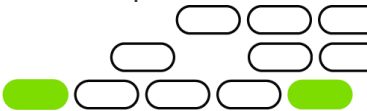


Sumário

1. CANVAS do Projeto Aplicado	4
1.1 Desafio	5
1.1.1 Análise de Contexto	5
1.1.2 Personas	6
1.1.3 Benefícios e Justificativas	7
1.1.4 Hipóteses	8
1.2 Solução	9
1.2.1 Objetivo SMART	9
1.2.2 Premissas e Restrições	11
1.2.3 Backlog de Produto	13
2. Área de Experimentação	15
2.1 Sprint 1	16
2.1.1 Solução	16
Evidência do planejamento:	16
Evidência da execução de cada requisito:	16
Evidência dos resultados:	16
2.1.2 Lições Aprendidas	16
2.2 Sprint 2	17
2.2.1 Solução	17
Evidência do planejamento:	17
Evidência da execução de cada requisito:	17
Evidência dos resultados:	17
2.2.2 Lições Aprendidas	17
2.3 Sprint 3	18
2.3.1 Solução	18
Evidência do planejamento:	18
Evidência da execução de cada requisito:	18
Evidência dos resultados:	18
2.3.2 Lições Aprendidas	18



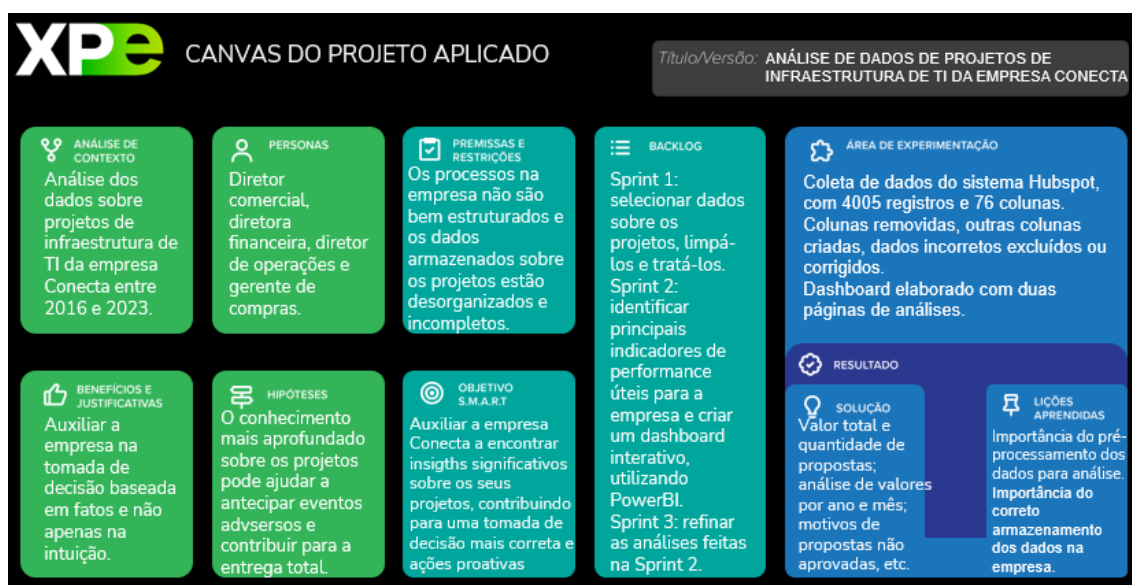
3. Considerações Finais	30
3.1 Resultados	19
3.2 Contribuições	19
3.3 Próximos passos	32



1. CANVAS do Projeto Aplicado

Este projeto aplicado compreende as etapas definidas pelo curso de pós-graduação (MBA) em Ciência de Dados da XP Educação, conforme Figura 1.

Figura 1 - CANVAS do Projeto Aplicado. Fonte: adaptado de IGTI (2023)



1.1 Desafio

1.1.1 Análise de Contexto

A ciência de dados vem sendo aplicada nos mais diversos contextos e tipos de organizações nos últimos tempos, como instituições públicas e privadas, de pequeno e grande porte, em áreas de marketing, saúde, vendas, economia, transportes, entre outras. Contudo, ainda existem muitas empresas que não fazem uso de nenhuma técnica de análise em seus dados para tomada de decisão ou o fazem de maneira incompleta ou insatisfatória.

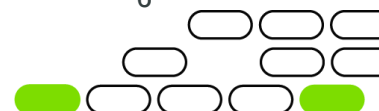
O objeto de estudo deste trabalho enquadra-se neste contexto. Trata-se de uma empresa familiar, de pequeno porte, da área de telecomunicações, aqui denominada de “Conecta” que carece de um intenso trabalho de organização processual, armazenamento de dados e mineração de informações, para conhecer melhor a própria empresa e conseguir extrair insights úteis para tomada de decisão e melhoria na gestão de marketing, pessoas e outros recursos.

A Conecta fornece produtos e serviços de infraestrutura de TI, como redes com cabeamento estruturado, data center, sistemas de monitoramento (CFTV), entre outros. A maior parte das decisões tomadas pelos gestores baseia-se em conhecimento tácito e em alguns dados disponíveis nos sistemas da empresa. Faltam relatórios mais consistentes que permitam entender o porquê determinadas situações acontecem e como fazer para evitá-las.

Atualmente é utilizada uma plataforma de relacionamento com cliente (*Customer Relationship Management - CRM*), o sistema HubSpot, para conectar marketing, vendas e atendimento ao cliente. Contudo, em sua versão gratuita, o sistema não disponibiliza os relatórios e análises necessários ao acompanhamento efetivo de todo o processo. Além disso, sabe-se que os processos dentro da Conecta ainda não estão bem definidos e estão em fase de reestruturação no momento da realização deste trabalho.

Todas as informações presentes neste documento foram coletadas a partir de reuniões com os gestores da empresa - diretor comercial, diretora financeira, gerente de compras e diretor de operações -, aplicando a técnica de *brainstorming*. Na primeira reunião foram levantadas as seguintes questões:

- Quais períodos do ano existe maior número de projetos vendidos?



- A Conecta vende mais projetos para o setor público ou privado? Há sazonalidade nesse critério?
- Quais os tipos de empresas mais atendidas?
- Qual o esforço da empresa em conquistar novos clientes *versus* esforço em vender para o mesmo cliente?
- Quais foram os projetos que mais geraram atrasos na entrega?
- Há relação entre o faturamento da empresa e a quantidade de mão-de-obra naquele período?
- Qual o canal de maior aquisição de clientes?
- Como aumentar o faturamento com os clientes já existentes?
- O retorno é diretamente proporcional ao investimento em Marketing?
- Existe uma relação entre o faturamento e o custo fixo da empresa ao longo dos anos?
- Qual a quantidade de materiais que a empresa já vendeu (por tipo)?

A Figura 2 reúne algumas certezas, dúvidas e suposições levantadas inicialmente. A Figura 3 apresenta um resumo das principais observações do tipo POEMS.

Figura 2 - Matriz CSD. Fonte: adaptado de IGTI (2023)



Figura 3 - Análise do contexto do problema (POEMS). Fonte: adaptado de IGTI (2023)



1.1.2 Personas

Os principais usuários da solução proposta por este trabalho são os gestores da empresa que fizeram parte das reuniões iniciais de brainstorming - diretor comercial, diretora financeira, diretor de operações e gerente de compras. Contudo, a partir da síntese das informações coletadas e priorização de problemas e dores, definiu-se o setor de projetos como foco inicial do trabalho e, portanto, a persona aqui especificada será o diretor comercial, que é a figura mais diretamente ligada a este setor e responsável por sua condução.

Trata-se de um profissional formado em Ciência da Computação, com pós-graduação em Redes de Computadores (Universidade Federal de Lavras), MBA em gerenciamento de projetos (IBMEC) e certificação PMP (PMI). Atua na Conecta desde 2005 e é atualmente diretor comercial e sócio da empresa. Ele é responsável pela maior parte das vendas, fazendo a supervisão das propostas e negociação com os clientes.

A Figura 4 apresenta o mapa de empatia, elaborado para se compreender melhor o diretor comercial da Conecta, no que diz respeito ao que ele pensa e sente, o que ele vê e ouve no dia a dia da empresa, bem como quais são seus principais desejos e dores.

