

Acidentes de Trabalho

Jacqueson de Lima - jacquesonbenevides@gmail.com

Palloma de Lira - pallomalira18@gmail.com

Fonte dos Dados:

<https://dados.gov.br/dados/conjuntos-dados/inss-comunicacao-de-acidente-de-trabalho-cat1>

Aplicação OLAP: <link para a aplicação OLAP>

Github: https://github.com/JacquesonLima/OLAP_AcidentesDeTrabalho

ETAPA 1 - PLANEJAMENTO

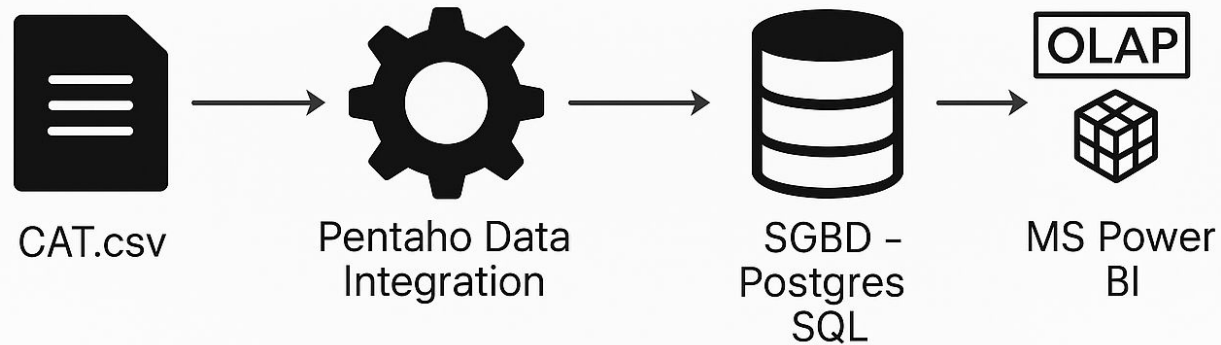
1. Contextualização

O Data Mart tem como foco analisar dados de acidentes de trabalho ocorridos dentro e fora das empresas, especificamente relacionados aos seus funcionários. A Comunicação de Acidente de Trabalho (CAT) é o documento emitido para registrar oficialmente os acidentes e doenças relacionadas ao trabalho, sendo essencial para análise e prevenção desses casos. Esses dados são de extensão nacional e abrangem empresas de todos os estados brasileiros.

2. Escopo/objetivo do Data Mart

O objetivo do Data Mart é analisar dados da CAT (Comunicação de Acidente de Trabalho), e a partir desses dados identificar padrões ou informações importantes sobre esses acidentes registrados, como o agente causador, o local do ocorrido, a ocupação do funcionário, as consequências do acidente, o setor do trabalho, região e a data do evento. Os dados analisados abrangem todo o território nacional, contemplando empresas de todos os estados do Brasil, e referem-se ao período de julho de 2018 a julho de 2020.

3. Arquitetura Tecnológica



4. Processo

1. Levantamento de requisitos
2. Definição das ferramentas
3. Coleta de dados a partir da fonte
4. Staging (padronização de campos) com ETL
5. Transformação de dados
6. Carga das Dimensões
7. Carga da Fato
8. Consumo OLAP

4. Abordagem

A abordagem utilizada foi a Bottom-up, onde construímos os dados do Data Mart a partir de fontes já existentes.

Nosso modelo de organização de dados foi o Star Schema, pois é mais simples e performático para nossa aplicação OLAP.

5. Usuários

Os usuários que utilizarão nosso Data Mart variam desde Analistas de Dados Corporativos, Gestores de Segurança de Trabalho, Gestores de RH ou Analistas de Órgãos Governamentais.

ETAPA 2 - LEVANTAMENTO DAS NECESSIDADES

6. Consultas de Apoio à Decisão

Quantos acidentes de trabalho ocorreram entre julho/2018 e setembro/2020?

Quais são os principais agentes causadores e as partes do corpo mais atingida?

