

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ROBSON DOUGLAS DA SILVA MARTINS

Learning Analytics no apoio, planejamento e avaliação de
metodologias ativas no Ensino de Física

São Carlos
2023

ROBSON DOUGLAS DA SILVA MARTINS

Learning Analytics no apoio, planejamento e avaliação de metodologias ativas no Ensino de Física

Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

Área de concentração: Física Biomolecular
Orientador: Prof. Dr. Sérgio Ricardo Muniz

Versão Corrigida
(versão original disponível na Unidade que aloja o Programa)

São Carlos
2023

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Martins, Robson Douglas da Silva

Learning Analytics no apoio, planejamento e avaliação de metodologias ativas no Ensino de Física / Robson Douglas da Silva Martins; orientador Sergio Ricardo Muniz - versão corrigida -- São Carlos, 2023.

84 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2023.

1. Learning analytics. 2. Ensino de física. 3. Metodologias ativas. 4. Data science na educação. I. Muniz, Sergio Ricardo, orient. II. Título.

À minha querida mãe **Ivone**, que abdicou de tudo para eu ser quem eu sou e seguir meus sonhos, ensinando o verdadeiro sentido de resiliência, respeito, caráter e amor incondicional.

Ao meu pai **Rubens** (*in memoriam*) que por pouco tempo, mas que muito me ensinou.

À minha esposa **Marcela** por estar sempre me apoiando, me desafiando a ser melhor cada dia e não me deixar perecer diante as dificuldades.

Aos meus irmãos **Rewrison** e **Rodolfo** que sempre me ajudaram, cuidaram e acreditaram em mim.

AGRADECIMENTOS

A **Deus** pela eterna companhia, guiando, protegendo e iluminando meu caminho, sempre me trazendo e mostrando motivos pelos quais vale a pena todo o esforço diário.

Ao professor **Sérgio Ricardo Muniz** por toda a contribuição e paciência no desenvolvimento deste trabalho e pelo exemplo de docência nos anos de graduação.

À **Marcela Fernandes de Lima Martins** pela ajuda com os textos lidos e dedicação a “nós”.

Ao **Instituto de Física de São Carlos (IFSC)** por me proporcionar o melhor ensino de física.

Aos amigos de faculdade e de vida **Geisiane Rosa da Silva, Daniel Cosmo Pizetta, Beatriz Saiani Bellini** por toda a ajuda e parceria nesse e em outras dezenas de projetos.

Ao amigo “Compadre” **Douglas Henrique Milanez** pela ajuda e orientações para o trabalho.

À **Comissão de Pós-Graduação do Instituto de Física de São Carlos (CPG-IFSC), (professores e funcionários)** por todo auxílio e ajuda durante o desenvolvimento deste trabalho.

Aos **funcionários da biblioteca do Instituto de Física de São Carlos (IFSC)** por todo o auxílio na correção deste trabalho.

***“Ando devagar
Porque já tive pressa
E levo esse sorriso
Porque já chorei demais
Hoje me sinto mais forte
Mais feliz, quem sabe
Só levo a certeza
De que muito pouco sei
Ou nada sei”***

Almir Sater, Renato Teixeira, 1990

RESUMO

MARTINS, R. D. S. **Learning analytics no apoio, planejamento e avaliação de metodologias ativas no ensino de física.** 2023. 84p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

Neste trabalho, foi desenvolvida uma análise exploratória de dados relativos às atividades e o desempenho de estudantes do curso de física básica matriculados em cursos de bacharelado do campus da USP de São Carlos. O objetivo da análise foi identificar padrões de correlação associados à aprovação na disciplina que possam auxiliar docentes nas tomadas de decisão em relação ao desenvolvimento da disciplina e no processo de retomada e/ou recuperação dos conteúdos. Foram analisados dados de nove turmas, com um total de 565 estudantes, coletados usando um ambiente virtual de aprendizagem (AVA) e analisados (de forma anonimizada) através de scripts em linguagem Python, com auxílio de ferramentas e técnicas matemáticas, computacionais e estatísticas que formam a base das técnicas utilizadas em *Learning Analytics* e *Data Science* aplicados a educação. Foram encontradas correlações positivas entre as notas de provas e aprovação na disciplina bem como poucas correlações entre a presença nas aulas e a nota final. Observou-se também uma relativa evasão de estudantes durante o período letivo, levantando questionamentos para estudos futuro a respeito do estímulo que possa ter levado a tal evasão.

Palavras-chave: Learning analytics. Ensino de física. Metodologias ativas. Data science na educação.

ABSTRACT

MARTINS, R. D. S. **Learning analytics in the support, planning and evaluation of active methodologies in physics teaching.** 2023. 84p. Dissertation (Master in Science) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2023.

In this work, exploratory data analysis was conducted on the activities and performance of students enrolled in introductory physics courses at the São Carlos campus of the University of São Paulo (USP). The goal of the analysis was to identify correlation patterns associated with course approval that could assist professors in decision-making regarding the course's development and the process of resuming and/or recovering educational content. In this study, data from nine classes, comprising a total of 565 students, were collected using a virtual learning environment (VLE) and analyzed (anonymized) using Python scripts, aided by mathematical, computational, and statistical tools and techniques that form the foundation of Learning Analytics and Data Science applied to education. Positive correlations were found between exam scores and course approval, while only a small correlation was observed between class attendance and final grade. A relative dropout rate of students was also observed during the academic period, prompting questions for future studies regarding the factors that may have contributed to such dropout.

Keywords: Learning analytics. Physics teaching. Active methodologies. Educacional data science.

LISTA DE FIGURAS

Figura 1 -	Crescimento do número de programas de Pós-Graduação na Área de Ensino (2000-2017)	20
Figura 2-	os 4 V's do Big Data	28
Figura 3 -	Abordagens tradicionais e de Big Data em pesquisa	29
Figura 4 -	Áreas constituintes do Data Science na Educação	30
Figura 5 -	tabela com Instituições e ferramentas de análise de aprendizado	31
Figura 6 -	Divisão de atividades por metodologias	34
Figura 7 -	Gráfico de relação de aprovados e não aprovados na disciplina	39
Figura 8 -	Histograma de distribuição das notas finais na disciplina	40
Figura 9 -	BoxPlot da nota final do curso	40
Figura 10 -	Gráfico de barras com distribuição das notas da prova 1 (P1)	41
Figura 11 –	BoxPlot das notas da prova 1 (P1)	41
Figura 12 -	Gráfico de barras com distribuição das notas da prova 2 (P2)	42
Figura 13 -	BoxPlot das notas da prova 2 (P2)	42
Figura 14 -	Gráfico de barras com distribuição das notas da prova 3 (P3)	43
Figura 15 -	BoxPlot das notas da prova 3 (P3)	43
Figura 16 -	Gráfico de barras com distribuição da quantidade de notas das prova 1, 2 e 3	44
Figura 17 -	Gráfico de barras com distribuição da quantidade de provas realizadas	44
Figura 18 -	BoxPlot com notas finais das provas	45
Figura 19 -	Gráfico de barras com distribuição das notas de ADS 1	46
Figura 20 -	Gráfico de barras com distribuição das notas de ADS2	46
Figura 21 -	Gráfico de barras com distribuição das notas de ADS3	47
Figura 22 -	Gráficos de barras com distribuição das notas de ADS4	47
Figura 23 -	Gráfico de barra com distribuição das notas de ADS5	47
Figura 24 -	Gráfico de Barras com distribuição das notas de ADS6	48

Figura 25 -	Gráfico de barras com distribuição das notas finais de ADS	48
Figura 26 -	Gráfico de barras com número de ADS realizadas	49
Figura 27 -	Gráfico de barras com distribuição das notas de TEC1	49
Figura 28 -	Gráfico de barras com distribuição das notas de TEC 2	50
Figura 29 -	Gráfico de barras com distribuição das notas de TEC 3	50
Figura 30 -	Gráfico de barras com distribuição das notas de TEC 4	50
Figura 31 -	Gráfico de barras com distribuição das notas de TEC 5	51
Figura 32 -	Gráfico de barras com distribuição das notas de TEC 6	51
Figura 33 -	Gráfico de barras com distribuição das notas finais de TEC	51
Figura 34 -	Gráfico de barras com número de TEC realizadas	52
Figura 35 -	Gráfico de barras com distribuição das notas finais das atividades online	52
Figura 36 -	Gráfico de barras com distribuição das notas finais das atividades totais	52
Figura 37 -	HeatMap com a correlação entre as principais atividades realizadas	53
Figura 38 -	HeatMap com correlação entre principais atividades para grupo com notas finais de 0 a 3	54
Figura 39 -	HeatMap com correlação entre principais atividades para grupo com notas finais de 3 a 5	54
Figura 40 -	HeatMap com correlação entre principais atividades para grupo com notas finais de 5 a 7	55
Figura 41 -	HeatMap com correlação entre principais atividades para grupo com notas finais de 7 a 10	55
Figura 42 -	Gráfico de dispersão entre nota da prova 1 (P1) e média final do curso.	56
Figura 43 -	Gráfico de dispersão entre nota da prova 1 e média final do curso de 0 a 3.	56
Figura 44 -	Gráfico de dispersão entre nota da prova 1 e média final do curso de 3 a 5.	57
Figura 45 -	Gráfico de dispersão entre nota da prova 1 e média final do curso de 5 a 7.	57
Figura 46 -	Gráfico de dispersão entre nota da prova 1 e média final do curso de 7 a 10.	58
Figura 47 -	Gráfico de dispersão entre nota da prova 2 (P2) e média final do curso.	58
Figura 48 -	Gráfico de dispersão entre nota da prova 2 e média final do curso de 0 a 3.	59
Figura 49 -	Gráfico de dispersão entre nota da prova 2 e média final do curso de 3 a 5.	59

Figura 50 -	Gráfico de dispersão entre nota da prova 2 e média final do curso de 5 a 7.	60
Figura 51 -	Gráfico de dispersão entre nota da prova 2 e média final do curso de 7 a 10.	60
Figura 52 -	Gráfico de dispersão entre nota da prova 3 e média final do curso.	61
Figura 53 -	Gráfico de dispersão entre nota da prova 3 e média final do curso de 0 a 3.	61
Figura 54 -	Gráfico de dispersão entre nota da prova 3 e média final do curso de 3 a 5.	62
Figura 55 -	Gráfico de dispersão entre nota da prova 3 e média final do curso de 5 a 7.	62
Figura 56 -	Gráfico de dispersão entre nota da prova 3 e média final do curso de 7 a 10.	63
Figura 57-	Gráfico de dispersão entre nota total de TEC e média final do curso	63
Figura 58 -	Gráfico de dispersão entre nota de TEC e média final do curso de 0 a 3	64
Figura 59 -	Gráfico de dispersão entre nota de TEC e média final do curso de 3 a 5	64
Figura 60 -	Gráfico de dispersão entre nota de TEC e média final do curso de 5 a 7	65
Figura 61 -	Gráfico de dispersão entre nota de TEC e média final do curso de 7 a 10	65
Figura 62 -	Gráfico de dispersão entre nota total de ADS e média final do curso	66
Figura 63 -	Gráfico de dispersão entre nota de ADS e média final do curso de 0 a 3	66
Figura 64 -	Gráfico de dispersão entre nota de ADS e média final do curso de 3 a 5	67
Figura 65 -	Gráfico de dispersão entre nota de ADS e média final do curso de 5 a 7	67
Figura 66 -	Gráfico de dispersão entre nota de ADS e média final do curso de 7 a 10	68
Figura 67 -	Gráfico de dispersão entre nota de Atv Online e média final do curso	68
Figura 68 -	Gráfico de dispersão entre nota de Atv Online e média final do curso de 0 a 3	69
Figura 69 -	Gráfico de dispersão entre nota de Atv Online e média final do curso de 3 a 5	69
Figura 70 -	Gráfico de dispersão entre nota de Atv Online e média final do curso de 5 a 7	70
Figura 71 -	Gráfico de dispersão entre nota de Atv Online e média final do curso de 7 a 10	70
Figura 72 -	Gráfico de dispersão entre nota de Atv Total e média final do curso	71
Figura 73 -	Gráfico de dispersão entre nota de Atv Total e média final do curso de 0 a 3	71
Figura 74 -	Gráfico de dispersão entre nota de Atv Total e média final do curso de 3 a 5	72
Figura 75 -	Gráfico de dispersão entre nota de Atv Total e média final do curso de 5 a 7	72
Figura 76 -	Gráfico de dispersão entre nota de Atv Total e média final do curso de 7 a 10	73

Figura 77 -	Gráficos de dispersão entre notas Total de Provas, ADS, TEC, Atividade online, Atividade total e nota final de curso	74
Figura 78 -	HeatMap entre as notas obtidas em Provas (maior peso) e média final em relação as presenças em ADS e presença total	75
Figura 79 -	Gráficos de dispersão entre notas Total de Provas, Presença em ADS e presença total no curso.	76

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Documentos oficiais e o Ensino de Física	19
1.2	Objetivo	21
2	CONTEXTO E REFERENCIAL TEÓRICO	23
2.1	Reformulação do currículo	23
2.2	Metodologias Ativas	24
2.3	Moodle – Ambiente Virtual de Aprendizagem (AVA)	25
2.4	Big Data, Data Science e Learning Analytics	26
2.5	Big Data	26
2.6	Data Science (Ciência de Dados)	28
2.7	Learning Analytics (Análítica de aprendizagem)	30
2.8	Big Data e Data Science na Educação	31
3	METODOLOGIA	33
3.1	Método para coleta de dados e a reestruturação curricular	33
3.2	Descrição do Dataset	37
4	RESULTADOS	39
5	CONCLUSÕES	77
	REFERÊNCIAS	83

1 INTRODUÇÃO

1.1 Documentos oficiais e o Ensino de Física

As Diretrizes Nacionais Curriculares (DNC) norteiam o currículo base dos cursos de graduação e estabelecem um conjunto de habilidades e competências desejadas nos egressos. Dessa forma, pretende-se uma estrutura base para a formação dos projetos pedagógicos dos cursos de graduação de cada Instituições de Ensino Superior (IES), sem destituir a autonomia das instituições neste processo. No caso específico dos cursos de bacharelado e licenciatura em Física, as DNC¹ indicam uma série de perfis formativos que devem estar presentes nos programas de graduação das IES. O documento apresenta e separa os perfis em quatro categorias: (I) Físico pesquisador (bacharelado em Física), (II) Físico educador (Licenciatura em Física), (III) Físico tecnólogo (Bacharelado ou Licenciatura em Física e Associada) e (IV) Físico interdisciplinar (Bacharelado em Física Aplicada). Além dos perfis citados, as Diretrizes Nacionais Curriculares também indicam algumas competências e habilidades, gerais e específicas, sendo as específicas elaboradas pelas IES com objetivo de atender às exigências do mercado, as quais compõem a formação do profissional graduado na área de Física.

Assim como no ensino superior, o ensino de física do núcleo básico (Ensino Médio), também segue parâmetros bem estruturados. Neste caso, os Parâmetros Curriculares Nacionais para o Ensino Médio – PCNEM² indicam as habilidades e competências necessárias para a formação do discente, às quais são divididas em três categorias: (I) Representação e Comunicação, (II) Investigação e Compreensão e (III) Contextualização sociocultural.

Embora as categorias norteadoras, tanto para o Ensino Superior como para o Ensino Básico, sejam bem estruturadas e pautadas em documentos, a atividade disciplinar em si, não é uma ciência exata, sendo de grande caráter fluido e dinâmico, não importando a área do conhecimento. Portanto, entende-se o Ensino como área de pesquisa translacional, pois ela faz a ponte entre o conhecimento gerado a partir da pesquisa teórica e experimental e suas aplicações no contexto da Educação Básica, Profissional, Tecnológica ou Superior.³

Têm-se na área de Ensino (área 46 do relatório), um número significativo de pesquisas e materiais desenvolvidos, com o objetivo de se maximizar a qualidade de ensino e nível de aprendizagem dos alunos. O relatório de avaliação quadrienal de 2017 da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), vinculada ao Ministério da Educação do Brasil, informou que, entre 2013 e 2016, foram publicados cerca de 13.100 artigos científicos

em periódicos qualificados entre A1 e B5. Esse montante relatado é consequência, pelo menos em parte, do expressivo crescimento do número de programas de pós-graduação na área de Ensino (área 46) entre 2000 e 2017, conforme apresentado na Figura 1.

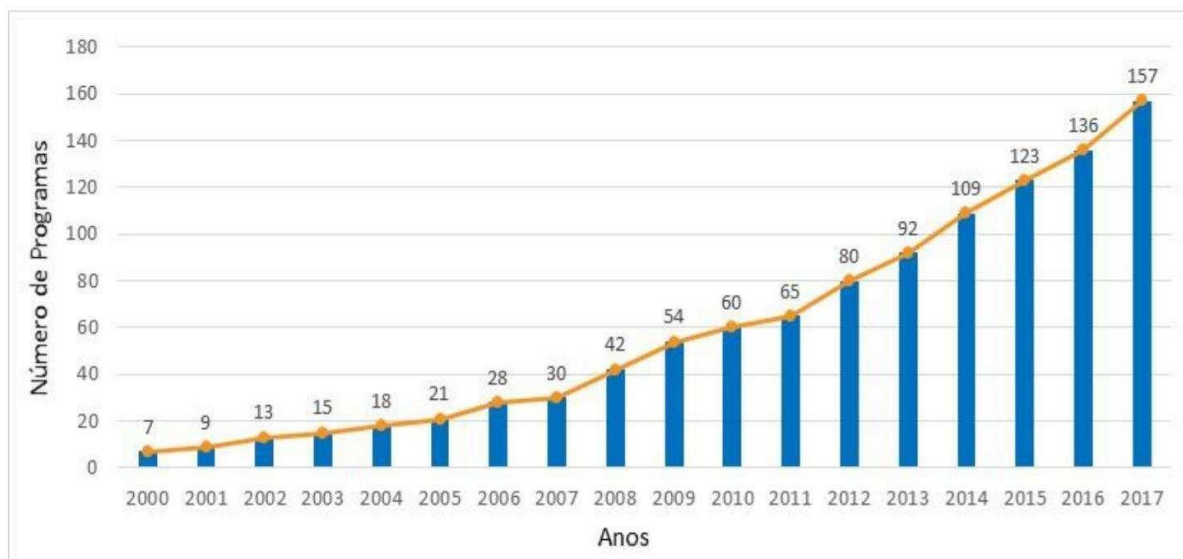


Figura 1 - Crescimento do número de programas de Pós-Graduação na Área de Ensino (2000-2017)

Fonte: Gráfico retirado do relatório de avaliação quadrienal de 2017 da CAPES³

Apesar de um extenso acervo, o processo de ensino-aprendizagem não é exato e unidirecional. Ele pode ser definido como um “complexo sistema de interações comportamentais entre professores e alunos.”⁴

Neste sentido, extrapola o simples processo de ensinar (por parte do docente) e aprender (por parte do discente), pois inclui as relações humanas existentes e criadas durante os variados aspectos pedagógicos e pessoais no ambiente educacional.

Embora são obtidos resultados positivos através de algumas estratégias, podendo ser aulas expositivas, práticas, interativas, utilizando mídias digitais etc., a reprodução e replicação exatas dessas metodologias em diferentes grupos não é garantia de sucesso de aprendizado, pelo contrário, a simples reprodução de uma estratégia em um novo ambiente ou público atinge de forma diferenciada e, muitas vezes, na contramão do resultado esperado. Assim, o comportamento intangível e dinâmico do processo de ensino-aprendizagem.

Sabemos, então, que a aplicação de uma ou outra metodologia de ensino não garante a eficácia do processo de ensino-aprendizagem no longo prazo. Diante deste cenário, Freire afirma, “Não há ensino sem pesquisa e pesquisa sem ensino”⁵, e desta forma, este é o ponto motivador chave deste trabalho, a compreensão do processo de ensino e aprendizagem,

buscando identificar como este processo de validação ocorre para cada discente, dentro de sua própria esfera educacional.

1.2 Objetivo

A presente pesquisa tem por objetivo identificar e avaliar qualitativa e quantitativamente a relação entre a estratégia de ensino-aprendizagem e o desempenho de discentes em um novo sistema de aprendizagem implementado no Instituto de Física de São Carlos, combinando um ambiente virtual de aprendizagem com atividades desenvolvidas em sala de aula, numa disciplina presencial de Física básica oferecida a diferentes turmas e cursos do campus da USP de São Carlos. Para tanto, estabeleceu-se os seguintes objetivos específicos:

- Revisar os aspectos teóricos e metodológicos relacionados ao binário ensino-aprendizagem para a área de Física;
- Realizar uma análise exploratória dos dados através de métodos e ferramentas de *Learning Analytics*;
- Identificar, se existente, padrões na relação entre o desempenho e o comportamento do discente na disciplina.
- Discutir a relevância do uso dessas ferramentas de análise objetiva no planejamento, diagnóstico e melhoria das disciplinas de graduação da USP.

2 CONTEXTO E REFERENCIAL TEÓRICO

2.1 Reformulação do currículo

O processo de ensino-aprendizagem, não apenas no Brasil, passou e passa por diversas mudanças ao longo do tempo, sejam elas no formato como também nas metodologias e ferramentas.

Há muito o ensino é realizado no formato expositivo, também conhecido como metodologia tradicional, em que o professor é o detentor do conhecimento e transmite aos discentes, os saberes necessários, e estes por sua vez, são meros expectadores do processo. Com o surgimento de novas tecnologias (tecnologias ativas), diversas novas ferramentas e metodologias também têm surgido e os docentes veem buscando essa inovação no seu dia a dia, com o objetivo de maximizar tanto a qualidade das aulas quanto atingir um maior engajamento dos alunos, percebendo que estes interagem de diferentes formas em metodologias e ferramentas distintas. “O trabalho docente caracteriza-se pela ação e interação entre sujeitos com vistas à construção de saberes na e para a realidade. A docência atual está inserida em mudanças significativas com a implementação de novos recursos tecnológicos”.⁶

Neste contexto, o Instituto de Física de São Carlos – IFSC, buscando adequar as disciplinas de seus cursos às metodologias inovadoras e atualizadas dentro da área de pesquisa em ensino, propôs um estudo de reformulação do currículo base dos cursos de graduação.

Após o estudo, realizado por parte de um grupo de docentes do Instituto, para a mudança da grade curricular, uma das mudanças mais significativa foi realizada na carga horária dentro de sala de aula, que sofreu uma redução, com o objetivo de permitir que o aluno abandone o comportamento de simples espectador e torne-se protagonista da construção do próprio ensino. O objetivo desta mudança, é permitir que o aluno faça parte do processo de ensino-aprendizagem, assim, o docente passa a ter um papel fundamental de não apenas transmissor de conteúdo, mas também de um mediador do processo do discente.

Com este objetivo, o Instituto de Física de São Carlos reformulou as disciplinas do ciclo básico (primeiro ano de graduação), oferecidas para diversos cursos do campus de São Carlos, o que possibilitou, assim, uma maior utilização de novas metodologias, como *Peer Instruction* (Instrução por pares), através de atividades de discussão e/ou trabalho em grupo, resolução de situações problemas com caráter conceitual e prático, buscando uma melhor contextualização

do tema para o discente além também da metodologia *Flipped Classroom* (Sala de aula invertida), com atividades online pré e pós aula.

A metodologia de sala de aula invertida, permite um processo de caráter formativo além de avaliativo, pois as atividades online pré aula possibilitam ao aluno uma melhor preparação para as aulas presenciais, enquanto as atividades pós-aula viabilizam, além do feedback imediato, uma verificação dos conceitos trabalhados.

A ideia de se utilizar as metodologias citadas é acelerar o processo de diagnóstico, tanto por parte do docente quanto do discente, em relação ao tópico específico estudado, em contraste com as metodologias tradicionais, onde este processo é lento e tardio.

2.2 Metodologias Ativas

“Metodologias são grandes diretrizes que orientam os processos de ensino e aprendizagem e que se concretizam em estratégias, abordagens e técnicas concretas, específicas, diferenciadas”⁷

Método, podendo-se entender como “maneira de proceder”, e logia como “estudo”, podemos traduzir Metodologias, alicerçando também nas palavras de⁷, como diversos métodos de estudo, resultando em diversificadas estratégias e ferramentas focadas no aprendizado. Desta forma e no caso desta pesquisa, tem-se as Metodologias Ativas.

Embora não seja um conceito inovador, nos últimos anos as metodologias ativas vem sendo destaque em pesquisa. De maneira frequente, o conceito de metodologia ativa é reconhecido e aplicado como uma abordagem em sala de aula, onde as atividades têm como principal objetivo engajar os estudantes durante o processo de aprendizagem.⁸

A maior parte da literatura brasileira trata as metodologias ativas como estratégias pedagógicas que colocam o foco do processo de ensino e aprendizagem no aprendiz, contrastando com a abordagem pedagógica do ensino tradicional, centrada no professor, que transmite informação aos alunos.⁹

Metodologias ativas, em resumo, caracterizam-se como estratégias pautadas na participação efetiva do discente como protagonista durante a construção do seu processo de aprendizagem e desenvolvimento das diversas expectativas e habilidades trabalhadas.

Desse modo, as metodologias ativas focam em desenvolver situações de aprendizagem que possibilitem e incentivem os aprendizes a fazerem coisas, colocarem os conhecimentos em ação, pensarem e conceituarem o que e como fazem, construir os conhecimentos sobre os

conteúdos envolvidos além de possibilitar que desenvolvam estratégias cognitivas e lógicas, dentre outras.

Assim sendo, essas estratégias podem ser abordadas em diferentes formatos e sistemas, podendo ser mais flexíveis, híbridas, utilizando de variadas ferramentas, tecnológicas ou não.

Nos últimos 8 anos, atuando como professor do ensino básico, tenho evidenciado no dia a dia em sala de aula, que os alunos tendem a ter um maior engajamento quando diferentes metodologias e ferramentas são aplicadas ao longo do processo de ensino e aprendizagem e que um número significativo destes alunos demonstra um maior interesse, ou até mesmo um interesse antes não evidenciado, quando ferramentas digitais são utilizadas.

2.3 Moodle – Ambiente Virtual de Aprendizagem (AVA)

As metodologias relatadas anteriormente, foram implementadas com o auxílio de um Ambiente Virtual de Aprendizagem (AVA), neste caso o MOODLE (*Modular Object-Oriented Dynamic Learning Environment*), por já ser uma ferramenta utilizada por alguns docentes da Universidade.

O sistema permite que o docente possa planejar e programar diferentes atividades através de uma variada lista de ferramentas de interação (fórum, tarefa, questionário, blog, wiki, chat, glossário, pesquisa de avaliação, links de videoaulas, além da base de dados), possibilitando, através de metodologias ativas de aprendizagem, que o discente experimente, teste e explore cada atividade e recurso disponibilizado, assim, trabalhando as possíveis habilidades e expectativas buscadas na disciplina em toda sua capacidade.

Através da abordagem, neste caso, de metodologias ativas e abordagem construtivista, o Moodle maximiza o desenvolvimento e pensamento crítico, dinâmico e autônomo do aluno. Além das funcionalidades citadas, o ambiente também permite que o aluno acompanhe o seu desenvolvimento e progresso no curso de forma simples e objetiva, evidenciando o que já foi realizado e o que ainda faltar realizar. Para o docente, também é permitido acompanhar o progresso dos estudantes de forma individualizada, em grupos etc., através de relatórios gerados pela própria plataforma.

