

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS ARARANGUÁ
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE (CTS)
TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO (TIC)

Mateus Fernando Arnold Trierveiler

Visualização de Dados Abertos Educacionais do PROUNI com ChartJS

Araranguá

2020

Mateus Fernando Arnold Trierveiler

Visualização de Dados Abertos Educacionais do PROUNI com ChartJS

Trabalho de Conclusão do Curso de Graduação em
Tecnologias da Informação e Comunicação do
Centro de Ciências Tecnologia e Saúde da
Universidade Federal de Santa Catarina como
requisito para a obtenção do Título de Bacharel em
Tecnologias da Informação e Comunicação.
Orientador: Prof. Cristian Cechinel

Araranguá

2020

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

Trierveiler, Mateus Fernando Arnold

Visualização de Dados Abertos Educacionais do PROUNI com
ChartJS / Mateus Fernando Arnold Trierveiler ; orientador,
Cristian Cechinel, 2020.

43 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Campus Araranguá,
Graduação em Tecnologias da Informação e Comunicação,
Araranguá, 2020.

Inclui referências.

1. Tecnologias da Informação e Comunicação. 2.
Visualização de Dados. 3. Dados Abertos Educacionais. I.
Cechinel, Cristian. II. Universidade Federal de Santa
Catarina. Graduação em Tecnologias da Informação e
Comunicação. III. Título.

Mateus Fernando Arnold Trierveiler

Visualização de Dados Abertos Educacionais do PROUNI com ChartJS

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

Araranguá, 04 de Dezembro de 2020.



Documento assinado digitalmente
Vilson Gruber
Data: 10/12/2020 20:19:31-0300
CPF: 175.317.788-07

Prof. Vilson Gruber
Coordenador do Curso

Banca Examinadora:



Documento assinado digitalmente
Cristian Cechinel
Data: 04/12/2020 18:26:44-0300
CPF: 454.628.449-72

Prof. Cristian Cechinel
Orientador
Universidade Federal de Santa Catarina



Documento assinado digitalmente
Robson Rodrigues Lemos
Data: 04/12/2020 20:02:50-0300
CPF: 613.256.949-91

Prof. Robson Rodrigues Lemos
Avaliador
Universidade Federal de Santa Catarina



Documento assinado digitalmente
Giovani Mendonça Lunardi
Data: 04/12/2020 19:28:35-0300
CPF: 520.394.559-49

Prof. Giovani Mendonça Lunardi
Avaliador
Universidade Federal de Santa Catarina

AGRADECIMENTOS

Primeiramente agradeço aos meus pais e familiares que tanto me apoiaram nesta jornada. Também agradeço a todos meus amigos de longa data, a aqueles obtidos durante o curso de graduação e ao meu orientador Cristian Cechinel por me aconselhar e ter papel fundamental na apresentação deste trabalho.

RESUMO

Vivemos em constante avanço tecnológico e com isso surgem novas tecnologias a todo momento, gerando um grande fluxo de dados diário por todo o mundo. Muitas dessas informações chegam ao usuário de diversas maneiras, que nem sempre sabe como lidar e interpretar, ou simplesmente não é o tipo de dado que deseja. Sendo assim, são necessárias maneiras de tratar estes dados de forma a torná-los compreensíveis ao usuário, e uma destas formas é a visualização de dados. Pensando nisso, foi elaborado este trabalho que tem como objetivo desenvolver uma plataforma Web responsável por gerar Visualizações de Dados Abertos Educacionais do PROUNI utilizando a ferramenta ChartJS. Os dados foram obtidos através do Portal Brasileiro de Dados Abertos e serviram como base para a elaboração da aplicação e a análise exploratória a partir das visualizações geradas. Constatou-se que a maior parte das bolsas estão localizadas nos estados de São Paulo e Minas Gerais e que a maior proporção de bolsas se dá no tipo Bolsa Integral. Também foi possível verificar que a região sudeste contém a maior parte absoluta de bolsas do PROUNI. Com este projeto pode-se comprovar a eficácia da utilização das técnicas de visualização de dados através das análises realizadas, algo que não poderia ser feito com a leitura em tabelas dos dados originais.

Palavras-chave: Visualização de dados. Tecnologias da Informação e Comunicação. ChartJS.

ABSTRACT

We live in constant technological progress and new technologies appear at every moment, generating a large flow of data all over the world. Much of this information reaches the user in different modes, but he sometimes don't know how to handle and interpret it, or simply isn't the type of data he want. Therefore, he needs a way to treat this data to make it understandable, and one of these ways is data visualization. Thinking about it, this work was elaborated with objective to develop a web platform responsible for generating Open Educational Data Views from PROUNI using the ChartJS framework. The data was obtained through the Portal Brasileiro de Dados Abertos and served as a basis for the elaboration of the application and the exploratory analysis based on the generated visualizations. It was found that most of the scholarships are located in the states of São Paulo and Minas Gerais and that the largest proportion of scholarships are Integral type. It was also possible to verify that the southeast region contains the majority of PROUNI scholarships. With this project, it is possible to prove the effectiveness of using data visualization techniques through the analyzes performed, something that could not be done with reading the original data in files.

Keywords: Data Visualization. Information and Communication Technologies. ChartJS.

LISTA DE FIGURAS

Figura 1 - Gráfico de barras – Chart.js Samples.....	17
Figura 2 - Gráfico de linhas – Chart.js Samples.....	22
Figura 3 – Modelo lógico dos dados.....	24
Figura 4 – Script de criação da tabela.....	25
Figura 5 – Comando SQL de inclusão dos dados.....	27
Figura 6 – Código responsável pela obtenção dos dados da proporcionalidade de bolsas.....	28
Figura 7 – Código responsável pela obtenção dos dados por estado brasileiro.....	28
Figura 8 – Código responsável pela obtenção dos dados por região brasileira.....	29
Figura 9 – Inclusão dos arquivos do Bootstrap.....	29
Figura 10 – Inclusão dos arquivos JS da primeira página.....	30
Figura 11 – Inclusão dos arquivos JS da segunda página.....	30
Figura 12 – Resultado final da interface front-end.....	31
Figura 13 – Trecho de código da visualização da proporcionalidade de bolsas.....	32
Figura 14 – Trecho de código da visualização de bolsas por estado.....	33
Figura 15 – Trecho de código da visualização por região brasileira.....	34
Figura 16 – Resultado da Visualização da Quantidade de Bolsas por Estado do Brasil.....	37
Figura 17 – Quantidade de bolsas em São Paulo.....	38
Figura 18 – Filtro por quantidade de bolsas por estado.....	38
Figura 19 – Visualização da Proporção de bolsas integrais e parciais no Brasil.....	39
Figura 20 – Visualização da Quantidade de bolsas por região do Brasil.....	40
Figura 21 – Filtro por região do Brasil.....	40
Figura 22 – Visualização por instituição de ensino.....	40

LISTA DE ABREVIATURAS E SIGLAS

PROUNI Programa Universidade Para Todos

PHP Hypertext Preprocessor

HTML HyperText Markup Language

WEB Rede mundial de computadores

JS JavaScript

JSON JavaScript Object Notation

GEOJSON Geographic JavaScript Object Notation

SQL Structured Query Language

CSV Comma Separated Values

CPF Cadastro de Pessoas Físicas

ENEM Exame Nacional do Ensino Médio

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

SISU Sistema de Seleção Unificada

MEC Ministério da Educação

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETIVOS.....	15
1.1.1	Objetivo Geral	15
1.1.2	Objetivos Específicos	15
1.1.3	Justificativa	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	VISUALIZAÇÃO DE DADOS	17
2.2	APLICAÇÃO WEB.....	18
2.2.1	A linguagem PHP.....	18
2.2.2	HTML (Hyper Text Markup Language).....	19
2.2.3	A linguagem JavaScript.....	19
2.2.4	O banco de dados MySQL.....	19
2.2.5	A ferramenta Bootstrap.....	20
2.2.6	A biblioteca de Visualização de Dados ChartJS	20
2.2.7	Os formatos de arquivos JSON e GEOJSON	21
2.3	Dados Abertos	21
2.3.1	Portal Brasileiro de Dados Abertos.....	22
2.3.2	Dados Abertos Educacionais	22
2.4	Bolsas no ENSINO SUPERIOR BRASILEIRO	22
2.4.1	PROUNI – Programa Universidade Para Todos	23
3	DESENVOLVIMENTO	24
3.1	Obtenção e Modelagem dos Dados	24
3.1.1	Caracterização dos Dados	24
3.1.2	Modelo Lógico.....	25
3.1.3	Script de Criação da Tabela no Banco de Dados.....	25
3.2	Métodos de Visualização de dados.....	26

3.2.1	Visualização da quantidade de bolsas por estado do Brasil	27
3.2.2	Visualização da proporção de bolsas integrais e parciais no Brasil	27
3.2.3	Visualização da quantidade de bolsas por região do Brasil	27
3.2.4	Visualização por instituição de ensino	27
3.3	Desenvolvimento da Aplicação Web.....	27
3.3.1	Criação e População do Banco de Dados	28
3.3.2	Desenvolvimento da conexão com o banco de dados	29
3.3.3	Desenvolvimento do Front-end.....	30
3.3.4	Desenvolvimento das Visualizações de Dados.....	32
4	RESULTADOS	36
4.1	Visualização da quantidade de bolsas por estado do Brasil	36
4.2	Visualização da proporção de bolsas integrais e parciais no Brasil	39
4.3	Visualização da quantidade de bolsas por região do Brasil	39
4.4	Visualização por instituição de ensino.....	41
5	CONCLUSÃO.....	42
	REFERÊNCIAS.....	43

1 INTRODUÇÃO

Vivemos em constante avanço tecnológico e com isso surgem novas tecnologias a todo momento, gerando um grande fluxo de dados diário por todo o mundo. Muitas dessas informações chegam ao usuário de diversas maneiras, que nem sempre sabe como lidar e interpretar, ou simplesmente não é o tipo de dado que deseja. Sendo assim, são necessárias maneiras de tratar estes dados de forma a torná-los compreensíveis ao usuário, e uma destas formas é a visualização de dados. Segundo Keim e Ward(2003), existe um grande número de técnicas de visualização de informação que foram desenvolvidas ao longo dos últimos anos para apoiar a exploração de grandes conjuntos de dados.

