

## Notificações de Casos de Dengue no Estado do Rio de Janeiro

Pedro Henrique Honorio João Pedro Silva de Sousa  
Saito 122122366  
122149392

Gabriel Henrici  
122137997

# Relatório 1

Data Warehousing e Inteligência de Negócio (ICP-361) - 2025/2

## 1. Descrição do Problema

A dengue permanece como um dos principais desafios de saúde pública no Brasil, com ciclos epidêmicos recorrentes e forte impacto sobre o sistema de vigilância e assistência. No Estado do Rio de Janeiro, a elevada densidade populacional, as variações climáticas e as condições socioambientais favorecem a disseminação do mosquito ***Aedes aegypti***, vetor do vírus da dengue. Apesar da existência de sistemas de notificação, como o Sinan, a análise detalhada dos registros ainda é limitada pela fragmentação dos dados e pela dificuldade de identificar padrões temporais e espaciais de ocorrência.

Diante desse cenário, torna-se fundamental compreender a evolução dos casos notificados ao longo dos últimos cinco anos (2021–2025), investigando tendências, sazonalidade, distribuição geográfica e possíveis associações com variáveis demográficas e ambientais. Essa análise pode subsidiar estratégias mais eficazes de prevenção, controle e resposta a surtos, contribuindo para o fortalecimento da vigilância epidemiológica no estado.

## 2. Dicionário de Dados

Esta seção descreve o dicionário de dados do projeto, apresentando as tabelas, campos e respectivas descrições utilizadas no banco de dados.

### **2.0.1. Dimensão DData**

| Campo                 | Descrição  |
|-----------------------|--|
| pk_tempo              | Chave surrogada da data  |
| data                  | Data completa  |
| dia                   | Dia  |
| mes                   | Mês  |
| ano                   | Ano  |
| trimestre             | Número do trimestre  |
| semestre              | Número do semestre   |
| estacao               | Estação do ano   |
| semana_epidemiologica | semana do calendário epidemiológico padronizado da notificação |
| ano_epidemiologico    | ano do calendário epidemiológico padronizado da notificação    |

**Tabela 1:** Dimensão de Tempo

## 2.0.2. Dimensão DPaciente

| Campo          | Descrição  |
|----------------|--|
| pk_paciente    | Chave surrogada  |
| idade          | Idade do paciente  |
| ano_nascimento | Ano de nascimento do paciente  |
| grupo_risco    | Flag binária que indica se o paciente apresenta risco de vida visto os seus sinais clínicos. |

**Tabela 2:** Dimensão com informações sobre o paciente a qual a notificação se refere

## 2.0.3. Dimensão DDemografia

| Campo         | Descrição   |
|---------------|---|
| pk_demografia | Id da combinação de valores de demografia   |
| faixa_etaria  | Faixa etária do paciente  |
| sexo          | Sexo do paciente  |
| raca_cor      | Raça ou cor do paciente   |
| escolaridade  | Grau de escolaridade do paciente  |
| gestante      | Flag binária para saber se o paciente é gestante ou não no momento da notificação |

**Tabela 3:** Dimensão com informações demográficas do paciente

## 2.0.4. Dimensão DSinais\_Clínicos

| Campo            | Descrição                                      |
|------------------|--|
| pk_sinal_clinico | Id da combinação de valores de sinais clínicos |
| cefalia          | Flag binária indicando cefalia                 |
| mialgia          | Flag binária indicando mialgia                 |
| vomito           | Flag binária indicando vômito                  |
| nausea           | Flag binária indicando náusea                  |
| febre            | Flag binária indicando febre                   |
| dor_retro        | Flag binária indicando ...                     |

**Tabela 4:** Dimensão que indica presença de sintomas do paciente da notificação

### 2.0.5. Dimensão DLocal

| Campo                  | Descrição                                      |
|------------------------|--|
| pk_local               | Id da localização                              |
| municipio_id           | Id do município                                |
| mesorregiao_id         | Id da mesorregião do estado do Rio de Janeiro  |
| mesorregiao_nome       | Nome da mesorregião                            |
| municipio_nome         | Nome do município                              |
| microrregiao_id        | Id da microrregião do estado do Rio de Janeiro |
| microrregiao_nome      | Nome da microrregião do estado                 |
| populacao_municipio    | Tamanho da população do município              |
| populacao_mesorregiao  | Tamanho da população da mesorregião            |
| populacao_microrregiao | Tamanho da população da microrregião           |