Estes relatórios, embora permitam que o docente obtenha diversas informações sobre o grupo, por si só, a visualização destas informações não é de simples e rápida interpretação, pois é gerado uma grande variedade de dados e informações de forma não estruturada, fazendo necessário o uso de técnicas e ferramentas específicas para a

limpeza, organização e análise dos dados. Tais técnicas e ferramentas são trabalhadas atualmente em diversas áreas de estudo, uma delas está relacionada à pesquisa em ensino, a *Learning Analytics*, que será objeto de estudo e ferramenta de análise neste trabalho.

2.4 Big Data, Data Science e Learning Analytics

Assim como ocorre com as atividades de comércio, entretenimento e comunicação, cada vez mais o processo educacional tem utilizado as Tecnologias de Informação e Comunicação (TIC) como uma plataforma de apoio ao processo de aprendizagem, e tem sido gerado uma grande quantidade de dados sobre esse processo.¹⁰

No sentido em que novas ferramentas são utilizadas e novos dados são gerados, têm-se a necessidade de utilizar novas ferramentas com capacidade de manipulação destes dados.

Essa capacidade, só tem sido possível pelo avanço científico e tecnológico, principalmente do setor de comunicação, especificadamente, na Internet e devido á isso, o surgimento do BIG DATA.

2.4.1 Big Data

O avanço da tecnologia e por consequência sua maior acessibilidade para a população, tem permitido que uma quantidade cada vez maior de dados seja gerada no mundo todo através da Internet. De acordo com,¹¹ nos últimos 20 anos, o conjunto de dados gerados e copiados ultrapassou a ordem de 10^{21} Bytes.

Todo acesso realizado nas redes, deixa rastros que são passíveis de serem capturados e posteriormente analisados, auxiliando empresas no processo de tomada de decisões em diferentes setores, como por exemplo, financiamento de créditos, campanhas de marketing etc.

A esse gigantesco volume de dados, dá-se o nome de Big Data.

Por definição, Big Data é o termo utilizado para definir um expressivo número de dados gerados, podendo estes serem estruturados ou, como na maioria dos casos, não estruturados, tornando-se inviável e extremamente complexo o processamento e análise destes pelos métodos tradicionais de processamento. De acordo com¹², credita-se à primeira utilização do termo Big Data à Michael Cox e David Ellsworth quando estes descrevem em seu artigo “*Application-Controlled Demand Paging for Out-of-Core Visualization*” o seguinte trecho:

A visualização oferece um desafio interessante para os sistemas de computadores: os conjuntos de dados gerados são geralmente bastante grandes, sobrecarregando as capacidades da memória principal, do disco local e até mesmo do disco remoto. Chamamos esse problema de Big Data.¹³

Com o surgimento de plataformas de *streaming* de dados, IoT (*Internet of things*), sensores de carros, indústrias, câmeras de segurança etc., estima-se que:

- Nos últimos 5 anos, têm-se gerado mais dados do que em todos os 2000 anos anteriores.

- Por dia, são gerados mais do que 2 quintilhões ($2 \cdot 10^{18}$ bytes) de bytes.

- Google estima que nos últimos 5 anos, a humanidade tenha gerado cerca de 300 exabytes.

Esse enorme volume e variedade de dados, em sua grande maioria, estão distribuídos em 3 categorias:

- Dados estruturados: dados em formato fixo, com tipos bem definidos e com as mesmas características e atribuições.

- Dados não estruturados: são armazenados em formatos que viabilizam a visualização pelos seres humanos, em contrapartida, dificulta o processo de leitura através de um computador, em geral, são dados de e-mail, documentos em formato de texto, imagens, vídeos etc.

- Dados semiestruturados: uma interseção das categorias anteriores, podendo conter alguma estrutura, mas não generalizada em todos os dados, como por exemplo, páginas web, redes sociais etc.

Dentre as três categorias acima, os dados estruturados representam a menor parcela dos dados gerados, sendo os dados não estruturados, a grande maioria.

Atualmente têm-se várias definições de Big Data que se atualiza constantemente e em alta velocidade. Uma das definições, feita pela empresa de tecnologia da informação IBM (*International Business Machines Corporation*) e apresentada por¹² definiu Big Data com 4 V's, como apresentado na Figura 2

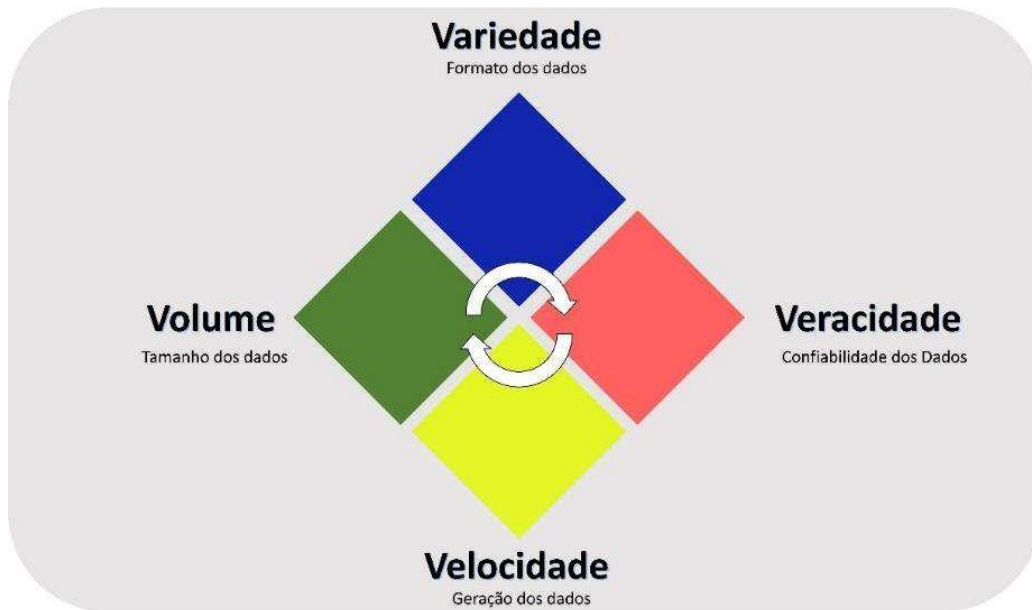


Figura 2- os 4 V's do Big Data
Fonte: Adaptada de FILATRO ¹²

Com esse extenso e variado número de dados, surgiu a necessidade de uma área especializada na análise e manipulação dessas informações, a *Data Science* (Ciência de Dados)

2.4.2 Data Science (Ciência de Dados)

Data Science, ou Ciência de Dados, pode ser definido como disciplina que fornece princípios, metodologias e orientações para transformação, validação, análise e criação de significado a partir de dados. O objetivo é extrair conhecimento de conjuntos de dados que, por sua vez, podem ser grandes demais (o chamado Big Data) para as análises estatísticas tradicionais.¹²

Para confirmar seu caráter científico, é necessário que a Ciência de Dados seja construída baseada no método científico, desta forma, construída baseada em hipóteses testáveis.

Entretanto, a Ciência de Dados, de acordo com,¹⁴ possui uma característica própria de análise orientada a dados, em comparação com a abordagem científica tradicional.

A Figura 3, adaptada de BALDASSARRE,¹⁴ ilustra a comparação da abordagem tradicional e a abordagem orientada a dados:

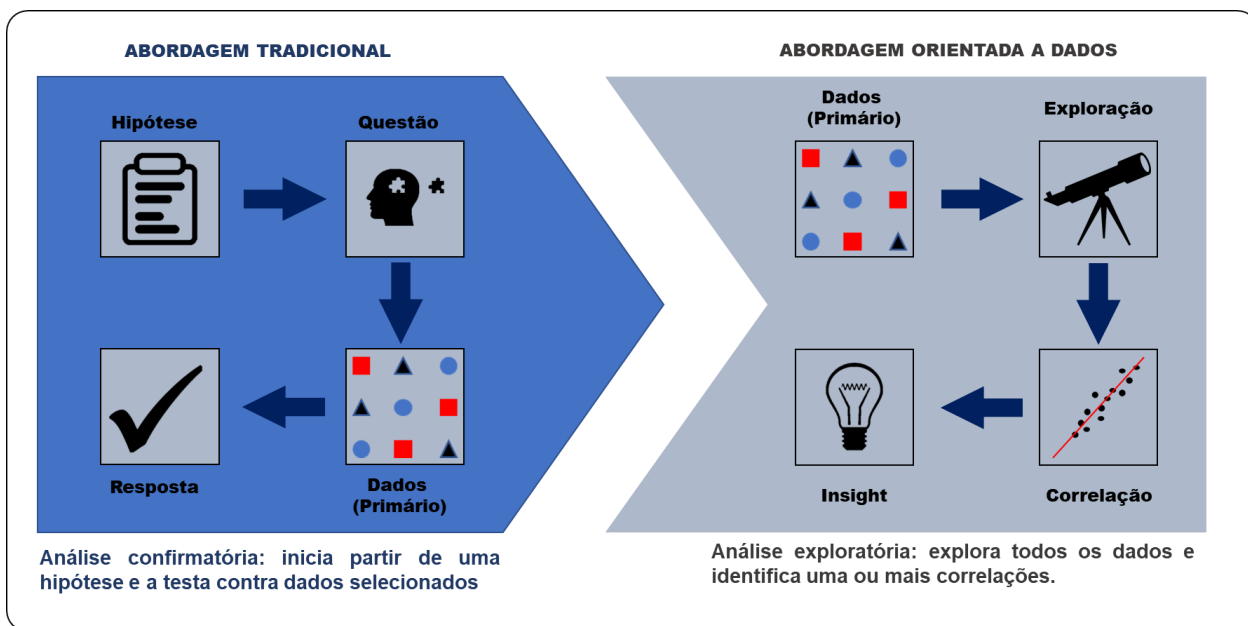


Figura 3 - Abordagens tradicionais e de Big Data em pesquisa
 Fonte: Adaptada de BALDASSARRE¹⁴

Em resumo, de acordo com ¹⁴ a Ciência de dados difere da abordagem científica tradicional pois tem como ponto de partida um banco de dados massivo para, assim, obter *insights* sobre o tema, por outro lado, a abordagem tradicional tem como ponto de partida uma hipótese.

Neste sentido, quando aplicado ao campo educacional, o *Data Science* reúne diversas áreas do conhecimento, como Estatística, Ciência da Computação, educação além de outras ciências sociais. Desta forma, surgiu uma subdisciplina dentro do *Data Science*, o ***Educational Data Science (EDS)***, “podendo ser definido como um campo orientado a dados, sistêmico, transdisciplinar e dinâmico, que combina habilidades técnicas e sociais à compreensão profunda da prática educacional em diferentes ambientes de aprendizagem”.¹²

Assim como o *Data Science*, de forma geral, reúne diversas áreas do conhecimento, o EDS também é pautado sobre algumas, como demonstrado na Figura 4 a seguir:

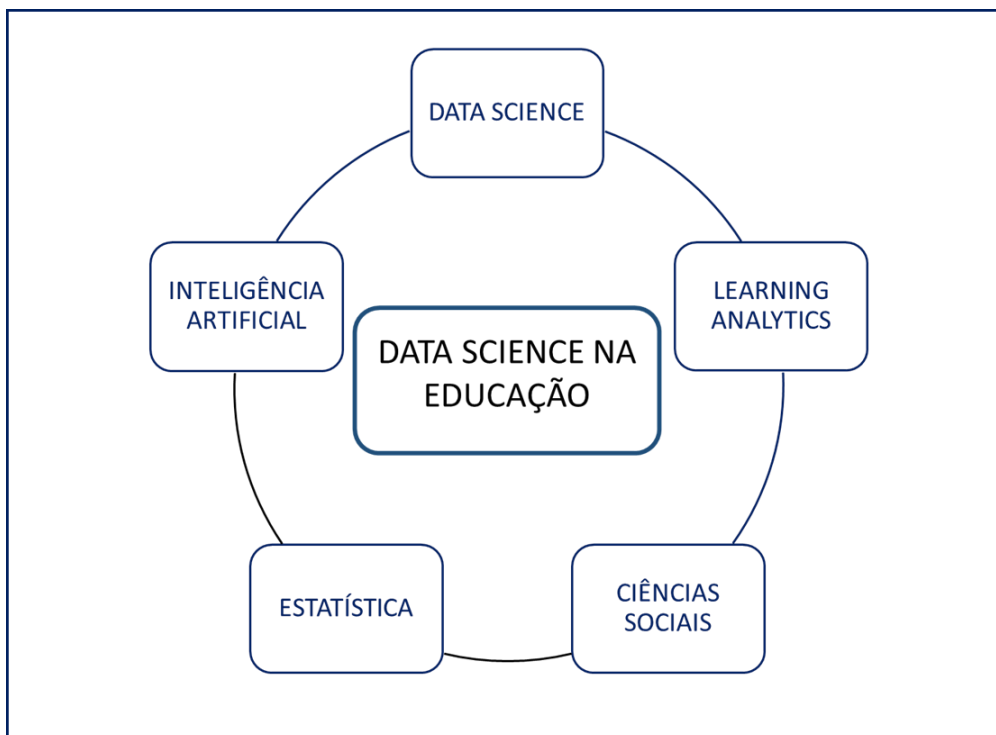


Figura 4 - Áreas constituintes do Data Science na Educação
Fonte: Adaptada de FILATRO¹²

2.4.3 Learning Analytics (Análítica de aprendizagem)

“Recentemente, Learning Analytics (LA) tem chamado a atenção de acadêmicos, pesquisadores e administradores. Esse interesse é motivado pela necessidade de entender melhor o ensino, aprendizagem, “conteúdo inteligente” e personalização e adaptação”.¹⁵

Segundo DIETZ¹⁶, diversas instituições de ensino vêm utilizando os conceitos e técnicas de Learning Analytics para melhorar o sucesso e aprovação dos seus estudantes, como ele apresenta em seu artigo *“Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective”*, uma tabela com alguns casos de sucesso, como demonstrado na Figura 5.

Tabela 1 - Com Instituições e ferramentas de analítica de aprendizagem

Institution	Learning Analytic Tool	Uses of Data
University of Central Florida	EIS (Executive Information System)	Data management
Rio Salado Community College	PACE (Progress and Course Engagement)	Track student progress in course; intervention
Northern Arizona University	GPS (Grade Performance System)	Student alerts for academic issues and successes
Purdue University	Course Signals System	Student alerts for academic issues; intervention
Ball State University	Visualizing Collaborative Knowledge Work	Enhance knowledge-building work
University of Michigan	E ² Coach	Student support and intervention
University of Maryland Baltimore County (UMBC)	Blackboard LCMS	Track performance and predict student success
Graduate School of Medicine, University of Wollongong	BIRT (Business Intelligence and Reporting Tools)	Reveal continuity of care issues

Fonte: *Using Learning ...*¹⁶

Atualmente, encontra-se diversas definições sobre *Learning Analytics* e uma dessas definições foi utilizada na 1ª Conferência Internacional de Learning Analytics (1st *International Conference on Learning Analytics*), como descreve Siemens¹⁵: “Learning Analytics é a medida, coleta, análise e apresentação de dados sobre os alunos e seus contextos, com o objetivo de entender e otimizar a aprendizagem e os ambientes em que ocorre”.¹⁵

Independentemente da forma como essas medidas e análises sejam realizadas, o objetivo é a partir do levantamento de dados sobre os alunos ou turmas, poder identificar possíveis eventos/motivos pelos quais o processo de solidificação dos conteúdos trabalhados sejam ou não os esperados.

2.4.4 Big Data e Data Science na Educação

Assim como nas demais áreas afetadas pelo Big Data, o campo educacional vem incorporando cenários dessas tecnologias, em virtude das diversas abordagens educacionais gerarem cada vez mais dados e também demandarem análises detalhadas e voltadas para um melhor planejamento e execução de ações na área da educação.¹⁷

Na última década, vários movimentos distintos surgiram em torno do uso de dados na educação.¹⁸

De acordo com PIETY¹⁸ surgiram nos últimos anos diversos movimentos baseados e alicerçados no uso de dados voltados a educação. Ele divide esses movimentos em áreas e são definidas como:

- Análise Acadêmica/Institucional,
- Análise de aprendizagem/ Mineração de dados educacionais,
- Análise/Personalização do aluno e
- Melhoria Instrucional Sistêmica.

Ainda segundo PIETY¹⁸, essas áreas estão diretamente conectadas com o avanço da tecnologia digital e a capacidade desta em coletar, analisar, compartilhar e representar um extenso número de dados de forma relativamente simples.

Analisando estes fatores e pelas características do tipo e formato do banco de dados utilizado neste trabalho, as informações a serem analisadas tornam-se inviáveis para uma análise tradicional, justificando o caráter de Big Data e Data Science desta pesquisa.

Além das características dos dados utilizados que relacionam esta pesquisa com os conceitos de Big Data e Data Science, a dinâmica pedagógica no curso integrado de física também justifica a utilização dos conceitos e técnicas abordados.

Como mostrado nos capítulos anteriores (capítulos 2.4.1 á capítulo 2.4.4), a utilização das técnicas adotadas nas áreas de Big Data e Data Science, além de possibilitar a análise de um massivo banco de dados, também permite acelerar o processo de análise, possibilitando uma intervenção no processo de ensino e aprendizagem mais dinâmica, quando e se desejado.

O curso integrado de física básica tem duração de 5 meses aproximadamente, ou seja, um curto período para se fazer de forma tradicional a coleta, análise, estudo e direcionamento de uma intervenção a ponto de caso necessário, uma recuperação do processo, a fim de torná-lo mais significativo e eficaz. Com as técnicas utilizadas nos conceitos abordados até aqui, todo esse processo pode ser acelerado e tornar-se mais dinâmico e significativo, tanto para o docente quanto para o discente.

3 METODOLOGIA

3.1 Método para coleta de dados e a reestruturação curricular

O objetivo desta pesquisa é realizar uma análise exploratória de dados relativos aos resultados de discentes da Universidade de São Paulo no campus de São Carlos dentro da disciplina de física integrada, oferecida pelo Instituto de Física de São Carlos (IFSC) para diversos cursos do campus.

Após a análise, busca-se identificar os diversos padrões de comportamento dos discentes, dentro das diversas atividades realizadas, e a partir dessas análises, evidenciar os padrões positivos e negativos, ou seja, padrões que corroborem para a aprovação ou não na disciplina estudada.

Como o foco desta pesquisa está em analisar os dados do desempenho dos discentes do curso integrado de física, faz-se necessário entender a reestruturação curricular do curso bem como as formas que os dados obtidos foram gerados.

Conforme já mencionado neste texto, os dados utilizados foram previamente coletados na disciplina “CURSO INTEGRADO DE FÍSICA I – 7600005 - Física I”, disciplina essa oferecida pelo Instituto de Física de São Carlos (IFSC) para diversos cursos do campus de São Carlos. No ano de 2018, a disciplina foi ministrada para as seguintes turmas do campus:

- Bacharelado em Ciências Física e Biomoleculares - IFSC.
- Bacharelado em Física – IFSC.
- Bacharelado em Física Computacional - IFSC.
- Bacharelado em Matemática Aplicada e Computação Científica - ICMC.
- Engenharia Aeronáutica - EESC.
- Engenharia de Computação – EESC/ICMC.
- Engenharia de Produção - EESC.
- Engenharia Elétrica - EESC.
- Engenharia Mecânica – EESC.
- Engenharia Mecatrônica – EESC.

O curso integrado de física I – 7600005 - Física I, foi uma das primeiras ações do corpo docente do IFSC responsável pela reestruturação curricular dos cursos do Instituto, visando a adequação das disciplinas às metodologias mais modernas e atuais no campo de ensino, feita em 2017. Tal reestruturação trouxe mudanças significativas às disciplinas dos cursos, em

específico, nas disciplinas do ciclo básico (primeiro ano de graduação), estas oferecidas aos diversos cursos do campus de São Carlos.

O objetivo da reestruturação, além do já mencionado acima, é também possibilitar que os discentes envolvidos desenvolvam um papel mais ativo, trabalhando com maior autonomia e permitindo serem os principais responsáveis pelo seu desenvolvimento na disciplina, dessa forma, não apenas desenvolvendo habilidades dentro da disciplina, mas também preparando o discente para a vida, pessoal e profissional.

O curso foi reestruturado para ser composto por diversas atividades através de diferentes estratégias e estímulos, dessa forma, também possibilitando o alcance ao maior número possível de discentes em suas especificidades e particularidades, o que também passa a ter um caráter de maior inclusão pedagógica. Neste sentido, o curso foi reestruturado com o formato representado de forma concisa pela Figura 6:



Figura 6 - Divisão de atividades por metodologias.
Fonte: Elaborada pelo autor

O diagrama mostrado na Figura 6 apresenta as metodologias e ferramentas de análise e avaliação em que o curso foi baseado sua reestruturação, neste caso metodologias tradicionais e metodologias ativas, constituída das seguintes atividades e estratégias abaixo:

- AE: aulas expositivas convencionais incluindo discussão dos temas principais, realizadas por um docente do curso, para um grupo de aproximadamente 60 alunos distribuídos de forma aleatória e mista entre todas as turmas, mantendo um caráter homogêneo entre as turmas e imparcial.

- AM: aulas de monitoria realizadas por alunos bolsistas do programa de pós-graduação PAE (Programa de Aperfeiçoamento de Ensino) realizadas semanalmente em salas pré-determinadas e em horário compatível com o cronograma de aula dos alunos.

- ADS e TEC: atividades de discussão supervisionada realizadas em grupo com duração de aproximadamente 50 minutos, em grupos de até 5 estudantes, composta de questões conceituais acerca do tema da aula, trabalho em grupo e Trabalho Extraclasse (TEC).

- EOL: exercícios on-line separados em pré-tópico e pós-tópico disponibilizados em formato de questionários e atividades no sistema e-Disciplinas (<https://edisciplinas.usp.br>) onde é capaz de utilizar recursos e ferramentas de correção automática e imediata, promovendo e possibilitando um feedback instantâneo ao discente.

Os questionários pré-tópicos são, em geral, mais curtos e objetivos, focando em conceitos e habilidades específicas, permitindo que o aluno mantenha um contato com o tópico antes mesmo da aula, o que possibilita um melhor e maior aproveitamento deste posteriormente nas discussões em aula

Os questionários pós-tópico são, em contraponto ao pré-tópico, mais extensões e com um nível de complexidade equiparado aos exercícios encontrados nos livros textos disponibilizados na ementa do curso. O *feedback* imediato em conjunto com a possibilidade de dezenas de realizações da mesma tarefa, possibilita que o aluno consiga verificar seu grau de entendimento daquele conteúdo, se ele realmente compreendeu o tema ou se é necessária alguma ação, deste ou do professor. As diversas tentativas, embora possam ser realizadas, sempre é sorteado um novo exercício com novos parâmetros, estratégia essa com o objetivo de mitigar os possíveis comportamentos errados ou pouco produtivos que visam apenas a nota alta e não o aprendizado significativo.

-LE: Listas de problemas sugeridos compostas dos conteúdos ministrados, retirado de livros-textos e/ou outros problemas propostos.

- Provas regulares: provas aplicadas em 3 momentos distintos do percurso da disciplina durante o semestre letivo, elaboradas por uma banca (professores da disciplina), previamente escolhida e definida pelos docentes da disciplina, assim, garante-se a imparcialidade e homogeneidades nas avaliações aplicadas para todas as turmas envolvidas neste processo de ensino aprendizagem.

A partir das diversas atividades relacionadas acima, a nota de cada discente foi calculada seguindo os critérios apresentados na Tabela 2 e detalhados abaixo:

NOTA_ADS é a média das notas das ADS e TEC propostas. A nota de cada ADS será calculada como apresentado na equação 1.

NOTA_EOL é a média aritmética das notas dos EOL (pré-tópicos).

NOTA_ATV é a média ponderada das notas das atividades trabalhadas por

metodologias ativas, apresentada na equação 2.

Nota_PROVA é a média aritméticas das notas obtidas nas provas dissertativas, apresentadas na equação 3.

NOTA_FINAL é a média ponderada final do discente, apresentada na equação 4.

Tabela 2 – Definições das notas das atividades

Atividade	Equação do cálculo da nota	Número da Equação
Nota_ADS	$[0,4 \times (NOTA_ADS) + 0,6 \times (NOTA_TEC)]$	(01)
Nota_EOL	Média aritmética das notas dos EOL avaliativos (pré-tópicos).	
Nota_ATV	$0,75 \times (NOTA_ADS) + 0,25 \times (NOTA_EOL)$	(02)
Nota_prova	$\frac{(P1 + P2 + P3)}{3}$	(03)
Nota-final	$0,15 \times (Nota_atv) + 0,85 \times (Nota_prova)$	(04)

Fonte: Elaborada pelo autor

As provas de recuperação seguiram o regimento da Universidade de São Paulo e foram aplicadas em data definida no final do semestre.

Os dados obtidos em todas as atividades acima citadas, foram registrados e coletados no AVA utilizado por todas as turmas da disciplina, organizados em uma planilha de dados a partir da qual é realizada esta pesquisa.

As turmas que constituíam a disciplina, como já mencionado neste texto, são oriundas de diversos cursos do campus de São Carlos, e foram distribuídas em 9 turmas, designadas de T1 a T9, onde cada turma possuía um discente diferente, o que também se notou que este não teve influência significativa nos resultados analisados das turmas.

Importante ressaltar que, as turmas foram separadas de forma aleatória, sem nenhum pré-requisito para qualquer turma que fosse sendo cada turma constituída por alunos de todos os cursos do campus participantes da disciplina.

Importante também lembrar que esta pesquisa (independente) não faz parte do trabalho oficial de reestruturação do curso do IFSC, sendo as informações contidas aqui fornecidas pelos docentes da disciplina após o fim da disciplina (aplicada em 2018), não tendo o autor desta

pesquisa participada dessa reestruturação e nem diretamente da coleta de dados. Este trabalho concentra-se, portanto, na análise dos dados fornecidos sobre o desempenho dos discentes durante a disciplina, desconsiderando também, informações pessoais dos discentes, turmas e ou qualquer informação pessoal de qualquer integrante do curso. Nossa pesquisa pautou-se apenas nos resultados fornecidos de forma anonimada dos dados, para preservar a privacidade dos discentes envolvidos.

3.2 Descrição do Dataset

Como já mencionado neste texto, atestar uma única metodologia que garanta a qualidade do aprendizado, principalmente atualmente, é de extrema dificuldade, sobretudo analisando os diversos fatores, muitas vezes intangíveis por sua variedade e complexidade.

A prática docente de mais de 8 anos atuando com alunos do ensino médio, nos possibilita a visualização de determinados padrões, no que diz respeito ao comportamento e/ou responsabilidade do aluno frente aos seus estudos.

Alguns padrões percebidos no decorrer dos anos são, em sua grande maioria:

- Melhor desempenho nas disciplinas que são dedicados mais tempo de estudo;
- Baixo desempenho daqueles que estudam em vésperas de atividades avaliativas;
- Baixo engajamento nas disciplinas em que possuem baixo rendimento;

Esses padrões, embora recorrentes nesses anos atuando como professor da rede básica de ensino, possuem maior caráter qualitativo, dessa forma, pouco auxiliam no processo de ensino e aprendizagem.

Nesse sentido, essa pesquisa busca por trazer um caráter quantitativo e tentar identificar, de forma mais objetiva, possíveis padrões para comportamentos negativos que possam prejudicar o processo de aprendizagem do discente bem com padrões de comportamento positivo e dessa forma, ser possível embasar as possíveis tomadas de decisão para a “recuperação” do aprendizado.