Freitas (2001) diz que as técnicas de visualização de informações procuram representar visualmente dados de uma certa área, de modo que a representação visual gerada permita melhorar a capacidade cognitiva do ser humano e este, a partir disso, interprete e compreenda as informações mostradas e, finalmente deduza novos conhecimentos.

Buscando aplicar técnicas de visualização de dados será desenvolvido neste trabalho uma aplicação web utilizando um banco de dados MySQL e a ferramenta Chart.js, responsável por interpretar os dados e produzir com os mesmos gráficos visuais, simplificando e facilitando a análise dos dados.

1.1 OBJETIVOS

Nas seções abaixo estão descritos o objetivo geral e os objetivos específicos deste trabalho.

1.1.1 Objetivo Geral

O objetivo geral deste trabalho é desenvolver uma aplicação Web para visualização os dados do PROUNI utilizando a biblioteca ChartJS.

1.1.2 Objetivos Específicos

- Realizar a revisão da literatura sobre a área de visualização de dados e trabalhos relacionados que podem ser utilizados na aplicação;

- Obter os dados e adicioná-los à um banco de dados MySQL.
- Aplicar e adequar as melhores metodologias de visualização de dados para cada tipo de dado obtido.
- Desenvolver a aplicação web utilizando as linguagens PHP e JavaScript e a biblioteca ChartJS. A aplicação deve ser dinâmica e permitir a adição de novos dados ao banco de dados MySQL;
- Analisar as visualizações geradas de forma a compreender melhor os dados apresentados.

1.1.3 Justificativa

A proposta deste trabalho é elaborar uma plataforma web que permita a visualização de dados do PROUNI, obtidos através do Portal Brasileiro de Dados Abertos, de forma gráfica, facilitando a análise dos mesmos. Isso se justifica devido ao crescimento da quantidade de dados presentes nesta era tecnológica. Além disso, muitos dados não podem ser analisados sem uma ferramenta ou técnica específica, e o ChartJS permite a representação gráfica destes dados coletados.

Com as visualizações geradas, pode-se analisar os dados facilmente e de forma intuitiva, para que então se possa tirar as conclusões sobre os mesmos, tais como o crescimento do número de bolsas ofertadas por ano ou então a comparação da quantidade de bolsas entre as regiões do Brasil.

2 FUNDAMENTAÇÃO TEÓRICA

Este tópico apresenta todo o embasamento teórico fundamental utilizado para compreender os conceitos de visualização de dados e as tecnologias e linguagens utilizadas para gerar tais visualizações.

2.1 VISUALIZAÇÃO DE DADOS

Segundo Card, Mackinlay e Schneiderman (1999), Visualização de dados é o uso de representações visuais de dados interativas com o suporte do computador para ampliar o conhecimento. Few (2009) afirma o seguinte:

Visualização da Informação é ideal para análise exploratória de dados. Nossos olhos são naturalmente atraídos por tendências, padrões e exceções que seriam difíceis ou impossíveis de achar utilizando abordagens tradicionais, tais como tabelas ou textos. Durante a exploração dos dados, até mesmo os melhores estatísticos muitas vezes colocam os seus cálculos de lado por um momento e deixam os seus olhos conduzirem a busca. (FEW, 2009)

Figura 1 - Gráfico de barras – Chart.js Samples



Fonte: ChartJS Samples.

Gristein e Kein (2005) classificam as 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.

2.2 APLICAÇÃO WEB

Para o desenvolvimento da aplicação web responsável por rodar as visualizações de dados serão utilizadas algumas linguagens e ferramentas, que estão presentes neste tópico.

2.2.1 A linguagem PHP

Segundo Niederauer (2011), “PHP é uma das linguagens mais utilizadas na Web. Milhões de sites no mundo inteiro utilizam PHP. A principal diferença em relação às outras linguagens é a capacidade que o PHP tem de interagir com o mundo web, transformando totalmente os websites que possuem páginas estáticas.” Ainda segundo o autor, seria inviável exibir um website de notícias apenas com HTML, pois as páginas seriam estáticas e a cada nova notícia que aparecesse a página deveria ser alterada manualmente e enviada ao servidor FTP (File Transfer Protocol) para que as novas notícias fossem mostradas. Com o PHP, tudo isso pode ser feito automaticamente com a ajuda de um banco de dados.

Dall’Oglio (2015) diz que ao longo de mais de uma década, o PHP vem adicionando mais e mais recursos e se consolida ano após ano como uma das linguagens de programação orientadas a objetos que mais crescem no mundo. Estima-se que o PHP seja utilizado em mais de 80% dos servidores web existentes, tornando-a disparadamente a linguagem mais utilizada para desenvolvimento web.

Segundo Freitas (2006), o PHP difere-se de outras linguagens, pois seu código é escrito embutido a um arquivo HTML. O que diferencia o PHP do JavaScript no lado do cliente é que o cliente recebe somente a resposta, não tendo acesso ao código que são interpretados no servidor.

2.2.2 HTML (Hyper Text Markup Language)

HTML significa Hyper Text Markup Language, e consiste numa linguagem de formatação de texto que possibilita a construção (programação/formatação) de ficheiros de texto, com a extensão htm ou html, que contêm pequenas marcas (tags) que indicam ao Web browser como apresentar o conteúdo incluído ou referenciado no documento, sob a forma de texto, imagens ou suportes multimídia. (Neves, 2004)

2.2.3 A linguagem JavaScript

Segundo Flanagan (2011), JavaScript é uma linguagem de alto nível, dinâmica, interpretada e não tipada, conveniente para estilos de programação orientados a objetos e funcionais. A sua sintaxe é derivada da linguagem Java, das funções de primeira classe de Scheme e da herança baseada em protótipos de Self.

JavaScript pode interagir com o código HTML, permitindo que autores da Web possam deixar seus sites mais robustos com conteúdo mais dinâmicos. JavaScript é apoiado por uma série de empresas de software e é uma linguagem aberta que qualquer um pode utilizar sem precisar adquirir uma licença.

2.2.4 O banco de dados MySQL

Segundo o Manual de Referência do MySQL, “Um banco de dados é uma coleção de dados estruturados. Ele pode ser qualquer coisa desde uma simples lista de compras a uma galeria de imagens ou a grande quantidade de informação da sua rede corporativa. Para adicionar, acessar, e processar dados armazenados em um banco de dados de um computador, você necessita de um sistema de gerenciamento de bancos de dados como o Servidor MySQL. Como os computadores são muito bons em lidar com grandes quantidades de dados, o

gerenciamento de bancos de dados funciona como a engrenagem central na computação, seja como utilitários independentes ou como partes de outras aplicações.”

“O MySQL 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, mas hoje atendendo a aplicações de grande porte e com mais vantagens que seus concorrentes.” (Milani, 2006)

2.2.5 A ferramenta Bootstrap

Bootstrap é uma ferramenta gratuita para desenvolvimento HTML, CSS e JS. Com ele pode-se criar protótipos rapidamente ou aplicações com sistemas de grid responsivo, componentes pré-construídos e poderosos plugins com jQuery. (BOOTSTRAP, 2020)

O Bootstrap também facilita a criação de projetos responsivos, ou seja, permite que o mesmo website se adapte também para dispositivos móveis sem que seja necessário escrever mais linhas de código.

