

ANALISTA DE DADOS
(PROJETO FINAL)

RELATÓRIO DESCRITIVO DO PROJETO FINAL DE CURSO NA
INSTITUIÇÃO JDL - JOVEM DATA LEARNING

EQUIPE 04

MARIA SALETE MACIEL RODRIGUES

ARTHUR GOMES LUZ

2024

ANALISTA DE DADOS

**ANÁLISE DE DESEMPENHO ACADEMICO VOLTADAS PARA O ENSINO
MÉDIO**

EQUIPE 04

MARIA SALETE MACIEL RODRIGUES

ARTHUR GOMES LUZ

2024

SUMÁRIO

1. Contexto geral, qual o tema geral do trabalho, onde, tem, e que necessidade é o foco das análises?
2. questão de pesquisa, quais são as questões de análise que se pretende responder com dados?
3. Hipóteses, qual a predição das respostas que temos antes de ver os dados?
4. Contexto da análise, por que esta questão é de interesse, quais informações básicas levou a hipótese, e por que isso é importante?
5. Conjunto de dados, quais dados estão sendo usados para responder as perguntas, descreva os conjuntos de dados.
6. Quais métodos foram utilizados para analisar os dados?
7. Conclusões, quais são os principais resultados da análise? Quais são as respostas as questões? Compare com as predições, quais respostas não conseguiram ser obtidas e por quê?
8. Recomendações, a partir do aprendizado obtido, quais são as ações que o grupo recomendaria a serem adotadas, o que precisaria ainda ser estudado?
9. Ética e privacidade, reconheça e trate quaisquer questões que envolvam ética e privacidade relacionada ao projeto.
10. Discussão, discuta o potencial e impacto cívico do projeto. Discuta todos os problemas encontrados e enumere as limitações do seu projeto, assim como novas questões que poderiam ser abordadas para complementar a análise do tema.
11. – 12. Gráficos mais importantes e suas análises

- **Contexto geral, qual o tema geral do trabalho, onde, tem, e que necessidade é o foco das análises?**

Tema geral: Análise de desempenho acadêmico voltadas para o ensino médio.

Onde: Em uma escola de ensino médio de Guaramiranga.

Foco da análise: Optamos por utilizar a escola/ensino como projeto, por ser algo fundamental e partindo da nossa análise recomendar o que for necessário para a instituição, preferimos tratar desse assunto por ser algo que está desde a infância em nosso cotidiano percebemos quanto é um assunto não valorizado. Com o foco em trazer melhorias para educação, partindo do estudo das notas dos alunos, não somente essa escola poderá agregar nossos métodos, mas qualquer instituição que esteja buscando resultados.

- **Questão de pesquisa, quais são as questões de análise que se pretende responder com dados?**

Partindo de nossas análises, queremos tratar de evolução dos alunos em relação as metas que a escola propõe, observar o andamento das notas durante os bimestres entre os anos de 2020 a 2023. Metas por exemplo em relação a média de nota que se faz necessária para não ser reprovado em uma prova, em que eles adotam a média 6 para não ficar de recuperação.

Nas relações entre os gráficos que fizemos queremos mostrar por exemplo: Média de notas por sexo (Feminino e Masculino), Disciplinas que possuem notas mais altas, saber se os alunos bateram a meta proposta pela escola, a diferença de quantidades de alunos referente a cada ano etc.

- **Hipóteses, qual a predição das respostas que temos antes de ver os dados?**

- Os alunos conseguiram ultrapassar as medias
- Haverá uma quantidade mista de alunos seja masculino ou feminino
- Os alunos vão tirar notas mais altas nas disciplinas das humanas e linguagens, diferente das exatas.
- No ano de 2020 é possível que os alunos tenham uma baixa nas notas, pelo período pandêmico

- No ano de 2021 é possível que nos últimos bimestres os alunos tenham notas mais altas, pois o governo instituiu que todas as escolas deveriam passar os alunos independente se tivessem na média estipulada ou não
- No ano de 2022 a 2023 é possível que aconteça uma falta de alunos também pelo fato da proposta do novo ensino médio, fazendo com que todas as escolas adotassem o ensino integral entre várias outras propostas em relação ao ensino, fazendo com que os jovens optassem por não concluir seu ensino por precisar estudar o dia inteiro.
- É possível que no período entre 2020 e 2021 aconteça uma diminuição na quantidade de alunos, porque muitos não se adaptaram ao ensino em EAD, e na volta as aulas híbridas muitos pais ficaram receosos em deixar seus filhos terem contato com outras pessoas.
- Em relação as correlações com disciplinas, achávamos que estariam agrupadas em linguagens, exatas e humanas.

