

Escola Politécnica
Banco de Dados II
Prof^a. Dr^a. Aline de Campos
Projeto Aplicado de Banco de Dados

Gabriel Kowaleski
Rafael Spumberg

1. Introdução

O Projeto Aplicado de Banco de Dados II tem como objetivo desenvolver um dashboard interativo utilizando a plataforma Power BI para implementar uma solução para inteligência de negócios utilizando uma base de dados. Através da construção de um dashboard interativo, intuitivo e funcional, pretende-se analisar dados relacionados a acidentes de trânsito no Brasil, nos anos de 2017, 2018 e 2019, para apoiar a tomada de decisão em questões relacionadas à segurança viária.

A temática foi escolhida devido à sua relevância social, considerando que os acidentes de trânsito deixando milhares de mortos e feridos por ano, é fundamental compreender quais estados e municípios possuem as maiores taxas de acidentes, como também compreender as possíveis causas dessas ocorrências, considerando o impacto dos acidentes tanto para a saúde pública quanto para a economia. Entender os padrões dos acidentes no período analisado tem como intuito auxiliar possíveis decisões de governos municipais, estaduais e federal a priorizar o investimento e qualificações de certas rodovias e tecnologias em detrimento de outros, aplicando da melhor forma o erário público.

2. Dataset

Os dados utilizados no projeto foram obtidos a partir do dataset *Federal Highway 2017-2019 Car Crashes*, disponível na plataforma Kaggle¹, uma das principais referências para compartilhamento de datasets. A fonte original dos dados é o Departamento Nacional de Trânsito (DENATRAN) do Brasil. Os dados foram disponibilizados em formato CSV, contendo registros de acidentes

¹ Disponível em: <https://www.kaggle.com/datasets/exatasmente/federal-highway-20172019-car-crashes?select=datatran2019.csv> Acesso em: 04 de dezembro de 2024.

ocorridos nas rodovias federais brasileiras entre os anos de 2017 até 2019. O conjunto é relevante para a análise de segurança viária devido ao seu alto nível de detalhamento e à abrangência nacional, permitindo um estudo consistente sobre a evolução dos acidentes ao longo do tempo e sua distribuição geográfica.

O dataset contém centenas de milhares de registros, representando cada ocorrência de acidente.

- Data: Data exata do acidente.
- Horário: Hora aproximada do ocorrido.
- BR: Código da rodovia onde o acidente ocorreu.
- UF: Unidade federativa em que o acidente aconteceu.
- Causa Principal: Motivo principal do acidente.
- Tipo de Acidente: Classificação do acidente (colisão, capotamento, etc.).
- Mortes: Número de mortes associadas ao acidente.
- Feridos: Número de feridos envolvidos.
- Entre outras variáveis que fornecem informações adicionais sobre o contexto do acidente.

Mediante análise do *dataset*, foi observado que não havia a necessidade de realizar a limpeza dos dados, como há dados repetidos, como número de mortos, seu houve ou não feridos, a data do acidente, a limpeza desses dados poderia colidir com essas informações. Para melhor visualizar algumas informações no *dashboard*, realizamos a inserção de três novas colunas: **Nome_UF_Completo**, **Ano** e **Mes**, como é possível observar na **Figura 1**, **Figura 2** e **Figura 3**.