Quais são as ocupações com maior incidência de acidentes?

Quais municípios ou CNAEs apresentam o maior número de acidentes?

7. Indicadores do <negócio>

Indicador	Descrição
Número total de acidentes	Serve para a contagem simples de todas as CATs registradas no período.
Taxa de óbitos por acidentes	Serve para obter um percentual de óbitos causados pelo acidente.
Taxa de acidentes por gênero	Serve para obter o percentual de acidentes de trabalho causados com um gênero específico.
Quantidade de acidentes por período	Exibe a quantidade total de acidentes em uma determinada data, podendo ser filtrada por ano, mês e dia.

<https://smartlabbr.org/sst/localidade/0?dimensao=perfilCasosAcidentes>

ETAPA 3 - MODELAGEM

9. Modelo Relacional

O modelo relacional, foi elaborado a partir dos dados disponibilizados na base da Comunicação de Acidente de Trabalho(CAT), buscando representar as principais entidades e atributos relacionados ao negócio. Foram identificados 4 entidades:

Entidades	Atributos
Empregado	ID do empregado, sexo, CBO (função), filiação segurado, parte do corpo Ating.,Especie de Benefício
Empresa	ID da empresa, CNAE2.0 (codigo), CNAE2.0(categoria), emitente CAT, local
Acidente	ID do acidente, agente causador, Natureza da lesão, Indicação de óbito, tipo de acidente
Tempo	ID do tempo, data, dia, mês, ano, trimestre e semestre

10. Modelo Dimensional

Área de Negócio	Processo do Negócio	Granularidade
Acidente de Trabalho	Análise dos acidentes de trabalho, considerando informações sobre o empregado, a empresa, o tipo do acidente, o agente causador, a região e o período do ocorrido.	Empregado -> empresa -> tipo do acidente -> data -> local

10. Modelo Dimensional

Hierarquia das Dimensões

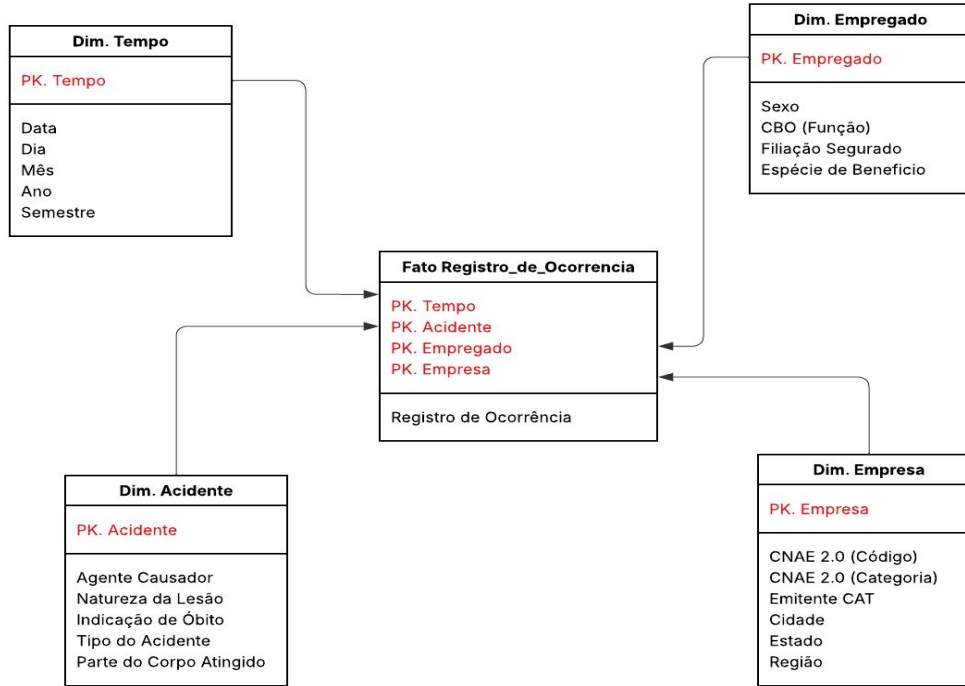
- **dim_tempo =**
 - ano -> semestre -> trimestre -> mês -> dia
- **dim_empresa =**
 - estado -> município -> empresa
- **dim_empregado =**
 - sexo, CBO(Função), filiação segurado, parte do corpo atingida, especie de benefício
- **dim_acidente =**
 - agente causador, Natureza da lesão, indicação de óbito, tipo de acidente

