

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**NÚCLEO DE EDUCAÇÃO A DISTÂNCIA

Pós-graduação *Lato Sensu* em Gestão e Análise Estratégica de Dados

**RELATÓRIO ANALÍTICO**

ANÁLISE DE PERÍODICIDADE DE ACIDENTES DE TRÂNSITO NO ESTADO DE SÃO PAULO

Ramon Mendes Santos

Belo Horizonte

2024

**SUMÁRIO**

[1. Introdução 3](#_Toc132204826)

[1.1. Contexto 3](#_Toc132204827)

[1.2. Objetivos 3](#_Toc132204828)

[1.3. Público alvo 3](#_Toc132204829)

[2. Modelos de Dados 4](#_Toc132204830)

[2.1. Modelo Dimensional 4](#_Toc132204831)

[2.2. Fatos e Dimensões 5](#_Toc132204832)

[3. Integração, Tratamento e Carga de Dados 12](#_Toc132204833)

[3.1. Fontes de Dados 12](#_Toc132204834)

[3.2. Processos de Integração e Carga (ETL) 14](#_Toc132204835)

[4. Camada de Apresentação 15](#_Toc132204836)

[4.1 Dashboard 15](#_Toc132204837)

[5. Registros de Homologação 21](#_Toc132204838)

[5. Conclusões 26](#_Toc132204839)

[6. Links 28](#_Toc132204840)

[REFERÊNCIAS 29](#_Toc132204841)

# 1. Introdução

## 1.1. Contexto

Neste trabalho apresento um projeto para análise de dados de acidentes de trânsito do estado de São Paulo.

A ferramenta poderá ser utilizada por empresas de consultoria de engenharia de trânsito para acompanhamento de quais municípios do estado, em qual período e evolução para traçar estratégia que ajude a diminuir tais ocorrências.

Para isso, foram utilizadas as ferramentas como Knime Analytics Plataform, para extração de dados da web, SQL Server para armazenar os dados e a ferramenta de visualização MS Power BI, onde será apresentado os dados de forma visual.

## 1.2. Objetivos

Este trabalho tem como objetivo expor municípios onde ocorre acidentes de trânsito com maior frequência no estado de São Paulo, em quais sazonalidades e situações climáticas a fim de se criar estratégias que evitem ocorrências de acidentes fatais e não fatais. Para isso, será criada, um dashboard com níveis estratégico, tático e operacional.

## 1.3. Público alvo

O relatório tem como público-alvo empresas e consultoria de engenharia de trânsito no estado de São Paulo. Serão utilizados por profissionais tecnicamente preparados. As personas são: gestores executivos, que utilizaram a página estratégica, engenheiros de trânsito e gestores nível tático que utilizaram a página tática e analistas e supervisores que utilizaram a página operacional.

# 

# 2. Modelos de Dados

## 2.1. Modelo Dimensional

Para atender a necessidade deste projeto, conforme imagem abaixo, foi utilizado o modelo de dados snowflake no software MS Power BI (Figura 1).

Figura 1 – Modelo de Dados SnowFlake

Interface gráfica do usuário

Descrição gerada automaticamente com confiança média

Fonte: Autoria própria do projeto desenvolvido (2024)

De acordo com Felipe Nery Rodrigues Machado (2020), o modelo de dados é um conjunto de conceitos que podem ser utilizados para descrever as estruturas lógicas e físicas de um banco de dados.

Conforme diz no site da Microsoft (2023), uma dimensão snowflake é um conjunto de tabelas normalizadas para uma única entidade de negócio. No qual, a uma entidade é atribuída a uma subcategoria, no caso do projeto apresentado, a dimensão feriado está atribuído a dimensão calendário.

Uma entidade é um agrupamento lógico de informações inter-relacionadas necessárias para a execução das atividades em um sistema, (Oliveira, Celso 2022).

## 2.2. Fatos e Dimensões

Tabela dimensão feriado, consta os principais feriados do ano, no qual são utilizados nas análises. São eles:

Carnaval, Confraternização Universal, Corpus Christi, Dia do Trabalho, Dia Nacional do Zumbi e da Consciência Negra, Dia de Finados, Independência do Brasil, Natal, Paixão de Cristo, Proclamação da República, Tiradentes, Nossa senhora Aparecida – Padroeira do Brasil.

|  |  |  |
| --- | --- | --- |
| **dim\_feriados:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Data | dia, mês e ano | data |
| Dia da Semana | Nome do dia da semana | texto |
| Feriado | Nome do feriado | Texto |

