



Juiz de Fora, 24 de julho de 2023

CONTROLE DE VERSÃO			
Autor	Versão	Data	Descrição
Marcell Felipe de Paula Oliveira	1.0	23/07/2023	Criação do documento

Sumário

Lista de Figuras	3
1 Introdução	4
2 Solicitação	4
3 Premissas da Solução	4
3.1 Origem e especificação dos dados:	4
3.2 Ambiente de Desenvolvimento:	4
4 Modelo da Arquitetura Sugerida	5
5 Dicionário de Dados	6
6 Processo e Desenvolvimento	7
7 Dashboard	8
7.1 Link do Dashboard no Power Bi	13

Lista de Figuras

1	Arquitetura Sugerida no Projeto.	5
2	StarSchema	7
3	Capa	8
4	Faturamento Gênero	9
5	Faturamento	10
6	Cohort	11
7	Pareto	12
8	Análise Geral	13

1 Introdução

Este documento visa detalhar todos os objetivos do projeto de Análise e Acompanhamento de Vendas e Clientes, sob uma perspectiva técnica. Ele lista as soluções, premissas e atividades de execução durante a elaboração do projeto, no contexto do processo seletivo do Fiesc.

2 Solicitação

Com o objetivo de otimizar o acompanhamento de clientes e vendas, o Fiesc solicitou um projeto no qual realizamos um mapeamento detalhado dessas duas áreas. O propósito é identificar padrões e tendências, organizando os dados por sexo, faixa etária, categoria de produto e produtos específicos, destacando as principais informações em uma linha do tempo.

3 Premissas da Solução

Nesta seção discutiremos as premissas da solução.

3.1 Origem e especificação dos dados:

Estabeleceremos uma conexão com um SGBD PostgreSQL. A primeira etapa envolverá a restauração da base de dados. Na segunda etapa, elaboraremos as views, criando tabelas de fato e dimensão para conexão com o Power BI. Na terceira etapa, procederemos à elaboração do dashboard. Os dados estão atualmente no formato de tabelas e devem aderir a um layout pré-definido, que está especificado neste documento.

3.2 Ambiente de Desenvolvimento:

PostgreSQL, Microsoft Power BI.

4 Modelo da Arquitetura Sugerida

A arquitetura proposta foi inicialmente desenvolvida a partir de um arquivo para restauração de banco de dados, onde os dados estão armazenados. Utilizaremos o PostgreSQL após a restauração do banco de dados, no qual criaremos as Views (tabelas virtuais). Continuando com a etapa anterior, as informações agora estruturadas serão armazenadas no PostgreSQL e, por fim, os dados serão modelados no Power BI para a elaboração de Dashboards correspondentes aos dados obtidos.

Vale ressaltar o uso do PostgreSQL, que foi crucial para toda a orquestração e desenvolvimento do projeto, pois foi através dele que criamos as tabelas de fatos e dimensões.

A figura abaixo apresenta a arquitetura da solução proposta, baseada no levantamento de requisitos e entendimento do negócio.



Figura 1: Arquitetura Sugerida no Projeto.

5 Dicionário de Dados

Abaixo se encontra o detalhamento dos dados necessários para aplicação da solução.

Campo	Descrição	Tipo	Restrição de domínio	Chave	Tamanho
<i>id_venda</i>	Identificador de venda	Int	not null	PK	-
<i>data_venda</i>	Data da venda	Date	not null	-	-
<i>desconto_venda</i>	Desconto da venda	Float	null	-	-
<i>status_venda</i>	Status da venda	Varchar	not null	-	max
<i>id_cliente</i>	Identificador de cliente	Int	not null	FK	-
<i>nome_cliente</i>	Nome do cliente	Varchar	not null	-	max
<i>email_cliente</i>	Email do cliente	Varchar	null	-	max
<i>sexo_cliente</i>	Sexo do cliente	Varchar	null	-	max
<i>data_nascimento_cliente</i>	Data de nascimento do cliente	Date	null	-	-
<i>id_produto</i>	Identificador do produto	Int	not null	FK	-
<i>nome_produto</i>	Nome do produto	Varchar	not null	-	max
<i>preco_produto</i>	Preço do produto	Float	not null	-	-
<i>id_categoria</i>	Identificador da categoria	Int	not null	FK	-
<i>nome_categoria</i>	Nome da categoria	Varchar	not null	-	max
<i>quantidade_itens</i>	Quantidade de itens na venda	Int	not null	-	-
<i>preco_itens</i>	Preço dos itens na venda	Float	not null	-	-

