



**Programa de Pós-Graduação em Epidemiologia em Saúde Pública**

**Disciplina: Desenho e Construção de Bases de Dados**

**Professores: Maria de Fátima Pina e Vanderlei Matos**

# **Banco de Dados Tuberculose no Brasil**

## **Trabalho Final**

### **Alunos**

**Alex Erickson**

**Maira Moreno**

**Maria Angela Esteves**

**Mércia Sacramento**

**Simone A. Borges Oliveira**

**Agosto de 2021**

## SUMÁRIO

<b>1. Contexto.....</b>	<b>3</b>
Tuberculose no Brasil.....	3
Tuberculose no Rio de Janeiro.....	5
<b>2. Objetivo.....</b>	<b>7</b>
<b>3. Modelo Conceitual.....</b>	<b>7</b>
<b>4. Modelo Lógico.....</b>	<b>10</b>
<b>5. Modelo Físico.....</b>	<b>13</b>
<b>6. Implementação e consultas.....</b>	<b>14</b>
<b>7. Problemas.....</b>	<b>22</b>
<b>8. Considerações finais.....</b>	<b>24</b>
Referências Bibliográficas.....	25

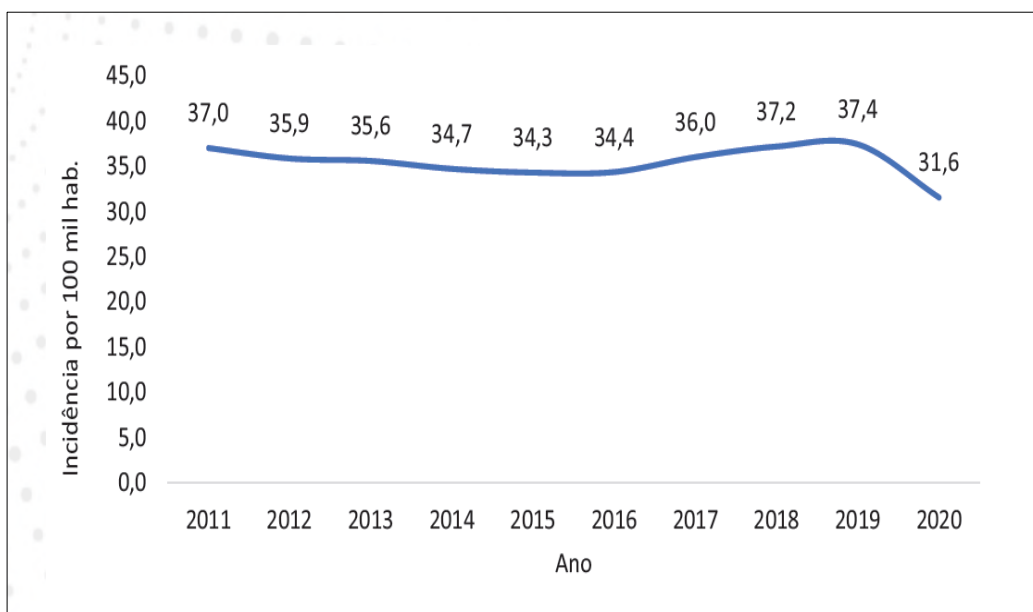
## 1. Contexto

### Tuberculose no Brasil

A Tuberculose (TB) é um problema de saúde pública no Mundo, agravado pela pandemia do Covid-19. O Brasil é considerado prioritário no controle da doença pela Organização Mundial de Saúde (OMS) por estar entre os 30 países de alta carga para a TB e para coinfeção TB-HIV. (WHO, 2020). Organizar os dados da doença é fundamental para analisá-los, subsidiando o planejamento de ações consistentes no controle da doença.

A Figura 1 apresenta a incidência de TB no Brasil, de 2011 a 2020. Observa-se a tendência de queda de 2011 a 2016, aliada a um aumento de incidência entre 2017 a 2019. Os dados de 2020 são sujeitos a mudança por trata-se de dados preliminares, mas apresentam tendência de queda.

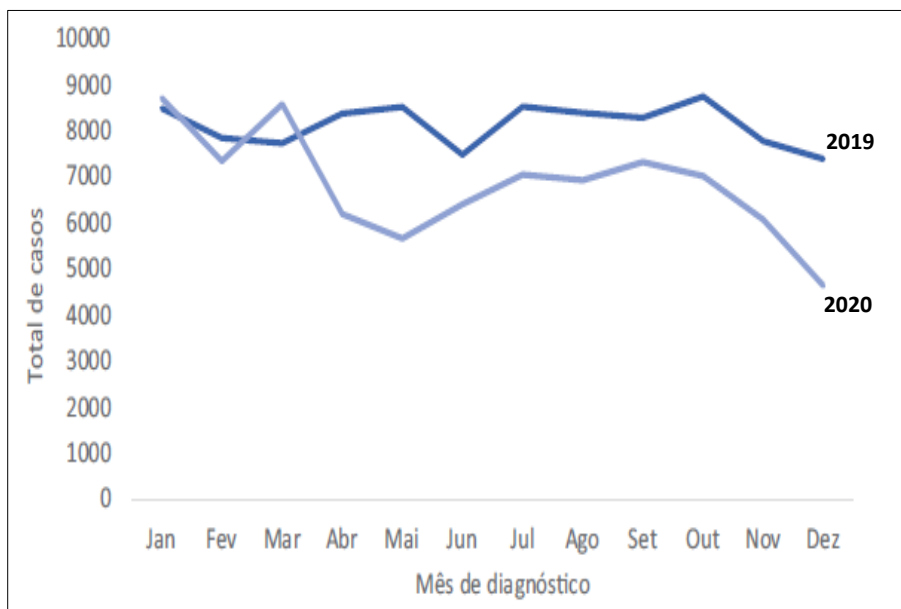
**Figura 1** - Coeficiente de incidência de tuberculose geral (por 100 mil hab.). Brasil, 2011 a 2020.



Fonte: Sistema de Informação de Agravos de Notificação/Secretarias Estaduais de Saúde/Ministério da Saúde; Instituto Brasileiro de Geografia e Estatística (BRASIL, 2021).

A Figura 2 apresenta o diagnóstico de TB no Brasil para 2019-2020, que evidencia a queda na notificação de casos em 2020 comparado a 2019. A redução foi mais pronunciada no mês de maio. O boletim epidemiológico do Ministério da Saúde (MS) também relata a redução (acima de 40%) do consumo de cartuchos de testagem no mesmo mês em 2020 (BRASIL, 2021).

**Figura 2** - Casos diagnosticados de tuberculose. Brasil, 2019 e 2020.

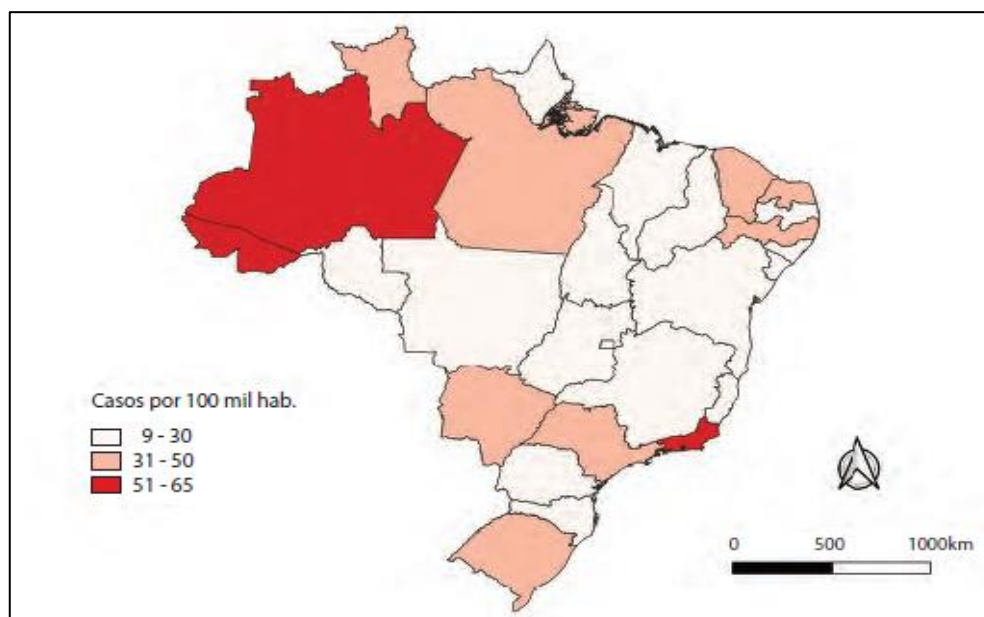


Fonte: Sistema de Informação de Agravos de Notificação (Sinan); Rede de Teste Rápido para Tuberculose/Coordenação Geral de Vigilância das Doenças de Transmissão Respiratória de Condições Crônicas/DCCI/SVS/MS (BRASIL, 2021).

