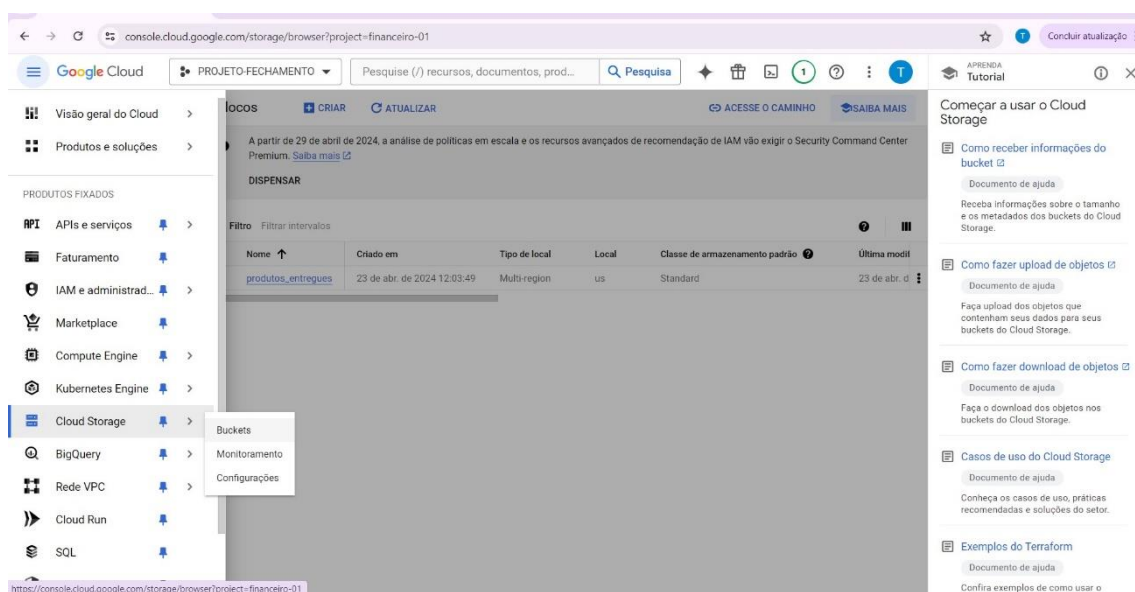
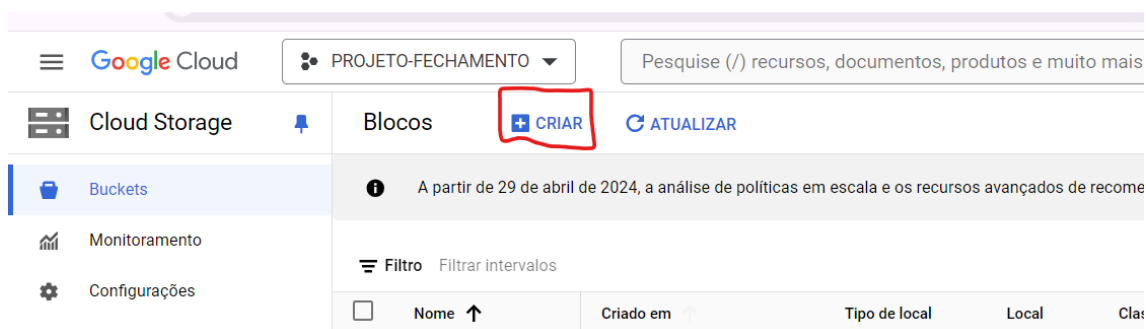


Criando um Delta Lake no GCP e subindo as informações no BIGQUERY

1- Criando um bucket na GCP Coud Storage -> Bucket



2 – Após acessar, iremos configurar algumas informações clicando em CRIAR e depois acessar a parte de configuração do bucket.



3 – Configurações informações adicionais. Nessa etapa apenas acrescentei o nome do bucket e cliquei em criar.