Os dados utilizados nesta pesquisa foram coletados pelos docentes da disciplina e fornecidos através de um *Dataset* disponibilizado em um arquivo Excel, com os dados pessoais dos discentes anonimados, uma vez que tais informações são irrelevantes para os resultados desta pesquisa. As identificações das turmas foram realizadas de forma aleatória, também buscando a imparcialidade no tratamento dos dados. Quando e se necessário, as turmas serão mencionadas nesta pesquisa através de numerações de 1 a 9

O *Dataset* construído possuía 62 colunas por 566 linhas quando disponibilizado. Após uma análise e tratamento desses dados, verificou-se a necessidade de inclusão e exclusão de algumas informações dentro do próprio arquivo, o *Dataset* final utilizado para a geração dos resultados possui 72 colunas e 566 linhas. A Tabela 3 apresenta alguns exemplos dos dados presentes no *Dataset* analisado.

Tabela 3 – Exemplos dos dados presentes no Dataset Utilizado

	Prova_P1	Prova_P2	Prova_P3	Prova_Sub	Provas_total	Grupos_ProvaTotal	ADS_1	ADS_2
0	6.95	6.5	5.4	-	6.28	3	7.8	9.6
1	4	7.5	3.4	-	4.97	2	9.8	8.3
2	4.3	7.5	3.5	-	5.1	2	8	6.1
3	2.1	4	1.9	1	2.37	1	8.2	8.5
4	4.6	9	3.3	-	5.63	2	7.7	9.2
...
560	-	-	-	-	-	0	-	-
561	1.4	4	1.7	-	2.37	1	8.48	5.7
562	2.2	6	3.8	-	4	2	9.6	8.4

Fonte: Elaborada pelo autor

Por se tratar de um *Dataset* relativamente grande, este foi analisado e trabalhado através de scripts de programação realizado em Python, por se tratar de uma linguagem de programação aberta e de fácil acesso, além de também ser uma das linguagens de programação mais utilizadas na área de *Data Science*, como já mencionado também no capítulo 2 desta pesquisa.

4 RESULTADOS

Partindo do conhecimento do *Dataset* utilizado, participaram da disciplina um total de 565 alunos, divididos em 9 turmas mistas, com alunos de diversos cursos misturados de forma mais ou menos aleatória (respeitando apenas a conveniência dos campi onde tinha aulas). Os Resultados gerais da disciplina estão apresentados na Figura 7 (relação de aprovados e não aprovados), que representa a relação de aprovados e não aprovados.

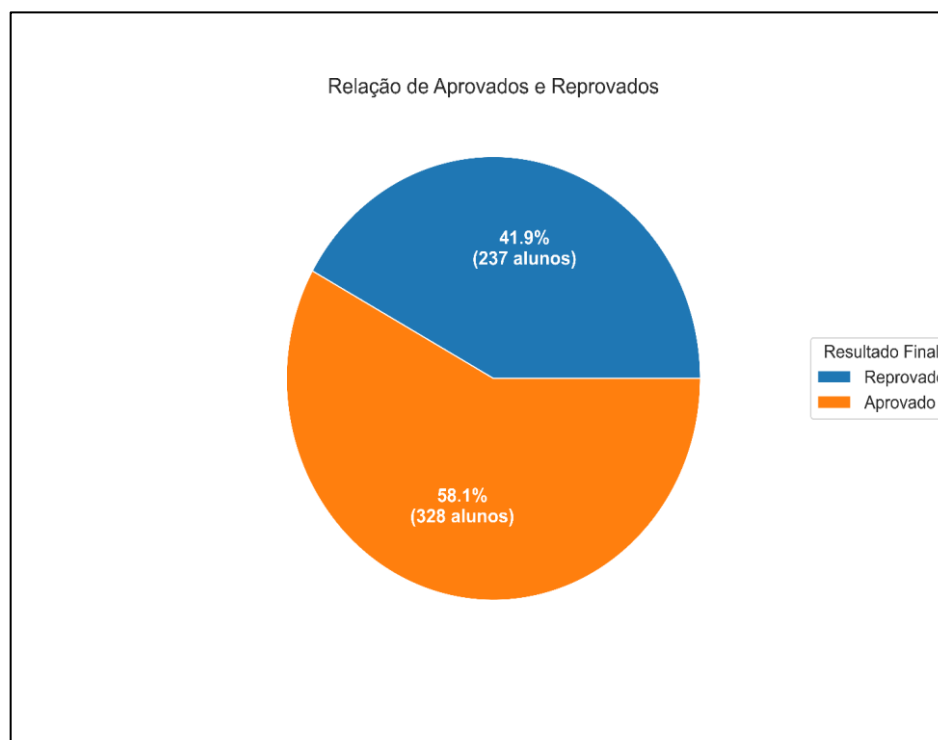


Figura 7 - Relação de aprovados e não aprovados na disciplina
Fonte: Elaborada pelo autor

Os dados considerados para construir o gráfico da Figura 7, se referem ao registro final no sistema de graduação, portanto, após considerar todas as atividades avaliativas mencionadas no capítulo 3 deste texto, sem excluir os estudantes que desistiram da disciplina (sem fazer o trancamento da matrícula). Assim, obteve-se na disciplina aproximadamente 58% de aprovação, como mostrado na figura.

A Figura 8 (histograma da distribuição das notas finais) logo abaixo, mostra a distribuição das notas finais de todas as turmas. O pico inicial corresponde a uma fração dos

desistentes. Portanto, se excluirmos esses 30 estudantes, o índice de aprovação é de cerca de 63%.

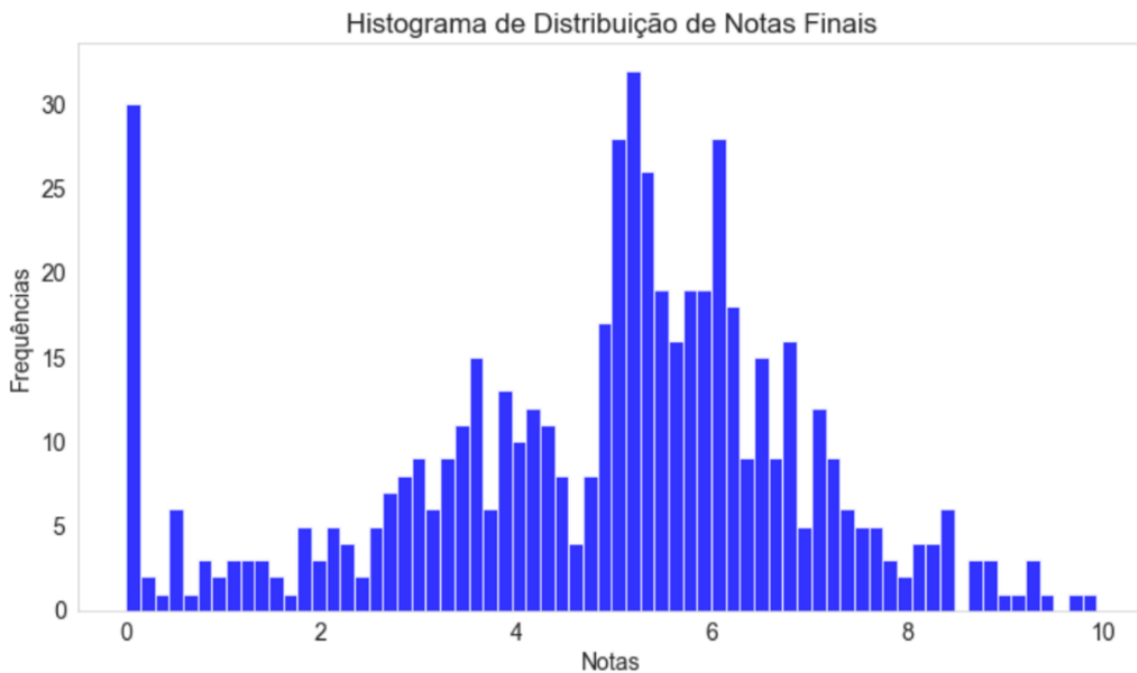


Figura 8 - Histograma de distribuição das notas finais na disciplina
Fonte: Elaborada pelo autor

O gráfico apresentado na Figura 8 indica uma distribuição maior na região de notas entre 4,5 e 7 pontos, o que corrobora com o número maior de aprovações.

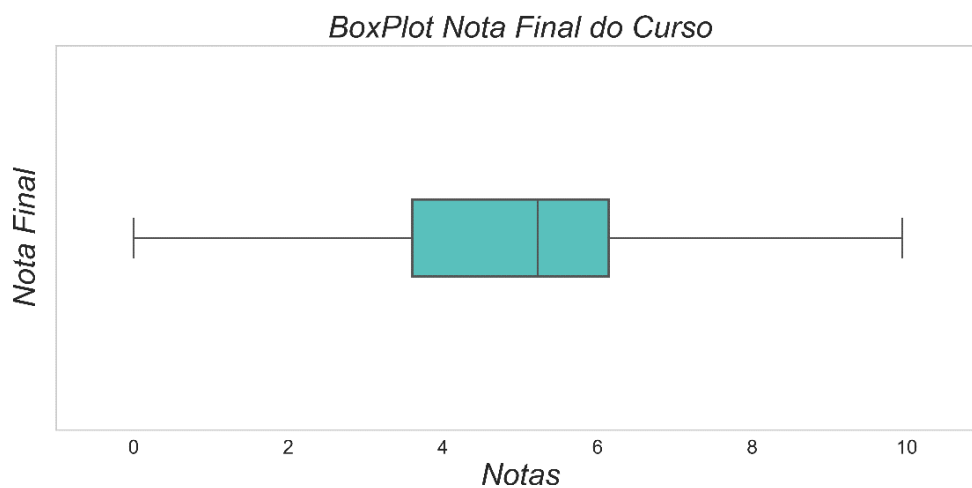


Figura 9 - BoxPlot da nota final do curso
Fonte: Elaborada pelo autor

É possível observar, através da figura 9 (boxplot com nota final do curso), que a mediana, calculada em 5,14 está levemente acima da média e da nota mínima para aprovação, que de acordo com o regimento da Universidade é de 5,0 pontos. Isso corresponde que pouco mais da metade (se não excluirmos os desistentes) dos discentes obtiveram notas totais acima da média mínima exigida. O que também é possível identificar olhando o número de aprovados.

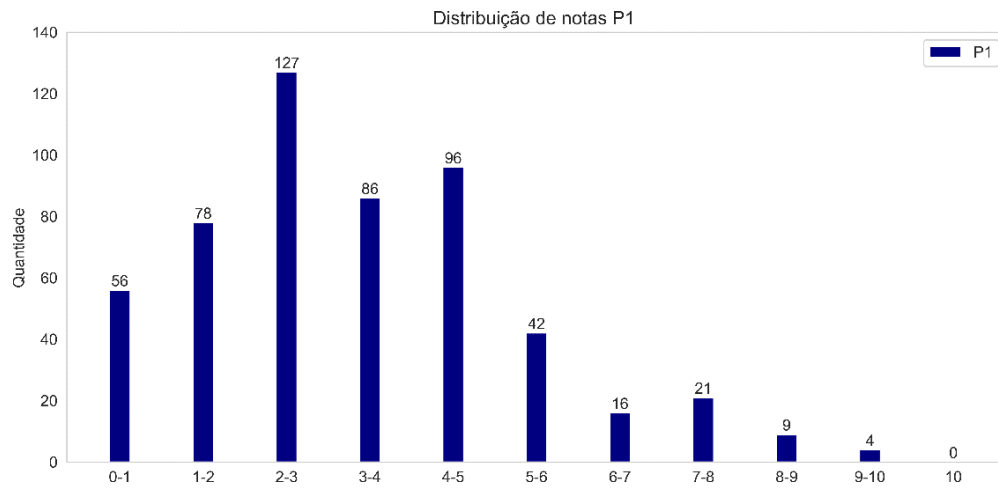


Figura 10 - Distribuição das notas da prova 1 (P1)
 Fonte: Elaborada pelo autor

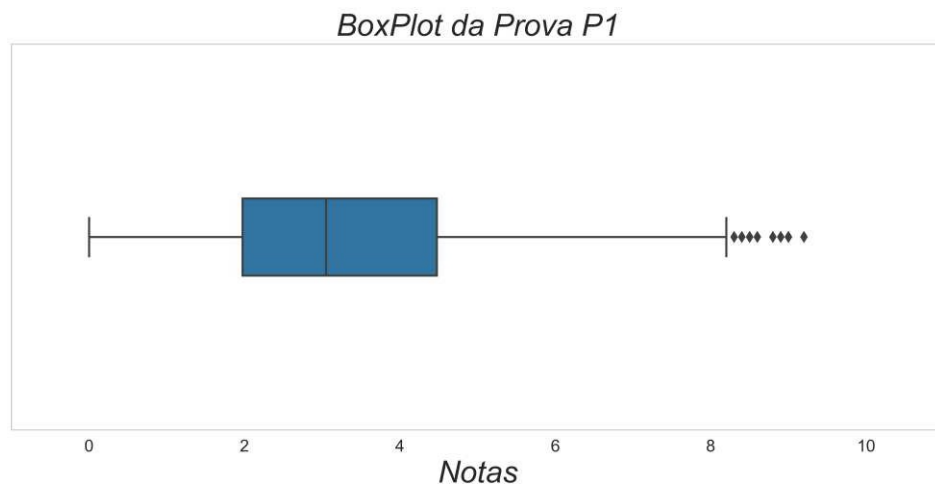


Figura 11 – BoxPlot das notas da prova 1 (P1)
 Fonte: Elaborada pelo autor

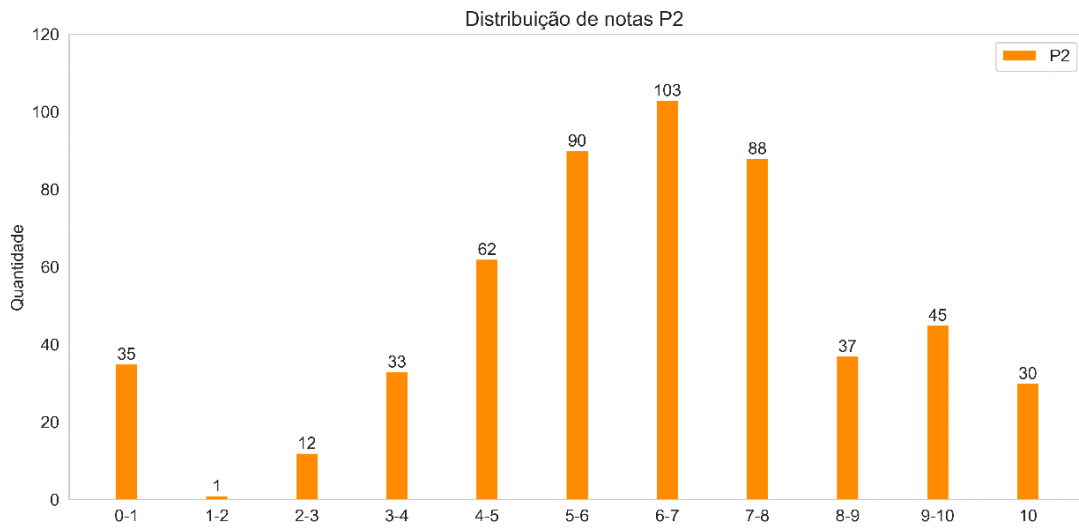


Figura 12 - Distribuição das notas da prova 2 (P2)
 Fonte: Elaborada pelo autor

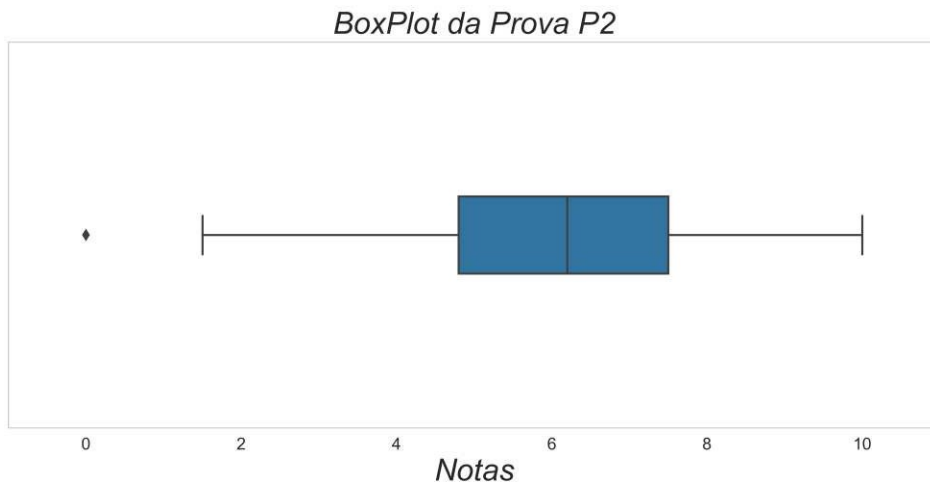


Figura 13 - BoxPlot das notas da prova 2 (P2)
 Fonte: Elaborada pelo autor

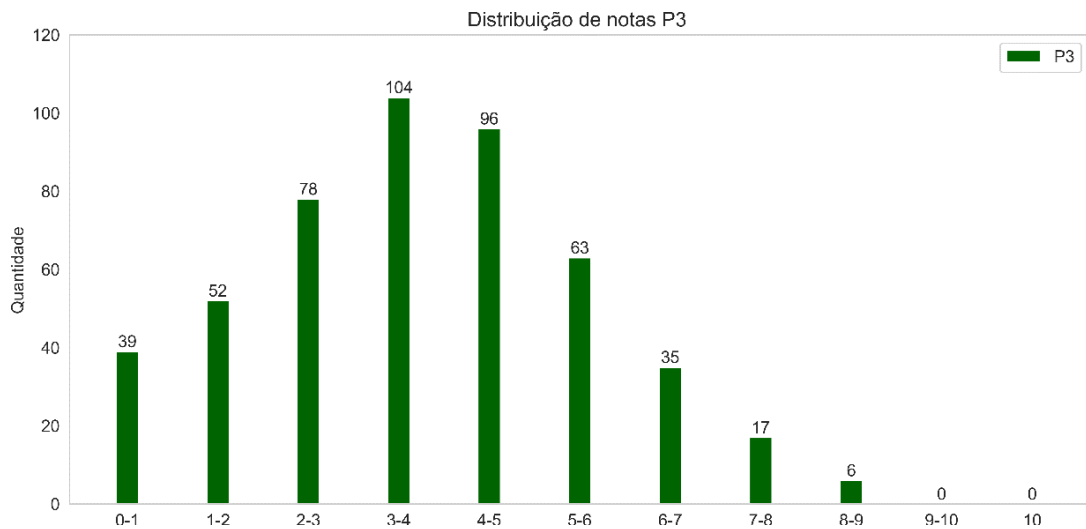


Figura 14 - Distribuição das notas da prova 3 (P3)
 Fonte: Elaborada pelo autor

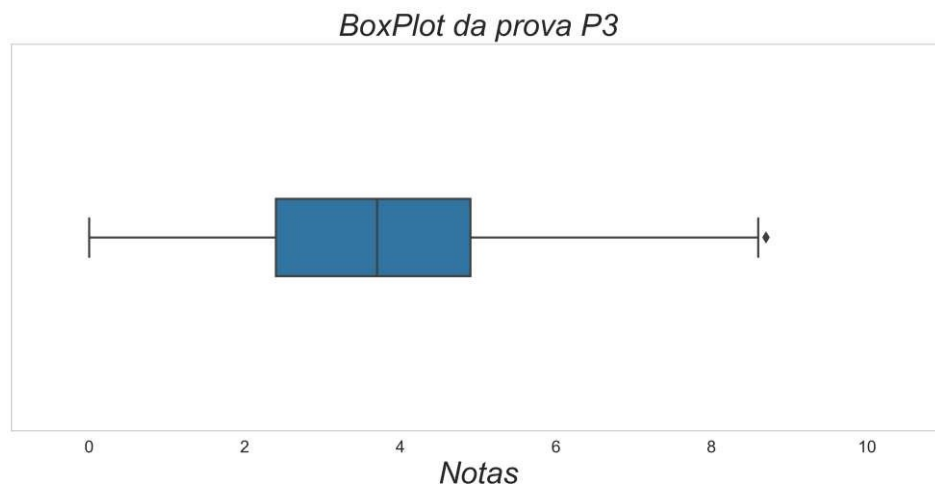


Figura 15 - BoxPlot das notas da prova 3 (P3)
 Fonte: Elaborada pelo autor

As Figuras 10, 12 e 14 mostram a distribuição das notas de cada prova (P1, P2, P3), respectivamente, relacionando-as. A distribuição de notas das provas P1 e P3, é semelhante, e estão mais centradas na região entre 3 e 5, o que não ocorre na prova P2, que está mais centrada na região entre 4 e 8. Isso indica uma possível recuperação no desempenho por parte dos alunos durante a prova P2 e, talvez, um possível pragmatismo de alguns estudantes com relação a P3 (especialmente estudantes de engenharia) após atingir a média para aprovação.

A Figura 16 mostra as distribuições apresentadas numa mesma visualização, para facilitar a comparação.

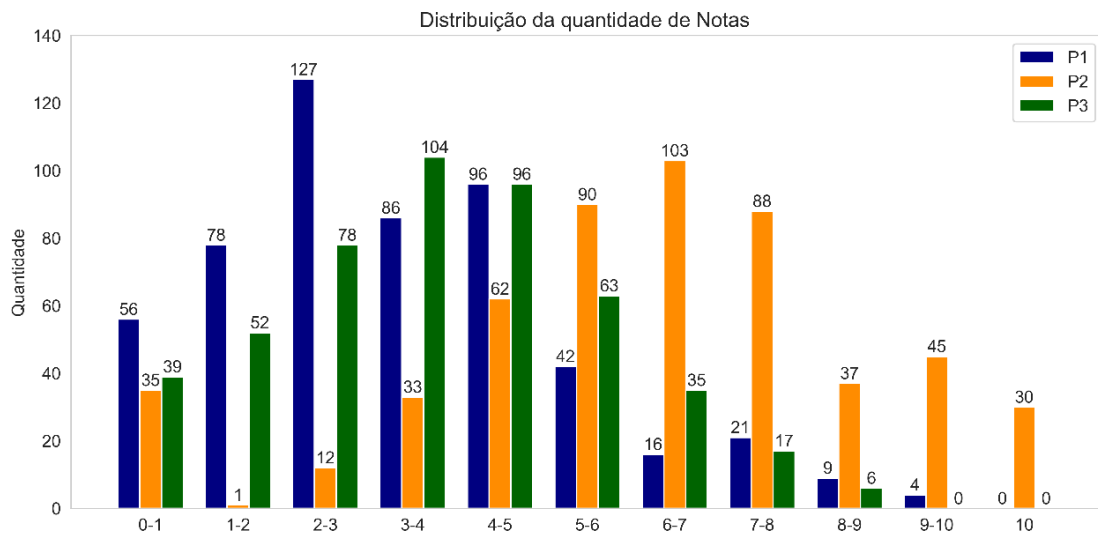


Figura 16 - Distribuição da quantidade de notas das provas 1, 2 e 3
 Fonte: Elaborada pelo autor

Importante observar que para notas abaixo de 5, houve um aumento da P3 em relação a prova P2, o que ocorre o inverso para notas acima de 5. Dentre várias possibilidades, uma causa recorrente é o número de missings (notas faltantes dentro do *dataset*) na P3, identificado pela Figura 17, indicando que muitos alunos desistem da disciplina ao longo do decorrer do semestre, também indicado pela Figura 17 com a distribuição da quantidade de notas, indicando mais de 40 desistentes no final.

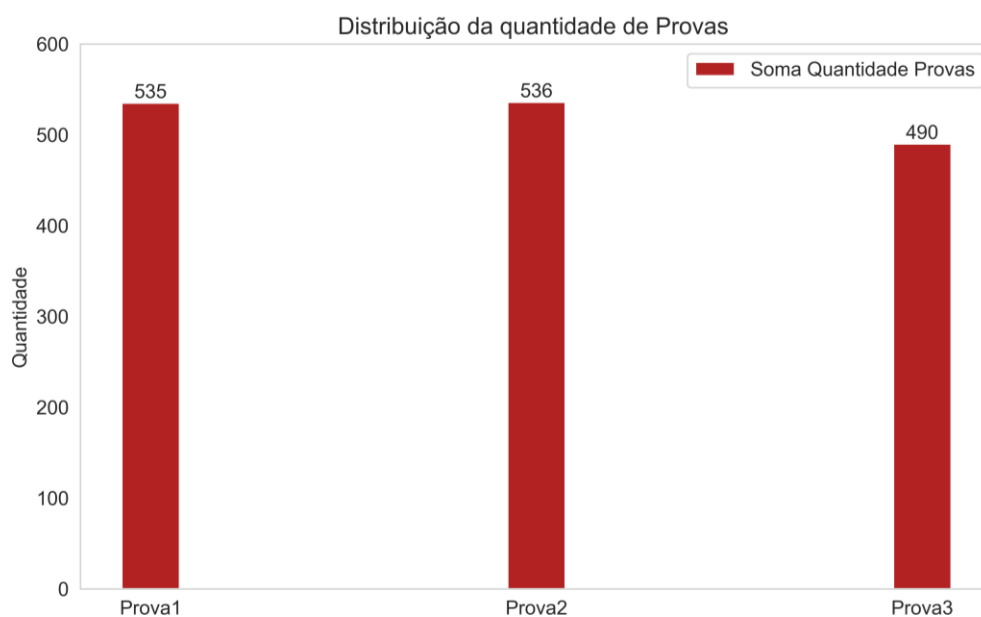


Figura 17 - Distribuição da quantidade de provas realizadas.
 Fonte: Elaborada pelo autor

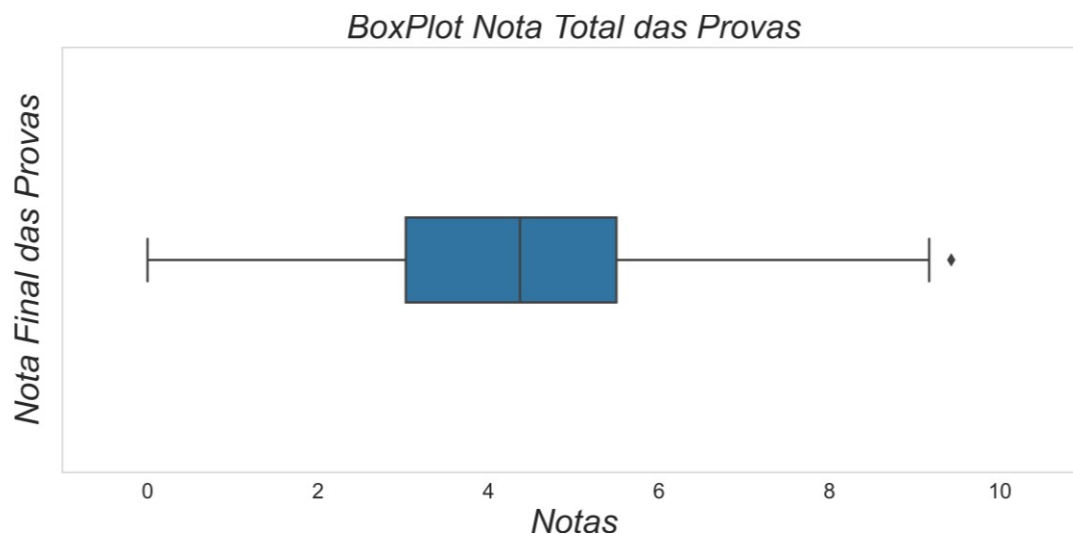


Figura 18 - BoxPlot com notas finais das provas
Fonte: Elaborada pelo autor

O boxplot acima (Figura 18) demonstra algumas medidas estatísticas sobre a nota final das provas. Observamos que a mediana, calculada em 4,37, está levemente acima da média da sala, calculada em 4,29, e um pouco abaixo da média da nota final, calculada em 4,83, e da nota mínima para aprovação, 5, de acordo com o regimento da Universidade. Isso significa que pouco mais da metade dos discentes obtiveram notas totais em prova abaixo da média, isso já indicando que, caso o critério adotado pelo docente, no início da disciplina fosse apenas as notas de provas, pouco mais da metade teriam reprovado na disciplina. Os dados indicados pelo BoxPlot acima, estão apresentados na Tabela 4, com algumas medidas estatísticas das principais atividades.

