



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO
TOCANTINS
CAMPUS DIANÓPOLIS
CURSO SUPERIOR LICENCIATURA EM COMPUTAÇÃO**

ISAAC PIMENTEL FERNANDES SOBRINHO

**O ESTUDO DA EVASÃO NO CURSO DE BACHARELADO EM ENGENHARIA
AGRONÔMICA NO IFTO *CAMPUS* DIANÓPOLIS, APLICANDO TÉCNICAS DE
MINERAÇÃO DE DADOS**

DIANÓPOLIS
AGOSTO / 2019

ISAAC PIMENTEL FERNANDES SOBRINHO

**O ESTUDO DA EVASÃO NO CURSO DE BACHARELADO EM ENGENHARIA
AGRONÔMICA NO IFTO *CAMPUS* DIANÓPOLIS, APLICANDO TÉCNICAS DE
MINERAÇÃO DE DADOS**

Monografia apresentada como requisito final para obtenção do grau de Licenciado em Computação do Instituto Federal de Educação, Ciência e Tecnologia do Tocantins – IFTO, *Campus* Dianópolis.

Orientador (a): Prof. Msc. Diego Rodrigues.

DIANÓPOLIS
AGOSTO / 2019

F363e Fernandes Sobrinho, Isaac Pimentel

O Estudo da evasão no curso de bacharelado em engenharia agrônômica no IFTO campus Dianópolis, aplicando técnicas de mineração de dados / Isaac Pimentel Fernandes Sobrinho – 2019.

72 f.

Orientador: Diego Rodrigues

Trabalho de Conclusão de Curso: (Licenciatura em Computação) – Instituto Federal de Educação Ciência e Tecnologia do Tocantins, campus Dianópolis. 2019.

1. Mineração de Dados . 2. Evasão. 3. Engenharia agrônômica. I. Rodrigues, Diego. II. Instituto Federal de Educação, Ciência e Tecnologia do Tocantins. III. Título.

Ficha catalográfica elaborada por: Carla dos Santos Figueiredo

Bibliotecária CRB5947

ISAAC PIMENTEL FERNANDES SOBRINHO

**O ESTUDO DA EVASÃO NO CURSO DE BACHARELADO EM ENGENHARIA
AGRONÔMICA NO IFTO *CAMPUS* DIANÓPOLIS, APLICANDO TÉCNICAS DE
MINERAÇÃO DE DADOS**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Licenciatura em Computação do Instituto Federal do Tocantins – *Campus* Dianópolis, como exigência à obtenção do grau de Licenciado em Computação.

Aprovado em: ____/____/____

BANCA AVALIADORA

DIEGO DE CASTILHO RODRIGUES

Mestre em Modelagem Computacional de Sistemas – Prof. Diego Rodrigues.
IFTO – *Campus* Dianópolis

Mestre em Engenharia Elétrica – Prof. Andrea Barboza Proto Sardi.
IFTO – *Campus* Dianópolis

Especialista em Desenvolvimento de Aplicações Para Dispositivos Móveis – Prof.
Marcos Dias da Conceição.
IFTO – *Campus* Dianópolis

À minha mãe Tânia pelo ensinamento pautado
no amor verdadeiro, à minha querida sobrinha
Ana Sophia e aos meus amigos Eric e Gabriel.

AGRADECIMENTOS

A Deus, pela força para superar todos os obstáculos e principalmente os que apareceram durante essa caminhada dolorosa.

À minha mãe Tânia, pelo incentivo pela profissão e dedicação na minha criação e formação, me ajudando nas situações que nunca pensei que eu fosse passar, e superei.

Ao meu pai Isaias, por acreditar e sempre confiar no que sempre fui e no que me tornei ao longo da vida.

Ao meu irmão Herbert e à minha sobrinha Ana Sophia, por me ajudar nos momentos precisos de forma singela, desde um simples sorriso aos abraços que curam.

Aos meus avós, Enedite, Madalena, Moises e Josué, por compreenderem a minha ausência e sempre acreditarem em mim e em tudo o que me proponho a fazer.

Ao meu Orientador, Prof. Diego Rodrigues, por toda confiança depositada desde o primeiro trabalho científico em que realizamos juntos até essa monografia, orientando e aconselhando acerca dos caminhos a serem seguidos.

A todos os meus amigos, com dedicação especial aos melhores: Eric Anderson e Gabriel Rosa, obrigado por me ajudarem em todos os momentos.

Mesmo quando tudo parece desabar, cabe a mim decidir entre rir ou chorar, ir ou ficar, desistir ou lutar; porque descobri, no caminho incerto da vida, que o mais importante é o decidir.

Cora Coralina

RESUMO

Este trabalho visa apresentar os conceitos do fenômeno Evasão no Ensino Superior, e identificar o perfil dos estudantes evadidos do curso de Bacharelado em Engenharia Agrônômica do Instituto Federal de Ciência e Tecnologia do Tocantins *Campus* Dianópolis, por meio de técnicas de mineração de dados, a fim de obter resultados significantes para a tomada de decisões dos gestores e todos os envolvidos no processo educacional dos alunos. O procedimento metodológico consistiu em analisar os dados fornecidos pela Coordenação de Registros Escolares (CORES) dos alunos ingressantes de 2014 a 2019, no qual consistem nas características gerais dos alunos do curso de Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis. Para obtenção dos resultados, após o tratamento dos dados, foi utilizado o algoritmo Apriori nas linguagens Python e R e na ferramenta WEKA, com o objetivo de detectar o padrão dos alunos evadidos. A mineração de dados se fez válida para o objetivo da pesquisa, pois foi detectado o perfil dos alunos evadido, apresentando que os alunos do curso não tem como fator o Sexo para evadirem, mesmo o curso tendo em sua maioria o público do sexo masculino.

Palavras-chave: Evasão, Engenharia Agrônômica, Mineração de Dados.

ABSTRACT

This paper aims to present the concepts of the phenomenon Evasion in Higher Education, and to identify the students profiles evaded from the Bachelor of Agricultural Engineering course of the Federal Institutes Science and Technology From Tocantins Campus Dianópolis, through data mining techniques, in order to achieve significant results for decision making by managers and all those involved in the students' educational process. The methodological process consisted in analyzing the data provided by the School Records Coordination (CORES) from incoming students from 2014 to 2019, which consist in the general characteristics of the students from the Bachelor of Agricultural Engineering - Campus Dianópolis. To obtain the results, after processing the data, the Apriori algorithm was used in Python and R languages and WEKA tool, in order to detect the pattern of evaded students. Data mining was valid for the research purpose, because was detected the students evaded profile, showing that the students of the course don't have the gender factor to evade, even though the course is having more male genders.

Key-words: Evasion, Agronomic Engineering, Data Mining.

LISTA DE TABELAS

Tabela 1 – Possíveis causas para a evasão.....	29
Tabela 2 – Tabela comparativa entre as duas linguagens.....	41
Tabela 3 – Padrão 1 (Python)	51
Tabela 4 – Padrão 2 (Python)	52
Tabela 5 – Padrão 3 (Python)	52
Tabela 6 – Padrão (R)	52
Tabela 7 – Padrão (WEKA).....	53

LISTA DE FIGURAS

Figura 1 – Número de Matrículas de Graduação na Rede Pública – 1980-2017	18
Figura 2 – Número de Universidades que ofertam o curso de Engenharia Agrônômica no Brasil.	21
Figura 3 – Exemplo de Regra de Associação.....	35
Figura 4 – Regiões do Brasil que mais utilizam a linguagem Python	37
Figura 5 – Interface da ferramenta Spyder (Python 3.7)	38
Figura 6 – Interface do RStudio.....	40
Figura 7 – Interface WEKA EXPLORER	44
Figura 8 – Código para extração das regras em Python	49
Figura 9 – Código para extração das regras em R.....	50

SUMÁRIO

1. INTRODUÇÃO	13
1.1 PROBLEMA	15
1.2 OBJETIVOS	15
1.2.1 Objetivo Geral	15
1.2.2 Objetivo Específico	15
1.3 JUSTIFICATIVA	15
2. CONTEXTUALIZAÇÃO DO ENSINO SUPERIOR NO BRASIL	16
3. O Instituto Federal de Educação, Ciência e Tecnologia do Tocantins – IFTO	19
3.2.1 Engenharia Agrônômica – <i>Campus</i> Dianópolis	19
4. EVASÃO NO ENSINO SUPERIOR	24
4.1 Causas Associadas à Evasão	28
5. O PROCESSO DE DESCOBERTA DE CONHECIMENTO EM BANCO DE DADOS – KDD	32
5.1 A Principal etapa do processo de KDD: Data Mining	33
5.2 Regras de Associação: Apriori	34
5.3 Linguagem Python	36
5.3.1 A Ferramenta Spyder	37
5.4 Linguagem R	38
5.4.1 Rstudio	39
5.5 Python e R	41
5.6 Weka	43
6. METODOLOGIA	44
6.1 Tipo de Pesquisa	44
6.2 Universo da Pesquisa	45
6.3 Coleta de dados	46
6.3.1 Tratamento dos Dados	47
6.3.2 Seleção:	47
6.3.3 Pre-Processamento:	47
6.3.4 Transformação:	47
6.3.5 Mineração de Dados:	48
6.4 Regras de Associação em Python:	48

6.5 Regras de Associação em R:.....	49
6.6 Regras de Associação no WEKA:.....	50
7. RESULTADOS.....	51
7.1 Padrões gerados em Python.....	51
7.2 Padrão gerado em R.....	52
7.3 Padrão gerado no WEKA.....	53
8. DISCUSSÕES.....	53
9. CONSIDERAÇÕES FINAIS.....	55
REFERÊNCIAS.....	56
APÊNDICES.....	62

1. INTRODUÇÃO

É indiscutível que a educação é o principal e mais significativo fator de uma sociedade, ela é responsável por todas as outras demandas que rodeiam e regem a população. Sem a educação, os grandes avanços e melhorias que chegaram a todas as áreas que se presencia hoje, não seriam possíveis.

O acesso ao ensino é direito de todos os cidadãos brasileiros, é o que consta na Constituição Federal (CF) de 1988, art. 6º, onde a educação é definida como um direito social e um dever do estado que deve ser bem desenvolvido para que não haja danos à soberania nacional. Além disso, ela é um direito essencial à formação e ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho, como consta no artigo 205 da Constituição.

Contudo, para que o indivíduo possa gozar desse direito, o estado deve também assegurá-lo através de políticas públicas que garantam os direitos fundamentais dos cidadãos, como a educação e os benefícios que a educação traz para ele, são questões que envolvem não apenas o cumprimento da lei ao ofertar os diferentes níveis de ensino, mas de assegurar os estudantes até o final do percurso:

Na contemporaneidade, ao tratarmos da relação entre educação, instituições de ensino e sociedade, inevitavelmente deparamo-nos com algumas questões conflitantes, dentre elas, a retenção e a evasão merecem destaque. Da educação básica à educação superior, em todos os níveis e modalidades de ensino, esses problemas estão presentes (MINISTÉRIO DA EDUCAÇÃO, 2014).

Entre os problemas presentes na educação superior, a evasão merece uma atenção especial e um estudo dedicado, pois se trata de um assunto delicado e preciso, e em comparação com outros fenômenos, ainda é pouco estudado no sistema educacional brasileiro.

O estudo da evasão é recente no Brasil, teve um marco institucional em 1995, quando o Ministério da Educação e Cultura (MEC) criou a Comissão Especial para o Estudo da Evasão nas Universidades Brasileiras, reunindo diferentes instituições de ensino superior – IES públicas para discutir a temática e organizar estudos que definissem uma fórmula de cálculo para a evasão para então, encontrar possíveis soluções. Ao finalizar o estudo, obtiveram dados importantes sobre as universidades públicas brasileiras, com altos índices de diplomação, retenção e evasão dos cursos de graduação. Portanto, a evasão é um problema recorrente em todo sistema de ensino:

Cabe destacar que esse é um problema recorrente em todo o mundo. Na África do Sul e nos EUA, cerca de 40% e 50% dos discentes abandonam o curso, respectivamente, já em países europeus, como Irlanda e Inglaterra, as taxas de evasão tendem a ser menores. O destaque fica para o Japão, com índice médio de 7%, país que apresenta a menor taxa de evasão do mundo (FURTADO E ALVEZ, 2012).

A importância de estudar a evasão vai além da obtenção de números e porcentagens, os estudos sobre evasão podem orientar políticas da educação para melhorar as vertentes institucionais:

[..] o estudo, por seu caráter, torna-se subsídio valioso à condução de uma avaliação objetiva dos resultados do sistema, avaliação esta indispensável para orientar políticas institucionais e governamentais mais eficazes, no sentido da melhoria do ensino de graduação (BRASIL/MEC, 1997).

Diante do exposto, este estudo de evasão foi realizado no curso de Engenharia Agrônoma do Instituto Federal de Tecnologia e Ciências do Tocantins – *Campus* Dianópolis, visto que há poucos estudos acerca da temática no *campus* e principalmente no curso citado, além disso, não só a região, mas como o país necessita de profissionais da engenharia, como defende Borges e Almeida (2013, p.72) “a educação em Engenharia no Brasil constitui um desafio diante de um cenário mundial que demanda uso intensivo de tecnologias e que exige, cada vez mais, um maior número de profissionais altamente qualificados”.

Não atender a essa demanda de profissionais altamente qualificados é deixar país atrasado no âmbito de desenvolvimento científico, abandonando oportunidades de competição do mercado e de produtos com tecnologias inovadoras. Portanto, é válido afirmar que um dos fatores que dificultam a inversão de novos profissionais no mercado de trabalho brasileiro é o alto número de evasão nos cursos superiores.

Lucchiari (2000) afirma que é preocupante e assustador o número expressivo de evasão no ensino superior, em um Brasil que necessita de profissionais em todas as áreas. A evasão nos cursos de Engenharia, por exemplo, acarreta de forma significativa ao atraso e não satisfação das necessidades do mercado interno, essa afirmação se fundamenta nos inúmeros estudos feitos por pesquisadores que apresentam esse fato e ainda mais, esse quadro se torna ainda mais agravante ao ser comparado com outros países.

A financiadora de Estudos e Projetos (FINEP) é um órgão governamental que patrocina programas de estímulo à formação de Engenheiros no Brasil, no entanto, as instituições de ensino não correspondem à demanda, deveriam ampliar

as vagas e formar mais profissionais. Essa situação não é preocupante ao analisar o número de vagas ofertadas, mas pelo alto índice de evasão: desistências, abandonos e trancamentos.

Contudo, as instituições de Ensino Superior precisam conhecer os desempenhos dos seus alunos, em especial no Ciclo Básico, e desenvolver mecanismos e/ou políticas institucionais que, não solucionarem ao menos minimizarem os problemas de retenção, reprovação e evasão nos cursos de engenharia (OLIVEIRA, et. al, 2010).

1.1 PROBLEMA

As ferramentas de mineração de dados podem ajudar a compreender o fenômeno da evasão no curso de Engenharia Agrônômica no Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, *campus* Dianópolis?

1.2 OBJETIVOS

1.2.1 Objetivo Geral

- Identificar o padrão dos estudantes evadidos do curso de Bacharelado em Engenharia Agrônômica no Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, *campus* Dianópolis, por meio de técnicas de Mineração de dados.

1.2.2 Objetivo Específico

- Analisar os dados dos acadêmicos do curso de Bacharelado em Engenharia Agrônômica;
- Gerar regras de associação por meio do algoritmo Apriori nas linguagens Python e R, e na ferramenta WEKA;
- Analisar as regras geradas e identificar o perfil do aluno evadido do curso de Engenharia Agrônômica.

1.3 JUSTIFICATIVA

De acordo com Santana (1996, p. 29)“A evasão é um dos maiores e mais preocupantes desafios do Sistema Educacional, pois é fator de desequilíbrio, desarmonia e desajustes dos objetivos educacionais pretendidos” (Santana 1996).

Durante o Congresso Brasileiro de Educação em Engenharia (COBENGE) foi apresentado um estudo aos participantes baseado nos dados fornecidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais – INEP, o estudo aponta a formação de engenheiros no Brasil e no mundo entre os anos de 1995 e 2010. Este levantamento apontou que cerca de 250 mil vagas são ofertadas, mas o número de alunos concluintes não ultrapassa os 16%, ou seja, existe muita evasão nas engenharias.

Outro fato interessante apontado pelo estudo é que, a quantidade de inscritos para prestar vestibular é equivalente ao dobro das vagas ofertadas pelas IES. O presidente do CONFEA, eng. civ. José Tadeu da Silva afirma que:

“os cursos de engenharia, hoje, são os que apresentam a maior concorrência no vestibular. Está superando a medicina. A engenharia na USP, em São Paulo, supera a medicina. É altamente procurada. E isso acontece porque há espaço. Há perspectiva de desenvolvimento e de mercado de trabalho com bons salários. O jovem percebe isso, e percebe que o país sinaliza que é um mercado promissor para ele poder trabalhar.” (COBENGE, 2013).

Diante do exposto, essa é uma contribuição científica em relação aos índices de evasão no curso superior de Engenharia Agrônômica ofertado no Instituto Federal de Educação, Ciências e Tecnologia do Tocantins (IFTO) – *Campus* Dianópolis, a fim de fornecer um aporte para que a Instituição (Diretoria de Ensino, Coordenador de Curso e demais setores) tenha uma perspectiva abrangente sobre o fenômeno da evasão e quais estratégias podem ser adotadas para o combate, aplicando técnicas de mineração de dados para identificar os padrões dos alunos evadidos.

2. CONTEXTUALIZAÇÃO DO ENSINO SUPERIOR NO BRASIL

A educação básica é formada por três níveis: educação infantil, ensino fundamental e ensino médio. Esses níveis ofertados para todos os cidadãos brasileiros visam assegurar todos às competências gerais que garantem a aprendizagem e o desenvolvimento de cidadãos éticos e atuantes. Ensino Superior, antes chamado terceiro grau é um nível de educação que vem após a educação básica, especificamente após o ensino médio.

