

ANÁLISE COMPARATIVA DO DESEMPENHO ACADÊMICO DOS ALUNOS FORMADOS PELA FATEC INDAIATUBA NOS CURSOS DE ANÁLISE E DESENVOLVIMENTO DE SISTEMAS E GESTÃO EMPRESARIAL UTILIZANDO TÉCNICAS DE BIG DATA

COMPARATIVE ANALYSIS OF ACADEMIC PERFORMANCE OF STUDENTS GRADUATED BY FATEC INDAIATUBA IN COURSES ON ANALYSIS AND DEVELOPMENT OF SYSTEMS AND BUSINESS MANAGEMENT USING BIG DATA TECHNIQUES

Gabriel Vieira Santello¹

Aldo Nascimento Pontes²

Maria das Graças Junqueira Machado Tomazela³

RESUMO

Os modelos analíticos para interpretação de dados extraídos de plataformas educacionais podem auxiliar na análise e visualização de informações, ajudando a prever o desempenho dos alunos, gerando recomendações, fornecendo feedback, dentre outras aplicações. Atualmente, muitas instituições de ensino usam a análise de dados para melhorar os serviços que fornecem, tanto direta quanto indiretamente. Nesse contexto, a implementação de técnicas de Big Data oferece uma grande oportunidade, pois possibilita trabalhar com um grande volume de dados e métricas. Sendo assim, este trabalho teve como objetivo utilizar ferramentas de Big Data para analisar o desempenho acadêmico dos alunos da FATEC Indaiatuba dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial, a partir de dados obtidos no cadastro das disciplinas no sistema SIGA das FATEC, comparando parâmetros relevantes entre os cursos, visando a apoiar a tomadas de decisão na faculdade em ações educacionais e para redução de evasão de alunos. Para alcançar o objetivo proposto, foi realizada uma pesquisa experimental, composta pela revisão bibliográfica e aplicação de técnicas de Big Data por meio das bibliotecas Pandas e Matplotlib para Python. Assim foi possível determinar a relação entre o resultado acadêmico dos alunos, as informações sobre quais semestres são mais propensos para o aluno desistir do curso e a influência da origem do aluno (escola pública ou particular) com os níveis de evasão, possibilitando maior engajamento do corpo discente e trazendo novas possibilidades de direcionamentos para as Faculdades de Tecnologia do Estado de São Paulo (FATEC).

Palavras chave: Análise de Dados; Dados Acadêmicos; Tecnologia

¹ Graduando do curso de Análise e Desenvolvimento de Sistema. Faculdade de Tecnologia de Indaiatuba

² Professor Doutor da Faculdade de Tecnologia de Indaiatuba.

³ Professora Doutora da Faculdade de Tecnologia de Indaiatuba. R. Dom Pedro I, 65 - Cidade Nova I, Indaiatuba - SP, 13334-100. Tel. (19) 3885-1923. e-mail: graca.tomazela@fatec.sp.gov.br.

ABSTRACT

Analytical models for interpreting data extracted from educational platforms can help in the analysis and visualization of information, helping to predict student performance, generating recommendations, providing feedback, among other applications. Today, many educational institutions use data analytics to improve the services they provide, both directly and indirectly. In this context, the implementation of Big Data techniques offers a great opportunity, as it makes it possible to work with a large volume of data and metrics. Therefore, this paper aimed to use Big Data tools to analyze the academic performance of FATEC Indaiatuba students in Systems Analysis and Development and Business Management courses, based on data obtained from the registration of disciplines in the SIGA system of FATECs, comparing relevant parameters between courses, aiming to support decision-making at the faculty in educational actions and to reduce student dropout. Experimental research was carried out, a bibliographical review and application of Big Data techniques through the Pandas and Matplotlib libraries for Python. It was possible to determine the relationship between the students' academic results, information about which semesters are more likely for students dropping out the course and the influence of the student's origin (public or private school) with dropout levels, enabling greater student engagement and bringing new possibilities of directions for the Faculties of Technology of the State of São Paulo (FATEC).

Keywords: Data analysis; Academic Data; Technology

¹ Graduando do curso de Análise e Desenvolvimento de Sistema. Faculdade de Tecnologia de Indaiatuba

² Professor Doutor da Faculdade de Tecnologia de Indaiatuba.

³ Professora Doutora da Faculdade de Tecnologia de Indaiatuba. R. Dom Pedro I, 65 - Cidade Nova I, Indaiatuba - SP, 13334-100. Tel. (19) 3885-1923. e-mail: graca.tomazela@fatec.sp.gov.br.

1 INTRODUÇÃO

A cada ano está havendo um aumento acentuado na quantidade de dados gerados nos mais diversos sistemas criados em todo o mundo. A disseminação de diferentes tipos de dispositivos eletrônicos e seus usos geram continuamente enormes quantidades de dados, como por exemplo dados de radiofrequência obtidos por RFID, dados de sensores, dados de interação de redes sociais e dados de Internet móvel. Dentre esses, encontram-se dados estruturados, semiestruturados (ou fracamente estruturados) e desestruturados, essenciais para significância dos estudos, mas que torna padronizações e análises difíceis com ferramentas até então tradicionais (HU *et al.*, 2016). Em essência, foram gerados aproximadamente 2,3 trilhões de gigabytes por dia em dados no mundo no ano de 2017 (LYNCH, 2017), sendo que entre 2015 e 2017, o volume de dados na Internet aumentou em 5 zettabytes em comparação com o ano anterior que atingiu 14,5 zettabytes (KOCHETKOV; PROKHOROV, 2017). Desse modo, surgiram novas técnicas para manipulação de quantidades massivas de dados, caracterizando-se assim o *Big Data*. Este termo, por sua vez, refere-se a quantidades de dados, que são muito grandes e/ou complexas para serem tratadas com eficácia e eficiência por teorias, tecnologias e ferramentas tradicionais.

