

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/transferecia-de-arquivos/>
Aplicação

OLAP

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

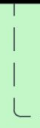
O projeto HIVision tem como objetivo desenvolver um Data Mart analítico voltado à integração, análise e visualização de dados epidemiológicos relacionados ao HIV em adultos, no período de 2007 a 2025. A iniciativa busca apoiar a tomada de decisão em políticas públicas de saúde, permitindo análises históricas,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

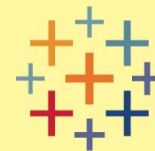


PostgreSQL

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

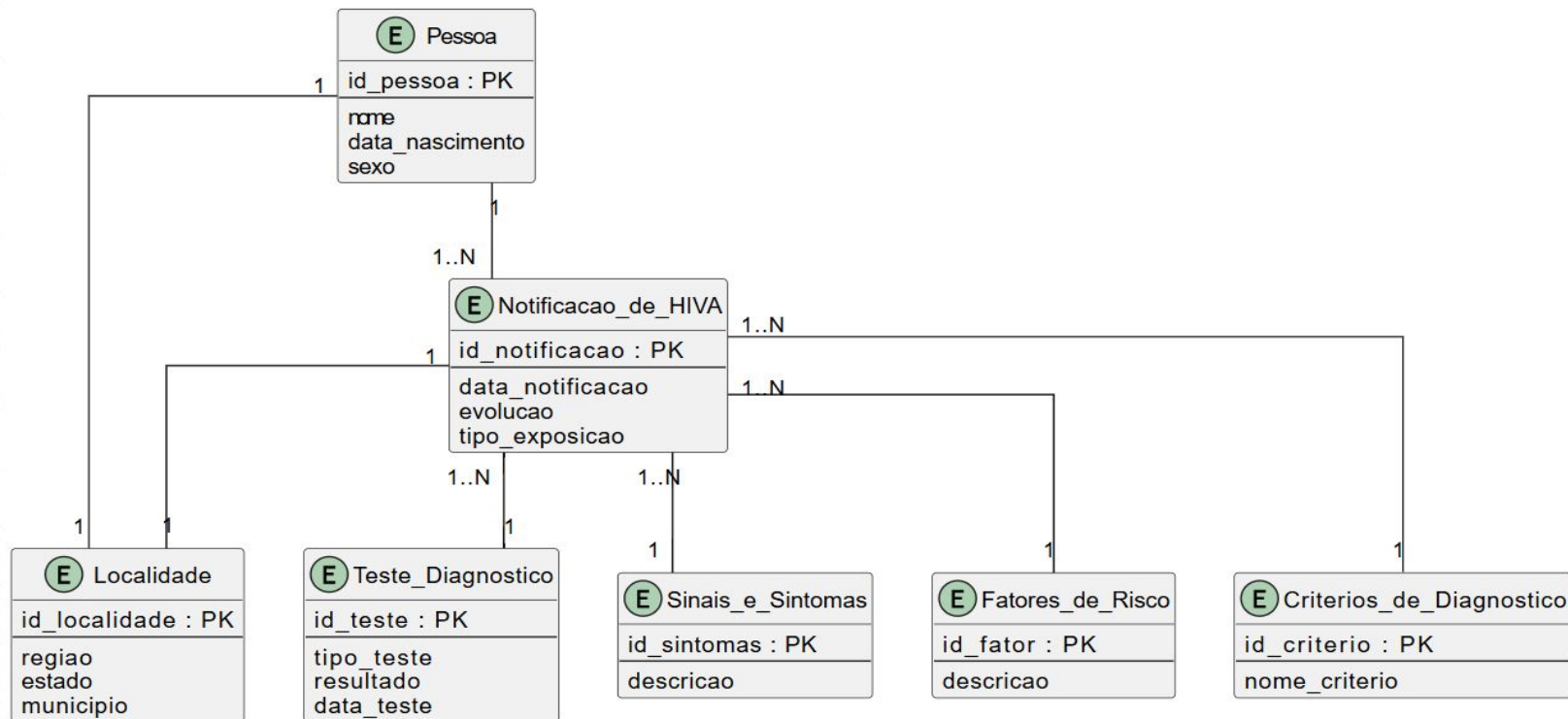
9.1) Descrição das entidades e relacionamentos

