

HIVision

Plataforma de Integração e Inteligência Epidemiológica

Leonardo dos Santos Silva (leonardo.ssilva2@ufrpe.br)
Gabriel Mesquita Gomes (gabriel.mesquitagomes@ufrpe.br)

Fonte dos Dados: <https://datasus.saude.gov.br/transferencia-de-arquivos/Aplicação>
OLAP:

<https://app.powerbi.com/view?r=eyJrIjoiZWQxNWFiNTYtZGQ2Yi00MzdhLTllODQtMGMyYmM1MTgyOWQ0IiwidCI6ImE2Zjc0ZWZhLTM2NzQtNDQ5YS1hY2FILWMwYTlzMzUwNCJ9&embedImagePlaceholder=true&pageName=445d82d2bad67f4b51b9>

Github: https://github.com/LeuhInt/SAD_project

ETAPA 1 - PLANEJAMENTO

1. Contextualização

O Data Mart HIVision está inserido na área de saúde pública, com foco na vigilância epidemiológica, que tem como finalidade o monitoramento contínuo de doenças e agravos de interesse coletivo, visando identificar tendências, fatores de risco e impactos sobre a população. A análise integrada desses dados é fundamental para subsidiar políticas de prevenção, controle e tratamento, permitindo que gestores e profissionais de saúde tomem decisões baseadas em evidências e direcionem recursos de forma mais eficiente.

2. Escopo/objetivo do Data Mart

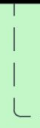
O projeto HIVision tem como objetivo desenvolver um Data Mart analítico voltado à integração, análise e visualização de dados epidemiológicos relacionados ao HIV em adultos, no período de 2007 a 2025. A iniciativa busca apoiar a tomada de decisão em políticas públicas de saúde, permitindo análises históricas, exploratórias e preditivas baseadas em informações provenientes do SINAN/DATASUS.

Arquitetura Tecnológica do Data Mart - HIVision

Fonte de Dados



e-SUS Sinan



ETL



pentaho®

Data Mart

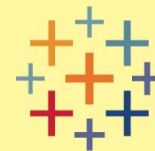


PostgreSQL

Apresentação



Power BI



+ a b | e a u®

Fluxo do Processo de Construção do Data Mart - HIVision

Process flow

● Design Lógico

● Implementação Física

● Desenvolvimento ETL

● Entrega

Design Lógico

Selecionar
Processo de Negócio

Definir Base
De Dados

Declarar
Granularidade

Identificar as
Dimensões

Análise de Requisitos
de Negócio

Identificar os Fatos
(Métricas)

Implementação Física

Selecionar
SGBD

Criação do Banco
Dados

Criação
De Tabelas

Desenvolvimento ETL

Selecionar
PDI

Extract

Transform

Load

Entrega

Selecionar
OLAP

Conectar o
Banco de Dados

Construir
Dashboards

Revisar e Validar
Métricas

Disponibilizar
Acesso

4. Abordagem

Botom-up:

- mais ágil;
- facilita incrementos futuros;
- integração com outros dados do DATASUS.

Star Schema

- simplicidade;
- uma tabela fato central com dados quantitativos;
- tabelas dimensão conectadas.

5. Usuários

- Gestores de Saúde Pública;
- Analistas e Técnicos de Vigilância Epidemiológica;
- Secretárias de Saúde;
- Pesquisadores e estudantes;

ETAPA 2 - LEVANTAMENTO DAS NECESSIDADES

6. Consultas de Apoio à Decisão

- A epidemia de HIV está aumentando, diminuindo ou estabilizada em nossa região?
- Quem é o perfil mais vulnerável hoje?
- Em quais municípios devemos concentrar esforços de testagem e envio de equipes de saúde?
- O número de diagnósticos em gestantes está caindo?
- A epidemia está afetando de forma desproporcional grupos racializados ou com menor escolaridade?

7. Indicadores do HIVision

Indicador	Descrição
Casos notificados de HIV	Número total de casos de HIV confirmados e registrados no sistema de vigilância em um período específico.
Novos casos de HIV	Número de pessoas diagnosticadas com HIV pela primeira vez durante um período definido
HIV em Gestantes	Contagem de gestantes diagnosticadas com HIV.
Taxa de mortalidade por HIV/AIDS	Número de óbitos por HIV/AIDS por 100.000 habitantes em um período específico.
Pessoas em TARV	Contagem de pessoas em tratamento antirretroviral (TARV) no sistema de saúde.
Casos por sexos	Distribuição de casos de HIV por sexo: masculino e feminino.
Distribuição etária dos casos	Percentual de casos por faixa etária
Testes realizados para HIV	Número total de testes realizados, útil para medir cobertura e rastreamento da população
Tempo médio entre diagnóstico e início do tratamento	Média de dias entre o diagnóstico de HIV e o início da terapia antirretroviral.