2.2.6 A biblioteca de Visualização de Dados ChartJS

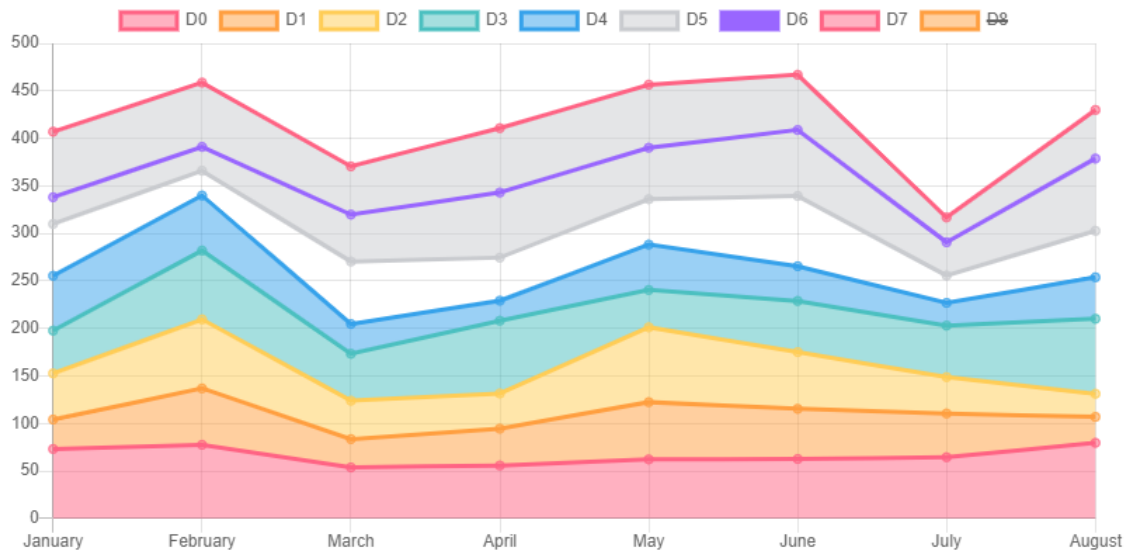
O ChartJS é uma biblioteca em JavaScript que auxilia na criação de gráficos utilizando apenas HTML, CSS e JS para mostrar os gráficos na tela do usuário.

Segundo Rocha (2019), o ChartJS vem com sete tipos básico de gráficos pré-configurados, o que torna mais limitado que algumas outras bibliotecas como o D3.js, porém mais fácil de usar. São eles:

- Barras;
- Linhas;
- Radar;
- Área polar;
- Rosquinha ou pizza;
- Bolhas;
- Dispersão.

Na figura 2 abaixo pode ser conferido um exemplo de Visualização de Dados em um gráfico de linhas apresentado nos exemplos da biblioteca ChartJS.

Figura 2 - Gráfico de linhas – Chart.js Samples



Fonte: ChartJS Samples.

2.2.7 Os formatos de arquivos JSON e GEOJSON

O JSON é uma formatação leve de troca de dados, sendo fácil de ler e escrever para humanos e máquinas, e está baseado em subconjuntos da linguagem de programação JavaScript (Introdução ao JSON, 2020). Um objeto em JSON é formado por pares de nomes e valores e é extremamente leve, e por isso indicado para tratamento de grandes volumes de dados.

GEOJSON é uma variação do formato JSON feito para codificar estruturas de dados geográficos. Utiliza um sistema de referência de coordenadas geográfica, Sistema Geodésico Mundial 1984 e unidades de graus decimais. (GEOJSON, 2020)

2.3 DADOS ABERTOS

Segundo a Open Knowledge International (2020), 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, a exigências que visem preservar sua proveniência e sua abertura.

2.3.1 Portal Brasileiro de Dados Abertos

“O Portal Brasileiro de Dados Abertos é a ferramenta disponibilizada pelo governo para que todos possam encontrar e utilizar os dados e as informações públicas. O portal preza pela simplicidade e organização para que você possa encontrar facilmente os dados e informações que precisa. O portal também tem o objetivo de promover a interlocução entre atores da sociedade e com o governo para pensar a melhor utilização dos dados, promovendo impactos positivos sob os pontos de vista social e econômico.” (Portal Brasileiro de Dados Abertos, 2020)

O Portal Brasileiro de Dados Abertos foi construído pelo governo para centralizar a busca e o acesso dos dados e informações públicas do Brasil, e segue a Lei De Acesso à Informação Pública (Lei 12.527/2011) que regula o acesso e informações detidas pelo governo.

2.3.2 Dados Abertos Educacionais

Segundo Bandeira et al.(2015), os dados abertos no contexto educacional são muito utilizados durante o desenvolvimento de estudos e pesquisas melhoram os processos e as ferramentas utilizadas pela sociedade.

Ainda segundo o autor, os dados educacionais possuem várias aplicações, sendo elas:

- Por gestores educacionais para o planejamento de metas e objetivos a serem alcançados pela educação;
- Para avaliar a efetividade de medidas adotadas no contexto educacional;
- Durante o desenvolvimento de pesquisas;
- Por soluções da indústria que atuam no campo educacional, pois o uso de dados educacionais abertos agrega valor e proporcionam inovação.

2.4 BOLSAS NO ENSINO SUPERIOR BRASILEIRO

As bolsas distribuídas pelo Governo Federal Brasileiro para ensino superior são divididas basicamente em dois programas: o Programa Universidade Para Todos (PROUNI), voltado para cêndidos de baixa renda que busca ingressar em universidade particulares, e o Sistema de Seleção Unificada (SISU), que disponibiliza vagas para instituições públicas de ensino superior. Ambos os programas disponibilizam suas vagas por disputa através do Exame

Nacional do Ensino Médio, prova realizada em todo o Brasil pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). (INEP,2020)

Além das bolsas descritas, também é possível obter financiamento para Instituições Superiores através do Programa de Financiamento Estudantil (FIES), programa disponibilizado pelo MEC para participantes da prova do ENEM.

2.4.1 PROUNI – Programa Universidade Para Todos

O Programa Universidade para Todos (Prouni) do Ministério da Educação é um programa que oferece bolsas de estudo, integrais e parciais (50%), em instituições particulares de educação superior. (PROUNI, 2020)

Para concorrer às bolsas integrais, o estudante deve comprovar renda familiar bruta mensal, por pessoa, de até 1,5 salário mínimo. Para as bolsas parciais (50%), a renda familiar bruta mensal deve ser de até 3 salários mínimos por pessoa.

O programa é destinado a candidatos brasileiros sem diploma de curso superior, que tenham participado da última edição do Exame Nacional do Ensino Médio (Enem), com no mínimo 450 pontos na média das notas e nota na redação que não seja zero.

3 DESENVOLVIMENTO

Neste capítulo serão descritas todas as etapas do processo de obtenção e modelagem dos dados, da criação e população do banco de dados e do desenvolvimento da plataforma WEB responsável por mostrar as visualizações ao usuário. Também serão especificadas neste capítulo as técnicas de visualização escolhidas para cada tipo de dado obtido.

O código completo da aplicação foi disponibilizado no GitHub, e está disponível no endereço <https://github.com/mateusfat/VisualizaPROUNI>.

3.1 OBTENÇÃO E MODELAGEM DOS DADOS

O conjunto de dados utilizado neste trabalho foi obtido no portal de dados abertos do Governo Federal do Brasil (disponível no site: <<http://dados.gov.br>>), onde é possível obter dados de diversos setores públicos do país. Com o objetivo de demonstrar visualizações de dados relevantes foi escolhido e obtido os dados abertos dos últimos cinco anos do PROUNI.

3.1.1 Caracterização dos Dados

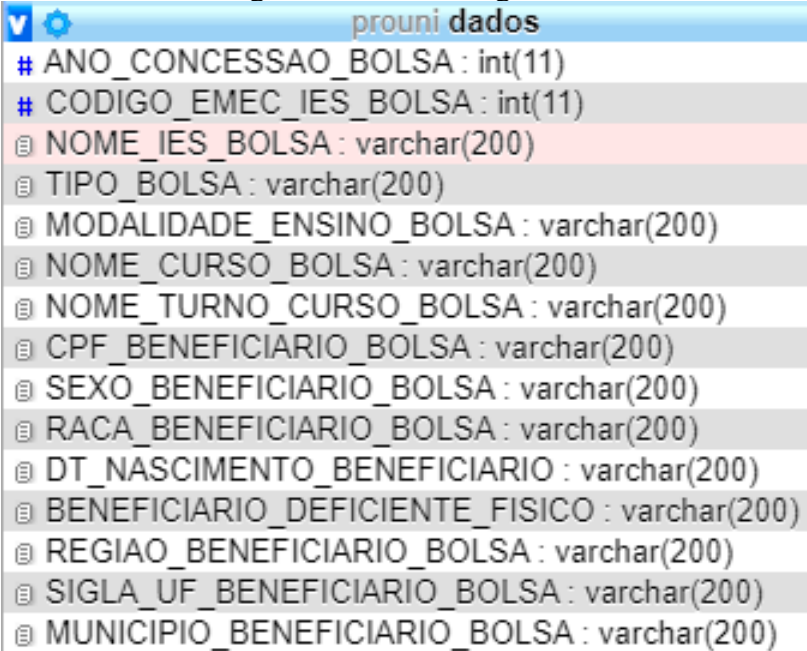
Verificou-se que os dados obtidos estavam em formato CSV, comumente utilizado em planilhas do Microsoft Excel, onde é disposto através de colunas, constando na primeira linha os títulos dos campos. Também é possível perceber que cada arquivo obtido contém apenas um ano de dados do PROUNI, portanto foram necessários cinco arquivos para representar os cinco anos desejados, cada um contendo cerca de 500000 registros de bolsas distribuídas.

