Luana CristinaCroft

luanacrystinne@msn.com

Resumo

Detalha desde o processo de ETL, criação do Star Schema ate a publicação no servidor para analisar o OLAP

Relatório datamart ocorrências da aeronáutica

Usando a suíte Pentaho e banco PostgreSQL

Sumário

[Instalação 2](#_Toc89766415)

[Análise 3](#_Toc89766416)

[ETL 3](#_Toc89766417)

[**Stages** 3](#_Toc89766418)

[Dimensões 4](#_Toc89766419)

[Modelo conceitual 5](#_Toc89766420)

[Hierarquias 6](#_Toc89766421)

[Sugestões 8](#_Toc89766422)

# Instalação

Foram instalados:

1. Pentaho Ketle (PDI)o 9.2 para ETL, transformações e Jobs

<https://sourceforge.net/projects/pentaho/>

1. PostgreSQL 14 para gravar os dados das stages e do dw.

<https://www.postgresql.org/download/>

1. PGAdmin 4 para administrar o banco

<https://www.postgresql.org/download/>

1. Schema Workbench (psw-ce-3.6.1) para modelagem e publicação do Data Mart;

<https://sourceforge.net/projects/pentaho/files/Pentaho-9.2/client-tools/psw-ce-9.2.0.0-290.zip/download>

1. Pentaho Server (C:\Pentaho\pentaho-server-9) com o Plugin Saiku (Saiku with Fix for Pentaho 9x) para criar a conexão, criar e acessar o OLAP, a modelagem (cubos);

<https://sourceforge.net/projects/pentaho/files/Pentaho-9.2/server/>

<https://github.com/ambientelivre/saiku-fix/tree/main/saiku-fix-pentaho-9/saiku>

1. Jar do postgreSql

<https://jdbc.postgresql.org/download.html>

1. SQL Power Architect Para modelagem da interação das tabelas;

<http://www.bestofbi.com/page/architect_download_os>

1. Chrome ou Firefox para acessar os cubos.
2. JDK e JRE

Download: <https://www.java.com/pt-BR/>

Set a variável e ambiente “%JAVA\_HOME\%bin” ou “PENTAHO\_JAVA\_HOME”

# Análise

São dados de um período de 10 anos contendo ocorrências aeronáuticas que foram registradas no solo brasileiro para investigar e prevenir.

# ETL

As origens dos dados foram do Portal Brasileiro de Dados Abertos. A coleta dos dados foi no formato .csv.

<https://dados.gov.br/dataset/ocorrencias-aeronauticas-da-aviacao-civil-brasileira>

## **Stages**

stage\_recomendacao

stage\_aeronave

stage\_fator\_contribuinte

stage\_ocorrencia

stage\_ocorrencia\_tipo

stage\_reportes

stage\_reportes\_especies

Observação:

**Os campos em branco ou null**

1. Os campos com datas foram substituído para 19000101 para manter um padrão.
2. Os de tempo ou time foram substituídos por 0.
3. Os campos int foram substituídos por 0.
4. Os campos strings foram substituídos por NI (Não informado);