Tabela 4 - Medidas estatísticas das principais atividades

Tipo de medida	Atividades				
	Prova	ADS	TEC	Ativ Pré	Nota Final
Média	4,29	7,14	8,2	7,89	4,83
Mediana	4,37	8,2	9,18	8,75	5,14
Min	0	0	0	0	0
25%	3,03	7,2	8,4	8	3,6
50%	4,37	8,2	9,1	8,75	5,22
75%	5,5	8,76	9,5	9,2	6,14
Máx	9,43	9,86	10	9,8	9,94

Fonte: Elaborada pelo autor

A Tabela 4 indica algumas medidas estatísticas sobre as principais atividades avaliativas utilizadas e sobre a nota final o curso:

Também podemos ver que os valores da média e mediana para as notas da Prova e Nota Final estão próximos ao valor mínimo exigido para aprovação na disciplina (5,0), uma vez que pelas regras da disciplina, a nota total da prova equivale a 85% da nota final.

Na Figura 18 e pela Tabela 4, também podemos notar que apenas 25% das notas das provas estão acima de 5,5, o que ainda pode ser um resultado preocupante, indicando que 75% dos discentes ainda possuem dificuldades na disciplina, obtendo nota mínima ou abaixo disso.

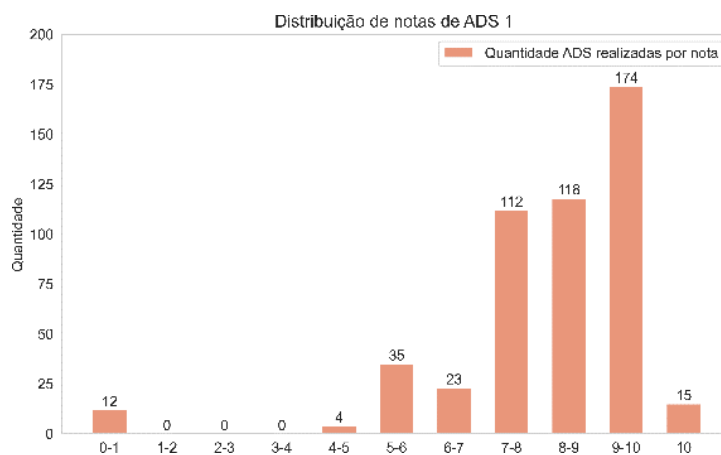


Figura 19 - Distribuição das notas de ADS 1
Fonte: Elaborada pelo autor

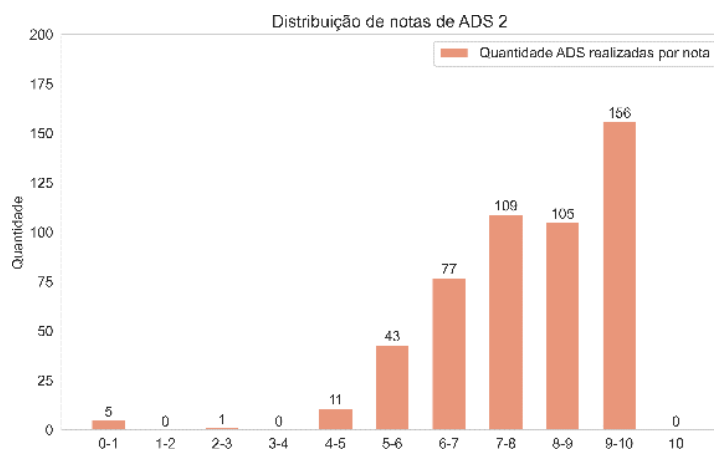


Figura 20 - Distribuição das notas de ADS2
Fonte: Elaborada pelo autor

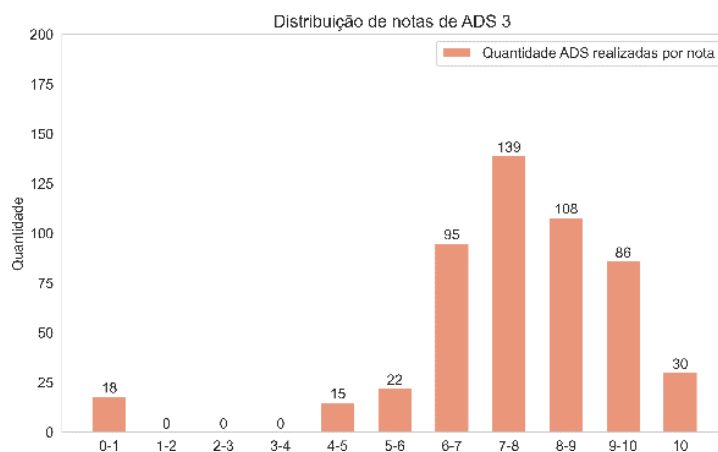


Figura 21 - Distribuição das notas de ADS3
 Fonte: Elaborada pelo autor

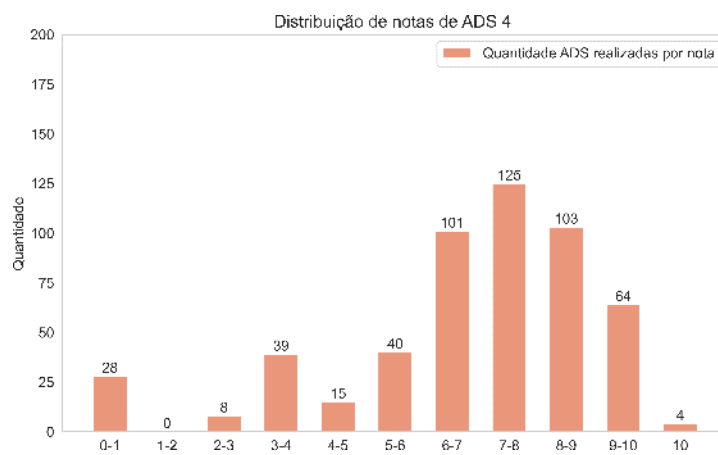


Figura 22 - Distribuição das notas de ADS4
 Fonte: Elaborada pelo autor

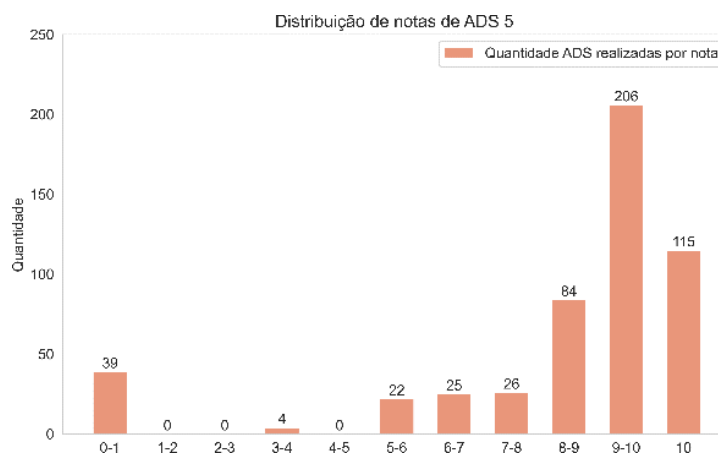


Figura 23 - Distribuição das notas de ADS5
 Fonte: Elaborada pelo autor

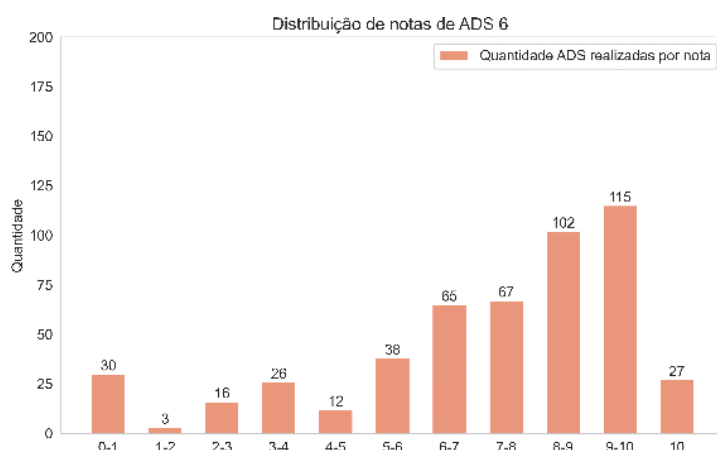


Figura 24 - Distribuição das notas de ADS6
 Fonte: Elaborada pelo autor

Os gráficos das Figuras de 19 a 24, demonstram o comportamento das notas obtidas nas atividades ADS, mantendo uma distribuição similar para todas as ADS's com pouca variação na região de maior intensidade. Abaixo, representamos através das Figuras 25 e 26, o comportamento da distribuição por nota das notas finais obtidas nas atividades ADS e a quantidade de atividades ADS realizadas, respectivamente. Também é possível notar um decaimento na realização das atividades ADS ao final do semestre.

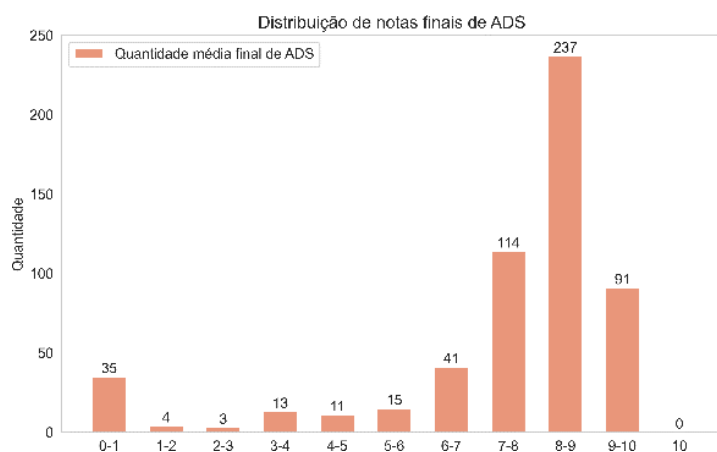


Figura 25 -Distribuição das notas finais de ADS
 Fonte: Elaborada pelo autor

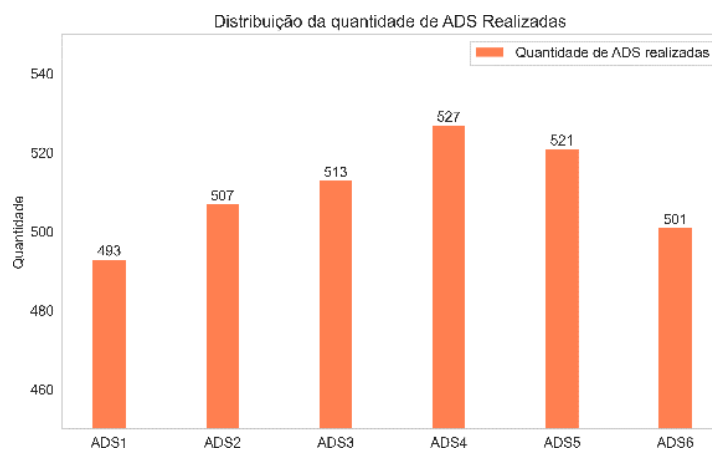


Figura 26 - Número de ADS realizadas
 Fonte: Elaborada pelo autor

O mesmo comportamento observado nas ADS também é percebido nas atividades de TEC e Atividade Online, identificadas pelos gráficos das figuras abaixo (Figuras de 27 a 36).

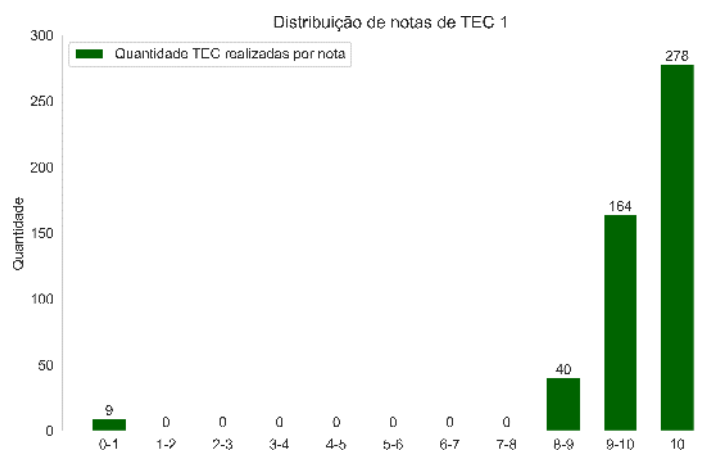


Figura 27 - Distribuição das notas de TEC1
 Fonte: Elaborada pelo autor

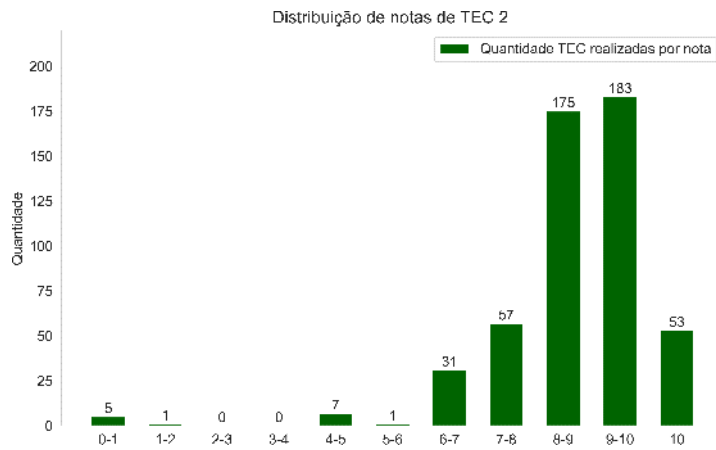


Figura 28 - Distribuição das notas de TEC 2
 Fonte: Elaborada pelo autor

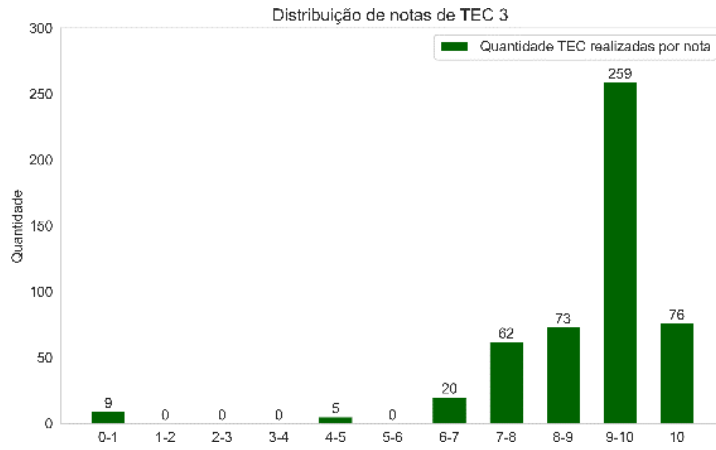


Figura 29 - Distribuição das notas de TEC 3
 Fonte: Elaborada pelo autor

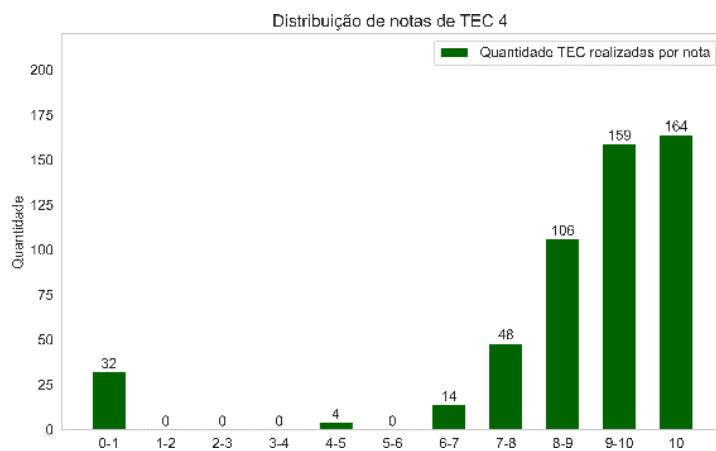


Figura 30 - Distribuição das notas de TEC 4
 Fonte: Elaborada pelo autor

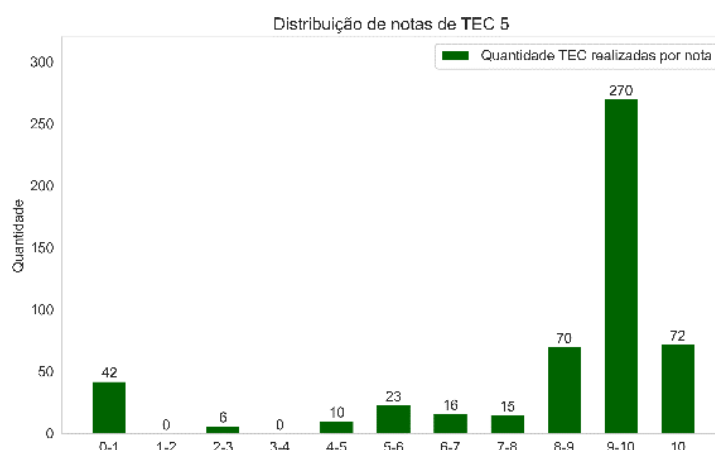


Figura 31 - Distribuição das notas de TEC 5
 Fonte: Elaborada pelo autor

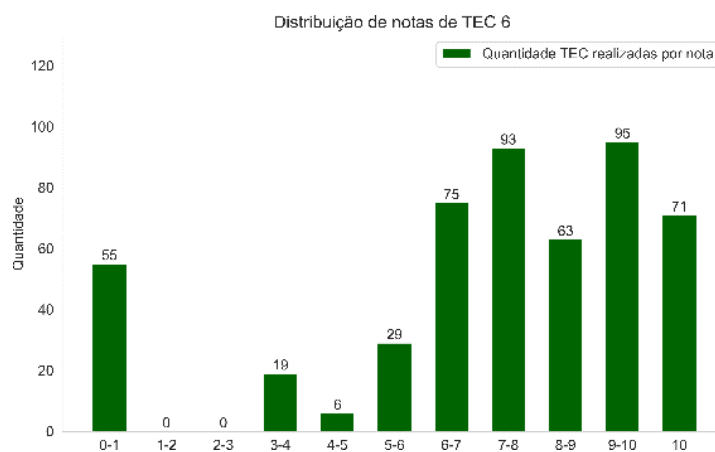


Figura 32 - Distribuição das notas de TEC 6
 Fonte: Elaborada pelo autor

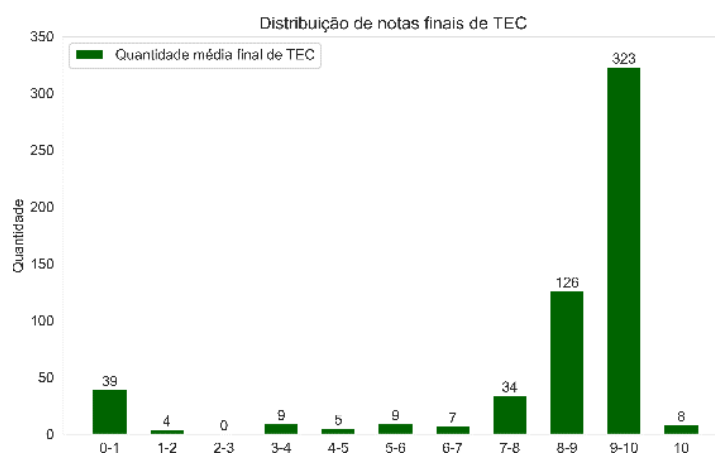


Figura 33 - Distribuição das notas finais de TEC
 Fonte: Elaborada pelo autor

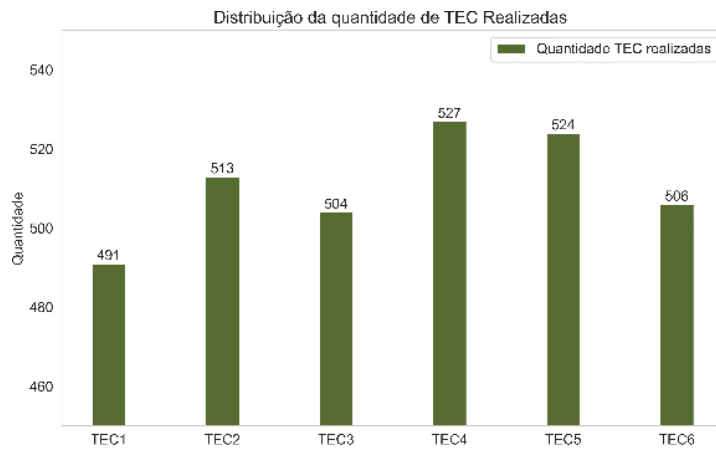


Figura 34 - Número de TEC realizadas
 Fonte: Elaborada pelo autor

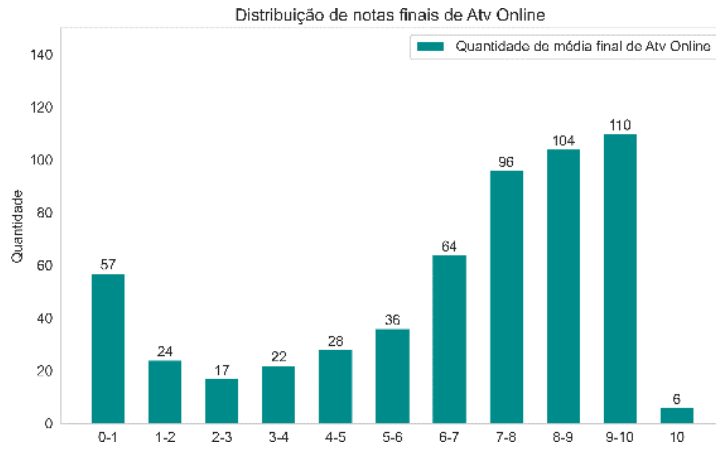


Figura 35 - Distribuição das notas finais das atividades online
 Fonte: Elaborada pelo autor

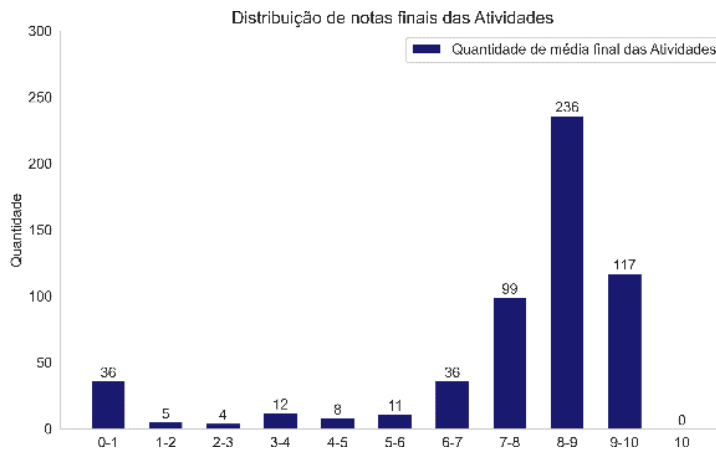


Figura 36 - Distribuição das notas finais das atividades totais.
 Fonte: Elaborada pelo autor

Para todas as atividades, sejam provas ou as atividades realizadas fora do período da aula, nota-se um decaimento na realização destas, o que além de comprometer a nota final, pois todas as atividades são parte integrante da média final, também indica uma tendência de comportamento de abandono da disciplina, por parte do discente. Esse abandono pode estar relacionado a uma série de fatores, internos ou externos, não avaliados neste trabalho.

Tendo em vista os resultados gerais das principais atividades, podemos tentar entender a influência de cada componente no comportamento nas notas das atividades. Para isso, a visualização do gráfico da Figura 37, mostrando correlações, torna-se importante.

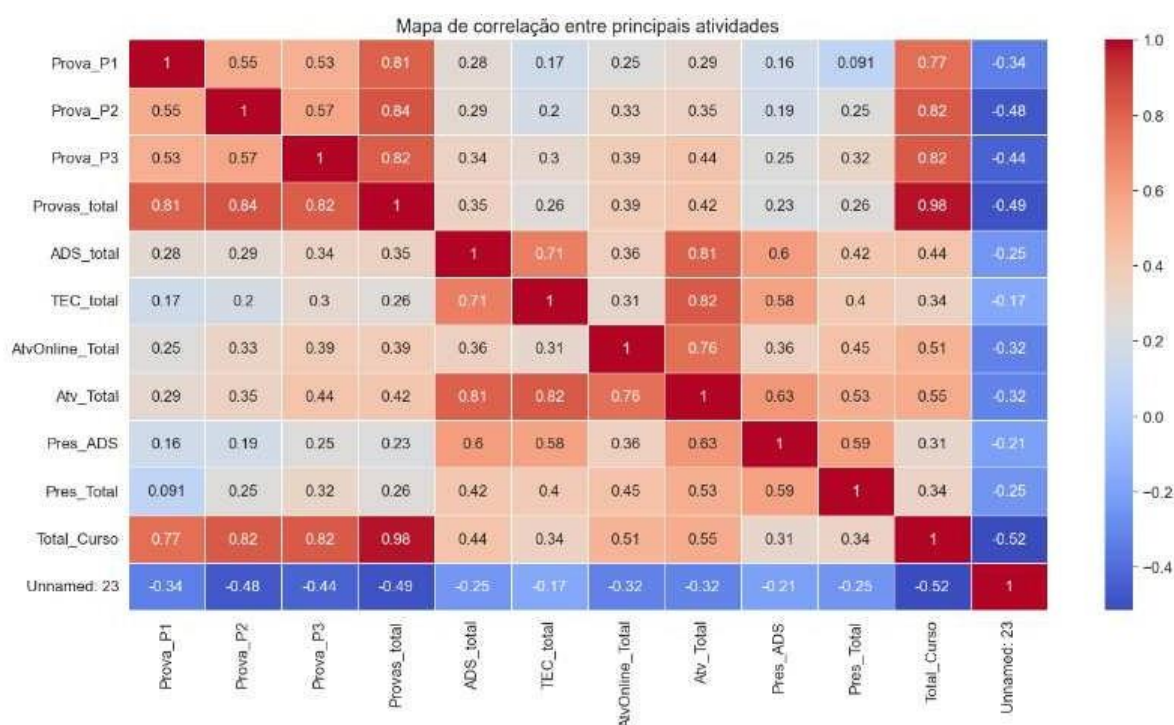


Figura 37 - *Heatmap* com a correlação entre as principais atividades realizadas
Fonte: Elaborada pelo autor

De acordo com as regras gerais da disciplina, apontadas no Capítulo 3, Seção 3.1, as provas possuem um peso significativo na nota final, sendo este de 85%. Neste sentido, observa-se pelo mapa de calor (*heatmap*) da Figura 37, uma correlação forte entre as notas das provas (Prova_P1, Prova_P2, Prova_P3, Provas_total) e a nota final do curso (Total_Curso), indicado por uma coloração tendendo ao vermelho, na legenda ao lado.