ETAPA 3 - MODELAGEM

9. Modelo Relacional

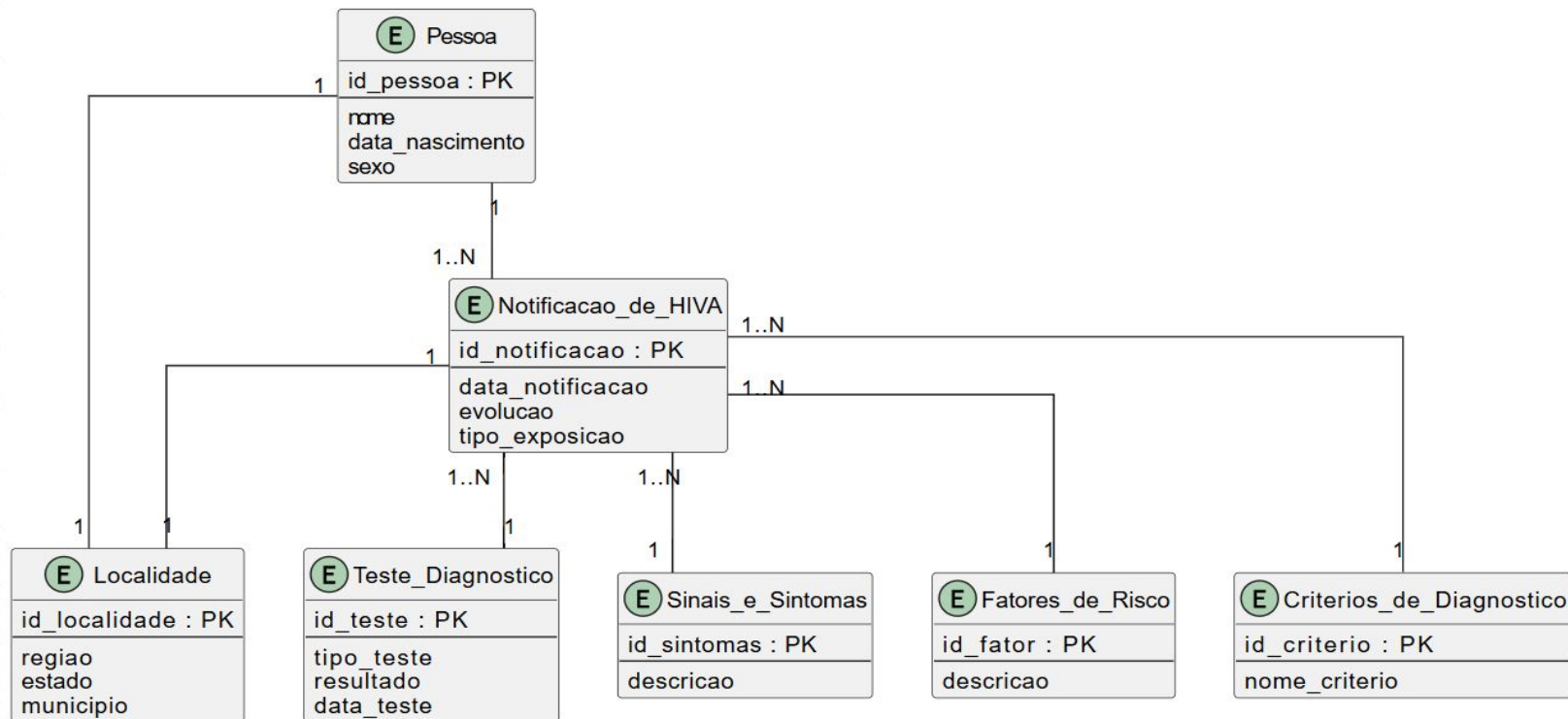
9.1) Descrição das entidades e relacionamentos

Relacionamento	Entidade A	Cardinalidade A	Ação realizada por A	Entidade B	Cardinalidade B	Coluna1 no diagrama
1	Pessoa	1	Notifica	Notificação de HIV/AIDS	1..N	FK_PACIENTE (em Notificação)
2	Notificação de HIV/AIDS	1	Local Notificação	Localidade	1	FK_LOCAL_NOTIFICACAO (em Notificação)
3	Pessoa	1	Local Residência	Localidade	1	FK_LOCAL_RESIDENCIA (em Pessoa)
4	Notificação de HIV/AIDS	1..N	Ação Determinada Por	Teste Diagnóstico	1	FK_TESTES (em Notificação)
5	Notificação de HIV/AIDS	1..N	Apresenta	Sinais e Sintomas	1	FK_ANTECEDENTES (em Notificação)
6	Notificação de HIV/AIDS	1..N	Possui	Fatores de Risco	1	FK_FATOR_RISCO (em Notificação)
7	Notificação de HIV/AIDS	1..N	Classificada Por	Critérios de Diagnóstico	1	FK_CRITERIO (em Notificação)

9. Modelo Relacional

9.2) Diagrama lógico ER

Modelo ER Lógico - Notificação de HIV em Adultos



10. Modelo Dimensional

A. Área de Negócios

a. Saúde Pública (Vigilância Epidemiológica)

B. Processo

a. Notificação de Agravos de HIV em Adultos (HIVA)

C. Granularidade

a. Uma notificação individual de um caso de HIV em adulto.

10. Modelo Dimensional

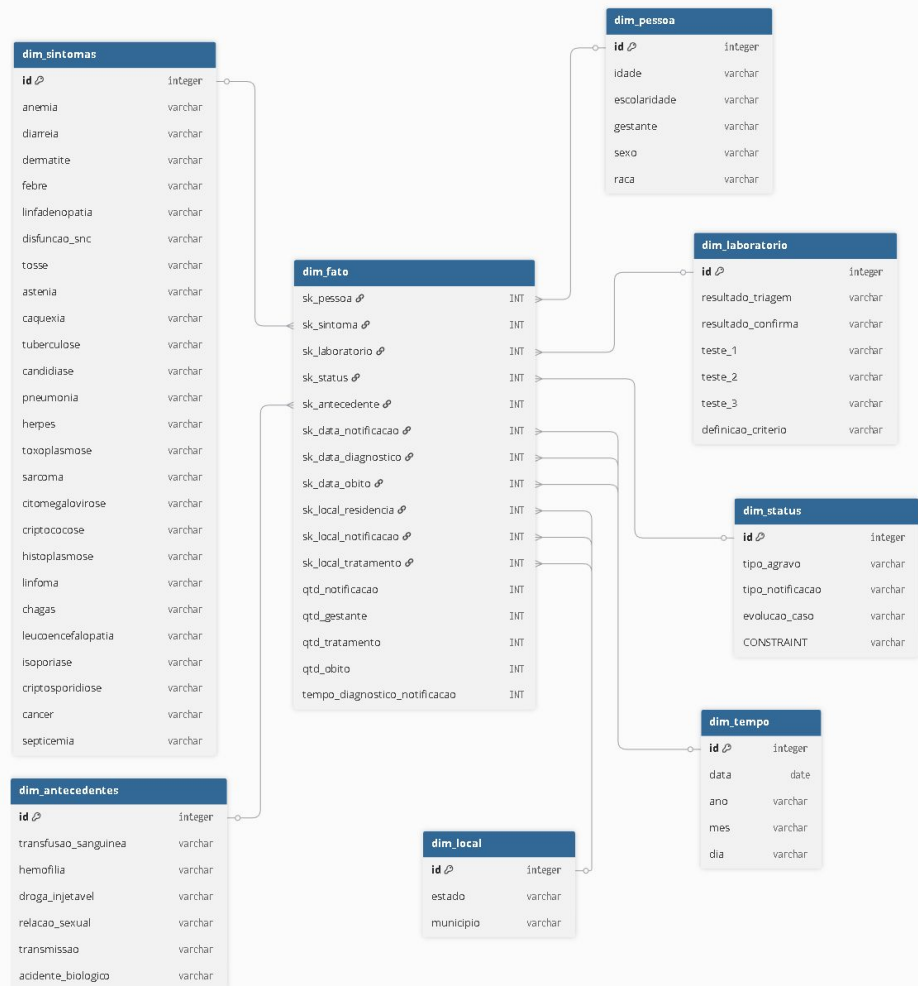
Dimensões	Atributos
DimTempo	Ano → mês → semana → dia
DimLocal	Região → UF → Municipio
DimPaciente	Idade, sexo, raca, escolaridade
DimAntecedentes	Descrição: Presença ou ausencia de outro fator
DimTestes	Descrição: Tipos de testes realizados
Dim FatorDeRisco	Descrição: Fatores de riscos para HIV
DimCritério	Descrição: Forma de diagnóstico

