

INFORMÁTICA NA EDUCAÇÃO

teoria & prática

Vol. xx | N° x | 20xx

ISSN digital ISSN impresso
1982-1654 1516-084X



Páginas xx-xx

**Na versão para avaliação,
NÃO INFORMAR os autores**

Autor(a) 1

Universidade/Faculdade Filiação 1
autor1@email.com

Autor(a) 2

Universidade/Faculdade Filiação 2
autor2@email.com

Autor(a) 3

Universidade/Faculdade Filiação 3
autor3@email.com

Autor(a) 4

Universidade/Faculdade Filiação 4
autor4@email.com



PORTO ALEGRE

RIO GRANDE DO SUL
BRASIL

Recebido em: xxxxxxxx de 20xx

Aprovado em: xxxxxxxx de 20xx

A importância do foco para o desempenho do aluno de programação no ensino superior

*The importance of focus for programming student
performance in higher education*

Resumo

A disciplina de programação tem figurado como uma das principais causadoras de desistência e reprovação em cursos superiores. Diante deste cenário, importa empregar uma estratégia que permita analisar os dados gerados pelos alunos a fim de compreender e aprimorar o processo de ensino-aprendizagem, processo conhecido por Análise de Aprendizagem. Para esta pesquisa, os dados foram coletados a partir da plataforma The Huxley e são relativos a uma turma de Programação do curso Engenharia de Computação. O objetivo deste trabalho foi o de identificar padrões comportamentais que estão relacionados com o desempenho final do aluno. Para tanto, foram definidas métricas para caracterizar os dados analisados. Com base na análise dos resultados, concluiu-se que adotar a estratégia de focar em uma questão até concluí-la traz melhores resultados para o aluno, independentemente do tempo investido para tal, seja na forma de um estudo contínuo ou intervalado.

Palavras-chave: Análise de aprendizagem. Educação superior. Disciplina de programação.

Abstract

The programming discipline has figured as one of the main causes of dropout and failure in higher education courses. Given this scenario, it is important to employ a strategy that allows analyzing the data generated by the students in order to understand and improve the teaching-learning process, a process known as Learning Analytics. For this research, data were collected from The Huxley platform and are related to a Programming class of the Computer Engineering course. The objective of this work was to identify behavioral patterns that are related to the final performance of the student. To this end, metrics were defined to characterize the analyzed data. Based on the analysis of the results, it was concluded that adopting the strategy of focusing on a question until it is completed brings better results for the student, regardless of the time invested in doing so, whether in the form of a continuous or interval study.

Keywords: Learning analytics. Higher education. Programming course.

1. Introdução

É notável que cursos da área de ciências exatas possuem uma alta taxa de desistência logo nos primeiros semestres. Em se tratando de cursos ligados à Computação, a disciplina de programação (seja ela introdutória ou não) figura como uma das principais causadoras de desistência e reprovação (GIRAFFA, 2013; CARVALHO *et al.*, 2016; CASTRO e TEDESCO, 2020). Os principais motivos apontados são as deficiências em disciplinas fundamentais, como português e matemática; os hábitos de estudo, uma vez que alunos advindos do médio costumam, de maneira geral, se comportarem de forma passiva, isto é, mais memorizando os assuntos do que de fato aprendendo; e a metodologia, muitas vezes inadequada, do professor.

Com a finalidade de contornar tais problemas, constata-se que o Ensino híbrido (*Blended Learning*) vem sendo amplamente aplicado no Ensino Superior (CARVALHO *et al.*, 2016; PEREIRA *et al.*, 2017; SCHUMACHER e IFENTHALER, 2021), sendo empregado também no <Nome protegido para revisão>. Tal metodologia permite que o professor realize o ensino presencial combinado com o uso de ferramentas online, gerando mais autonomia no aluno, além de aumentar a interação entre alunos e professor e entre os próprios alunos, sendo estes últimos os pontos mais fortes dos ambientes online (ULFA e FATAWI, 2021).

Dentre o universo de ferramentas online usadas em cursos de Computação, destacam-se os Juizes Online. Nesta categoria se encontram as ferramentas *The Huxley*¹ e *Beecrowd*² (antigo *URI Online Judge*), apenas para citar as mais conhecidas. Estas ferramentas possibilitam que o professor elabore listas online de exercícios de programação para que o aluno possa responder ao seu tempo, de forma assíncrona. Assim que é feita a submissão, o Juiz Online indica para o aluno de forma automatizada, e consequentemente sem a intervenção do professor, se a resposta está correta ou não. Na medida em que os alunos interagem com a ferramenta, ocorre a geração de dados (YI *et al.*, 2017).

O ponto de particular interesse para esta pesquisa é que esses Juizes Online acabam por registrar um conjunto de dados sobre as submissões, as quais podem ser analisados *a posteriori*, tais como o número de submissões por questão, data e hora de cada submissão, o código submetido, se determinada submissão está correta etc., abrindo assim um leque de possibilidades para pesquisa e análise. Esse processo de coletar e analisar dados gerados por estudantes é conhecido por Análise de Aprendizagem (SIEMENS, 2012).

Diante do exposto, as questões de pesquisa que se deseja responder são: A partir das submissões, é possível identificar algum padrão comportamental que justifique o desempenho final? Padrões apontados por

pesquisadores de outras instituições se repetem nos dados aqui coletados?

Na busca por respostas, estabeleceu-se como objetivo deste trabalho identificar padrões comportamentais que estão relacionados com o desempenho final do aluno. Para tanto, realizou-se uma pesquisa exploratória de caráter quantitativo. Definiu-se um conjunto de métricas a fim mensurar um conjunto de características para serem analisadas e correlacionadas com a nota final do aluno. Os dados foram coletados a partir da ferramenta *The Huxley*, com base nos dados gerados por uma turma de Introdução à Programação do <Nome protegido para revisão>.

Este estudo se justifica pela sua relevância científica e acadêmica. Pesquisas recentes apontam que a Análise de Aprendizagem tem produzido uma série de benefícios aos estudantes e professores do Ensino Superior (SOKOUT *et al.*, 2020; QUADRI e SHUKOR, 2021; ULFA e FATAWI, 2021). Esses benefícios incluem a melhora no rendimento dos alunos, a melhora na metodologia do professor e uma redução na desistência e reprovação. Ademais, o estudo em Análise de Aprendizagem ainda não alcançou todo o seu potencial, sendo apontado como o caminho para transformar as relações de ensino-aprendizagem (DAWSON *et al.*, 2019). O estudo também se justifica por contribuir para a unidade de análise de estudo, sendo a pesquisa inédita no âmbito do <Nome protegido para revisão>.

O artigo está organizado em seis seções: a primeira corresponde a esta introdução; a segunda envolve a dimensão teórica deste estudo; a terceira discute os trabalhos relacionados; a quarta corresponde aos procedimentos metodológicos adotados na pesquisa para atender aos objetivos propostos; a quinta contempla a análise e discussão dos resultados e por fim as considerações finais.

2. Dimensão teórica – Análise de Aprendizagem

A definição mais aceita para Análise de Aprendizagem (*Learning Analytics*) foi dada por Siemens (2012) e consiste no processo de coletar e analisar dados gerados por estudantes com o objetivo de compreender e melhorar o processo de aprendizagem. Conforme apontado por Nistor e Garciac (2018), esse processo de coleta e análise contribui para que pesquisadores e professores se aproximem ainda mais da essência do aprendizado, melhorando o processo de ensino-aprendizagem tanto do ponto de vista do professor quanto do aluno.

De maneira geral, o processo de Análise de Aprendizagem pode ser organizado em quatro etapas. Primeiro, os estudantes geram os dados. Segundo, os dados são coletados e armazenados. Terceiro, os dados são analisados e visualizados. Por fim, como resultado, busca-se a compreensão e o melhoramento do processo de aprendizagem, que pode envolver alguma adequação na metodologia do professor, no ambiente de aprendizagem ou simplesmente ter uma

¹ Acessível em <https://thehuxley.com/>.

² Acessível em <https://www.beecrowd.com.br/>.

melhor visualização do progresso do aluno (KNOBBOUT e STAPPEN, 2020; VIBERG *et al.*, 2018; NISTOR e GARCIAC, 2018).

