

APLICAÇÃO PARA GESTÃO ESTRATÉGICA DOS DADOS DO CENSO DA EDUCAÇÃO SUPERIOR

THIAGO FRAMBACH MUNGARD

Centro Universitário Augusto Motta
thiagomungard@hotmail.com

ALEX AVELLAR DE OLIVEIRA

Centro Universitário Augusto Motta
contato@aavellar.com

NATHIELLY DE SOUZA CAMPOS

Centro Universitário Augusto Motta
nathIELLY@gmail.com

REINALDO VIANA ALVARES

Centro Universitário Augusto Motta
reinaldoviana@gmail.com

ABSTRACT

The census of college education, an activity managed by the National Institute of Studies and Educational Research Anísio Teixeira (INEP), represents the most complete data collection on Brazilian higher education. The study aims to build a business intelligence application for data analysis from the census of higher education, allowing support to managers of institutions to developed data-based decision-making activities. An application was built using Qlik, a Sense Desktop tool. That's free, easy to use, and supports access to large volumes of data. In this stage of the research were incorporated to the microdata related to the census from 2012 to 2015. The use of the application makes it possible preview and attainment information about the situation and trends in the Brazilian higher education area.

RESUMO

O censo da educação superior, atividade gerenciada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), representa a coleta de dados mais completa sobre a educação superior brasileira. O estudo apresentado teve como objetivo construir uma aplicação de inteligência de negócios para análise dos dados do censo da educação superior, permitindo apoio aos gestores das instituições no desenvolvimento de atividades de tomada de decisões com base em dados. A aplicação foi construída usando a ferramenta Qlik Sense Desktop, a qual é livre, de fácil uso, além de suportar acesso a grandes volumes de dados. Nesta etapa da pesquisa foram incorporados os microdados relativos às edições de 2012 a 2015 do censo. O uso da aplicação possibilita a visualização e obtenção de informações sobre a situação e as tendências do setor do ensino superior brasileiro.

Palavras chave: censo, educação superior, inteligência de negócios.

1. INTRODUÇÃO

Realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o Censo da Educação Superior é a pesquisa mais completa que existe sobre as Instituições de Ensino Superior (IES) no Brasil, reunindo grande quantidade de informações sobre as IES, alunos, cursos e docentes de todo o território nacional. Conforme informado pelo INEP em seu portal, o censo tem como objetivo oferecer à sociedade informações detalhadas sobre a situação da educação superior no país e subsidia o planejamento e a avaliação de políticas públicas, além de contribuir em cálculos de indicadores da educação tais como o Conceito Preliminar de Curso (CPC) e Índice Geral de Cursos (IGC).

Esse trabalho apresenta uma aplicação que possibilita obter informações do Censo Superior da Educação de forma prática e objetiva, fornecendo subsídios para uma análise dos dados disponíveis. A aplicação se fez necessária porque a forma a qual o INEP disponibiliza os dados é de difícil leitura e as informações podem se tornar obscuras para uma pessoa não habituada a analisar dados estatísticos em planilhas NASCIMENTO (2014, p. 7). A aplicação usa como fonte de dados os microdados do Censo da Educação Superior de 2012 a 2015. Como ferramenta tecnológica, foi escolhido o aplicativo *Qlik Sense Desktop*, uma ferramenta livre, com diversos recursos gráficos para facilitar a visualização de informações e a construção de painéis de dados.

O estudo está organizado da seguinte forma: a seção dois apresenta os trabalhos relacionados. A seção três discorre sobre o funcionamento do Censo da Educação Superior, bem como caracteriza a escolha da ferramenta usada para criar a aplicação. A seção quatro detalha as atividades realizadas para a criação da aplicação. A seção seis apresenta algumas funcionalidades da aplicação.