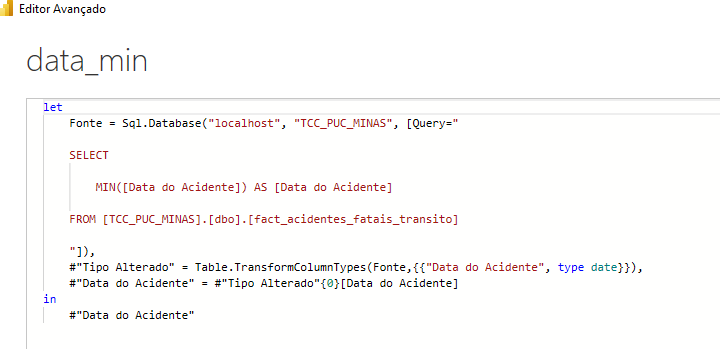
Tabela dimensão calendário, que é usado para filtrar períodos nas análises

E compara períodos distintos.

|  |  |  |
| --- | --- | --- |
| **dim\_calendario:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Data | dia mês e ano | data |
| Dia da Semana | número do dia da semana | inteiro |
| Nome do Dia | nome do dia da semana | texto |
| Mês | número do mês | inteiro |
| Nome do Mês | nome do mês por extenso | texto |
| Início do Mês | data inicial do mês | data |
| Fim do Mês | data final do mês | data |
| Trimestre | trimestre do ano | texto |
| Ano | ano | inteiro |

Para dinamizar a tabela de calendário, foi utilizado a linguagem M, nativa no MS Power BI, para encontrar a data mínima (Figura 1.1). Para a data máxima (Figura 1.2), no qual é o ano atual e inserido ambos em uma função personalizada, conforme imagem abaixo (Figura 1.3).

Figura 1.1 – Menor data



Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.2 – Ano atual

Texto

Descrição gerada automaticamente com confiança média

Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.2 – dimensão calendário

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Abaixo, dependências das consultas das tabelas no Power Query (Figura 2) após automação da tabela dim\_calendario.

Figura 2 – Dependência das consultas



Fonte: Autoria própria do projeto desenvolvido (2024)

Dimensão local acidente, consulta direta no banco de dados local. Nele contém valores únicos dos locais onde ocorreram as ocorrências de trânsito no estado de São Paulo, contendo o nome do município, logradouro também, latitude e longitude para criação de mapas.

|  |  |  |  |
| --- | --- | --- | --- |
| **dim\_local\_acidente:** |  |  |  |
| Nome da coluna | Descrição | Tipo |  |
| PK\_LOCAL | chave primaria da tabela | Inteiro |  |
| MUNICIPIO | nome do município | Texto |  |
| LOGRADOURO | nome do logradouro | texto |  |
| LATITUDE | latitude do local | número decimal |  |
| LONGITUDE | longitude do local | número decimal |  |

Tabela tipo ocorrência, é uma tabela auxiliar para criação de medida DAX, sendo possível alternar entre medidas através da coluna código com o uso da função switch.

O tipo de ocorrência refere-se a fatais e não fatais, onde ocorrências fatais são acidentes onde houve vítimas que foram a óbito, e a não fatais são onde as vítimas tiveram ou não lesões físicas, mas sobreviveram ao acidente.

|  |  |  |
| --- | --- | --- |
| **TIPO\_OCORRENCIA:** |  |  |
| Nome da coluna | Descrição | Tipo |
| TIPO OCORRENCIA | tipo de ocorrência | texto |
| CODIGO | código do tipo de ocorrência | Inteiro |

Tabela dimensão hora do acidente criada via linguagem M nativa do software, onde é feito a junção das duas tabelas, colunas horas registradas na tabela fato com acidentes fatais e não fatais, conforme imagem abaixo (Figura 3.1), com valores únicos.

Figura 3.1 – Dim hora

Interface gráfica do usuário, Texto, Aplicativo

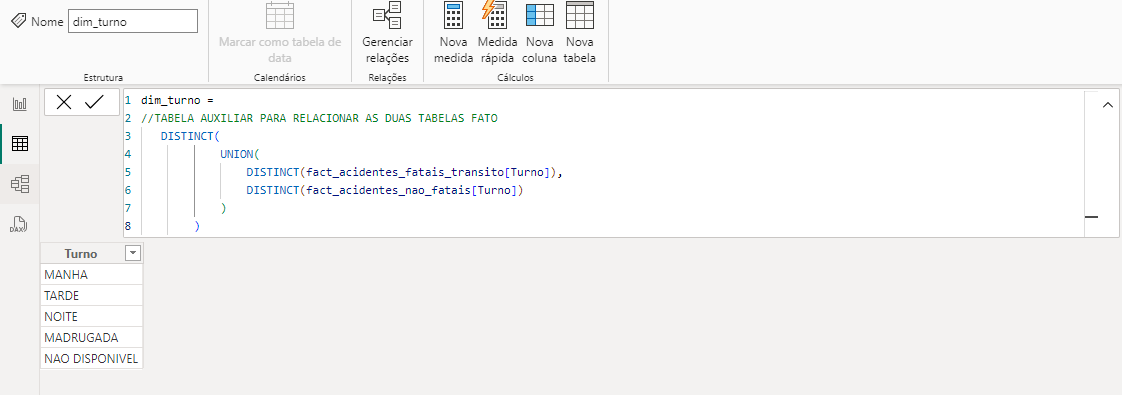
Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

|  |  |  |
| --- | --- | --- |
| **dim\_hora\_do\_acidente** |  |  |
| Nome da coluna | Descrição | Tipo |
| Hora do Acidente | hora do acidente em registro | Hora |
| Turno | período do dia | Texto |

Tabela dimensão turno, criada com medidas DAX, de forma dinâmica, para contabilizar as ocorrências entre as tabelas fato. Contendo uma coluna que descreve o período do dia informado na ocorrência conforme segue abaixo (Figura 3.2).

