

# Um estudo focado ao PROUNI através da análise de dados abertos: período de 2005 até 2016

*A study focused on PROUNI by open data analysis: period from 2005 to 2016*

**Vanderlei Benedito da Silva Filho**

Campus Bragança Paulista, Instituto Federal de São Paulo  
[vanderleibsf@gmail.com](mailto:vanderleibsf@gmail.com)

**Letícia Souza Netto Brandi**

Campus Bragança Paulista, Instituto Federal de São Paulo  
[prof.leticia@ifsp.edu.br](mailto:prof.leticia@ifsp.edu.br) | [fsn.leticia@gmail.com](mailto:fsn.leticia@gmail.com)

## Resumo

O número de informações geradas em todo o mundo tem aumentado exponencialmente, muitos dos dados gerados têm sido disponibilizados seguindo os princípios de dados abertos, a análise destes dados permite que eles possam ser utilizados de forma benéfica, garantindo transparência, participação e colaboração. Neste contexto, através de uma pesquisa descritiva, este artigo tem como objetivo a utilização da ferramenta de visualização de dados *Tableau Public* para gerar gráficos interativos possibilitando analisar os dados referentes ao período entre 2005 e 2016 do Programa Universidade Para Todos – PROUNI, disponibilizado pelo Governo Federal Brasileiro através de seu portal de dados abertos. A análise dos gráficos gerados permitiu verificar o desempenho do programa desde sua criação, traçar o perfil de seus beneficiários e realizar um comparativo com as informações da população brasileira.

**Palavras-chave:** Dados abertos, Visualização de dados, Governo Eletrônico, PROUNI.

## Abstract

*The number of information generated around the world has increased exponentially, many data generated are made available to respect the open data principles. These data analysis can be used in a beneficial way, to guarantee transparency, participation and collaboration. In this context, this paper, which is based in a descriptive research, aims to use data visualization tool *Tableau Public* to generate interactive graphs that makes it possible to analyse the data of the University for All Program - PROUNI between 2005 and 2016, which is available on open data portal of the Brazilian federal government. The graphs analysis allowed to verify the program performance since its creation, draw the profile of its beneficiaries and make a comparison with information about Brazilian people.*

**Keywords:** *Open Data, Data Visualization, Electronic Government, PROUNI.*

## 1. Introdução

Para Davenport e Patil (2012) saber trabalhar com um grande número de dados vindos de diversos locais e com diferentes formatos é uma das habilidades mais desejadas na última década. Isso ocorre pelo aumento exponencial dos dados gerados pela sociedade e a necessidade de minerar as informações obtidas.

No entanto, estes dados, em sua maioria, não estão estruturados para facilitar sua compreensão mesmo por aqueles que podem acessá-los e manipulá-los. Desta forma, a extração de informações e produção de conhecimentos que poderiam ser úteis para sociedade não ocorrem com a agilidade e eficácia necessárias (Isotani e Bittencourt, 2015).

Os governos produzem muita informação, se houver boa utilização das tecnologias existentes e iniciativas da sociedade, os dados governamentais poderão ser cada vez mais benéficos a todos, garantindo transparência, participação e colaboração (CGI, 2011). Por exemplo, o Programa Universidade para Todos – PROUNI, concede bolsas para estudantes brasileiros que não possuem diploma de nível superior. O Governo Federal brasileiro disponibiliza os dados do perfil dos beneficiários deste programa em seu portal de dados abertos. Apesar dos dados estarem disponíveis não há uma análise minuciosa sobre eles.

Este trabalho tem como objetivo geral utilizar um *software* de visualização de dados para gerar gráficos interativos a partir da base de dados do PROUNI e através destes gráficos analisar a concessão das bolsas desde sua criação, a fim de verificar quais estudantes estão sendo atendidos pelo programa. Além do objetivo geral há objetivos específicos listados a seguir:

- a) Comparar o número de bolsas concedidas por raça/cor, sexo e deficiência em relação ao total da população brasileira;
- b) Distinguir as regiões com mais beneficiários;
- c) Apresentar o tipo de bolsa mais procurada, assim como a modalidade de ensino e o turno;
- d) Identificar as instituições e cursos que possuem um maior número de bolsistas;
- e) Verificar se houve crescimento do programa.

## 2. Fundamentação teórica

### 2.1. Dados abertos

A atividade, conectividade e o crescimento de serviços na *Web* aumentaram consideravelmente a quantidade de informações disponíveis, exigindo meios mais complexos e rápidos para acessar, pesquisar, classificar e fornecer informações (CGI, 2009).

A *Open Definition* (2018) define dados abertos como dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer indivíduo. São sujeitos, no máximo, à atribuição à fonte original e compartilhamento com as mesmas licenças em que foram apresentados, permitindo que pessoas físicas e jurídicas os explorem. O Manual dos Dados Abertos: Desenvolvedores, lista alguns princípios que caracterizam os dados abertos, conforme CGI (2011):

- a) **acesso livre:** qualquer pessoa deverá ter acesso aos documentos, sem discriminação de grupos ou pessoa;
- b) **separar os dados ao máximo:** os dados devem ser separados ao máximo, em estruturas distintas;
- c) **responsabilidade:** deve-se promover a responsabilidade, para que os dados possam ser publicados com qualidade e da maneira mais rápida e segura;
- d) **rápida integração:** deve-se oferecer ferramentas e especificações técnicas para que as integrações sejam realizadas de forma rápida e constante;
- e) **compartilhamento de boas práticas:** para que exista uma rápida integração, é necessário o compartilhamento de informações e experiências relacionadas a boas práticas;
- f) **formatos de arquivos:** as informações devem estar em documentos em formatos de fácil manipulação por humanos e por máquinas, formatos proprietários podem criar dependência tecnológica para o uso das informações gerando restrições ao acesso dos dados;
- g) **serialização de dados:** não se deve disponibilizar um único formato aberto de arquivos, isso prejudica a utilização por um grupo de pessoas, por falta de conhecimento ou por falta de estruturação para manipular os arquivos.