**Tabela 5:** Dimensão geográfica que exerce o papel da residência do paciente e local de origem da notificação

### 2.0.6. Dimensão DTipo\_Notificacao

| Campo               | Descrição   |
|---------------------|---|
| pk_tipo             | Id da combinação de valores dos tipos                                   |
| flag_obito          | Flag binária para indicar se a notificação é um óbito                   |
| flag_hospitalizacao | Flag para indicar se é uma notificação de hospitalização                |
| num_estado          | Código do estado do paciente  |
| descricao_estado    | Descrição do estado (cura, óbito por dengue ou por outras causas, etc.) |

**Tabela 6:** Dimensão com informações adicionais sobre a natureza da notificação

### 2.0.7. Fato FNotificação

| Campo                | Descrição  |
|----------------------|--|
| pk_notificacao       | Chave surrogada da notificação                       |
| fk_data              | Chave estrangeira da data da notificação             |
| fk_demografia        | Chave estrangeira da demografia do paciente          |
| fk_paciente          | Chave estrangeira do paciente                        |
| fk_local_notificacao | Chave estrangeira do local de emissão da notificação |
| fk_local_residencia  | Chave estrangeira do local de residência do paciente |

**Tabela 7:** Tabela Fato com informações sobre cada notificação emitida

## 2.1. Dados Removidos

Antes da modelagem dimensional, foi feita uma análise preliminar para eliminar variáveis com alta taxa de nulos ou baixa variância. Identificaram-se 64 colunas com mais de 70% valores ausentes. No momento, optamos por descartar essas variáveis, porém elas poderão retornar em etapas futuras do projeto.

Dentre as colunas removidas por falta de variância (todos valores idênticos), podemos destacar as seguintes variáveis de sinais clínicos:

|            |            |            |
|------------|------------|------------|
| EXANTEMA   | LEUCOPENIA | ARTRALGIA  |
| DOR_COSTAS | LACO       | PETEQUIA_N |
| CONJUNTVIT | DIABETES   | HIPERTENSA |
| ARTRITE    | HEMATOLOG  | ACIDO_PEPT |
| RENAL      | HEPATOPAT  | AUTO_IMUNE |

O restante das variáveis desconsideradas podem ser encontradas no [repositório](#).

## 2.2. Dados Derivados

Entre as colunas apresentadas, algumas foram derivadas de outras ou resultaram do cruzamento com bases externas. São elas:

| Dimensão DPaciente | Dimensão DLocal        | Dimensão DDemografia |
|--------------------|------------------------|----------------------|
| grupo_risco        | mesorregiao_id         | faixa_etaria         |
|                    | mesorregiao_nome       |                      |
|                    | microrregiao_id        |                      |
|                    | microrregiao_nome      |                      |
|                    | populacao_microrregiao |                      |
|                    | populacao_municipio    |                      |
|                    | populacao_mesorregiao  |                      |

## 2.3. Observações

- Por ora, a tabela fato não considera nenhuma métrica associada diretamente
- Embora improvável, é possível que as chaves estrangeiras da tabela fato não formem uma chave única de forma a constituir a chave primária do fato. Desse modo, optou-se por adicionar uma coluna pk\_notificacao que se comporta como chave substituta para compor a chave primária do fato.

### 3. Análises

Para cada dimensão, foram definidas as análises apresentadas a seguir, com o objetivo de identificar padrões e fatores associados à ocorrência de casos de dengue.

| Dimensão    | Objetivo   |
|-------------|--|
| Clínica     | Compreender padrões de sintomas e gravidade dos casos.         |
| Espacial    | Identificar áreas prioritárias para vigilância e controle.     |
| Temporal    | Detectar surtos, sazonalidades e tendências ao longo do tempo. |
| Demográfica | Detectar grupos populacionais mais vulneráveis.                |

**Tabela 8:** Tabela descrevendo as dimensões e os objetivos de cada dimensão.

#### 3.1. Análises Geográficas

##### Residência e Notificação

- As notificações serão analisadas considerando tanto o local de residência do paciente quanto o local de notificação, permitindo observar fluxos entre municípios.

