

MVP de Engenharia de Dados

Prof. Victor Almeida

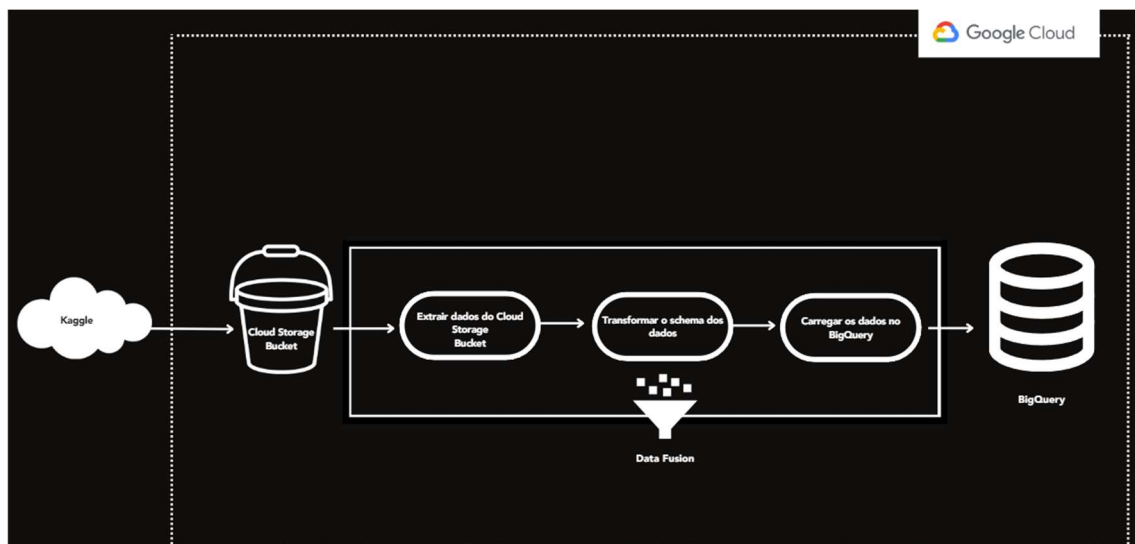
Aluna: Carmem Maria Marinho de Souza Vasconcelos

Introdução

Este documento apresenta parte do MVP de Engenharia de Dados, referente a Sprint 3 do curso de Pós-Graduação em Ciências de Dados e Analytics, da PUC-Rio.

O MVP consiste em construir um pipeline de dados utilizando tecnologias na nuvem. O pipeline irá envolver a busca, coleta, modelagem, carga e análise dos dados.

Nesta Parte 1, serão expostos a busca, coleta, modelagem e carga, conforme imagem abaixo.



A Análise (qualidade dos dados e solução do problema) e Autoavaliação estão na Parte 2 do MVP, apresentado no notebook do Colab, que pode ser acessado no link abaixo:

https://github.com/CarmemMarinho/poscienciasdedados/blob/main/MVP_Srint3_EngenhariaDeDados_Parte2.ipynb.

1. Objetivo

Uma das etapas mais desafiadoras na jornada do empreendedorismo reside no planejamento inicial. Estabelecer o produto/serviço a ser oferecido e identificar o público-alvo demanda uma extensa pesquisa e uma atenção minuciosa, uma vez que esses são fatores críticos para determinar o êxito da empresa.

Todos os dias, muitas pessoas procuram o Sebrae buscando apoio na fase de planejamento do negócio. Há pouco tempo, durante meu trabalho no Sebrae, um dos clientes comentou que gostaria de trabalhar com gênero masculino e perguntou qual seria o melhor produto/serviço para empreender.

A partir desse questionamento senti a falta de um trabalho no Sebrae focado na análise do comportamento do consumidor, que pudesse responder à pergunta do cliente e oferecer suporte na tomada de decisão durante o planejamento do empreendimento. Então, defini como objetivo desse MPV analisar o perfil dos consumidores masculinos e encontrar o segmento ideal para investir na abertura de uma empresa. Para ajudar nessa análise, procurou-se responder as seguintes perguntas:

1. Como a idade do cliente influencia no seu comportamento de compra?
2. Existem padrões discerníveis nas vendas em diferentes períodos de tempo?
3. Quais são as relações entre idade e gastos dos clientes?

Especificamente, vamos buscar responder também as questões abaixo:

- Qual a faixa etária dos clientes por categoria do produto?
- Qual a idade dos clientes que compram mais em quantidade por categoria?
- Qual a idade dos clientes que compram mais em valor por categoria?
- Qual o mês com maior faturamento por categoria?
- Qual o mês com maior quantidade de produtos vendidos por categoria?
- Quais meses do ano apresentaram queda no volume de vendas (quantidade)?
- Quais meses do ano apresentaram queda no volume de vendas (faturamento)?
- Qual a categoria de produto mais vendido em quantidade no período total?
- Qual a categoria de produto com maior faturamento no período total?

2. Busca e Coleta de Dados

Os dados utilizados neste projeto estão presentes no Dataset Retails, obtidos no Kaggle (www.kaggle.com), de domínio público. Ele apresenta informações de varejo e as interações com os clientes, o que possibilita revelar tendências de varejo, por meio da análise dos padrões de vendas e perfis de clientes.

Embora seja um dataset com dados de um cenário fictício, ele reflete cenários de varejo do mundo real, permitindo praticar a análise em um contexto familiar.

Inclui os seguintes atributos: ID da transação, data, ID do cliente, sexo, idade, categoria do produto, quantidade, preço por unidade e valor total. Esses atributos permitem uma exploração de tendências de vendas, influências demográficas e comportamentos de compra.

Para mais detalhes sobre este dataset, consulte: Retails Sales: <https://www.kaggle.com/datasets/mohammadtalib786/retail-sales-dataset>

O dataset possui 1000 linhas e 9 atributos.

Informações dos Atributos:

Transaction_ID: Um identificador exclusivo para cada transação

Date: Data em que a transação ocorreu

Customer_ID: Identificador exclusivo para cada cliente

Gender: O gênero do cliente

Age: Idade do cliente

Product_Category: Categoria do produto adquirido

Quantity: Quantidade de unidades do produto adquiridas

Price_per_Unit: Preço de uma unidade do produto

Total_Amount: Valor monetário total da transação

3. Modelagem de Dados

Os dados estão dispostos em um modelo flat, numa única tabela em um arquivo com formato csv.

Catálogo de Dados - Foi realizado utilizando os recursos do BigQuery. Durante a etapa de transformação, conforme apresentado mais abaixo, foram realizados filtro e exclusões de atributos. Dessa forma o catálogo de dados foi elaborado com as informações a seguir:

Date: Data em que a transação ocorreu.

Customer_ID: ID do cliente. Identificador exclusivo para cada cliente.

Age: Idade do cliente. Os valores são entre 18 e 64 anos.