2. TRABALHOS RELACIONADOS

O “Resumo técnico: Censo da educação superior 2014” de autoria do INEP, apresenta uma consolidação dos dados do Censo da Educação Superior de 2014 em formato de relatório, que envolve, de maneira abrangente, os principais resultados do censo, tendo como objetivo fornecer estatísticas do sistema de ensino superior à sociedade.

No trabalho “Um Aplicativo para Análise dos Dados do ENADE” de GOMES (2014), foi construída uma aplicação para consulta dos dados dos cursos avaliados pelo ENADE, usando como base os dados disponibilizados pelo INEP e a ferramenta *QlikView*.

No artigo “Uma Aplicação para os dados do ENADE usando Ferramenta *QlikSense*” de NASCIMENTO (2014), uma aplicação de inteligência de negócios foi construída usando a ferramenta *Qlik Sense Desktop*, a qual acessa os resultados do ENADE e permite consultas de forma interativa.

3. METODOLOGIA

Esta seção apresenta o funcionamento e a importância do censo da educação superior e discorre sobre conceitos relativos à construção de sistemas de inteligência de negócios.

3.1 O Censo da Educação Superior Brasileira

O Censo da educação superior é uma das ferramentas mais importantes e úteis para o setor da educação superior brasileira, conforme noticiado pelo Ministério da Educação em uma coletiva de imprensa realizada para a divulgação do Censo da Educação Superior de 2015. O INEP, ao desenhar a trajetória dos estudantes na educação superior, identificou alta taxa de desistência em alguns cursos. A seguir, trecho dessa coletiva, publicado no portal do Ministério da Educação:

“Pela primeira vez, o Censo da Educação Superior traçou um perfil dos estudantes ao longo da graduação, considerando as taxas de permanência, conclusão e desistência. Os dados relativos ao ano de 2015, divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) nesta quinta-feira, 6, revelam um acréscimo desordenado na taxa de desistência do curso de ingresso, na avaliação da trajetória dos alunos entre 2010 e 2014. Em 2010, 11,4% dos alunos abandonaram o curso para o qual foram admitidos. Em 2014, esse número chegou a 49%.” *Fonte: Ministério da Educação*

Na mesma coletiva de imprensa, o então Ministro da Educação José Mendonça Bezerra Filho comentou:

“Este Censo da Educação Superior reforça a tese de que há uma necessidade muito grande de reforma do ensino médio no Brasil. A mudança, proposta pela Medida Provisória nº 746/2016, terá um impacto direto nos indicadores do ensino superior”
Fonte: Ministério da Educação

“O Brasil tem apenas 8% dos alunos do ensino médio em programas vocacionais. A falta de orientação contribui para que haja uma desistência significativa dos jovens que ingressam no nível superior” *Fonte: Ministério da Educação*

De acordo com o INEP, as informações contidas no censo da educação superior são fornecidas pelas próprias IES, através do preenchimento de questionários eletrônicos e carga de dados no sistema e-MEC - sistema do Ministério da Educação criado para a tramitação eletrônica dos processos. Durante o período de preenchimento dos questionários, o Pesquisador Institucional (PI) – interlocutor responsável pelas informações da IES junto ao INEP – pode realizar alterações nos dados da IES, de forma a prover as informações corretas ao INEP. Posteriormente, o INEP verifica a consistência dos dados e disponibiliza o sistema do censo (CENSUP) para que as IES possam conferir e validar os dados. Após o período de validação dos dados pelas IES, o INEP realiza rotinas de análise nos dados do censo para conferir as informações e o Censo da Educação Superior é finalizado.

Uma vez finalizado, os dados do censo são divulgados à sociedade e a sinopse estatística é publicada. Após a divulgação, os dados não podem mais ser modificados, pois passam a fazer parte das estatísticas oficiais do governo brasileiro, conforme informado pelo próprio INEP em seu portal:

“Os dados são divulgados e a Sinopse Estatísticas é publicada. Depois disso não pode mais haver alteração nas informações, pois passam a ser estatísticas oficiais.”
Fonte: Ministério da Educação

A sinopse estatística é a consolidação dos dados coletados pelo sistema do censo junto às IES, em formato de planilha eletrônica, sendo composta por diversas tabelas divididas por modalidade de ensino, faixa etária, cor/raça do aluno ou docente, localização do curso ou IES, categoria

administrativa, organização acadêmica, entre outras. A Sinopse Estatística de 2015, por exemplo, é composta por 51 tabelas.