Figura 3.2 – Dim turno



Fonte: Autoria própria do projeto desenvolvido (2024)

|  |  |  |
| --- | --- | --- |
| **Tabela dim\_turno:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Turno | período do dia | texto |

Tabela dimensão condições climáticas, criada com medidas DAX, de forma dinâmica, para contabilizar as ocorrências entre as tabelas fato. Contendo uma coluna

que descreve o clima no dia da ocorrência para auxiliar na análise descritiva (Figura 3.3).

Figura 3.3 – Dim condições climáticas

Interface gráfica do usuário, Texto, Aplicativo, Email

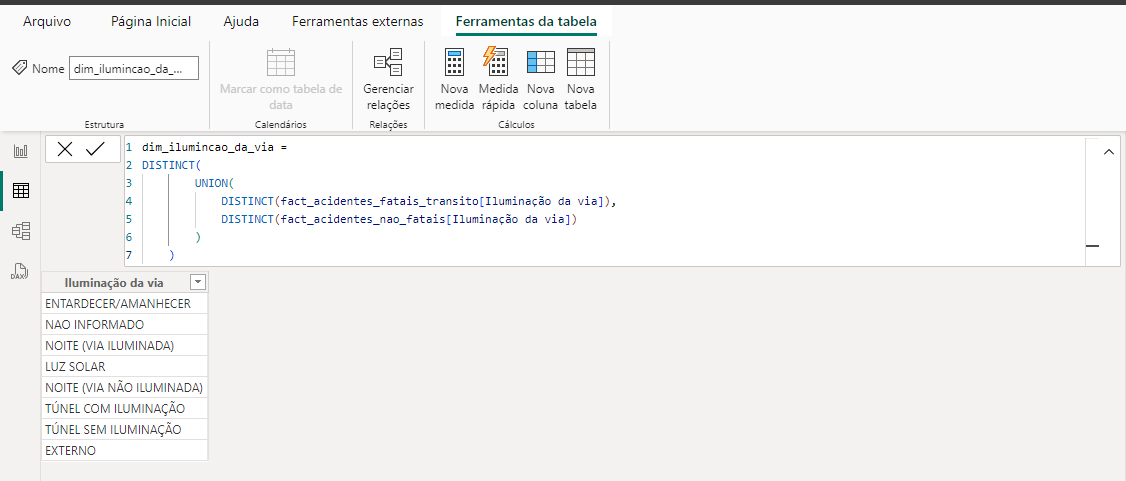
Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

|  |  |  |
| --- | --- | --- |
| **dim\_condicoes\_climaticas:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Condições Climáticas | clima do dia | texto |

Tabela dimensão iluminação da via, criada com medidas DAX, de forma dinâmica, para contabilizar as ocorrências entre as tabelas fato. Contendo uma coluna que descreve a iluminação da via no dia da ocorrência para auxiliar na análise descritiva, conforme segue abaixo (Figura 3.4):

Figura 3.4 – Tabela dimensão iluminação da via



Fonte: Autoria própria do projeto desenvolvido (2024)

|  |  |  |
| --- | --- | --- |
| **dim\_ilumincao\_da\_via:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Iluminação da via | visibilidade da via | texto |

Tabela fato acidentes fatais de trânsito, traz dados descritivos sobre ocorrências de vítimas fatais. Extraído do banco de dados local.

|  |  |  |
| --- | --- | --- |
| **fact\_acidentes\_fatais\_transito:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Data do Acidente | data do dia do acidente | date time |
| Hora do Acidente | hora do acidente | texto |
| Região Administrativa | região de administração | texto |
| Numeral / KM | quilometragem do carro | texto |
| Jurisdição | jurisdição do acidente | texto |
| Administração | administração da via | texto |
| Conservação | administração da via | texto |
| Tipo do Local do Acidente | local do acidente | texto |
| Outro Veículo Envolvido | tipo de veículo envolvido | texto |
| Tipo de via | tipo da via | texto |
| Quantidade de vítimas | quantidade de vítimas | inteiro |
| Sentido da Via | sentido regional | texto |
| Limite da velocidade da via | limite da velocidade da via | texto |
| Número BO | número de boletim de ocorrência | texto |
| Superfície da Via | estado da superfície da via | texto |
| Tipo de pista | quantidade de sentido da via | texto |

