"off"}]},
  {featureType:"poi.place_of_worship",elementType:"labels.text",stylers:[{visibility:"off"}]},
  {featureType:"poi.place_of_worship",elementType:"labels.icon",stylers:[{visibility:"off"}]},
  {featureType:"road",elementType:"geometry",stylers:[{visibility:"simplified"}]},
  {featureType:"water",stylers:[{visibility:"on"},{saturation:40},{gamma:0},{hue:"#00aeef"}]},
  {featureType:"administrative.neighborhood",elementType:"labels.text.fill",stylers:[{color:"#333333"}]},
  {featureType:"road.local",elementType:"labels.text",stylers:[{weight:0.5},{color:"#333333"}]},
  {featureType:"transit.station",elementType:"labels.icon",stylers:[{gamma:1},{saturation:50}]},
  mapTypeId: google.maps.MapTypeId.ROADMAP
});
```

Fonte: Elaborado pelo autor (2017).

Outra informação importante, nesse tipo de visualização, são os pontos, mais conhecidos como marcadores. Cada ponto no mapa identifica um domínio registrado na respectiva cidade. O código apresentado na Figura 31 mostra que é criada um *overlay* (camada, em tradução livre), criando uma variável chamada “*marker*”, que será o ponto e o círculo a ser desenhado no mapa. Na chamada para o desenho do círculo, existe, ainda, uma função “*style*”, que é responsável por colocar a cor no círculo, de acordo com a sua região. O resultado de um marcador é apresentado na Figura 32.

Figura 31 - Trecho do código em que são criados os marcadores e os círculos com o D3.

```

overlay.draw = function() {
  var projection = this.getProjection(),
      padding = 10;

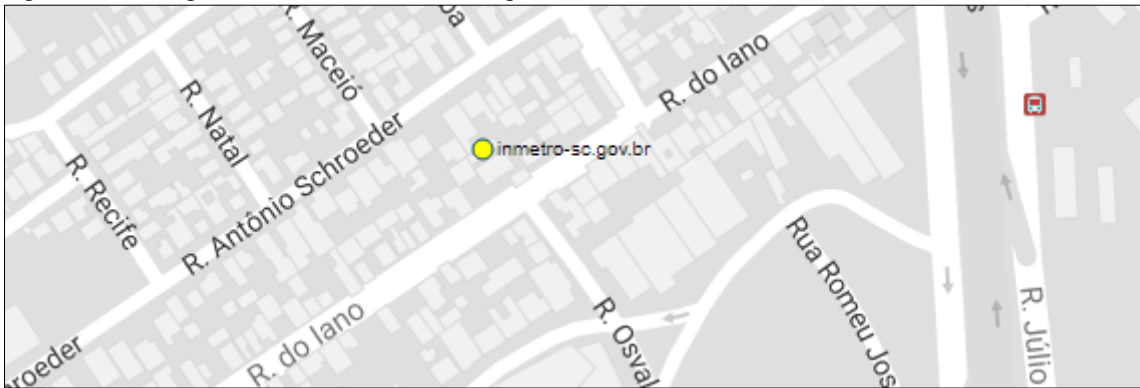
  var marker = layer.selectAll("svg")
    .data(d3.entries(data))
    .each(transform) // update existing markers
    .enter().append("svg")
    .each(transform)
    .attr("class", "marker");

  marker.append("circle")
    .attr("r", 5)
    .attr("cx", padding)
    .attr("cy", padding)
    .style("fill", function(d) {
      if (d.value[2] == "Norte") {return "green"}
      if (d.value[2] == "Nordeste") {return "blue"}
      if (d.value[2] == "Centro Oeste") {return "purple"}
      if (d.value[2] == "Sudeste") {return "red"}
      if (d.value[2] == "Sul") {return "yellow"}
    });
}

```

Fonte: Elaborado pelo autor (2017).

Figura 32 - Exemplo de marcador criado no mapa.



Fonte: Elaborado pelo autor (2017).

Os parâmetros requeridos para completar esse tipo de visualização podem ser vistos na Figura 33 e o código para a criação destes resultados, no formato JSON, é apresentado na Figura 34.

Figura 33 - Exemplo de valores esperados pelo D3 para iniciar a visualização dos dados.