Tabela 1: Tabela de Vendas

Campo	Descrição	Tipo	Restrição de domínio	Chave	Tamanho
<i>id_cliente</i>	Identificador do cliente	Int	not null	PK	-
<i>nome_cliente</i>	Nome do cliente	Varchar	not null	-	max
<i>email_cliente</i>	Email do cliente	Varchar	null	-	max
<i>sexo_cliente</i>	Sexo do cliente	Varchar	null	-	max
<i>data_nascimento_cliente</i>	Data de nascimento do cliente	Date	null	-	-

Tabela 2: Tabela de Clientes

Campo	Descrição	Tipo	Restrição de domínio	Chave	Tamanho
<i>id_produto</i>	Identificador do produto	Int	not null	PK	-
<i>nome_produto</i>	Nome do produto	Varchar	not null	-	max
<i>preco_produto</i>	Preço do produto	Float	not null	-	-
<i>id_categoria_produto</i>	Identificador da categoria do produto	Int	not null	FK	-
<i>nome_categoria</i>	Nome da categoria	Varchar	not null	-	max

Tabela 3: Tabela de Produtos

Campo	Descrição	Tipo	Restrição de domínio	Chave	Tamanho
$id_{categoria}$	Identificador da categoria	Int	not null	PK	-
$nome_{categoria}$	Nome da categoria	Varchar	not null	-	max

Tabela 4: Tabela de Categorias

Campo	Descrição	Tipo	Restrição de domínio	Chave	Tamanho
id_{item}	Identificador do item	Int	not null	PK	-
$id_{venda_{item}}$	Identificador da venda do item	Int	not null	FK	-
$id_{produto_{item}}$	Identificador do produto do item	Int	not null	FK	-
$quantidade_{item}$	Quantidade do item	Int	not null	-	-
$preco_{item}$	Preço do item	Float	not null	-	-
$nome_{produto}$	Nome do produto	Varchar	not null	-	max
$preco_{produto}$	Preço do produto	Float	not null	-	-
$nome_{categoria}$	Nome da categoria	Varchar	not null	-	max

Tabela 5: Tabela de Itens Vendidos

6 Processo e Desenvolvimento

As atividades da pipeline estão enumeradas a seguir:

- 1 O acesso para coleta é feito através de um restore do banco Postgres.
- 2 Em seguida, os dados serão armazenados no Postgres, a partir do qual as informações serão utilizadas para armazenamento. As tabelas serão modeladas em views no SQL.
- 3 Com toda a modelagem feita para um formato Star Schema, os dados serão enviados para o Power BI em formato de cinco tabelas.
- 4 Com o tratamento dos dados pela ferramenta anterior concluído, os dados serão modelados no Power BI, com o objetivo de obter estatísticas relevantes para o projeto em forma de dashboards.

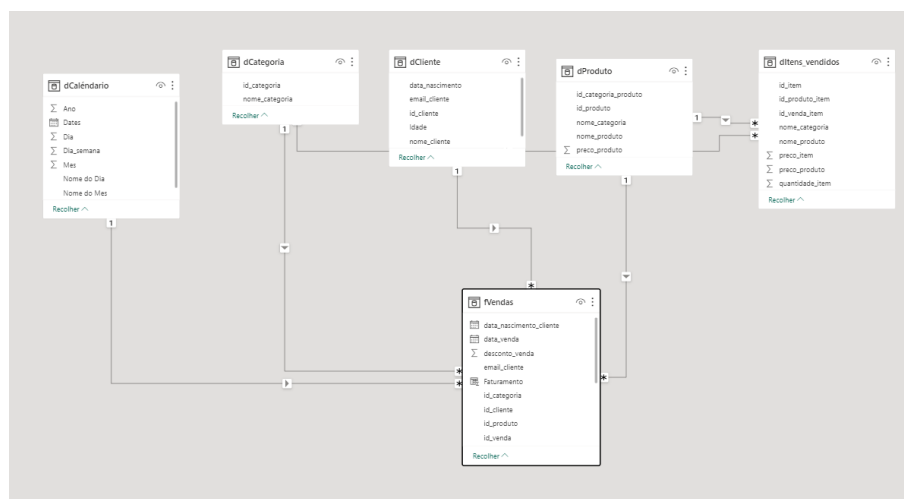


Figura 2: StarSchema

7 Dashboard

Na Figura 3, temos a capa na qual pensamos em toda a experiência do usuário. Procuramos tornar a interface o mais interativa possível, com pop-ups informando a ação que cada respectivo botão irá realizar.