Dentre as inúmeras possíveis aplicações, a tecnologia de análise de *Big Data* pode fornecer análise de dados para o gerenciamento de assuntos educacionais. Um dos objetivos da aplicação de *Big Data* no sistema de educação inteligente é a visualização de dados, que permite aos usuários observar diretamente os resultados do processamento de dados. Outro objetivo é fornecer parâmetros para tomada de decisão, dando um julgamento prospectivo em um certo grau de acordo com os dados visualizados, que é onde o presente trabalho foca seus esforços. A tecnologia de *Big Data* pode melhorar a eficiência operacional das instituições de ensino, minerando a informação e o conhecimento escondidos em uma vasta quantidade de dados e fornecendo a base para atividades sociais e econômicas de seres humanos. A análise de grandes volumes de dados também pode ser aplicada para melhorar a qualidade do ensino nas universidades, analisar comportamentos de usuários na rede universitária e prever o comportamento de determinado grupo de alunos em vários aspectos (HU *et al.*, 2016).

As pesquisas sobre Big Data são recentes, assim vários trabalhos trazem os principais conceitos e definições da área. Por exemplo os de trabalhos Kochetkov e Prokhorov (2017) e Lynch (2017) discutem as abordagens para o uso de *Big Data* no processo educacional de instituições de ensino superior. O trabalho de Larson e Chang (2016) aborda sobre a importância das técnicas de visualização como um componente chave para análise rápida de

Big Data.

Algumas pesquisas possuem caráter aplicado, como a pesquisa de O'Neil e Schult (2013) que apresenta conceitos e técnicas sobre *Data Science* utilizando modelos estatísticos e algoritmos em linguagens de programação como o R. Também no conjunto das pesquisas aplicadas, têm-se os trabalhos de Wassan (2015) e West (2012), essas pesquisas mostram que várias instituições educacionais também desenvolveram softwares de *dashboard* e armazenamento de dados que permitem acompanhar os problemas de aprendizado, desempenho e comportamento dos alunos.

Vyas e Gulwani (2017) tratam sobre a utilização de algoritmos computacionais para prever quais alunos em uma universidade precisam de apoio diferenciado para que não abandonem seu curso. Para isso, foi utilizada a técnica CART, que implementa árvores de decisão binárias, uma estrutura estatística capaz de detectar alunos com problemas ou deficiências em matérias-chave.

Hu *et al.* (2017) analisaram de maneira sistemática os diversos dados relacionados à educação, provenientes de duas universidades públicas, com o objetivo de auxiliar os alunos na tomada de decisão, usando de informações sobre suas escolhas futuras ao decorrer do curso.

Os recursos das instituições de ensino, para serem bem aplicados, devem ser utilizados de acordo com as respostas que a análise de dados feita para cada uma das escolas traz, assim os responsáveis pela gestão decidem pontos focais de atuação para evitar a evasão escolar e o melhor aproveitamento dos alunos. O índice de evasão e a diversidade de alunos são fatores que estão inseridos nos objetivos centrais das escolas, faculdades e universidades, de acordo com os trabalhos citados anteriormente. Os dados coletados diariamente por essas instituições podem ser úteis para sua própria gestão. O impacto da pesquisa se dá no auxílio da tomada de decisão da faculdade, pela destinação de recursos financeiros e educacionais.

Diante do contexto apresentado, o objetivo deste trabalho foi aplicar ferramentas de *Big Data* para analisar as informações dos alunos da Faculdade de Tecnologia (FATEC) de Indaiatuba dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial, a partir de dados obtidos no cadastro das disciplinas no sistema SIGA das FATEC, comparando parâmetros relevantes entre os cursos. Sendo assim, esse trabalho apresenta a hipótese que se as informações dos alunos da FATEC Indaiatuba dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial forem analisadas por meio de técnicas de *Big Data*, será possível auxiliar a gestão da faculdade nos processos educacionais e diminuir a evasão escolar durante a graduação.

Ressalta-se que, conforme afirma Marquesone (2016), embora o volume de dados seja

um dos fatores que impulsionou o surgimento do *Big Data*, aplicações que utilizam conjuntos de dados em uma escala menor também podem se beneficiar das tecnologias de *Big Data* para extrair valor dos dados. Além disso a autora afirma que, apesar da crescente necessidade de estruturas mais flexíveis que os modelos relacionais, dados estruturados continuam sendo muito utilizados em soluções de *Big Data*. Dessa forma, justifica-se a aplicação da tecnologia de *Big Data* no presente trabalho.