Vale salientar que, no contexto de Análise de Aprendizagem em que os pesquisadores fazem a coleta dos dados a partir de ferramentas online, os dados podem ser obtidos, basicamente, de duas formas. Na primeira forma, a própria ferramenta online disponibiliza os dados para download, seja em *JSON (JavaScript Object Notation)* ou qualquer outro formato. Na segunda forma, por outro lado, os dados estão disponíveis através da própria ferramenta online, isto é, através das suas páginas *Web*, mas não podem ser obtidos diretamente, sendo necessário o auxílio de algum *script* que realize o trabalho de extração. É desta forma que se encontram os dados na plataforma *The Huxley*.

3. Trabalhos relacionados

A Análise de Aprendizagem tem sido aplicada por vários pesquisadores em seus respectivos contextos, e algumas destas aplicações foram de suma importância para o desenvolvimento desta pesquisa, assim como da metodologia adotada.

Em especial, tem-se o trabalho de Araújo *et al.* (2013). Os autores investigaram como o hábito de estudo do aluno influencia no resultado final da disciplina de Introdução à Programação. Em sua metodologia, os autores definiram uma métrica baseada em Sessão de Estudo, aproveitada por este trabalho, e que será apresentada em detalhes na seção de Procedimentos Metodológicos. Após analisar os dados, os autores encontraram uma forte correlação entre o número de exercícios corretos e a nota final da disciplina, além de concluir que há um forte indicativo de que o estudo diário representa uma forte influência no sucesso dos alunos, muito embora não seja a frequência que determina o sucesso, mas a qualidade do estudo (ULFA e FATAWI, 2021).

Gaudencio *et al.* (2013a) propuseram uma metodologia para auxiliar o professor na elaboração e reformulação de listas de exercícios para a disciplina de Introdução à Programação, usando como base a análise automática das respostas dos alunos. Na mesma linha, Souza *et al.* (2021) empregaram a Análise de Aprendizagem para encontrar métricas capazes de identificar problemas com exercícios de programação em disciplinas introdutórias de programação.

Gaudencio *et al.* (2013b) desenvolveram uma ferramenta chamada *TSTView* que, com base na análise de dados parciais, permite detectar antecipadamente dificuldades individuais dos estudantes, em grupos de estudantes ou dificuldades gerais da turma, permitindo ao professor intervir em tempo hábil para sanar as dificuldades.

Pereira *et al.* (2017) propuseram um método para inferir o grau de aprendizagem de alunos em turmas de Introdução à Programação. A ideia central dos pesquisadores consistiu em identificar antecipadamente o desempenho do aluno a partir de dados parciais. Em sua metodologia, os autores estabeleceram uma lista de parâmetros, grande parte relacionados ao código

submetido para Juízes Online. Seus resultados mostraram que o modelo preditivo apresentou uma taxa de acerto próximo aos 80%. Com efeito, conforme mencionado pelos autores, identificar previamente o desempenho do aluno permite uma atuação direcionada do professor para com o aluno.

No trabalho desenvolvido por Oliveira *et al.* (2021), os autores correlacionaram a nota de um questionário respondido no início da disciplina de Introdução à Programação com a nota final da disciplina. O objetivo era avaliar se a habilidade em resolução de problemas (mensurada por meio do questionário) tinha alguma relação com o desempenho final na disciplina, para que servisse como um preditor de desempenho final. Tal estudo, como relatado pelos próprios autores, já havia sido realizado em outros contextos por outros autores e apresentado uma forte correlação. Contudo, os autores obtiveram um resultado diverso do esperado, e acreditam que o ensino remoto durante a pandemia da COVID-19 tenha influenciado o resultado.

Ulfa e Fatawi (2021), por sua vez, analisaram como as interações dos alunos em um Ambiente Virtual de Aprendizagem podem ajudar a prever o desempenho final de alunos em uma disciplina de Programação *Web*. Dados como número de *logins*, tempo de interação, atividades de leitura, exercícios enviados e outros registros foram coletados e correlacionados. O objetivo era encontrar um padrão comportamental que servisse como forma de prever o resultado de alunos de futuras turmas.

Knobbout e Stappen (2020) realizaram uma massiva revisão de literatura sobre o tema Análise de Aprendizagem. Foram aproximadamente 2000 artigos considerados. Os autores analisaram como o campo de análise de dados aplicado à educação tem impactado no processo de aprendizagem. Eles observaram que, apesar da grande quantidade de pesquisa realizada dentro dessa temática, muito mais foi feito para compreender o processo de aprendizagem do que para melhorá-lo, algo que já havia sido observado por Viberg *et al.*, (2018). Isso mostra apenas que a pesquisa em Análise de Aprendizagem ainda não alcançou todo o seu potencial, mas está avançando continuamente para transformar as relações de ensino-aprendizagem (DAWSON *et al.*, 2019), sobretudo no Ensino Superior (QUADRI e SHUKOR, 2021).

4. Procedimentos metodológicos

Com a finalidade de atender aos objetivos desta pesquisa, o método se estruturou como estudo de natureza exploratória e caráter quantitativo. A pesquisa foi conduzida por meio de quatro etapas fundamentais, a saber: Definição de Métricas e Variáveis, Coleta de Dados, Processamento dos Dados, e Análise dos Dados; os quais serão descritos em detalhes nas subseções a seguir.

4.1 Definição de Métricas e Variáveis

Uma Métrica é um conceito abstrato que representa aquilo que se quer analisar. A Variável, por sua vez, é

uma característica concreta e mensurável de uma Métrica. Assim, uma Métrica só é mensurável por meio de uma ou mais de suas Variáveis.

As três métricas definidas para este trabalho estão apresentadas na Tabela 1, juntamente com uma breve explicação.

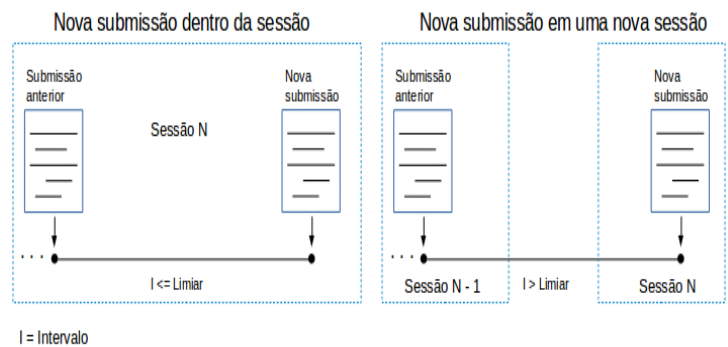
Tabela 1 – Métricas.

Métrica	Explicação
Exercícios Corretos	Identifica a quantidade total de exercícios que o aluno conseguiu resolver ao longo da disciplina.
Sessão de Estudo	Segue a definição dada por Araújo <i>et al.</i> (2013), que consiste em um período ininterrupto de estudo, caracterizado por submissões contínuas de respostas para a plataforma online.
Foco na Questão	Esta métrica é uma das contribuições deste trabalho. Considera-se que um aluno teve Foco na Questão quando iniciou uma questão da lista de exercícios e permaneceu nela, sem alternar para outra questão, independente do tempo entre as submissões, até encontrar uma solução para a questão.

Fonte: Elaborado pelos autores.

Em relação à métrica Sessão de Estudo, é importante mencionar que foi utilizada a metodologia *threshold-based session* (sessão baseada em limiar) de Geiger e Halfaker (2013) para delimitar o escopo uma sessão de estudo. Seu objetivo é definir, baseando-se em um limiar de tempo, se uma nova submissão faz parte ou não de uma mesma sessão de estudo. O método consiste em distinguir um intervalo curto, que caracteriza uma nova submissão dentro da mesma sessão, de um intervalo longo, que caracteriza o fim de uma sessão e o início de uma nova. Neste trabalho, empregando-se metodologia sobre os dados coletados, foi encontrado o limiar de 40 minutos. A Figura 1 ilustra a metodologia.

Figura 1 – Metodologia de sessão baseada em limiar.



Fonte: Elaborado pelos autores.