Figura 3: Capa

Na Figura 4, temos um total de três painéis onde um informa a quantidade de clientes, a quantidade de produtos e uma segmentação de dados, fornecendo as idades dos possíveis compradores. Utilizamos um gráfico Treemap para analisar os dados de idade e faturamento. Um gráfico de cascata é usado para analisar os gêneros dos compradores e quanto eles renderam em termos de faturamento. Por último, empregamos outro gráfico de cascata para analisar o faturamento por produto.

Justificativa de Design: Escolhemos um gráfico Treemap porque ele proporciona uma excelente visualização hierárquica, tornando fácil ver os detalhes dos dados, como idade e faturamento. O gráfico de cascata é ideal para representar a contribuição de diferentes categorias (neste caso, gênero dos compradores e faturamento por produto) para um total.

Explicação de Insights: A partir do gráfico Treemap, podemos identificar quais faixas etárias contribuem mais para o faturamento. O gráfico de cascata permite entender qual gênero traz mais lucro e quais produtos são mais rentáveis. Estas informações podem ajudar a direcionar estratégias de marketing e vendas.

Processo de Tomada de Decisões: Decidimos explorar a idade dos compradores, o gênero e o faturamento por produto, pois acreditamos que esses aspectos são vitais para entender o comportamento do cliente. Essa compreensão é essencial para personalizar as ofertas e melhorar a eficiência das estratégias de vendas.

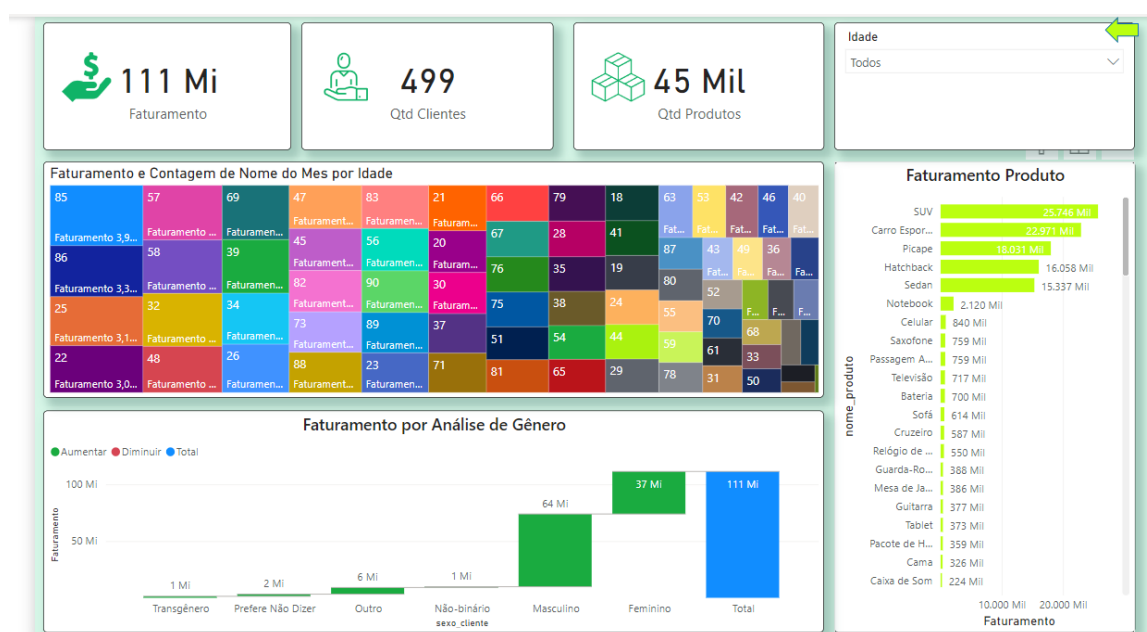


Figura 4: Faturamento Gênero

Na Figura 5, implementamos dois gráficos de colunas agrupadas. O primeiro é utilizado para a análise do faturamento por produto e o segundo para a análise do faturamento por categoria. Adicionamos um total de cinco segmentações de dados para uma análise mais detalhada: ano, mês, trimestre, nome do dia da semana e o status da compra (se é cancelada, concluída ou pendente).

