



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO

CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA

ESCOLA DE INFORMÁTICA APLICADA

Desenvolvimento de indicadores para análise de desempenho e evasão de alunos da
UNIRIO com a utilização de Self-Service BI

BIANCA DE PAIVA ORSI

LAURA CALANCA GÓES

Orientador

FLÁVIA MARIA SANTORO

RIO DE JANEIRO, RJ – BRASIL

NOVEMBRODE2016

Desenvolvimento de indicadores para análise de desempenho e evasão de alunos da
UNIRIO com a utilização de Self-Service BI

BIANCA DE PAIVA ORSI
LAURA CALANCA GÓES

Projeto de Graduação apresentado à Escola de
Informática Aplicada da Universidade Federal do
Estado do Rio de Janeiro (UNIRIO) para obtenção do
título de Bacharel em Sistemas de Informação.

Aprovada por:

[NOME DO ORIENTADOR] (UNIRIO)

[NOME DO PRIMEIRO INTEGRANTE DA BANCA]

[NOME DO SEGUNDO INTEGRANTE DA BANCA]

RIO DE JANEIRO, RJ – BRASIL.

NOVEMBRODE2016

Agradecimentos

Gostaria de agradecer ao apoio da minha família e amigos durante esses meses que precisei me concentrar no TCC, a parceria e dedicação da minha amiga Bianca Orsi que abraçou a idéia junto comigo e tornou esse trabalho mais divertido, aos nossos usuários Ângelo e Armando que trouxeram a demanda e nos deram a chance de realizar um projeto útil para a nossa universidade, e a nossa orientadora Flavia Santoro e nosso professor Tanaka, que acompanharam o desenvolvimento do nosso trabalho desde o início e nos ajudaram a pensar sempre da melhor forma.

Laura Góes.

Gostaria de agradecer à minha família que sempre me apoiou e me deu todo o suporte nos meus estudos, a minha amiga Laura Góes, por sua dedicação e companheirismo nesse trabalho e na faculdade, aos professores do curso de Sistemas de Informação da UNIRIO, principalmente a nossa orientadora Flávia Santoro e o professor Tanaka, que me fez despertar o interesse e dedicação pela área de BI e aos nossos usuários, Ângelo e Armando, que nos ajudaram e tornaram esse trabalho possível.

Bianca Orsi.

RESUMO

Neste trabalho é apresentada uma análise do ingresso de alunos na Universidade Federal do Estado do Rio de Janeiro. No Brasil, a evasão escolar, entendida como interrupção no ciclo de estudo, causa prejuízos significativos sob o aspecto econômico, social e humano em qualquer que seja o nível de educação. O ambiente entre as instituições de educação superior é competitivo e a viabilidade econômica tem forte ligação com os programas de permanência do aluno. Utilizando os micro dados do SIE - Sistema de Informação para Ensino, e do SiSU - Sistema de Seleção Unificada, foi criada uma solução de visualização na ferramenta Tableau, com a definição de métricas que geraram indicadores para auxiliar o entendimento da instituição sobre possíveis causas que geram evasão, desistência e mobilidade de alunos entre cursos, viabilizando a tomada de decisão da Pró-Reitoria de Graduação da Universidade. Por se tratar de uma ferramenta de Self-Service BI, os usuários finais não precisarão do suporte da TI após a configuração das fontes de dados no Tableau. O objetivo desse trabalho é fornecer uma solução, e não uma análise precisa dos dados referentes ao passado, uma vez que os responsáveis pela base de dados do SIE nem sempre mantêm um padrão em algumas informações, pois as mesmas não passavam por nenhum tipo de análise. Com a adoção da nossa solução, será necessária uma atenção maior no preenchimento desses dados para que os resultados gerados estejam de acordo com a realidade.

Palavras-chave: INDICADORES, BI, TABLEAU, DESEMPENHO, COTAS.

ABSTRACT

This thesis presents an analysis of students' admission to the Federal University of the State of Rio de Janeiro. In Brazil, school dropout, understood as interruption in the study cycle, causes significant losses under the economic, social and human aspects, regardless of the level of education. The environment among institutions of higher education is competitive and the economic viability has a strong link with the permanence of students in the programs. Using the micro data from the SIE - Information System for Teaching, and the SiSU - Unified Selection System, a visualization solution was created in the Tableau Desktop tool, with the metrics definition that generated indicators to help the institution's understanding possible causes that can generate avoidance, withdrawal and mobility of students between courses, making possible the decision of the Pro-Rector's Office of Graduation.

Because it is a Self-Service BI tool, the final users will not need IT support after setting up data sources in Tableau. The purpose of this paper is to provide a solution, not an accurate analysis of past data, since those responsible for the SIE database do not always maintain a standard in some information, and they did not undergo for any type of analysis. With our solution, it will be necessary to pay more attention when filling this data, so the results generated are in line with the reality.

Key words: INDICATORS, BI, TABLEAU, PERFORMANCE, QUOTES.

Índice

1	Introdução	9
1.1	Motivação.....	9
1.2	Objetivos	10
1.3	Organização do texto.....	11
2	Contexto de Estudo	12
2.1	Sobre a UNIRIO.....	12
2.2	Sobre a necessidade da Pró-Reitoria de Graduação da UNIRIO	15
2.3	Definição de tópicos para análise.....	17
3	Conceitos de Inteligência do Negócio	18
3.1	Introdução	18
3.2	Data Warehouse	19
3.2.1	ETL : Extract, Transform, Load	19
3.3	Modelagem Dimensional	20
3.3.1	Data Warehouse.....	21
3.3.2	Data Warehouse.....	21
3.3.3	OLAP – On-Line Analytical Processing	22
3.4	Self-Service BI: o conceito e suas vantagens.....	23
3.5	Perfil do Usuário Final	24
3.6	Tableau.....	25
4	Proposta de Solução.....	30
4.1	Fonte de Dados.....	30
4.2	Arquitetura da Solução.....	35
4.3	Definição de Métricas e Dimensões.....	35
4.4	Análise de Resultados	38
5	Conclusão.....	47
5.1	Conclusões sobre os Indicadores obtidos.....	47

5.2	Considerações Finais.....	47
5.3	Limitações do Projeto	48
5.4	Trabalhos Futuros	49
6	Apêndice	51
6.1	Introdução	51
6.2	Reunião para entendimento do processo de ingresso na UNIRIO.....	51
6.3	Reunião para definição de requisitos da solução	52
6.4	Reunião para dúvidas	53

Índice de Tabelas

Tabela 1 - Campos dos relatórios do SIE	34
Tabela 2 - Campos dos relatórios do SiSU.....	34
Tabela 3 - Campos chave para ligar tabelas do SiSU e SIE	35

Índice de Figuras

Figura 1 - Percentual de Pretos e Pardos entre os concluintes (Fonte: http://educacao.uol.com.br/noticias/2013/05/20/negros-ainda-sao-minoria-entre-formados-no-ensino-superior.htm)	14
Figura 2 - Conexão de dados no Tableau Desktop	26
Figura 3 - Tipos de Join	27
Figura 4 - União de fontes de dados no Tableau	28
Figura 5 - Tipos de dados representados no Tableau Desktop	28
Figura 6 - Painel de Dados no Tableau Desktop	29
Figura 7 - Modelo de Dados	32
Figura 8 - Exemplo de relatório do SIE, referente ao CCET	32
Figura 9 - Exemplo de Planilha extraída do SiSU utilizando o SiSU Gestão	33
Figura 10 - Exemplo de Planilha de chaves ID_PESSOA (SIE) e CPF_CODIFICADO (SiSU), gerada pela DTIC.	33
Figura 11 - Arquitetura da Solução	35
Figura 12 - Configuração de Dimensões no Tableau	36
Figura 13 - Configuração de Metricas no Tableau	37
Figura 14 - Indicador de Evasão por Curso	39
Figura 15 - Índice de Evasão filtrado por cursos específicos	40
Figura 16 - Índice de Evasão de um curso específico	40
Figura 17 - Indicador de Evasão por Formas de Ingresso	41
Figura 18 - Formas de Ingresso agrupadas	42
Figura 19 - Indicador de evasão por localidade, com visualização em mapa	43
Figura 20 - Informações de uma região específica escolhida	43
Figura 21 - Média do Coeficiente de rendimento de alunos que evadem	44
Figura 22 - Comparação de evasão por CR entre Medicina e Ciências Biológicas	45
Figura 23 - Comparação de Evasão por CR entre turnos diferentes de um mesmo curso	45
Figura 24 - Relação da evasão com o Coeficiente de Rendimento e o município de origem do aluno	46

1 Introdução

1.1 Motivação

Segundo o Censo da Educação Superior de 2014, o número de vagas ociosas nas instituições federais de ensino no Brasil está em torno de 114 mil. Um dos problemas relacionados às vagas ociosas é a evasão (Ministério da Educação, 2015). A evasão possui consequências para a uma universidade, pois as vagas que são desocupadas após o término das chamadas do SiSU - o Sistema de Seleção Unificada¹ para acesso a vagas em universidades no Brasil, não são ocupadas novamente, e a universidade precisa abrir editais de transferência interna ou externa para talvez preencher algumas.

A Universidade Federal do Estado do Rio de Janeiro, Fundação Pública, instituída nos termos da Lei nº 6.655, de 05 de junho de 1979, com sede e foro na cidade do Rio de Janeiro, popularmente conhecida como UNIRIO, é uma destas instituições de ensino superior. O índice de vagas ociosas na UNIRIO é bem alto. Em 2012, segundo o jornal O Globo, havia 1.273, onde 19 eram de Medicina, curso mais concorrido no Brasil, em qualquer universidade que o possui.

Durante uma entrevista com a Pró-Reitoria de Graduação (PROGRAD) da UNIRIO, detalhada no Apêndice 1 desse trabalho, foram feitos relatos sobre a quantidade de vagas ociosas na universidade, por conta do alto índice de evasão de alunos. Portanto, é de interesse desse setor, entender motivos que levam os alunos a evadirem, e em quais centros de ensino isso ocorre com maior frequência, para tomar medidas preventivas e corretivas. Desta forma, de acordo com essa demanda, foi desenvolvida uma pesquisa utilizando dados do SiSU e do SIE (Sistema de Informações para o Ensino), como base para a criação de indicadores, a fim de obter respostas para as questões de interesse da Pró-Reitoria de Graduação da UNIRIO, facilitando assim o processo de tomada de decisão dos responsáveis desse setor.

¹<http://sisu.mec.gov.br/>

Do ponto de vista de Sistemas de Informação, a solução proposta envolve análise de dados seguindo métricas, com a utilização de uma ferramenta que torna mais fácil a visualização de informações para os usuários finais, sem dependência de uma equipe especializada em TI no dia-a-dia. A partir de uma demanda real, foram definidos os indicadores, de forma que possam ser alimentados por planilhas com campos estabelecidos de acordo com os requisitos. Os usuários poderão incluir semestralmente as planilhas seguindo esse modelo e fazer as análises desejadas de forma simples e rápida. Isso é possível, pois foi seguido o conceito de Self-Service BI, que dentre as opções de BI do mercado, é o que torna o usuário final mais independente possível de uma equipe especializada em TI.

As ferramentas mais indicadas nesse conceito são QlikSense, Tableau e o Power BI da Microsoft. As três ferramentas possuem a mesma finalidade, porém a Tableau Desktop é uma tecnologia inovadora que transforma dados em gráficos ou tabelas de forma rápida e prática e pode ajudar a visualizar padrões e identificar tendências. Apresenta facilidade na configuração e é de operação bastante amigável ao nível do usuário departamental. Foi escolhida também por já possuir credenciamento com a UNIRIO, o que possibilita a obtenção de licenças gratuitas.

1.2 Objetivos

A partir da demanda da Pró-Reitoria de graduação, foi decidido desenvolver uma solução que auxilie o levantamento de informações semestrais acerca de ingressos e abandonos nos cursos da UNIRIO. Por se tratar de uma equipe pequena, onde nem todos conhecem Tecnologias de Informação (TI), foi escolhida uma ferramenta de Self-Service BI (Business Intelligence), a fim de possibilitar o usuário final de operar sem a necessidade de suporte da TI, como ocorre na maioria das ferramentas de BI.

Esta solução pretende prover suporte para os responsáveis pelo ingresso de alunos nas tomadas de decisões, a fim de evitar o aumento do número de evasões na Universidade e conseqüentemente, diminuir o número de vagas ociosas.

1.3 Organização do texto

O presente trabalho está estruturado em capítulos e, além desta introdução, será desenvolvido da seguinte forma:

- Capítulo II: Contexto de Estudo: Contextualização do ambiente que será estudado e suas principais particularidades, tais como formas de ingresso e formas de evasão da Universidade, detalhamento da demanda dos responsáveis institucionais e os tópicos definidos para análise.
- Capítulo III: Escolha da Ferramenta: Perfil do usuário final da solução, fundamental para a escolha da ferramenta. Paralelo entre BI e Self-Service BI, justificando a preferência do segundo para este trabalho, e a devida ferramenta escolhida. Apresentação do dicionário de dados que são analisados.
- Capítulo IV: Solução: Detalhamento de métricas e indicadores desenvolvidos, gráficos escolhidos para a visualização dos dados e análise de alguns resultados obtidos.
- Capítulo V: Conclusões – Apresentação das considerações finais, dificuldades encontradas ao longo do desenvolvimento do projeto, contribuições da pesquisa e possíveis trabalhos futuros.
- Capítulo VI: Apêndice – Detalhamento de todas as reuniões feitas com a PROGRAD para entendimento da necessidade e levantamento de requisitos para a solução.

2 Contexto de Estudo

2.1 Sobre a UNIRIO

A UNIRIO é uma instituição de ensino superior composta por alguns centros de estudo, são eles: Centro de Ciências Exatas e Tecnológicas, Centro de Ciências Humanas, Centro de Letras e Artes, Centro de Ciências Biológicas e da Saúde, Centro de Ciências Jurídicas e Políticas, e Instituto Biomédico. Para ingressar na universidade existem algumas opções de convênios e exames. Atualmente, o exame oficial para ingresso é o ENEM - Exame Nacional do Ensino Médio, uma prova elaborada pelo Ministério da Educação para verificar o domínio de competências e habilidades dos estudantes que concluíram o ensino médio, composto por quatro provas de múltipla escolha, com 45 questões cada, e uma redação. Desde 2009, a nota obtida nesse exame pode ser utilizada para inscrição no SiSU- Sistema de Seleção Unificada², que é um sistema informatizado, gerenciado pelo Ministério da Educação (MEC), no qual instituições públicas de ensino superior oferecem vagas para candidatos participantes do Enem, sendo assim possível concorrer a uma vaga em um curso na UNIRIO. Os alunos que ingressam dessa forma, podem concorrer às seguintes vagas, mediante a documentação para comprovação quando necessário:

- Ação afirmativa - professor de rede pública: Professores possuem cotas em cursos de licenciatura;
- Ampla Concorrência: Alunos que se inscrevem sem participar de cotas;
- Cota para estudantes de escola pública independente de renda;
- Cota étnica (alunos pretos/pardos/indígenas) para estudantes de escola pública independente de renda;

²<http://sisu.mec.gov.br/>

- Cota étnica (alunos pretos/pardos/indígenas) para estudantes de escola pública com renda familiar de até 1,5 salário mínimo;
- Cota para alunos de Escola Pública com renda familiar de até 1,5 salário mínimo.

