

Análise Geoespacial das Ocorrências de Acidentes nas Rodovias Federais do Sudeste Aplicando Técnicas de Business Intelligence

Amanda M. Colácio¹, Camila Luiza R. Reis¹, Leonardo Vilela Cardoso¹

¹Bacharelado em Sistemas de Informação

Instituto de Ciências Exatas e Informática (ICEI)

Pontífica Universidade Católica de Minas Gerais (PUC MINAS) - Contagem, MG - Brazil

{acolacio, camila.luiza, lvcardoso}@sga.pucminas.br

Abstract. *The federal road network plays a crucial role in the integration of regions, movement of goods and flow of people across the country. However, the study focuses on the Southeast region, which has a high incidence of traffic accidents and is a challenge for road safety. Therefore, the objective of the study is to carry out a detailed geospatial analysis of accident occurrences on federal highways in states in the southeast region, using Business Intelligence techniques. The Federal Highway Police (PRF) public accident database was used from January 2017 to September 2023, carrying out the development from data collection to the creation of visualizations on a dashboard. Among the main results, the analysis revealed that São Paulo, specifically on BR-116 at km 219, has the highest incidence of traffic accidents. However, Minas Gerais is the most dangerous state, with BR-40 having the highest proportion of fatalities and BR-381 recording the highest number of deaths. The study concludes that the data analyzed can be used to inform strategies for preventing and mitigating traffic accidents in the future.*

Keywords: *Geospatial Analysis. Accident Occurrences. Federal Highways. Southeast region. Business Intelligence. Dashboard*

Resumo. *A malha rodoviária federal desempenha um papel crucial na integração de regiões, movimentação de mercadorias e fluxo de pessoas por todo o país. Contudo, o presente estudo tem foco na região do Sudeste, a qual apresenta alta incidência de acidentes de trânsito, sendo um desafio para a segurança viária. Portanto, o objetivo do estudo é realizar uma análise geoespacial detalhada das ocorrências de acidentes nas rodovias federais dos estados da região sudeste, utilizando técnicas de Business Intelligence. Utiliza-se a base de dados pública de acidentes da Polícia Rodoviária Federal (PRF) de janeiro de 2017 a setembro de 2023, envolvendo etapas a partir da coleta de dados até a criação de visualizações em um dashboard. Dentre os principais resultados, a análise revelou que São Paulo, especificamente na BR-116 no km 219, tem a maior incidência de acidentes de trânsito. No entanto, Minas Gerais é o estado mais perigoso, com a BR-40 apresentando a maior proporção de vítimas fatais e a BR-381 registrando o maior número de óbitos. O estudo conclui que os dados analisados podem ser utilizados para informar estratégias de prevenção e mitigação de acidentes de trânsito no futuro.*

Palavras-chaves: *Análise Geoespacial. Ocorrências de Acidentes. Rodovias Federais. Região Sudeste. Business Intelligence. Dashboard*

1. Introdução

No contexto nacional, a malha rodoviária federal desempenha um papel fundamental na integração de regiões, na movimentação de mercadorias e no fluxo de pessoas por todo o país, consequentemente, de acordo com os dados da pesquisa realizada pela Confederação Nacional do Transporte (CNT), apresentadas no CNT 2022a, a malha rodoviária federal em 2021 teve o total de 120.939,80 km. Desta extensão, 12,4% é pavimentada, totalizando 213.452,0 km. 9,1% é planejada, somando 157.309,0 km. O restante, 78,5%, é não pavimentada, totalizando 1.349.938,0 km.

Este estudo se concentra especificamente na região Sudeste do Brasil, notadamente reconhecida pela CNT 2021 como uma área que apresenta elevada incidência de acidentes com vítimas. Além disso, de acordo com CNT 2018, o estado de Minas Gerais foi líder nesse cenário, com os mais altos índices de fatalidades em rodovias federais, bem como à frente do ranking de custos com acidentes. Portanto, o objetivo do estudo é examinar de maneira abrangente diversos aspectos relacionados aos acidentes, com ênfase na aplicação de técnicas de *Business Intelligence* (BI) através da base de dados pública disponibilizada pela Polícia Rodoviária Federal.

