

Projeto Final - BIG QUERY

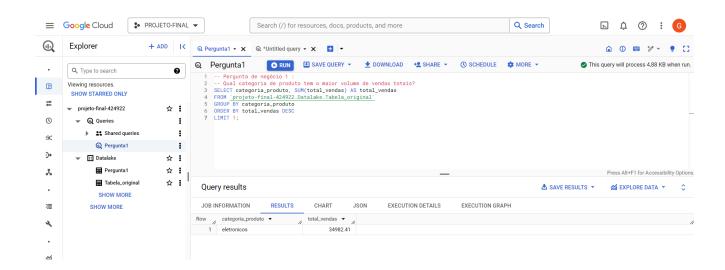
Professora: Franciane Rodrigues

Equipe:

- 1. Ana Carolina
- 2. Carina Mayara
- 3. David Ben
- 4. Giovanna Shiguemori
- 5. Marcos Neves
- 6. Rafael Souza

SELECT categoria_produto, SUM(total_vendas) AS total_vendas FROM `projeto-final-424922.Datalake.Tabela_original` GROUP BY categoria_produto ORDER BY total_vendas DESC LIMIT 1;

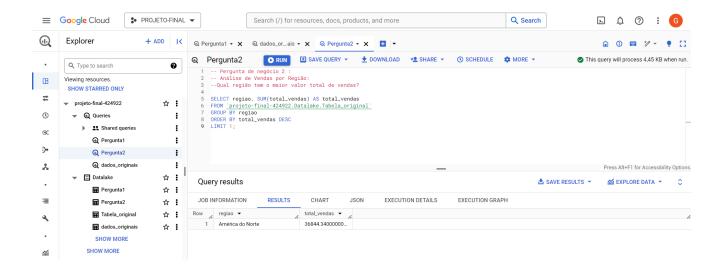
- **SELECT** categoria_produto, **SUM**(total_vendas) AS total_vendas: Seleciona a categoria do produto e a soma das vendas totais para cada categoria.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica o dataset e a tabela que contém os dados.
- GROUP BY categoria_produto: Agrupa os dados por categoria de produto.
- ORDER BY total_vendas DESC: Ordena os resultados em ordem decrescente de vendas totais.
- LIMIT 1: Retorna apenas a categoria de produto com o maior volume de vendas.



```
2) -- Pergunta de negócio 2 :
    -- Análise de Vendas por Região:
    -- Qual região tem o maior valor total de vendas?

SELECT regiao, SUM(total_vendas) AS total_vendas
FROM `projeto-final-424922.Datalake.Tabela_original`
GROUP BY regiao
ORDER BY total_vendas DESC
LIMIT 1;
```

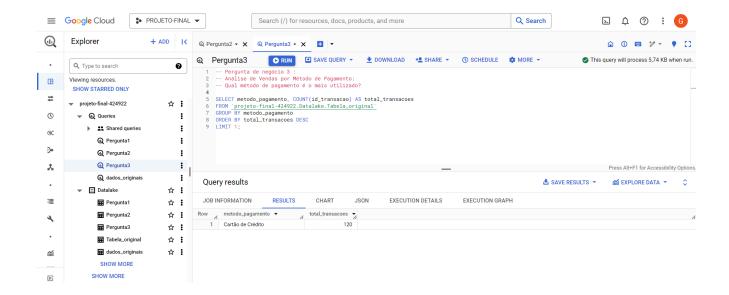
- **SELECT** regiao, **SUM**(total_vendas) **AS** total_vendas: Seleciona a região e a soma das vendas totais para cada região.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica o dataset e a tabela que contém os dados.
- GROUP BY regiao: Agrupa os dados por região.
- ORDER BY total_vendas DESC: Ordena os resultados em ordem decrescente de vendas totais.
- LIMIT 1: Retorna apenas a região com o maior valor total de vendas.



```
3) -- Pergunta de negócio 3 :
    -- Análise de Vendas por Método de Pagamento:
    -- Qual método de pagamento é o mais utilizado?

SELECT metodo_pagamento, COUNT(id_transacao) AS total_transacoes
FROM `projeto-final-424922.Datalake.Tabela_original`
GROUP BY metodo_pagamento
ORDER BY total_transacoes DESC
LIMIT 1;
```

