



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

YURI SOUZA DOS SANTOS

**ESTRATÉGIA DE MONITORAMENTO E ANÁLISE DE
ENGAJAMENTO DE ESTUDANTES EM CURSOS DE BANCO DE
DADOS**

CAMPINA GRANDE - PB

2022

YURI SOUZA DOS SANTOS

**ESTRATÉGIA DE MONITORAMENTO E ANÁLISE DE
ENGAJAMENTO DE ESTUDANTES EM CURSOS DE BANCO DE
DADOS**

**Trabalho de Conclusão Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharel em
Ciência da Computação.**

Orientador : Cláudio Elízio Calazans Campelo

CAMPINA GRANDE - PB

2022

YURI SOUZA DOS SANTOS

**ESTRATÉGIA DE MONITORAMENTO E ANÁLISE DE
ENGAJAMENTO DE ESTUDANTES EM CURSOS DE BANCO DE
DADOS**

**Trabalho de Conclusão Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharel em
Ciência da Computação.**

BANCA EXAMINADORA:

**Cláudio Elízio Calazans Campelo
Orientador – UASC/CEEI/UFCG**

**Carlos Eduardo Santos Pires
Examinador – UASC/CEEI/UFCG**

**Francisco Vilar Brasileiro
Professor da Disciplina TCC – UASC/CEEI/UFCG**

Trabalho aprovado em: 02 de Setembro de 2022.

CAMPINA GRANDE - PB

ABSTRACT

Monitoring students engagement makes it possible to enrich the education system. With it, it is possible to obtain information related to the dedication and performance of students in certain activities. However, regarding Database courses, there is a shortage of strategies to monitor student engagement, making it difficult to conduct research and analyze student performance. Faced with this problem, we developed a mechanism capable of monitoring the students' interaction with the database and generating a set of data involving the use of the Data Definition Language and Data Manipulation Language commands performed. Subsequently, with this mechanism, we've monitored 95 students enrolled in the Database Course of the Computer Science course at the Federal University of Campina Grande. With this, we've carried out a case study, in which we used, in addition to the information collected from the databases, other information related to the participation of students in the discipline, such as the grades obtained and the submission of practical scripts. In this way, we have contributed with a tool capable of monitoring the practical activities of students and extracting information that can be useful in the educational system.

Estratégia de Monitoramento e Análise de Engajamento de Estudantes em Cursos de Banco de Dados

Yuri Souza dos Santos
Departamento de Sistemas e Computação
Universidade Federal de Campina Grande
Campina Grande, Paraíba - Brasil
yuri.santos@ccc.ufcg.edu.br

Cláudio Elízio Calazans Campelo
Departamento de Sistemas e Computação
Universidade Federal de Campina Grande
Campina Grande, Paraíba - Brasil
campelo@computacao.ufcg.edu.br

RESUMO

O monitoramento do engajamento dos estudantes possibilita um enriquecimento do sistema educacional. Através dele é possível obter informações relacionadas a dedicação e desempenho dos alunos em determinadas atividades. Entretanto, tratando-se dos cursos de Banco de Dados, há uma escassez de estratégias para monitorar o engajamento dos discentes, dificultando a realização de pesquisas e análises do desempenho do estudante. Diante dessa problemática, desenvolvemos um mecanismo capaz de monitorar a interação dos estudantes com o banco de dados e gerar um conjunto de dados envolvendo o uso dos comandos *Data Definition Language* e *Data Manipulation Language* efetuados. Posteriormente, com esse mecanismo, monitoramos 95 discentes da disciplina Banco de Dados, do curso de Ciência da Computação da Universidade Federal de Campina Grande. Com isso, realizamos um estudo de caso, no qual utilizamos, além das informações colhidas dos bancos de dados, outras informações relacionadas à participação dos alunos da disciplina, como as notas obtidas e a submissão dos roteiros práticos. Dessa forma, contribuímos com uma ferramenta capaz de monitorar as atividades práticas dos estudantes e extrair informações que podem ser úteis no sistema educacional.

PALAVRAS-CHAVE

Engajamento dos Alunos, Monitoramento em Banco de Dados, *Data Definition Language*, *Data Manipulation Language*, DML, DDL

LINKS ÚTEIS

- **Repositório da ferramenta Web - FRONTEND:** <https://github.com/yurisd/monitoring-mechanism-Frontend>.
- **Repositório com o código do mecanismo de monitoramento em banco de dados e da ferramenta Web - BACKEND:** <https://github.com/yurisd/monitoring-mechanism-Backend>.
- **Análises no Google Colab:** https://colab.research.google.com/drive/1rzjwrr-om3_1ApxTGTgGcjKNA01EWZp?usp=sharing.

1 INTRODUÇÃO

Durante a execução de cursos voltados para área tecnológica, a realização de monitoramentos das atividades dos alunos pode auxiliar

no seu desenvolvimento intelectual, além de ser importante para o trabalho dos professores em tratando-se de verificar o engajamento e a assimilação dos conteúdos por parte dos seus alunos, pois, como Coates [2] defendeu, os dados sobre o envolvimento dos alunos têm a vantagem de fornecer informações sobre o que eles estão realmente fazendo. Podendo assim, encontrar possíveis relações entre o esforço estabelecido pelo aluno e o seu desempenho na disciplina, além de poder identificar vulnerabilidades na forma como o aluno desenvolve as suas atividades. Nesse sentido, podemos enquadrar os cursos nas áreas de Banco de Dados, que são cursos voltados para as áreas tecnológicas, com atividades práticas, além de fornecer uma boa base para a carreira do estudante.