Google Cloud

PROJETO-FECHAMENTO

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

Pesquisa

Cloud Storage

Buckets

Monitoramento

Configurações

Marketplace

Notas de lançamento

Nomeie seu bucket

Escolha um nome definitivo exclusivo. [Diretrizes de nomenclatura](#)

produtos_entregues

Dica: não inclua informações confidenciais

MARCADORES (OPCIONAL)

CONTINUAR

Escolha onde armazenar seus dados

Local: us (várias regiões nos Estados Unidos)

Tipo de local: Multi-region

Escolha uma classe de armazenamento para seus dados

Classe de armazenamento padrão: Standard

Escolha como controlar o acesso a objetos

Prevenção de acesso público: Ativada

Controle de acesso: Uniforme

Escolher como proteger os dados do objeto

Política de exclusão reversível: Ativado

Controle de versões de objetos: Desativado

Política de retenção de buckets: Desativado

Retenção de objetos: Desativado

Tipo de criptografia: Gerenciado pelo Google

CRIAR

CANCELAR

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	\$0.020 por GB gravado

ESTIMAR SEU CUSTO MENSAL

5- Bucket criado, o próximo passo será criar a camada bronze,silver e gold, assim irei utilizando o modelo Delta Lake para armazenar os dados

Cloud Storage

Buckets

Monitoramento

Configurações

Blocos

CRIAR

ATUALIZAR

ACESSO O CAMINHO

SAIBA MAIS

A partir de 29 de abril de 2024, a análise de políticas em escala e os recursos avançados de recomendação de IAM vão exigir o Security Command Center Premium. [Saiba mais](#)

DISPENSAR

Filtro

Filtrar intervalos

<input type="checkbox"/> Nome	Criado em	Tipo de local	Local	Classe de armazenamento padrão	Última modificação	Acesso público	Controle de acesso	Proteção	Ref
<input type="checkbox"/> produtos_entregues	23 de abr. de 2024 12:03:49	Multi-region	us	Standard	23 de abr. de 2024 12:03:49	Não público	Uniforme	Fazer a exclusão reversível	Nen

6 - Basta agora clicar em cima do bucket, acessar ele e criar as pastas, na opções “CRIAR PASTA” e logo em seguida especificar o nome da cada pasta.

Google Cloud

PROJETO-FECHAMENTO

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

Pesquisa

Cloud Storage

Buckets

Monitoramento

Configurações

Detalhes do bucket

ACESSO O CAMINHO

ATUALIZAR

SAIBA MAIS

produtos_entregues

Local

us (várias regiões nos Estados Unidos)

Classe de armazenamento

Standard

Acesso público

Não público

Proteção

Fazer a exclusão reversível

OBJETOS

CONFIGURAÇÃO

PERMISSÕES

PROTEÇÃO

CICLO DE VIDA

OBSERVABILIDADE

RELATÓRIOS DE INVENTÁRIO

OPERAÇÕES

Intervalos > produtos_entregues

FAZER UPLOAD DE ARQUIVOS

CARREGAR PASTA

CRIAR PASTA

TRANSFERIR DADOS

GERENCIAR RETENÇÕES

EDITAR RETENÇÃO

FAZER O DOWNLOAD

EXCLUIR

Criar pasta

O Cloud Storage usa um namespace simples. Portanto, fazer upload de objetos para uma pasta simplesmente anexa um prefixo de caminho aos nomes dos objetos. [Saiba mais](#)