Tabela fato acidentes não fatais de trânsito, traz dados descritivos sobre ocorrências de vítimas sobreviventes aos acidentes, desde pessoas ilesas as pessoas em estado mais grave. Extraído do banco de dados local.

|  |  |  |
| --- | --- | --- |
| **fact\_acidentes\_nao\_fatais:** |  |  |
| Nome da coluna | Descrição | Tipo |
| Data do Acidente | data de ocorrência do acidente | data |
| Hora do Acidente | hora da ocorrência do acidente | texto |
| Batalhão de Trânsito | batalhão de trânsito | texto |
| Companhia de Trânsito | companhia de trânsito | texto |
| Região Administrativa | região administrativa | Texto |
| Número/KM | quilometragem | Texto |
| Jurisdição | jurisdição | Texto |
| Administração | consercionária ou prefeitura | Texto |
| Conservação | conservação da via | Texto |
| Veículos Envolvidos - Bicicleta | quantidade de envolvidos por tipo de veículo | Inteiro |
| Veículos Envolvidos - Caminhão | quantidade de envolvidos por tipo de veículo | Inteiro |
| Veículos Envolvidos - Automóvel | quantidade de envolvidos por tipo de veículo | Inteiro |
| Veículos Envolvidos - Motocicleta | quantidade de envolvidos por tipo de veículo | Inteiro |
| Veículos Envolvidos - Ônibus | quantidade de envolvidos por tipo de veículo | Inteiro |
| Veículos Envolvidos - Pedestre | quantidade de envolvidos por gravidade | Inteiro |
| Pessoas Envolvidas - Grave | quantidade de envolvidos por gravidade | Inteiro |
| Pessoas Envolvidas - Ileso | quantidade de envolvidos por gravidade | Inteiro |
| Pessoas Envolvidas - Leve | quantidade de envolvidos por gravidade | Inteiro |
| Tipo de Via | responsável pela via (município, estado, federação) | Texto |
| Condições Climáticas | condição do clima no dia do acidente | Texto |
| Mão de direção | quantidade de direção na via | Texto |
| Obras na pista | status de obras na via | Texto |
| Relevo | planície da via | Texto |
| Superfície da via | status de limpeza da via | Texto |
| Tipo de pavimento | tipo de pavimentação da via | Texto |
| Tipo de pista | quantidade de direção na via | Texto |
| Traçado | informa se há cruzamento na via | texto |

# 3. Integração, Tratamento e Carga de Dados

## 3.1. Fontes de Dados

Os arquivos utilizados para a análise dos dados foram extraídos da página web do site do estado de São Paulo. Sendo a base de acidentes fatais em formato .xlsx, conforme segue abaixo (Figura 1.1).

Figura 1.1 – Base acidentes fatais

Interface gráfica do usuário, Aplicativo, Tabela

Descrição gerada automaticamente

E, a base de acidentes não fatais em formato de .csv (Figura 1.2).

Figura 1.2 – Base acidentes não fatais

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Os arquivos são tratados no pipeline de dados desenvolvido na ferramenta Knime Analytics e inserido em um banco de dados local, conforme imagem abaixo (Figura 1.3).

Figura 1.3 – banco de dados local – SQL Server

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

## 3.2. Processos de Integração e Carga (ETL)

Para a extração dos dados da web foi criado um pipeline de dados utilizando a ferramenta Knime Analytics.

Na primeira parte do processo foi criado um meta node onde ele faz o download dos arquivos por meio de uma web link (Figura 1) baixando direto na pasta downloads do sistema operacional (Figura 1.1). O arquivo de feriados nacionais foi baixado manualmente e importado via power query no software MS Power BI (Figura 1.1).