2 MATERIAL E MÉTODO

O desenvolvimento deste estudo se deu pela abordagem experimental que, segundo Gil (2017), consiste em determinar um objeto de estudo, selecionar as variáveis capazes de influenciá-lo e definir meios de controle e observação dos efeitos que essa variável produz nesse objeto.

Durante o desenvolvimento da análise, foram definidas as variáveis para tornar a coleta e a análise de dados relevante:

1. Dados pessoais dos alunos da FATEC Indaiatuba quanto a identificação e sua caracterização;
2. Categorias de dados a serem analisados para caracterizar o desempenho de cada aluno;
3. Dados para a comparação dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial da FATEC Indaiatuba, durante sua coexistência.

2.1 Dados utilizados na análise

Para a realizar este trabalho, foi feita a análise de dados coletados por meio dos sistemas de gerenciamento dos alunos da FATEC Indaiatuba (SIGA), com foco nos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial. Os dados para a realização desta pesquisa foram obtidos e cedidos pela FATEC Indaiatuba no formato padrão de planilhas do Microsoft Office Excel (.xlsx). A coleta se deu pelos desenvolvedores da plataforma SIGA, com a autorização do diretor geral desta instituição e repassado pela professora orientadora deste projeto.

A tabela concedida pela FATEC Indaiatuba continha inicialmente dados de todos os cursos oferecidos pela instituição, totalizando 19380 linhas. Como filtragem, realizou-se

seleções buscando somente pelos dados dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial. Como se trata de uma análise comparativa entre esses cursos, foi necessário padronizar o período temporal utilizado nas análises. Desse modo, optou-se pelo período de coexistência dos cursos na FATEC Indaiatuba, sendo que todos os dados considerados são de alunos do período noturno. Foram selecionados então dados partindo do primeiro semestre de 2012 até o segundo semestre de 2018.

Como a planilha foi cedida em formato estruturado (.xlsx), foi realizada a conversão para o formato .csv, suportado pela biblioteca Pandas para a manipulação dos dados. É importante notar que, no geral, as ferramentas utilizadas em *Big Data* foram desenvolvidas para trabalhar com dados desestruturados.

Inicialmente foi feita uma varredura geral para definição de quais seriam as colunas, presentes na tabela, relevantes para as respostas buscadas pelo projeto. As colunas foram então mapeadas e descritas para que fosse possível serem utilizadas durante a codificação nas etapas posteriores.

Para a coluna CURSO foi feita a filtragem visando a selecionar apenas os cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial. Quanto ao desempenho acadêmico, foram selecionados os campos DISCIPLINA, NOTA, FREQUENCIA, STATUS_ALUNO, SEMESTRE_ANO e CONCEITO (de aprovação ou reprovação na disciplina). Ademais, a tabela apresenta, dentre outros fatores, as colunas TURNO (turno em que o aluno estuda), RA (número de registro acadêmico do aluno – utilizado como chave primária em certas análises), NOME (nome do aluno), ESCOLA_PUBLICA (expõe se o aluno é proveniente de uma escola pública ou particular) e NOTA_VESTIBULAR (desempenho do aluno no vestibular).

2.2 Tecnologias e ferramentas utilizadas

A partir da seleção e limpeza inicial dos dados, estes foram importados na nuvem por meio do *Google Colab Notebook*, não sendo necessário criar virtualizações do sistema, uma vez que essa ferramenta já funciona como um sistema isolado.

A limpeza e processamento inicial dos dados foi feita utilizando também a biblioteca *Pandas* para Python, sendo esses então armazenados em um arquivo .csv com todos os dados padronizados e corretamente distribuídos entre as colunas, eliminando campos com dados faltantes. Uma vez que os dados estavam prontos para uso, organizados e processados, esses foram armazenados na nuvem em servidores Google e trabalhados por meio da ferramenta

Google Colab Notebook. Para a visualização dos resultados, a biblioteca *Matplotlib* para Python foi essencial, uma vez que com ela é possível a geração de gráficos e adaptação destes às necessidades do trabalho, tais como modificações de estilo e cor.

2.3 Indicadores de desempenho

Buscando sanar as perguntas decorrentes da hipótese deste projeto, definiu-se quais indicadores de desempenho seriam utilizados para comparar os resultados dos cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial da FATEC Indaiatuba. Assim, foram definidos os seguintes índices: medianas das notas do vestibular; quantidade de alunos que concluíram o curso (3 anos ou mais de curso); número de semestres para conclusão do curso e índice de evasão; porcentagem inicial de alunos vindos de escola pública e particular; porcentagem de alunos egressos vindos de escola pública e particular; desistências por semestre; aprovação em disciplinas por categoria (Humanas, Exatas e Tecnológicas).

3 RESULTADO E DISCUSSÃO

3.1 Notas do Vestibular e Dados Gerais

A comparação entre os cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial da FATEC Indaiatuba, apresentados a partir deste ponto pelas siglas “ADS” e “GE” respectivamente, iniciou-se pela obtenção de dados relevantes entre os dois cursos. Primeiramente, optou-se por descobrir qual era o total de alunos analisados para cada um dos cursos. Utilizando o método *describe* da biblioteca Pandas, concluiu-se que seriam analisados os dados de 1047 alunos do curso de Gestão Empresarial e 700 alunos do curso de Análise e Desenvolvimento de Sistemas.

