



Wallacy Rezende SILVA*

 <https://orcid.org/0000-0002-6927-2181>

Evandro de Araújo JARDINI**

 <https://orcid.org/0000-0002-6034-2171>

Recebido em: 19 de dezembro de 2017

Aprovado em: 07 de maio de 2018

POTENCIALIZAÇÃO DA ANÁLISE DOS MOTIVOS DE EVASÃO DOS ALUNOS DO INSTITUTO FEDERAL DE SÃO PAULO – CÂMPUS VOTUPORANGA POR MEIO DE BUSINESS INTELLIGENCE

ANALYSIS MAXIMIZATION OF SCHOOL EVASION REASONS FROM SÃO PAULO FEDERAL INSTITUTE –VOTUPORANGA CAMPUS - THROUGH BUSINESS INTELLIGENCE

RESUMO

Atualmente, o Instituto Federal de São Paulo – Campus Votuporanga, por meio da Coordenadoria Sociopedagógica, busca instrumentos que facilitem o acompanhamento da evasão dos alunos dos cursos. Este trabalho apresenta contribuições ao processo de gerenciamento e de análise das informações de evasão dos alunos, proporcionando padronização na redação das justificativas de evasão, tabulação mais fácil e aperfeiçoamento dos métodos que disponibilizam essas informações para todas as coordenadorias. Essas melhorias serão aplicadas por meio de conceitos e ferramentas referentes à tecnologia de Business Intelligence, possibilitando uma análise mais acurada das informações pertinentes. Descreve também as etapas a serem seguidas para a implantação de uma solução tecnológica baseada em ferramenta de Business Intelligence, utilizando, juntamente com base de dados, planilhas eletrônicas com a intenção de contribuir com informações para tomadas de decisões. Por meio de tais conceitos e ferramentas, geraram-se gráficos e dashboards que contribuíram para a tomada de decisão da Coordenadoria Sociopedagógica.

Palavras-chave: Business Inteligence. Evasão de alunos. Análise de dados.

ABSTRACT

Nowadays, São Paulo Federal Institute- Votuporanga Campus, through its Socio-educational Coordinating Body looks for instruments to make the students program evasion monitoring easier. This paper presents contributions to management and analysis of students evasion data, providing standardization when writing evasion justification, an easier tabulation and method improvements that provides those data to all coordinating body. Such improvements are applied to concepts and tools relating to Intelligence Business technology, making a more accurate analysis of relevant data possible. It also describes all the steps followed to implement a technological solution based on Intelligence Business tools, together to database electronic spreadsheets were used, aiming to contribute with data to make decisions. Through such concepts and tools, it was generated some graphics and dashboards in order to contribute to Socio-educational Coordinating Body decision making.

Keywords: Business Intelligence. School evasion. Data Analysis.

*Graduado em Análise e Desenvolvimento de Sistemas Instituto Federal de São Paulo - Campus Votuporanga, wallacyrezende@gmail.com

**Graduado em Tecnologia em Processamento de Dados pelo Centro Estadual de Educação Tecnológica Paula Souza, mestre em Ciências da Computação pela Universidade de São Paulo e doutorado em Engenharia Elétrica São Carlos pela Universidade de São Paulo, eajardini@ifsp.edu.br

1 INTRODUÇÃO

Dentre os objetivos das instituições de ensino, um dos mais importantes é formar o maior número possível de alunos, possibilitando-lhes a conclusão dos estudos. Entretanto, devido a diversos motivos, a diminuição da evasão ainda é um desafio a ser vencido na maioria das escolas.

O Instituto Federal de Educação, Ciência e Tecnologia não foge à regra, uma vez que se observa um número significativo de evasões. De acordo com Daros (2015), estima-se que o índice de conclusão dos cursos nas diversas modalidades ofertadas pelos Institutos Federais não chega a 50%, enquanto as metas do Plano Nacional de Educação (PNE-2011-2020) (BRASIL, 2011) preveem, em média, 90% para a taxa de conclusão dos cursos.

Para mitigar o problema, inicialmente, deve-se entendê-lo, motivo pelo qual é imperioso conhecer as causas que levam os alunos a evadirem, já que estão estudando em uma instituição pública de credibilidade como o Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP) e sem precisar despende qualquer quantia financeira. A partir desse conhecimento, torna-se possível traçar estratégias para contornar a evasão e diminuir seus impactos. Porém, para viabilizar organizações e automações dos motivos de evasão, deve-se haver uma padronização de redação nos registros desses motivos.

Segundo a Organização Didática aprovada pela Resolução nº 859 de 07 de maio de 2013 e alterada pela Resolução nº 94 de 29 de setembro de 2015 (MINISTÉRIO DA EDUCAÇÃO, 2015), o aluno que deixar de efetuar a renovação de matrícula dentro dos prazos estabelecidos será considerado desistente. Para fins de estatística, esse abandono é caracterizado como evasão. O registro dos motivos das evasões é realizado, nos Institutos Federais, por Coordenadorias Sociopedagógicas (CSP), compostas por uma equipe técnica que envolve pedagogos, psicólogos, assistentes sociais e técnicos em assuntos educacionais.