Figura 1 – Tabela Nome_UF_Completo

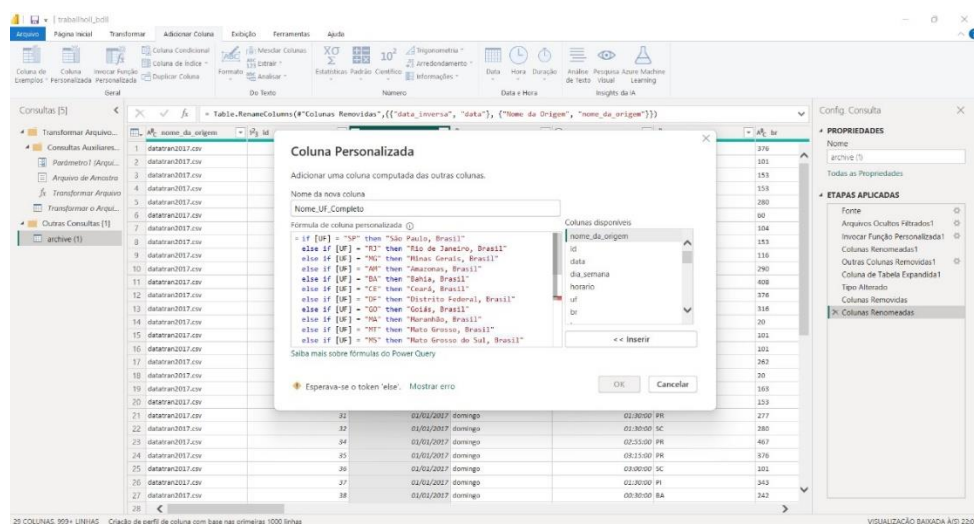


Figura 2 – Tabela Mes

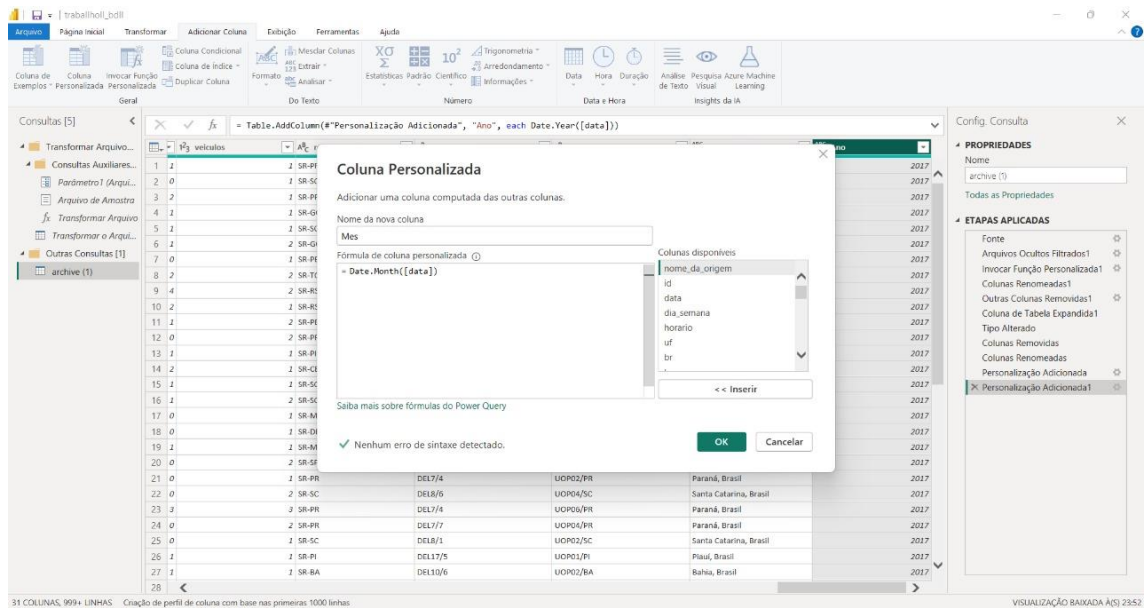
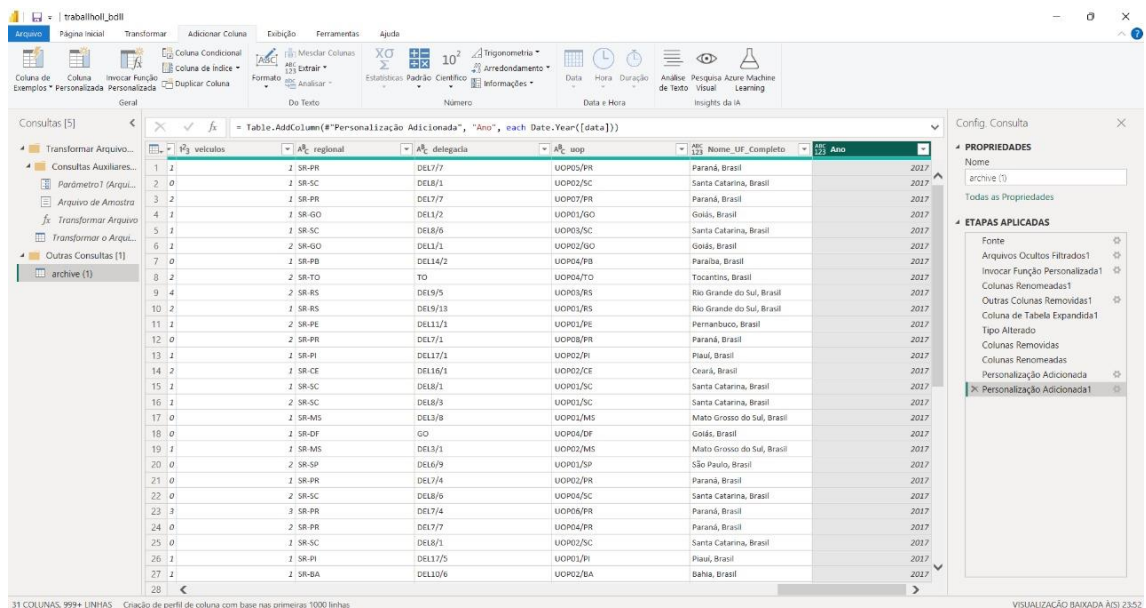


