

**Universidade Estácio de Sá
Nova América**

BASE DE DADOS NA VENDA DE BEBIDAS

Alexandre Faria da Fonseca - 202304375372
Gabriel Fellipe Venâncio de Oliveira – 202303535252
Jonas Oliveira Silva Campos - 202304322741
Nicolas da Cunha Pinto Rabelo - 202304269611
Riguel Nascimento Marins da Costa - 202304270473

Professor: ANDRÉ LUIZ CARMO EPPINGHAUS

2025
Rio de Janeiro RJ

Sumário

1. DIAGNÓSTICO E TEORIZAÇÃO	3
1.1. Identificação das partes interessadas e parceiros	3
1.2. Problemática e/ou problemas identificados	3
1.3. Justificativa	3
1.4. Objetivos/resultados/efeitos a serem alcançados (em relação ao problema identificado e sob a perspectiva dos públicos envolvidos)	4
1.5. Referencial teórico (subsídio teórico para propositura de ações da extensão)	4
2. PLANEJAMENTO E DESENVOLVIMENTO DO PROJETO	4
2.1. Plano de trabalho (usando ferramenta acordada com o docente)	4
2.2. Descrição da forma de envolvimento do público participante na formulação do projeto, seu desenvolvimento e avaliação, bem como as estratégias pelo grupo para mobilizá-los.....	5
2.3. Grupo de trabalho (descrição da responsabilidade de cada membro).....	5
2.4. Metas, critérios ou indicadores de avaliação do projeto	5
2.5. Recursos previstos	5
2.6. Detalhamento técnico do projeto	6
3. ENCERRAMENTO DO PROJETO.....	10
3.1. Relato Coletivo:	10
3.1.1. Avaliação de reação da parte interessada	10
3.2. Relato de Experiência Individual.....	10
3.2.1. CONTEXTUALIZAÇÃO	10
3.2.2. METODOLOGIA.....	11
3.2.3. RESULTADOS E DISCUSSÃO:.....	12
3.2.4. REFLEXÃO APROFUNDADA	13
3.2.5. CONSIDERAÇÕES FINAIS	13
4. ANEXOS	14

1. DIAGNÓSTICO E TEORIZAÇÃO

1.1. Identificação das partes interessadas e parceiros

Alexandre Faria da Fonseca - Aluno, 23 anos, classe média baixa, cursando Ciência da computação.

Gabriel Fellipe Venâncio de Oliveira – Aluno, 24 anos, classe média baixa, cursando Análise e desenvolvimento de sistemas.

Jonas Oliveira Silva Campos - Aluno, 23 anos, classe média baixa, cursando Ciência da Computação.

Nicolas da Cunha Pinto Rabelo - Aluno, 21 anos, classe média baixa, cursando Ciência da Computação.

Riguel Nascimento Marins da Costa - Aluno, 26 anos, classe média baixa, cursando Sistemas de informação.

André Luiz Carmo Eppinghaus – Docente, orientador do projeto.

1.2. Problemática e/ou problemas identificados

Observa-se uma baixa agilidade na visualização dos dados relacionados a distribuição de bebidas. Essa lentidão compromete a tomada de decisões em tempo hábil, impactando diretamente a eficiência das operações logísticas e a satisfação dos clientes. A gerente de logística e a equipe de análise de dados enfrentam dificuldades ao precisarem dedicar um tempo excessivo ao processamento e interpretação manual dos dados, reduzindo a produtividade e eleva o risco de erros operacionais.

A identificação desse problema ocorreu por meio de relatos da própria gerente e de colaboradores envolvidos nas operações, além da análise dos tempos de resposta, na visualização de dados, pesquisas internas, estudo de dados históricos e observações diretas do processo logístico.

1.3. Justificativa

O projeto melhora a eficiência operacional e a capacidade de resposta da empresa, impactando positivamente a satisfação do cliente e a competitividade no mercado. Academicamente, contribui para o avanço do conhecimento em *Big Data* em que os participantes desenvolverão habilidades em análise de dados, uso de ferramentas de *Big Data* e gestão de projetos. Economicamente, melhora a eficiência e reduz custos. Socialmente, pode melhorar a qualidade do serviço prestado aos consumidores. Podemos citar benefícios

como maior agilidade na visualização de dados, melhor tomada de decisões, aumento da eficiência operacional e satisfação do cliente, como resultado, a gerente de logística terá o acesso mais rápido e preciso aos dados, permitindo decisões mais informadas e rápidas.