Figura 4 - Mapa de empatia. Fonte: adaptado de IGTI (2023).



1.1.3 Benefícios e Justificativas

Conforme já descrito na contextualização do problema, a empresa Conecta não possui os processos bem estruturados e os sistemas utilizados pela empresa não permitem a elaboração de relatórios concisos sobre as atividades, o que leva aos gestores a tomada de decisão baseada em dados mais recentes e no conhecimento tácito dos envolvidos (o chamado “*feeling*”).

Dessa forma, faz-se necessário o uso de técnicas de ciência de dados para extrair informações úteis para que a empresa possa se conhecer melhor e tomar decisões baseadas em fatos e não em suposições.

A análise dos dados pode ajudar a entender questões como oscilações no faturamento, motivos que levam à perda de projetos para o concorrente ou a projetos que não finalizam, se o tamanho da equipe influencia na execução correta e completa de um projeto, quais períodos do ano geralmente a empresa vende mais, se vende mais para o setor público ou privado, qual o esforço da empresa para conquistar novos clientes e como tudo isso pode ser melhorado.



Para entender os detalhes do problema, a Tabela 1 apresenta o mapeamento das ações do diretor comercial da Conecta, de acordo com a metodologia *Blueprint*.

Tabela 1 - Detalhamento do problema (*Blueprint*). Fonte: elaborada pela autora.

Itens	Detalhamento
Objetivos	Encontrar padrões nos projetos
Atividades	Buscar clientes com características comuns, projetos semelhantes.
Questões	Projetos semelhantes, em clientes com características comuns tendem a ter o mesmo comportamento?
Barreiras	Poucas informações sobre clientes e projetos armazenadas nos sistemas da empresa.
Ações do cliente	
Funcionalidades	Visualizar os resultados da análise e dos indicadores especificados na solução.
Interação	Interação com filtros disponíveis no dashboard.
Mensagem	Interpretação das informações, reportando os “achados” para funcionários da empresa, diretamente envolvidos com os projetos.
Onde ocorre	Na sede da empresa.
Tarefas pendentes	Medidas tomadas após a análise dos dados.
Tarefas escondidas	Estudo dos dados para definir principais análises.
Processos de suporte	Dashboard interativo via ferramenta Power BI.
Saída desejável	Insights significativos para a empresa.

A análise de dados proposta neste trabalho busca trazer valor para empresa Conecta, no sentido de permitir tomada de decisão mais assertiva, baseada em informações agregadas, corretas e completas acerca dos projetos que já foram desenvolvidos. Dessa forma, a Figura 5 resume os principais benefícios e a proposta de valor deste estudo.

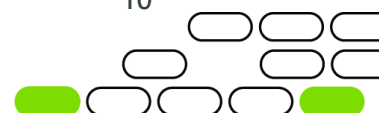


Figura 5 - Canvas de Proposta de Valor. Fonte: adaptado de IGTI (2023)



1.1.4 Hipóteses

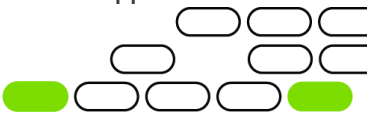
De acordo com o que foi discutido pelos diretores da Conecta na primeira reunião, uma das principais causas do atraso na entrega dos projetos é de origem externa, diz respeito a obras civis que não finalizam dentro do prazo estimado e acabam atrasando as demais atividades de outras empresas envolvidas.

Além disso, há uma falta de profissionais qualificados no mercado, o que vem impactando as atividades na Conecta, que muitas vezes precisa terceirizar alguns de seus serviços por falta de funcionários em número, disponibilidade e qualidade suficientes.

Outra hipótese que justifica a execução deste trabalho é a de que encontrar padrões nos projetos não-entregues pode ajudar no planejamento mais adequado para sua execução. Estas e outras hipóteses estão resumidas na Tabela 2.

Tabela 2 - Observações e hipóteses. Fonte: elaborado pela autora.

Observações	Hipóteses
Projetos atrasam	Obras civis de terceiros impactam nas atividades da Conecta.
Projetos não entregues	Estimativas incorretas; proprietário do projeto que viaja mais, impacta no monitoramento.



Vendas para clientes atuais é mais fácil	Falha nas estratégias de aquisição de novos clientes.
O redesenho dos processos é necessário	A melhoria nos processos e na documentação dos projetos pode ajudar na identificação de padrões.

Conforme mencionado na contextualização deste trabalho, a empresa Conecta não possui nenhum tipo de relatório sobre suas operações, tomando na maioria das vezes ações reativas e não proativas. São muitas áreas que carecem de análises aprofundadas dos dados, desde a parte financeira, de marketing, até o setor de projetos. Contudo, após realizar o agrupamento de diferentes ideias junto aos gestores, fez-se uma análise rápida de viabilidade de todas elas e criou-se a Matriz de Priorização de Ideias, apresentada na Tabela 4. Esta tabela foi baseada no método de Matriz de Priorização BASICO (Benefícios, Abrangência, Satisfação, Investimento, Cliente, Operações), cujas notas, seguiram os critérios apresentados na Tabela 3.

Tabela 3 - Balizadores para notas da Matriz BASICO.

Escala	B - Benefícios	A - Abrangência	S - Satisfação	I - Investimentos	C - Cliente	O - Operacionalidade
5	De vital importância	Total (de 70 a 100%)	Muito grande	Pouquíssimo investimento	Nenhum impacto	Muito fácil
4	Significativo	Muito grande (de 40 a 70%)	Grande	Algum investimento	Impacto pequeno	Fácil
3	Razoável	Razoável (de 20 a 40%)	Média	Médio investimento	Médio impacto	Média facilidade
2	Poucos benefícios	Pequena (de 5 a 20%)	Pequena	Alto investimento	Impacto grande	Difícil
1	Algum benefício	Muito pequena	Quase não é notada	Altíssimo investimento	Impacto muito grande no cliente	Muito difícil

Fonte: <https://engenhariaexercicios.com.br/gestao-de-qualidade/matriz-gut-basico-conceito-aplicacao-das-matrizes-priorizacao/>. Acesso em 20 de mar. de 2023.

Tabela 4 - Matriz BASICO da empresa Conecta. Fonte: elaborada pela autora.

Ideias	B	A	S	I	C	O	Total	Priorização
Construir dashboard interativo com principais indicadores sobre os projetos	5	5	5	4	3	4	26	1º
Encontrar padrões em projetos não concluídos	5	5	5	4	4	3	26	2º
Encontrar padrões em projetos concluídos	5	5	5	4	4	3	26	3º
Identificar sazonalidade nas vendas	3	3	4	4	5	4	23	4º
Identificar relação entre equipe e faturamento	4	3	4	3	4	3	21	5º
Segmentar clientes em grupos similares	4	3	3	3	4	3	21	6º



1.2 Solução

1.2.1 Objetivo SMART

O objetivo deste trabalho é auxiliar uma empresa de serviços de infraestrutura de TI a encontrar *insights* significativos sobre os projetos que ela desenvolve, de forma a identificar padrões que possam contribuir para uma tomada de decisão mais correta e ações proativas.

1.2.2 Premissas e Restrições

A solução proposta neste trabalho foi definida de acordo com algumas restrições encontradas na etapa inicial de levantamento de necessidades da organização.

A primeira delas diz respeito aos dados disponíveis na empresa Conecta. Não há um sistema integrado, nem um banco de dados contendo informações completas de clientes, vendas, projetos, etc., o que dificulta a extração de informações. Tem-se um sistema CRM, chamado Hubspot, porém este começou a ser utilizado apenas em 2015 (a empresa existe desde 2003) e os dados inseridos não estão padronizados e organizados. Isso acarreta a necessidade de um pré-processamento não tão trivial.

Outra restrição refere-se ao prazo para o desenvolvimento da solução e sua avaliação, de apenas 2 meses, conforme regras estabelecidas pelo programa de Pós-graduação da XP Educação.

Importante elencar uma terceira restrição que é a ausência de qualquer análise de dados por parte da organização. Isto é, a empresa nunca investiu recursos com foco em armazenamento de dados e extração de informações estratégicas para tomada de decisão. Dessa forma, há muitos caminhos a seguir, tornando difícil realizar o filtro dessa aplicação, que deve ser pontual, dadas as restrições citadas anteriormente.