Figura 3 – Tabela Ano



A tabela **Nome_UF_Completo** foi criada, mesmo havendo no *dataset* uma coluna UF, justamente pela vontade do grupo de inserir um mapa no dashboar para facilitar a visualização dos dados e seus filtros, porém o Power BI interpretava algumas siglas das Unidades Federativas como sendo de outros países, exemplos: SC – South Caroline. Utilizamos o seguinte código para filtrar as Unidades Federativas do Brasil:

```
if [UF] = "AC" then "Acre, Brasil"
else if [UF] = "AL" then "Alagoas, Brasil"
else if [UF] = "AM" then "Amazonas, Brasil"
else if [UF] = "AP" then "Amapá, Brasil"
else if [UF] = "BA" then "Bahia, Brasil"
else if [UF] = "CE" then "Ceará, Brasil"
else if [UF] = "DF" then "Distrito Federal, Brasil"
else if [UF] = "ES" then "Espírito Santo, Brasil"
else if [UF] = "GO" then "Goiás, Brasil"
else if [UF] = "MA" then "Maranhão, Brasil"
else if [UF] = "MG" then "Minas Gerais, Brasil"
else if [UF] = "MS" then "Mato Grosso do Sul, Brasil"
else if [UF] = "MT" then "Mato Grosso, Brasil"
else if [UF] = "PA" then "Pará, Brasil"
else if [UF] = "PB" then "Paraíba, Brasil"
else if [UF] = "PE" then "Pernambuco, Brasil"
else if [UF] = "PI" then "Piauí, Brasil"
else if [UF] = "PR" then "Paraná, Brasil"
else if [UF] = "RJ" then "Rio de Janeiro, Brasil"
else if [UF] = "RN" then "Rio Grande do Norte, Brasil"
else if [UF] = "RO" then "Rondônia, Brasil"
else if [UF] = "RR" then "Roraima, Brasil"
else if [UF] = "RS" then "Rio Grande do Sul, Brasil"
else if [UF] = "SC" then "Santa Catarina, Brasil"
else if [UF] = "SE" then "Sergipe, Brasil"
else if [UF] = "SP" then "São Paulo, Brasil"
else if [UF] = "TO" then "Tocantins, Brasil"
else "UF desconhecida"
```

Já para as tabelas **Ano** e **Mes**, utilizamos o mesmo código, alterando somente algumas variáveis, separando as informações da tabela **data**. O código utilizado foi o seguinte:

```
= Date.Month([data])
= Date.Year([data])
```

3. Questão de Negócio

O *dashboard* que o grupo construiu possui como objetivo abordar e responder questões críticas relacionadas aos acidentes de trânsito no Brasil, fornecendo informações relevantes para suportar à tomada de decisões de gestores e políticos em relação e estratégias a serem implementadas. As principais questões de negócio que o nosso trabalho pretende responder são:

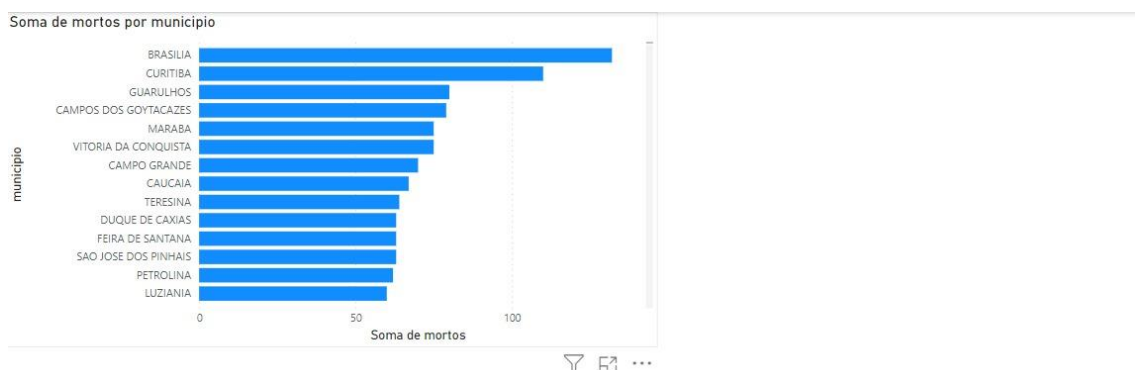
- Identificar os meses com maior número de acidentes fatais.
- Entender a relação entre o clima (como chuva ou neblina) e a ocorrência de acidentes.
- Entender a relação entre o clima (como chuva ou neblina) e a ocorrência de acidentes.
- Identificar as rodovias federais com maior número de acidentes fatais.
- Analisar a distribuição de acidentes por estado ou região.
- Verificar tendências temporais no número de mortos e feridos.

Como problemas ou oportunidades a serem exploradas, destacamos a possibilidade de identificar infraestrutura precária em determinadas rodovias, falta de políticas adaptativas a condições meteorológicas. Mas também a possibilidade de desenvolver campanhas preventivas e melhorar a alocação de recursos públicas para a criação, desenvolvimento e melhoramento das infraestruturas em questão. Essas questões visam ajudar a criar uma visão mais abrangente do problema das rodovias públicas no Brasil, permitindo que autoridades públicas tomem decisões baseadas em dados e evidências, visando salvar vidas e otimizar recursos.

4. Construção da aplicação

Para a construção do *dashboard*, pensamos, primeiramente, na criação de duas páginas, para não poluir a primeira página com muitas informações. O primeiro gráfico a ser inserido foi um gráfico de barras, sendo município X soma de mortes, como é possível observar na **Figura 4**, porém ele logo foi descartado devido a mudança do foco do projeto de município para as Unidades Federativas.

Figura 4 – Primeiro Gráfico



Após partirmos para a delimitação dos estados, utilizamos um Treemap para destacar a Unidade Federativa com maior número de fatalidades em acidentes, ressaltando também os tipos de acidentes com maior número de episódios. Também utilizamos um gráfico em formato de pizza, para visualizar as fatalidades em uma totalidade, nos três anos analisados, percebendo as maiores causas dos acidentes. Podemos observar ambos os gráficos na **Figura 5** em sua concepção inicial e na **Figura 6** em sua forma final.