```
{
  "eletrobras.gov.br": [
    "-22.9017435",
    "-43.1798759",
    "Sudeste"
  ],
  "premioconservacaoenergia.gov.br": [
    "-22.9017435",
    "-43.1798759",
    "Sudeste"
  ],
  "procel.gov.br": [
    "-22.9017435",
    "-43.1798759",
    "Sudeste"
  ],
  ],
}
```

Fonte: Elaborado pelo autor (2017).

Figura 34 - Código em PHP para entregar as informações requeridas do banco de dados em um JSON válido.

```
include('conecta.php');

$result = "SELECT dominio, latitude, longitude, setor from dominios";
$exec = mysql_query($result);

if ( ! $exec ) {
    echo mysql_error();
    die;
}

$conteudo = array();

while ($row = mysql_fetch_array($exec)) {
    $lat = $row['latitude'];
    $long = $row['longitude'];
    $dominio = $row['dominio'];

    $monta = array($lat, $long);
    $conteudo[$dominio] = $monta;
}

header('Content-Type: application/json');

echo json_encode($conteudo, JSON_PRETTY_PRINT);

mysql_close($link);
```

Fonte: Elaborado pelo autor (2017).

Com a apresentação do *back-end* e do *front-end* da aplicação, seguimos para os resultados obtidos com este trabalho de conclusão de curso.

## 7 RESULTADOS

Toda aplicação WEB necessita estar hospedada em um servidor, para que possa estar disponível para os usuários. Primeiramente, fez-se o uso de um servidor local, instalado no computador do autor e configurado com a ferramenta XAMPP. Posteriormente, foi solicitado ao orientador deste trabalho um espaço no servidor do LabData, para melhor performance e disponibilidade da aplicação. O orientador fez o pedido ao SETIC (Superintendência de Governança Eletrônica e Tecnologia da Informação e Comunicação), o qual foi prontamente aceito. A aplicação VIEW.GOV\* é acessível para qualquer usuário, dentro ou fora da UFSC.

Comparando a ferramenta VIEW.GOV com a visualização atual, no Portal de Dados Abertos do Governo Federal, a ferramenta permite uma visualização e exploração dos dados, de uma forma mais intuitiva. A visualização com ferramenta auxilia o usuário para uma compreensão mais rápida dos dados em questão e agrega informações adicionais, baseadas no conjunto baixado.