##### Notificações de Contratação da Doença

- Distribuição espacial por município e região.
- Identificação de municípios e regiões com maior concentração de casos.
- Comparação entre local de residência e local de notificação.

##### Notificações de Óbito

- Distribuição geográfica de óbitos por município e região.
- Identificação de áreas críticas com maior letalidade.

#### 3.2. Análises de Sazonalidade

- Identificação de épocas do ano com picos de notificação e óbitos.
- Análise da sazonalidade por estação, trimestre e semestre.
- Avaliação de variações interanuais (2021-2025) para detecção de ciclos epidêmicos.

#### 3.3. Análises Demográficas

- Distribuição das notificações por sexo, faixa etária, raça/cor, escolaridade e ocupação.
- Identificação de grupos vulneráveis, incluindo gestantes e faixas etárias com maior incidência ou letalidade.
- Avaliação da diferença de padrões clínicos entre grupos demográficos.

#### 3.4. Análises de Sintomas

- Identificação de sintomas mais relacionados a óbitos ou casos graves.
- Distribuição de sintomas por faixa etária e sexo.
- Associação entre conjuntos de sintomas e evolução clínica (cura, óbito, caso grave).

#### 3.5. Análises de Proporção

- Cálculo da incidência de notificações e óbitos por 100 mil habitantes.
- Taxas de óbito e letalidade por município e grupo demográfico.
- Proporção de casos graves e leves em relação ao total de notificações, inferida a partir dos sintomas e evolução.

## 4. Modelo

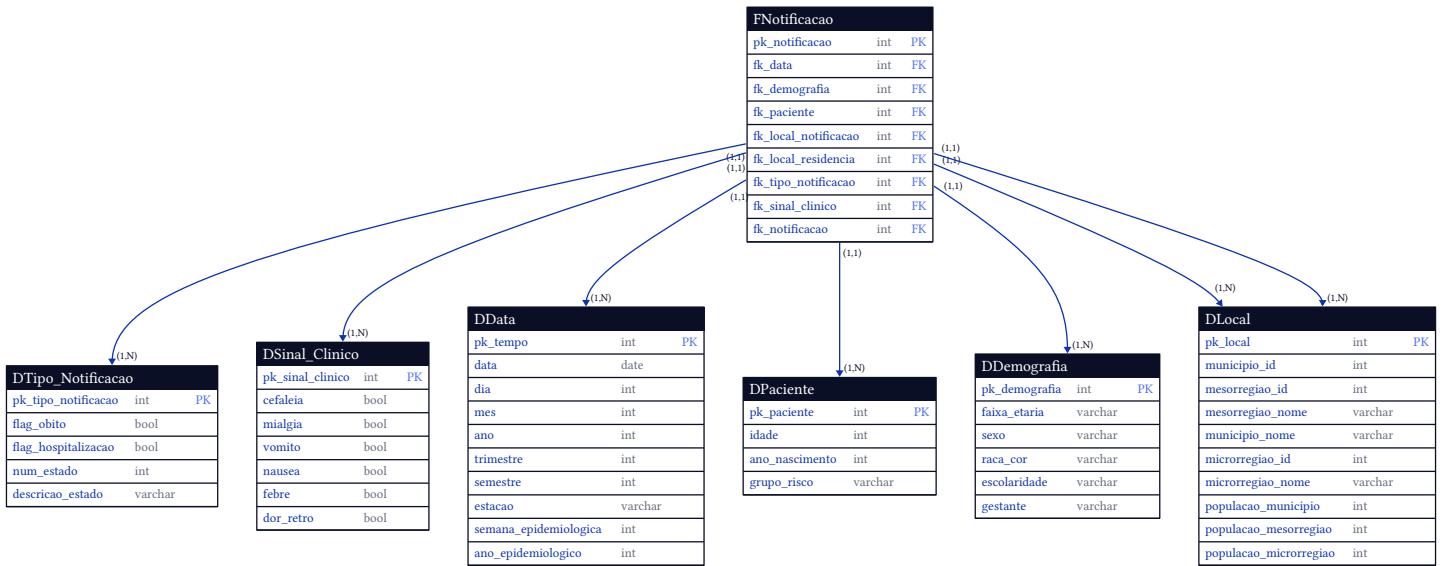


Figura 1: Modelo dimensional atual.