

# Desenvolvimento de uma Ferramenta Web para Migração e Visualização dos Dados Abertos do DATASUS

Marcelo Henrique Casali, Cristiano Bertolini, Guilherme Bernardino da Cunha

<sup>1</sup>Departamento de Tecnologia da Informação  
Universidade Federal de Santa Maria (UFSM)  
Campus Frederico Westphalen - Frederico Westphalen - RS - Brazil

marcelohcasali@gmail.com, cristiano.bertolini@ufsm.br, guilherme@ufsm.br

**Abstract.** *Health data is complex and needs efficient visualization systems. The main objective of this work is to analyze, implement and migrate open government databases in the health area according to DATASUS data. The data extracted will refer to the single health system (SUS). The main contribution is the information visualization and migration system, represented in maps and graphs by queries in a MySQL database, bringing all available information from each municipality and state regarding SUS hospital morbidity due to external causes.*

**Resumo.** *Dados da área da saúde são complexos e precisam de sistemas de visualização eficientes. O principal objetivo deste trabalho é analisar, implementar e migrar as bases de dados governamentais abertas na área da saúde de acordo com os dados do DATASUS. Os dados extraídos serão referentes ao sistema único de saúde (SUS). A maior contribuição é o sistema de visualização e migração das informações, representadas em mapas e gráficos por consultas em um banco de dados MySQL, trazendo todas as informações disponíveis de cada município e estado referente à morbidade hospitalar do SUS por causas externas.*

## 1. Introdução

A lei de acesso à informação pública (Lei 12.527/2011) [Planalto.gov.br 2011] regulamenta o acesso a dados e informações mantidas pelo governo, constitui um marco para a democratização da informação nacional pública [Federal 2015]. O maior benefício da reutilização dos dados abertos é a facilidade de combinar dados nos formatos entregues (*JSON, XML, CSV, RDF*) com múltiplas fontes de dados, interligando-se com outras iniciativas de dados abertos na Web.

A principal fonte de informação sobre dados abertos no Brasil é o Portal Brasileiro<sup>1</sup> de Dados Abertos [Perreira 2012]. O portal está sendo construído e estruturado de forma colaborativa entre diferentes instituições - públicas, privadas e sociedade civil. Nota-se também que o DATASUS<sup>2</sup>, com a base de dados sobre as atividades do Sistema Único de Saúde (SUS), vinculado ao Ministério da Saúde, também tornou-se uma importante ferramenta para a divulgação e publicação de dados abertos. Assim, além dos dados com base na área da saúde, também é possível obter informações financeiras e contábeis sobre o SUS.

---

<sup>1</sup>data.gov.br

<sup>2</sup><http://www2.datasus.gov.br/DATASUS/index.php>

Dados abertos governamentais são produzidos e colocados à disposição das pessoas para tornar possível não apenas sua leitura e acompanhamento, mas também sua reutilização em novos projetos, sites e aplicativos. Uma dificuldade para se trabalhar com os dados abertos disponibilizados pelos portais brasileiros é o trabalho na modelagem dos dados, que realizam a limpeza, transformação e reorganização de todo o conteúdo que será utilizado. O principal desafio é equilibrar as necessidades da aplicação com os dados e formatos disponibilizados.

A partir do ano de 2009 o governo brasileiro sancionou a lei de acesso à informação que determina que os órgãos públicos disponibilizem para a sociedade informações a respeito da gestão pública, e a partir de então os dados abertos passaram a ser disponibilizados. “São citados ainda pelo W3C<sup>3</sup> cerca de três benefícios que atingem diretamente ao cidadão a partir da disponibilização dos dados abertos: inclusão, transparência e responsabilidade” [Opendatamanual.org 2015].

Não somente no Brasil, mas no mundo inteiro, a criação de projetos que utilizam dados abertos é utilizada para melhorar a vida da população como um todo. Diante deste cenário, o presente trabalho se propôs a definir um sistema único, mas completo com os dados disponibilizados pelo DATASUS referente à morbidade hospitalar dos pacientes atendidos pelo sistema único de saúde (SUS) por causas externas.

Segundo as normas do SIH/SUS, as internações provocadas por causas externas devem ser classificadas no diagnóstico principal como:

- Acidentes englobam as quedas;
- Envenenamento;
- Afogamento;
- Queimaduras;
- Acidente de trânsito;

Este trabalho está organizado da seguinte forma: a Seção 2 apresenta um breve referencial teórico composto por sistema de visualização, dados abertos e banco de dados; a Seção 3 descreve a solução implementada que será o desenvolvimento de uma aplicação Web denominada WebSUS; a Seção 4 apresenta os trabalhos relacionados. Por fim, são apresentadas as conclusões e referências bibliográficas.

## **2. Referencial Teórico**

Esta seção apresenta um breve referencial teórico sobre as áreas de dados abertos, banco de dados e bibliotecas de visualização e análise de dados. Dividiu-se em 3 (três) etapas: sistema de visualização, dados abertos e banco de dados.

### **2.1. Sistema de Visualização**

Nesta seção serão apresentadas informações sobre sistemas de visualização, frisando a importância e benefícios dos mesmos para os usuários.

Sistemas de visualização consistem em comunicar a informação de maneira clara e efetiva utilizando meios gráficos. Não significa que a visualização de dados necessita ter um visual muito sofisticado. Para transmitir ideias efetivamente, tanto a forma estética

---

<sup>3</sup><https://www.w3.org/>

quanto as necessidades funcionais precisam estar equilibradas, promovendo a compreensão de um complexo conjunto de dados, comunicando seus principais aspectos de uma forma mais intuitiva [Freitas 2001].

A ferramenta *GoogleCharts*<sup>4</sup> desenvolvida pela *Google*<sup>5</sup> fornece uma maneira adequada de visualizar dados incorporados em um site, os gráficos e mapas são expostos como classes em *JavaScript*, gráficos de linhas simples a mapas de árvore hierárquicos complexos.

A ferramenta *GeoCharts*<sup>6</sup> é um mapa geográfico, podendo assim trabalhar dados com países, continentes ou regiões, com cores e valores atribuídos a partes específicas. Os valores são exibidos como uma escala de cores, podendo personalizar um gráfico ou mapa para se adequar à aparência de um site. O mapa é processado no navegador usando um *plugin* incorporado, aquele não é rolável ou móvel, mas pode ser configurado para permitir o zoom em alguns estados, porém o Brasil não está habilitado para o opção no momento.

## 2.2. Dados Abertos

Nesta seção serão abordados os principais conceitos sobre dados abertos, informando os principais portais brasileiros e os formatos disponíveis, comentando sobre suas principais características.

Esse cenário de manipulação e criação de aplicativos com dados abertos ocorre entre o setor público e privado, permitindo que o cidadão saia do papel de mero receptor da informação e passe a ser o atuante que interpreta os dados abertos em seu benefício e de sua comunidade [Opendatamanual.org 2015].

O termo Dados Abertos refere-se a dados que podem ser livremente usados, reusados e distribuídos por qualquer pessoa [Okfn.org 2015]. Os principais formatos utilizados são:

- JSON (*JavaScript Object Notation*) que é uma formatação de troca de dados, pois usa convenções que são as linguagens C e familiares, incluindo *C++*, *Java*, *JavaScript*, *Perl*, *Python* e muitas outras [Okfn.org 2015].
- XML (*Xtensible Markup Language*) recomendada pela W3C para a criação de documentos com dados organizados hierarquicamente, tais como textos, banco de dados ou desenhos vetoriais [Libreoffice.org 2015].
- CSV (*Comma Separated Values*) é um formato de arquivo de texto que pode ser usado para trocar dados de uma planilha entre aplicativos [Symphony 2015].
- RDF (*Resource Description Framework*) é uma sintaxe padrão para representar um grafo dirigido em XML desenvolvida pelo W3C- *World Wide Web Consortium* [W3C, 2001], é um modelo padrão para o intercâmbio de dados naWeb. Suas características é a facilidade de fusão de dados [W3C 2015].

Visando promover de forma mais organizada e produtiva o desenvolvimento de aplicativos que utilizem dados abertos, as organizações públicas e privadas têm organizado movimentos onde desenvolvedores, analistas e profissionais de diversas áreas tra-

---

<sup>4</sup><https://developers.google.com/chart/>

<sup>5</sup><https://developers.google.com/chart/interactive/docs/>

<sup>6</sup><https://developers.google.com/chart/interactive/docs/gallery/geochart>

lham juntos na análise e tratamento dos dados para posterior desenvolvimento dos aplicativos. Esses eventos são chamados de *hackathon* [Opendatamanual.org 2015].