- 1.4. Objetivos/resultados/efeitos a serem alcançados (em relação ao problema identificado e sob a perspectiva dos públicos envolvidos)
 1. Aumentar a agilidade e a precisão na visualização dos dados dos serviços de entrega, proporcionando uma resposta mais rápida às demandas logísticas e facilitando a tomada de decisões estratégicas.
 2. Desenvolver e implementar uma solução baseada em *Big Data* para centralizar e automatizar a visualização dos dados operacionais, além de capacitar a equipe envolvida para o uso eficiente da nova ferramenta.
 3. Reduzir significativamente o tempo de resposta para acesso e interpretação dos dados, aumentando a confiabilidade das análises e contribuindo para uma gestão logística mais eficiente e proativa.
- 1.5. Referencial teórico (subsídio teórico para propositura de ações da extensão)

GEEKSFORGEES. *Data Science with Python Tutorial*. 2025. Disponível em: <https://www.geeksforgeeks.org/data-science-with-python-tutorial/> (Acesso em: 28 de março de 2025).

REAL PYTHON. *Python Data Cleaning With pandas and NumPy*. 2025. Disponível em: <https://realpython.com/python-data-cleaning-numpy-pandas/> (Acesso em: 1 de abril de 2025).

W3SCHOOLS. *Matplotlib Tutorial*. 2025. Disponível em: https://www.w3schools.com/python/matplotlib_intro.asp (Acesso em: 7 de abril de 2025).

2. PLANEJAMENTO E DESENVOLVIMENTO DO PROJETO

2.1. Plano de trabalho (usando ferramenta acordada com o docente)

O que fizemos	Data	participação
Coleta dos dados e contato com as partes interessadas	18/02/2025	Alexandre
Limpeza e tratamento dos dados	28/03/2025	Gabriel e Riguel
Análise dos dados	07/04/2025	Jonas e Nicolas
Construção de Relatórios e <i>Dashboards</i>	23/04/2025	Jonas e Nicolas

Documentação	06/05/2025	Gabriel
Revisão do projeto	11/05/2025	Riguel
Correção e finalização	25/05/2025	Todos

- 2.2. Descrição da forma de envolvimento do público participante na formulação do projeto, seu desenvolvimento e avaliação, bem como as estratégias pelo grupo para mobilizá-los.

O problema foi identificado a partir de conversas e reuniões com a gerente de logística e analista de dados da empresa. Relatos sobre a dificuldade na visualização e análise dos dados de entrega foram fundamentais para definir o escopo do projeto. As etapas do projeto foram divididas em coleta de dados, limpeza e transformação dos dados, desenvolvimento da solução de visualização, e implementação da solução. O cronograma de execução das atividades foi definido com base na complexidade de cada etapa e na disponibilidade dos recursos. Os desafios enfrentados na implementação consistiram em integração dos dados de diferentes fontes. Durante o desenvolvimento do projeto, ajustes no cronograma e na alocação de recursos foram realizados para gerenciar imprevistos. Para assegurar a continuidade do projeto, adaptamos as ferramentas utilizadas e ajustamos o treinamento da equipe.

- 2.3. Grupo de trabalho (descrição da responsabilidade de cada membro)

Alexandre Faria – Coleta de dados e comunicação com a parte interessada.
Gabriel Felipe – Documentação e limpeza dos dados.
Jonas Campos – Análise dos dados, visualização dos dados e elaboração dos relatórios.
Nicolas Rabelo – Análise dos dados, visualização dos dados e elaboração dos relatórios.
Riguel Nascimento - Tratamento dos dados e revisão de conteúdo.

- 2.4. Metas, critérios ou indicadores de avaliação do projeto

Metas: Implementar uma ferramenta de visualização de dados baseada em *Big Data*, reduzir o tempo médio de acesso e interpretação dos dados pelos gestores e capacitar os usuários na utilização da nova ferramenta.

Crítérios: Efetividade da solução implementada na visualização dos dados, grau de satisfação dos usuários com a ferramenta e qualidade e clareza dos *dashboards* desenvolvidos.

Indicadores: Tempo médio de visualização de dados antes e depois da solução, *feedback* qualitativo dos gestores e redução no tempo de resposta para tomadas de decisão.

- 2.5. Recursos previstos

Recursos materiais: Computadores pessoais dos alunos e da Universidade e a plataforma *Google Colab*.

Recursos institucionais: Ambiente físico e Virtual da Universidade e base de dados da empresa.

Recursos Humanos: Alunos participantes do projeto, docente e a funcionária da empresa.