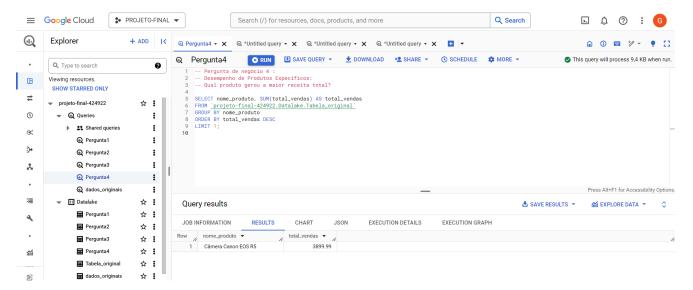
- **SELECT** metodo_pagamento, **COUNT**(id_transacao) **AS** total_transacoes: Seleciona o método de pagamento e a contagem total de transações para cada método.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica o dataset e a tabela que contém os dados.
- GROUP BY metodo_pagamento: Agrupa os dados por método de pagamento.
- ORDER BY total_transacoes DESC: Ordena os resultados em ordem. decrescente de contagem de transações.
- LIMIT 1: Retorna apenas o método de pagamento mais utilizado.



```
4) -- Pergunta de negócio 4 :
    -- Desempenho de Produtos Específicos:
    -- Qual produto gerou a maior receita total?

SELECT nome_produto, SUM(total_vendas) AS total_vendas
FROM `projeto-final-424922.Datalake.Tabela_original`
GROUP BY nome_produto
ORDER BY total_vendas DESC
LIMIT 1;
```

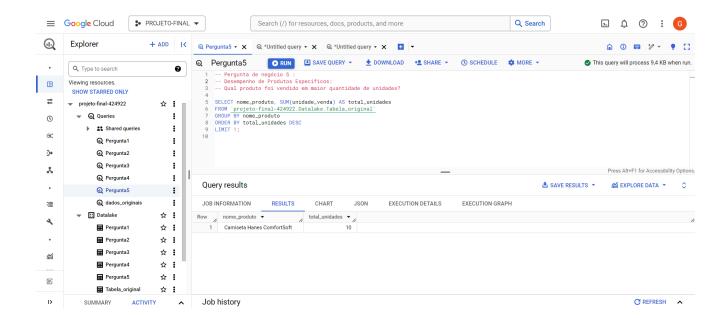
- **SELECT** nome_produto, **SUM**(total_vendas) **AS** total_vendas: Seleciona o nome do produto e a soma das vendas (total_vendas) para cada produto.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica a tabela de dados a ser usada.
- GROUP BY nome_produto: Agrupa os resultados por nome de produto, para calcular a soma das vendas de cada produto individualmente.
- ORDER BY total_vendas DESC: Ordena os resultados em ordem decrescente de total_vendas, colocando o produto com a maior receita no topo.
- LIMIT 1: Retorna apenas o primeiro resultado, que é o produto com a maior receita.



```
5) -- Pergunta de negócio 5 :
    -- Desempenho de Produtos Específicos:
    -- Qual produto foi vendido em maior quantidade de unidades?

SELECT nome_produto, SUM(unidade_venda) AS total_unidades
FROM `projeto-final-424922.Datalake.Tabela_original`
GROUP BY nome_produto
ORDER BY total_unidades DESC
LIMIT 1;
```

- SELECT nome_produto, SUM(unidade_venda) AS total_unidades: Seleciona o nome do produto e a soma das unidades vendidas (total_unidades) para cada produto.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica a tabela de dados a ser usada.
- **GROUP BY** nome_produto: Agrupa os resultados por nome de produto, para calcular a soma das unidades vendidas de cada produto individualmente.
- ORDER BY total_unidades DESC: Ordena os resultados em ordem decrescente de total_unidades, colocando o produto com a maior quantidade de unidades vendidas no topo.
- LIMIT 1: Retorna apenas o primeiro resultado, que é o produto com a maior quantidade de unidades vendidas.

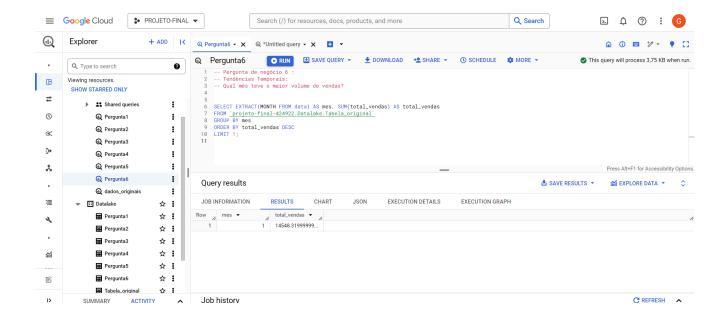


```
6) -- Pergunta de negócio 6 :
    -- Tendências Temporais:
    -- Qual mês teve o maior volume de vendas?

SELECT EXTRACT(MONTH FROM data) AS mes, SUM(total_vendas) AS total_vendas
FROM `projeto-final-424922.Datalake.Tabela_original`
GROUP BY mes

ORDER BY total_vendas DESC
LIMIT 1;
```