Sabendo disso, o próximo passo dado buscou comparar o desempenho desses alunos no vestibular, por meio do agrupamento de dados utilizando a coluna `NOTA_VESTIBULAR`.

Buscando obter respostas mais relevantes, decidiu-se então por calcular as medianas das notas de vestibular de cada um dos cursos. A nota mediana de GE foi de 57,75 pontos, enquanto a mediana de ADS foi de 61,875 pontos. Esse dado não corrobora intuitivamente com a concorrência histórica entre os dois cursos, pois GE possui mais candidatos/vagas do que ADS.

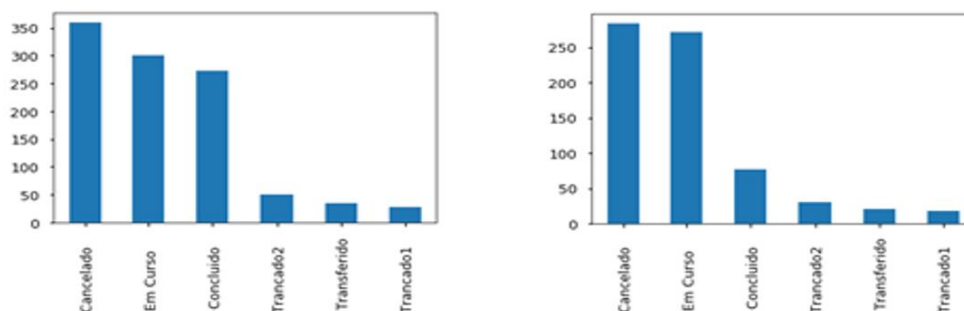
A partir deste ponto de corte, algumas análises subsequentes foram realizadas com

dados totais e dados apenas de alunos que estavam abaixo da mediana na nota do vestibular. Isso porque, intuitivamente, espera-se que alunos com menores notas tenham mais dificuldades no decorrer do curso. Havia então dados de 517 alunos abaixo da mediana de notas no vestibular para GE e de 349 alunos de ADS. Quando a análise exige que sejam feitas comparações apenas entre alunos que concluíram os cursos, tem-se dados de 273 alunos para GE (183 abaixo da mediana) e 77 para ADS (44 abaixo da mediana).

A figura 1 mostra a situação dos alunos de cada curso no momento desta análise. Os gráficos foram obtidos por meio do agrupamento de linhas do banco de dados considerando o STATUS_ALUNO e RA, para que cada um dos RAs fossem contados apenas uma vez. É importante notar que, apesar de um alto número de matrículas canceladas e baixo número de conclusões de curso, ainda há um grande número de alunos em curso, que podem ter sua evasão evitada se medidas forem tomadas.

Os status a seguir, presentes na Figura 1, se dividem entre Cancelado (GE 34,4%; ADS 40,5%), Em Curso (GE 28,7; ADS 38,7%), Concluído (GE 26%; ADS 11%), Transferido (para outras instituições de ensino) (GE 4,9%; ADS 3%), Trancado 1 (alunos que estão no primeiro trancamento de matrícula) (GE e ADS 2,6%) e Trancado 2 (alunos que estão no segundo trancamento de matrícula) (GE 3,3%; ADS 4,3%). Uma vez que um aluno retorna de um trancamento de matrícula, ele passa a figurar no banco de dados como “Em Curso”, assim não é mais possível saber se ele já foi um aluno em trancamento ou não para aquele momento da análise.

Figura 1: Status da matrícula dos alunos dos cursos de GE e ADS, respectivamente



Fonte: Elaborado pelos autores

3.2 Índice de Evasão e Tempo de Formação

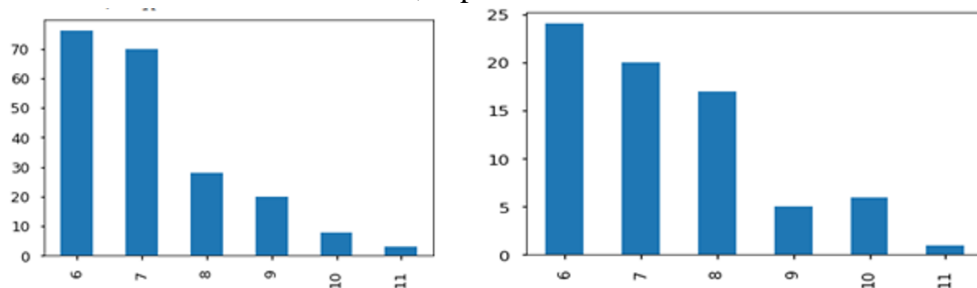
Com os dados acima obtidos, viu-se que a evasão do curso de Gestão Empresarial ficou em 34,38% e de Análise e Desenvolvimento de Sistemas em 40,57%, desconsiderando os

possíveis cancelamentos que ainda podem ocorrer dos alunos classificados como “Em Curso”.