Nome *

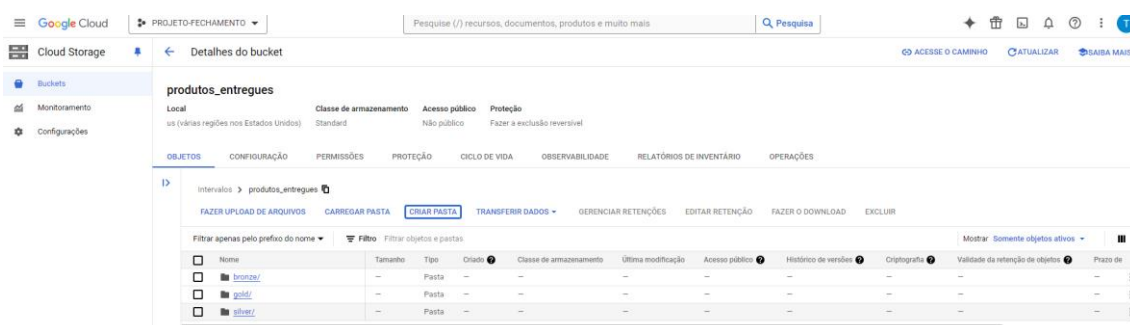
CANCELAR CRIAR

7 – Camadas dentro do Bucket criada.

Camada bronze – Dados Brutos

Camada Silver – Dados tratados

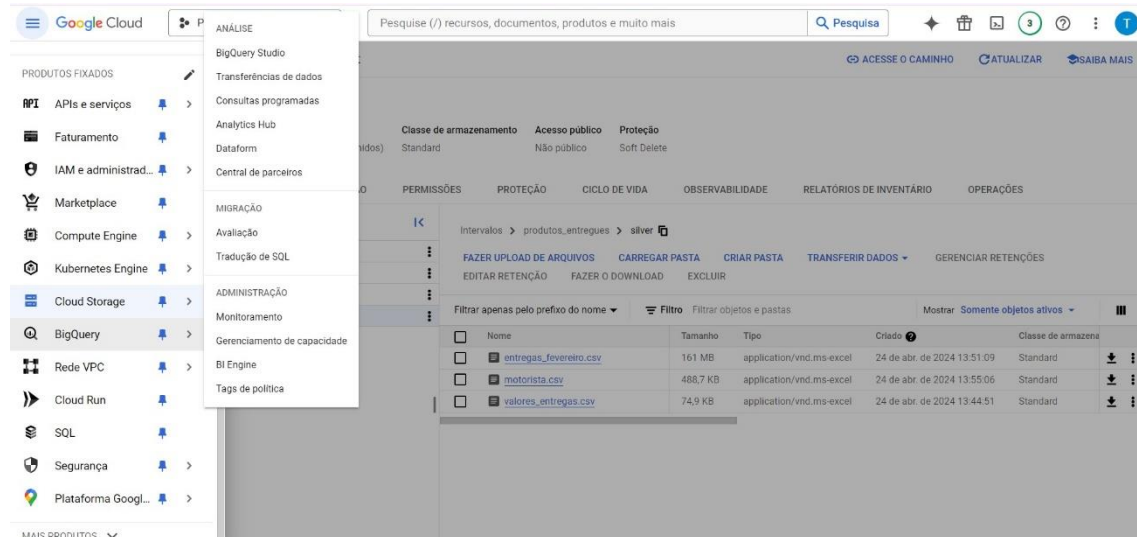
Camada Gold – Relatórios, query salvas, dashboard...



8 – Com o Storage criado para receber, a próxima etapa é tratar os dados no jupyter notebook e salvar diretamente no bucket. Todo os trabalho de tratamento dos dados está na pasta do github