## 2.2. Portal Brasileiro de dados abertos

A disseminação das tecnologias de informação e comunicação (TIC) tem provocado diversas mudanças sociais, alterando hábitos e expectativas, há um acesso cada vez maior a um volume gigantesco de dados. Isso fez com que, nas duas últimas décadas, a administração pública no Brasil recebesse pressões fortes e contínuas da sociedade visando melhorias de desempenho (CGI, 2018).

De acordo com o CGI (2009) desde o final da década de 1990 os governos buscam encontrar formas melhores de manter contato com os cidadãos pela *Web*. Ao disponibilizar informações do governo de forma *on-line* e que sejam acessíveis, compreensíveis e utilizáveis, os cidadãos podem interagir com seu governo, constituindo o chamado governo eletrônico ou e-Governo.

Do ponto de vista do Executivo Federal Brasileiro, há pelo menos oito anos, estão sendo desenvolvidas iniciativas de inovação tecnológica para a participação social (Filho e Martins, 2017). Uma das ações do Governo Brasileiro foi a criação o Portal Brasileiro de Dados Abertos ([dados.gov.br](http://dados.gov.br)), uma ferramenta para que todos possam encontrar e utilizar os dados e as informações públicas, seguindo os princípios de dados abertos. Os dados abertos governamentais fazem parte da política da lei de acesso à informação do governo federal, sancionada em 2011. O portal objetiva disponibilizar dados relativos às diversas temáticas da administração pública, como dados do sistema de transporte, de segurança pública, indicadores de educação, gastos governamentais, processo eleitoral, entre outros. Os dados abertos podem ser encontrados com facilidade, porém o seu uso requer um nível básico de conhecimento técnico. (Brasil, 2018).

## 2.3. PROUNI

Criado pelo Governo Federal brasileiro em 2004 e institucionalizado pela Lei nº 11.096, de 13 de janeiro de 2005, o Programa Universidade para Todos - PROUNI é um programa do Ministério da Educação

que concede bolsas de estudo integrais e parciais em cursos de graduação e sequenciais de formação específica, em instituições de ensino superior privadas. Podem participar do PROUNI os estudantes brasileiros que não possuam diploma de curso superior e que atendam a pelo menos uma das condições a seguir:

- a) ter cursado o ensino médio completo em escola da rede pública;
- b) ter cursado o ensino médio completo em escola da rede privada, na condição de bolsista integral da própria escola;
- c) ter cursado o ensino médio parcialmente em escola da rede pública e parcialmente em escola da rede privada, na condição de bolsista integral da própria escola privada;
- d) ser pessoa com deficiência;
- e) ser professor da rede pública de ensino, no efetivo exercício do magistério da educação básica e integrando o quadro de pessoal permanente da instituição pública e concorrer a bolsas exclusivamente nos cursos de licenciatura. Nesses casos não há requisitos de renda.

Para concorrer às bolsas integrais a renda familiar bruta mensal do candidato deve ser de até um salário mínimo e meio por pessoa. Para as bolsas parciais de 50%, essa renda deve ser de até três salários mínimos por pessoa.

No segundo semestre de 2008 foram concedidas bolsas complementares de 25% sendo os 75% restantes financiados pelo FIES (Fundo de Financiamento Estudantil).

Há um processo seletivo composto por duas fases:

- a) **processo regular:** o candidato deve ter participado da edição do Exame Nacional do Ensino Médio (Enem) do ano imediatamente anterior e ter obtido, ao menos, 450 pontos na média das notas das provas e nota acima de 0 na redação;
- b) **processo de ocupação das bolsas remanescentes:** pode-se inscrever o candidato que seja professor da rede pública de ensino, no efetivo exercício do magistério da educação básica e integrando o quadro de pessoal permanente da instituição pública, para os cursos com grau de licenciatura destinados à formação do magistério da educação básica; ou tenha participado do Enem, a partir da edição de 2010, e tenha obtido, em uma mesma edição, média das notas nas provas igual ou superior a 450 pontos e nota superior a 0 na redação.

São realizados dois processos seletivos por ano, em ambos os processos as inscrições são gratuitas e efetuadas exclusivamente pela internet, por meio da página do PROUNI (PROUNI, 2018).

### 3. Materiais e métodos

Segundo Oliveira (2011, p. 28) a abordagem qualitativa da pesquisa é caracterizada como:

*Sendo um processo de reflexão e análise da realidade através da utilização de métodos e técnicas para compreensão detalhada do objeto de estudo em seu contexto histórico e/ou segundo uma estruturação. Esse processo implica em estudos segundo a literatura pertinente ao tema, observações, aplicação de questionários, entrevistas e análise de dados, que deve ser apresentado de forma descritiva.*

Portanto este trabalho buscou descobrir a natureza, características, causas e relações dos dados estudados, caracterizando uma pesquisa descritiva (Barros e Lehfeld, 2007), neste tipo de pesquisa se observa, registra, analisa e correlaciona os fatos sem manipulá-los, possuindo como tarefa característica a coleta de dados (Cervo *et al.*, 2007).