O sistema de cotas nas instituições de ensino superior previsto na lei nº 12.711/2012, implementado na UNIRIO em 2013, ocorre basicamente por dois motivos, segundo informações coletadas na entrevista detalhada no apêndice: pagamento de dívida histórica, uma vez que os grupos mencionados acima sofreram racismo em diversos momentos na história, e a tentativa de mudança do lugar social desses grupos nas universidades. Segundo o Portal Brasil³, o Percentual de negros no ensino superior é metade dos brancos. A porcentagem de pretos e pardos concluintes nas instituições é ainda menor, segundo dados fornecidos pelo INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, uma entidade pública federal vinculada ao MEC, do ENADE, Exame Nacional de Desempenho de Estudantes que avalia o rendimento dos alunos dos cursos de graduação, ingressantes e concluintes, em relação aos conteúdos programáticos dos cursos em que estão matriculados.

³ <http://portal.mec.gov.br/ultimas-noticias/212-educacao-superior-1690610854/32581-ministerio-anuncia-novo-sistema-para-preencher-vagas-ociosas>

PERCENTUAL DE PRETOS E PARDOS ENTRE CONCLUINTES

Medicina	2,66%
Design	3,58%
Odontologia	3,8%
Medicina Veterinária	4,26%
Farmácia	5,02%
Direito	5,03%
Comunicação Social	5,11%
Administração	5,26%
Psicologia	5,38%
Ciências Econômicas	5,5%

Figura 1 - Percentual de Pretos e Pardos entre os concluintes (Fonte: <http://educacao.uol.com.br/noticias/2013/05/20/negros-ainda-sao-minoria-entre-formados-no-ensino-superior.htm>)

Vale ressaltar que, alunos se autodeclararam pretos, pardos ou indígenas no momento da inscrição no Enem e na vaga do curso desejado, não havendo qualquer tipo de comprovação. Já as cotas que envolvem renda familiar, devem ser comprovadas, e se encontrada qualquer informação por parte da comissão de avaliação que comprove o contrário, o aluno não terá direito à vaga.

Para diminuir esse índice, os diretores da Coordenadoria de Acompanhamento e Avaliação do Ensino de Graduação (CAEG) e da Diretoria de Políticas, Normatização e registros acadêmicos de graduação (DIPRAG) relataram que existem editais de programas de permanência para alunos que possuem dificuldades de arcar com despesas de moradia e transporte, e editais de transferência externa, onde alunos de outras

universidades podem ocupar uma vaga ociosa na UNIRIO se obtiverem nota na avaliação desse processo. No entanto, muitas vagas ainda permanecem ociosas, uma vez que o abandono de curso é algo frequente e os editais de transferência externa não.

2.2 Sobre a necessidade da Pró-reitoria de Graduação da UNIRIO

A Pró-Reitoria de Graduação da UNIRIO é composta por várias coordenadorias e diretorias, dentre elas, a Coordenadoria de Acompanhamento e Avaliação do Ensino de Graduação (CAEG) e a Diretoria de Políticas, Normatização e Registros Acadêmicos de Graduação (DIPRAG).

A Diretoria de Políticas, Normatização e Registros Acadêmicos de Graduação ⁴tem as seguintes atribuições:

1. Manter um fluxo contínuo de informações atualizadas sobre legislação do ensino superior;
2. Assessorar os setores envolvidos com a administração e desenvolvimento dos cursos de graduação da UNIRIO no que tange a legislação e normas;
3. Gerar apoio técnico, administrativo e pedagógico aos coordenadores de curso, visando a dinamização e efetivação do papel de articulador do projeto pedagógico que desempenham;
4. Atualizar informações sobre os cursos e serviços, desenvolvidos no âmbito do ensino de graduação e aperfeiçoar os instrumentos que a vinculam;
5. Estimular a reflexão sobre a avaliação dos processos de ensino e de aprendizado, procurando evidenciar, na complexidade e peculiaridade que os caracteriza, a importância de sua função como recurso auxiliar para a melhoria do ensino.

A Coordenadoria de Acompanhamento e Avaliação de Ensino de Graduação é órgão diretamente ligado à Pró-Reitoria de Graduação e tem as seguintes atribuições:⁵

1. Assessorar os setores envolvidos com o desenvolvimento e a avaliação permanente dos cursos de graduação da UNIRIO;
2. Gerar apoio técnico, administrativo e pedagógico aos coordenadores de curso, visando à dinamização e efetivação do papel de articulador do projeto pedagógico que desempenham;

⁴ <http://www.unirio.br/prograd/unidades-administrativas>

⁵ <http://www.unirio.br/prograd/unidades-administrativas>

3. Desenvolver programas e projetos voltados para a melhoria do ensino de graduação ofertado pela UNIRIO;
4. Manter atualizadas as informações sobre os cursos desenvolvidos no âmbito do ensino de graduação e aperfeiçoar os instrumentos que as veiculam: guia acadêmico, catálogo de cursos de graduação, entre outros;
5. Acompanhar as matrículas especiais de estudantes nos cursos de graduação da UNIRIO, bem como de alunos regulares.

Em visita Pró-Reitoria de graduação da UNIRIO, os diretores da CAEG e da DIPRAG, relataram a falta de informação mais detalhada sobre os números de evasão em cada curso, e o perfil dos alunos que evadem (mais detalhes da demanda podem ser vistos nos relatos sobre as entrevistas realizadas no Apêndice 1). Essas informações ajudariam a entender possíveis motivos que levam alunos a evadir. Dentre várias possibilidades, os motivos mais comuns para evadir são a dificuldade de locomoção, a escolha por um curso que não é do interesse do aluno, e a falta de recursos para arcar com despesas de material didático.

Os cursos de graduação são constantemente avaliados pelo MEC segundo um “Índice de Sucesso” que está diretamente relacionado com o Índice de Evasão do mesmo. Se o Índice de Evasão está muito alto em um curso, o Índice de Sucesso conseqüentemente está baixo e a universidade não pode receber novos alunos no mesmo até corrigir esse problema. Se o problema não for corrigido, o curso pode até mesmo ser fechado. Sendo assim, a universidade possui o desafio de aplicar diversas medidas que podem ou não diminuir o problema, porém, a causa em si não é tratada e pode diminuir o índice de evasão em uma universidade ao mesmo tempo que aumenta em outra.

O cadastro do aluno, sua forma de ingresso na universidade, forma de evasão e coeficiente de rendimento, ficam armazenados no SIE – Sistema de Informações para o Ensino, sistema que possui cadastro dos alunos, forma de ingresso e evasão se houver, tal como o desempenho acadêmico. Outras informações relevantes como o endereço do aluno antes de se mudar para perto da universidade (quando há mudança), na base de dados do SiSU.

Com a obtenção de uma informação mais concreta, a partir dessas fontes de dados e até mesmo possíveis cruzamentos entre elas, seria possível acompanhar os números semestralmente e aplicar medidas preventivas e corretivas sobre a causa real do

problema do alto índice de vagas ociosas a fim de manter o Índice de Sucesso dos cursos.

2.3 Definição de tópicos para análise

Após o entendimento da demanda, conforme detalhamento no Apêndice1, foram selecionados, juntamente com a PROGRAD, alguns tópicos para análise:

- Índice de Abandono de Curso por Curso e por Centro, a fim de entender quais cursos e centros possuem maior índice de abandono para poder corrigir possíveis problemas no ensino dos mesmos.
- Forma de Ingresso por Forma de Evasão, a fim de analisar se são alunos cotistas (e tipos de cota) ou ampla concorrência os que mais evadem ou concluem os cursos.
- Perfil Geográfico de Alunos que abandonam cursos, a fim de entender se a distância da moradia para a universidade tem influência na evasão.
- Coeficiente de Rendimento de Alunos que abandonam um Curso, a fim de entender se o baixo desempenho influencia na evasão.

Para essas análises, serão levadas em conta as formas de ingresso: Siso - Ampla Concorrência e Siso – Cotas, e Forma de Evasão: Abandono de Curso, Cancelamento e Jubilamento. Conforme reuniões detalhadas no Apêndice 1, a diferença entre essas formas de evasão é que no Abandono de Curso, o aluno simplesmente não vai mais às aulas e não renova sua matrícula. No Cancelamento, ele solicita o cancelamento da mesma. O jubilamento ocorre quando a universidade não permite que o aluno renove sua matrícula devido ao coeficiente de rendimento abaixo de 4.0 ou quatro reprovações em uma mesma matéria. Essas são as formas de evasão mais relevantes pois resultam em vagas ociosas.

Cursos que envolvem Teste de Habilidade específica como avaliação complementar ao ENEM não serão estudados.

A próxima etapa foi a escolha de uma tecnologia de informação que melhor atendesse à necessidade dos usuários. Para isso, analisamos o perfil destes usuários, seus recursos, conhecimentos e rotina, para propor e desenvolver uma solução que não trouxesse nenhum tipo de dificuldade. Por isso, escolhemos uma ferramenta de Self-Service BI.

3 Conceitos de Inteligência do Negócio

3.1 Introdução

Apesar de ser um termo recente, o princípio de Business Intelligence é usado há milhares de anos. Fenícios, persas, egípcios e outros povos do Oriente Médio cruzavam informações obtidas junto à natureza em benefício próprio. Observar e analisar o comportamento das marés, os períodos de seca e chuvas, a posição dos astros, entre outras, eram formas de obter informações que eram utilizadas para tomar as decisões que permitissem a melhoria de vida de suas respectivas comunidades (BARBIERI – 2001).

O termo Business Intelligence foi utilizado então pela primeira vez na década de 50 por Hans Peter Luhn em um artigo intitulado “A Business Intelligence System” (ELENA 2011), que propunha o desenvolvimento de um Sistema automático, baseado em máquinas de processamento de dados. Segundo Luhn (1958), o sistema abordaria coleta ou aquisição de novas informações; disseminação; armazenamento; recuperação e transmissão de informações.

Podemos caracterizar Business Intelligence como o processo de coleta, organização, transformação, análise, compartilhamento e monitoramento de dados e informações que oferecem suporte para a gestão de negócios. Basicamente, BI é um conjunto de teorias, metodologias, processos e tecnologias que transformam uma grande quantidade de dados brutos em informações úteis para tomadas de decisões estratégicas.

3.2 Data Warehouse

De acordo com Barbieri (2001), um Data Warehouse, que pode ser traduzido como armazém de dados, é um banco de dados destinado a sistemas de apoio à decisão e cujos dados foram armazenados em estruturas lógicas dimensionais, permitindo o seu processamento analítico por ferramentas especiais. Esse tipo de banco de dados é projetado para suportar grandes demandas de processamento, isso porque manipula elevadas quantidades de dados provenientes de diversas fontes.

Em um ambiente de Data Warehouse, as informações analíticas e estratégicas são tratadas por meio de uma modelagem dimensional a fim de centralizar e otimizar as consultas requeridas.

3.2.1 ETL : Extract, Transform, Load

ETL (extract, transform, load) é um processo de qualquer projeto de BI. A extração é a etapa do processo no qual os dados das fontes são lidos e gravados em um banco de dados denominado Staging. Este é considerado um banco de dados sujo, pois as informações gravadas nele possuem exatamente o mesmo formato das informações das fontes.

Na transformação, os dados do banco de dados Staging são lidos e sofrem alterações e adaptações em seus formatos para serem adequados ao Data Warehouse.

Por fim, na carga, os dados preparados na etapa transformação são escritos nas tabelas fato e nas tabelas de dimensão. Os dados são carregados efetivamente no Data Warehouse e prontos para serem consumidos por relatórios, dashboards e serviços de análise.

Além dessas três etapas comuns do processo, outros procedimentos também são importantes, como a limpeza dos dados primeiramente, evitando o carregamento de dados contraditórios e irrelevantes, garantindo que somente o que é relevante será carregado no modelo Data Warehouse. O segundo procedimento é a integração de dados de múltiplas fontes, permitindo quantidades significativas de cruzamento de dados, proporcionando resultados mais relevantes para uma futura tomada de decisão.

Por fim, algo muito relevante é a atualização dos dados e a manutenção do histórico de dados, já que para a tomada de decisão é importante confrontar dados atuais com dados do passado.

3.3 Modelagem Dimensional

A modelagem dimensional é uma técnica de concepção e visualização de um modelo de dados de um conjunto de medidas que descrevem aspectos comuns de negócios. É utilizada especialmente para sumarizar e reestruturar dados e apresentá-los em visões que suportem a análise dos valores desses dados [MACHADO 2004].

Em um sistema de processamento analítico, o grande volume de dados necessários para as consultas de planejamento estratégico, devem ser processadas de forma rápida. Para garantir esse desempenho, existe uma redundância planejada de dados, compensando gastos com armazenamento e atualização das informações. Dessa maneira, configura-se uma estrutura simples, com tabelas de dados históricos em séries temporais, descritos através de tabelas de dimensões, de modo que o modelo reflita o processo de análise do negócio.

A modelagem dimensional é representada por tabelas fatos e dimensões, onde tabelas fatos contemplam os cálculos, indicadores ou quantidades dos dados que serão analisados, como por exemplo, em uma tabela que registre as vendas de uma empresa pode conter a quantidade de vendas do negócio. Uma tabela fato é composta por uma chave primária (formada por uma combinação única de valores de chaves da dimensão) e pelas métricas de interesse para o negócio.

Já as tabelas que representam as dimensões apresentam os atributos e contêm a descrição do negócio. Seus atributos configuram-se como as restrições das consultas ou agrupamento dos resultados, por exemplo, tipo de produto vendido por uma empresa.

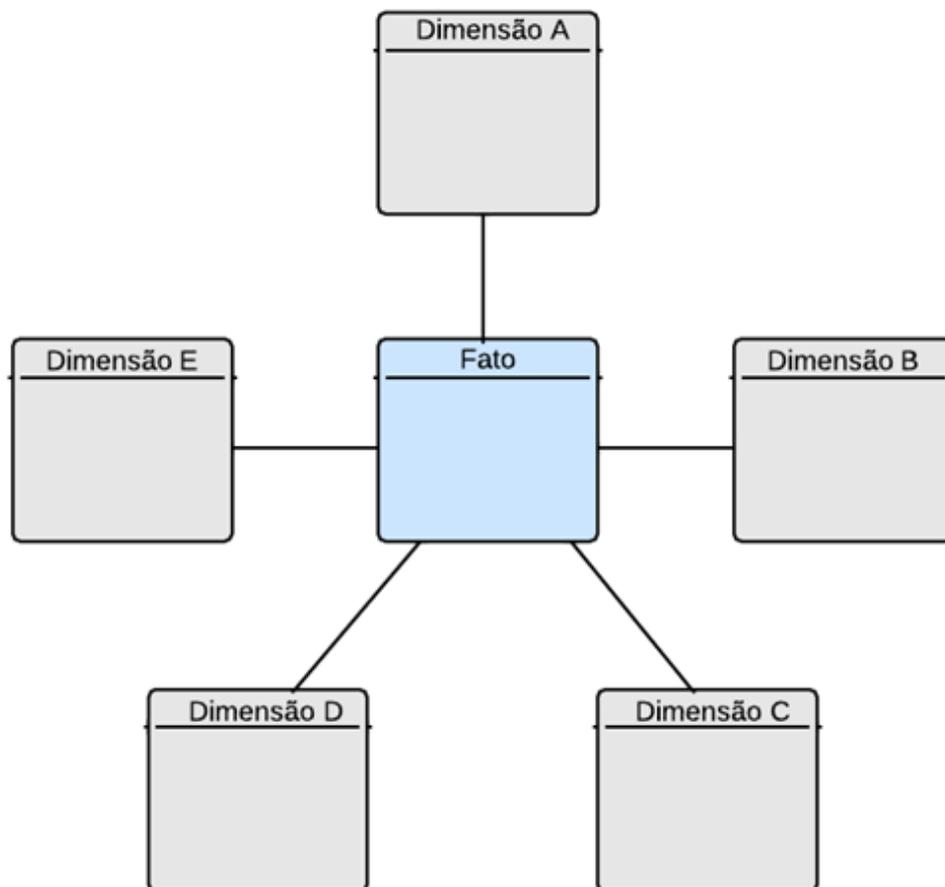
Outro componente importante em um modelo dimensional são as métricas. Estas são as informações gravadas nas tabelas fato que permitem quantificar ou medir o desempenho do negócio. As métricas geralmente são numéricas e podem ser ou não agregadas, sendo na maioria das vezes do tipo aditivas, isto é, permitem operações como adição, subtração e médias. Além de aditivas, as métricas podem ser semi-aditivas ou não aditivas. Esta última não pode ser manipulada livremente, como por exemplo, valores

percentuais ou relativos, enquanto que as métricas semi-aditivas representam os valores que não podem ser somados em todas as dimensões.