Diante dessas situações, a Tabela 5 apresenta a matriz de riscos do projeto, que destaca os impactos gerados caso algo ocorra fora do esperado e quais ações serão realizadas para minimizar os riscos.



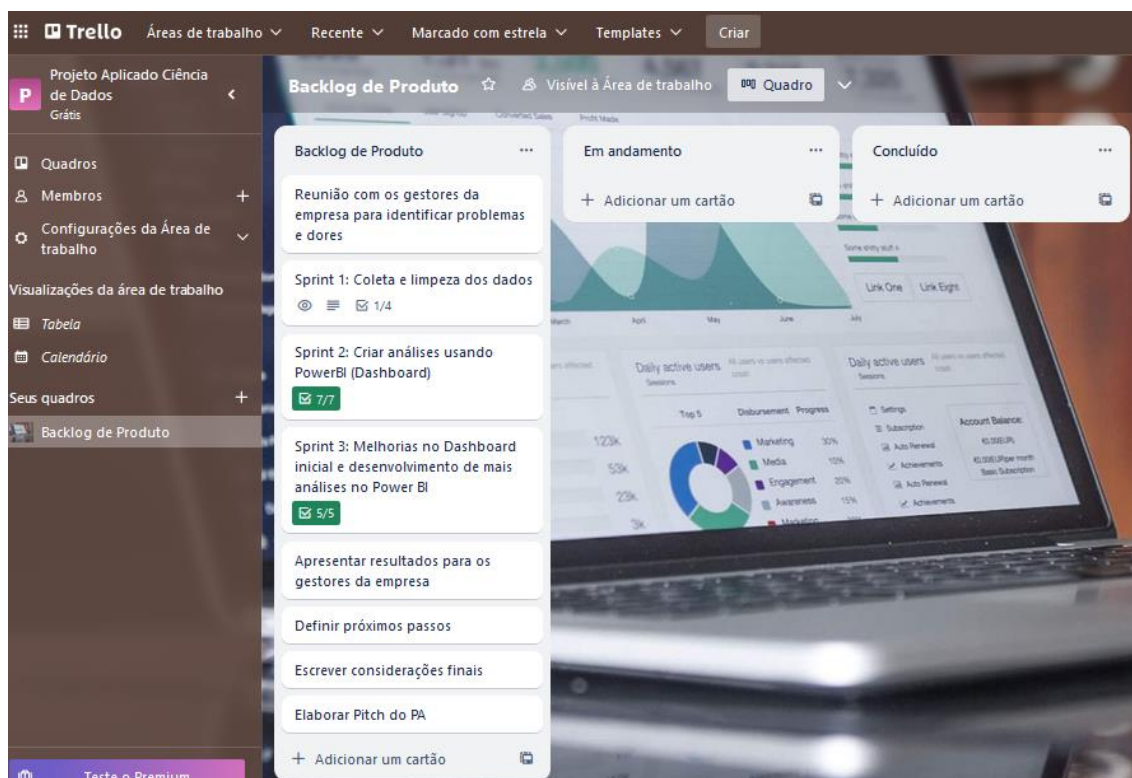
Tabela 5 - Matriz de riscos. Fonte: elaborado pela autora.

Risco identificado	Impacto Potencial	Ações preventivas	Ações corretivas
Dados serem insuficientes para análises	Atraso no cronograma do projeto aplicado	Realizar análises preliminares	Buscar mais dados junto à empresa
Não encontrar padrão nos projetos	Não atender um dos objetivos do trabalho	Realizar análises preliminares	Aplicar métodos de análise diferentes
Dashboard não fornecer insights significativos para a empresa	Não atender um dos objetivos do trabalho	Apresentar alguns gráficos iniciais aos gestores	Buscar mais dados junto à empresa

1.2.3 Backlog de Produto

Esta seção descreve os passos necessários para a realização da solução proposta neste trabalho, que é a análise dos dados sobre os projetos da empresa Conecta. Assim, a Figura 6 apresenta o backlog do produto, com o planejamento de todas as etapas, utilizando o aplicativo Trello.

Figura 6 - Backlog de Produto. Fonte: elaborado pela autora.



2. Área de Experimentação

Esta seção apresenta a descrição da execução deste projeto, detalhando todas as etapas realizadas para atingir o objetivo SMART, conforme Figura 6, apresentada na seção anterior.

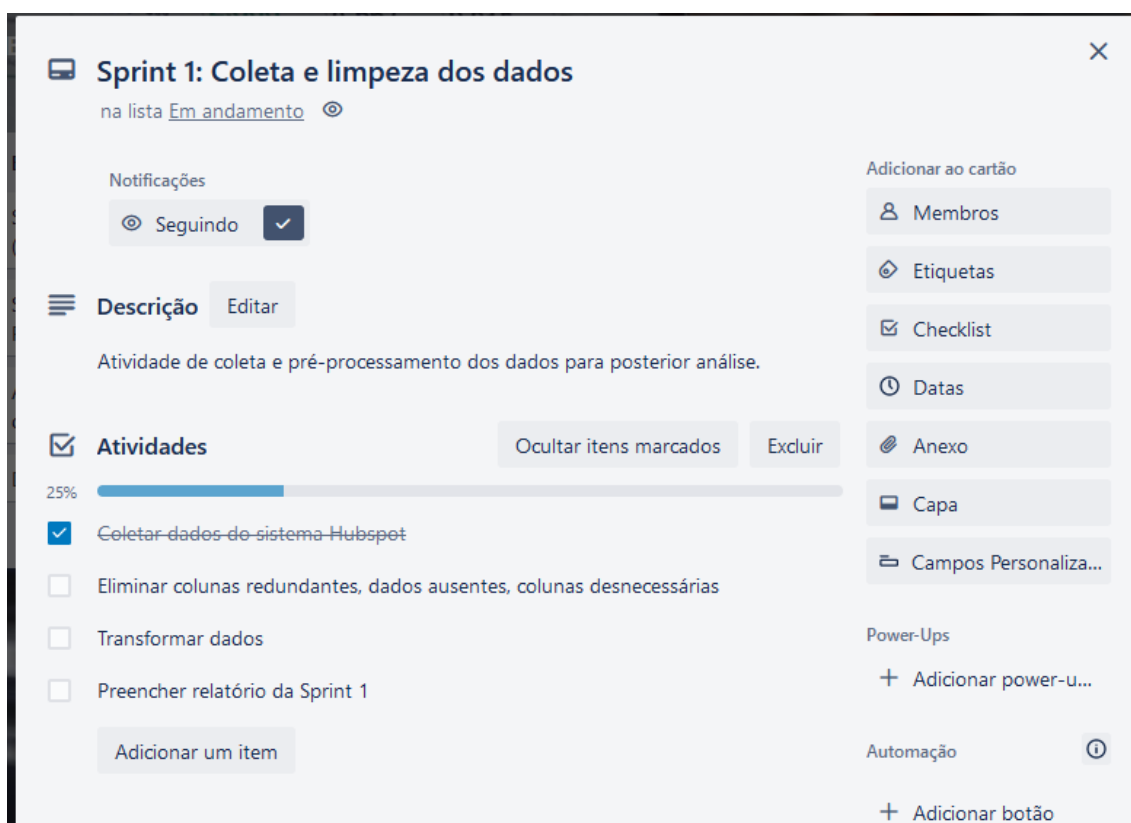
2.1 Sprint 1

2.1.1 Solução

- **Evidência do planejamento:**

Definiu-se como atividades da Sprint 1 a coleta de dados do sistema Hubspot, limpeza e tratamento dos dados sobre os projetos da Conecta, conforme descrito no sistema Trello e ilustrado pela Figura 7.