Diante disso, o trabalho analisou a base de dados disponibilizada pelo Ministério da Educação no Portal Brasileiro de Dados Abertos referente às bolsas concedidas e perfil dos beneficiários do PROUNI. Esta base foi publicada em formato CSV (*Comma-Separated Values*), definido pela Cartilha Técnica para Publicação de Dados Abertos no Brasil v1, publicada pela Secretaria de Logística e Tecnologia da Informação – SLTI do Ministério do Planejamento Orçamento e Gestão – MP (Brasil, 2012), como um formato para armazenamento de dados tabulares em texto. Sua codificação é muito simples: cada linha do arquivo representa uma linha na tabela, e as colunas são separadas por vírgula. CSV é recomendado para representação de estrutura de dados mais simples, de natureza tabular, onde não existem subpropriedades ou listas, gerando um arquivo menor e mais leve para processamento. Os arquivos CSV podem ser abertos e manipulados em qualquer editor de planilhas, mesmo por uma pessoa que não tenha conhecimentos de programação.

A base possui dados de 2005 a 2016 detalhando o quantitativo das bolsas concedidas pelo PROUNI por ano, segmentadas por região, unidade federativa, município, instituição de educação superior, curso, modalidade de ensino (presencial ou à distância), turno e tipo de bolsa. Detalhamento do perfil dos beneficiários do PROUNI por sexo, raça/cor, faixa etária e pessoas com deficiência.

Para analisar a base de dados do PROUNI foi utilizado a ferramenta de visualização de dados *Tableau Desktop*, na sua versão gratuita (*Public Edition*). Com esta ferramenta é possível criar e compartilhar rapidamente gráficos interativos, mapas e painéis em tempo real, sendo de fácil uso para qualquer pessoa (*Tableau*, 2018). O processo de transformação de dados analíticos em dados visuais torna possível a compreensão e interpretação destes dados por agentes humanos (Andrade *et al.*, 2017), o que auxilia na tomada de decisões importantes (LEA *et al.*, 2018). A visualização de dados de forma interativa permite a leitura das informações, podendo filtrar os dados que se deseja visualizar, além de navegar em textos complementares (PEDROZA *et al.*, 2013).

#### 4. Preparação dos dados

Visando alcançar resultados consistentes, os dados estudados passaram inicialmente por um processo de limpeza e organização. Vasco (2015) define a limpeza dos dados como uma tarefa cujo objetivo é detectar e corrigir erros e inconsistências nos dados, de forma a melhorar a qualidade da informação. Ainda segundo o autor as anomalias encontradas em bancos de dados podem ser divididas em três categorias:

- a) **sintaxe** - condizem ao formato e valores adotados para a representação do dado;
- b) **semântica** - condizem ao não entendimento do dado registrado, podem conter contradições, duplicidades e registros inválidos;
- c) **cobertura** - condizem à ausência de informação.

Os dados referentes a cada ano estavam em um arquivo CSV próprio, desta forma os 12 arquivos CSV (2005-2016) foram carregados no *Tableau*. Após a carga foi necessário realizar a união entre os arquivos, gerando uma base de dados única com um grande volume de dados, possuindo um total de 1.989.317 linhas. Todos os arquivos possuíam as mesmas 15 colunas não sendo necessário modificações quanto a isso. O Quadro 1 define cada campo existente nos arquivos CSV analisados.

**Quadro 1 – Descrição dos Campos**

<b>Campo</b>	<b>Tipo</b>	<b>Descrição</b>
ANO_CONCESSAO_BOLSA	Numérico	Ano da concessão da bolsa PROUNI
CODIGO_EMEC_IES_BOLSA	Numérico	Código e-MEC referente a Instituição de Ensino Superior que concedeu a bolsa.
NOME_IES_BOLSA	Alfanumérico	Nome/Razão Social da Instituição de Ensino Superior que concedeu a bolsa.
TIPO_BOLSA	Alfanumérico	Descrição do tipo da bolsa concedida (integral – 100%, parcial – 50% ou complementar - 25%).
MODALIDADE_ENSINO_BOLSA	Alfanumérico	Descrição da modalidade de ensino da bolsa concedida (presencial ou à distância – EAD).
NOME_CURSO_BOLSA	Alfanumérico	Nome do curso do beneficiário da bolsa
NOME_TURNO_CURSO_BOLSA	Alfanumérico	Descrição do turno do curso do beneficiário da bolsa.
CPF_BENEFICIARIO_BOLSA	Alfanumérico	CPF do beneficiário da bolsa.
SEXO_BENEFICIARIO_BOLSA	Alfanumérico	Sexo informado pelo beneficiário da bolsa.
RACA_BENEFICIARIO_BOLSA	Alfanumérico	Raça/Cor informado pelo beneficiário da bolsa.
DT_NASCIMENTO_BENEFICIARIO	Data	Data de nascimento do beneficiário da bolsa.
BENEFICIARIO_DEFICIENTE_FISICO	Alfanumérico	Indicação se o beneficiário da bolsa é portador de algum tipo de deficiência (sim ou não).
REGIAO_BENEFICIARIO_BOLSA	Alfanumérico	Nome da região de residência do beneficiário da bolsa.
SIGLA_UF_BENEFICIARIO_BOLSA	Alfanumérico	Sigla da unidade federativa de residência do beneficiário da bolsa.
MUNICIPIO_BENEFICIARIO_BOLSA	Alfanumérico	Nome do Município de residência do beneficiário da bolsa.

Fonte: adaptado de Dicionário – Programa Universidade Para Todos (Brasil, 2016)

Com a base única gerada as colunas denominadas “CODIGO\_EMEC\_IES\_BOLSA”, “CPF\_BENEFICIARIO\_BOLSA” e “MUNICIPIO\_BENEFICIARIO\_BOLSA” foram removidas devido ao fato de seus dados serem irrelevantes para o estudo. A coluna “DT\_NASCIMENTO\_BENEFICIARIO” também foi removida, neste caso, devido às inconsistências dos dados. Haviam anos de nascimento iguais ao ano de concessão da bolsa, ou mesmo de anos futuros, como 2968, claramente causado por erros de digitação. A conferência dos dados desta coluna em outras bases apresentou-se inconveniente para o estudo.