3.3.1 Data Warehouse

Modelo Estrela (*Star Schema*)

O modelo estrela de um Data Warehouse configura-se por dados desnormalizados, onde uma tabela fato centralizada se relaciona com as dimensões, como mostra a figura a seguir:

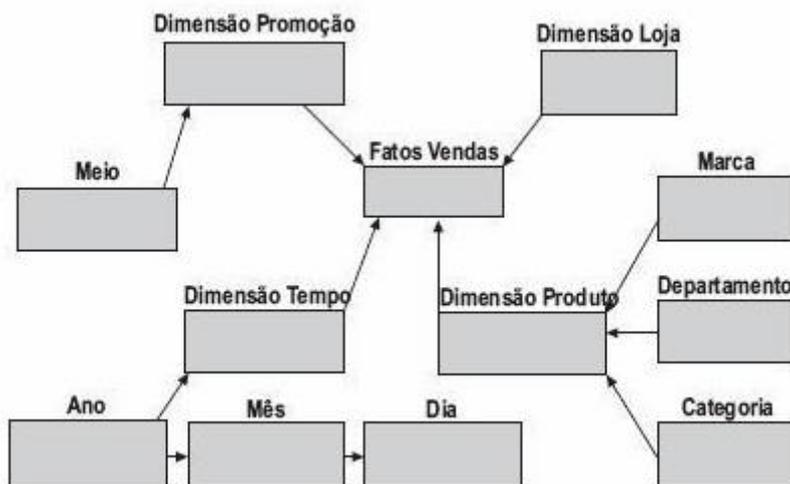


3.3.2 Data Warehouse

Modelo Floco de Neve (*Snowflake Schema*)

No modelo floco de neve as dimensões se relacionam com a tabela fato, porém algumas dimensões se relacionam apenas entre elas, o que ocorre para fins de normalização das tabelas dimensões. Essa técnica visa diminuir o espaço ocupado por estas tabelas.

A figura a seguir exemplifica um modelo floco de neve:



3.3.3 OLAP – On-Line Analytical Processing

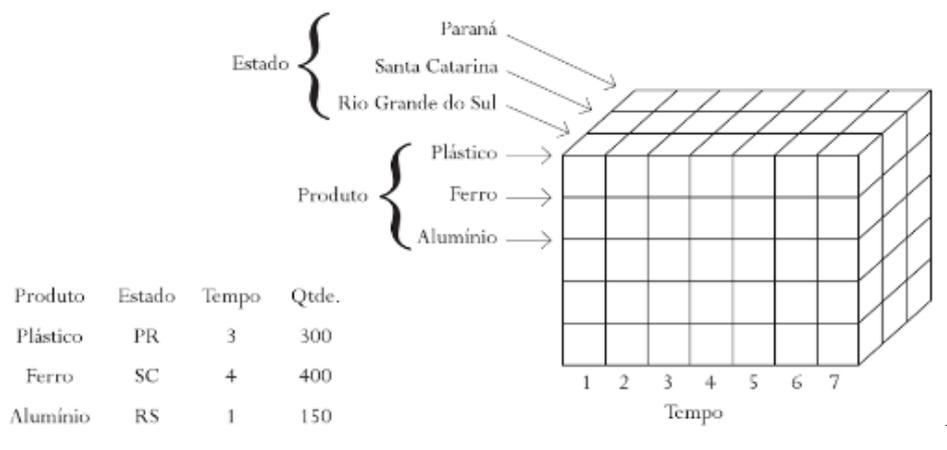
Os sistemas do ambiente operacional que geralmente apoiam os usuários nas funções cotidianas são chamados OLTP (On Line Transaction Processing). Seu principal objetivo é executar o maior número de transações possíveis no menor tempo de processamento.

Apesar do poder de processamento de uma grande quantidade de transações em pouco tempo, esses sistemas se apresentam pouco flexíveis para o desenvolvimento de relatórios e consultas complexas, devido às limitações impostas pelo modelo de dados e pela linguagem SQL.

Para atender a necessidade de um sistema de suporte a decisão, onde há um grande volume de dados e uma demanda complexa de informações, são adotados os sistemas OLAP (On Line Analytical Processing).

Segundo Inmon (1999), OLAP é uma tecnologia de software que permite a analistas, gerentes e executivos, a obtenção de dados de uma forma rápida, consistente e com acesso interativo para uma diversidade de possíveis visões da informação na empresa. Resumidamente, OLAP possui um conjunto de funcionalidades que facilita a análise multidimensional. Ferramentas OLAP se caracterizam por possuir uma grande

capacidade de responder rapidamente às solicitações dos usuários e a facilidade de trabalhar com quantidades volumosas de dados, além da interatividade que oferecem. A análise dos dados é feita de forma multidimensional. Ao utilizar o recurso de banco de dados, ocorre a limitação pelo fato do modelo ser dimensional, por meio da busca de informações nas linhas e colunas de uma tabela. Portanto, uma aplicação multidimensional tende a retornar ao tomador de decisão informações mais valiosas. As análises são definidas em formato de cubos, conforme figura X



3.4 Self-Service BI: o conceito e suas vantagens

Levando em consideração a necessidade das organizações por mais praticidade e agilidade na obtenção de informações, foi criada a abordagem SS BI – Self-Service BI, onde é disponibilizado aos usuários finais um ambiente intuitivo, eficiente, onde eles conseguem gerar relatórios facilmente, onde a equipe de TI apenas configura as fontes de dados no ambiente, e o restante pode ser feito pelo usuário final sem a necessidade de suporte como nas ferramentas de BI tradicional.

Até pouco tempo atrás, BI era principalmente para especialistas. As ferramentas eram difíceis de operar e exigiam um nível de conhecimento semelhante ao dos administradores de banco de dados. Os gerentes de negócios que precisavam de informações sobre o negócio enviavam pedidos de consulta a especialistas de BI que trabalhavam no departamento de TI ou, esperavam por um relatório de BI mensal/trimestral que o mesmo liberava, preenchido com consultas padrão difíceis de manipular. Tudo isso mudou porque os avanços na tecnologia de banco de dados e consulta (assim como ferramentas de front-end intuitivamente redesenhadas tornaram

mais fácil para as pessoas do negócio usar ferramentas de BI da categoria Self-Service BI. A ideia por trás desse conceito, é que os executivos e gerentes de negócios sejam capazes de se instalar e operar rapidamente com essas ferramentas, sem ter um diploma em estatísticas ou um histórico de análise de dados. Eles também devem ser capazes de gerar suas próprias consultas em cima de conjuntos de dados corporativos sem ter conhecimentos de TI. O objetivo é que essas ferramentas de SS BI, traduzidas como BI de auto-atendimento, sejam tão fáceis de usar quanto planilhas.

O BI de auto-atendimento de forma nenhuma invalida gerenciamento de banco de dados tradicional ou profissionais de TI, afinal as técnicas que esses profissionais fornecem são complexas de gerar e úteis para muitas organizações. Porém, o BI de auto-atendimento tenta gerar novos *insights* através de perspectivas do negócio, realizando novos valores de dados através de análises mais informais.

O QlikSense é uma ferramenta bastante similar ao Tableau, ambas possuem interface bastante amigável e possibilidade de visualização de dados de forma rápida e prática. Ambos possuem também muitos conectores nativos para a maioria dos bancos de dados, além de conectores para mídias sociais, fontes externas e uma série de outras possibilidades. Porém, o QlikSense se destaca no item “integração dos dados” em ambientes em que existem muitos sistemas e estes acabam sendo integrados e modelados dentro da ferramenta. Para o Tableau, a base de dados precisa estar mais “pronta”, com o ETL sendo feito fora da ferramenta. Como a etapa de integração não é relevante para este projeto, e a base de dados da PROGRAD vêm em forma de planilha do Excel, essa qualidade do QlikSense se torna irrelevante.

O Power BI da Microsoft oferece um amplo suporte a análises de BI e Self-Service BI. São diversas aplicações integradas que possibilitam uma solução robusta de tratamento e análise dados. Porém, por possuir uma complexidade um pouco maior e considerando o trabalho para o usuário final, não se adequaria para a solução proposta, tornando o usuário mais dependente de treinamentos e suporte.

3.5 Perfil do Usuário Final

Os usuários que demandaram a solução proposta nesse trabalho foram o diretor da Coordenadoria de Acompanhamento e Avaliação do Ensino de Graduação da UNIRIO – CAEG, na Pró-Reitoria de Graduação – PROGRAD, e o Professor associado

responsável pela Diretoria de políticas, normatização e registros acadêmicos de Graduação – DIPRAG, também na PROGRAD.

A PROGRAD é composta por várias pequenas equipes, e nenhuma delas é de TI. As possibilidades de soluções mais comuns necessitariam de suporte terceirizado de profissionais de TI com uma certa frequência. Por se tratar de uma equipe pequena, onde uma das pessoas possui noções básicas de Lógica e TI, a solução ideal é aquela que não gera dependência de suporte e manutenção. Essa possibilidade é viável com a utilização de ferramentas de Self-Service BI.

3.6 Tableau

A proposta da ferramenta Tableau⁶, conforme publicada no seu site é: *“A Tableau ajuda as maiores organizações do mundo a extraírem o máximo de seus ativos mais valiosos: seus dados e pessoas. Se você estiver tomando decisões que afetam toda a sua organização ou inserindo informações em seu software, aplicativo ou site, escolha o software de análise que trabalha da maneira como as pessoas pensam”* – De acordo com o seu fabricante, para obter o verdadeiro auto-atendimento, é necessária uma plataforma de análise comercial que as pessoas realmente usem na prática, fácil de implantar e gerenciar, que garanta a governança e a segurança dos dados. O Tableau é exatamente essa plataforma: uma solução empresarial que atende às necessidades de toda a empresa.

A escolha dessa plataforma aconteceu devido a facilidade de entendimento das funcionalidades e na forma intuitiva de visualizar dados, criar métricas e obter respostas, sem precisar da manutenção de uma pessoa de TI. O Tableau Desktop pode ser instalado em ambiente Windows e Mac e sua licença é oferecida gratuitamente a estudantes e professores de instituições credenciadas.

3.7 Funcionalidades básicas do Tableau Desktop

A ferramenta possibilita conexão com diversas fontes de dados, sendo elas arquivos de planilha ou texto, banco de dados relacionais, dimensionais ou com big data e até mesmo com dados disponíveis na Web e banco de dados em nuvem. As conexões

⁶<http://www.tableau.com/>

podem ser locais ou em servidores. É possível fazer combinação dessas fontes de dados no próprio Tableau, conforme mostrado na Figura 2 e na Figura 4.

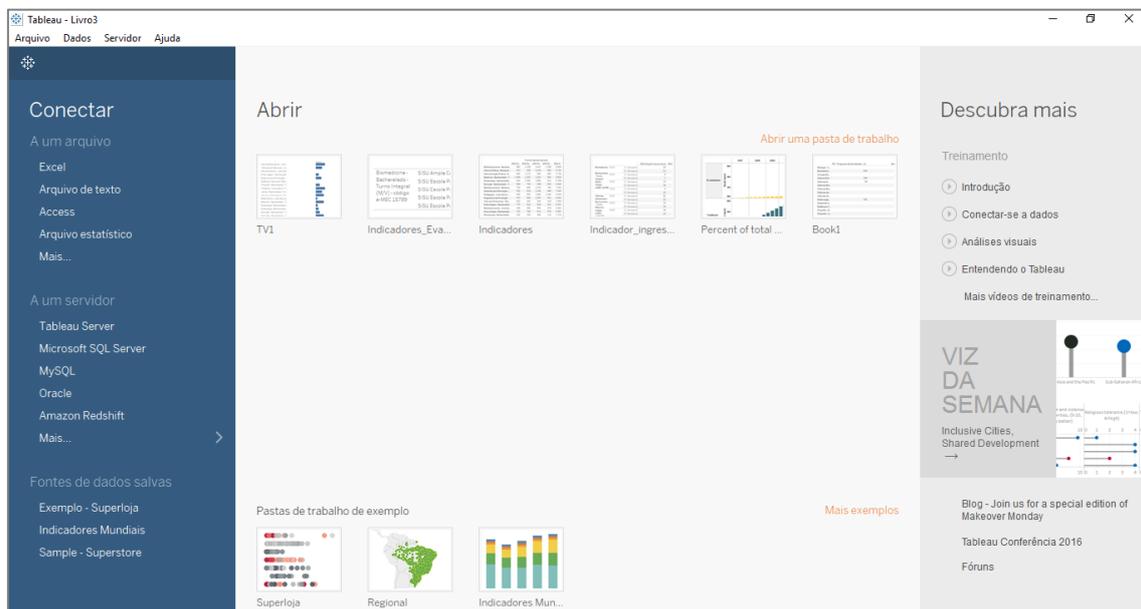
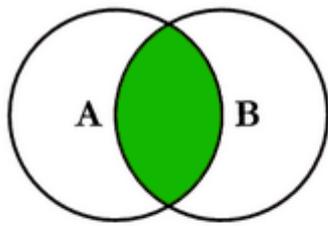


Figura 2 - Conexão de dados no Tableau Desktop

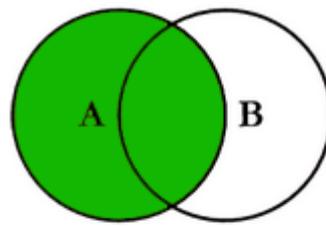
Ao combinar mais de uma fonte de dados, o usuário escolhe o tipo de combinação de dados que deseja fazer por meio de um campo de ligação em comum entre as fontes de dados: junção interna; junção à esquerda; junção à direita e junção externa completa.

Esses tipos de combinações funcionam exatamente como os *joins* da linguagem SQL (Structured Query Language): considerando duas tabelas A e B, por exemplo, o join simples retornará apenas os resultados em comum dessas duas tabelas – que são relacionadas por algum campo de ligação. O *leftjoin*, por sua vez, retorna os resultados em comum e resultados que existem em A, mas não em B, enquanto que o *rightjoin* busca resultados em comum e resultados que existem em B e não em A. Por último, o *fullouterjoin* retorna todos os resultados possíveis – resultados em comum de A e B e resultados de A que não existem em B e vice-versa.

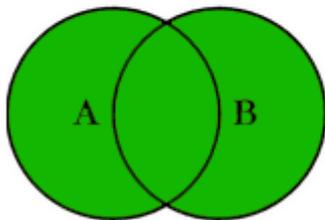
Esse tipo de exemplo pode ser visualizado pela representação de conjuntos matemáticos. Na Figura 3, duas tabelas são representadas por A e B e a parte colorida representa o resultado da junção dessas duas tabelas (nas figuras constam o *join* de SQL e o respectivo nome da junção no Tableau Desktop). Nota-se que a nomenclatura do Tableau utiliza o termo “união” para se referir à junção da linguagem SQL.