Alguns estudos foram reportados na literatura visando melhorar a assimilação do conteúdo em disciplinas da área de Banco de Dados. Por exemplo, Martinhago et al. [3] aplicaram, no âmbito da disciplina de Banco de Dados no ensino superior, o conceito da Computação Desplugada. Conceito esse que a princípio foi criado para amenizar o problema de entendimento de conceitos de ciência da computação pelos alunos do ensino básico sem requerer a utilização direta do computador. Eles desenvolveram atividades desplugadas para alguns conceitos chave da disciplina, como modelo relacional e consultas *Structured Query Language* (SQL). O resultado do experimento indicou que o método não teve impacto significativo na aprendizagem dos conceitos, mas a aceitação das atividades foi positiva e os alunos se mostraram muito interessados e participativos. Outra iniciativa foi a de Atcharyachanvanich et al. [1] que propuseram o desenvolvimento de um sistema para processamento de instruções SQL, o MySQL Sandbox. A ideia foi construir uma ferramenta de apoio que ajudasse a eliminar as limitações existentes em sistemas *e-learning* desenvolvidos para o ensino de cursos de Banco de Dados online, por exemplo, as limitações no número de instruções SQL suportado; além de outras preocupações como, a vulnerabilidade de injeção de SQL e o risco dos alunos executarem alguma instrução SQL sensível, especialmente em instruções *Data Definition Language* DDL.

Porém, tratando-se de monitorar o engajamento dos alunos em disciplinas de Banco de Dados, ainda existe uma carência de trabalhos. Um dos motivos pode ser a escassez de opções para monitorar comandos como os de DDL e *Data Manipulation Language* (DML) no banco de dados (BD). Quanto a isso, existem iniciativas como a do Santos [4] que sugeriu uma arquitetura com mecanismos que proporcionassem o monitoramento das alterações ocorridas por eventos DDL no BD monitorado. O seu objetivo foi, com base na análise do histórico dos eventos DDL monitorados nesse banco, utilizar um sistema de recomendação que recomendasse a criação de outras possíveis tabelas.

Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

Neste trabalho, desenvolvemos um mecanismo de monitoramento de banco de dados. Ele possui a capacidade de supervisionar o BD e coletar as informações referentes aos comandos DDL e DML executados pelo aluno. Sua implementação foi efetuada utilizando recursos nativos do próprio Sistema de Gerenciamento de Banco de Dados, como gatilhos, funções e procedimentos. Além das informações coletadas no mecanismo, é possível coletar outras informações, por exemplo, as notas das atividades práticas no qual os alunos utilizaram o BD, as notas de provas, entre outras. A junção dessas informações potencializa a realização de análises descritivas e preditivas. No nosso estudo, estruturamos esses dados em um conjunto de 47 atributos. Esses dados inclui a quantidade de inserções realizadas por um determinado aluno no seu BD, a nota obtida por esse aluno na prova, a quantidade de comandos executados no período noturno, dentre outros.

Além disso, concebemos uma abordagem para análise descritiva baseado nos 47 atributos previamente estabelecidos. Com o objetivo de encontrar possíveis correlações, *insights* e tendências. Outrossim, desenvolvemos uma ferramenta Web capaz de reunir as informações provenientes do mecanismo de monitoramento, além de outros dados que podem ser coletados pelo professor da disciplina.

Feito isso, realizamos um estudo de caso utilizando informações colhidas dos discentes da disciplina de Banco de Dados, do curso de Ciência da Computação da Universidade Federal de Campina Grande (UFCG), durante os semestres letivos 2021.1 e 2021.2. Esse estudo não trouxe resultados significativos em termo de correlação, pelo fato de possuímos uma base muito pequena. Além disso, existe um possível enviesamento dos dados, pelo fato de que houve adaptações na metodologia da disciplina para que ela pudesse ser ofertada de forma remota no cenário da pandemia do COVID-19 e de forma híbrida no retorno das aulas presenciais. No entanto, o principal objetivo dessa pesquisa, foi prover um ferramental capaz de realizar análises mais robustas em um maior conjunto de dados. Podendo ser reutilizado em outros contextos, como outras instituições de ensino.

2 METODOLOGIA

Esta seção descreve a metodologia adotada para concepção da estratégia de monitoramento e análise de engajamento dos alunos em cursos de Banco de Dados. Desta forma, descrevemos nas seções 2.1 e 2.2, os mecanismos desenvolvidos para monitorar, armazenar, coletar e manipular os comandos *Structured Query Language* (SQL) realizados por esses acadêmicos durante as atividades práticas da disciplina. Além disso, descrevemos na seção 2.3, as estratégias utilizadas para coletar informações relacionadas à participação dos estudantes durante as aulas.

2.1 Mecanismo de monitoramento e coleta

Desenvolvemos um mecanismo capaz de monitorar, armazenar e coletar os comandos *Data Definition Language* (DDL) e *Data Manipulation Language* (DML) do banco de dados. Para o seu desenvolvimento, fizemos uso de recursos nativos do próprio Sistema de Gerenciamento de Banco de Dados (SGBD), como gatilhos e funções, utilizando a linguagem de programação em banco de dados PL/pgSQL¹.

O mecanismo assume as atividades práticas consiste de comandos que são executados dentro de um SGBD PostgreSQL e que cada aluno possui um BD dentro desse SGBD. Esse BD possui o esquema *auditable* que, por sua vez, possui a tabela *TAB_EVENT_LOGS*. É nessa tabela que o mecanismo armazena todos os comandos executados por cada aluno. Essa tabela possui as seguintes colunas:

- **ID**: Identificador único da tabela.
- **DATE_TIME**: Contém a informação do horário em que o comando foi executado.
- **EVENT_NAME**: Contém o nome do tipo de comando executado, por exemplo: CREATE TABLE, INSERT.
- **TABLE_NAME**: Contém o nome da tabela que está sendo criada ou manipulada pelo usuário.
- **ROW_NEW**: Possui uma tupla com o novo conteúdo inserido na tabela.
- **ROW_OLD**: Possui uma tupla com o conteúdo já existente na tabela.
- **CURRENT_QUERY**: Salva o comando SQL executado pelo usuário.