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

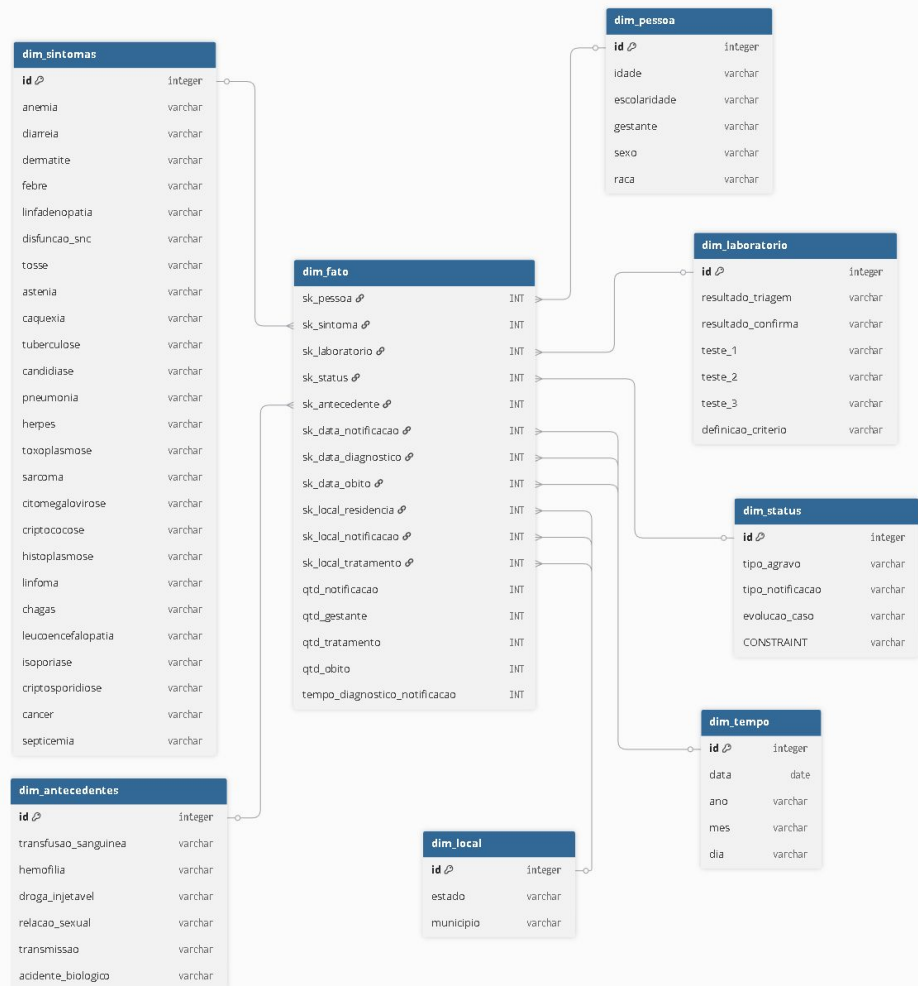
Dimensões	Atributos
DimTempo	Ano → mês → semana → dia
DimLocal	Região → UF → Municipio
DimPaciente	Idade, sexo, raca, escolaridade
DimAntecedentes	Descrição: Presença ou ausencia de outro fator
DimTestes	Descrição: Tipos de testes realizados
Dim FatorDeRisco	Descrição: Fatores de riscos para HIV
DimCritério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