No IFSP – Campus Votuporanga, o efetivo registro da evasão ocorre por meio de contato da CSP com o discente, momento em que é registrado o motivo da desistência. Esse registro é anotado em uma folha de papel e, posteriormente, é transferido para uma planilha eletrônica. Essa metodologia de registros em planilha eletrônica apresenta algumas desvantagens diante da necessidade de se analisar os dados. Como não existe um padrão de redação da(s) causa(s) da evasão, os motivos podem ser inseridos de formas linguísticas diferentes, dificultando o processo de análise e classificação dos dados.

No caso das instituições federais de ensino, as verbas a elas destinadas estão diretamente relacionadas à quantidade de alunos matriculados (matrículas ativas). Nas instituições de natureza privada, cada matrícula não trancada ou cancelada representa percepção de receita. Dessa forma, independentemente da natureza jurídica, a minimização da evasão torna-se necessária para a existência e manutenção da instituição, porquanto aluno matriculado significa a garantia de aumento de benefícios e/ou recursos acadêmicos.

Diante dessas constatações, propõe-se a utilização de uma ferramenta de análise que emprega conceitos de Business Intelligence (BI) no IFSP - Câmpus Votuporanga, a fim de auxiliar os servidores das instituições em atividades de análise e tomadas de decisão relacionadas ao fenômeno da evasão.

Optou-se por estudar as ferramentas Pentaho, Power BI e o Sistema de Gerenciamento de Banco de dados (SGBD) Postgres, que, combinados, possibilitam aplicar os conceitos de BI na base de dados acadêmica. Essa combinação proporciona a extração de informações relevantes sobre o que se deseja analisar e as armazena em uma base de dados analítica denominada base Data Warehouse, a qual viabiliza a visualização das informações em dashboards e indicadores.

1.1 Objetivo do trabalho

Conforme apresentado, o IFSP – Campus Votuporanga enfrenta um problema comum às instituições de ensino: a evasão de alunos. Para combater essa adversidade, deve-se conhecer suas causas. Contudo, a maneira atual de se registrar os motivos das evasões não é padronizada a ponto de facilitar recuperações de dados, análises e a elaboração de políticas de combate a este problema. Diante dessa problemática, este trabalho tem o objetivo de propor uma metodologia para a padronização do registro dos motivos que levam os alunos a evadirem, o que gerará dados para estudo por meio de ferramentas de análise de dados, como as encontradas na tecnologia de Business Intelligence. Esse instrumental proporcionará, ademais, acesso instantâneo às informações pertinentes.

Dessa forma, acredita-se ser possível tornar disponíveis e de forma mais viável aos servidores da CSP, informações sobre a evasão como: evasão por faixa etária; por sexo; pelas cidades onde os alunos residem, conseqüentemente, que apresentam maior índice de evasão e pelos anos em que os alunos mais se evadiram.

A utilização da ferramenta Power BI viabilizará, pois, acesso instantâneo aos motivos de evasão, disponibilidade dos dados pela web, análise temporal dos motivos das evasões, possibilidade de compartilhamento das informações para a CSP, diretores, coordenadores, entre outros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Business Intelligence (BI)

Uma das características do Business Intelligence é oferecer mecanismos que facilitem e viabilizem a análise meticulosa dos dados operacionais de uma base de dados para que, posteriormente, favoreça a obtenção de informações relevantes em processos de análise e de tomada de decisão.

De acordo com Barbieri (2001) representa a habilidade de se estruturar, acessar e explorar informações, normalmente guardadas em DW/DM (Data Warehouse/Data Mart), com o objetivo de desenvolver percepções, entendimentos, conhecimentos, os quais podem produzir um melhor processo de tomada de decisão.

Dessa maneira, o BI otimiza o desempenho nas tomadas de decisões, já que possibilita a obtenção de informações de maneira cada vez mais célere.

2.2 Conceitos e ferramentas em BI

O Data Warehouse (DW) é a base de dados que disponibiliza informações multidimensionais para apoio na tomada de decisões. É constituída por dados que são copiados e analisados da base de dados On-Line Transaction Processing (OLTP). Para criá-la, é necessário elencar todas as necessidades de uma empresa e das pessoas que ali trabalham, motivo pelo qual não há um modelo a ser seguido.

O conceito de Data Mart (DM) é, praticamente, o mesmo do DW, com algumas diferenças. Uma delas seria o volume de dados que as constituem. Enquanto os DMs são voltados para uma determinada área específica em um negócio, sendo um subconjunto de informações dos DWs, priorizando o desempenho em relação à velocidade das consultas de informações, o DW relaciona todo o negócio em questão, ou seja, envolve todos os dados sobre

a empresa e sua operação (VALENTIM et al., 2017).