A Tabela 2 apresenta cada uma das variáveis definidas para este trabalho. Note que cada variável está relacionada à sua métrica. Para cada variável há uma breve descrição.

Tabela 2 – Variáveis.

Métrica associada	Variável	Descrição
Exercícios Corretos	Acertos	Quantidade total de questões que o aluno acertou ao longo da disciplina.
Sessão de Estudo	Número de sessões	Quantidade de sessões de estudo do aluno ao longo do período.
	Tamanho da sessão	Média do tamanho das sessões de estudo do aluno.
	Dias de atividade	Razão entre os dias que o aluno enviou pelo menos uma submissão e os 100 dias letivos do período.
	Tempo total	Somatória do tempo empreendido pelo aluno em todas as suas sessões de estudo.
Foco na Questão	Número de focos	Quantidade de focos do aluno ao longo do período.
	Tempo médio focado	Tempo médio focado, em minutos, do aluno.
	Tempo total focado	Somatória do tempo registrado em todos os focos do aluno.

Fonte: Elaborado pelos autores.

Vale salientar que as duas primeiras métricas foram definidas tomando por base o observado na literatura. Em especial, a métrica Sessão de Estudo e suas variáveis foram consideradas a partir do trabalho de Araújo *et al.* (2013).

4.2 Coleta de Dados

Uma vez realizada a definição das métricas e suas variáveis, a pesquisa foi conduzida para a etapa de Coleta de Dados. Como unidade de análise, envolveu-se o <Nome protegido para revisão>. Como unidade de observação, envolveu-se os discentes de uma turma de Introdução à Programação (estudantes do 1º período) do curso de Engenharia da Computação.

Para fins deste estudo, os dados foram coletados diretamente da ferramenta online *The Huxley*, considerando uma turma do período letivo 2019.1, já finalizada durante a coleta. Duas razões pautaram a escolha. A primeira, por se tratar de uma turma pré-pandemia, entendendo-se que as condições adversas proporcionadas pelo ensino remoto poderia distorcer os resultados, como observado na pesquisa de Oliveira *et al.* (2021). A segunda, por ser a turma mais recente com a maior quantidade de alunos – 53 ao todo.

Uma vez que a plataforma *The Huxley* não disponibiliza o *download* dos dados gerados pelos alunos, foi necessário desenvolver um pequeno *script* (S1) em *JavaScript* a fim de extrair os dados por meio da *API* (*Application Programming Interface*) *REST* (*Representational State Transfer*) da plataforma e salvá-los no formato *JSON*, gerando assim o que se chamou de dados brutos.

Por razões éticas, apenas a equipe de professores envolvidos no projeto teve contato direto com a plataforma *The Huxley* e com os dados brutos, visto que tais dados traziam informações sobre outros alunos da instituição.

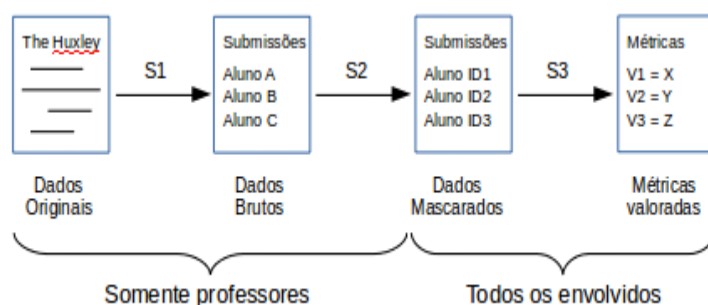
4.3 Processamento dos Dados

A etapa de Processamento dos Dados envolveu uma sequência de procedimentos sobre os dados brutos. Primeiramente foi necessário mascarar as informações referente aos alunos para que o aluno participante desta pesquisa pudesse trabalhar sobre os dados, gerando-se os dados brutos mascarados. Para tanto, foi desenvolvido um *script* (S2) em *Python* que realizasse tal tarefa, atribuindo, aos nomes dos alunos e matrículas, uma identificação genérica individualizada.

Finalmente, um outro *script* (S3) *Python* foi desenvolvido para extrair, dos dados brutos mascarados, os dados relevantes e processá-los com o objetivo de valorar as métricas (por meio de suas variáveis) estabelecidas. Os resultados foram salvos em um arquivo no formato *CSV* (*Comma-Separated Values*), viabilizando, assim, o carregamento do arquivo por meio de qualquer planilha eletrônica.

A Figura 2 sintetiza os passos envolvidos nas etapas de Coleta de Dados e Processamento dos Dados descritos anteriormente.

Figura 2 – Diagrama completo.



Fonte: Elaborado pelos autores.

4.4 Análise dos Dados

Na sequência, os dados foram tabulados, organizados e analisados. O arquivo *CSV* gerado ao fim da etapa de processamento foi aberto por meio de uma planilha eletrônica. Cada uma das métricas foram caracterizadas experimentalmente por meio de gráficos de dispersão e de pizza. Por fim, todas as métricas foram correlacionadas com as notas finais dos alunos usando como método o coeficiente de correlação de Spearman. O método foi selecionado por ser apropriado para variáveis ordinais contínuas e discretas (ULFA e FATAWI, 2021). A análise dos dados será apresentada em detalhes na sessão seguinte, de Resultados e Discussão.

5. Resultados e discussão

Inicialmente, procurou-se caracterizar experimentalmente cada métrica estabelecida neste trabalho com vistas a traçar o perfil de estudo dos alunos da disciplina de Introdução à Programação, conforme apresentado na subseção Análise do Perfil. Posteriormente, as diferentes métricas foram correlacionadas com o desempenho de seu respectivo aluno, como será mostrado na subseção Análise das Correlações. Os resultados obtidos são apresentados a seguir.

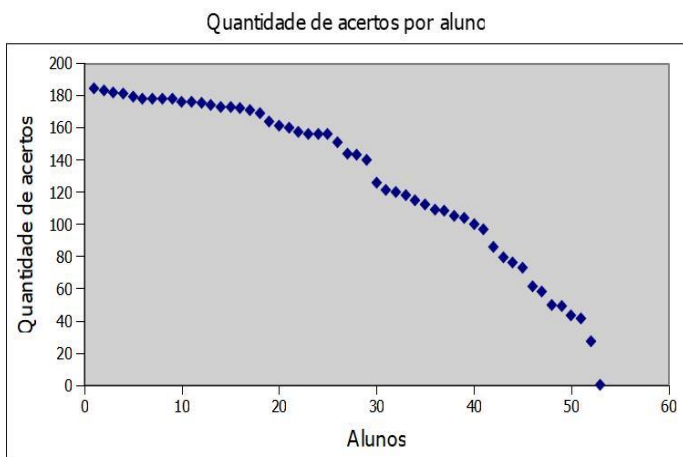
5.1 Análise do Perfil

Antes de apresentar as análises, convém mencionar que os gráficos de dispersão foram gerados organizando os dados de cada variável em ordem decrescente. Para a elaboração dos gráficos de pizza, os dados foram agrupados em quatro grupos (A, B, C e D), os quais representam intervalos-quartis. Cada um abrange um intervalo de valores de 25% do total de valores daquela variável, iniciando do intervalo de maiores valores (75-100% - 4º quartil) até o de menores valores (0-25% - 1º quartil), a fim de se ter uma melhor compreensão sobre como os dados estão distribuídos.

5.1.1 Métrica 1: Exercícios corretos - Acertos

O Gráfico 1 apresenta a quantidade total de acertos de cada aluno, ou seja, o número de questões que o aluno submeteu e obteve êxito ao longo da disciplina. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 184 e o menor valor foi de 0. A mediana da amostra apresentada é de 144 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foi de 178, com 4 repetições. A média geral para esta variável foi de aproximadamente 129 números de acertos.