Justificativa de Design: Optamos pelos gráficos de colunas agrupadas por sua eficácia na comparação de múltiplas categorias. Eles oferecem uma visão clara das diferenças entre as categorias, neste caso, produtos e tipos de produtos. As segmentações de dados foram escolhidas para permitir um desdobramento das informações ao longo do tempo e pelo status da venda, oferecendo assim, uma visão mais detalhada do desempenho das vendas.

Explicação de Insights: A partir dos gráficos de colunas agrupadas, podemos identificar quais produtos e categorias geram mais faturamento. Com as segmentações, podemos ver como o faturamento se altera

ao longo do tempo, nos diferentes dias da semana e como o status da compra influencia o faturamento. Esses insights são vitais para planejar ações de marketing, gerenciar estoques e prever tendências de vendas.

Processo de Tomada de Decisões: Decidimos explorar o faturamento por produto e categoria, a influência do tempo e o status da compra. Acreditamos que essas são áreas chave para entender o desempenho das vendas e para tomar decisões informadas que possam aprimorar as estratégias de negócio.

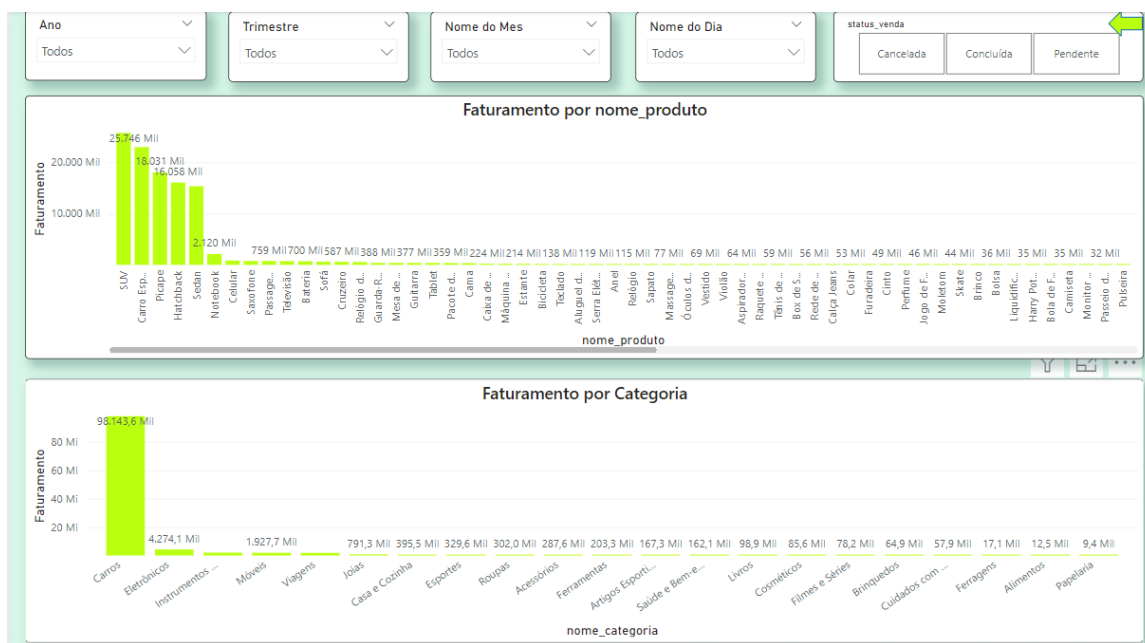


Figura 5: Faturamento

Na Figura 6, implementamos duas segmentações de dados. A primeira armazena os nomes das categorias e a segunda, os anos. Ambas foram aplicadas com o objetivo de aprimorar nossa análise de cohort.

Justificativa de Design: A escolha por incluir essas duas segmentações de dados decorre da necessidade de entender o comportamento do usuário ou do cliente ao longo do tempo (análise de cohort). Segmentar os dados por categorias e anos permite identificar tendências e padrões específicos ao longo do tempo e por categorias de produtos.