- SELECT EXTRACT(MONTH FROM data) AS mes, SUM(total_vendas) AS total_vendas:
 Seleciona o mês da data da transação e a soma das vendas totais para cada
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica o dataset e a tabela que contém os dados.
- GROUP BY mes: Agrupa os dados por mês.
- ORDER BY total_vendas DESC: Ordena os resultados em ordem decrescente de vendas totais.
- LIMIT 1: Retorna apenas o mês com o maior volume de vendas.

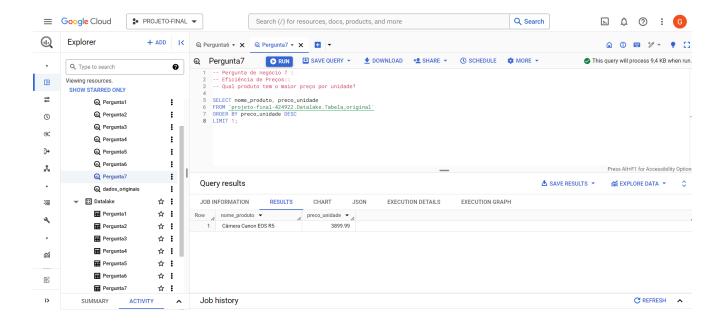


```
    7) -- Pergunta de negócio 7 :

            Eficiência de Preços::
            Qual produto tem o maior preço por unidade?

    SELECT nome_produto, preco_unidade
    FROM `projeto-final-424922.Datalake.Tabela_original`
    ORDER BY preco_unidade DESC
    LIMIT 1;
```

- **SELECT** nome_produto, preco_unidade: Seleciona o nome do produto e o preço por unidade.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica o dataset e a tabela que contém os dados.
- ORDER BY preco_unidade DESC: Ordena os resultados em ordem decrescente de preço por unidade.
- LIMIT 1: Retorna apenas o produto com o maior preço por unidade.



```
-- Top 10 dos produtos mais vendidos na Europa no mês de Janeiro de 2024

SELECT

nome_produto,
SUM(total_vendas) AS total_vendas_janeiro

FROM
   `projeto-final-424922.Datalake.Tabela_original`
WHERE

EXTRACT(MONTH FROM data) = 1 -- Janeiro

AND EXTRACT(YEAR FROM data) = 2024 -- Ano de 2024

AND regiao = 'Europa'

GROUP BY

nome_produto

ORDER BY

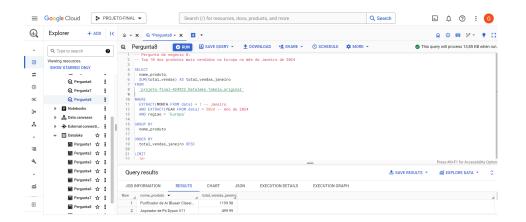
total_vendas_janeiro DESC

LIMIT

10;
```

8) -- Pergunta de negócio 8:

- **SELECT** nome_produto, **SUM**(total_vendas) **AS** total_vendas_janeiro: Seleciona o nome do produto e a soma das vendas como "total_vendas_janeiro".
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica de qual tabela e dataset os dados serão extraídos.
- WHERE EXTRACT(MONTH FROM data) = 1: Filtra os dados para incluir apenas vendas realizadas em Janeiro (mês 1).
- AND EXTRACT(YEAR FROM data) = 2024: Filtra os dados para incluir apenas vendas realizadas no ano de 2024.
- AND regiao = 'Europa': Filtra os dados para incluir apenas vendas na região 'Europa'.
- GROUP BY nome_produto: Agrupa os resultados pelo nome do produto, permitindo a soma das vendas para cada produto.
- ORDER BY total_vendas_janeiro DESC: Ordena os resultados em ordem decrescente com base no total de vendas em Janeiro.
- LIMIT 10: Limita os resultados aos 10 primeiros, mostrando apenas os produtos mais vendidos.



```
9) -- Pergunta de negócio 9:
    -- Qual foi o dia que mais teve produtos vendidos?

SELECT
    data,
    SUM(unidade_venda) AS total_transacoes

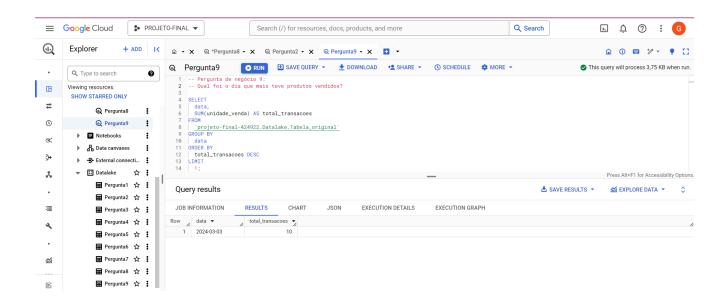
FROM
    `projeto-final-424922.Datalake.Tabela_original`

GROUP BY
    data

ORDER BY
    total_transacoes DESC

LIMIT
    1;
```

