PERGUNTA 2

Descrição: Crie um dashboard que forneça insights sobre o engajamento dos apoiadores através de dados de abertura de e-mail, petições assinadas, contribuições financeiras para cada campanha, bem como informações demográficas dos apoiadores.

Insights sobre engagamento dos apoiadores por:

- 1. Aberturas de e-mail
- 2. Petições assinadas
- 3. Contribuições financeiras
- 4. Informações demográficas
- 5. Engajamento de e-mail em cada uma das petições
- 6. Distribuição demográfica desse engajamento

Tabela de E-mails (email events.csv)

Esquema

Field Name	Mode	Туре	Description	
id	NULLABLE	INTEGER	Identificador único para cada entrada de registro de atividade.	
email	NULLABLE	STRING	Endereço de e-mail associado à ação.	
action	NULLABLE	STRING	Tipo de ação realizada, como "login", "logout" ou "redefinição de senha".	
timestamp	NULLABLE	TIMESTAMP	Data e hora em que a ação foi realizada.	
email_id	NULLABLE	INTEGER	Chave estrangeira que referencia o ID do usuário associado à ação.	

- Email_Id como chave estrangeira.
- Total de registros na tabela (300.000)
- Número de registros únicos de e-mail (8.000)
- Registros mais recentes e mais antigos na tabela (2023-01-01 a 2024-12-30)

Tabela de Petições (petition_names.csv)

Esquema

Field Name	Mode Type		Description	
petition_name	NULLABLE	STRING	O nome das petições assinadas.	
associated_mail	NULLABLE	STRING	Todos os id's de e-mail que estão associados à petição.	

- Email_Id como chave estrangeira.
- Total de registros na tabela originalmente (25)

• Total de registros após tratamento (20.000)

Informações Demográficas (contacts.csv)

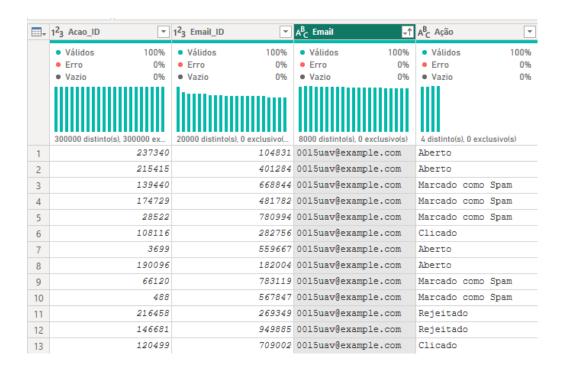
Esquema

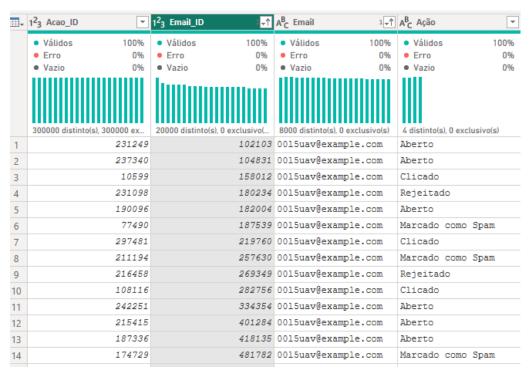
Field Name	Mode	Type	D escription
email	NULLABLE	STRING	Endereço eletrônico do apoiador.
gender	NULLABLE	STRING	Identificação do sexo do apoiador.
birth_date	NULLABLE	DATE	Dsata de nascimento do apoiador.
phone_number	NULLABLE	STRING	Número de contato do apoiador.

- Email como chave estrangeira.
- Total de registros na tabela (8000)