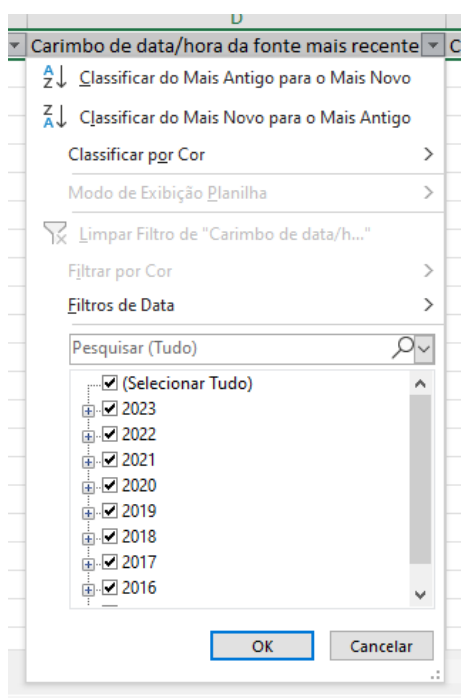
Figura 7 - Planejamento da Sprint 1. Fonte: elaborado pela autora.



- Evidência da execução de cada requisito:

Os dados foram exportados para uma planilha do Excel (extensão .xlsx), a partir da própria ferramenta Hubspot, contendo informações sobre propostas de serviços fechadas e não fechadas, concluídas e não concluídas, de março de 2016 a março de 2023, contemplando, portanto, 7 anos de atividades, com 4005 registros inseridos. A Figura 8 mostra uma captura de tela do filtro de data para evidenciar o período em questão.

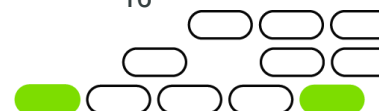
Figura 8 - Intervalo de tempo dos dados coletados. Fonte: elaborado pela autora.



A base de dados extraída do sistema Hubspot continha inicialmente 76 colunas, muitas delas não preenchidas e algumas redundantes, conforme apresenta a Tabela 6.

Tabela 6 - Estrutura inicial da base de dados. Fonte: Elaborado pela autora.

Record ID	Descrição do negócio	Shipment IDs
Abandoned cart URL	Discount savings	Source store
Acompanhamento	Ecommerce deal	Tax price
Campanha	Equipe da HubSpot	Taxa de câmbio
Campanha do último agendamento na ferramenta de reuniões	Etapas do negócio	Tipo de fonte original
Canal	Fonte de dados original 2	Tipo de negócio



Canal de aquisição	Fonte mais recente	Último contato
Canal de atendimento	Grupo de Anúncios	Valor
Carimbo de data/hora da fonte mais recente	Mídia do último agendamento na ferramenta de reuniões	Valor do contrato anual
Categoria prevista	Nome do negócio	Valor do negócio projetado
Closed Lost Reason	Novo Cliente?	Valor do negócio projetado na moeda local
Closed Won Reason	Número de atividades de vendas	Valor na moeda da empresa
Colaboração do negócio	Número de contatos associados	Valor previsto
Criado por ID do usuário	Número de contatos efetuados	Valor total do contrato
Dados da fonte mais recente 1	Order number	Associated Contact
Dados da fonte mais recente 2	Origem do último agendamento na ferramenta de reuniões	Associated Company
Dados da fonte original 1	Originado do Marketing Digital?	Associated Ticket
Data da próxima atividade	Pipeline	Associated Payment
Data da última atividade	Prioridade	Associated Subscription
Data da última modificação	Probabilidade	Associated Contact IDs
Data da última reunião agendada na ferramenta de reuniões	Probabilidade da fase de negócios	Associated Company IDs
Data de atribuição do proprietário	Probabilidade de previsão	Associated Ticket IDs
Data de criação	Proprietário do negócio	Associated Payment IDs
Data de Faturamento	Próxima etapa	Associated Subscription IDs
Data de fechamento	Receita anual recorrente	
Data do Conversor	Receita recorrente mensal	
<p>Legenda:</p> <p>Colunas vazias</p> <p>Colunas desnecessárias e ou incompletas</p> <p>Colunas redundantes</p>		

Alguns dados precisaram ser transformados, da seguinte forma, utilizando o Microsoft Power Query:

- novo cliente? - a coluna continha a frase “Sim, Cliente nunca fechou serviço antes” ou “Não, Cliente já fechou serviço anteriormente”. Os dados foram transformados para sim ou não, apenas.
- segmento do cliente - nova coluna criada, dado o nome do cliente na coluna “Associated Company” (Figura 9);
- Proposta aprovada - nova coluna criada, com os valores “sim” e “não”, de acordo com os dados das colunas “Etapas do Negócio”. Os registros com



proposta aprovada, a faturar, faturada, negócio aprovado e comissão foram definidos como propostas aprovadas;

- nome do negócio - o dado dessa coluna era formado por um código, o tipo de projeto, o nome do cliente e uma breve descrição. Foi mantido apenas o tipo de projeto;
- etapa do negócio - foram excluídos os dados das propostas ainda em aberto, cujo conteúdo nesta coluna era “Em negociação/Proposta enviada”, “Follow-up”, e “Contrato Fechado”. Todos estes registros estavam incompletos, sem valor de negócio, pois muitos deles eram apenas tentativas de vendas que não resultaram em propostas;
- data de criação - formato alterado para apenas data (estava como data e hora);
- pipeline - foram excluídos os registros cujo conteúdo da coluna pipeline (funil de vendas) eram ioRedes e prospecção pipeline, pois não dizem respeito ao setor de projetos.

Algumas colunas também foram renomeadas para facilitar o entendimento dos relatórios posteriormente, como as colunas “Closed lost reason”, “Closed Won Reason”, entre outras.

Figura 9 - Coluna adicionada “Segmento do cliente”. Fonte: elaborado pela autora.

Adicionar Coluna Condicional

Adicionar uma coluna condicional que é calculada das outras colunas ou valores.

Nome da nova coluna

Segmento do Cliente

Nome da Coluna	Operador	Valor	Saída
Se Associated Comp...	contém	ABC 123 Unidas	Então ABC 123 Aluguel de carros
Senã... Associated Comp...	contém	ABC 123 Prefeitura	Então ABC 123 Prefeitura
Senã... Associated Comp...	contém	ABC 123 InterCement	Então ABC 123 Fábrica de cimento
Senã... Associated Comp...	contém	ABC 123 Havan	Então ABC 123 Varejo
Senã... Associated Comp...	contém	ABC 123 Cimed	Então ABC 123 Indústria farmacêutica
Senã... Associated Comp...	contém	ABC 123 Sucafina	Então ABC 123 Fábrica de café

Adicionar Cláusula

Senão

ABC 123 Outros

OK Cancelar

Figura 10 - Coluna adicionada “Projeto Ganho”. Fonte: elaborado pela autora.

Adicionar Coluna Condicional

Adicionar uma coluna condicional que é calculada das outras colunas ou valores.

Nome da nova coluna

Projeto Ganho Correto

	Nome da Coluna	Operador	Valor ①		Saída ②
Se	Etapa do negócio	contém	ABC 123 Faturado	Então	ABC 123 Sim
Senã...	Etapa do negócio	contém	ABC 123 Faturar	Então	ABC 123 Sim
Senã...	Etapa do negócio	contém	ABC 123 Aprovada	Então	ABC 123 Sim
Senã...	Etapa do negócio	contém	ABC 123 Aprovado	Então	ABC 123 Sim
Senã...	Etapa do negócio	contém	ABC 123 Comissão	Então	ABC 123 Sim
Senã...	Etapa do negócio	contém	ABC 123 Perdido	Então	ABC 123 Não