On-Line Transaction Processing (OLTP) são aplicações que utilizam dados operacionais em seu contexto. São, por exemplo, aqueles sistemas em que os usuários cadastram um pedido ou emitem uma nota fiscal, os quais não visam, prioritariamente, a análise de informações, mas o cadastro de dados importantes para a empresa.

A On-Line Analytic Processing (OLAP) é uma aplicação de processamento analítico online de dados utilizada por usuários finais para analisar os dados obtidos de um DW. As consultas feitas pelos usuários retornam dados agregados, não sendo dados operacionais, que serão analisados e, posteriormente, utilizados para tomar decisões estratégicas. Por meio da ferramenta OLAP é possível criar relatórios, cruzar dados, analisar gráficos e dashboards, gerenciar e atribuir permissões de acesso a usuários ou grupos de usuários.

As aplicações OLAPs possuem, em sua estrutura, alguns conceitos, dentre os quais o de multidimensionalidade, que seria empregar cubos, dimensão, hierarquia, membro e medida.

Em Anzanello (2017), são definidos os seguintes termos:

- Cubo é uma estrutura que armazena dados e informações cruzadas, tornando-as mais fáceis de visualizar e analisar.
- Dimensão é uma unidade de análise que agrupa dados de negócio relacionados. Tornam-se cabeçalho de colunas e linhas, como por exemplo, produtos, regiões de venda, clientes ou períodos de tempo.
- Hierarquia são os níveis hierárquicos a serem explorados em uma dimensão.
- Membro é o subconjunto de uma dimensão. Cada nível hierárquico possui membros designados àquele nível, como por exemplo, uma dimensão geográfica, que possui um nível país, em que seus membros seriam Brasil, China, Estados Unidos, etc.
- Medida é uma dimensão que é utilizada para realizar comparações. Inclui membros como: desconto, preço, quantidade, lucros ou taxas.

Em BI, é comum utilizar o termo Key Performance Indicator (KPI) para definir um indicador de desempenho. Em suma, KPI pode ser entendido como uma ferramenta que exibe uma medição ou o desempenho de determinado processo.

O termo KPI vem do inglês Key Performance Indicator. Os Indicadores-Chave de Performance (ICP) contemplam uma profundidade maior na análise dos dados. Em linhas gerais, os ICP contemplam os objetivos principais de um negócio, a razão pela qual ele existe (PORTO, 2014, p.54).

Essas definições conceituam BI e são utilizadas por várias ferramentas, a fim de analisar

o desempenho de um determinado processo.

2.3 Extract Transform Load (ETL)

Em sua concepção, Extract Transform Load (ETL) é o processo de extração de dados externos para uma base de dados Data Warehouse, por meio de uma ferramenta em que a base de dados analítica é alimentada. Essa ferramenta possui a capacidade de se comunicar com uma base que dispõe de diversos formatos de dados podendo inseri-los em um ambiente Data Warehouse. O processo é iniciado pela extração dos dados, seguido pela limpeza, ajustes e consolidação (ou também chamada transformação) e, por fim, realiza-se a entrega ou carga dos dados (ARONSON et al., 2009).

2.4 Ferramentas OLAP

Pentaho é uma plataforma de código aberto que foi desenvolvida por meio de APIs para promover o crescimento das soluções em BI, o que possibilita criar, rapidamente, soluções de BI. Portanto, é possível aumentar o valor das soluções feitas pelas empresas e, ao mesmo tempo, facultar um custo mais baixo por isso.

Um dos motivos pelos quais os desenvolvedores elegem essa ferramenta de BI é por ser um software gratuito. Por este motivo, a utilização do Pentaho vem crescendo nos últimos tempos (FREITAS, 2015).

O Power BI é constituído por um conjunto de ferramentas de análise de negócios, cujo intuito é analisar e compartilhar dados. Com ele é possível criar uma grande quantidade de painéis com vários aspectos, mostrando os dados de uma maneira dinâmica e de fácil entendimento para o usuário corporativo. Nesse ambiente, o usuário tem a possibilidade de indicar filtros com base no que deseja visualizar, possibilitando que consiga elaborar relatórios com o objetivo de exportar esses dados (MICROSOFT, 2017).

3 DESENVOLVIMENTO

Para a criação da base de dados analíticos, inicialmente, utilizou-se o SGBD Postgres.



Porém, para facilitar a utilização e o cadastro das informações por usuários não técnicos, optou-se por não utilizar mais o SGBD Postgres e inserir os dados em planilhas eletrônicas, visto que os servidores da CSP têm fácil acesso a elas, eliminando, dessa forma, a necessidade da criação de um sistema com interface para cadastro das informações relacionadas à evasão.