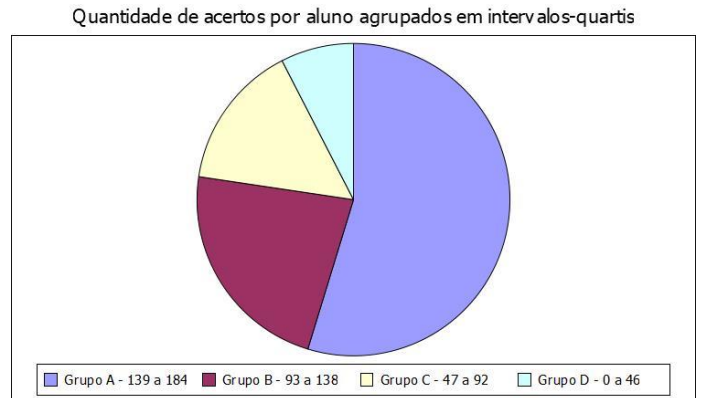
Gráfico 1 – Dispersão da variável "Acertos"



Fonte: Elaborado pelos autores.

O Gráfico 2 apresenta a quantidade total de acertos dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo A, sendo 29 alunos, correspondentes a 54,72% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Acertos", no intervalo do 4º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo D com 4 alunos, que representam 7,55% em relação ao total de alunos, representando o 1º quartil. O grupo B teve 12 alunos, correspondentes a 22,64% e representando o 3º quartil e o grupo C teve 8 alunos, correspondendo a 15,09% e representando o 2º quartil.

Gráfico 2 – Pizza da variável "Acertos"

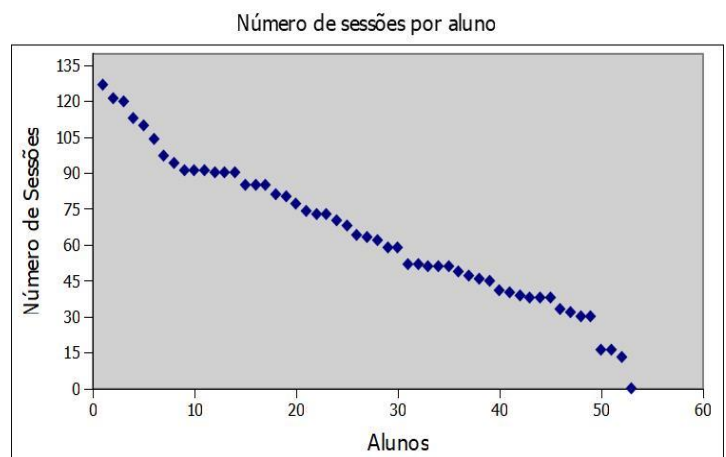


Fonte: Elaborado pelos autores.

5.1.2 Métrica 2: Sessão de Estudo – Número de sessões

O Gráfico 3 apresenta o número de sessões de estudo de cada aluno, ou seja, a quantidade total de sessões de estudo em minutos realizadas ao longo da disciplina. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 127 e o menor valor foi de 0. A mediana da amostra apresentada é de 63 para esta variável, enquanto a moda, ou seja, os valores que mais se repetem na amostra, foram 38, 51, 85, 90 e 91, todos com 3 repetições. A média geral para esta variável foi de aproximadamente 65 sessões de estudo.

Gráfico 3 – Dispersão da variável "Número de sessões"

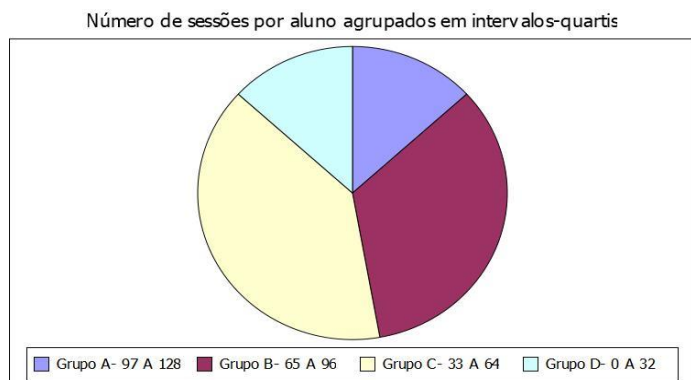


Fonte: Elaborado pelos autores.

O Gráfico 4 apresenta o número total de sessões de estudo dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo C, sendo 21 alunos, correspondentes a 39,62% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Número de sessões", no intervalo do 2º quartil,

enquanto que o menor grupo de alunos se apresenta no intervalo dos grupos A e D, com 7 alunos cada, que representam cada um deles 13,21% em relação ao total de alunos, representando o 4º e 1º quartil. O grupo B teve 18 alunos, correspondentes a 33,96% e representando o 3º quartil.

Gráfico 4 – Pizza da variável "Número de sessões"

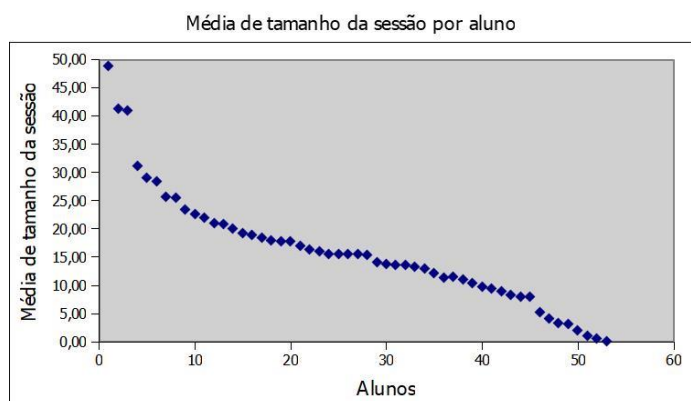


Fonte: Elaborado pelos autores.

5.1.3 Métrica 2: Sessão de Estudo – Tamanho da sessão

O Gráfico 5 apresenta a média do tamanho das sessões de cada aluno, ou seja, representa a soma do tamanho de todas as sessões do aluno dividido pelo número de sessões de estudo do mesmo. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 48.73 e o menor valor foi de 0. A mediana da amostra apresentada é de 15.48 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foi de 15.52, com 2 repetições. A Média geral para esta variável foi de aproximadamente 16 minutos de tempo médio das sessões de estudo.

Gráfico 5 – Dispersão da variável "Tamanho da sessão"

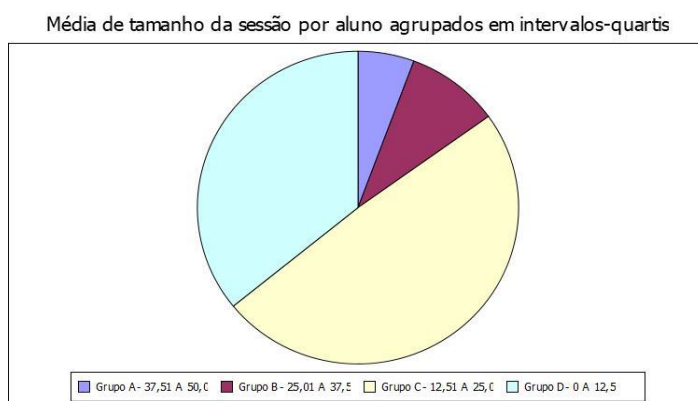


Fonte: Elaborado pelos autores.

O Gráfico 6 apresenta a média de tamanho da sessão dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É

perceptível que há uma maior parcela dos alunos inseridos no grupo C, sendo 26 alunos, correspondentes a 49,06% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Tamanho da sessão", no intervalo do 2º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo A com 3 alunos, que representam 5,66% em relação ao total de alunos, representando o 1º quartil. O grupo B teve 5 alunos, correspondentes a 9,43% e representando o 3º quartil e o grupo D teve 19 alunos, correspondendo a 35,85% e representando o 1º quartil.

Gráfico 6 – Pizza da variável "Tamanho da sessão"

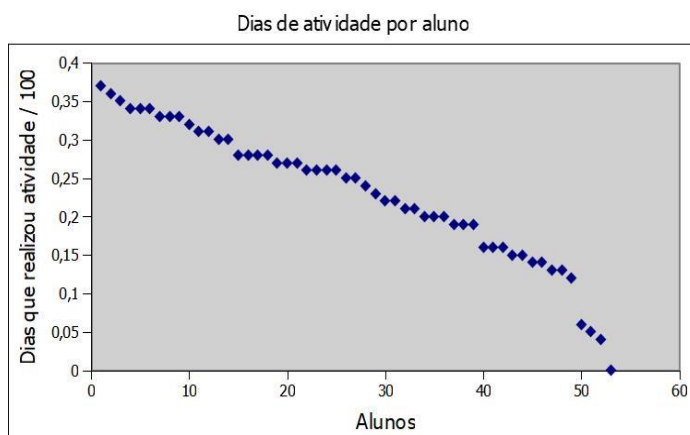


Fonte: Elaborado pelos autores.