Os dados do censo são disponibilizados em arquivos no formato separado por vírgula (CSV), em conjunto com o material de suporte necessário para a leitura desses arquivos totalizando cinco arquivos por edição.

Mesmo com as sinopses estatísticas trazendo diversas informações sobre a situação do ensino superior no Brasil, há limitações. Por exemplo, o usuário fica limitado às tabelas criadas pelo INEP e não é possível visualizar informações sob outra ótica. Uma outra limitação é que as sinopses estatísticas são de difícil leitura, pois representam grande quantidade de números em diversas tabelas, sem recurso visual que facilite a interpretação dos dados.

Pode ser de interesse do gestor avaliar os dados e tirar suas próprias conclusões. Uma contribuição deste trabalho é a possibilidade de visualizar os dados do censo sob diversas perspectivas.

3.2 Sistemas de Inteligência de Negócios

A construção de sistemas de inteligência de negócios comumente envolve quatro etapas:

(i) obtenção e estudo das fontes de dados, a qual envolve estudo e entendimento dos dados disponíveis. Aqui é comum a detecção de problemas de inconsistência de dados, ou mesmo informações ausentes.

(ii) planejamento do escopo da aplicação, etapa em que são projetadas as funcionalidades da aplicação, dividindo-as em assuntos ou categorias de análise.

(iii) Procedimento de extração, transformação e carga de dados (do inglês, Extraction, Transform and Load - ETL), onde ocorre a extração de dados das fontes de origem, possíveis transformações envolvendo os mesmos, preparando o ambiente a partir do qual a aplicação fará acesso. É comum que essa etapa consuma 70% do tempo dedicado ao projeto da aplicação.

(iv) Construção de painéis de acesso aos dados, com objetivo de agregar conhecimento útil em processos de tomada de decisão com base em dados.

A ferramenta escolhida para a construção do aplicativo foi o *Qlik Sense Desktop*, pela facilidade de uso, além do fato de ser gratuita e possuir recursos necessários para a realização da etapa de extração, transformação e carga de dados.

4. APLICATIVO PARA ANÁLISE DOS DADOS DO CENSO DA EDUCAÇÃO SUPERIOR

A aplicação foi construída seguindo as etapas apresentadas na seção anterior. A seguir, são apresentados os principais desafios encontrados em cada fase.

4.1 Obtenção e estudo das fontes de dados

A fonte de dados usada corresponde aos microdados do censo da educação superior, referente às edições de 2012 a 2015. Os dados do Censo da Educação Superior são disponibilizados em arquivos no formato CSV delimitados pelo sinal “[”], totalizando cinco tabelas, nomeadas conforme seu conteúdo: Aluno, Curso, Docente, IES e Local_Oferta. Uma vez que as tabelas Aluno, Curso, Docente e IES já contêm todas as informações referentes ao local, a tabela Local_Oferta não foi usada.

Junto aos dados o INEP disponibiliza um material de suporte, sendo o mais importante deles o “ANEXO I”, que é o dicionário de dados das tabelas citadas. O “ANEXO I” contém uma lista de todos

os campos das tabelas, com a descrição, tipo e tamanho de cada campo, além de observações adicionais em alguns casos.

4.2 Planejamento do escopo da aplicação

O planejamento do escopo da aplicação foi realizado após o estudo das fontes de dados do censo. Como resultado, foram escolhidos cenários para investigação levando em conta assuntos de interesse tais como receita da IES *versus* investimentos em pesquisa, quantitativo de alunos por IES, oferta de cursos, dentre outros.