Este trabalho tem como objetivo geral realizar uma análise geoespacial abrangente das ocorrências de acidentes nas rodovias federais dos estados da região sudeste, utilizando técnicas de Business Intelligence, fornecendo *insights* significativos e estratégicos para aprimorar a segurança viária, por meio de um dashboard interativo utilizando a ferramenta *Power BI*. Além disso, como objetivos específicos, o estudo busca analisar áreas de maior incidência de acidentes, investigar padrões temporais, classificar e analisar categorias de acidentes mais proeminentes, avaliar a severidade de acordo com vítimas, modelar uma base dimensional com *ETL* (*extract, transform, load*) utilizando as bases disponibilizadas em conjunto com o Power BI e, por fim, realizar a criação de um dashboard interativo para visualização e análise.

Diante desse cenário, o presente estudo busca não apenas compreender os fatores que contribuem para a alta incidência de acidentes, mas também oferecer soluções práticas e estratégicas. Por essa razão, considera-se a relevância de empregar o Business Intelligence, juntamente com a análise geoespacial, visando a criação de um dashboard dinâmico para visualização e interpretação dos dados em busca de disponibilizar uma perspectiva crítica para análise, permitindo a identificação de áreas de maior incidência e outros fatores associados aos acidentes.

O estudo é organizado em seis seções distintas. A primeira seção apresenta uma visão geral, abordando o contexto, o problema, objetivo geral e os específicos, justificativa e como é dividido. Na segunda, é apresentado o referencial teórico, explorando conceitos das técnicas e ferramentas de BI utilizadas, bem como a modelagem de dados. Na terceira seção, são discutidos os trabalhos relacionados, destacando aspectos relevantes que possam contribuir com o desenvolvimento do estudo. Na quarta, é falado sobre a metodologia do trabalho, fornecendo clareza sobre os passos e abordagens utilizados. Na quinta, os resultados são apresentados e explicados através do *dashboard* finalizado. Por fim, na sexta seção, observa-se as conclusões recapitulando o que foi desenvolvido ao longo do estudo e sobre os trabalhos futuros. Em seguida, as referências utilizadas para fundamentar a pesquisa.



Figura 1. Os Pilares de Business Intelligence

2. Referencial Teórico

Nesta seção são apresentados os principais conceitos e técnicas utilizados para a realização deste trabalho, bem como as ferramentas utilizadas para construir a proposta de solução do problema apresentado.

2.1. Business Intelligence

Segundo Barbieri 2011, o conceito de Business Intelligence pode ser compreendido de maneira abrangente como o processo de utilizar diversas fontes de informações para desenvolver estratégias competitivas nos negócios de uma empresa. Isso envolve a utilização de estruturas de dados, como bancos de dados tradicionais, data warehouses e data marts, criados com o objetivo de organizar e tratar informações de maneira relacional e dimensional. Além disso, Business Intelligence é um termo que abrange diversas traduções, como inteligência de negócios, inteligência competitiva ou inteligência empresarial. Fundamentalmente, refere-se a uma abordagem destinada a auxiliar as organizações na tomada de decisões informadas, utilizando dados e informações coletadas por meio de diferentes sistemas de informação.

Os pilares de BI consistem na coleta de dados, armazenamento (Data Warehouse), processamento e transformação dos dados (ETL), visualização e relatórios, tendo como objetivo final a tomada de decisão, conforme apresentado na Figura 1.

Nesse contexto, destaca-se a relevância do Power BI. Conforme um artigo disponibilizado por Mihart et al. 2023, o Power BI é um conjunto abrangente de serviços, aplicativos e conectores que têm como objetivo transformar dados diversos em informações visualmente envolventes e interativas. Funcionando como uma ferramenta de relatórios dentro do campo de BI, fundamentada em modelos de dados, o Power BI permite a conexão com várias fontes de dados, incluindo planilhas do Excel e data warehouses na nuvem e em locais. Isso viabiliza a visualização e a descoberta de insights relevantes, promovendo facilidade no compartilhamento com colegas e colaboradores. O Power BI é, portanto, uma materialização prática do conceito de BI, oferecendo um conjunto abrangente de ferramentas que possibilitam a realização de análises de dados de maneira

completa e abrangente.