O Ensino Superior é o mais elevado dos sistemas educativos, é uma educação ofertada nas universidades, faculdades, institutos politécnicos, escolas superiores e outras instituições com graus acadêmicos ou diplomas. Desde 1966 o

mundialmente conhecido Pacto Internacional dos Direitos Econômicos, Sociais e Culturais, em seu artigo 13º diz que “a educação superior deverá tornar-se de acesso igualitário para todos, com base na capacidade, por todos os meios apropriados e, em particular, pela introdução progressiva da educação gratuita”.

Com a chegada da Família Real Portuguesa ao Brasil, registra-se a primeira criação da escola de ensino superior no Brasil: em 1808 as Faculdades de Cirurgia, em Salvador – Bahia e a de Medicina do Rio de Janeiro (CURI, 2011). A partir de 1930, com o avanço e o aumento da população no qual a industrialização se concentrava intensamente nos centros urbanos, fez com que houvesse exigência maior em relação à educação. Principalmente na Região Sudeste do país, sobretudo em São Paulo e no Rio de Janeiro, onde houve maiores demandas por ensino superior (CACETE, 2014).

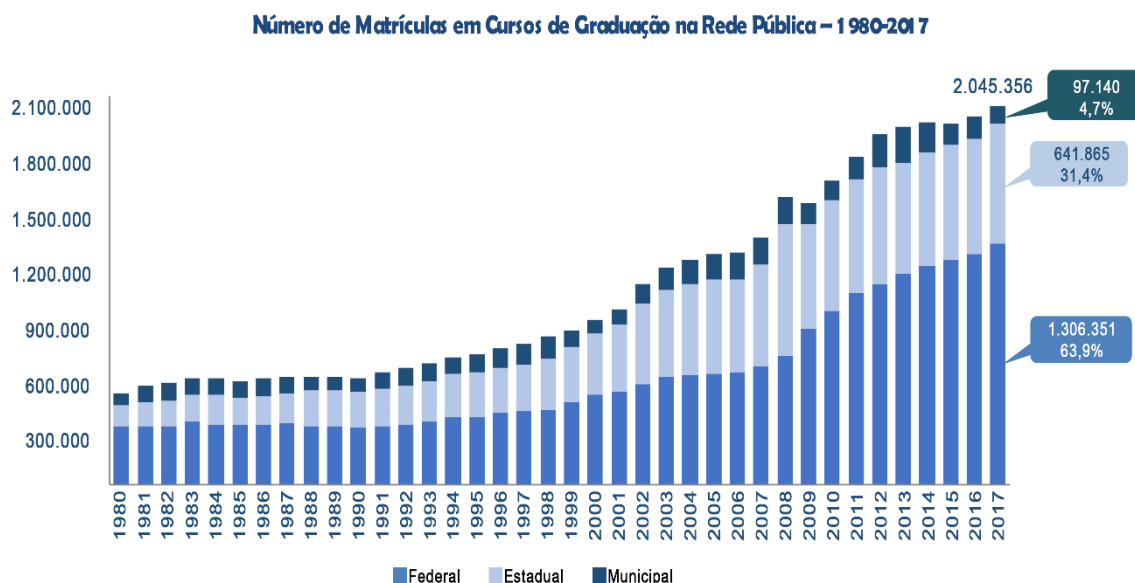
Segundo Cacete (2014, p. 1063) “o surgimento do ensino superior no país deu-se inicialmente sob a forma de cadeiras que foram sucedidas por cursos, posteriormente por escolas e por faculdades de medicina, direito, engenharia, agronomia etc., todas eminentemente profissionais”.

O processo de estruturação do ensino superior no Brasil teve uma forte e inegável expansão de 2003 para 2014, no qual, o número de universidades federais saiu de 43 para 63. Já o número de municípios atendidos pelo ensino superior saiu de 114 para 289, no mesmo período (BRASIL, 2014).

Dados divulgados do INEP sobre o censo da educação superior de 2017 mostram que o Brasil possui 2.448 instituições de ensino, sendo 2.152 privadas e 296 públicas. Das 296 públicas, 109 são da rede Federal, 124 da rede Estadual e 63 da rede municipal.

O Panorama do Ensino Superior no Brasil elaborado pelo Sindicato das Entidades Mantenedoras de Estabelecimentos de Ensino Superior (SEMESP) apresenta os números de matrículas por região, nos cursos presenciais constam na região Norte 423,6 mil matrículas, no Nordeste 1,3 milhão de matrículas, na região Centro-Oeste 578,5 mil matrículas, na região Sudeste 2,9 milhões e no sul do país 963,4 mil matrículas, sendo no Sudeste o maior número de matriculados e o Norte com o menor.

Na rede pública, os números de matrículas cresceram entre 2016 e 2017 2,8%, principalmente na rede federal com mais de 1,3 milhão de alunos em 2017. Como apresenta nos dados do INEP (2018):

Figura 1 – Número de Matrículas de Graduação na Rede Pública – 1980-2017

Fonte: INEP (2018)

A região norte do Brasil, em específico o estado do Tocantins (onde se localiza o IFTO *campus* Dianópolis, local de estudo deste projeto) tem uma população estimada de 1,5 milhão, formado por 139 municípios, possui 26 instituições de ensino superior. A mesorregião Ocidental do estado possui 93 municípios, e a Oriental possui 46. A primeira foi responsável por cerca de 26 mil matrículas e a outra cerca de 22 mil (SEMESP, 2015).

As ofertas de educação superior alcançaram além dos setores público e privado, as modalidades de ensino a distância, onde os alunos estudam por meio de plataformas digitais, a faculdade fornece os conteúdos ao estudante, sendo em alguns casos, necessária a presença do aluno apenas no dia da realização da prova presencial, facilitando o acesso em relação à disponibilidade de tempo do discente para com as aulas.

De 2009 a 2013 as matrículas nos cursos ofertados a distância no Brasil registraram um acréscimo de 37,5%, um aumento de 50% na rede privada e uma queda de 10,5% na rede pública. No estado do Tocantins, as matrículas em Cursos Presenciais privados obtiveram nos últimos 13 anos um acréscimo de 1.000% em relação ao número de matrículas, já o setor público apresentou um aumento de 247% (SEMESP, 2015).

Os dados apresentados acerca do avanço da educação superior e as vagas ofertadas não contêm os alunos nas instituições de ensino, surge a partir disso um grande problema enfrentado pelas IES do mundo todo, a evasão escolar.

3. O Instituto Federal de Educação, Ciência e Tecnologia do Tocantins – IFTO

Com todo o avanço no ensino superior no Brasil, o Instituto Federal de Educação, Ciência e Tecnologia do Tocantins - IFTO vem se destacando na qualidade do ensino dos cursos ofertados, se preocupando com o mercado de trabalho em que os profissionais qualificados atuarão e também com desenvolvimentos que contribuem para o avanço e melhoria da população tocaninense.

São mais de sessenta cursos, nos níveis básicos, superior e pós-graduação, que atendem mais de 13 mil estudantes. O IFTO traz à sociedade brasileira a experiência e o dever cumprido de um centenário em educação técnica, com o objetivo de avançar na integração do ensino, pesquisa e extensão, atendendo às demandas do mundo do trabalho, em consonância com os arranjos produtivos regionais e locais, gerando melhoria de vida para os tocaninenses, proporcionando desenvolvimento educacional, científico e tecnológico do Estado (IFTO, 2014, p.7-8).

O IFTO foi criado por meio da Lei nº 11.892/2008, que instituiu a Rede Federal de Educação Profissional, Científica e Tecnológica. Possui sete campi em pleno funcionamento distribuídos por todo o Estado, localizados em Palmas, Porto Nacional, Paraíso do Tocantins, Gurupi, Araguaína, Araguatins e Dianópolis, um *campus* em fase de implantação localizado em Colinas do Tocantins, três campi avançados em implantação, localizados nos municípios de Lagoa da Confusão, Formoso do Araguaia e Pedro Afonso, além da educação a distância, que oferece educação técnica em 16 polos (IFTO, 2014, p.7).

3.2.1 Engenharia Agrônômica – *Campus* Dianópolis

A agronomia foi oficializada no ano de 1848, mesmo ano da fundação do Instituto Nacional de Versailles, na França. A primeira escola de agronomia foi fundada em 1802 na Alemanha, por Samuel Tessedik, no Brasil a primeira escola a ofertar agronomia foi só em 1875, na cidade de São Bento das Lajes, Bahia. Mesmo com a criação das escolas citadas, não havia um curso de Agronomia, o reconhecimento da profissão só foi regulamentado em 1993, pelo decreto 23.196/1933.

A agronomia é uma ciência agrária que estuda os aspectos físicos, químicos e biológicos salientados pela agricultura, com isso, é encarregado ao engenheiro agrônomo um importante papel em todo o processo, desde o plantio à colheita, processamento e comercialização dos produtos agrícolas, e na parte administrativa de propriedades, monitorando e preparando o solo, entre outras importantes funções. A presença do profissional na área garante a qualidade e produtividade nos serviços voltados para o desenvolvimento dos municípios e qualidade de vida da população. Para exercer a agronomia, o profissional deve ser formado no curso de bacharelado em agronomia ou engenharia agrônômica e obter o registro no Conselho Regional de Engenharia e Agronomia (CREA) do estado.

É indispensável a importância da agronomia para o desenvolvimento econômico do país, o aumento de engenheiros agrônomos no mercado de trabalho apresenta resultados cada vez mais positivos, segundo Paiva (2017):

Este aumento da produção tem levado o setor do agronegócio a representar 21% do PIB nacional, no ano de 2015, com um valor de 1,27 trilhões de reais. Com certeza, podemos afirmar que o engenheiro agrônomo tem contribuído significativamente para estes números. A participação dos engenheiros agrônomos está atrelada ao planejamento e a supervisão do uso dos princípios e processos básicos da produção agrícola, que combina os conhecimentos de biologia, química, física e matemática aos estudos específicos sobre a planta, o solo e ao clima.

É válido ressaltar que a agronomia também acompanha os avanços tecnológicos, o que contribui de forma significativa para o avanço do país além dos aspectos econômicos, sociais e tecnológicos, como afirma VANDERLÍ e NIVAL (2010, p.22):

À medida que a tecnologia vai se tornando mais complexa, em termos de necessidade de conhecimentos de base matemática, física, química, expressão gráfica, entre outros, para solucionar problemas e projetar soluções, torna-se objeto de estudo e aplicação do campo da Engenharia.

Nessa mesma visão, BACHA (2012, p.24) discorre sobre os papéis da agropecuária no desenvolvimento econômico:

A literatura sobre desenvolvimento econômico considera que a agropecuária tem, basicamente, cinco funções importantes no processo de desenvolvimento econômico, as quais são:

- Fornecer alimentos para a população total;
- Fornecer capital para a expansão do setor não agrícola;
- Fornecer mão de obra para o crescimento e diversificação de atividades na economia;
- Fornecer divisas para a compra de insumos e bens de capitais necessários ao desenvolvimento de atividades econômicas;
- Constituir-se em mercado consumidor para os produtos do setor não agrícola.

A essas funções pode-se acrescentar, no caso do Brasil, a importância da agropecuária em fornecer matéria-prima necessária ao processo de desenvolvimento industrial.

Com isso, é possível refletir sobre a importância da formação de Engenheiros Agrônomos para o país, mas se comparado a outros países, os resultados não alcançam as necessidades, afirma VANDERLÍ e NIVAL (2010, p. 65):

Nesse cenário, a Engenharia tem uma importância estratégica para o desenvolvimento socioeconômico do país, pois está intimamente ligada à sua capacidade de desenvolver ou absorver novas tecnologias. Entretanto, se comparada a outros países, a formação em Engenharia no Brasil não tem alcançado uma dimensão proporcional à sua necessidade de desenvolvimento.

No Brasil cerca de 7 mil agrônomos(as) se formam por ano e aproximadamente 150 mil estão em atividades. Há cerca de 223 universidades de ensino distintas que ofertam o curso de Engenharia Agrônômica na modalidade presencial no Brasil, segundo o Ranking Universitário Folha (2017) como apresentado no seguinte mapa da figura 2:

Figura 2 – Número de Universidades que ofertam o curso de Engenharia Agrônômica no Brasil.



Fonte: O autor

O estado do Acre (AC) é contemplado por uma instituição que oferta o curso, o estado de Alagoas (AL) também apenas por uma universidade. No estado do Amapá (AP) até o ano da divulgação desses dados em 2017 não constava

nenhuma universidade ofertante do curso de Engenharia Agrônômica (mas em 2018 uma Universidade passou a ofertar).

No Amazonas (AM) há uma universidade ofertante, já na Bahia (BA) há 12 universidades que ofertam, o Ceará (CE) possui 4, no Distrito Federal (DF) 3 como também no Espírito Santo (ES). No Goiás (GO) há 13 universidades ofertantes do curso de engenharia, no Maranhão (MA) há apenas 3, já o estado do Mato Grosso (MT) possui 11 universidades ofertantes e no Mato Grosso do Sul (MS) 9.

O estado de Minas Gerais (MG) é contemplado por 33 universidades que ofertam o curso de Engenharia Agrônômica, já no Paraná (PR), 25 instituições. Na Paraíba (PB) possui apenas 2 universidades ofertantes do curso e no Pará (PA) 6. No Pernambuco (PE) há 5, e no Piauí (PI) apenas 2, assim como no estado do Rio de Janeiro (RJ) e Rio Grande do Norte (RN) que também possuem apenas 2 universidades ofertantes do curso de Agronomia.

No Rio Grande do Sul chega a 18 o número de universidades que ofertam o curso de Engenharia Agrônômica. Já em Rondônia (RO) contém apenas 6 e em Roraima (RR) este número cai para 3. Em Santa Catarina (SC) há 11 instituições de ensino que ofertam o curso supracitado, o que difere do estado de Sergipe (SE) que possui apenas 1. No estado de São Paulo (SP) esse número aumenta para 40 universidades ofertantes, e no Tocantins (TO) o número cai para 5.

Os primeiros engenheiros agrônomos são datados nos registros de 1942 no Rio de Janeiro e São Paulo, mas no Tocantins foi registrado apenas em 1991. A demora aconteceu pela emancipação do estado, que só veio ocorrer em 1998.

A Secretaria de Agricultura, Pecuária e Aquicultura do Tocantins aponta o estado do Tocantins como o “novo” polo agrícola do Brasil, e salienta que o estado se destaca como o maior produtor de grãos da região Norte do Brasil, sobretudo de soja, arroz, milho e feijão, isso porque metade do território do estado possui potencial para a agricultura, “são terras férteis, de valor competitivo no mercado e de topografia plana, o que favorece o processo de mecanização agrícola.” (SEAGRO, 2019)

Com isso, o estado do Tocantins contribui de forma significativa na economia do país, como consta no relatório sobre a agropecuária do estado do Tocantins pela SETEC (2019):

O Estado de Tocantins caracteriza-se por ter na agricultura o setor que absorve a maior parte da população economicamente ativa. A agropecuária

contribui com 60% do PIB, e é a principal atividade econômica do Estado, com destaque para a pecuária.

Contudo, percebe-se a relevância do curso de Engenharia Agrônômica no Tocantins, pois o papel do Engenheiro Agrônomo em todo o processo de desenvolvimento é de importância inegável:

Se aprofundarmos um pouco mais e observarmos o histórico da produtividade média de grãos no Tocantins desde a sua criação em 1988, crescemos mais de 54%, saímos de pouco mais de 2200 kg/há para quase 3400 kg/há nesta safra. Pode parecer pouco olhando de maneira direta, contudo, esses números expressivos são reflexos de anos de trabalho de Agrônomos e seus colaboradores que vão desde avaliação de doses de fertilizantes, manejo, controle fitossanitário, variedades geneticamente modificadas e que acredito ser o principal, a sustentabilidade do sistema de produção. São muitos os exemplos que estão sendo trabalhos a fim de se alcançar maiores incrementos de produtividade, consumindo cada vez menos recursos naturais e gerando com isso conseqüentemente, um maior equilíbrio para o sistema produtivo (JÚNIOR, 2019).

O autor JÚNIOR (2019) ainda aborda sobre a necessidade da pesquisa, desenvolvimento e da inovação para alcançar os objetivos da agronomia, o que o engenheiro adquire também durante sua formação acadêmica:

O grande volume do que produzido através da pesquisa, do desenvolvimento e da inovação e que tem como fim chegar a esse equilíbrio, passa pelas Autarquias municipais, estaduais e federais, Organizações Estaduais de Pesquisa, Universidades, Embrapa e ONGs (JÚNIOR, 2019).

O *Campus* Dianópolis está sediado no município de Dianópolis, localizado no sudeste do estado do Tocantins. Dianópolis possui cerca de 21 mil habitantes e é contemplada por diversas questões culturais por se tratar de uma cidade histórica com importante relevância para o estado do Tocantins.

O antigo Instituto de Menores de Dianópolis, hoje *Campus* Dianópolis, apresenta uma vocação naturalmente agrícola, pois surgiu a partir da doação, por parte de Estado do Tocantins, de uma área rural de aproximadamente 593ha, onde funcionou, por muitos anos, a Fundação Agroindustrial São José, mais conhecido como Instituto de Menores de Dianópolis. Um grande potencial para agricultura, agropecuária e aquicultura pode e dever aproveitado, pois a área do *Campus* conta com represa e tanques para cultivo aquícola, um pivô para irrigação com capacidade para cobrir até 10ha, rebanho de gado leiteiro e de corte, suínos, equinos, ovinos e aves (IFTO, 2014, p. 12).

O *campus* oferece dois cursos superiores, sendo Licenciatura em Computação e Engenharia Agrônômica. O curso de Engenharia Agrônômica na região contribui para a agropecuária e para expansão da produção através da formação de profissionais qualificados para o trabalho prático. Além disso, o Instituto

Federal insere ao longo da formação do aluno atividades de pesquisa, ensino e extensão e desenvolvimento da ciência.

O curso a nível superior é da área de Ciências Agrárias, ofertadas 40 vagas anuais na modalidade de bacharelado, sendo presencial no turno integral. A duração do curso é de quatro anos (oito semestres), com máximo de oito anos (dezesseis semestres), a carga horária é de 4.850 horas, cada hora-aula é de 40 minutos. Dentro da carga horária total, constam 240 horas de estágio supervisionado, 110 horas para atividades complementares e 100 horas obrigatórias para o Trabalho de Conclusão de Curso.