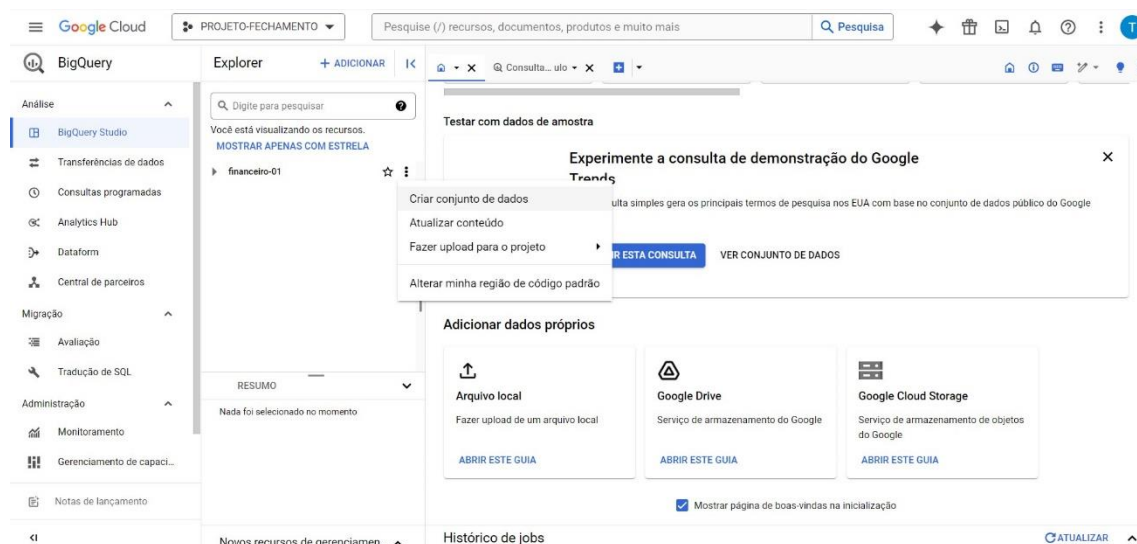
9 – Agora é o momento de pegar os dados que estão na camada Silver e subir no bigquery, mas antes irei criar um conjunto de dados para receber as informações.

Acessando o bigquery – Bigquery -> Bigquery Studio

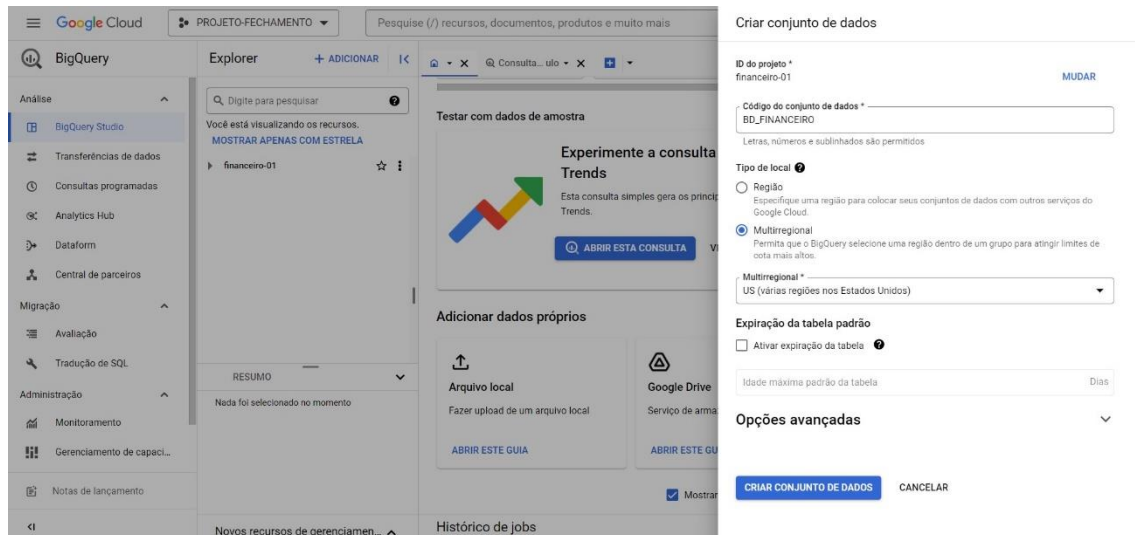


10 – Criando um conjunto de dados depois criar as tabelas com os dados.

Clicando em três pontinhos e depois em criar conjunto de dados.