Na estratificação por Unidade Federada (UF), evidencia-se uma importante heterogeneidade no país, com os maiores coeficientes de incidência de TB (acima de 51 casos/100 mil hab.) observados nos estados do Rio de Janeiro, Amazonas e Acre (Figura 3).

A queda de incidência e da notificação de TB em 2020, considerando a série histórica da doença do Brasil são atípicas. Não é possível avaliar ainda o impacto da pandemia na epidemiologia da doença, uma vez que estes resultados tanto podem significar a possibilidade de aumento da gravidade da doença como ser reflexo da sobrecarga operacional do Sistema Único de Saúde (SUS) por conta da Covid 19.

**Figura 3** - Coeficiente de incidência de tuberculose (por 100 mil hab.). Unidade Federada, 2020.

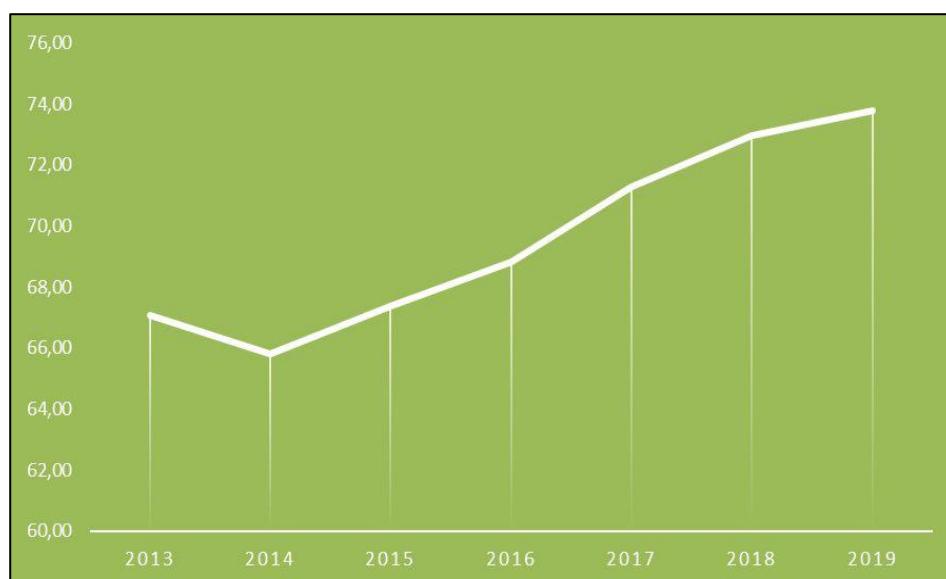


Fonte: Sistema de Informação de Agravos de Notificação/Secretarias Estaduais de Saúde/Ministério da Saúde; Instituto Brasileiro de Geografia e Estatística (BRASIL, 2021).

### Tuberculose no Rio de Janeiro

No Brasil, o Rio de Janeiro é um dos estados que apresentam as maiores taxas de incidência de TB. De fato, para o ano de 2019, a incidência da doença no estado foi de 72,4 casos por cada 100 mil habitantes (Figura 4), quase o dobro da incidência reportada para o Brasil inteiro durante o mesmo período (37,4 casos/100 mil hab. em 2019) (RIO DE JANEIRO, 2020, 2021).

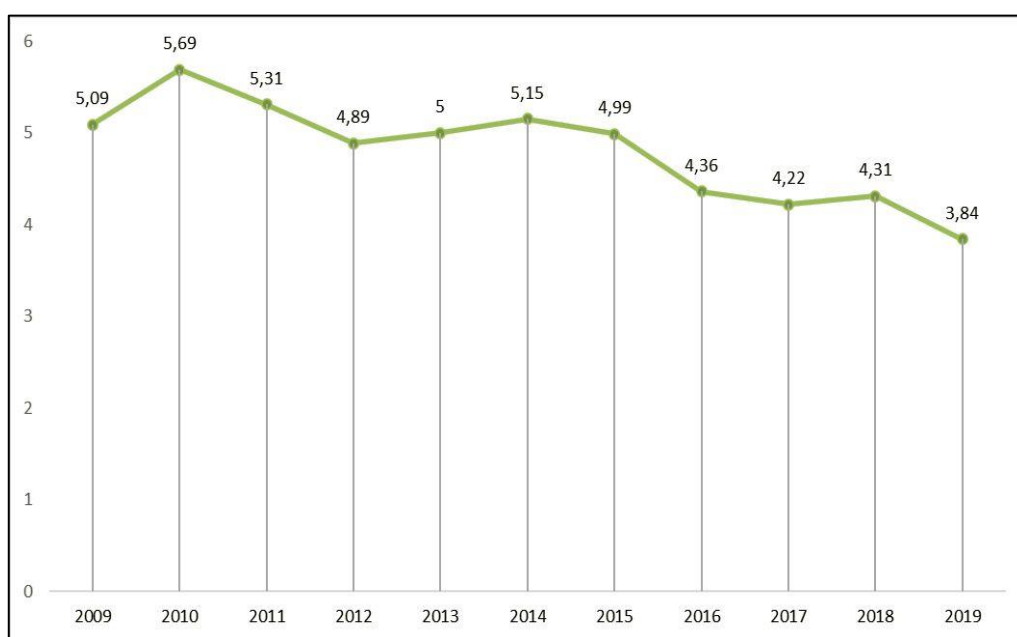
**Figura 4** - Série temporal da taxa de incidência de tuberculose no Estado do Rio de Janeiro, 2013-2019.



Fonte: SINAN GT/SES-RJ 10/12/2020; POP IBGE TCU (RIO DE JANEIRO, 2020, 2021).

Na Figura 5 observa-se uma tendência de redução da taxa de mortalidade por TB no Estado do Rio de Janeiro desde 2015. Contudo, esses valores são relativamente altos (3,84 óbitos/ 100 mil habitantes em 2019) no Sistema de Informação sobre Mortalidade (SIM), se for considerado que os demais indicadores, como incidência e cura da doença que vêm tendo piora a cada ano, parece que esses números não estariam refletindo a realidade do prognóstico da TB no estado, exigindo uma avaliação mais aprimorada da notificação compulsória da doença na região.

**Figura 5** - Série temporal da taxa de mortalidade por tuberculose no Estado do Rio de Janeiro, 2009 - 2019.



Fonte: TABNET 10/12/2020, SIM / SVS / SES-RJ; POP IBGE TCU; Ano 2019 - SIM SES-RJ 10/12/2020 (RIO DE JANEIRO, 2020, 2021).

Desta forma, tornar-se evidente que a TB é um agravo transmissível de grande relevância no Brasil e, particularmente, no estado do Rio de Janeiro, onde fatores externos como aumento da desigualdade social, a precariedade do sistema prisional no estado, aliado à crise de saúde pública pela atual pandemia da Covid-19 só pioraram a situação da doença na região.

## 2. Objetivo

O objetivo do trabalho foi o de desenhar e construir banco de dados para recuperar informações sobre a situação sócio-demográfica dos casos notificados e dos óbitos por tuberculose no Brasil, para dessa forma, facilitar posteriores análises dos dados epidemiológicos da doença, e contribuir com as ações de prevenção e controle.

## 3. Modelo Conceitual

É o modelo em que os objetos, suas características e relacionamentos representam o ambiente observado ou seja, a realidade observada é o ponto de partida para qualquer processo de modelagem.

Na década de 70, Peter Chen propôs a abordagem Entidade-Relacionamento (E-R), baseada nos modelos: Conceitual, Lógico e Físico.

A abordagem E-R propõe a elaboração de um modelo que represente os objetos observados e seus relacionamentos, independentemente de preocupações com a implementação.

Cabe lembrar, que a caracterização dos relacionamentos deve ser feita baseada em cardinalidade (ou grau do relacionamento), número de elementos que participam do relacionamento, condição de participação dos elementos no relacionamento e condição de estabelecimento do relacionamento.

