

Tarefa 06

Banco de Dados Avançado

Aluna: Letícia Trein Medeiros

Escolha da Base

- O sistema **156** unificado da prefeitura de Curitiba recebe **informações e demandas** provindas dos cidadãos, ela é interessante e fornecesse informações de demandas da população de forma eficaz e possibilita o **direcionamento adequado** ao órgão responsável para melhor resposta.
- A presença de dados de **geolocalização** de bairros e ruas fornece informações específicas e exclusivas por **bairros e regionais**, que podem ser **diferentes** por questões socioeconômicas e **mapeamento social** da cidade.
- Por ser um sistema de atualização **diária**, é possível fazer um acompanhamento temporal das demandas
- Ter uma boa noção das demandas da população pode ser muito útil para tomada de **decisões políticas e sociais**, redirecionamento de recursos e tomada de decisões de prevenções.

Detalhamento Base de Dados

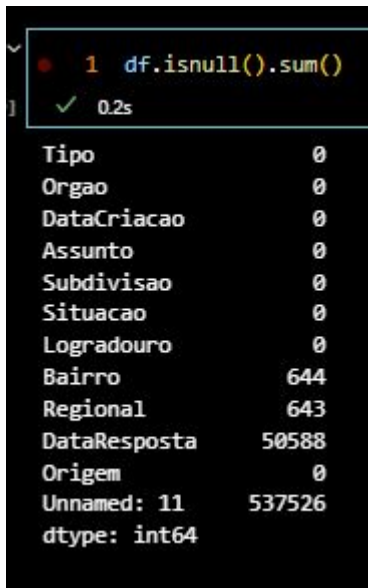
- Sistema Integrado de Atendimento ao Cidadão - **SIAC 156**
 - Descrição: Base de Dados contendo as solicitações geradas na Central 156, principal **canal de comunicação** entre o **cidadão** e a **Prefeitura Municipal de Curitiba**. Inclui todas as demandas direcionadas às Secretarias e Órgãos da Administração Municipal. Estes dados são oriundos do Sistema Integrado de Atendimento ao Cidadão - SIAC
 - Órgão Responsável: Governo Municipal de Curitiba
 - Frequência de Atualização: Diária
 - Banco de Dados: Dados abertos da Cidade de Curitiba
 - Formato: CSV
 - Disponível em: <https://dadosabertos.c3sl.ufpr.br/curitiba/156/>
-
- Data de dowload: 01/10/2024 -> 2024-09-30_156_-_Base_de_Dados.csv

Descrição Atributos - Dicionário

- Tipo: Tipo do pedido (Solicitação, Denúncia, Reclamação, etc)
- Orgao: Órgão responsável pelo atendimento (Secretarias)
- DataCriacao: Data de criação do pedido (dd/mm/aaaa)
- Assunto: Nome do assunto do serviço
- Subdivisao: Nome da subdivisão do serviço (detalhamento do assunto)
- Situacao: Situação do pedido (Concluído, aberto, etc)
- Logradouro: Endereço de atendimento do pedido
- Bairro: Nome do bairro de atendimento do pedido
- Regional: Nome da regional de atendimento do pedido
- DataResposta: Data da resposta do atendimento (dd/mm/aaaa)
- Origem: Origem de entrada do pedido (Mobile, Telefone, Portal ou Chat)
- ID: Coluna criada para Primary Key (como valor único por tupla)

Desafios para inserção - Pré-processamento: Tabela original

1 - Avaliação existência de valores nulos e colunas inválidas no dataset.



```
1 df.isnull().sum()
```

Tipo	0
Orgao	0
DataCriacao	0
Assunto	0
Subdivisao	0
Situacao	0
Logradouro	0
Bairro	644
Regional	643
DataResposta	50588
Origem	0
Unnamed: 11	537526
dtype:	int64

2 - Retirada as colunas com maior número de valores nulos

```
1 df_new = df.drop(columns=['Unnamed: 11'])
```