De posse dos dados disponibilizados pela CSP, foram utilizadas as justificativas de evasão elencadas por Daros (2015, p.2):

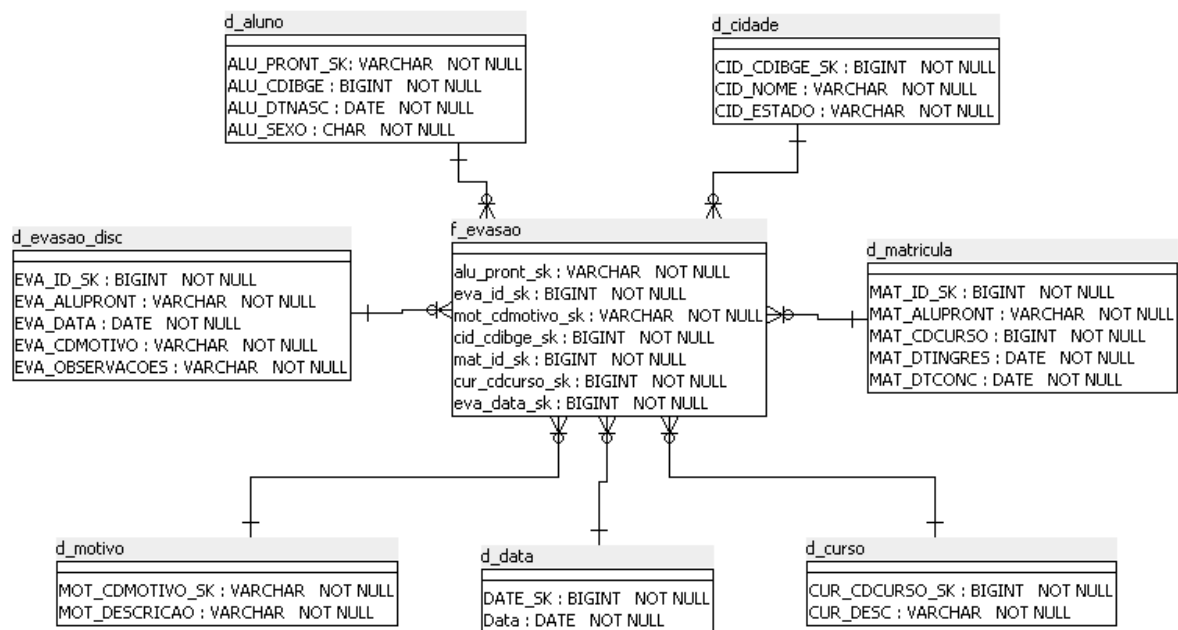
- Dificuldades na relação professor-aluno;
- Horário de trabalho incompatível;
- Passou no vestibular em outras instituições;
- Não se identificou com o curso;
- Não conseguiu conciliar trabalho com o curso;
- Dificuldade em acompanhar o curso;
- Mudança para outros cursos.

Com base nas justificativas mais frequentes, procedeu-se à classificação dos motivos, adequando-os à realidade do Campus Votuporanga:

- Não se identificou ou não se adaptou ao curso.
- Ingressou ou retornou a outro colégio ou curso.
- Interferência do emprego.
- Família.
- Transporte.
- Saúde pessoal.
- Particularidade.
- Problemas com horário.
- Reprovação.
- Outros motivos.

Primeiramente, os dados extraídos do sistema OLTP utilizado pelos servidores do IFSP, juntamente com as informações fornecidas pela CSP, resultaram em dimensões para o modelo de dados estrela (Figura 1), por meio do procedimento ETL, que foi feito manualmente analisando os dados relevantes a serem utilizados e armazenando-os em planilhas eletrônicas usadas como base de dados analítica.