Destacamos que a cardinalidade, de acordo com a quantidade de objetos envolvidos em cada lado do relacionamento, pode ser classificada basicamente de três formas:

- Relacionamento (1,1) - (um para um): cada uma das duas entidades envolvidas referênciam obrigatoriamente apenas uma unidade da outra.
- Relacionamento (1, n) - (um para muitos): uma das entidades envolvidas pode referenciar várias unidades da outra, porém, do outro lado cada uma das várias unidades referenciadas só pode estar ligada a uma unidade da outra entidade.
- Relacionamento (n, n) - (muitos para muitos): neste tipo de relacionamento cada entidade, de ambos os lados, pode referenciar múltiplas unidades da outra.

Foi utilizado o programa **br.Modelo 3.31** para construir esses relacionamentos.

Com o intuito de alcançar nosso objetivo foram pensadas algumas perguntas sobre as pessoas envolvidas neste estudo - os casos e os óbitos de TB:

- Quem são (eram)?
- Onde moram (moravam)?
- Onde se tratam (tratavam)?
- Tem (tinham) outras doenças além de TB?

Sob esse ponto de vista foi possível construir o modelo conceitual que apresentamos na Figura 6.

O modelo apresenta as entidades (tabelas) – **obitos, casos\_notif e municipios** – com relacionamentos entre si tendo cardinalidades de (0,n) e (0,1). Já as entidades - **obitos, casos\_notif e doencas** -, relacionadas a **causas e comorbidades**, apresentam cardinalidades de (0,n).

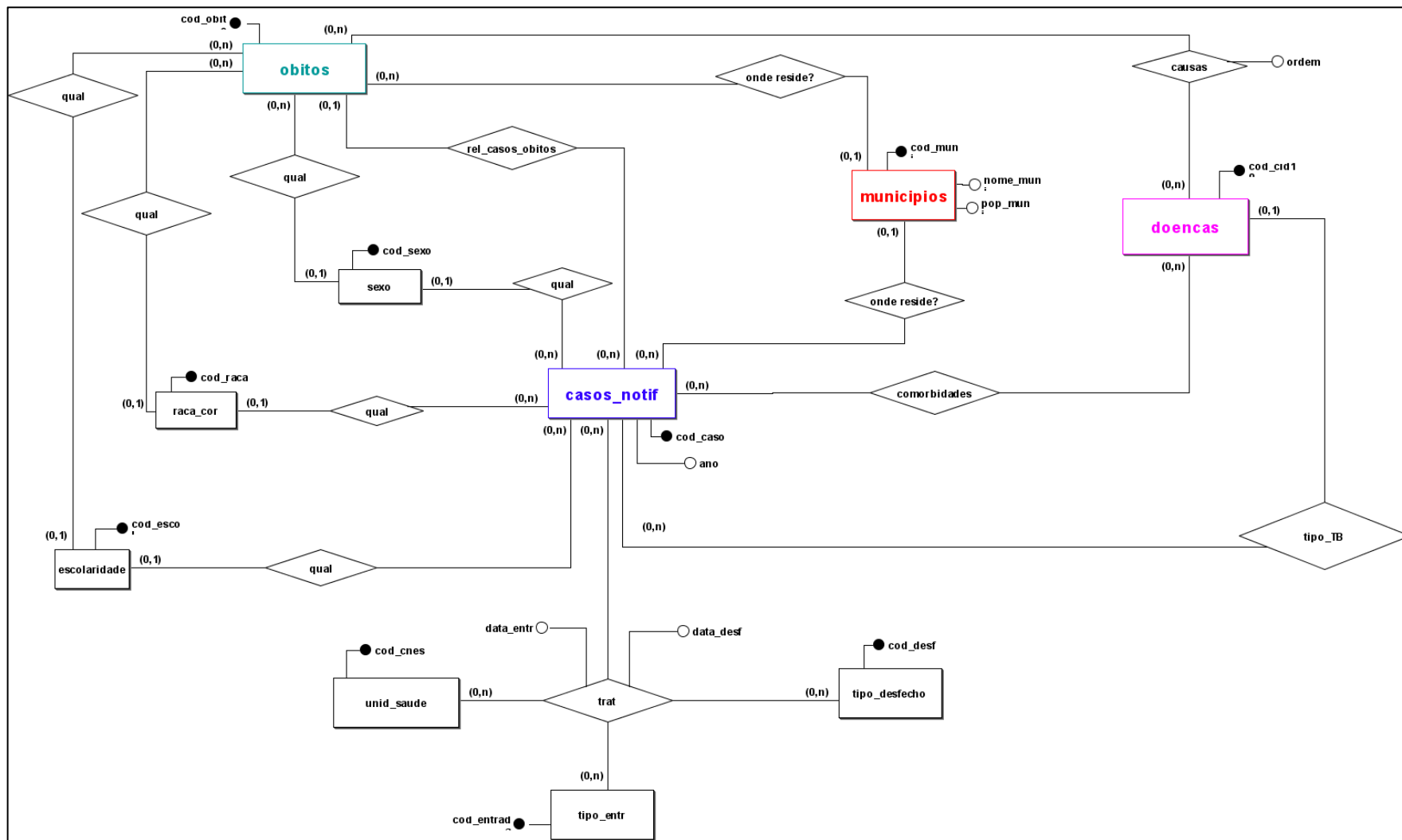
As entidades (tabelas) – **sexo, escolaridade e raca\_cor** – estão relacionadas as entidades **obitos e casos\_notif**, com cardinalidades de (0,n) e (0,1).

A entidade **casos\_notif** apresenta também um relacionamento terciario (trat) abrangendo o tratamento do paciente referente a **unidade de saúde, tipo de entrada e tipo de desfecho**, com cardinalidades de (0,n).

A versão de modelo conceitual, assim como no pgAdmin4 e no Access apresentada neste trabalho escrito, já acolhem todas as sugestões realizadas pela Prof.<sup>a</sup> Fátima durante a apresentação do grupo.



Figura 6 - Modelo Conceitual – Tuberculose.



#### 4. Modelo Lógico

No modelo lógico, os objetos, suas características e relacionamentos são representados de acordo com as regras de implementação e restrições impostas pelo *software* que se irá utilizar. No presente trabalho, optou-se pela utilização do Postgres (versão 13.1) em conjunto com o pacote Microsoft Access®.

A conversão do modelo conceitual em modelo lógico exige que sejam determinados os tipos de dados das variáveis (entidades e atributos) inseridas no modelo conceitual. Esta escolha deve ser realizada considerando não apenas a natureza do dado, mas também a sua extensão estimada.

Adicionalmente se faz importante a construção de um dicionário de variáveis, de forma a permitir que todos aqueles que venham utilizar o banco de dados, em qualquer das suas etapas (construção, implementação ou execução da coleta) tenham conhecimento sobre a natureza, características e conteúdo de cada variável presente que constitui o modelo em uso.

Na Figura 7, apresentamos o dicionário das variáveis do banco de tuberculose, de forma detalhada.

**Figura 7 - Dicionário das Variáveis.**

Tabela	Nome da variável	Tipo	Tamanho	Descrição da variável
casos_notif	cod_caso	numerico	12	código do caso notificado
	ano	numerico	4	ano do caso notificado
causas	ordem	numerico	6	ordem das causas da doença no sistema
doencas	cod_cid10	Texto Curto	255	código da Classificação Internacional de Doenças 10
escolaridade	cod_escol	numerico	2	código da escolaridade
municipios	cod_muni	numerico	12	código do municipio (IBGE)
	nome_muni	Texto Curto	255	nome do municipio
	pop_muni	numerico	12	estimativa da população do municipio
obitos	cod_obito	numerico	12	código da óbito
raca_cor	cod_raca	numerico	2	código de raça
sexo	cod_sexo	numerico	2	código do sexo
tipo_desfecho	cod_desf	numerico	2	código do desfecho do caso notificado
tipo_entr	cod_entrada	numerico	2	código do tipo de entrada do caso notificado
unid_saude	cod_cnes	numerico	20	código do Cadastro Nacional de Estabelecimentos de Saúde

O modelo lógico gerado para a tuberculose possui treze tabelas, três a mais que a quantidade de entidades existentes no modelo conceitual, decorrente da conversão de relacionamentos múltiplos em tabelas. O modelo lógico para tuberculose é apresentado na Figura 8.

As entidades que configuram as principais variáveis do modelo são: óbitos, casos notificados, doenças e municípios. Seus relacionamentos são descritos a seguir.

O relacionamento entre as entidades unidade de saúde, tipo de entrada e tipo de despecho deu origem a tabela “trat\_tipo\_despecho\_casos\_notif\_tipo\_entr\_unid\_saude”, que foi renomeada para “tratamento”, que por sua vez se relaciona com a tabela dos casos notificados. Por este motivo, a tabela tratamento recebe as chaves estrangeiras de todas as entidades com as quais se relaciona (unid\_saude, tipo\_entr, tipo\_despecho e casos\_notif).