O *hackathon*<sup>7</sup> é um tipo de evento que conta com a parceria da *Open Knowledge Foundation*<sup>8</sup> e a W3C Brasil<sup>9</sup>. Durante estes eventos a instituição realiza, organiza e disponibiliza a massa de dados aos participantes para análise e as equipes participantes têm o prazo geral de 48 horas para a análise e modelagem dos dados para o desenvolvimento, teste e entrega do projeto. No final deste período, os projetos são expostos e passam por uma votação onde os melhores projetos são premiados e disponibilizados ao público [Opendataday.org 2015].

A criação de projetos que utilizam dados abertos visa melhorar a vida da população global, auxiliando as pessoas a programarem melhor suas atividades e até mesmo para definir a localização de um futuro relacionando às suas condições de vida.

Diante deste cenário, o manual dos dados abertos governamentais<sup>10</sup> cita algumas características a respeito:

- Transparência e controle democrático;
- Participação popular;
- Melhores ou novos produtos e serviços privados;
- Inovação;
- Melhora na eficiência e efetividade de serviços governamentais;
- Medição do impacto das políticas;
- Conhecimento novo a partir da combinação de fontes de dados e padrões.

### 2.3. Banco de Dados

Nesta seção serão abordados os conceitos iniciais do *MySQL*, comentando um pouco sobre suas principais características.

Um mundo onde a globalização está cada vez mais presente na vida do cidadão, os processos estão, cada vez mais, automatizados, a necessidade de armazenamento de dados e informações se torna o primeiro passo para iniciar um projeto com dados abertos.

O *MySQL*<sup>11</sup> é um servidor e gerenciador de banco de dados (SGBD) relacional, de licença dupla (sendo uma delas de software livre), projetado inicialmente para trabalhar com aplicações de pequeno e médio portes, principalmente com aplicações para interface web, possuindo todas as características que um banco de dados de grande porte precisa, sendo reconhecido por algumas entidades como banco de dados *open-source* [PachevSasha 2007]. O *MySQL* além de ser extremamente rápido pelo fato de armazenar os dados em tabelas no modo *ISAM* (código de baixo nível), o *MySQL* é altamente confiável. Por possuir essas características, a ferramenta é indicada para o uso em aplicações em todas as áreas, independentemente do tamanho da aplicação.

---

<sup>7</sup><http://opendataday.org/>

<sup>8</sup><http://br.okfn.org/>

<sup>9</sup><http://www.w3c.br/>

<sup>10</sup>[http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual\\_Dados\\_Abertos\\_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf)

<sup>11</sup><https://www.mysql.com/>

### 3. Solução Implementada

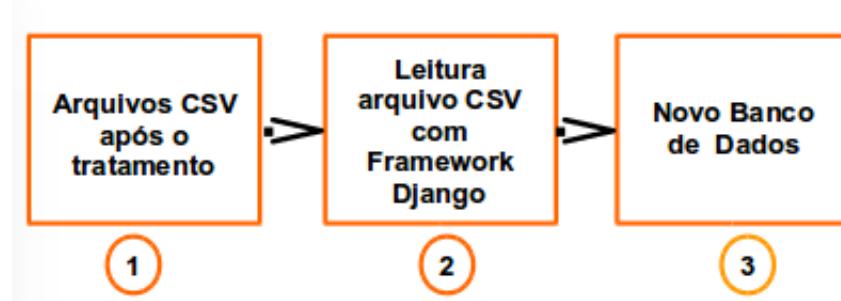
Nesta seção será apresentado o sistema de visualização de dados para o DATASUS denominado WebSUS. Tem-se como solução implementada a criação de uma aplicação que proporcionará ao usuário a visualização em mapas e gráficos dos dados abertos disponibilizados no DATASUS sobre a morbidade hospitalar dos pacientes atendidos no Sistema Único de Saúde - SUS com causas externas.

#### 3.1. Base de Dados

Para a construção do banco de dados, foram utilizados os arquivos disponíveis no DATA-SUS, com formato CSV.

Desenvolveu-se um algoritmo em *Python* com o *Framework Django*, que faz a importação dos arquivos CSV para o banco de dados *MySQL*, permitindo assim a inserção dos dados.

O processo de tratamento dos arquivos CSV e inserção dos dados para o banco de dados *MySQL* foi desenvolvido em três etapas diferentes conforme mostra a Figura 1.



**Figura 1. Etapas do processo de criação do Banco de Dados**

As 3 (três) etapas do processo de criação do novo banco de dados são descritas nas subseções a seguir.

#### 3.2. Primeira Etapa

Na primeira etapa do processo foi realizado o tratamento do arquivo CSV oriundo do DATASUS, onde é possível visualizar o arquivo original e o arquivo após o tratamento dos dados, para a leitura e a sua inserção manualmente no banco de dados *MySQL*.

A Figura 2 apresenta o arquivo CSV sem tratamento dos dados. Cada linha corresponde a uma cidade do Brasil, e as colunas apresentam informações das cidades, como quantidade de óbitos, internações, dias de permanência, taxa de mortalidade, valores e população. As colunas também representam a data, formada por ano e mês.

| A  | B        | C        | D        | E        | F        | G        | H        | I        | J        | K        | L        |
|--|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 Morbidade Hospitalar do SUS por Causas Externas - por local de internação - Brasil |          |          |          |          |          |          |          |          |          |          |          |
| 2 Internações por Município e Ano/mês atendimento                                    |          |          |          |          |          |          |          |          |          |          |          |
| 3 Período:Jan/2008-Set/2016  |          |          |          |          |          |          |          |          |          |          |          |
| 4 Município  | 2007/Jan | 2007/Fev | 2007/Mai | 2007/Jun | 2007/Jul | 2007/Ago | 2007/Sep | 2007/Out | 2007/Nov | 2007/Dez | 2008/Jan |
| 5 110001 Alta Floresta D'Oeste   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 6 110037 Alto Alegre dos Parecis   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 7 110040 Alto Paraisópolis   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 8 110034 Alvorada D'Oeste  | -        | -        | -        | -        | -        | -        | -        | -        | -        | 3        | 2        |
| 9 110002 Ariquemes   | -        | -        | -        | -        | -        | -        | -        | -        | -        | 12       | 39       |
| 10 110045 Buritis  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 11 110003 Cabixi   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 12 110060 Cacaulândia  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 13 110004 Cacoal   | -        | -        | -        | -        | -        | -        | -        | -        | -        | 8        | 45       |
| 14 110070 Campo Novo de Rondônia   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 15 110005 Cerejeiras   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 16 110008 Colorado do Oeste  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 17 110007 Corumbá  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 18 110008 Costa Marques  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 19 110094 Cujubim  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 20 110009 Espigão D'Oeste  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | 1        |
| 21 110100 Governador Jorge Teixeira  | -        | -        | -        | -        | -        | -        | -        | -        | -        | 2        | -        |
| 22 110010 Guajará-Mirim  | -        | -        | -        | -        | -        | -        | -        | -        | -        | 6        | 4        |
| 23 110110 Itapuã do Oeste  | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        |
| 24 110011 Jaru   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | 5        |
| 25 110012 Ji-Paraná  | -        | -        | -        | -        | -        | -        | -        | -        | 4        | 14       | 101      |
| 26 110013 Machadinho D'Oeste   | -        | -        | -        | -        | -        | -        | -        | -        | -        | -        | 1        |

**Figura 2. CSV antes de ser tratado**

A Figura 3 apresenta como ficou estruturado o arquivo CSV para a inserção dos dados no banco de dados MySQL. O primeiro processo para o tratamento dos dados foi remover as letras, deixando a primeira coluna com o código do Instituto Brasileiro de Geografia e Estatística (IBGE) referente aos municípios brasileiros, onde cada coluna representa o mês de um determinado ano, iniciando em janeiro de 2008 sucessivamente até dezembro de 2015.

Esse arquivo contém 5.565 (cinco mil quinhentos e sessenta e cinco) linhas e cada linha corresponde a um município do Brasil. Após a primeira coluna, os dados representam as informações como: quantidade de óbitos, internações, dias de permanência, taxa de mortalidade, valores e população.

| A         | B  | C  | D  | E  | F  | G  | H  | I  | J  | K  | L  | M  | N  | O  | P  | Q  |
|-----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1 110001  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 6  | 3  | 5  | 4  | 3  | 8  | 15 |
| 2 110037  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 3 110040  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 2  | 2  | 0  | 3  | 2  |    |
| 4 110034  | 2  | 0  | 0  | 0  | 1  | 0  | 1  | 1  | 0  | 1  | 1  | 3  | 3  | 1  | 1  | 3  |
| 5 110002  | 39 | 42 | 37 | 41 | 38 | 51 | 56 | 63 | 66 | 60 | 54 | 39 | 49 | 40 | 50 | 35 |
| 6 110045  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 7 110003  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 8 110060  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 9 110004  | 45 | 53 | 47 | 62 | 87 | 92 | 82 | 65 | 74 | 78 | 63 | 71 | 61 | 68 | 69 | 55 |
| 10 110070 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 11 110005 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 3  | 1  | 3  | 4  | 1  | 3  | 0  | 2  | 2  |
| 12 110006 | 0  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 2  | 1  | 2  |
| 13 110007 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |

**Figura 3. CSV após o tratamento**

### 3.3. Segunda Etapa

Na segunda etapa, desenvolveu-se um algoritmo em *Python* com o *Framework Django* para tratar especificamente dos arquivos CSV, possibilitando a leitura do arquivo CSV, até a inserção dos dados diretamente na tabela específica do banco de dados *MySQL*.