3 - Retirada de valores nulos

```
[43] 1 df_new = df_new.dropna()
```

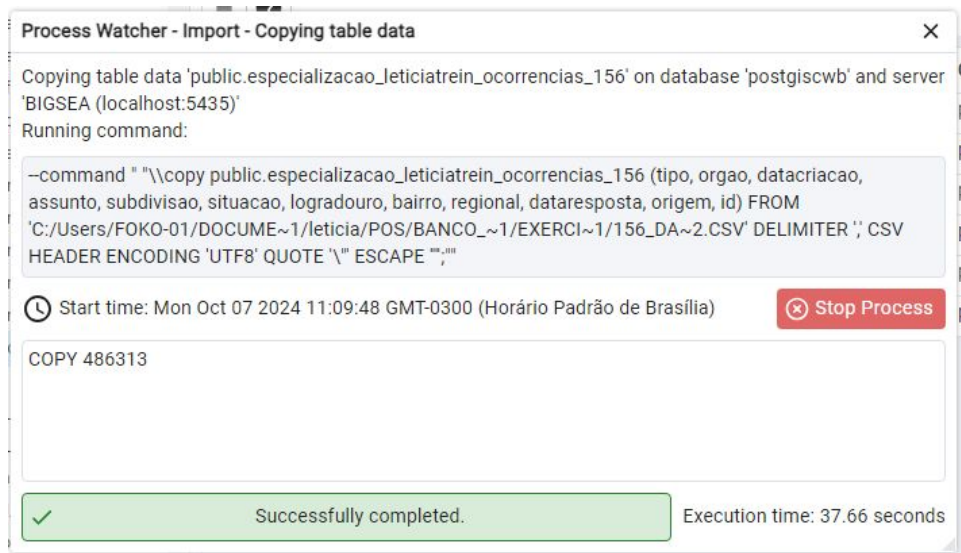
4 - Criação de coluna ID para chave primária de valor único

```
1 df_new['ID'] = range(1, len(df_new) + 1)
```

Desafios para inserção - Criação de tabela e importação

- 1 - Pelo formato original das colunas com DATAS, foi necessário classificar a variável como VARCHAR, uma vez que o formato DATE estava apresentando erro para inserção
- 2 - Foi necessário especificar o separador para ',' para inserção corretas de dados.

```
CREATE TABLE ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156 (  
  Tipo          VARCHAR(20),  
  Orgao         VARCHAR(200),  
  DataCriacao   VARCHAR(15),  
  Assunto       VARCHAR(100),  
  Subdivisao    VARCHAR(100),  
  Situacao      VARCHAR(50),  
  Logradouro    VARCHAR(100),  
  Bairro        VARCHAR(30),  
  Regional      VARCHAR(50),  
  DataResposta  VARCHAR(15),  
  Origem        VARCHAR(10),  
  ID            INT PRIMARY KEY  
);
```



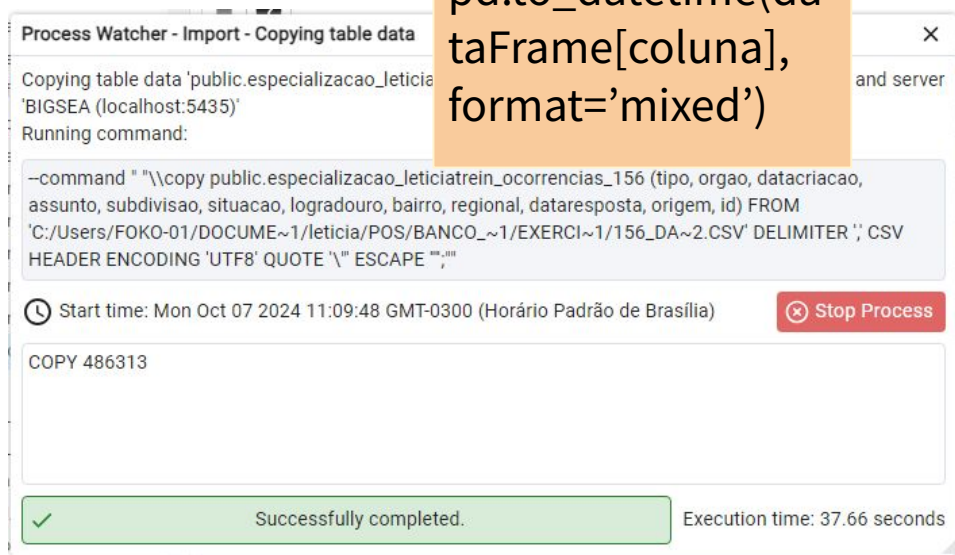
Desafios para inserção - Criação de tabela e importação