4.3 Procedimento de extração, transformação e carga de dados

Após o planejamento do escopo da aplicação, foram realizados procedimentos de extração, transformação e carga de dados, visando estruturar o ambiente de dados para uso a partir da aplicação.

Na etapa de ETL, os dados foram importados e tratados usando o módulo *Data Load Editor* do próprio *Qlik Sense Desktop*. Os dados de 2012 a 2015 foram consolidados em quatro tabelas seguindo os mesmos nomes dos arquivos originais (Aluno, Curso, Docente e IES), concatenando uma edição do censo ao final da posterior (e.g. tabela Aluno inicia com dados dos alunos de 2015, seguido por 2014, 2013 e 2012). Adicionalmente, em cada uma das tabelas foi criado o campo “ANO” para identificar o ano da edição do Censo da Educação Superior.

Foram identificadas diversas mudanças na estrutura das tabelas de diferentes edições do censo. Todas as mudanças foram tratadas e os ajustes necessários realizados para criar uma base de dados consolidada e consistente. Nesse tratamento o uso do dicionário de dados foi indispensável, pois apenas com ele foi possível identificar tais mudanças e tratá-las adequadamente. Como exemplo é possível citar o campo “TP_PROCEDE_EDUC_PUBLICA” da tabela Aluno, que contém informações sobre o tipo de escola que o aluno concluiu o ensino médio. A partir de 2014 esse campo passou a se chamar “CO_TIPO_ESCOLA_ENS_MEDIO”, então o campo “TP_PROCEDE_EDUC_PUBLICA” foi renomeado ao concatenar as tabelas de 2013 e 2012.

4.4 Construção dos painéis

Realizadas as etapas anteriores, painéis de dados foram construídos a partir dos assuntos definidos no escopo da aplicação. Foram criados quatro painéis gerais, um para cada tabela do censo (IES, Aluno, Docente e Curso), de forma que o usuário possa aplicar os filtros necessários para visualizar as informações desejadas. Além de filtros, medidas e dimensões foram criadas a partir dos dados originais para possibilitar a visualização de certas informações. Um exemplo é a medida “INSCRITOS_TOTAL” que consolida os inscritos em cursos presenciais.

Cada painel fornece visão geral conforme o tema (e.g. o painel IES fornece uma visão geral sobre as IES), permitindo visualizar em uma mesma tela informações sobre quantidades, valores, distribuição por regiões, categorias administrativas, cor/raça do aluno/docente, tipos de cursos, dentre outras. Cabe ressaltar que nem todas as informações contidas nos painéis são obrigatoriamente da tabela de mesmo nome. Por exemplo, o painel IES contém as quantidades de alunos, docentes e cursos que são dados oriundos das tabelas Aluno, Docente e Curso respectivamente. Dessa forma, ao visualizar as informações das IES é possível também visualizar quantos alunos estudam, quantos docentes lecionam e quantos cursos são ministrados nas IES, tudo em um único painel.

5. RESULTADOS

A aplicação permite realizar análises com o uso de filtros para seleção de dados de interesse. Nas próximas seções são apresentados exemplos de análises disponíveis.

5.1 – Análise quantitativa das IES

Aplicando o filtro “ANO” (campo criado para identificar a edição do censo), é possível observar, conforme figura 01, que existiam 2.416 IES em 2012. Adicionalmente, nota-se que existiam aproximadamente 8,6 milhões de alunos, 328 mil docentes e 32 mil cursos no mesmo ano.



Figura 01: Painel IES contendo as quantidades de IES, alunos, docentes e cursos no ano de 2012.

Alterando o filtro “ANO” para as demais edições do censo, identificou-se redução nas quantidades de IES a cada ano, chegando ao menor número em 2015, com 2.364 IES. Em contrapartida, houve aumento nos quantitativos de alunos, docentes e cursos ao longo do período em questão, chegando aos maiores números em 2015, conforme figura 03.



