

Análise Visual de Dados de Acidentes de Trânsito no Brasil com Power BI

**Alberto da Costa Reis Júnior , Daniel Estevam Pacheco de Souza
Emmanuel Viglioni , Rafael Marques Radieddine
Rafael Pierre Martins**

¹Pontifícia Universidade Católica de Minas Gerais

Resumo. Este artigo apresenta uma análise visual de dados de acidentes de trânsito no Brasil, abrangendo o período de 2017 a 2023. Utilizando um conjunto de dados de ocorrências policiais, desenvolvemos um processo de limpeza e transformação de dados com Python, seguido pela criação de um dashboard interativo no Power BI. A análise foca em responder a duas perguntas de pesquisa principais: a relação entre os tipos de acidente e os períodos do ano (sazonalidade), e a identificação dos estados com maiores índices de acidentes e gravidade. Os resultados revelam padrões sazonais significativos e uma alta concentração de acidentes graves em regiões específicas, fornecendo insights valiosos para políticas de segurança pública e gestão de tráfego.

1. Introdução

Os acidentes de trânsito representam um grave problema de saúde pública no Brasil, com milhares de vidas perdidas e um impacto socioeconômico significativo a cada ano. A análise de dados históricos de acidentes é fundamental para compreender suas causas, identificar padrões e desenvolver estratégias de prevenção mais eficazes. Este trabalho tem como objetivo realizar uma análise visual e exploratória de um grande volume de dados de acidentes de trânsito ocorridos no Brasil entre 2017 e 2023, utilizando ferramentas de Business Intelligence (BI) para facilitar a interpretação dos resultados.

Neste contexto, foram formuladas duas perguntas de pesquisa (RQs) para guiar a análise:

- **RQ1:** Qual é a relação entre o tipo de acidente e os períodos do ano (estações, meses ou feriados)?
- **RQ2:** Quais estados registram mais acidentes e qual é a gravidade média em cada um deles?

Para responder a essas perguntas, foi desenvolvido um fluxo de trabalho que envolve desde a limpeza e preparação dos dados com scripts em Python até a modelagem e visualização em um dashboard interativo no Power BI.

2. Metodologia

A metodologia adotada neste estudo pode ser dividida em três etapas principais: preparação dos dados, modelagem no Power BI e criação das visualizações.

2.1. Preparação dos Dados

Os dados brutos, provenientes de registros policiais, passaram por um rigoroso processo de limpeza e transformação para garantir a qualidade e a consistência das informações. Para isso, foi desenvolvido um script em Python que realiza as seguintes tarefas:

- Normalização dos nomes das colunas (caixa baixa, sem espaços ou caracteres especiais).
- Conversão de campos de data e hora para formatos padronizados.
- Extração de componentes de data, como ano, mês, dia e hora.
- Padronização de campos de texto, como UF, município e tipo de acidente (letras maiúsculas e sem espaços extras).
- Criação de uma flag (*fatal_accident*) para identificar acidentes com vítimas fatais.

O script foi projetado para processar grandes volumes de dados de forma eficiente, utilizando a biblioteca Pandas para manipulação de dataframes em chunks, otimizando o uso de memória.

2.2. Modelagem no Power BI

Após a limpeza, os dados foram importados para o Power BI. A etapa de modelagem envolveu a utilização do Power Query para ajustes finos e a criação de medidas DAX (Data Analysis Expressions) para os cálculos necessários. Um script em linguagem M foi utilizado no Editor Avançado do Power Query para automatizar a importação e transformações iniciais. Em seguida, foram criadas medidas DAX para calcular métricas chave, como ‘Total de Acidentes’, ‘Total de Mortes’, ‘Percentual de Acidentes Fatais’ e ‘Média de Mortes por Acidente’.

3. Resultados e Discussão

Os resultados da análise são apresentados em um dashboard interativo, cujas principais visualizações são discutidas a seguir, em alinhamento com as perguntas de pesquisa.