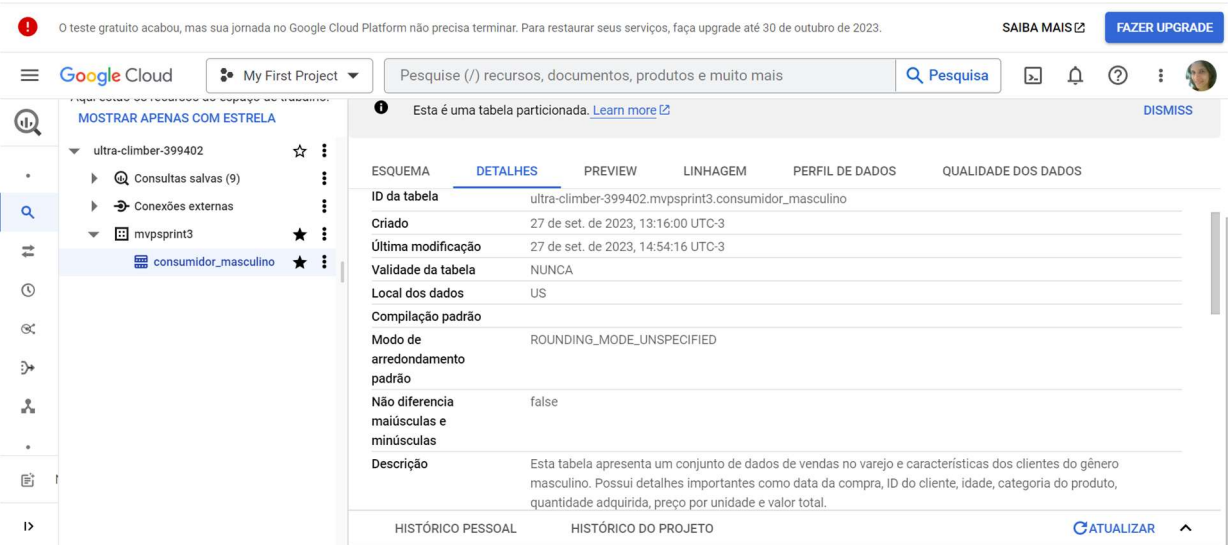
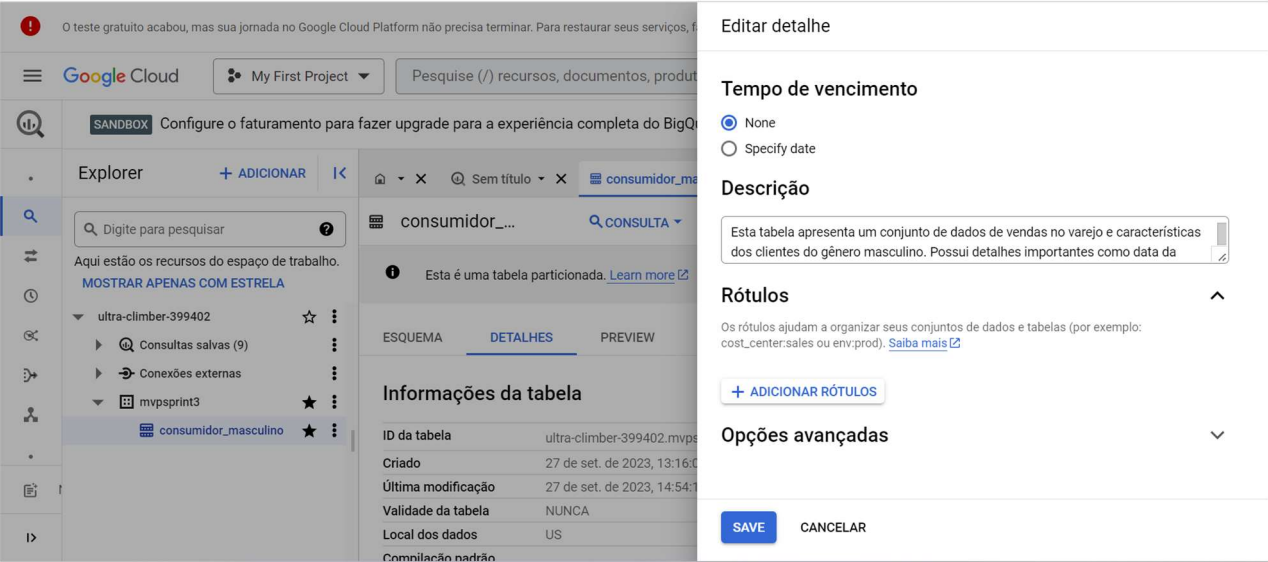
Product_Category: Categoria do produto adquirido. São os segmentos aos quais os produtos pertencem: Beauty, Clothing e Electronics.

Quantity: Quantidade de unidades do produto adquiridas. São valores positivos e diferentes de zero. Os valores são entre 1 e 4 unidades.

Price_per_Unit: Preço unitário do produto. São valores positivos e diferentes de zero.

Total_Amount: Valor monetário total da transação. São valores positivos e diferentes de zero.

A seguir, a imagem apresenta a inserção da descrição da tabela, para compor o catálogo de dados -



Abaixo, a imagem apresenta a inserção da descrição dos atributos, para compor o catálogo de dados -

Google Cloud

My Project

Pesquisa (/) recursos, documentos, produtos e muito mais

Pesquisa

SANDBOX

Configure o faturamento para fazer upgrade para a experiência completa do BigQuery

Saiba mais

DISPENSAR

FAZER UPGRADE

Explorer

Sem título

consumidor_masculino

CONSULTA

COMPARTILHAR

COPIAR

SNAPSHOT

EXCLUIR

EXPORTAR

ATUALIZAR

Esta é uma tabela particionada

Learn more

DISMISS

| ESQUEMA | DETALHES | PREVIEW | LINHAGEM | PERFIL DE DADOS | QUALIDADE DOS DADOS | | | |
|--------------------------|------------------|---------|----------|-----------------|---------------------|--------------|-------------------|--|
| <input type="checkbox"/> | Nome do campo | Tipo | Modo | Chave | Compilação | Valor padrão | Tags de políticas | Descrição |
| <input type="checkbox"/> | Date | STRING | NULLABLE | | | | | Data em que a transação ocorreu. |
| <input type="checkbox"/> | Customer_ID | STRING | NULLABLE | | | | | ID do cliente. Identificador exclusivo para cada cliente |
| <input type="checkbox"/> | Age | INTEGER | NULLABLE | | | | | Idade do cliente. Os valores são entre 18 e 64 anos. |
| <input type="checkbox"/> | Product_Category | STRING | NULLABLE | | | | | Categoria do produto adquirido. São os segmentos aos quais os produtos pertencem: Beauty, Clothing e Electronics. |
| <input type="checkbox"/> | Quantity | INTEGER | NULLABLE | | | | | Quantidade de unidades do produto adquiridas. São valores positivos e diferentes de zero. Os valores são entre 1 e 4 unidades. |
| <input type="checkbox"/> | Price_per_Unit | INTEGER | NULLABLE | | | | | Preço unitário do produto. São valores positivos e diferentes de zero. |
| <input type="checkbox"/> | Total_Amount | INTEGER | NULLABLE | | | | | Valor monetário total da transação. São valores positivos e diferentes de zero. |