A Figura 4 representa o pseudo-código do algoritmo desenvolvido em *Python* com o *Framework Django*, possibilitando então a leitura e inserção dos registros em arquivos

CSV diretamente no banco de dados *MySQL*.

---

**Algorithm 1** Algoritmo

---

```
1: procedure def LERCSV(request):
2:     coluna = 1
3:     mes = 1
4:     anoInicial = 2008
5:     anoFinal = 2017
6:     while (anoInicial < anoFinal) : do
7:         whith open('arquivo.csv', 'rb')as csvfile :
8:             spamreader = csv.reader(csvfile, delimiter = ',', quotechar = '|')
9:             for row in spamreader: do
10:                 if row[0] is not None: then
11:                     t = Internacao()
12:                     t.idmunicipio_id = row[0]
13:                     t.quantidade = row[coluna]
14:                     t.data = str(ano) + '-' + str(mes) + '-01'
15:                     t.save(force_insert = True)
16:                     printt.idmunicipio_id, t.data, t.quantidade
17:                     mes = mes + 1
18:                     coluna = coluna + 1
19:                     print mes
20:                     if coluna %12 == 1 : then
21:                         print anoInicial
22:                         mes = 1
23:                         anoInicial = anoInicial + 1
```

---

**Figura 4.** Algoritmo desenvolvido para a inserção dos arquivos CSV

O Algoritmo apresentado na Figura 4, verifica o conteúdo no arquivo CSV, após a verificação, insere todos os registros encontrados diretamente na tabela específica no banco de dados *MySQL*.

### 3.4. Terceira Etapa

Na terceira etapa, pode-se visualizar a estrutura do banco de dados e o número de registros. Após a inserção de todos os dados, a Figura 5 apresenta as informações do novo banco de dados com dados inseridos do DATASUS, morbidade hospitalar dos pacientes atendidos no Sistema Único de Saúde - SUS com causas externas entre 2008 e 2015.

| Tabela            | Ação  | Linhos           | Tipo          | Colação                | Tamanho         | Sobrecarga     |
|-------------------|---|------------------|---------------|------------------------|-----------------|----------------|
| aih               | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 341,190          | InnoDB        | utf8_general_ci        | 15 MB           | -              |
| causas            | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 110,385          | InnoDB        | utf8_general_ci        | 20.1 MB         | -              |
| django_migrations | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 0                | InnoDB        | utf8_general_ci        | 16 KB           | -              |
| estados           | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 27               | InnoDB        | utf8_general_ci        | 16 KB           | -              |
| internacao        | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 341,190          | InnoDB        | utf8_general_ci        | 22 MB           | -              |
| municios          | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 5,565            | InnoDB        | utf8_general_ci        | 560 KB          | -              |
| obitos            | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 149,634          | InnoDB        | utf8_general_ci        | 8 MB            | -              |
| permanencia       | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 341,190          | InnoDB        | utf8_general_ci        | 17 MB           | -              |
| populacao         | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 44,520           | InnoDB        | utf8_general_ci        | 3.7 MB          | -              |
| taxa              | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 149,634          | InnoDB        | utf8_general_ci        | 8 MB            | -              |
| valor             | ★ <a href="#">Visualizar</a> <a href="#">Estrutura</a> <a href="#">Procurar</a> <a href="#">Inserir</a> <a href="#">Limpar</a> <a href="#">Eliminar</a> | 341,190          | InnoDB        | utf8_general_ci        | 17 MB           | -              |
| <b>11 tabelas</b> | <b>Soma</b>   | <b>1,824,525</b> | <b>InnoDB</b> | <b>utf8_general_ci</b> | <b>111.6 MB</b> | <b>0 Bytes</b> |

**Figura 5. Banco de dados após a inserção de todos os dados**

A Tabela 1 apresenta a quantidade de registros com que a base de dados se encontra atualmente.

| Tabela            | Linhos/Números de registros | Tamanho        |
|-------------------|-----------------------------|----------------|
| AIH               | 341.190                     | 15MB           |
| Causas            | 110.385                     | 20.1MB         |
| Estados           | 27                          | 16KB           |
| Internações       | 341.190                     | 22MB           |
| Municípios        | 5.565                       | 560KB          |
| Óbitos            | 149.634                     | 8MB            |
| Permanência       | 341.190                     | 17MB           |
| População         | 44.520                      | 3.7MB          |
| Taxa              | 149.634                     | 8MB            |
| Valor             | 341.190                     | 17MB           |
| <b>11 Tabelas</b> | <b>1.824.525</b>            | <b>111.6MB</b> |

**Tabela 1. Número de informações do banco de dados**

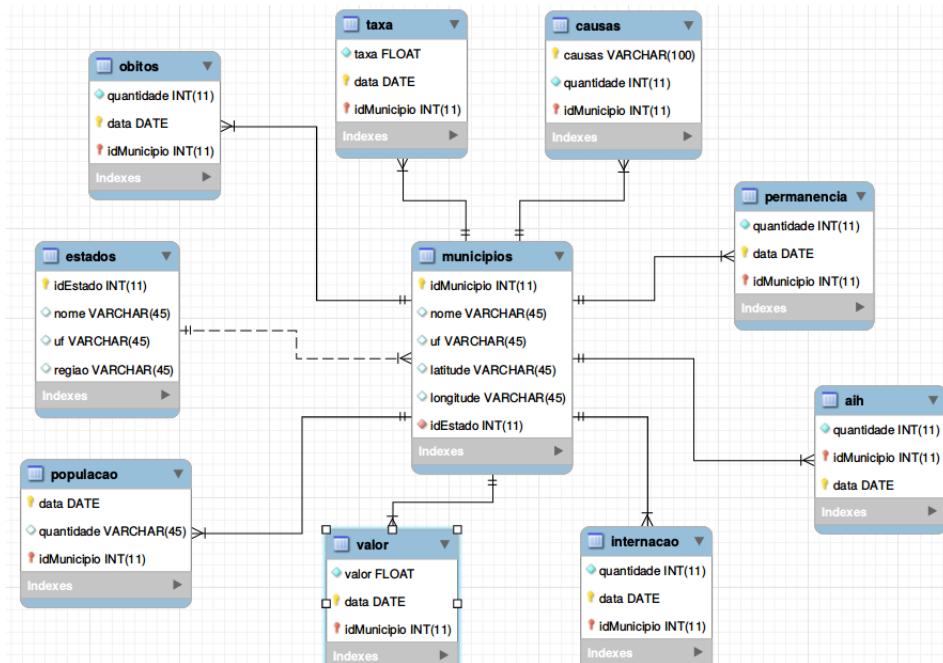
### 3.5. Consultas a Base de Dados

As consultas são baseadas na morbidade hospitalar do Sistema Único de Saúde - SUS com causas externas entre janeiro de 2008 a dezembro de 2015, podendo então interagir e relacionar os dados para futuras análises:

- **Número de óbitos:** designa a morte da pessoa que estava internada;
- **Número de internações:** qualquer internação motivada por uma causa externa, que por sua gravidade ou complexidade, exija um tratamento clínico ou cirúrgico;
- **Valor total gasto:** valores gastos com médicos, hospitais, internações, medicamentos e serviços em gerais;
- **Taxa de mortalidade:** a taxa de mortalidade é um índice demográfico obtido pela relação entre o número de mortos de uma população e um determinado espaço de tempo, neste caso, um ano;
- **Autorização de internações hospitalares:** arquivo do SUS conhecido como "AIH". A AIH é originada quando um hospital ou uma unidade de saúde gera uma solicitação de internação hospitalar e tem o objetivo de pré-validar os dados de internação;
- **Dias de permanência:** relação entre o total de pacientes/dia e o total de internados que tiveram saída do hospital em determinado período, incluindo os óbitos. Representa o tempo médio em dias que os pacientes ficaram no hospital.

- **População:** Número de habitantes referente ao ano informado. Podendo assim fazer futuras análises para verificar a média de óbitos pela quantidade de habitantes.

A Figura 6 apresenta o modelo Entidade Relacionamento (*ER*) do banco de dados, onde é possível visualizar o modelo dos dados implementados.