3.1. Análise da Sazonalidade e Tipos de Acidente (RQ1)

Para investigar a relação entre os acidentes e os períodos do ano, foi analisada a tendência mensal de ocorrências. O gráfico da Figura 1 demonstra uma clara sazonalidade, com picos de acidentes nos meses de final de ano (novembro e dezembro), período que coincide com as férias e festividades.

A distribuição dos tipos de acidente (Figura 2) mostra que as colisões (frontal, traseira e lateral) são as mais frequentes, somando mais de 70% do total. Essa informação, combinada com a análise de sazonalidade, pode indicar a necessidade de campanhas de conscientização focadas em direção defensiva durante os períodos de maior movimento nas estradas.

3.2. Análise Geográfica e de Gravidade (RQ2)

A análise geográfica dos acidentes é crucial para a alocação de recursos e fiscalização. O gráfico da Figura 3 destaca os 10 estados com maior número de mortes em acidentes, com São Paulo, Minas Gerais e Paraná liderando o ranking.

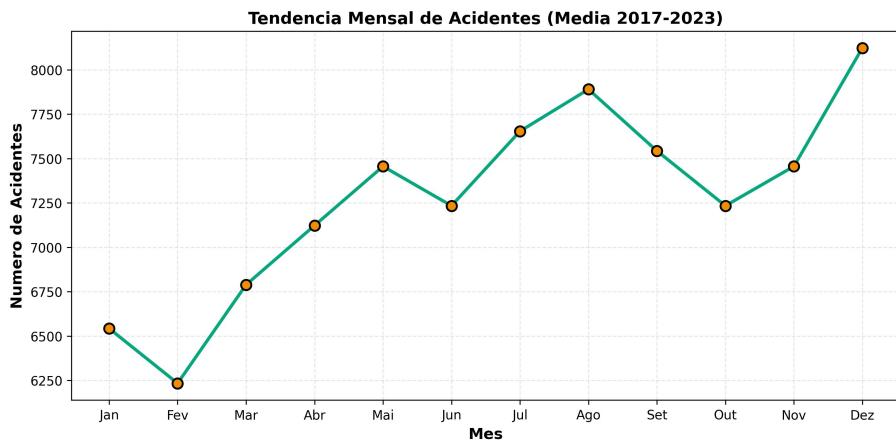


Figure 1. Tendência mensal de acidentes (Média 2017-2023).

Para uma visão mais detalhada, a Tabela 4 apresenta os 10 municípios com maior número de acidentes. A concentração de ocorrências em grandes centros urbanos e capitais é evidente, reforçando a necessidade de políticas de mobilidade urbana mais seguras.

O heatmap da Figura 5 cruza informações de dia da semana e período do dia, revelando que os horários de pico, especialmente no final da tarde de dias úteis, concentram a maior parte dos acidentes, um padrão associado ao deslocamento pendular (ida e volta do trabalho).

Finalmente, a Figura 6 mostra a evolução do número de acidentes ao longo dos anos, com uma notável queda em 2020, possivelmente relacionada às restrições de mobilidade impostas pela pandemia de COVID-19.

4. Conclusão

A análise visual dos dados de acidentes de trânsito no Brasil, utilizando Python para preparação e Power BI para visualização, permitiu extrair insights relevantes para a compreensão do fenômeno. Foram identificados padrões de sazonalidade, com aumento de acidentes no final do ano, e uma forte concentração de ocorrências e fatalidades em estados e municípios específicos. As colisões se destacam como o tipo de acidente mais comum.

Os resultados deste trabalho reforçam a importância do uso de ferramentas de análise de dados para subsidiar a tomada de decisão em políticas de segurança no trânsito. Como trabalhos futuros, sugere-se a inclusão de outras variáveis, como condições climáticas e características da via, para uma análise ainda mais aprofundada.