As colunas restantes tiveram seus nomes modificados para melhor visualização no *Tableau*, conforme exibido no Quadro 2.

**Quadro 2 - Nomenclatura Modificada das Colunas**

Coluna Original	Nome para visualização
ANO_CONCESSAO_BOLSA	Ano de Concessao da Bolsa
NOME_IES_BOLSA	Instituicao de Ensino
TIPO_BOLSA	Tipo da Bolsa
MODALIDADE_ENSINO_BOLSA	Modalidade Ensino da Bolsa
NOME_CURSO_BOLSA	Curso da Bolsa
NOME_TURNO_CURSO_BOLSA	Turno do Curso da Bolsa
SEXO_BENEFICIARIO_BOLSA	Sexo do Beneficiario da Bolsa.
RACA_BENEFICIARIO_BOLSA	Raca do Beneficiario da Bolsa
BENEFICIARIO_DEFICIENTE_FISICO	Beneficiario Deficiente Fisico
REGIAO_BENEFICIARIO_BOLSA	Regiao do Beneficiario da Bolsa
SIGLA_UF_BENEFICIARIO_BOLSA	UF do Beneficiario da Bolsa

Fonte: os autores

Algumas linhas da base de dados não tinham as informações regionais, porém foram mantidas devido a importâncias das demais informações, desta forma os dados regionais são exibidos como “Nulo” nos gráficos gerados.

Em relação as colunas utilizadas no estudo, o *Tableau* reconheceu todas como sendo do tipo “Cadeia de Caracteres”, sendo necessário alterar a coluna “UF do Beneficiário da Bolsa” para “Função Geográfica - Estado/Província” e a coluna “Ano de concessão da Bolsa” para o tipo “Data”.

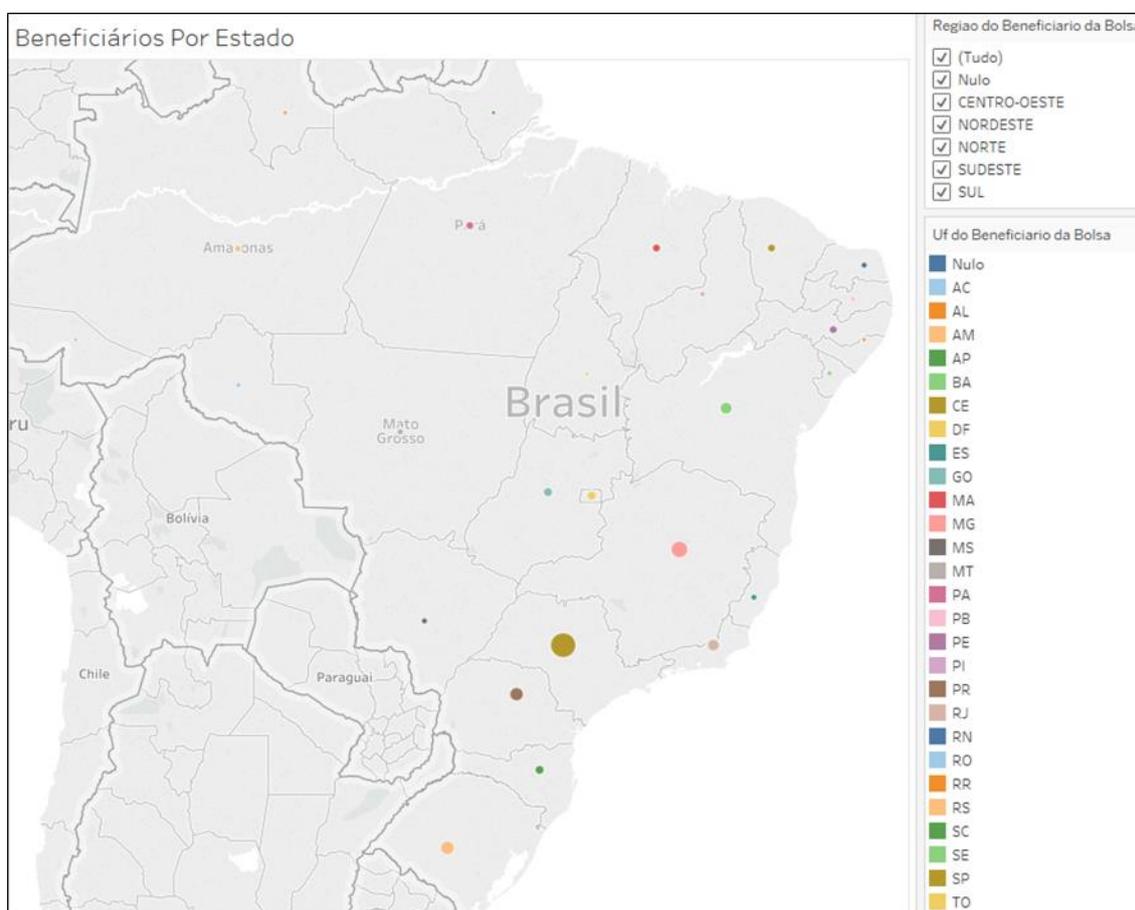
## 5. Análise dos dados

No *Tableau* a pasta de trabalho é dividida em planilhas, estas planilhas podem ser usadas em conjunto para a montagem de painéis, na base estudada, após as adequações, foram criadas 5 planilhas: “Beneficiários por Estado”, “Modalidade de Ensino”, “Instituições com mais Beneficiários”, “Nuvem Cursos” e “Total”. Cada uma com suas particularidades.