Figura 03: Painel IES contendo as quantidades de IES, alunos, docentes e cursos no ano de 2015.

Com base nos dados observados, nota-se que houve redução na quantidade de IES. No entanto, não significa que houve redução do ensino superior no Brasil. Pelo contrário, as informações analisadas mostram um aumento de alunos, docentes e cursos ao longo de cada ano.

5.2 – Análise dos investimentos em pesquisas

Ao avaliar a relação entre receitas e investimentos em pesquisa, notou-se que o valor investido em pesquisas é muito pequeno em relação ao das receitas. Tal padrão se mantém ao longo do período analisado, conforme figura 03, referente ao ano de 2012.

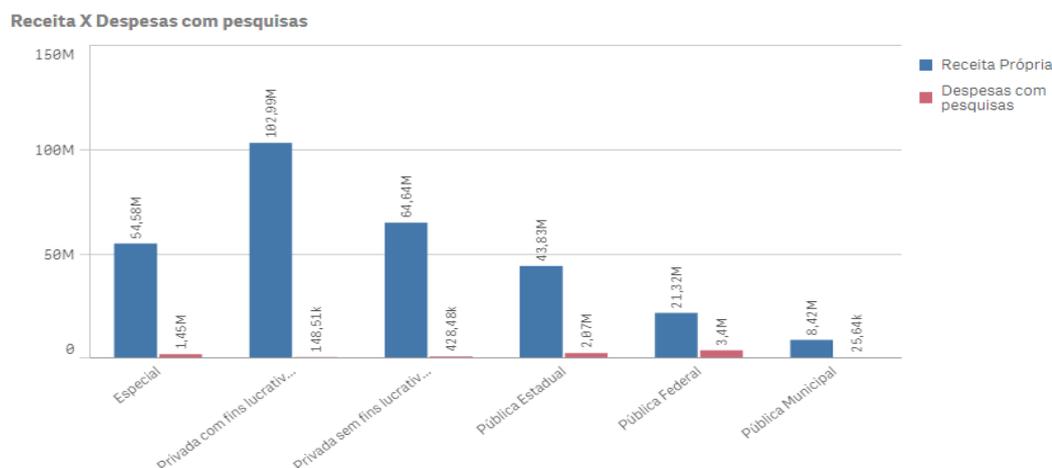


Figura 03: Receitas e despesas com pesquisas das IES, por categoria administrativa, no ano de 2012

Seguindo o mesmo padrão, na figura 04 a seguir são observados os dados referentes ao ano de 2013.

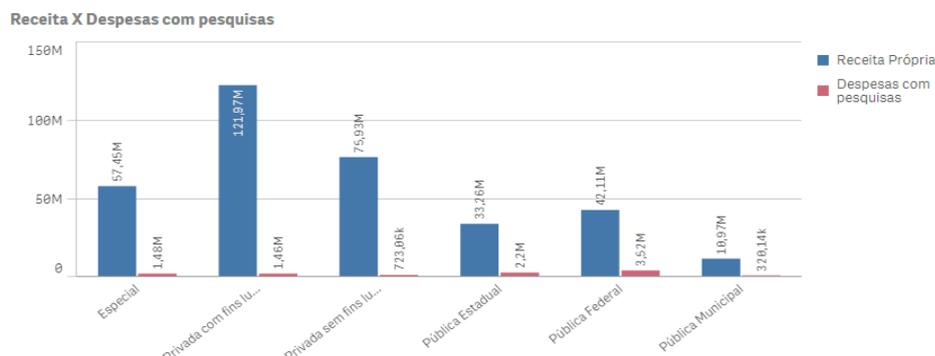


Figura 04: Receitas e despesas com pesquisas das IES, por categoria administrativa, no ano de 2013

Com base nas informações observadas com a aplicação, observa-se que a quantidade de recursos destinados às pesquisas representa um valor muito baixo em relação à receita das IES.