Posto isso, foram desenvolvidos quatro tipos de funções. Elas possuem gatilhos de eventos, que são ativadas quando ocorre os seguintes comandos DDL:

- **CREATE TABLE**: Quando o estudante executa esse tipo de comando, são coletadas as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, o nome da tabela e o comando SQL executado pelo aluno. Além disso, após o usuário criar a tabela, o mecanismo adiciona um gatilho responsável por monitorar os comandos DML do tipo INSERT, DELETE e UPDATE que, a posteriori, serão executados nesta tabela. Sempre que esse gatilho é acionado, são gerados outros três gatilhos (um para cada comando DML) atrelados a essa tabela, responsáveis por monitorar a ocorrência de cada um desses comandos:
 - **INSERT**: Na ocorrência desse evento, coletamos as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, o nome da tabela no qual está ocorrendo o evento, a tupla com os dados inseridos na tabela, além do SQL executado pelo estudante.
 - **DELETE**: Quando o estudante executa esse tipo de comando, o mecanismo coleta as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, o nome da tabela no qual está ocorrendo o evento, a tupla com os dados removidos da tabela, além do SQL executado pelo estudante.
 - **UPDATE**: Nesse tipo de comando, o mecanismo coleta as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, a tupla com os novos dados inseridos na tabela, a tupla com os antigos dados da tabela e o comando SQL
- **DROP TABLE**: Na ocorrência desse tipo de comando, são coletadas as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, o nome da tabela removida e o comando SQL executado pelo estudante. Ademais, após o usuário remover a tabela, o mecanismo também remove os gatilhos atrelados essa tabela, evitando assim possíveis problemas de inconsistência.

¹Disponível em: <<https://www.postgresql.org/docs/current/plpgsql.html>>

Tabela 1: Demonstração do que é salvo de cada comando na tabela de logs.

Tipo	Possui registro?					
	event_name	table_name	data_time	row_old	row_new	current_query
INSERT	Yes	Yes	Yes	No	Yes	Yes
DELETE	Yes	Yes	Yes	Yes	No	Yes
UPDATE	Yes	Yes	Yes	Yes	Yes	Yes
CREATE TABLE	Yes	Yes	Yes	No	No	Yes
DROP TABLE	Yes	Yes	Yes	No	No	Yes
ALTER TABLE	Yes	Yes	Yes	No	No	Yes
Outros eventos DDL	Yes	No	Yes	No	No	Yes

- **ALTER TABLE:** Nesse tipo de comando, o mecanismo coleta as informações referentes ao horário da ocorrência do comando, o tipo de comando executado, o nome da tabela alterada e o comando SQL executado pelo aluno.

Além desses comandos, o mecanismo também coleta as informações de outros tipos de eventos DDL, como os comandos *CREATE FUNCTION* e *DROP OWNED*. Nesse caso, são colhidas informações referentes ao horário da ocorrência do comando, o tipo de comando executado e o comando SQL executado pelo estudante.

A Tabela 1 contém o que é salvo de cada comando na tabela `TAB_EVENT_LOGS`. Apenas as colunas definidas com *Yes* são preenchidas com o conteúdo dos comandos efetuados.

A Figura 1 mostra a arquitetura do mecanismo desenvolvido para capturar comandos DDL. Todo comando passa pela função `FN_LOG_EVENT`, que decidirá qual das quatro funções de monitoramento serão executadas. Caso a função executada seja, por exemplo, a de *CREATE TABLE*, o mecanismo cria uma função responsável por monitorar todos os comandos DML que passarem por aquela tabela criada.

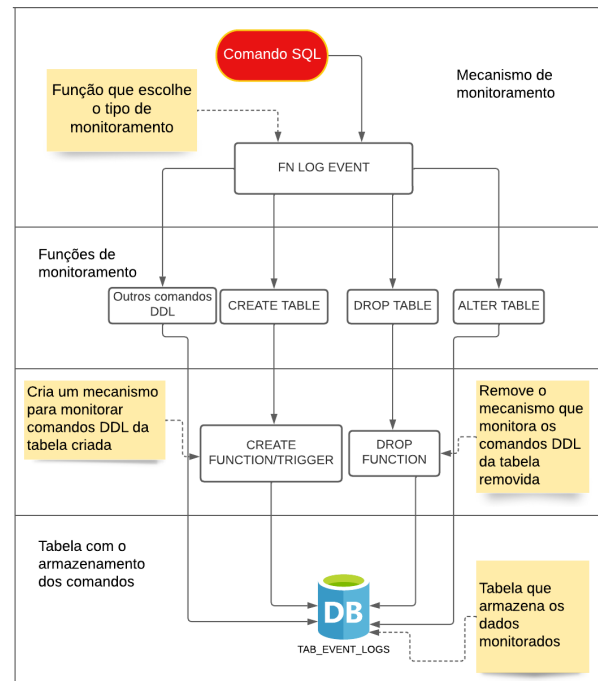
2.2 Estruturação dos dados

Realizamos uma estruturação dos dados visando derivar as informações armazenadas na tabela `TAB_EVENT_LOGS` de cada aluno em novos componentes, de forma que essas informações ficassem no formato apropriado para a realização das análises. Para isso, realizamos uma conexão com o BD de cada estudante, utilizando recursos da linguagem Javascript e da biblioteca `pg`². Com isso, conseguimos manter uma comunicação contínua com os bancos. Dessa maneira, conseguimos resgatar as informações, além de realizar filtros por tipo de comando e a sua data de ocorrência.