10. Modelo Dimensional

Métrica	Tipo	Descrição
dt_notificacao	Não-aditiva	Data em que o caso foi notificado
dt_diagnostico	Não-aditiva	Data em que o diagnóstico foi confirmado
qtd_notificacao	Aditiva	Quantidade de notificações registradas
dias_entre_diag_notif	Semi-aditiva	Diferença em dias entre diagnóstico e notificação
qtd_testes_realizados	Aditiva	Total de testes realizados em um período ou local
obito_confirmado	Aditiva	Número de óbitos confirmados
gestante_confirmado	Aditiva	Número de gestantes confirmadas com a doença
total_tarv	Semi-aditiva	Total de pacientes em tratamento antirretroviral
qtd_fatores_risco	Semi-aditiva	Número de fatores de risco associados

10. Modelo Dimensional

Disponível em: [Star Schema](#)



10. Modelo Dimensional

	sk_pessoa integer	sk_sintoma integer	sk_laboratorio integer	sk_status integer	sk_antecedente integer	sk_data_notificacao integer	sk_data_diagnostico integer	sk_data_obito integer	sk_local_residencia integer	sk_local_notificacao integer	sk_loca integer
1	361	403523	334	1	3	5773	5773	0	5669	5669	
2	536	403523	26	1	2	5773	5773	0	0	0	
3	4964	403521	183	9	156	5773	5773	0	660	660	
4	4964	403523	26	1	2	5773	5773	0	5381	5381	
5	480	403523	808	1	3	5773	5773	0	5491	5491	
6	3600	403523	5	1	2	5773	5773	0	6010	6010	
7	10433	403578	808	1	2	5773	5773	0	6279	6279	
8	32292	403523	183	9	3	5773	5773	0	5740	5740	
9	1755	403523	808	1	2	5773	5773	0	6675	6675	
10	707	403521	183	1	70	5773	5773	0	660	660	
11	82955	403523	808	1	2	5773	5773	0	5680	5722	
12	53	403523	808	1	2	5773	5773	0	660	660	
13	1261	403523	808	1	3	5773	5773	0	6413	6413	
14	22203	403578	5	1	270	5773	5773	0	0	0	
15	1054	404111	334	1	3	5773	5773	0	4598	4598	
16	1569	403521	5	1	2	5773	5773	0	5710	5710	

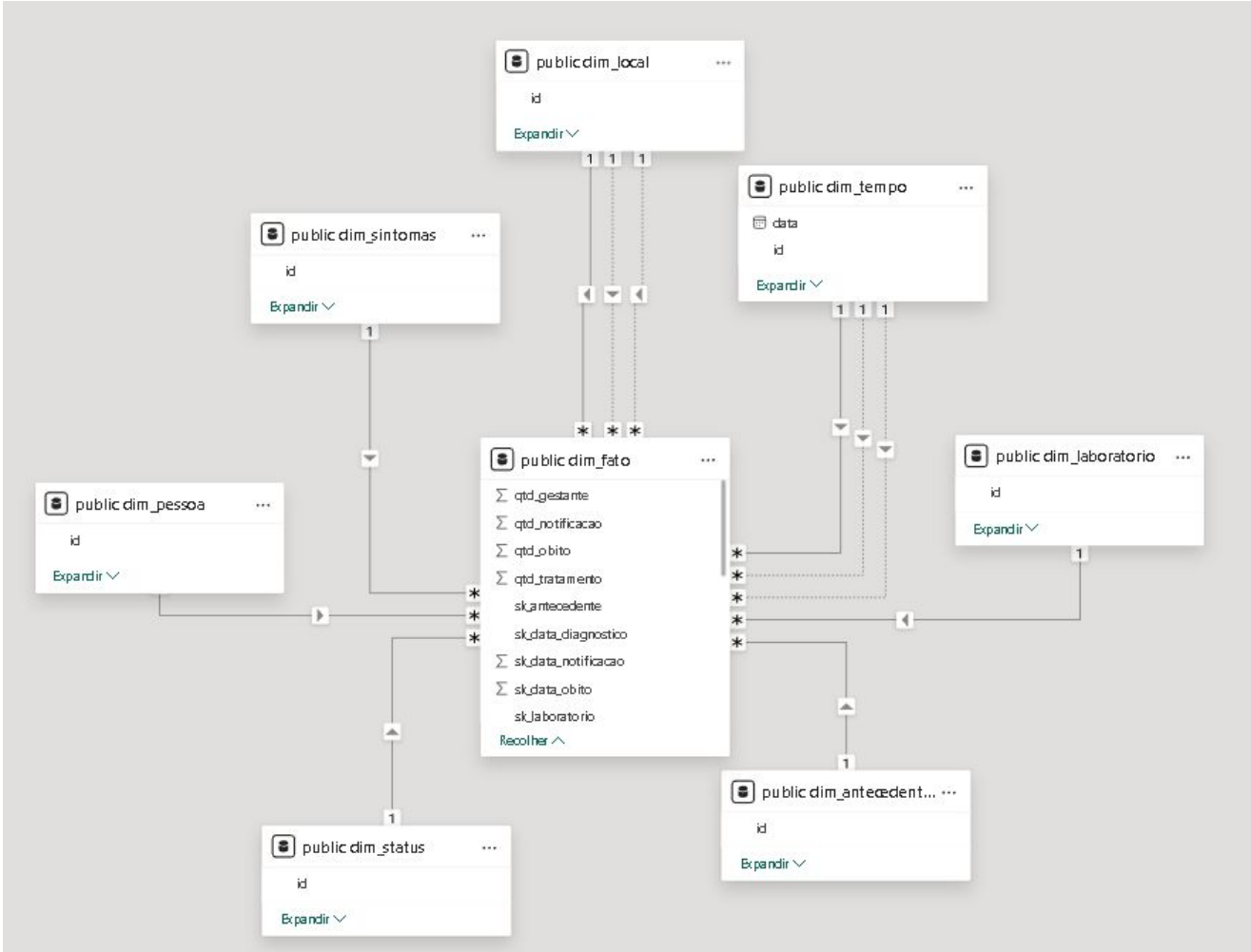
10. Modelo Dimensional do Data Mart (lógico)