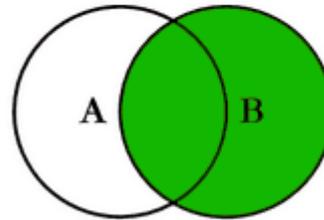
JOIN - UNIÃO INTERNA



LEFT JOIN - UNIÃO À ESQUERDA



FULL OUTER JOIN -
UNIÃO EXTERNA COMPLETA



RIGHT JOIN - UNIÃO À DIREITA

Figura 3 - Tipos de Join

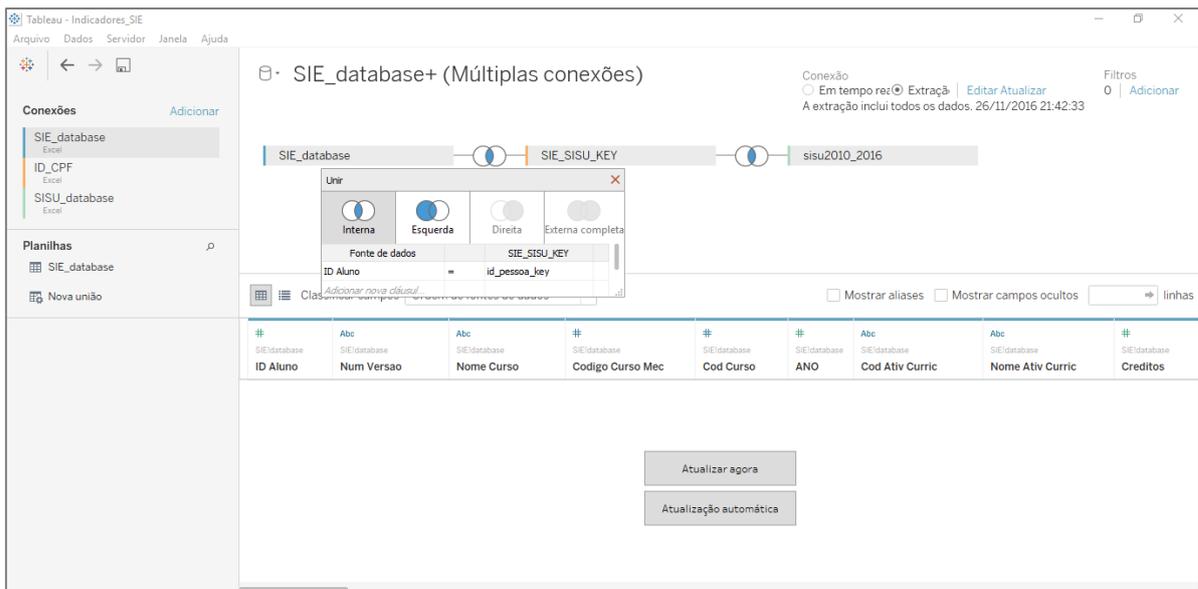


Figura 4 – Combinação de fontes de dados no Tableau

Ao se conectar a uma ou mais fontes de dados no Tableau, a ferramenta interpreta os tipos de dados dos campos, atribuindo os formatos a cada um. Porém, isso pode ser alterado manualmente, caso exista a necessidade.

Os tipos de dados são representados conforme a Figura 5.

Ícone	Tipo de dados
Abc	Valores de texto (cadeia de caracteres)
📅	Valores da data
🕒	Valores de data e hora
#	Valores numéricos
T F	Valores booleanos (apenas relacionais)
🌐	Valores geográficos (usados com mapas)

Figura 5 - Tipos de dados representados no Tableau Desktop

Os dados são organizados, então, no Painel de Dados, como mostra a Figura 6.

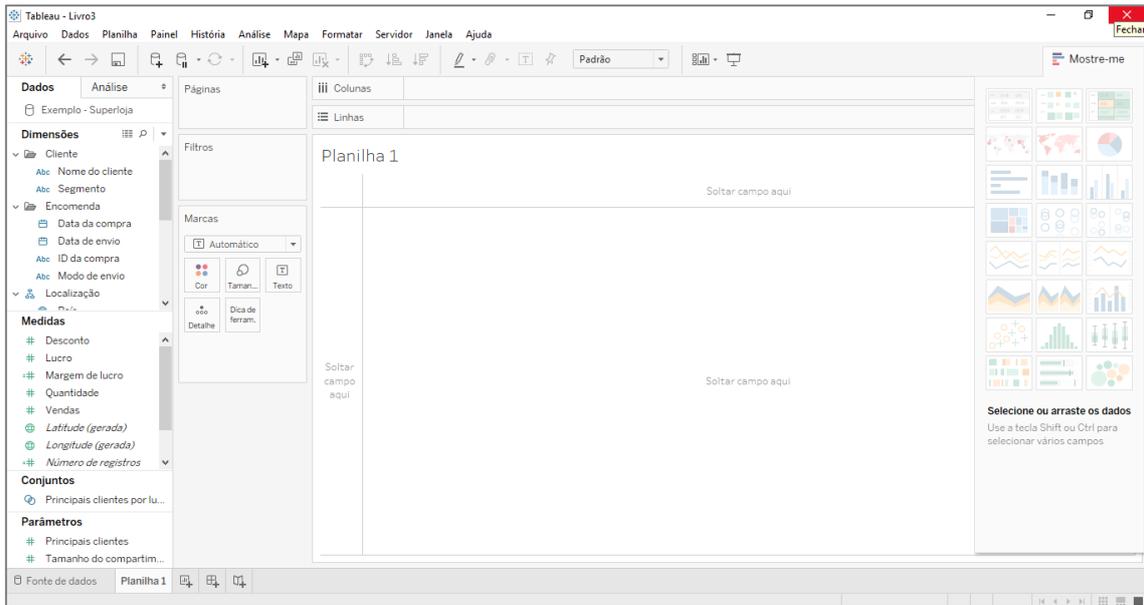


Figura 6 - Painel de Dados no Tableau Desktop

No painel de dados, o Tableau atribui cada campo na fonte de dados com uma de duas funções de dados possíveis: dimensão ou medida e dessa forma organiza esses dados da seguinte maneira:

- **Dimensões:** campos que geralmente contêm dados qualitativos discretos. O Tableau interpreta como dimensão campos que apresentam cadeias de caracteres ou valores booleanos. Por exemplo, datas, forma de ingresso na faculdade, forma de evasão.
- **Medidas:** campos que contêm dados numéricos que podem ser agregados. Por exemplo, quantidade de alunos ingressantes, quantidade de cursos na universidade.
- **Conjuntos:** área opcional que permite armazenar campos personalizados com base nas dimensões disponíveis e critérios especificados pelo usuário.
- **Parâmetros:** variáveis dinâmicas que podem ser criadas pelo usuário.

É possível transformar dimensões em medidas e vice-versa. Além desses tipos de campos, é permitido criar campos calculados a partir de um outro campo por meio de fórmulas que são disponibilizadas pela ferramenta.

4 Proposta de Solução

O presente trabalho se baseia no desenvolvimento de uma solução prática e flexível para análise de indicadores que podem ser gerados a partir dos dados dos sistemas SIE e SiSU da UNIRIO. Para gerar tais indicadores e desenvolver uma solução para o usuário final, a PROGRAD, utilizou-se uma ferramenta de Self-Service BI, onde os cálculos dos indicadores foram previamente desenvolvidos, possibilitando a independência de uma equipe de TI para manipular esses dados constantemente. Basicamente, os dados são extraídos em planilhas pré-definidas pelo usuário dos sistemas SIE e SiSU e são conectadas e processadas no Tableau Desktop, onde os indicadores já definidos poderão ser analisados e manipulados pelos usuários, trazendo respostas e permitindo tomadas decisões de impacto positivo no índice de evasão da UNIRIO.

4.1 Fonte de Dados

As fontes de dados usadas neste trabalho são baseadas em planilhas geradas pela DTIC – Diretoria de Tecnologia de Informação e Comunicação da UNIRIO, de acordo com os campos necessários para análise; planilha gerada pelo sistema do SiSU que cada Instituição de Ensino Superior possui acesso através do SiSU gestão, e um arquivo gerado pela DTIC para possibilitar o cruzamento dos dados do SIE e do SiSU, por meio de codificação do CPF do aluno associado a um código de identificação.

Por questões de segurança, a DTIC não fornece acesso ao banco de dados do SIE para outros setores, então atualmente a única forma de se obter essas informações é através de relatórios que podem ser gerados no próprio SIE pelos responsáveis institucionais. A extração dos dados e a criptografia dos mesmos foi desenvolvida pela DTIC.

Foram utilizadas duas bases de dados (SIE e SiSU) para se obter dados referentes à vida acadêmica do aluno e também informações fornecidas na inscrição do SiSU. Para cruzar essas duas fontes de dados, utilizou-se a planilha de ligação (ID aluno – CPF) para

associar o código do aluno ao CPF. Nos dados extraídos do SIE, apenas o código de identificação do aluno foi disponibilizado e na base do SiSU, o CPF criptografado foi disponibilizado.

O processo de extração das planilhas do SIE feito pela DTIC ocorre a partir da extração de relatórios no próprio SIE, a partir de queries executadas no banco de dados do sistema. O relatório das chaves ID_PESSOA e CPF precisaram ser solicitados a DTIC pois não existe esse modelo de relatório no SIE. A extração dos relatórios do SiSU, também em formato de planilhas, é feita pela própria PROGRAD através de uma ferramenta do MEC que as instituições de ensino superior tem acesso chamada SiSU Gestão. A partir da demanda que originou esse projeto, foi então definido um modelo de dados para um entendimento melhor da solução proposta e para mapear todos os campos e seus respectivos tipos de dados. Como mostra a figura 5, as três planilhas enviadas pelo usuário (uma do SIE, uma do SiSU, e a de chaves que também é original do SIE) dispostas em um modelo simples de dados. Esse será o modelo adotado para os relatórios semestrais necessários para alimentar a solução do Tableau. Na figura 6 podem ser vistas três planilhas, uma extraída do SIE, uma do SiSU, e a que faz a ligação entre elas através das chaves. O modelo mostrado na figura 5 foi criado com base nos campos dessas planilhas, que foram os selecionados como relevantes para as análises. Nas bases de dados do SIE e no SiSU existem muitos outros campos, mas para a análise vigente, não são relevantes. O modelo abaixo considera também a possibilidade de cruzamento de dados permitida pelo Tableau para gerar informações específicas, detalhadas mais a diante.

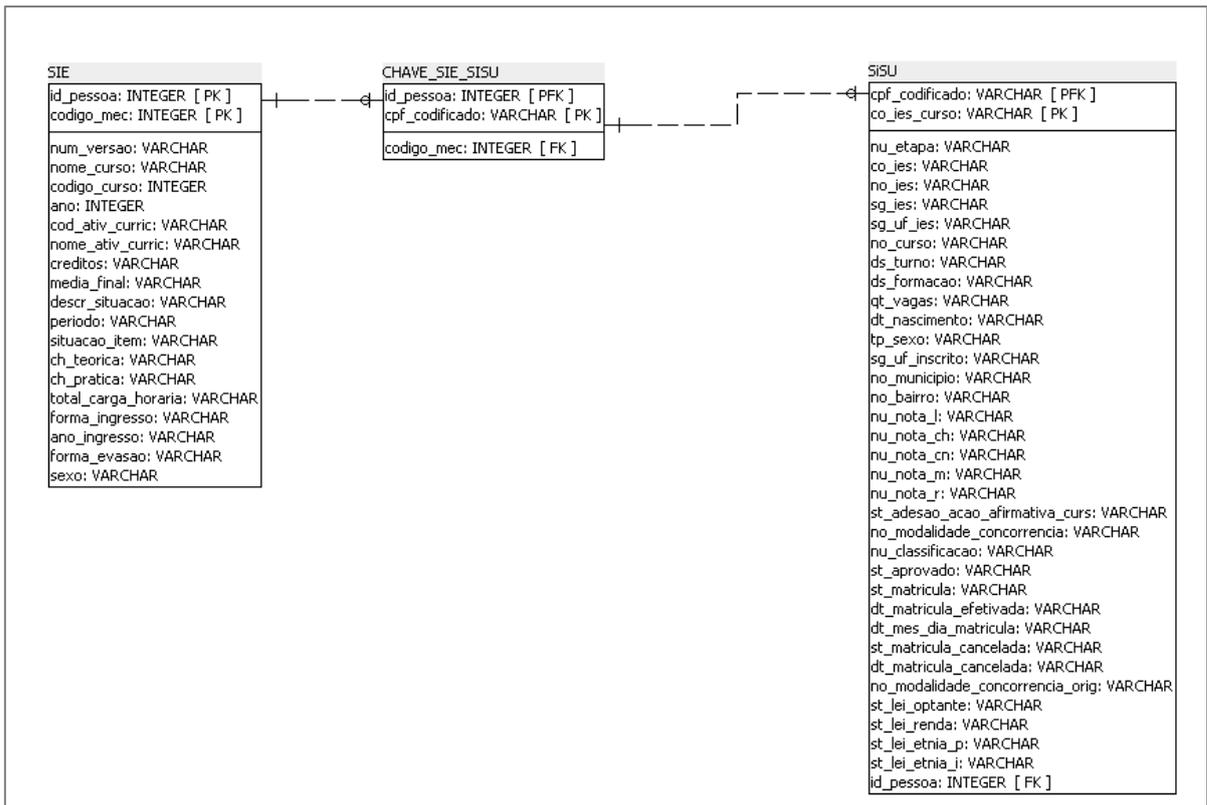


Figura 7 - Modelo de Dados

ID_PESSOA	NUM_VERSAO	NOME_CURSO	COD_CURSO	ANO	COD_ATIV_CURRIC	NOME_ATIV_CURRIC	CREDITOS	MEDIA_FINAL	DESCR_SITUACAO
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	TEP0001	DESENHO E EXPRESSÃO GRÁFICA I		3	0	REF - Reprovado por falta
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	TEP0045	FÍSICA I MECÂNICA NEWTONIANA		3	0	REF - Reprovado por falta
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	HFC0016	INTRODUÇÃO À ADMINISTRAÇÃO		4	0	REF - Reprovado por falta
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	TEP0006	INTRODUÇÃO À ENGENHARIA DE PRODUÇÃO		2	0	REF - Reprovado por nota/conceito
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	TIN0001	PROGRAMAÇÃO I		3	0	REF - Reprovado por falta
100686	2010/1	Engenharia de Produção - Bac - Hab. Prod. Cultura - Turno Noturno - código e MEC 1101776 220	2014	SCN0084	QUÍMICA GERAL		3	0	REF - Reprovado por falta
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2014	HFE0002	EDUCAÇÃO AMBIENTAL E CIDADANIA	4	8	APV - Aprovado
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2014	TME0071	GEOMETRIA EUCLIDIANA	6	0	REF - Reprovado por falta
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2014	TIN0001	INTRODUÇÃO À CIÊNCIA DA COMPUTAÇÃO	3	0	REF - Reprovado por falta
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2015	TME0071	GEOMETRIA EUCLIDIANA	6	0	REF - Reprovado por falta
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2015	TIN0001	INTRODUÇÃO À CIÊNCIA DA COMPUTAÇÃO	3	0	REF - Reprovado por falta
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2015	TME0019	MATEMÁTICA BÁSICA	6	0	ASC - Matrícula
106843	2014/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2016	TR10001	Terceramento Total	0	0	TR1 - Terceramento geral do Curso
56709	2010/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2010	SCA0001	BIOLOGIA GERAL I	4	0	REF - Reprovado por falta
56709	2010/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2010	TME0020	GEOMETRIA ANALÍTICA PLANA	4	0	REF - Reprovado por falta
56709	2010/1	Matemática - Licenciatura - Turno Noturno - código e MEC 1101770	230	2010	TIN0072	INTRODUÇÃO À CIÊNCIA DA COMPUTAÇÃO	2	0	REF - Reprovado por falta