- a) **Beneficiário por Estado:** Utilizando-se da coluna “Uf do Beneficiário da Bolsa” e um cálculo da soma do número de registros, foi gerado um mapa, onde cada estado brasileiro foi representado por uma cor específica em círculos que variam de tamanho de acordo o número de beneficiários de cada estado.

Devido à ausência de informações sobre a região dos beneficiários em algumas linhas da base de dados a planilha exibe no filtro a opção “Nulo”, referente aos beneficiários que não possuíam informações regionais. A Figura 1 apresenta esta planilha.

**Figura 1 - Planilha Beneficiários por Estado**



Fonte: os autores.

- b) **Modalidade de Ensino:** O cruzamento dos dados de “Tipo da Bolsa”, “Modalidade Ensino da Bolsa” e “Turno do Curso da Bolsa” juntamente com o cálculo da soma do número de registros gerou um gráfico de barras horizontais onde é possível visualizar os tipos de bolsa mais

procurados, a qual modalidade de ensino se enquadra e em qual turno é ministrado, conforme pode ser visto na Figura 2.

**Figura 2 - Planilha Modalidade de Ensino**

Modalidade de Ensino		Tipo da Bolsa		
Modalidade..	Turno do Curso d..	BOLSA COMPLEMENTAR 25%	BOLSA INTEGRAL	BOLSA PARCIAL 50%
EDUCAÇÃO ..	Curso a distância	28	275.640	48.039
PRESENCIAL	Integral	23	48.486	11.222
	Matutino	396	244.457	118.687
	Noturno	1.815	786.456	397.077
	Vespertino	68	37.248	19.675

Beneficiário Deficiente Físico

(Tudo)

NÃO

SIM

---

Raça do Beneficiário da Bolsa

(Tudo)

Amarela

Branca

Indígena

Não Informada

Parda

Preta

---

Sexo do Beneficiário da Bolsa

(Tudo)

Feminino

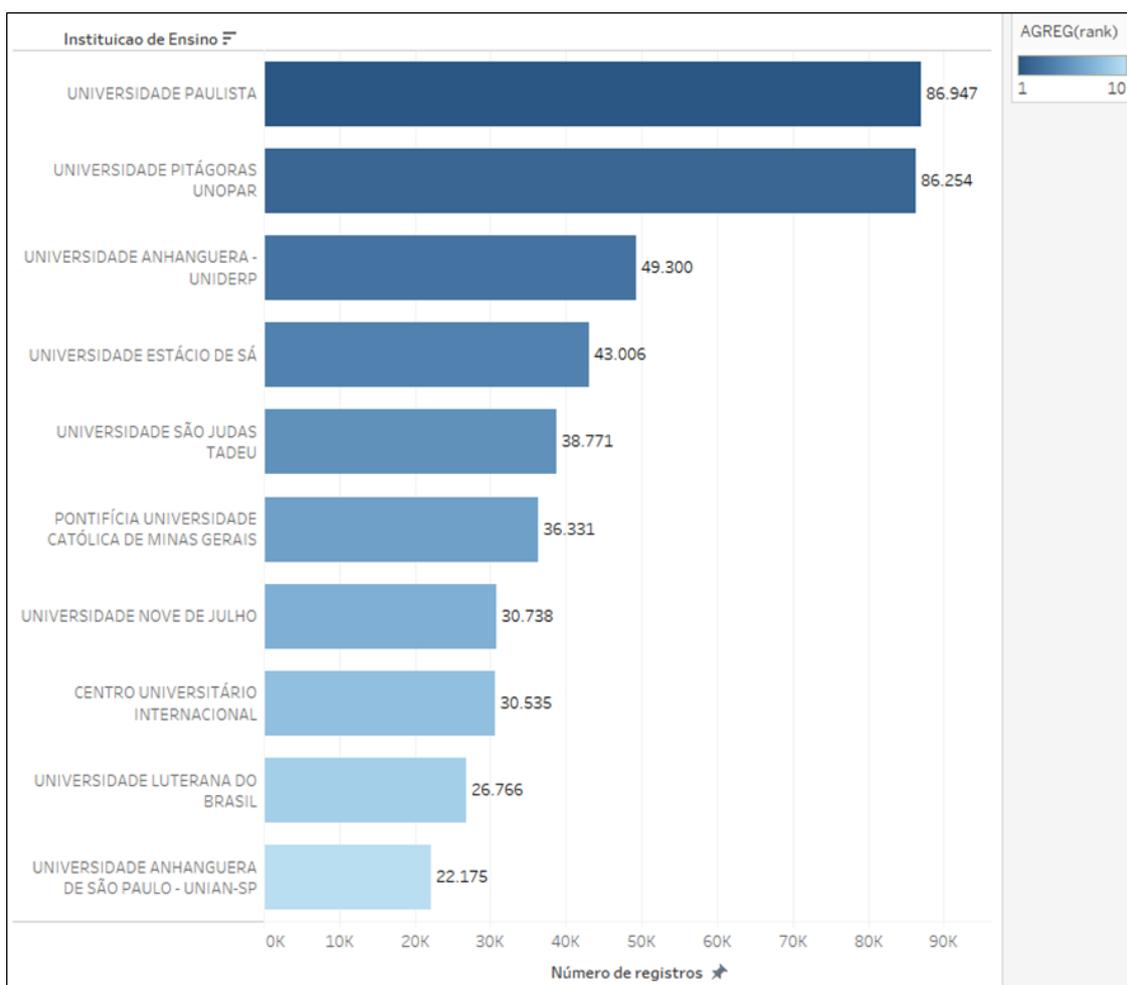
Masculino

Fonte: os autores.