Observamos também a pouca correlação na presença do aluno na disciplina (Pres_Total) e a nota final do curso (Total_Curso), observado por uma coloração tendendo ao azul.

Também notamos uma correlação significativa entre as atividades ADS, TEC e Atividade Online em relação à nota final do curso, correlação de aproximadamente 0,51, indicado pela variável “Atv_Total”, que compõem a nota das atividades em uma nota apenas.

Tendo em vista esses resultados gerais apresentados pelo HeatMap acima, dividiu-se os grupos para cada atividade por notas, sendo separados em 4 grupos com notas finais entre 0-3, 3-5, 5-7, e 7-10, fazendo novamente a matriz de correlação entre as principais atividades para cada grupo. Os resultados são mostrados nas Figuras 38-41.

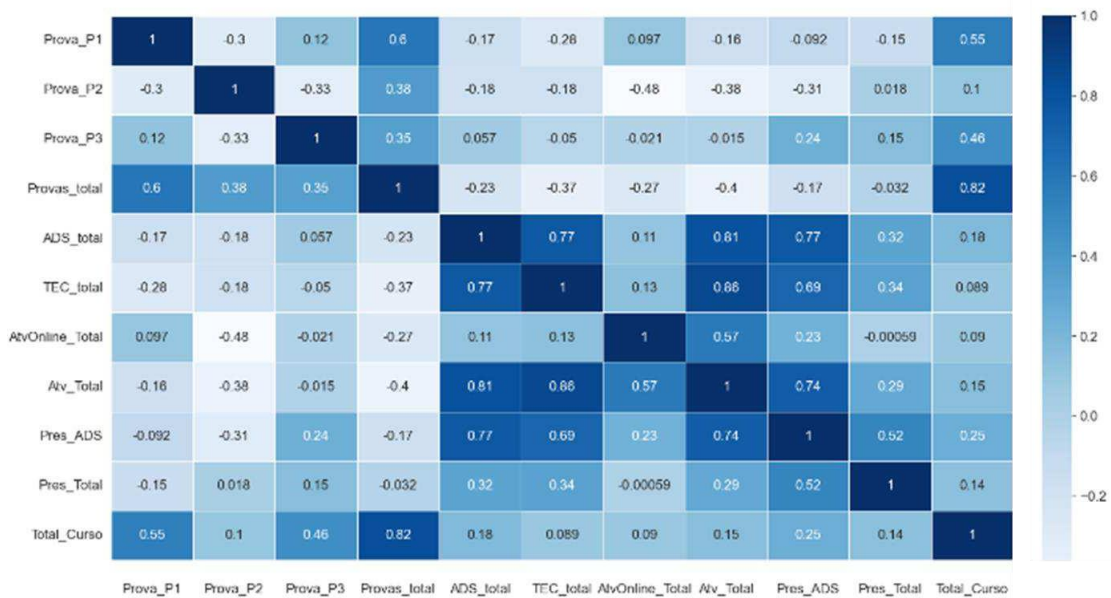


Figura 38 - Heatmap com correlação entre principais atividades para grupo com notas finais de 0 a 3
Fonte: Elaborada pelo autor

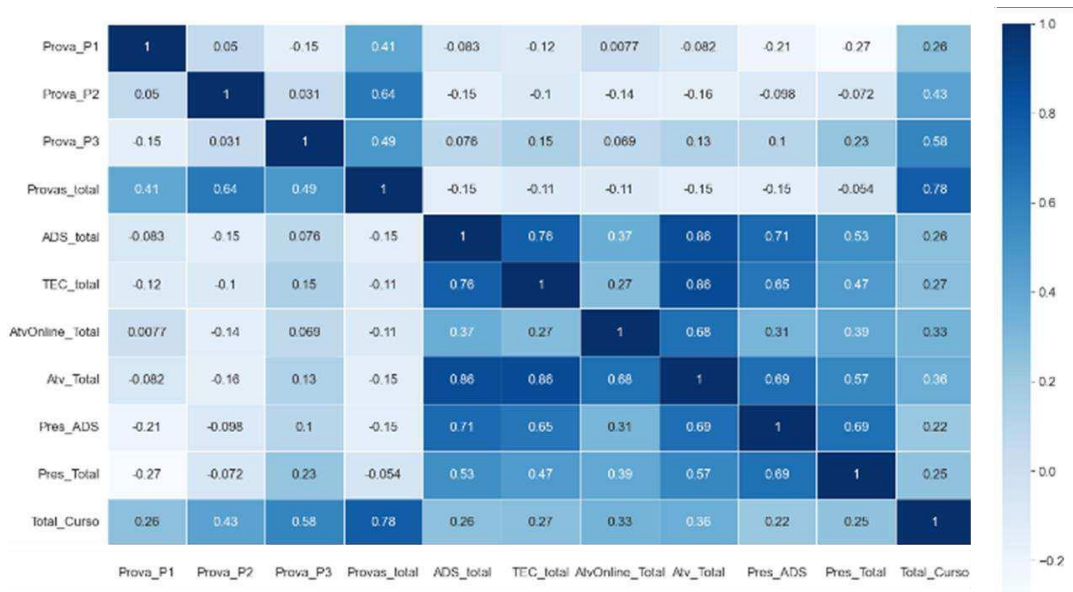


Figura 39 - Heatmap com correlação entre principais atividades para grupo com notas finais de 3 a 5
Fonte: Elaborada pelo autor

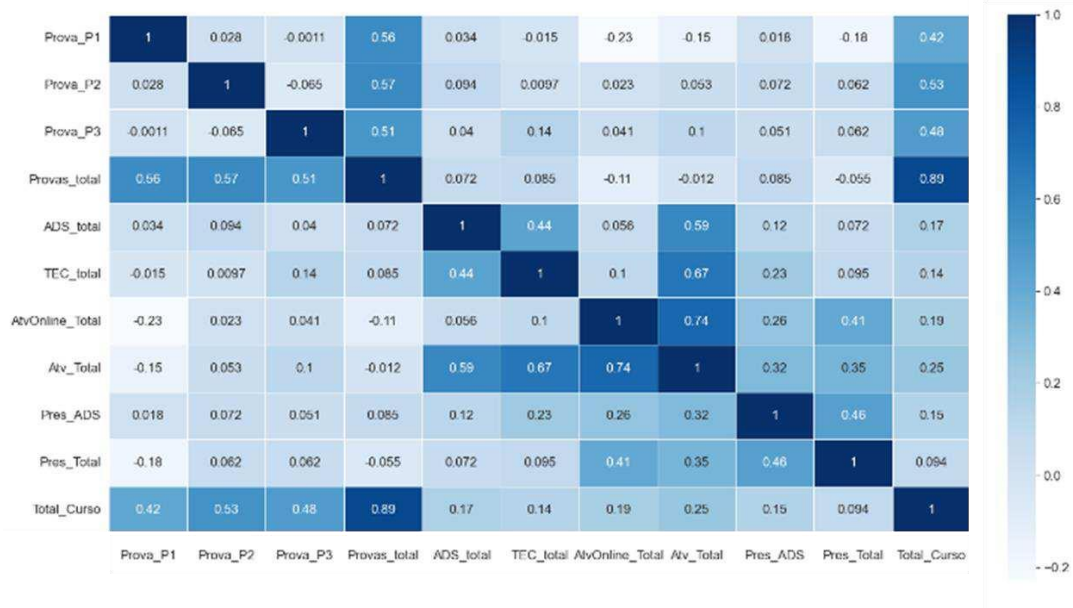


Figura 40 - *Heatmap* com correlação entre principais atividades para grupo com notas finais de 5 a 7
 Fonte: Elaborada pelo autor

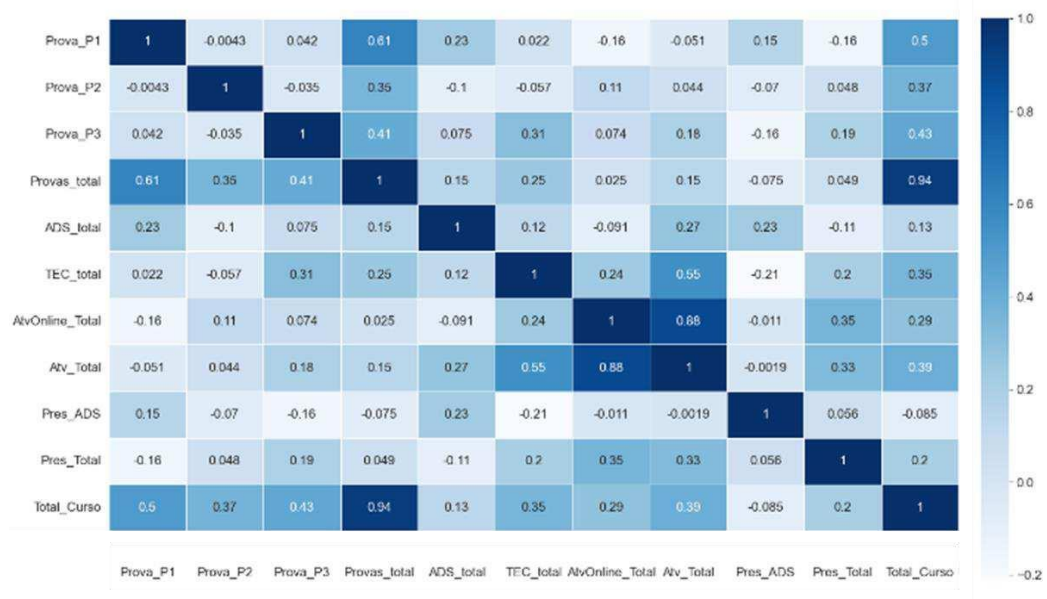


Figura 41 - *Heatmap* com correlação entre principais atividades para grupo com notas finais de 7 a 10
 Fonte: Elaborada pelo autor.

É possível perceber que ao separar os alunos nos respectivos grupos, as correlações entre algumas atividades diminuem conforme as notas finais aumentam, evidenciado pela diminuição da cor nos gráficos.

As evidências da diminuição das correlações mais significativas ficam na região entre as atividades ADS, TEC e Atv online em conjunto com as presenças e a nota final, indicando que para os grupos com notas abaixo da média, onde a correlação é mais forte, o aproveitamento das notas nas atividades foi importante para o resultado.

Para entender de forma mais detalhada esse comportamento, foram realizados gráfico de dispersão entre cada atividade. Inicialmente entre as notas de prova e nota total do curso

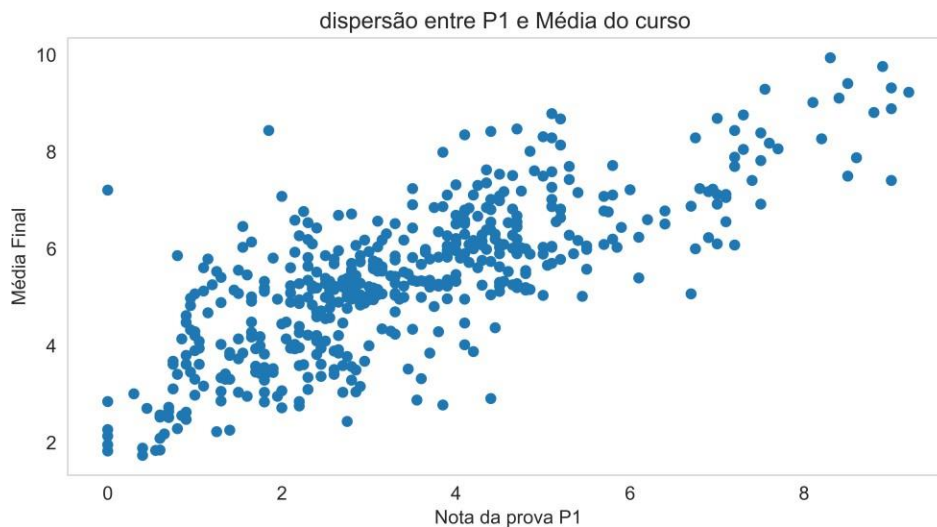


Figura 42 - Dispersão entre nota da prova 1 (P1) e média final do curso
Fonte: Elaborada pelo autor.

O gráfico apresentado na Figura 42, indica uma correlação positiva aproximadamente linear entre a nota da prova 1 (Prova_P1) com a nota final do semestre (Total_Curso), com uma distribuição de notas maior na região abaixo de 5, o que também ficou evidenciado pelo gráfico da Figura 10.

Como realizado nos HeatMaps, dividiu-se os resultados em 4 grupos de notas finais e criou-se o gráfico de dispersão para cada grupo em relação a nota final do curso, demonstrado pelos gráficos das figuras 43, 44, 45 e 46, abaixo.

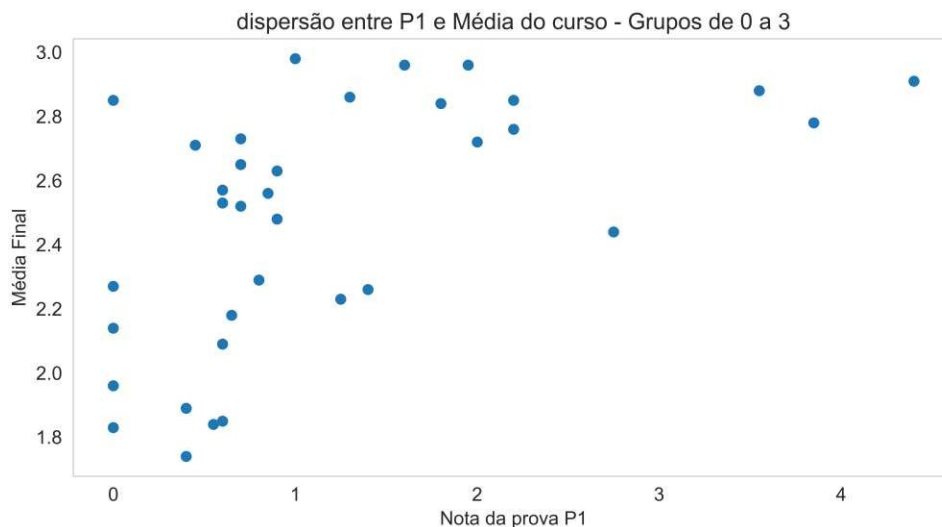


Figura 43 - Dispersão entre nota da prova 1 e média final do curso de 0 a 3
Fonte: Elaborada pelo autor.

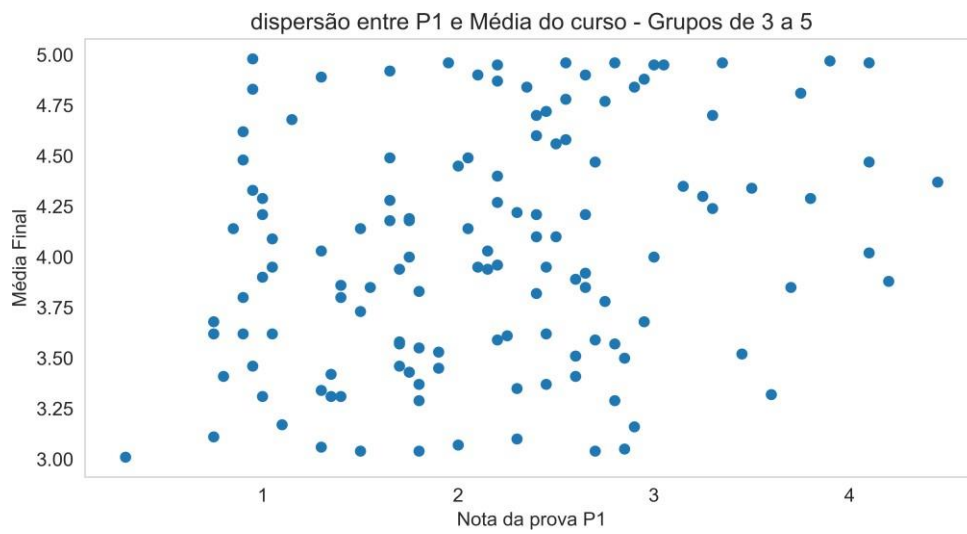


Figura 44 - Dispersão entre nota da prova 1 e média final do curso de 3 a 5
 Fonte: Elaborada pelo autor.

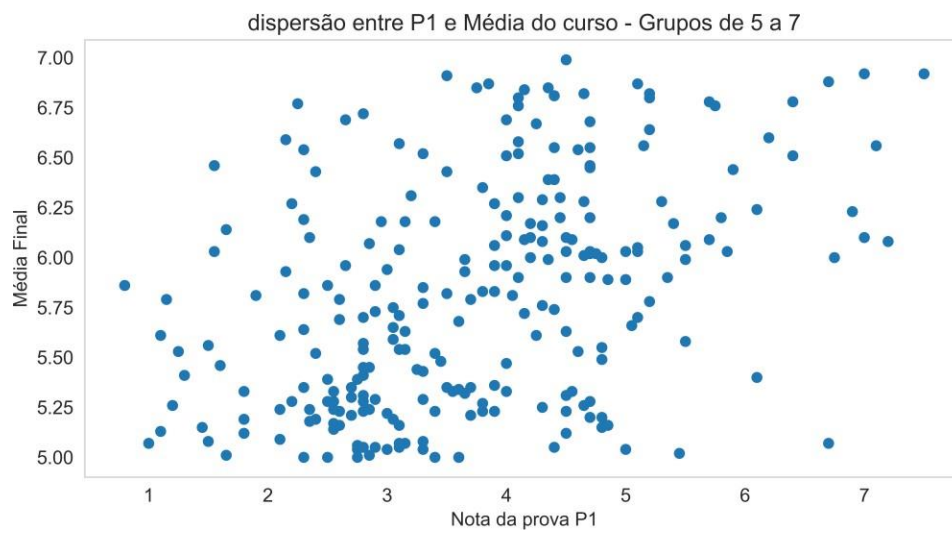


Figura 45 - Dispersão entre nota da prova 1 e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor.

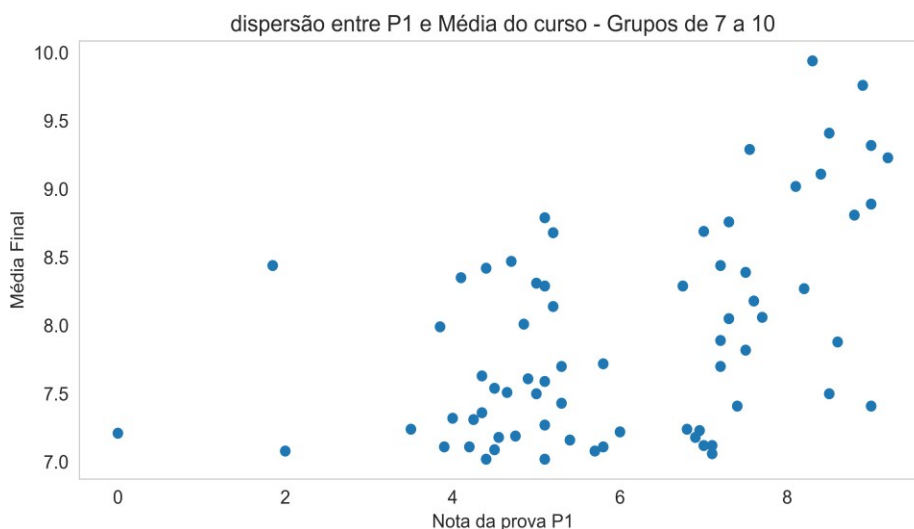


Figura 46 - Dispersão entre nota da prova 1 e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor

Pelos gráficos notamos que praticamente todos os alunos que obtiveram nota final abaixo da média mínima exigida para aprovação, obtiveram também nota abaixo de 5 na prova P1, evidenciado pelos gráficos das Figuras 43 e 44. O gráfico da Figura 46 mostra que a grande maioria dos alunos que finalizaram o curso com nota acima de 7, tiveram nota na prova P1 acima de 5.0.

Também foram realizadas análises de dispersão para as notas de outras provas, como seguem abaixo.

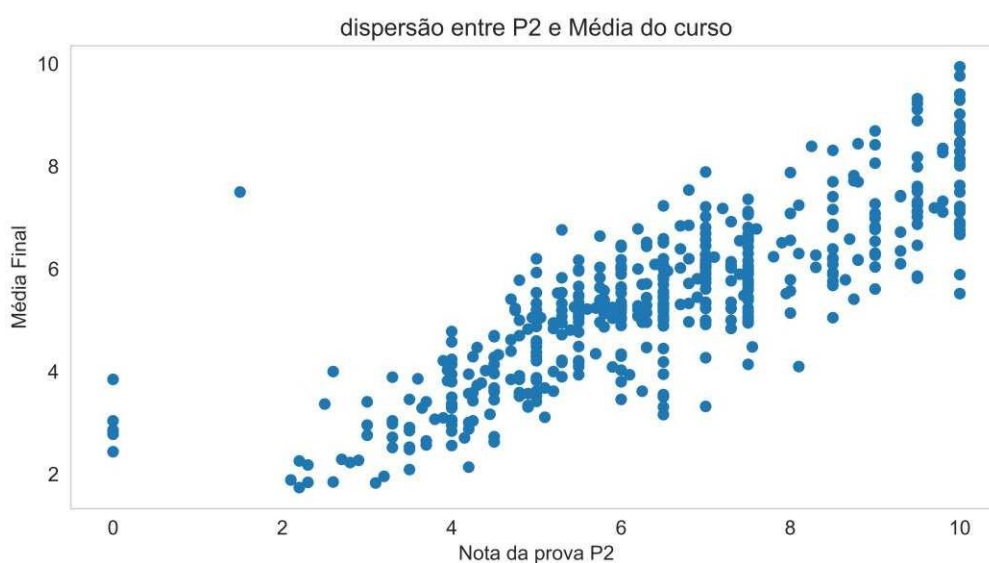


Figura 47 - Dispersão entre nota da prova 2 (P2) e média final do curso
 Fonte: Elaborada pelo autor

Para a nota na prova P2, a linearidade segue semelhante a prova P1, mostrando uma correlação positiva entre a nota da prova P2 e a nota final.

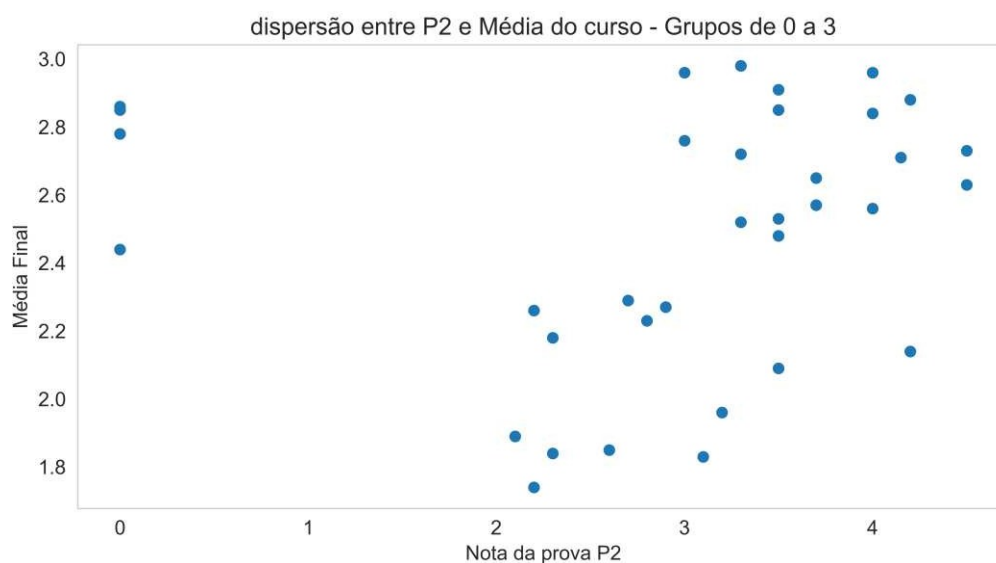


Figura 48 - Dispersão entre nota da prova 2 e média final do curso de 0 a 3
Fonte: Elaborada pelo autor.

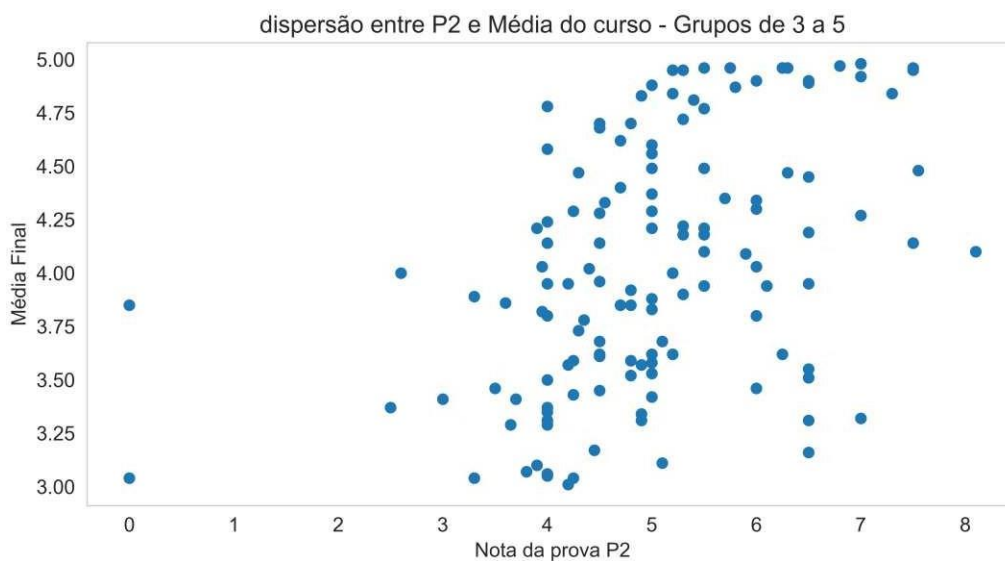


Figura 49 - Dispersão entre nota da prova 2 e média final do curso de 3 a 5
Fonte: Elaborada pelo autor.

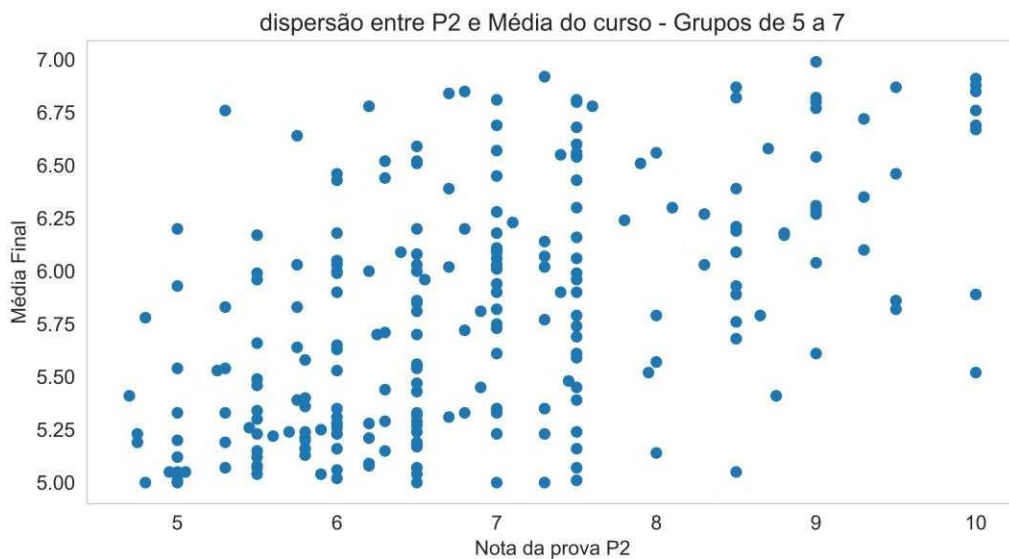


Figura 50 - Dispersão entre nota da prova 2 e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor.