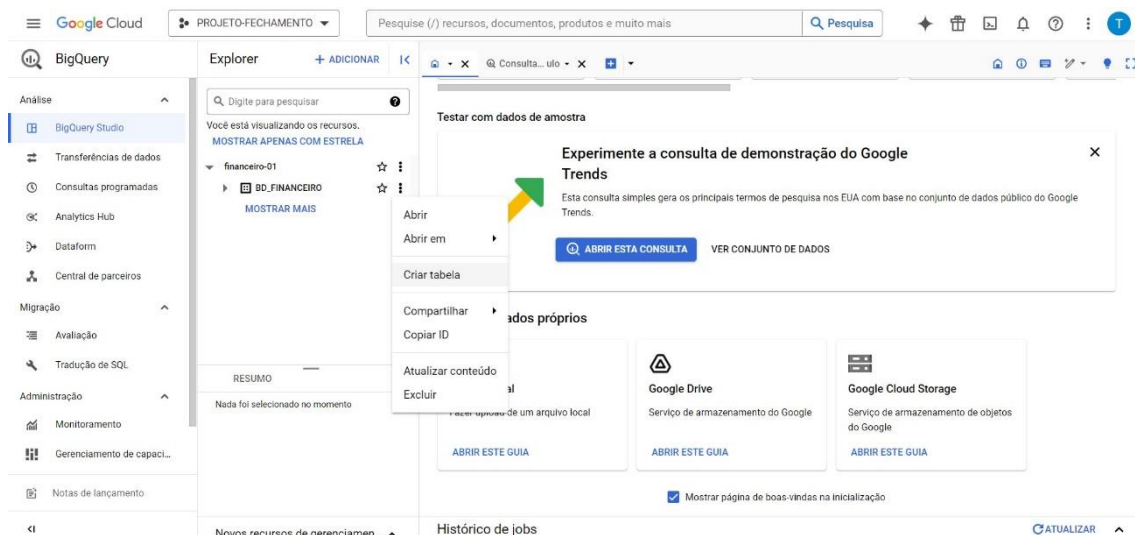


11 – Será necessário especificar o nome do conjunto de dados, a região. Somente essas duas configurações que coloquei e cliquei em Criar conjunto de dados.



12 – Após a criação do conjunto de dados, o bigquery já está preparado para a criação das tabelas, onde iremos realizar consulta. Então iremos para a criação.

Três pontinhos -> criar tabela



13 – Nessa etapa, iremos especificar o local que estão os dados, o tipo de arquivo que os dados estão armazenados , e em qual conjunto de dados ele irá fazer parte e também o schema da tabela (que pode ser de forma automática ou manual)

Esse procedimento irei fazer para a criação de 3 tabelas.

Origem

Criar tabela de Google Cloud Storage

Selecione o arquivo do bucket do GCS ou use um padrão de URI

produtos_entregues/silver/entregas_fevereiro.csv

Formato do arquivo

CSV

☐ Particionamento de dados de origem

Destino

Projeto *

financeiro-01

Conjunto de dados *

BD_FINANCEIRO

Tabela *

PRODUTOS_ENTREGUES

O tamanho máximo do nome é de 1.024 bytes UTF-8. Letras Unicode, marcas, números, conectores, traços e espaços são permitidos.

Tipo de tabela

Tabela nativa

Esquema

CRIAR TABELA CANCELAR

14 - Após esse procedimentos, temos um conjunto de dados com 3 tabelas, onde irei utilizar query SQL para responder algumas perguntas sobre essa base de dados.