1 - Pelo formato original das colunas com DATAS, foi necessário classificar a variável como VARCHAR, uma vez que o formato DATE estava apresentando erro para inserção

2 - Foi necessário especificar o separador para ',' para inserção corretas de dados.

Sugestão recebida:
comando
`pd.to_datetime(dataFrame[coluna],
format='mixed')`

```
CREATE TABLE ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156 (  
  Tipo          VARCHAR(20),  
  Orgao         VARCHAR(200),  
  DataCriacao   VARCHAR(15),  
  Assunto       VARCHAR(100),  
  Subdivisao    VARCHAR(100),  
  Situacao      VARCHAR(50),  
  Logradouro    VARCHAR(100),  
  Bairro        VARCHAR(30),  
  Regional      VARCHAR(50),  
  DataResposta  VARCHAR(15),  
  Origem        VARCHAR(10),  
  ID            INT PRIMARY KEY  
);
```



Localização e número de Tuplas

- Caminho: postgiscwb.public.especializacao_leticiatrein_ocorrencias_156
- Número total de Tuplas:
 - 486.313



The screenshot shows a PostgreSQL query editor interface. At the top, a SQL query is entered: `SELECT COUNT(*) FROM ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156`. Below the query, there are tabs for 'Data Output', 'Messages', and 'Notifications'. The 'Data Output' tab is active, displaying a table with the results of the query. The table has two columns: 'count' (type 'bigint' with a lock icon) and a single row with the value '486313'.

	count bigint 
1	486313

Perguntas Tarefa 3

Sugestão recebida:
adicionar um
comentário sucinto
em cada uma das
consultas



Pergunta 1: Quantos pedidos foram criados por tipo?

SQL Input:

```
SELECT Tipo, COUNT(*) AS QuantidadePedidos  
FROM  
ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156  
GROUP BY Tipo;
```

	tipo character varying (20)	quantidadepedidos bigint
1	Denúncia	6074
2	Elogio	6621
3	Informação	741
4	Recadastro	2188
5	Reclamação	45119
6	Solicitação	424253
7	Sugestão	1317

Temos que o tipo com mais número de ocorrências seria “Solicitações”, seguido por “Reclamações”, “Elogios” e “Denúncias”

Pergunta 2: Quais são os 10 principais assuntos das ocorrências no bairro Centro?

SQL Input:

```
SELECT Assunto, COUNT(*) AS  
QuantidadeOcorrencias  
  
FROM  
ESPECIALIZACAO_LETICIATREIN_OCORRE  
NCIAS_156  
  
WHERE Bairro = 'CENTRO'  
  
GROUP BY Assunto  
  
ORDER BY QuantidadeOcorrencias DESC  
  
LIMIT 10;
```

	assunto character varying (100)	quantidadeocorrencias bigint
1	Abordagem social de rua - Adulto e Idoso	5978
2	Trânsito	3770
3	Coleta	1566
4	Iluminação pública - Via pública	1553
5	Central Saúde Já	1401
6	Hospitais ensino/contratualizados	639
7	Limpeza	502
8	Passeio	460
9	Motoristas, cobradores e demais funcionários do transpor...	449

Foi observado um maior número de ocorrências relacionadas a abordagem social de rua, seguido por trânsito e coleta no bairro Centro.