- c) **Instituições com Mais Beneficiários:** Através da coluna “Instituições de Ensino” e do cálculo soma do número de registros foi gerado um gráfico de barras horizontais exibindo o número de beneficiários de cada instituição. O gráfico exibe as instituições de forma decrescente, começando da mais procurada para a menos procurada. Devido ao grande número de instituições foi necessário criar um filtro “rank” que permite a exibição de um número determinado de dados.

Para a criação do “rank” foi necessário adicionar um novo cálculo de tabela selecionando a coluna “Instituição de Ensino” de acordo com a soma do número de registros, depois foi adicionado um filtro para exibir apenas os 10 primeiros resultados. A coloração das barras modifica de acordo com o “rank” da instituição, quanto mais procurada mais escura a tonalidade do azul, opostamente quanto menos procurado mais claro a tonalidade do azul. A Figura 3 ilustra esta planilha.

### 3 - Planilha Instituições com Mais Beneficiários



Fonte: os autores.

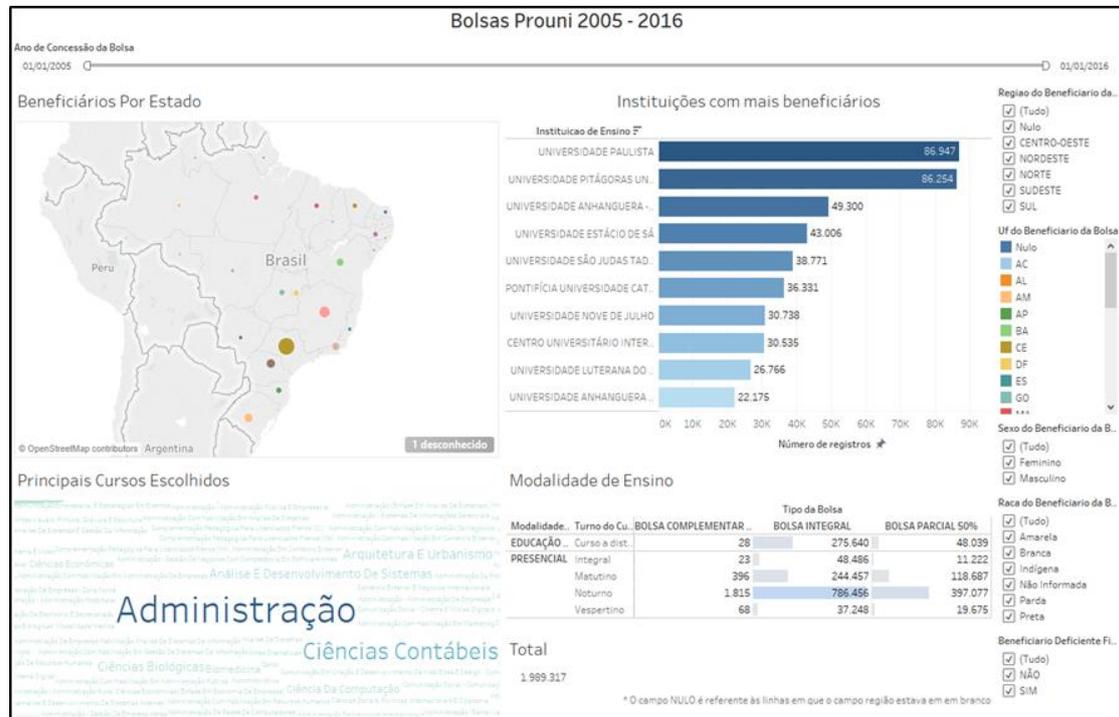
- d) **Nuvem Cursos:** Buscando visualizar os cursos mais procurados foi gerado, a partir da coluna “Curso da Bolsa” e do cálculo soma do número de registros, uma nuvem de palavras, onde os cursos com maior número de beneficiários apresentam-se em tamanho maior que os demais e com uma tonalidade diferente da cor azul. Novamente devido ao grande número de cursos foi criado um filtro “rank” no qual são exibidos os 200 cursos mais procurados. Esta planilha pode ser visualizada na Figura 4.



f) Beneficiário Deficiente Físico.

Além dos filtros criados no painel, exibido na Figura 6, cada planilha pode ser utilizada como filtro para as demais.

**Figura 6 - Painel Bolsas Prouni 2005 - 2016**



Fonte: os autores.

## 6. Resultados

Entre 2005 e 2016 1.939.317 estudantes se tornaram beneficiários do PROUNI, recebendo bolsas integrais ou parciais. No primeiro ano foram 92.629 beneficiários, em 2006 houve um aumento de 14% neste número. Até o ano de 2013 este valor sofreu poucas variações, aumentando ou diminuindo em relação ao ano anterior, já em 2014 houve um aumento significativo, 26% em relação a 2013, totalizando 223.598 beneficiários. Os anos de 2015 e 2016 tiveram poucas variações, fechando em 239.262 em 2016. Desta forma o programa demonstrou um crescimento considerável no número de bolsas concedidas anualmente em relação aos seus primeiros anos.

Segundo o Instituto Brasileiro de Geografia e Estatística – IBGE (2017), através da Pesquisa Nacional por Amostra de Domicílios Contínua, o Brasil possuía em 2016 uma população de 205,5 milhões. Nas pesquisas domiciliares do IBGE, a cor dos moradores é definida por autodeclaração, ou seja, o entrevistado escolhe uma opção entre 5 dadas no questionário: branco, pardo, preto, amarelo ou indígena. Nos dados do PROUNI a coluna “Raça do Beneficiário da Bolsa” possui as mesmas opções de autodeclaração. O Quadro 3 apresenta a participação da população de acordo com a sua cor em um comparativo entre os dados do IBGE e do PROUNI.