Explicação de Insights: A partir dessas segmentações, somos capazes de identificar tendências de compras em diferentes categorias ao longo dos anos. Esses insights podem fornecer informações valiosas sobre as preferências dos clientes, permitindo a realização de campanhas de marketing mais eficazes e a otimização do sortimento de produtos.

Processo de Tomada de Decisões: Decidimos explorar a análise de coorte por acreditar que ela pode fornecer insights valiosos sobre o comportamento do cliente ao longo do tempo. Ao entender como os clientes interagem com diferentes categorias de produtos ao longo dos anos, podemos tomar decisões informadas que ajudem a melhorar a experiência do cliente e aumentar as vendas.

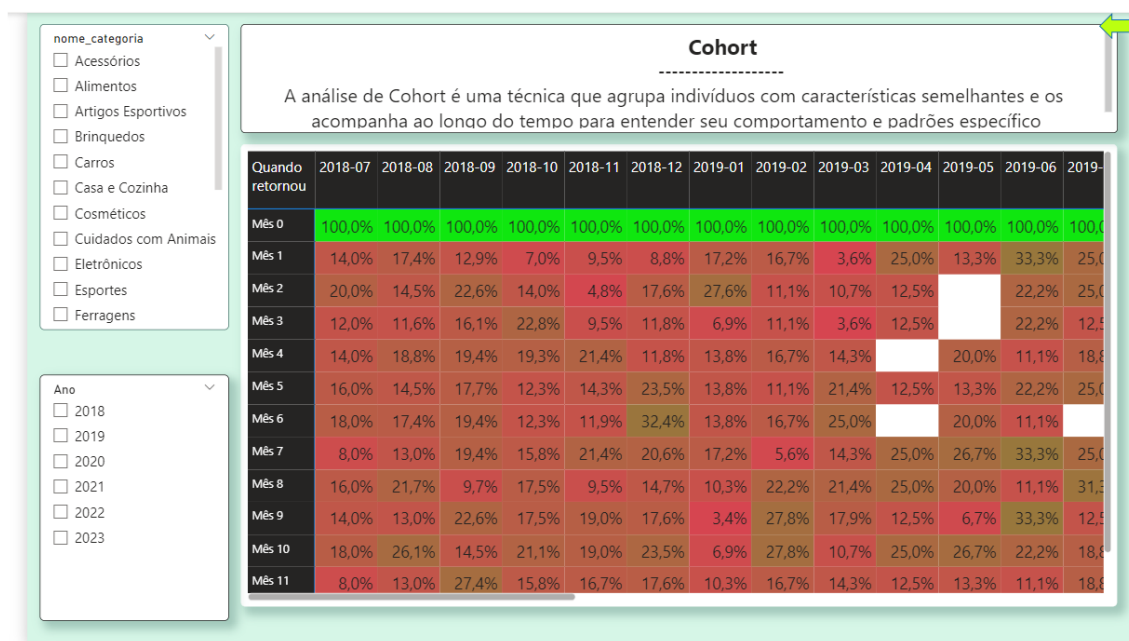


Figura 6: Cohort

Na Figura 7, temos um gráfico de Pareto acompanhado da curva ABC, ambos auxiliando na mensuração da evolução e identificação dos principais contribuintes. Além disso, recorreremos a uma matriz para ranquear o produto com a maior vendabilidade.

Justificativa de Design: O gráfico de Pareto, juntamente com a curva ABC, foi utilizado para identificar os produtos que mais contribuem para as vendas totais. Este tipo de análise é útil para priorizar esforços de marketing ou de otimização de estoque. A matriz foi empregada para classificar os produtos de acordo com a sua vendabilidade, fornecendo um ranking claro que pode orientar estratégias de negócios.

Explicação de Insights: O gráfico de Pareto, combinado com a curva ABC, pode revelar quais produtos estão gerando a maior parte da receita, permitindo que a empresa se concentre em melhorar ainda mais esses produtos ou promovê-los adequadamente. Além disso, a matriz de vendabilidade oferece uma representação visual clara dos produtos que têm o melhor desempenho, o que pode influenciar decisões de marketing e de gestão de estoque.

Processo de Tomada de Decisões: Optamos por essas formas de análise porque fornecem uma visão clara e quantificável do desempenho do produto. Acreditamos que entender quais produtos são mais vendidos e quais contribuem mais para a receita total é crucial para tomar decisões estratégicas informadas que possam melhorar a eficiência e a lucratividade da empresa.