k_pacient	fk_tempo	cal_notif	cal_reside	fator_ris	fk_criterio	antecedentes	fk_testes	notificac	diagnostid	notificac	entre_diag	testes_realiz	confirmnte	confirm	total_tary	fatores_r
PAC001	2025W01	LOC001	RES001	FR01	CR01	ANT01	TST01	2025-01-0	2025-01-0	1	2	2	0	0	3	1
PAC002	2025W02	LOC002	RES002	FR02	CR02	ANT02	TST02	2025-01-1	2025-01-0	1	2	3	0	1	4	2
PAC003	2025W02	LOC001	RES003	FR03	CR01	ANT03	TST03	2025-01-1	2025-01-1	1	2	2	0	0	5	1
PAC004	2025W03	LOC003	RES004	FR02	CR03	ANT02	TST01	2025-01-2	2025-01-1	1	2	1	1	0	4	2
PAC005	2025W03	LOC002	RES005	FR01	CR02	ANT01	TST02	2025-01-2	2025-01-2	1	2	3	0	1	5	1
PAC006	2025W04	LOC001	RES006	FR03	CR03	ANT03	TST03	2025-01-2	2025-01-2	1	2	2	0	0	3	3
PAC007	2025W05	LOC003	RES007	FR02	CR01	ANT02	TST01	2025-02-0	2025-02-0	1	2	1	1	0	6	2
PAC008	2025W05	LOC002	RES008	FR01	CR02	ANT01	TST02	2025-02-0	2025-02-0	1	2	2	0	1	4	1
PAC009	2025W06	LOC001	RES009	FR03	CR03	ANT03	TST03	2025-02-1	2025-02-0	1	2	3	0	0	5	3
PAC010	2025W06	LOC003	RES010	FR02	CR02	ANT02	TST01	2025-02-1	2025-02-1	1	2	1	1	0	6	2