Buscou-se saber então o comportamento dos alunos que concluíram os cursos, considerando o semestre de formação. Na Figura 2 é possível notar que a maior parte dos alunos de GE se formam entre o 6º e o 7º semestres. Como os cursos da FATEC Indaiatuba possuem 6 semestres, esse resultado mostra que grande parte desses alunos se forma permanecendo geralmente apenas mais um semestre para conclusão do curso. Como contraponto, há o curso de ADS, em que a maior parte de alunos formados se encontra entre o 6º e 8º semestres. Vê-se então que esses últimos alunos demandam mais recursos da faculdade, uma vez que demoram mais tempo para concluírem o curso.

Nota-se que, para ambos cursos analisados, há alunos que concluíram o curso em até 11 semestres e em menos de 6 semestres. Os primeiros podem ser explicados por trancamentos realizados, em que posteriormente o aluno voltou ao curso e o concluiu. Já os últimos são explicados por alunos provenientes de transferência, já que, provavelmente, houve aproveitamento de matérias para conclusão do curso.

Figura 2: Tempo para conclusão de curso dos alunos dos cursos de GE e ADS, respectivamente



Fonte: Elaborado pelos autores

3.3 Escola Pública e Particular

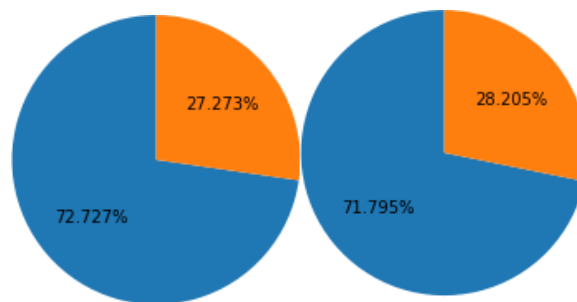
Uma vez que se obteve os índices de evasão dos alunos e seu comportamento para conclusão do curso, decidiu-se então por comparar a esfera da escola de ensino médio da qual tais alunos eram provenientes. As comparações foram feitas buscando correlações entre a proporção de alunos que ingressavam na FATEC Indaiatuba e destes que concluíam o seu respectivo curso, sendo que parte vinha de escolas públicas e outra parte de escolas particulares.

As análises foram feitas por meio do agrupamento de alunos pelo RA, ESCOLA_PUBLICA e STATUS_ALUNO e a contagem feita pelo número de linhas obtidas

e agrupadas para a coluna ESCOLA_PUBLICA. No total, o curso de Análise e Desenvolvimento de Sistemas teve cerca de 74% de seus alunos ingressantes vindos de escola pública (azul), sendo que os 26% restantes provêm de escolas particulares (laranja). Quando se olha para os alunos que concluíram o curso, nota-se que a maior desistência vem dos alunos de escolas particulares, que passam à proporção de cerca de 22% dos concluintes.

Em uma análise mais profunda, fez-se o mesmo experimento apenas com os alunos que concluíram o curso de ADS que estavam abaixo da mediana nas notas do vestibular. Tal resultado é apresentado pela Figura 3. Dessa vez, surpreendentemente a proporção entre alunos de escola pública (azul) e particular (laranja) praticamente não se modifica, permanecendo com cerca de 72% e 28%, respectivamente. Esse resultado sugere que, uma vez que um aluno com maiores dificuldades didáticas passa no vestibular e estuda para permanecer no curso até sua conclusão, o fator Escola Pública x Escola Particular não influencia na sua capacidade de concluir o curso, ou seja, seu *background* não se faz mais relevante do que os estudos no decorrer de sua formação universitária.

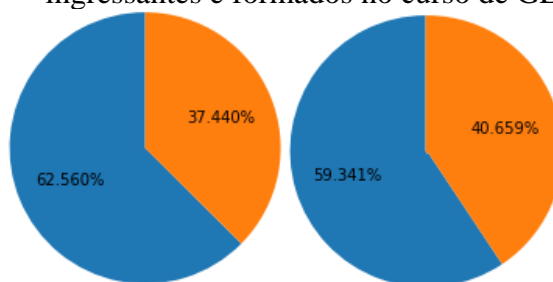
Figura 3: Porcentagem de alunos abaixo da mediana nas notas do vestibular, provenientes de escola pública e escola particular ingressantes e formados no curso de ADS



Fonte: Elaborado pelos autores

Em seguida, trabalhou-se no mesmo tipo de comparação para os alunos do curso de Gestão Empresarial. Primeiramente, a comparação foi feita considerando o total de alunos ingressantes neste curso e do total de alunos que o concluíram. Dessa vez, nota-se por meio da Figura 4 que, em relação ao curso de ADS, GE possui mais alunos provenientes de escola particular, sendo estes cerca de 37% do total de alunos ingressantes. Quando se olha para os alunos que concluíram esse curso, vê-se que há uma desistência maior dos alunos vindos de escola pública, já que os alunos de escola particular passam a ser quase 41% do total.