Com os resultados obtidos nas visualizações fornecidas pela aplicação, podemos analisar que:

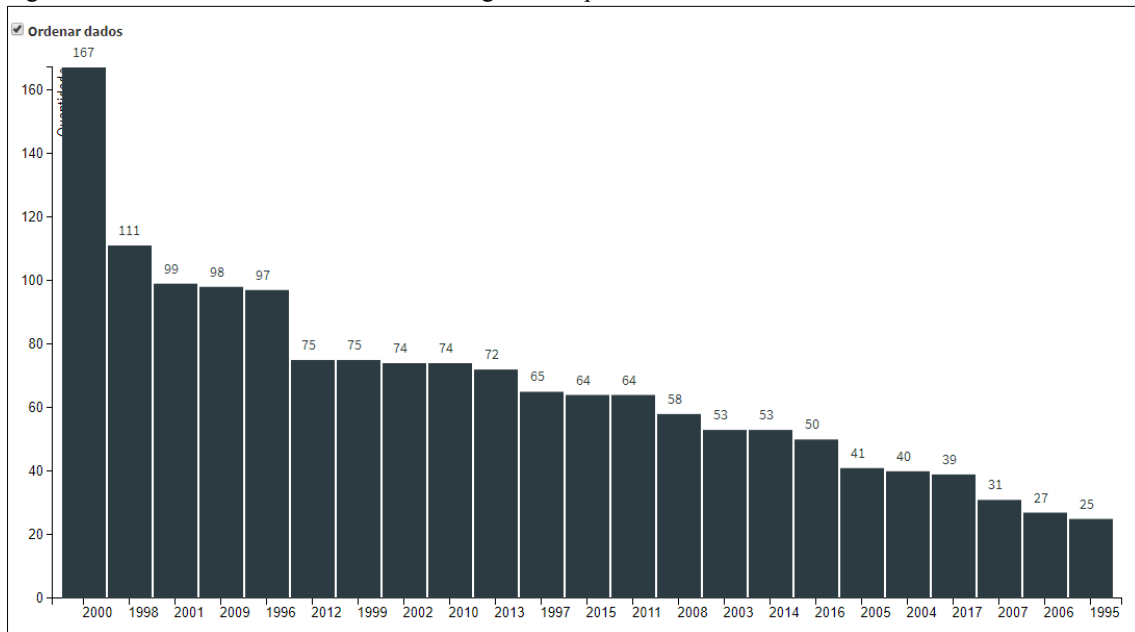
O ano 2000 foi o ano em que teve mais registros de domínios do Governo Federal, com 167 ocorrências, ao passo que 1995 foi o que teve menos registros, com apenas 25, como mostra a Figura 35.

---

\* <http://labdata.sites.ufsc.br/viewgov>



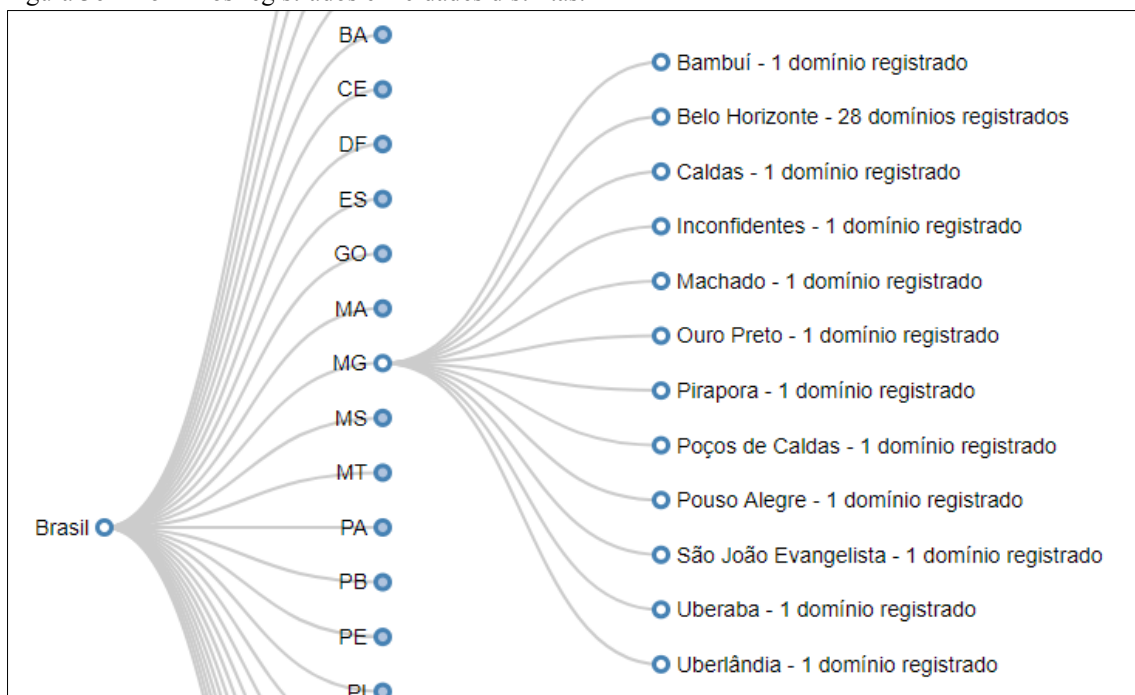
Figura 35 - Domínios do Governo Federal registrados por ano.



Fonte: Elaborado pelo autor (2017).

Pode-se observar, também, que Minas Gerais é o estado com mais cidades distintas a registrar domínios, sendo doze cidades no total. Todos os estados têm pelo menos um domínio registrado em alguma cidade (Figura 36).

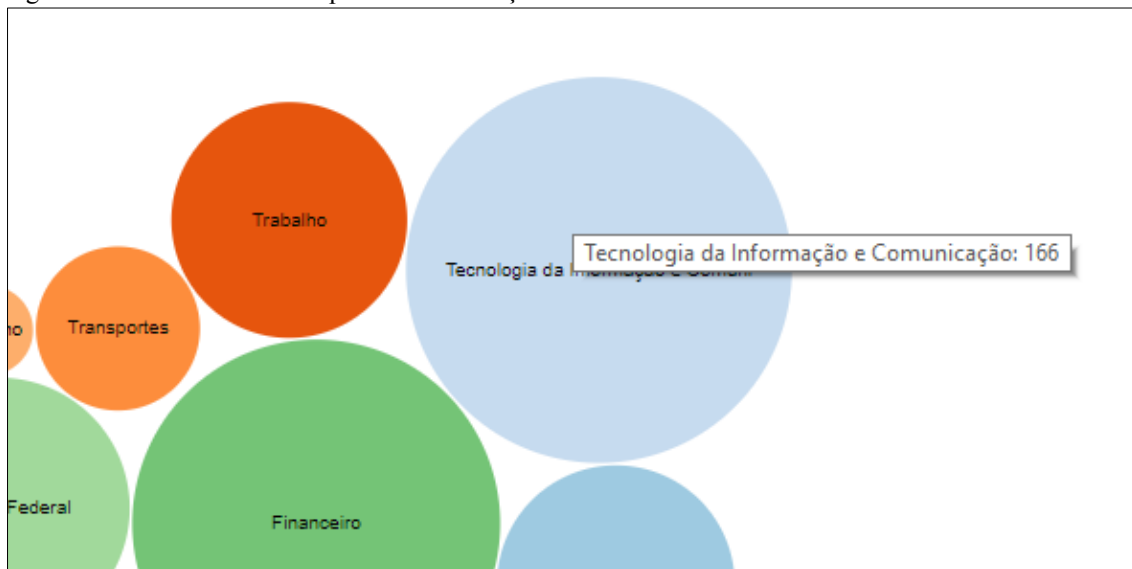
Figura 36 - Domínios registrados em cidades distintas.



Fonte: Elaborado pelo autor (2017).

Na visualização por setor de atuação, é possível analisar que, o setor que mais registrou domínios no Governo Federal foi o setor de Tecnologia da Informação e Comunicação. Obtemos essa informação a partir da visualização dos dados e podemos perceber que este setor tem o círculo maior que os outros. Além disso, através da movimentação do mouse sob o círculo, conseguimos obter a informação a respeito da quantidade de domínios (Figura 37).

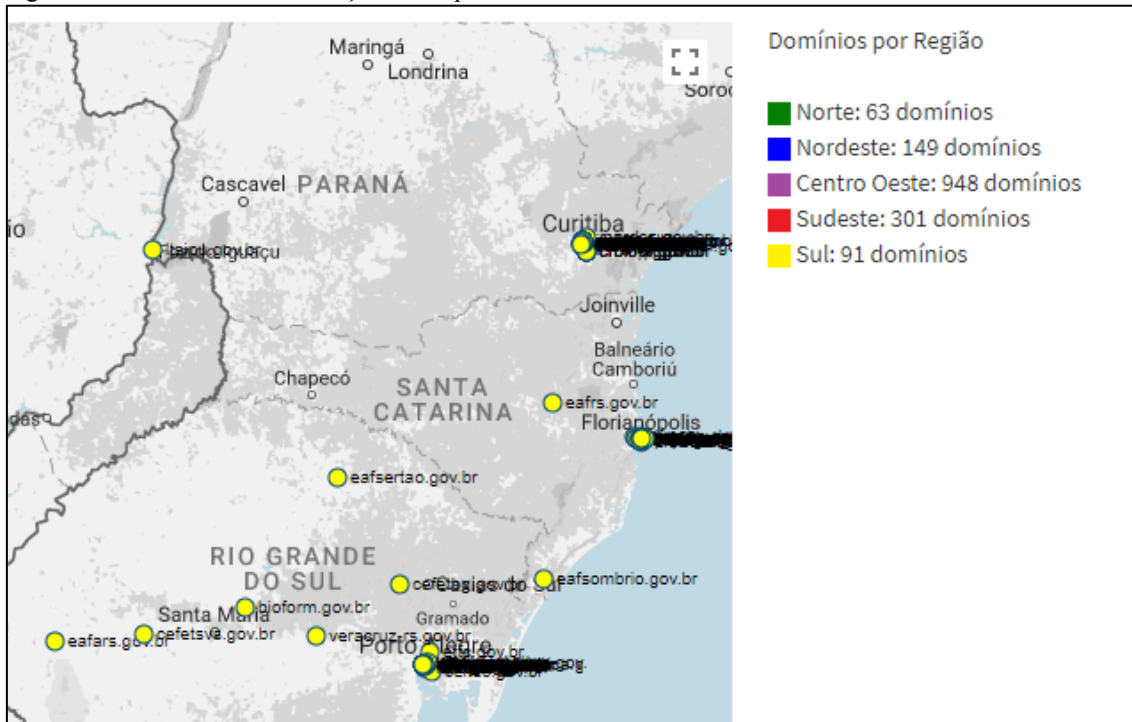
Figura 37 - Total de domínios por setor de atuação.



Fonte: Elaborado pelo autor (2017).

A visualização no mapa permite identificar em quais cidades os domínios estão registrados. Como mostra a Figura 38, o mapa ainda diferencia os marcadores, de acordo com as regiões brasileiras e, usando o recurso de “zoom” do próprio mapa, podemos observar exatamente o local de onde partiu o registro do domínio.

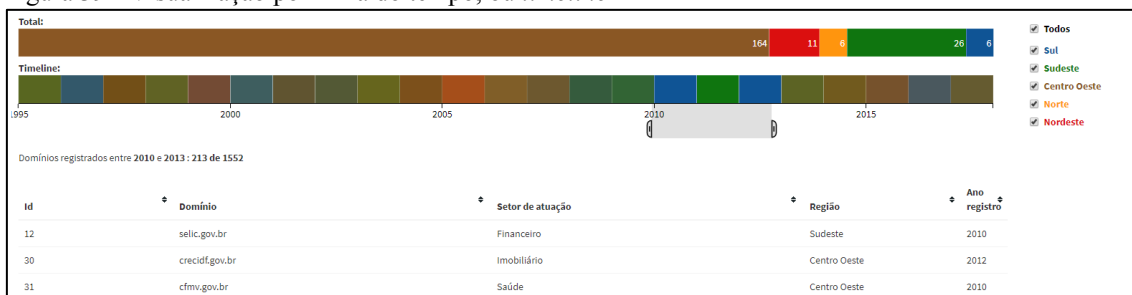
Figura 38 - Detalhe de visualização no mapa usando o recurso de “zoom”.