A tabela de causas foi originada do relacionamento entre as entidades óbitos e doenças, que permite múltiplas entradas e apresenta um atributo de ordem para estas inserções. Desta forma, a tabela de causas recebe a chave estrangeira da tabela de doenças e da tabela de óbitos.

A tabela comorbidades é gerada a partir do relacionamento (n-m) entre as entidades casos notificados e doenças. Desta forma, a nova tabela criada recebe as chaves estrangeiras das entidades com as quais se relaciona (casos\_notif e doenças).

As tabelas municípios, sexo, escolaridade e raça/cor tiveram suas chaves primárias migradas como chaves estrangeiras para as tabelas de óbitos e casos notificados.

A migração das chaves entre as tabelas propicia que as tabelas receptoras concentrem as informações importantes sobre os dados aos quais se referem, de forma a facilitar a realização de associações e consultas posteriores.



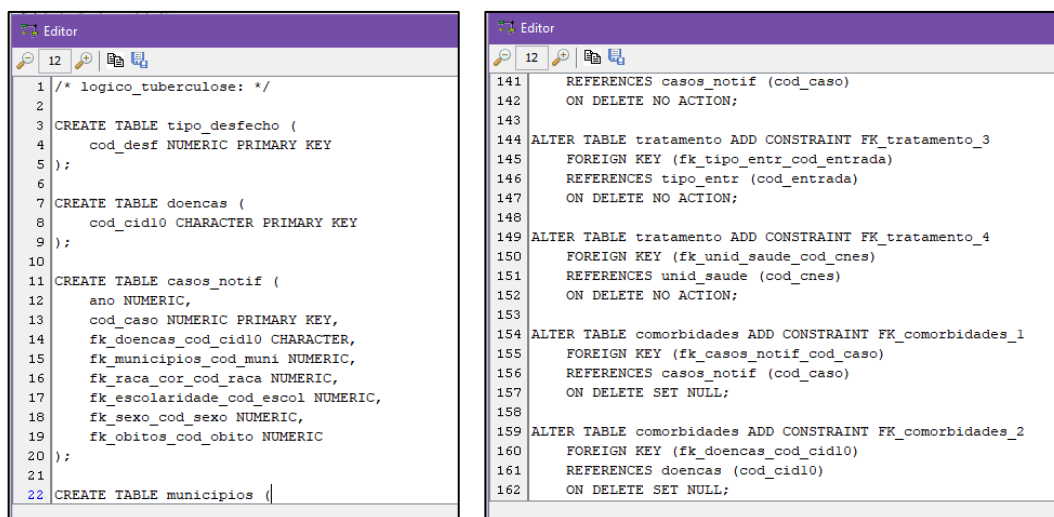
## 5. Modelo Físico

O modelo físico é a última etapa na modelagem de dados. Essa etapa é derivada do modelo lógico onde serão detalhadas a estrutura física do banco de dados, incluindo as tabelas, campos de valores, etc.

A conversão do modelo lógico para o físico é feita, também, pelo BrModelo, porém, antes de concluir essa etapa deve-se definir cada tipo de variável como descrito no dicionário das variáveis (Figura 7).

O modelo físico para a tuberculose foi gerado com 162 linhas de comando (Figura 9).

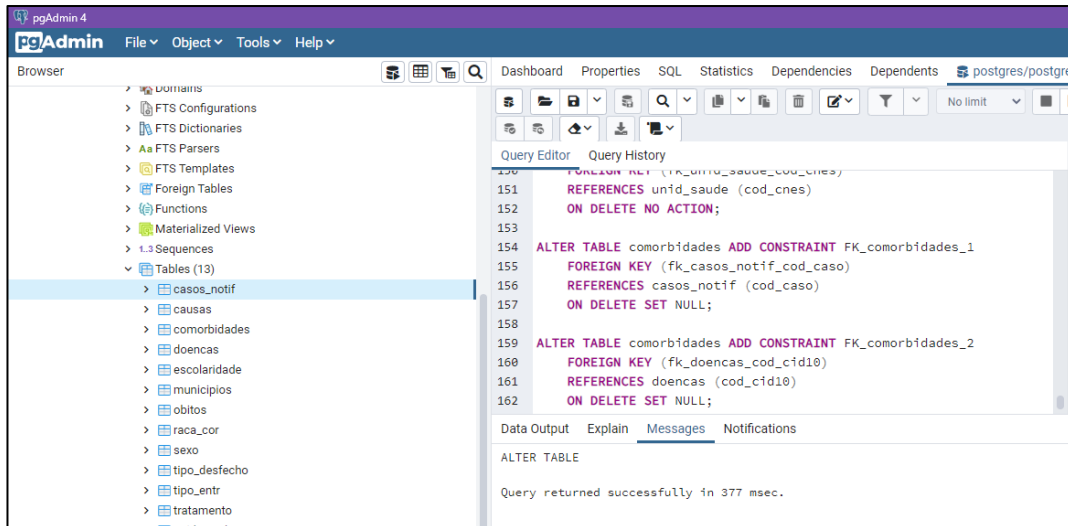
**Figura 9 - Modelo Físico - Tuberculose (Script com as linhas de comando).**



```
1 /* logico_tuberculose: */
2
3 CREATE TABLE tipo_desfecho (
4     cod_desf NUMERIC PRIMARY KEY
5 );
6
7 CREATE TABLE doencas (
8     cod_cid10 CHARACTER PRIMARY KEY
9 );
10
11 CREATE TABLE casos_notif (
12     ano NUMERIC,
13     cod_caso NUMERIC PRIMARY KEY,
14     fk_doencas_cod_cid10 CHARACTER,
15     fk_municipios_cod_muni NUMERIC,
16     fk_raca_cor_cod_raca NUMERIC,
17     fk_escolaridade_cod_escol NUMERIC,
18     fk_sexo_cod_sexo NUMERIC,
19     fk_obitos_cod_obito NUMERIC
20 );
21
22 CREATE TABLE municipios (
23
24 );
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141 REFERENCES casos_notif (cod_caso)
142 ON DELETE NO ACTION;
143
144 ALTER TABLE tratamento ADD CONSTRAINT FK_tratamento_3
145 FOREIGN KEY (fk_tipo_entr_cod_entrada)
146 REFERENCES tipo_entr (cod_entrada)
147 ON DELETE NO ACTION;
148
149 ALTER TABLE tratamento ADD CONSTRAINT FK_tratamento_4
150 FOREIGN KEY (fk_unid_saude_cod_cnes)
151 REFERENCES unid_saude (cod_cnes)
152 ON DELETE NO ACTION;
153
154 ALTER TABLE comorbidades ADD CONSTRAINT FK_comorbidades_1
155 FOREIGN KEY (fk_casos_notif_cod_caso)
156 REFERENCES casos_notif (cod_caso)
157 ON DELETE SET NULL;
158
159 ALTER TABLE comorbidades ADD CONSTRAINT FK_comorbidades_2
160 FOREIGN KEY (fk_doencas_cod_cid10)
161 REFERENCES doencas (cod_cid10)
162 ON DELETE SET NULL;
```

A conversão final apresenta os códigos necessários para materializar o banco de dados. Esses códigos são copiados e aplicados num outro programa pgAdmin4 (Figura 10) que é uma plataforma de desenvolvimento e administração para desenvolvimento de bancos de dados em recursos SQL. Os códigos copiados são rodados nesse programa que, caso não ocorra erros graves, geram bancos de dados no formato ODBC que podem ser transportados para outros programas de gerenciamento de banco de dados, por exemplo, o Access do Windows. Essa etapa é importante também, porque permite a detecção de possíveis erros persistentes na elaboração dos modelos.