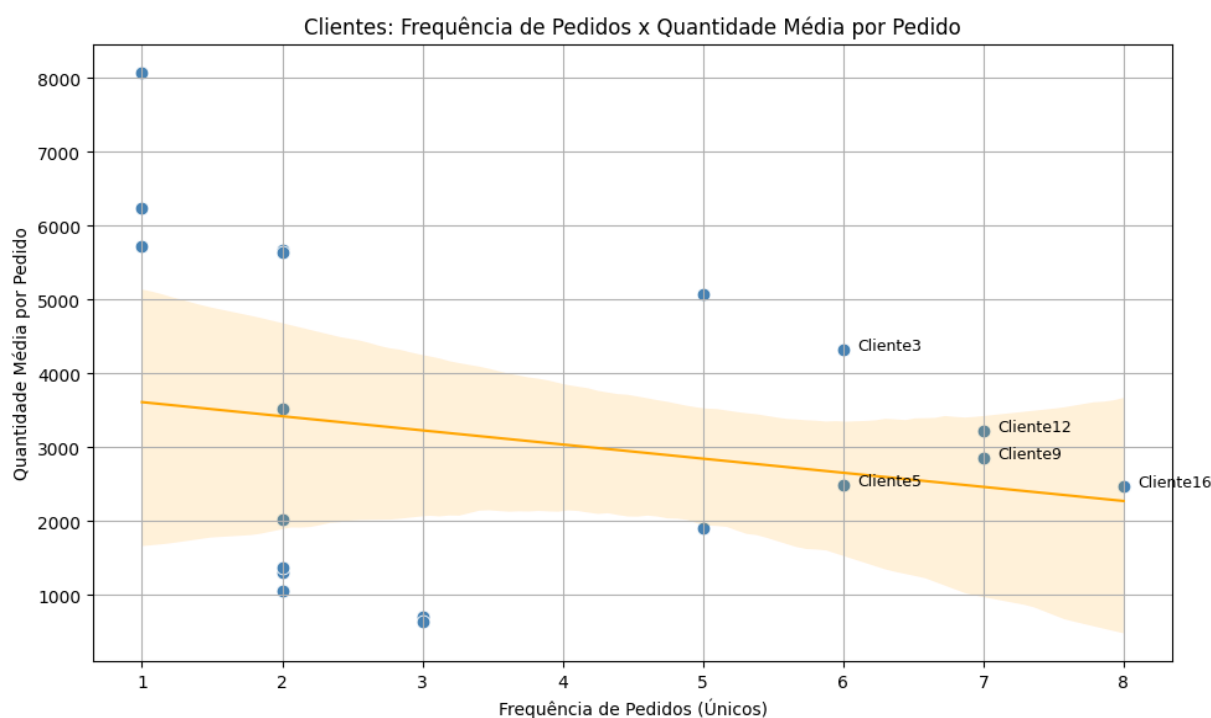
2.6. Detalhamento técnico do projeto

O projeto teve como foco o desenvolvimento de uma solução analítica voltada à melhoria da visualização e interpretação de dados logísticos da empresa. A solução está alinhada aos objetivos socioeducacionais da disciplina e foi construída com participação ativa da parte interessada.

Durante o desenvolvimento da solução, foram construídos gráficos com base nas vendas realizadas entre os dias 01/01/2025 e 08/04/2025 e, após a análise, foram identificados pontos em que podem ser aplicados melhorias e, também, planos de ação para o aumento da eficiência nas vendas dos produtos e melhor relação com os clientes.

Utilizamos a linguagem de programação *Python* para realizar a tratativa e limpeza dos dados, refinando os dados para análise com a remoção de dados nulos e duplicados, além de definirmos a tipagem dos dados. Os gráficos foram construídos utilizando as bibliotecas *Pandas*, *Matplotlib* e *Seaborn*. As análises foram estruturadas em *scripts* modulares, possibilitando reaproveitamento de código e reprodutibilidade das análises. Para o ambiente de visualização, foi usado a plataforma *Google Colab*, com comentários orientando cada etapa. Isso viabilizou que a parte interessada pudesse executar e entender os relatórios de forma autônoma.

Relatório 1: Perfil de Compra dos Clientes (Frequência vs. Volume Médio por Pedido)



A análise dos dados de compra dos clientes revelou padrões importantes de comportamento. Ao cruzar a frequência de pedidos com a quantidade média por pedido, foi possível identificar dois grupos predominantes: clientes que compram com alta frequência e baixo volume por pedido, e clientes com baixa frequência, mas alto volume médio.

Embora a intuição possa sugerir que os clientes mais frequentes são os mais valiosos, o gráfico indica que isso nem sempre se confirma. A linha de tendência demonstra uma correlação negativa leve entre os dois fatores, o que significa que, à medida que a frequência aumenta, o volume médio por pedido tende a diminuir.