Constam nos arquivos disponibilizados dados referentes ao detalhamento quantitativo das bolsas concedidas por ano, segmentadas por região, unidade federativa e município, instituição de educação superior, nome do curso, modalidade de ensino (presencial ou a distância – EAD), turno e tipo de bolsa. Também constam dados referentes aos favorecidos das bolsas, que tiveram os últimos dígitos de seu CPF ocultados por se tratar de dados de acesso aberto.

3.1.2 Modelo Lógico

Com o objetivo de tornar a aplicação extensível e dinâmica verificou-se a necessidade de popular um banco de dados com os dados obtidos e para isso foi realizada a modelagem dos mesmos. Para a inserção em um banco de dados MySQL foi realizado a modelagem lógica apresentada na Figura 3 abaixo, que demonstra as interações e campos presentes no banco, assim como seus tipos de dados. Por se tratar de um banco com uma única tabela, não é possível verificar nenhuma interação no modelo.

Figura 3 – Modelo lógico dos dados



	prouni dados
#	ANO_CONCESSAO_BOLSA : int(11)
#	CODIGO_EMEC_IES_BOLSA : int(11)
#	NOME_IES_BOLSA : varchar(200)
#	TIPO_BOLSA : varchar(200)
#	MODALIDADE_ENSINO_BOLSA : varchar(200)
#	NOME_CURSO_BOLSA : varchar(200)
#	NOME_TURNO_CURSO_BOLSA : varchar(200)
#	CPF_BENEFICIARIO_BOLSA : varchar(200)
#	SEXO_BENEFICIARIO_BOLSA : varchar(200)
#	RACA_BENEFICIARIO_BOLSA : varchar(200)
#	DT_NASCIMENTO_BENEFICIARIO : varchar(200)
#	BENEFICIARIO_DEFICIENTE_FISICO : varchar(200)
#	REGIAO_BENEFICIARIO_BOLSA : varchar(200)
#	SIGLA_UF_BENEFICIARIO_BOLSA : varchar(200)
#	MUNICIPIO_BENEFICIARIO_BOLSA : varchar(200)

Fonte: Elaborado pelo autor (2020).

3.1.3 Script de Criação da Tabela no Banco de Dados

A partir do modelo lógico acima foi criado um *script* responsável pela criação da tabela que será utilizada para a aplicação. O resultado final pode ser conferido na Figura 4 abaixo.

Figura 4 – Script de criação da tabela

```

1  SET SQL_MODE = "NO_AUTO_VALUE_ON_ZERO";
2  SET AUTOCOMMIT = 0;
3  START TRANSACTION;
4  SET time_zone = "+00:00";
5
6
7  DROP TABLE IF EXISTS `dados`;
8  CREATE TABLE IF NOT EXISTS `dados` (
9      `ANO_CONCESSAO_BOLSA` int(11) NOT NULL,
10     `CODIGO_EMEC_IES_BOLSA` int(11) NOT NULL,
11     `NOME_IES_BOLSA` varchar(200) NOT NULL,
12     `TIPO_BOLSA` varchar(200) NOT NULL,
13     `MODALIDADE_ENSINO_BOLSA` varchar(200) NOT NULL,
14     `NOME_CURSO_BOLSA` varchar(200) NOT NULL,
15     `NOME_TURNO_CURSO_BOLSA` varchar(200) NOT NULL,
16     `CPF_BENEFICIARIO_BOLSA` varchar(200) NOT NULL,
17     `SEXO_BENEFICIARIO_BOLSA` varchar(200) NOT NULL,
18     `RACA_BENEFICIARIO_BOLSA` varchar(200) NOT NULL,
19     `DT_NASCIMENTO_BENEFICIARIO` varchar(200) NOT NULL,
20     `BENEFICIARIO_DEFICIENTE_FISICO` varchar(200) NOT NULL,
21     `REGIAO_BENEFICIARIO_BOLSA` varchar(200) NOT NULL,
22     `SIGLA_UF_BENEFICIARIO_BOLSA` varchar(200) NOT NULL,
23     `MUNICIPIO_BENEFICIARIO_BOLSA` varchar(200) NOT NULL
24 ) ENGINE=MyISAM DEFAULT CHARSET=latin1;
25 COMMIT;

```

Fonte: Elaborado pelo autor (2020).

3.2 MÉTODOS DE VISUALIZAÇÃO DE DADOS

Os métodos de visualização utilizadas foram escolhidas conforme as informações contidas nos arquivos e buscando trazer a melhor e mais intuitiva forma de apresentação gráfica ao usuário. Utilizando essencialmente a técnica de Dados Multivariados foram escolhidos quatro tipos diferentes de Visualização de Dados utilizando, sendo cada um deles para uma determinada característica e necessidade dos dados.

3.2.1 Visualização da quantidade de bolsas por estado do Brasil

A visualização da quantidade de bolsas por estado do Brasil apresenta um mapa geográfico do país dividido em unidades federativas (estados), e demonstra a quantidade de bolsas integrais e parciais para cada um deles. O mapa é colorido conforme a quantidade de bolsas e é possível filtrar por quantidade, ocultando então as cores não presentes no filtro selecionado.

3.2.2 Visualização da proporção de bolsas integrais e parciais no Brasil

A visualização da proporção de bolsas integrais e parciais no Brasil apresenta um gráfico de pizza, contendo a proporção entre as bolsas integrais e parciais do PROUNI. É possível verificar a quantidade de cada tipo de bolsa ao posicionar o cursor do mouse sobre o tipo de bolsa para o qual deseja a informação.

3.2.3 Visualização da quantidade de bolsas por região do Brasil

A visualização da quantidade de bolsas por região do Brasil apresenta um gráfico de barras contendo dados das cinco regiões brasileiras e a quantidade de bolsas distribuídas para cada uma delas entre os cinco últimos anos. É possível filtrar, ocultando ou exibindo a região que desejar, facilitando assim a comparação entre regiões e anos.

3.2.4 Visualização por instituição de ensino

A visualização por instituição de ensino apresenta um gráfico de pizza contendo a proporção entre bolsas integrais e parciais da instituição durante os últimos cinco anos. As quantidades podem ser observadas ao posicionar o cursor do mouse sobre os gráficos.

3.3 DESENVOLVIMENTO DA APLICAÇÃO WEB

Para o desenvolvimento da aplicação WEB responsável por mostrar ao usuário final as visualizações gráficas é necessário a utilização de diversas linguagens e ferramentas. Para o banco de dados foi escolhido o *software* MySQL, que traz todas as funcionalidades necessárias

no momento para esta aplicação, e para fazer as conexões com o mesmo serão utilizadas funções em PHP. Já para o desenvolvimento do *front-end*, o framework escolhido foi o Bootstrap, que juntamente com a linguagem de marcação HTML trás muita facilidade para estruturar uma interface agradável.

Os códigos responsáveis por tratar os dados e apresentar as visualizações ao usuário final foram desenvolvidos em JavaScript, linguagem utilizada também pela biblioteca ChartJS, responsável por gerar os gráficos. Durante os capítulos seguintes serão descritas todas as etapas realizadas para o desenvolvimento da aplicação final.

3.3.1 Criação e População do Banco de Dados

A criação do banco de dados necessário se deu através do script apresentado pela Figura 4 - Script de criação da tabela, apresentada na seção 3.1.3, que descreve o script final de criação obtido através do modelo lógico.

Os dados abertos do PROUNI obtidos no Portal Brasileiro de Dados Abertos estavam em formato CSV, e para a inclusão dos mesmos no banco de dados foi necessário rodar um comando SQL diretamente no PHPMyAdmin, software livre responsável por facilitar o acesso aos dados do MySQL como consultas, inclusão, exclusão, criação de base dados e tabelas, e diversas outras operações. O comando visualizado na Figura 5 abaixo foi executado cinco vezes, sendo cada uma delas para um ano de dados contidos em cada arquivo do PROUNI.

Figura 5 – Comando SQL de inclusão dos dados



Fonte: Elaborado pelo autor.

3.3.2 Desenvolvimento da conexão com o banco de dados

Após a criação e população dos dados no banco, viu-se necessário o desenvolvimento de arquivos PHP, responsáveis pela conexão, pela obtenção destes dados e pela decodificação em formato JSON para que sejam aproveitados pela biblioteca ChartJS. Visando facilitar o desenvolvimento dos mesmos, foram criados quatro arquivos diferentes, sendo cada um deles para uma determinada visualização.

O primeiro arquivo, chamado “tipo.php” é responsável por buscar os dados que contenham a proporção entre bolsas integrais e parciais no Brasil. A Figura 6 abaixo demonstra o código final desenvolvido, que trará a contagem da quantidade das bolsas, agrupadas por ano escolhido na aplicação e por tipo de bolsa (integral ou parcial 50%).