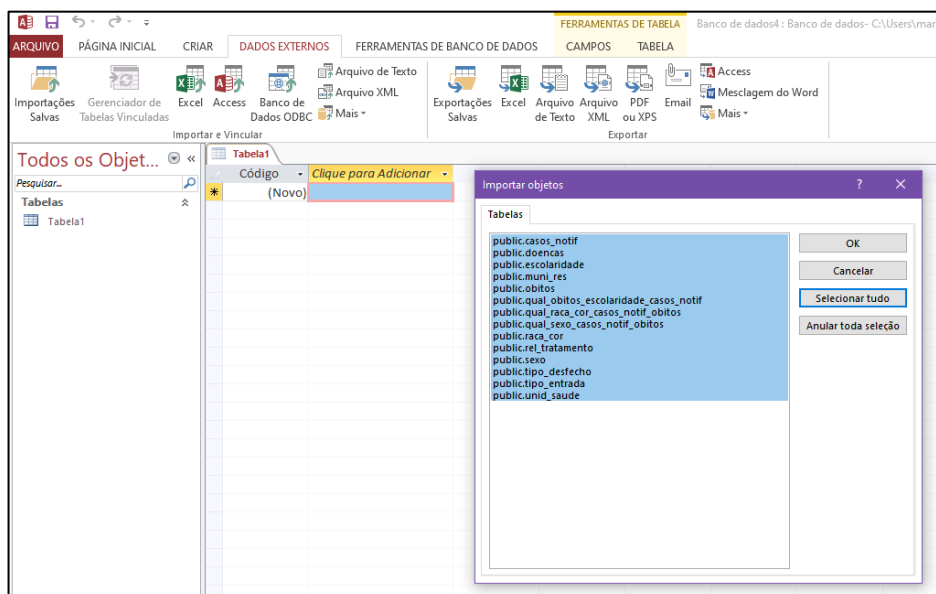
Figura 10 - Script do modelo físico tuberculose executado no pgAdmin4.



## 6. Implementação e consultas

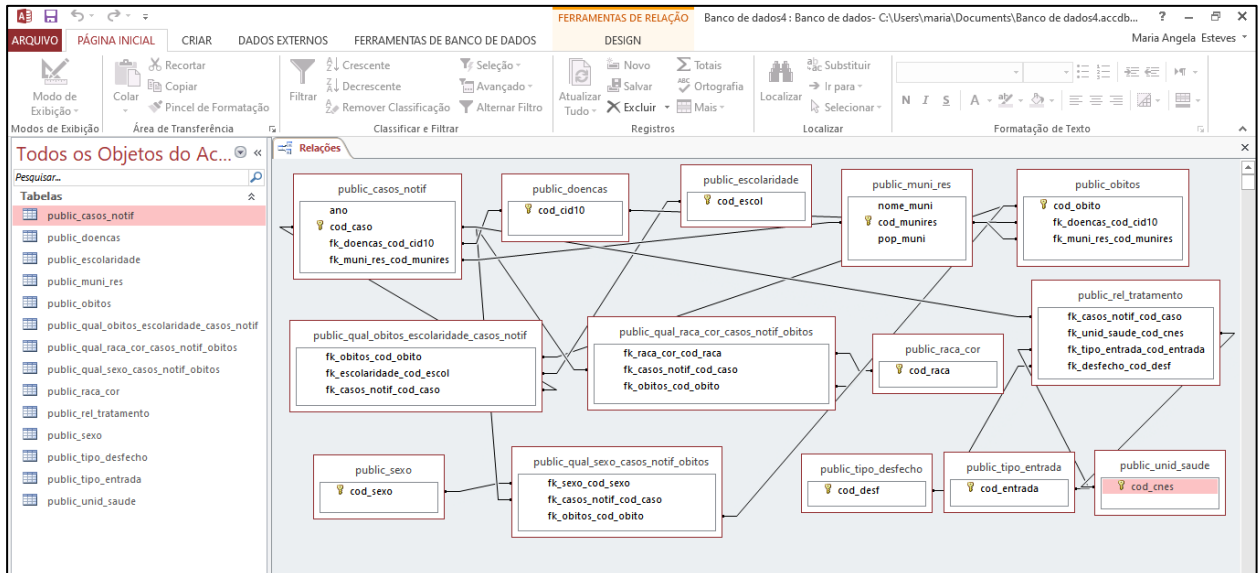
O programa de gerenciamento nesse projeto foi o Access do Windows. O banco de dados gerado no programa PgAdmin4 é importado para o Access e gerenciado com as ferramentas disponibilizadas deste programa (Figura 11).

Figura 11 - Importação do banco de dados no Access.



Após a importação, foi possível efetivar uma série de procedimentos e consultas do banco de dados. Para efeito de teste foram inseridos alguns dados nas tabelas. Inicialmente, fez-se o relacionamento manualmente entre as tabelas (Figura 12), conforme definido anteriormente no modelo conceitual, o que permitiu a elaboração de várias consultas.

**Figura 12** - Relacionamento gerado manualmente no Access entre as tabelas.



Nas Figuras de 13 a 18, apresentamos alguns exemplos de consultas que foram realizadas, com vários bancos de dados, em fase de testes.

Figura 13 - Consulta: relação entre as tabelas casos notificados e sexo.

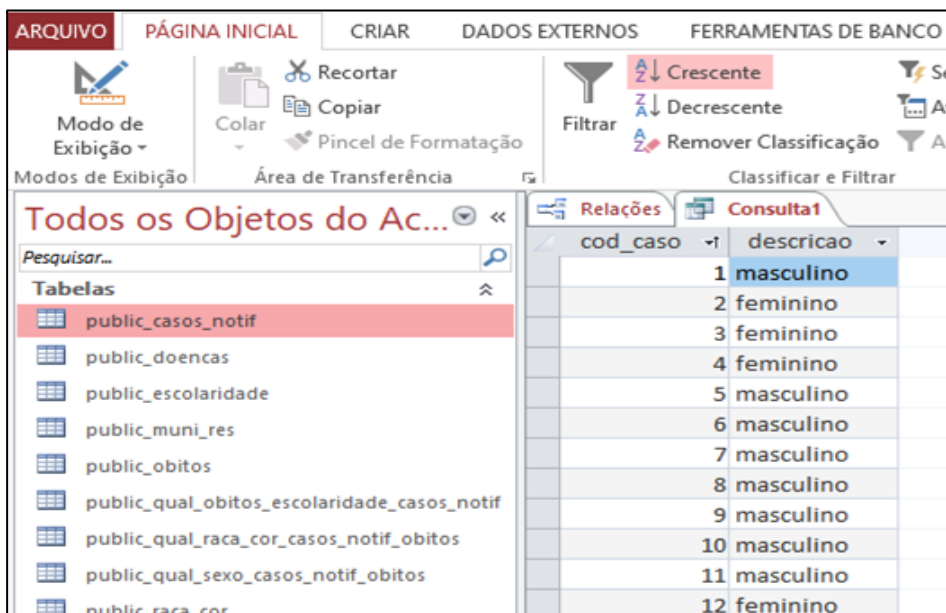
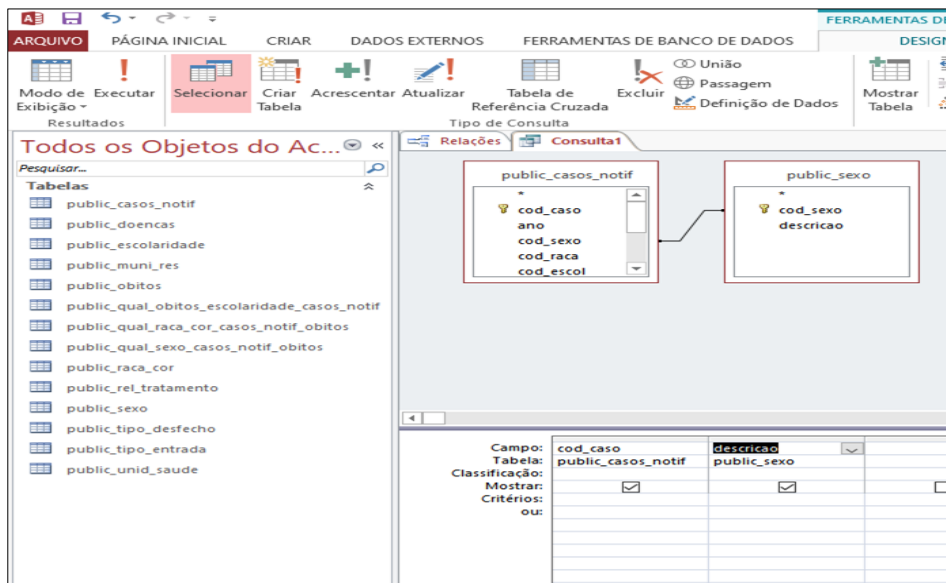
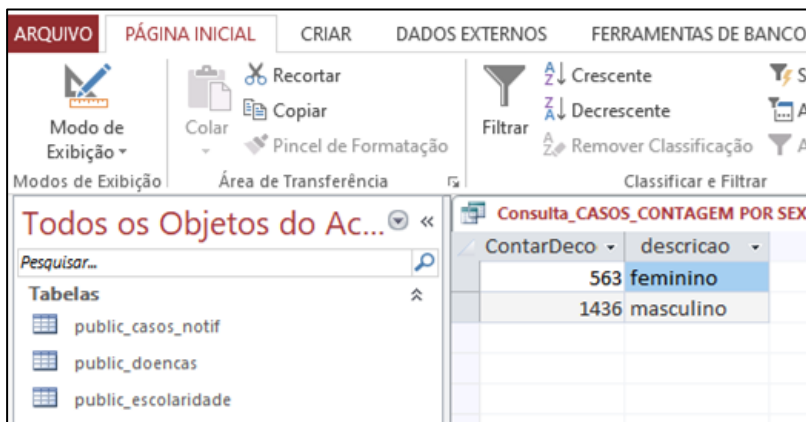
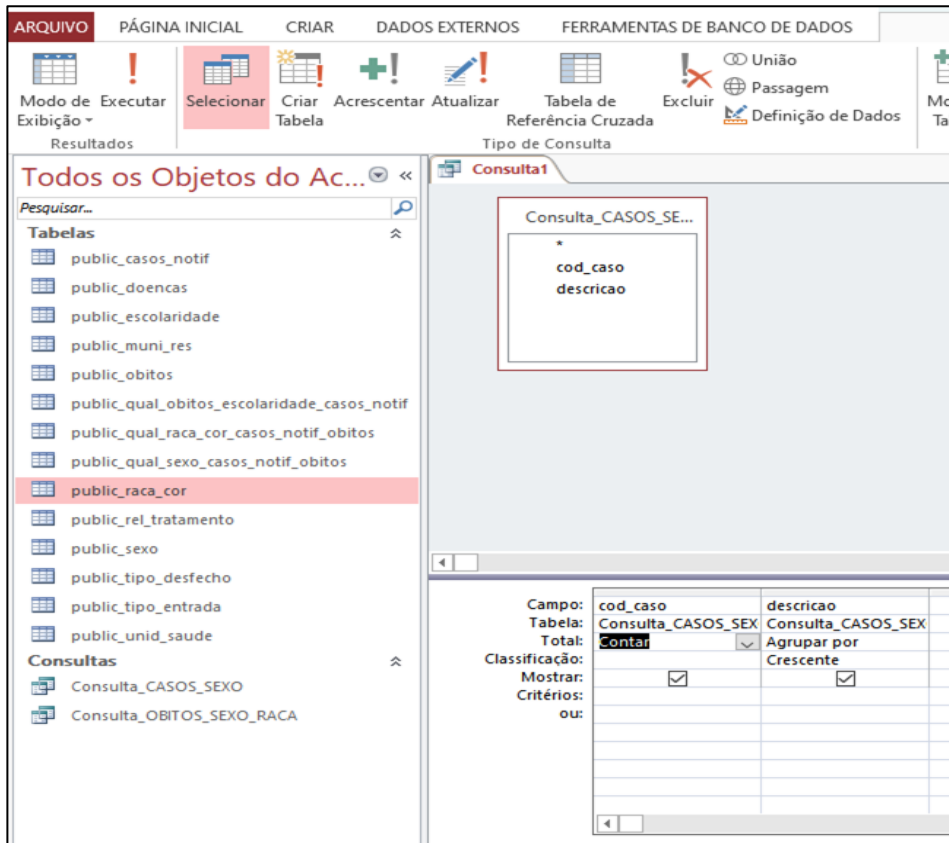




