

# Análise dos Dados Ambulatórios em Pernambuco

Alan Victor Da Silva Torres  
av.torres@ufrpe.br

Fonte dos Dados: <https://datasus.saude.gov.br/>

Aplicação OLAP: <link para a aplicação OLAP>

Github: <link do projeto no github>

# ETAPA 1 - PLANEJAMENTO

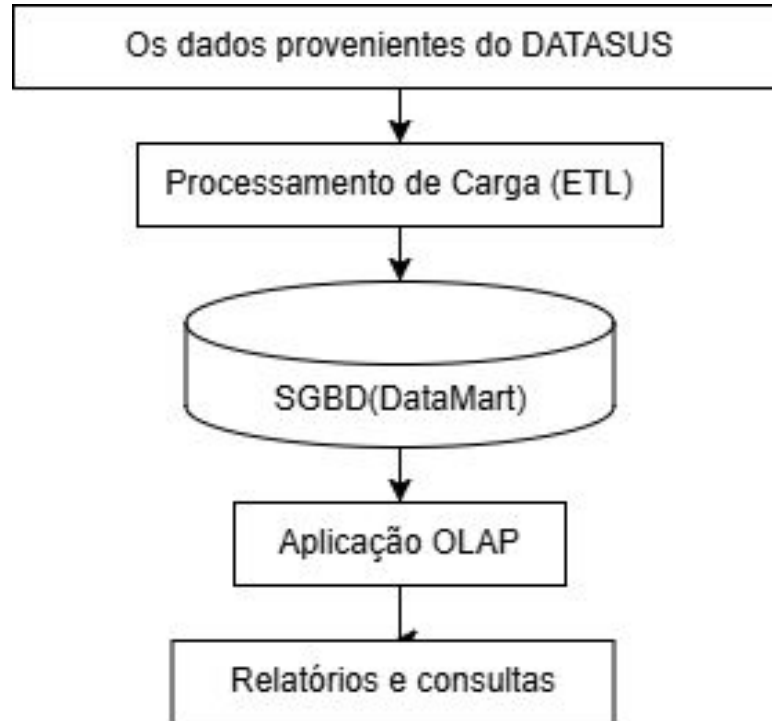
# 1. Contextualização

Os dados se tornaram extremamente importantes na sociedade atual. Quando falamos de dados relacionados à saúde, essa importância se torna ainda mais evidente. No entanto, nem sempre esses dados são utilizados de maneira eficiente. Por isso, este projeto visa criar um data mart com informações ambulatoriais do estado de Pernambuco. O objetivo é obter insights relevantes, como identificar épocas com maior incidência de ocorrências e regiões com maiores necessidades. Essas informações são essenciais para aprimorar a gestão da saúde nos municípios pernambucanos.

## 2. Escopo/objetivo do Data Mart

Como objetivo principal, este data mart busca criar insights que auxiliem na gestão estadual do sistema público de saúde, permitindo a formulação de estratégias mais eficientes para aprimorar a saúde pública no estado. Para isso, foram analisados dados ambulatoriais do estado de Pernambuco, gerados pelo DATASUS no período de 2020 a 2022.

### 3. Arquitetura Tecnológica



## 4. Processo

Para construir o Data Mart, seguiremos um processo estruturado. Primeiro, definiremos os requisitos provenientes dos dados obtidos no DATASUS. Em seguida, extraímos os dados relevantes para a análise, selecionando as informações específicas da região de Pernambuco. Realizaremos transformações para limpeza e padronização, como a exclusão de registros nulos, a fim de obter uma base de dados mais consistente. Após aplicar o processo ETL, os dados poderão ser carregados no Data Mart. Em seguida, utilizaremos esses dados na aplicação OLAP para gerar análises e métricas necessárias, avaliando os relatórios.

## 4. Abordagem

Para a abordagem, escolhemos o método Bottom-Up, proposto por Ralph Kimball, pois ele nos proporciona uma visão mais específica do estado de Pernambuco, com a possibilidade de expansão para uma visão regional e nacional. Além disso, optamos pelo esquema 'Star Schema' devido à sua simplicidade e eficácia. Esse esquema permite a criação de uma tabela de fatos central, contendo métricas relevantes, e várias tabelas de dimensões conectadas a ela. Essa estrutura facilita a análise dos dados e a geração de relatórios significativos para gestores e analistas.