**Quadro 3 - Participação da População em Relação a Cor**

Cor	IBGE	PROUNI
Branco	44,2%	44,68%
Pardo	46,6%	39,65%
Preto	8,2%	12,58%
Amarelos, indígenas e não declarados	Menos de 1%	3%

Fonte: os autores

Ainda de acordo com a pesquisa do IBGE, há grandes diferenças regionais na distribuição da população por cor ou raça, o que pode ser explicado pelo processo de ocupação do território. No Sul 76,8% da população se autodeclarou branca, 18,7% parda e apenas 3,8% preta. Por outro lado, na região Norte, 72,3% da população se declarou parda, 19,5% branca e 7,0% preta. Analisando os dados do PROUNI na região sul, 76,9% dos beneficiários são brancos, já na região norte, apenas 15,25% se autodeclarou da cor branca. A análise destes dados revela que a distribuição das bolsas segue números muito semelhantes a proporção da população total levando em consideração a cor de autodeclaração dos habitantes.

Em relação ao sexo, o IBGE (2018) apresenta que pouco mais de 51% da população brasileira é do sexo feminino, número novamente acompanhado pelos beneficiários do PROUNI, onde 53,4% dos bolsistas é do sexo feminino.

O quarto volume da Pesquisa Nacional de Saúde (PNS) 2013 (IBGE, 2016), aponta que o percentual de pessoas de 14 anos ou mais de idade que se declarou ter deficiência física foi de 1,5%. Apenas 0,71% dos beneficiários do PROUNI possuem deficiência física. Este valor a baixo da média da população pode ser explicado pelos diferentes graus da deficiência, pela dificuldade de locomoção até as instituições ou mesmo o não preparo das instituições para receber estes alunos. O número de beneficiários por ano vem aumentando, porém não acompanha a proporção total de crescimento do programa. No ano de 2005, 955 beneficiários possuíam deficiência física, 0,99% do valor total, já no ano de 2016 o número de bolsistas subiu para 1.770, 0,73% do total deste ano, na soma de todos os anos do programa este número cai ainda mais, para 0,71%.

Desde a criação do programa a região Sudeste possui o maior número de beneficiários, sendo que mais da metade deste número está no estado de São Paulo. 402 linhas da base de dados não possuíam dados da região do beneficiário, porém este número não interferiu na análise dos dados, visto que representa 0,02% do total. O Quadro 4 apresenta o percentual de beneficiários de cada região.

**Quadro 4 – Percentual de Beneficiários por Estado**

Região	Porcentagem de Beneficiários
Sudeste	49%
Sul	18,5
Nordeste	17%
Centro-Oeste	9,3%
Norte	6,2%
Nulos	0,02%

Fonte: os autores

A interpretação para esta situação pode estar relacionada ao número de instituições existentes em cada região, a infraestrutura para se locomover até a instituição, a concentração da população, entre outros fatores. O Brasil possui 26 estados e o Distrito Federal, 7 destes estados estão nas regiões Sudeste e Sul, consideradas as mais desenvolvidas do país, nestas duas regiões se concentram 67,5% dos alunos que recebem bolsas. Esta divisão se demonstrou proporcional ao longo de todos os anos.

Em todos os estados os cursos presenciais são mais procurados do que os cursos à distância, porém nas regiões norte e centro-oeste encontra-se a menor diferença entre estas modalidades. O que reforça a ideia que estas regiões possuem maior dificuldade de locomoção ou menor número de instituições.

A maioria das bolsas concedidas é integral, atendendo à população com menor renda, o turno mais procurado é o noturno, provavelmente devido ao fato de muitos alunos conciliarem os estudos a um trabalho diurno. As bolsas complementares de 25% foram concedidas apenas no ano de 2008, desta forma atenderam a um número muito pequeno de estudantes.

A Universidade Paulista é a instituição com maior número de bolsistas do PROUNI, devido aos seus inúmeros polos e a oferta de cursos à distância a instituição distribui bolsas em todos os estados brasileiros, seu destaque também está relacionado ao fato de sua sede estar na região Sudeste, onde há um maior número de beneficiários.

A segunda instituição com mais bolsas concedidas é a Universidade Pitágoras Unopar, com números totais semelhantes a Universidade Paulista, porém esta instituição se destaca na modalidade de ensino à distância, responsável por 95,3% de suas bolsas. Apesar de sua sede estar no estado do Paraná, a maioria dos seus bolsistas estão nos estados de Minas Gerais e da Bahia.

Em nenhum estado a instituição mais procurada se repetiu. E com exceção da Bahia onde a Universidade Pitágoras Unopar do Paraná tem os maiores números, todos os estados tem como sua principal instituição uma com sede na própria região.

O curso mais procurado é o de Administração, principalmente na modalidade presencial, se destacando nas Universidades Pitágoras Unopar e Universidade Paulista. Em relação a modalidade à distância, o curso mais procurado é o de Pedagogia, em sua maioria na Universidade Pitágoras Unopar, seguido de perto pelo curso de Administração. O curso de Administração, inclusive, é o mais procurado no país e também em cada região.

Cursos como Ciências Contábeis, Arquitetura e Urbanismo e Ciências Biológicas também se destacam entre todos os estudantes, mostrando que diferentes perfis de alunos estão sendo contemplados pelo programa. Cursos como Análise e Desenvolvimento de Sistemas e Ciência da Computação tem mais procura por estudantes do sexo masculino, já o curso de Publicidade e Propaganda se destaca entre as mulheres.

## 7. Conclusão

A utilização de uma ferramenta de visualização de dados permite uma análise ampla de uma base com muitos dados. Os 12 arquivos utilizados possuíam juntos quase 2 milhões de linhas, sendo impossível uma análise sem a utilização de uma ferramenta.