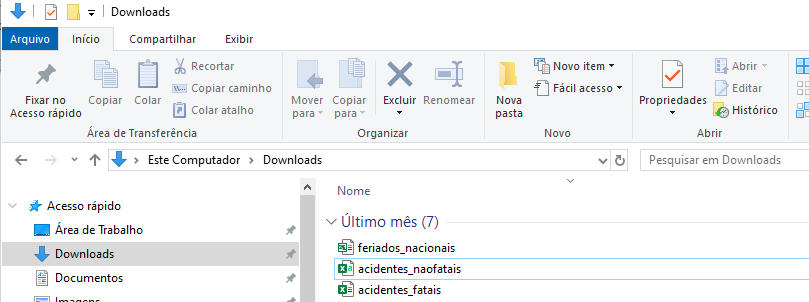
Figura 1 – Meta node no knime

Uma imagem contendo Interface gráfica do usuário

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.1 – Pasta downloads

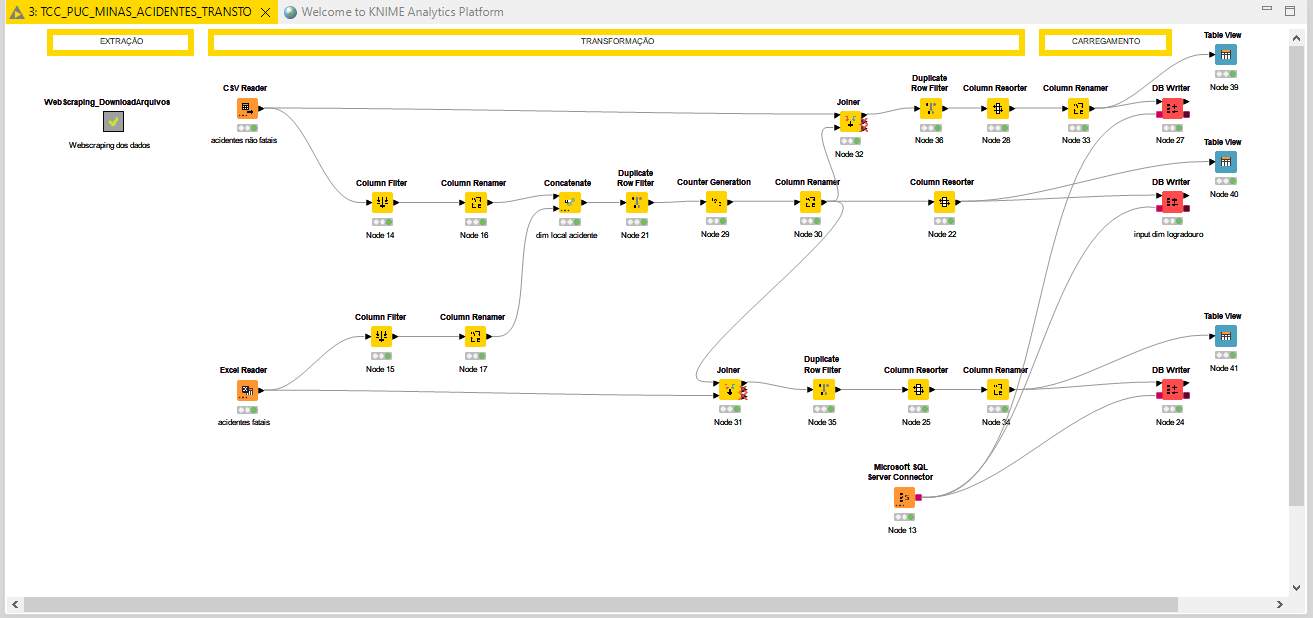


Fonte: Autoria própria do projeto desenvolvido (2024)

Na imagem a seguir (Figura 1.2), vemos o processo de transformação dos dados até o momento do carregamento em um banco de dados local, SQL Server.

Foi criado a tabela dimensão logradouro a partir das duas tabelas fatos a fim de normalizar os dados. Também, foi removido linhas duplicadas pela coluna de ID da tabela fato de acidentes não fatais, reduzindo assim a quantidade de linhas.

Figura 1.2 – Transformação dos dados



Fonte: Autoria própria do projeto desenvolvido (2024)

A partir do momento em que o banco de dados é populado pelo fluxo de dados, fica disponível o acesso para alimentação de tabelas em ferramentas OLAP.

# 4. Camada de Apresentação

## 4.1 Dashboard

No desenvolvimento do dashboard foi criado três telas principais onde estão dispostas as análises para acompanhamento e tomada de decisão. São elas: estratégico, tático e operacional.

Na tela estratégica (Figura 1.1) mostra um panorama geral para acompanhamento de evolução das ocorrências de acidentes de trânsito no estado de São Paulo. A página possui 5 cartões com os totais de ocorrência fatal e não fatal e o percentual que cada um possui na somatória das duas situações. Abaixo dos cartões, foi disposto um gráfico de barras com os totais de ocorrência por ano.

No gráfico de linhas ocorrências por ano, trago uma previsão de ocorrência para os próximos 3 anos com o intervalo de confiança de 95%.

No segundo gráfico de linhas, está demonstrado o acumulado de ocorrências do ano em comparação ao ano anterior.

Ao lado direito do painel está disposto um gráfico de mapa onde mostra de forma visual a densidade do local onde ocorreu maior incidência de acidentes.

Figura 1.1 – tela estratégica

Interface gráfica do usuário

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Na tela tática (Figura 1.2) é apresentado uma análise climática, iluminação da via e temporal. É apresentado as quantidades de vítimas por sua gravidade em cartões, um gráfico de funil para dimensionar a quantidade de acidentes por clima, um gráfico de barras por iluminação da via e um gráfico de barras para dimensionar em qual dia da semana ocorrem mais acidentes.

Na tabela está disposto os tipos de gravidade de acidente por turno do dia com a possibilidade de expandir a linha para mostrar por município.

No último gráfico de barras na parte inferior do painel está exposto o grau de densidade por horário, mostrando os picos e maior concentração.