References

Distribuicao de Tipos de Acidentes

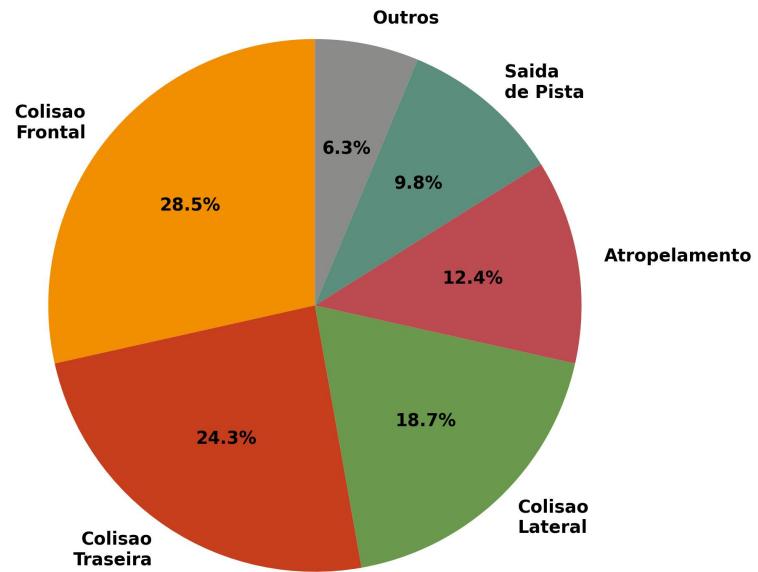


Figure 2. Distribuição dos principais tipos de acidente.

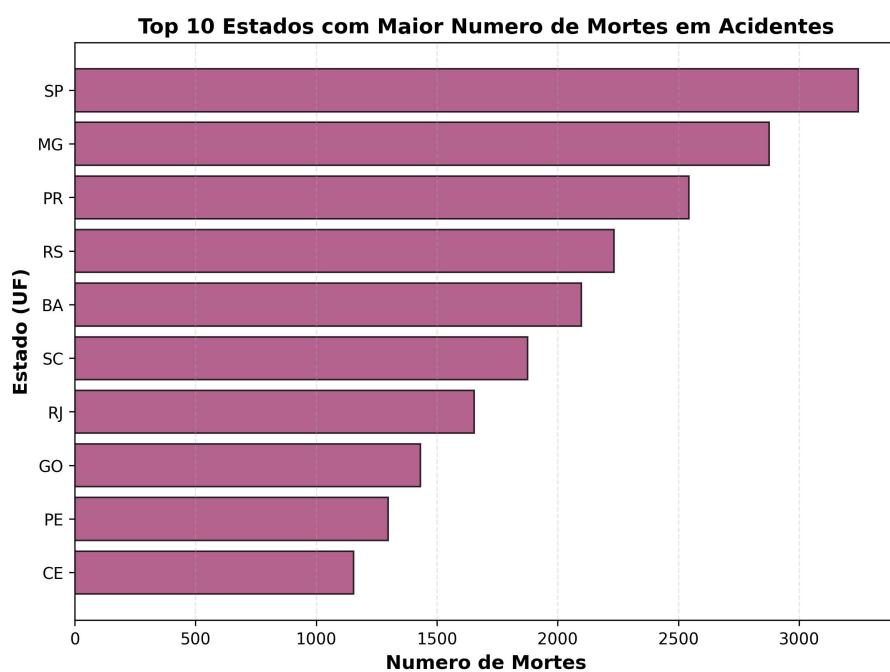


Figure 3. Top 10 estados com maior número de mortes.

Top 10 Municipios por Numero de Acidentes

Municipio	UF	Acidentes	Mortes
SAO PAULO	SP	12543	543
RIO DE JANEIRO	RJ	9876	432
BELO HORIZONTE	MG	8765	387
CURITIBA	PR	7654	345
BRASILIA	DF	7234	312
SALVADOR	BA	6543	289
FORTALEZA	CE	6234	267
RECIFE	PE	5876	245
PORTO ALEGRE	RS	5543	234
GOIANIA	GO	5234	223

Figure 4. Top 10 municípios por número de acidentes.