Para a estruturação dos dados, também utilizamos a linguagem Javascript. Dessa forma, transformamos as informações em um conjunto de dados, sendo que cada coluna representa um atributo estruturado e cada linha equivale ao conteúdo estruturado de um aluno. Esse conjunto possui 37 colunas com valores quantitativos e pode ser armazenado em um arquivo no formato *Comma Separated Values* (CSV). Abaixo listamos algumas dessas colunas:

- **create_table:** Quantidade de comandos *CREATE TABLE* realizado no BD.
- **insert:** Quantidade de comandos *INSERT* realizado no BD.
- **varchar type:** Quantidade de colunas com o tipo `VARCHAR`.
- **worked_21h_24h:** Quantidade de comandos efetuados no período entre às 21 horas e 24 horas.
- **drop_columns:** Quantidade de colunas removidas do BD.

²Disponível em: <<https://node-postgres.com/>> Acesso em 11 ago, 2022

**Figura 1: Arquitetura do mecanismo para monitoramento do Banco de Dados.**

- **add_constraint:** Quantidade de restrições adicionadas no BD através de um comando *ALTER TABLE*.

2.3 Outros dados coletados

Além dos 37 dados estruturados na seção 2.2 que julgamos serem úteis para as nossas análises, existem outros tipos de atributos úteis, que podem ser coletados manualmente pelo professor, para a realização das análises. Atributos como as notas das provas, se o aluno esteve presente nas aulas, dentre outros.

No nosso caso, coletamos, de forma manual, outros dados promissores para a realização do estudo de caso. Esses dados contabilizam 10 novos atributos, contendo informações como as notas obtidas nos roteiros práticos, provas, e *quizzes*; horário do envio das atividades; entre outras. Ao juntar as informações coletadas pelo mecanismo com esses novos dados, geramos um novo conjunto de dados, no formato de arquivo CSV, totalizando um total de 47 atributos.

Para realizar a coleta desses dados, realizamos um acompanhamento dos alunos durante dois semestres letivos e armazenamos essas informações em uma planilha. Após isso, integramos os dados dessa planilha com os dados estruturados do mecanismo de monitoramento. Para isso, utilizamos recursos da linguagem Javascript e da biblioteca `google-spreadsheet`³.

³Disponível em: <<https://theophrain.github.io/node-google-spreadsheet>> Acesso em 11 ago, 2022

3 MECANISMO DE ANÁLISE DOS DADOS

Nesta seção explicamos os recursos utilizados para realizar as análises descritivas dos dados. Nessas análises usamos recursos como Histogramas, Boxplots e Agrupamentos. Sendo assim, descrevemos as estratégias realizadas para visualizar e interpretar essas análises. Esses métodos serão utilizados no nosso estudo de caso, na seção 4.

Para ser possível visualizar as análises, nós utilizamos dois procedimentos: o Google Colab 3.1 e o desenvolvimento de uma Ferramenta Web 3.2.

3.1 Google Colaboratory

Com o auxílio do Google Colab e utilizando recursos da linguagem Python, realizamos uma análise descritiva dos dados. Para isso, esses dados passaram por um pré-processamento, visando desfragmentar os dados em novas colunas, além de criar novos grupos de dados para serem utilizados no estudo de caso.

3.2 Ferramenta Web

Desenvolvemos uma ferramenta web capaz de reunir informações visuais sobre as análises realizadas. Esse mecanismo foi desenvolvido utilizando Node.js no Backend e React.js no Frontend, com o auxílio da biblioteca Plotly⁴ para a geração dos gráficos e da biblioteca Ant Design⁵ para a estilização das telas. Ele possui conexão com os bancos de dados, além de se comunicar com uma planilha que contém as notas dos alunos.

A Figura 4 mostra uma tela da ferramenta visual, sendo possível visualizar os comandos efetuados no BD, os principais tipos de dados utilizados e as notas do aluno responsável pelo BD.

4 ESTUDO DE CASO

No nosso estudo de caso, utilizamos dados de uma turma de dois semestres letivos (2021.1 e 2021.2) dos discentes da disciplina de Banco de Dados, do curso de Ciência da Computação da UFCG. A turma do semestre 2021.1 possuiu 45 alunos e funcionou de forma remota, devido à pandemia do COVID-19. Enquanto isso, A turma do semestre 2021.2 possuiu 50 alunos e funcionou no formato híbrido, em que parte do conteúdo era trabalhado em sala de aula e a outra parte era trabalhado em outro ambiente. Nesse estudo, utilizamos informações envolvendo as atividades práticas realizadas pelos estudantes. Por exemplo, os roteiros práticos, em que eram definidos algumas orientações para a realização das atividades e era esperado que o aluno as seguisse. Cabe salientar que os dados podem estar enviesados. Ressaltamos também, que o nome dos alunos foram substituídos por nomes fictícios.

4.1 Relações entre o uso dos comandos DDL

Realizamos um experimento, utilizando os números de cada comando CREATE TABLE, ALTER TABLE e DROP TABLE, efetuados pelos alunos do semestre 2021.2. A Figura 2 representa uma amostra de 20 alunos, retirados do resultado desse experimento. Podemos perceber uma tendência entre os alunos em realizar mais comandos de ALTER TABLE (na cor verde) do que comandos CREATE TABLE

(na cor laranja). As colunas em azul, representam os comandos DROP TABLE.