Figura 14 - Frequência simples: utilizando a consulta anterior nos casos notificados.



**Figura 15 - Consulta: utilizando a relação entre as tabelas: óbito, sexo e raça.**

Diagram illustrating the relationship between three tables: **public\_obitos**, **public\_sexo**, and **public\_raca\_cor**.

- public\_obitos** fields: **cod\_obito** (primary key), **data\_ob**, **idade**, **cod\_sexo**, **cod\_raca**.
- public\_sexo** fields: **cod\_sexo** (primary key), **descricao**.
- public\_raca\_cor** fields: **cod\_raca** (primary key), **descricao**.

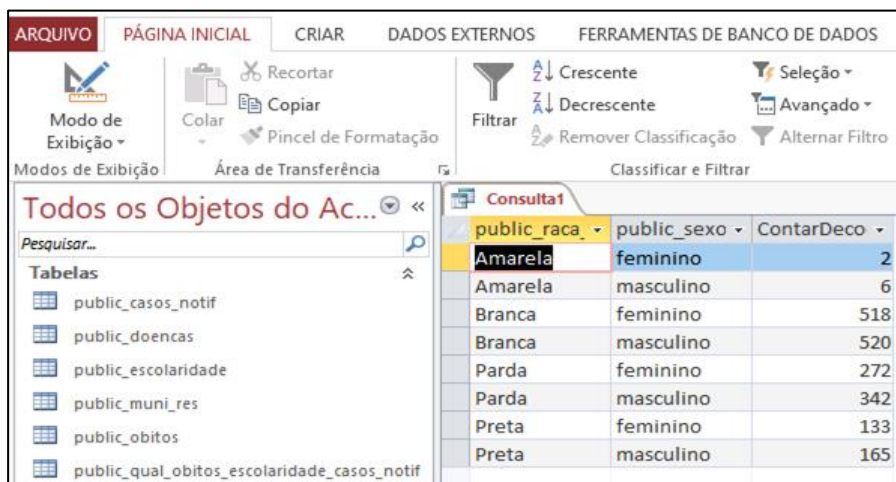
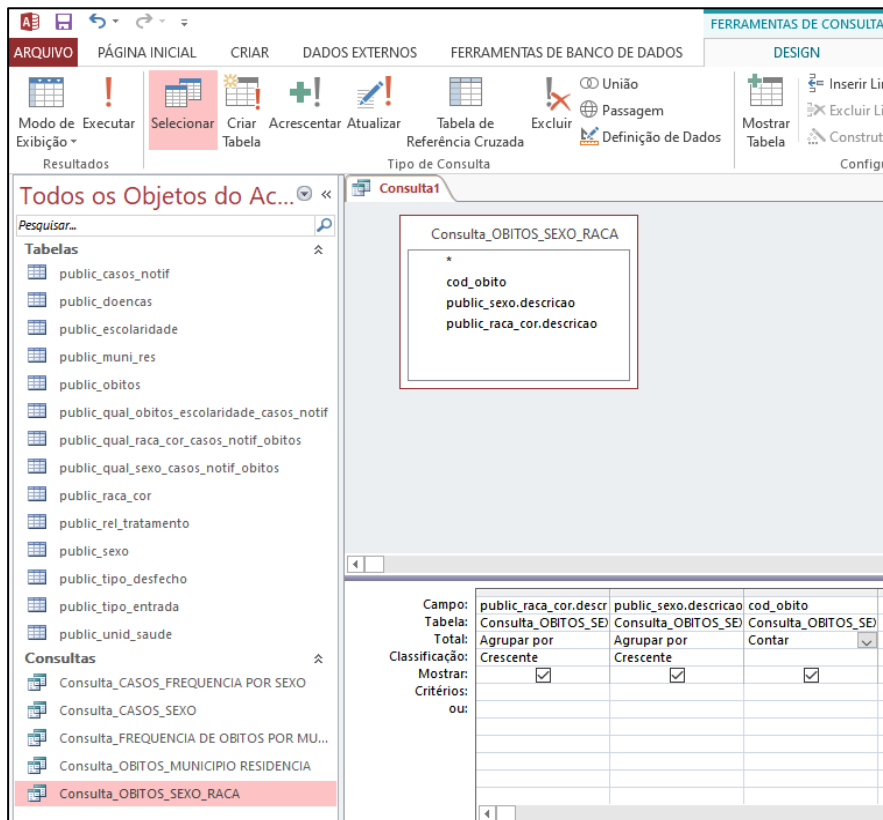
Relationships: **cod\_obito** (public\_obitos) is linked to **cod\_sexo** (public\_sexo). **cod\_sexo** (public\_sexo) is linked to **cod\_raca** (public\_raca\_cor).