Fonte: Elaborado pelo autor (2017).

Quando é obtida a informação com uma visualização dos dados em forma temporal ou do tipo *timeline*, utilizando um intervalo de anos de 2010 a 2013 por exemplo, conseguimos observar que temos um total de 213 domínios registrados para esse período em todas as regiões brasileiras. Dispostos em uma tabela abaixo da visualização, estão os dados domínios cobertos pelo intervalo de tempo selecionado. Os dados incluem: um número identificador único, o domínio, o setor de atuação, a região e o ano em que foi registrado. A Figura 39 apresenta o exemplo citado acima.

Figura 39 - Visualização por linha do tempo, ou *timeline*



Fonte: Elaborado pelo autor (2017).

O tráfego vindo de *smartphones* e *tablets* já é muito grande hoje em dia e vem crescendo exponencialmente. A aplicação VIEW.GOV se adapta a qualquer dispositivo e proporciona ao usuário mais conforto para o acesso pela plataforma que desejar.

Figura 40 – Visualização do VIEW.GOV a partir de um Smartphone.



Fonte: Elaborado pelo autor (2017).

## 8 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Com o advento do *big-data* na área de Tecnologia da Informação e Comunicação, cada vez mais surge a necessidade de se ter ferramentas robustas e intuitivas para o processamento e a visualização de dados.

A trajetória do curso e a vivência acadêmica foram de suma importância para desenvolvimento da aplicação de visualização de dados intitulada VIEW.GOV.

O conjunto de dados original foi apresentado em forma de uma lista bruta sem algum tipo de processamento ou interação. A eficiência da aplicação é comprovada quando comparada com a lista, podemos visualizar os dados de uma forma intuitiva e mais objetiva.

Outro ponto importante no desenvolvimento do projeto e no decorrer do trabalho de conclusão de curso, é que a aplicação é ferramenta que estará sempre disponível na WEB para qualquer usuário e pode ser acessada de forma facilitada e por ser responsivo, pode ser acessado por qualquer dispositivo.

