

Relatório 2 – Projeto Integrador III

Alunos: Caio Siquara, Filipe Freitas, Gustavo Calanzani, Pedro Almeida e Rafael Marques

Prof.: Howard Cruz Roatti

Pergunta a ser respondida: Quais fatores influenciam no sucesso ou fracasso acadêmico de um aluno?

O QUE ESTÁ SENDO ESTUDADO

Primeiramente é necessário compreender o que está sendo analisado durante este projeto.

Foi obtido um arquivo CSV no Kaggle com notas dos alunos de uma faculdade em 3 disciplinas (matemática, leitura e escrita). Junto de suas notas, haviam 5 colunas com informações sobre: gênero do aluno, raça/etnia do aluno, nível de escolaridade de seus pais (maior nível), tipo de alimentação e se haviam feito um curso preparatório.

Diante disso, foi realizado um tratamento utilizando o Jupyter Notebook para garantir que todos os alunos possuem notas nas 3 disciplinas e que as notas são válidas (valores entre 0 e 100). Logo após, foram criadas as duas tabelas, uma fato e uma dimensão, em um arquivo “.db” utilizando o próprio Jupyter Notebook e criando as estruturas no SQLite. A primeira tabela é composta apenas pelas notas dos estudantes e a segunda pelos outros campos do CSV.

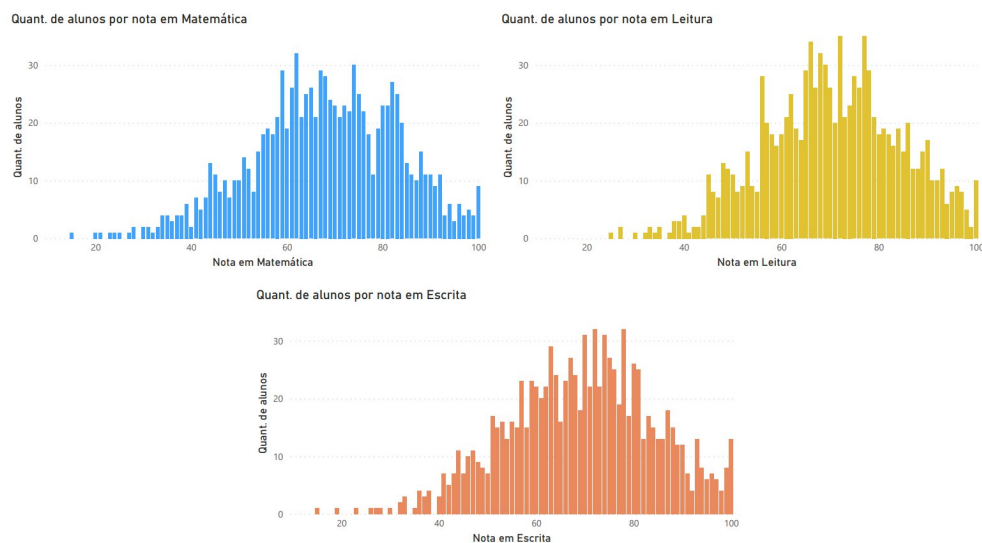
Já para a parte análise exploratória, o arquivo “.db” foi carregado para o PowerBI para a criação de tabelas, gráficos e histogramas que nos permitisse compreender e tirar conclusões das informações.

