

Tarefa 03

Banco de Dados Avançado

Aluna: Letícia Trein Medeiros

Escolha da Base

- O sistema 156 unificado da prefeitura de Curitiba recebe informações e demandas providas 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


- Dados abertos da Cidade de Curitiba: 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()
```

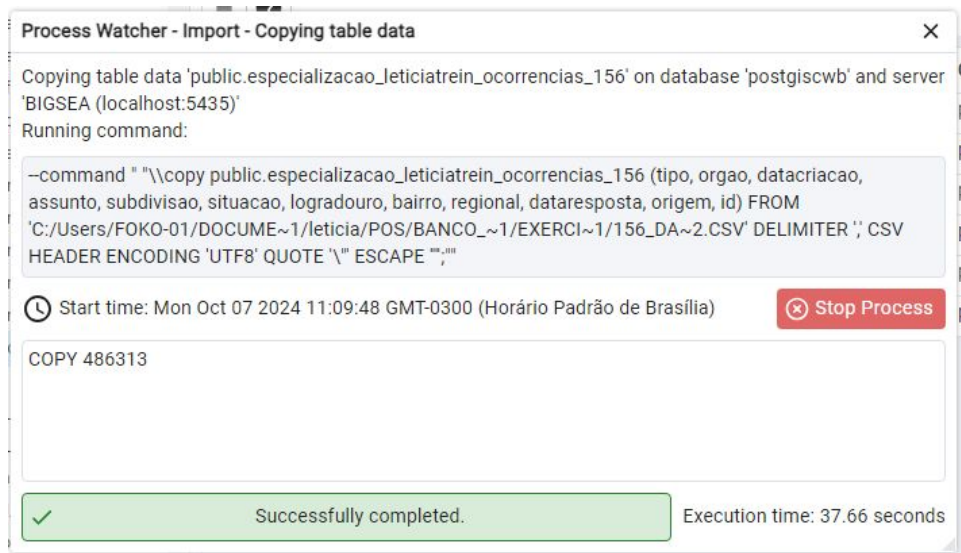
4 - Criação de coluna ID para chave primária de valor único

```
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  
);
```



Localização e número de Tuplas

- Caminho:
 - Servers ⇒ Databases ⇒ Schemas ⇒ public ⇒ Tables ⇒ especializacao_leticiatrein_ocorrencias_156
- Número total de Tuplas:
 - 486.313



```
19 SELECT COUNT(*) FROM ESPECIALIZACAO_LETICIATREIN_OCORRENCIAS_156
20
```

Data Output		Messages	Notifications
	count bigint		
1	486313		

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

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
10	Poluição	392

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
11	AGUA VERDE	11856
12	PINHEIRINHO	11482
13	PORTAO	10518
14	PILARZINHO	10111
15	SANTA FELICIDADE	10068

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
7	Maus tratos - Comércio de animais de estimação
8	Atendimento profissional

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
7	BIGORRILHO	53
8	UBERABA	49
9	CAJURU	49
10	XAXIM	47