CREDITOS	MEDIA_FINAL	DESCR_SITUACAO	PERIODO	SITUACAO_ITEM	CH_TEORICA	CH_PRACTICA	TOTAL_CARGA_HORARIA	FORMA_INGRESSO	ANO_INGRESSO	FORMA_EVASAO	ANO_EVASAO	SEXO
3	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
3	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
4	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
2	0	REF - Reprovado por nota/conceito	1º Semestre	2	0	0	30	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
3	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
3	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
4	8	APV - Aprovado	2º Semestre	1	0	0	60	SISU Escola Pública até 1.5 S.M. Preto, Pardo, Indígena	2014	JUB - Jubilamento	2015	F
6	0	REF - Reprovado por falta	2º Semestre	3	0	0	90	SISU Ampla Concorrência	2014	Sem evasão	0	M
3	0	REF - Reprovado por falta	2º Semestre	3	0	0	60	SISU Ampla Concorrência	2014	Sem evasão	0	M
6	0	REF - Reprovado por falta	1º Semestre	3	0	0	90	SISU Ampla Concorrência	2014	Sem evasão	0	M
3	0	REF - Reprovado por falta	1º Semestre	3	0	0	60	SISU Ampla Concorrência	2014	Sem evasão	0	M
6	0	ASC - Matrícula	1º Semestre	10	0	0	90	SISU Ampla Concorrência	2014	Sem evasão	0	M

Figura 8 - Exemplo de relatório do SIE, referente ao CCET

NU_ETAP	CO_IES	NO_IES	SG_IES	SG_UF	IE_CO	ES_CNO	CURS	D5	TURN	D5	FORM	DT_VAGA	CPF_CODI	DT_NASC	TP_SEXO	SG_UF	IF_NO	MUN_NO	BAIR	NU_NOTA1	NU_NOTA2	NU_NOTA3	NU_NOTA4	NU_NOTA5	ST_ADESI	NO_MOD	NU_CLAS	ST_APPRO	ST_MATR	DT_MATR	DT_MES	ST_MATR	DT_MATR	NO_MOD	ST_LEI_O	ST_LEI	RIS	LEI	LEI	ETNIA					
1	693	UNIVERS	UNIRIO	RJ	15782	ENFERM	Integral	Bacharel	30	1,5E+10	#####	M	RJ	RIO DE J	Barra da	618,9	698,5	597,2	737	760	NBo	Ampla Cc	25	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo						
1	693	UNIVERS	UNIRIO	RJ	1103343	CÊNCIAS	Noturno	Licenciati	4	2,7E+09	#####	F	RJ	RIO DE J	Vila Isai	630,7	662	500,8	618,1	640	Sim	Candidat	3	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo				
1	693	UNIVERS	UNIRIO	RJ	1103343	CÊNCIAS	Noturno	Licenciati	12	3,4E+09	#####	M	RJ	RIO DE J	Recreio	639,4	708	598,2	656,1	660	NBo	Ampla Cc	5	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo				
1	693	UNIVERS	UNIRIO	RJ	1103343	CÊNCIAS	Noturno	Licenciati	12	1,3E+10	#####	M	RJ	RIO DE J	Barra da	622,9	666,5	546,5	656,4	720	NBo	Ampla Cc	6	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo			
1	693	UNIVERS	UNIRIO	RJ	1106953	BIOLÓGIA	Noturno	Licenciati	3	1,5E+10	#####	F	RJ	RIO DE J	Olarina	643,7	631	532	583,7	720	Sim	Candidat	3	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo			
1	693	UNIVERS	UNIRIO	RJ	15788	DIREITO	Noturno	Bacharel	35	1,3E+10	#####	F	RJ	RIO DE J	Copacab	687,2	775,1	719,7	758,4	760	NBo	Ampla Cc	20	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	15788	DIREITO	Noturno	Bacharel	35	1,4E+10	#####	M	RJ	RIO DE J	Portugue	624	744,1	685,3	755,4	880	NBo	Ampla Cc	29	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	15788	DIREITO	Noturno	Bacharel	35	5,8E+09	#####	M	RJ	NITEROI	Fonseca	682,2	774,6	696,9	733,4	800	NBo	Ampla Cc	30	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	15	1,3E+10	#####	F	RJ	RIO DE J	Olarina	620,1	704	609,8	675	800	NBo	Ampla Cc	1	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	15	9,8E+10	#####	M	RJ	RIO DE J	Rocha MI	570,6	774,5	595,2	678,6	760	NBo	Ampla Cc	2	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	3	8,3E+09	#####	F	RJ	NOVA IG	Austin	607,9	640,6	563,6	593,9	680	Sim	Candidat	2	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	3	6,6E+10	#####	F	RJ	RIO DE J	Copacaba	623,9	668	510,2	700,4	660	Sim	Candidat	2	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	4	6E+10	#####	F	RJ	RIO DE J	Botafogo	600,1	668,2	493,6	653,1	740	Sim	Candidat	2	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	3	2,6E+09	#####	F	RJ	RIO DE J	Del Casti	573,4	668	560,5	603,7	740	Sim	Candidat	3	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	3	1,7E+10	#####	F	RJ	RIO DE J	Bonuco	653,5	659	544,1	561,1	680	Sim	Candidat	3	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	5	1,6E+10	#####	M	RJ	RIO DE J	Ricardo	603,6	709,2	623,9	550,7	580	Sim	Candidat	3	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	4	1,4E+10	#####	M	RJ	RIO DE J	Caju	558,1	601,1	533	639,4	720	Sim	Candidat	4	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	25	8,7E+09	#####	M	RJ	RIO DE J	Vila Isai	567,5	720,5	642,8	708,2	680	NBo	Ampla Cc	11	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101738	SERVIÇO	Noturno	Bacharel	25	1,7E+10	#####	F	RJ	RIO DE J	Sagua	670,5	734,8	585,3	722,3	600	NBo	Ampla Cc	12	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101776	ENGENH	Noturno	Bacharel	12	1,3E+10	#####	F	ES	VITORIA	Senta Lui	711,1	687,7	666,4	841,8	760	NBo	Ampla Cc	8	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo		
1	693	UNIVERS	UNIRIO	RJ	1101776	ENGENH	Noturno	Bacharel	12	1,3E+10	#####	M	RJ	RIO DE J	Recreio	639,1	631,5	706,3	776,5	960	NBo	Ampla Cc	9	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1101776	ENGENH	Noturno	Bacharel	12	1,4E+10	#####	M	RJ	RIO DE J	Campo G	676,7	761	684,9	779,1	780	NBo	Ampla Cc	10	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1103343	CÊNCIAS	Noturno	Licenciati	12	1,2E+10	#####	M	RJ	RIO DE J	Rocha	543,6	654,9	591,3	537,8	880	NBo	Ampla Cc	7	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	1103343	CÊNCIAS	Noturno	Licenciati	12	1,5E+10	#####	F	RJ	RIO DE J	Jardim B	603	674,2	597,7	678,9	640	NBo	Ampla Cc	9	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	
1	693	UNIVERS	UNIRIO	RJ	15788	DIREITO	Noturno	Bacharel	35	1,5E+10	#####	F	RJ	RIO DE J	Tijuca	657,4	716,7	653,3	765,5	900	NBo	Ampla Cc	25	Sim	#####	jun/16	NBo	Ampla Cc	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	NBo	

Figura 9 - Exemplo de Planilha extraída do SiSU utilizando o SiSU Gestão

ID_PESSOA	CPF_CODIFICADO
564	8217712751
7	9274563725
8	7775861794
10	7407964724
11	5356177736
15	5288557804
17	91939836724
18	8666203730
22	9178574730
23	9261592783
24	8492558745
26	5610398821
27	7050004751

Figura 10 - Exemplo de Planilha de chaves ID_PESSOA (SIE) e CPF_CODIFICADO (SiSU), gerada pela DTIC.

Detalhamento dos dados:

Planilha de dados do SIE:

ID_PESSOA	código do curso de graduação registrado na UNIRIO
NUM_VERSAO	ano e versão do curso
NOME_CURSO	nome do curso
CODIGO_MEC	código do curso de graduação registrado no MEC
COD_CURSO	código do curso de graduação registrado na UNIRIO
ANO	ano correspondente à atividade curricular
COD_ATIV_CURRIC	código da atividade curricular
NOME_ATIV_CURRIC	descrição da atividade curricular
CREDITOS	quantidade de créditos da atividade curricular
MEDIA_FINAL	média obtida na disciplina cursada
DESCR_SITUACAO	descrição da situação na disciplina cursada
PERIODO	período referente à atividade curricular
SITUACAO_ITEM	descrição da situação da atividade curricular
CH_TEORICA	carga horária teórica

CH_PRATICA	carga horária prática
TOTAL_CARGA_HORARIA	total da carga horária
FORMA_INGRESSO	forma de ingresso na UNIRIO
ANO_INGRESSO	ano de ingresso na UNIRIO
FORMA_EVASÃO	forma de evasão
ANO_EVASÃO	ano de evasão
SEXO	sexo do aluno

Tabela 1 - Campos dos relatórios do SIE

Planilha de Dados do SiSU:

NU_ETAPA	etapa de chamada classificatória do SiSU
CO_IES	código da Instituição de Ensino Superior
NO_IES	nome da Instituição de Ensino Superior
SG_IES	sigla da Instituição de Ensino Superior
SG_UF_IES	sigla do Estado da Instituição de Ensino Superior
CO_IES_CURSO	código do curso de graduação registrado no MEC
NO_CURSO	nome do curso de graduação
DS_TURNO	descrição do turno do curso de graduação
DS_FORMACAO	tipo de formação do curso de graduação
QT_VAGAS	quantidade de vagas ofertadas no curso de graduação
CPF_CODIFICADO	CPF criptografado do aluno
DT_NASCIMENTO	data de nascimento do aluno
TP_SEXO	sexo do aluno
SG_UF_INSCRITO	sigla do Estado de origem do aluno
NO_MUNICIPIO	nome do município de origem do aluno
NO_BAIRRO	nome do bairro de origem do aluno

Tabela 2 - Campos dos relatórios do SiSU

Planilha de identificadores (chaves):

ID_PESSOA	código de identificação do aluno
------------------	----------------------------------

Tabela 3 - Campos chave para ligartabelas do SiSU e SIE

4.2 Arquitetura da Solução

A solução não necessita de uma configuração complexa nem depende de instalação pelo desenvolvedor. Trata-se de um arquivo local que pode ser conectado a outros arquivos – fontes de dados – no mesmo ambiente.

**Figura 11 - Arquitetura da Solução**

Inicialmente, foi entregue ao usuário, o arquivo do Tableau com as fontes de dados configuradas, bem como os indicadores previamente especificados.

Para processar dados futuros, o usuário adicionará as planilhas, no mesmo formato das anteriores, nas conexões do Tableau e os dados serão atualizados e incluídos.

4.3 Definição de Métricas e Dimensões

Ao conectar as fontes de dados ao arquivo do Tableau, a ferramenta já atribui formatos aos campos. Essa atribuição pode ser alterada de acordo com a necessidade do usuário ou ao modelo especificado, como citado na Seção 3.4.

De acordo com a solução proposta, a configuração de métricas e dimensões foram configuradas no Tableau conforme figuras 12 e 13.

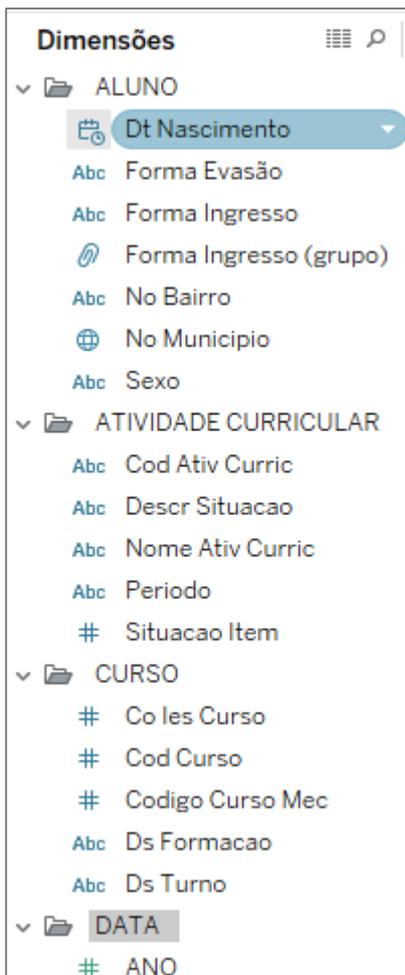


Figura 12 - Configuração de Dimensões no Tableau

Os campos de dados que foram definidos como dimensões foram agrupados por assunto: todas as datas referentes a uma atividade curricular ou do curso foram colocadas na pasta DATA; dados referentes ao aluno, como data de nascimento, forma de ingresso na faculdade, forma de evasão, município de origem e sexo foram agrupados na pasta ALUNO; dados de descrição do curso, bem como seu código, turno e tipo de formação foram alocados em CURSO; dados referentes às atividades curriculares, como a descrição da atividade, nome, situação e período de referência foram agrupados em ATIVIDADE CURRICULAR, como mostra a figura 12.

As métricas, que correspondem aos campos numéricos e calculados, conforme figura 13, também foram agrupadas por assunto.

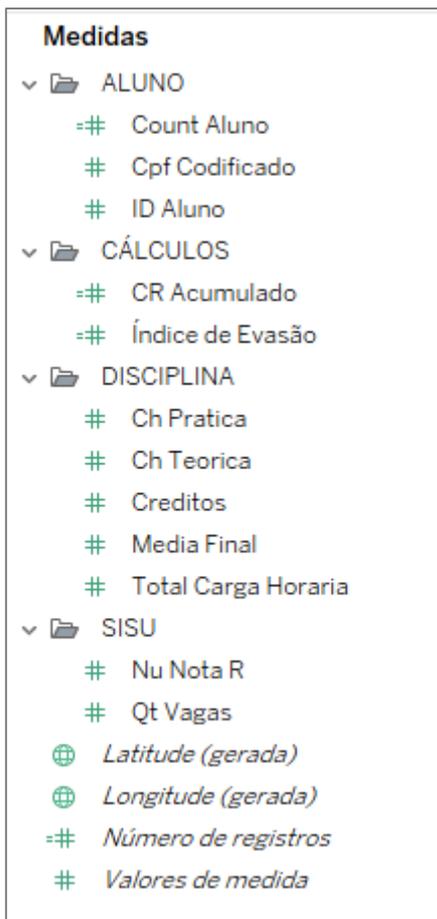


Figura 13 - Configuração de Metricas no Tableau

Os campos calculados que foram criados são:

- Contagem distinta de alunos: $COUNTD([ID\ Aluno])$
- Índice de Evasão: $COUNT([Forma\ Evasão])/COUNTD([ID\ Aluno])$
- CR Acumulado: $(SUM([Media\ Final])*[Creditos])/SUM([Creditos])$

Os indicadores são gerados a partir do cálculo simples, o que os diferencia são os filtros das dimensões utilizadas:

- Ano de Evasão
- Curso de graduação
- Forma de Ingresso
- Município de origem
- Forma de evasão

Na solução apresentada neste trabalho, utilizamos apenas esses filtros, porém os usuários terão disponíveis todos os campos enviados nas planilhas.