Figura 1 – Modelo de dados Estrela

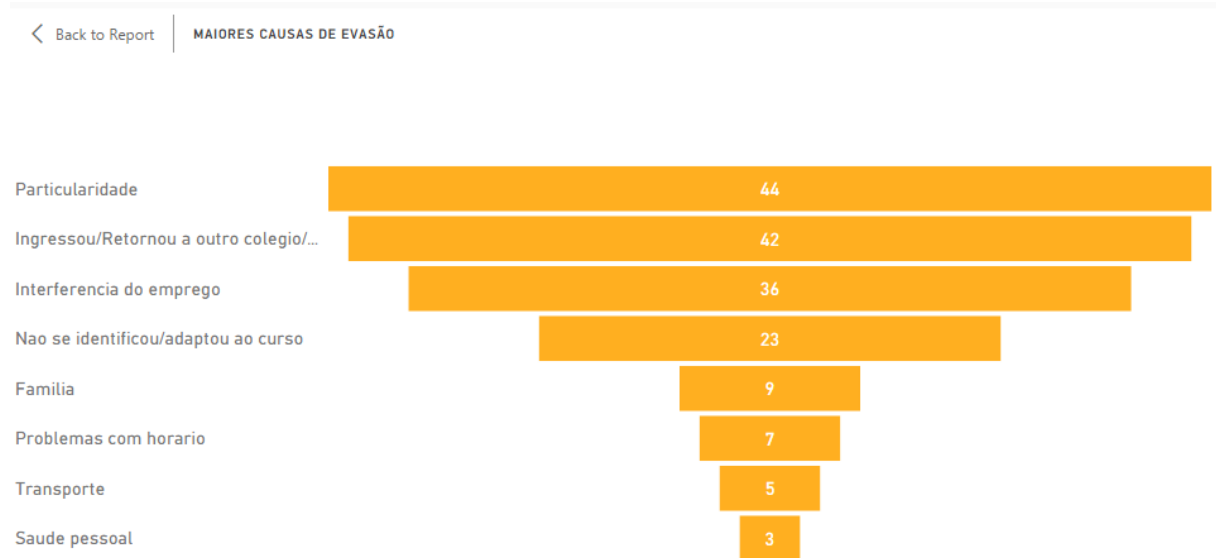


Fonte: Do próprio autor.

Por meio desses dados, criaram-se as dimensões aluno (d_aluno), curso (d_curso), matrícula (d_matricula), cidade (d_cidade), evasão discente (d_evasao_disc), motivos de evasão (d_motivo) e data (d_data), que foram tabeladas para a base de dados analítica (os dados armazenados em planilhas eletrônicas). Em seguida, as dimensões foram analisadas e normalizadas para se adequarem ao modelo. As dimensões são constituídas por: d_aluno (alu_pront_sk, alu_cdibge, alu_dtnasc e alu_sexo), d_curso (cur_cdcurso_sk e cur_desc), d_matricula (mat_id_sk, mat_alupront, mat_cdcurso, mat_dtingres e mat_dtconc), d_cidade (cid_cdibge_sk, cid_nome e cid_estado), d_evasao_disc (eva_id_sk, eva_alupront, eva_data, eva_cdmotivo e eva_observacoes que contêm as especificações do motivo que levou o aluno a evadir), d_motivo (mot_cdmotivo_sk e mot_descricao) e d_data (date_sk e data). Por meio dessas dimensões, foi criada a fato evasão (f_evasao), que possui o relacionamento entre todas as dimensões, contendo os campos (alu_pront_sk, eva_id_sk, mot_cdmotivo_sk, cid_cdibge_sk, mat_id_sk, cur_cdcurso_sk e eva_data_sk)

Por fim, utilizou-se a ferramenta Power BI para a criação de dashboards que vão auxiliar os profissionais da CSP em suas tomadas de decisões. Foram implementados os seguintes relatórios gráficos:

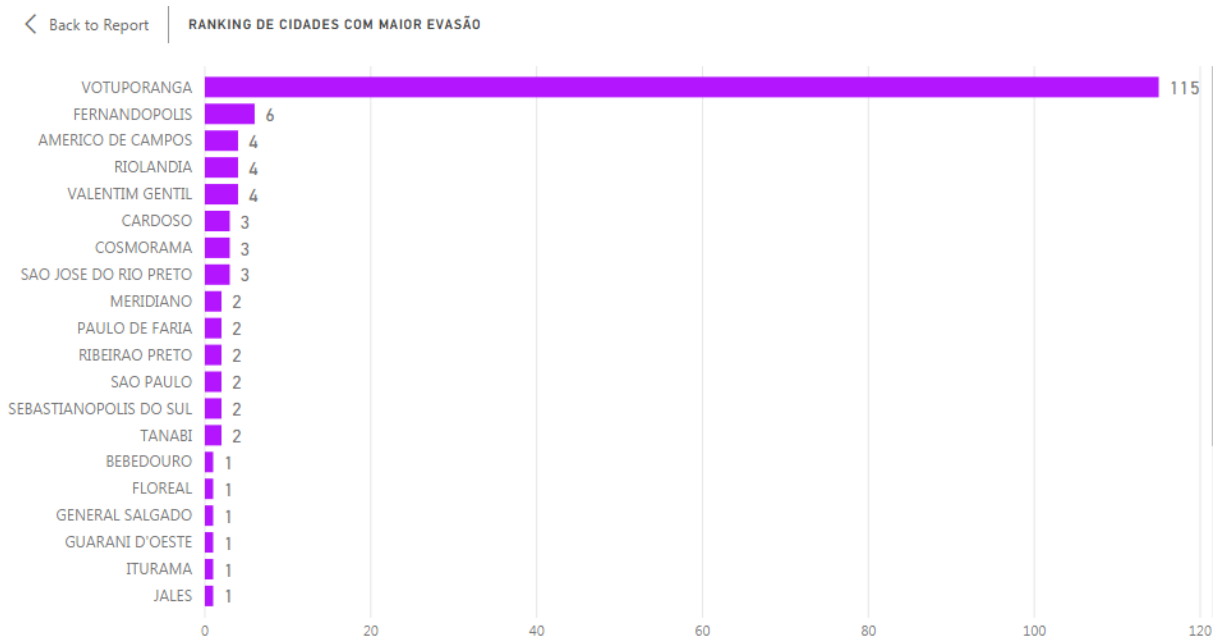
Gráfico 1 – Maiores Causas de Evasão



Fonte: Do próprio autor.