10. Modelo Dimensional do Data Mart (lógico)

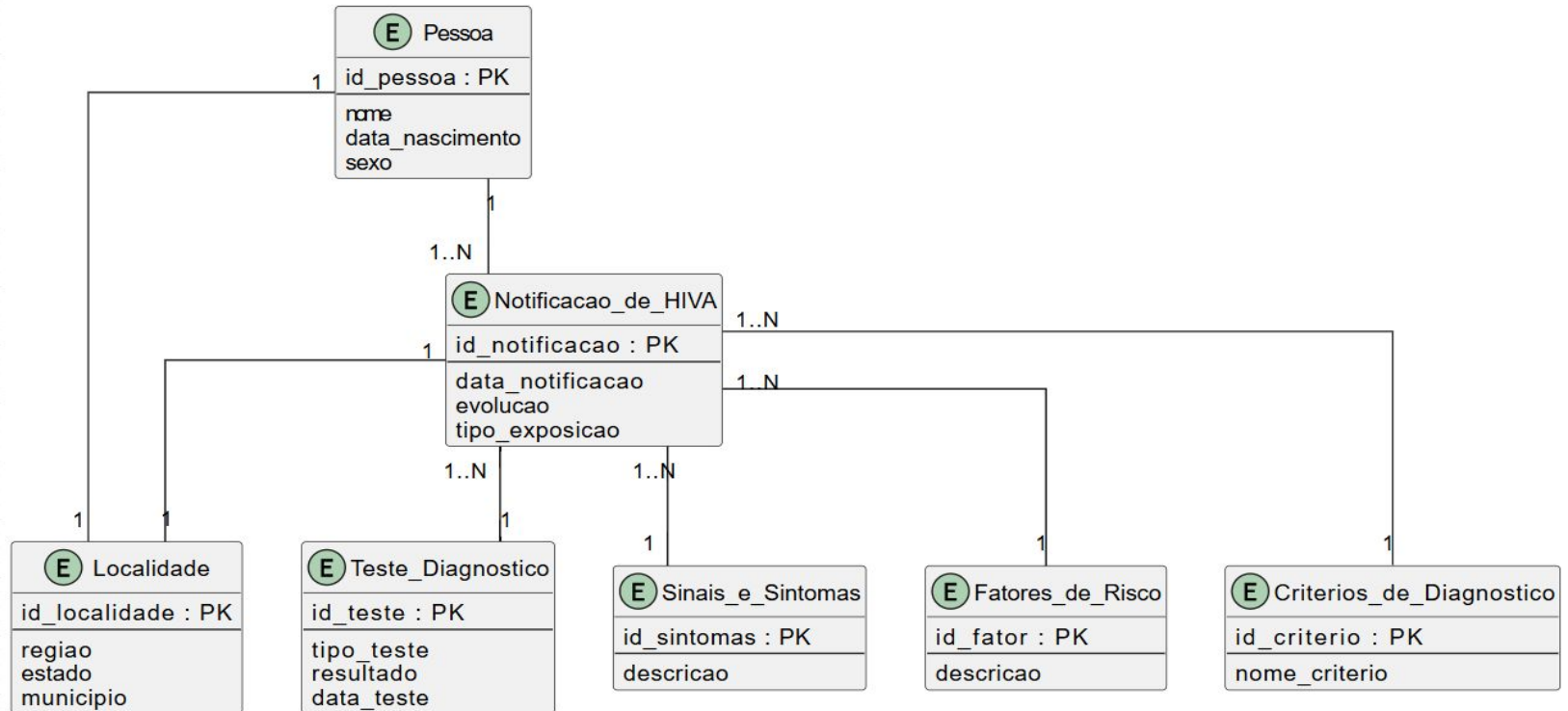
H. Estimativa de espaço

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

ETAPA 4 - PROJETO FÍSICO DO BD

11. Modelo Relacional do Data Mart (físico)

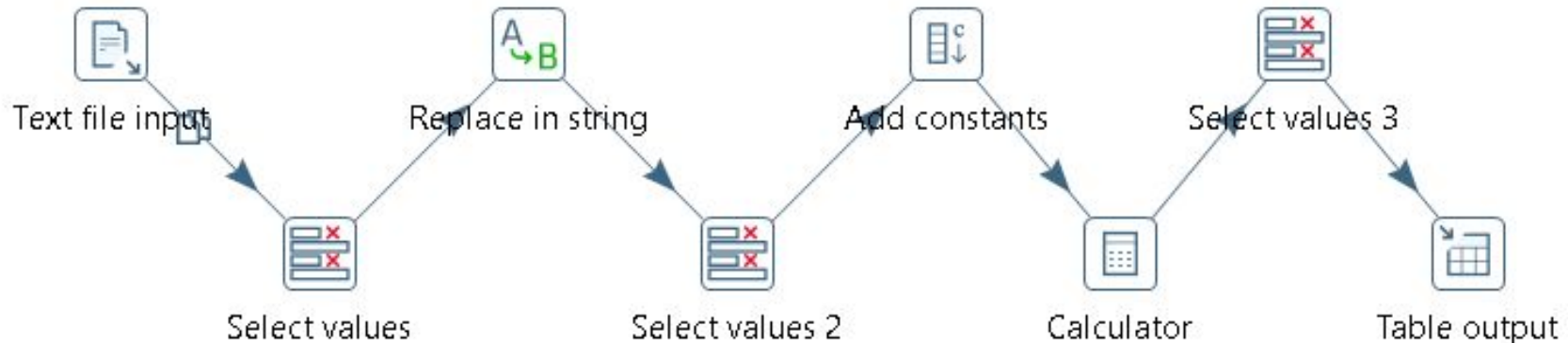
Modelo ER Lógico - Notificação de HIV em Adultos



ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

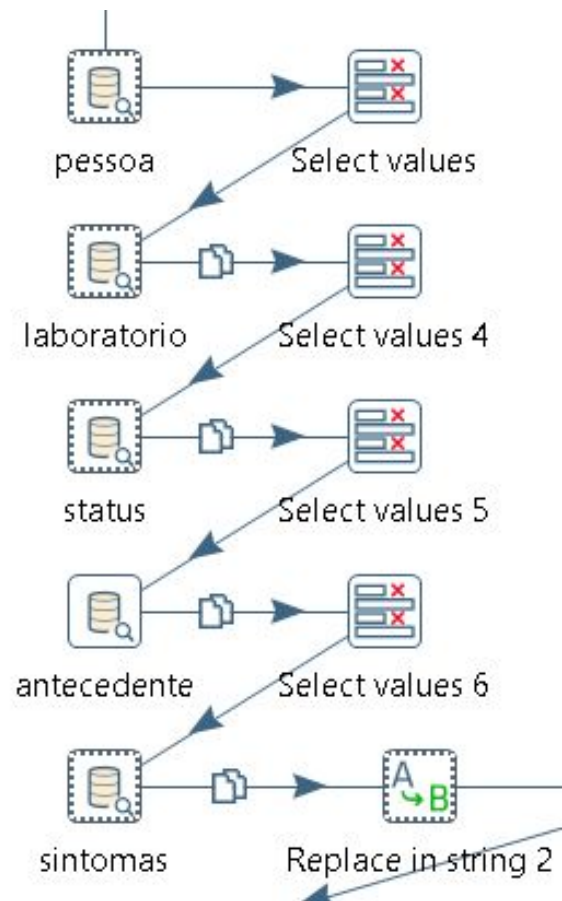
12. Plano de Carga da Dimensão Pessoa

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



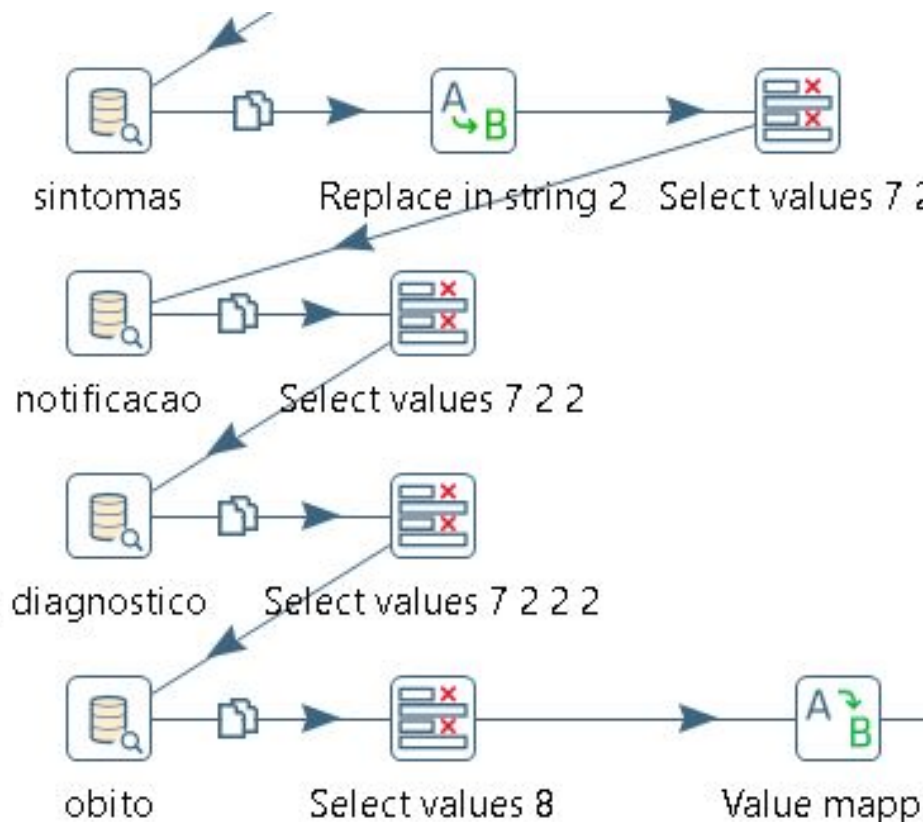
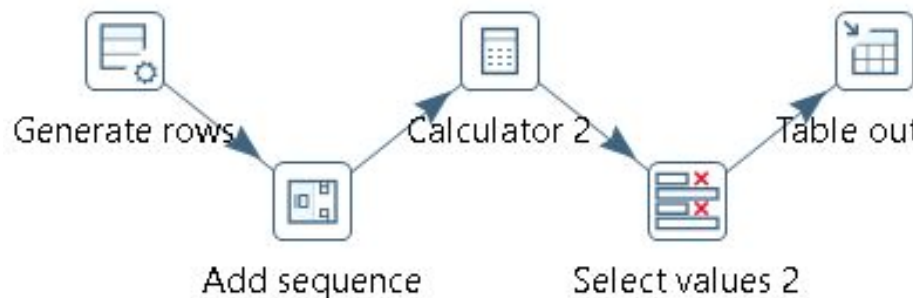
12. Plano de Carga das Dimensões laboratorio, status, antecedentes, sintomas