4.4 Análise de Resultados

O primeiro indicador gerado é o mais generalizado. Ele mostra a porcentagem de evasão de alunos de Ampla Concorrência em cada curso por semestre, desde 2010.

Para obter esse resultado, o cálculo do Índice de Evasão foi relacionado às seguintes dimensões: Forma de Ingresso, Forma de Evasão, Curso de Graduação e o Ano de Evasão como filtros.

Para gerar esse indicador, devemos:

- selecionar o cálculo Índice de Evasão, do grupo de métricas, e arrastar até a área de *Marcas*;
- selecionar o Ano de Evasão e arrastar até a área de colunas e área de filtro;
- arrastar o campo Forma Evasão até a área de filtro e editar o filtro, selecionando as opções: ABA – abandono de curso, CAN – cancelamento geral de curso e JUB – jubramento;
- arrastar o campo Forma de Ingresso até a área de filtro e selecionar as opções referentes a ENEM, Vestibular e SiSU.
- Arrastar o campo Nome Curso para a área filtro e Linhas do gráfico.

Realizados esses passos, o gráfico apresentado na figura 14 será exibido.

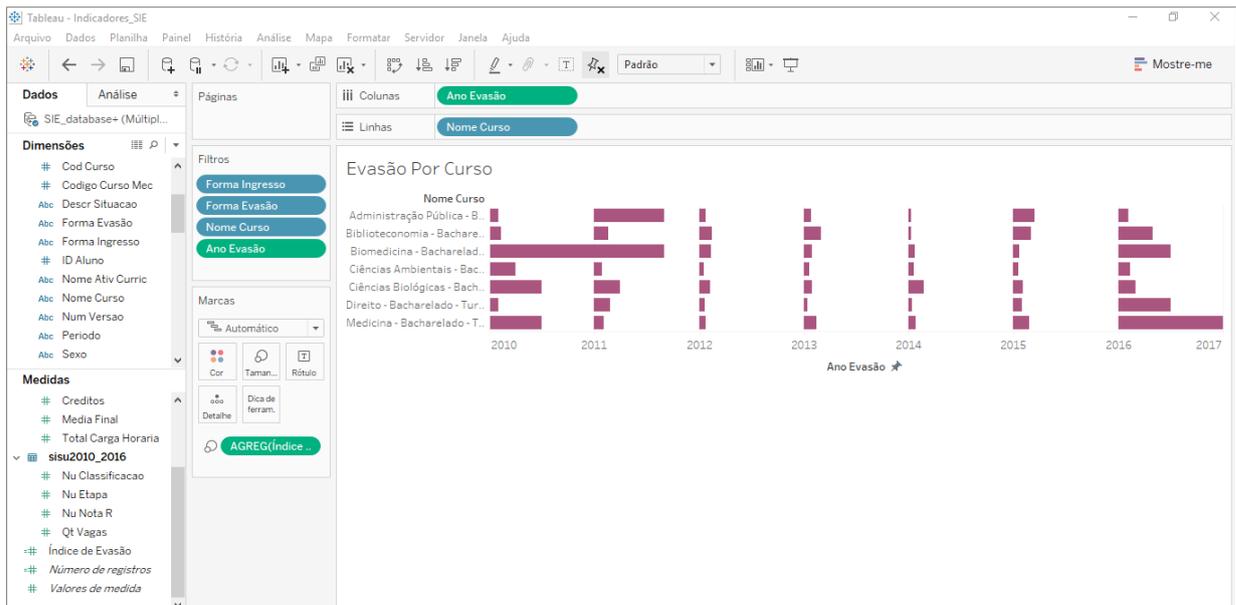


Figura 14 - Indicador de Evasão por Curso

As dimensões foram dispostas nos filtros para dinamizar e flexibilizar os resultados. Dessa maneira, pode-se selecionar um curso específico ou uma forma de evasão específica, por exemplo. Além disso, o tipo de gráfico pode ser alterado dinamicamente sem nenhum impacto nas métricas e dimensões.

A partir do mesmo indicador exibido na figura 15, podemos comparar resultados específicos. Na figura 16, comparamos o índice de evasão entre os cursos do CCET (Centro de Ciência e Tecnologia). Em relação ao gráfico da figura 14, mudando apenas o filtro Curso.

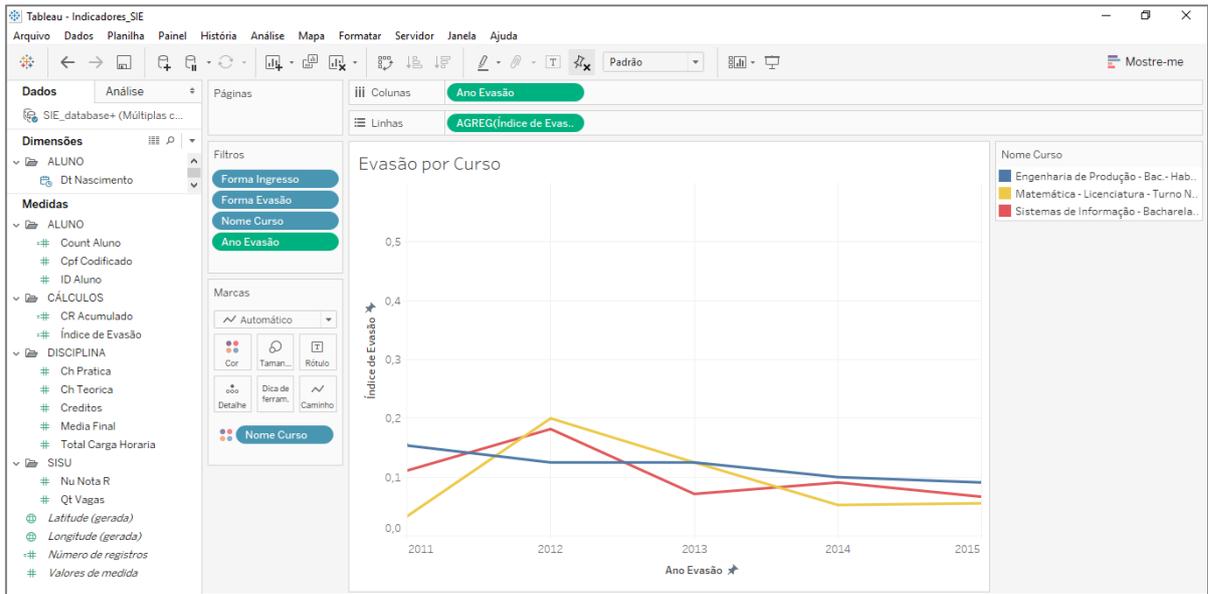


Figura 15 - Índice de Evasão filtrado por cursos específicos

A figura 16 exibe o índice de evasão no período de 2010 a 2016 do curso de graduação Bacharelado em Sistemas de Informação, considerando o SiSU como forma de ingresso. Podemos perceber um pico de evasão em 2012.

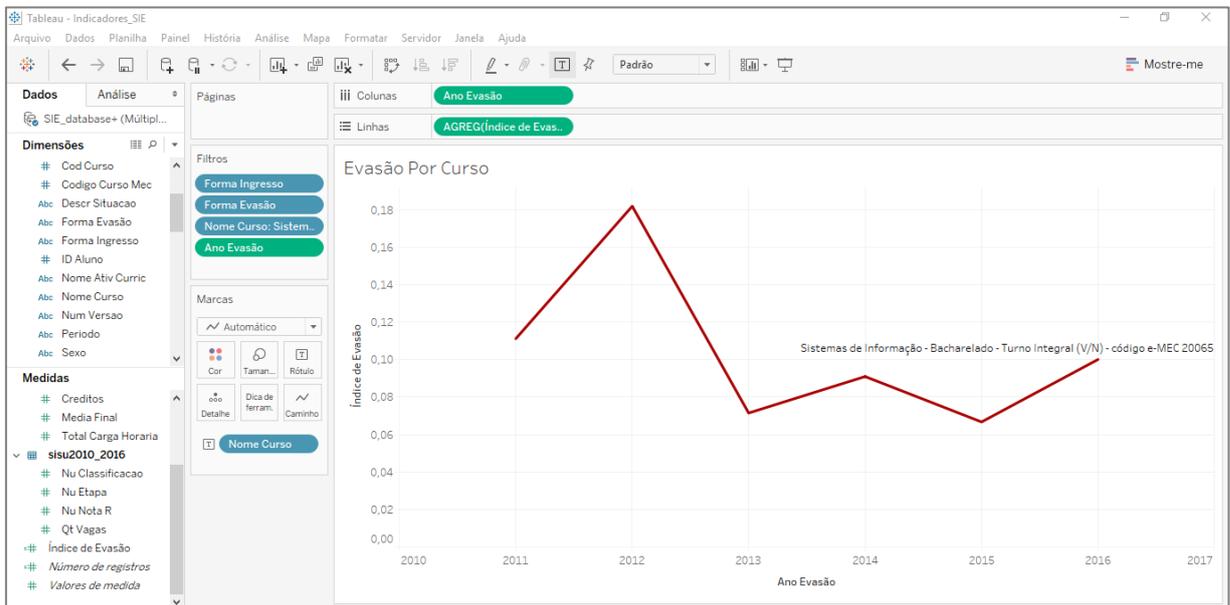


Figura 16 - Índice de Evasão de um curso específico

O segundo indicador gerado, apresentado na figura 17, considera os ingressos a partir de 2013, quando foram implementadas as cotas. Sendo assim, utiliza como filtro as formas de ingresso Ampla Concorrência e Cotas, e formas de evasão Cancelamento,

Jubilamento e Abandono de Curso. Com esse indicador é possível comparar o índice de evasão em cada semestre de acordo com a forma de ingresso, e analisar a cada período se o índice aumentou, diminuiu ou está estável.

Esse indicador possui dois modelos, um com os tipos de cota segregados, e outro com os mesmos agrupados.

Os mesmos filtros do primeiro indicador foram utilizados, porém a Forma de Ingresso foi colocada como “Linha” no gráfico do Tableau.

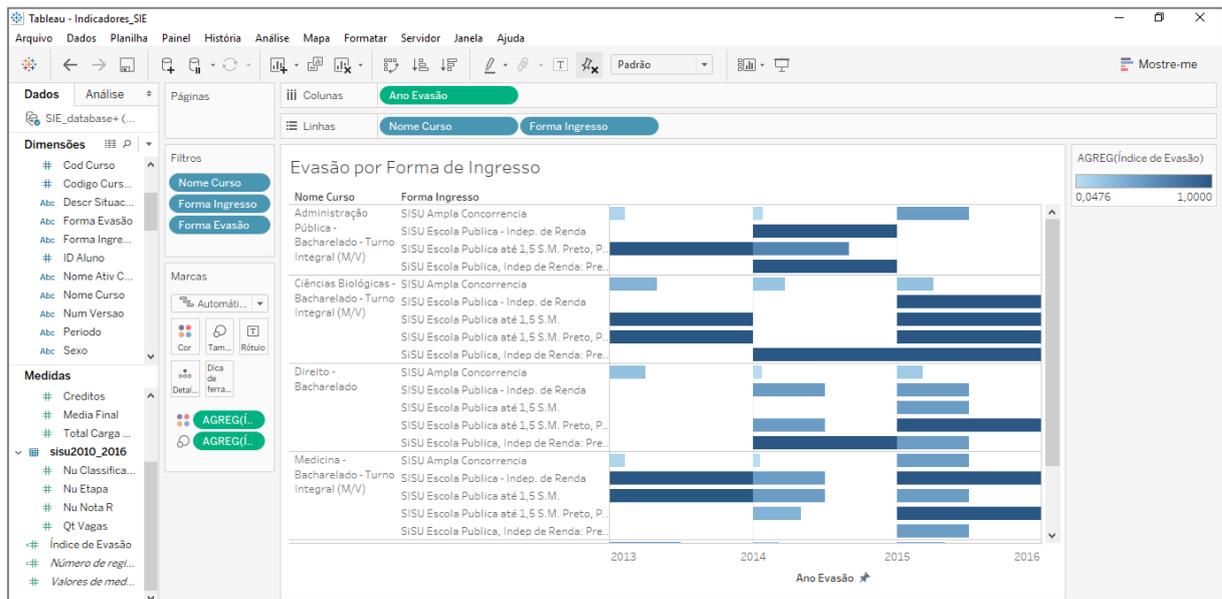


Figura 17 - Indicador de Evasão por Formas de Ingresso

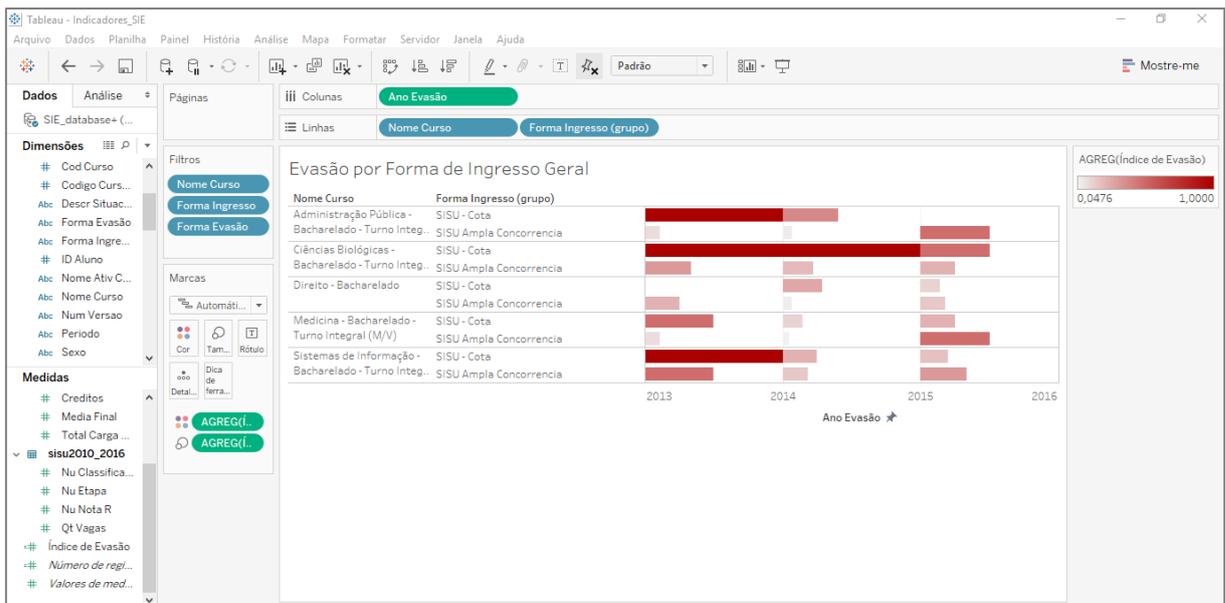


Figura 18 - Formas de Ingresso agrupadas

O terceiro indicador mostra o perfil geográfico dos alunos que evadem. A localização utilizada para essa informação é oriunda da base de dados do SiSU, pois se refere ao endereço de origem do aluno, e não seu endereço após migração para o Rio de Janeiro (em casos de migração). Sendo assim, é possível analisar possíveis fatores sociais e geográficos que levaram o aluno a evadir.