**Figura 6. Banco de dados utilizado no projeto**

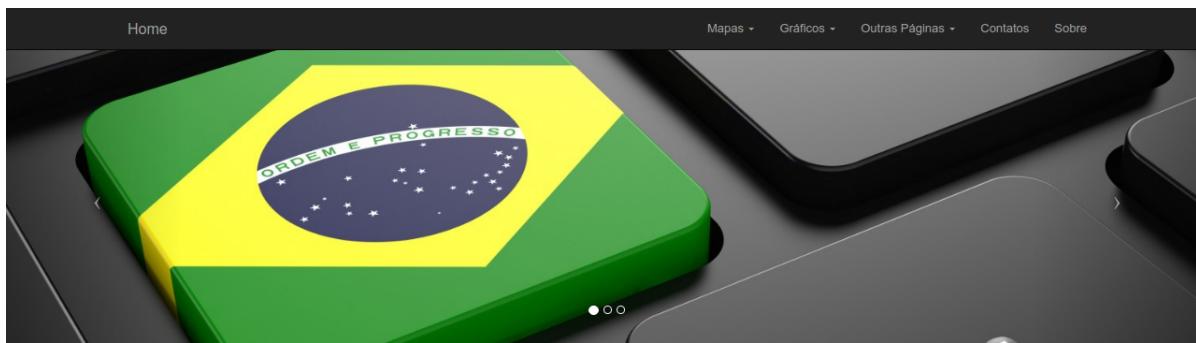
### 3.6. Sistema de Visualização

Nesta seção serão apresentadas informações que orientam o desenvolvimento do sistema denominado WebSUS. Explicando a criação da aplicação, que proporcionará ao usuário a visualização gráfica dos dados abertos da morbidade hospitalar do Sistema Único de Saúde - SUS com causas externas disponibilizados pelo DATASUS [Datasus 2015].

Os dados deste trabalho são fornecidos em forma de arquivo CSV, que dificulta a interpretação dos dados sem um tratamento para este projeto. [Traina 2001] enfatiza que os seres humanos não são eficientes para “interpretar” grande volume de dados em forma numérica ou textual, especialmente em espaços de altas dimensões, mas têm uma percepção muito boa quando esses dados são apresentados em forma de gráficos.

A visualização gráfica nos permite inferências sobre os dados e possibilita a comparação entre fontes diferentes. Sendo assim, este trabalho apresenta o desenvolvimento de uma aplicação web que utiliza os dados do DATASUS. Inicia-se desde a obtenção dos dados nas bases relacionais até sua consulta e visualização na interface web. Para o desenvolvimento do projeto foram utilizadas as seguintes tecnologias: *PHP* com *JavaScript* e *HTML* interagindo com a biblioteca *Google Charts* e o banco de dados *MySQL*.

A Figura 7 permite a visualização da interface Web desenvolvida.



## Bem Vindo

✓ Sobre o Datasus

O Departamento de Informática do Sistema Único de Saúde (DATASUS) surgiu em 1991 com a criação da Fundação Nacional de Saúde (Funasa), pelo Decreto 100 de 16.04.1991, publicado no D.O.U. de 17.04.1991.

[Ler mais](#)

❖ Origem dos dados

Os dados disponíveis são oriundos do Sistema de Informações Hospitalares do SUS - SIH/SUS, gerido pelo Ministério da Saúde, através da Secretaria de Assistência à Saúde, em conjunto com as Secretarias Estaduais de Saúde.

[Ler mais](#)

⌚ Dados abertos

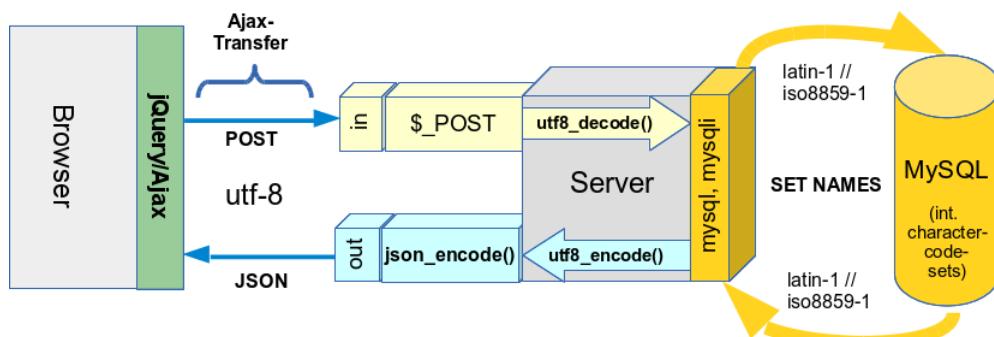
Dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, as exigências que visem preservar sua proveniência e sua abertura.

[Ler mais](#)

**Figura 7. Interface Web desenvolvida**

O objetivo principal foi o tratamento dos dados para a criação de um banco de dados relacional e assim podendo trabalhar com os dados na biblioteca *GoogleCharts* para indicar a localização e consequentemente os dados da mesma em forma de gráficos e mapas. Esses dados são obtidos através de uma requisição em *JSON* a partir de consultas no banco de dados.

A Figura 8 mostra a comunicação entre as tecnologias que foram utilizadas no projeto WebSUS para realizar a comunicação dos dados com a ferramenta *GoogleCharts*, que funciona da seguinte forma:



**Figura 8. Comunicação entre as tecnologias adotadas no projeto.**

De acordo com a Figura 8 tem-se:

- *Browser* envia dados de um formulário pelo método *\$\_POST* (envio de dados);
- *PHP* recebe os dados enviados pelo método *\$\_POST* e trabalha juntamente com *MySQL* realizando a consulta diretamente no servidor;
- *Json\_encode()* envia para o *Browser* os dados obtidos na consulta por um *array*;

- *Google Charts* trabalha juntamente com o *Browser* utilizando o *JSON* recebido.

A Figura 9 apresenta um exemplo de JSON, onde AIH representa o número de autorização de internação hospitalar, o nome que representa a cidade, e o valor representa o valor total gasto. Permitindo que se façam análises e trabalhar com a biblioteca Google-Charts. Nesse caso foi feito uma divisão do valor com o número de AIH, assim obtendo o valor médio de cada internação.



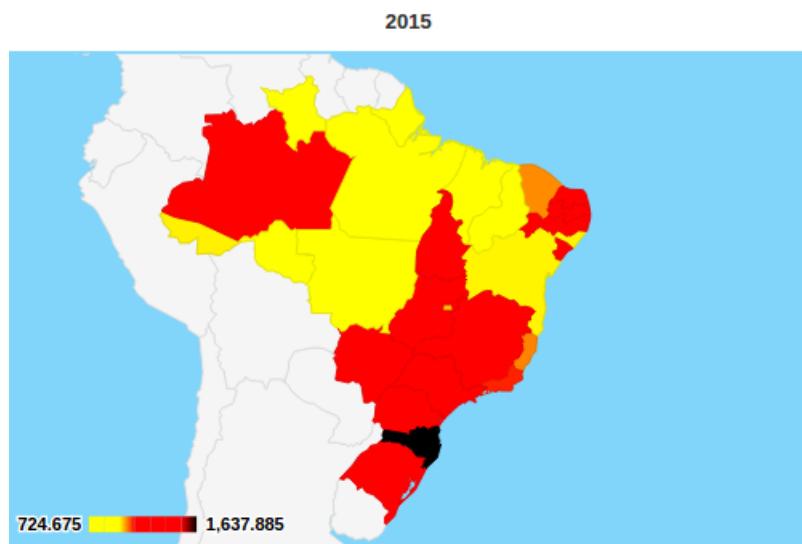
The screenshot shows a browser window with the URL `localhost/valor/valor.2014.php`. The page content displays a large block of JSON data representing internation statistics for 2014. The data is a list of objects, each containing an AIH number, a city name, and a total value. The JSON structure is as follows:

```
[{"aih": "4678", "nome": "Acre", "valor": "4126272.4800720215"}, {"aih": "13593", "nome": "Alagoas", "valor": "13742680.124191284"}, {"aih": "1994", "nome": "Amapá", "valor": "1329802.1221733093"}, {"aih": "10020", "nome": "Amazonas", "valor": "9781477.04921341"}, {"aih": "68265", "nome": "Bahia", "valor": "68122334.15128326"}, {"aih": "51493", "nome": "Ceará", "valor": "51212976.524513245"}, {"aih": "17495", "nome": "Distrito Federal", "valor": "17228890"}, {"aih": "24343", "nome": "Espírito Santo", "valor": "24819432.56194687"}, {"aih": "41918", "nome": "Goiás", "valor": "48724016.550457"}, {"aih": "35515", "nome": "Maranhão", "valor": "25961064.15840912"}, {"aih": "25924", "nome": "Mato Grosso", "valor": "21572378.423171997"}, {"aih": "19944", "nome": "Mato Grosso do Sul", "valor": "22309506.458946228"}, {"aih": "126446", "nome": "Minas Gerais", "valor": "164306427.9190445"}, {"aih": "51256", "nome": "Pará", "valor": "40725003.95681763"}, {"aih": "15225", "nome": "Paraíba", "valor": "18562851.41984558"}, {"aih": "80017", "nome": "Paraná", "valor": "110249288.33564758"}, {"aih": "55191", "nome": "Pernambuco", "valor": "71236711.13385773"}, {"aih": "21013", "nome": "Piauí", "valor": "19917831.895137787"}, {"aih": "65780", "nome": "Rio de Janeiro", "valor": "70042158.13788223"}, {"aih": "12102", "nome": "Rio Grande do Norte", "valor": "13828167.586532593"}, {"aih": "62484", "nome": "Rio Grande do Sul", "valor": "79554368.58679962"}, {"aih": "14642", "nome": "Rondônia", "valor": "11403203.085609436"}, {"aih": "29118", "nome": "Roraima", "valor": "2371435.660003662"}, {"aih": "44767", "nome": "Santa Catarina", "valor": "68237952.10068512"}, {"aih": "235557", "nome": "São Paulo", "valor": "296687549.46819305"}, {"aih": "7841", "nome": "Sergipe", "valor": "8401461.533565521"}, {"aih": "13554", "nome": "Tocantins", "valor": "14781669.150894165}]]
```