Adicionar Cláusula

Senão ①

ABC 123

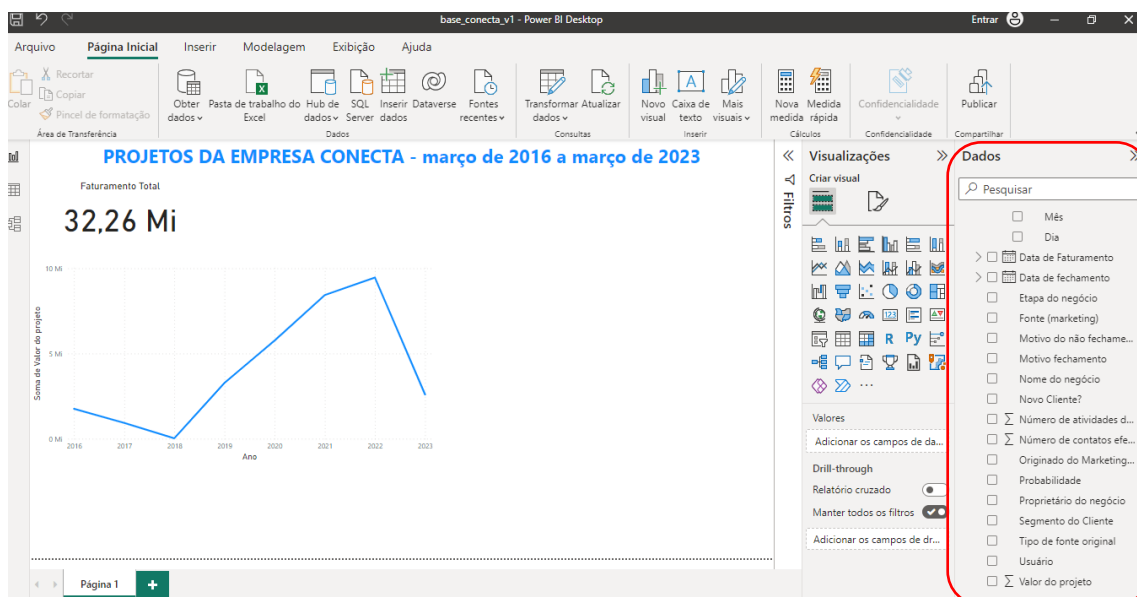
OK

Cancelar

- Evidência dos resultados:

Todo o trabalho de limpeza e tratamento dos dados foi realizado com a colaboração do Diretor Comercial, persona definida na seção 1.1.2 deste trabalho, pois foi necessário compreender cada atributo e identificar quais realmente eram úteis para as análises e quais estavam realmente sendo preenchidos e utilizados pela empresa. Depois de selecionar as colunas, transformar dados e renomear atributos, a base de dados final foi carregada no Microsoft Power BI Desktop para as análises que serão descritas na Sprint 2 (Figura 11).

Figura 11 - Base de dados pré-processada e carregada no Power BI Desktop. Fonte: elaborado pela autora.



2.1.2 Lições Aprendidas

Estima-se que a etapa de pré-processamento dos dados corresponda a 70 ou 80% de todo o trabalho de análise de dados. Foi possível perceber bem essa questão nesta primeira Sprint do projeto aplicado, sendo necessárias várias conversas informais com o diretor comercial da empresa, assim como muitas horas e dias de trabalho para conseguir adequar a base de dados aos requisitos necessários à extração de informações.

Ainda assim, ao iniciar a Sprint 2, que consiste no desenvolvimento do Dashboard com os principais indicadores de desempenho, ainda poderemos identificar mais ajustes e retornar a etapa de pré-processamento, sendo esta tarefa bastante natural e importante, dada a natureza cíclica do processo de análise de dados e a necessidade de se ter dados de qualidade.

2.2 Sprint 2

2.2.1 Solução

- **Evidência do planejamento:**

Definiu-se como atividade da Sprint 2 a elaboração do *Dashboard* com indicadores de desempenho do setor de projetos da empresa Conecta. O *checklist* com os principais itens a serem incluídos está representado na Figura 12.

Figura 12 - Planejamento da *Sprint 2*. Fonte: elaborado pela autora.

The screenshot shows a Jira board for 'Sprint 2: Criar análises usando PowerBI (Dashboard)'. It includes a 'Notificações' section with a 'Seguir' button, a 'Descrição' section with a text input field, and a 'Tarefas' section with a checklist of tasks. The tasks are:

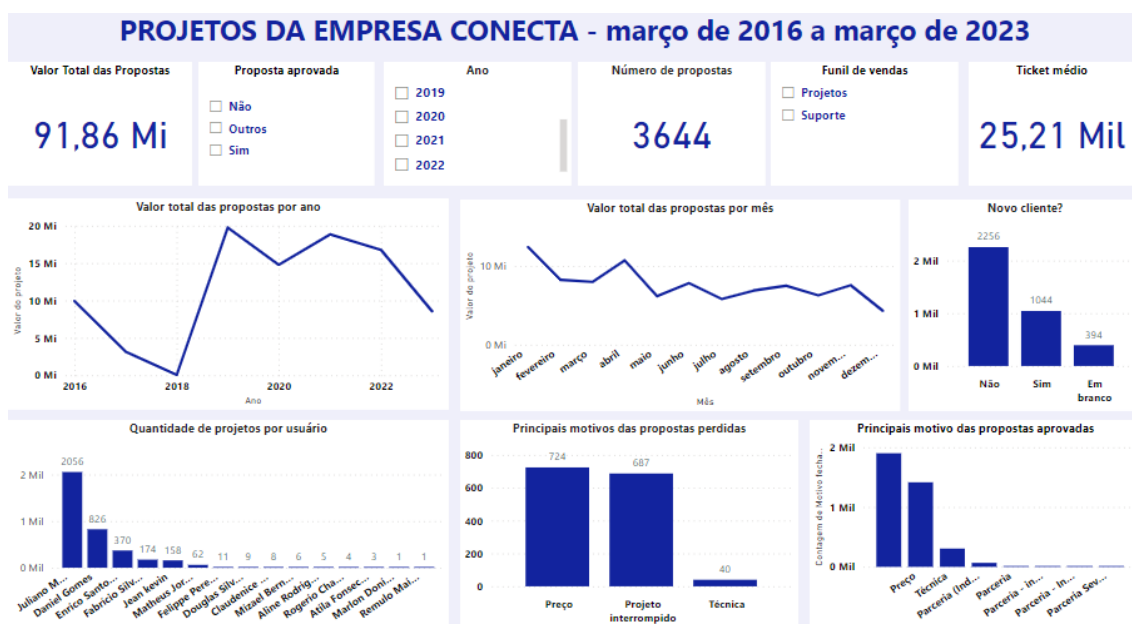
- ☐ Criar cartões com indicadores (valor total das propostas, quantidade de propostas e ticket médio)
- ☐ Criar gráfico de faturamento x ano
- ☐ Criar gráfico de quantidade de projetos por usuário
- ☐ Criar gráfico de faturamento x meses do ano
- ☐ Criar gráfico com número de novos clientes
- ☐ Criar gráficos com principais motivos de propostas aprovadas e não aprovadas
- ☐ Criar filtros (ano, propostas aprovadas, funil de vendas)

Buttons for 'Excluir' and 'Adicionar um item' are also visible.

- **Evidência da execução de cada requisito:**

Todos os requisitos dessa etapa do trabalho foram elaborados utilizando o Software Microsoft Power BI Desktop, como mostra a Figura 13.

Figura 13 - Dashboard de projetos da empresa Conecta. Fonte: elaborado pela autora.



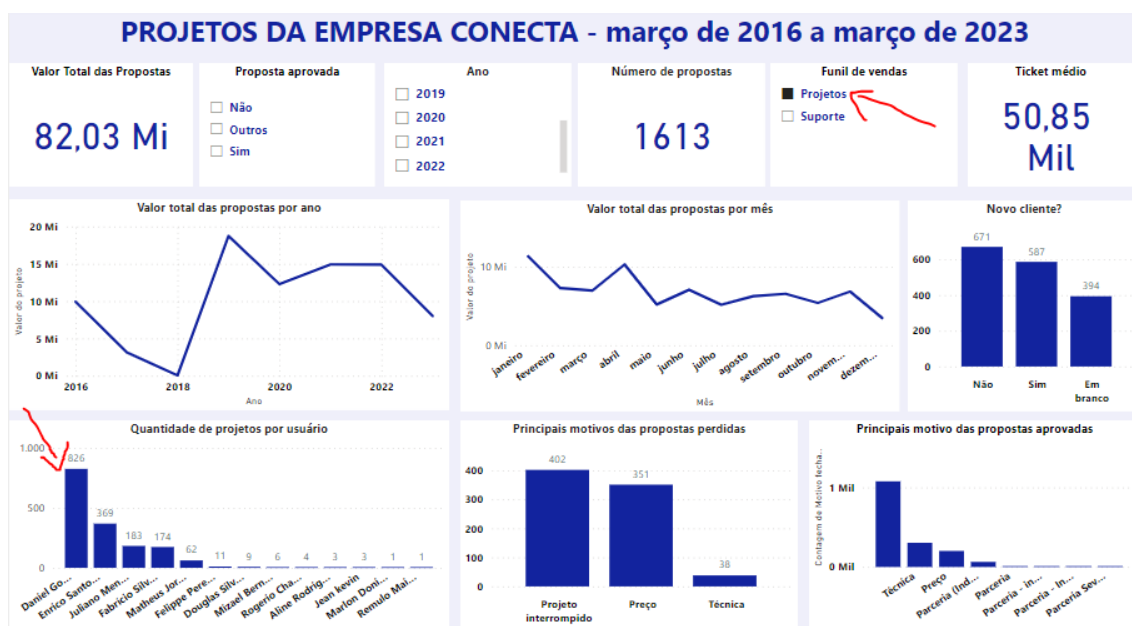
● Evidência dos resultados:

Observa-se pelo Dashboard da Figura 13 que após a etapa de limpeza dos dados, ficaram 3644 registros na base de dados. Através desses gráficos e indicadores, pôde-se observar as seguintes situações:

- o valor total das propostas por ano apresentou uma queda significativa em 2018. Contudo, verificou-se que foram registradas apenas 4 propostas no sistema neste ano. Os envolvidos não souberam o motivo;
- analisando o total de propostas por mês, em cada ano, não pôde ser observada nenhuma tendência específica. Cada ano teve oscilações em diferentes meses;
- outra observação importante é que a maioria das propostas aprovadas foram para instituições que já são clientes da Conecta, informação que pode trazer diretrizes para o marketing da empresa;
- dentre os motivos que levam à aprovação das propostas, o preço foi o que mais se destacou;
- no filtro de proposta aprovada aparece o valor “outros”, pois há propostas que ainda estão em negociação (2021, 2022 e 2023);

- a quantidade de registros por usuário foi questionada pelo Diretor Comercial, contudo, ao filtrar esses registros pelo funil de projetos, o gráfico passou a fazer mais sentido para ele, como mostra a Figura 14.

Figura 14 - Dashboard com filtro de projetos (funil de vendas). Fonte: elaborado pela autora.



2.2.2 Lições Aprendidas

O maior desafio na execução da Sprint 2 foi corrigir várias inconsistências encontradas na base de dados da Conecta durante a criação do Dashboard. Havia muitos registros com dados incompletos, dados sem padronização e campos vazios. Além disso, havia inconsistências de preenchimento, pois alguns registros estavam preenchidos como projetos faturados, sendo incluídos no filtro de propostas aprovadas, porém, havia uma descrição do motivo da proposta não ter sido aprovada. Estas constatações só foram possíveis após a criação dos gráficos e filtros.

Estas inconsistências demandaram bastante tempo para ajustes, pois foram necessárias algumas reuniões com Diretor Comercial, para que estes dados fossem corrigidos e não excluídos da base, de forma a manter uma quantidade de registros satisfatória para análises. Contudo, percebe-se que ainda há correções a serem feitas, para que as informações representem corretamente a realidade da empresa. Essas últimas correções serão feitas na Sprint 3 e antes da conclusão do trabalho.

2.3 Sprint 3

2.3.1 Solução

- Evidência do planejamento:

Conforme descrito no final da Sprint anterior, para o desenvolvimento da Sprint 3 foram realizadas reuniões com o diretor comercial da empresa Conecta para apresentação das primeiras análises feitas na Sprint 2. Durante essas reuniões, foram identificadas algumas inconsistências na base de dados e detectada necessidade de mais algumas análises. Dessa forma, o planejamento da Sprint 3 está ilustrado na Figura 15.

Figura 15 - Evidência de planejamento da Sprint 3. Fonte: elaborado pela autora.

Sprint 3: Melhorias no Dashboard inicial e desenvolvimento de mais análises no Power BI
na lista [Em andamento](#)

Notificações
[Seguir](#)

Adicionar ao cartão
Membros
Etiquetas
Checklist
Datas
Anexo
Capa
Campos Personaliz...

Descrição
Adicione uma descrição mais detalhada...

Tarefas
0%
Reunião com diretor comercial da empresa
Correções na base de dados
Melhoria do dashboard inicial
Desenvolvimento de outras análises
Criação de nova página de dashboard
Adicionar um item

Power-Ups
+ Adicionar power-...

Automação
+ Adicionar botão

Ações
Mover
Copiar
Criar template

Atividade
Mostrar Detalhes
Escrever um comentário...

- **Evidência da execução de cada requisito:**

Para a realização de novas análises, precisou-se criar novas colunas de dados agregados, como mostra a Figura 16. A pedido do diretor comercial, definiu-se como faixa de valores das propostas a seguinte escala:

- Pequeno - projetos de até R\$100.000;
- Médio - projetos de R\$100.000 a R\$ 300.000;
- Grande - projetos acima de R\$ 300.000.

E para a coluna de “Tempo de Faturamento”:

- Até 30 dias;
- De 30 a 60 dias;
- De 60 a 90 dias;
- Acima de 90 dias.

Figura 16 - Outras colunas adicionadas para novas análises. Fonte: elaborado pela autora.

The screenshot shows the Microsoft Power Query Editor interface. The ribbon at the top includes 'Arquivo', 'Página Inicial', 'Transformar', 'Adicionar Coluna', 'Exibição', 'Ferramentas', and 'Ajuda'. The 'Transformar' tab is active, showing options like 'Fechar e Aplicar', 'Nova Fonte', 'Fontes Recentes', 'Inserir Dados', 'Configurações da fonte de dados', 'Gerenciar Parâmetros', 'Atualizar Visualização', 'Editor Avançado', 'Gerenciar', 'Escolher Colunas', 'Remover Colunas', 'Reduzir Linhas', 'Dividir Coluna', 'Agrupar por', and 'Tipo de Dados: Data'. The main area displays a table with the following columns: 'company', 'Segmento do Cliente', 'Proposta aprovada', 'Data faturamento - criação', 'Faixa valores', and 'Tempo faturamento'. The table contains 15 rows of data.

	company	Segmento do Cliente	Proposta aprovada	Data faturamento - criação	Faixa valores	Tempo faturamento
1		Varejo	Sim		0 Pequeno	até 30 dias
2		Aluguel de carros	Em aberto		null Pequeno	não se aplica
3		Aluguel de carros	Em aberto		null Pequeno	não se aplica
4		Outros	Em aberto		null Pequeno	não se aplica
5		Outros	Em aberto		null Pequeno	não se aplica
6		Varejo	Em aberto		null Pequeno	não se aplica
7	os	Outros	Sim		null Pequeno	não se aplica
8		Aluguel de carros	Sim		null Pequeno	não se aplica
9		Varejo	Sim		-3 Pequeno	até 30 dias
10		Outros	Em aberto		null Pequeno	não se aplica
11		Outros	Em aberto		null Pequeno	não se aplica
12	ipal de Lavras	Prefeitura	Sim		5 Pequeno	até 30 dias
13		Aluguel de carros	Sim		0 Pequeno	até 30 dias
14		Aluguel de carros	Sim		0 Pequeno	até 30 dias
15		Aluguel de carros	Sim		0 Pequeno	até 30 dias

As correções feitas na base após a Sprint 2 foram as seguintes:

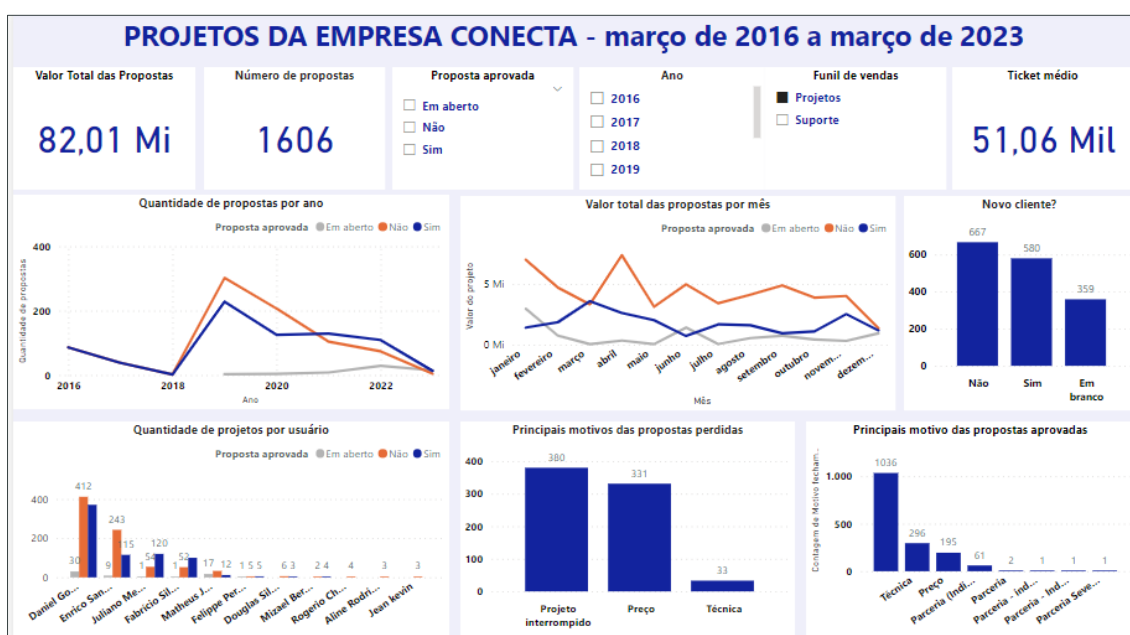
- propostas aprovadas e faturadas, porém, com a coluna de “motivo de perda” preenchida;
- propostas com etapa do negócio como “faturada”, porém, sem data de faturamento;
- propostas com data de faturamento, porém com etapa do negócio como “proposta perdida”;

- registros com “etapa do negócio” em branco;
- removidos 54 registros com valor 0, 1 ou nulo;
- 356 registros preenchidos como “não informado” na coluna “originado do marketing digital”.

• **Evidência dos resultados:**

A Figura 17 apresenta as melhorias realizadas na primeira página do Dashboard, com dados inconsistentes removidos e apresentação de propostas aprovadas e não aprovadas nos mesmos gráficos.

Figura 17 - Primeira página de análises do Dashboard. Fonte: elaborado pela autora.

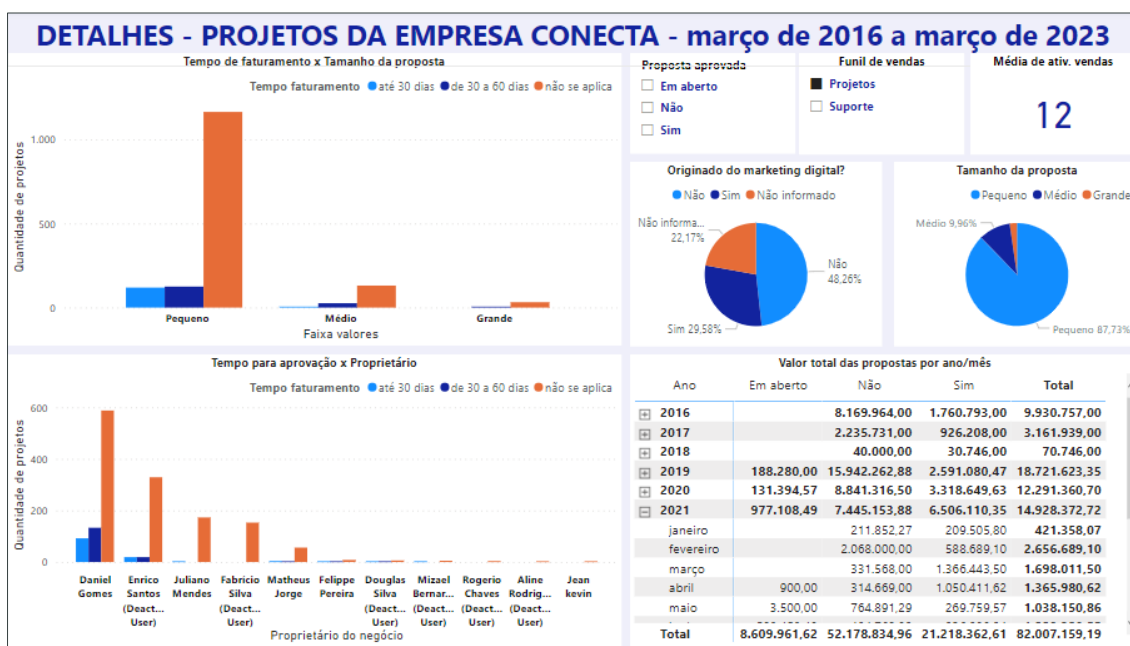


A Figura 18 mostra a segunda página desenvolvida com mais análises, a pedido do diretor comercial, contendo:

- Gráfico de barras relacionando o tempo para faturamento com o tamanho do projeto, no sentido de verificar se o tempo para faturamento estava condicionado ao tempo para sua execução e finalização (projetos grandes demoram mais para serem negociados e executados, por exemplo);

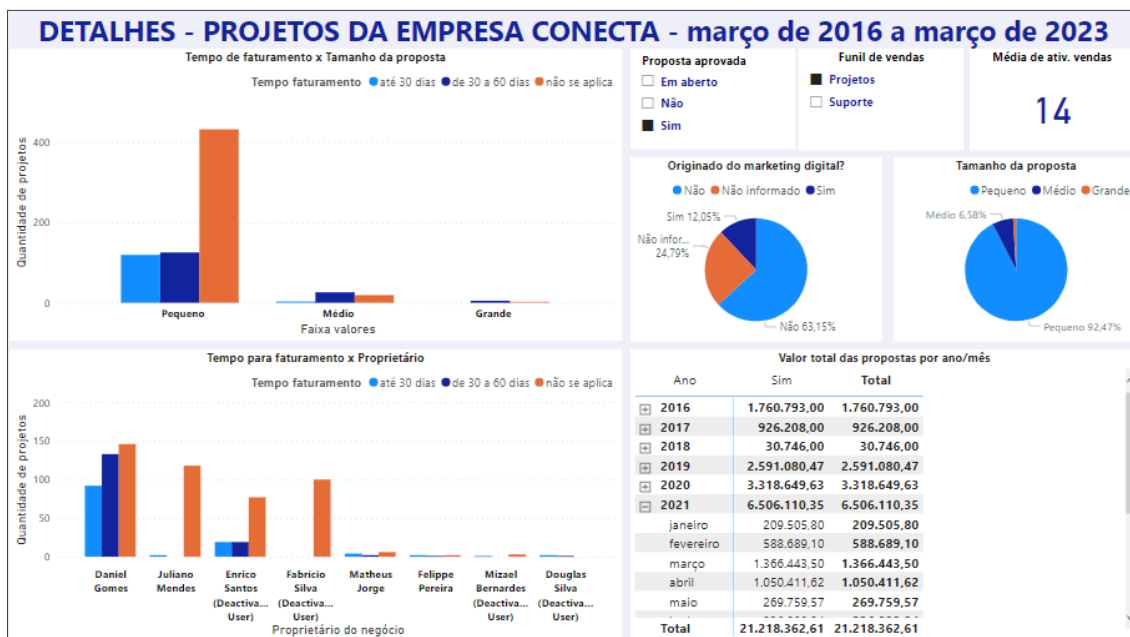
- Gráfico de pizza mostrando a proporção de propostas originadas do marketing digital;
- Gráfico de pizza apresentando a proporção de projetos de acordo com seu tamanho;
- Gráfico de barras relacionando o tempo para faturamento com o proprietário do negócio, no sentido de analisar se o proprietário do negócio exerce alguma influência na demora na aprovação e faturamento do projeto;
- Cartão com a média de atividades de vendas na negociação da proposta;
- Matriz com todos os valores condensados, para melhor visualização pelos gestores.

Figura 18 - Segunda página de análises do Dashboard. Fonte: elaborado pela autora.



Pode-se observar pela figura que os gráficos apresentam grande número de propostas com o valor “não se aplica” para o tempo de faturamento. Isso porque o filtro de aprovação não foi selecionado, então o tempo de faturamento não se aplica às propostas perdidas. Contudo, ao se aplicar o filtro, ainda aparecem registros dessa forma, como mostra a Figura 19.

Figura 19 - Segunda página de análises com filtro aplicado. Fonte: elaborado pela autora.