15 – Abaixo estão os atributos de cada tabela.

financeiro-01 X *PERGUNTAS X *query_01 X

MOTORISTA CONSULTA COMPARTILHAR

ESQUEMA DETALHES VISUALIZAÇÃO LINHAGEM

Filtro Insira o nome ou o valor da propriedade

	Nome do campo	Tipo	Modo	Chave
<input type="checkbox"/>	PLACA	STRING	NULLABLE	-
<input type="checkbox"/>	NOME_BASE	STRING	NULLABLE	-
<input type="checkbox"/>	VEICULO	STRING	NULLABLE	-
<input type="checkbox"/>	AGREGADO	STRING	NULLABLE	-
<input type="checkbox"/>	CPF_CNPJ	INTEGER	NULLABLE	-

financeiro-01

*PERGUNTAS

*query_01

*query_0

PRODUTOS_ENTREGUESCONSULTACOMPARTILHAR

ESQUEMADETALHESVISUALIZAÇÃOLINHAGEMPERFIL

Filtro

Insira o nome ou o valor da propriedade

<input type="checkbox"/>	Nome do campo	Tipo	Modo	Chave	Cor
<input type="checkbox"/>	NUMERO_PEDIDO	STRING	NULLABLE	-	-
<input type="checkbox"/>	FRANQUEADO	STRING	NULLABLE	-	-
<input type="checkbox"/>	NOME_BASE	STRING	NULLABLE	-	-
<input type="checkbox"/>	AGREGADO	STRING	NULLABLE	-	-
<input type="checkbox"/>	ASSINATURA	STRING	NULLABLE	-	-
<input type="checkbox"/>	VALOR_MERCADORIA	FLOAT	NULLABLE	-	-
<input type="checkbox"/>	CEP	FLOAT	NULLABLE	-	-
<input type="checkbox"/>	ENDERECO_COMPLETO	STRING	NULLABLE	-	-
<input type="checkbox"/>	CIDADE	STRING	NULLABLE	-	-
<input type="checkbox"/>	BAIRRO	STRING	NULLABLE	-	-
<input type="checkbox"/>	DESTINATARIO	STRING	NULLABLE	-	-
<input type="checkbox"/>	HORARIO_ENTREGA	DATE	NULLABLE	-	-
<input type="checkbox"/>	CPF_CNPJ	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	PESO_COBRAVEL	FLOAT	NULLABLE	-	-
<input type="checkbox"/>	ESTADO_REMETENTE	STRING	NULLABLE	-	-
<input type="checkbox"/>	CIDADE_DE_ORIGEM	STRING	NULLABLE	-	-

EDITAR ESQUEMAVER POLÍTICAS DE ACESSO DE LINHA

Histórico de jobs

financeiro-01

*PERGUNTAS

*query_01

*c

VALORESCONSULTACOMPARTILHARCOF

ESQUEMADETALHESVISUALIZAÇÃOLINHAGEM

Filtro

Insira o nome ou o valor da propriedade

<input type="checkbox"/>	Nome do campo	Tipo	Modo	Chave	Compilaç
<input type="checkbox"/>	NOME_BASE	STRING	NULLABLE	-	-
<input type="checkbox"/>	CIDADE	STRING	NULLABLE	-	-
<input type="checkbox"/>	BAIRRO	STRING	NULLABLE	-	-
<input type="checkbox"/>	VEICULO	STRING	NULLABLE	-	-
<input type="checkbox"/>	VALOR	FLOAT	NULLABLE	-	-