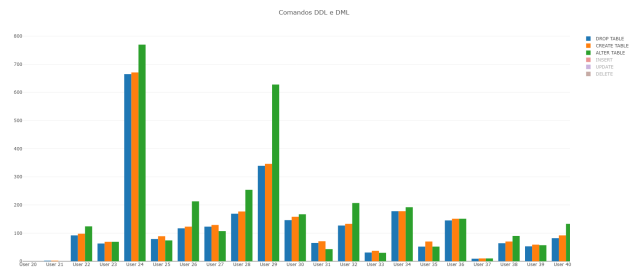


Figura 2: Distribuição dos comandos DDL realizados pelo estudante

Decidimos então calcular a razão entre o número de ALTER TABLE e o número de CREATE TABLE. Posteriormente, realizamos um agrupamento utilizando essa relação e as notas obtidas pelo aluno na prova 1. Nessa prova, para alguns assuntos era necessário o entendimento do comando CREATE TABLE, como também o entendimento dos tipos de dados e definições das restrições no uso desse comando. Além disso, foram abordados assuntos envolvendo o uso dos comandos INSERT e SELECT. A Figura 3 representa esse agrupamento, sendo o eixo x a distribuição das razões e o eixo y a distribuição das notas. Um fato que chama a atenção é o número de alunos que estavam com a razão menor que 0.8 e obtiveram notas inferiores a 6 (pontos laranjas). Por se tratar de um único semestre, ainda é cedo para tirar alguma conclusão, especialmente pelo fato desse assunto não ter sido o foco principal da prova. Porém, os resultados indicam que pode existir uma tendência dos alunos que fazem mais ALTER TABLE do que CREATE TABLE estarem com alguma dificuldade para assimilar o conteúdo da disciplina de forma geral, resultando em uma maior dificuldade na realização da prova.

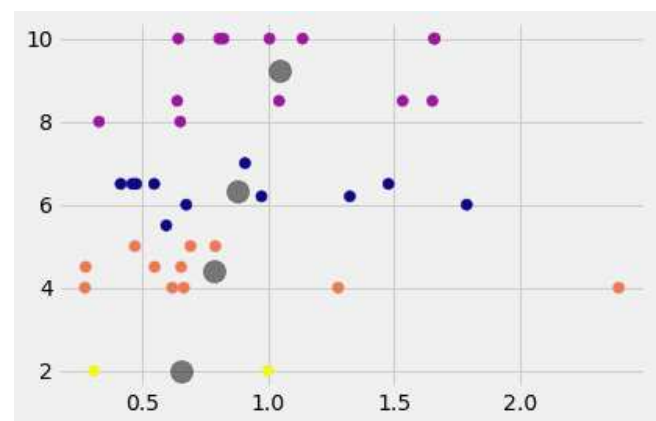


Figura 3: Agrupamento executado utilizando os seguintes dados: razão entre as quantidades de comandos CREATE TABLE e ALTER TABLE (eixo x); e notas da Prova 1 (eixo y).

⁴Disponível em: <<https://plotly.com/>> Acesso em 14 ago, 2022

⁵Disponível em: <<https://ant.design/>> Acesso em 22 ago, 2022

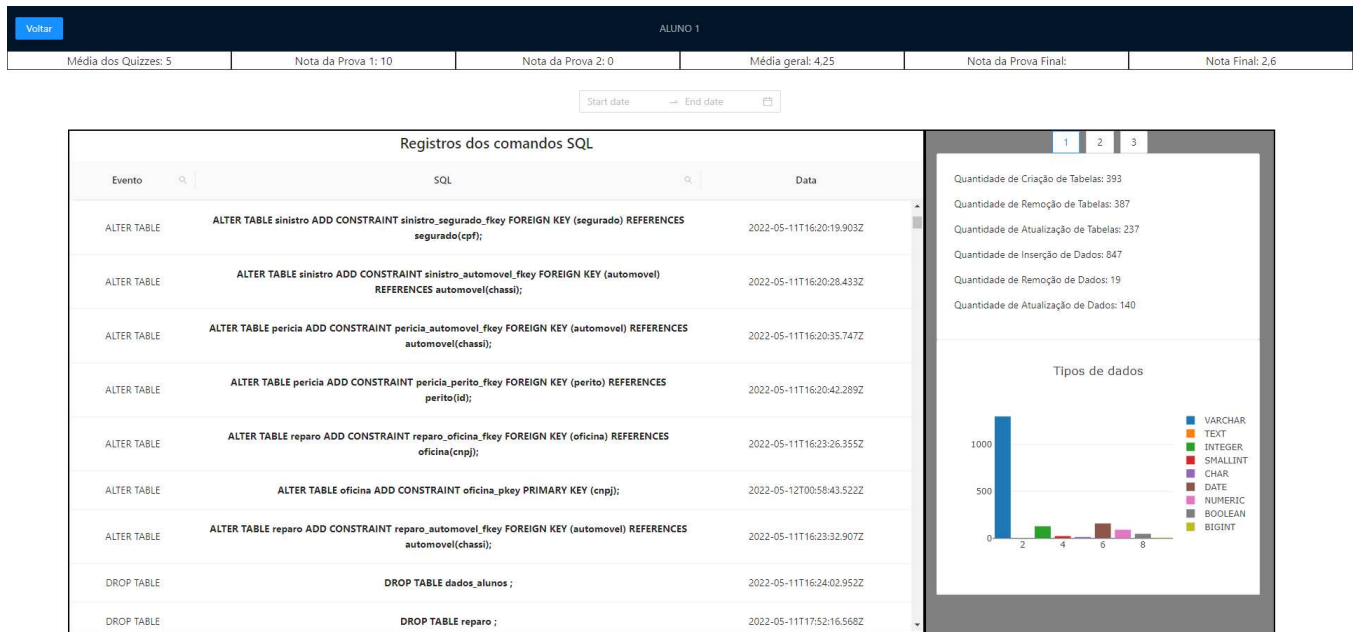


Figura 4: Tela da ferramenta visual onde mostra as informações do aluno.