Entre todos os clientes, cinco se destacam por apresentar alto volume e alta frequência, combinando os dois fatores mais desejáveis para o negócio. Esses são os principais clientes (Top 5), que aparecem em posições superiores no gráfico e representam um potencial estratégico para ações de fidelização e crescimento conjunto.

Essa segmentação permite desenhar estratégias específicas:

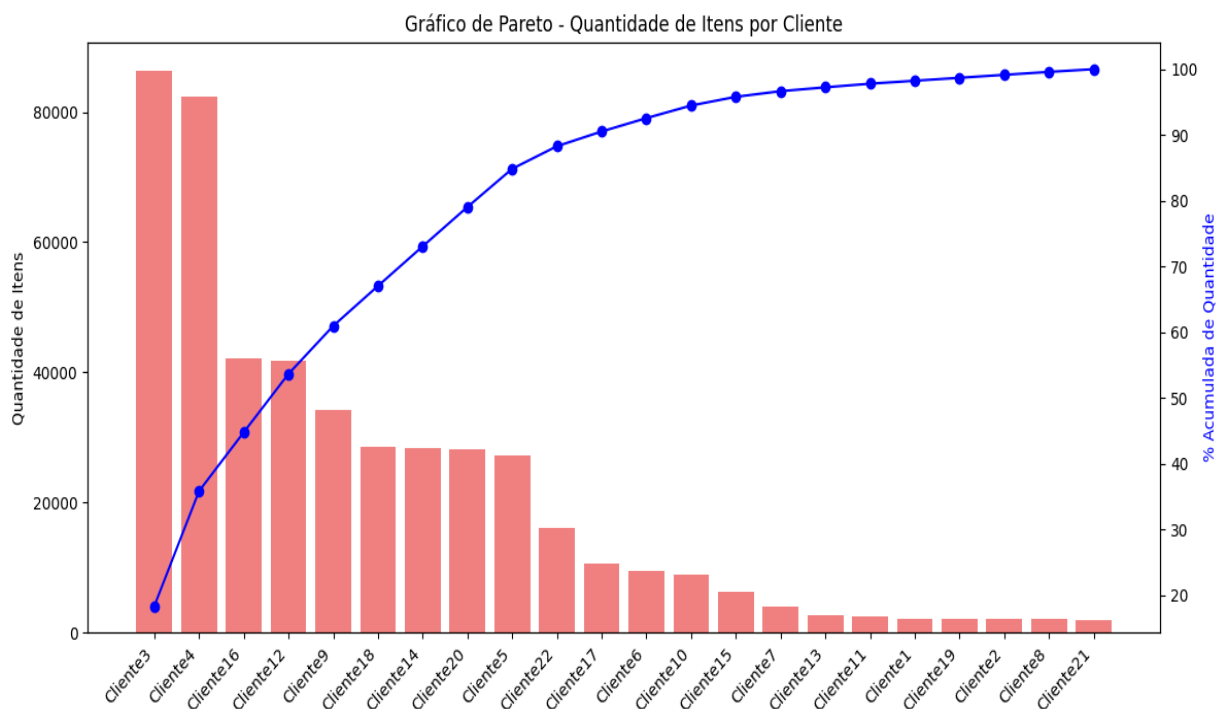
Clientes com baixa frequência e alto volume podem ser estimulados com ações de relacionamento ou ofertas personalizadas, para aumentar sua recorrência.

Já os clientes com alta frequência e pedidos pequenos devem ser acompanhados para otimizar os custos operacionais e incentivados a consolidar pedidos maiores, reduzindo o impacto logístico.

Os principais clientes (Top 5) devem ser priorizados com programas de retenção, atendimento consultivo e políticas comerciais diferenciadas.

Esses dados orientam uma abordagem segmentada e estratégica, capaz de aumentar a rentabilidade e a eficiência comercial da empresa.

Relatório 2: Análise de Quantidade de itens por Cliente via Gráfico de Pareto



A análise de quantidade de itens comprados por cliente revelou uma alta concentração de vendas em poucos clientes. O gráfico de Pareto demonstra que aproximadamente 20% dos clientes são responsáveis por cerca de 80% da quantidade total de itens vendidos, confirmando o princípio clássico de Pareto (80/20). Isso significa que uma minoria da base de clientes concentra a maior parte do volume movimentado.