ANALISANDO OS DADOS

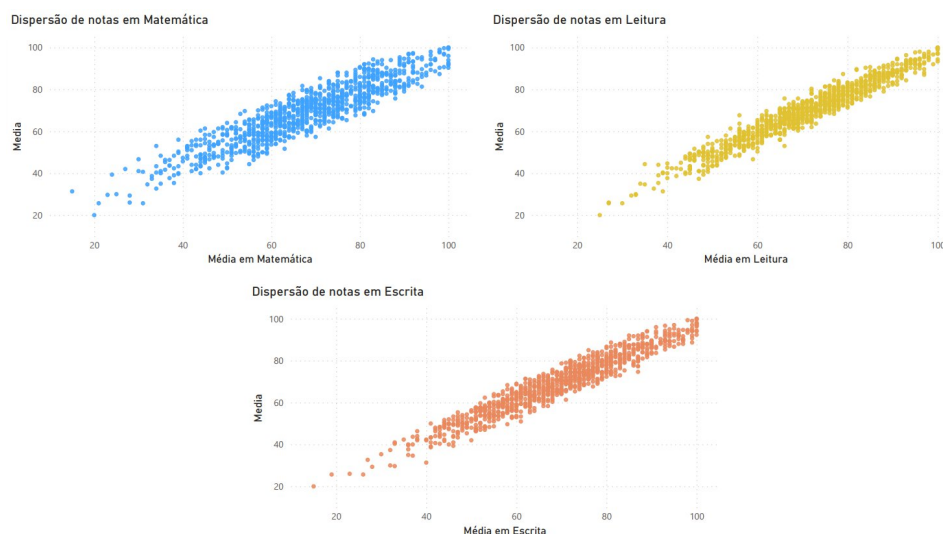
DISTRIBUIÇÃO DOS DADOS

Antes de se discutir as conclusões da análise exploratória, é necessário compreender a distribuição dos dados dos alunos.

Na tabela abaixo há a distribuição da quantidade de estudantes por sua nota em cada uma das três disciplinas:

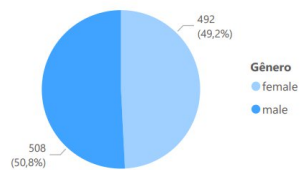


É possível perceber através desses histogramas que a maior parte dos alunos possui nota entre 50 e 90 pontos. Ademais, com o auxílio do gráfico de dispersão abaixo, infere-se também que em geral o nível de conhecimento de um determinado estudante é equivalente entre as 3 disciplinas.

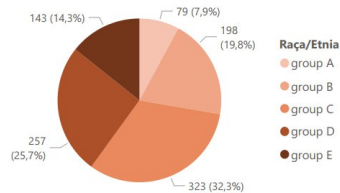


Junto das informações anteriores, também é importante entender a distribuição de alunos por seu perfil. Os gráficos abaixo informam sobre a quantidade de alunos por gênero, raça/etnia, escolaridade dos pais, tipo de alimentação e se fez curso preparatório dessa faculdade.