10. Modelo Dimensional

Métricas da fato

- Registro de Ocorrência

10. Modelo Dimensional



10. Modelo Dimensional

	PK_Tempo	PK_Acidente	PK_Empregado	PK_Empresa	Reg. de Ocorrência
1	1	101	1001	501	1
2	2	102	1002	501	1
3	1	101	1003	502	1
4	3	103	1001	501	1
5	4	104	1004	503	1
6	5	101	1005	502	1
7	6	105	1006	501	1
8	7	102	1007	503	1
9	8	106	1008	504	1
10	9	107	1009	505	1

10. Modelo Dimensional do Data Mart (lógico)

Tabela Fato:

A Tabela Fato contém cerca de 990.870 registros, cada um composto por 5 colunas de 4 bytes, totalizando 20 bytes por linha (5×4). Multiplicando esse valor pela quantidade total de registros, obtém-se 19.817.400 bytes, equivalentes a 19,82 MB.

Considerando que as dimensões ocupam aproximadamente 25% do tamanho da Tabela Fato (4,955 MB), o tamanho total estimado do Data Mart é de 24,78 MB, somando fatos e dimensões.

ETAPA 4 - PROJETO FÍSICO DO BD

11. Modelo Relacional do Data Mart (físico)

```
CREATE TABLE dim_tempo|  
(  
    id_tempo SERIAL PRIMARY KEY NOT NULL,  
    data date NOT NULL,  
    dia integer NOT NULL,  
    mes integer NOT NULL,  
    ano integer NOT NULL,  
    trimestre integer NOT NULL,  
    semestre integer NOT NULL  
);
```

11. Modelo Relacional do Data Mart (físico)

```
CREATE TABLE dim_empresa(  
    id_empresa SERIAL PRIMARY KEY,  
    cnae_codigo INTEGER NOT NULL,  
    cnae_categoria VARCHAR(255) NOT NULL,  
    emitente_cat VARCHAR(255) NOT NULL,  
    cidade VARCHAR(255) NOT NULL,  
    estado VARCHAR(255) NOT NULL  
);
```

```
CREATE TABLE dim_empregado(  
    id_empregado SERIAL PRIMARY KEY,  
    sexo VARCHAR(20) NOT NULL,  
    cbo_funcao VARCHAR(255) NOT NULL,  
    filiacao_de_segurado VARCHAR(255) NOT NULL,  
    especie_de_beneficio VARCHAR(255) NOT NULL  
);
```

11. Modelo Relacional do Data Mart (físico)

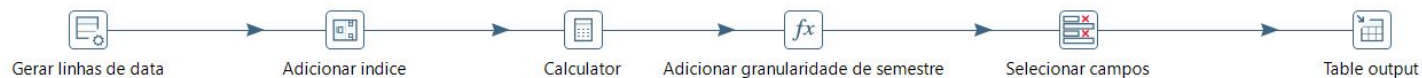
```
CREATE TABLE dim_acidente (  
    id_acidente SERIAL PRIMARY KEY,  
    agente_causador      VARCHAR(255) NOT NULL,  
    natureza_lesao       VARCHAR(255) NOT NULL,  
    obito                VARCHAR(10)  NOT NULL,  
    tipo_acidente        VARCHAR(255) NOT NULL,  
    parte_corpo_atingida VARCHAR(255) NOT NULL  
);
```