A curva acumulada mostra uma ascensão rápida no início, indicando que poucos clientes compram volumes elevados. A partir de certo ponto, a curva se estabiliza, evidenciando que a maioria dos clientes contribui com volumes baixos e dispersos. Esses dados evidenciam uma forte dependência de um pequeno grupo de clientes estratégicos, os quais são fundamentais para manter a operação com estabilidade e escala.

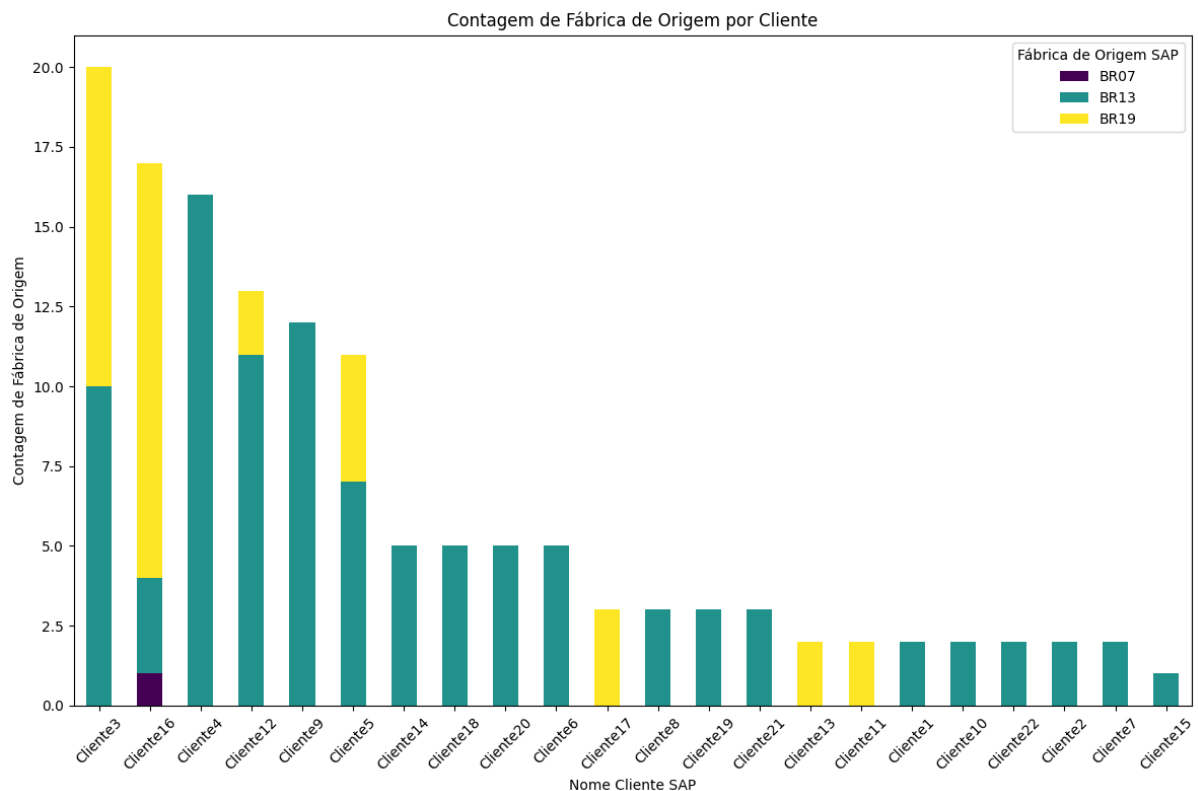
Esse cenário reforça a importância de manter e fidelizar os principais clientes, que aparecem no topo da curva. Eles impactam diretamente o desempenho da empresa em termos de receita, logística e planejamento de estoque. Por outro lado, há uma parcela significativa de clientes com compras reduzidas, que, embora menos relevantes individualmente, podem representar potencial de crescimento se estimulados com ações comerciais específicas.

Diante disso, surgem duas frentes estratégicas:

Fidelização e relacionamento com os grandes compradores, garantindo sua permanência e explorando oportunidades de *cross-sell* e *upsell*.

Iniciativas para ativar e expandir a base menos ativa, por meio de políticas de incentivo, promoções ou bonificações por volume, para reduzir a dependência dos principais clientes e diluir o risco comercial.

Relatório 3: Análise de Origem de Fábrica por Cliente



A análise da origem de fábrica por cliente revelou diferentes perfis logísticos entre os compradores, com variações significativas no número de unidades de produção que abastecem cada um. O gráfico mostra claramente que alguns clientes são atendidos por várias fábricas distintas, enquanto outros têm seus pedidos concentrados em apenas uma única unidade.