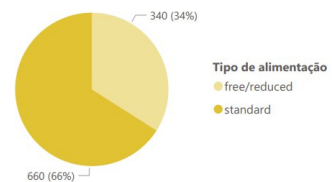
Alunos por Gênero



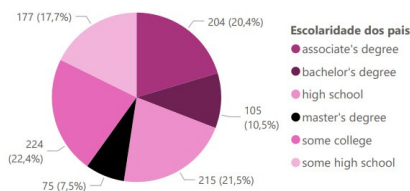
Alunos por Raça/Etnia



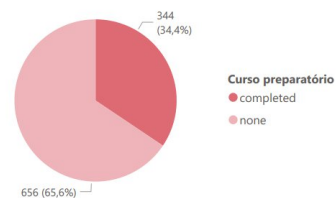
Alunos por Tipo de alimentação



Alunos por Escolaridade dos pais



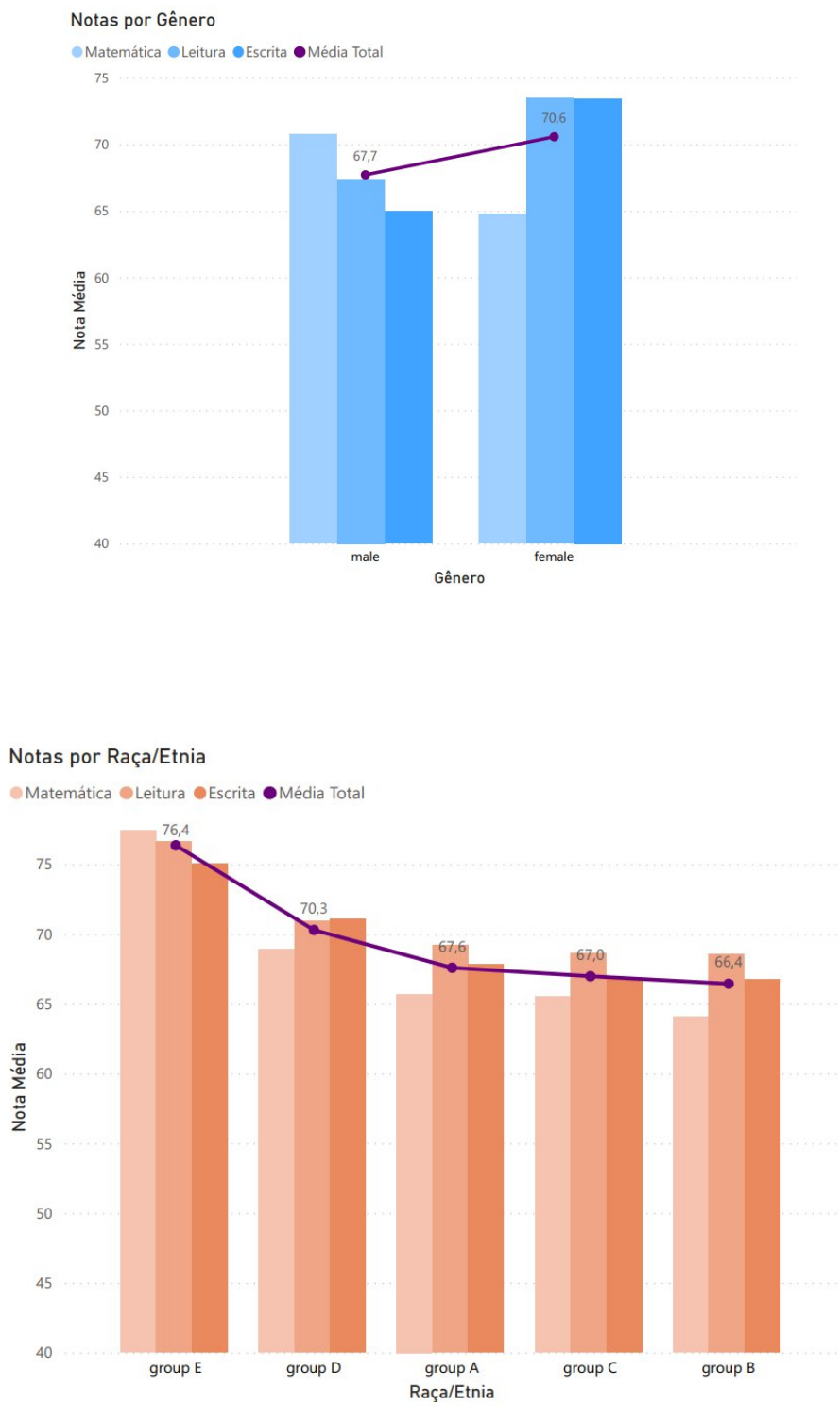
Alunos por Curso preparatório



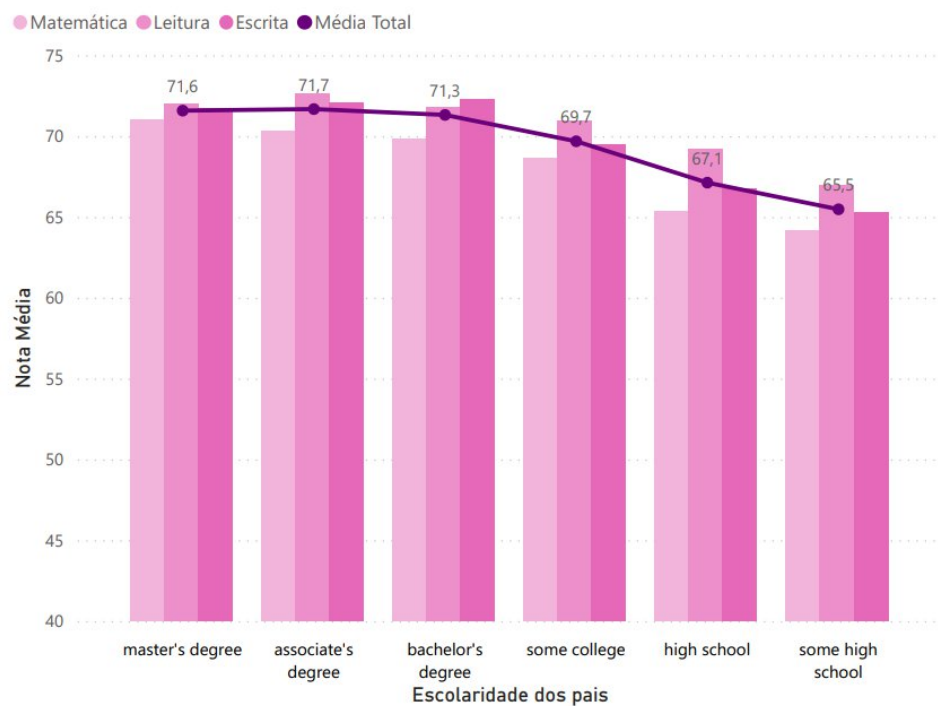
Um ponto interessante a se notar dos gráficos de pizza acima é que a quantidade de alunos que fez curso