Figura 1.2 – tela tática

Interface gráfica do usuário

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Na tela operacional (Figura 1.3), é possível acompanhar a evolução das estratégias adotadas a fim de diminuir a quantidade de ocorrência de acidentes de trânsito no estado de São Paulo.

O painel trás comparativo de quantidade de ocorrência na linha do tempo com o intuito de ter de forma visual a eficiência da estratégia adotada para determinados feriados, dia da semana, municípios e ruas.

Está disposto também, quais vias tem maior participação de acidentes e a proporção do tipo de gravidade dos acidentes.

Figura 1.3 – tela operacional

Tela de jogo de vídeo game

Descrição gerada automaticamente com confiança média

Fonte: Autoria própria do projeto desenvolvido (2024)

Para analisar os dados e dispor em diferentes tipos de cenários nos painéis, foram criadas medidas DAX (data analisys expressions) para contextualizar os dados.

Quantidade de vítimas fatais: somatório da coluna quantidade de vítimas da tabela fact acidentes fatais trânsito.

QTD DE VITIMAS FATAIS = SUM(fact\_acidentes\_fatais\_transito[Quantidade de vítimas])

Quantidade de vítimas grave: somatório da coluna pessoas envolvidas grave, da tabela fact acidentes não fatais.

QTD VITIMAS - GRAVE = SUM(fact\_acidentes\_nao\_fatais[Pessoas Envolvidas - Grave])

Quantidade de vítimas ilesas: somatório da coluna pessoas envolvidas ileso da tabela fact acidentes não fatais.

QTD VITIMAS - ILESO = SUM(fact\_acidentes\_nao\_fatais[Pessoas Envolvidas - Ileso])

Quantidade de vítimas leve: somatório da coluna pessoas envolvidas leve da tabela fact acidentes não fatais.

QTD VITIMAS - LEVE = SUM(fact\_acidentes\_nao\_fatais[Pessoas Envolvidas - Leve])

Quantidade de ocorrências fatais: contagem de linhas da tabela acidentes fatais de trânsito.

QTDE DE OCORRENCIAS FATAIS = COUNTROWS(fact\_acidentes\_fatais\_transito)

Quantidade de ocorrências não fatais: contagem de linhas da tabela acidentes não fatais de trânsito.

QTDE DE OCORRENCIAS NAO FATAIS = COUNTROWS(fact\_acidentes\_nao\_fatais)

Tipo ocorrência: seleciona a métrica a partir da tabela suporte Tipo de ocorrência, sendo o valor 1 para ocorrências com acidentes fatais, 2 para quantidade de ocorrências não fatais e uma terceira opção com o total de ocorrência.

TIPO OCORRENCIA =

SWITCH(TRUE(),

    SELECTEDVALUE(TIPO\_OCORRENCIA[CODIGO]) = 1,

        [QTDE DE OCORRENCIAS FATAIS],

    SELECTEDVALUE(TIPO\_OCORRENCIA[CODIGO]) = 2,

        [QTDE DE OCORRENCIAS NAO FATAIS],

    [TOTAL DE OCORRENCIAS]

)

Total de ocorrência: é a somatória das medidas de quantidade de ocorrências fatais e quantidade de ocorrências não fatais.

TOTAL DE OCORRENCIAS = [QTDE DE OCORRENCIAS FATAIS] + [QTDE DE OCORRENCIAS NAO FATAIS]

Tipo ocorrência ano anterior: calcula a quantidade de ocorrência no mesmo período da seleção de datas referente ao ano anterior.

TIPO OCORRENCIA ANO ANTERIOR = CALCULATE([TIPO OCORRENCIA],

    DATEADD(dim\_calendario[Data], -1, YEAR)

)

Ytd ocorrência: year to date é uma medida que acumula as ocorrências em um determinado período.

YTD OCORRENCIA = CALCULATE([TIPO OCORRENCIA],

    DATESYTD(dim\_calendario[Data])

)

Ytd ocorrência último ano: year to date ocorrências do último ano é uma medida que acumula as ocorrências em um determinado período referente ao ano anterior ao ano selecionado.

YTD OCORRENCIA ULTIMO ANO =

 CALCULATE([YTD OCORRENCIA],

    DATEADD(dim\_calendario[Data], -1, YEAR)

)

% evolução de ocorrência: compara a evolução na linha do tempo dividindo a ocorrência do mês e ano atual ao mesmo mês do ano anterior subtraindo -1 para trazer apenas o percentual de evolução.

% EVOLUÇÃO DE OCORRENCIAS = DIVIDE ([TIPO OCORRENCIA], [TIPO OCORRENCIA ANO ANTERIOR], 0) - 1

% ocorrências fatais: medida utilizada para identificar a participação de cada atributo do total. Divide-se a quantidade do atributo pelo total geral.