No Gráfico 1 - “Maiores causas de Evasão” podem-se analisar as razões mais frequentes que levaram os alunos a se evadirem dos cursos. As informações utilizadas nesse gráfico foram disponibilizadas pela CSP.

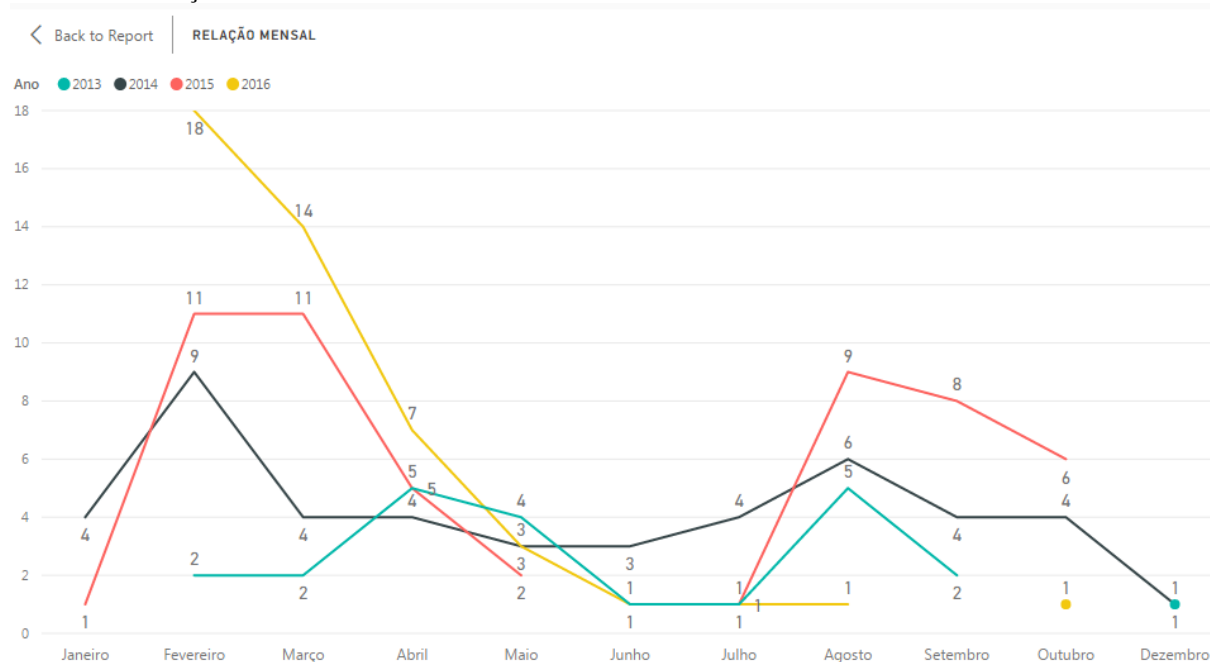
Gráfico 2 – Ranking de Cidades com Maior Evasão



Fonte: Do próprio autor.

O Gráfico 2 exibe as cidades onde residem os alunos evadidos. Os dados utilizados foram disponibilizados pelo sistema OLTP do IFSP.