- **Contexto da análise, por que esta questão é de interesse, quais informações básicas levou a hipótese, e por que isso é importante?**

Por que essa questão é de interesse? Ao tratarmos de educação após um período pandêmico é de grande valia, pois observaremos o quanto os alunos ainda sim se empenharam nos estudos independentes da situação que estava no mundo e se foram muito prejudicados ao longo dos anos por essa situação.

Informações básicas que levaram a hipótese? Porque fomos estudantes no período da pandemia e pudemos ver de perto o quanto isso nos prejudicou e principalmente afetou todo um método de ensino que as escolas já estavam preparando para os alunos, mas tiveram que rever tudo novamente aprimorando de acordo com o que estava ao alcance das escolas.

Por que isso é importante? É importante pois depois desse cenário pandêmico a educação passou a ser bastante desvalorizada, ao ponto de muitos jovens não darem devida importância para os estudos, mesmo sabendo que é algo crucial ter uma boa base de ensino.

- **Conjunto de dados, quais dados estão sendo usados para responder as perguntas, descreva os conjuntos de dados.**

Entramos em contato com uma escola, pedindo inicialmente o histórico de notas dos alunos no primeiro bimestre de 2024, porém ao apresentar esses dados para o professor ele observou que eram poucas informações. Então optamos por buscar as notas de primeiro bimestre de 2022 e 2023 também. Partindo disso criamos um kaban para nos conduzir ao longo do processo e ter uma noção do que ainda precisava ser feito e nos atentar aos prazos. Ainda sim eram poucos dados e preferimos então ir pessoalmente na escola em busca das notas nos anos de 2020 a 2023 de todos os bimestres e fazer uma breve reunião com o diretor escolar, para entendermos as necessidades da escola e as principais perguntas a serem respondidas com dados. Desenvolvemos um pequeno formulário com perguntas, feitas pelo grupo a serem respondidas pelo diretor. Nossa reunião aconteceu via google Meet e durou mais ou menos 20 minutos. O Arthur por morar mais próximo da escola se disponibilizou, ao entrar no sistema a secretaria da instituição conseguiu encontrar todos os dados que precisávamos, porém só era transformado em arquivo PDF. Mesmo assim, baixamos todo o conteúdo necessário e tivemos que transformar os dados de PDF para Excel, utilizamos um site gratuito na internet chamado “IlovePDF”, após isso adicionamos todas as notas em outra planilha do Excel com o intuito de fazer somente uma base com todas as disciplinas e para isso escolhemos as principais matérias escolares mantendo o padrão para os três anos. Dividimos os dados em três bases (1º, 2º e 3º ano) para cada um da equipe fazer uma parte e depois mesclarmos as bases para iniciar outros processos, também optamos por fazer novas variáveis como sexo e número do aluno para não ser necessário expor os nomes dos alunos. Para a variável sexo foi necessário os nomes de todos os alunos para diferenciarmos entre feminino e masculino, depois utilizamos o PROCX para associar esses dados a base e em relação a coluna (Nº do aluno) utilizamos somente a fórmula “ALEATORIO” do próprio Excel também para criação de vários números diferentes para cada estudante de forma mais rápida. Após esse processo precisamos transformar todas as disciplinas em uma só coluna as empilhando, pois cada disciplina era uma coluna, fizemos o empilhamento no Power bi, já que faríamos a maior parte da análise por esse programa.

Sobre as correlações consultamos um machine learning para nos ajudar a fazer um programa onde mostra por grupos de exatas, humanas e linguagens como acontece a correlação das notas dos alunos diante destes grupos e partindo disso, criamos médias de cada grupo que foi citado acima para analisarmos com gráficos essas relações.

2024

- **Quais métodos foram utilizados para analisar os dados?**
 - Inicialmente criamos gráficos no Excel para observar como as disciplinas se relacionavam com as outras variáveis criadas e ter mais ou menos uma previsão de como seria no Power BI.
 - No Power BI fizemos a formatação das variáveis disciplinas e notas de linhas para uma só coluna com o intuito de criarmos uma variável que mostrasse as disciplinas e notas em duas colunas diferentes.
 - Vamos trazer os dados analisados tanto no Power BI com dashboards quanto no Excel para as correlações que foram criadas.