EDITAR ESQUEMA

VER POLÍTICAS DE ACESSO DE LINHA

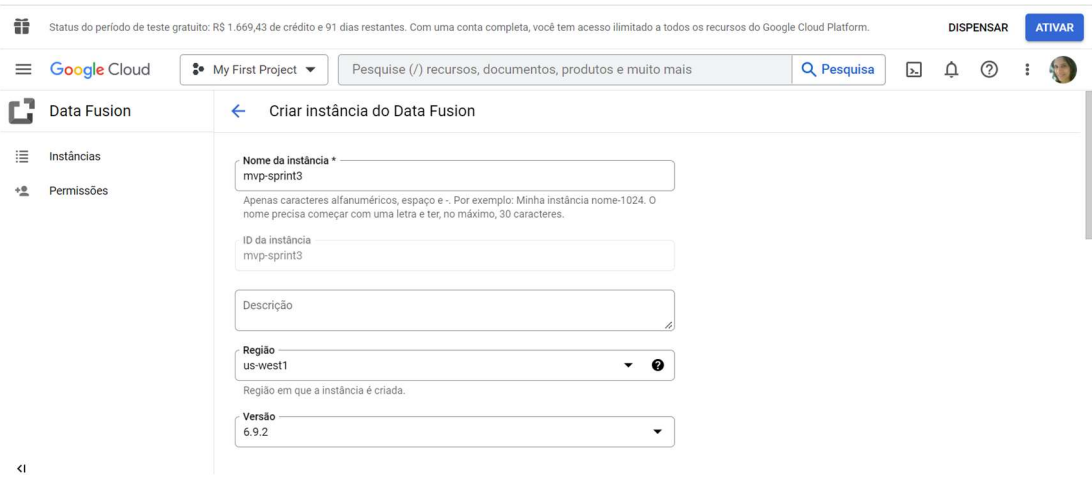
Linhagem dos Dados - Os dados foram obtidos no Kaggle (www.kaggle.com), realizando download do arquivo com formato csv, para o meu computador.

4. Carga

Para execução do trabalho foram utilizadas as ferramentas da Plataforma Google Cloud (GCP). Na extração, transformação e carga para o banco de dados, utilizei a ferramenta de ETL Data Fusion e para o Data Lake utilizei a plataforma BigQuery.

Iniciei criando a instância mvp-sprint3 no Data Fusion, conforme imagens abaixo.

Configurando a instância -



The screenshot shows the Google Cloud Data Fusion console. The main pane is titled 'Criar instância de Data Fusion'. It contains a form with the following fields: 'Nome da instância *' (mvp-sprint3), 'ID da instância' (mvp-sprint3), 'Descrição' (empty), 'Região' (us-west1), and 'Versão' (6.9.2). The 'Região' field has a dropdown arrow and a help icon. The 'Versão' field has a dropdown arrow. The 'Nome da instância' field has a note: 'Apenas caracteres alfanuméricos, espaço e -. Por exemplo: Minha instância nome-1024. O nome precisa começar com uma letra e ter, no máximo, 30 caracteres.'

Instância criada -

Status do período de teste gratuito: R\$ 1.669,43 de crédito e 91 dias restantes. Com uma conta completa, você tem acesso ilimitado a todos os recursos do Google Cloud Platform.

DISPENSAR

ATIVAR

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Data Fusion

Instâncias

criar instância

atualizar

excluir

fazer upgrade

Como foi sua experiência?

Instâncias

Permissões

Selecione a instância do Cloud Data Fusion que você quer ver.

| | | Nome da instância | Ação | Edição | Região | Zona | Versão | Criptografia | Criado |
|--------------------------|-------------------------------------|-------------------|-------------------------------|--------|----------|------|------------------------|------------------------|----------------|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | mvp-sprint3 | Ver instância | Basic | us-west1 | -- | 6.9.2 (latest version) | Gerenciada pelo Google | 17 de 20 23:50 |

Em seguida, para inserir os dados coletados na Plataforma Google Cloud, no serviço Cloud Storage, criei o bucket mvp-sprint3bd, para onde fiz o upload do dataset coletado, e o bucket temporário mvp-sprint3bd-temp.

Criando os buckets -

Status do período de teste gratuito: R\$ 1.669,43 de crédito e 91 dias restantes. Com uma conta completa, você tem acesso ilimitado a todos os recursos do Google Cloud Platform.

DISPENSAR

ATIVAR

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Cloud Storage

criar um bucket

Buckets

Monitoramento

Configurações

Marketplace

Notas de lançamento

☒ Nomeie seu bucket

Nome: mvp-sprint3bd

☐ Escolha onde armazenar seus dados

Essa escolha define a posição geográfica dos dados e afeta o custo, o desempenho e a disponibilidade. Ela não pode ser alterada. [Saiba mais](#)

Tipo de local

☒ Multi-region

Disponibilidade mais alta entre áreas maiores

us (várias regiões nos Estados Unidos)

☐ Dual-region

Alta disponibilidade e baixa latência em 2 regiões

☐ Region

Latência mais baixa em uma única região

Importante

Preços do local

As taxas de armazenamento variam dependendo da classe de armazenamento dos dados e da localização dos buckets. [Detalhes do preço](#)

Configuração atual: Multi-region / Standard

| Item | Custo |
|--|--------------------|
| us (várias regiões nos Estados Unidos) | \$0.026 por GB/mês |
| Com replicação padrão gravado | \$0.020 por GB |

ESTIMAR SEU CUSTO MENSAL

Status do período de teste gratuito: R\$ 1.669,43 de crédito e 91 dias restantes. Com uma conta completa, você tem acesso ilimitado a todos os recursos do Google Cloud Platform.

DISPENSAR

ATIVAR

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Cloud Storage

Detalhes do bucket

atualizar

saiba mais

Buckets

Monitoramento

Configurações

Marketplace

Notas de lançamento

mvp-sprint3bd

Local

us (várias regiões nos Estados Unidos)