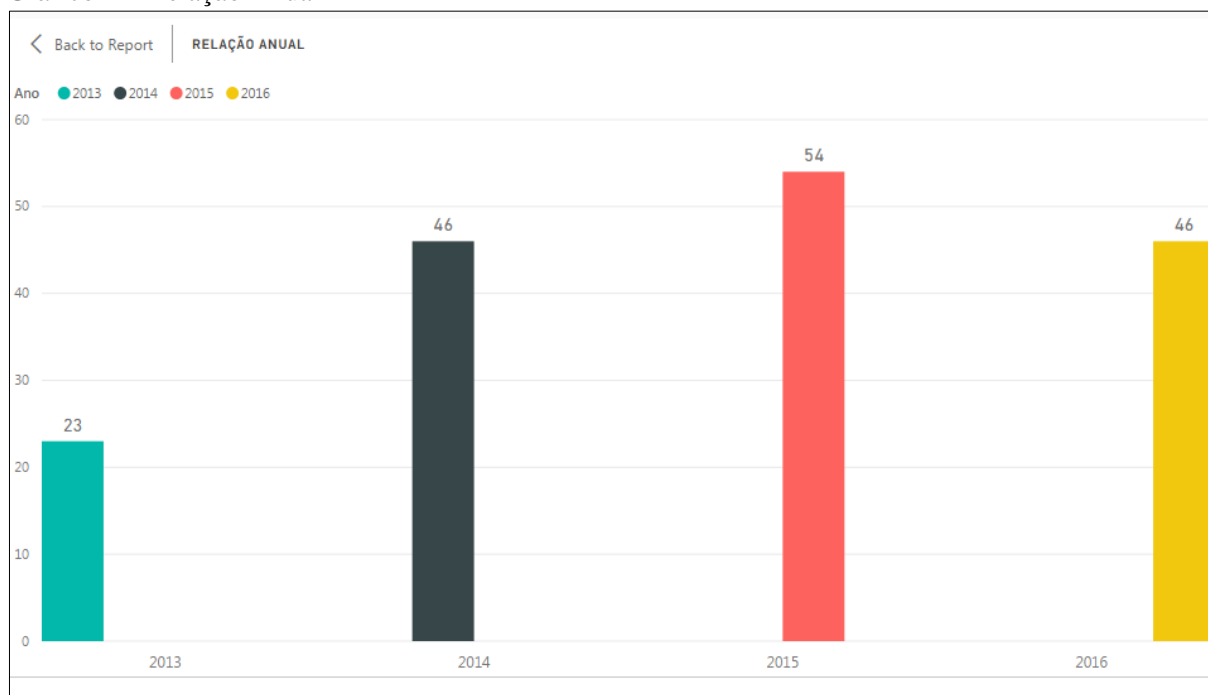
Gráfico 3 – Relação Mensal



Fonte: Do próprio autor.

O Gráfico 3 permite visualização das evasões, mensalmente, durante os anos 2013, 2014, 2015 e 2016, conforme as informações da CSP.

Gráfico 4 – Relação Anual



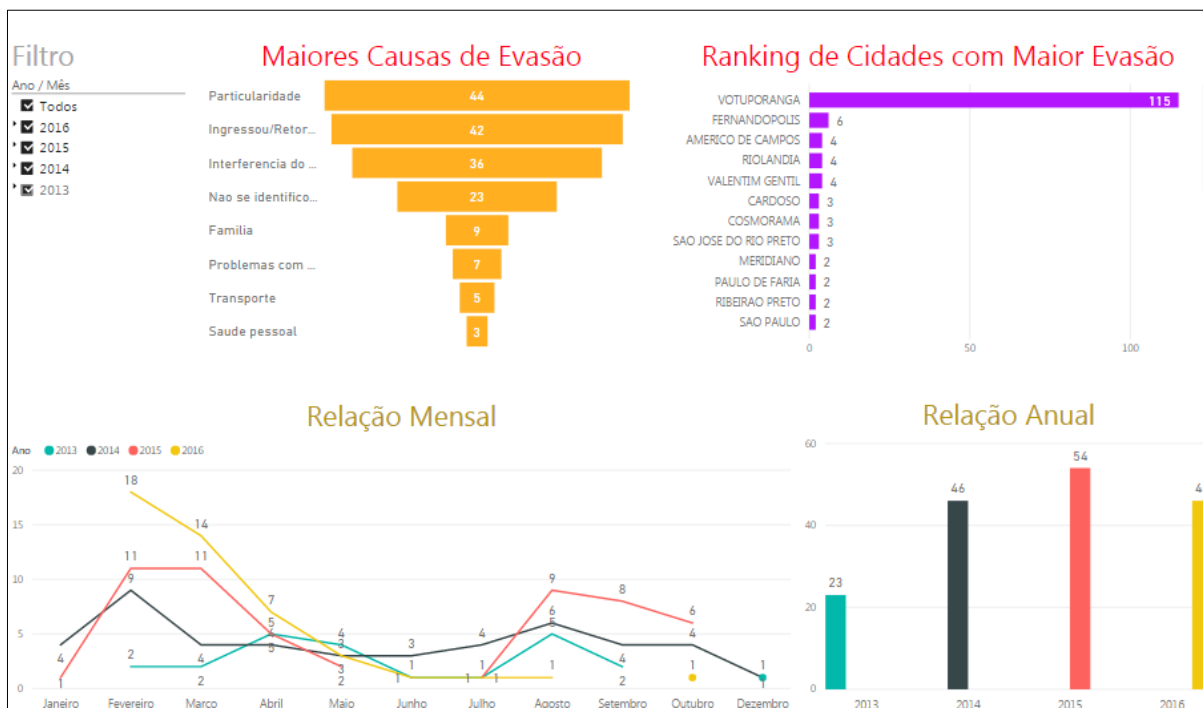
Fonte: Do próprio autor.

O Gráfico 4 demonstra a quantidade de evasões durante os anos 2013, 2014, 2015 e 2016, com base nas informações da CSP.