Campo:	cod_obito	descricao	descricao
Tabela:	public_obitos	public_sexo	public_raca_cor
Classificação:			
Mostrar:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Critérios:			
ou:			

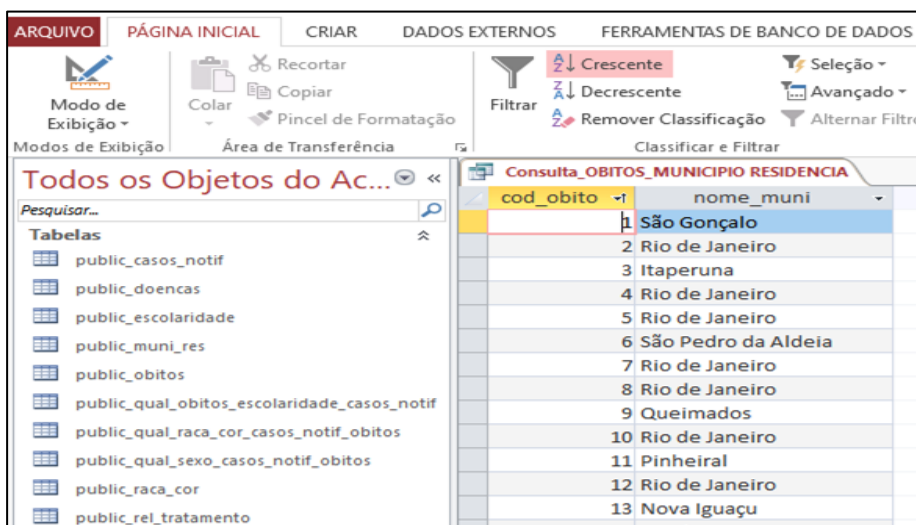
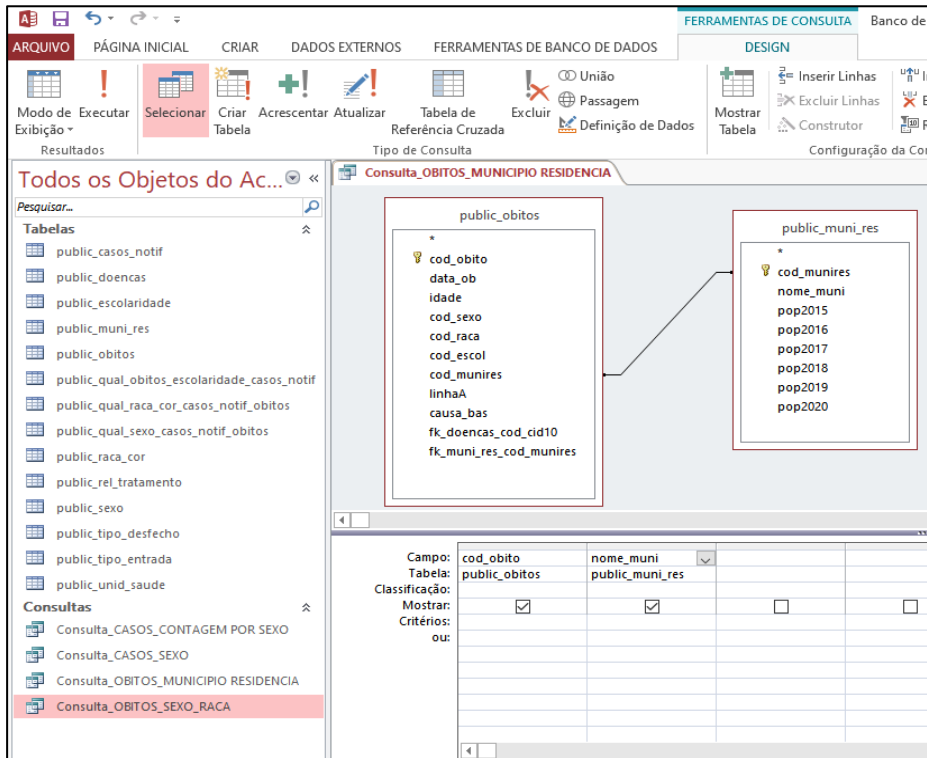
Query Name: **Consulta\_OBITOS\_SEXO\_RACA**

cod_obito	public_sexo	public_raca
1	feminino	Branca
2	masculino	Branca
3	masculino	Parda
4	masculino	Branca
5	feminino	Branca

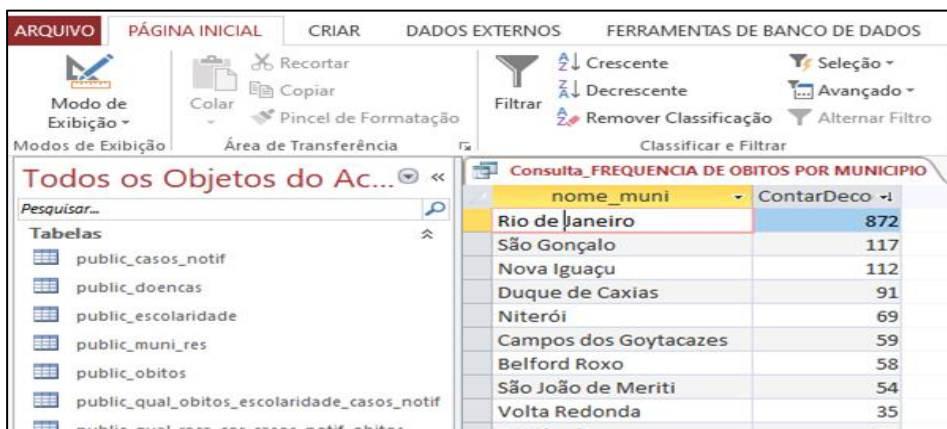
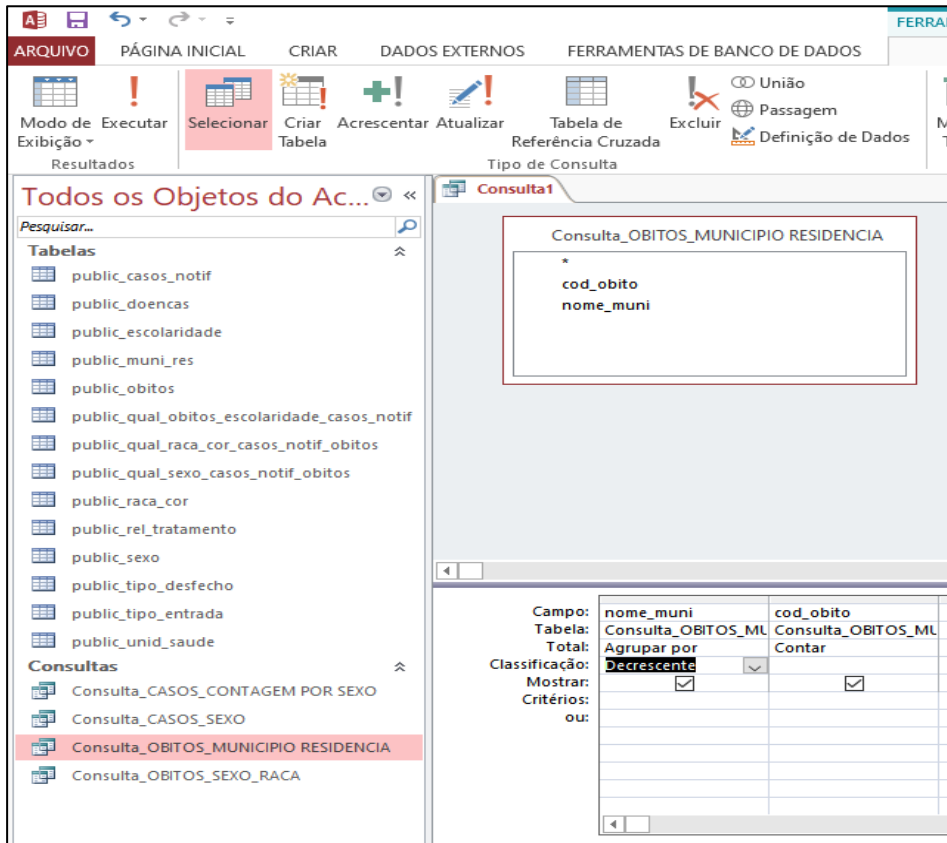
**Figura 16** - Frequência simples: utilizando a consulta anterior com os óbitos por raça e sexo.