Classe de armazenamento

Standard

Acesso público

Não público

Proteção

Nenhum

OBJETOS

CONFIGURAÇÃO

PERMISSÕES

PROTEÇÃO

CICLO DE VIDA

OBSERVABILIDADE

RELATÓRIOS DE INVENTÁRIO

Intervalos

mvp-sprint3bd

FAZER UPLOAD DE ARQUIVOS

CARREGAR PASTA

CRIAR PASTA

TRANSFERIR DADOS

GERENCIAR RETENÇÕES

FAZER O DOWNLOAD

EXCLUIR

Filtrar apenas pelo prefixo do nome

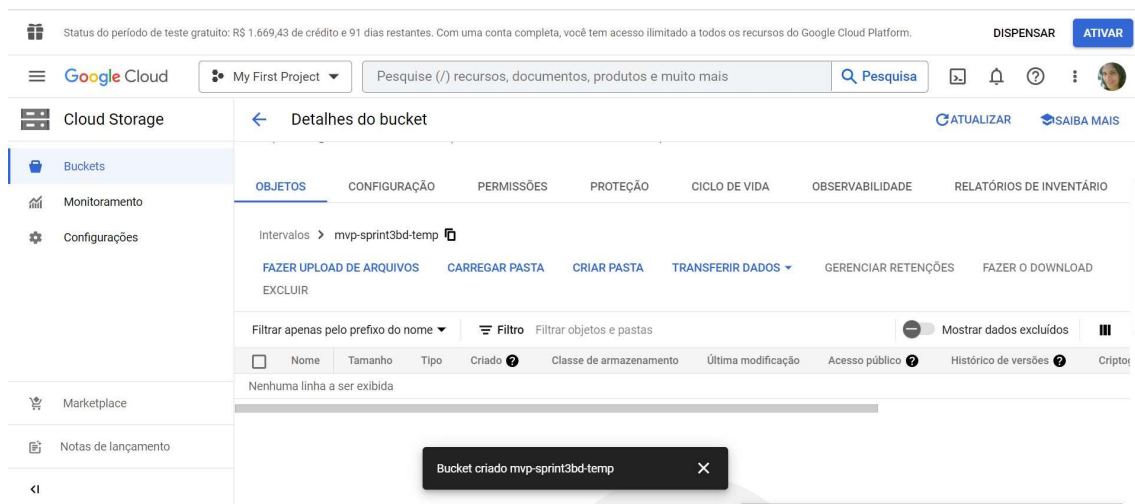
Filtro

Filtrar objetos e pastas

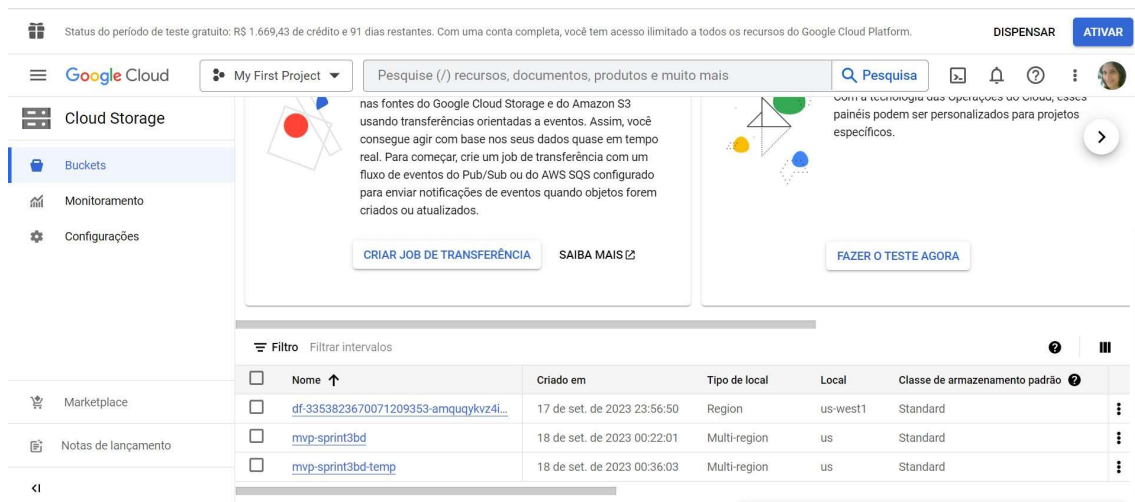
Mostrar dados excluídos

| <input type="checkbox"/> | Nome | Tamanho | Última modificação | Acesso público | Histórico de versões | Criptografia |
|-----------------------------|------|---------|--------------------|----------------|----------------------|--------------|
| Nenhuma linha a ser exibida | | | | | | |

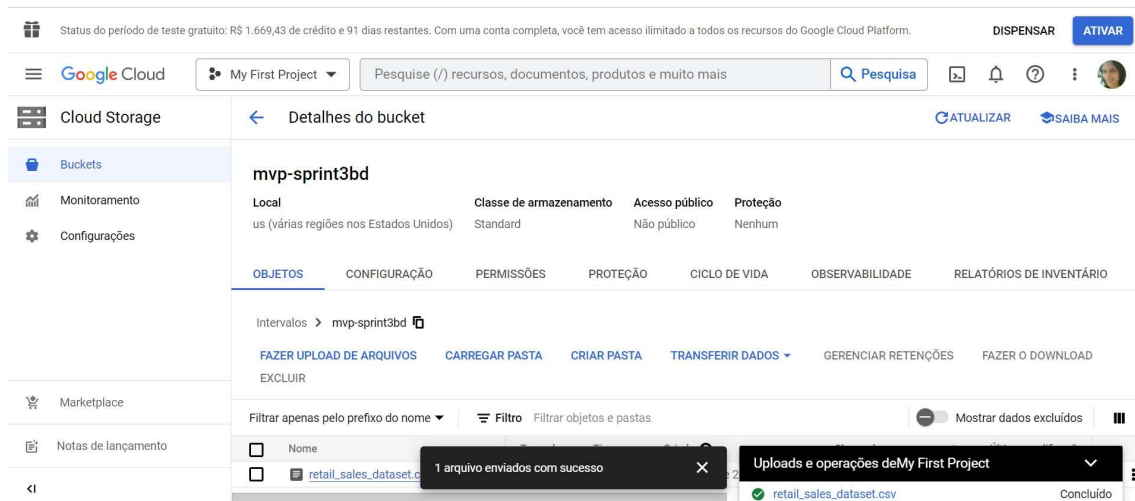
Bucket criado mvp-sprint3bd



Buckets criados -



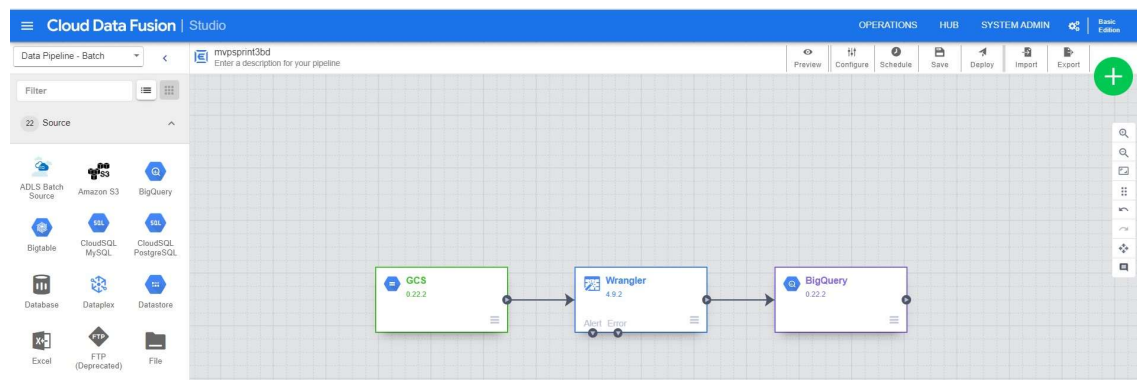
A imagem abaixo mostra o upload do arquivo “reatails_sales_dataset” para o bucket mvp-sprint3bd.