O corpo docente atual do curso pretendido é composto por profissionais qualificados, como especialistas, mestres e doutores. Como supracitado, o compromisso do curso ofertado pelo *Campus* Dianópolis abrange uma visão social e profissional, como consta no Projeto Pedagógico do Curso (PPC):

O compromisso do curso de Engenharia Agrônoma do *Campus* Dianópolis, do IFTO, com a sociedade brasileira traduz-se na formação de profissionais dotados de uma visão interdisciplinar e que sejam capazes de contribuir para o desenvolvimento da ciência e da tecnologia como cidadãos partícipes e comprometidos com a construção de uma sociedade justa, equilibrada e autossustentável. Esse compromisso está em estreita consonância com a filosofia que tem norteado as atividades de ensino, pesquisa e extensão praticadas no Instituto Federal de Educação, Ciência e Tecnologia do Tocantins (IFTO, 2014, p. 12).

Portanto, percebe-se a importância e a responsabilidade do IFTO – *Campus* Dianópolis não apenas com a formação do acadêmico de Engenharia Agrônoma, mas com as atividades que serão desenvolvidas por ele no mercado de trabalho enquanto profissional altamente qualificado. Mas, além da oferta do ensino de qualidade, é necessário se atentar às questões relacionadas à permanência do estudante na Instituição, pois a evasão do aluno se traduz em perdas irreparáveis tanto sociais quanto de recursos, que todos envolvidos no processo de ensino e aprendizagem tendem a sofrer.

4. EVASÃO NO ENSINO SUPERIOR

A evasão é considerada um fracasso escolar, ligado aos elementos que resultam no desligamento do aluno do ensino superior seja abandono, desistência, transferência, trancamento etc., tendo como consequência diversos fatores, como afirma Batista, Sousa e Oliveira (2009, p.4):

O abandono à escola é composto então pela conjugação de várias dimensões que interagem e se conflitam no interior dessa problemática.

Dimensões estas de ordem política, econômica, cultural e de caráter social. Dessa maneira, o abandono escolar não pode ser compreendido, analisado de forma isolada. Isto porque, as dimensões socioeconômicas, culturais, educacionais, históricas e sociais entre outras, influenciam-se mutuamente.

Os estudos científicos sobre Evasão escolar começaram em 1975, pelas teorias elaboradas com o objetivo de explicar a permanência dos estudantes nas Instituições de Ensino, baseadas em análises voltadas à finalidade estudante-instituição, por meio de variáveis ligadas aos estudantes, instituição e temas como a integração dos estudantes.

Antes dessas, prevaleciam teorias descritivas, procuravam oferecer explicações baseadas em variáveis demográficas ou psicológicas. A partir dessas vertentes, surgiram os grandes autores Ajzen (1975) e Ethington (1990) que foram precursores dos estudos de Vincent Tinto (reconhecido como o maior especialista no assunto).

Para Ajzen, o aluno observa e analisa as interações do dia a dia, e age de acordo com o sentido que essa análise do cotidiano fornece a ele, podendo ser um resultado de aceitação ou de rejeição da ideologia acerca da formação superior para sua vida futura. Para Ethington, o estudante estabelece metas para si mesmo, metas essas que já estavam inclusas em seu modelo psicológico sobre a temática.

Para os dois autores, questões como influências, origem demográfica e aspectos psicológicos afetam nas ideias e expectativas dos estudantes, influenciando assim, diretamente na decisão de permanecer na IES ou evadir.

Spady, em 1975 estudou e apresentou um modelo no qual se baseou na teoria do suicídio, de Durkheim. A teoria de Durkheim defende que o indivíduo comete suicídio quando há o rompimento dos laços sociais, pela falta de integração e interação com a sociedade.

A partir disso, Spady traçou sua teoria sobre o estudante que evade, defendendo que as relações entre amigos influenciam diretamente na permanência, bem como o apoio familiar, quando não há apoio das amigas e no seio da família, o aluno tende a evadir, sendo a integração um fator essencial para a permanência. Em traço com a teoria de Durkheim, quando o indivíduo rompe seus laços sociais e não consegue atender as demandas acadêmicas, ele evade, de forma semelhante aos motivos que levam um suicida a tirar a própria vida.

O autor mais conhecido do tema Vincent Tinto expandiu o trabalho de Spady e incluiu na teoria do suicídio a teoria da troca, no qual apresenta que o

indivíduo foge de comportamentos muito custosos e busca status, relações, interações e estados emocionais satisfatórios a si mesmo.

Para Tinto, a teoria da troca se aplica ao estudante para definir sua integração na academia e na sociedade. Se o aluno perceber que tudo aquilo que ele está vivendo na Universidade está lhe ofertando benefícios e esses benefícios são maiores que os custos, ele permanece, caso contrário, ele evade.

No Brasil, a evasão é um objeto de estudo recente, surgiu em 1995 através de uma avaliação institucional com o Programa de Avaliação Institucional das Universidades Brasileiras (PAIUB), onde foi realizado o primeiro seminário sobre a temática, estudo esse dirigido totalmente às 61 instituições públicas brasileiras, o que representava 77,2% da educação superior pública do país. Ao término do evento, o MEC elaborou objetivos para encontrar soluções para o combate à evasão, instituindo a Comissão Especial de Estudos sobre a Evasão.

Visando o estabelecimento de um panorama de evasão na educação superior no país, Silva Filho *et al* (2007) realizaram estudos com os dados que foram fornecidos pelo MEC de 2000 a 2007, concluindo que a taxa anual média de evasão era de 22%. Esse estudo também informou que a taxa de evasão no Brasil é menor do que em outros países da América Latina. No entanto, se comparado com outros países como a Itália, Suécia e França, as taxas são maiores. Com isso, é provado que a evasão é um problema recorrente nas Universidades de ensino de todo o mundo, sendo então, um problema internacional.

Atualmente, a evasão é definida como um termo polissêmico, podendo ser caracterizada como abandono, desistência, fracasso, transferência, saída definitiva e/ou suspensão temporária do curso, da instituição e/ou do sistema escolar. É a situação em que o estudante deixa de renovar a matrícula do curso superior e interrompe o seu ciclo de estudos em um curso.

Para o MEC (1997), a evasão é compreendida como a saída definitiva do estudante do curso de origem sem o concluir. No entanto, essa definição não abrange as transferências de curso interna e nem externas. Além dos diversos conceitos sobre evasão, nem todas as pesquisas se identificam com o termo evasão, passando a adotar o antônimo: permanência. Schmitt (2014) destaca que a relação entre os dois termos leva a uma mesma unidade de pesquisa, pois as questões de vulnerabilidade, preditivos e preventivos se cruzam. Neste contexto, os autores Baggi e Lopes (2011, p.371) consideram que:

A evasão tem múltiplas razões, dependendo do contexto social, cultural, político e econômico em que a instituição está inserida. Pode estar relacionada, por exemplo, a má qualidade de ensino oferecido pela IES, provocando a perda definitiva do aluno.

Há três tipos de evasão classificadas pela Comissão de Estudos sobre a Evasão nas Universidades Públicas Brasileiras, são eles: Evasão dos Sistemas, Evasão da Instituição e Evasão do Curso. Evasão dos Sistemas acontece quando o discente chega à desistência, ou seja, abandona a IES na qual está matriculado, de forma permanente ou temporária.

A evasão da Instituição ocorre quando o aluno passa pela transferência externa ou aprovação no vestibular para cursar em outra instituição, ou seja, quando ocorre o abandono do discente que está matriculado na IES para outra.

A evasão do Curso acontece quando o discente passa por transferência interna ou é aprovado no vestibular para outro curso na mesma instituição, ou seja, quando o aluno se desliga e não conclui o curso em que primeiro se matriculou.

Os níveis citados estão ligados a fatores que interferem na evasão como os fatores internos e externos às instituições e aos fatores individuais dos alunos. Os fatores internos estão ligados a deficiências estruturais físicas na IES, má atuação dos professores, má assistência socioeducacional, ações como atividades de pesquisa e extensão, grade curricular/turno e assistência aos alunos de baixa renda, ou seja, práticas voltadas para a integração do aluno com a IES, que visam a permanência e o bom desenvolvimento do acadêmico (MEC/SESU, 1997).

Os fatores externos estão ligados ao que ocorre fora da instituição, mas que influenciam diretamente na permanência do aluno, como tomada de decisões em relação ao curso, falta de orientação sobre o curso, imaturidade, curso de segunda opção, busca pela herança profissional, pressão advinda da família, escolha do curso de baixa concorrência visando à facilidade no egresso, descontentamento com o curso e sua futura profissão, razões socioeconômicas, distância entre a universidade e a residência do aluno, problemas pessoais e outros fatores externos contribuem para a evasão do aluno (MEC/SESU, 1997).

Já os fatores individuais do aluno contribuem para que ele venha a evadir porque estão ligados a habilidades de estudo, personalidade, formação escolar anterior, escolha da profissão, dificuldades pessoais, dificuldades na adaptação da mudança em relação à vida universitária, dificuldades em relação às notas e reprovações (MEC/SESU, 1997).

4.1 Causas Associadas à Evasão

Conhecer o perfil do estudante que evade no Ensino Superior é de válida importância para previsão daqueles que apresentam ser propensos a evadirem, com as características pode se pensar e elaborar estratégias para diminuir as chances desse abandono.

Diversos autores identificam inúmeros aspectos característicos dos estudantes que tendem a evadir, Santos (2014), Baggi e Lopes (2011) identificaram que a escolha da profissão é uma variável do perfil do estudante, e salientam sobre a importância dos programas de orientação profissional como um auxílio para o estudante na escolha da profissão. Os autores ainda informam que problemas como descontentamento com a carreira, com o curso e/ou com a profissão acontece com mais frequência nos cursos de licenciatura: "falta de identidade com o curso, escolha errada da carreira, desencanto com a universidade, baixa demanda pelo curso, possivelmente associada ao baixo prestígio do curso escolhido, entre eles, as licenciaturas (BAGGI E LOPES, 2011, p.361)".

Entre tantos fatores que destinam à evasão, alguns estão relacionados a vertentes pessoais, como a falta de identificação com o curso (LOPES, 2014), a descoberta de novos interesses e/ou áreas do conhecimento (AMARAL, 2013) e à falta de compromisso com a graduação (CASTRO, 2012).

A falta do bom relacionamento entre discentes e docentes também pode provocar o abandono do ensino, visto que esses possuem os principais papéis em todo o processo de ensino-aprendizagem, como é apresentado nas análises de Sobrinho (2013) sobre as relações entre professores e alunos.

Há também características voltadas para condições socioculturais dos alunos, como baixo capital cultural e econômico das famílias (PARENTE, 2014 e LOURENÇO, 2014), pois estudantes universitários que não possuem nenhum auxílio financeiro estudantil e não se envolvem nas atividades extracurriculares têm mais probabilidade à evasão (RAMOS, 2013).

Essa corrente gerada pelos resultados dos estudos feitos pelos autores citados relaciona fortemente o capital cultural familiar aos resultados escolares. Lima (2013) ao analisar as possibilidades estatísticas relacionadas à conclusão do curso em uma instituição pública demonstrou por meio dos resultados que os acadêmicos no qual os pais possuíam formação superior ou pós-graduação se formavam mais

rápido, e os que os pais não possuíam formação superior, demoravam mais para se formar.

O autor observa que por suposição, os acadêmicos que vem de classes populares e de elite têm as mesmas condições de evadirem, mas o que diferencia é que a evasão dos alunos na sociedade elite acontece em direção aos cursos mais aclamados, já nas classes populares está ligada diretamente ao fracasso escolar.

Os estudos indicam que a evasão tem razões e probabilidades associadas a diversos aspectos, podendo ser pessoais, culturais, sociais ou econômicos. Esses aspectos podem interferir de forma direta ou indireta no percurso do indivíduo, por isso, é de fundamental importância compreender os motivos que levam o acadêmico a interromper os estudos, mas esse processo não é algo tão simples:

Compreender o que leva um estudante a interromper seus estudos ou deixar de concluir o curso de graduação é uma tarefa semelhante à montagem de um tangram. Esse jogo é uma espécie de quebra-cabeça de origem milenar, composto por sete peças geométricas que formam um quadrado na posição original. As partes podem ser dispostas de diferentes formas, gerando inúmeros desenhos, mas elas serão sempre as mesmas (LIMA e ZANGO, 2018).

Essa metáfora referida por Lima e Zago (2018) demonstra de modo figurativo de como a evasão se conceitua e apresenta suas causas por diferentes vieses, mostrando os possíveis arranjos que se pode ter ao montar um desenho com tangram, destacando a convergência do jogo, assim a evasão se constrói por diferentes modelos, mas todos basicamente parecidos ou iguais.

Na literatura, grandes autores apresentam abordagens e motivos para o fenômeno, como resumido na Tabela 1, por Hoffmann, Nunes e Muller (2019):

Tabela 1 – Possíveis causas para a evasão.

Autor	Abordagem	Motivos
Lobo (2012)	Abordagem ensino público e privado.	Os principais problemas estão relacionados à gestão universitária.
Oliveira (2009)	Construir a motivação do aluno em virtude da necessidade de retê-lo na instituição a fim de moldá-lo, orientá-lo, formá-lo, encaminhando cidadãos conscientes para a formação da sociedade.	As IES não conhecem o perfil de seus alunos.
Tigrinho (2008)	Abordagem no ensino público e privado e formação de ações para combater o fenômeno.	Os principais problemas estão relacionados à gestão da IES com poucas ações de combate à evasão.
	Preocupação com a	

Dias (2008)	responsabilidade social do cidadão.	Imaturidade dos estudantes.
Silva et al. (2007)	Ambiente nacional e internacional	Econômico-financeira
Costa & Gouvinhas (2005)	Panorama da evasão no ensino superior privado.	Quando IES: Infraestrutura, corpo docente, matriz curricular. Quando Aluno: situação econômico-financeira, incompatibilidade de horário e desempenho do acadêmico.

Fonte: HOFFMANN, NUNES e MILLER (2019).

Para dar apoio aos estudantes, a partir de 2010 houve uma mudança no perfil da Educação Superior brasileira, com a criação das políticas públicas que dão suporte aos discentes, como o Sistema de Seleção Unificada – SISU, Programa Universidade para Todos – PROUNI, e o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais – Reuni, ampliando o acesso e visando a diminuição da evasão e aumentando a inclusão social nas instituições.

O SISU é uma plataforma criada pelo Ministério da Educação brasileiro para os alunos que fizeram o Exame Nacional do Ensino Médio (ENEM) como uma forma de ingresso em uma Instituição de educação superior, a nota do ENEM substitui o vestibular, facilitando aos alunos o acompanhamento das notas de cortes durante o processo de inscrição.

O PROUNI tem por objetivo fornecer bolsas de estudos para alunos também por meio da nota do ENEM nas faculdades públicas e privadas, as bolsas podem ser integrais ou parciais para os cursos de graduação e sequenciais de formação específica.

O REUNI é um programa do Governo Federal do Brasil que visa a ampliação das vagas no ensino superior, para melhor aproveitamento das estruturas físicas e dos recursos humanos nas Universidades Federais.

Contudo, trouxeram respostas esperadas para o problema da evasão, visto que esses programas garantem o acesso dos alunos ao ingresso na faculdade, mas devem ser levadas em considerações as questões que envolvem a permanência deles nas IES.

Com a adoção do SISU, houve mudanças relevantes no acesso à Educação Superior, entre elas, o aumento significativo na mobilidade de alunos

ingressantes, o que elevou a taxa de evasão nas instituições analisadas (SZERMAN, 2015 e GOMEZ, 2014).

Andrade (2014) realizou uma pesquisa em cursos recém-ofertados e os resultados apresentam que, após três anos de implementação dos cursos, já representavam uma porcentagem de 30% de evasão dos discentes. Nesse estudo, ele analisou sobre a expansão da educação superior e aborda que com o aumento das vagas e com a variedade de cursos, há também possibilidades para o estudante experimentar as diferentes situações na construção da trajetória acadêmica, como a migração de curso, contribuindo para os números de evasão.

Se antes o estudante ingressava em uma IES e não vislumbrava mudar para outra, hoje ele pode se inscrever para qualquer instituição do Brasil (desde que haja vagas ofertadas) através do SisU, para as públicas, ou do Prouni, para as comunitárias ou particulares, caso em que já ingressa com bolsa de estudos (LIMA e Zago, 2018).

Com todas as definições e possíveis causas para a evasão, há também estudos que demonstram que a taxa de evasão no primeiro ano do curso é de duas a três vezes, maior que a dos anos seguintes.

Segundo a Associação Nacional dos dirigentes das Instituições Federais de Ensino Superior (ANDIFES), no Brasil isso acontece após o (SISU). Consta que o principal motivo tem sido a possibilidade de o aluno escolher aleatoriamente pela média obtida (ANDIFES, 2013).

Apesar dessas informações, não se sabe o motivo principal para os índices de evasão ser tão elevados nas instituições de ensino superior brasileiras, pois são questões peculiares de cada instituição que contribuem para esse fenômeno. Contudo, se sabe que a não conclusão do curso representa uma perda social, de recursos e de tempo de todas as pessoas envolvidas no processo de ensino e aprendizagem:

O abandono do aluno sem a finalização dos seus estudos apresenta uma perda social, de recursos e de tempos de todos os envolvidos no processo de ensino, pois perdeu o aluno, seus professores, a instituição de ensino, o sistema de educação e toda a sociedade (LOBO, 2012).

Lima e Zago (2018) relatam que as causas associadas à evasão não possibilitam análises isoladas, pois cada razão está ligada a um contexto institucional e pessoal, como figurado ao exemplo do tangram.