- SELECT data, SUM(unidade_venda) AS total_transacoes: Esta parte seleciona duas informações:
- data: Recupera a data da coluna "data".
- SUM(unidade_venda) AS total_transacoes: Calcula a soma dos valores na coluna "unidade_venda" (provavelmente representando unidades vendidas) e nomeia essa soma calculada como "total_transacoes" (total de transações).
- FROM `projeto-final-424922.Datalake.Tabela_original: Especifica a origem dos
 dados, que é a tabela "Tabela_original" localizada dentro do conjunto de
 dados "Datalake" do projeto "projeto-final-424922".
- GROUP BY data: Agrupa os resultados com base em datas únicas. Portanto, você obterá uma linha para cada data distinta em seu conjunto de dados.
- ORDER BY total_transacoes DESC: Ordena os resultados em ordem decrescente com base no "total_transacoes" calculado. O dia com o maior número de transações estará no topo.
- LIMIT 1: Limita a saída apenas à primeira linha após a ordenação.
 Basicamente, você está pedindo a data com o maior total de transações.



```
10) -- Pergunta de negócio 10:
   -- Em qual continente foi mais usado o método de pagamento menos usado no geral?
WITH MetodoPagamentoContagem AS (
  SELECT
    metodo_pagamento,
    COUNT(*) AS total_transacoes
    `projeto-final-424922.Datalake.Tabela_original`
  GROUP BY
    metodo_pagamento
),
MenosUsado AS (
  SELECT
    metodo_pagamento
    MetodoPagamentoContagem
  ORDER BY
    total_transacoes ASC
  LIMIT 1
SELECT
  regiao,
  COUNT(*) AS total_transacoes
FROM
  `projeto-final-424922.Datalake.Tabela_original`
  metodo_pagamento = (SELECT metodo_pagamento FROM MenosUsado)
GROUP BY
  regiao
ORDER BY
  total_transacoes DESC
LIMIT 1:
```

- 1. Subquery MetodoPagamentoContagem:
- SELECT metodo_pagamento, COUNT(*) AS total_transacoes: Seleciona o método de pagamento e conta a quantidade de vezes que ele aparece.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Define a tabela de onde os dados serão extraídos.
- GROUP BY metodo_pagamento: Agrupa os resultados pelo método de pagamento, permitindo a contagem.
- 2. Subquery MenosUsado:
- SELECT metodo_pagamento: Seleciona apenas a coluna do método de pagamento.
- FROM MetodoPagamentoContagem: Utiliza os resultados da subquery anterior.
- ORDER BY total_transacoes ASC: Ordena os métodos de pagamento em ordem crescente de acordo com a quantidade de transações.
- **LIMIT 1**: Seleciona apenas o primeiro resultado, que representa o método de pagamento menos usado.

- 3. Query Principal:
- SELECT regiao, COUNT(*) AS total_transacoes: Seleciona a região e conta o número de transacões.
- FROM `projeto-final-424922.Datalake.Tabela_original`: Especifica a tabela de dados.
- WHERE metodo_pagamento = (SELECT metodo_pagamento FROM MenosUsado): Filtra
 as transações para incluir apenas aquelas que utilizam o método de pagamento
 menos usado, encontrado na subquery MenosUsado.
- GROUP BY regiao: Agrupa os resultados por região.
- ORDER BY total_transacoes DESC: Ordena as regiões em ordem decrescente de acordo com o número de transações.
- LIMIT 1: Seleciona a primeira região da lista, representando aquela onde o método de pagamento menos usado é mais frequente.