H. Estimativa de espaço

$(539.535 \text{ linhas}) \times (91 \text{ bytes}) = 49.097.685 + 20\% = 58.917.222 \text{ bytes} = 58,9 \text{ Mb}$

ETAPA 4 - PROJETO FÍSICO DO BD

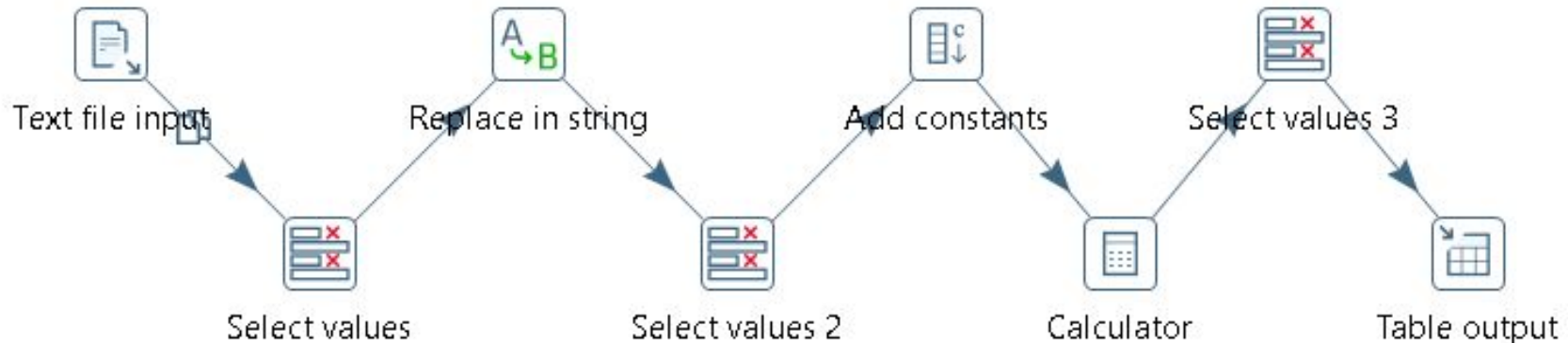
11. Modelo Relacional do Data Mart (físico)



ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

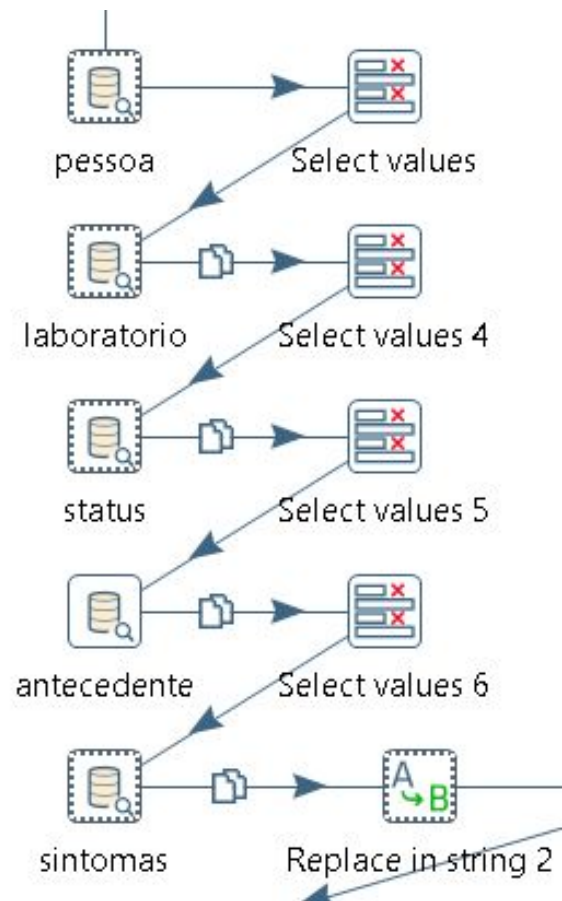
12. Plano de Carga da Dimensão Pessoa

- input data
- select fields and rename
- replace NA (4000)
- convert idade to integer
- add constant 4000
- calculate (A-B)
- idade real
- select fields (comparador)



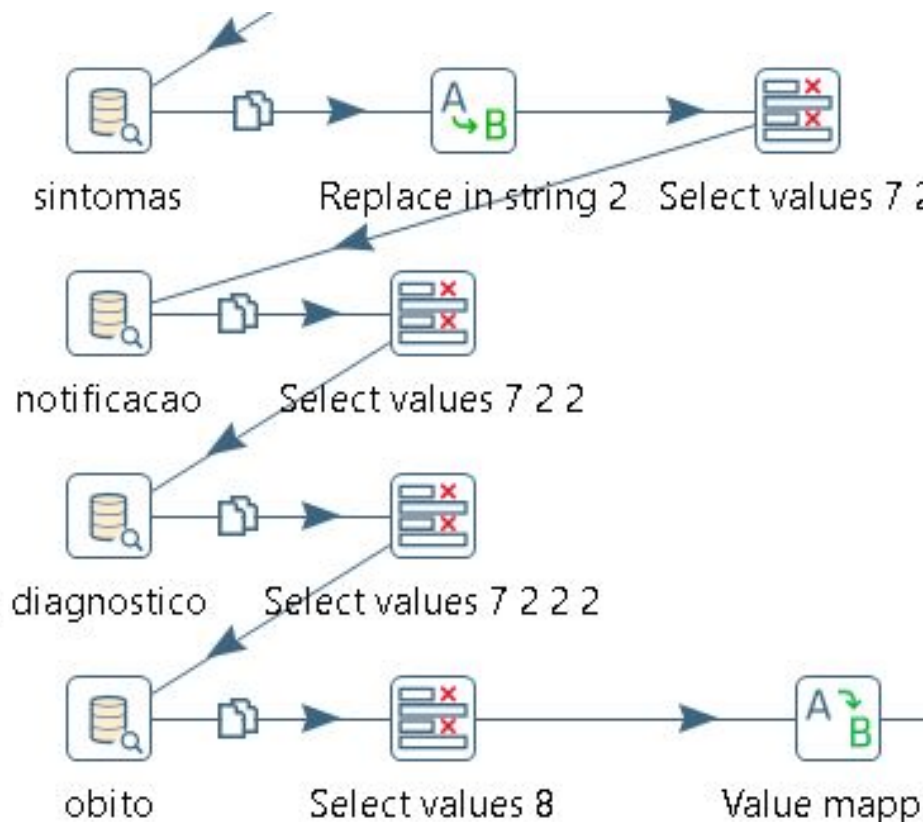
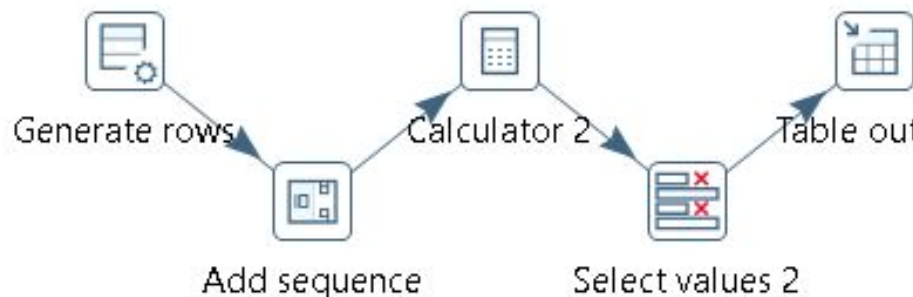
12. Plano de Carga das Dimensões laboratorio, status, antecedentes, sintomas