2.2. Data Warehouse

Segundo Kimball 1998, um Data Warehouse (DW) é visto como um conjunto de Data Marts que representam áreas específicas de análise de forma intuitiva, tornando mais fácil a filtragem e agregação de dados gerenciais para fins analíticos. Por outro lado, Inmon 1997 define o DW como uma base de dados orientada pelas seguintes características:

- Assunto: Um DW se concentra em um domínio de assunto específico, fornecendo informações detalhadas sobre, em vez das operações atuais de uma empresa. Tipos de assunto: clientes, vendas, fornecedores, etc.
- Integrado: O processo de armazenamento de dados integra dados de várias fontes, como mainframe, bancos de dados relacionais, arquivos simples, etc.
- Tempo variável: Os dados mantidos em um data warehouse são acessados com um intervalo de tempo específico e oferecem informações a partir de um ponto de vista histórico.
- Não volátil: Os dados no data warehouse são somente leitura, não podendo ser removidos ou alterados.

Além disso, Turban et al. 2008 descreve que um Data Warehouse é também um repositório de dados atuais e históricos de possível interesse aos gerentes de toda a organização, bem como um conjunto de dados produzidos para oferecer suporte à tomada de decisão sobre um determinado assunto.

Nesse sentido, os conceitos citados visam facilitar a análise de dados empresariais, mas abordam a estrutura e a abrangência do DW de maneiras distintas.

2.3. Modelagem *Star Schema*

O data warehouse se baseia no conceito de modelagem dimensional e uma das formas para desenvolver é com o modelo Star Schema, também conhecido como esquema estrela. Conforme Turban et al. 2008:

A modelagem dimensional é um sistema baseado em recuperação que suporta acessos com alto volume de consultas. O esquema estrela é o meio pelo qual a modelagem dimensional é implementada. Ele contém uma tabela de fatos central cercada por diversas tabelas de dimensão. A tabela de fatos contém uma grande quantidade de linhas que correspondem aos negócios ou fatos observados.

O esquema de estrela é eficiente para armazenar dados, manter o histórico e atualizar os dados, reduzindo a duplicação de definições de negócios repetitivas. Isso torna rápido para agregar e filtrar dados no data warehouse. Além disso, o esquema de estrela é simples de entender e implementar, tornando fácil para os usuários finais encontrar os dados de que precisam

2.4. Dashboard

De acordo com Turban et al. 2008, o uso de tecnologias de visualização pode ser uma ferramenta valiosa para facilitar a identificação rápida de padrões e tendências nos dados, auxiliando na tomada de decisão. Essa abordagem não apenas proporciona vantagens competitivas em análises preditivas, mas também revela problemas que poderiam passar despercebidos. As ferramentas de visualização podem ser estruturadas em vários formatos, como planilhas, relatórios, consultas e dashboards.

Os Dashboards representam painéis visuais que utilizam gráficos e tabelas para comunicar instantaneamente a performance organizacional. Ao empregar essa ferramenta, os gestores obtêm um acesso contínuo e inteligente à situação atual do desempenho de seus negócios, de maneira eficaz e imediata.

3. Trabalhos relacionados

Esta seção se dedica à exploração de estudos que utilizam técnicas de Business Intelligence e como elas podem ser aplicadas para identificar padrões e tendências, bem como análises em relação às ocorrências de acidentes viários

No trabalho realizado por Barroso Junior et al. 2019, foi utilizado dados abertos de acidentes rodoviários disponibilizados em 2016 pela Polícia Rodoviária Federal (PRF). Para a realização da análise, foi selecionado 13 variáveis de 28 variáveis disponíveis dos dados abertos coletados, utilizando o modelo binomial de regressão logística. Embora este estudo tenha contribuído significativamente para a análise de acidentes rodoviários, não foram realizadas aplicações com ferramentas específicas de Business Intelligence (BI). O uso de ferramentas de BI poderia ter auxiliado na criação de visualizações interativas e painéis de controle, melhorando a compreensão dos resultados.

Em relação à Business Intelligence foram empregadas no estudo de Souza Júnior et al. 2015 técnicas de gerenciamento de dados, incluindo a limpeza de dados (Data Cleaning), bem como a extração, transformação e carregamento (ETL) de informações provenientes de um backup da base legada do SAMU de Maceió-AL. A criação do dashboard por meio desses dados permitiu a visualização dos atendimentos relacionados a ocorrências, desempenhando um papel fundamental no planejamento estratégico das operações das ambulâncias do SAMU. Como exemplo, o dashboard desenvolvido contribuiu em melhorias significativas na gestão e na tomada de decisões das ambulâncias quanto à localização das bases descentralizadas do SAMU, especialmente nas áreas mais propensas a acidentes