Figura 6 – Código responsável pela obtenção dos dados da proporcionalidade de bolsas

```
data > tipo.php
1  <?php header('Content-Type: text/html; charset=UTF-8');
2
3  $mysql = new mysqli('localhost', 'root', '', 'prouni');
4  $mysql->set_charset('utf8');
5
6  $sql = 'SELECT COUNT(*) AS value, TIPO_BOLSA as name, ANO_CONCESSAO_BOLSA as ano FROM `dados` GROUP BY ANO_CONCESSAO_BOLSA, TIPO_BOLSA';
7
8  $data = $mysql->query($sql)->fetch_all(MYSQLI_ASSOC);
9  echo json_encode($data);
10
11  <?>
```

Fonte: Elaborado pelo autor.

O segundo arquivo, chamado “data.php”, contém o código responsável por buscar os dados que contenham a quantidade de bolsas por estado brasileiro e por ano escolhido na aplicação. O código pode ser conferido na Figura 7 abaixo.

Figura 7 – Código responsável pela obtenção dos dados por estado brasileiro

```
data > data.php
1  <?php header('Content-Type: text/html; charset=UTF-8');
2
3  $mysql = new mysqli('localhost', 'root', '', 'prouni');
4  $mysql->set_charset('utf8');
5
6  $sql = 'SELECT SIGLA_UF_BENEFICIARIO_BOLSA AS name, COUNT(*) AS value, ANO_CONCESSAO_BOLSA as ano, TIPO_BOLSA as tipo FROM `dados` GROUP BY SIGLA_UF_BENEFICIARIO_BOLSA, ANO_CONCESSAO_BOLSA, TIPO_BOLSA';
7
8
9  $data = $mysql->query($sql)->fetch_all(MYSQLI_ASSOC);
10 echo json_encode($data);
11
12  <?>
```

Fonte: Elaborado pelo autor.

O terceiro arquivo desenvolvido, chamado “regiao.php”, obtém os dados da quantidade de bolsas por região do Brasil nos últimos cinco anos. A Figura 8 abaixo apresenta o código descrito.

Figura 8 – Código responsável pela obtenção dos dados por região brasileira

```
data > regiao.php
1 <?php header('Content-Type: text/html; charset=UTF-8');
2
3
4 $mysql = new mysqli('localhost', 'root', '', 'prouni');
5 $mysql->set_charset('utf8');
6
7
8 $sql = 'SELECT COUNT(*) AS value, REGIAO_BENEFICIARIO_BOLSA as name, ANO_CONCESSAO_BOLSA as ano FROM `dados` GROUP BY ANO_CONCESSAO_BOLSA, REGIAO_BENEFICIARIO_BOLSA';
9 $data = $mysql->query($sql)->fetch_all(MYSQLI_ASSOC);
10
11 echo json_encode($data);
12
13 >
```

Fonte: Elaborado pelo autor.

O último arquivo, nomeado como “pesquisa.php”, é responsável por obter os dados da quantidade de bolsas por ano e tipo da bolsa da Instituição de Ensino pesquisada pelo usuário na aplicação. Seu código é semelhante aos anteriores, pois apresentará como resultado final as mesmas visualizações, porém somente da instituição pesquisada.

3.3.3 Desenvolvimento do Front-end

Conforme descrito anteriormente, para o desenvolvimento do *front-end* foi utilizado o *framework* Bootstrap que facilita muito a obtenção de uma interface agradável visualmente ao usuário. Para isso, foram criadas duas páginas HTML, sendo uma delas para as visualizações da proporção entre bolsas integrais e parciais e para a visualização do número de bolsas por estado, e outra para a quantidade de bolsas por região brasileira.

Figura 9 – Inclusão dos arquivos do Bootstrap

```
7 <script src="https://ajax.googleapis.com/ajax/libs/jquery/3.5.1/jquery.min.js"></script>
8 <script type="text/javascript" src="https://cdn.jsdelivr.net/npm/echarts/dist/echarts.min.js"></script>
9 <link rel="stylesheet" href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.0/css/bootstrap.min.css"
10 integrity="sha384-9aIt2nRc12Uk9gS9baDl411NQApFmC26EwAOH8WgZl5MYYYxFfc+NcPb1dKGj7Sk" crossorigin="anonymous">
11 <link rel="stylesheet" href="css/style.css" crossorigin="anonymous">
```

Fonte: Elaborado pelo autor.

A figura 9 descreve o código responsável pelo carregamento dos arquivos CSS do Bootstrap, necessários para o desenvolvimento da estilização da aplicação. Já a Figura 10 abaixo apresenta o trecho de código que carrega os arquivos JS responsáveis pelo tratamento dos dados e apresentação das visualizações de dados Proporção de bolsas integrais e parciais no Brasil e Brasil – Número de bolsas PROUNI por estado, contidas na primeira página chamada ‘index.html’.

Figura 10 – Inclusão dos arquivos JS da primeira página

```

75 <script src="js/search.js"></script>
76 <script src="js/script.js"></script>
77 <script src="https://cdn.jsdelivr.net/npm/popper.js@1.16.0/dist/umd/popper.min.js" integrity="sha384-Q6E9RHvbIyZFJoft+2mJbHaEWldlvI9IOYy5n3zV9zzTtmI3UksdQRVvoxMfooAo"
78 crossorigin="anonymous"></script>
79 <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.5.0/js/bootstrap.min.js" integrity="sha384-OgVRvuATP1z7JjHLku0U7W704+h835Lr+6QL9UvYjZE3Ipu6Tp75j7Bh/kR0JkI"
80 crossorigin="anonymous"></script>

```

Fonte: Elaborado pelo autor.

A figura 11 abaixo descreve a inclusão dos arquivos JS responsáveis pela visualização de dados Quantidade de Bolsas por Região do Brasil apresentada na segunda página, chamada “regiao.html”.

Figura 11 – Inclusão dos arquivos JS da segunda página

```

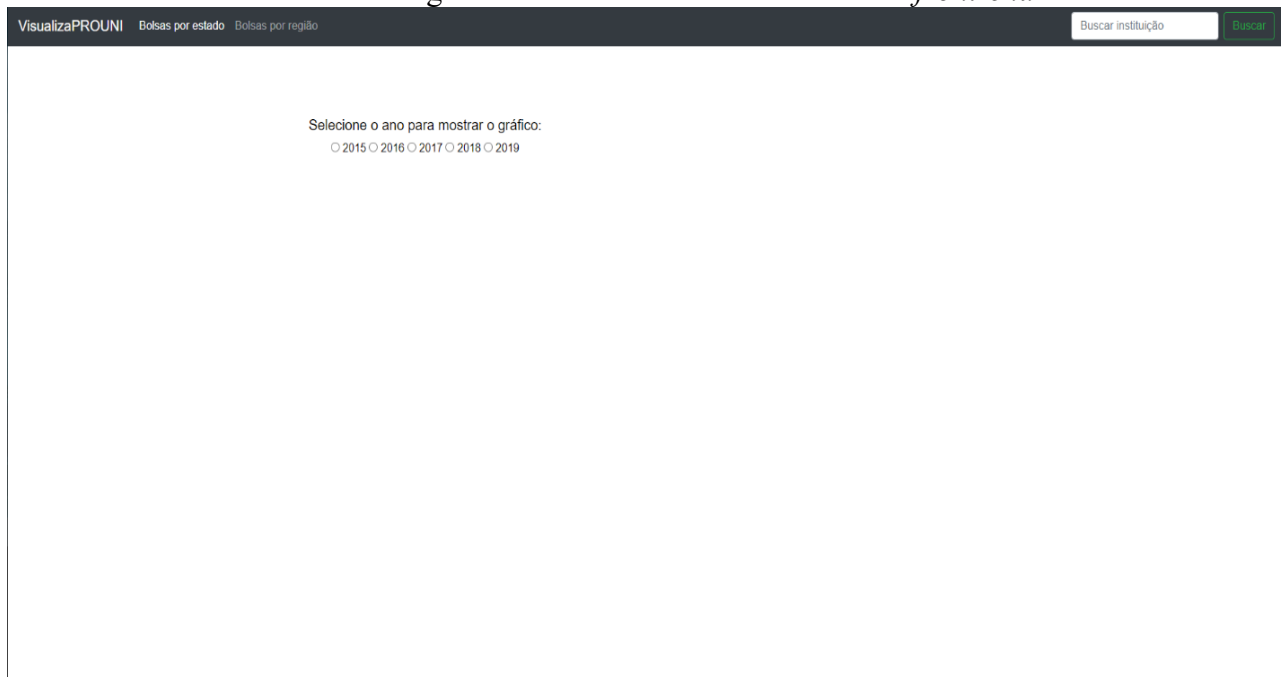
42 <script src="js/search.js"></script>
43 <script src="js/script3.js"></script>
44 <script src="https://cdn.jsdelivr.net/npm/popper.js@1.16.0/dist/umd/popper.min.js" integrity="sha384-Q6E9RHvbIyZFJoft+2mJbHaEWldlvI9IOYy5n3zV9zzTtmI3UksdQRVvoxMfooAo"
45 crossorigin="anonymous"></script>
46 <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.5.0/js/bootstrap.min.js" integrity="sha384-OgVRvuATP1z7JjHLku0U7W704+h835Lr+6QL9UvYjZE3Ipu6Tp75j7Bh/kR0JkI"
47 crossorigin="anonymous"></script>

```

Fonte: Elaborado pelo autor.