O *Tableau*, mesmo na sua versão gratuita demonstrou possuir inúmeros recursos e ter uma curva de aprendizagem bastante rápida, gerando diversos gráficos que possibilitaram chegar às diversas conclusões.

A análise dos dados demonstrou que o PROUNI vem atendendo a população e o número de beneficiários vem aumentando consideravelmente desde sua criação, havendo diminuição apenas quanto as pessoas com deficiência, uma análise mais ampla cruzando informações com outras bases poderia entender esta questão e colaborar em uma ampliação deste número, melhorando as instituições, o transporte e a divulgação dos cursos.

O maior número de bolsas concedidas é para o custeamento integral das mensalidades do curso e além disso destinadas para o período noturno de estudos, atendendo a população com menor renda e os cursos procurados são de áreas variadas, atendendo diversos perfis de alunos.

As regiões Sul e Sudeste possuem o maior número de bolsistas, o cruzamento destes dados com informações de transporte, número de instituições em cada estado entre outras, poderia auxiliar no entendimento desta distribuição e em ações para atender melhor as regiões com menos beneficiários.

Pode-se concluir que o governo eletrônico através da adoção de tecnologias da informação e comunicação e a efetiva disponibilização de dados abertos pelos gestores públicos corroboram com a transparência das informações de forma a viabilizar a participação e a fiscalização da sociedade no controle eficaz do uso dos recursos públicos.

## Referências Bibliográficas

---

- ANDRADE, P. F. L. GODOY, L. A. GIANNO, M. A. CUNHA, C. B. YOSHIZAKI, H. T. Y. (2017) Análise e visualização de dados de rastreamento para caracterização da logística urbana. *Revista Transportes*, 25(3), 24-35. DOI:10.14295/transportes.v25i3.1353
- BARROS, A. J. S., LEHFELD, N. A. S. (2007) *Fundamentos de metodologia científica*. São Paulo: Pearson Prentice Hall.
- BRASIL, Ministério do Planejamento, Orçamento e Gestão. (2012) *Cartilha Técnica Para Publicação De Dados Abertos No Brasil*. Disponível em: <http://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>
- BRASIL, Ministério do Planejamento, Desenvolvimento e Gestão. (2018) *Portal Brasileiro de Dados Abertos*. Disponível em: <http://dados.gov.br>
- BRASIL, Ministério da Educação. (2016) *Dicionário – Programa Universidade para Todos (ProUni)*. Disponível em: [http://informacao.mec.gov.br/bilibs/PDA/PROUNI/Dicionario\\_ProUni.pdf](http://informacao.mec.gov.br/bilibs/PDA/PROUNI/Dicionario_ProUni.pdf)
- CERVO, A. L., BERVIAN, P. A., DA SILVA, R. (2007) *Metodologia científica*. São Paulo: Pearson Prentice Hall.
- CGI. (2011) Comitê Gestor de Internet No Brasil. *Manual dos dados abertos: Desenvolvedores*. São Paulo: CGI.
- CGI. (2009) Comitê Gestor de Internet No Brasil. *Melhorando o acesso ao governo com o melhor uso da web*. São Paulo: CGI.
- CGI. (2018) Comitê Gestor de Internet No Brasil. *TIC Governo Eletrônico: Pesquisa sobre o uso das tecnologias de informação e comunicação no setor público brasileiro 2017*. São Paulo: CGI.
- DAVENPORT, T. H. PATIL, D. J. (2012) Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, 90(10), 70-76.
- FILHO, H.C.P.P, MARTINS, R. A. P. (2017) *Governança digital como vetor para uma nova geração de tecnologias de participação social no Brasil*. *Liinc*. 13(1), 223-236. Disponível em: <http://dx.doi.org/10.18617/liinc.v13i1.3895>
- IBGE. (2018) Instituto Brasileiro de Geografia E Estatística. *Distribuição percentual da População por Sexo - Brasil - 1980 a 2010*. Disponível em: <https://brasilemsintese.ibge.gov.br/populacao/distribuicao-da-populacao-por-sexo.html>
- IBGE. (2016) Instituto Brasileiro de Geografia E Estatística. *Pesquisa Nacional de Saúde 2013: indicadores de saúde e mercado de trabalho: Brasil e grandes Regiões*. Rio de Janeiro: IBGE.
- IBGE. (2017) Instituto Brasileiro de Geografia E Estatística. *População chega a 205,5 milhões, com menos brancos e mais pardos e pretos*. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/18282-pnad-c-moradores>
-

ISOTANI, S. BITTENCOURT, I, I. (2015) *Dados Abertos Conectados*. São Paulo: Novatec.

LEA, B. R. YU, W. MIN, H. (2018) Data visualization for assessing the biofuel commercialization potential within the business intelligence framework. *Journal of Cleaner Production*, 188, 921-941.

OPEN DEFINITION. (2018) *The Open Definition*. Disponível em: <http://opendefinition.org/>

PEDROZA, N. P., BEZERRA, E. P., NICOLAU, M. (2013) *Ferramentas para elaboração de infográficos: Um estudo de caso*. X EVIDOSOL e VII CILTEC - online. Disponível em: <http://evidosol.textolivre.org>

PROUNI. (2018) *Programa Universidade para Todos*. Disponível em: <http://prouniportal.mec.gov.br>

OLIVEIRA, M. M. (2011) *Como fazer projetos, relatórios, monografias, dissertações e teses*. Rio de Janeiro: Elsevier.

TABLEAU. (2018) *Tableau Public*. Disponível em: <https://public.tableau.com/pt-br>

VASCO, D. O. (2013) *Identificação de anomalias contextuais*. Porto: Universidade do Porto, Faculdade de Economia.