- input data
- select fields (comparador)

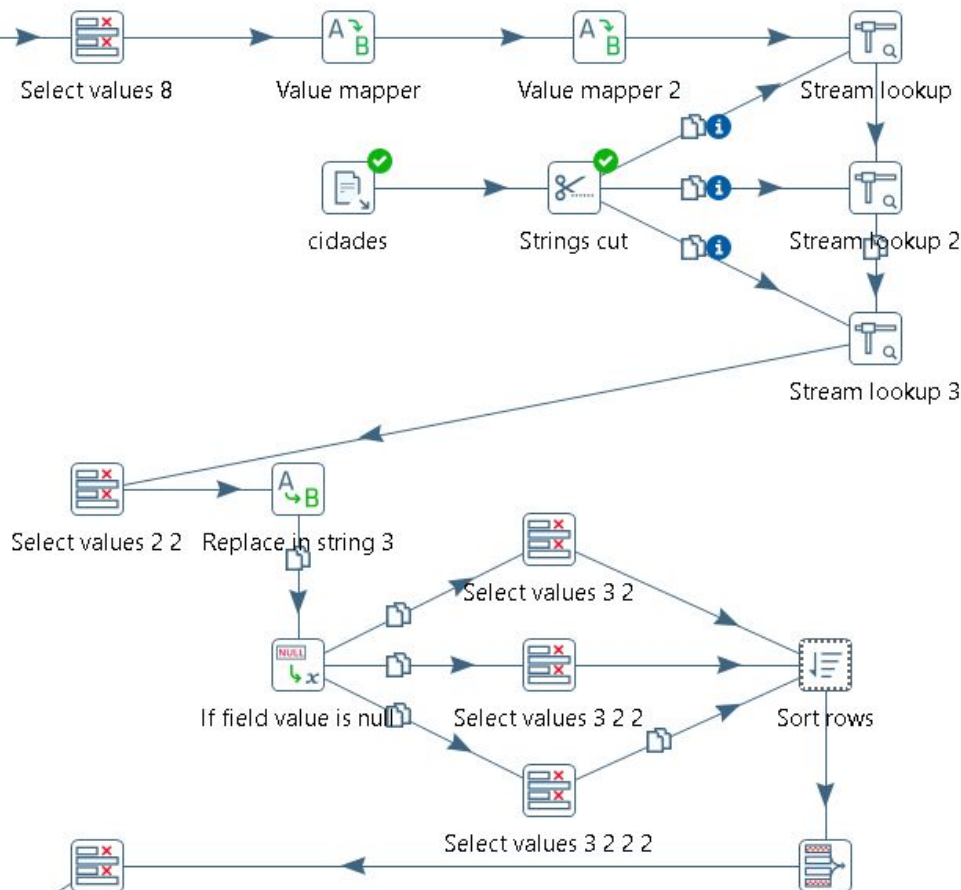


13. Plano de Carga da Dimensão Tempo

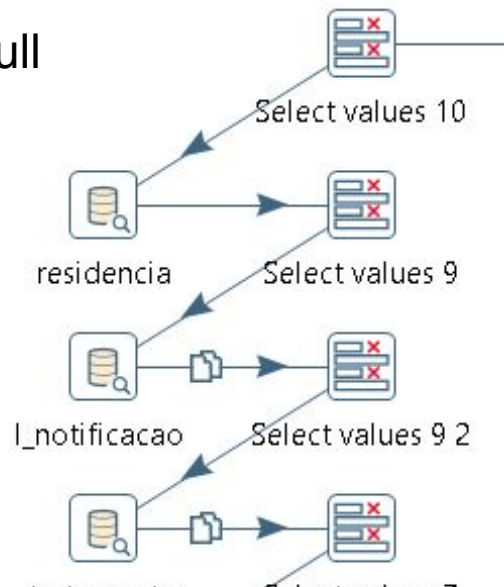
- input data
- dim_tempo genérica (2006 - 2025)
- replace NA with vazio
- select datas (notif, diag, obito)
- link generico-input data
- default 0