4.2 Variação das notas nos roteiros

Analisamos a distribuição das notas nos roteiros 1 a 4 do semestre 2021.1, representados na Figura 5. O que chama a atenção é a crescente nas notas 0 (ZERO), enquanto houve uma diminuição nas notas 10 (DEZ). Uma justificativa pode ser o abandono da disciplina por parte dos estudantes.

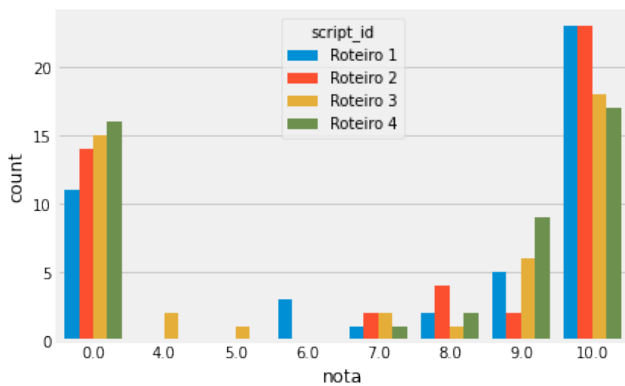


Figura 5: Distribuição das notas nos roteiros 1 a 4 durante o semestre letivo 2021.1

Resolvemos analisar a distribuição das notas, dessa vez, apenas dos alunos que interagiram com o seu BD, representados na Figura 6. Com esse gráfico conseguimos confirmar que boa parte das notas 0 (ZERO), vistos na Figura 5, tratava-se de alunos que não haviam interagido com o seu BD. Porém, outra informação que nos chamou

atenção, foi um grupo de alunos que tiraram 10 (DEZ) no roteiro 2, sem ter interagido com o seu respectivo banco.

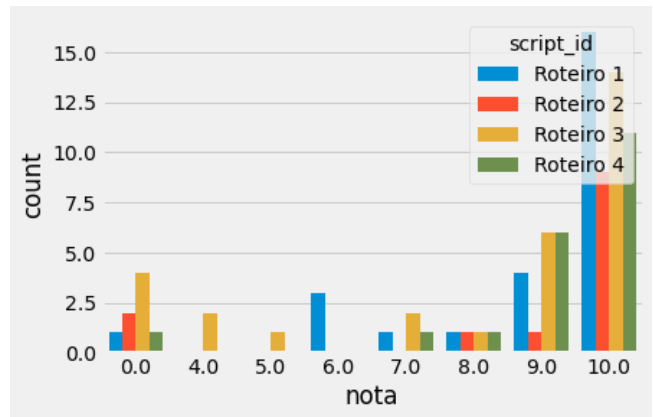


Figura 6: Distribuição das notas dos alunos que interagiram com o seu BD nos roteiros 1 a 4 durante o semestre letivo 2021.1

4.3 Correlação entre os roteiros

Realizamos uma correlação entre os comandos DDL e DML com a nota obtida nos roteiros 1, 2 e 3 do grupo de alunos do semestre 2021.2. Nosso intuito era observar o impacto do uso dos comandos com a nota do roteiro.

A Figura 7 representa a correlação entre esses comandos durante o roteiro 1. Nesse roteiro era esperado que o aluno trabalhasse

com comandos DDL, como *CREATE TABLE*, *ALTER TABLE* e *DROP TABLE*. Observamos uma correlação de 31% para *CREATE TABLE*, 28% para *ALTER TABLE* e 30% para *DROP TABLE* em relação com as notas obtidas nesse roteiro.

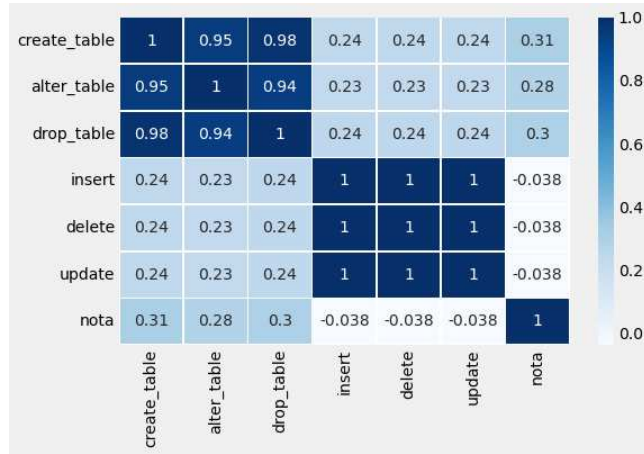


Figura 7: Correlação dos comandos DDL e DML no roteiro 1

A Figura 8 representa a correlação entre esses comandos durante o roteiro 2. Nesse roteiro era esperado que o aluno continuasse trabalhando com os comandos DDL, além de iniciar os estudos dos comandos DML, como *INSERT*, *DELETE* e *UPDATE*. Conseguimos visualizar uma maior correlação dos comandos DML em relação com as notas obtidas nesse roteiro, quando comparados aos comandos DDL.



Figura 8: Correlação dos comandos DDL e DML no roteiro 2

Já no roteiro 3, representado pela Figura 9 era esperado que o aluno realizasse alguns comandos DDL mais elaborados, de forma que, ao utilizar os comandos DML, as regras definidas pelos comandos DDL fossem respeitadas. Regras como, valor único na tabela e tipos de valores que a coluna aceita. Notamos que, apesar dos comandos DML possuírem maior correlação em relação às notas obtidas nesse roteiro, os comandos DDL também possuíram uma

alta correlação, demonstrando a importância do aluno realizar uma boa modelagem na construção das suas tabelas.