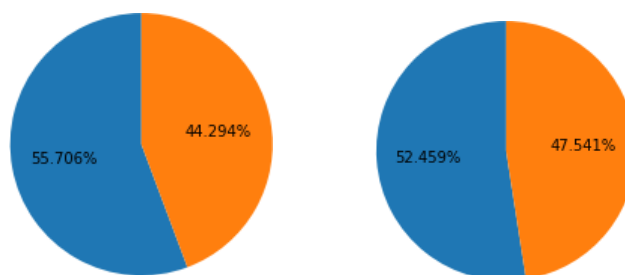
Figura 4: Porcentagem de alunos provenientes de escola pública e escola particular ingressantes e formados no curso de GE



Fonte: Elaborado pelos autores

Observa-se que o mesmo ocorre quando considera-se alunos que foram aprovados no vestibular com notas abaixo da mediana para este curso. A Figura 5 mostra que a proporção de alunos de escola particular (laranja) com desempenho inferior no vestibular aumenta para cerca de 44% entre os ingressantes e 47% entre os que concluíram o curso. Com este resultado, conclui-se que, ao contrário do esperado, o desempenho de alunos de escola particular no vestibular não é maior do que o desempenho dos alunos de escola pública (azul), porém esses são os que mais permanecem no curso até sua conclusão. Desse modo, vê-se que é importante focar recursos da instituição de ensino para permanência de alunos provenientes de escolas públicas até o final do curso.

Figura 5: Porcentagem de alunos abaixo da mediana nas notas do vestibular, provenientes de escola pública e escola particular ingressantes e formados no curso de GE



Fonte: Elaborado pelos autores

3.4 Desistências

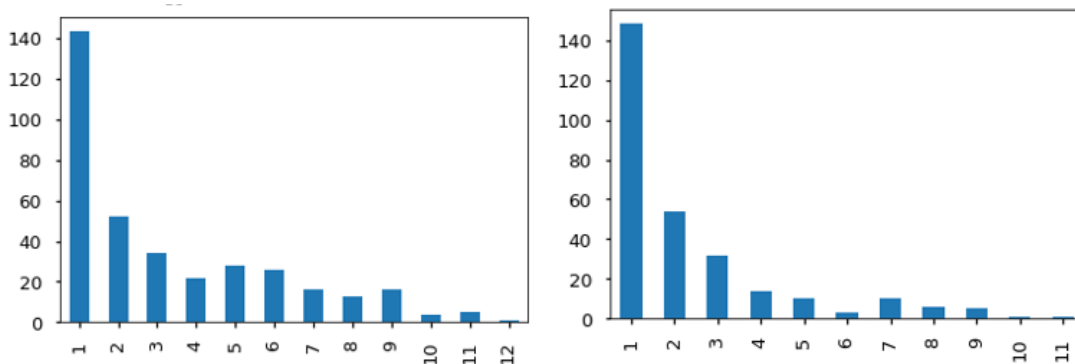
Quando se volta o olhar para as desistências dos cursos de Gestão Empresarial e Análise e Desenvolvimento de Sistemas, vê-se que é importante saber quando essa grande quantidade de alunos decide por abandonar o curso e seguir novos caminhos. Buscando por tais respostas, decidiu-se filtrar os dados a fim de obter a quantidade de alunos desistentes por semestre, contando “1” como o primeiro semestre de aulas do aluno, “2” o segundo semestre

e assim sequencialmente. Nota-se por meio da Figura 6 que GE apresentou alunos desistentes até o 12º semestre, ou seja, houve casos em que o aluno ficou vinculado à faculdade por 6 anos e acabou por abandonar o curso. Já para ADS obteve-se comportamento semelhante nesse quesito, com alunos desistindo até o 11º semestre.

Em análise mais minuciosa, observa-se que a maior parte dos alunos que optam por abandonar o curso tomam essa decisão já no 1º semestre de aulas. Isso pode ocorrer por diversos fatores, tais como oportunidades em outras universidades, decepção com a escolha feita quanto ao curso, decepção quanto ao conteúdo do curso/professores, dificuldade no acompanhamento das aulas desde o primeiro momento, falta de recursos ou apoio para seguir com os estudos, dentre outros. Uma vez que isso foi observado, faz-se necessário descobrir os motivos principais de desistência e atuar como instituição para coibir a desistência desde o início do aluno na faculdade.

Uma vez que analisou-se o primeiro semestre de ambos os cursos separadamente, vemos que o restante das desistências se distribui pelos demais semestres e que o comportamento se dá pela diminuição de abandonos semestre a semestre.

Figura 6: Alunos desistentes por semestre dos cursos de GE e ADS, respectivamente



Fonte: Elaborado pelos autores

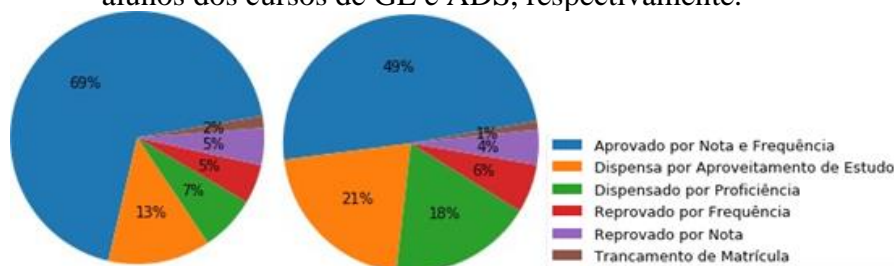
3.4 Índices de Aprovação (Humanas, Exatas e Tecnológicas)

Reunindo todos os dados de desistências e evasão escolar obtidos até o momento, definiu-se que o próximo e último passo seria descobrir se as disciplinas de alguma área específica estariam contribuindo para esses casos. Para isso, distribuiu-se as matérias de cada curso em três grandes áreas: Humanas, Exatas e Tecnológicas.