Os dados de domínios do Governo Federal que foram carregados na base de dados para serem usados na aplicação são datados até o dia 07/11/2017. Porém, da forma em que a aplicação foi desenvolvida, possibilita a um usuário intermediário com alguns conhecimentos em bancos de dados em MySQL e na linguagem PHP, carregar um novo arquivo com novos dados no futuro, podendo assim sempre manter as informações atualizadas.

O *framework* D3 tem uma comunidade ativa que todos os dias disponibilizam novas formas de visualização explorando diferentes tipos de algoritmos de visualização de dados, além de já contar com milhares de técnicas prontas e livres para o uso.

Pode-se afirmar que os objetivos de se obter uma forma mais intuitiva na visualização de dados foram alcançados e futuramente, novos dados podem ser carregados a fim de manter uma ferramenta sempre atualizada. Os códigos-fonte estão bem comentados facilitando a compreensão para uma futura manutenção e implementação de novas formas de visualização de dados.

## 9 REFERÊNCIAS BIBLIOGRÁFICAS

ALMSAEED, Abdullah. AdminLTE. 2017. Disponível em:

<<https://github.com/almasaeed2010/AdminLTE>>. Acesso em: 28 out. 2017.

AMARAL, Fernando. Introdução à Ciência de Dados: mineração de dados e big data. Rio de Janeiro: Alta Books, 2016. 320 p.

BOSTOCK, Mike. Force-Directed Graph. 2017. Disponível em:

<<https://bl.ocks.org/mbostock/4062045>>. Acesso em: 28 out. 2017.

CARRIL, Marly. HTML: Passo a passo .1ª ed. Joinville: Editora Clube de Autores, 2012. 80p.

CONVERSE, Tim; PARK, Joyce. PHP: A bíblia. 2. ed. São Paulo: Editora Campus, 2003. 904 p. 604 v.

D3 - Data-Driven Documents. 2017. Disponível em: <<https://d3js.org>>. Acesso em: 28 out. 2017.

GIL, Antonio Carlos. Como elaborar projetos de pesquisa. 5. ed. São Paulo: Atlas, 2008.

JOBSTRAIBIZER, Flávia. Criação de sites com o CSS. São Paulo: Universo dos Livros Editora Ltda, 2009. 144 p.