preparatório é muito próxima da quantidade de alunos que se alimenta com uma refeição gratuita na instituição de ensino.

ANÁLISE EXPLORATÓRIA DOS DADOS

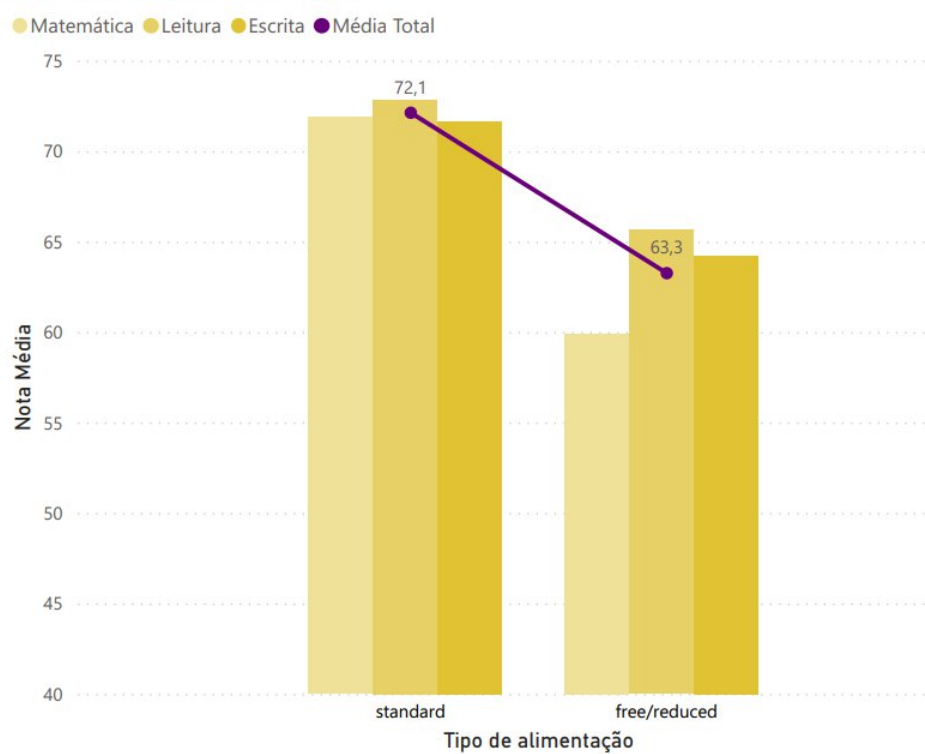
Uma vez compreendida a distribuição do perfil dos alunos e das notas, deu-se início ao trabalho de análise exploratória dos dados. Nos gráficos abaixo estão os resultados de cada disciplina e a média total dos discentes divididos por cada tipo de perfil.