Os gráficos das figuras 7, 8 e 9, a seguir, mostram o comportamento das aprovações e reprovações nas disciplinas quando compara-se os cursos de Gestão Empresarial e Análise e

Desenvolvimento de Sistemas. Intuitivamente, espera-se que o curso da área de Humanas, como Gestão Empresarial, tenha maior afinidade com a área de Humanas, bem como Análise e Desenvolvimento de Sistemas tenha maior propensão a ter um bom desempenho nas matérias Tecnológicas e Exatas.

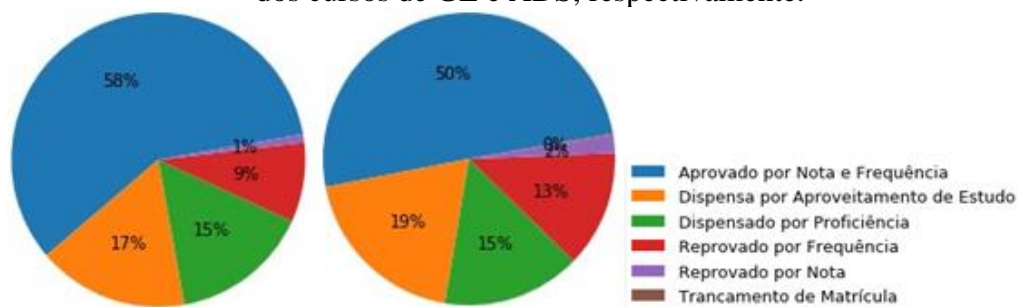
Figura 7: Índices de aprovação, dispensa e reprovação nas disciplinas de Humanas dos alunos dos cursos de GE e ADS, respectivamente.



Fonte: Elaborado pelos autores

Primeiramente, nota-se que GE possui melhor desempenho nas matérias denominadas como Humanas quando comparados com os alunos de ADS. Tem-se um índice de aprovação de 69% para GE frente a apenas 49% para alunos de ADS. Desse modo, sugere-se que alunos de ADS necessitam de maior empenho e cuidado durante o desenvolvimento de tais disciplinas. Ao mesmo tempo, a instituição de ensino tem a oportunidade de criar alternativas que tenham como objetivo trazer o aluno de ADS de maneira mais efetiva para os conteúdos abordados em disciplinas de Humanas. Ressalta-se entretanto que, se for considerada a porcentagem de dispensados por aproveitamento de estudos e por proficiência, o índice de aprovação nos dois cursos fica bastante próximo, 89% para GE e 88% para ADS. De qualquer forma, acredita-se que é necessária a criação de mecanismos de melhor aproveitamento das matérias da área de humanas aos alunos de ADS que efetivamente cursam essas disciplinas durante a realização do curso.

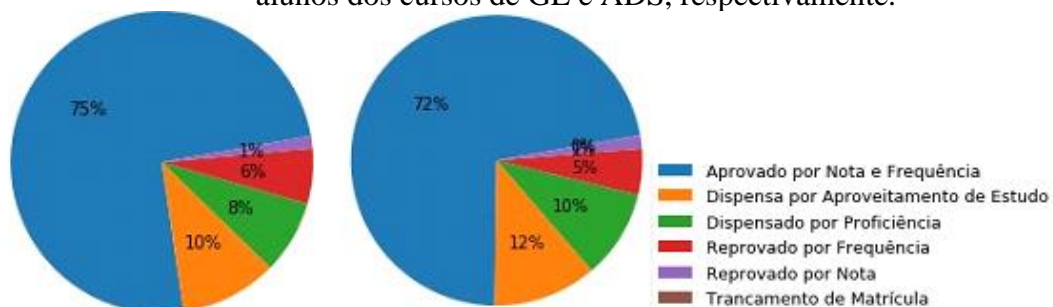
Figura 8: Índices de aprovação, dispensa e reprovação nas disciplinas de Exatas dos alunos dos cursos de GE e ADS, respectivamente.



Fonte: Elaborado pelos autores

Já para matérias denominadas como Exatas, os alunos de GE possuem desempenho também melhor quando comparados com os alunos de ADS, porém com comportamento mais próximo. Dessa vez, o índice de aprovação é de 58% para GE contra 50% para alunos de ADS. Mais uma vez vê-se que alunos de ADS precisam de um acompanhamento maior durante os estudos de disciplinas desta categoria, assim é necessário trabalhar mais a base de conhecimento dos alunos e fornecer melhores condições para monitoria e acompanhamento discente. Para as matérias de Exatas, ao se considerar no total de aprovados as dispensas por aproveitamento de estudos e por proficiência, a diferença do índice de aprovação entre os dois cursos se mantém similar, 90% para GE e 84% para ADS.

Figura 9: Índices de aprovação, dispensa e reprovação nas disciplinas de Tecnológicas dos alunos dos cursos de GE e ADS, respectivamente.



Fonte: Elaborado pelos autores

Por fim, tem-se as matérias denominadas como “Tecnológicas”, que mostram índice de aprovação elevado para ambos cursos. O índice de aprovação se apresenta em 75% para GE e 72% para alunos de ADS, desconsiderando alunos já aprovados por dispensa por proficiência ou aproveitamento. Nota-se então que as matérias dessa categoria provavelmente não impactam significativamente na desistência de alunos, quando comparadas com

“Humanas” ou “Exatas”.