## 5. Usuários

Dessa forma, esse projeto se torna útil tanto para gestores de saúde estadual, profissional de saúde dos municípios e gestores de políticas públicas, que podem utilizar esses dados para tomar decisões estratégicas, quanto para análises, que têm a oportunidade de explorar os dados e identificar tendências em diversos fatores.



# ETAPA 2 - LEVANTAMENTO DAS NECESSIDADES

## 6. Consultas de Apoio à Decisão

Para esta etapa, realizou-se um estudo para identificar os indicadores mais relevantes para a gestão da saúde pública em documentos governamentais e trabalhos relacionados à vivência dos profissionais de saúde nos pontos de atendimento. Dessa forma, foram identificados indicadores essenciais para orientar a distribuição de recursos, priorizar a alocação de profissionais e serviços, e medir diversos fatores relacionados à situação da saúde pública.

## 7. Indicadores do <negócio>

- Quantidade de procedimentos realizados, para medir o total de procedimentos realizados em um período específico. Ajuda a monitorar a carga de trabalho e eficiência operacional e avalia a produtividade e a capacidade de atendimento da organização.
- Número de pacientes atendidos, para avaliar a demanda e a carga de trabalho.
- Percentual de procedimentos por Sexo, Faixa Etária, Raça, para Identificar tendências de gênero, idade e possíveis necessidades específicas de atendimento. Além de ajudar a entender quais grupos etários são mais atendidos e pode orientar políticas de atendimento.

# ETAPA 3 - MODELAGEM

## 9. Modelo Relacional

Para gerar o modelo relacional, foram identificadas quatro tabelas de dimensões essenciais. A **dim\_local** captura as informações geográficas e de localização dos estabelecimentos de saúde, com o atributo principal sendo o `id_municipio`. A **dim\_tempo** cobre a dimensão temporal, permitindo análises ao longo do tempo, com atributos como `mes` e `ano`. A **dim\_procedimento** detalha os procedimentos médicos realizados, categorizados por regras contratuais e códigos de doenças, com atributos como `cid_principal_categoria` e `tipo_regra_contratual`. Por fim, a **dim\_demografico** abrange as características demográficas dos pacientes, com atributos como `idade`, `raça` e `sexo`. Além disso, foi criada a tabela de fatos **fato\_atendimento**, que armazena métricas quantitativas fundamentais para a geração de relatórios significativos.