- **Conclusões, quais são os principais resultados da análise? Quais são as respostas as questões? Compare com as predições, quais respostas não conseguiram ser obtidas e por quê?**
 - Os alunos ultrapassaram a média proposta pela escola.
 - Realmente na área das exatas os alunos não obtiveram notas elevadas, principalmente o sexo masculino.
 - As correlações de certa forma foram semelhantes com os grupos disciplinares dentro da instituição.
 - Nos anos citados devido ao decreto do governo, os alunos deveriam passar de ano independente de suas notas finais, com isso concluímos que não houve baixo nível nas notas, por justa causa.

- **Recomendações, a partir do aprendizado obtido, quais são as ações que o grupo recomendaria a serem adotadas, o que precisaria ainda ser estudado?**

Recomendações para os painéis:

- Criar conexão direta com os sistemas de notas para a extração automática das bases de dados
- Adaptar os painéis para atualização automática dos gráficos
- Publicar os painéis para os interessados (coordenadores, professores e pais de alunos) para que possam ser usados em fóruns de melhoria de ensino.
- Pedir feedback sobre os painéis para os interessados para que melhorias sejam implementadas neles.
- Expandir os painéis para novas instituições de ensino.

Recomendações para o ensino:

- Currículo Diversificado, trazendo disciplinas que não ensinem somente conteúdos que já estão habituados.
- Trabalhar com novas tecnologias em sala de aula para os alunos se engajarem e serem mais interativos sobre as disciplinas ensinadas.
- Coletar feedbacks constantes, de funcionários, pais e alunos para identificar áreas de melhoria e implementar mudanças necessárias.
- Monitores (os próprios alunos que tiverem afinidade com as disciplinas) dentro dos grupos (Exatas, linguagens, Humanas) nas salas.
- Ter um maior aprofundamento em relação ao emocional dos alunos.

- **Ética e privacidade, reconheça e trate quaisquer questões que envolvam ética e privacidade relacionada ao projeto.**

Fizemos um termo de confidencialidade e ética para tratarmos dos dados com os nomes dos alunos, para deixar claro que não iríamos expor nenhum nome e isso foi mencionado na reunião com o diretor escolar.

Termo:

Este Termo de Fidelização tem como objetivo formalizar a relação de fidelidade entre [Zélia de Matos Brito] e o grupo, oferecendo uma análise de melhoria de gestão.

1. Definições

Para os fins deste Termo, aplicam-se as seguintes definições:

Cliente Fidelizado: Pessoa física ou jurídica que adere ao programa de fidelidade da [Zélia de Matos Brito].

Benefícios: uma melhor visão de dados que serão trabalhados dentro da gestão, incluindo, mas não se limitando a dúvidas sobre algum gráfico.

1.1. Adesão ao Programa de Fidelidade

A adesão ao programa de fidelidade da [Zélia de Matos Brito] é voluntária e gratuita.

2. Responsabilidades do Cliente Fidelizado

1.2. Manter seus dados em sigilo conforme foi utilizado pelo grupo com permissão da Empresa [Zélia de Matos Brito]

2. Disposições Gerais

2.1. Este Termo de Fidelização entra em vigor na data de publicação do projeto, que só será mencionado o nome da escola com autorização do cliente.

[Pacoti-CE], [28/06/2024]

[Zélia de Matos Brito]

[Diógenes de Sousa Luz – Diretor Escolar]

2024

- **Discussão, discuta o potencial e impacto cívico do projeto. Discuta todos os problemas encontrados e enumere as limitações do seu projeto, assim como novas questões que poderiam ser abordadas para complementar a análise do tema.**
 - Nosso projeto tem o potencial e o impacto cívico no âmbito de trazer melhorias para as escolas de ensino médio, ajudando a gestão escolar a ter novos métodos de análise em relação as notas dos alunos e a criação de novos conceitos estratégicos para o melhor desenvolvimento, não somente da instituição que obtivemos os dados para o projeto, mas sim para todas as instituições que necessitem de uma melhora positiva em sua educação.
 - Encontramos problemas no formato dos dados que foram enviados para nós, por conta disso tivemos que fazer a transformação de todos os dados e somente depois fazer a criação dos gráficos
 - Na criação de novas variáveis, tratando do sexo feminino e masculino não conseguimos esses dados no sistema da escola pois não tinha como enviar como arquivo, então tivemos que fazer a divisão de sexo utilizando somente os nomes dos alunos e depois enviando os dados para a base principal com o PROCX.