Assim, as ações a serem tomadas pela instituição de ensino devem buscar para ambos os cursos um enfoque maior em causas dos índices de reprovadas e soluções que amenizem isso frente às disciplinas de “Humanas” e “Exatas”, sendo que ADS demonstra maior dificuldade em ambas as áreas. Ao mesmo tempo, retomando os resultados obtidos com as notas dos vestibulares, vê-se que os alunos com melhor desempenho no vestibular (ADS) não são aqueles que apresentam melhor desempenho nas disciplinas do curso. Os motivos a serem analisados, dentre outros, passam pelo vestibular não estar selecionando adequadamente alunos para este curso até disciplinas dos cursos de ADS serem relativamente mais difíceis do que as do curso de GE.

5. CONCLUSÃO

O comparativo entre os cursos de Análise e Desenvolvimento de Sistemas e Gestão Empresarial da FATEC Indaiatuba trouxe resultados importantes para ajudar a instituição no planejamento futuro visando a possíveis ações de melhoria educacional que possam evitar a evasão do aluno, principalmente nos primeiros períodos dos cursos.

Ao trabalhar-se as notas de vestibular dos alunos de ambos os cursos, nota-se que o desempenho dos melhores alunos no vestibular não se mantém após a entrada na instituição. Adicionalmente, alunos oriundos de escola pública, concluem em maior proporção o curso de ADS, quando comparados aos alunos de escola particular, sendo que o inverso ocorre para o curso de GE.

Explorando o tempo para formação dos alunos e os semestres com maior desistência, notou-se que predominantemente se desiste dos cursos logo nos primeiros semestres, o que não dá tempo hábil à instituição comprovar a importância da formação acadêmica para a vida pessoal e profissional. Os motivos para evasão podem ser estudados em trabalhos posteriores, porém afirma-se que, independentemente de quais sejam, devem ser trabalhados e evitados desde o princípio.

Por fim, completando o objetivo deste trabalho e afirmando a hipótese inicial do projeto, viu-se que os alunos de ADS possuem maiores dificuldades em todas as grandes classes de disciplinas (Humanas, Exatas e Tecnológicas), sendo que as matérias de Humanas são as que mais contribuem com a reprovação do aluno. Dessa maneira, foram encontrados pontos em que a instituição de ensino pode focar seus esforços visando a uma maior quantidade de alunos formados e que um menor número de desistências ocorra.

Trabalhos futuros devem ser desenvolvidos para encontrar quais são as dificuldades que a instituição de ensino superior tem que mais afetam a permanência dos alunos nos cursos, sendo aconselhado o *benchmark* com outras FATECs e faculdades públicas que ofereçam cursos de tecnologia. Assim, espera-se que os recursos possam ser melhor alocados e que a satisfação da comunidade seja maior com os serviços prestados e disponibilizados à população

REFERÊNCIAS

GIL, A. C. **Como elaborar projetos de pesquisa**. 6 ed. São Paulo. Atlas, 2017.

HU, Q., POLYZOU A., KARYPIS, G., RANGWALA, H. Enriching Course-Specific Regression Models with Content Features for Grade Prediction. In: **ICDSA, International Conference on Data Science and Advanced Analytics**, Fairfax/Minneapolis, 2017. p. 504-513.

HU, Z. Research on Big Data Application in Smart Education System. In: **Proceedings of the 2016 4Th International Conference on Sensors, Mechatronics and Automation (Icsma 2016)**, v. 136, p. 306–309, 2016.

KOCHETKOV, O. T., PROKHOROV, I. V. **The reseach of approaches of applying the results of big data analysis in higher education**. Conference Proceedings. v. 1797, n. 020008, 2017.

LARSON, D., CHANG, V. **A review and future direction of agile, business intelligence, analytics and data science**. *Jornal International Journal of Information Management*. v. 36, p. 700-710, 2016.

LYNCH, C. F. **Who prophets from big data in education? New insights and new challenges**. *Theory and Research in Education*. v. 15, n. 3, p. 249–271, 2017.

MARQUESONE, Rosangela. **Big data: técnicas e tecnologias para extração de valor de dados**. [S. l.]: Casa do Código, 2016. 245 p.

O'NEIL, C., SCHUTT, R. **Doing Data Science: Straight Talk from the Frontline**. 1 ed. California. O'Reilly, 2013.

SIN, K. MUTHU, L. Application of Big Data in education data mining and learning analytics: a literature review. **ICTACT journal on soft computing**, v. 5, n. 4, 2015.

VYAS, M. S., GULWANI, R. Predicting Student's Performance using CART approach in Data Science. In: **ICECA 2017, International Conference on Electronics, Communication and Aerospace Technology**, Mumbai, 2017. p. 58-61

WASSAN J. T. **Discovering Big Data Modelling for Educational World**. *Procedia - Social and Behavioral Sciences*. v. 176, p. 642 – 64, 2015.

WEST, D. M. **Big Data for education: Data mining, data analytics, and web dashboards**. *Governance Studies at Brookings*. v. 9, p. 1 – 10, 2012.