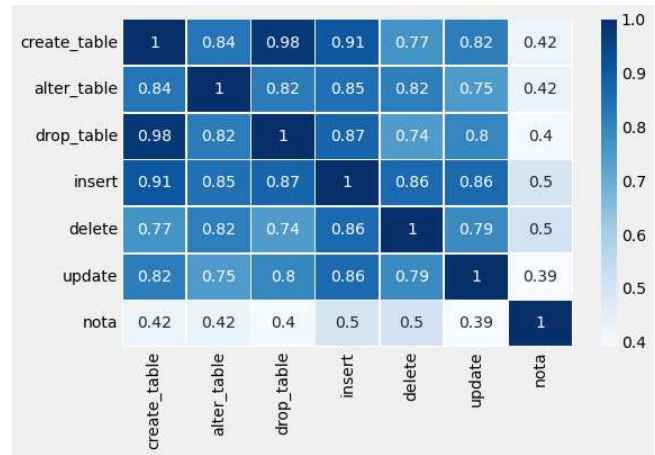


Figura 9: Correlação dos comandos DDL e DML no roteiro 3

4.4 Sessões de utilização do banco de dados

Por último, observamos a quantidade de sessões que os alunos dedicaram para o uso do BD durante o semestre 2021.2. Uma sessão era tida como finalizada quando o estudante passava, pelo menos uma hora, sem realizar nenhum comando. Posteriormente realizamos um agrupamento entre essas sessões e as notas obtidas na prova 1, representados na Figura 10. Para encontrar o número de grupos utilizamos a técnica do cotovelo e, conseqüentemente, encontramos 5 grupos. Porém, não conseguimos identificar uma relação entre a quantidade de vezes que o estudante dedicou para os estudos e a sua conseqüente nota na prova.

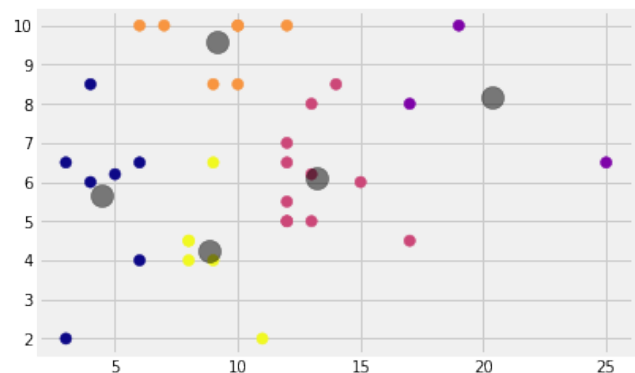


Figura 10: Agrupamento da sessão de estudos com as notas obtidas na prova

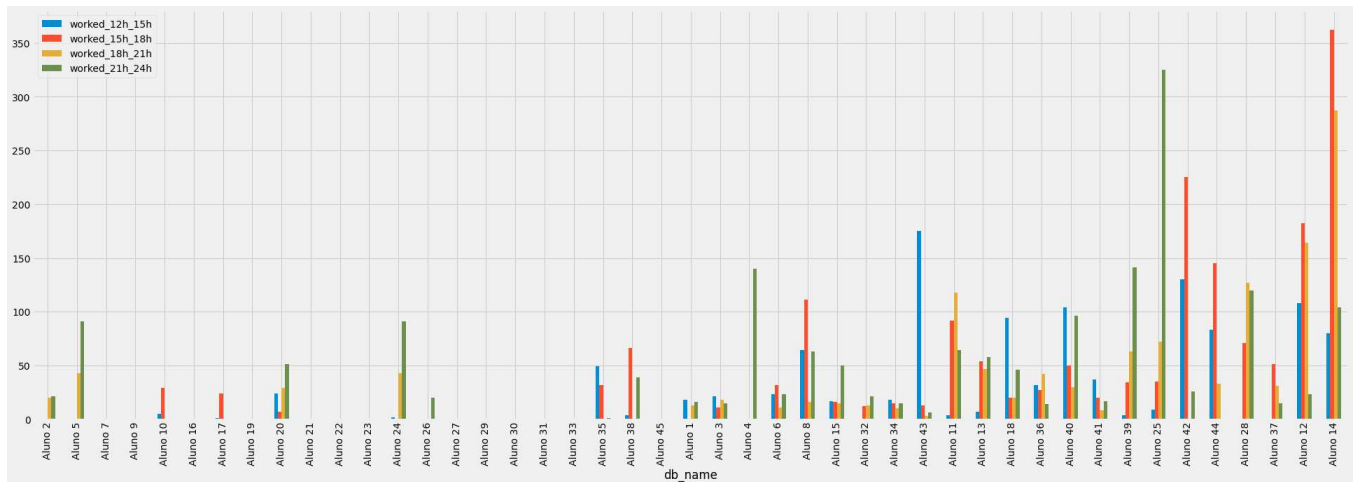


Figura 11: Frequência de interação no horário noturno no semestre 2021.1.