5.1.4 Métrica 2: Sessão de Estudo – Dias de atividade

O Gráfico 7 apresenta a proporção de cada aluno, em relação aos 100 dias letivos (100% ou 1,0), de dias que o aluno submeteu questões da disciplina. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 0.37 ou 37 dias de atividade e o menor valor foi de 0. A mediana da amostra apresentada é de 0.25 ou 25 dias de atividade para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foram 0.28 e 0.26, com 2 repetições. A média geral para esta variável foi de 0.23, ou 23 dias de atividade.

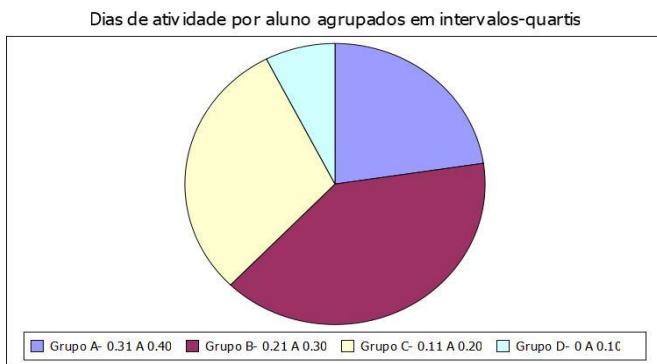
Gráfico 7 – Dispersão da variável "Dias de Atividade"



Fonte: Elaborado pelos autores.

O Gráfico 8 apresenta a proporção de dias de atividade por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo B, sendo 21 alunos, correspondentes a 39,62% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Dias de atividade", no intervalo do 3º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo D com 4 alunos, que representam 7,55% em relação ao total de alunos, representando o 1º quartil. O grupo A teve 12 alunos, correspondentes a 22,64% e representando o 4º quartil e o grupo C teve 16 alunos, correspondendo a 30,19% e representando o 2º quartil.

Gráfico 8 – Pizza da variável "Dias de atividade"

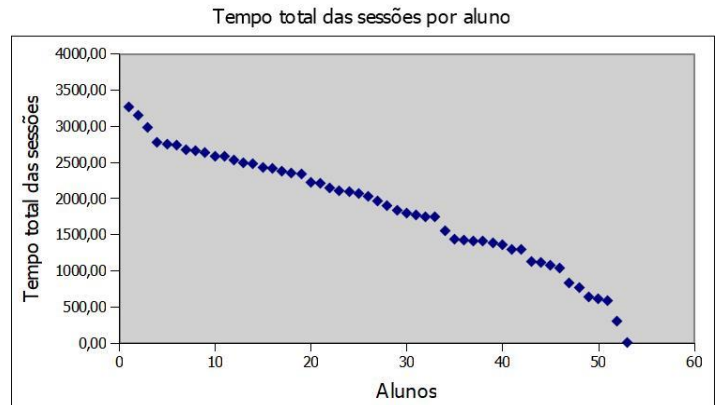


Fonte: Elaborado pelos autores.

5.1.5 Métrica 2: Sessão de Estudo – Tempo total

O Gráfico 9 apresenta o tempo total de estudo de cada aluno, ou seja, a soma do tempo de todas as sessões de estudo de cada aluno. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 3257.62 minutos e o menor valor foi de 0. A mediana da amostra apresentada é de 1956.85 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foram todos, com 1 repetição cada. A média geral para esta variável foi de aproximadamente 1853 minutos de tempo total de estudo.

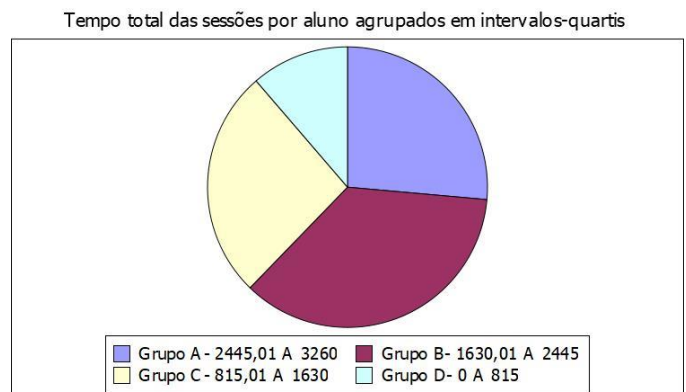
Gráfico 9 – Dispersão da variável "Tempo total"



Fonte: Elaborado pelos autores.

O Gráfico 10 apresenta o tempo total dedicado às sessões de estudo dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo B, sendo 19 alunos, correspondentes a 35,85% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Tempo total", no intervalo do 3º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo D com 6 alunos, que representam 11,32% em relação ao total de alunos, representando o 1º quartil. O grupo A e C, tiveram 14 alunos cada, correspondentes a 26,41% cada, e representando o 4º e o 2º quartil.

Gráfico 10 – Pizza da variável "Tempo total"



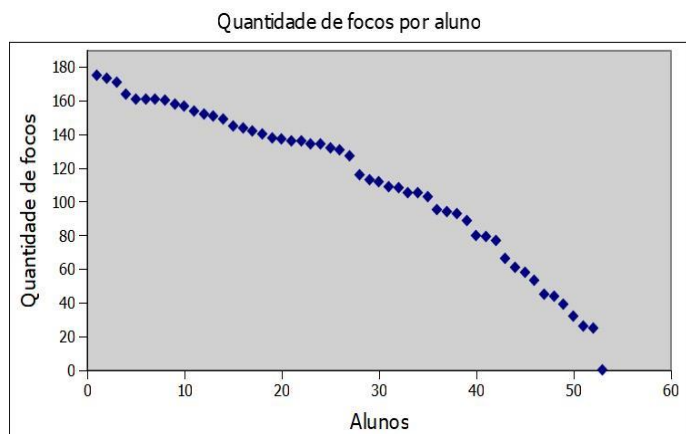
Fonte: Elaborado pelos autores.

5.1.6 Métrica 3: Foco na Questão – Número de focos

O Gráfico 11 apresenta a quantidade de focos, ou seja, quantas vezes um aluno iniciou uma questão e permaneceu nela até concluí-la, sem tentar outra. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 175 focos e o menor valor foi de 0. A mediana da amostra apresentada é de 127 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra,

foi de 161 focos, com 3 repetições. A média geral para esta variável foi de aproximadamente 112 focos.

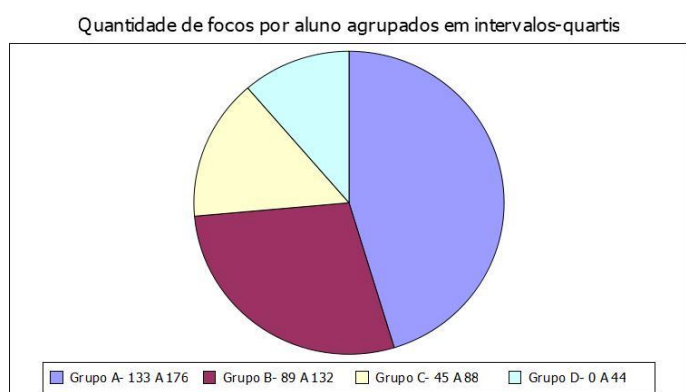
Gráfico 11 – Dispersão da variável "Número de focos"



Fonte: Elaborado pelos autores.

O Gráfico 12 apresenta a quantidade de focos dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo A, sendo 24 alunos, correspondentes a 45,29% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Focos", no intervalo do 4º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo D com 6 alunos, que representam 11,32% em relação ao total de alunos, representando o 1º quartil. O grupo B teve 15 alunos, correspondentes a 28,30% e representando o 3º quartil e o grupo C teve 8 alunos, correspondendo a 15,09% e representando o 3º quartil.