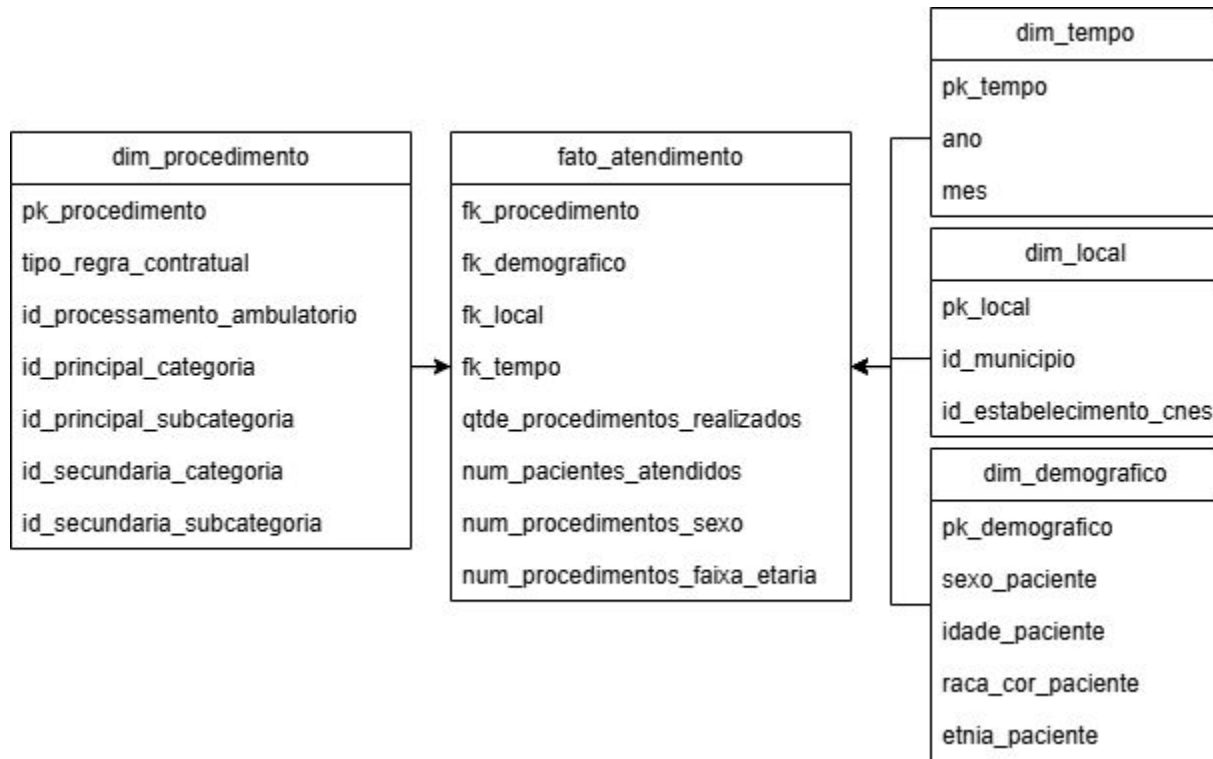
## 10. Modelo Dimensional

Esse Modelo Dimensional foi desenvolvido especificamente para a gestão de saúde pública, visando facilitar a tomada de decisões estratégicas e operacionais. A área de negócios engloba o gerenciamento de atendimentos e procedimentos médicos, com foco em otimizar os recursos e melhorar a qualidade dos serviços de saúde. O processo central é a análise granular dos atendimentos, onde cada fato registrado corresponde a uma interação única no sistema de saúde. A granularidade é definida no nível mais detalhado possível: o atendimento individual.

# 10. Modelo Dimensional

As hierarquias dentro das dimensões são claras: na dimensão Local, temos a hierarquia município > estabelecimento; na dimensão Tempo, a hierarquia é ano > mês > dia; na Demográfico, a hierarquia pode ser idade > faixa etária, além de outras combinações de características como sexo e raça. Essas estruturas permitem uma análise flexível e detalhada, suportando uma gestão eficiente e orientada a dados.

## 10. Modelo Dimensional





# 10. Modelo Dimensional

| fk_procedimento | fk_demografico | fk_local | fk_tempo | qtde<br>procedimentos<br>realizados (2020) | num pacientes<br>atendidos (2020) | num<br>procedimentos<br>sexo (F) | num<br>procedimentos<br>faixa etária<br>(10-20) |
|-----------------|----------------|----------|----------|--|-----------------------------------|----------------------------------|---|
| 1               | A1             | Rec      | 18J2023  | 150  | 120                               | 80                               | 30  |
| 2               | B2             | Cen      | 22F2023  | 200  | 180                               | 110                              | 50  |
| 3               | C3             | Nor      | 05M2023  | 250  | 200                               | 140                              | 70  |
| 4               | A1             | Sul      | 15A2023  | 300  | 250                               | 160                              | 90  |
| 5               | B2             | Est      | 28J2023  | 180  | 150                               | 100                              | 40  |
| 6               | C3             | Rec      | 12A2023  | 220  | 190                               | 130                              | 60  |
| 7               | A1             | Nor      | 07M2023  | 170  | 140                               | 90                               | 45  |
| 8               | B2             | Cen      | 20J2023  | 230  | 200                               | 120                              | 55  |
| 9               | C3             | Sul      | 18F2023  | 260  | 210                               | 140                              | 65  |
| 10              | A1             | Est      | 25M2023  | 190  | 160                               | 110                              | 50  |

# ETAPA 4 - PROJETO FÍSICO DO BD

## 11. Modelo Relacional do Data Mart (físico)

# ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

## 12. Plano de Carga da Dimensão Tempo

## 13. Plano de Carga da Dimensão XXXXX

## 14. Plano de Carga da Dimensão XXXXX

## 15. Plano de Carga da Fato



# ETAPA 6 - APLICAÇÃO OLAP e PAINEL DE BORDO

## 16. Consulta OLAP 1

## 16. Consulta OLAP 2

## 16. Consulta OLAP 3

16. Painel de Bordo XXXX (ponto extra)

# Referências

1. Dados fontes (abertos)

# Anexos

# Apêndices