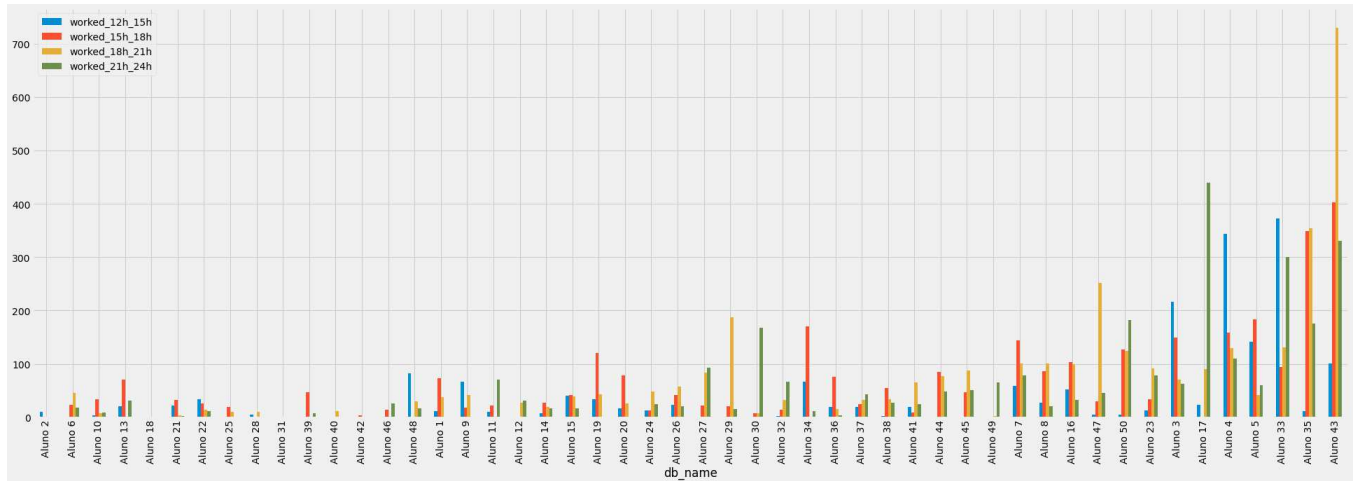


Figura 12: Frequência de interação no horário noturno no semestre 2021.2.

4.5 Interação entre os alunos e os seus respectivos bancos de dados durante as suas atividades no horário noturno

Aqui visualizamos a frequência de interação entre os alunos e os seus respectivos bancos de dados durante as suas atividades no horário noturno. A Figura 11 representa a distribuição no semestre 2021.1, enquanto a Figura 12 representa a distribuição no semestre 2021.2. Abaixo listamos as cores e a sua representação no gráfico.

- **Azul:** Representam a frequência de interações nos horários entre 12h e 15h.
- **Vermelho:** Representam a frequência de interações nos horários entre 15h e 18h.
- **Amarelo:** Representam a frequência de interações nos horários entre 18h e 21h.
- **Verde:** Representam a frequência de interações nos horários entre 21h e 24h.

Observamos que durante o semestre 2021.1, alguns alunos não interagiram com o seu BD. Diferente do semestre 2021.2, no qual apenas 2 alunos não interagiram com o BD. Observamos também uma preferência no período entre 15h e 21h.

5 CONCLUSÕES E TRABALHOS FUTUROS

Nesse artigo, realizamos um estudo envolvendo as informações colhidas dos alunos de Banco de Dados, com o objetivo de encontrar correlações e tendências. Mais especificamente, utilizamos dados relacionados as atividades práticas, desenvolvidas pelos alunos. Para isso, foi necessário desenvolver um mecanismo capaz de monitorar e coletar informações a respeito dos comandos SQL realizado pelos estudantes. Além de outras informações, como a nota dos alunos nos roteiros da disciplina.

Após isso, concebemos um mecanismo de análise de dados, utilizando conceitos de ciência de dados descritiva, a fim de poder

realizar análises envolvendo os dados dos alunos. Com isso, realizamos um estudo de caso, utilizando informações dos alunos de Banco de Dados da UFCG. Nele não encontramos grandes achados, porém, com o mecanismo desenvolvido, possibilitamos a continuação desse estudo. Esse estudo de caso pode ser visualizado por meio do Google Colab e por uma ferramenta Web desenvolvida por nós.

Acreditamos que o nosso trabalho traz grandes contribuições nesse cenário de monitoramento do engajamento dos alunos. Principalmente tratando-se de engajamento na área de Banco de Dados, que ainda é pouco explorada na literatura. Com a construção do mecanismo de monitoramento, permitimos a realização de análises em uma base de dados mais robusta, podendo subsidiar outras pesquisas.

Porém, tivemos uma limitação quanto ao monitoramento de comandos do tipo *SELECT*. Podendo ser tema de um trabalho futuro. Além disso, pretendemos aumentar a nossa base de dados, utilizando

informações dos próximos semestres letivos. Por sua vez, este maior volume de dados vai viabilizar a implementação de modelos preditivos, além da realização de análises exploratórias/descriptivas mais profundas.

REFERÊNCIAS

- [1] Kanokwan Atcharyachanvanich, Srinual Nalintippayawong, and Tanasab Permpool. 2017. Development of a MySQL Sandbox for processing SQL statements: Case of DML and DDL statements. In *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. 1–6. <https://doi.org/10.1109/JCSSE.2017.8025930>
- [2] Hamish Coates. 2005. The value of student engagement for higher education quality assurance. *Quality in Higher Education* 11 (04 2005), 25–36. <https://doi.org/10.1080/13538320500074915>
- [3] Adriana Martinhago, Rodrigo Smarzarro, Iasmini Lima, and Leidiane Guimarães. 2014. Computação Desplugada no Ensino de Bancos de Dados na Educação Superior. In *Anais do XXII Workshop sobre Educação em Computação*. SBC, Porto Alegre, RS, Brasil, 21–29. <https://sol.sbc.org.br/index.php/wei/article/view/10956>
- [4] Paulo Henrique Santos. 2014. Uma arquitetura para monitoramento de banco de dados e recomendações utilizando sistema de banco de dados ativos.