Figura 7: Pareto

Na Figura 8, utilizamos um gráfico de linha para analisar a evolução do faturamento ao longo do tempo. Para uma análise mais detalhada, incluímos segmentações por ano, mês, trimestre e nome do dia.

Justificativa de Design: O gráfico de linha é uma das melhores maneiras de visualizar a evolução de uma métrica ao longo do tempo devido à sua clareza e simplicidade. A inclusão das segmentações por ano, mês, trimestre e nome do dia permite um exame mais detalhado da performance do faturamento em diferentes períodos.

Explicação de Insights: Com esse gráfico, podemos entender como o faturamento muda e identificar possíveis padrões ou tendências. Por exemplo, podemos descobrir se existe um mês específico em que as vendas sempre aumentam ou se há um dia da semana que tende a ter um desempenho melhor. Esses insights podem ajudar a orientar as estratégias de marketing e vendas.

Processo de Tomada de Decisões: Ao decidir que aspectos dos dados explorar e destacar, nossa motivação foi permitir uma visão mais granular do desempenho do faturamento. Ao entender como o faturamento muda em diferentes períodos, podemos identificar oportunidades para otimizar as vendas e a estratégia de negócios. Por exemplo, se identificarmos que as vendas são geralmente mais altas em um determinado mês, podemos intensificar as atividades de marketing nesse período para maximizar o faturamento.

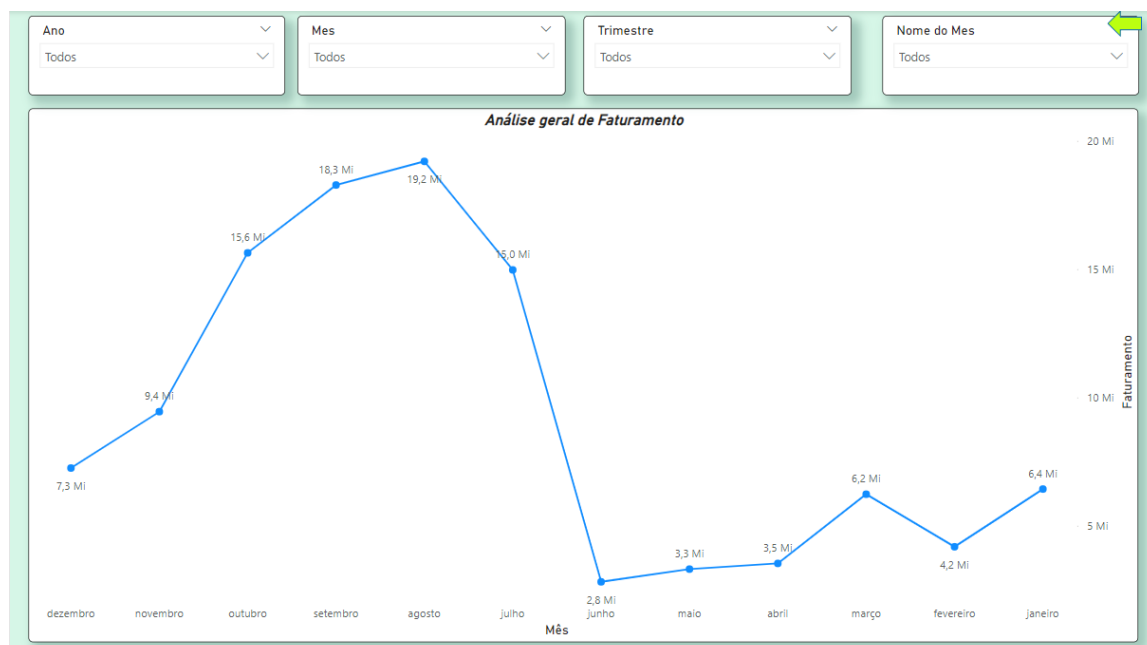


Figura 8: Análise Geral

7.1 Link do Dashboard no Power Bi

https://app.powerbi.com/links/NsXRRpBqCU?ctid=5105be30-82bb-4e62-a669-660388b53927pbi_source=linkSharebookmarkGuid=c7268ddd-4525-4026-95c8-a6cd856554d9