

# 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 engajamento 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	Type	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	Description
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

- Dados faltantes
  - Não há o campo “região” (mencionada na parte de requisitos técnicos);
  - Não há dados sobre contribuições financeiras (mencionada na descrição da pergunta).
- 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.
- 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.
- 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).
- 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 significa 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.

	1 <sup>2</sup> <sub>3</sub> Acao_ID	1 <sup>2</sup> <sub>3</sub> Email_ID	A <sup>B</sup> <sub>C</sub> Email	A <sup>B</sup> <sub>C</sub> Ação
	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>300000 distinto(s), 300000 ex...</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>20000 distinto(s), 0 exclusivo(s)</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>8000 distinto(s), 0 exclusivo(s)</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>4 distinto(s), 0 exclusivo(s)</div>
1	237340	104831	0015uav@example.com	Aberto
2	215415	401284	0015uav@example.com	Aberto
3	139440	668844	0015uav@example.com	Marcado como Spam
4	174729	481782	0015uav@example.com	Marcado como Spam
5	28522	780994	0015uav@example.com	Marcado como Spam
6	108116	282756	0015uav@example.com	Clicado
7	3699	559667	0015uav@example.com	Aberto
8	190096	182004	0015uav@example.com	Aberto
9	66120	783119	0015uav@example.com	Marcado como Spam
10	488	567847	0015uav@example.com	Marcado como Spam
11	216458	269349	0015uav@example.com	Rejeitado
12	146681	949885	0015uav@example.com	Rejeitado
13	120499	709002	0015uav@example.com	Clicado

	1 <sup>2</sup> <sub>3</sub> Acao_ID	1 <sup>2</sup> <sub>3</sub> Email_ID	A <sup>B</sup> <sub>C</sub> Email	A <sup>B</sup> <sub>C</sub> Ação
	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>300000 distinto(s), 300000 ex...</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>20000 distinto(s), 0 exclusivo(s)</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>8000 distinto(s), 0 exclusivo(s)</div>	<div> <div>Válidos 100%</div> <div>Erro 0%</div> <div>Vazio 0%</div> </div> <div>4 distinto(s), 0 exclusivo(s)</div>
1	231249	102103	0015uav@example.com	Aberto
2	237340	104831	0015uav@example.com	Aberto
3	10599	158012	0015uav@example.com	Clicado
4	231098	180234	0015uav@example.com	Rejeitado
5	190096	182004	0015uav@example.com	Aberto
6	77490	187539	0015uav@example.com	Marcado como Spam
7	297481	219760	0015uav@example.com	Clicado
8	211194	257630	0015uav@example.com	Marcado como Spam
9	216458	269349	0015uav@example.com	Rejeitado
10	108116	282756	0015uav@example.com	Clicado
11	242251	334354	0015uav@example.com	Aberto
12	215415	401284	0015uav@example.com	Aberto
13	187336	418135	0015uav@example.com	Aberto
14	174729	481782	0015uav@example.com	Marcado como Spam

## [Link para o dashboard](#)

### Código consolidado

```
-- DEFINE a Tabela de Eventos
WITH
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,
CASE
  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;

```