Finalizando a etapa de upload dos dados, segui para a criação do dataset mvpsprint3 no BigQuery, conforme imagem a seguir:

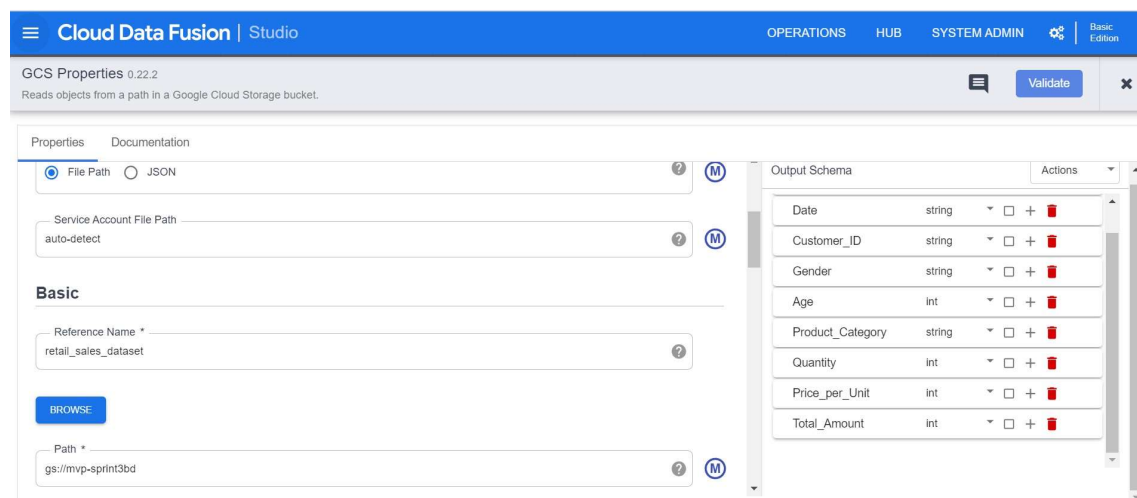


O próximo passo, através do Data Fusion Studio, foi criar as etapas do pipeline mvpsprint3bd: GCS, Wrangler e BigQuery



As imagens abaixo mostram as definições das propriedades das etapas do fluxo, GCS, Wrangler e BigQuery.

Primeiro, foi realizada a configuração do GCS para extrair os dados da fonte, que no caso foi o arquivo "reatails_sales_dataset" do bucket "mvp-sprint3bd".

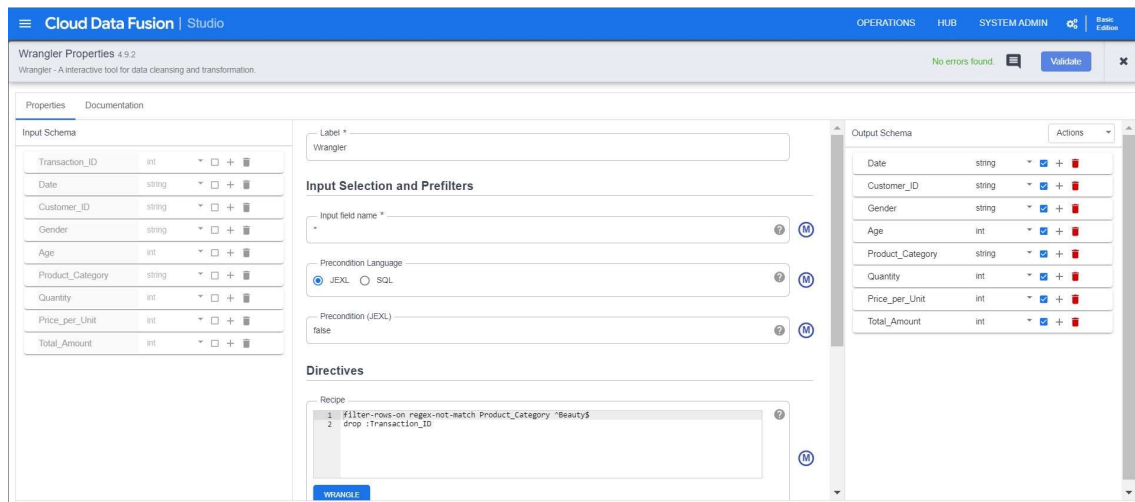


Em seguida, no Wrangler, foram realizadas as transformações dos dados:

1. Exclusão do atributo “Transaction_ID”
2. No atributo “Gender”, foi feito o filtro “Male”
3. Exclusão do atributo “Gender”

Não foi necessária nenhuma ação de limpeza de dados, pois, como será apresentado no tópico de qualidade de dados, os dados não apresentavam erros, valores ausentes, nem duplicatas.

Configuração do wrangler -



Realizando as transformações dos dados no wrangler -

Cloud Data Fusion | Studio

Wrangle

Cloud Storage Default - mvp-sprint3bdr/retail_sales_dataset.csv
retail_sales_dataset.csv Columns: 7 | Rows: 490

Columns (7)

| # | Transformation |
|---|--|
| 1 | filter-rows-on regex-not-match Gender 'Male\$' |
| 2 | drop :Transaction_ID |
| 3 | drop :Gender |

Transformation steps (3)

[Apply] [More]

| # | Date | Customer_ID | Age | Product_Category | Quantity | Price_per_Unit | Total_Amount |
|----|------------|-------------|-----|------------------|----------|----------------|--------------|
| 1 | 2023-11-24 | CUST001 | 34 | Beauty | 3 | 50 | 150 |
| 2 | 2023-01-13 | CUST003 | 50 | Electronics | 1 | 30 | 30 |
| 3 | 2023-05-21 | CUST004 | 37 | Clothing | 1 | 500 | 500 |
| 4 | 2023-05-06 | CUST005 | 30 | Beauty | 2 | 50 | 100 |
| 5 | 2023-03-13 | CUST007 | 46 | Clothing | 2 | 25 | 50 |
| 6 | 2023-02-22 | CUST008 | 30 | Electronics | 4 | 25 | 100 |
| 7 | 2023-12-13 | CUST009 | 63 | Electronics | 2 | 300 | 600 |
| 8 | 2023-02-14 | CUST011 | 23 | Clothing | 2 | 50 | 100 |
| 9 | 2023-10-30 | CUST012 | 35 | Beauty | 3 | 25 | 75 |
| 10 | 2023-08-05 | CUST013 | 22 | Electronics | 3 | 500 | 1500 |
| 11 | 2023-01-17 | CUST014 | 64 | Clothing | 4 | 30 | 120 |
| 12 | 2023-02-17 | CUST016 | 19 | Clothing | 3 | 500 | 1500 |

Após as transformações, a tabela ficou com 490 linhas e 7 atributos.

E por último, foi configurado os parâmetros necessários para carregamentos dos dados transformados no banco de dados BigQuery. A tabela carregada foi nomeada de “consumidor_masculino”, conforme imagem abaixo:

Cloud Data Fusion | Studio

BigQuery Properties 0.22.2

This sink writes to a BigQuery table. BigQuery is Google's serverless, highly scalable, enterprise data warehouse. Data is first written to a temporary location on Google Cloud Storage, then loaded into BigQuery from there.