% OCORRENCIAS FATAIS = DIVIDE ([QTDE DE OCORRENCIAS FATAIS], [TOTAL DE OCORRENCIAS], 0)

% ocorrências não fatais: medida utilizada para identificar a participação de cada atributo do total. Divide-se a quantidade do atributo pelo total geral.

% OCORRENCIAS NÃO FATAIS = DIVIDE([QTDE DE OCORRENCIAS NAO FATAIS], [TOTAL DE OCORRENCIAS], 0)

% total climático: medida utilizada para identificar a participação de cada atributo do total. Divide-se a quantidade do atributo pelo total geral.

% TOTAL CLIMATICO =

 VAR TOTAL = CALCULATE([TIPO OCORRENCIA],

        ALL(dim\_condicoes\_climaticas)

)

RETURN

DIVIDE([TIPO OCORRENCIA],

        TOTAL,

         0)

% total feriados: medida utilizada para identificar a participação de cada atributo do total. Divide-se a quantidade do atributo pelo total geral.

% TOTAL FERIADOS =

 VAR TOTAL = CALCULATE([QTDE DE OCORRENCIAS FATAIS],

        ALL(dim\_feriados[Feriado])

)

RETURN

DIVIDE([QTDE DE OCORRENCIAS FATAIS],

        TOTAL,

         0)

% total superfície da via: medida utilizada para identificar a participação de cada atributo do total. Divide-se a quantidade do atributo pelo total geral.

% TOTAL SUPERFICIE DA VIA =

 VAR TOTAL = CALCULATE([QTDE DE OCORRENCIAS FATAIS],

        ALL(fact\_acidentes\_fatais\_transito[Superfície da Via])

)

RETURN

DIVIDE([QTDE DE OCORRENCIAS FATAIS],

        TOTAL,

         0)

Texto ocorrência: seleciona o tipo de ocorrência no texto no cabeçalho.

TEXTO - OCORRENCIA = "OCORRÊNCIA(S) " & SELECTEDVALUE (TIPO\_OCORRENCIA[TIPO OCORRENCIA])

Texto sublinhado linha do tempo: usado como subtítulo mostrando o tipo de medida selecionada.

TEXTO - SUBTITULO LINHA DO TEMPO = "OCORRÊNCIA(S) " & SELECTEDVALUE(TIPO\_OCORRENCIA[TIPO OCORRENCIA]) & " - PREVISÃO"

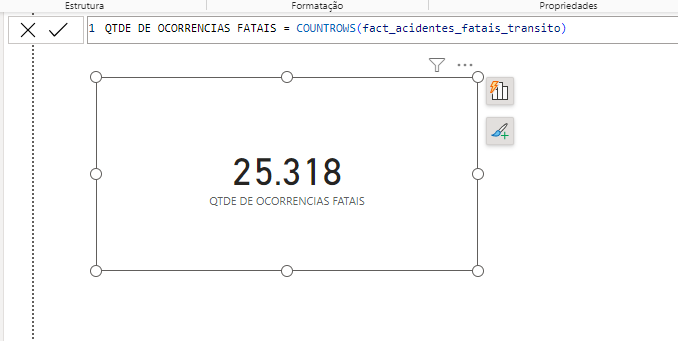
# 5. Registros de Homologação

Teste de homologação dos dados apresentados no dashboard desenvolvido para a análise.

Validação da tabela fato de acidentes fatais.

A figura 1 é a medida DAX que conta a quantidade de linhas, no qual é a quantidade de ocorrências fatais utilizada nas análises, em comparação a figura 1.1.1 onde é apresentada a consulta em linguagem SQL contando a quantidade de linhas da tabela. Na figura 1.1.2 é apresentado a fonte dos dados imputados no banco de dados, contendo a mesma quantidade de linhas das figuras 1 e 1.1.1

Figura 1 – Medida DAX contagem de linhas acidentes fatais



Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.1 – consulta SQL contagem de linhas acidentes fatais

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.1.2 – fonte de dados tabela acidentes fatais

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Validação da tabela fato de acidentes não fatais.

A figura 1.2 é a medida DAX que conta a quantidade de linhas, no qual é a quantidade de ocorrências não fatais utilizada nas análises, em comparação a figura 1.2.1 onde é apresentada a consulta em linguagem SQL contando a quantidade de linhas da tabela de acidentes não fatais no banco de dados. Ambas apresentam a mesma quantidade de linhas. Porém, a tabela fonte, no processo de tratamento foi aplicado filtro de linhas para remover ids duplicados para ter a integridade dos dados, apresentado assim uma quantidade menor de linhas no banco de dados quando comparado a fonte de dados.

Figura 1.2 – Medida DAX contagem de linhas acidentes não fatais

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.2.1 – consulta SQL contagem de linhas acidentes não fatais

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

Figura 1.2.2 – tabela fonte acidentes não fatais

Tela de computador

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

A fonte dos dados da tabela de acidentes não fatais (Figura 1.2.3) foi testada a remoção de duplicadas de forma manual pela função do MS Excel no qual totalizou a mesma quantidade de linhas no banco de dados.