O resultado final da interface desenvolvida pode ser observado na Figura 12, onde o usuário poderá navegar entre os tipos de visualização e pesquisar por uma instituição no menu superior, ou escolher um determinado ano para apresentar as visualizações “Visualização da proporção de bolsas integrais e parciais no Brasil” e “Visualização da quantidade de bolsas por região do Brasil”.

Figura 12 – Resultado final da interface *front-end*



Fonte: Elaborado pelo autor.

3.3.4 Desenvolvimento das Visualizações de Dados

A primeira visualização de dados, intitulada de “Proporção de Bolsas Integrais e Parciais no Brasil”, foi desenvolvida com o intuito de trazer informações pertinentes a proporcionalidade entre os tipos de bolsas distribuídas em nosso país, sendo eles integrais ou parciais 50%. A visualização desenvolvida apresenta um gráfico do tipo pizza contendo o número e a porcentagem dos tipos de bolsas, e o trecho de código responsável pela geração da mesma pode ser conferido na Figura 13 abaixo.

Figura 13 – Trecho de código da visualização da proporcionalidade de bolsas

```

128     option2 = {
129         title: {
130             text: 'Proporção de bolsas Integrais e Parciais no Brasil',
131             subtext: 'Dados obtidos em dados.gov.br/dataset/mec-prouni/',
132             sublink: 'http://www.dados.gov.br/dataset/mec-prouni/',
133             left: 'center'
134         },
135         tooltip: {
136             trigger: 'item',
137             formatter: '{a} <br/>{b} : {c} ({d}%)'
138         },
139         legend: {
140             orient: 'vertical',
141             left: 'left',
142             data: ['Integral', 'Parcial']
143         },
144         series: [
145             {
146                 name: 'Quantidade e proporção das bolsas no país',
147                 type: 'pie',
148                 radius: '55%',
149                 center: ['50%', '60%'],
150                 data: filtrados2,
151                 emphasis: {
152                     itemStyle: {
153                         shadowBlur: 10,
154                         shadowOffsetX: 0,
155                         shadowColor: 'rgba(0, 0, 0, 0.5)'
156                     }
157                 }
158             }
159         ]
160     };
161     myChart2.setOption(option2);
162 });
163

```

Fonte: Elaborado pelo autor.

A segunda visualização desenvolvida foi chamada de “Brasil – Quantidade de Bolsas PROUNI por estado” e tem como objetivo apresentar informações sobre o número de bolsas

distribuídas em cada estado brasileiro em cada um dos últimos cinco anos. Para o desenvolvimento desta visualização também foi utilizado um arquivo GEOJSON contendo dados geográficos de cada estado brasileiro, necessários para que o ChartJS apresente o mapa do país. O trecho de código contido na Figura 14 abaixo apresenta uma parte do desenvolvimento da visualização.

Figura 14 – Trecho de código da visualização de bolsas por estado

```

261 echarts.registerMap('Brasil', geoJson);
262 option = {
263   title: {
264     text: 'Brasil - Número de bolsas PROUNI por estado',
265     subtext: 'Dados obtidos em dados.gov.br/dataset/mec-prouni/',
266     sublink: 'http://www.dados.gov.br/dataset/mec-prouni/',
267     left: 'center'
268   },
269   tooltip: {
270     trigger: 'item',
271     showDelay: 0,
272     transitionDuration: 0.2,
273     formatter: function (params) {
274       var value = (params.value + '').split('.');
275       value = value[0].replace(/(\d{1,3})(?=(?:\d{3})+(?!\d))/g, '$1,');
276       return 'Número de Bolsas PROUNI: ' + '<br/>' + params.name + ': ' + value;
277     }
278   },
279   visualMap: {
280     left: 'right',
281     min: 300,
282     max: 61000,
283     inRange: {
284       color: ['#52569F', '#74ADD1', '#DFB277', '#D73027', '#A50026']
285     },
286     text: ['Maior', 'Menor'],
287     calculable: true
288   },
289   toolbox: {
290     show: false,
291     orient: 'vertical',
292     left: 'left',
293     top: 'top',
294     feature: {
295       dataView: {readOnly: false},
296       restore: {},
297       saveAsImage: {}
298     }
299   },
300   series : [
301     {
302       name: 'Brasil',
303       type: 'map',
304       roam: true,
305       map: 'Brasil',
306       itemStyle:{
307         emphasis:{label:{show:true}}
308       },
309       textFixed : {
310       },
311       data: filtrados
312     }
313   ]
314 };
315 myChart.setOption(option);

```

Fonte: Elaborado pelo autor.

A terceira visualização, chamada de “Quantidade de Bolsas por região do Brasil”, traz informações sobre o número de bolsas distribuídas por região do Brasil nos últimos cinco anos de PROUNI. O código final pode ser conferido na Figura 15 abaixo.

Figura 15 – Trecho de código da visualização por região brasileira

```

30 //Gráfico por região
31 $.getJSON('data/regiao.php').done(function (dados) {
32     data3 = dados;
33     const filtrados3 = data3.filter(dado =>{
34         return dado.ano === anoRecebido;
35     });
36     console.log(filtrados3);
37
38     myChart3.hideLoading();
39
40     option3 = {
41         title: {
42             text: 'Quantidade de bolsas por região do Brasil',
43             subtext: 'Dados obtidos em dados.gov.br/dataset/mec-prouni/',
44             sublink: 'http://www.dados.gov.br/dataset/mec-prouni/',
45             left: 'center'
46         },
47         color: ['#3398DB'],
48         tooltip: {
49             trigger: 'axis',
50             axisPointer: {
51                 type: 'shadow'
52             }
53         },
54         grid: {
55             left: '3%',
56             right: '4%',
57             bottom: '3%',
58             containLabel: true
59         },
60         xAxis: [
61             {
62                 type: 'category',
63                 data: ['Centro-Oeste', 'Nordeste', 'Norte', 'Sudeste', 'Sul'],
64                 axisTick: {
65                     alignWithLabel: true
66                 }
67             }
68         ],
69         yAxis: [
70             {
71                 type: 'value'
72             }
73         ],
74         series: [
75             {
76                 name: 'Quantidade',
77                 type: 'bar',
78                 barWidth: '70%',
79                 data: [filtrados3[0], filtrados3[1], filtrados3[2], filtrados3[3], filtrados3[4]]
80             }
81         ]
82     };
83
84     myChart3.setOption(option3, true);
85

```

Fonte: Elaborado pelo autor.

A última visualização, proveniente da pesquisa por instituição de ensino, apresentará a quantidade de bolsas distribuídas para aquela instituição buscada, assim como um gráfico do tipo pizza contendo a proporção entre os tipos de bolsas dos últimos cinco anos.

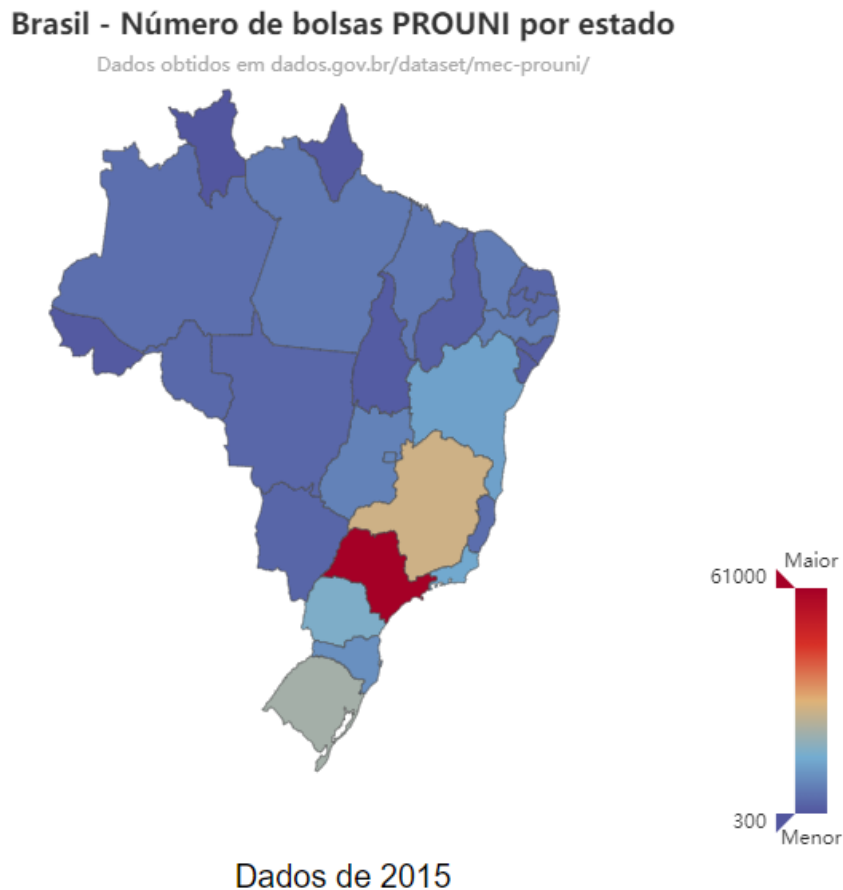
4 RESULTADOS

Como resultado final da aplicação Web desenvolvida com o auxílio da biblioteca ChartJS buscou-se trazer visualizações relevantes ao usuário para a análise dos dados abertos educacionais do PROUNI. Para isso foram desenvolvidas quatro visualizações, as quais serão demonstradas neste capítulo.