O Tableau disponibiliza uma função geográfica, onde é atribuído a um campo, por exemplo, cidade, sua localização geográfica. Esse tipo de visualização é gerado somente via conexão a Internet. Para gerar um campo com localização geográfica, basta clicar sobre o mesmo e selecionar a propriedade de função geográfica, definindo se é Estado, País, Cidade ou outros.

O mapa foi gerado da seguinte maneira:

- Definiu-se o campo *No Município* como função geográfica de cidade e arrastou o campo para a área de gráfico do Tableau. Automaticamente o mapa foi gerado.
- Os filtros foram selecionados e arrastados para a área de Filtros, no caso, Ano de Evasão, Forma de Evasão e Forma de Ingresso.

Dessa maneira, é possível identificar o índice de evasão por localidade, como é apresentado na figura 19:

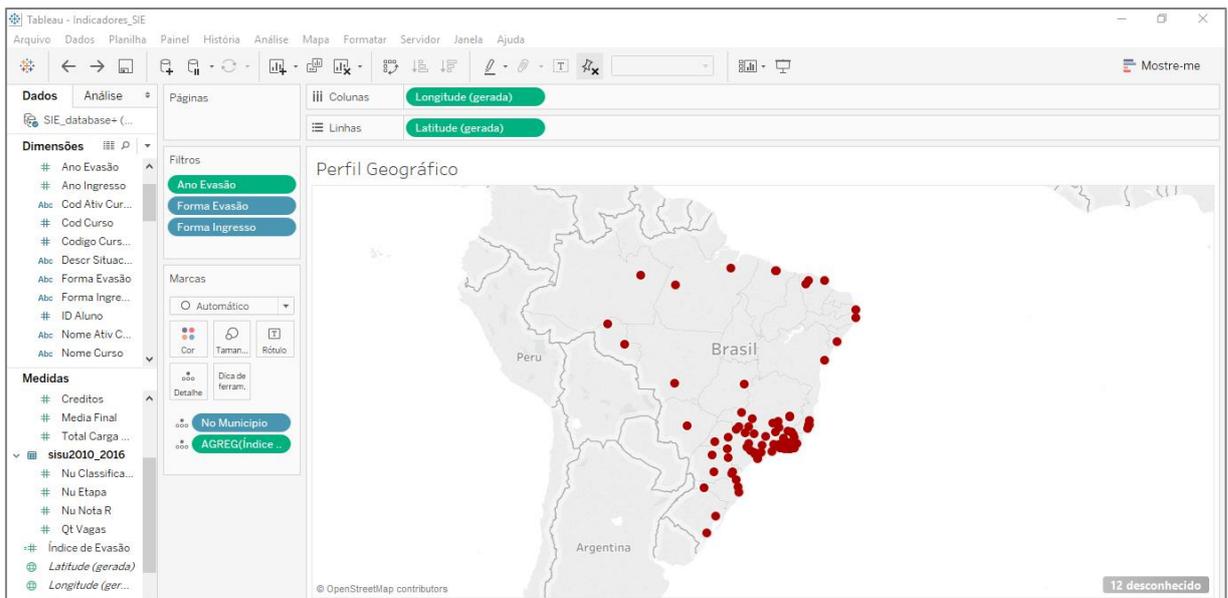


Figura 19 - Indicador de evasão por localidade, com visualização em mapa

O mapa é interativo e dinâmico. Uma região pode ser destacada e ao passar o cursor do mouse pelo mapa, os dados são exibidos conforme figura 20:

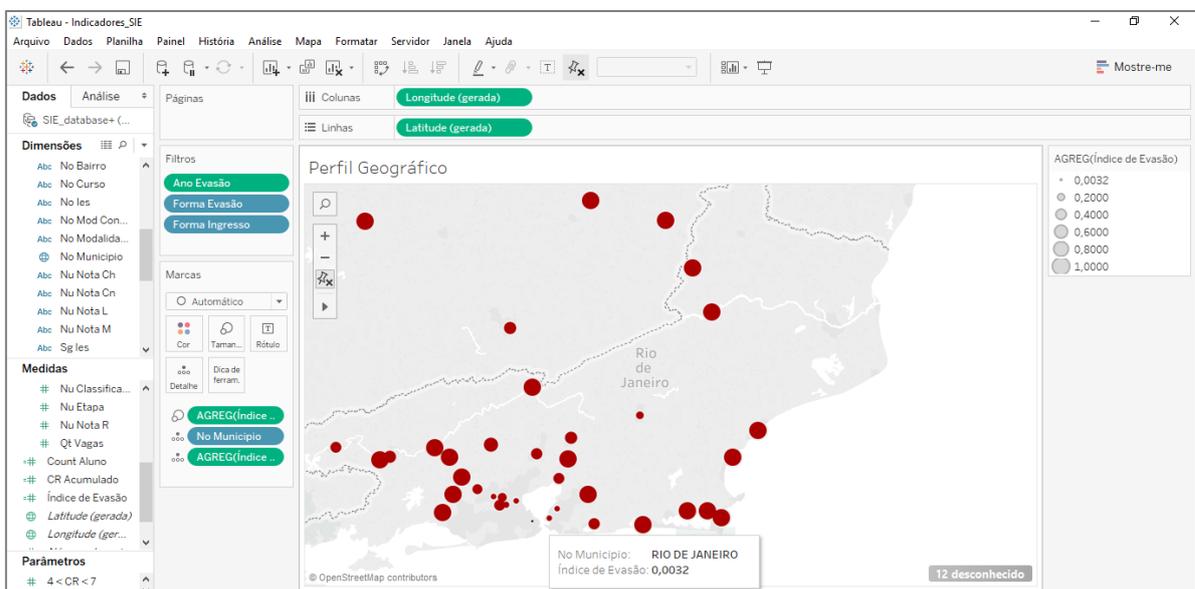


Figura 20 - Informações de uma região específica escolhida

Esse tipo de visualização geográfica também pode ser manipulado por diversos filtros.

O quarto indicador mostra a média do coeficiente de rendimento dos alunos que evadem. Para isso a média do Coeficiente de Rendimento Acumulado foi calculada e comparada com o índice de evasão, como mostra a figura 21:

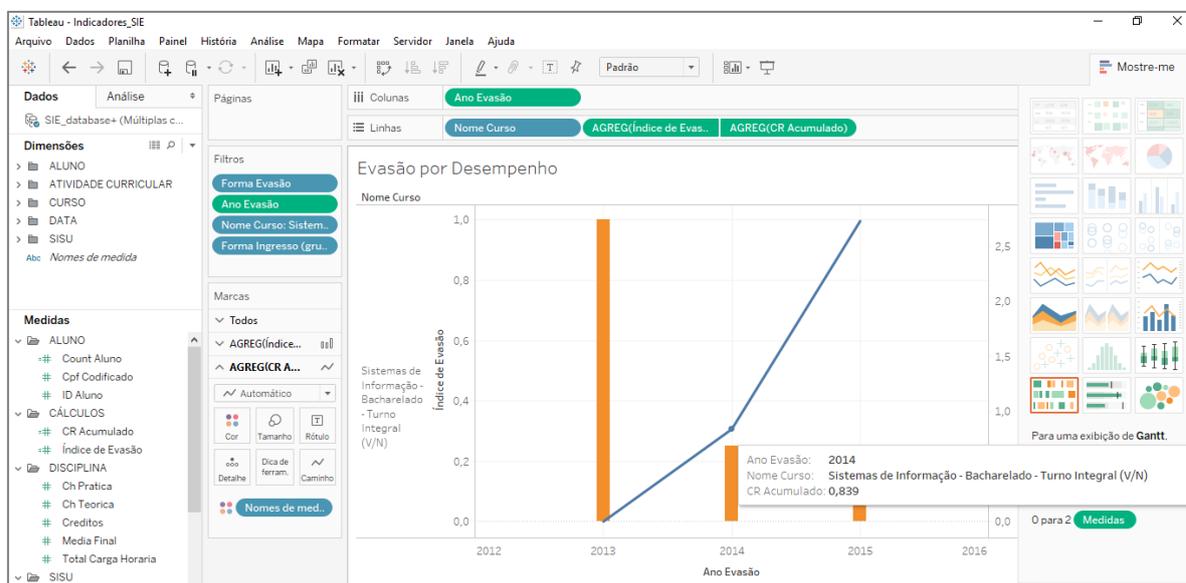


Figura 21 – Média do Coeficiente de rendimento de alunos que evadem

Nesse exemplo, filtramos o resultado pelo curso de Sistemas de Informação e a forma de ingresso por cotistas. O gráfico revela que de 2013 a 2015 quanto maior a média de CR Acumulado menor é o índice de evasão.

Na figura 22, comparamos esse mesmo índice filtrando cursos diferentes – Medicina e Ciências Biológicas Bacharelado (turno integral) - e comparando os resultados entre eles. Pode-se perceber que o comportamento do CR Acumulado em relação à evasão é bem diferente entre os dois cursos.

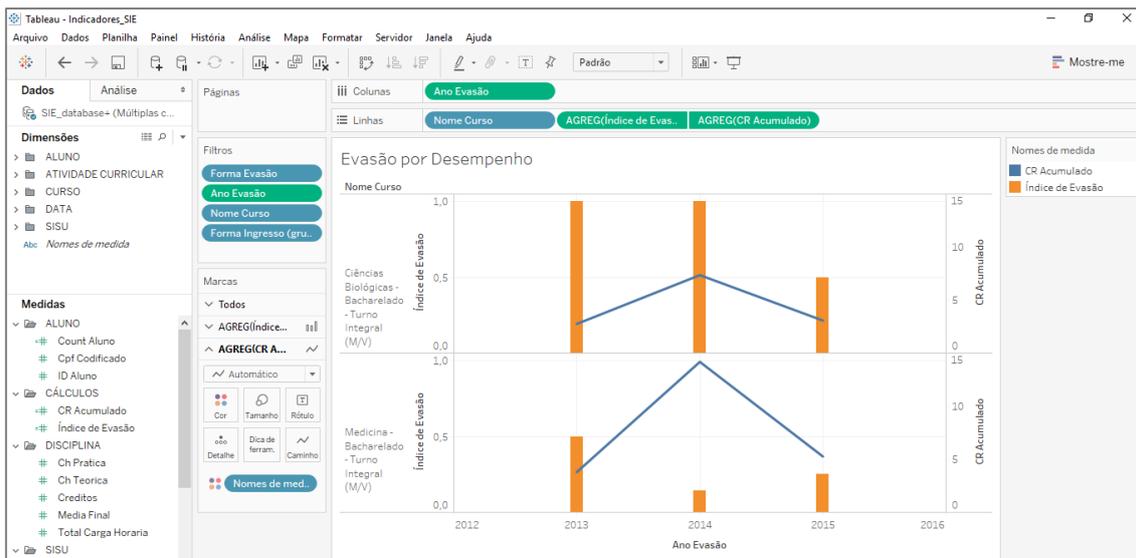


Figura 22 - Comparação de evasão por CR entre Medicina e Ciências Biológicas

Outra forma significativa de analisar esses resultados, é comparar esse mesmo indicador entre cursos de turnos diferentes, conforme figura 23, por exemplo, turno integral X turno noturno. Para isso, comparamos o curso de Pedagogia Vespertino X Pedagogia Noturno e constatamos um comportamento parecido. Nesse resultado, utilizamos as formas de ingresso do SiSU de cotistas e não cotistas:

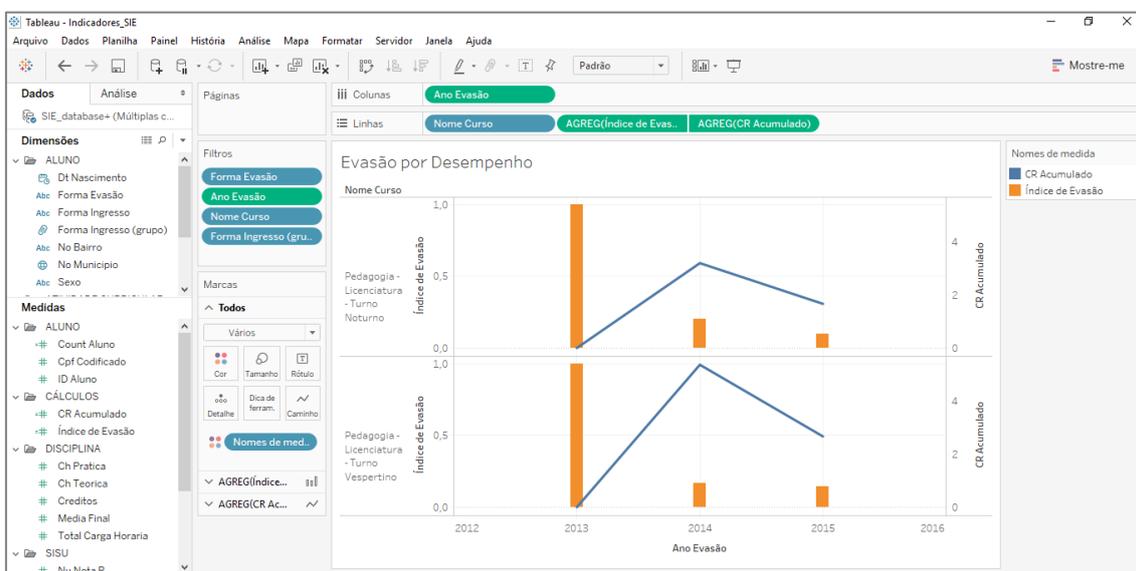


Figura 23 - Comparação de Evasão por CR entre turnos diferentes de um mesmo curso

Outra visualização que pode ser gerada, conforme figura 24, é a relação da evasão com

o Coeficiente de Rendimento e o município de origem do aluno. Para exibir esse indicador, utilizou-se o gráfico do indicador geográfico, aproveitando os mesmos filtros. Para a relação com o Coeficiente de Rendimento, o campo do cálculo de CR foi colocado na função COR da área Marcas. Dessa forma, os CRs mais baixos são exibidos com uma coloração mais clara, enquanto que os mais altos, uma coloração mais escura.

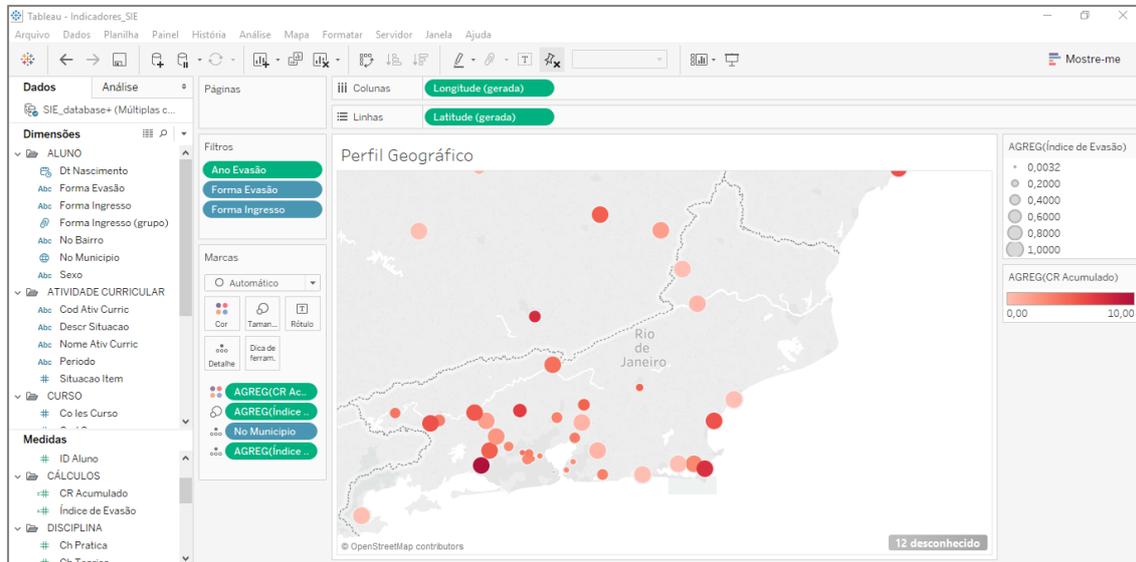


Figura 24 - Relação da evasão com o Coeficiente de Rendimento e o município de origem do aluno

5 Conclusão

5.1 Conclusões sobre os Indicadores obtidos

Embora o objetivo desse trabalho não seja encontrar respostas para questões que a faculdade busca entender, é possível observar resultados relevantes para a universidade. O índice de evasão na área de exatas é muito maior do que nos demais. Alguns motivos citados por alunos podem explicar isso, como por exemplo: dificuldade com matemática, falta de afinidade com métodos de ensino adotados. Outro motivo observado que gera a mesma consequência, é a nota de corte do SiSu para ingressar na UniRio. A nota obtida nos exames de ciências exatas não possui um peso, ou mínimo necessário, para o aluno ingressar nos cursos dessa área. Muitas vezes os alunos escolhem o curso pois possuem a nota de corte compatível com a do SiSu, porém não tiveram bom desempenho nas disciplinas de exatas. Esse cenário costuma resultar em evasão.