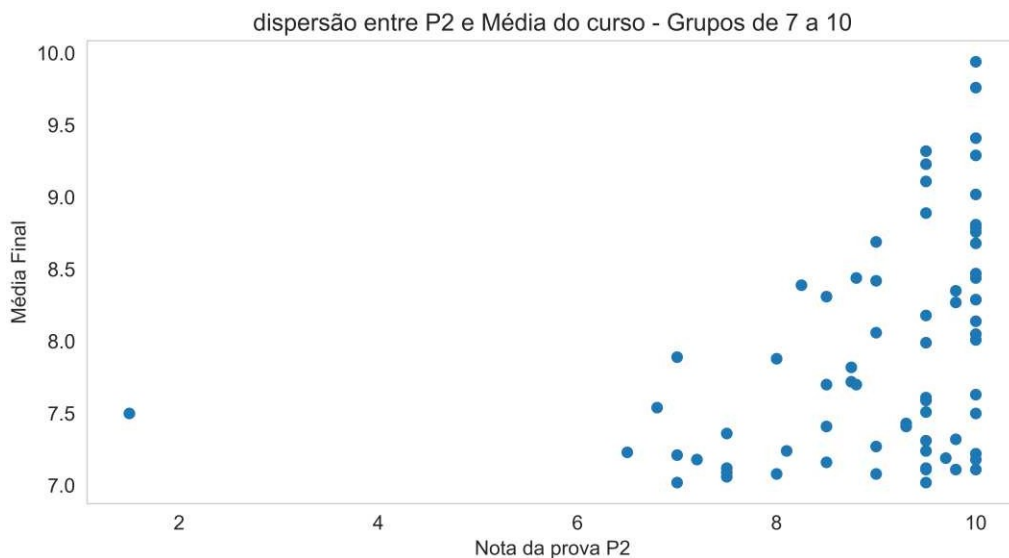


Figura 51 - Dispersão entre nota da prova 2 e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor.

Para as notas dos grupos divididos, nota-se algumas diferenças em relação a nota da prova P1. Na Figura 48, todos os alunos que obtiveram nota final menor do que 3, também tiveram nota na prova P2 abaixo da média mínima exigida, similar ao ocorrido na prova P1. Para as Figuras 49 e 50, nota-se uma distribuição mais homogênea, bem mais acentuada na Figura 50, porém também levemente presente na Figura 49, indicado que, na Figura 49, boa parte dos alunos tiveram nota na P2 perto da média ou acima dela. Já na Figura 50, quase que a totalidade dos alunos que passaram na disciplina (dentro desse grupo) obtiveram nota maior que

5 na prova 2. Na Figura 51, um dado importante indicando que todos os alunos que tiveram nota maior do que 7 na prova P2, tiraram nota maior do que 6 na média final, com exceção de apenas 1 caso. Isso já indicava uma forte correlação entre a prova P2 e a média final.

Abaixo seguem as análises de correlações entre a prova 3 (P3) e média final do curso.

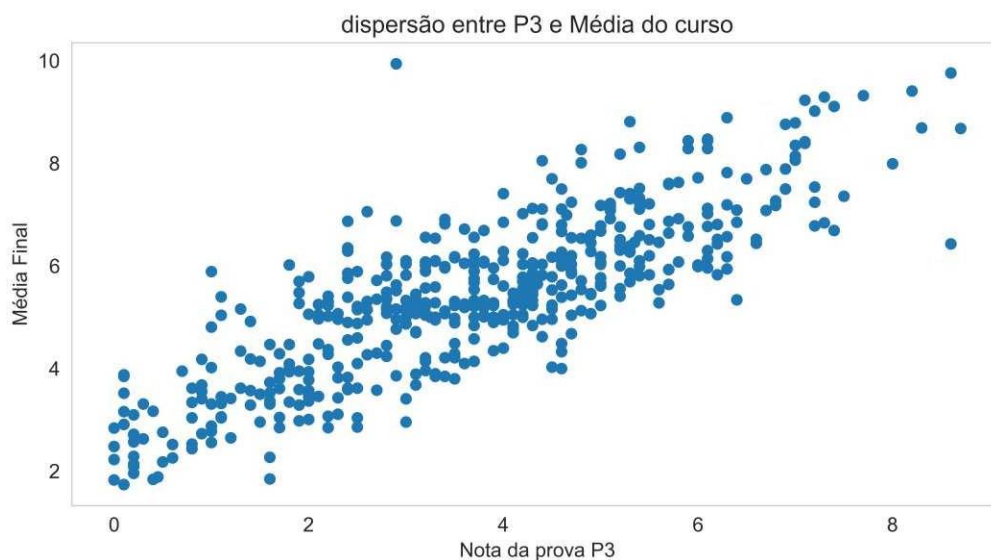


Figura 52 - Dispersão entre nota da prova 3 e média final do curso
Fonte: Elaborada pelo autor.

Assim como na P2, a prova P3 apresenta uma correlação linear positiva com a média final, indicado pela Figura 52. Abaixo estão as correlações para os grupos de notas.

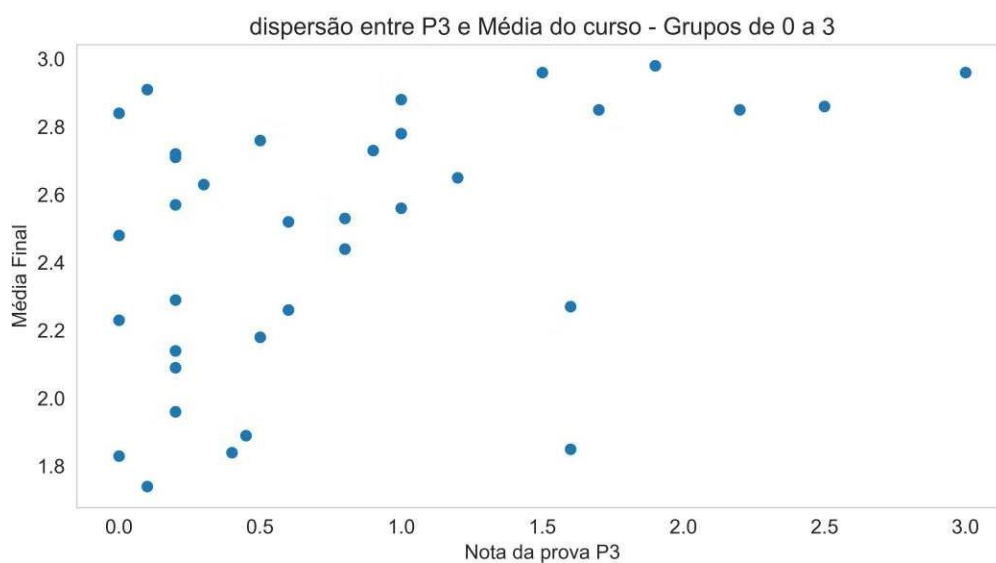


Figura 53 - Dispersão entre nota da prova 3 e média final do curso de 0 a 3.
Fonte: Elaborada pelo autor.

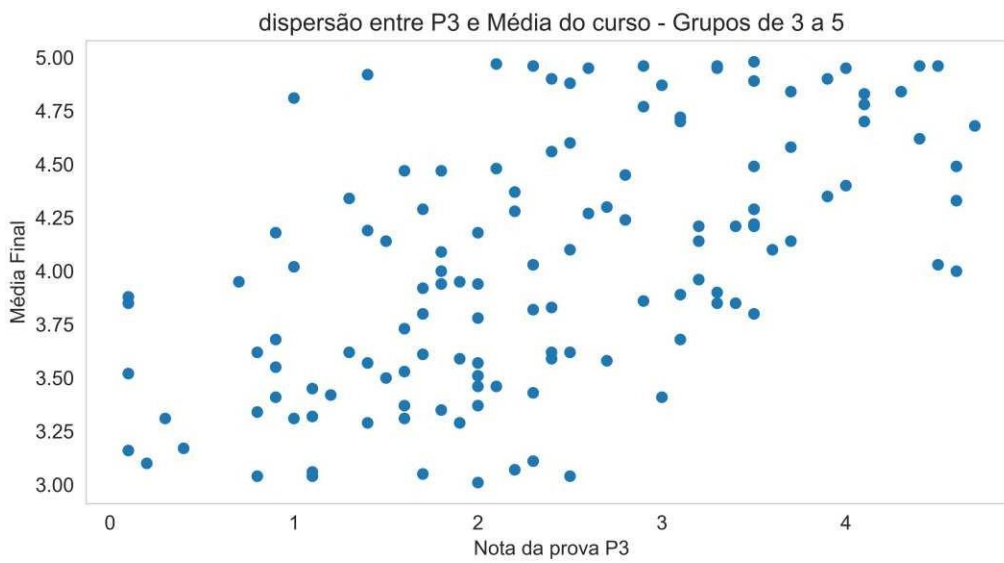


Figura 54 - Dispersão entre nota da prova 3 e média final do curso de 3 a 5
 Fonte: Elaborada pelo autor.

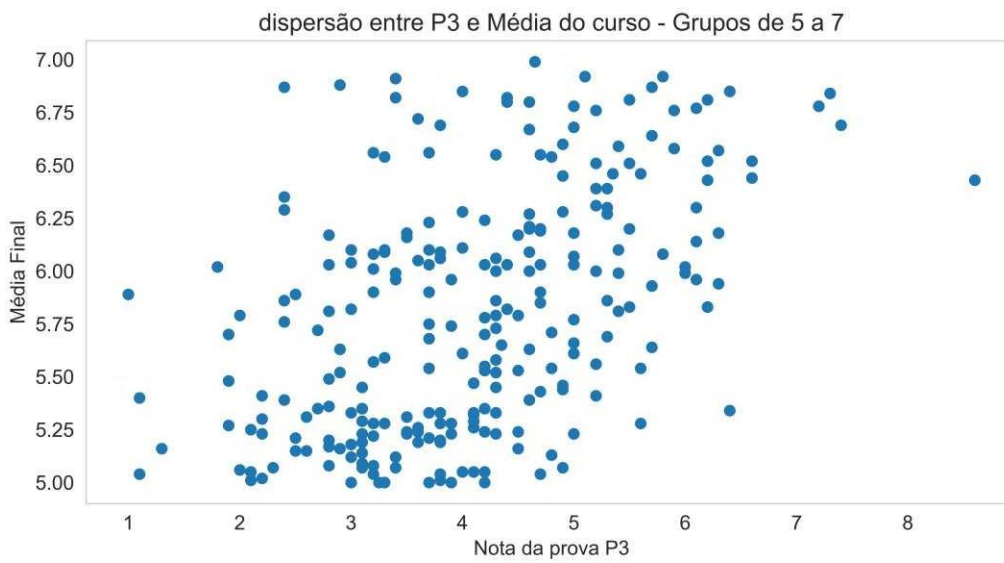


Figura 55 - Dispersão entre nota da prova 3 e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor.

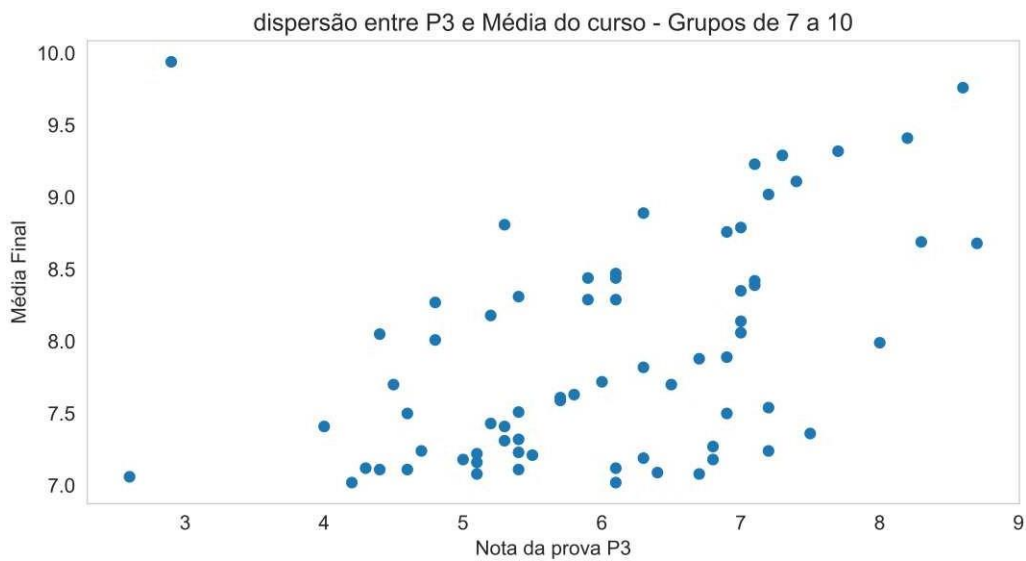


Figura 56 - Dispersão entre nota da prova 3 e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor.

As correlações para a prova P3 e a nota final, obteve os maiores casos de dispersão da nota final em várias regiões. Nas Figuras 53, 54 e 55, temos alunos em diferentes regiões com diferentes notas, com pouquíssimas ocorrências com notas acima de 5 aparecendo somente nos grupos de notas de 5 – 7 (Figura 55) e acima de 7 (Figura 56).

Em seguida, analisamos analogamente as atividades TEC, ADS e ATV online em relação a nota final do semestre.

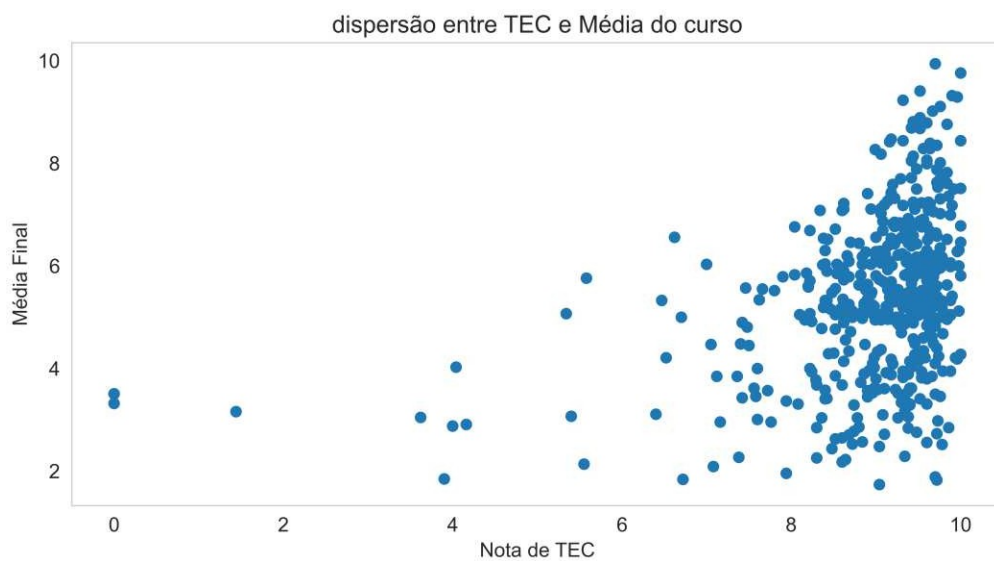


Figura 57- Dispersão entre nota total de TEC e média final do curso
 Fonte: Elaborada pelo autor.

Para o caso em que estão todos os alunos dentro de um único grupo, a relação de Tec e nota final é positiva, mesmo com um significativo número de ocorrências para as notas finais abaixo da média.

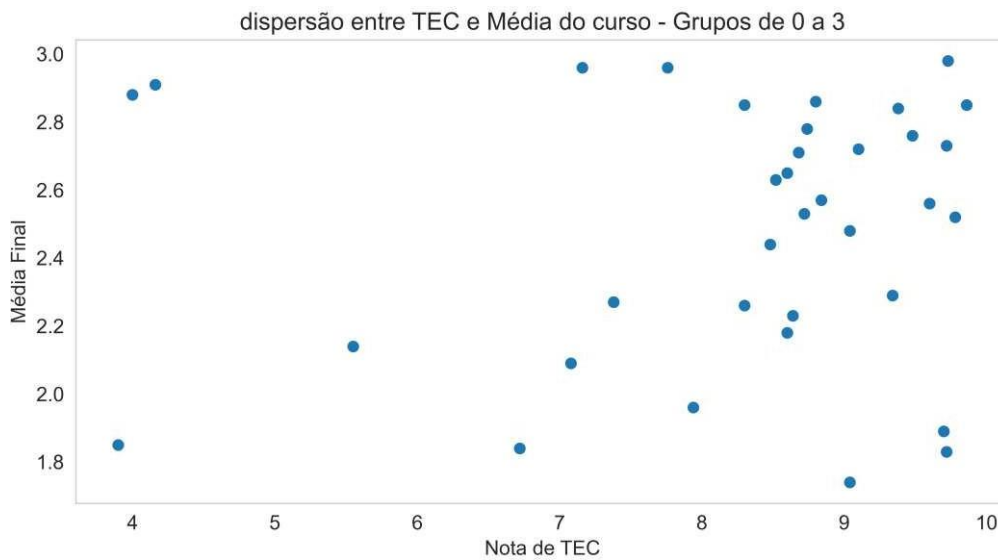


Figura 58 - Dispersão entre nota de TEC e média final do curso de 0 a 3
Fonte: Elaborada pelo autor.

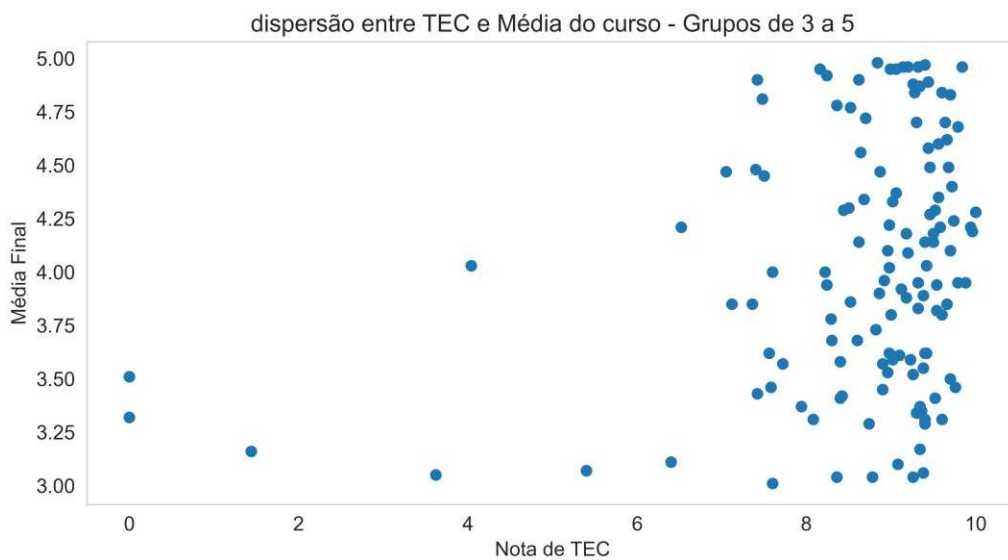


Figura 59 - Dispersão entre nota de TEC e média final do curso de 3 a 5
Fonte: Elaborada pelo autor.

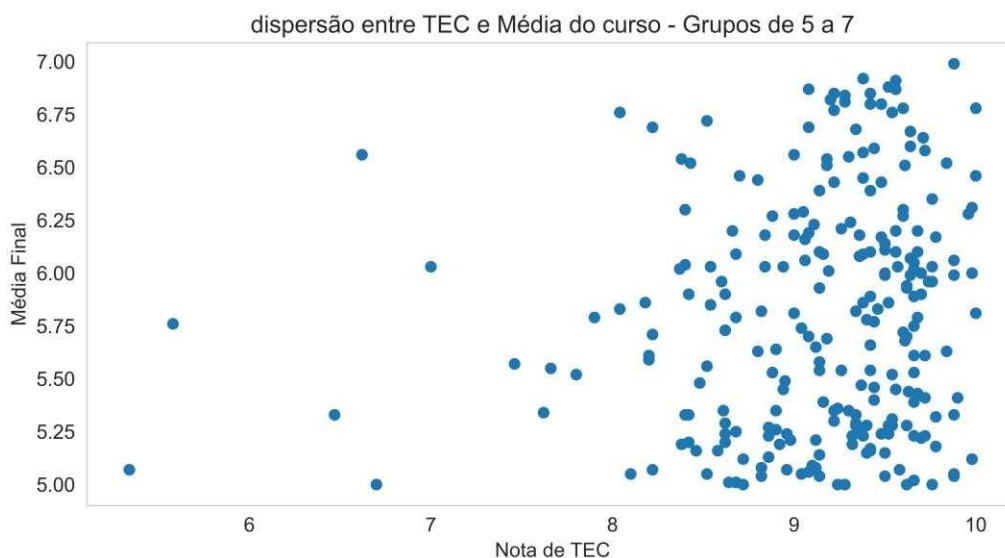


Figura 60 - Dispersão entre nota de TEC e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor.

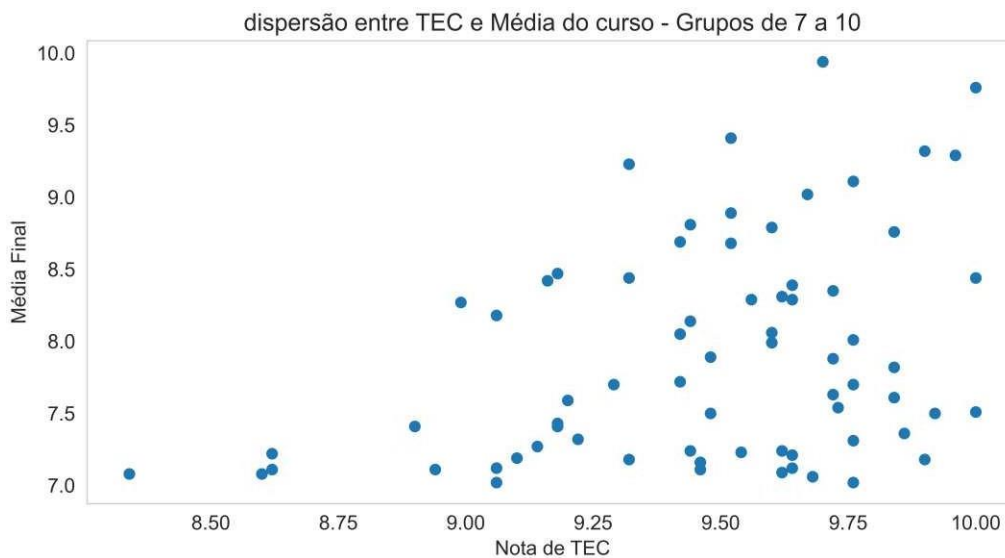


Figura 61 – Dispersão entre nota de TEC e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor.

Também separando as correlações por grupos de notas, percebemos que a grande maioria das ocorrências está concentrado no grupo de notas de 5 a 7 (Figura 60) e com a maioria desse grupo com notas acima de 8 nas atividades, tendo poucos casos abaixo disso e nenhum caso abaixo de 5. Para o grupo da Figura 61, todos os casos estão com notas nas atividades acima de 8.

Analogamente, seguimos para os casos relacionando as atividades de ADS e a nota final

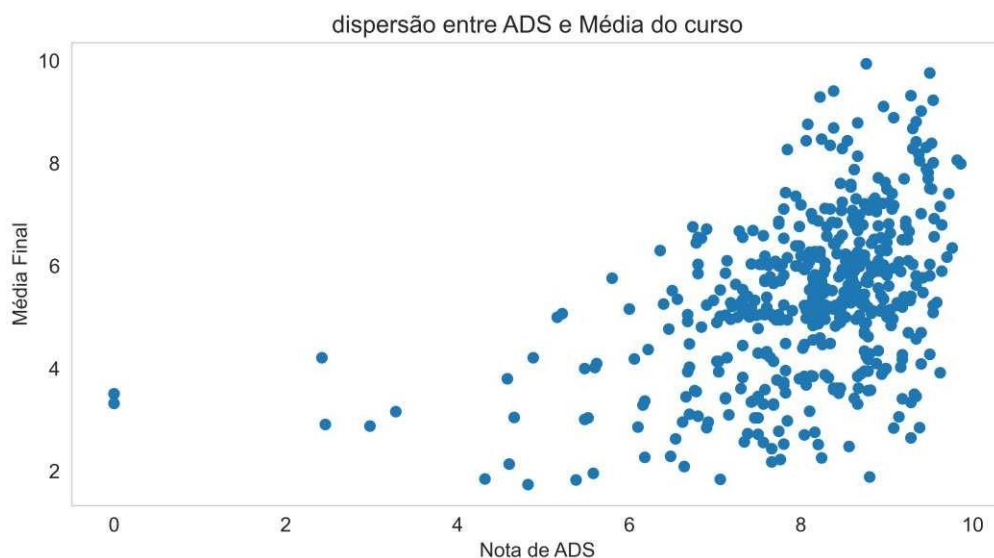


Figura 62 - Dispersão entre nota total de ADS e média final do curso
 Fonte: Elaborada pelo autor.

Gráficos de dispersão entre ADS e Nota total para todos os alunos matriculados com notas de 0 a 10.

Repetindo o padrão observado na atividade de TEC, a grande maioria das notas de ADS estão concentradas acima de 7, sendo uma boa parte para os alunos com nota final acima da nota mínima para aprovação.

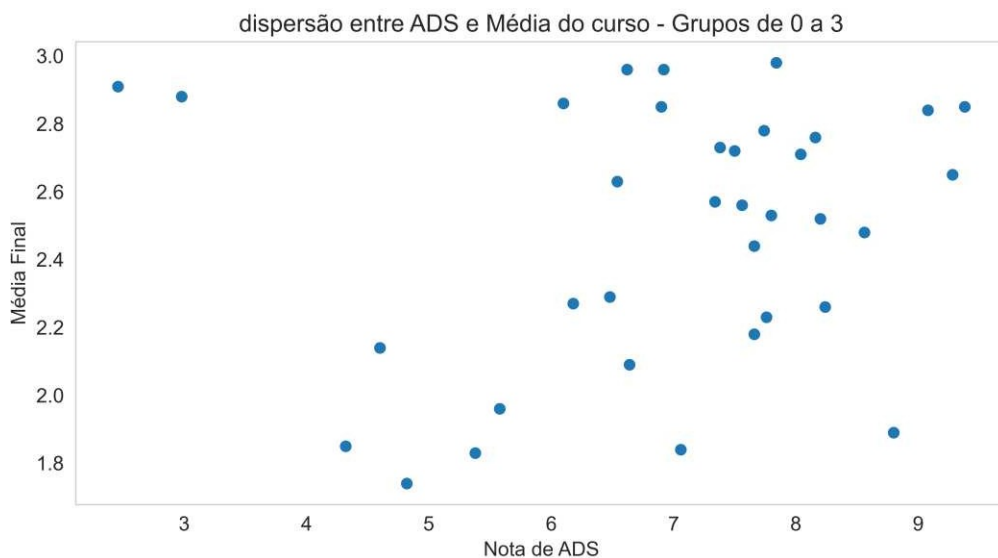


Figura 63 - Dispersão entre nota de ADS e média final do curso de 0 a 3
 Fonte: Elaborada pelo autor.