11. Modelo Relacional do Data Mart (físico)

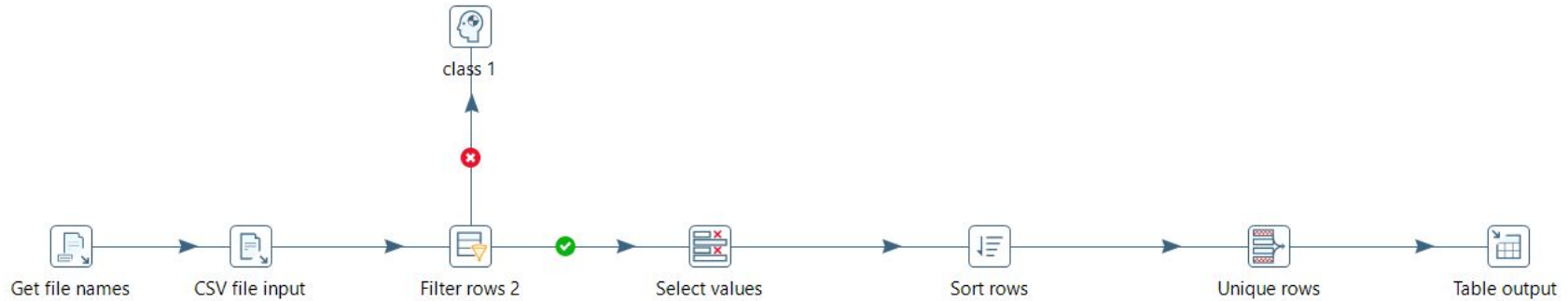
```
CREATE TABLE fato (  
    tempo_sk INTEGER  
    NOT NULL REFERENCES dim_tempo (id_tempo),  
  
    empregado_sk INTEGER  
    NOT NULL REFERENCES dim_empregado (id_empregado),  
  
    empresa_sk INTEGER  
    NOT NULL REFERENCES dim_empresa (id_empresa),  
  
    acidente_sk INTEGER  
    NOT NULL REFERENCES dim_acidente (id_acidente),  
  
    registro_ocorrencia INTEGER  
    NOT NULL DEFAULT 1,  
  
    CONSTRAINT fato_pk  
        PRIMARY KEY (tempo_sk, empregado_sk, empresa_sk, acidente_sk)  
);
```


ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

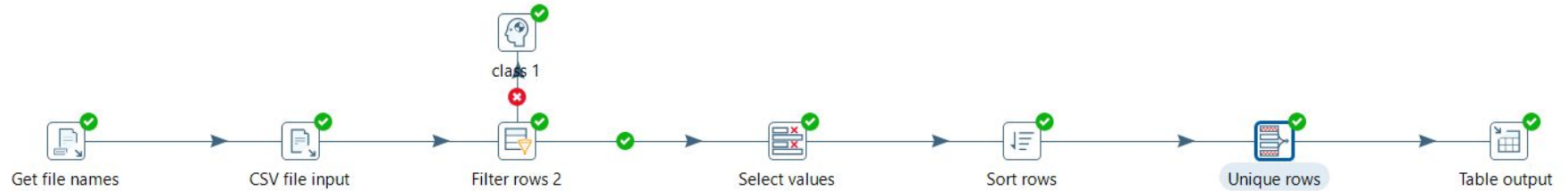
12. Plano de Carga da Dimensão Tempo



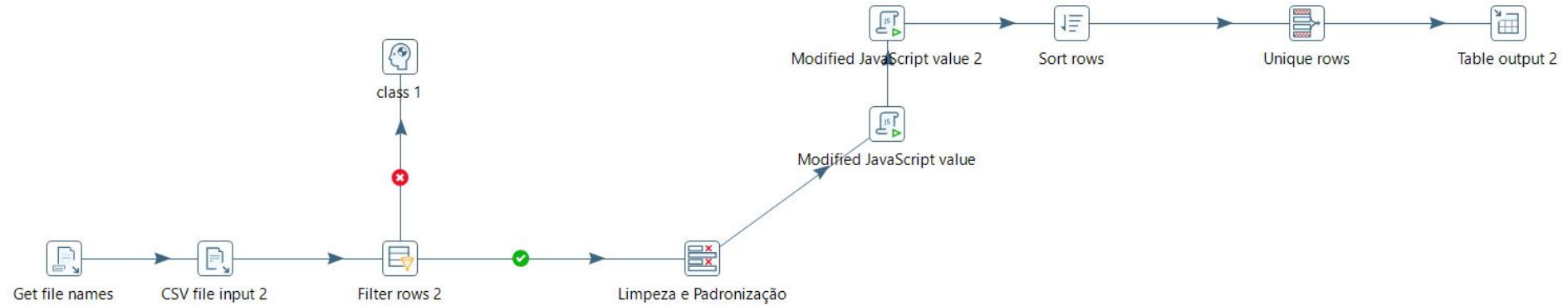
13. Plano de Carga da Dimensão Empresa



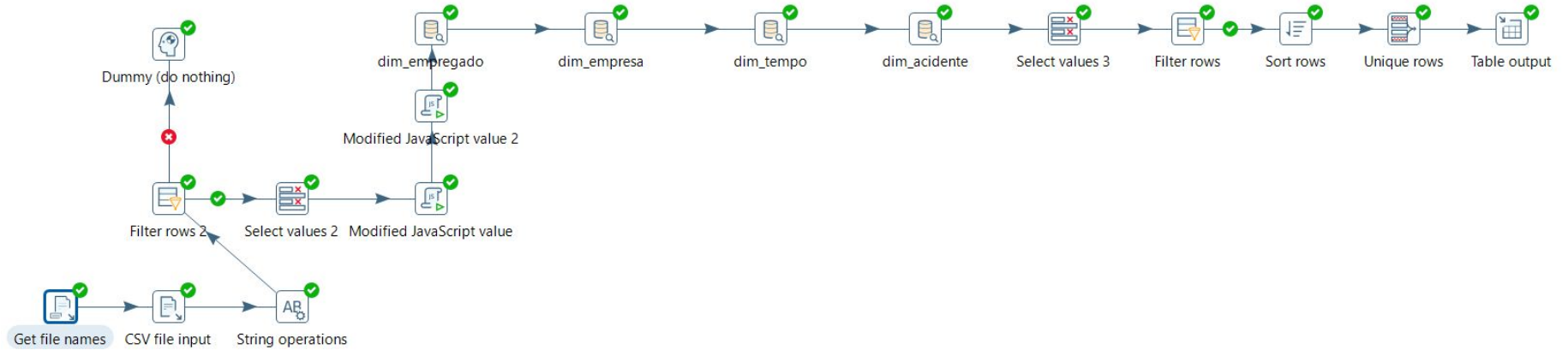
14. Plano de Carga da Dimensão Empregado