- input data
- select fields (comparador)

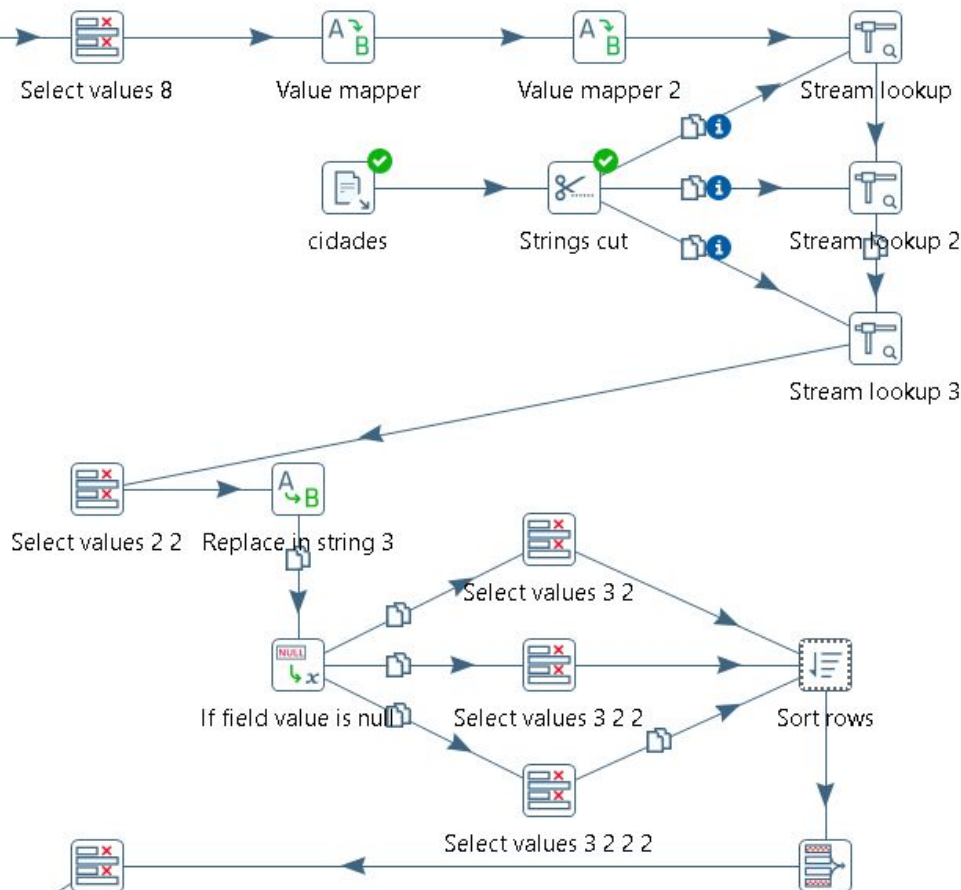


13. Plano de Carga da Dimensão Tempo

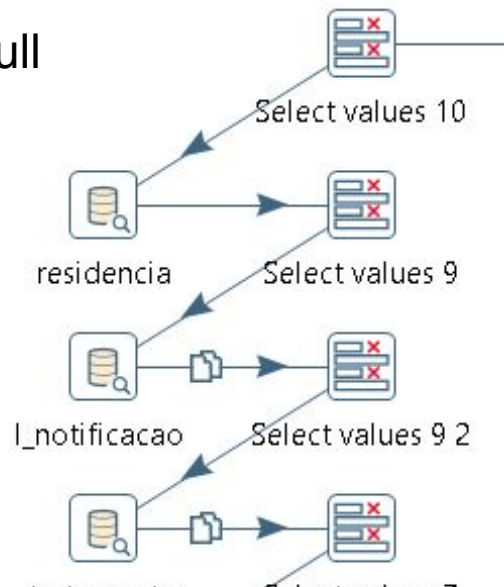
- input data
- dim_tempo genérica (2006 - 2025)
- replace NA with vazio
- select datas (notif, diag, obito)
- link generico-input data
- default 0