Figura 1.2.3 – tabela fonte acidentes não fatais

Interface gráfica do usuário, Aplicativo, Tabela, Excel

Descrição gerada automaticamente

Fonte: Autoria própria do projeto desenvolvido (2024)

# 

# 5. Conclusões

Ao analisarmos ao longo dos 5 anos é possível ver que houve um aumento de 5,56% de acidentes de trânsito no estado de São Paulo. Em comparação nos anos de 2019 e 2020 observa-se uma diminuição de ocorrências em 9,37% no qual pode-se atribuir as restrições de circulação devido a pandemia, quando comparado ao ano de 2022 contra 2023 houve um aumento de 8,43%, quando começou a normalizar a circulação. OMS classifica coronavírus como pandemia. Gov.br, 11/03/2020. Disponível em: <<https://www.gov.br/pt-br/noticias/saude-e-vigilancia-sanitaria/2020/03/oms-classifica-coronavirus-como-pandemia>>.

Quanto as condições climáticas, no ano de 2023 é possível identificar que 54,4% dos casos ocorreram em condições climáticas documentadas como bom e cerca de 39,3% a luz do dia que ocorre com maior frequência entre sexta-feira e sábado. Nestes dados apontados, podemos concluir que entre outros motivos, o excesso de confiança pode ter causado os acidentes.

Em feriados, o Carnaval atrai a atenção com um total no ano de 2022 de 807 casos, seguido pelo Natal com 618 casos registrados. Entre 2019 a 2023 o feriado de Carnaval se posiciona em primeiro no ranking de acidentes de trânsito.

Quando olhamos para o tipo de via, as vias municipais ocorreram o maior índice de acidentes, sendo de 78,1% no ano de 2021, enquanto rodovias sendo de 17,7% e não informados de 4,2%. Acredita-se que devido a ter um grande fluxo de pessoas em vias municipais à maior incidência de casos. Porém, com 2,7% ocorrendo vítimas fatais enquanto em rodovias 9,6%, que podemos atribuir por ter um limite de velocidade maior do que em vias municipais.

A proposta para este projeto é de conscientização para a população e acesso aos dados de acidentes de trânsito no estado de São Paulo de forma resumida e clara. A fim de que seja possível traçar estratégias de prevenção à acidentes, em pontos específicos e em determinada sazonalidade de forma em que os recursos humanos, materiais e financeiros, do estado ou da iniciativa privada, sejam aplicados com maior eficiência e eficácia no objetivo final de prevenir acidentes e salvar vidas.

Durante o processo de desenvolvimento do projeto foi observado algumas limitações, como: acesso à informação, embora esteja em um site, a informação tratada no trabalho deve ser de interesse público, ou seja, são dados que devem ser divulgados e ter forma mais acessível a fim de conscientizar a população. Outro ponto no qual é possível se trabalhar são os dados categóricos do tipo idade, sexo, renda no intuito de se traçar um perfil que seja mais recorrente, porém, não está documentado nos arquivos.

# 6. Links

Base de dados de acidentes não fatais:

<http://painelderesultados.infosiga.sp.gov.br/bases/acidentes_naofatais.csv>

Base de dados de acidentes fatais:

<http://painelderesultados.infosiga.sp.gov.br/bases/acidentes_fatais.xlsx>

Base de dados datas de feriados:

<https://www.anbima.com.br/feriados/arqs/feriados_nacionais.xls>

Início da pandemia:

<https://www.gov.br/pt-br/noticias/saude-e-vigilancia-sanitaria/2020/03/oms-classifica-coronavirus-como-pandemia>

Fim da pandemia:

<https://www.gov.br/saude/pt-br/assuntos/noticias/2022/abril/ministerio-da-saude-declara-fim-da-emergencia-em-saude-publica-de-importancia-nacional-pela-covid-19>

Dashboard:

<https://app.powerbi.com/view?r=eyJrIjoiMDFlYzgzMTgtMTI2Mi00OTE0LThhYTYtZTc3YmU4OTBhMzg1IiwidCI6IjE0Y2JkNWE3LWVjOTQtNDZiYS1iMzE0LWNjMGZjOTcyYTE2MSIsImMiOjh9&embedImagePlaceholder=true&pageName=ReportSection>

# REFERÊNCIAS

MICROSOFT. **Entenda o esquema em estrela e a importância para o Power BI.** MICROSOFT, 2023. Disponível em: https://learn.microsoft.com/pt-br/power-bi/guidance/star-schema

NERY RODRIGUES MACHADO, Felipe. **Banco de Dados Projeto e Implementação.** São Paulo: Saraiva, 2020.

PODEROSO DE OLIVEIRA, Celso Henrique. **SQL Curso Prático.** São Paul: Novatec Editora 2002.