Pergunta 3: Quais os 15 bairros com mais Solicitações?

SQL Input:

```
SELECT Bairro, COUNT(*) AS QuantidadePedidos
FROM
ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_1
56
GROUP BY Bairro
ORDER BY QuantidadePedidos DESC
LIMIT 15;
```

	bairro character varying (30)	quantidadepedidos bigint
1	CIDADE INDUSTRIAL	32328
2	CENTRO	22745
3	CAJURU	22152
4	SITIO CERCADO	20490
5	BOQUEIRAO	20431
6	UBERABA	16075
7	XAXIM	14930
8	NOVO MUNDO	14297
9	BAIRRO ALTO	12863
10	CAPAO RASO	12787

Vemos que os bairros com maior número de ocorrência foram: CIC, Centro, Cajuru, Sítio Cercado, Boqueirão, Uberaba.

Pergunta 4: Quais são as Subdivisões do Assunto “Animais”?

SQL Input:

```
SELECT DISTINCT Subdivisao
```

```
FROM
```

```
ESPECIALIZACAO_LETICIATREIN_OCORR  
ENCIAS_156
```

```
WHERE Assunto = 'Animais'
```

	subdivisao character varying (100)
1	Cão sem guia e sem focinheira
2	Cães bravos ou avançando em via pública - Fiscalizaç...
3	Maus tratos - Animais silvestres
4	Abelhas
5	Dejetos gerados pelos animais de estimação
6	Maus tratos - Animais domésticos

As subdivisões de animais são: “Cão sem guia e sem focinheira”, “Cães bravos”, “Maus tratos - animais silvestres”, “Abelhas”, etc.

Pergunta 5: Quais são os bairros que apresentam mais demandas relacionadas a violência?

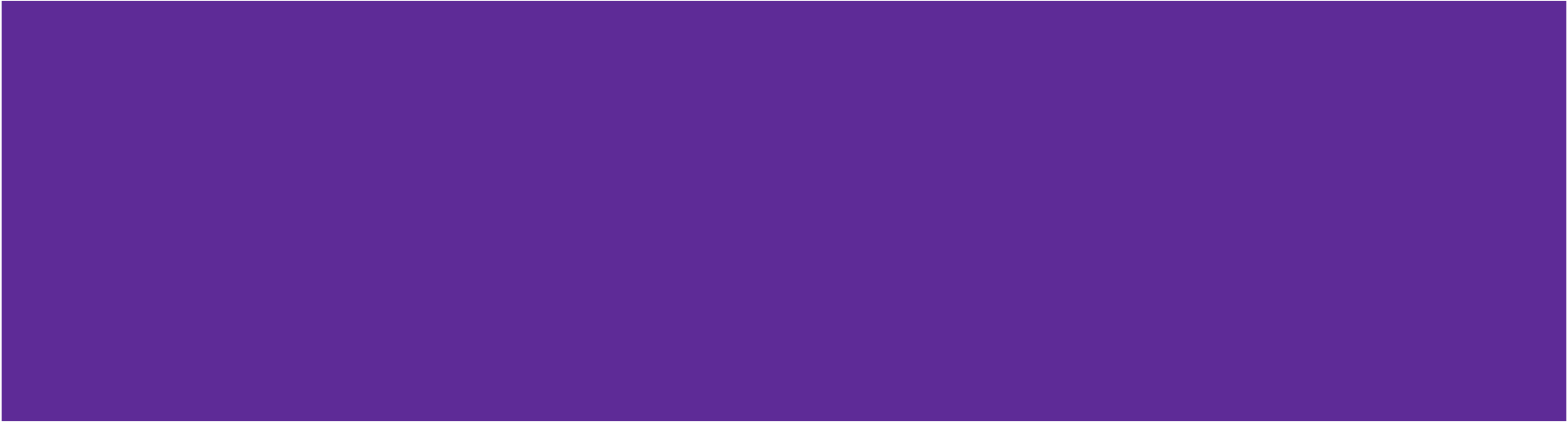
SQL Input:

```
SELECT Bairro, COUNT(*) AS TotalDenuncias  
FROM  
ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156  
WHERE Assunto like '%Violência%'  
GROUP BY Bairro  
ORDER BY TotalDenuncias DESC  
LIMIT 10;
```