Figura 5 – Treemap e Pizza

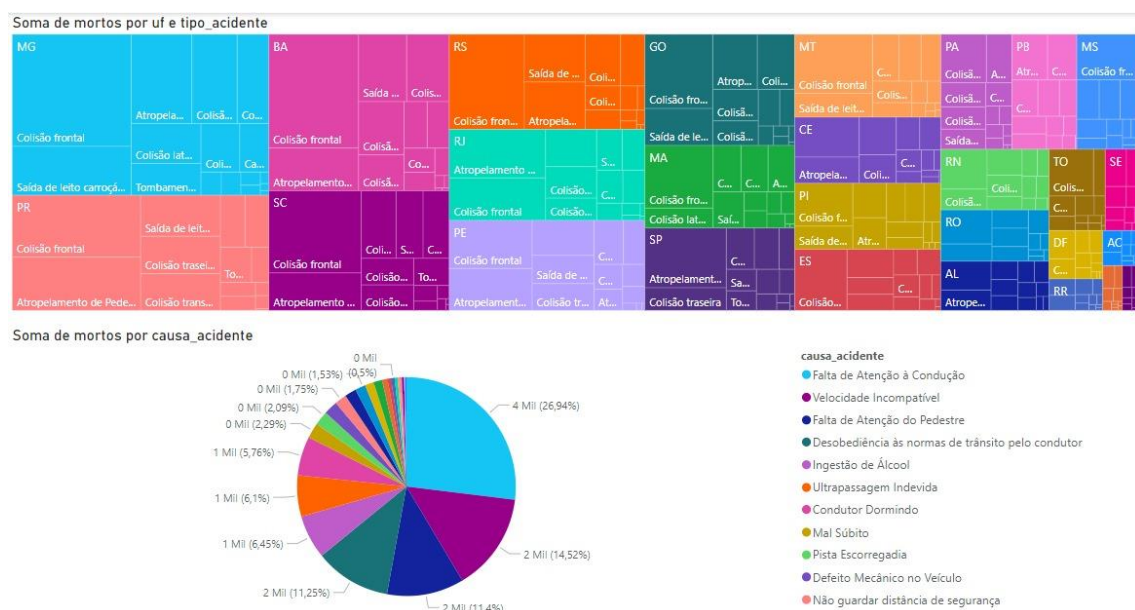
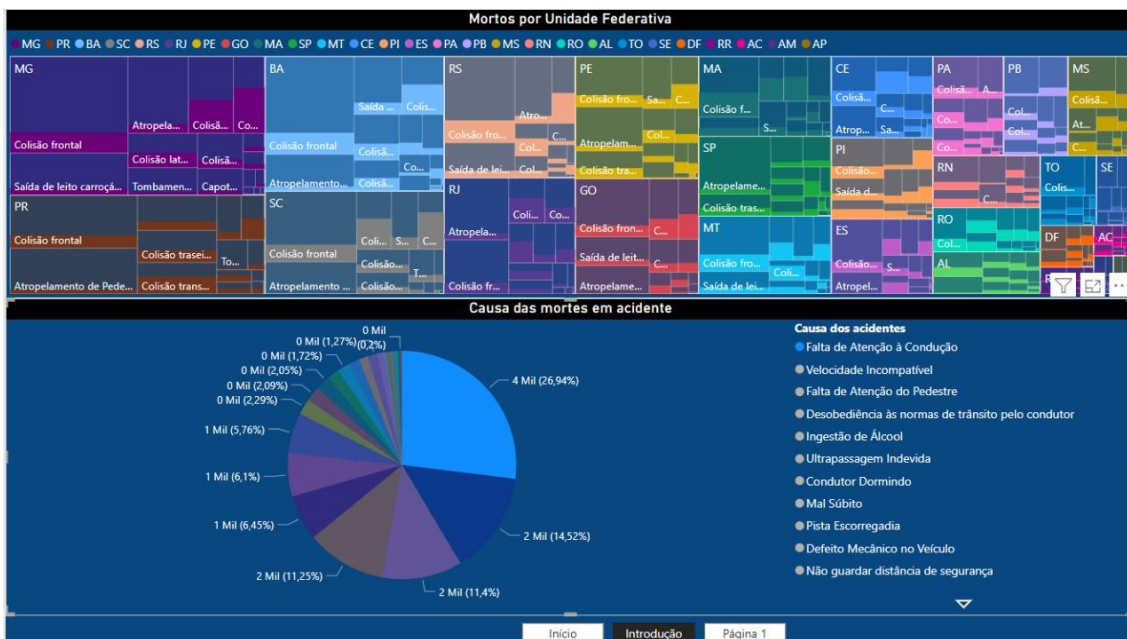


Figura 6 – Treemap e Pizza finalizados



Pensando na primeira página do *dashboard*, pensamos em incluir alguns gráficos para ressaltar as informações mais relevantes. Pensamos em colocar um mapa do Brasil, com a separação de todas as Unidades Federativas, assim o usuário pode clicar em cada uma delas para obter as informações referentes ao estado em questão. Há também um gráfico sobre o total de mortes por anos, no qual podemos observar a tendencia decrescente de mortes no período analisado. Inserimos também um gráfico pizza, o qual revela o total de fatalidades levando em consideração as condições climáticas. Há mais dois gráficos de barra, o primeiro identifica o total de mortos em cada rodovia federal e o segundo o total de fatalidades por dia de semana. Por último, apresentamos dois cards: o primeiro com o total de mortos e o segundo com o total de feridos. Podemos observar a evolução do *dashboard*, na **Figura 7**, **Figura 8** e **Figura 9**.