5.3 – Análise do corpo técnico-administrativo

A coleta de dados sobre funcionários técnico-administrativos permitiu a construção de um painel para investigação de informações a respeito desse universo. A figura 05 a seguir mostra dados relativos à edição de 2014 do censo, com informações a respeito dos funcionários das instituições localizadas no estado do Rio de Janeiro.



Figura 05: Informações sobre funcionários técnico-administrativos, referente ao ano de 2014

No ano de 2014, as instituições do Rio de Janeiro empregavam mais de 43 mil pessoas. A maior parte dos funcionários técnico-administrativos possuía o ensino médio completo. O número de funcionários com mestrado ultrapassou o de funcionários com o ensino fundamental incompleto.

5.4 – Análise do corpo docente

A coleta de dados sobre os docentes permitiu a construção de um painel para investigação de informações a respeito desse universo. A figura 06 a seguir mostra dados relativos à edição de 2014 do censo, com informações a respeito dos docentes de todas as instituições participantes do censo.



Figura 06: Informações sobre docentes, referente ao ano de 2014

No ano de 2014, o ensino superior brasileiro empregava mais de 396 mil docentes. Cerca de 48% dos docentes estavam trabalhando sob o regime de tempo integral. Além disso, os docentes do sexo masculino superavam os do sexo feminino em aproximadamente dez pontos percentuais.

5.5 – Panorama das IES

Na figura 07 a seguir é exibido panorama das instituições participantes do censo de 2014, sob a ótica de informações sobre categoria administrativa, organização acadêmica, além de receita por região e por unidade federativa.

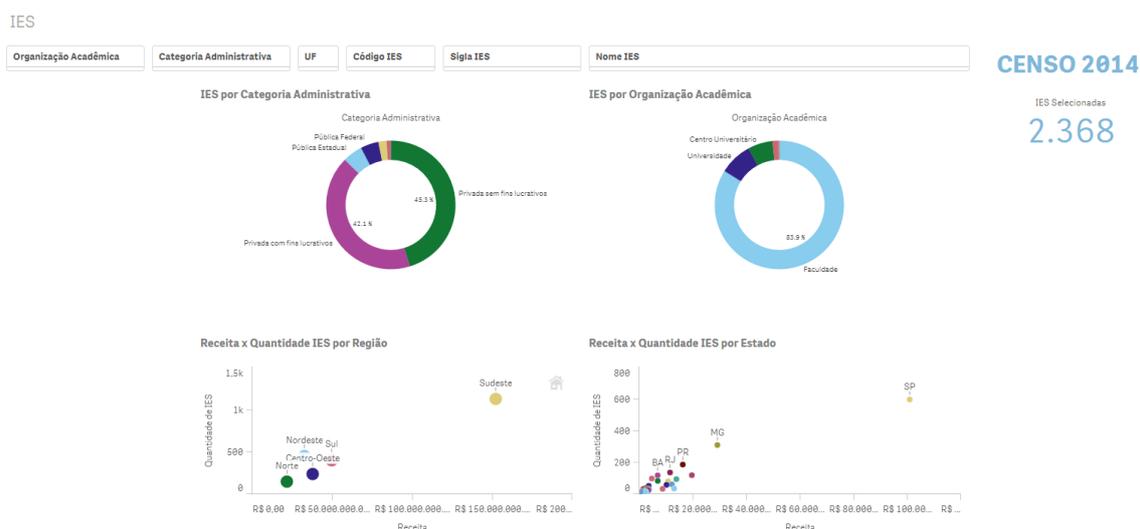


Figura 07: Panorama das instituições, referente ao ano de 2014

É possível observar que no Brasil, cerca de 47% das instituições são privadas. Além disso, mais de 80% representam faculdades. A maior parte das instituições está concentrada na região sudeste do país, o que faz com que essa região possua o maior quantitativo de

receita. Considerando a região sudeste, o estado de São Paulo apresentou o maior quantitativo de instituições, concentrando o maior valor de receita.