Contudo, é inegável a necessidade de se estudar o tema nas instituições de ensino, para que possam procurar medidas de intervenção, visto que a evasão trás perdas irreparáveis para os evadidos e para as instituições. São questões que

envolvem financeiro, recursos que deixarão de ser recebidos e desfrutados e mais, sonhos e possíveis mudanças de vidas que são abandonados, é sempre desvantagem para todos envolvidos no processo de ensino e aprendizagem:

Seja a evasão definida, temporária ou por mobilidade, trata-se de um fenômeno que tem consequências econômicas e sociais, sem contar as implicações pessoais para aqueles que interrompem seus cursos. Os reflexos negativos da evasão na IES estão presentes na ociosidade de vagas que deixam de ser preenchidas, no desuso das estruturas físicas e dos equipamentos, na ociosidade dos professores e dos funcionários, tudo isso ocasiona muitas vezes a redução do quadro funcional e outras tantas perdas para as IES. Para o estudante, a evasão pode representar o atraso ou a desistência de um sonho, de uma carreira profissional, de um crescimento pessoal ou a redução de chances de uma melhoria na renda (TONTINI e WALTER, 2014).

5. O PROCESSO DE DESCOBERTA DE CONHECIMENTO EM BANCO DE DADOS – KDD.

Com o avanço tecnológico em diversas áreas, o processo de Knowledge Discovery in Databases (Descoberta de Conhecimento em Banco de Dados) – KDD surgiu pela necessidade de se trabalhar com dados brutos de sistemas de diversos órgãos de distintas áreas, inclusive no descobrimento de padrões de evasão em instituições de ensino.

Para melhor discorrer sobre o processo KDD, Fayyad (1996) define como “um processo, de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grande conjunto de dados.” A Descoberta de Conhecimento em Banco de Dados pode auxiliar de forma significativa no estudo da evasão, pois, ao analisar os dados de forma precisa e automática, se obtém padrões de alunos que tendem a evadir e dependendo do algoritmo, até fazer análises preditivas por meio de suas características relevantes.

O processo em si possui duas características relevantes: é interativo e iterativo. Interativo, pois o usuário pode intervir e controlar o curso das atividades. Iterativo, por ser uma sequência finita de operações onde o resultado de cada uma é dependente dos resultados das que a precedem (PRASS, 2016).

Dentre as várias etapas sequenciais do processo de KDD, as principais são: Seleção, Pré-Processamento, Transformação, Mineração de Dados (sendo essa a principal) e Avaliação. Alguns autores definem o KDD como todo o processo de descoberta de conhecimento, e o processo de Mineração de Dados como uma

das etapas do KDD, já outros, defendem os dois termos como sendo sinônimos da inteligência artificial.

A etapa de Seleção é a escolha da massa de dados dentre os dados brutos selecionados para se trabalhar. Podendo ser um conjunto ou subconjunto de dados onde se realiza a extração, no qual são selecionados apenas os atributos interessantes ao objetivo do usuário.

O Pré-Processamento é a forma que se escolhe para organizar os dados para o tratamento deles. Transformação é a etapa em que o dado é transformado em algo em que os algoritmos de mineração de dados consigam entender, ou seja, são formas apropriadas para a mineração, ocorre de forma com que com que o algoritmo possa minerar na etapa seguinte.

O processo de Mineração de Dados é a fase mais importante de todo o processo (será abordada na próxima **seção**). Essa etapa é acontece a construção de modelos com as informações úteis, com a geração de agrupamento, os elementos que são semelhantes chegam a algum padrão de acordo critérios do usuário.

A última fase do processo é a de Avaliação, ocorre ao analisar o modelo gerado pela etapa anterior, se realiza a etapa de avaliação dos dados para obtenção de conhecimentos (informação). Esse é o último passo do processo, no qual se interpreta cada um dos padrões gerados e chega-se aos resultados.

5.1 A Principal etapa do processo de KDD: Data Mining

Com o excesso de dados disponíveis e com a falta de utilização deles, surgiu a necessidade de transformar (minerar) esses dados em informação útil. Para atingir o objetivo de se obter conhecimento por meio dos dados, desenvolveram então as ferramentas de Data Mining – Mineração de Dados. Como definição, DANTAS et al., (2019) aborda:

Data Mining ou Mineração de Dados é o processo de pesquisa em grandes quantidades de dados para extração de conhecimento utilizando técnicas de Inteligência Computacional para procurar relações de similaridade ou discordância entre dados, com objetivo de encontrar padrões, irregularidades e regras, com o intuito de transformar dados, aparentemente ocultos, em informações relevantes para a tomada de decisão e/ou avaliação de resultados.

O processo de Mineração de Dados auxilia na tomada de decisões úteis e estratégicas para melhoria de qualquer órgão em que se aplica o estudo, pois ao

realizar a descoberta de padrões, é também possível obter previsões e realizar análises traduzidas em práticas de melhorias. Esse processo busca uma grande quantidade de dados que são informações desconhecidas.

Todo processo é praticamente inviável de se realizar manualmente, pois é feito na base de dados armazenados eletronicamente nos bancos de dados, sendo algumas informações ocultas em que só o algoritmo consegue extrair de forma útil. Além disso, se realizado de forma manual, seria um processo passível de erros e extremamente custoso de tempo.

Os tipos de padrões que podem ser minerados são: Regras de Associações, Classificação e Predição, Análise de Clusters, Análise de Outliers e Sequências (ou Análise Evolutiva) podendo o banco de dados ser relacional, transacional, orientado-objeto, relacional-estendido e *data warehouse*.

As técnicas a serem utilizadas na Mineração de Dados podem ser: Técnicas de aprendizado de máquina (Machine Learning), estatística, redes neurais, algoritmos genéticos e outros.

5.2 Regras de Associação: Apriori

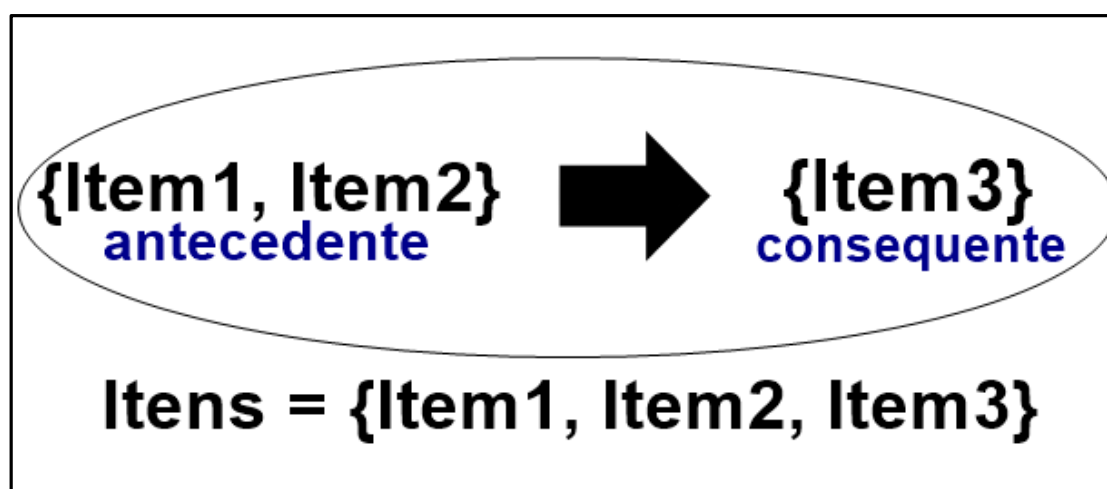
No processo de Mineração de Dados, a técnica de Regras de Associação é bastante utilizada, ela serve para encontrar padrões entre os itens de uma base de dados por meio da extração de regras geradas pelo algoritmo, no qual os resultados são apresentados de acordo com a co-ocorrência e não casualidade.

Por meio das regras geradas pode-se obter resultados que definem padrões e/ou perfis de acordo com as informações da base de dados, esses resultados auxiliam na tomada de decisões de determinadas organizações, pois ao definir esses padrões, facilita nas estratégias dos objetivos que se deseja alcançar.

A aplicação do algoritmo para regras de associação pode ser utilizada em diversas organizações como mercados, lojas físicas e online, planos de vendas, combate à evasão escolar (estudo deste trabalho) e afins.

As regras são geradas e interpretadas pelos itens antecedentes e consequentes, como no exemplo da figura abaixo, *itens* é a lista de todos os itens no antecedente e no consequente:

Figura 3 – Exemplo de Regra de Associação



Fonte: O autor

O Item1 e o Item2 como antecedente e o item3 como consequente, ou seja, se ocorreu o item1 e item 2, ocorrerá o item3. Para ajudar a entender a força da associação, existem métricas a serem definidas, sendo elas: Suporte, Confiança e Lift:

- Suporte: é a frequência em que o conjunto de itens aparece nas transações, servindo para ajudar a identificar as regras válidas;
- Confiança: é a probabilidade em que os itens ocorrem, é encontrado de acordo com a quantidade dos itens que aparecem em conjunto, divididos pelo número total de transações;
- Lift: é a probabilidade dos itens aparecerem juntos e pode ser calculado dividindo a Confiança pelo Suporte.

Para lidar com grandes conjuntos de dados o algoritmo apriori extrai regras para combinações possíveis de itens, sendo um dos mais famosos existentes para mineração, podendo ser implementado em diversas linguagens e ferramentas.

O algoritmo faz uso de duas funções: a função Apriori_gen, para gerar os candidatos a eliminar aqueles que não são frequentes, e a função Genrules, utilizada para extrair as regras de associação (ROMÃO et al, 2019). Apriori foi proposto por Agrawal em 1994, sendo um modelo que assume dados não categóricos, ou seja, não é aplicado a dados numéricos, tendo o princípio de que se um itemset é frequente, todos os seus subconjuntos também são frequentes, isso acontece devido à propriedade do suporte calculado pelo algoritmo.

O Apriori gera um conjunto de regras de acordo com o suporte e confiança, sua estrutura é baseada em antecedente e consequente, sendo o conjunto A o Antecedente e o conjunto B o Consequente, basicamente apresentado da seguinte forma: **A => B**.

Os valores de Suporte Mínimo (min_sup) e Confiança Mínima (min_conf) são parâmetros importantes a serem definidos ao algoritmo para que as regras possam ser reduzidas e extraídas como proposto nas definições paramétricas. A definição delas ajuda na extração e análise das regras, pois quanto maior a confiança, maior a confiabilidade da informação gerada.

5.3 Linguagem Python

Python é uma linguagem de programação criada por Guido van Rossum, simples utilizada em diversas aplicações, desde projetos simples aos grandes. A linguagem surgiu em 1991 e além de ser uma linguagem de alto nível e de tipagem dinâmica e forte, é multiparadigma: Orientada a objetos, Programação imperativa e Programação funcional. A linguagem foi criada com o objetivo de ser produtiva e legível.

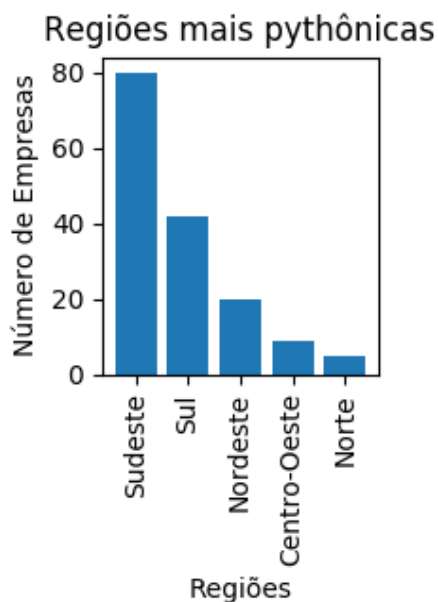
Algumas características relevantes na linguagem devem ser levadas em consideração, como apresenta em sua documentação:

- Utiliza uma sintaxe elegante, facilitando a leitura dos programas;
- É uma linguagem fácil de usar que facilita o funcionamento do programa. Isso torna o Python ideal para o desenvolvimento de protótipos e outras tarefas de programação ad-hoc, sem comprometer a manutenção;
- Vem com uma grande biblioteca padrão que suporta muitas tarefas comuns de programação, como conectar-se a servidores da Web, pesquisar texto com expressões regulares, ler e modificar arquivos;
- O modo interativo do Python facilita o teste de pequenos trechos de código. Há também um ambiente de desenvolvimento em pacote chamado IDLE;
- É facilmente estendido adicionando novos módulos implementados em uma linguagem compilada, como C ou C ++;
- Também pode ser incorporado a um aplicativo para fornecer uma interface programável;
- É executado em qualquer lugar, incluindo Mac OS X, Windows, Linux e Unix, com versões não oficiais também disponíveis para Android e iOS (PYTHON, 2017).

No Brasil, o Python é utilizado por inúmeras empresas de pequeno e grande porte que aderiram a aderiram pelas suas especificações que o difere de muitas outras linguagens, no Site Python Brasil, dados em ranking foram divulgados

por região, no qual se percebe que a Região Sudeste do Brasil apresenta o maior número de utilização da linguagem nas suas empresas:

Figura 4 – Regiões do Brasil que mais utilizam a linguagem Python



Fonte: Python Brasil (2019)

O Python está na sua versão 3.7.4, de 08 de junho de 2019, e é uma linguagem poderosa, comparável a Perl, Ruby, Scheme ou Java. Possui uma comunidade extensa e ativa, por ser uma linguagem livre, os usuários contribuem voluntariamente pelo mundo todo.

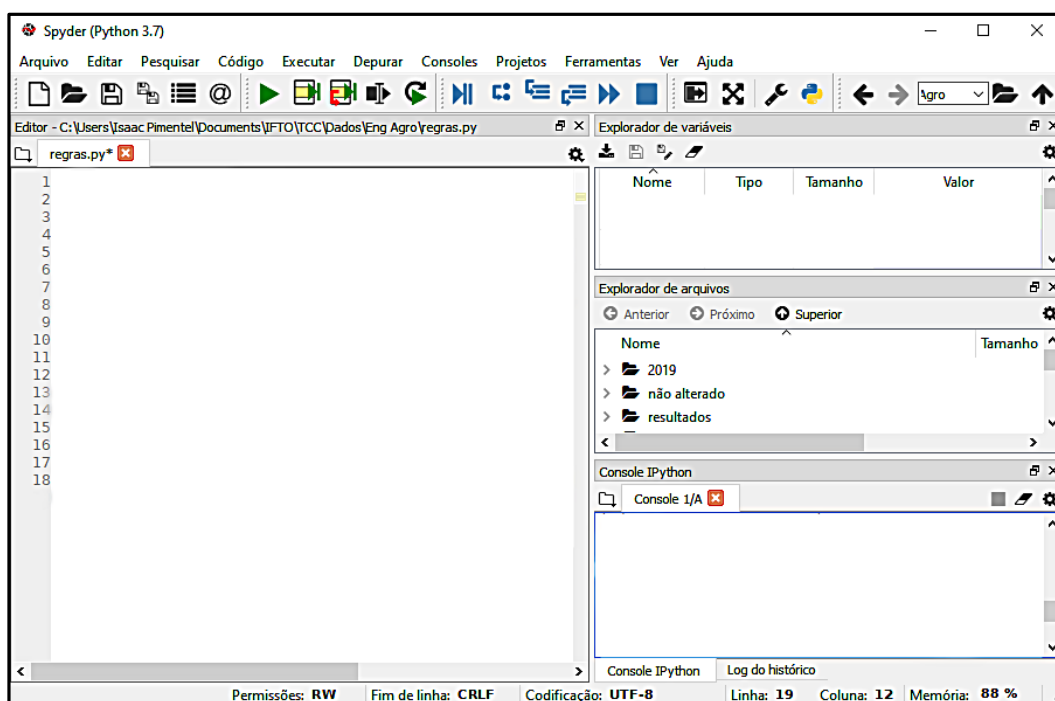
5.3.1 A Ferramenta Spyder

Spyder – The Scientific Python Development Environment é um ambiente gráfico “que permite a utilização do Python num ambiente interativo, facilitando a edição de *scripts*, teste, *debugging* e visualização gráfica (CUSTÓRIO, 2017). A ferramenta possui diversas interações que visam a facilidade durante a utilização, com:

- Editor, que ajuda durante toda a escrita do código;
- Console Python que apresenta os resultados no próprio ambiente;
- Explorador de variáveis, no qual possibilita a modificação e interação das variáveis em tempo real;
- Perfil, onde é possível encontrar e eliminar gargalos para liberar o desempenho do código;

- Depurador, servindo para rastrear cada etapa da execução do código;
 - Socorro (ajuda), que apresenta a visualização instantânea dos documentos de qualquer objeto;
 - E os Plugins, que é toda a funcionalidade na interface,
- como na figura abaixo:

Figura 5 – Interface da ferramenta Spyder (Python 3.7)



Fonte: O autor

5.4 Linguagem R

Sob influência da linguagem S, o ambiente computacional e linguagem de programação R surgiu em 1993, por Ross Ihaka e Robert Gentleman. Trata-se de uma linguagem multi-paradigma, sequencial, orientada a objeto, imperativo, dinâmico, processual e recíproco. É conhecida como um produto estatístico especializado justamente por conter fins estatísticos para manipulação de dados de tamanho médio.

Com o passar dos anos, o R foi ganhando popularidade principalmente entre os analistas de dados e estatísticos para análise exploratória de dados interativos, é um software livre que pode ser compilado e rodado na ampla variedade das plataformas: Mac, Windows e Unix/Linux, sob a licença General Public License – GNU, a sua última versão é a 3.5.0, de 23 de abril de 2018.

O R possui diversas vantagens e destaques, desde a parte técnica à parte gráfica, como definido em sua documentação oficial:

Um dos pontos fortes de R é a facilidade com que gráficos de qualidade de publicação bem projetados podem ser produzidos, incluindo símbolos de fórmulas matemáticas quando necessário. Houve um grande cuidado com os padrões das pequenas opções do design em gráficos, mas o usuário mantém o controle total (R, 2019).

Além disso, o R ainda apresenta uma variedade de técnicas para o estudo e análise da estatística por meio das bibliotecas (pacotes) para as diversas áreas de conhecimento relacionado a estatística e suas áreas, esses pacotes são disponíveis de forma gratuita. Mesmo com a sua similaridade com o S, a implementação é distinta e com relevantes diferenças, mas muitos códigos em S podem ser executados no R, isso pode ocorrer de forma contrária também.