4 RESULTADOS E DISCUSSÃO

Os dados importados do esquema estrela em Planilhas eletrônicas foram desenvolvidos na ferramenta Power BI dashboards para análise, o que pode ser visto na Figura 2, logo abaixo. A figura retrata quatro gráficos criados: Maiores Causas de Evasão, Ranking de Cidades com Maior Evasão, Relação Mensal e Anual. Além disso, foi projetado um filtro por ano, possibilitando a filtragem de dados conforme seleções do usuário.

Figura 2 – Painel Evasão de Alunos



Fonte: Do próprio autor.

4.1 Painel evasão de alunos

No gráfico “Majores causas de Evasão” podem-se analisar os motivos de evasão mais frequentes no Campus. Percebe-se que “Particularidade” (que seriam motivos particulares do discente) é a principal causa, tendo 44 casos, representando 26% do total de 169. Com base nessas informações, a CSP pode realizar entrevistas com o intuito de intervir nessa situação.

O gráfico “Ranking de Cidades com maior Evasão” demonstra as cidades que possuem maior índice de evasão. A análise permite concluir que a cidade de Votuporanga está na primeira posição do gráfico: visto que a instituição está instalada nessa cidade, isso se torna comum, levando-se em consideração a proporção de alunos.

O gráfico “Relação Mensal” permite analisar a quantidade de evasão mensalmente de 2013 a 2016. Os meses de maior frequência de evasão são fevereiro, março e agosto. Uma hipótese seria que, no mês de março, são realizadas as primeiras provas dos cursos, o que leva os discentes com notas baixas a evadirem.

Por fim, o gráfico “Relação Anual” exhibe a quantidade de evasões separadas por ano. No ano de 2015, houve a maior quantidade de evasão, porém houve uma queda no ano seguinte.

5 CONCLUSÃO

Este trabalho abordou uma proposta de desenvolvimento de uma metodologia que auxilie a CSP a acompanhar as evasões do IFSP por meio de ferramentas de Business Intelligence. Também foram descritas as etapas a serem seguidas na implantação dessa ferramenta a fim de contribuir com informações para trabalhos afins.

Como sugestões de melhorias para aprofundamento no assunto deste trabalho: disponibilizar indicadores em porcentagens nos gráficos; substituir a planilha eletrônica pelo SGBD Postgres e propor uma solução de integração entre a base de dados do sistema acadêmico e a base de dados analítica; criar um KPI que indique quando o número de evasão do ano atual for superior ao anterior.

REFERÊNCIAS

ANZANELLO, C. **OLAP**: conceitos e utilização. Disponível em: <http://www.softsystemit-ead.com.br/phocadownload/BI/Conceitos%20OLAP.pdf>. Acesso em: 08 de fev. 2017.

BARBIERI, C. Business Intelligence: modelagem e tecnologia. Rio de Janeiro: Axcel Books, 2001.

DAROS, M. A. **Evasão no IFSP**: um convite à reflexão. Sertãozinho, I CONEPT, 2015. Disponível em: <http://ocs.ifsp.edu.br/submissao/index.php/conept/i-conept/paper/viewFile/193/65>. Acesso em: 12 abr. 2017

FREITAS, A. A. Pentaho BI: conhecendo a ferramenta open source **BI.SQL Magazine**, v.128, p.22-7, 2015. Disponível em: <http://www.devmedia.com.br/pentaho-bi-conhecendo-a-ferramenta-open-source-bi/32095>. Acesso em 23 de abr. de 2017.

MICROSOFT. **O que é Power BI?**2017. Disponível em: <https://powerbi.microsoft.com/pt-br/what-is-power-bi/>. Acesso em: 12 abr. 2017.

ARONSON, J. E. et al. **Intelligence**: um enfoque gerencial para a inteligência do negócio. São Paulo:Bookman, 2009.

VALENTIM, M. L. P. et al. O processo de inteligência competitiva em organizações. **DataGramZero -Revistade Ciência da Informação**, v.3, n.4, 2003. Disponível em: http://moodle.fgv.br/cursos/centro_rec/docs/o_processo_inteligencia_competitiva.pdf . Acesso em: 08 de fev. 2017.

BRASIL. Ministério da Educação. **O PNE 2011-2020**: metas e estratégia. http://fne.mec.gov.br/images/pdf/notas_tecnicas_pne_2011_2020.pdf

MINISTÉRIO DA EDUCAÇÃO. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo. **Resolução nº 94/2015**, de 29 de setembro de

2015.https://arq.ifsp.edu.br/portal/phocadownload/cae/2017/documentos/organizacao_didatica_dos_cursos_ofertados_pelo_ifsp.pdf

PORTO, C. **Facebook marketing**: tudo que você precisa saber para gerar negócios na maior rede social do mundo. São Paulo: Novatec,
2014.<https://books.google.com.br/books?isbn=8575224972>