EDITAR ESQUEMAVER POLÍTICAS DE ACESSO DE LINHA

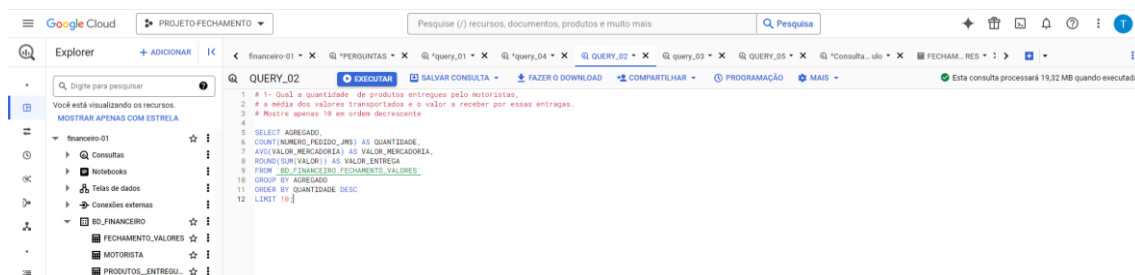
15 - Realizando a consulta com os joins específicos para criar uma view e salvar essa consulta para realizar novas consultas em cima dessa base salva.

```
SELECT PE.* ,
M.VEICULO,
V.VALOR
FROM `BD_FINANCEIRO.PRODUTOS__ENTREGUES` AS PE
JOIN `BD_FINANCEIRO.MOTORISTA` AS M
ON PE.CPF_CNPJ = M.CPF_CNPJ
LEFT JOIN `BD_FINANCEIRO.VALORES` AS V
ON PE.CIDADE = V.CIDADE
AND PE.BAIRRO = V.BAIRRO
AND PE.NOME_BASE = V.NOME_BASE
AND M.VEICULO = V.VEICULO
```

Perguntas para serem respondidas :

1- Qual a quantidade de produtos entregues pelo motoristas,
a média dos valores transportados e o valor a receber por essas entregas.
Mostre apenas 10 em ordem decrescente

```
SELECT AGREGADO,
COUNT(NUMERO_PEDIDO_JMS) AS QUANTIDADE,
AVG(VALOR_MERCADORIA) AS VALOR_MERCADORIA,
ROUND(SUM(VALOR)) AS VALOR_ENTREGA
FROM `BD_FINANCEIRO.FECHAMENTO_VALORES`
GROUP BY AGREGADO
ORDER BY QUANTIDADE DESC
LIMIT 10;
```



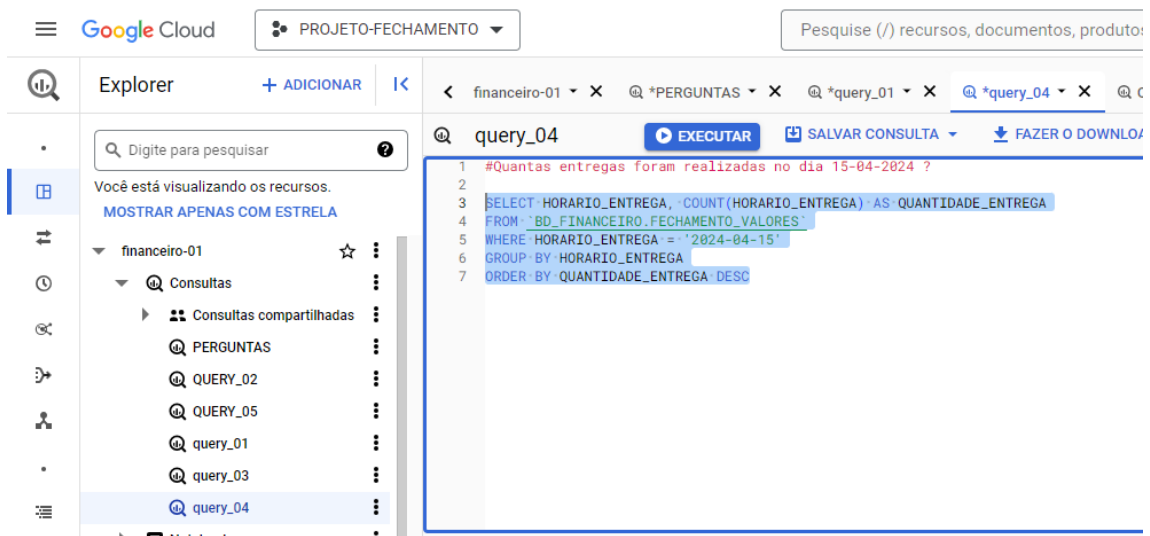
#Quais são os cinco franqueados que têm o maior valor total de mercadorias
#entregues nas cidades?

```
SELECT F.FRANQUEADO, F.CIDADE, ROUND(SUM(F.VALOR_MERCADORIA)) AS VALOR_TOTAL
FROM `BD_FINANCEIRO.FECHAMENTO_VALORES` AS F
GROUP BY F.FRANQUEADO, F.CIDADE
ORDER BY VALOR_TOTAL DESC
LIMIT 5
```