Gráfico 12 – Pizza da variável "Número de focos"



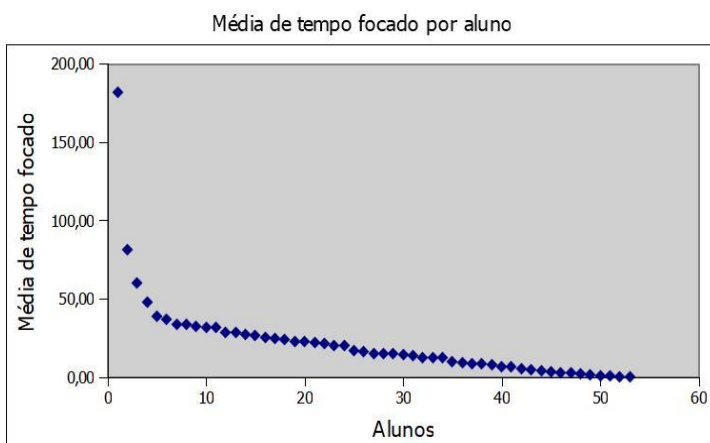
Fonte: Elaborado pelos autores.

5.1.7 Métrica 3: Foco na Questão – Tempo médio focado

O Gráfico 13 apresenta a média de tempo focado em minutos de cada aluno, ou seja, a soma de tempo focado dividido pela quantidade de vezes em que o

aluno permaneceu focado. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 182.04 e o menor valor foi de 0. A mediana da amostra apresentada é de 15.18 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foram todos, com 1 repetição cada. A média geral para esta variável foi de aproximadamente 21 minutos e 39 segundos de tempo médio focado.

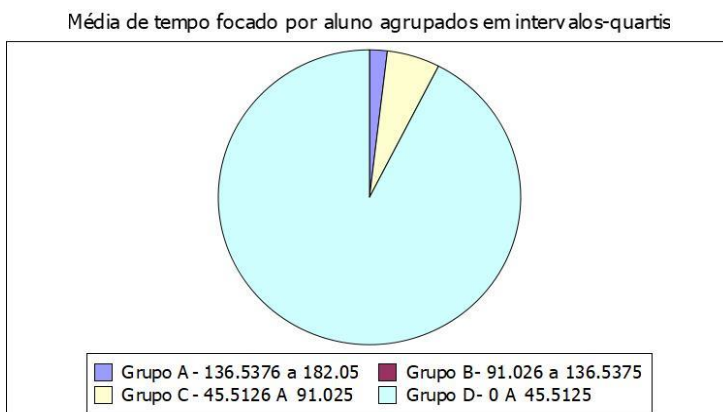
Gráfico 13 – Dispersão da variável "Tempo médio focado"



Fonte: Elaborado pelos autores.

O Gráfico 14 apresenta a média de tempo focado dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo D, sendo 49 alunos, correspondentes a 92,45% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Tempo médio focado", no intervalo do 1º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo B com 0 alunos, que representam 0% em relação ao total de alunos, representando o 3º quartil. O grupo A teve 1 aluno, correspondente a 1,89% e representando o 4º quartil e o grupo C teve 3 alunos, correspondendo a 5,66% e representando o 2º quartil.

Gráfico 14 – Pizza da variável "Tempo médio focado"

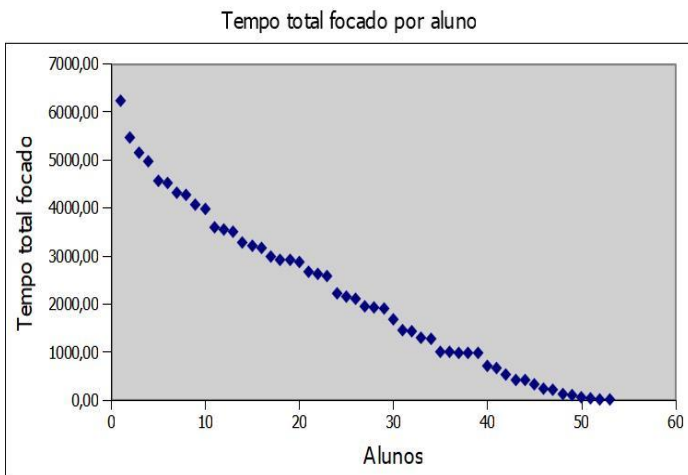


Fonte: Elaborado pelos autores.

5.1.8 Métrica 3: Foco na Questão – Tempo total focado

O Gráfico 15 apresenta a quantidade total de tempo focado em minutos de cada aluno. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 6232.12 e o menor valor foi de 0. A mediana da amostra apresentada é de 1934.15 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foram todos, com 1 repetição cada. A média geral para esta variável foi de aproximadamente 2173 minutos e 6 segundos de tempo total focado.

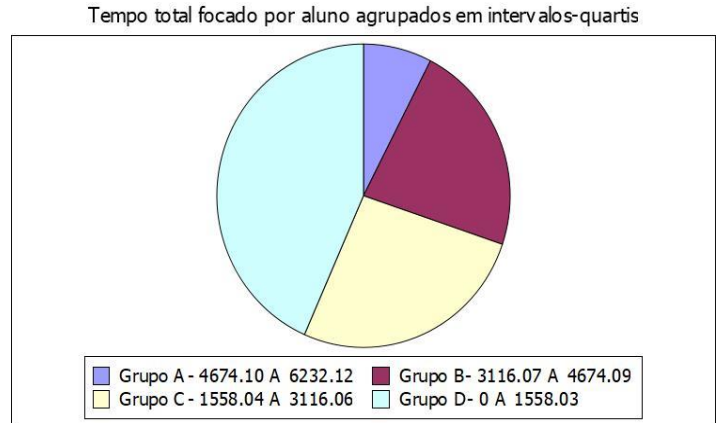
Gráfico 15 – Dispersão da variável "Tempo total focado"



Fonte: Elaborado pelos autores.

O Gráfico 16 apresenta a quantidade total de tempo focado dos alunos por meio de grupos. Cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo D, sendo 23 alunos, correspondentes a 43,40% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Tempo total focado", no intervalo do 1º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo A com 4 alunos, que representam 7,55% em relação ao total de alunos, representando o 4º quartil. O grupo B teve 12 alunos, correspondentes a 22,64% e representando o 3º quartil e o grupo C teve 14 alunos, correspondendo a 26,41% e representando o 2º quartil.

Gráfico 16 – Pizza da variável "Tempo total focado"

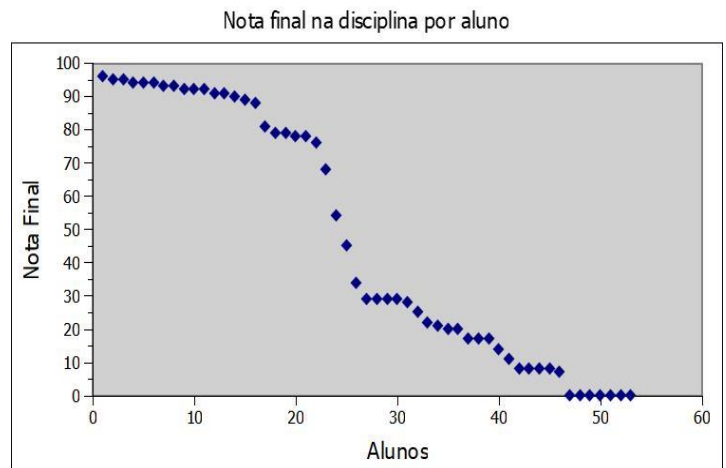


Fonte: Elaborado pelos autores.

5.1.9 Nota Final

O Gráfico 17 apresenta a nota final de cada aluno obtida na disciplina. Através do gráfico é possível observar que o alcance, ou seja, o maior valor obtido da amostra foi de 96 e o menor valor foi de 0. A mediana da amostra apresentada é de 29 para esta variável, enquanto a moda, ou seja, o valor que mais se repete na amostra, foi de 0, com 7 repetições. A Média geral para esta variável foi de aproximadamente 47.5 de nota final na disciplina.