Este Termo de Fidelização tem como objetivo formalizar a relação de fidelidade entre [Zélia de Matos Brito] e o grupo, oferecendo uma análise de melhoria de gestão.

1. Definições

Para os fins deste Termo, aplicam-se as seguintes definições:

Cliente Fidelizado: Pessoa física ou jurídica que adere ao programa de fidelidade da [Zélia de Matos Brito].

Benefícios: uma melhor visão de dados que serão trabalhados dentro da gestão, incluindo, mas não se limitando a dúvidas sobre algum gráfico.

1.1. Adesão ao Programa de Fidelidade

1.2. A adesão ao programa de fidelidade da [Zélia de Matos Brito] é voluntária e gratuita.

2. Responsabilidades do Cliente Fidelizado

2.1. Manter seus dados em sigilo conforme foi utilizado pelo grupo com permissão da Empresa [Zélia de Matos Brito]

2.2. Disposições Gerais

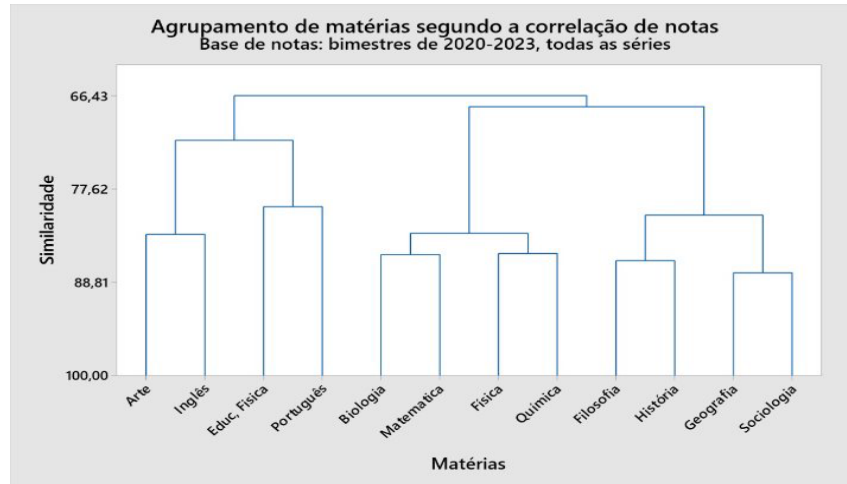
3. Este Termo de Fidelização entra em vigor na data de publicação do projeto, que só será mencionado o nome da escola com autorização do cliente.

[Pacoti-CE], [28/06/2024]

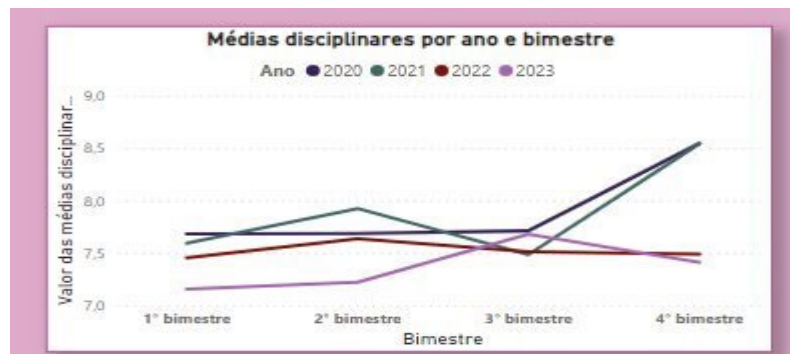
[Zélia de Matos Brito]

[Diógenes de Sousa Luz – Diretor Escolar]

2024



- Pedimos o auxílio de um especialista em métodos estatísticos para agrupar as disciplinas conforme as correlações a partir das notas e o resultado foi compatível com os grupos disciplinares que as escolas se baseiam (Exatas&Biologia, Linguagens e Humanas)
- Partindo dessas análises desenvolvemos médias dos grupos disciplinares para criações de gráficos que mostraremos a seguir.



- Nesse gráfico podemos analisar que no ano de 2020 e 2021 as médias disciplinares estão maiores que nos outros anos.
- Porém podemos observar que o grupo disciplinar de exatas & bio, está bem abaixo de todos os outros, deixando claro que o período pandêmico afetou no desenvolvimento dos alunos.