**Figura 17** - Consulta: utilizando a relação existente entre as tabelas: óbitos e municípios.



**Figura 18** - Frequência simples: utilizando a consulta anterior com os óbitos e municípios.

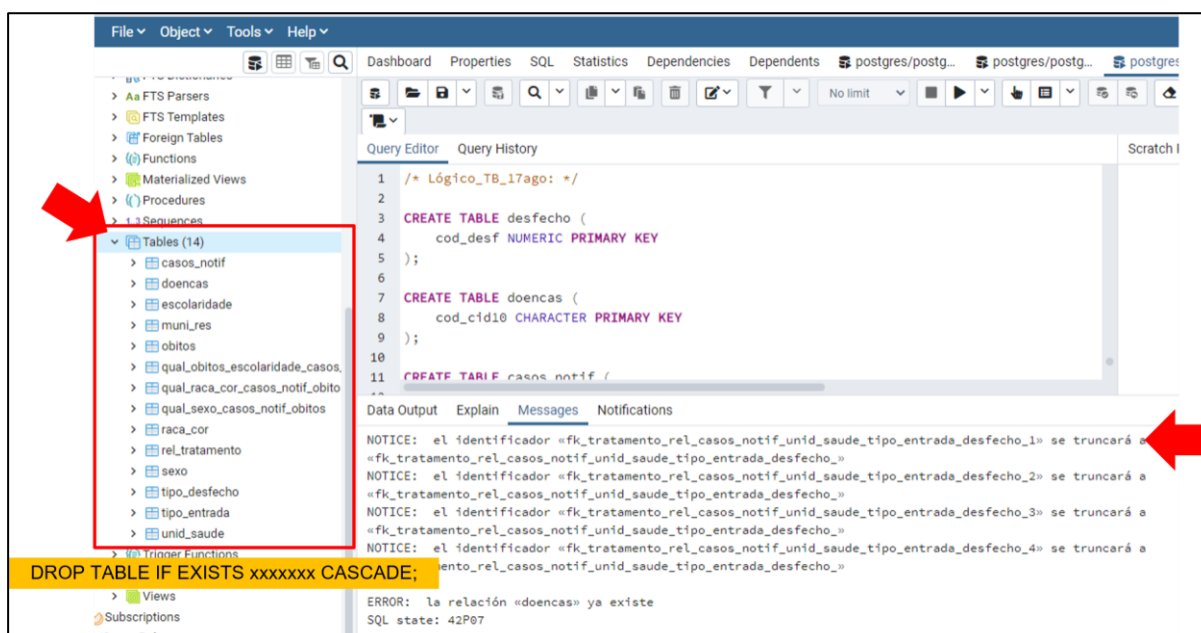


## 7. Problemas

Relatamos aqui alguns problemas encontrados durante a implementação do modelo físico no pgAdmin4 e como foram solucionados.

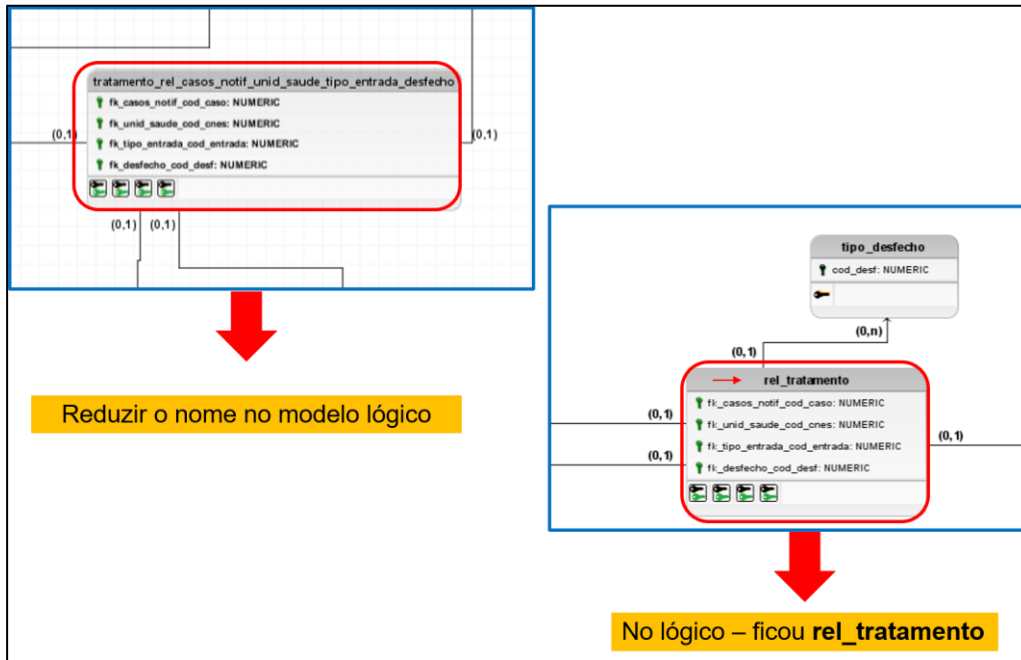
O **Erro #1: “relation already exist”** foi causado porque as tabelas de modelos de exercícios anteriores não foram removidas do servidor. Para solucionar o erro, foi preciso utilizar o script condicional “DROP TABLE IF EXISTS xxxxxxx CASCADE;” no editor Query do pgAdmin (Figura 19). Com a ajuda deste comando, o sistema foi instruído a descartar todas as tabelas, desde que estivessem vinculadas ou relacionadas a outras, e desta forma, foi eliminado o erro.

Figura 19 - Erro #1, “relation already exist”.



O **Erro #2: “name truncation”** ocorreu porque o nome de uma das tabelas (especificamente a tabela de relacionamento agregado) era muito grande. Para resolver o problema, bastou reduzir o tamanho do nome no modelo lógico, conforme mostrado na Figura 20.

Figura 20 - Erro #2, "name truncation".



## 8. Considerações Finais

A recuperação de informações após o relacionamento de dados de diferentes fontes aportou de forma rápida e prática novas informações sobre TB, pois, a partir dos relacionamentos, foi possível identificar por exemplo, a ocorrência de casos de TB por sexo e raça; a ocorrência de mortalidade por TB por sexo e raça, e a frequência de casos de TB e de óbitos por TB por local de ocorrência, todos eles indicadores epidemiológicos relevantes.

O relacionamento de dados através da modelagem de banco de dados, pode ser uma ferramenta para a melhoria das informações dos agravos quanto à completude, consistência e confiabilidade dos sistemas de notificação compulsória de TB, pois facilitaria a implementação de sistemas de informação que possam se comunicar entre eles e que reflitam a realidade da doença.

O uso de dados secundários que, embora sejam provenientes dos Sistemas de Informação Nacionais Brasileiros, apresentam com frequência inconsistências, omissões e com completude ruim (ROMERO & CUNHA, 2006).

A necessidade de uso de diferentes programas para o desenho e implementação de banco de dados pode ser uma limitação importante para a adoção da metodologia de modo amplo nas atividades de pesquisa ou de gestão.

De todas as etapas do processo de desenho e construção de base de dados, a etapa de elaboração do modelo conceitual é a mais crítica e onde se deve dedicar o maior tempo de reflexão. Importa nesta etapa, elaborar bem as perguntas de pesquisa que se quer responder com o banco de dados, bem como conhecer a dinâmica da realidade estudada e suas relações com os dados, na prática.



## Referências Bibliográficas

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Boletim Epidemiológico Especial - Tuberculose**, Número Especial | mar. 2021. Disponível em: <http://www.aids.gov.br/pt-br/pub/2021/boletim-tuberculose-2021>. Acesso em: 09ago.2021.

RIO DE JANEIRO. Boletim Epidemiológico de Tuberculose no estado do Rio de Janeiro Nº 01/2020. Secretaria de Estado de Saúde do Rio de Janeiro. 31 de outubro de 2020. Disponível em: <http://www.riocomsaude.rj.gov.br/Publico/MostrarArquivo.aspx?C=j%2FgTIADlpqE%3D>. Acesso em: 10ago.2021.

RIO DE JANEIRO. Cenário epidemiológico: Tuberculose no Estado do Rio de Janeiro. Análise do ano de diagnóstico – 2019. Secretaria de Estado de Saúde do Rio de Janeiro. Subsecretaria de Vigilância em Saúde. RJ. **Boletim Epidemiológico Tuberculose** Nº 01/2021. Disponível em: <http://www.riocomsaude.rj.gov.br/Publico/MostrarArquivo.aspx?C=j%2FgTIADlpqE%3D>. Acesso em: 10ago.2021.

ROMERO, ED; CUNHA, CB. Avaliação da qualidade das variáveis epidemiológicas e demográficas do Sistema de Informações sobre Nascidos Vivos, 2002. **Cad Saude Publica**. v. 23, n. 3, p. 701-714, 2007.

WHO. World Health Organization. **Global Tuberculosis Report 2020**. Geneva: WHO. 2020. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf>. Acesso em: 09ago.2021.