Gráfico 17 – Dispersão da Nota Final dos alunos



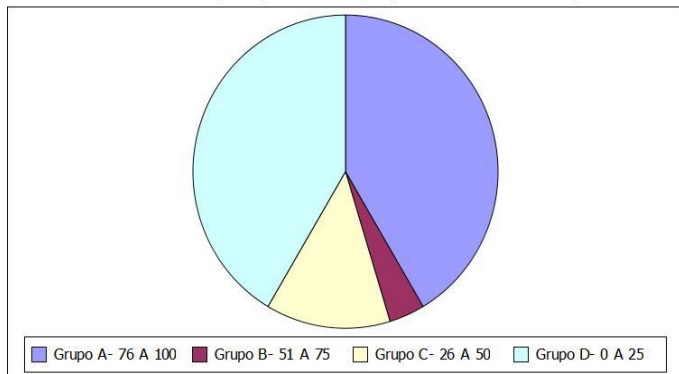
Fonte: Elaborado pelos autores.

O Gráfico 18 apresenta a nota final obtida na disciplina dos alunos por meio de grupos, onde cada grupo representa um intervalo de 1 quartil em relação ao valor mais alto da variável, onde esses grupos somados chegam a sua completude de 100% ou 4 quartis. É perceptível que há uma maior parcela dos alunos inseridos no grupo C, sendo 21 alunos, correspondentes a 39,62% do total de alunos, o que implica dizer que o maior grupo de alunos se apresenta, para a variável "Nota final", no intervalo do 2º quartil, enquanto que o menor grupo de alunos se apresenta no intervalo do grupo A com 3 alunos, que representam 5,66% em

relação ao total de alunos, representando o 4º quartil. O grupo B teve 15 alunos, correspondentes a 28,30% e representando o 3º quartil e o grupo D teve 14 alunos, correspondendo a 26,42% e representando o 1º quartil.

Gráfico 18 – Pizza da Nota Final dos alunos

Nota final na disciplina por aluno agrupados em intervalos-quartis

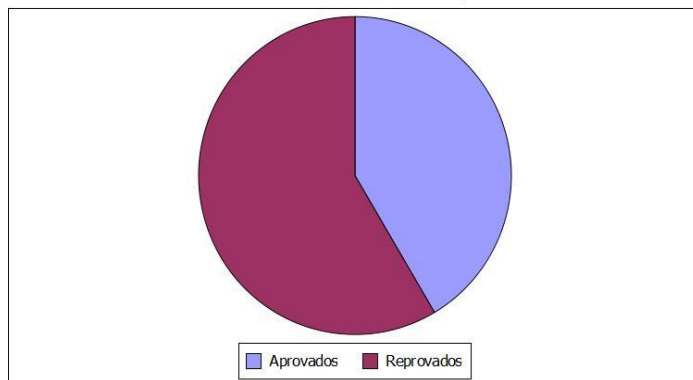


Fonte: Elaborado pelos autores.

Por fim, o Gráfico 19 apresenta os alunos aprovados (nota final igual ou superior a 70) e os alunos reprovados (nota final menor que 70) ao final da disciplina por meio de grupos. O maior grupo foi o de reprovados, com 31 alunos, que correspondem a 58,49% do total da turma e o menor grupo foi o de aprovados, com 22 alunos, que correspondem a 41,51% do total da turma, que é de 53 alunos.

Gráfico 19 – Pizza dos alunos aprovados e não aprovados

Resultado dos alunos na disciplina



Fonte: Elaborado pelos autores.

5.2 Análise das Correlações

A correlação entre cada uma das variáveis e a nota final dos alunos foi realizada através do coeficiente de correlação de Spearman. Para a interpretação e análise dos coeficientes, considerou-se o seguinte: de 0-25%, uma associação muito fraca ou inexistente; de 25-50%, uma associação fraca; de 50-75%, uma associação moderada; e de 75-100%, uma associação forte ou perfeita. Coeficientes positivos indicam que as duas variáveis se relacionam na mesma direção e sentido, enquanto que coeficientes negativos indicam que as variáveis se relacionam na mesma direção, mas em sentidos opostos. A Tabela 3 apresenta o resultado das correlações.

Tabela 3 – Correlações de Spearman.

Métrica associada	Variáveis									
		Acertos	Número de sessões	Tamanho da sessão	Dias de atividade	Tempo total	Número de focos	Tempo médio focado	Tempo total focado	Nota final
Exercícios Corretos	Acertos	1,00								
Sessões de Estudo	Número de sessões	0,71	1,00							
	Tamanho da sessão	0,00	-0,40	1,00						
	Dias de atividade	0,75	0,92	-0,35	1,00					
	Tempo total	0,83	0,78	0,04	0,80	1,00				
Foco na Questão	Número de focos	0,98	0,67	-0,02	0,70	0,76	1,00			
	Tempo médio focado	-0,22	0,05	-0,08	0,03	-0,02	-0,21	1,00		
	Tempo total focado	0,34	0,54	-0,16	0,51	0,49	0,34	0,64	1,00	
Nota Final	Nota final	0,87	0,68	-0,04	0,63	0,75	0,86	-0,19	0,24	1,00

Fonte: Elaborado pelos autores.

Da tabela de correlações, é possível depreender uma série de informações. Considerando-se primeiramente a variável “Acertos”, tem-se que o seu grau de associação se mostra muito forte com as variáveis “Dias de Atividade”, “Tempo total”, “Número de Focos” e “Nota final”. À exceção de “Número de Focos”, já era esperado uma forte relação com as demais variáveis, afinal a frequência de estudo e o tempo empreendido estão intuitivamente relacionados com o desempenho final do aluno. A surpresa ficou por conta da correlação com “Número de Focos”. Observa-se uma relação perfeita entre as duas variáveis, o que leva invariavelmente a concluir que os alunos focados acertaram mais questões e, conseqüentemente, apresentaram um bom desempenho ao final da disciplina. Ademais, a variável “Acertos” apresenta um grau de associação mediano com a variável “Número de sessões”, mostrando que o número de sessões de estudo também influenciou na quantidade de questões resolvidas. Porém, o “Tamanho da Sessão” já não possui relevância, uma vez que não possui associação alguma com a quantidade de acertos.

A variável “Número de sessões” apresenta forte correlação com “Dias de atividade” e “Tempo total”, bem como uma correlação moderada com “Tempo total focado”, o que faz completo sentido em se considerando a natureza das variáveis. Possui também uma correlação moderada com “Número de Focos”, uma vez que o aluno pode se manter em apenas uma questão e permear várias sessões de estudo. Com a variável “Nota Final”, apresenta uma correlação moderada, pois a frequência de estudo pôde contribuir com o desempenho final.

Em relação à variável “Tamanho médio das sessões”, vale destacar apenas que o tamanho médio das sessões apresenta correlação inexistente com a “Nota final”, sendo possível concluir que o tamanho das sessões não importa para o desempenho final.

A variável “Dias de atividade” apresenta uma correlação forte com a variável “Tempo total”, o que faz sentido em razão da natureza das variáveis. Apresenta uma correlação moderada com as variáveis “Número de focos” e “Tempo total focado”, visto a possibilidade de o aluno poder ter se mantido em uma mesma questão por vários dias. Observa-se também uma correlação moderada com a “Nota final”.

Por fim, a variável “Tempo total” mostra-se fortemente relacionada com as variáveis “Número de focos” e “Nota Final”. As demais colunas da tabela de correlações ou não possuem correlações relevantes ou possuem correlações já contempladas.

Em síntese, considerando apenas a variável “Nota Final”, pode-se destacar que esta apresentou uma correlação forte com:

- A variável “Acertos” da métrica 1, como já era esperado;
- A variável “Tempo total” da métrica 2, como esperado;
- A variável “Número de Focos” da métrica 3, estabelecida nesta pesquisa.

Apresentou uma correlação moderada entre:

- A variável “Número de sessões” da métrica 2;

- A variável “Dias de atividade” da métrica 2.

E apresentou correlações fracas e/ou praticamente nulas entre as demais variáveis do estudo.

6. Considerações finais

Neste trabalho, foi realizado um estudo baseado no conceito de Análise de Aprendizagem. Foram considerados os dados gerados por estudantes da disciplina de uma turma Introdução à Programação do <Nome protegido para revisão> com o objetivo de identificar os padrões comportamentais que justificaram o desempenho final do aluno. Para tanto, foram estabelecidas algumas métricas com o intuito de caracterizar os dados coletados. Dentre as métricas estabelecidas, esta pesquisa apresentou como contribuição uma nova métrica chamada de Foco na Questão. Após a valoração das métricas, elas foram correlacionadas com a nota final do aluno e analisadas.