Properties | Documentation

Input Schema

| Field | Type | Required | Nullable | Drop |
|------------------|--------|----------|----------|------|
| Date | string | ✓ | ✓ | ✕ |
| Customer_ID | string | ✓ | ✓ | ✕ |
| Age | int | ✓ | ✓ | ✕ |
| Product_Category | string | ✓ | ✓ | ✕ |
| Quantity | int | ✓ | ✓ | ✕ |
| Price_per_Unit | int | ✓ | ✓ | ✕ |
| Total_Amount | int | ✓ | ✓ | ✕ |

Basic

Reference Name: (Name used to identify this sink for lineage)

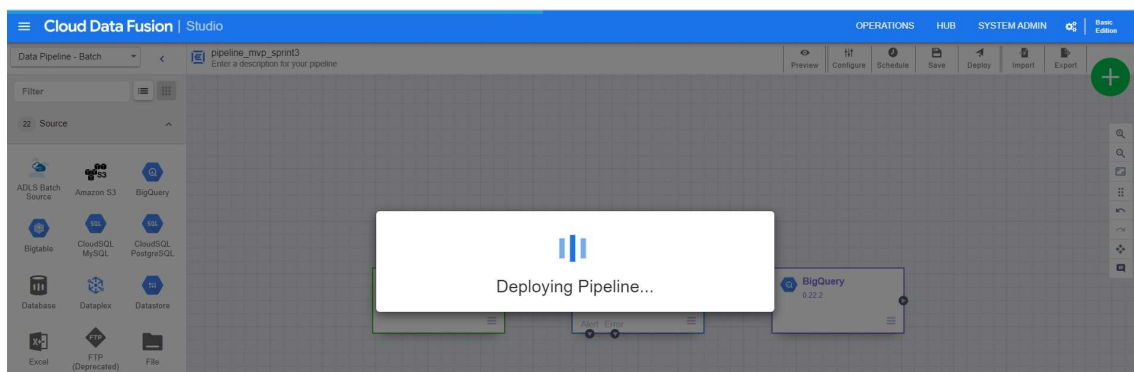
BROWSE

Dataset: * mvpsprint3

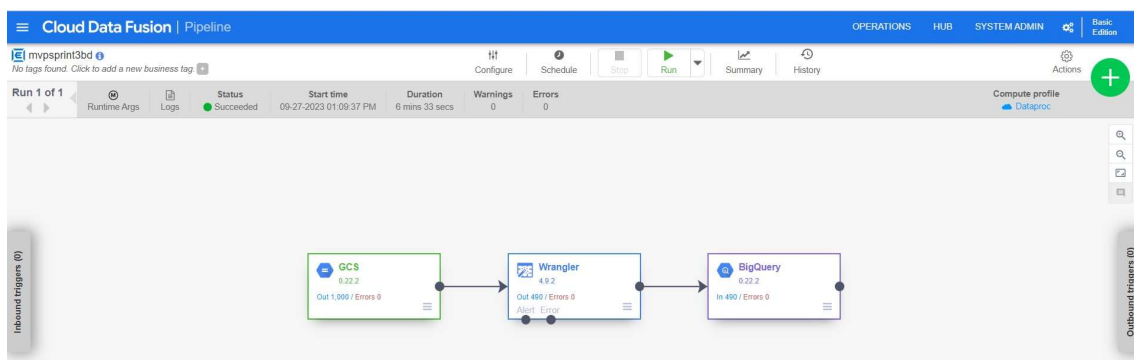
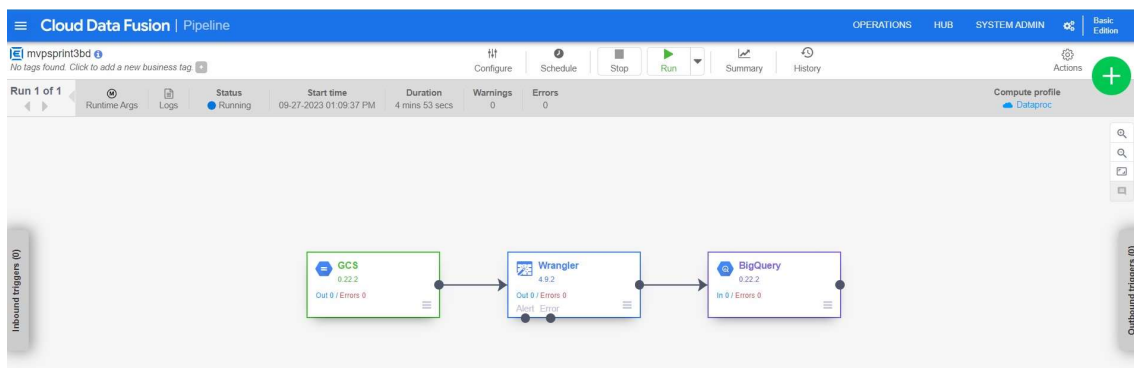
Table: * consumidor_masculino

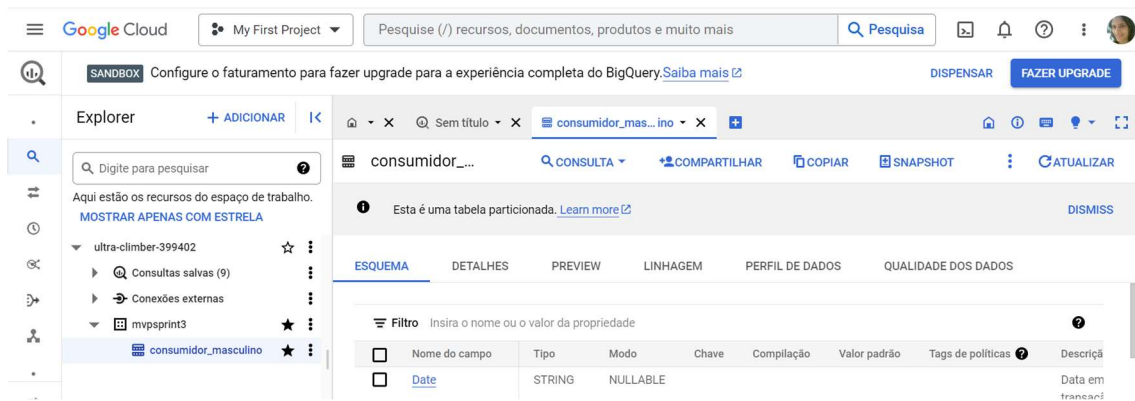
Temporary Bucket Name: mvpsprint3bd-temp

O passo seguinte foi fazer o “deploy” do fluxo, conforme imagem abaixo:



E, finalizado o deploy, foi acionado do RUN para realizar a carga dos dados para o BigQuery.

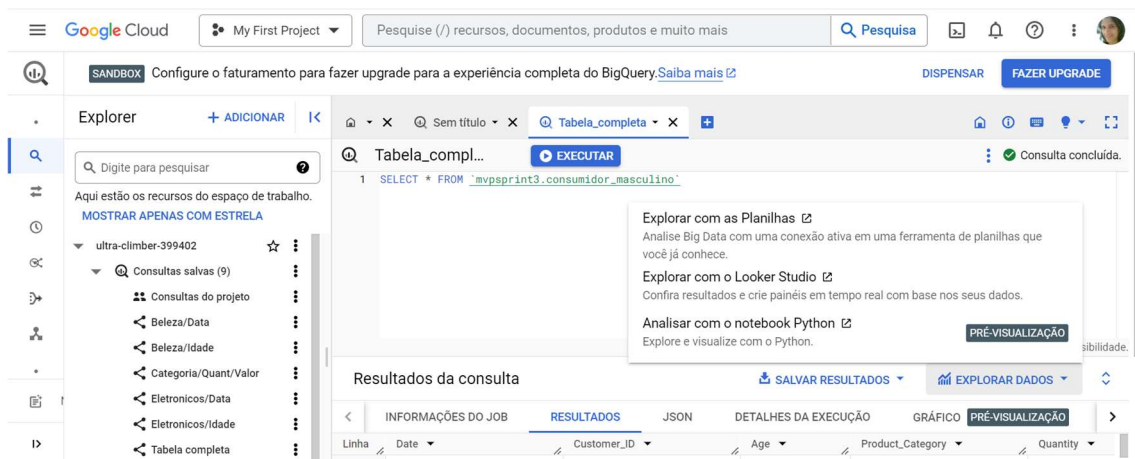




As etapas de Análise (qualidade dos dados e solução do problema) e Autoavaliação estão na Parte 2 do MVP, apresentado no notebook do Colab, que pode ser acessado no link:

https://github.com/CarmemMarinho/poscienciasdedados/blob/main/MVP_Srint3_EngenhariaDeDados_Parte2.ipynb

Para acessar os dados no Colab, utilizei o recurso disponível no BigQuery de “Analisar com notebook Python”, como pode ser visto na imagem abaixo.