O R fornece uma ampla variedade de técnicas estatísticas (modelagem linear e não linear, testes estatísticos clássicos, análise de séries temporais, classificação, agrupamento, ...) e técnicas gráficas e é altamente extensível. A linguagem S é frequentemente o veículo de escolha para pesquisa em metodologia estatística, e R fornece uma rota de código aberto para a participação nessa atividade (R,2019).

Segundo a sua documentação, “o termo “ambiente” pretende caracterizá-lo como um sistema totalmente planejado e coerente, em vez de um crescimento incremental de ferramentas muito específicas e flexíveis” (R, 2019). A linguagem/ambiente possui um conjunto de instalações de software para a manipulação de dados, cálculo e exibição gráfica, como:

- Instalação eficaz de manipulação e armazenamento de dados;
- Conjunto de operadores para cálculos e matrizes, em particular matrizes;
- Coleção grande, coerente e integrada de ferramentas intermediárias para análise de dados;
- Instalações gráficas para análise de dados e exibição na tela ou em cópia impressa e;
- Uma linguagem de programação bem desenvolvida, simples e eficaz, que inclui condicionais, loops, funções recursivas definidas pelo usuário e recursos de entrada e saída (R, 2019).

Os criadores do sistema defendem o R como um ambiente em que as técnicas de estatísticas são implementadas, e não como um sistema propriamente estatístico, como definido na linguagem:

Muitos usuários pensam no R como um sistema estatístico. Preferimos pensar nele como um ambiente no qual as técnicas estatísticas são implementadas. R pode ser estendido (facilmente) através de pacotes. Existem cerca de oito pacotes fornecidos com a distribuição R e muitos outros estão disponíveis através da família de sites da CREN na internet, cobrindo uma ampla variedade de estatísticas modernas (R.2019).

5.4.1 Rstudio

O RStudio é um ambiente desenvolvido para R. É software livre com interface interativa que visa a facilidade em seus blocos fundamentais: interface, importação de dados e comandos, os objetivos da ferramenta é capacitar os usuários a serem produtivos com o R:

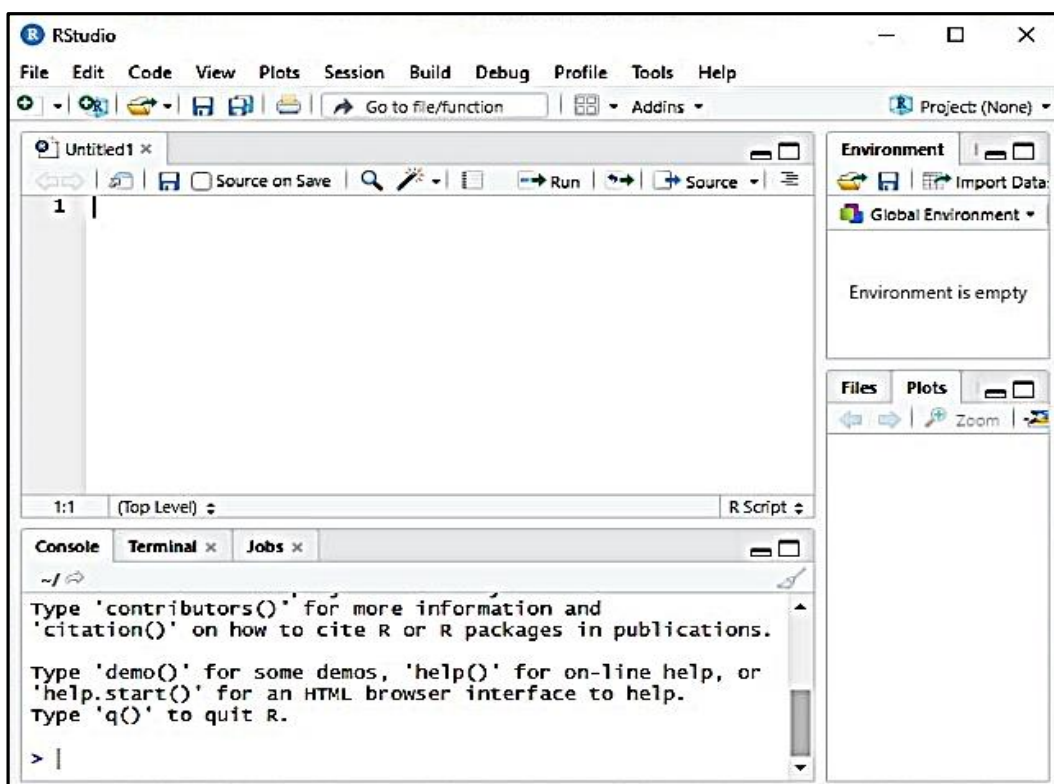
Inspirado pelas inovações dos usuários de R na ciência, educação e indústria, o RStudio desenvolve ferramentas gratuitas e abertas para produtos profissionais prontos para empresas, para equipes escalarem e compartilharem o trabalho (RStudio, 2019).

Além de ser um software de código aberto, disponível para Windows, Mac OS e Linux, o RStudio se preocupa com a interface e interação para o usuário, é de fácil acesso e entendimento das funções fornecidas:

O RStudio tem a missão de fornecer o software profissional de código aberto e profissional mais amplamente usado para o ambiente de computação estatística R. Essas ferramentas promovem a causa de equipar todos, independentemente de meio, a participar de uma economia global que recompensa cada vez mais a alfabetização de todos (RStudio, 2019).

A figura abaixo apresenta a interface do software e suas divisões, contendo em cada parte uma função diferente:

Figura 6 – Interface do RStudio



Fonte: O autor

A interface conta com quatro painéis de trabalho, o painel superior esquerdo o local é onde se cria e abre um arquivo contendo o código R. O painel

interior esquerdo é onde se localiza o console, destinado para digitar os comandos em R.

Outro painel disponível é o superior direito, no qual contém a Guia Área de Trabalho e o Guia Histórico, o primeiro, mostra a lista que o R criou durante a execução R e o segundo apresenta todo o histórico dos comandos utilizados. O painel inferior direito possui o Guia Arquivos (mostra os arquivos do diretório), Guia Gráficos (mostra todos os gráficos que foram utilizados na seção) e Guia Pacotes (exibe os pacotes externos do R disponíveis no sistema).

O software ainda possui seus principais produtos profissionais, que atendem os trabalhos de grande escala:

O software profissional RStudio de código aberto e pronto para a empresa combina análise de dados robusta e reproduzível com ferramentas para compartilhar efetivamente produtos de dados. Nossos principais produtos profissionais, RStudio, Server Pro, RStudio Connect e RStudio Package Manager, equipam profissionais de ciências de dados para desenvolver e compartilhar seu trabalho em grande escala (RStudio, 2019).

5.5 Python e R

As duas linguagens possuem particularidades que as diferem em diversos âmbitos, algumas informações podem ser analisadas no quadro abaixo:

Tabela 2 – Tabela comparativa entre as duas linguagens.

	Python	R
Data de Criação	1991	1995
Versão atual	3.7.4 (8 de julho de 2019)	3.6.4 (23 de abril de 2018)
Origem	Python foi inspirada na linguagem C, Modula-3 e ABC.	R foi uma implementação da linguagem S da Bell Labs
Usabilidade	Codificação e Debugging é mais fácil, pois a sintaxe é bem simples.	Modelos estatísticos podem ser escritos com poucas linhas de código.
Padronização	Python possui um padrão bem definido, permitindo que diferentes tipos de funcionalidades sejam escritos da mesma forma.	A mesma funcionalidade pode ser escrita de diversas formas diferentes em R.
Flexibilidade	Python é bastante flexível e permite facilmente escrever algo a partir do zero.	É fácil escrever fórmulas complexas em R. Praticamente todos os tipos de testes e modelos estatísticos estão disponíveis para o uso na linguagem.
Repositórios	PyPi é o repositório do Python, com diversas bibliotecas. Usuários podem contribuir com novos pacotes.	CRAN (Comprehensive R Archive Network) é o repositório da linguagem R em que cada usuário pode contribuir com novos pacotes

		(que são coleções de funções em R com código compilado). Esses pacotes podem ser facilmente instalados com uma linha de código.
Uso em análise de Dados	Python é principalmente usada quando a análise de dados precisa ser integrada com aplicativos web ou se o código estatístico precisa ser integrado em um servidor em ambiente de produção, que vai servir muitos usuários.	R é principalmente usada quando as atividades de análise de dados requerem computação standalone (em um único computador) ou análise em servidores individuais.
Tratamento de dados	Python não foi criada inicialmente para análise de dados, mas a linguagem tem crescido rapidamente e cada vez oferece mais recursos.	R é extremamente eficiente em análise de dados, devido seu grande número de pacotes com modelos, fórmulas e testes estatísticos.
IDE	Python IDE, Spyder e IPython Notebook	RStudio
Pacotes mais populares	<ul style="list-style-type: none"> • SciPY / NumPy (computação científica); • Pandas (manipulação de dados); • Matplotlib (gráficos); • Scikit-learn (Machine Learning). 	<ul style="list-style-type: none"> • Dplyr, plyr e data.table (manipulação de dados); • Stringr (manipulação de strings); • Zoo (time-series); • Ggvis, lattice e ggplot2 (gráficos); • Caret (Machine Learning).
Pontos positivos	<ul style="list-style-type: none"> • Python é ótima para criação de scripts e automatização de regras para mineração de dados; • Se integra facilmente em um fluxo de trabalho de produção de desenvolvimento; • Pode ser usada em diferentes partes no processo de desenvolvimento de software (back-end, front-end); • IPython Notebook é uma poderosa ferramenta para análise exploratória e apresentações. 	<ul style="list-style-type: none"> • R é ótimo para prototipagem e para análise estatística; • R tem um enorme conjunto de bibliotecas disponíveis para análise de diferentes tipos de estatística; • RStudio IDE é definitivamente uma grande vantagem. Ele facilita tarefas tediosas e facilita o fluxo de trabalho;
Pontos Negativos	<ul style="list-style-type: none"> • Não é tão complexo para análise estatística como R, mas já percorreu um longo caminho nos últimos anos; • Na minha opinião, a curva de aprendizado é mais acentuada do que R, por se tratar de uma linguagem muito mais 	<ul style="list-style-type: none"> • A sintaxe pode ser obscura, às vezes; • É mais difícil de integrar a um fluxo de desenvolvimento de produção; • Na minha opinião, é mais adequado para tarefas do tipo “consultoria”.

	completa.	
Velocidade	Python é considerada uma das linguagens de programação mais velozes atualmente.	R foi desenvolvida para fazer análise de dados, não para ser rápida. Entretanto, o código pode ser otimizado. O problema, é que a linguagem é bastante utilizada por quem não possui conhecimentos de programação, tornando o código mais lento do que deveria.

Fonte: MATOS (2018)

5.6 Weka

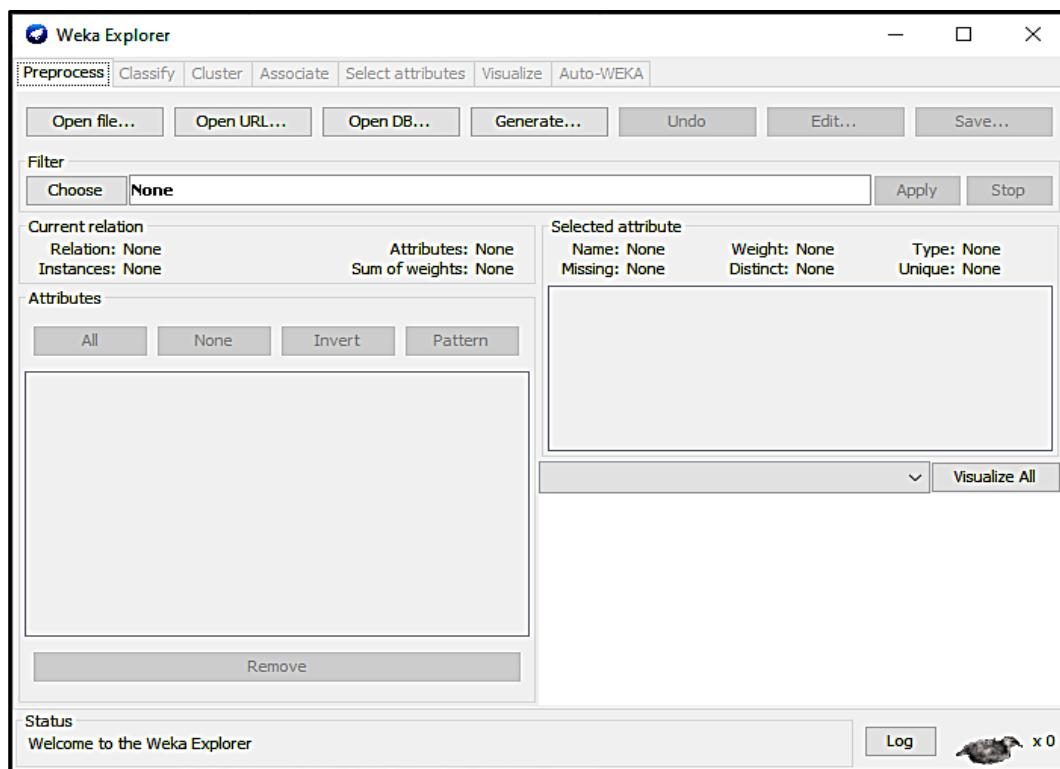
Weka (Waikato Environment for Knowledge Analysis) é uma ferramenta com pacotes para mineração de dados desenvolvida na Universidade de Waikato – Nova Zelândia. Sob licença da GPL – General Public Licença, é uma ferramenta bastante utilizada no meio acadêmico, a sua tarefa principal é a classificação, mas em suas funções há outros métodos para utilização, sobretudo as Regras de Associação. “Weka é uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Ele contém ferramentas para preparação, classificação, clustering, mineração de regras de associação e visualização de dados” (WEKA, 2019).

O seu potencial é grande por se tratar de uma de um software livre de Data Mining open source, com interface interativa e de fácil manipulação, segundo Rosa, Rodrigues e Sobrinho (2017), a ferramenta se destaca pelos principais métodos embutidos:

- Métodos de classificação;
 - Métodos para predição numérica;
 - Métodos de agrupamento;
 - Métodos de associação
- (ROSA, RODRIGUES e SOBRINHO, 2017).

Para que os dados sejam utilizados na ferramenta, devem possuir o formato ARFF – Attribute Relation File Format, no qual consistem em um arquivo de texto com algumas informações de cabeçalho para importação dos campos das informações da base de dados, onde há os domínios dos atributos. A ferramenta possui duas versões, a versão estável 3.6.11 (24 de abril de 2014) e a versão em teste 3.7.11 (24 de abril de 2014) e é um sistema de multiplataforma.

Figura 7 – Interface WEKA EXPLORER



Fonte: O autor

Na interface Explorer, na parte superior são apresentados em sessões (algoritmos): Preprocess, Classify, Cluster, Associate, Select Attributes, Visualize e Auto-WEKA. A principal seção é a de preprocess, onde é selecionado o caminho da base de dados para que nas outras sessões ocorra a mineração dos dados de acordo com a necessidade de escolha.

6. METODOLOGIA

A metodologia é um passo fundamental para melhor apresentação e interpretação dos resultados, por meio dela, se trilha os passos utilizados para o alcance do objetivo proposto, como afirma WAZLAWICK (2009, p.39):

O método consiste na sequência de passos necessários para demonstrar que o objetivo proposto foi atingido, ou seja, se os passos definidos no método forem executados, os resultados obtidos deverão ser convincentes.

6.1 Tipo de Pesquisa

Esta pesquisa estatística exploratória tem como fundamentação a observação, pois os elementos utilizados para o estudo não são afetados e consiste nas características gerais dos alunos do curso de Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis.

As variáveis utilizadas são qualitativas e quantitativas, a primeira pois os dados são de categoria nominal sem hierarquia, e a segunda, os resultados assumem quantidades contínuas ao apresentar a confiança e o lift (itens abordados no decorrer da metodologia).

As técnicas estatísticas deste estudo são descritivas, pois visam organizar, demonstrar e resumir os dados do estudo; probabilidade, pois analisa as situações geradas por meio das regras de associação; e inferência, ao visar à obtenção de respostas sobre o fenômeno Evasão de uma determinada população por meio de dados representativos.

Na população fornecida para a realização do estudo contém o registro dos alunos que foram matriculados no curso de Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis, no período de 2014 a 2019. A amostra selecionada para o estudo foram os dados de 2014 a 2018.

Os métodos da pesquisa adquirem características documental, bibliográfica e telematizada. Documental ao utilizar dados fornecidos pela Instituição; Bibliográfica no que tange à pesquisa para coleta de informações que serviram de base para a construção desse trabalho; e Telematizada, pois foram utilizadas pesquisas atreladas ao uso da internet e das telecomunicações.

6.2 Universo da Pesquisa

A população da pesquisa são os dados relacionados à matrícula de todos os acadêmicos do curso Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis no período de 2014 a 2019. Os dados foram fornecidos pela Coordenação de Registros Escolares (CORES) do *campus* em estudo, neles constam o total de 321 registros de matrículas desde a implementação do curso. A amostra utilizada foi de registros de ingresso de 2014 a 2018, reduzindo para 283 registros, o que representa 88% da população. A amostra foi selecionada pela objetividade de identificar os padrões/perfis dos alunos evadidos ingressantes do período de 2014 a 2018 por meio do algoritmo Apriori na Linguagem Python, R e na ferramenta Weka. Os dados de 2019 não foram utilizados, pois a coleta foi realizada no primeiro semestre do ano de 2019, portanto não continham os dados significativos para identificação.

6.3 Coleta de dados

Durante todo o processo de construção desta monografia, foram realizadas pesquisas em diversos meios de informações como em livros, revistas, artigos científicos, teses de mestrado, sites oficiais, documentos internos (PPC) e outros. Toda essa etapa serviu para fundamentação válida acerca da temática Evasão e Mineração de Dados, incluindo as diversas visões dos autores sobre o assunto e observando a criticidade sobre as ideias expostas.