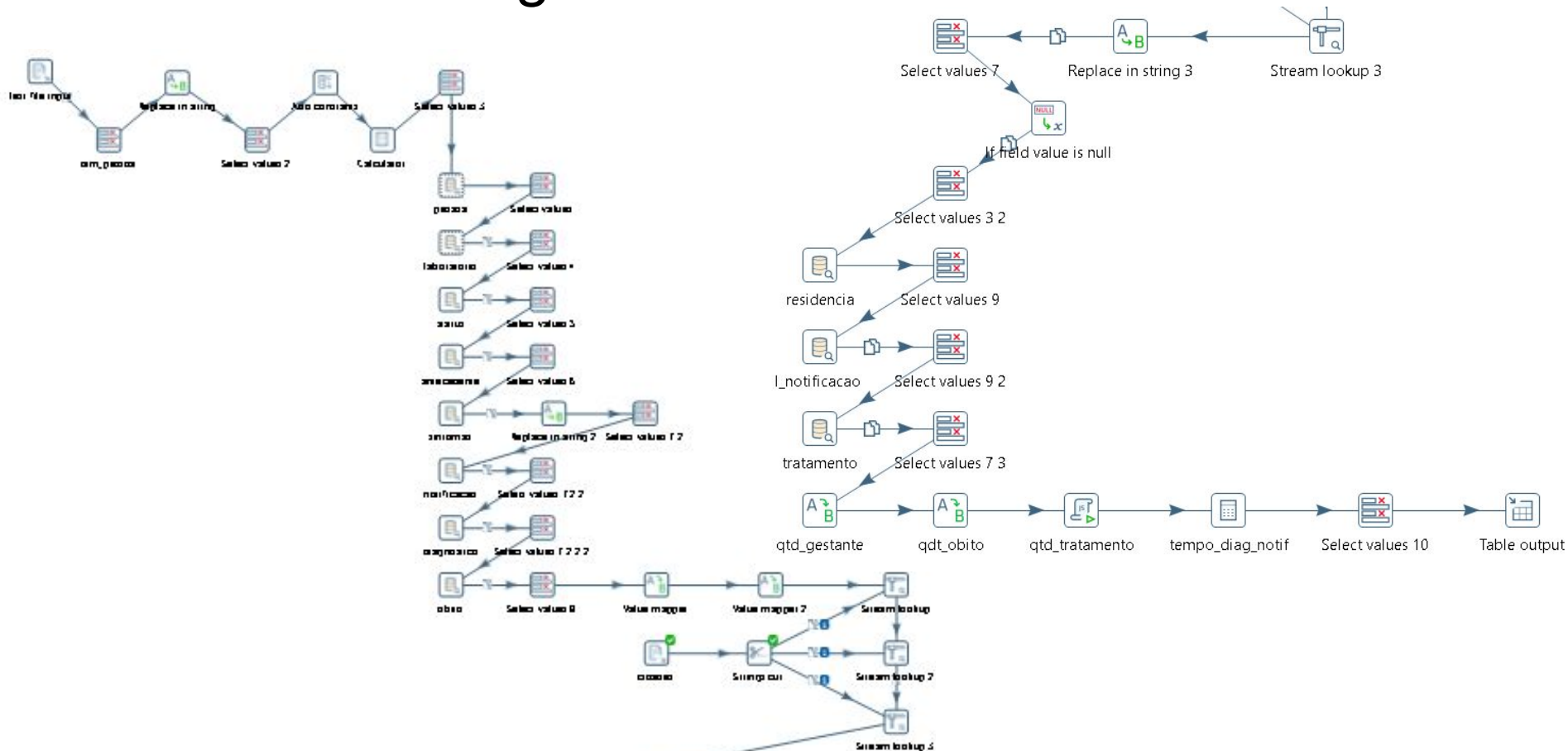
14. Plano de Carga da Dimensão Local



- select fields
- add estados para os códigos
- add cidades para os códigos
- código - nome/sigla
- unir dados
- replace NA e null
- default 0
- unir estados e municípios
- sort e remover duplicados



18. Plano de Carga da Fato



ETAPA 6 - APLICAÇÃO OLAP e PAINEL DE BORDO

Universidade Federal Rural de Pernambuco

Unidade Acadêmica de Serra Talhada

Bacharelado em Sistemas de Informação



[Acessar Painéis](#)

Professora: Dr. Ellen Polliana Ramos Souza Pereira
Disciplina: Sistemas de Apoio à Decisão

Leonardo dos Santos Silva

Este Datamart foi desenvolvido para apoiar a análise epidemiológica, o monitoramento de casos e a tomada de decisão em saúde pública, utilizando dados do Sistema de Informação de Agravos de Notificação (SINAN).

Autor: Leonardo Silva

Tecnologias utilizadas:

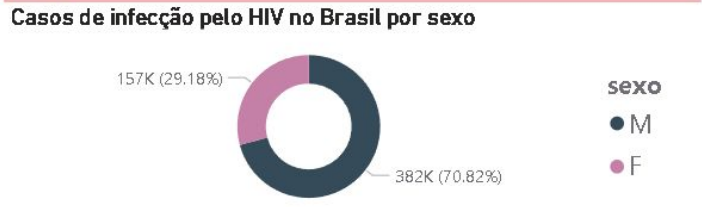
- Pentaho Data Integration (ETL)
- PostgreSQL (Data Mart)
- Power BI (Visualização)

Fonte dos dados:

Sistema de Informação de Agravos de Notificação – SINAN HIV/AIDS
Ministério da Saúde – Brasil

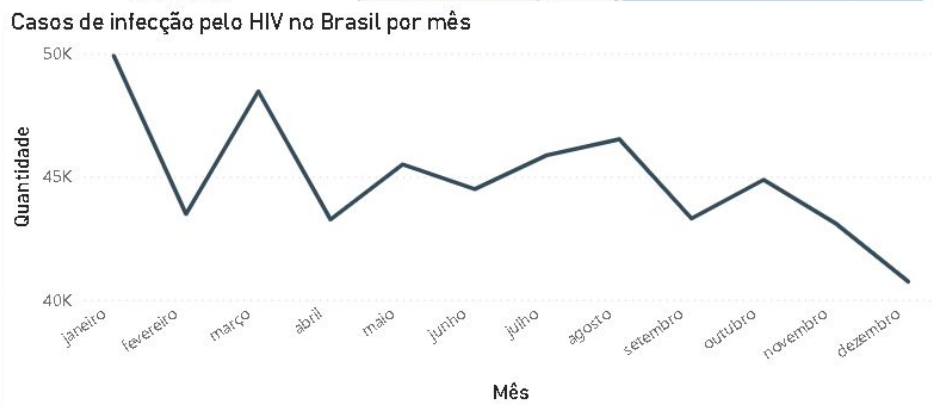
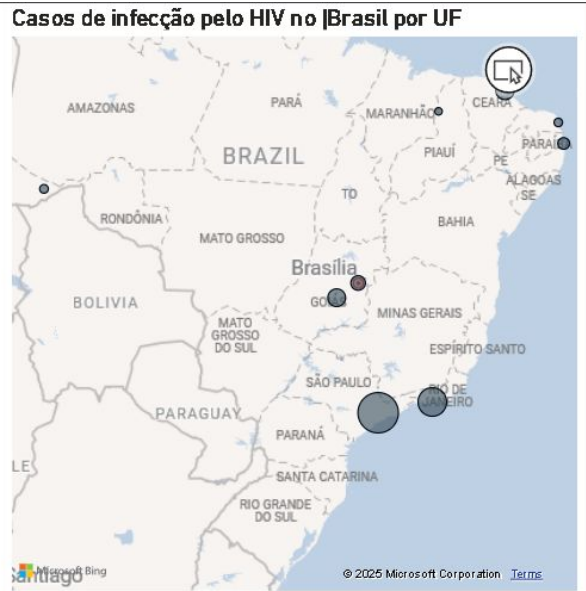
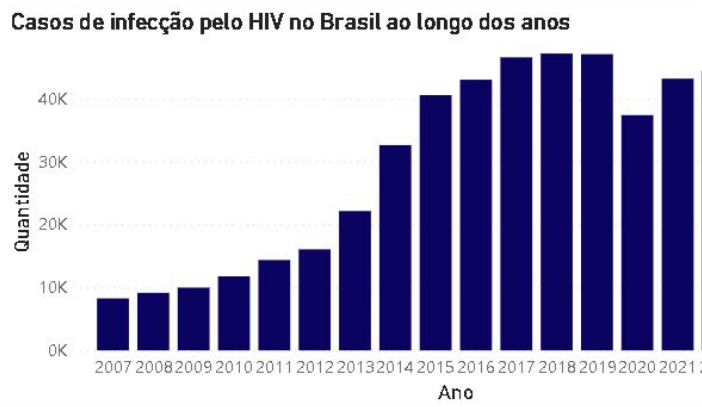