Entretanto, para aprendizado e complemento do trabalho, também fiz algumas análises no BigQuery utilizando SQL, conforme apresentado a seguir.

Consulta 1 – Quantidade total de produtos vendidos e faturamento total por categoria

Faça upgrade da sua conta para evitar interrupções no serviço. Você tem R\$ 274,04 de crédito e 80 dias restantes na seu teste gratuito.

SAIBA MAIS

FAZER UPGRADE

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Explorer

+ ADICIONAR

Q Digite para pesquisar

Aqui estão os recursos do espaço de trabalho.

MOSTRAR APENAS COM ESTRELA

ultra-climber-399402

Consultas salvas (9)

Consultas do projeto

Beleza/Data

Beleza/Idade

Categoria/Quant/Valor

Eletronicos/Data

Eletronicos/Idade

Tabela_completa

Categoria/Quant/lor

Eletronicos/Data

Vestuário/Data

Categoria/Qu...

EXECUTAR

1 SELECT Product_Category, SUM(Quantity) AS Quantidade ,SUM(Total_Amount) AS Valor FROM `mvpsprint3.consumidor_masculino` GROUP BY Product_Category

Pressione Alt+F1 para abrir as opções de acessibilidade.

Resultados da consulta

SALVAR RESULTADOS

EXPLORAR DADOS

INFORMAÇÕES DO JOB

RESULTADOS

JSON

DETALHES DA EXECUÇÃO

GRÁFICO

PRÉ-VISUALIZAÇÃO

| Linha | Product_Category | Quantidade | Valor |
|-------|------------------|------------|-------|
| 1 | Beauty | 353 | 68685 |
| 2 | Clothing | 453 | 74305 |
| 3 | Electronics | 410 | 80170 |

Consulta 2 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por idade, para a categoria “Beleza” (Beauty)

Faça upgrade da sua conta para evitar interrupções no serviço. Você tem R\$ 274,04 de crédito e 80 dias restantes na seu teste gratuito.

SAIBA MAIS

FAZER UPGRADE

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Explorer

+ ADICIONAR

Q Digite para pesquisar

Aqui estão os recursos do espaço de trabalho.

MOSTRAR APENAS COM ESTRELA

ultra-climber-399402

Consultas salvas (9)

Consultas do projeto

Beleza/Data

Beleza/Idade

Categoria/Quant/Valor

Eletronicos/Data

Eletronicos/Idade

Tabela completa

Tabela_completa

Vestuário/Data

Beleza/Idade

EXECUTAR

SALVAR CONSULTA

COMPARTILHAR

PROGRAMAÇÃO

Consulta

1 SELECT DISTINCT Age, COUNT(Customer_ID) AS Quant_Clientes, SUM (Quantity) AS Quant_Produtos, SUM (Total_Amount) AS Valor_total

2 FROM `mvpsprint3.consumidor_masculino`

3 WHERE Product_Category = 'Beauty'

4 GROUP BY Age

5 ORDER BY Valor_total DESC

Pressione Alt+F1 para abrir as opções de acessibilidade.

Resultados da consulta

SALVAR RESULTADOS

EXPLORAR DADOS

INFORMAÇÕES DO JOB

RESULTADOS

JSON

DETALHES DA EXECUÇÃO

GRÁFICO

PRÉ-VISUALIZAÇÃO

| Linha | Age | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|-----|----------------|----------------|-------------|
| 1 | 46 | 6 | 19 | 6160 |
| 2 | 21 | 4 | 13 | 4700 |
| 3 | 22 | 4 | 11 | 4230 |
| 4 | 45 | 6 | 15 | 3840 |
| 5 | 27 | 6 | 15 | 3455 |
| 6 | 57 | 4 | 14 | 3200 |

Consulta 3 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por idade, para a categoria “Vestuário” (Clothing)

Faça upgrade da sua conta para evitar interrupções no serviço. Você tem R\$ 274,04 de crédito e 80 dias restantes na seu teste gratuito.

SAIBA MAIS

FAZER UPGRADE

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Explorer

+ ADICIONAR

Q Digite para pesquisar

Aqui estão os recursos do espaço de trabalho.

MOSTRAR APENAS COM ESTRELA

ultra-climber-399402

Consultas salvas (9)

Consultas do projeto

Beleza/Data

Beleza/Idade

Categoria/Quant/Valor

Eletronicos/Data

Vestuário/Idade

EXECUTAR

Esta consulta processará 20,81 KB quando executada.

1 SELECT DISTINCT Age, COUNT(Customer_ID) AS Quant_Clientes, SUM (Quantity) AS Quant_Produtos, SUM (Total_Amount) AS

2 FROM `mvpsprint3.consumidor_masculino`

3 WHERE Product_Category = 'Clothing'

4 GROUP BY Age

5 ORDER BY Valor_total DESC

Pressione Alt+F1 para abrir as opções de acessibilidade.

Resultados da consulta

SALVAR RESULTADOS

EXPLORAR DADOS

INFORMAÇÕES DO JOB

RESULTADOS

JSON

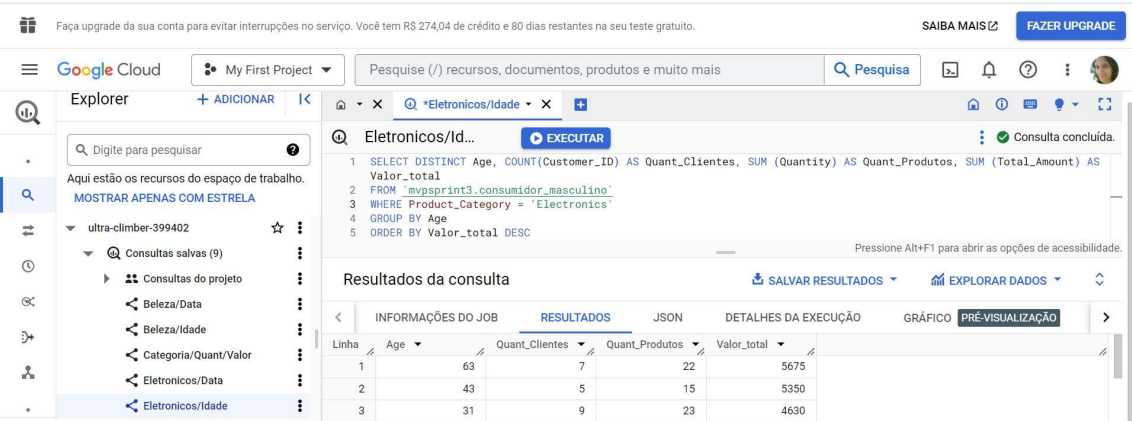
DETALHES DA EXECUÇÃO

GRÁFICO

PRÉ-VISUALIZAÇÃO

| Linha | Age | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|-----|----------------|----------------|-------------|
| 5 | 54 | 6 | 19 | 2950 |
| 6 | 32 | 7 | 15 | 2700 |