Figura 7 – Página inicial

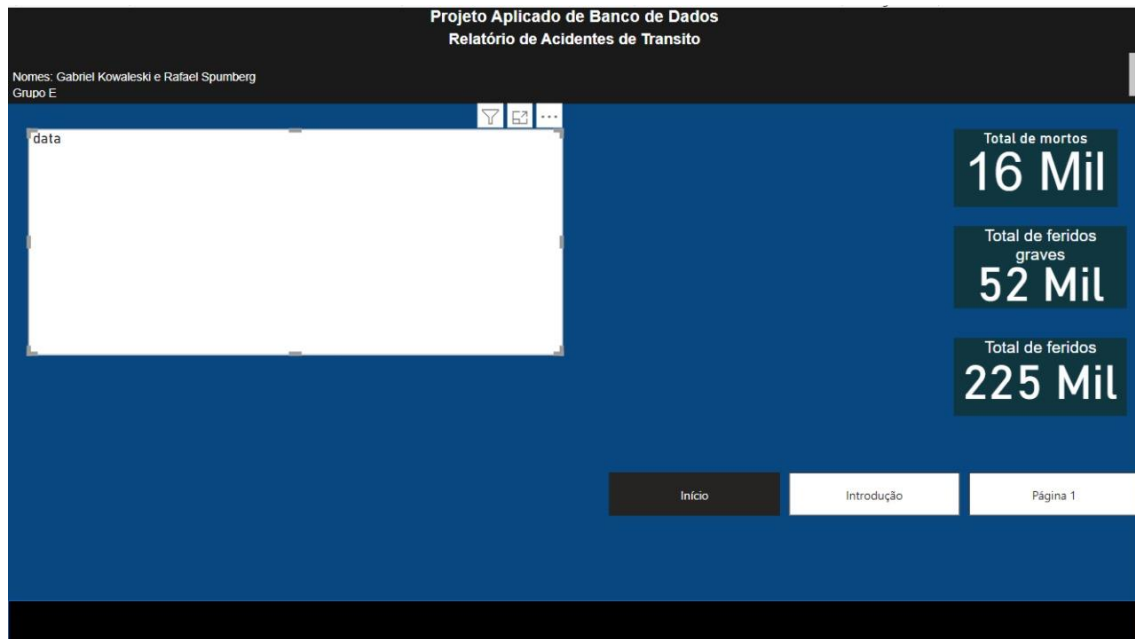


Figura 8 – Primeira formulação da página inicial

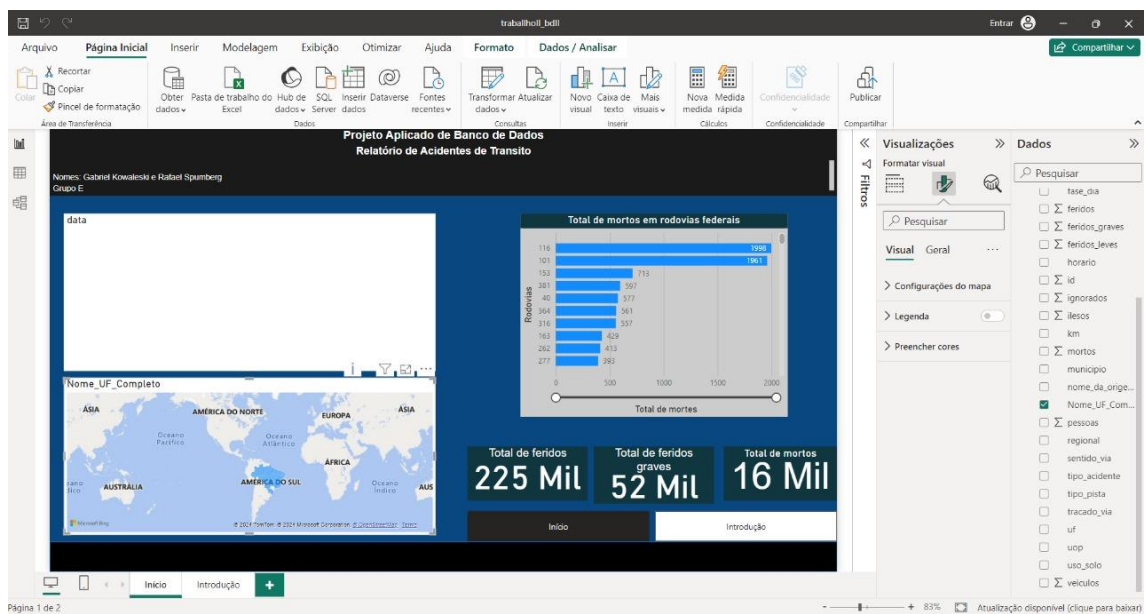
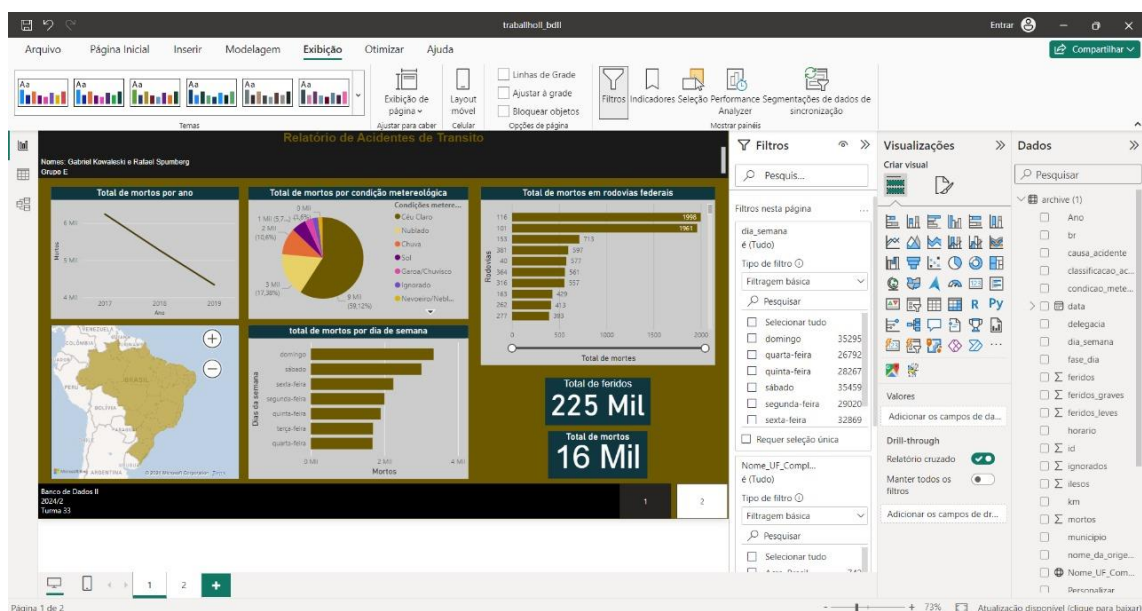


Figura 9 -



5. Conclusão

Após a finalização do *dashboard*, podemos constatar algumas questões sobre os dados utilizados. A mais relevante é que houve uma diminuição significativa de mortes no período analisado, sendo essa redução de quase 2 mil mortes nos três anos analisados. Um ponto que vale o destaque, mas que não surpreende é que as duas rodovias com maior número de morte são as BRs 116 e 101, porém devido a sua extensão no território nacional é compreensível serem as com maior incidência de acidentes fatais.

Cabe destacar mais três pontos. O primeiro é que os dias com maior incidência de acidentes fatais são sábado e domingo. O segundo é que o estado com maior número de mortos é Minas Gerais, com a maior causa sendo colisão frontal, seguido pelo Paraná, com causa principal também sendo colisão frontal. O Rio Grande do Sul aparece quem quinto lugar, possuindo também como causa principal colisão frontal. O último ponto, porém, um dos mais relevantes, é que a causa principal para os acidentes é a falta de atenção à direção, que incluem ações como excesso de velocidade, ultrapassagens perigosas e desatenção ao volante. Esses fatores estão diretamente associados às colisões frontais, que foram identificadas como a principal causa de acidentes fatais nos estados analisados.

Esses dados evidenciam a necessidade de campanhas educativas mais incisivas e contínuas para conscientizar os motoristas sobre práticas seguras no trânsito. Além disso, reforçam a importância de fiscalizações mais rigorosas, especialmente em rodovias de alta incidência, como as BRs 116 e 101, e nos estados com maiores números de acidentes fatais, como Minas Gerais e Paraná.

Por fim, a redução significativa no número de mortes ao longo dos anos analisados sugere que algumas políticas de segurança no trânsito têm surtido efeito, mas ainda há muito trabalho a ser feito para reduzir ainda mais esses números e salvar vidas. A análise detalhada dessas tendências fornece subsídios importantes para a criação de estratégias mais direcionadas e eficazes para a melhoria da segurança viária no Brasil.