**Figura 9. Exemplo de um JSON**

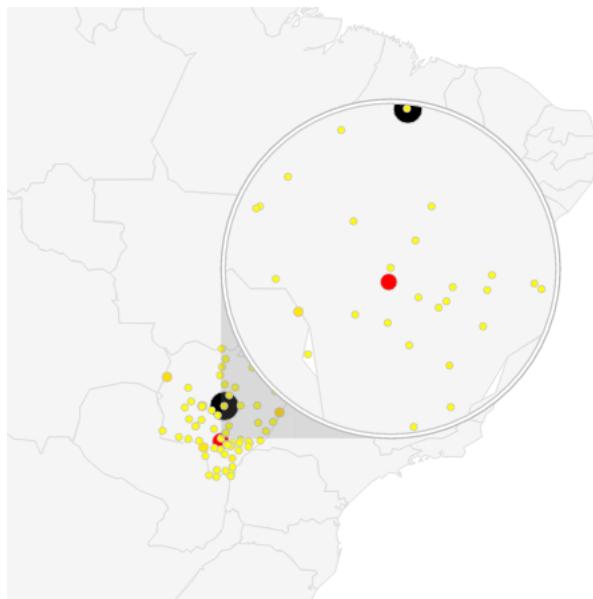
Para construção do mapa, foi utilizado o pacote *GeoCharts* do *GoogleCharts*. O serviço permite incorporar um mapa em uma página web usando a tecnologia *JavaScript*; bem como o tratamento dos dados e a forma de visualização.

A Figura 10 apresenta um mapa, referente ao valor gasto e um valor médio de cada internação por causas externas do SUS no ano de 2015. Pode-se observar que o estado de Santa Catarina apresentou um maior custo por internação em relação à quantidade de internações, sendo um custo médio por internação de R\$ 1.637,88 e gasto total de R\$ 74.394,35. Os estados com menos gastos por internações são: Amapá com custo médio por internação de R\$ 763,34 e gasto total de R\$ 1.664,83 e Maranhão com custo médio por internação de R\$ 724,67 gasto total de R\$ 26.376,71.



**Figura 10. Mapa indicador de estados.**

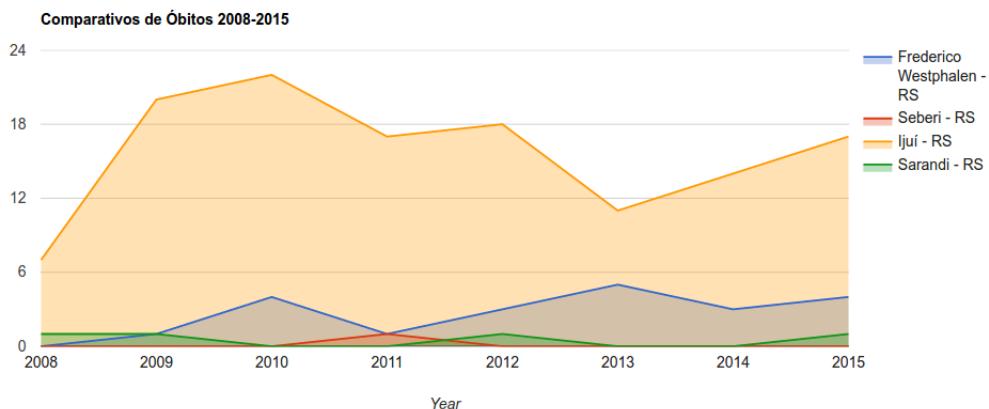
A Figura 11 apresenta os dados analisados em Mato Grosso do Sul e seus municípios, referente à quantidade de internações por causas externas do SUS. Pode-se notar que os municípios com mais internações foram Campo Grande-MS (na cor preta) com 10.202 internações no ano de 2015, e em seguida Dourados-MS (na cor vermelha) com 2.114 internações no ano de 2015. Neste mapa o usuário poderá escolher o estado, data inicial e final e também como conteúdo desejado.



**Figura 11. Mapa indicador de municípios.**

A Figura 12 apresenta um gráfico comparativo de óbitos com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos. Na vertical, mostra-se a quantidade de óbitos dos municípios, e no eixo da horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou du-

rante o período amostrado, uma quantidade de óbitos maior que as outras 3 (três) cidades apresentadas.



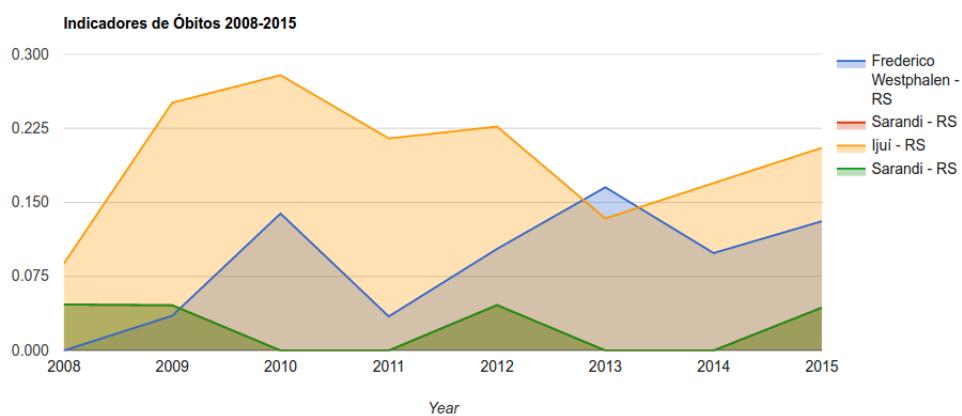
**Figura 12. Gráfico comparativo de óbitos.**

A Tabela 2 apresenta comparativos de óbitos de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, os valores destacados na tabela representam os maiores números de óbitos das respectivas cidades apresentadas na Figura 12. De acordo com a Tabela 2, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos maior que as outras 3 (três) cidades apresentadas.

| Municípios              | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|-------------------------|------|------|------|------|------|------|------|------|
| Frederico Westphalen-RS | 0    | 1    | 4    | 1    | 3    | 5    | 3    | 4    |
| Seberi-RS               | 0    | 0    | 0    | 1    | 0    | 0    | 0    | 0    |
| Ijuí-RS                 | 7    | 20   | 22   | 17   | 18   | 11   | 14   | 17   |
| Sarandi-RS              | 1    | 1    | 0    | 0    | 1    | 0    | 0    | 1    |

**Tabela 2. Tabela com comparativos de óbitos entre 2008-2015**

A Figura 13 apresenta um gráfico indicador de óbitos com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos por mil habitantes. Na vertical, mostra-se o índice de óbitos dos municípios por mil habitantes, e no eixo da horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2013 onde a cidade com maior índice foi Frederico Westphalen.