Observações

- 1. Dados faltantes
 - a. Não há o campo "região" (mencionada na parte de requisitos técnicos);
 - b. Não há dados sobre contribuições financeiras (mencionada na descrição da pergunta).
- 2. A integração e transformações dos dados foram realizadas com o BigQuery. Fiz download dos arquivos localmente e fiz upload dos arquivos no Google Cloud Storage. Poderia ter conectado o Power BI diretamente ao Google Drive, mas não seria possível usar SQL, apenas a linguagem M/Power Query, na qual tenho conhecimento avançado, no entanto, essa opção não estava disponível para a atividade em questão, então optei por seguir o caminho do BigQuery acima, onde poderia usar SQL e criar views.
- 3. Ao unir a tabela de e-mails (eventos) com a tabela de petições e a de contatos em uma só *visualização* (considerando esta como uma view ou tabela unificada), a contagem de eventos e a contagem de petições se tornam a mesma. Talvez a ação do tipo "aberto" significa "assinado", mas não ficou claro essa equivalência no enunciado.
- 4. A tabela de petições possui 20000 IDs de emails distintos, cada peticoes tem exatamente 800 Emails IDs distintos (25 peticoes * 800 emails distintos = 20000 peticoes).
- 5. Ao realizar a junção da tabela de emails (eventos) com a tabela de contatos (demografia), usando Email como chave estrangeira, encontrei múltiplos Email_IDs associados ao mesmo Email e vice-versa. Isso signfica que a mesma ação de e-mail, como "Aberto", pode parecer estar associada a diferentes usuários, o que pode inflar métricas e complicar a análise. O ideal é que garantir que cada Email_ID represente uma único conta de e-mail.





Link para o dashboard

Código consolidado

```
-- DEFINE a Tabela de Eventos
 Eventos AS (
  SELECT
    id AS Acao_ID,
    email_id AS Email_ID,
    email AS Email,
    DATE(timestamp) AS DATA,
    FORMAT_TIMESTAMP('%Y%m', timestamp) AS Data_ID,
    EXTRACT (YEAR
    FROM
      timestamp) AS Ano,
      WHEN action = 'click' THEN 'Clicado'
      WHEN action = 'spam' THEN 'Marcado como Spam'
      WHEN action = 'bounce' THEN 'Rejeitado'
      WHEN action = 'open' THEN 'Aberto'
    ELSE
    'Desconhecido'
  END
    AS `Ação`
 FROM
    greenpeace-0524.greenpeace0524.eventos ),
  -- DEFINE a Tabela de Petições Peticoes AS (
SELECT
  string_field_0 AS `Nome Petição`,
  CAST(TRIM(associated_email) AS INT64) AS Email_ID
FROM (
 SELECT
    string_field_0,
    associated_email.
    ROW_NUMBER() OVER (PARTITION BY string_field_0 ORDER BY OFFSET )
AS rownumber
  FROM
    greenpeace-0524.greenpeace0524.peticoes,
    UNNEST(SPLIT(REGEXP_REPLACE(string_field_1, r'\s+', ''), ',')) AS
associated_email
 WITH
  OFFSET
    )),
  -- DEFINE a Tabela de Contatos Contatos AS (
SELECT
  email AS Email,
  CASE
    WHEN gender = 'female' THEN 'Feminino'
    WHEN gender = 'male' THEN 'Masculino'
    WHEN gender = 'non-binary' THEN 'Não-Binário'
 ELSE
  'Desconhecido'
END
```

```
AS `Gênero`,
  TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE), YEAR) AS
Idade.
  CASE
    WHEN TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE),
YEAR) <= 12 THEN 'Até 12 anos'
    WHEN TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE),
YEAR) BETWEEN 13
  AND 19 THEN '13-19 anos'
   WHEN TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE),
YEAR) BETWEEN 20 AND 39 THEN '20-39 anos'
   WHEN TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE),
YEAR) BETWEEN 40
  AND 59 THEN '40-59 anos'
    WHEN TIMESTAMP_DIFF(CURRENT_DATE(), CAST(birth_date AS DATE),
YEAR) BETWEEN 60 AND 74 THEN '60-74 anos'
  ELSE
  '75+ anos'
END
 AS `Faixa Etária`
FROM
  greenpeace-0524.greenpeace0524.contatos ) -- Consulta Principal
SELECT
  e.Acao_ID,
  e.Email_ID,
  e.Email,
  e. `Ação`,
  p.`Nome Petição`,
 c.`Gênero`,
  c.Idade,
  c. `Faixa Etária`,
  e.Data.
  e.Data_ID,
  e.Ano
FROM
  Eventos e
LEFT JOIN
 Peticoes p
ON
  e.Email_ID = p.Email_ID
LEFT JOIN
  Contatos c
ON
  e.Email = c.Email
ORDER BY
  e.Email_ID,
  e.Email;
```