Notas por Escolaridade dos pais

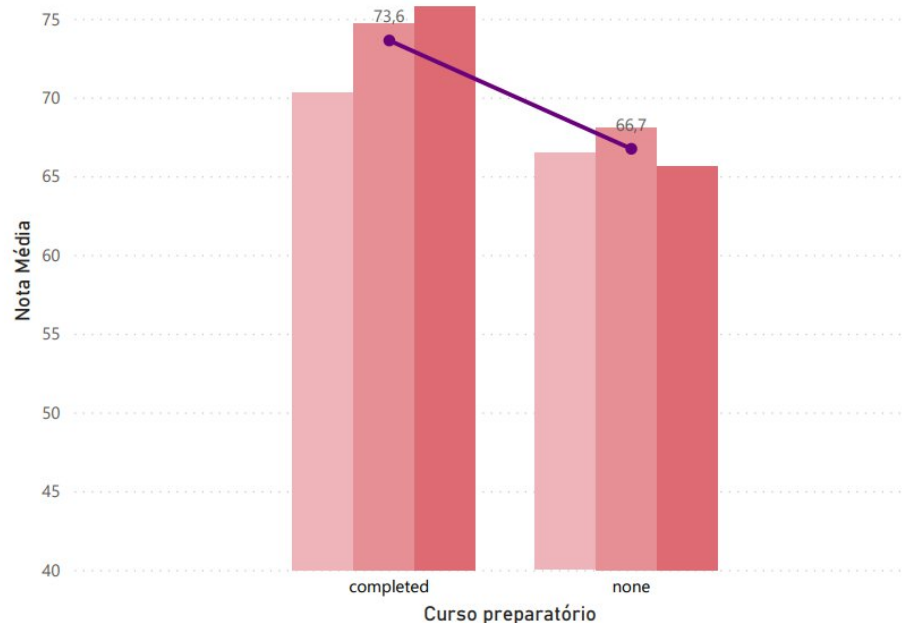


Notas por Tipo de alimentação



Notas por Curso preparatório

● Matemática ● Leitura ● Escrita ● Média Total



Diante da análise dos dados expostos acima, é possível identificar padrões em cada tipo de perfil dos estudantes.

Referente ao gênero, é possível observar que em geral, as mulheres se saem melhor nas disciplinas de leitura e escrita e os homens se saem melhor em matemática. Apesar disso, o gênero parece ter pouca influência sobre o resultado de aprovação, uma vez que a diferença entre eles é inferior a 3 pontos.

Tratando da etnia, é possível observar que, em geral, o grupo E se saiu consideravelmente melhor que o resto dos estudantes, sobretudo na área de matemática.

Já quando observamos os dados relativos ao nível de escolaridade dos pais, é possível notar um claro crescimento das notas dos alunos que possuem família com um nível maior de instrução. Logo, quanto maior o grau de instrução dos pais, maior a tendência de que seu filho vá bem nos testes.

No que se refere ao tipo de alimentação, observamos uma tendência de que alunos que se alimentam de forma completa ao invés de alimentação

gratuita conseguem se sobressair sobre os outros estudantes que se alimentam com a opção reduzida.

Por último, no que diz respeito aos universitários que realizaram curso preparatório, notamos que estes, em geral, se saíram melhor nos testes do que aqueles que não fizeram/completaram o curso.

Com base no que foi analisado, é possível perceber que diversos fatores podem afetar o desempenho acadêmico dos alunos como gênero, etnia, nível de escolaridade dos pais, nutrição e cursos preparatórios. Entretanto, esses são apenas algumas dos fatores e que estão disponíveis na base de dados, mas existem inúmeras outras variáveis no mundo real que não estão sendo consideradas nas análises.