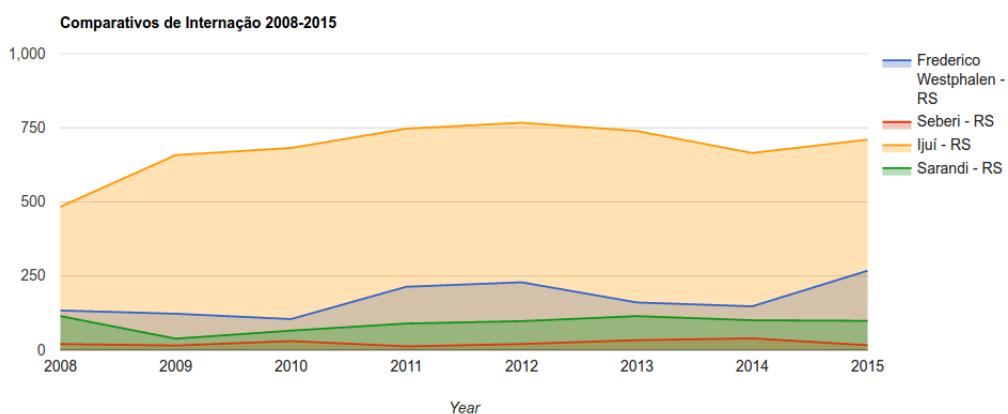
**Figura 13. Gráfico indicador de óbitos.**

A Tabela 3 apresenta o indicador de óbitos de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, o valor destacado na tabela representa os maiores números de óbitos das respectivas cidades apresentadas na Figura 13. De acordo com a Tabela 3, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2013 onde a cidade com maior índice foi Frederico Westphalen.

| Municípios              | 2008  | 2009  | 2010  | 2011  | 2012  | 2013  | 2014  | 2015  |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Frederico Westphalen-RS | 0     | 0,035 | 0,139 | 0,034 | 0,103 | 0,165 | 0,099 | 0,131 |
| Seberi-RS               | 0     | 0     | 0     | 0,092 | 0     | 0     | 0     | 0     |
| Ijuí-RS                 | 0,088 | 0,251 | 0,279 | 0,215 | 0,227 | 0,134 | 0,170 | 0,205 |
| Sarandi-RS              | 0,047 | 0,046 | 0     | 0     | 0,046 | 0     | 0     | 0,043 |

**Tabela 3. Tabela com indicador de óbitos entre 2008-2015**

A Figura 14 apresenta um gráfico comparativo de óbitos com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de internações. Na vertical, mostra-se a quantidade de internações dos municípios, e no eixo da horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou durante o período amostrado, uma quantidade de internação maior que as outras 3 (três) cidades apresentadas.



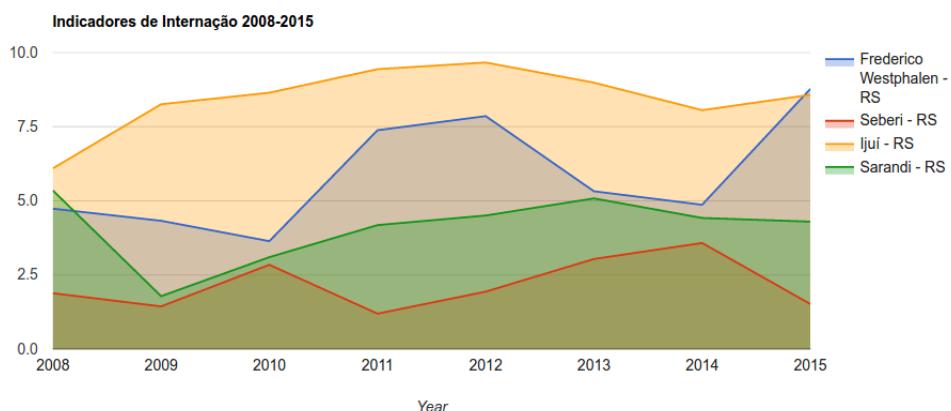
**Figura 14. Gráfico comparativo de internação.**

A Tabela 4 apresenta comparativos de internação de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, os valores destacados na tabela representam os números de internações das respectivas cidades apresentadas Figura 14. De acordo com o Tabela 4, o município de Ijuí apresentou durante o período amostrado, uma quantidade de internação maior que as outras 3 (três) cidades apresentadas.

| Municípios              | 2008       | 2009       | 2010       | 2011       | 2012       | 2013       | 2014       | 2015       |
|-------------------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Frederico Westphalen-RS | 134        | 123        | 105        | 214        | 229        | 161        | 148        | 268        |
| Seberi-RS               | 21         | 16         | 31         | 13         | 21         | 34         | 40         | 17         |
| Ijuí-RS                 | <b>484</b> | <b>658</b> | <b>682</b> | <b>747</b> | <b>767</b> | <b>739</b> | <b>665</b> | <b>710</b> |
| Sarandi-RS              | 115        | 39         | 66         | 90         | 98         | 115        | 101        | 99         |

**Tabela 4. Tabela com comparativos de internação entre 2008-2015**

A Figura 15 apresenta um gráfico indicador de internações com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de internações por mil habitantes. Na vertical, mostra-se o índice de internações dos municípios por mil habitantes, e no eixo das horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2015 onde a cidade com maior índice foi Frederico Westphalen.



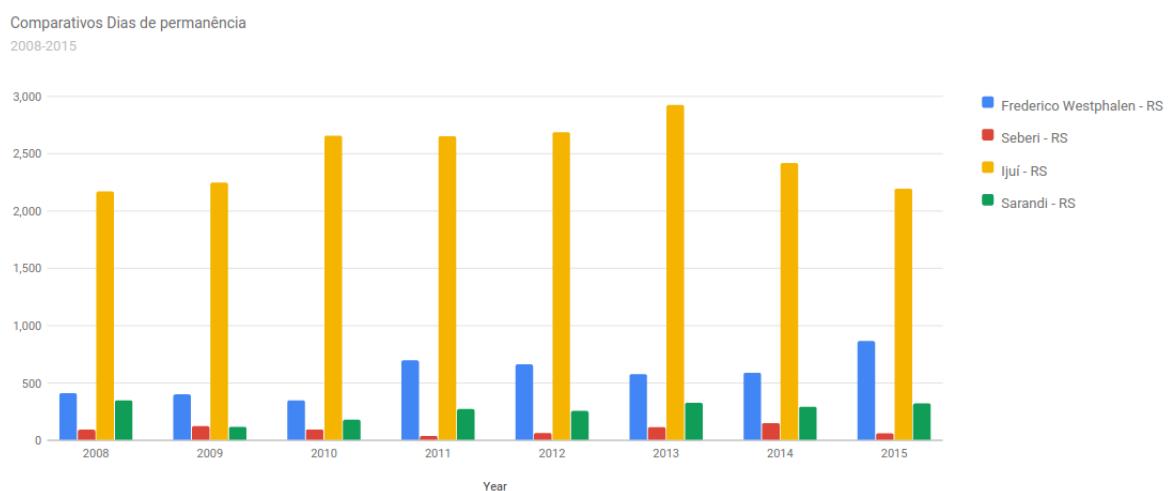
**Figura 15. Gráfico indicador de internação**

A Tabela 5 apresenta indicadores de internação de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, os valores destacados na tabela representam as maiores quantidades de internações hospitalares das respectivas cidades apresentadas na Figura 15. De acordo com a Tabela 5, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2015 onde a cidade com maior índice foi Frederico Westphalen

| Municípios              | 2008         | 2009         | 2010         | 2011         | 2012         | 2013         | 2014         | 2015         |
|-------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Frederico Westphalen-RS | 4,736        | 4,327        | 3,640        | 7,379        | 7,854        | 5,322        | 4,867        | <b>8,770</b> |
| Seberi-RS               | 1,887        | 1,442        | 2,845        | 1,197        | 1,939        | 3,039        | 3,579        | 1,523        |
| Ijuí-RS                 | <b>6,095</b> | <b>8,254</b> | <b>8,642</b> | <b>9,437</b> | <b>9,660</b> | <b>8,982</b> | <b>8,054</b> | 8,571        |
| Sarandi-RS              | 1,887        | 1,442        | 2,845        | 1,197        | 1,939        | 3,039        | 3,579        | 1,523        |

**Tabela 5. Tabela com indicadores de internação entre 2008-2015**

A Figura 16 apresenta um gráfico comparativo em dias de permanência hospitalar com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015. Na vertical, mostra-se a quantidade de dias permanecidos no hospital dos municípios, e no eixo da horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou durante o período amostrado, uma quantidade de permanência maior que as outras 3 (três) cidades apresentadas.