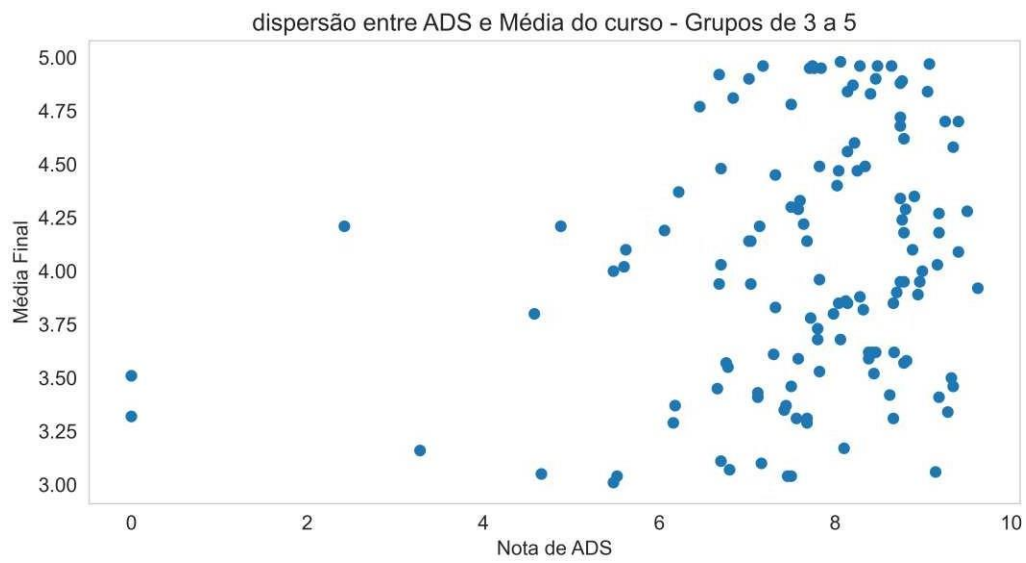


Figura 64 - Dispersão entre nota de ADS e média final do curso de 3 a 5
 Fonte: Elaborada pelo autor.

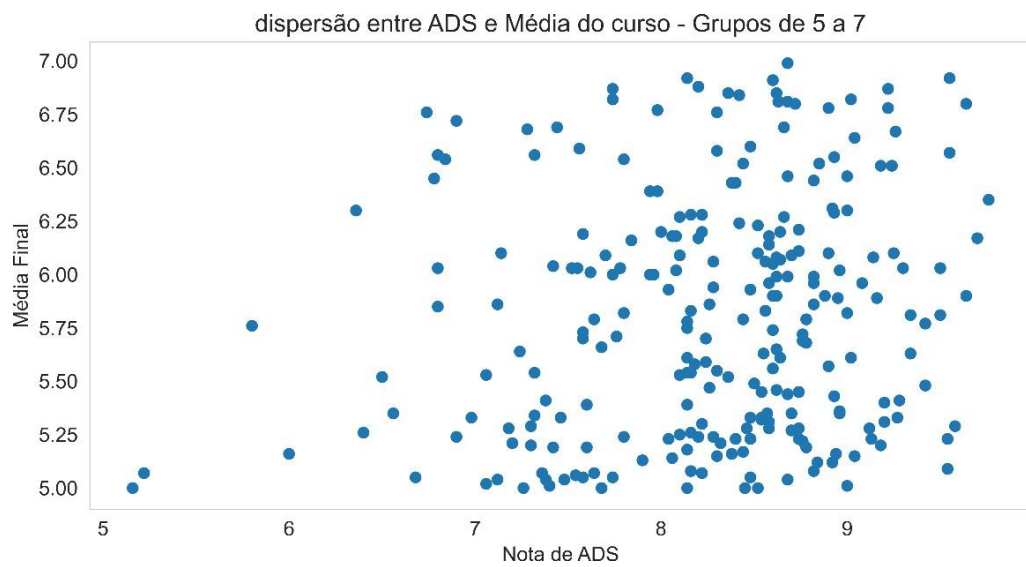


Figura 65 - Dispersão entre nota de ADS e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor.

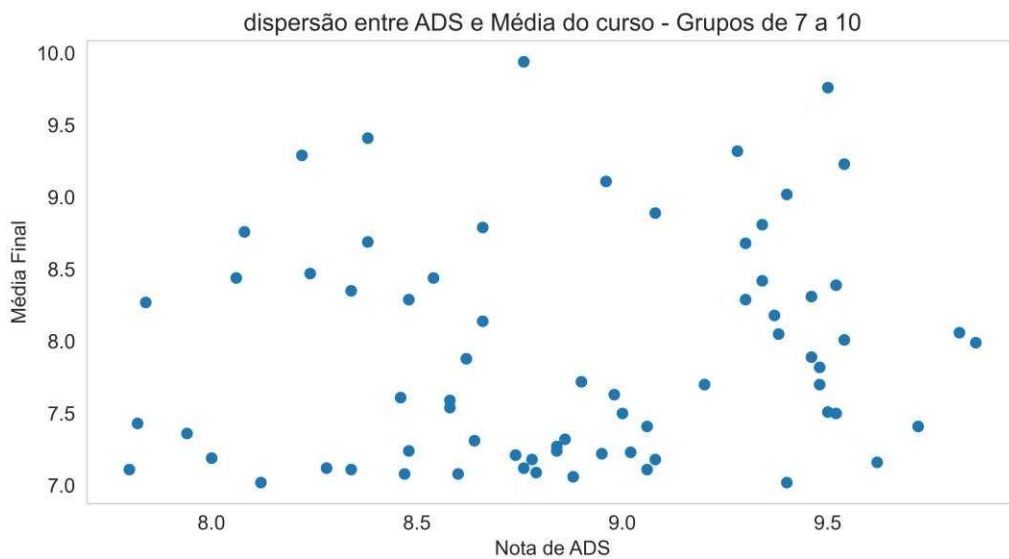


Figura 66 - Dispersão entre nota de ADS e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor.

Gráficos de dispersão entre ADS total e Nota total, separados em grupos de notas de 0-3, 3-5, 5-7, 7-10

Para os grupos das Figuras 63 e 64, observamos uma dispersão das notas entre 3 e 10 para as atividades ADS. Já para o grupo das Figuras 65 e 66, grupos dos alunos aprovados na disciplina, a grande maioria obteve nota na atividade ADS acima de 7, com alguns casos entre 5 e 7, porém nenhum caso abaixo disso.

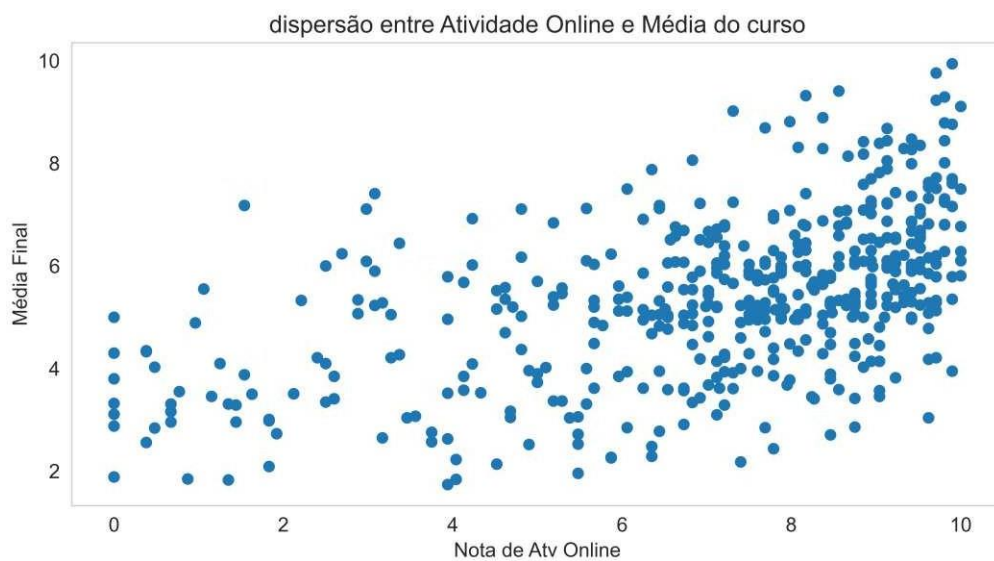


Figura 67 - Dispersão entre nota de Atv Online e média final do curso
 Fonte: Elaborada pelo autor.

As Atv Online seguem um padrão mais linear em relação as atividades TEC e ADS, observado pela Figura 67, onde contempla todos os alunos da disciplina.

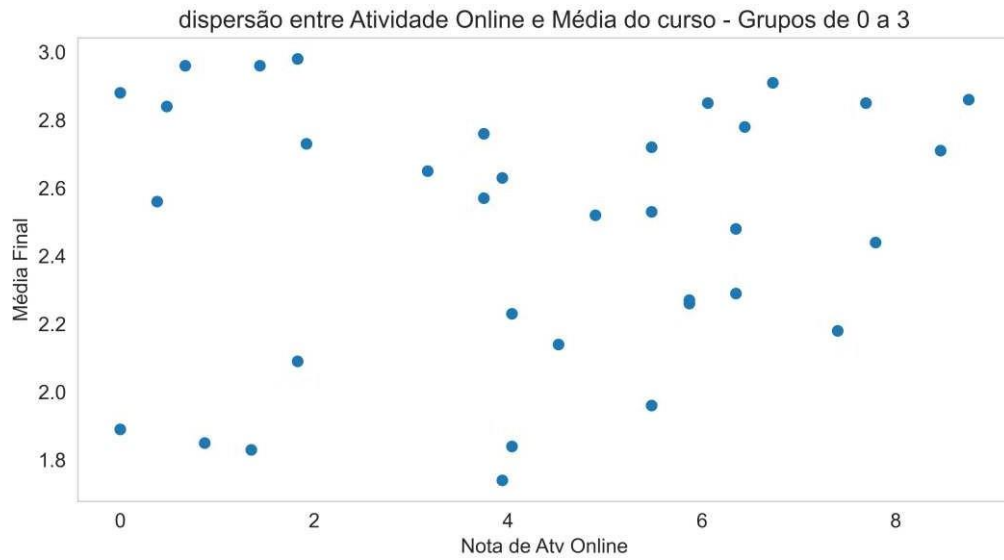


Figura 68 - Dispersão entre nota de Atv Online e média final do curso de 0 a 3
Fonte: Elaborada pelo autor.

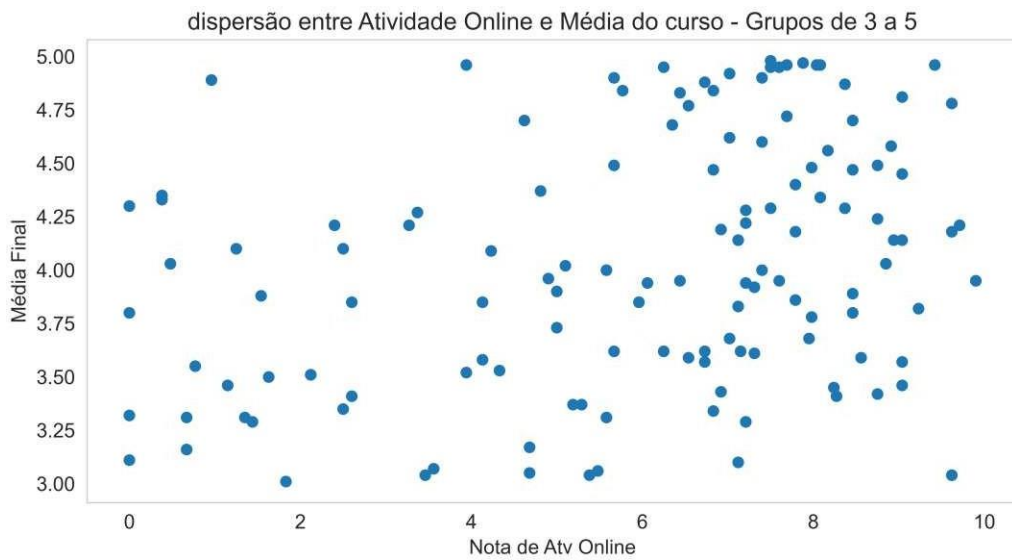


Figura 69 - Dispersão entre nota de Atv Online e média final do curso de 3 a 5
Fonte: Elaborada pelo autor.

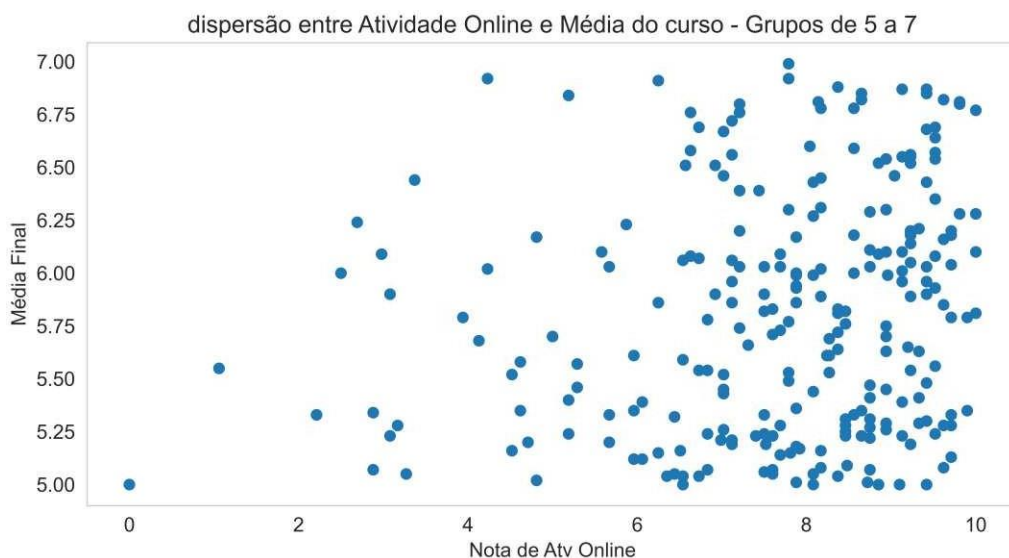


Figura 70 - Dispersão entre nota de Atv Online e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor

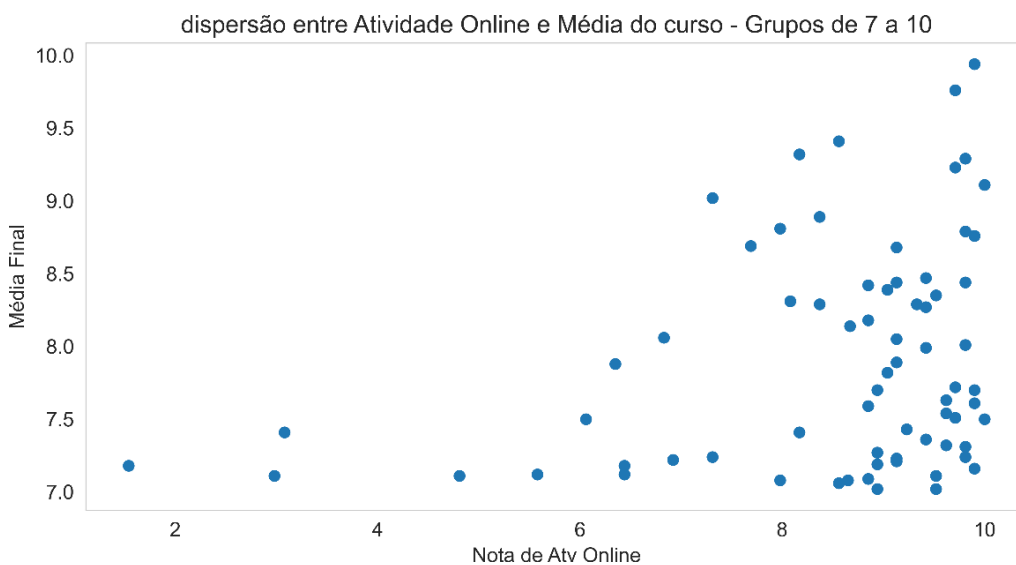


Figura 71 - Dispersão entre nota de Atv Online e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor

Observa-se que, diferentemente das outras atividades, nas atividades online há uma ocorrência maior nos grupos das Figuras 68 e 69, com notas nas atividades abaixo de 5. Já nos grupos das Figuras 70 e 71, a grande maioria dos alunos obtiveram nota nas atividades acima de 5, com poucos casos abaixo.

Em seguida, apresentamos o mesmo padrão de correlações, agora para as notas das atividades como um todo, após aplicado os critérios estabelecidos e apresentados no Capítulo 3, Seção 3.1.

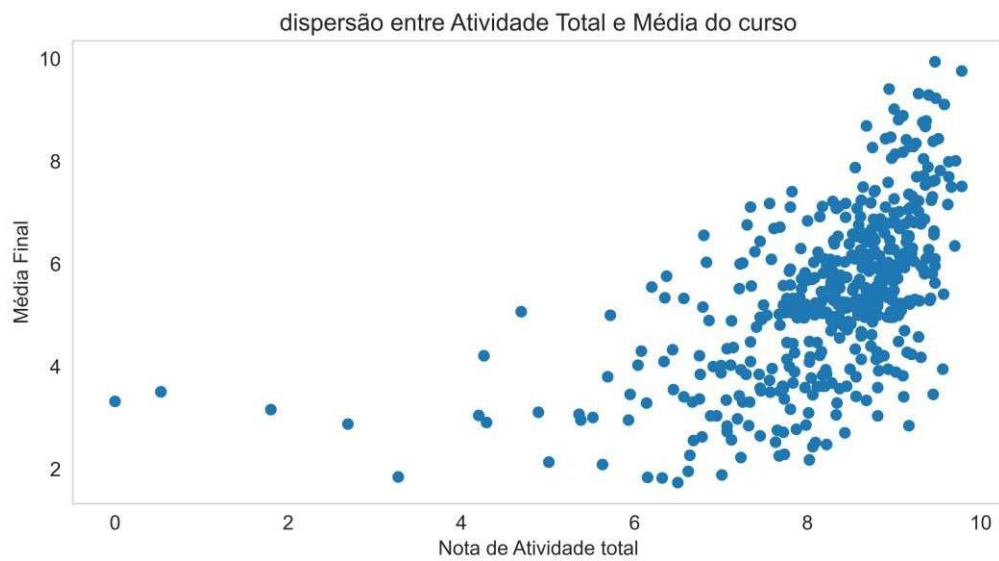


Figura 72 - Dispersão entre nota de Atv Total e média final do curso
 Fonte: Elaborada pelo autor

Pela Figura 72, nota-se que a maioria dos alunos obtiveram notas nas atividades acima de 6, com poucos casos abaixo.

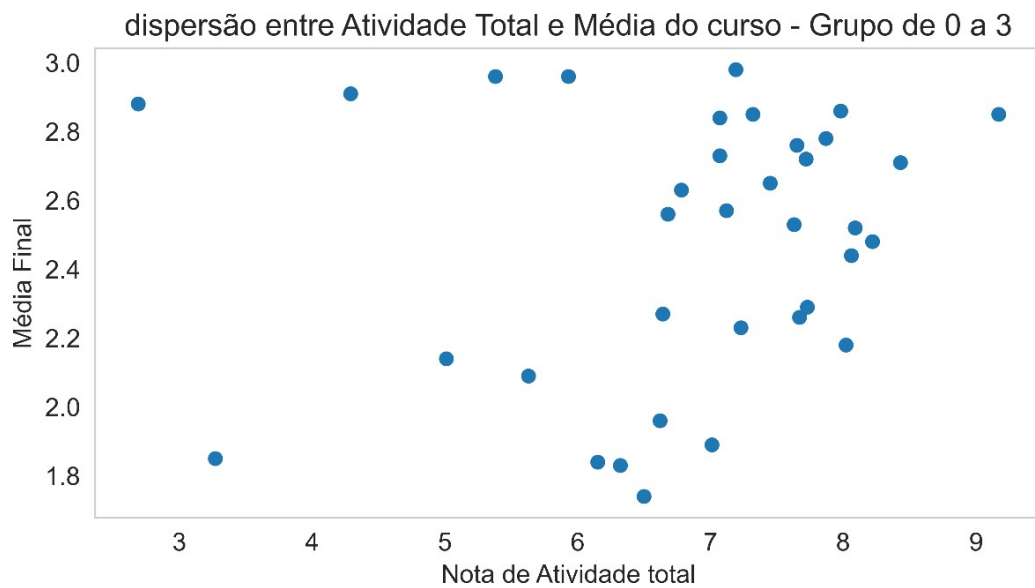


Figura 73 - Dispersão entre nota de Atv Total e média final do curso de 0 a 3
 Fonte: Elaborada pelo autor.

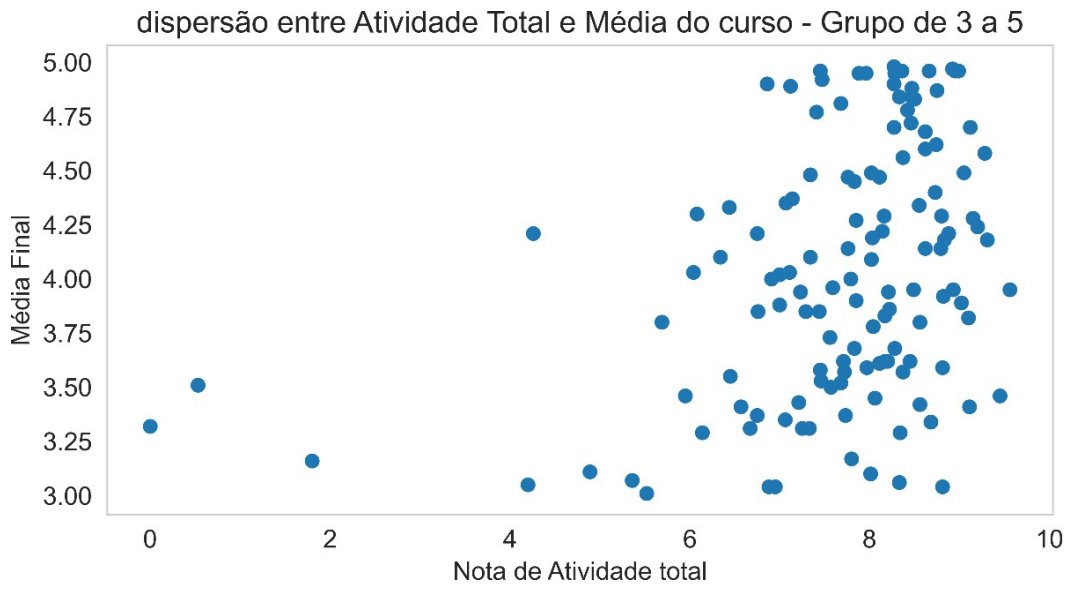


Figura 74 - Dispersão entre nota de Atv Total e média final do curso de 3 a 5
 Fonte: Elaborada pelo autor

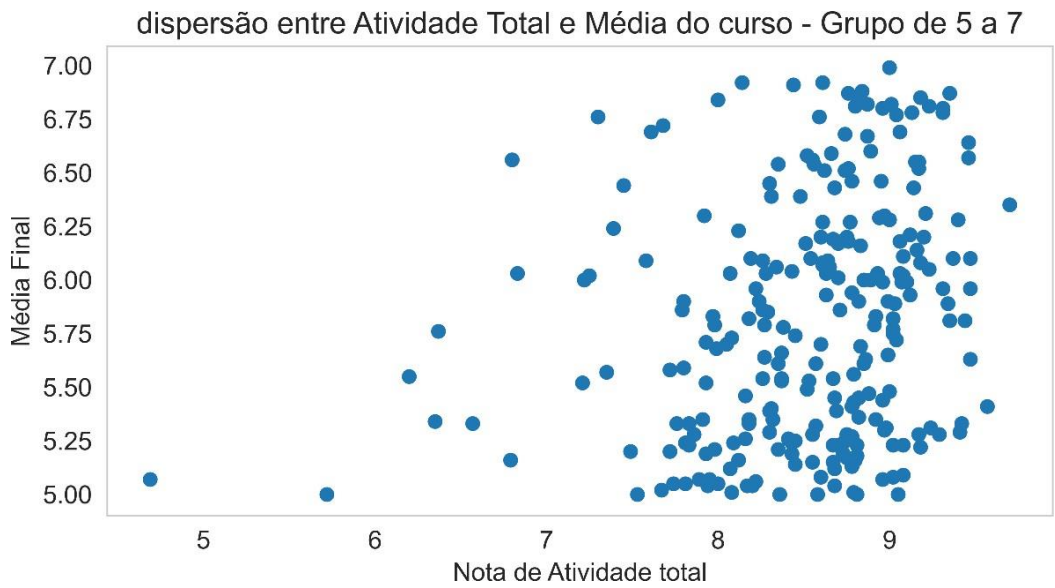


Figura 75 - Dispersão entre nota de Atv Total e média final do curso de 5 a 7
 Fonte: Elaborada pelo autor

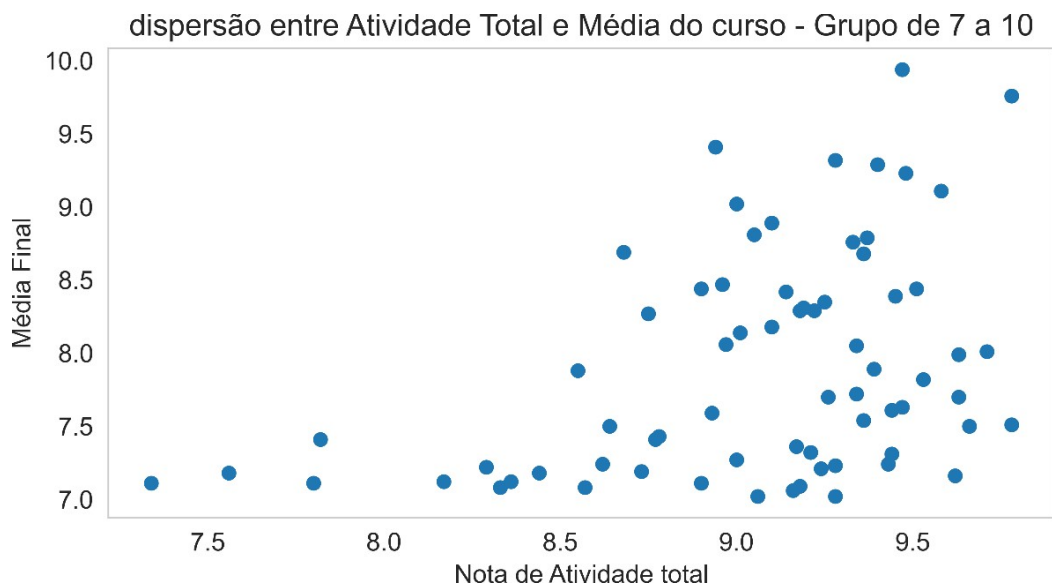


Figura 76 - Dispersão entre nota de Atv Total e média final do curso de 7 a 10
 Fonte: Elaborada pelo autor

Para os grupos das Figuras 73 e 74, encontra-se ocorrências em todas as notas, porém um grupo menor de ocorrências. Para os grupos das Figuras 75 e 76, têm-se ocorrências quase que para as regiões acima de 5, com algumas poucas abaixo disso no grupo 2 e nenhuma no grupo 3. No grupo 4, todos os casos estão com notas nas atividades acima de 7.

Na Figura 77, apresentamos um resumo dessas dispersões contabilizando todos os alunos, entretanto, foi utilizado a média das provas e atividades juntamente com a média final do curso.