Essa informação é essencial para entender a complexidade logística envolvida no atendimento de cada cliente. Aqueles que recebem produtos de diferentes fábricas exigem maior coordenação na cadeia de suprimentos, por envolverem múltiplos pontos de expedição, transporte e sincronização. Esse tipo de atendimento, embora mais desafiador, também pode indicar um nível mais alto de importância estratégica, já que esses clientes demandam volumes variados e diversidade de produtos.

Por outro lado, clientes com concentração de origem em uma única fábrica apresentam uma logística mais simples e direta. No entanto, essa centralização aumenta o risco operacional, pois qualquer interrupção nessa unidade pode comprometer todo o atendimento.

A segmentação apresentada no gráfico permite mapear com clareza:

Cientes de alta complexidade logística, que demandam múltiplas fábricas.

Cientes de baixa complexidade, com pedidos centralizados em uma única origem.

Esse entendimento orienta decisões estratégicas para mitigar riscos e otimizar a operação. É essencial garantir resiliência na logística multisite, desenvolvendo alternativas produtivas e planos de contingência para clientes com dependência de múltiplas fábricas. Ao mesmo tempo, é necessário monitorar os clientes mais centralizados para que não se tornem pontos de vulnerabilidade em caso de indisponibilidade de produção.

3. ENCERRAMENTO DO PROJETO

3.1. Relato Coletivo:

O projeto de extensão “Base de Dados na Venda de Bebidas” teve como principal objetivo desenvolver uma solução de visualização de dados capaz de aumentar a agilidade e a eficiência das decisões logísticas da empresa parceira. Ao longo do trabalho, o grupo enfrentou desafios técnicos relacionados à integração e ao tratamento dos dados, mas conseguiu superá-los com planejamento, divisão de tarefas e aplicação de conhecimentos teóricos em *Python* e *Big Data*.

A participação ativa da parte interessada desde o diagnóstico até a avaliação final foi fundamental para alinhar expectativas e validar as soluções propostas. A equipe utilizou ferramentas como *Google Colab* e bibliotecas *Python* (*Pandas*, *NumPy*, *Matplotlib* e *Seaborn*) para transformar dados brutos em informações estratégicas, visualmente compreensíveis por meio de *dashboards*.

A interdisciplinaridade entre programação, análise de dados, logística e comunicação foi essencial para o sucesso da proposta. Além disso, todos os alunos desenvolveram habilidades técnicas e interpessoais, contribuindo com um aprendizado significativo tanto no aspecto acadêmico quanto profissional.

3.1.1. Avaliação de reação da parte interessada

A parte interessada demonstrou grande satisfação com os resultados apresentados. Em entrevista realizada ao final do projeto, a gerente de logística afirmou que a nova ferramenta de visualização reduziu o tempo de acesso e interpretação dos dados, facilitando significativamente a tomada de decisões. A clareza dos *dashboards* e os *insights* extraídos dos relatórios foram considerados úteis e aplicáveis imediatamente nas operações da empresa.

A gerente destacou a relevância do projeto e elogiou o comprometimento e profissionalismo do grupo de alunos. A interação constante e a escuta ativa durante o desenvolvimento foram citadas como pontos positivos. Com isso, consideramos que os objetivos socioeducacionais foram plenamente atingidos.

3.2. Relato de Experiência Individual

3.2.1. CONTEXTUALIZAÇÃO

Alexandre Faria da Fonseca - Participei ativamente da coleta de dados e da comunicação com a parte interessada. Minha atuação esteve diretamente relacionada ao entendimento do problema enfrentado pela empresa no contexto logístico. O projeto foi desenvolvido no

âmbito de um componente curricular de extensão, com foco na análise de dados reais fornecidos por uma empresa parceira. Desde o início, mantive contato constante com a empresa, o que garantiu um alinhamento entre as expectativas da empresa e as soluções propostas pelo grupo.

Gabriel Fellipe Venâncio de Oliveira - Minha principal atuação no projeto foi relacionada à limpeza e preparação da base de dados, além da organização da documentação técnica. O projeto surgiu como parte de uma disciplina extensionista com foco na aplicação prática de *Big Data* e análise de dados em *Python*. Desde o início, percebi a importância de uma base de dados limpa e bem estruturada para garantir análises confiáveis.

Jonas Oliveira Silva Campos - Fui responsável pela criação de *dashboards* e relatórios visuais durante o desenvolvimento do projeto. Trabalhei com análise de padrões de compra, origem dos produtos e segmentação de clientes, sempre visando tornar os dados mais compreensíveis para os interessados. A proposta do projeto surgiu a partir de dados logísticos reais, e tivemos liberdade para aplicar técnicas analíticas relevantes.