14. Plano de Carga da Dimensão Acidente

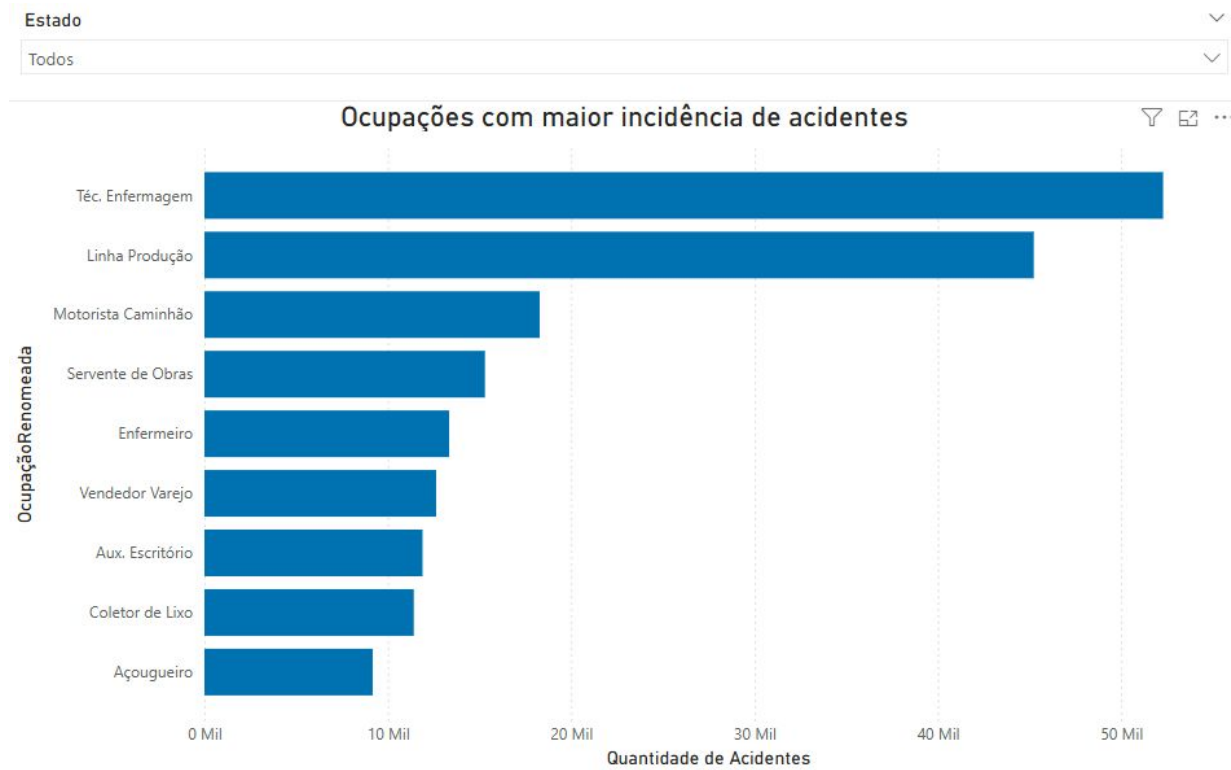


15. Plano de Carga da Fato

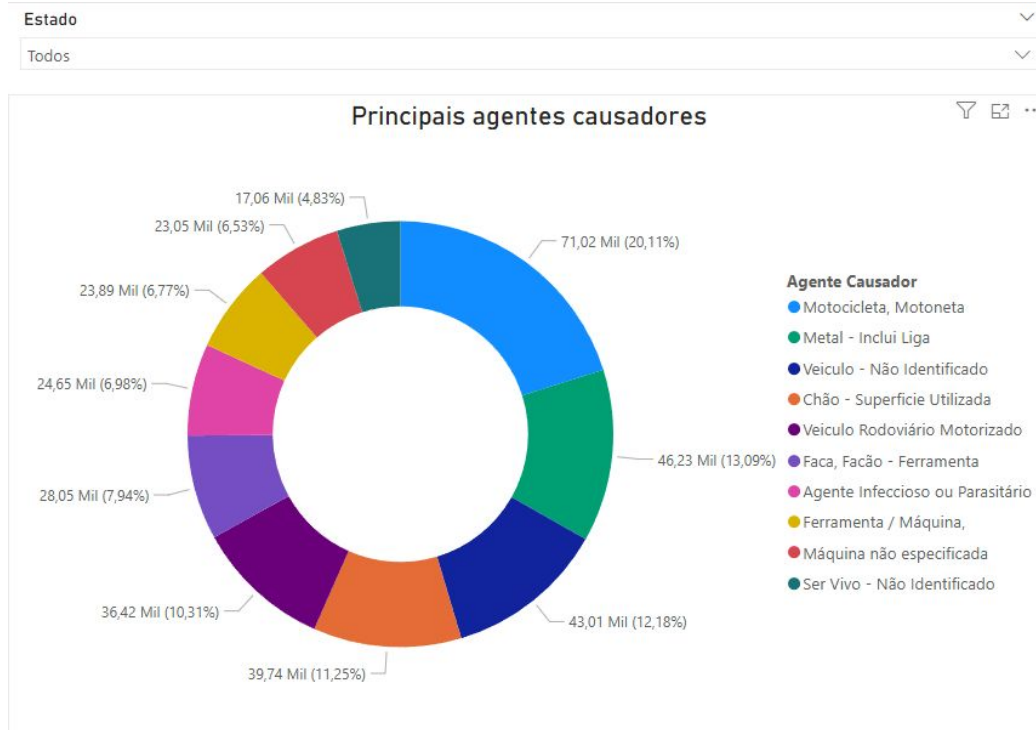


ETAPA 6 - APLICAÇÃO OLAP e PAINEL DE BORDO