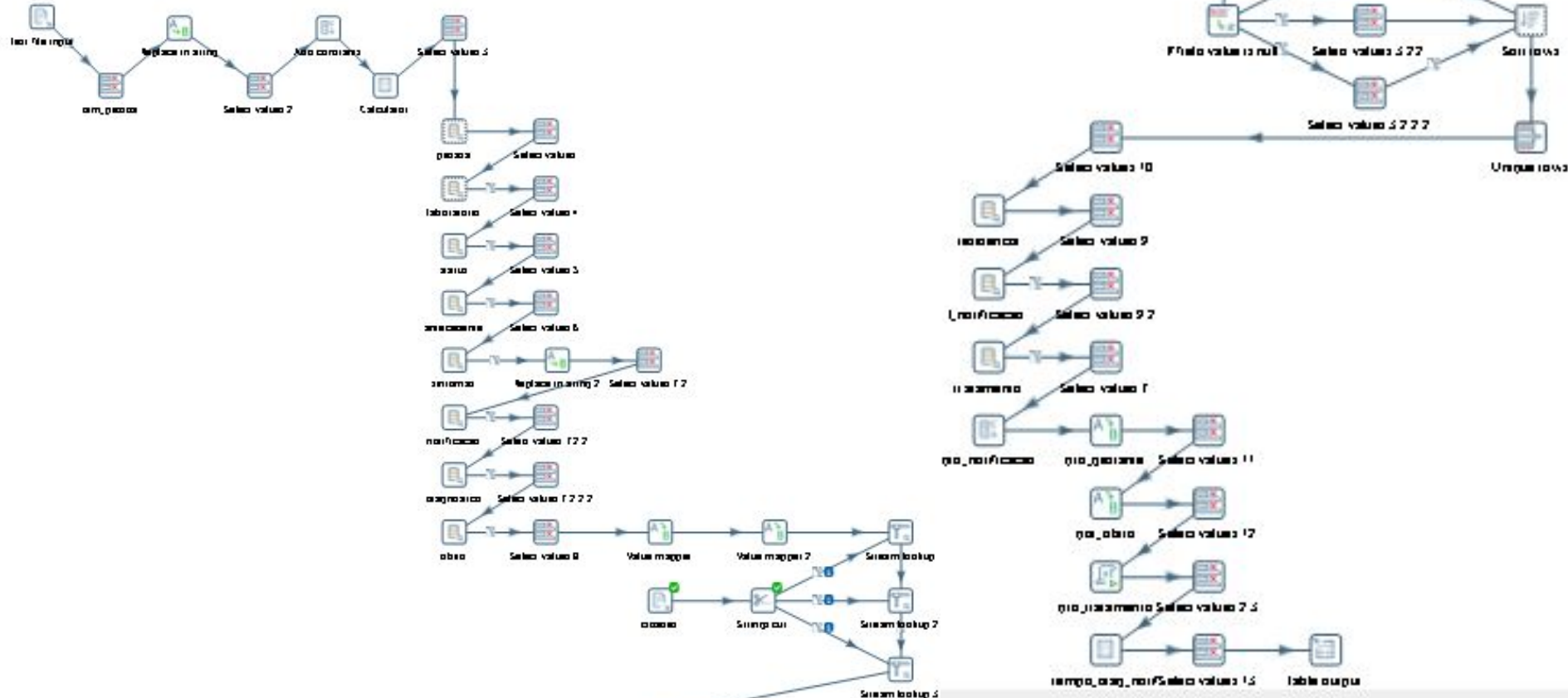
14. Plano de Carga da Dimensão Local



- 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

19. Consulta OLAP 1

Como evoluíram as notificações de HIV ao longo dos anos em cada estado?

20. Consulta OLAP 2

Qual é a taxa de óbitos por HIV em cada município?

21. Consulta OLAP 3

De todos os notificados, quantos efetivamente iniciaram o tratamento?

21. Consulta OLAP 4

Quais sintomas estão mais associados a resultados positivos nos testes laboratoriais?

22. Consulta OLAP 5

Estamos ficando mais rápidos em notificar os casos ao longo dos anos ou o sistema está piorando?

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.