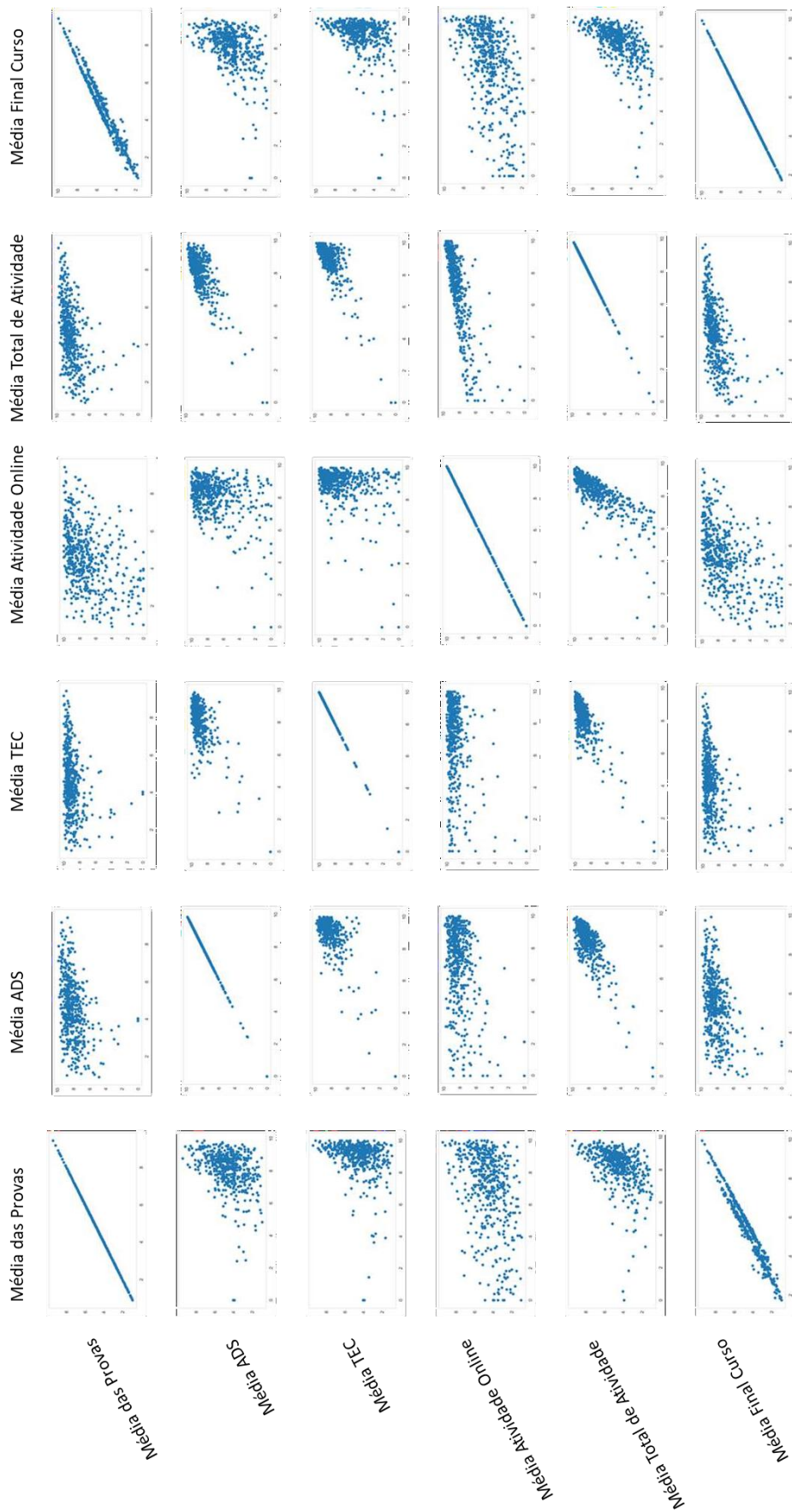


Figura 77 - Dispersão das notas Total de Provas, ADS, TEC, Ativ. online, Atividade total e nota final de curso
 Fonte: Elaborada pelo autor

A Figura 77 mostra as dispersões entre notas Total de Provas, ADS, TEC, Atividade online, Atividade total e nota final de curso e ascorrelações mais positivas e lineares ficam entre as medias de prova e média final, uma vez que as provas equivalem a 85% da média final.

Também percebemos que para a média das atividades, as ADS e TEC tem uma correlação mais positiva, o que corrobora com o que foi apresentado no Capítulo 4, Seção 4.1. O gráfico acima também apresenta uma correlação positiva entre as atividades e a média final, onde quase a totalidade dos alunos que passaram na disciplina, obtiveram nota maior do que 5 nas atividades.

A Figura 78 apresenta um estudo de correlações entre a média final da disciplina, notas em provas e as presenças em ADS e presença total na disciplina.

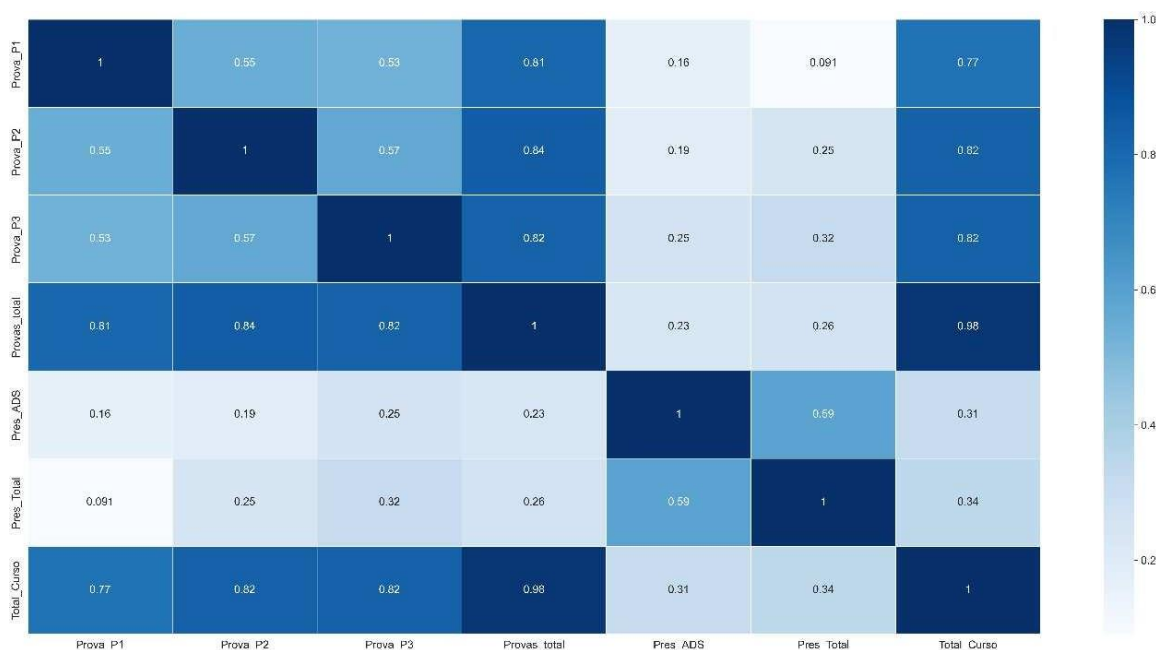


Figura 78 - Heatmap entre as notas obtidas em Provas (maior peso) e média final em relação as presenças em ADS e presença total.

Fonte: Elaborada pelo autor

O HeatMap apresentado na Figura 78 apresenta uma correlação baixa entre as presenças e as notas em prova e média final. Nota-se pelo gráfico, que coloração das colunas e linhas entre presenças e notas é significativamente clara, indicando pouca correlação entre esses parâmetros.

Abaixo analisamos a relação de dispersão entre as atividades, e assim como no heatmap acima, as relações de dispersão entre as frequências e as notas de provas e nota total são muito baixas, indicando pouca correlação direta entre esses parâmetros.

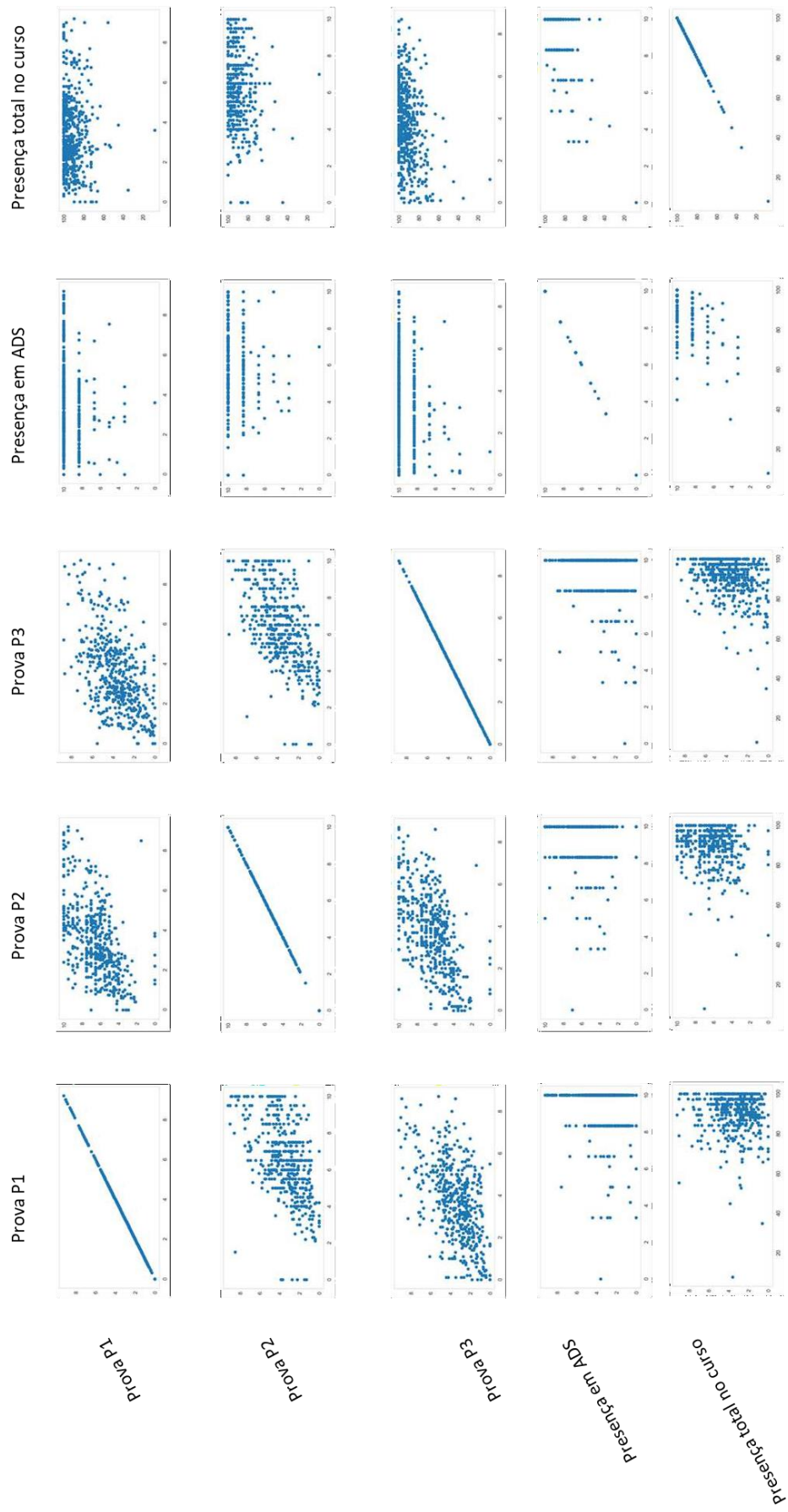


Figura 79 - Dispersão entre notas Total de Provas, Presença em ADS e presença total no curso.
 Fonte: Elaborada pelo autor.

5 CONCLUSÕES

A motivação deste trabalho teve como ponto de partida entender o desempenho de estudantes em uma disciplina do curso básicas de física para, eventualmente, poder criar modelos que possam ajudar a compreender e contribuir para maior eficácia em outras disciplinas, tanto no ciclo básico (Ensino Médio) como nos cursos de graduação. Isso não é fácil, e exige muitos dados, análises e tentativas e erros. Este trabalho é apenas um primeiro passo neste sentido.

Partindo desta motivação, definiu-se como um primeiro objetivo analisar os resultados dos discentes do curso integrado de física no Instituto de Física de São Carlos-USP, após as mudanças desta disciplina devido a mudança na grade curricular do IFSC, ocorrida em 2017. Essa nova disciplina introduziu novas metodologias de ensino concomitantemente a metodologia tradicional já adotada anteriormente na disciplina. Embora seja uma análise (parcial) de um processo de ensino, o foco deste trabalho foi realizar uma análise além de qualitativa, mas também quantitativa, seguindo uma metodologia científica e analítica.

Foram realizadas, como já apresentado no Capítulo 3, uma densa coleta de dados referentes às atividades e o desempenho de estudantes, nas mais diversas atividades propostas pelos docentes da disciplina ao longo do semestre. As atividades foram realizadas em diferentes formatos e plataformas de aplicação, baseados nas metodologias adotadas.

Esse banco de dados construído, relativamente extenso, necessitou da utilização de ferramentas e técnicas de análise mais profundas que as convencionais e simples estatísticas descritivas básicas (média, desvio padrão, etc), assim utilizou-se conceitos e métodos das áreas de *Big Data*, *Data Science* e *Learning Analytics* (LA). Com relação à LA, explicitamente, é preciso admitir que pouco foi explorado, na prática, por limitações de tempo do autor (por exemplo, no sentido de criação de modelos e previsões de comportamento), mas, ainda assim, o espírito de análise quantitativa e o uso de ferramentas com foco nos resultados de aprendizagem estão presentes aqui e esperamos que este primeiro trabalho estimule mais estudos nessa área no futuro.

Todas as análises foram direcionadas para responder, a princípio, duas questões iniciais, relacionadas ao desempenho dos alunos, que por sua vez estão relacionados a problemas clássicos do componente curricular que é o elevado número de reprovações, provocando um segundo problema: o número de evasões no curso.

As perguntas norteadoras para as análises foram:

- Quais atividades possibilitam um maior (ou menor) engajamento?
- Quais os perfis dos alunos aprovados e reprovados na disciplina?

Neste sentido, foram realizadas diversas análises, a princípio de resultados mais simples, porém significativos sobre as turmas analisadas.

Verificou-se, por exemplo, que este grupo obteve uma taxa de aprovação de 58% dos alunos, o que corresponde a 327 alunos aprovados e 237 alunos reprovados, quando são considerados todos os estudantes matriculados na disciplina.

Importante ressaltar que nessa estatística de aprovação/reprovação está sendo contabilizado os alunos evadidos (desistências) durante o semestre. Isto é, aqueles que não realizaram o pedido de trancamento da matrícula, mesmo tendo desistido de participar na disciplina ainda no início do curso. Do ponto de vista do número de alunos matriculados, estes estão, portanto, na estatística dos matriculados no sistema da USP. Como os dados indicam, talvez esse não seja o melhor número para indicar a taxa de sucesso dos alunos, dado que muitos desistiram ainda no início do curso, deixando de fazer atividades, ou não realizando as provas. As razões dessa desistência são variadas e não foi objeto de estudo neste primeiro trabalho, pois envolve muitas variáveis e acompanhamento individual dos alunos, que nem sempre querem colaborar. Portanto, o foco foi observar os dados que permaneceram. Ainda assim, é importante ter em mente o número significativo de estudantes que parece ter simplesmente desistido da disciplina em algum momento. Uma estimativa simples, baseada por exemplo na desistência em fazer a P3, resulta em cerca de 45 alunos, que corresponde a cerca de 8% do número total. Se adotássemos esse critério, por exemplo, o índice de aprovação seria cerca de 66%, apenas para comparação de alguns limites extremos.

Esse resultado sobre a taxa de aprovação, simplesmente descritivo da turma, não indica evidências significativas a respeito do comportamento ou perfil dos alunos, entretanto, deixa claro que boa parte dos alunos ainda não conseguem aprovação na disciplina.

Analisando algumas medidas estatísticas a respeito dos dados coletados, foi observado informações a respeito do grupo de estudantes analisado. No sentido de tentar entender melhor o comportamento dos alunos em diferentes momentos da disciplina, foram realizadas análises dos resultados das provas, identificando que pouco mais de 70% dos discentes obtiveram nota inferior a 5,0 nas provas P1 e P3, enquanto que na prova P2 foram quase que 75% de discentes com nota maior que 5 pontos.

Este resultado levantou questionamentos a respeito do comportamento dos alunos durante o semestre, mais propriamente em relação à variação de desempenho entre as provas P1 e P2 e as provas P2 e P3. O que pode ter acontecido para uma melhora significativa entre as duas primeiras provas e também o que pode ter ocorrido para uma regressão na nota entre a segunda e terceira prova. Outro questionamento foi sobre o número de provas realizadas, uma vez que nas provas P1 e P2, teve-se 536 provas realizadas e na prova P3, somente 490, uma queda de 46 provas, pouco menos de 10% do número total, um decaimento significativo que, por falta de tempo e dados, não foi estudado para entender o motivador da queda nas realizações das provas. Este resultado, a princípio, parece ser preocupante, pois a prova possuía o maior peso na nota final da disciplina.

Esse aumento significativo da prova 1 em relação à prova 2, foi acompanhado também de um aumento significativo do número de atividades TEC e ADS realizadas, levantando um questionamento sobre a influência das atividades no desempenho da prova.

Ainda analisando o resultado das provas, agora a nota total das provas incluindo a prova substitutiva, também se observou que 25% dos alunos estão na região de notas abaixo de 3,0 pontos. Região que, pelo regimento da universidade, não dá o direito de pleitear uma prova de recuperação ao final do semestre.

Alguns docentes comentaram ter a impressão de observar um certo pragmatismo, principalmente nos estudantes de engenharia, que têm um número muito grande de créditos no semestre e, portanto, priorizam outras disciplinas no final do semestre, depois de ter conseguido garantir a nota mínima para aprovação na disciplina. Esse comportamento, porém, precisa ser estudado com mais atenção e mais dados, o que não foi feito neste primeiro estudo.

O número de atividades, ADS, TEC e Atividades online entregues também caiu significativamente ao longo do semestre. Foram 501 atividades ao final do semestre, em um total de 6 aplicações, sendo o quinto menor número de participações e ficando acima apenas da primeira aplicação. Em contrapartida, foram 527 participações em seu máximo. O número de participações em TEC's foi mais homogêneo, porém no final do semestre também teve uma redução de atividades realizadas, sendo 527 o máximo e 506 o número na última aplicação.

Essa diminuição nas participações, tanto em provas como nas atividades, pode indicar um abandono da disciplina por parte dos discentes, seja por ter abandonado o curso como um todo ou até mesmo um abandono apenas da disciplina ou, talvez, numa certa fração de estudantes o pragmatismo mencionado antes.

Os fatores que podem influenciar e/ou motivar este abandono não foram estudados neste trabalho e cabe aqui uma sugestão de questão de análise para um próximo trabalho. Abaixo, elencamos algumas sugestões:

- Baixo rendimento nas provas 1 e 2, impossibilitando a recuperação na prova 3 para uma aprovação direta, seguida de desistência. E até que ponto isso é uma percepção realista, considerando que há a possibilidade de uma prova substitutiva e ainda a recuperação.
- Desistência na etapa final, pelo acúmulo de provas com outras disciplinas.
- Não adaptação à rotina de estudos, com vários elementos avaliativos.
- Será que há mesmo o pragmatismo relatado de forma anedótica por alguns docentes?

Os fatores elencados acima cabem numa futura análise, além da investigação com mais detalhamento sobre os possíveis fatores para o abandono.

Foram observadas também pequenas correlações entre as notas das atividades e provas com a presença dos alunos. Usualmente, na metodologia tradicional, com aula puramente expositiva, têm-se uma correlação positiva de presença e nota, uma vez que todo o conteúdo é apresentado somente em sala de aula. Nas metodologias ativas, um dos objetivos é justamente a variação da exposição dos conteúdos além da forma como são trabalhados, estimulando que o aluno seja o protagonista do seu aprendizado. Dessa forma, possibilita o discente desenvolver seu aprendizado em diferentes momentos fora da sala de aula. Neste sentido, a baixa correlação observada talvez não seja um indicador necessariamente negativo, indicando que boa parte dos discentes estão conseguindo desenvolver autonomia de aprendizado, fora da aula. Isso pode, também, ser um efeito isolado desta amostra e, portanto, mais estudos são necessários para entender se essa é mesmo uma variável relevante na análise.

No âmbito do estudo fora do horário de aula presencial, as atividades TEC, ADS e Atividades online têm papel significativo no processo de aprendizagem dos alunos, de acordo com a metodologia utilizada. Essa relevância também pode ser verificada nos dados que mostram as notas das atividades divididas por grupos. Nas atividades ADS e TEC, nos grupos com notas maiores, grupo 3 ($5 > \text{nota final} < 7$) e grupo 4 ($\text{nota final} > 7$), observou-se que todos os alunos também tinham notas acima de 5 nessas atividades. Importante destacar que somente essas atividades não garantem o sucesso dos discentes na disciplina, porém estes resultados mostram um importante papel destas no processo de evolução do discente.

Um resultado não esperado foi a baixa correlação das presenças em aula ou atividades com a nota final de aprovação. Como neste primeiro estudo, não foi analisado caso a caso de alunos separadamente, mas sim o grupo como um todo, essa baixa correlação de presenças pode ter sofrido influência do grupo geral, ficando uma sugestão de análise para um próximo estudo.

Também notou-se que o número de participações nas atividades Online (ADS, TEC, etc.), durante quase todo o semestre, não coincide com o número de provas realizadas, ficando sempre abaixo do número de provas realizadas, indicando que alguns alunos não realizam as atividades embora estivessem frequentando as aulas presenciais e realizando as provas.

Este comportamento, entretanto, inverte ao final do semestre, onde o número de atividades ADS, por exemplo, é superior ao número de provas realizadas, 506 ADS realizadas contra 490 provas realizadas, evidenciando mais uma vez o abandono por parte dos alunos.

O principal objetivo deste trabalho era analisar os dados e observar, quando possível e se existente, padrões no desempenho dos discentes. Alguns padrões identificados foram:

- O abandono da disciplina ou curso ao final do semestre, evidenciado menor número de provas realizadas em comparação com as provas anteriores.

- Aumento significativo do resultado na prova 2, fator a ser analisado em um próximo estudo, buscando identificar qual ou quais fatores colaboraram para a melhora no desempenho.

- Diminuição do desempenho na prova 3. Qual ou quais fatores colaboraram para uma queda no desempenho após a prova 2.

- Baixa correlação das presenças em aula e atividades com a nota final.

Embora não sendo o objetivo deste trabalho propor soluções para os padrões encontrados, o que exige uma análise mais bem complexa e multifatorial, a caracterização destes, através de análises quantitativas, é e foi de extrema importância para uma visão mais detalhada do comportamento da disciplina além de indicar novas possibilidades e caminhos para a continuação desta pesquisa em um trabalho futuro.

Neste sentido, alguns pontos a serem observados em um próximo estudo são:

- Analisar o comportamento individual dos alunos durante o desenvolvimento da disciplina, por exemplo, fazendo uso de algoritmos específicos na análise dos dados do AVA.

- Identificar possíveis comportamentos que corroborem a diminuição do engajamento ao longo do semestre. Quais fatores contribuíram para a diminuição das notas na prova 3?

- Buscar correlações de presença e desempenho individual dos alunos.

- Identificar o nível de engajamento dos alunos nas atividades Online.

REFERÊNCIAS

- 1 BRASIL. Ministério da Educação. Conselho Nacional de Educação. PARECER CNE/CES 1.304/2001-Homologado Despacho do Ministro em 4/12/2001. Dispõe sobre Diretrizes Nacionais Curriculares para os Cursos de Física. **Diário Oficial da União**, Brasília, DF, 7 dec. 2001. Seção 1, p. 25.
- 2 BRASIL. Ministério da Educação. Secretaria de Educação Básica. **Orientações curriculares para o ensino médio**. Brasília, 2006.
- 3 BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Diretoria de Avaliação. **Relatório de Avaliação 2013-2016: ensino**. Disponível em: https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/avaliacao/relatorio_quadrienal_ensino.pdf. Acesso em: 23 jan. 2022
- 4 KUBO, O. M.; BOTOMÉ, S. P. Ensino-aprendizagem: uma interação entre dois processos comportamentais. **Interação em Psicologia**, v. 5, n. 1, 31 Dez. 2001.
- 5 FREIRE, P. **Pedagogia da autonomia: saberes necessários à prática educativa**. São Paulo: Paz e Terra, 1996.
- 6 BEZERRA, D. P. *et al.* A evolução do ensino da física – perspectiva docente. **Scientia Plena**, v. 5, n. 9, 2009.
- 7 MORAN, J. Metodologias ativas e modelos híbridos na educação. **Novas tecnologias digitais: reflexões sobre mediação, aprendizagem e desenvolvimento**. Curitiba: CRV, 2017.p.23-35.
- 8 MÜLLER, M. G. *et al.* Uma revisão da literatura acerca da implementação da metodologia interativa de ensino Peer Instruction (1991 a 2015). **Revista Brasileira de Ensino de Física**, v. 39, n. 3, Mar. 2017.
- 9 VALENTE, J. A.; BIANCONCINI DE ALMEIDA, M. E.; FLOGI SERPA GERALDINI, A. Metodologias ativas: das concepções às práticas em distintos níveis de ensino. **Revista Diálogo Educacional**, v. 17, n. 52, p. 455, Junho 2017.
- 10 SCAICO, P. D.; DE QUEIROZ, R. J. G. B.; SCAICO, A. **O conceito big data na educação**. In: WORKSHOP DE INFORMÁTICA NA ESCOLA (WIE 2014), 20., 2014. **Anais [...]**. Dourados: SBC, 2014.
- 11 CHEN, M.; MAO, S.; LIU, Y. Big data: a survey. **Mobile Networks and Applications**, v. 19, n. 2, p. 171–209, Abr. 2014.
- 12 FILATRO, A. **Data science na educação: presencial, a distância e corporativa**. São Paulo: Saraiva Educação, 2021.

13 COX, M.; ELLSWORTH, D. **Application-controlled demand paging for out-of-core visualization**. 1997. Disponível em: <https://ieeexplore-ieee-org.ez67.periodicos.capes.gov.br/stamp/stamp.jsp?tp=&arnumber=663888>. Acesso em: 23 jan. 2022.

14 BALDASSARRE, M. Think big: learning contexts, algorithms and data science. **Research on Education and Media**, v. 8, n. 2, p. 69–83, Dec. 2016.

15 SIEMENS, G. Learning analytics: the emergence, **American Behavioral Scientist**, v. 57, n. 10, p. 1380– 1400, Oct. 2013.

16 DIETZ-UHLER, B.; HURN, J. E. Using learning analytics to predict (and Improve) student success: a faculty perspective. **Journal of Interactive Online Learning**, v. 12, n. 1, p. 17–26, 2013.

17 SILVA, L. A. *et al.* Ciência de dados educacionais: definições e convergências entre as áreas de pesquisa. *In*: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 4., 2017, Recife. **Anais** [...]. Recife: 2017. DOI: 10.5753/cbie.wcbie.2017.764.

18 PIETY, P. J.; HICKEY, D. T.; BISHOP, M. J. Educational data sciences. *In*: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE, 4., 2014, New York. **Proceedings** [...]. New York: ACM, 2014