	bairro character varying (30)	totaldenuncias bigint
1	CENTRO	148
2	CIDADE INDUSTRIAL	128
3	BATEL	99
4	SITIO CERCADO	82
5	AGUA VERDE	58
6	TATUQUARA	54

Os bairros que apresentam maior número de ocorrências relacionadas a violência são: Centro, CIC, Batel, Sítio Cercado, Água Verde e Tatuquara

Geoprocessamento



Preparo dos dados

Para avaliação geoespacial da base de dados, foi adicionada um campo GEOM como multi polígono.

- `SELECT AddGeometryColumn('public', 'especializacao_leticiatrein_ocorrencias_156', 'geom', 4326, 'MULTIPOLYGON', 2);`

Para georeferenciamento, foram utilizadas as geometrias dos bairros disponibilizadas na base de limites_legais. Foi necessária adequação de dois bairros para total compatibilidade entre os dados

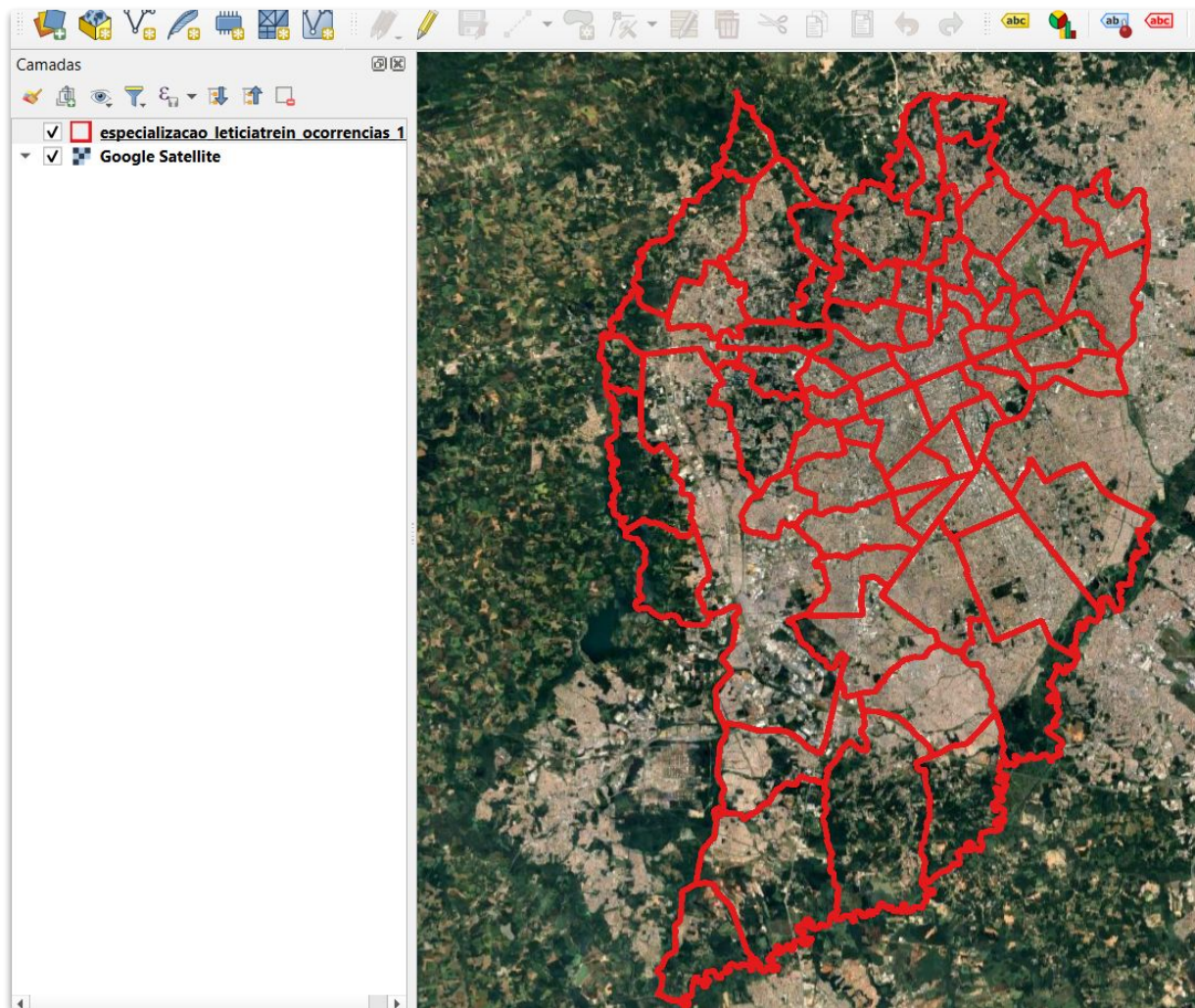
- `UPDATE public.especializacao_leticiatrein_ocorrencias_156 SET bairro = CASE WHEN bairro = 'AGUA VERDE' THEN 'ÁGUA VERDE' WHEN bairro = 'REBOUCAS' THEN 'REBOUÇAS' ELSE bairro ;`
- `UPDATE public.especializacao_leticiatrein_ocorrencias_156 b SET geom = a.geom FROM limites_legais.divisa_de_bairros a WHERE soundex(a.nome) = soundex(b.bairro);`