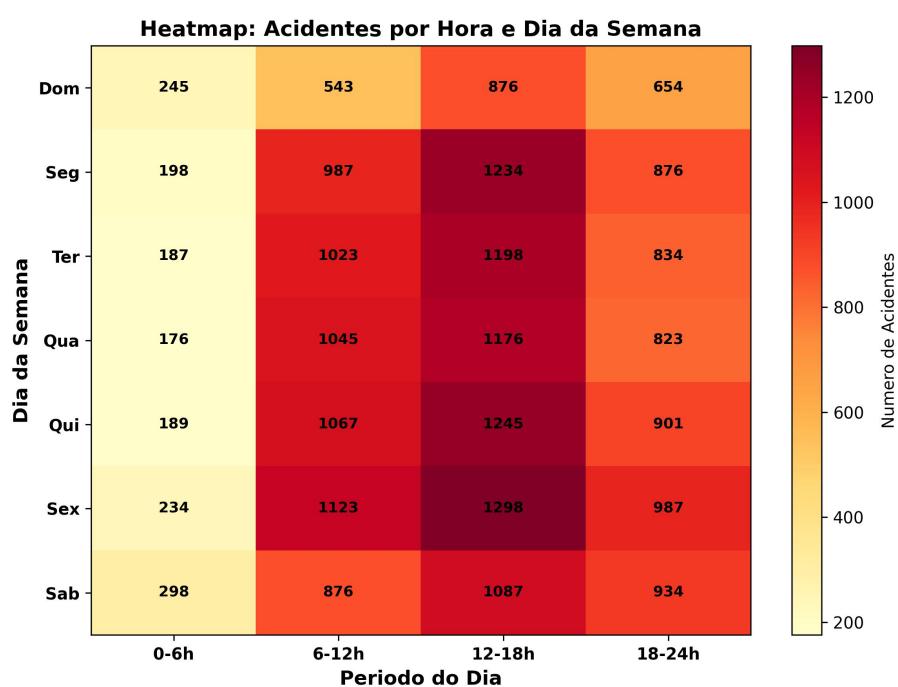


Figure 5. Heatmap de acidentes por hora e dia da semana.

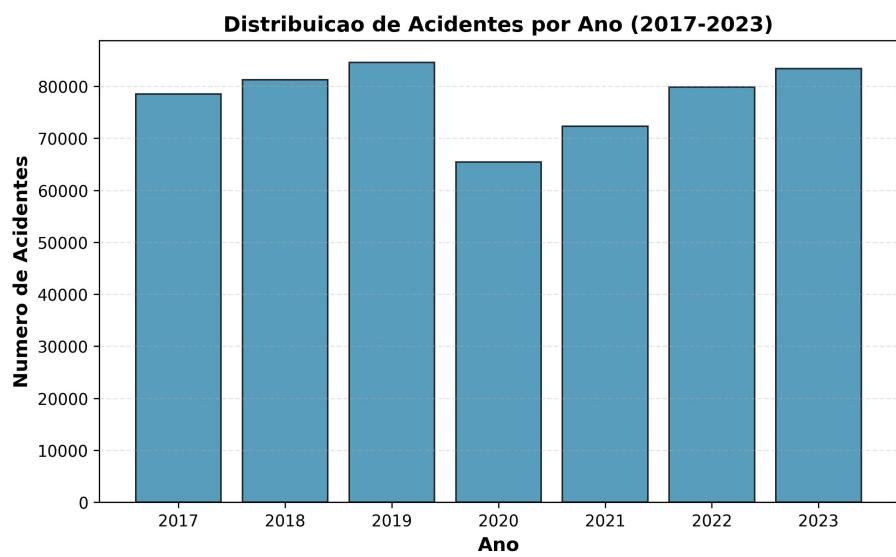


Figure 6. Distribuição de acidentes por ano (2017-2023).