Isso deve-se ao fato de que há muitas propostas em aberto (algumas demoram meses em negociação), outras que não foram devidamente atualizadas no sistema (tornando alguns registros ainda incorretos) e algumas que acabaram de ser inseridas (relativas aos meses de fevereiro e março de 2023), que não foram removidas das análises a pedido do diretor comercial. Ainda assim, muitos insights foram possíveis e estão descritos nas considerações finais deste trabalho.

2.3.2 Lições Aprendidas

Nesta etapa do trabalho ficou ainda mais evidente quão árduo é o trabalho do cientista de dados, principalmente em se tratando de empresas que não possuem a cultura de registro de informações bem enraizada, resultando em dados faltosos, confusos e muitas vezes incoerentes.

Havia a intenção de se aplicar mineração de dados nesse projeto, contudo, a base de dados ainda carece de muitos ajustes e, portanto, essa tarefa ficará no planejamento dos próximos passos. Para evidenciar essa informação, foi feito o carregamento dos dados na ferramenta Google Colab, para facilitar a visualização dos dados faltantes (Figura 20).

Figura 20 - Dataset da Conecta no Google Colab. Fonte: elaborado pela autora.

Colunas, quantidade de registros e tipos de dados				Quantidade de valores nulos	
<pre><class 'pandas.core.frame.DataFrame'> RangeIndex: 3636 entries, 0 to 3635 Data columns (total 28 columns): # Column Non-Null Count Dtype --- - 0 Movido não aprovação 1379 non-null object 1 Motivo aprovação 1738 non-null object 2 Usuário 2873 non-null object 3 Fonte (marketing) 3494 non-null object 4 Data de criação 3636 non-null datetime64[ns] 5 Data de Faturamento 1553 non-null datetime64[ns] 6 Etapa do negócio 3636 non-null object 7 Fonte mais recente 3494 non-null object 8 Grupo de Anúncios 358 non-null object 9 Nome do negócio 3636 non-null object 10 Novo Cliente? 3636 non-null object 11 Número de atividades de vendas 3143 non-null float64 12 Número de contatos associados 3636 non-null int64 13 Número de contatos efetuados 3143 non-null float64 14 Originado do Marketing Digital? 3636 non-null object 15 Funil de vendas 3636 non-null object 16 Probabilidade 3515 non-null object 17 Probabilidade da fase de negócios 3636 non-null float64 18 Proprietário do negócio 3636 non-null object 19 Tipo de fonte original 3516 non-null object 20 Tipo de negócio 93 non-null object 21 Valor 3636 non-null float64 22 Associated Company 3466 non-null object 23 Segmento do Cliente 3636 non-null object 24 Proposta aprovada 3636 non-null object 25 Data faturamento - criação 1553 non-null float64 26 Faixa valores 3636 non-null object 27 Tempo faturamento 3636 non-null object dtypes: datetime64[ns](2), float64(5), int64(1), object(20) memory usage: 795.5+ KB</pre>				<pre>df.isna().sum() Movido não aprovação 2257 Motivo aprovação 1898 Usuário 763 Fonte (marketing) 142 Data de criação 0 Data de Faturamento 2083 Etapa do negócio 0 Fonte mais recente 142 Grupo de Anúncios 3278 Nome do negócio 0 Novo Cliente? 0 Número de atividades de vendas 493 Número de contatos associados 0 Número de contatos efetuados 493 Originado do Marketing Digital? 0 Funil de vendas 0 Probabilidade 121 Probabilidade da fase de negócios 0 Proprietário do negócio 0 Tipo de fonte original 120 Tipo de negócio 3543 Valor 0 Associated Company 170 Segmento do Cliente 0 Proposta aprovada 0 Data faturamento - criação 2083 Faixa valores 0 Tempo faturamento 0 dtype: int64</pre>	



3. Considerações Finais

3.1 Resultados

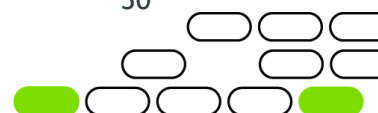
O objetivo deste trabalho foi analisar dados sobre os projetos de infraestrutura de TI da empresa Conecta. Trata-se de uma empresa familiar, de médio porte, cuja sede encontra-se em Lavras-MG. A principal persona definida no trabalho é o diretor comercial, que atua diretamente na venda e supervisão dos projetos em estudo.

Foram extraídos dados do sistema Hubspot, exportados para uma planilha, contendo inicialmente 4005 registros com 76 colunas, com informações como: data de criação, proprietário do negócio (quem fez a venda), nome do cliente, se foi originado do marketing digital ou não, se é novo cliente ou não, qual o valor da proposta, o status da proposta (aprovada, perdida, em negociação, etc.), data do faturamento, entre outras.

Foi necessário um intenso trabalho de limpeza e tratamento dos dados, pois a autora deste estudo não atua na empresa e precisou compreender todo o processo de venda e execução dos projetos na Conecta para realizar essa tarefa, o que foi feito através de várias reuniões com o diretor comercial. Além disso, havia muitos registros incompletos, incorretos e incoerentes, demandando análises praticamente individuais dos mesmos, o que foi feito, em sua maioria, na primeira sprint do trabalho.

Nas Sprints 2 e 3 foi criado um Dashboard com duas páginas de análises, contendo o valor total das propostas, o número total de propostas, ticket médio da empresa, média do número de atividades de vendas, uma matriz com o resumo dos principais valores, bem como diferentes gráficos. Pôde-se observar que:

- 2019 foi o ano com maior número total de propostas, dentre os anos analisados;
- Em 2021, teve-se mais propostas aprovadas do que não aprovadas, mesmo considerando as propostas em aberto, o que é raro de acontecer;
- Houve uma queda significativa no valor dos projetos em 2018, porém, ao analisar os registros constatou-se que havia apenas 5 registros nesta data, o que não pôde ser explicado pelo diretor comercial;
- A empresa vende mais para os atuais clientes do que para novos, como já se sabia;



- Os principais motivos que levam à aprovação das propostas pelos clientes são a técnica e o preço;
- Os principais motivos que levam à perda das propostas são “projeto interrompido” e também o preço;
- A premissa de que o tempo de faturamento seria proporcional ao tamanho do projeto foi refutada, pois não há um padrão nesse caso (tempo decorrido desde o primeiro contato com o cliente até a data de faturamento da proposta já aceita e executada ou em execução);
- A maior parte das propostas não é originada do marketing digital, como se imaginava inicialmente;
- Mais de 85% de todas os projetos que a empresa elabora possui valor de até 100 mil Reais;
- A média de atividades de vendas das propostas aprovadas é 14, e das propostas em aberto 18.

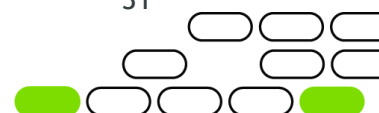
Outros insights também foram identificados com a aplicação dos filtros inseridos no dashboard, mas não serão detalhados aqui.

3.2 Contribuições

Este trabalho pôde demonstrar para a empresa em estudo quão importante se faz o correto registro das atividades de vendas dos projetos, dado que muitas inconsistências só puderam ser identificadas a partir do desenvolvimento deste. Isso corroborou a necessidade da empresa em padronizar seus processos internos que, como já descrito na contextualização, essa atividade já se encontra em andamento.

As análises aqui desenvolvidas também trouxeram informações novas e conseguiram validar ou refutar algumas hipóteses, como por exemplo, a questão do tempo de faturamento de cada proposta não ser relativa ao tamanho do projeto, a maioria das vendas serem para organizações que já são clientes da Conecta, entre outras.

Este projeto aplicado também contribuiu para que a empresa entendesse o grande potencial do trabalho de análise de dados e como este pode ser útil para a organização, que agora pretende manter ativo este setor de forma permanente.



3.3 Próximos passos

Como próximos passos, sugere-se a aplicação de mineração de dados nessa base, de forma a encontrar mais padrões que possam auxiliar a empresa de alguma forma, como previsão de aprovação de uma proposta, segmentação de clientes em grupos, entre outros.

Também pode-se realizar mais análises em dashboard, incluindo mais insights relacionados aos clientes. Essa parte não foi realizada neste trabalho, pois a base de dados disponibilizada não possui informações sobre os clientes, apenas o nome.