Nicolas da Cunha Pinto Rabelo - Trabalhei em conjunto com Jonas na análise dos dados e construção dos *dashboards*. A partir dos dados fornecidos pela empresa, nosso foco foi extrair informações estratégicas que apoiassem a tomada de decisão na área de logística e vendas.

Riguel Nascimento Marins da Costa - No projeto, assumi a responsabilidade pela normalização e estruturação dos dados, além da revisão final dos produtos gerados pelo grupo. Essa experiência me permitiu aplicar conhecimentos técnicos adquiridos em aula em um projeto com dados reais e objetivos práticos.

3.2.2. METODOLOGIA

Alexandre Faria da Fonseca - A experiência foi vivenciada remotamente, com reuniões periódicas entre os membros do grupo e os representantes da empresa. Os sujeitos envolvidos foram os alunos participantes do projeto, o docente orientador e a empresa parceira. Durante o período de desenvolvimento, uma das minhas funções principais foram o levantamento de requisitos e a validação das informações coletadas. Isso envolveu entender os fluxos logísticos, questionar pontos ambíguos e garantir que a base de dados estivesse condizente com a realidade da operação. Essa etapa foi essencial para nortear as demais fases do projeto.

Gabriel Fellipe Venâncio de Oliveira - O projeto foi realizado de forma colaborativa, com encontros online para definição de tarefas e troca de informações. Com base nos dados fornecidos pela empresa parceira, trabalhei diretamente com bibliotecas como *Pandas* e *NumPy* para tratar valores nulos, padronizar colunas e identificar inconsistências.

Jonas Oliveira Silva Campos - Utilizamos bibliotecas como *Matplotlib* e *Seaborn* para a criação de gráficos e *dashboards*. Durante o processo, nos dividimos em duplas e grupos para otimizar a construção das análises. Junto ao colega Nicolas, desenvolvemos representações visuais que auxiliassem na identificação de gargalos e oportunidades no processo logístico da empresa. As reuniões com o orientador nos ajudavam a validar as escolhas técnicas.

Nicolas da Cunha Pinto Rabelo - Com base nas orientações recebidas, mergulhamos nas técnicas de visualização usando bibliotecas como *Seaborn* e *Matplotlib*. O processo envolveu segmentação de dados, definição de métricas e construção de gráficos relevantes. As decisões visuais foram sempre embasadas nas necessidades apresentadas pela parte interessada.

Riguel Nascimento Marins da Costa - Trabalhei com a padronização de tipagens, eliminação de ruídos e preparação dos dados para a análise estatística. Após essa etapa, revisei os *dashboards*, gráficos e relatórios gerados pelos colegas, garantindo que todas as entregas estivessem coerentes técnica e visualmente. A atuação em equipe foi essencial para o sucesso do processo.

3.2.3. RESULTADOS E DISCUSSÃO:

Alexandre Faria da Fonseca - A experiência superou minhas expectativas, especialmente pela possibilidade de trabalhar com dados reais e interagir com profissionais da área. A comunicação com a parte interessada exigiu postura profissional e clareza técnica. Aprendi como o bom levantamento de requisitos pode evitar retrabalho e influenciar positivamente nos resultados do projeto. Um dos maiores desafios foi garantir a fidelidade da base de dados ao processo real. Como resultado, a base ajustada serviu como base sólida para as análises desenvolvidas posteriormente.

Gabriel Fellipe Venâncio de Oliveira - A experiência prática foi extremamente válida, pois me permitiu aplicar conceitos aprendidos em aula em um contexto real. A limpeza dos dados apresentou desafios, como valores duplicados e formatações diversas, mas contribuiu significativamente para o sucesso das análises. Aprendi a importância da organização e da clareza na documentação, o que é essencial para qualquer projeto técnico. Enfrentei dificuldades com algumas inconsistências nos dados, mas, com o apoio da equipe, conseguimos solucioná-las.

Jonas Oliveira Silva Campos - O uso de visualizações facilitou muito a compreensão das informações por parte da empresa. O gráfico de Pareto, por exemplo, ajudou a identificar quais clientes representavam maior volume de vendas. A experiência foi bastante gratificante, pois consegui colocar em prática minha habilidade em *Python* e entender o valor estratégico de uma boa apresentação de dados. Enfrentei algumas dificuldades com a escolha dos tipos ideais de gráficos, mas o aprendizado foi constante.

Nicolas da Cunha Pinto Rabelo - Durante o desenvolvimento, percebi como a forma de apresentar os dados influencia diretamente na sua interpretação. A visualização dos perfis de clientes, por exemplo, revelou padrões que não seriam perceptíveis apenas com números. Foi uma experiência enriquecedora, especialmente por entender a importância do contexto na análise. A maior dificuldade foi alinhar os gráficos ao perfil do público que iria consumi-los.