Os resultados se mostram pertinentes, sobretudo pelo fato de coadunarem com o observado por Araújo *et al.* (2013), excluindo-se, é claro, a métrica Foco na Questão, que foi definida neste trabalho. Observa-se uma forte correlação com o Tempo Total de estudo e uma relação moderada com o Número de Sessões e Dias de Atividade. Em relação à variável Número de Focos, esta apresentou uma correlação perfeita com a variável Acertos e uma forte correlação com a Nota Final.

Os pesquisadores deste trabalho entendem que a métrica Foco se constitui como uma métrica mais pertinente para avaliar o comportamento do estudante se comparada com a métrica Sessão de Estudo. Os resultados mostram que o Número de Focos de um aluno apresentou uma correlação mais forte do que todas as variáveis da métrica Sessão de Estudo. Ademais, depreende-se dos resultados que o tempo focado não tem relevância para o resultado final, reforçando a ideia de que não é o tempo de estudo, mas a qualidade deste que leva a melhores resultados (ULFA e FATAWI, 2021). Assim, entende-se que o aluno que se mantém focado em apenas uma questão por vez apresenta uma qualidade de estudo melhor do que aquele que fica alternando entre questões, justamente pelo fato de que manter o foco leva a um maior número de acertos ao longo da disciplina, mesmo que o aluno consuma muito tempo para chegar à solução. É provável que a persistência em uma determinada questão leve o aluno a maturar em sua mente o problema, inquietando-o até se chegar a uma solução.

Finalmente, respondendo às questões de pesquisa com base nos resultados, conclui-se que: (i) existem sim padrões que determinam de forma previsível o desempenho final do aluno e (ii) os padrões observados por outros pesquisadores aqui se repetem. Outrossim, conforme constatado nesta pesquisa, dentre os padrões observados, um destacou-se: alunos que se mantêm focados na solução de uma questão por vez, independente do tempo, conseguem acertar mais questões, e alunos que acertam mais questões alcançam, portanto, melhores resultados.

Os resultados desta pesquisa abrem caminhos para identificar antecipadamente alunos com

dificuldades, tornando-se possível que o professor intervenha individualmente e em tempo hábil na expectativa de corrigir a tendência de desempenho dos futuros alunos. Em função disso, espera-se maximizar o aproveitamento do aluno e, sobretudo, minimizar a taxa de evasão e reprovação em disciplinas de programação no âmbito do <Nome protegido para revisão>. Portanto, como proposta de trabalhos futuros, planeja-se realizar uma pesquisa para criar um modelo que prevê o desempenho do aluno por meio de um indicador baseando-se apenas na coleta de dados parciais. A ideia é compartilhar esse indicador com o aluno durante o progresso da disciplina, para que ele possa se autoavaliar e fazer os devidos ajustes em seu perfil de estudo.

Referências

ARAÚJO, E. C. *et al.* O papel do hábito de estudo no desempenho do aluno de programação. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, nº 33, 2013, Maceió. Anais do evento. Maceió: Editora SBC, 2013, p. 730-738.

CARVALHO, L. S. G.; OLIVEIRA, D. B. F.; GADELHA, B. F.. Juiz online como ferramenta de apoio a uma metodologia de ensino híbrido em programação. In: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, nº5, 2016, Uberlândia. Anais do Evento. Uberlândia: Editora SBC, 2016, p. 140-149.

CASTRO, F. e TEDESCO, P. Promoting Reflection on Error in Higher Education Introductory Programming Courses. *Brazilian Journal of Computers in Education (Revista Brasileira de Informática na Educação – RBIE)*, S.I., v. 28, p. pp. 150-165, 2020.

DAWSON, S. *et al.* Increasing the Impact of Learning Analytics. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK19)*. Association for Computing Machinery, New York, NY, USA, 446–455, 2019.

GAUDENCIO, M. *et al.* Eu Sei o que Vocês Fizeram (Agora e) na Aula Passada: o TSTView no Acompanhamento de Exercícios de Programação. In: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, nº 2, 2013, Campinas. Anais do Evento. Campinas: Editora SBC, 2013b, p. 204-213.

GAUDENCIO, M.; DANTAS, A.; GUERRERO, D. D. S. Análise Automática de Exercícios de Programação como Forma de Avaliar a Cobertura de Tópicos da Disciplina. In: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, nº 2, 2013, Campinas. Anais do Evento. Campinas: Editora SBC, 2013a, p. 617-626.

GEIGER, R. S. e HALFAKER, A. (2013). Using edit sessions to measure participation in wikipedia. In *Proceedings of the 2013 conference on Computer*

supported cooperative work, CSCW '13, pp. 861–870, New York, NY, USA. ACM.

GIRAFFA, M.; MORA, M. C. Evasão na disciplina de algoritmo e programação: um estudo a partir dos fatores intervenientes na perspectiva do aluno. In: CONFERENCIA LATINOAMERICANA SOBRE EL ABANDONO EN LA EDUCACIÓN SUPERIOR, nº 3, 2013, Madrid. E.U.I.T de Telecomunicación. Madrid: Dpto. de Publicaciones de la E.U.I.T. de Telecomunicación, 2013, p. 188-197.

KNOBBOUT, J.; STAPPEN, E. Where is the Learning in Learning Analytics? A Systematic Literature Review on the Operationalization of Learning-Related Constructs in the Evaluation of Learning Analytics Interventions. *IEEE Transactions on Learning Technologies*, v. 13, nº 3, p. 631-645, 2020.

NISTOR, N.; GARCÍAC A. H. What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior*, 2018, v. 89, pp. 335-338.

OLIVEIRA, J. C. *et al.* Correlação entre habilidade de resolução de problemas e desempenho em disciplina introdutória de programação. In: ARTIGOS RESUMIDOS - SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 32, 2021, On-line. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021, p. 15-20.

PEREIRA, F. D.; OLIVEIRA, E. H. T.; OLIVEIRA, David Fernandes. Predição de Zona de Aprendizagem de Alunos de Introdução à Programação em Ambientes de Correção Automática de Código. In: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, nº 6, 2017, Recife. Anais do Evento. Recife: Editora SBC, 2017, p. 1507-1516.

QUADRI, A. T.; SHUKOR, N. A. The Benefits of Learning Analytics to Higher Education Institutions: A Scoping Review. *International Journal of Emerging Technologies in Learning (iJET)*, [S. l.], v. 16, n. 23, p. pp. 4–15, 2021.

SCHUMACHER, C. e IFENTHALER, D. Investigating prompts for supporting students' self-regulation – A remaining challenge for learning analytics approaches?. *The Internet and Higher Education*, 2021, v. 49, pp. 1-12.

SIEMENS, G. Learning analytics: envisioning a research discipline and a domain of practice. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS, nº 2, 2012, New York. *Proceedings*. New York: ACM Digital Library, 2012, p. 4-8.

SOKOUT, H.; USAGAWA, T.; MUKHTAR, S. Learning Analytics: Analyzing Various Aspects of Learners' Performance in Blended Courses. The Case of Kabul Polytechnic University, Afghanistan. *International Journal of Emerging Technologies in Learning (iJET)*, [S. l.], v. 15, n. 12, p. pp. 168–190, 2020.

SOUZA, L. M. *et al.* I know what you coded last summer. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 32, 2021, Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021, p. 909-920.

ULFA, S. e FATAWI, I. Predicting Factors that Influence Students' Learning Outcomes Using Learning Analytics in Online Learning Environment. International Journal of Emerging Technologies in Learning (iJET), [S. l.], v. 16, n. 01, p. pp. 4–17, 2021.

VIBERG O. *et al.* The current landscape of learning analytics in higher education. Computers in Human Behavior, 2018, v. 89, pp. 98-110.

YI, B. *et al.* Learning Analytics-Based Evaluation Mode for Blended Learning and Its Applications. International Symposium on Educational Technology (ISET), 2017, pp. 147-149