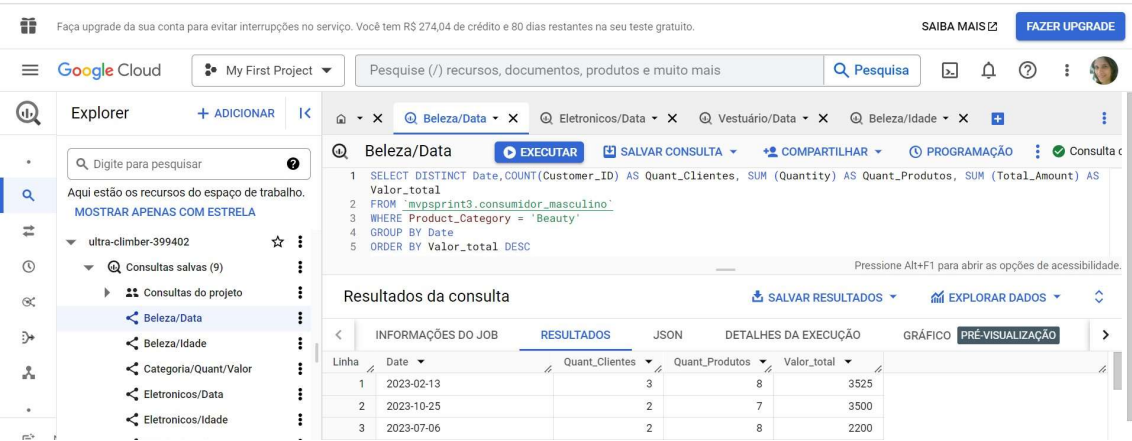
Consulta 4 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por idade, para a categoria “Eletrônicos” (Electronics)



Google Cloud interface showing the results of Query 4. The query filters for the 'Electronics' category and groups results by age. The results table shows three rows for different age groups: 63, 43, and 31.

| Linha | Age | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|-----|----------------|----------------|-------------|
| 1 | 63 | 7 | 22 | 5675 |
| 2 | 43 | 5 | 15 | 5350 |
| 3 | 31 | 9 | 23 | 4630 |

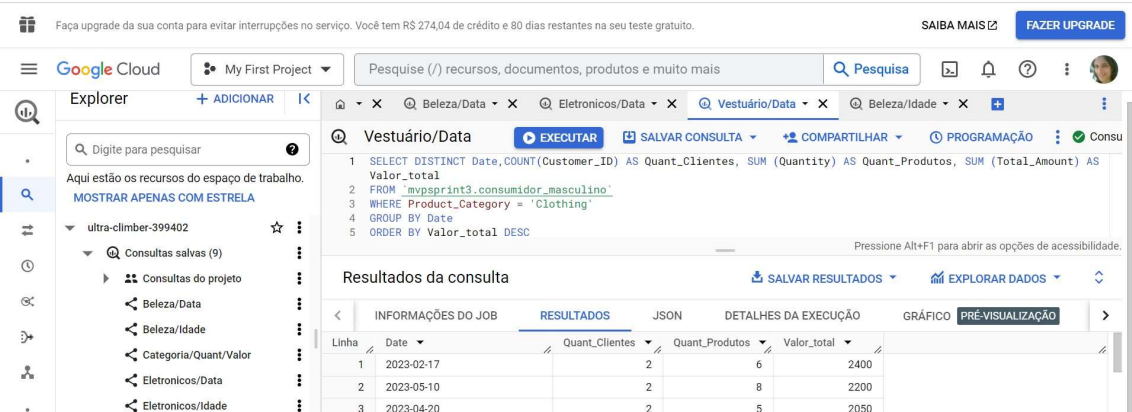
Consulta 5 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por data, para a categoria “Beleza” (Beauty)



Google Cloud interface showing the results of Query 5. The query filters for the 'Beauty' category and groups results by date. The results table shows three rows for different dates: 2023-02-13, 2023-10-25, and 2023-07-06.

| Linha | Date | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|------------|----------------|----------------|-------------|
| 1 | 2023-02-13 | 3 | 8 | 3525 |
| 2 | 2023-10-25 | 2 | 7 | 3500 |
| 3 | 2023-07-06 | 2 | 8 | 2200 |

Consulta 6 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por data, para a categoria “Vestuário” (Clothing)



Google Cloud interface showing the results of Query 6. The query filters for the 'Clothing' category and groups results by date. The results table shows three rows for different dates: 2023-02-17, 2023-05-10, and 2023-04-20.

| Linha | Date | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|------------|----------------|----------------|-------------|
| 1 | 2023-02-17 | 2 | 6 | 2400 |
| 2 | 2023-05-10 | 2 | 8 | 2200 |
| 3 | 2023-04-20 | 2 | 5 | 2050 |

Consulta 7 – Quantidade total de produtos vendidos, quantidade total de clientes e faturamento total por data, para a categoria “Eletrônicos” (Electronics)

Faça upgrade da sua conta para evitar interrupções no serviço. Você tem R\$ 274,04 de crédito e 80 dias restantes na seu teste gratuito.

SAIBA MAIS

FAZER UPGRADE

Google Cloud

My First Project

Pesquise (/) recursos, documentos, produtos e muito mais

Pesquisa

Explorer

+ ADICIONAR

1

Digite para pesquisar

Aqui estão os recursos do espaço de trabalho.

MOstrar APENAS COM ESTRELA

ultra-climber-399402

Consultas salvas (9)

Consultas do projeto

Beleza/Data

Beleza/Idade

Categoria/Quant/Valor

Eletronicos/Data

Eletronicos/Idade

Tabela completa

Tabela_completa

Vestuário/Data

Beleza/Data

Eletronicos/Data

Vestuário/Data

Beleza/Idade

Eletronicos/D...

EXECUTAR

Consulta concluída.

1 SELECT DISTINCT Date,COUNT(Customer_ID) AS Quant_Clientes, SUM (Quantity) AS Quant_Produtos, SUM (Total_Amount) AS Valor_total

2 FROM `mvpsprint3.consumidor_masculino`

3 WHERE Product_Category = 'Electronics'

4 GROUP BY Date

5 ORDER BY Valor_total DESC

Pressione Alt+F1 para abrir as opções de acessibilidade.

Resultados da consulta

SALVAR RESULTADOS

EXPLORAR DADOS

INFORMAÇÕES DO JOB

RESULTADOS

JSON

DETALHES DA EXECUÇÃO

GRÁFICO

PRÉ-VISUALIZAÇÃO

| Linha | Date | Quant_Clientes | Quant_Produtos | Valor_total |
|-------|------------|----------------|----------------|-------------|
| 1 | 2023-06-24 | 2 | 8 | 4000 |
| 2 | 2023-05-16 | 3 | 11 | 3100 |
| 3 | 2023-05-08 | 2 | 7 | 2075 |
| 4 | 2023-05-17 | 2 | 5 | 2050 |
| 5 | 2023-12-09 | 2 | 5 | 2025 |
| 6 | 2023-12-05 | 1 | 4 | 2000 |
| 7 | 2023-05-08 | 1 | 4 | 2000 |