4.1 VISUALIZAÇÃO DA QUANTIDADE DE BOLSAS POR ESTADO DO BRASIL

Esta visualização buscou mostrar a quantidade de bolsas por estado brasileiro, sendo assim possível tirar conclusões sobre quais estados possuem a maior ou a menor quantidade de bolsas do país. Para isso, ela apresenta um mapa geográfico brasileiro colorido gradualmente conforme a quantidade de bolsas de cada estado, sendo azul a menor quantidade e vermelho, a maior. O resultado da visualização pode ser conferido na Figura 16.

Figura 16 – Resultado da Visualização da Quantidade de Bolsas por Estado do Brasil



Fonte: Elaborado pelo autor.

A visualização também permite que usuário verifique o número exato de bolsas PROUNI de cada estado brasileiro, basta que o mesmo posicione o cursor do mouse sobre o local desejado. A figura 17 apresenta o resultado ao posicionar o mouse sobre o estado de São Paulo, apresentando o total de 68485 bolsas distribuídas naquele ano.

Figura 17 – Quantidade de bolsas em São Paulo



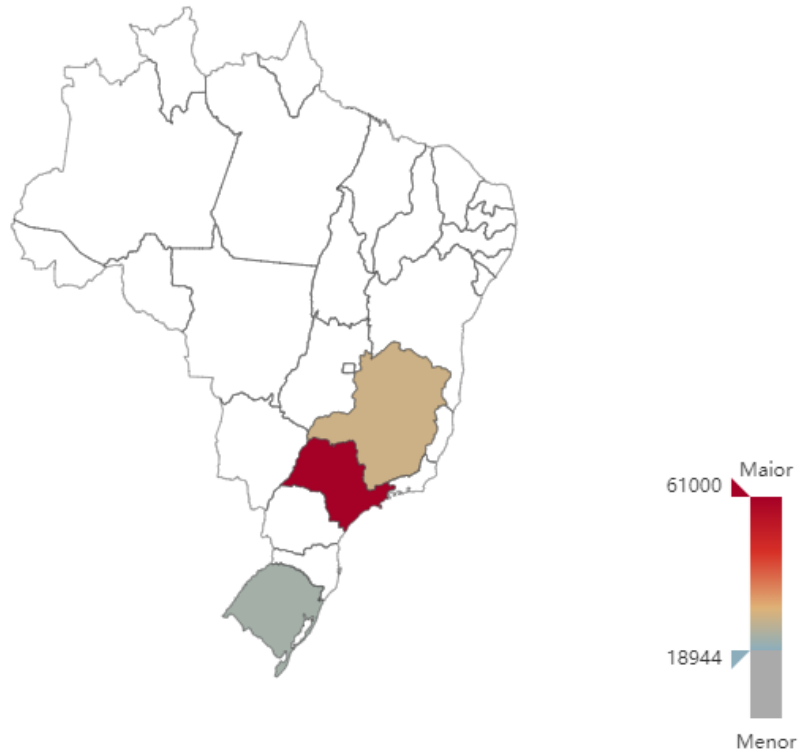
Fonte: Elaborado pelo autor.

Também é possível filtrar uma determinada faixa de quantidade de bolsas através da legenda na lateral direita inferior do gráfico, para que então somente os estados contidos nesta faixa de valores sejam apresentados coloridos no mapa. Confira a Figura 18, que demonstra os estados contidos na faixa de valores de 18944 bolsas até 61000 bolsas no ano de 2015.

Figura 18 – Filtro por quantidade de bolsas por estado

Brasil - Número de bolsas PROUNI por estado

Dados obtidos em dados.gov.br/dataset/mec-prouni/



Dados de 2015

Fonte: Elaborado pelo autor.

4.2 VISUALIZAÇÃO DA PROPORÇÃO DE BOLSAS INTEGRAIS E PARCIAIS NO BRASIL

Buscando demonstrar a proporção entre bolsas integrais e parciais (50%) em nosso país foi desenvolvida esta visualização do tipo pizza. Nela o usuário pode verificar a quantidade de bolsas e a porcentagem de cada tipo ao posicionar o cursor do mouse sobre o dado desejado. O resultado da visualização pode ser visto na Figura 19.

Figura 19 – Visualização da Proporção de bolsas integrais e parciais no Brasil



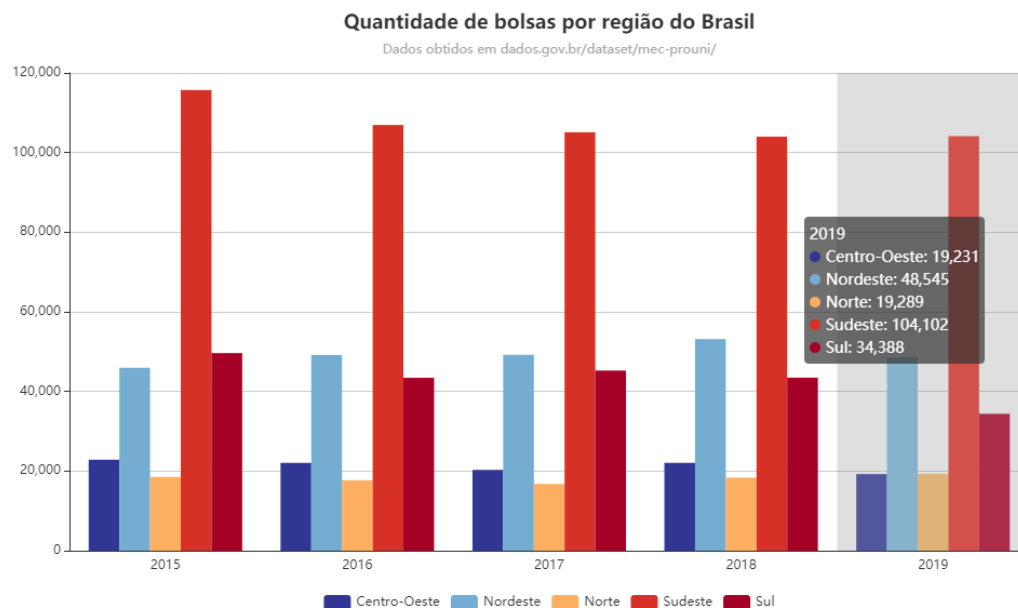
Fonte: Elaborado pelo autor.

4.3 VISUALIZAÇÃO DA QUANTIDADE DE BOLSAS POR REGIÃO DO BRASIL

Esta visualização busca mostrar ao usuário a quantidade de bolsas por região brasileira durante os últimos cinco anos em forma de gráfico de barras. No gráfico, cada região é colorida com uma determinada cor e separada por ano, facilitando a comparação entre as mesmas.

Igualmente às demais visualizações, o número exato de bolsas por região e por ano pode ser conferido ao posicionar o cursor do mouse sobre o ano desejado. A visualização pode ser vista na Figura 20.

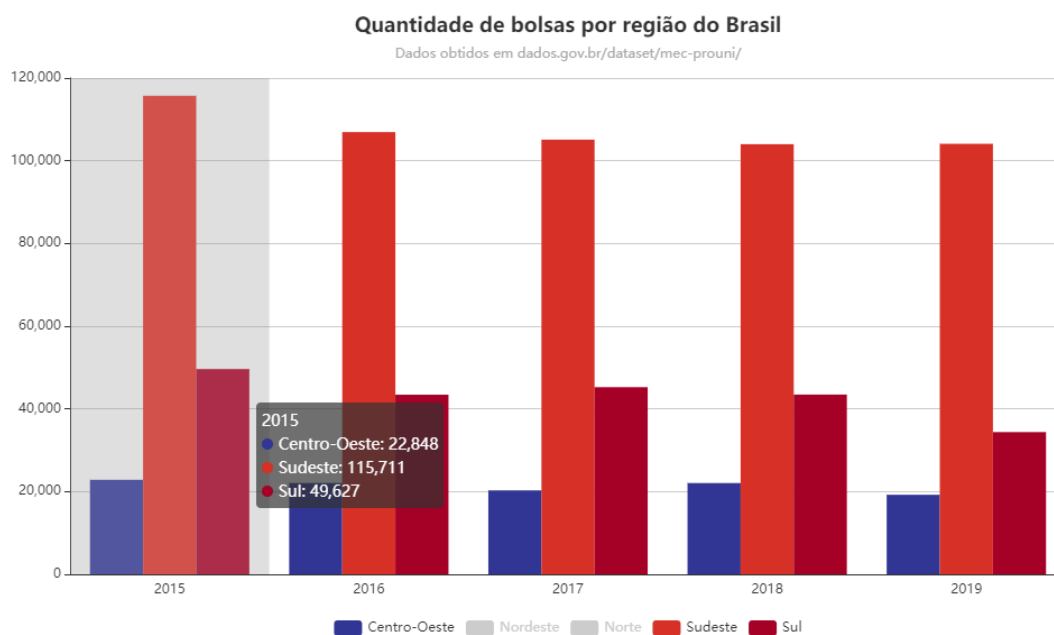
Figura 20 – Visualização da Quantidade de bolsas por região do Brasil



Fonte: Elaborado pelo autor.