A etapa da pesquisa documental foi realizada por meio do conjunto de dados fornecido pela CORES. Os dados foram entregues no formato Comma-separated values (CSV) com os registros dos ingressantes do curso de Agronomia do IFTO – *Campus* Dianópolis no período de 2014 a 2019. É importante salientar que todos os dados e informações obtidas nesta pesquisa tiveram tratamento estritamente sigiloso, ou seja, não obtiveram informações que permitem a identificação de qualquer participante, sendo analisados os dados de forma conjunta/agrupada. A pesquisa não possui nenhum interesse no resultado individual, ou seja, o principal interesse está relacionado no resultado coletivo.

Os seguintes dados foram solicitados para realização da pesquisa:

- Raça;
- Sexo;
- Renda;
- Ano de Nascimento (Faixa Etária);
- Percentual de Conclusão;
- Status (Evadido, em Curso, Concluído).

Alguns dados sofreram tratamento para melhor interpretação durante a análise das regras, como Ano de Nascimento que foi convertido para Faixa Etária, agrupados em: jovem, adulto e idoso. Considerado “Jovem” os que possuem idade entre 15 aos 29 anos, “Adulto” entre 30 e 59 anos e “Idoso” os que possuem idade igual ou superior a 60 anos.

Os dados de Percentual de Conclusão obtiveram duas abordagens distintas, uma para os “Evadidos” no qual foram definidos três quartis com intervalo de vinte, como: primeiro quartil, menos de 20%, outra abordagem para os “Em curso” que foram divididos em cinco quartis, com intervalos de vinte, exceto o primeiro que teve intervalo de 10%.

O conceito de evasão utilizado nesta pesquisa foi entendido como todos os aspectos de desistência do curso, como: trancamento, desligamento, evadido e transferência para outra IES. Com isso, os registros de matrículas fornecidos pela CORES que continham essas informações foram convertidos para “Evadido”.

A utilização dos dados coletados se deu pela intenção de identificar os padrões/perfis dos alunos evadidos no curso de Engenharia Agrônômica do *campus* Dianópolis desde o primeiro ano de oferta (2014) até os últimos registros completos (2018).

Além da pesquisa bibliográfica e documental, a pesquisa telematizada obteve um papel indispensável para construção das várias etapas do trabalho, tanto na teoria quanto na técnica/prática.

6.3.1 Tratamento dos Dados

Além dos tratamentos citados na **seção** anterior, os dados foram reorganizados para posteriormente serem classificados no algoritmo Apriori na Linguagem Python e na Linguagem R. Para a ferramenta WEKA os dados foram convertidos para formato ARFF - Attribute Relation File Format, obedecendo às etapas de todo o processo de Mineração de Dados:

6.3.2 Seleção:

A escolha se deu dentre os dados brutos fornecidos, foi realizada uma seleção e escolha dos registros de 2014 a 2018.

6.3.3 Pre-Processamento:

A etapa de pré-processamento foi a organização para realização do tratamento. Aconteceu durante a separação na tabela e formatação de alguns dados inválidos e não preenchidos entre os registros como em alguns casos, no campo renda não havia registro, foi identificado como “Não Definido”.

6.3.4 Transformação:

Esta parte do processo aconteceu ao transformar as palavras para serem lidas pelo algoritmo, removendo os espaços e os caracteres especiais, apropriando-os para a mineração.

6.3.5 Mineração de Dados:

Nesta etapa foram geradas as regras de associação nas linguagens Python e R, e na ferramenta WEKA.

6.4 Regras de Associação em Python:

Para a mineração no Python, foi utilizada a ferramenta computacional Anaconda, para ambientes de trabalho e módulos do Python. Anaconda contém mais de 720 pacotes open-source e corre nos sistemas operativos Windows, Mac OSX e Linux (CUSTÓDIO, 2017).

Para gerar as regras em um ambiente gráfico interativo, foi utilizado o ambiente científico Spyder – The Scientific Python Development Environment com a linguagem Python, foi utilizado o algoritmo apriori, importado em forma de biblioteca apyori, consiste em gerar padrões associativos de acordo com os parâmetros definidos no script e identificar relações subjacentes entre itens diferentes.

Os parâmetros são importantes para as regras que são geradas de acordo o Suporte, Confiança e Lift. O suporte dá uma ideia do quão frequente é um conjunto de itens em todas as transações, o seu valor ajuda a identificar as regras que valerão a pena considerar. A confiança define a probabilidade de ocorrência e o Lift indica que quanto mais frequente se torna um item, o outro item ocorre.

Os parâmetros utilizados no algoritmo foram os seguintes: Suporte ($\text{min_support} = 0.3$) e Confiança ($\text{min_confidence} = 0.8$). Esses parâmetros foram escolhidos para facilitar a análise e precisão das regras. Além desses parâmetros, o próprio algoritmo define o lift. Para utilização de todos os dados, foi definido na estrutura do código o $\text{range}(0,6)$ considerando todas as colunas da tabela com os dados para a mineração.

Após a execução do algoritmo apriori no Python, foram geradas 136 regras de associação. Para análise, foram consideradas apenas as regras que possuíam como conseqüente “Evadido”, sendo o total de 27 regras selecionadas, analisadas e expostas na **seção** resultados. As regras selecionadas foram separadas em três grupos com base no sexo (Masculino Feminino e Indefinido). O item “Indefinido” foi considerado para os indivíduos cujas regras não definiram o sexo específico, ou o campo vazio. A figura abaixo apresenta o código utilizado no Spyder (Python) para extração das regras na linguagem:

Figura 8 – Código para extração das regras em Python

```

1 import pandas as pd
2 from apyori import apriori
3
4 dados = pd.read_csv('dadostcc.csv', header = None)
5 transacoes = []
6 for i in range(0,283):
7     transacoes.append([str(dados.values[i, j]) for j in range(0, 6)])
8
9
10 regras = apriori(transacoes, min_suporte = 0.3, min_confidence = 0.8)
11
12 resultados = list(regras)
13 resultados

```

Fonte: O autor

Na linha 1 é importada a biblioteca Pandas que fornece ferramentas para análise de dados, na linha 2 é feita a importação da biblioteca Apriori. Na linha 4 é apresentado o formato dos dados no formato CSV que o algoritmo vai receber e o caminho dele. Na Linha 5 é criada uma variável em formato de lista e na linha seguinte é feito um laço de repetição para percorrer todas as linhas do registro. Na linha 7 é adicionada uma função para transformar os dados em String e um laço para adicionar dentro da variável transações, o valor dentro do range é o tanto de colunas a serem consideradas.

A linha 10 recebe a função apriori e a variável transações seguido com os parâmetros de suporte mínimo e confiança mínima para as regras a serem extraídas. Na linha 12 a variável recebe a lista e na linha seguinte é para impressão das regras no console do Spyder.

Para definir o padrão dos alunos evadidos gerados pelas regras na linguagem Python foi utilizada a média que é uma medida estatística que representa um conjunto com informações de determinada base de dados, usada para medir a confiança dos resultados.

6.5 Regras de Associação em R:

A segunda etapa do processo foi gerar as regras de associação na linguagem R, no ambiente RStudio por meio do algoritmo Apriori, importado pelo pacote Arules, no qual oferece um ambiente para exploração de técnicas para regras de associação, apresentados e descritos abaixo:

Figura 9 – Código para extração das regras em R

```

1  install.packages("arules")
2  library(arules)
3
4
5
6  transacoes = read.transactions(file.choose(), format="basket", sep=",")
7  transacoes
8  inspect(transacoes)
9
10
11 regras = apriori(transacoes, parameter=list(supp=0.03, conf=0.8, minlen=6))
12 regras
13 inspect(regras)

```

Fonte: O autor

Na linha 1 é feita a instalação do pacote “arules” responsável por executar o algoritmo apriori, na linha 2 é feito o carregamento da biblioteca do projeto. Na linha 6 a variável transações recebe os dados, o parâmetro read.transactions faz a leitura do arquivo, seguido pelo file.choose() que recebe o caminho do arquivo, o formato “basket” faz a leitura no formato CSV.

Na linha 7 a variável é executada para verificação do carregamento do arquivo. A linha 8 mostra todos os dados do arquivo, linha por linha. Na linha 11 a variável “regras” executa o algoritmo apriori, “transações” é a base de dados que será analisado, seguido pela definição dos parâmetros de suporte, confiança e o número de colunas a serem consideradas.

A linha 12 apresenta o número de regras geradas (mas ainda não mostra as regras), seguindo pela linha 13 o inspect(regras) mostra todas as regras geradas de acordo com os parâmetros definidos.

Como na linguagem Python, no R também foram definidos os mesmos parâmetros para Suporte (supp=0.03) e Confiança (conf=0.8), e definido na estrutura do código o minlen=6 para a utilização de todos os campos da base de dados selecionada, gerando o total de 11 regras em torno dos parâmetros definidos, mas apenas uma com o consequente Evadido, que será apresentado na **seção** de Resultados.

6.6 Regras de Associação no WEKA:

A terceira etapa foi a mineração dos dados para a regras de associação na ferramenta WEKA, no qual foi utilizado também o algoritmo apriori. No Weka Explorer os dados foram carregados no formato ARFF, o processo para gerar as regras foi na seção Associate no qual foi escolhido o algoritmo apriori.

Para gerar as regras de associação, foi utilizado o parâmetro `metricType=Confidence`, que define o padrão confiança para gerar as regras, e a confiança foi a mesma dos scripts anteriores (`minMetric=0.8`). Na ferramenta é possível definir o número de regras a serem geradas para apresentar o máximo possível, foi definido o total de 283 regras, esse número é total de linhas de informações da base de dados, porém, de acordo com o número de confiança definido, a ferramenta apresentou 40 regras.

Com o conseqüente Evadido, o WEKA gerou sete regras que serão expostas e analisadas na **seção** seguinte.

7. RESULTADOS

Os resultados obtidos na extração das regras nas diferentes linguagens e ferramenta são apresentados por subseções abaixo.

7.1 Padrões gerados em Python

Na análise dos dados coletados e nas regras geradas pelo Apriori na linguagem Python, alguns critérios foram levados em consideração, nas 26 regras selecionadas os resultados obtidos foram: o campo “Raça” obteve resultado “Parda” em todos os casos. Já o campo “Sexo” foi dividido nas três categorias: Masculino, Feminino e Indefinido. A “Renda” obteve “Não Declarado” em todos os seus campos, a “Faixa Etária” foi “Jovem”, e o “Percentual de Conclusão” em todos os casos foram “Menos de 20%”.

Dessa forma, foi possível esquematizar os dados para definir o perfil dos alunos do curso de Engenharia Agrônômica do *campus* Dianópolis, com base na recorrência de cada característica, como apresentado nas tabelas abaixo:

Tabela 3 – Padrão 1 (Python)

ANTECEDENTE					=	CONSEQUENTE	
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)	Status	Confiança	
Padrão 1							
Parda	Feminino	Não declarada	Jovem	Menos de 20	= Evadido	68%	

Fonte: O autor

A tabela acima foi definida pelo Sexo Feminino, em que ao todo obtiveram quatro regras de associação para chegar ao Padrão 1. A média foi calculada de acordo a confiança de 4 regras anteriores, obtendo o resultado de 68%.

Tabela 4 – Padrão 2 (Python)

ANTECEDENTE					CONSEQUENTE		
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)	=	Status	Confiança
Padrão 2							
Parda	Masculino	Não declarado	Jovem	Menos de 20	=	Evadido	68%

Fonte: O autor

Na tabela 2, pode se observar o padrão do Sexo Masculino definido de acordo com as 12 regras de associação, semelhantemente ao Padrão 1, todas as características se repetem, exceto o Sexo. A média de confiança desse padrão também foi de 68%.

Tabela 5 – Padrão 3 (Python)

ANTECEDENTE					CONSEQUENTE		
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)	=	Status	confiança
Padrão 3							
Parda	Indefinido	Não declarada	Jovem	Menos de 20	=	Evadido	82%

Fonte: O autor

O terceiro e último padrão foi definido por meio das 11 regras geradas. A diferença é que o algoritmo apriori não definiu o sexo nessa categoria, concluindo que esse campo poderia assumir duas variáveis: Masculino e Feminino. A confiança do Padrão 3 obteve maior porcentagem, sendo ela 82% de confiança, também definido pela média das regras selecionadas.

7.2 Padrão gerado em R

Nas 11 regras geradas pelo algoritmo Apriori na Linguagem R, apenas uma apresentou o consequente como “Evadido, com isso, obtém se apenas um padrão de alunos evadidos do curso de Bacharelado em Engenharia Agrônômica, *campus* Dianópolis, como apresentado na tabela abaixo:

Tabela 6 – Padrão (R)

ANTECEDENTE					CONSEQUENTE		
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)	=	Status	confiança
Padrão							
Parda	Masculino	Não declarada	Jovem	Menos de 20	=	Evadido	100%

Fonte: O autor

No padrão único gerado em R, pode-se observar que os alunos de cor “Parda”, do sexo “Masculino”, “Jovem” e com percentual de conclusão “Menos de 20%) se encontram com o Status de “Evadido”, a confiança para essa regra é de 100% gerada pela própria regra.

7.3 Padrão gerado no WEKA

Na ferramenta WEKA das 40 regras geradas, apenas 7 constavam resultados de interesse para esta pesquisa, por apresentar como consequentes “Evadido”. O campo “Cor” obteve em todas as regras selecionadas o resultado “Parda”, o “Sexo” obteve resultado “Indefinido”, a “Renda” apresentou “Não Declarado” em todas as regras e a “Faixa Etária” foi “Jovem”, o “Percentual de Conclusão” apresentou em todos os casos “Menos de 20%”, dessa forma, foi definido a média entre os resultados para apresentar a confiança, como exposto na tabela abaixo:

Tabela 7 – Padrão (WEKA)

ANTECEDENTE					=	CONSEQUENTE	
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)		Status	confiança
Padrão							
Parda	Indefinido	Não declarada	Jovem	Menos de 20	=	Evadido	89%

Fonte: O autor

A média foi calculada de acordo a confiança das 7 regras com consequente evadido, obtendo o resultado de 89%.

8. DISCUSSÕES

Após as pesquisas realizadas e com o processo de mineração de dados adotado para este estudo, foi possível identificar o padrão e/ou perfis dos alunos que tendem a evadir no curso de Bacharelado em Engenharia Agrônômica do IFTO, *campus* Dianópolis. Esses resultados foram obtidos por meio de regras de associação, uma importante tarefa realizada nas linguagens Python e R e na ferramenta Weka, todos com o algoritmo apriori.

Os padrões (2 e 3) gerados em Python evidenciam os mesmos padrões encontrados pela linguagem R e ferramenta Weka, o que as difere é o valor da

confiança, os mesmos parâmetros de configuração foram aplicados ao R e Weka e mesmo assim a linguagem e a ferramenta encontraram apenas um padrão cada.

Vale salientar que a proposta desta pesquisa não é definir qual é a melhor ferramenta entre as três estudadas, e sim discutir qual se aplicou melhor ao conjunto de dados analisados. Sendo assim, o Python se mostrou mais versátil para a análise desta base de dados, pois pôde localizar três padrões de alunos evasores sendo que dois destes padrões foram confirmados por outras ferramentas.

Os padrões gerados apresentam que os jovens do curso em estudo, independem do sexo para evadirem, mesmo o curso tendo um público maior do sexo masculino, tanto homens quanto mulheres possuem a mesma chance de evadirem, com isso, conclui-se que a evasão não é um fator que tem como principal fator o sexo.

Muitos estudos apresentam maiores taxas de evasão no primeiro ano do curso, as regras deste estudo também apontam que os alunos que possuem menos de 20% de conclusão de curso tendem a desistir, isso pode ser pela escolha imatura da profissão ou até imaturidade pessoal:

Uma das principais questões que a literatura levanta na evasão do ensino superior é sobre a escolha do curso. Essa, pode se tornar um motivo para a desistência do aluno, por falta de incompatibilidade acadêmica (ALVES et al, 2017).

Os anos iniciais são primordiais para a permanência dos alunos do curso, pois o discente não está acostumado com o ambiente universitário, com professores e colegas, metodologia das aulas e outros fatores.

As regras geradas apontam também motivos voltados para as dificuldades iniciais que o primeiro contato com o curso causa no aluno, visto que o campo “% conclusão” em todas as regras apresentam “menos de 20”, esse “estranhamento” com o curso ocorre por diversos fatores como quando os alunos possuem má formação da educação básica, o que conseqüentemente interfere na trajetória acadêmica.

Durante a análise das regras percebeu-se a presença de informações importantes e relevantes para gestores da universidade, pois, ao descobrir os padrões, pode-se refletir ainda mais sobre o problema, tendo em vista a tomada de decisões para que o fenômeno possa ser amenizado.

O IFTO possui o Plano Nacional de Assistência Estudantil – PNAES que tem por objetivos no Art. 2ª do DECRETO Nº 7.234, DE 19 de JULHO de 2010:

- I – democratizar as condições de permanência dos jovens na educação superior pública federal;
- II - Minimizar os efeitos das desigualdades sociais e regionais na permanência e conclusão da educação superior;
- III – reduzir as taxas de retenção e evasão; e
- IV – contribuir para a promoção da inclusão social pela educação (BRASIL, 2010).

Com esses objetivos, é notória a preocupação do IFTO em sanar as dificuldades dos seus alunos, visando além do ensino, assistência para assegurá-los na rede de ensino com ações desenvolvidas nas áreas de moradia, alimentação, transporte e outros que visam e refletem na qualidade do ensino e de vida do estudante. Tinto (1982) defende que tudo o que é voltado para políticas, estudos e ações que ajudam no processo de aprendizagem e inclusão do aluno está diretamente ajudando e combatendo à evasão.

9. CONSIDERAÇÕES FINAIS

Com a realização desse estudo, desde o processo de pesquisa que englobou os seus diversos tipos, aos resultados, foi notório o enriquecimento do saber que foi obtido. Com ele foi possível refletir e analisar sobre o perfil dos alunos evadidos do curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis.

Todas as etapas de mineração de dados foram essenciais para se obter os resultados baseados nos objetivos. O processo de KDD se fez muito válido para encontrar as regras de associação e fazer uma análise reflexiva sobre elas, o que acarretou resultados satisfatórios à pesquisa.