19. Consulta OLAP 1



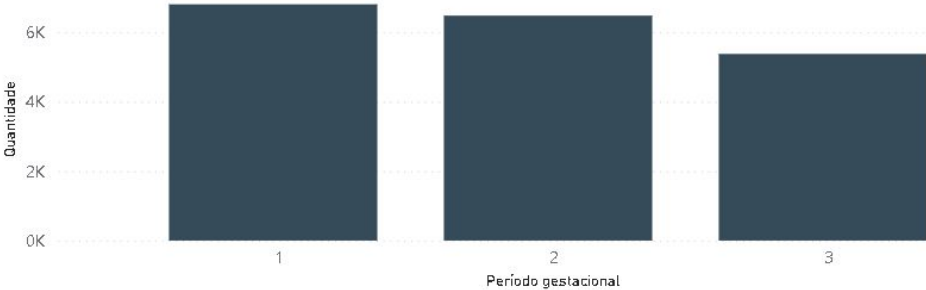
Casos de infecção pelo HIV no Brasil por faixa etária e sexo

sexo	13-19	20-29	30-39	40-49	50-59	60-69	70+	Ignorado	Total
M	19372	157270	106874	56720	27184	9150	2440	2929	381939
F	10824	43804	44175	32008	17760	6073	1391	1357	157392
Total	3019	20107	15104	8872	4494	15223	383	4286	539331
	6	4	9	8	4		1		

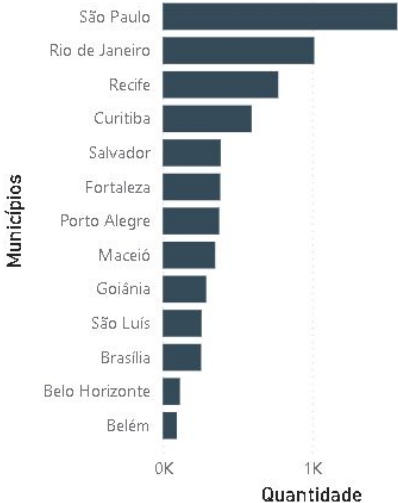


20. Consulta OLAP 2

Casos de infecção pelo HIV em gestantes no Brasil por período gestacional



Casos de infecção pelo HIV em gestantes no Brasil por município



Casos de infecção pelo HIV em gestantes no Brasil por UF

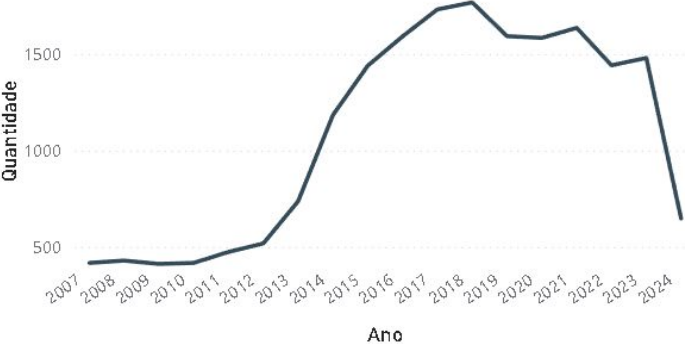


Casos de infecção pelo HIV em gestantes no Brasil

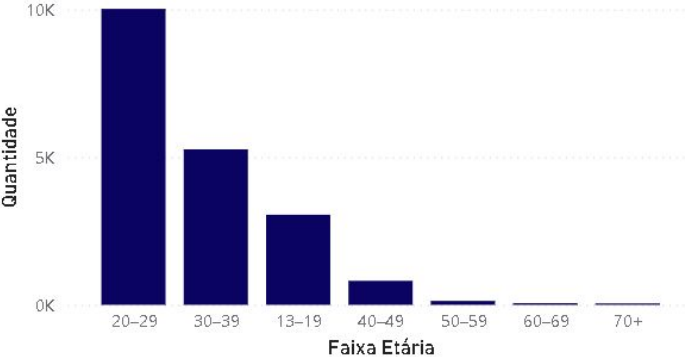
(2017 - 2024)

20K

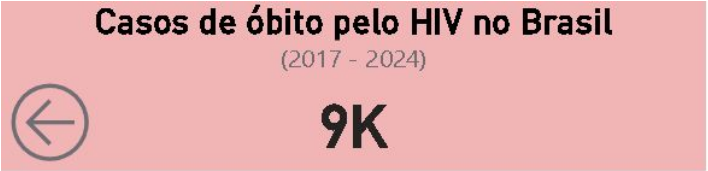
Casos de infecção pelo HIV em gestantes no Brasil por ano



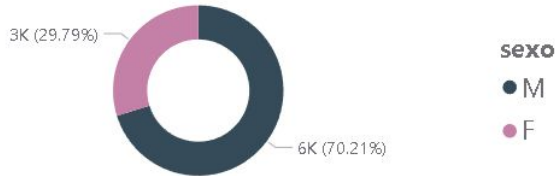
Casos de infecção pelo HIV em gestantes no Brasil por faixa etária



21. Consulta OLAP 3



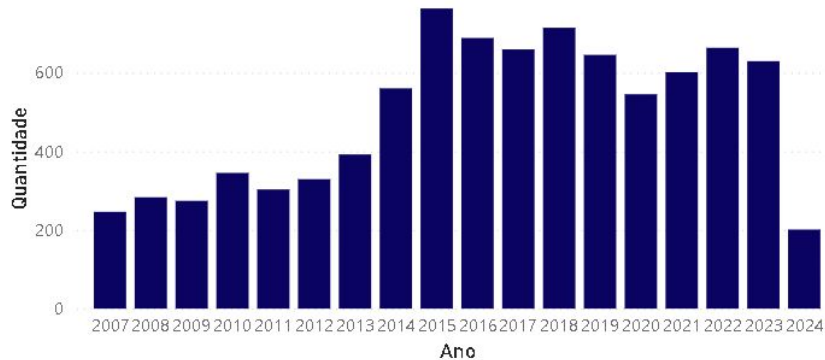
Casos de óbito pelo HIV no Brasil por sexo