Marr, D. (1982). Vision - A Computational Investigation into the Human Representation and Processing of Visual Information. W. H. Freeman and Company.

MARQUESONE, Rosangela. Big Data: Técnicas e tecnologias para extração de valor dos dados. São Paulo: Casa do Código, 2016. 235 p

MATTHES, Eric. Curso Intensivo de Python: Uma introdução prática e baseada em projetos à programação. São Paulo: Novatec Editora Ltda, 2016. 656 p.

OSHIRO, Adriane Kaori. IVIEW: infra-estrutura de visualização de informação evolucionária na web. 2003. 91 f. Dissertação (Mestrado) - Curso de Ciências de Computação, Instituto de Ciências Matemáticas e de Computação, São Carlos, 2003. Cap. 1. Disponível em:

<[http://www.teses.usp.br/teses/disponiveis/55/55134/tde-21082015-095444/publico/AdrianeKaoriOshiro\\_ME.pdf](http://www.teses.usp.br/teses/disponiveis/55/55134/tde-21082015-095444/publico/AdrianeKaoriOshiro_ME.pdf)>. Acesso em: 27 out. 2017.

PARTNERS, Novus. NVD3: Re-usable charts for d3.js. 2017. Disponível em: <<http://nvd3.org>>. Acesso em: 28 out. 2017.

PORTAL BRASILEIRO DE DADOS ABERTOS. 2017. Disponível em: <<http://dados.gov.br>>. Acesso em: 31 out. 2017.

QUIERELLI, Davi Antonio. Criando Sites com HTML, CSS, PHP. Leme: Clube de Autores, 2012. 91 p.

RÊGO, Bergson Lopes. Gestão e Governança de Dados: Promovendo dados como ativo de valor nas empresas. Rio de Janeiro: Brasport, 2013. 312 p.

RUTTER, Jake. Smashing jQuery: Interatividade Avançada Com JavaScript Simples. Porto Alegre: Brookman, 2012. 321 p. Tradução de Igor Vianna.

SMITH, Ben. JSON básico. São Paulo: Novatec Editora Ltda, 2015. 400 p.

TOUCANO. 3D Bar Chart. 2016. Disponível em: <<https://toucano.uk/#gallery-3dbar>>. Acesso em: 28 out. 2017.

Ward, M., Grinstein, G., Kein, D. Interactive data visualization: foundations, techniques, and applications, CRC Press, 2015.

YIN, Robert K. Estudo de caso – planejamento e métodos. 2. ed. Porto Alegre: Bookman, 2001. 205 p. Tradução de: Daniel Grassi.