Por meio do algoritmo apriori utilizado na linguagem Python, R e na ferramenta WEKA foi possível analisar as regras obtidas pelos dados de 2014 a 2018 do curso supracitado.

Os resultados servem para nortear os gestores do *campus* a continuarem o trabalho de combate à evasão, pois, esse fenômeno traz resultados negativos em diversas vertentes, tanto institucionais quanto sociais. Sendo assim, esse processo é voltado para todos os envolvidos na Instituição de ensino, sendo um papel didático acadêmico que visa à permanência do aluno, pois a educação é a única prática capaz de mudar os mais diversos âmbitos existentes, pois “A sociedade e cada meio social particular estabelece o ideal que educação desenvolve” – Émile Durkheim.

REFERÊNCIAS

ALVEZ, Maria do Carmo Maracaja. RAMOS, Josefa Edileide Santos. BORBA, Marcelo da Costa. MOUTINHO, Lucia Maria Goés. CABRAL, Romilson Marques.

CAUSAS PARA EVASÃO NO PRIMEIRO PERÍODO DOS CURSOS DAS ENGENHARIAS AGRARIAS. 2017. Disponível em <<https://ojs.franca.unesp.br/index.php/caminhos/article/view/2207/2042>> Acesso em: 09 set. 2019.

ANDIFES - ASSOCIAÇÃO NACIONAL DOS DIRIGENTES DAS INSTITUIÇÕES FEDERAIS DE ENSINO SUPERIOR. **UFSM – Estudo apresenta dados sobre a evasão nas universidades federais.** Disponível em: <<http://www.andifes.org.br/ufsm-estudo-apresenta-dados-sobre-a-evasaonasuniversidades-federais/>>. Acesso em: 10 jul. 2019.

AMARAL, João B. do. ***Evasão discente no ensino superior: estudo de caso no Instituto Federal de Educação, Ciência e Tecnologia do Ceará (Campus Sobral). Dissertação (Mestrado Profissional em Políticas Públicas e Gestão da Educação Superior) Universidade Federal do Ceará.*** Fortaleza, 2013.

BACHA, Carlos José Caetano. **Economia e política agrícola no Brasil.** – 2. Ed. – São Paulo: Atlas, 2012

BAGGI, Cristiane A. S.; LOPES, Doraci, A. L. **Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica.** *Avaliação*, Campinas, v.16, n.2, p.355-374, julho, 2011.

BATISTA, S. D.; SOUZA, A. M.; OLIVEIRA, J. M da S. **A Evasão escolar no ensino médio: um estudo de caso.** *Revista Profissão Docente*, Uberaba, v.9, n.19, 2009.

BORGES, M. N.; ALMEIDA, N. N. **Perspectivas para a Engenharia Nacional - Desafios e Oportunidades.** Disponível em: <<http://ptdocz.com/doc/330551/baixar-este-arquivo-pdf>> Acesso em: 08 jun. 2019.

BRASIL. (2010). Decreto n. 7.234, de 19 de jul. de 2010. **Programa Nacional de Assistência Estudantil - PNAES.** Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7234.htm> Acesso em: 09 ago. 2019

BRASIL. CONSTITUIÇÃO FEDERAL (1988). Título II, capítulo II, **dos direitos sociais.**

BRASIL - MINISTÉRIO DA EDUCAÇÃO (MEC). **Instituições de Educação Superior e Cursos Cadastrados.** Disponível em: <<http://emec.mec.gov.br/>>. Acesso em: 3 abr. 2019.

BRASIL/MEC/INEP – **Senso da Educação Superior 2018** – Disponível em: <<http://portal.mec.gov.br/docman/setembro-2018-pdf/97041-apresentac-a-o-censo-superior-u-ltimo/file>> Acesso em: 11 jul. 2019.

BRASIL/MEC/SESu/ABRUEM/ANDIFES. **Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas.** Brasília-DF, 1997. Disponível em: <http://www.udesc.br/arquivos/id_submenu/102/diplomacao.pdf>. Acesso em: 07 jul. 2019.

CREA-TO. **Apesar do alto índice de procura, evasão dos cursos de engenharia é grande, revela estudo.** COBENGE, 2013. Disponível em: <http://www.crea->

to.org.br/noticia-1422-apesar-do-alto-o-ndice-de-procura-evas-o-dos-cursos-de-engenharia-grande-revela-estudo. Acesso em: mar. 2019.

CURI, L. R. L. **O paroxismo da qualidade: avaliação do ensino superior no Brasil**. Campinas, SP: Universidade Estadual de Campinas, 2011.

CACETE, N. H. **Breve história do ensino superior brasileiro e da formação de professores para a escola secundária**. *Educação e Pesquisa*, v. 40, n. 4, p. 1061–1076, 2014.

CASTRO, Alexandre K. dos S. S. de,. **Evasão No Ensino Superior: um estudo no curso de Psicologia da UFRGS**. Dissertação (Programa de Pós-Graduação em Psicologia). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.

CUSTÓRIO, Susana. **Python, Anaconda e Spyder**. Fev. 2017. Disponível em: <https://fenix.ciencias.ulisboa.pt/downloadFile/281612415664146/1_Python.pdf>. Acesso em 07 set. 2019.

DURKHEIM, E. – **O suicídio, Estudo Sociológico**. Editora Presença – 7ª edição (2001)

DANTAS, Eric Rommel G. Dantas. JÚNIOR, José Carlos Almeida Patrício. LIMA, Daniel Silva de. AZAVEDO, Ryan Ribeiro de. **O Uso da Descoberta de Conhecimento em Base de Dados para Apoiar a Tomada de Decisões**. Disponível em: <https://www.aedb.br/seget/arquivos/artigos08/331_331_Artigo_SEGET_EJDR_Versa_o_Final_010808.pdf> Acesso em: 23 de Jul. 2019.

FURTADO, V. V. A.; ALVES, T.W. **Fatores determinantes da evasão universitária: uma análise com alunos da UNISINOS**. *Revista Contemporânea de Economia e Gestão*, v.10, n.2, 2012. Disponível em: <http://www.contextus.ufc.br/2014> Acesso em: 01 jul. 2019.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery: An overview. In: **Advances in Knowledge Discovery and Data Mining**, AAAI Press/The MIT Press, England, 1996, p.1-34.

HOFFMANN, Ivan Londero; NUNES, Raul Ceretta; MULLER, Felipe Martins. **As informações do Censo da Educação Superior na implementação da gestão do conhecimento organizacional sobre evasão**. 2019. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2019000200209> Acesso em: 17 de jul. 2019.

JÚNIOR, Thadeu Teixeira. **O dia do Engenheiro Agrônomo: Produção agrícola x Sustentabilidade**. 2019. Disponível em: <<https://central3.to.gov.br/arquivo/373199/>> Acesso em: 28 ago. 2019.

LUCCHIARI, Dulce Helena Penna Soares. **As Diferentes Abordagens em Orientação Profissional em Ação: Formação e Prática de Orientandos**. In: *Orientação Profissional em Ação: formação e prática de orientadores*. São Paulo: Summus, 2000 p.[28]-225.

LOURENÇO, Ana V. M. **O fenômeno da evasão no ensino superior no curso de Administração no Estado do Rio De Janeiro nos anos de 2006 a 2012: um**

estudo de caso UNIGRANRIO. Dissertação (Mestrado em Administração). Rio de Janeiro: Universidade do Grande Rio., 2014.

LIMA JUNIOR, Paulo. **Evasão do ensino superior de Física segundo a tradição disposicionalista em Sociologia da Educação.** Tese (Programa de Pós-Graduação em Ensino de Física). Universidade Federal do Rio Grande do Sul. Porto Alegre, 2013.

LIMA, Franciele Santos de; Zago, Nadir. **EVASÃO NA EDUCAÇÃO SUPERIOR: tendências e resultados de pesquisa.** 2018.

LOPES, João C. de S. **Evasão nos cursos de graduação em Ciências Contábeis em instituições de ensino superior da Região Sul do Brasil.** Dissertação (Programa de Pós-Graduação em Ciências Contábeis) - Universidade do Vale do Rio dos Sinos. São Leopoldo, 2014.

LOBO, M. B. D. C. M. **Panorama da Evasão no Ensino Superior Brasileiro: Aspectos Gerais das Causas e Soluções.** Instituto Lobo para Desenvolvimento da Educação, ciência e da Tecnologia, v. 25, n. 08780 220, p. 1–23, 2012.

MATOS, David. **R ou Python para análise de dados?. Ciência de Dados.** 2018. Disponível em: < <http://www.cienciaedados.com/r-ou-python-para-analise-de-dados/>> Acesso em: 14 set. 2019.

MEC, Ministério da Educação. **Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras. *Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas.*** Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (ANDIFES), Associação Brasileira dos Reitores das Universidades Estaduais e Municipais (ABRUEM), Ministério de Educação e Cultura. Secretaria de Ensino Superior. Brasília, 1997. 152 p.

MEC/SESU. **Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras.** Brasília: ANDIFES/ABRUEM/SESU/MEC. 1997.

_____**MINISTÉRIO DA EDUCAÇÃO. Documento Orientador para a Superação da Evasão e Retenção na Rede Federal de Educação Profissional, Científica e Tecnológica.** 52 p. Brasília, 2014.

OLIVEIRA, V. F. (Org.) **Trajetória e estado da arte da formação em Engenharia, Arquitetura e Agronomia. Volume I: Engenharias.** 1. ed. Brasília: INEP/MEC, 2010. v.1. 304p. Disponível em: <<http://www.ufjf.br/observatorioengenharia/files/2012/01/vol1.pdf>.> Acesso em: 5 jul. 2019.

PROJETO PEDAGÓGICO DO CURSO DE GRADUAÇÃO DE ENGENHARIA AGRÔNOMICA. Dianópolis – TO, 2014. Disponível em: <<http://www.ifto.edu.br/ifto/colegiados/consup/documentos-aprovados/ppc/campus-dianopolis/bacharelado-em-engenharia-agronomica/ppc-engenharia-agronomica-campus-dianopolis-1.pdf/view>>. Acesso em: 14 mar. 2019.

PIVA, Adriana. **A IMPORTÂNCIA DA AGRONOMIA PARA O DESENVOLVIMENTO ECONÔMICO DO PAÍS.** 2017. Disponível em: <<http://revistaecoturismo.com.br/turismo-sustentabilidade/importancia-da-agronomia-para-o-desenvolvimento-economico-do-pais/>> Acesso em: 27 ago. 2019.

PARENTE, Nória N. **As condições de acesso e permanência dos estudantes do curso de licenciatura em Física do IFCE, Campus De Sobral**. Dissertação (Mestrado Profissional em Políticas Públicas e Gestão da Educação Superior). Universidade Federal do Ceará. Fortaleza, 2014.

PRASS, Fernando Sarturi. **KDD – UMA VISÃO GERAL DO PROCESSO**, 2016. Disponível em: <http://fp2.com.br/blog/wp-content/uploads/2012/07/kdd_uma_visao_geral_do_processo.pdf> Acesso em: 23 jul. 2019.

PYTHON. **Guia para iniciantes**. 2017. Disponível em: <<https://wiki.python.org/moin/BeginnersGuide/Overview>> Acesso em: 10 ago. 2019.

PYTHON BRASIL – Empresas que usam Python. 2019. Disponível em: <<https://python.org.br/empresas/>> Acesso em: 20 set. 2019.

RAMOS, Lilian das G. **Dois ensaios sobre Educação Superior no Brasil. Dissertação (Programa de Pós-Graduação em Economia do Desenvolvimento**. Faculdade de Administração, Contabilidade e Economia). Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, 2013.

ROMÃO, Wesley. et al. **EXTRAÇÃO DE REGRAS DE ASSOCIAÇÃO EM C&T: O ALGORITMO APRIORI**. 2019. Disponível em: <http://www.abepro.org.br/biblioteca/ENEGEP1999_A0901.pdf> Acesso em: 8 ago. 2019.

R. **What is R?**. 2019. Disponível em: <<https://www.r-project.org/about.html>> Acesso em: 14 set. 2019.

SANTANA, A. P.; PEROSSO, J. E. C.; MACEDO, K. L. O.; FARIAS, S. P. D de. **Evasão escolar em escolas públicas municipais rurais localizadas em Montes Claros**. Trabalho de Conclusão de Curso. Universidade Estadual de Montes Claros. Montes Claros: 1996. Acesso em: mar. 2019.

RStudio. **Why RStudio**. 2019. Disponível em: <<https://www.rstudio.com/about/>> Acesso em: 15 set. 2019.

ROSA, Gabriel; SOBRINHO, Isaac P. F.; RODRIGUES, Diego. **O Uso de Ferramentas de Mineração de dados como auxílio na Prevenção da Evasão nas Universidades**. 8ª JICE, 2017. Disponível em: <<http://propi.ifto.edu.br/ocs/index.php/jice/8jice/paper/view/8190>> Acesso em 15 set. 2019.

SEAGRO – TO. **Agricultura**. 2019. Disponível em: <<https://seagro.to.gov.br/agricultura/>> Acesso em: 29 de ago. 2019.

SETEC. **A AGROPECUÁRIA DO ESTADO DE TOCANTINS**. 2019. Disponível em: <<http://portal.mec.gov.br/setec/arquivos/>> Acesso em: 29 ago. 2019.

SEMESP – **MAPA DO ENSINO SUPERIOR 2015**. Disponível em: <<http://convergenciacom.net/pdf/mapa-ensino-superior-brasil-2015.pdf>>. Acesso em: 11 jul. 2019.

SPADY, W.G. **Dropouts from higher education: An interdisciplinary review and synthesis**. *Interchange*. Vol. 1, nº 1. (1970)

SCHIMITT, Rafael E.. **A evasão na educação superior: uma compreensão ecológica do fenômeno como estratégia para a gestão da permanência estudantil**. In: X Anped Sul -Reunião Científica da ANPED, 2014. Textos Completos. Florianópolis: UDESC, Outubro, 2014 p 1-21. Disponível em: <http://xanpedsul.faed.udesc.br/arq_pdf/690-0.pdf>. Acesso em: set. 2019.

SILVA FILHO, R. L.L., MOTEJUNAS, P. R., HIPÓLITO, O.; LOBO, M. B. **A evasão no ensino superior brasileiro**. *Cadernos de Pesquisa*, Maranhão, v. 37, n. 132, p. 641-659, setembro/dezembro. 2007.

SANTOS, Priscila K. **Abandono na Educação Superior: um estudo do tipo Estado do Conhecimento**. *Educação Por Escrito*. Porto Alegre, v. 5, n. 2, p. 240-255, julho/dezembro, 2014.

SOBRINHO, José F. de O. **Evasion the top distance education: a case study in the UFC Virtual Institute**. Dissertação (Mestrado Profissional em Políticas Públicas e Gestão da Educação Superior). Universidade Federal do Ceará. Fortaleza, 2013.

TINTO, V., **Limits of Theory and Practice in Student Attrition**, *The Journal of Higher education*, Vol. 53 (1982).

TINTO, V., **Student Success and the Construction of Inclusive Educational Communities, American Association of State Colleges and Universities - AASCU**, (2005).

TONTINI, Gérson; WALTER, Silvana. **Pode-se identificar a propensão e reduzir a evasão de alunos? Ações estratégicas e resultados táticos para instituições de ensino superior**. *Avaliação*, Campinas, v. 19, n. 1, p. 89-110, março. 2014.

UOL. **O Ranking Universitário Folha**. 2017. Disponível em: <<https://ruf.folha.uol.com.br/2017/ranking-de-cursos/agronomia/>> Acesso em: 25 ago. 2019.

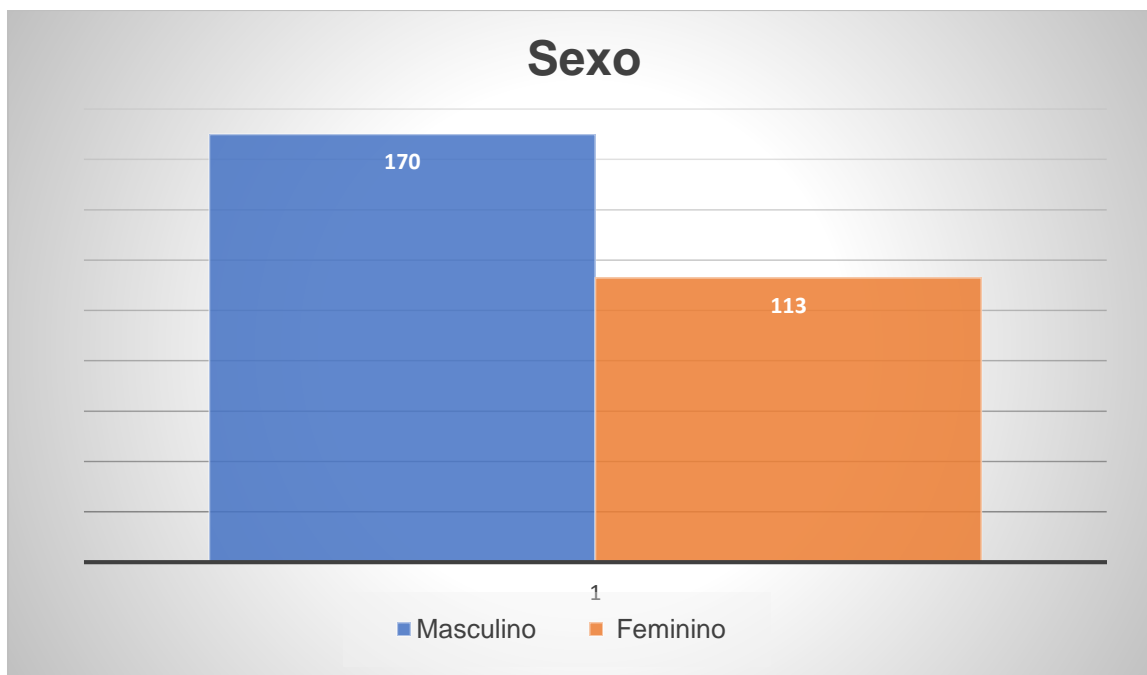
VANDERLÍ, Fava de Oliveira. NIVAL, Nunes de Almeida. **Trajetória e estado da arte da formação em Engenharia, Arquitetura e Agronomia – volume I: Engenharias / Organizador: Vanderlí Fava de Oliveira**. – Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Conselho Federal de Engenharia, Arquitetura e Agronomia, 2010.