Na figura 08 a seguir é exibido panorama das instituições públicas federais participantes do censo de 2014.



Figura 08: Panorama das instituições públicas federais, referente ao ano de 2014

Observa-se que quase 60% das instituições são universidades. Convém ressaltar que, apesar da maioria das instituições estarem localizadas na região sudeste, as instituições localizadas no Paraná concentravam maior receita.

6. CONSIDERAÇÕES FINAIS

O censo da educação superior, realizado anualmente pelo INEP, é fundamental para o entendimento do cenário educacional no país e serve como instrumento para o desenvolvimento de políticas públicas do setor.

Usando os microdados das edições de 2012 a 2015, foi construída uma aplicação que permite a análise dos dados de forma interativa, prática e com recursos que facilitam a visualização e comparação das informações. Como ferramenta tecnológica foi usado o *Qlik Sense Desktop*, que é uma ferramenta gratuita e de fácil uso.

A aplicação permitiu uma análise dos dados do INEP em que foi identificado um crescimento no ensino superior no Brasil ao longo dos anos. Porém, ainda faltam investimentos na área de pesquisa científica. Outra utilidade, por exemplo, refere-se ao uso da aplicação pelas IES para avaliar a oferta de cursos de forma a identificar padrões e ajustar a oferta de vagas conforme a demanda. Cabe ressaltar que o modelo criado não é restrito apenas às análises apresentadas nesse trabalho, sendo possível realizar diversas outras.

Como possíveis trabalhos futuros, a aplicação pode ser incrementada com as demais edições do censo (o censo é realizado desde 1995), além da inclusão dos dados de futuras edições, permitindo análises do sistema da educação considerando todo o seu histórico.

REFERÊNCIAS

ELIAS, Diego. **Qual é a Melhor Ferramenta de BI?**, 2017.

GARTNER Group. **Gartner IT Glossary**, Connecticut, Stamford, 2017.

GOMES, André Viana, MOTTA, Paulo Cezar Pereira. **Um Aplicativo para Análise dos Dados do ENADE**, Trabalho de conclusão de curso, Rio de Janeiro, UNISUAM, 2014.

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Microdados do Censo da Educação Superior**, Brasília, 2017.

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Resumo técnico: Censo da educação superior 2014**, Brasília, 2014. <http://portal.inep.gov.br/resumos-tecnicos1>

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Sinopses Estatísticas da Educação Superior 2015**, Brasília, 2017. <http://portal.inep.gov.br/web/guest/sinopses-estatisticas-da-educacao-superior>

NASCIMENTO, Luciano Silva, KIRMES FILHO, Claudemir. **Uma Aplicação para os dados do ENADE usando Ferramenta QlikSense**, Trabalho de conclusão de curso, Rio de Janeiro, UNISUAM, 2014.

OLIVEIRA, Grimaldo. **Qlik Sense - A Sensação de Experimentar Dados**, 2016.

PORTAL INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior**, Brasília, 2017. <http://portal.inep.gov.br/microdados>

PORTAL MEC Ministério da Educação. **Altos índices de desistência na graduação revelam fragilidade do ensino médio, avalia ministro**, 2016.

PRIMAK, Fabio Vinicius. **Decisões com BI (Business Intelligence)**, 2008.

RIBEIRO, Viviane. **O que é Data Warehouse?**, 2011.

RIBEIRO, Viviane. **O que é ETL?**, 2011.

SANTOS, Maribel Yasmina, RAMOS, Isabel. **Business Intelligence : tecnologias da informação na gestão de conhecimento**, 2006.

TURBAN, Efraim, SHARDA, Ramesh, ARONSON, Jay E., KING, David. **Business Intelligence: Um enfoque gerencial para a inteligência do negócio**, 2009.