#Quantas entregas foram realizadas no dia 15-04-2024 e exibe em ordem decrescente ?

```
SELECT HORARIO_ENTREGA, COUNT(HORARIO_ENTREGA) AS QUANTIDADE_ENTREGA
FROM `BD_FINANCEIRO.FECHAMENTO_VALORES`
WHERE HORARIO_ENTREGA = '2024-04-15'
GROUP BY HORARIO_ENTREGA
ORDER BY QUANTIDADE_ENTREGA DESC
```



#Quais são os dez agregados que têm o maior
#número de entregas para destinatários diferentes em cidades distintas?
#sql

```
SELECT AGREGADO,
CIDADE,
DESTINATARIO,
COUNT(DISTINCT(NUMERO_PEDIDO_JMS)) AS TOTAL_PEDIDOS
FROM `BD_FINANCEIRO.FECHAMENTO_VALORES`
GROUP BY AGREGADO, CIDADE, DESTINATARIO
ORDER BY TOTAL_PEDIDOS DESC
LIMIT 10
```

Google Cloud PROJETO-FECHAMENTO

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

Explorer + ADICIONAR

Você está visualizando os recursos. [MOSTRAR APENAS COM ESTRELA](#)

- financeiro-01
 - Consultas
 - Consultas compartilhadas
 - PERGUNTAS
 - QUERY_02
 - QUERY_05
 - query_01
 - query_03
 - query_04
 - Notebooks
 - Telas de dados
 - Conexões externas

QUERY_05

```

1 #Quais são os dez agregados que têm o maior
2 #número de entregas para destinatários diferentes em cidades distintas?
3 #sql
4
5 SELECT AGREGADO,
6        CIDADE,
7        DESTINATARIO,
8        COUNT(DISTINCT(NUMERO_PEDIDO_JMS)) AS TOTAL_PEDIDOS
9 FROM `BD_FINANCEIRO.FECHAMENTO_VALORES`
10 GROUP BY AGREGADO, CIDADE, DESTINATARIO
11 ORDER BY TOTAL_PEDIDOS DESC
12 LIMIT 10

```

#Qual é o valor médio de mercadorias entregues em cada estado remetente (ESTADO_REMETENTE)?

```

SELECT ESTADO_REMETENTE , ROUND(AVG(VALOR_MERCADORIA)) AS MEDIA_VALOR
FROM `BD_FINANCEIRO.FECHAMENTO_VALORES`
GROUP BY ESTADO_REMETENTE
ORDER BY MEDIA_VALOR DESC

```

Google Cloud PROJETO-FECHAMENTO

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

Explorer + ADICIONAR

Você está visualizando os recursos. [MOSTRAR APENAS COM ESTRELA](#)

- financeiro-01
 - Consultas
 - Consultas compartilhadas
 - PERGUNTAS
 - QUERY_02
 - QUERY_05
 - query_01
 - query_03
 - query_04
 - Notebooks
 - Telas de dados
 - Conexões externas
 - BD_FINANCEIRO
 - FECHAMENTO_VAL...
 - MOTORISTA

query_03

```

1 #Quais são os cinco franqueados que têm o maior valor total de mercadorias
2 #entregues nas cidades?
3
4 SELECT F.FRANQUEADO, F.CIDADE, ROUND(SUM(F.VALOR_MERCADORIA)) AS VALOR_TOTAL
5 FROM `BD_FINANCEIRO.FECHAMENTO_VALORES` AS F
6 GROUP BY F.FRANQUEADO, F.CIDADE
7 ORDER BY VALOR_TOTAL DESC
8 LIMIT 5

```