**Os campos String com acentos**

1. Retirado as acentuações.
2. Configurado para que os textos fiquem em letra maiúscula.

# Dimensões

1. Quando foi registrado

Dim\_tempo

Dim\_data

1. Qual tipo de ocorrência

Dim\_ocorrencia

1. Quem

Dim\_aeronaves

Dim\_especies

1. O quê

Fato\_ocorrencia

1. Onde

Dim\_Localidade

# Transformações criadas

## Stages

stage\_recomendacao

stage\_aeronave

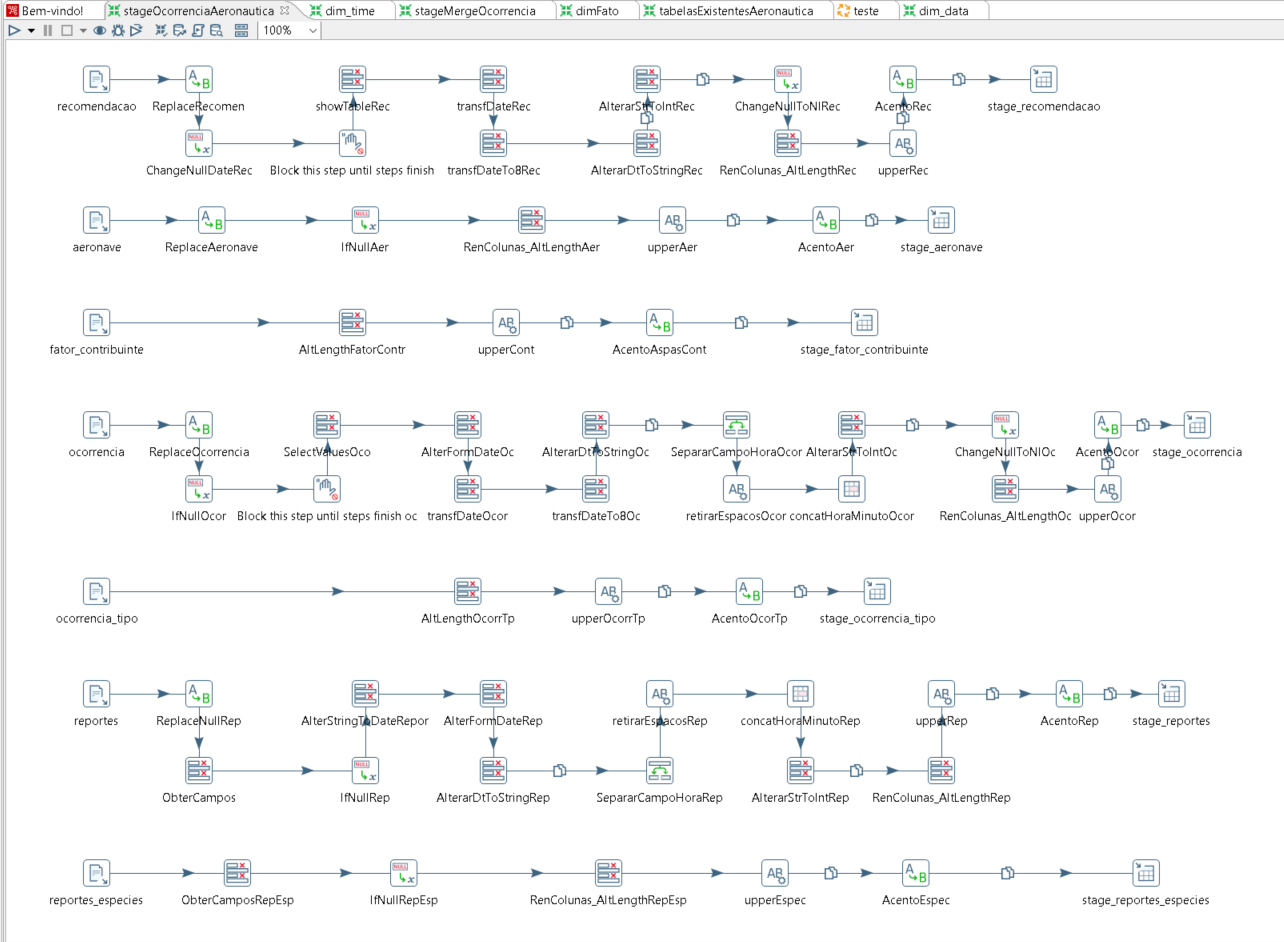
stage\_fator\_contribuinte

stage\_ocorrencia

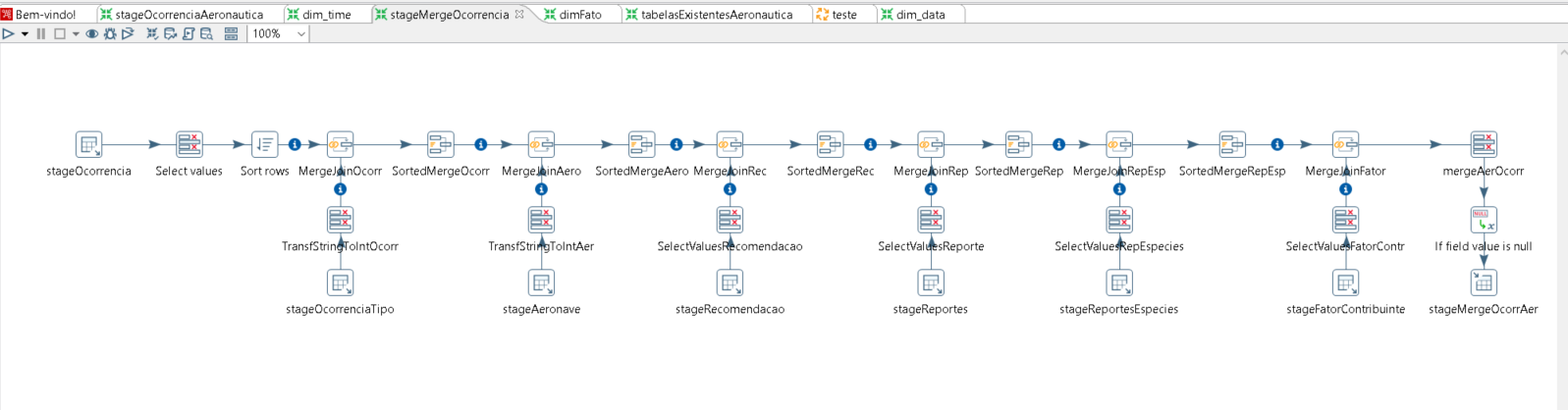
stage\_ocorrencia\_tipo

stage\_reportes

stage\_reportes\_especies