QGIS

Com a utilização do complemento QuickMapService, foi plotado como fundo o Google Satellite

Os Dados do 156 foram plotados com a adição de camada ligada ao banco de dados

Foi necessária adequação do SRC para SIRGAS 2000 / UTM Zone 22S, tanto no projeto quanto nas camadas para visualização no mapa corretamente



Pergunta Tarefa 6:



Pergunta:

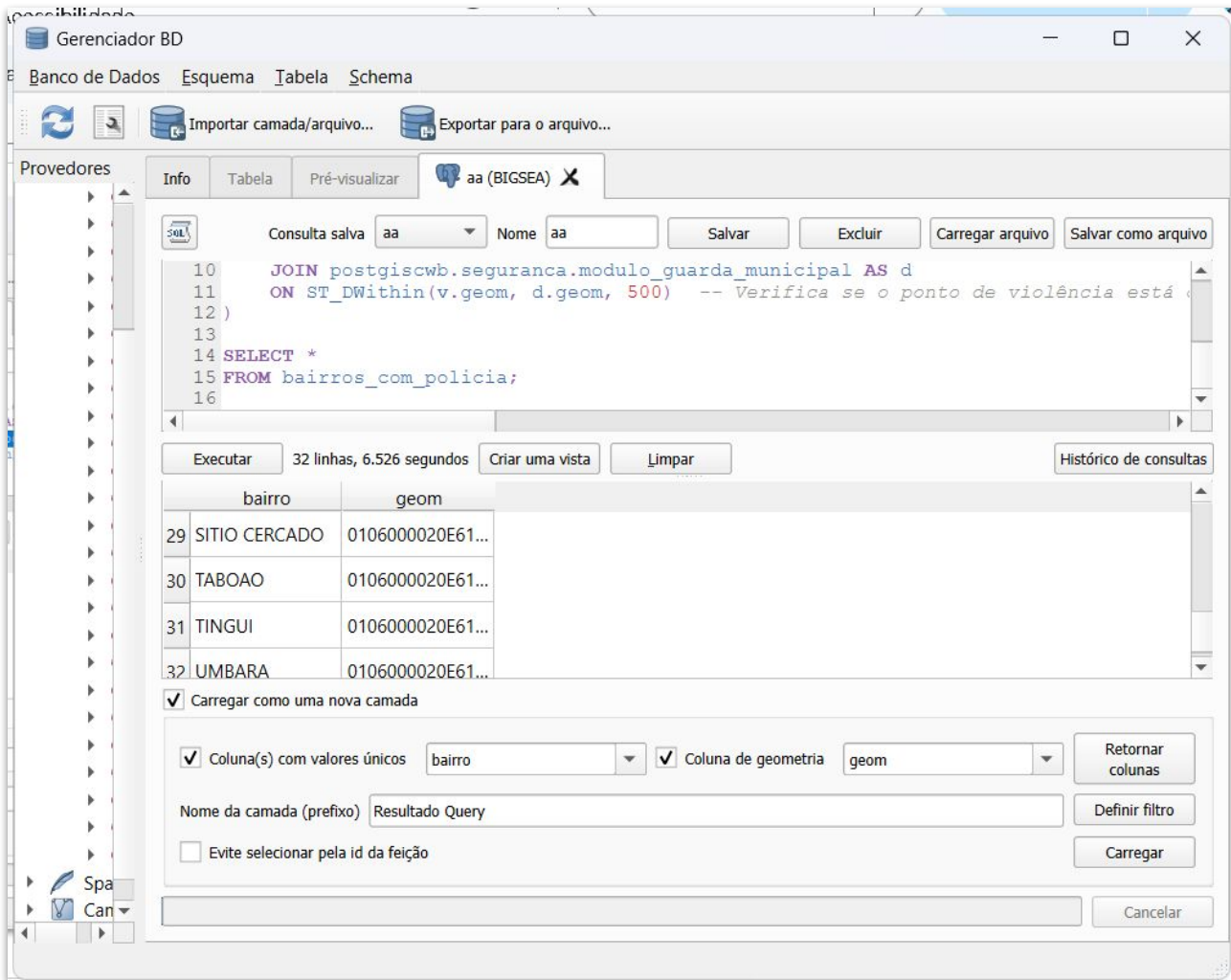
Quais bairros registraram ocorrências de violência dentro de um raio de 500 metros de um módulo policial?

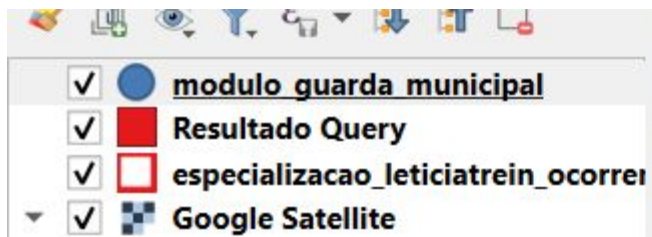
Dados Adicionais:

Para ter acesso a localização das guardas municipais, foram utilizados os dados disponibilizados no caminho: `postgiswcb.seguranca.modulo_guarda_municipal`

Query Input

```
WITH violencia_ocorrencias AS ( SELECT DISTINCT bairro, geom FROM  
especializacao_leticiatrein_ocorrencias_156 WHERE assunto LIKE '%Violência%'),  
  
bairros_com_policia AS (SELECT DISTINCT v.bairro, v.geom FROM violencia_ocorrencias  
AS v JOIN postgiscwb.seguranca.modulo_guarda_municipal AS d ON  
ST_DWithin(v.geom, d.geom, 500)),  
  
SELECT * FROM bairros_com_policia;
```





Apenas um dos módulos policiais não apresentou registros relacionados à Violência ao redor do módulo (raio de 500m)

