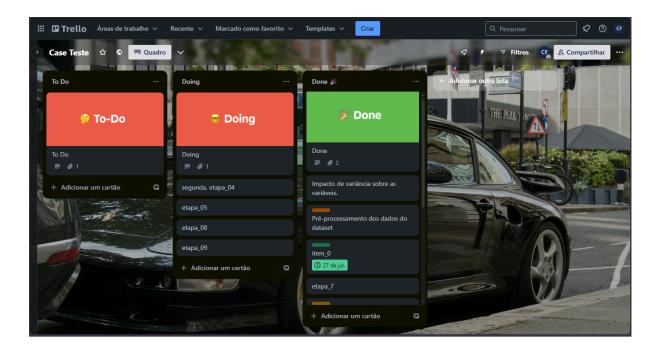
#### Item 0



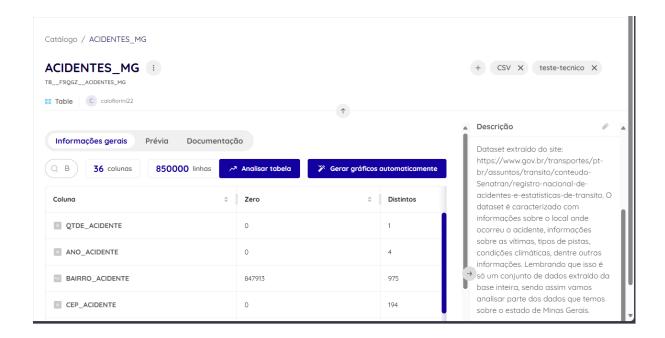
https://trello.com/invite/b/66a4364d00ce4ef66f41fe72/ATTle81a89f216ea57749f277d 9650fca39dC49389B9/case-teste

#### Item 1

#### Dataset extraído do site:

https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-Senatran/registro-na cional-de-acidentes-e-estatisticas-de-transito. O dataset é caracterizado com informações sobre o local onde ocorreu o acidente, informações sobre as vítimas, tipos de pistas, condições climáticas, dentre outras informações. Lembrando que isso é só um conjunto de dados extraído da base inteira, sendo assim vamos analisar parte dos dados que temos sobre o estado de Minas Gerais.

Upei 3 dataset, 1 completo, 2 Com apenas pequenas modificações para ajudar na visualização e o último do google sheet.



Item 2



O dataset está carregado e tem a visualização dos dados nele.

• Item 3

Nome: Dataset de Acidentes em Minas Gerais Descrição:

Esse dataset contém dados extraídos do governo federal, dos anos de 2018 à 2021, onde contém informações sobre os tipos de acidentes e informações sobre o acidente. A ideia do uso deste dataset, é pegar insights sobre os locais onde acontecem mais acidentes e demonstrar isso através de informação simples para que as pessoas possam entender que naquele local devemos ter atenção dobrada, independente dos fatores externos. Uma possível aplicação futura desse tipo de dataset, seria em redes móveis, onde podemos criar formas de alertas para

motoristas estejam atentos por meio de aplicativos do governo ou uma aplicação embarcada, onde podemos utilizar os sensores dos carros para evitar mais acidentes.

Dataset extraído do site:

https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-Senatran/registro-na cional-de-acidentes-e-estatisticas-de-transito.

Esse site contém dados desde 2018 até esse mês de 2024, contém uma atualização constante, e o seu formato é em .csv

O nome e o tipo das colunas são:

uf\_acidente: int
ano\_acidente: int
mes\_acidente: int
dia\_semana: string
fase\_dia: string
tp\_acidente: string
end\_acidente: string
km\_via\_acidente: int
hora\_acidente: time
tp\_rodovia: string
cond\_pista: string
tp\_cruzamento: string
tp\_pavimento: string
tp\_curva: string

lim velocidade: string

tp\_pista: string gtde acidente: int

qtde\_acid\_com\_obitos: int

qtde\_envolvidos: int qtde\_feridosilesos: int

qtde obitos: int

O dataset está desbalanceado, não está normalizado e nem temos validações dos dados

Disponibilizada no mesmo dataset Acidentes MG v1

• Item 4

Não consegui fazer

Item 5

Não consegui fazer

• Item 6

## 1. Identificação dos Processos de Negócio

 Análise de acidentes de trânsito: Este é o principal processo que estamos interessados em analisar.

### 2. Definição do Grão

• **Nível de detalhe**: Cada registro no dataset representa um acidente específico.

## 3. Identificação das Dimensões

Vamos identificar as dimensões a partir do seu dataset:

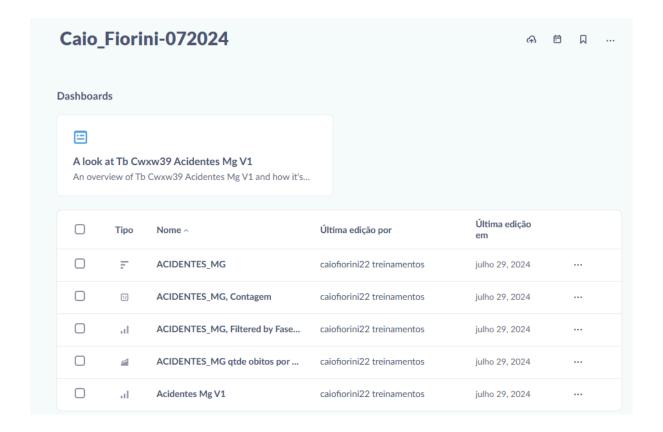
- **Dimensão de Tempo**: ano\_acidente, mes\_acidente, dia\_semana, fase\_dia, hora acidente
- Dimensão de Localização: uf\_acidente, end\_acidente, km\_via\_acidente
- **Dimensão de Tipo de Acidente**: tp\_acidente, tp\_rodovia, cond\_pista, tp\_cruzamento, tp\_pavimento, tp\_curva, lim\_velocidade, tp\_pista

# 4. Identificação dos Fatos

Os fatos são as medidas quantitativas:

 Medidas de Acidente: qtde\_acidente, qtde\_acid\_com\_obitos, qtde envolvidos, qtde feridosilesos, qtde obitos

### Item 7



Item 08

não tive acesso a área de inteligencia

• Item 09

não consegui

Link Github - caioFiorini/CAIO\_FIORINI\_DDF\_TECH\_072024: Repositório dedicado à teste técnico na empresa Dadosfera. (github.com)