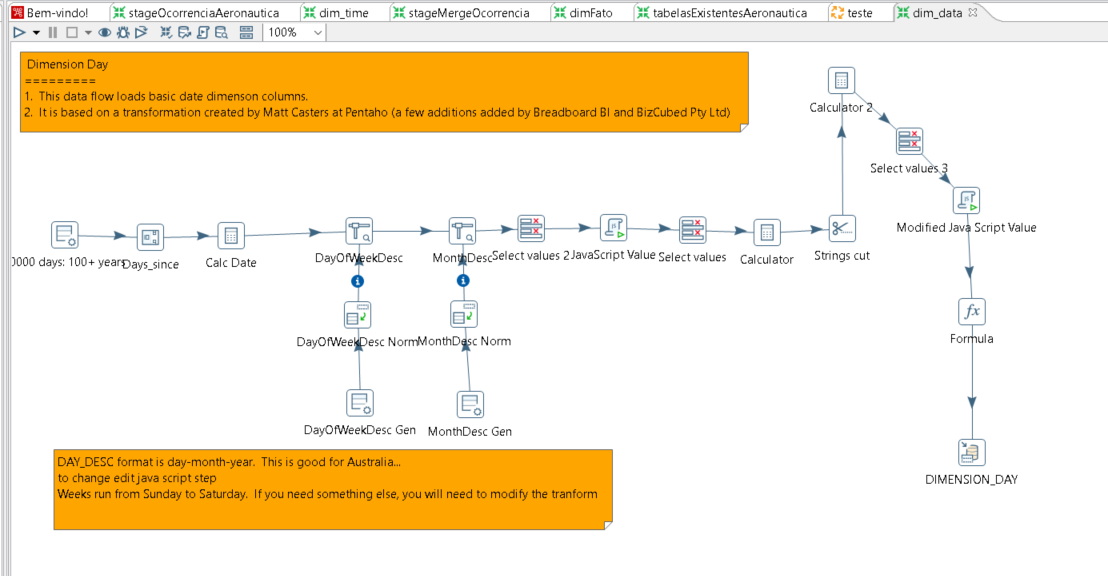


Stage única - StageMergeOcorrAer

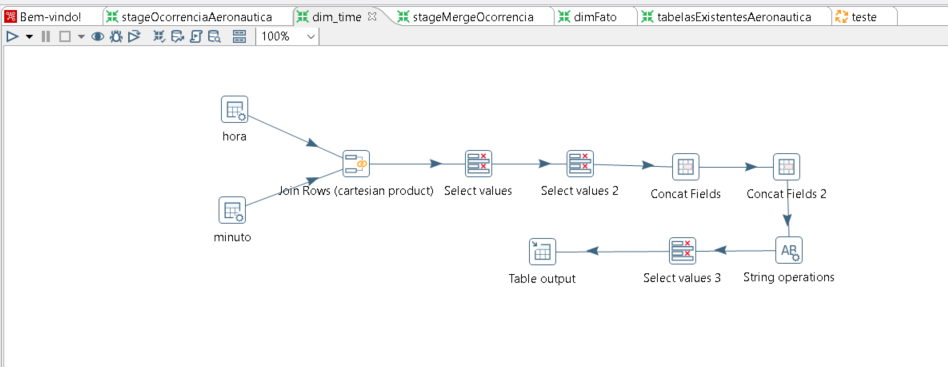


## Fato e Dimensões

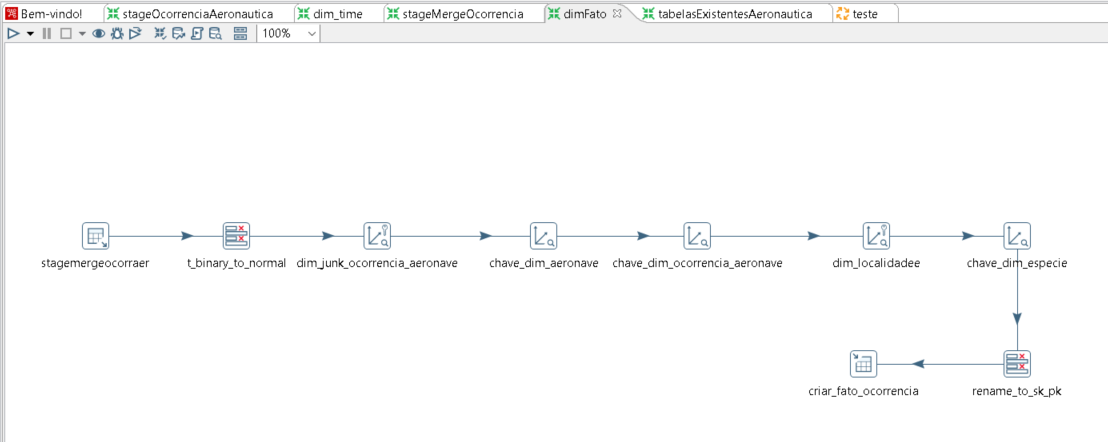
Dim\_data



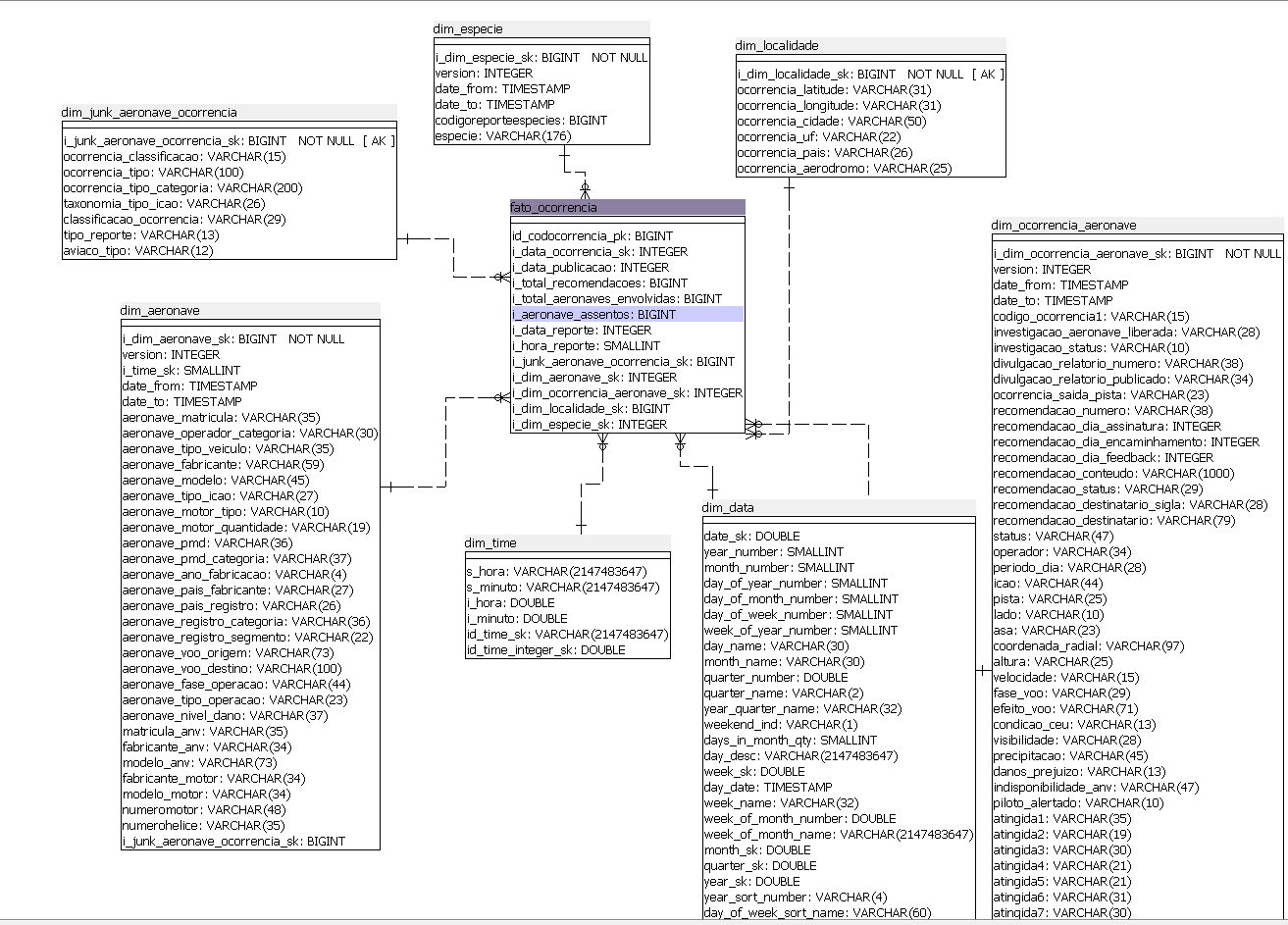
Dim\_time



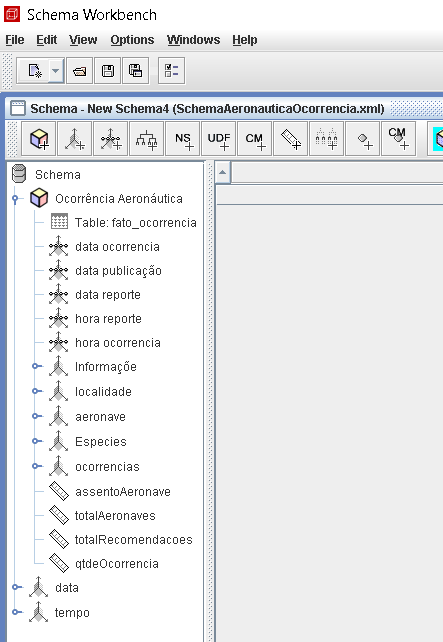
Todas as demais dimensões e fato:

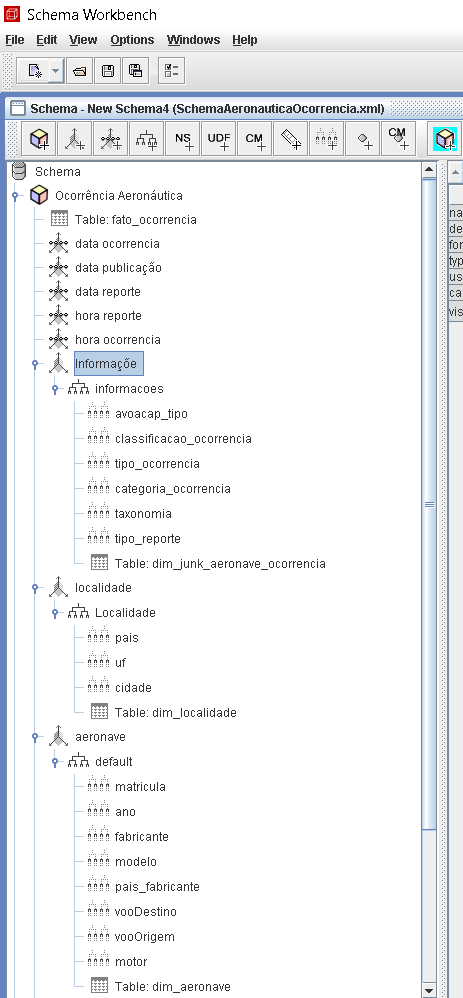


# Modelo conceitual



# Hierarquias





# Sugestões

Adicionar dados meteorológicos, migração ou comportamento dos animais em determinados períodos, manutenção das aeronaves e pistas e dados dos pilotos sem interferir na Lei de proteção dos dados.