Quando há evasão por localidade, entende-se que o aluno pode ter voltado para a sua cidade de origem por diversos motivos, mas algum deles poderiam ser evitados com a implementação de programas de auxílio moradia e transporte, ou até mesmo a construção de dormitórios na UniRio. Alunos que vem de outras cidades tendem a preferir universidades que possuem essas facilidades.

5.2 Considerações Finais

Os usuários da solução precisam da informação gerada pelos indicadores desenvolvidos para mostrar fatos às coordenações dos cursos da UNIRIO. A partir desses fatos, é dever das coordenações discutir com seus colegiados o que pode estar ocasionando-os. Nesse trabalho foi dada prioridade ao fato “Evasão”, mas ao entender a ferramenta, os usuários conseguem analisar muitos outros, como mobilidade entre cursos, transferências externas, entre outros.

Vale ressaltar que para ocorrer uma interpretação real dos dados, o cadastro dos alunos no SIE precisa ser o mais correto possível. Hoje, existem informações incorretas no sistema. A partir do momento em que as informações passarão por análise, é preciso treinar os responsáveis para cadastrarem corretamente as informações do aluno, e esclarecer o significado de situações que até então não recebiam muita importância, como por exemplo: confusão entre formas de evasão como “Jubilamento”, “Cancelamento” e “Desligamento”.

5.3 Limitações do Projeto

Sobre o problema dos cadastrados incorretamente no SIE mencionado na sessão anterior, um forte exemplo é o caso da forma de evasão, onde há confusão dos funcionários entre Cancelamento, Jubilamento, Abandono e Desligamento. As três primeiras, são discutidas nesse trabalho, já o Desligamento ocorre quando um aluno consegue ingressar na faculdade por questões judiciais, mas a liminar é caçada e depois ele precisa se desligar. Como ele ocupou uma vaga extra, esse desligamento não gera evasão, portanto esse tipo de informação não foi analisada, mas obtivemos informações da PROGRAD que alguns alunos desligados estão cadastrados com cancelamento de vaga, ou seja, estão fazendo parte do nosso índice de evasão incorretamente. Portanto, é necessário alertar os funcionários que realizam cadastro de alunos no sistema sobre o que significam essas informações, a fim de manter um cenário de análise o mais próximo possível da realidade.

O fato das cotas existirem apenas a partir de 2013, torna a análise de evasão de cotistas ainda pouco conclusiva. Mas como o objetivo era criar uma solução e não apenas verificar dados do passado, uma vez que a mesma continuará sendo utilizada, esse estudo continuará sendo feito semestralmente.

Outra limitação é o fato de não ser permitido o acesso ao Banco de Dados. Solicitar relatórios à DTIC é demorado, e se não estiverem no modelo estabelecido, há um trabalho extra de modificação de planilhas, além disso, extração de relatórios da base de dados do SIE também possui limitações, pois não é possível gerar relatórios com mais de 65536 linhas. Isso faz necessário dividir a extração, e depois unir essas planilhas. Todos esses problemas poderiam ser resolvidos com uma integração direta do Tableau com o banco de dados do SIE.

Algumas inconsistências de dados foram encontradas no SIE, como a média geral do aluno: foram lançadas no sistema notas com a casa decimal e sem casa decimal, por exemplo 8.3 e 83. Isso gerou um grande impacto nos resultados da média de CR Acumulado.

5.4 Trabalhos Futuros

No trabalho atual, existe um retrabalho para gerar os relatórios e importá-los para o Tableau, uma vez que os usuários da solução não possuem acesso ao Banco de Dados do SIE, então precisam solicitar a eles os relatórios, informando quais dados precisam. Dessa forma, a atual fonte de dados é baseada em planilhas. Seria interessante em trabalhos futuros realizar a conexão do Tableau com o Banco de Dados do SIE, de forma que fossem criadas *views* apenas com as informações necessárias, já que há restrições de segurança e o acesso das tabelas não é liberado a todos. Assim, a fonte de dados seria o próprio banco de dados.

Outra possibilidade é o desenvolvimento de outros indicadores importantes, como o número de vagas ociosas em um curso, ou a relação de cursos doadores e cursos receptores, que são cursos que perdem seus alunos para outros cursos que o aluno não possuía nota suficiente para passar nas primeiras tentativas, e o próprio usuário poderia realizar este trabalho. Uma vez criada a solução por nós, ele mesmo consegue criar novos indicadores, pois daremos um treinamento básico sobre a ferramenta, e o conhecimento dos cálculos e métricas eles já possuem, uma vez que construímos os indicadores com a validação de cada passo com eles, e o que faltava para eles era apenas o conhecimento sobre o Tableau em si.

Não é considerada a hipótese de substituir o Tableau em trabalhos futuros, uma vez que a UNIRIO não possui licença de outras ferramentas similares, e com as fontes de dados atuais, o Tableau atende perfeitamente a demanda.

Essas possibilidades contribuiriam para o enriquecimento da informação na UNIRIO, e a tomada de decisões de impacto positivo em outras questões fora a evasão. Se isso tudo puder acontecer com uma integração direta do Tableau com o Banco de Dados do SIE, haverá um ganho significativo de tempo. Seria de suma importância um alinhamento com os responsáveis pelo SIE para que o responsável pela solução de indicadores possa ter acesso direto a base de dados do sistema.

Referências Bibliográficas

Portal Brasil (2014) “Percentual de negros no ensino superior é metade de brancos”
<http://www.brasil.gov.br/cidadania-e-justica/2011/05/percentual-de-negros-no-ensino-superior-e-metade-do-de-brancos>, Maio.

Strom, David (2016) “The Best Self-Service Business Intelligence (BI) Tools of 2016”
<http://www.pcmag.com/article2/0,2817,2491954,00.asp>, Novembro.

“Análise e business intelligence | Tableau Software”
<http://www.tableau.com/pt-br>

Portal Mec (2015) “Ministério anuncia novo sistema para preencher vagas ociosas”
<http://portal.mec.gov.br/ultimas-noticias/212-educacao-superior-1690610854/32581-ministerio-anuncia-novo-sistema-para-preencher-vagas-ociosas>, Dezembro.

Rodrigues, Mateus – G1 DF (2016) “MEC preenche apenas 2,5% de vagas ociosas no ensino superior public”
<http://g1.globo.com/educacao/noticia/2016/05/mec-preenche-apenas-25-de-vagas-ociosas-no-ensino-superior-publico.html>, Maio.

Datastorm “Conceitos e Ferramentas de BI”
<http://datastorm.com.br/conceitos-ferramentas-de-bi/>.

BHS (2016) “Tableau Desktop: use o Business Intelligence a favor do seu negócio”
<http://bhs.com.br/negocio/tableau-desktop-use-o-business-intelligence-favor-do-seu-negocio>, Junho.

BUTLER (2016) “Tableau e QlikSense: características, critérios de avaliação e análises”
<http://www.tiespecialistas.com.br/2016/08/tableau-e-qlik-sense-carateristicas-criterios-de-avaliacao-e-analises/>, Agosto.

“SQL Join: Entenda como funciona o retorno dos dados”
<http://www.devmedia.com.br/sql-join-entenda-como-funciona-o-retorno-dos-dados/31006>

“Power BI | Ferramentas de BI de Visualização de Dados Interativa”
<https://powerbi.microsoft.com/pt-br/>

Elmasri, R. e Navathe, S.B. (2005), “Sistemas de Bancos de Dados”, 4º ed. Addison Wesley Longman, Inc

Machado, F.N.R. (2004), “Tecnologia e Projeto de Data Warehouse”, 1º ed. São Paulo

BARBIERI, Carlos. **BI – BUSINESS INTELLIGENCE – Modelagem & Tecnologia**, Ed. Axcel Books – 2001

6 Apêndice

6.1 Introdução

Nesse apêndice constam os detalhes de todas as informações coletadas em reuniões com o Armando e Ângelo, diretores da CAEG e DIPRAG respectivamente. Todos os indicadores foram calculados com base na necessidade que eles apresentaram, e o cálculo utilizado foi definido de acordo com as explicações que foram dadas sobre o ingresso na UNIRIO.

6.2 Reunião para entendimento do processo de ingresso na UNIRIO

Nessa reunião eles relataram que para ingressar na UNIRIO pelo SiSU, 50% das vagas são ampla concorrência, e 50% cotas. As cotas são para alunos que ao menos cursaram todo o Ensino Médio em escola pública, e são divididas da seguinte forma:

- 1) Estudantes de Escola Pública com Renda Familiar de até 1,5 salário mínimo;
- 2) Estudantes de Escola Pública independente de Renda Familiar;
- 3) Estudantes de Escola Pública, preto/pardo/indígena, com Renda Familiar de até 1,5 salário mínimo;
- 4) Estudantes de Escola Pública, preto/pardo/indígena, independente de Renda Familiar.

O sistema de cotas foi implementado em 2013 na UNIRIO. Então entre 2010 e 2013, todos os alunos que entravam pelo SiSU estavam concorrendo a vagas de ampla concorrência.

Foi explicado que apesar das cotas possuírem o objetivo de eliminar dívidas históricas e mudar o lugar social dessas etnias nas universidades, muitos burlam o sistema, uma vez que a etnia é comprovada por auto declaração, sem qualquer tipo de verificação da mesma. Além disso, há alunos que burlam também a comprovação da renda familiar, omitindo a renda de alguns parentes, ou mostrando documentos inválidos. Porém, a

comitiva avaliadora tem um tempo para pesquisar mais informações sobre a pessoa e sua família, e muitas vezes são surpreendidos com a real situação financeira nas redes sociais ou Google.

A cada chamada do SiSU (por ordem de desempenho na prova do ENEM), os alunos precisam ir até a universidade fazer a inscrição, e os cotistas passam pela parte burocrática no processo. Se ao final das cinco chamadas sobram vagas, a universidade tem um tempo para realizar novas chamadas onde irão os alunos interessados em ingressar até o auditório, onde podem ter a oportunidade de ocupar vagas de ampla concorrência ou cota.

É possível ingressar na UNIRIO por transferência externa também, um programa que a universidade possui para ocupar vagas ociosas devido ao abandono de curso, e alunos de outras universidades fazem uma prova e apresentam o histórico (só é possível a transferência entre cursos de currículos compatíveis), e se obtiverem o resultado necessário, transferem seu curso para a UNIRIO. Para evitar a evasão também são abertos programas de permanência para alunos que precisam de um auxílio com transporte e moradia.

Os cursos de uma universidade são avaliados pelo MEC segundo um índice de sucesso, onde quanto menos vagas ociosas, maior o índice de sucesso. Quando este índice está baixo, a universidade não pode mais receber matrículas naquele curso até que a situação se regularize, e caso isso não aconteça, o curso pode ser fechado.

6.3 Reunião para definição de requisitos da solução

Nessa reunião, foram definidos os tópicos mais importantes para análise. Entender o que leva um aluno a evadir é o ponto principal a ser estudado, então foi solicitada a análise do perfil de aluno que evade, considerando: endereço, coeficiente de rendimento e forma de ingresso. O endereço considerado precisa ser da base de dados do SiSU, onde as pessoas cadastram seus endereços de origem, pois no SIE muitas vezes já está cadastrado o endereço que a pessoa se alocou no Rio de Janeiro. O fato da pessoa ter mudado de cidade ou estado já pode favorecer o abandono.

O coeficiente de rendimento é analisado pelos responsáveis conforme três intervalos: entre 0 e 4,0, que pode levar o aluno a ser jubilado, entre 4,0 e 7,0 que é um desempenho razoável em que os alunos estão conseguindo ter aprovações, e acima de

7,0 que é um bom desempenho. Porém, o cadastro das notas no SIE não segue um padrão a princípio, o que pode atrapalhar essa definição na solução.

A forma de ingresso será levada em conta a partir de 2013, ano em que o sistema de cotas foi implementado. É preciso entender o desempenho dos alunos cotistas para analisar a eficiência do sistema, apesar de que o mesmo não garante que o aluno de fato se enquadre nas características, já que a auto declaração da cota étnica por lei, não pode ser contestada (lei nº 12.711/2012).

As formas de evasão consideradas devem ser: Abandono de Curso, que é quando o aluno simplesmente sai do curso e não refaz sua matrícula nem oficializa seu cancelamento junto a secretaria, Jubilamento, que é quando a faculdade não permite mais que o aluno renove matrícula pois está com coeficiente de rendimento abaixo de 4,0 e 4 reprovações em uma mesma disciplina, e Cancelamento, que é quando o aluno solicita o cancelamento da sua matrícula. Essas formas de evasão são as que deixam vagas ociosas, por isso devem ser analisadas. Formas de evasão como “Desligamento” por exemplo, não devem ser analisadas. Esta se trata de quando um aluno ingressa por meios judiciais mas a liminar cai e ele precisa ser desligado. Como ele não ocupou uma vaga ociosa e sim uma “vaga extra”, esse dado não deve entrar no índice de evasão.

A análise será feita por curso e por centro.

6.4 Reunião para dúvidas

Após o recebimento dos dados, surgiram algumas dúvidas após algumas verificações.

Os dados analisados são a partir do ano de 2010, ano em que o SiSU se tornou a forma oficial de ingresso na UNIRIO. Reparamos que entre 2010 e 2013, a forma de ingresso está “ENEM – Vestibular”. De fato, só em 2013, junto com a implementação das cotas, o SIE ganhou as formas de ingresso “SiSU - ampla concorrência” e “SiSU – cotas”. Sendo assim, no período mencionado, a forma de ingresso “ENEM – Vestibular” será substituída por “SiSU - ampla concorrência”.

Cursos com forma de ingresso “THE – Teste de Habilidade Específica” não serão analisados pois foge do objetivo de entendimento de evasão a partir de alunos que utilizam apenas o SiSU. Entende-se que o THE por dificultar o processo de ingresso e diminuir a porcentagem de alunos interessados em apenas ingressar em um curso qualquer, dificulta a evasão nos cursos que o utilizam.

Quando as cotas foram implementadas, pretos e pardos constavam em um tipo de cotas, e indígenas em outra. Depois, foi unificado. Para este trabalho não será relevante a segregação, então foi feito um agrupamento nas cotas mencionadas para “Cota Étnica”.