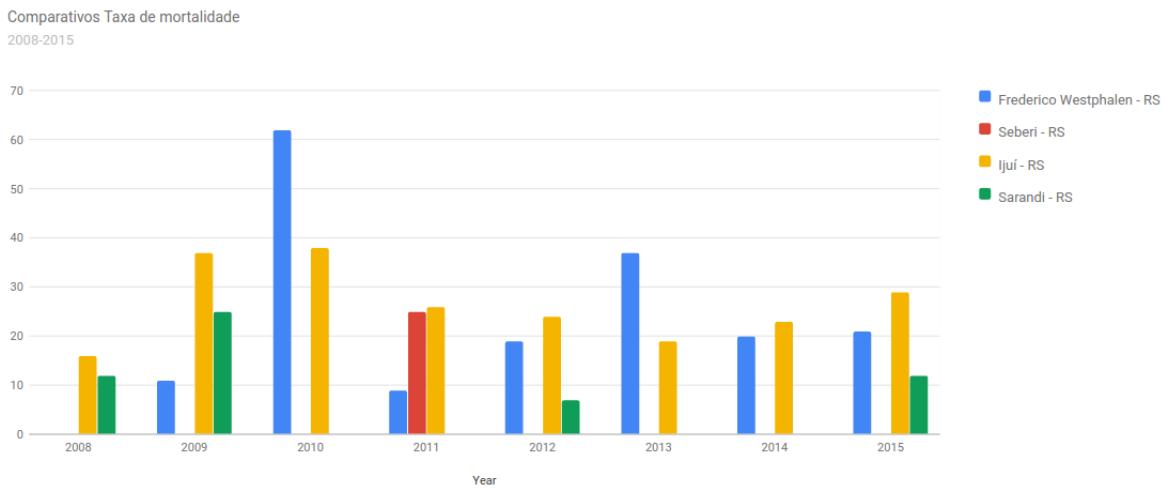
**Figura 16. Gráfico comparativo em dias de permanência hospitalar.**

A Tabela 6 apresenta dados correspondentes aos dias de permanência hospitalar de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, os valores destacados na tabela representam os maiores números de permanência hospitalar das respectivas cidades apresentadas na Figura 16. De acordo com a Tabela 6, o município de Ijuí apresentou durante o período amostrado, uma quantidade de permanência maior que as outras 3 (três) cidades apresentadas.

| Municípios              | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|-------------------------|------|------|------|------|------|------|------|------|
| Frederico Westphalen-RS | 415  | 405  | 353  | 702  | 667  | 581  | 593  | 870  |
| Seberi-RS               | 97   | 128  | 99   | 43   | 68   | 120  | 155  | 65   |
| Ijuí-RS                 | 2175 | 2252 | 2662 | 2657 | 2692 | 2929 | 2423 | 2200 |
| Sarandi-RS              | 353  | 122  | 184  | 277  | 261  | 332  | 297  | 326  |

**Tabela 6. Tabela comparativa em dias de permanência hospitalar entre 2008-2015**

A Figura 17 apresenta um gráfico comparativo da taxa de mortalidade com 4 (quatro) municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos e o número de AIH aprovadas, computadas como internações, no período, multiplicada por 100 (cem). Na vertical, mostra-se a taxa de mortalidade, e no eixo das horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2010 e 2013 onde a cidade com maior taxa foi Frederico Westphalen.



**Figura 17. Gráfico comparativo da taxa de mortalidade referente a quantidade de óbitos e o número de AIH aprovadas.**

A Tabela 7 apresenta dados correspondente à taxa de mortalidade, referente às AIH aprovadas de 4 (quatro) municípios por causas externas do SUS entre 2008-2015, o valor destacado na tabela representa as maiores taxas de mortalidades das respectivas cidades apresentadas na Figura 17. De acordo com a Tabela 7, o município de Ijuí apresentou durante o período amostrado, uma quantidade de óbitos por mil habitantes maior que as outras 3 (três) cidades apresentadas, exceto em 2010 e 2013 onde a cidade com maior taxa foi Frederico Westphalen.

| Municípios              | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|-------------------------|------|------|------|------|------|------|------|------|
| Frederico Westphalen-RS | 0    | 11   | 62   | 9    | 19   | 37   | 20   | 21   |
| Seberi-RS               | 0    | 0    | 0    | 25   | 0    | 0    | 0    | 0    |
| Ijuí-RS                 | 16   | 37   | 38   | 26   | 24   | 19   | 23   | 29   |
| Sarandi-RS              | 12   | 25   | 0    | 0    | 7    | 0    | 0    | 12   |

**Tabela 7. Tabela comparativa da taxa de mortalidade referente a quantidade de óbitos e o número de AIH aprovadas.**

No sistema desenvolvido o usuário poderá realizar diversas consultas aos dados do DATASUS de qualquer município ou estado do Brasil. A análise de comparativos e indicadores será visualizado em gráficos, sendo que todos os dados dos indicadores foram trabalhados com um índice de 1.000 (mil) habitantes, podendo assim ter uma análise melhor dos dados e do sistema. Na área de mapas pode-se visualizar os estados e municípios.

#### 4. Trabalhos Relacionados

A ação do Governo Federal em propiciar dados relacionados a múltiplos serviços, envolvendo dados da saúde, idealiza muitas possibilidades para visualização de dados e planejamento. Nesta seção são apresentados alguns trabalhos relacionados.

Segundo [Ribeiro 2011] várias ações estão sendo realizadas para incentivar o trabalho com dados abertos, não somente no Brasil mas em nível mundial. Estas ações estão facilitando o intercâmbio de dados e a descoberta de informações relevantes que levam a

melhora dos serviços como, por exemplo, o saúde pública. Ribeiro cita que existe um projeto em desenvolvimento, chamado DadosGov desenvolvido pelo Comitê de Organização de Informações da Presidência da República (COI-PR), o projeto permite a criação de um catálogo de informações abertas, aprimorando a gestão pública e facilitando o acompanhamento de todos os cidadãos.

Segundo [Traina 2001], os seres humanos possuem uma capacidade inata para interpretar informações apresentadas em formato gráfico, conseguindo reconhecer mais facilmente o conhecimento associado às informações.[Traina 2001] apresenta uma ferramenta para a visualização de dados de forma gráfica onde os dados armazenados ficam em uma base relacional, a qual é independente de qualquer processo de sumarização dos mesmos. Entretanto, a maioria das ferramentas permitem somente a visualização das informações estáticas sobre os dados e poucas permitem a visualização em forma dinâmica.[Traina 2001] frisa o futuro desenvolvimento de uma técnica para a visualização de dados a partir de base de dados composta por qualquer tipo de informação, assim implementando uma ferramenta prática para a reutilização de base de dados.

[Canban 2015], informa que existem novas interfaces interativas na área da saúde em desenvolvimento, pois o volume de dados clínicos é grande. Assim, existe a necessidade de apoiar novos projetos para a saúde e, desta forma, médicos podem encontrar maneiras mais eficazes para tratamentos clínicos e futuras tomadas de decisões. O trabalho apresenta uma análise visual da saúde em forma de oportunidades no mercado tecnológico e desafios para pesquisas, gerando dados com pacientes em forma de relatórios onde os médicos podem comparar casos de pacientes semelhantes, facilitando assim o tratamento.[Canban 2015] cita um sistema chamado *HARVEST*, que tem uma interface interativa para registros de cada paciente, com isso fornecendo um amplo conteúdo para ser analisado, assim os clínicos e pesquisadores podem gerar hipóteses e obter uma compreensão mais profunda de cada caso. Em particular percebeu que algumas áreas como a saúde, devem ter mais atenção quando o assunto é tecnologia, devido à falta de métodos padronizados de avaliação, validação e eficácia das ferramentas de análise visual.

[Germano 2008], o presente estudo realizado foi parte de uma pesquisa realizada em aglomerados de municípios brasileiros com precariedade das informações de óbito, visando a estimativa da mortalidade infantil nessas áreas no ano 2000. O artigo discute as principais fontes de informação do óbito infantil nesses municípios, assim como alguns dos problemas relacionados à operacionalização do Sistema de Informações sobre Mortalidade (SIM). A metodologia adotada foi a de busca ativa de óbitos em fontes oficiais e não oficiais de informações previamente cadastradas. A busca ativa identificou 344 óbitos não processados no SIM (66% do total de óbitos). Foi observado um baixo percentual de óbitos processados no SIM oriundos das fontes oficiais: cartórios e estabelecimentos de saúde, o que assinalou para problemas tanto na notificação do óbito e emissão da declaração de óbito quanto no fluxo de informações. A elevada contribuição do Programa Agentes Comunitários de Saúde/Programa Saúde da Família na identificação de óbitos infantis aponta para a necessidade de busca ativa de informações nesta fonte como rotina nos municípios com precariedade dos dados. Paulo Germano [Germano 2008] frisa que a insuficiente capacitação dos profissionais e a alta rotatividade de técnicos foram questões-chave para a deficiente operacionalização do sistema.