Riguel Nascimento Marins da Costa - Foi muito satisfatório ver como o trabalho de preparação e revisão contribuiu para a qualidade do resultado. Enfrentei desafios relacionados à heterogeneidade dos dados, mas com estudo e colaboração conseguimos superá-los. A experiência me mostrou como a revisão técnica é essencial para a entrega de projetos confiáveis e profissionais.

3.2.4. REFLEXÃO APROFUNDADA

Alexandre Faria da Fonseca - A vivência prática proporcionou uma conexão clara com os conceitos teóricos abordados nas aulas. Pude perceber que, além das ferramentas, o sucesso de um projeto depende da comunicação eficaz e do entendimento do contexto do negócio. A etapa de levantamento de requisitos, muitas vezes negligenciada, se mostrou fundamental.

Gabriel Felipe Venâncio de Oliveira - A prática evidenciou que, sem uma base de dados consistente, qualquer análise pode ser comprometida. A teoria apresentada nas aulas se confirmou na prática, mostrando que a fase de preparação dos dados é, muitas vezes, a mais trabalhosa, porém a mais crítica para garantir a qualidade do projeto.

Jonas Oliveira Silva Campos - A prática de visualização de dados se demonstrou muito mais complexa e rica do que aparenta na teoria. Entender o público-alvo, o contexto da informação e a clareza visual foram aspectos que só se consolidaram com a prática. A experiência evidenciou a importância da estética e da funcionalidade em *dashboards*.

Nicolas da Cunha Pinto Rabelo - Essa experiência prática reforçou os conceitos aprendidos sobre *storytelling* com dados e análise exploratória. A teoria sobre visualização se confirmou como um componente essencial para transformar dados brutos em *insights* estratégicos.

Riguel Nascimento Marins da Costa - O aprendizado teórico sobre tipos de dados e normalização se mostrou essencial na prática. A vivência reforçou a importância de revisar os dados de forma crítica antes de iniciar as análises. Também ficou claro o quanto o trabalho colaborativo e a divisão de responsabilidades contribuem para a qualidade do projeto.

3.2.5. CONSIDERAÇÕES FINAIS

Alexandre Faria da Fonseca - A experiência demonstrou o potencial transformador da ciência de dados em ambientes corporativos. Como perspectiva futura, seria interessante trabalhar com *dashboards* interativos que pudessem ser utilizados diretamente pela empresa, ou até

mesmo implantar soluções automatizadas de análise com alertas e *KPIs* em tempo real. Tecnologias como *Power BI* ou integração com bancos de dados em nuvem também poderiam agregar valor ao projeto.

Gabriel Fellipe Venâncio de Oliveira - Como proposta de melhoria, poderia ser interessante explorar técnicas de validação automática e *scripts* de limpeza que possam ser reutilizados em outros projetos. A experiência também abre espaço para o uso de pipelines de dados mais robustos com ferramentas como *Apache Spark* ou *Airflow* em trabalhos futuros.

Jonas Oliveira Silva Campos - No futuro, o projeto poderia se beneficiar da criação de *dashboards* interativos com ferramentas como *Plotly*, *Dash* ou *Power BI*. Além disso, a inclusão de filtros dinâmicos para segmentações por período, cliente ou SKU aumentaria o valor estratégico das análises.

Nicolas da Cunha Pinto Rabelo - Como proposta futura, seria interessante incluir visualizações com mapas de calor ou georreferenciamento, especialmente considerando que parte dos dados incluíam códigos geográficos. Isso abriria novas possibilidades para análises espaciais.

Riguel Nascimento Marins da Costa - Como evolução futura, o uso de pipelines automatizados e o versionamento dos dados poderiam otimizar o processo. Além disso, explorar soluções de *Big Data* mais robustas, como *Apache Spark*, traria ganho de escala e eficiência para projetos com grandes volumes de dados.

4. ANEXOS

Base de dados:
https://docs.google.com/spreadsheets/d/1OCAo_DuFLgV1HKkYBYxtczE8gZEOa9/edit?usp=sharing&ouid=101586588073646713128&rtpof=true&sd=true

Colab:
<https://colab.research.google.com/drive/1DtsptBhi2IMafhFySClw5s6Mvm2iuoJY?usp=sharing>

Apresentação:
<https://drive.google.com/file/d/1erzHcM59ebklzINng18rh52EQumL5zv/view?usp=sharing>