WAZLAWICK, Raul Sidnei. **Metodologia de pesquisa para ciência da computação/Raul Sidnei Wazlawick**. Rio de Janeiro, 2009.

WEKA. *Weka 3: Machine Learning Software in Java*. 2019. Disponível em: <<https://www.cs.waikato.ac.nz/~ml/weka/>> Acesso em: 15 set. 2019.

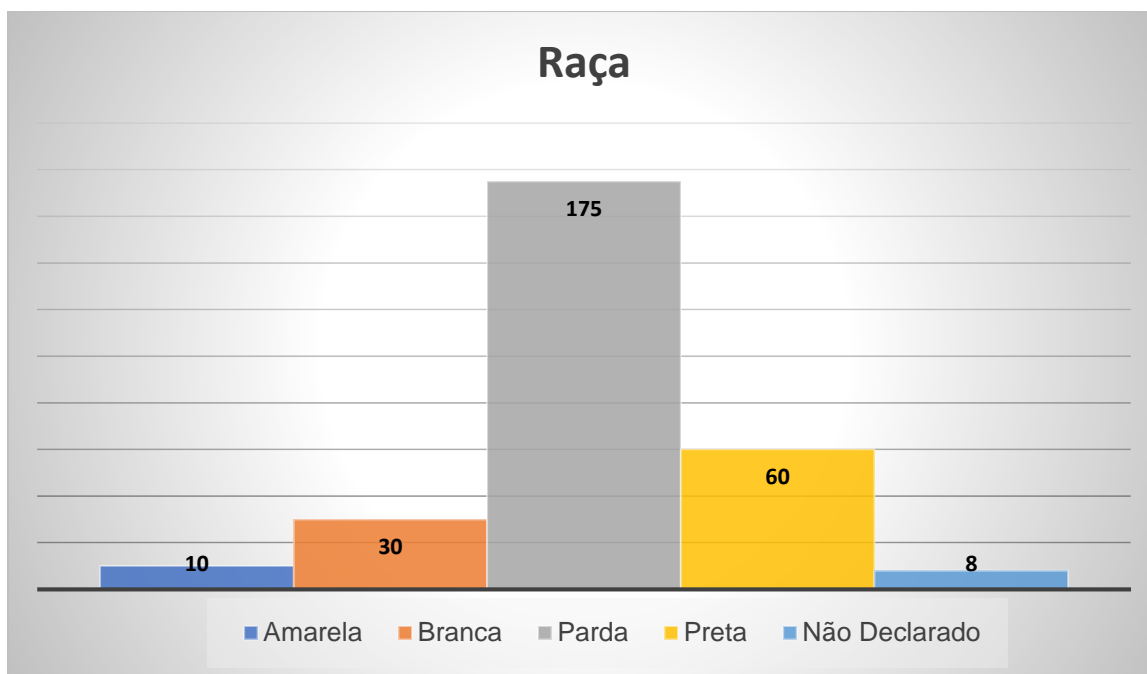
APÊNDICES

APÊNDICE A – Número de alunos por Sexo no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis, do período de 2014 a 2018.



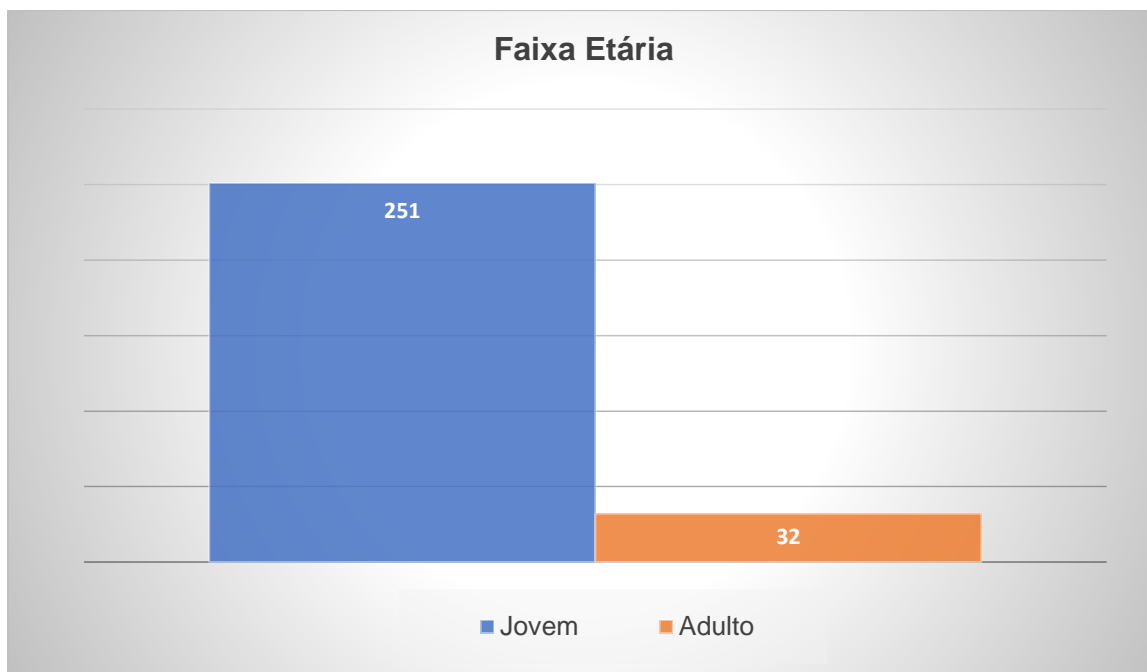
Fonte: O autor

APÊNDICE B – Número de alunos por Raça no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis, do período de 2014 a 2018.



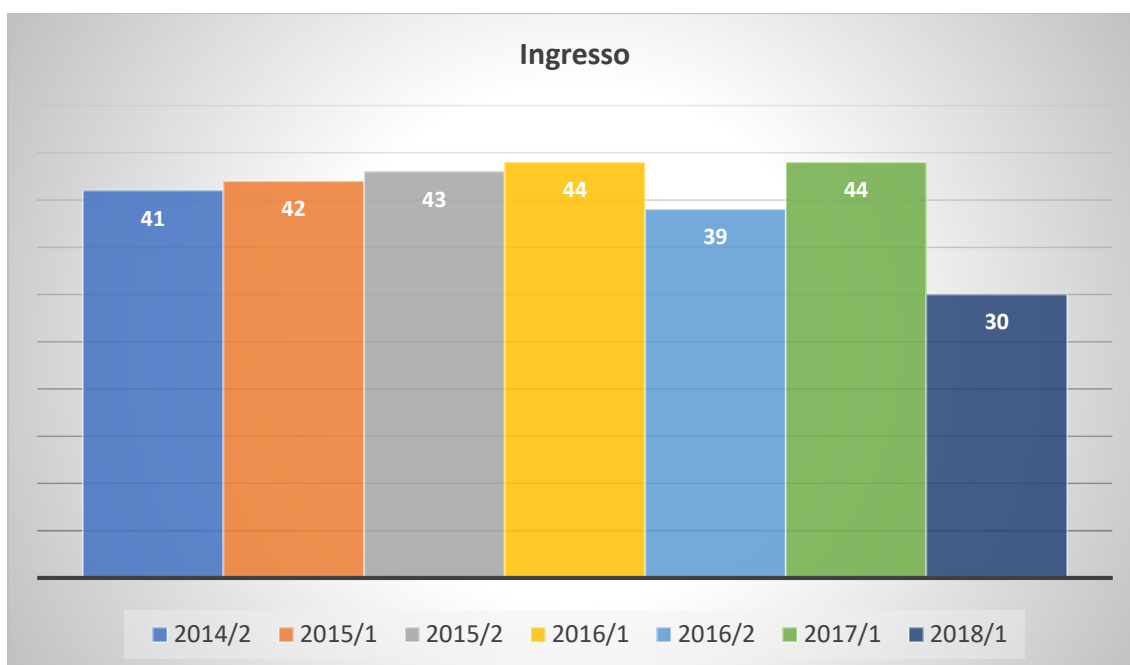
Fonte: O autor

APÊNDICE C – Número de alunos por Faixa Etária no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis, do período de 2014 a 2018.



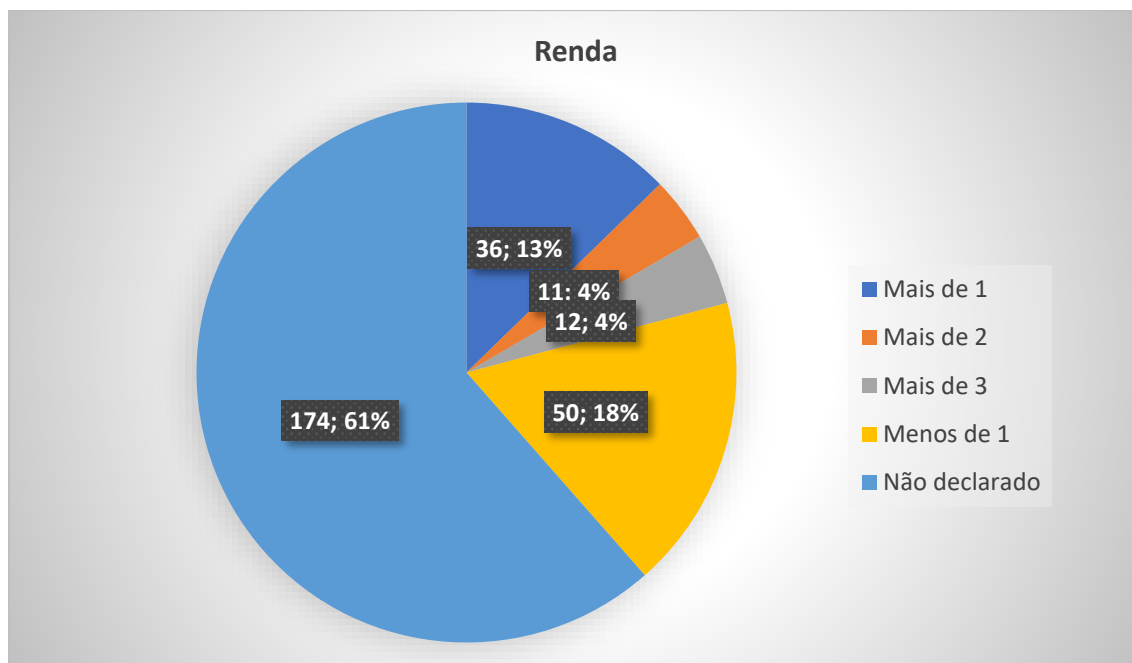
Fonte: O autor

APÊNDICE E – Número de alunos por ano de Ingresso no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis.



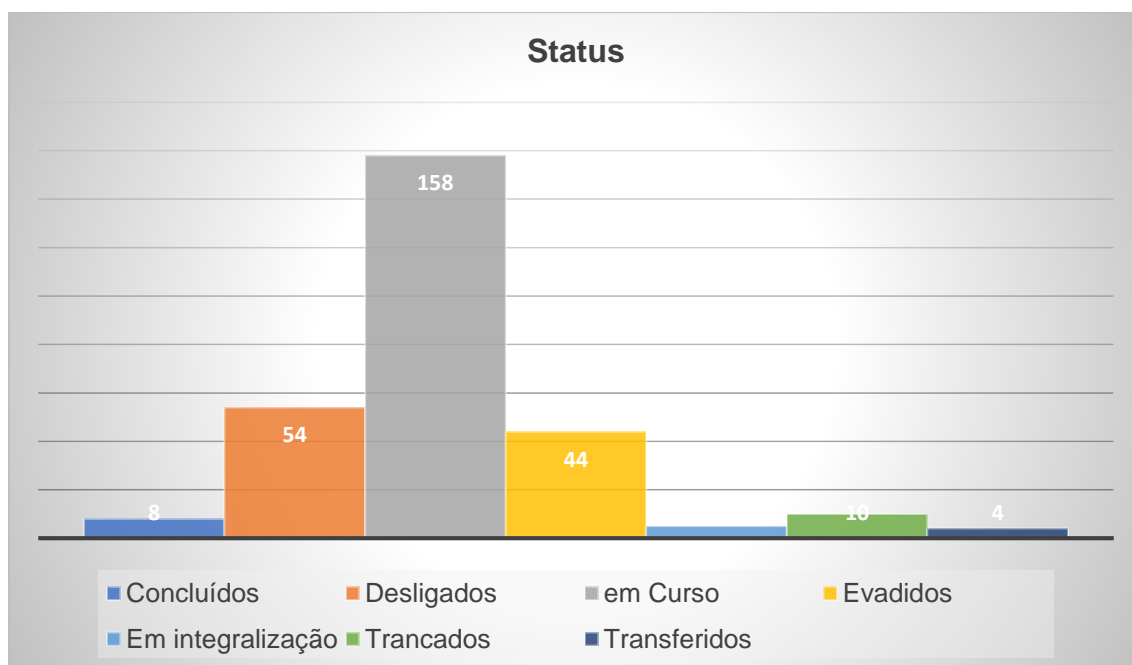
Fonte: O autor

APÊNDICE F – Número e porcentagem de alunos por Renda no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis, do período de 2014 a 2018.



Fonte: O autor

APÊNDICE G – Número de alunos por Status no curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis, do período de 2014 a 2018.



Fonte: O autor

SEI/IFTO - 0682245 - Ofício

<https://sei.iftto.edu.br/sei/controlador.php?acao...>

MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO TOCANTINS
CAMPUS DIANÓPOLIS
GERÊNCIA DE ENSINO
COORDENAÇÃO DE REGISTROS ESCOLARES

OFÍCIO Nº 107/2019/CORES/GE/DNO/REI/IFTO

Dianópolis, 21 de maio de 2019.

Ao Sr. Gerente de Ensino do *Campus* Dianópolis/IFTO.Assunto: **Solicitação de Dados Trabalho de Conclusão**

1. Cumprimentando-o, tendo em vista cumprir com as obrigações de expediente desse setor, encaminho processo 23481.013598/2019-55, sobre pedido dos alunos Gabriel Rosa e Isac Pimentel Fernandes.
2. Neste sentido, encaminho o requerimento anexo, para ser analisado e decidido pela Gerência de Ensino.
3. Sem mais, colocamo-nos à disposição.
Atenciosamente,

Roselaine Gusson Mendes
Coordenadora de Registros Escolares

Documento assinado eletronicamente por **Roselaine Gusson Mendes**, **Coordenadora**, em 21/05/2019, às 16:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.iftto.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0682245** e o código CRC **19A811CC**.

SEI/IFTO - 0682245 - Ofício

<https://sei.ifto.edu.br/sei/controlador.php?acao...>

Rodovia TO - 040 - Km 349
Lote 01 - Loteamento Rio Palmeiras
CEP 77.300-000 Dianópolis - TO
(63) 99947-3511
www.ifto.edu.br - cores.dianopolis@ifto.edu.br

Referência: Processo nº 23481.013598/2019-55

SEI nº 0682245

SEI/IFTO - 0686397 - Despacho

<https://sei.ifto.edu.br/sei/controlador.php?acao...>

MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO TOCANTINS
CAMPUS DIANÓPOLIS
DIREÇÃO-GERAL
GERÊNCIA DE ENSINO

DESPACHO nº 107/2019/GE/DNO/REI/IFTO

Dianópolis, 24 de maio de 2019.

Processo n.º: **23481.029457/2018-73**Interessado: **Instituto Federal do Tocantins**Assunto: **Solicitação de Dados Trabalho de Conclusão**

Ao Senhor Coordenador de Pesquisa, Extensão, Pós-graduação e Inovação do Campus Dianópolis.

1. Em atenção ao OFÍCIO Nº 107/2019/CORES/GE/DNO/REI/IFTO (0682245), encaminhamos o processo em tela, que trata do pedido de utilização dos dados do SIGA para pesquisa de TCC do Curso de Licenciatura em Computação (0682240), para a referida coordenação orientar a formalização dos procedimentos para utilização dos dados institucionais.
2. Solicitamos que as presentes informações sejam encaminhadas aos interessados para fins de conhecimento e encaminhamentos necessários.
3. Atenciosamente,

Ítalo Cordeiro Silva Lima
Gerente de Ensino

Documento assinado eletronicamente por **Italo Cordeiro Silva Lima, Gerente**, em 26/05/2019, às 20:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.ifto.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0686397** e o código CRC **F87620FD**.

Rodovia TO - 040 - Km 349
Lote 01 - Loteamento Rio Palmeiras
CEP 77.300-000 Dianópolis - TO
(63) 99947-3511
www.ifto.edu.br - geren.dianopolis@ifto.edu.br

SEI/IFTO - 0686397 - Despacho

<https://sei.iftto.edu.br/sei/controlador.php?acao...>

Referência: Processo nº 23481.013598/2019-55

SEI nº 0686397

SEI/IFTO - 0725846 - E-mail

<https://sei.iftto.edu.br/sei/controlador.php?acao...>**E-mail - 0725846****Data de Envio:**

05/07/2019 12:13:58

De:

IFTO/Coordenação de Pesquisa e Extensão <copex.dianopolis@iftto.edu.br>

Para:

cores.dianopolis@iftto.edu.br

Assunto:

Utilização de Dados Pesquisa TCC

Mensagem:

Prezada Coordenadora da CORES, Roselaine Gusson Mendes,

entramos em contato através deste, solicitando o fornecimento dos dados do SIGA para pesquisa de TCC do Curso de Licenciatura em Computação (processo 23481.013598/2019-55, dos alunos Gabriel Rosa e Isaac Pimentel Fernandes), desde que durante a transmissão dos dados, em hipótese alguma ocorra a identificação do usuário e o acesso as informações pessoais, prontuários, banco de dados não acessível ao público, etc.

Respeitosamente,
Wanessa de Sousa Rocha

Coordenadora Substituta de Pesquisa, Extensão, Pós-graduação e Inovação
Portaria n.º 198/2019/DNO/REI/IFTO, de 25 de junho de 2019.