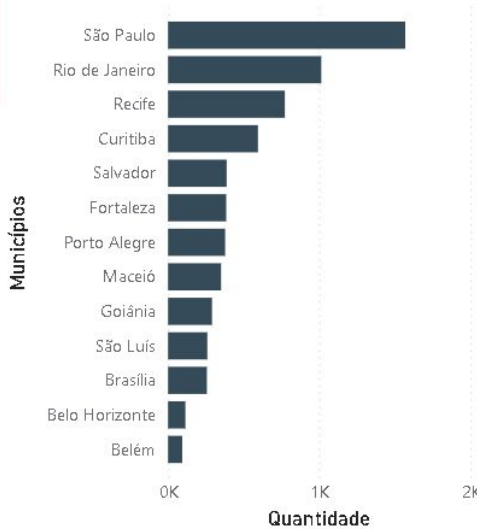
Casos de infecção pelo HIV no Brasil por faixa etária e sexo

sexo	13-19	20-29	30-39	40-49	50-59	60-69	70+	Ignorado	Total
F	67	328	720	733	473	212	79	21	2633
M	82	1102	1765	1587	1034	432	168	36	6206
Total	149	1430	2485	2320	1507	644	247	57	8839

Casos de óbito pelo HIV no Brasil ao longo dos anos



Casos de óbito pelo HIV no Brasil por município



Casos de óbito pelo HIV no Brasil por UF



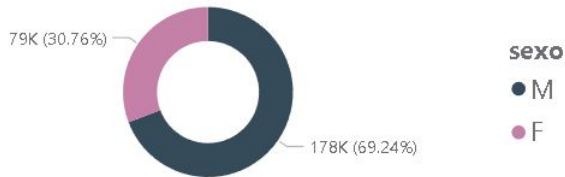
Casos de óbito pelo HIV no Brasil por mês



21. Consulta OLAP 4

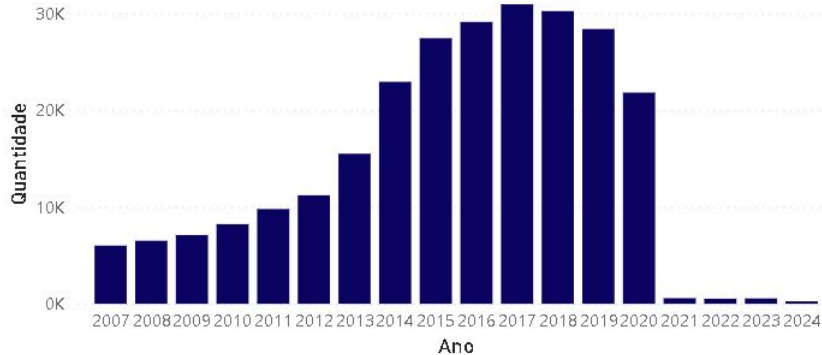


Casos de TARV/HIV no Brasil por sexo

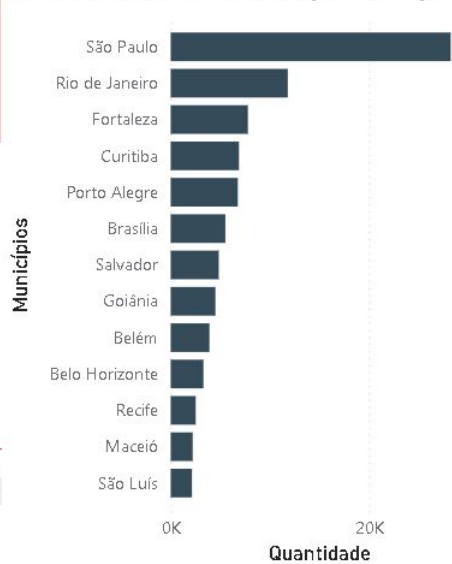


sexo	13-19	20-29	30-39	40-49	50-59	60-69	70+	Ignorado	Total
F	5977	22894	22703	15368	8285	2649	559	651	79086
M	9432	73848	50141	26378	12181	3838	959	1250	178027
Total	15409	9674	7284	4174	2046	6487	1518	1901	257113

Casos de TARV/HIV no Brasil ao longo dos anos



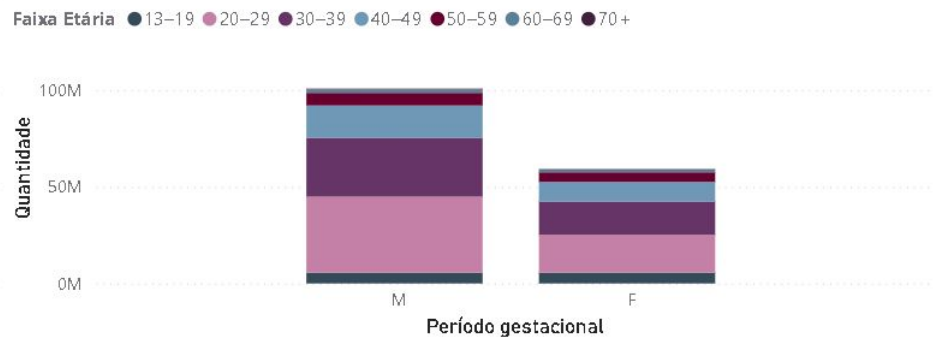
Casos de TARV/HIV no Brasil por município



Casos de TARV/HIV no Brasil por UF



Diferença entre tempo de diagnóstico e notificação pelo HIV no Brasil por faixa etária e sexo



Referências

1. UNAIDS. **Global HIV & AIDS statistics — Fact sheet**. Genebra: Joint United Nations Programme on HIV/AIDS, 2024. Disponível em:
<https://www.unaids.org/en/resources/fact-sheet>. Acesso em: 08 nov. 2025.
2. BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de HIV, Aids, Tuberculose, Hepatites Virais e Infecções Sexualmente Transmissíveis. **Boletim Epidemiológico: HIV e Aids**. Brasília: Ministério da Saúde, Número Especial, dez. 2023. Disponível em:
<https://www.gov.br/aids/pt-br/central-de-conteudo/boletins-epidemiologicos/2023/hiv-aids/boletim-epidemiologico-hiv-e-aids-2023.pdf>. Acesso em: 08 nov. 2025.