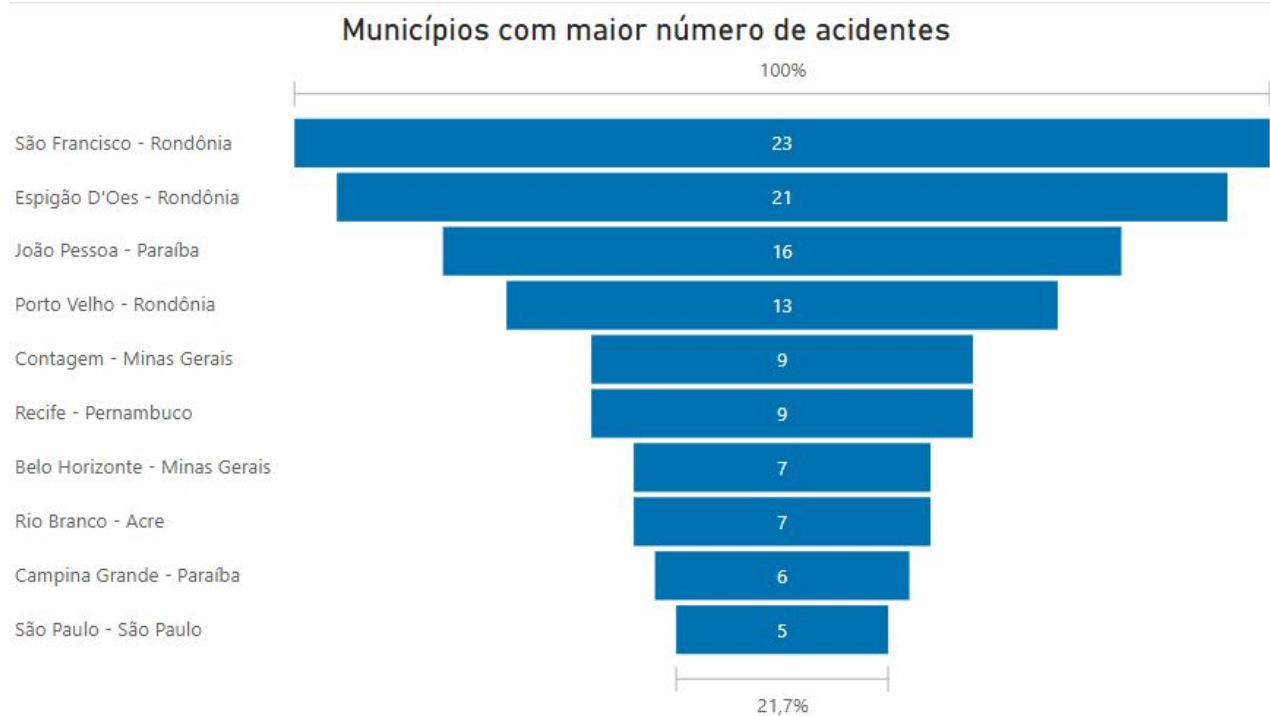
16. Consulta OLAP 1



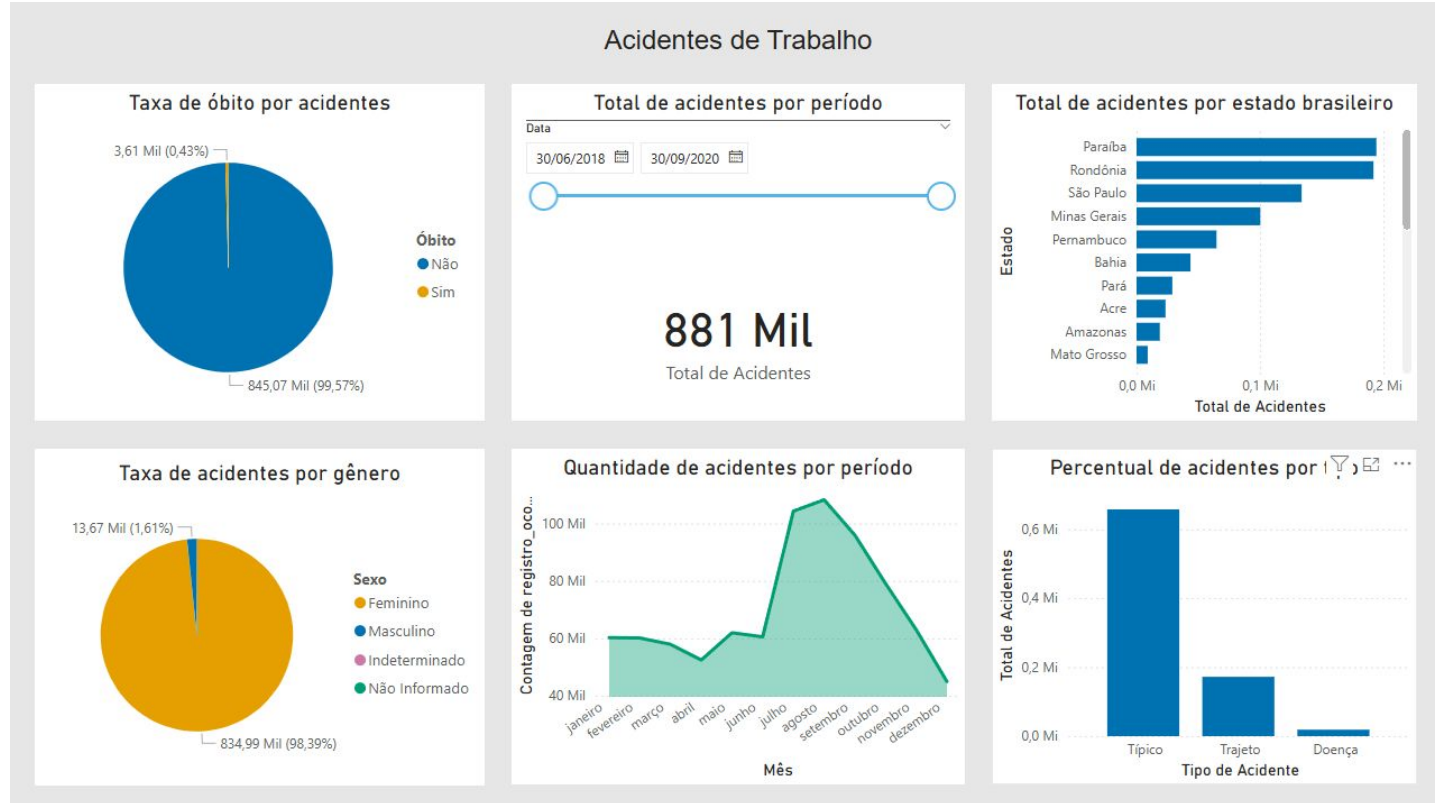
16. Consulta OLAP 2



16. Consulta OLAP 3



16. Painel de Bordo XXXX (ponto extra)



Referências

INSTITUTO NACIONAL DO SEGURO SOCIAL (INSS). *Comunicação de Acidente de Trabalho – CAT*. Portal Dados Abertos. Disponível em:

<https://dados.gov.br/dados/conjuntos-dados/inss-comunicacao-de-acidente-de-trabalho-cat1>.