Ainda nesta visualização, o usuário pode filtrar as regiões que deseja, basta que o mesmo clique na região que deseja ocultar/mostrar. A figura 21 demonstra o resultado após ocultar as regiões Nordeste e Norte.

Figura 21 – Filtro por região do Brasil

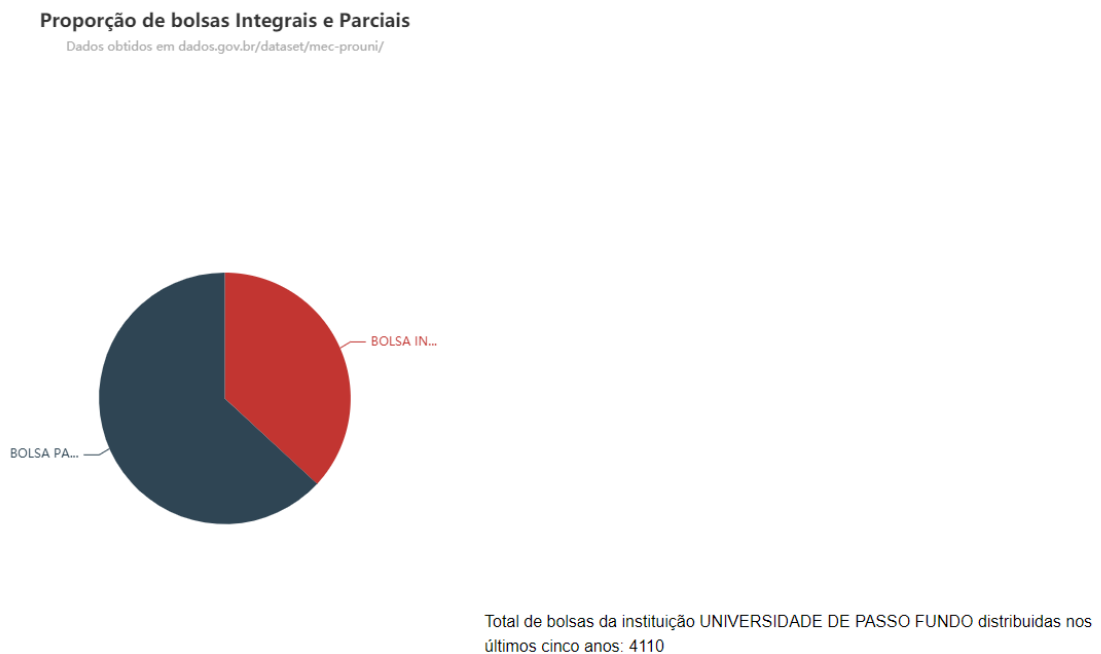


Fonte: Elaborado pelo autor.

4.4 VISUALIZAÇÃO POR INSTITUIÇÃO DE ENSINO

A visualização por instituição de ensino mostra a proporção entre os tipos de bolsa e a quantidade total de bolsas distribuídas pela instituição pesquisada pelo usuário nos últimos cinco anos de PROUNI. Para que a visualização e os dados sejam mostrados, é necessário que o usuário preencha o campo de pesquisa localizado na parte superior direita da interface da aplicação, e caso informe uma instituição inexistente no banco de dados será apresentado o erro “Nenhuma instituição de ensino com esse nome encontrada.”. O resultado pode ser conferido na Figura 22.

Figura 22 – Visualização por instituição de ensino



Fonte: Elaborado pelo autor.

5 CONCLUSÃO

A aplicação desenvolvida atingiu seu objetivo, que era desenvolver uma aplicação Web para visualização dos dados abertos do PROUNI utilizando a ferramenta ChartJS, e também demonstrou a importação da visualização de dados para a análise dos dados. Sem a mesma não seria possível obter informações como os estados com maior quantidade de bolsas, a proporção entre os tipos de bolsas ou então as regiões com o maior número de bolsas distribuídas.

Nas visualizações apresentadas neste trabalho, as quais utilizaram técnicas de Visualização para Dados Multivariados, é possível verificar que o estado com maior quantidade de bolsas do país é o estado de São Paulo, com 68.485 bolsas em 2015, 64.982 em 2016, 64.159 em 2017, 60.713 em 2018 e 59.441 em 2019. Também se verifica que o estado com a menor quantidade de bolsas é o de Roraima com apenas 373 em 2015, 487 em 2016, 320 em 2017, 344 em 2018 e 326 em 2019.

A proporção entre as bolsas integrais e parciais se manteve parecida durante os cinco anos, sendo a maioria das bolsas do tipo integral com algo em torno de 73,26%, representando 185.076 das bolsas no ano de 2015 contra 67.559 (26,74%) de bolsas parciais.

As regiões Sudeste e Nordeste apresentam as maiores quantidades de bolsas do país, e mantiveram esta posição durante os cinco anos de dados apresentados. A região com o menor número de bolsas distribuídas foi o do Norte, com números bem parecidos com a região Centro-Oeste.

A aplicação cumpre com o seu propósito e permite a análise dos dados obtidos de forma a obter relevantes conclusões, algo que dificilmente seria possível somente observando a planilha de dados originais contidas no Portal Brasileiro de Dados Abertos. A aplicação também permite que o usuário inclua os anos seguintes de dados no banco de dados, fazendo com que a mesma seja escalável.

Como sugestões de trabalhos futuros surge a ideia do desenvolvimento de mais visualizações, aproveitando o restante das variáveis presentes nos dados e a opção de incluir os arquivos dos anos seguintes diretamente pela aplicação, sem a necessidade da utilização de um comando SQL ou da ferramenta PHPMyAdmin.

REFERÊNCIAS

KEIM, D., WARD, M. **Visual Data Mining Techniques**, Johnson/Hansen: **The Visualization Handbook**, 2003.

FREITAS, C. M. D. S. et al. **Introdução à visualização de informações**. **Revista de informática teórica e aplicada**. Porto Alegre. Vol. 8, n. 2 (out. 2001), p. 143-158, 2001.

Card, Mackinlay, Shneiderman. **Readings in information visualization: using vision to think**. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1999.

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

Few, S. **Now you see it: simple visualization techniques for quantitative analysis**. Analytics Press, 2009.

D3 - Data-Driven Documents. 2019. Disponível em: <<https://d3js.org/>> Acesso em: 10 nov. 2019.

Chart.js Bar Chart - Stacked. Disponível em: <<https://www.chartjs.org/samples/latest/charts/bar/stacked-group.html>> Acesso em 11 out. 2020.

Niederauer, J. **Desenvolvendo Websites com PHP**. Novatec Editora. 2 ed. São Paulo, 2011.

Milani, A. **MySQL: guia do programador**. Novatec Editora. São Paulo, 2006.

Flanagan, D. **JavaScript: The Definitive Guide**. O'Reilly Media, Inc., Sebastopol, 2011.

Dall'Oglio, P. **PHP Programando com Orientação a Objetos**. Novatec Editora. 3 ed. São Paulo, 2015.

MySQL 8.0 Reference Manual. Disponível em <<https://dev.mysql.com/doc/refman/8.0/en/>> Acesso em 12 out. 2020.

Neves, P. M. C. **O Guia Prático da HTML**. Centro Atlântico Lda, 1 ed. Lisboa, 2004.

Bootstrap Documentação. Disponível em <<https://getbootstrap.com.br/docs/4.1/getting-started/introduction/>> Acesso em 24 out. 2020.

Sobre o dados.gov.br. Disponível em <<https://dados.gov.br/pagina/sobre>> Acesso em 24 out. 2020.

Open Knowledge Foundation. Disponível em <<https://okfn.org/>> Acesso em 24 out. 2020.

Introdução ao JSON. Disponível em <<https://www.json.org/json-pt.html>> Acesso em 26 out. 2020.

GEOJSON. Disponível em <<https://geojson.org>> Acesso em 26 out. 2020.

PROUNI. Disponível em <<http://prouniportal.mec.gov.br/>> Acesso em 22 nov. 2020.

INEP. Disponível em <<http://portal.inep.gov.br/>>. Acesso em 09 dez. 2020.