Segundo [Zacarias et al. 2015], devido à falta de dados disponíveis e incompletos

dificulta-se realizar comparações em bases de dados. [Zacarias et al. 2015] percebeu que os dados referentes a acidentes de trabalho não tem a devida importância, pois além de poucos dados, eles não são atualizados periodicamente. Com uma maior quantidade de dados, acredita-se que a aplicação (Análise Comparativa dos Acidentes de Trabalho no Brasil a partir de Dados Abertos) desenvolvida poderia ser estendida para o uso de técnicas estatísticas, assim prevendo determinadas situações e até mesmo para servir de apoio a decisões do governo. Este trabalho estende o proposto em [Zacarias et al. 2015] e utiliza, além de um volume de dados maior, uma maior precisão na visualização dos mesmos.

Na tabela 8, apresentam-se as diferenças e contribuições em relação ao trabalho proposto.

| Trabalhos         | Software Utilizados  | Banco de Dados | Plataforma | Importação de arquivos CSV | Principais contribuições  |
|-------------------|--|----------------|------------|----------------------------|---|
| Ribeiro (2011)    | X  | X              | X          | X                          | Contextualizar a temática de dados abertos<br>Pesquisa de dados abertos na área da saúde  |
| Traina (2011)     | C++ Builder  | Oracle         | Desktop    | Não informado              | Algoritmo para visualização do banco de dados<br>Reutilização de banco de dados<br>Contextualizar a temática de dados abertos   |
| Canban JJ (2015)  | X  | X              | X          | X                          | Pesquisa de tecnologia na área da saúde<br>Possíveis desafios encontrados em pesquisas<br>Análise visual de dados na área da saúde  |
| Germano (2008)    | X  | X              | X          | X                          | Revelar informações sobre óbitos infantis<br>Pesquisa sobre óbitos infantis   |
| Zacarias (2015)   | HTML<br>CSS<br>JavaScript<br>Python - Django<br>API Highcharts       | Não informado  | Web        | Não informado              | Pesquisa de dados abertos na área da saúde<br>Contextualizar a temática de dados abertos<br>Desenvolvimento WEB   |
| Trabalho proposto | HTML<br>CSS<br>JavaScript<br>Python - Django<br>PHP<br>Google Charts | MySQL          | Web        | Python - Django            | Pesquisa de dados abertos na área da saúde<br>Pesquisa de dados do DATASUS<br>Contextualizar a temática de dados abertos<br>Desenvolvimento WEB<br>Algoritmo de inserção de dados CSV em banco de dados MySQL |

**Tabela 8. Tabela comparativa dos trabalhos relacionados**

## 5. Conclusões

O presente artigo teve como enfoque principal o estudo de bases de dados abertas na área da saúde, do Sistema Único de Saúde (SUS), extraíndo os dados do site DATASUS, convertendo-os em uma forma mais prática e atrativa para a sua interpretação, conseguindo assim demonstrar em forma gráfica e mapeada como ocorreu a evolução da morbidade hospitalar do SUS por causas externas, podendo saber em que locais houve maior uso deste serviço.

Os dados disponibilizados pelo DATASUS são mensalmente atualizados. Percebe-se que os dados, mesmo sendo atualizados, não têm a devida importância que deveriam ter, pois o arquivo fornecido dificulta comparações e análises, pois os arquivos disponibilizados são no formato CSV.

A utilização de dados abertos com informações integradas a ferramentas para visualização de dados, facilitam o dia-a-dia dos cidadãos e da sociedade. Acessando os mapas e gráficos, o usuário poderá ter as informações com mais clareza e visibilidade, sem que seja necessário o acesso a tabelas e arquivos em forma de texto. Isso permite o usuário visualizar as informações com mais facilidade para entender e ter uma visão geral do Brasil.

É importante frisar que a principal contribuição deste trabalho foi a migração dos dados obtidos do DATASUS para um banco de dados MySQL. A partir desta base

de dados, ela poderá ser reutilizada por outros projetos, podendo trazer ainda mais benefícios à população brasileira, permitindo o cidadão a economizar tempo na busca de informações, interpretando os dados de uma forma mais atrativa e significativa, ainda assim beneficiando-se de novos conhecimento na área da saúde.

Os gráficos facilitam a análise e interpretação de um conjunto de dados. Neste trabalho, foi possível visualizar a importância dos gráficos indicadores. Relacionando os dados obtidos com o número de habitantes, podemos ter uma real visualização e situação do Sistema Único de Saúde (SUS) com morbidade hospitalar por causas externas.

A aplicação web desenvolvida poderia ser estendida para usar mais formas estatísticas e implementar algum método para atualizar a base de dados mensalmente, então podendo servir como apoio à decisões do governo.

Como trabalhos futuros, serão realizados os refinamentos e ajustes, podendo aproveitar o máximo possível do banco de dados e abordando mais a implementação da parte estatística dos dados disponíveis.

A aplicação desenvolvida pode ser acessada através do endereço <http://200.132.38.214:8003/datasus>.

## Referências

- [Canban 2015] Canban, D. G. (2015). Visual analytics in healthcare – opportunities and research challenges. *Oxford University - JAMIA*. [http://www.skateboardingalice.com/papers/2015\\_Caban.pdf](http://www.skateboardingalice.com/papers/2015_Caban.pdf).
- [Datasus 2015] Datasus (2015). Portal da saúde. <http://www2.datasus.gov.br/DATASUS/index.php>. Acessado em 05/12/2016.
- [Federal 2015] Federal, B. G. (2015). Portal brasileiro de dados abertos. <http://dados.gov.br>. Acessado em 05/12/2016.
- [Freitas 2001] Freitas, C. M. D. S. F. (2001). Introdução à visualização de informações. *RITA*, pages 2–16. <https://goo.gl/uJ19kJ>.
- [Germano 2008] Germano, P. G. F. (2008). Sistema de informações sobre mortalidade: estudo de caso em municípios com precariedade dos dados. *SCPRJ*, pages 2–10. <http://www.scielo.br/pdf/csp/v24n10/07.pdf>.
- [Libreoffice.org 2015] Libreoffice.org (2015). Formatos de arquivos xml. [https://help.libreoffice.org/Common/XML\\_File\\_Formats/pt-BR](https://help.libreoffice.org/Common/XML_File_Formats/pt-BR). Acessado em 05/12/2016.
- [Okfn.org 2015] Okfn.org (2015). Open knowledge. <https://okfn.org/>. Acessado em 05/12/2016.
- [Opendataday.org 2015] Opendataday.org (2015). International open data hackathon. <http://opendataday.org/>. Acessado em 05/12/2016.
- [Opendatamanual.org 2015] Opendatamanual.org (2015). Manual of open data. [http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual\\_Dados\\_Abertos\\_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf). Acessado em 05/12/2016.

- [PachevSasha 2007] PachevSasha, S. P. (2007). *Understanding MySQL Internals*. O'Reilly. <http://bootskii.net/books/Technical.Ebooks.Pack.2/OReilly%20Understanding%20MySQL%20Internals.pdf>.
- [Perreira 2012] Perreira, C. R. P. (2012). Elementos para construção de um modelo de análise de iniciativas de publicações de dados governamentais abertos. *EACH*, pages 32–63. <https://goo.gl/CF7DSY>.
- [Planalto.gov.br 2011] Planalto.gov.br (2011). Lei nº 12.527, de 18 de novembro de 2011. [http://www.presidencia.gov.br/ccivil\\_03/\\_Ato2011-2014/2011/Lei/L12527.html](http://www.presidencia.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.html). Acessado em 05/12/2016.
- [Ribeiro 2011] Ribeiro, C. J. S. R. (2011). Dados abertos governamentais (open government data): Instrumento para exercício da cidadania pela sociedade. *Enancib*, pages 124–135. <https://core.ac.uk/download/pdf/11890148.pdf>.
- [Symphony 2015] Symphony, I. L. (2015). Abrindo e salvando arquivos csv de texto. [http://infolib.lotus.com/resources/symphony/3.0.0/sym20abd014/pt\\_br/text/scalc/guide/csv\\_files.html](http://infolib.lotus.com/resources/symphony/3.0.0/sym20abd014/pt_br/text/scalc/guide/csv_files.html). Acessado em 05/12/2016.
- [Traina 2001] Traina, A. J. M. T. (2001). Dados governamentais abertos no brasil. *USP*, pages 2–15. <http://www.lbd.dcc.ufmg.br/colecoes/sbbd/2001/007.pdf>.
- [W3C 2015] W3C (2015). W3c rdf. <https://www.w3.org/RDF/>. Acessado em 05/12/2016.
- [Zacarias et al. 2015] Zacarias, I., Vitalli, R. A., Prediger, D., Moerschbacher, J., and Bertolini, C. (2015). Análise comparativa dos acidentes de trabalho no brasil a partir de dados abertos. *Anais do Computer on the Beach*, pages 229–238. <http://siaiap32.univali.br/seer/index.php/acotb/article/view/7038>.