Contribuindo para a análise dos acidentes de trânsito em rodovias federais brasileiras, da Costa Júnior and Júnior 2022 focam em analisar a ocorrência de vítimas fatais em acidentes de trânsito nas rodovias federais brasileiras. Para isso, utilizam-se de dados disponibilizados pela Polícia Rodoviária Federal (PRF), entre os anos de 2017 e 2019, além do modelo de regressão logística em linguagem R. Os principais resultados deste estudo incluem o registro de 184.635 acidentes nas rodovias federais brasileiras, no qual 11.340 acidentes apresentaram ocorrência de morte (vítimas fatais) e 173.295 de acidentes sem vítimas fatais (feridos e ilesos), representando um percentual de, respectivamente, 6,14% e, 93,86% dos acidentes. Estes resultados possibilitam a elaboração e adoção de políticas de prevenção por parte da esfera pública, a fim de reduzir o número de óbitos em acidentes ocorridos em rodovias federais.

Por outro lado ainda com o foco em vítimas, o estudo de Andrade and Antunes 2019 investiga a tendência de redução nas vítimas de acidentes de trânsito ocorridos nas rodovias federais do Brasil, sendo realizada a análise após a implementação da Década de Ação pela Segurança no Trânsito (DAST). Os resultados deste estudo demonstram uma tendência favorável de redução no número de vítimas fatais e feridas em acidentes de trânsito, em consonância com o objetivo geral da DAST de diminuição da morbimortalidade nas estradas do país. Este estudo, que analisa a tendência do número de mortos, feridos graves e feridos leves por acidentes de trânsito nas rodovias federais brasileiras, segundo macrorregião, antes e depois do início da DAST, utiliza dados sobre acidentes com vítimas, fatais ou feridas, disponibilizados pela Polícia Rodoviária Federal para o período de 2007 a 2017. Os métodos utilizados incluem o cálculo da variação percentual mensal (VPM) do número de mortos, feridos graves e feridos leves, utilizando o método de Prais-Winsten. Contudo, este estudo contribui para a compreensão do impacto das políticas de segurança no trânsito nas rodovias federais brasileiras e enfatiza a importância de medidas contínuas para aprimorar a segurança viária.

Em geral, todos os trabalhos relacionados apresentam contribuições valiosas para a análise geoespacial de acidentes rodoviários. Embora que os artigos mencionados utilizem a base pública da PRF, o presente estudo se concentra em uma análise aprofundada dos acidentes ocorridos na região sudeste entre os anos de 2017 a setembro de 2023. Consequentemente, utilizando informações mais recentes para realizar uma análise atualizada e detalhada dos incidentes.

4. Metodologia

Este trabalho apresenta uma pesquisa de modo qualitativa, que realiza uma análise geoespacial das ocorrências de acidentes nas rodovias federais do sudeste, utilizando a base de dados pública de acidentes da Polícia Rodoviária Federal (PRF), referente aos anos de 2017 até setembro de 2023, com abrangência nacional e disponibilizada para download no site do governo gov.br. Nas bases, estão disponíveis informações dos acidentes, como data, dia da semana, classificação do acidente, tipo do acidente, além de uma extensa gama de fatores associados ao cenário em que os acidentes ocorrem. O esquema representando as etapas dessa metodologia é visto na Figura 2

Obsevando o esquema da metodologia (Figura 2), processo de análise de dados no Power BI envolve várias etapas, a partir da coleta de dados até a criação de visualizações em um dashboard. O primeiro passo é a coleta de bases da PRF, que é seguida pela técnica ETL (Extract, Transform, Load) no Power BI. Esta técnica é iniciada pela extração e preparação dos dados para análise. A etapa de transformação envolve a limpeza dos dados, que é o processo de identificar e corrigir erros e inconsistências nos conjuntos de dados. Além disso, a transformação também pode envolver a criação de novas colunas, a modificação de dados existentes e a reorganização dos dados. O último passo da técnica ETL é o carregamento, onde os dados transformados são carregados de volta para o Power BI para análise. Na segunda etapa, os dados são modelados para atender às necessidades de análise. Isso envolve a criação de relações entre diferentes conjuntos de dados, a definição de medidas e a criação de hierarquias. Por fim, na última etapa, com os dados preparados e modelados, é feita a criação de visualizações de dados em um dashboard. Isso inclui gráficos e outros elementos visuais que ajudam na interpretação dos dados.

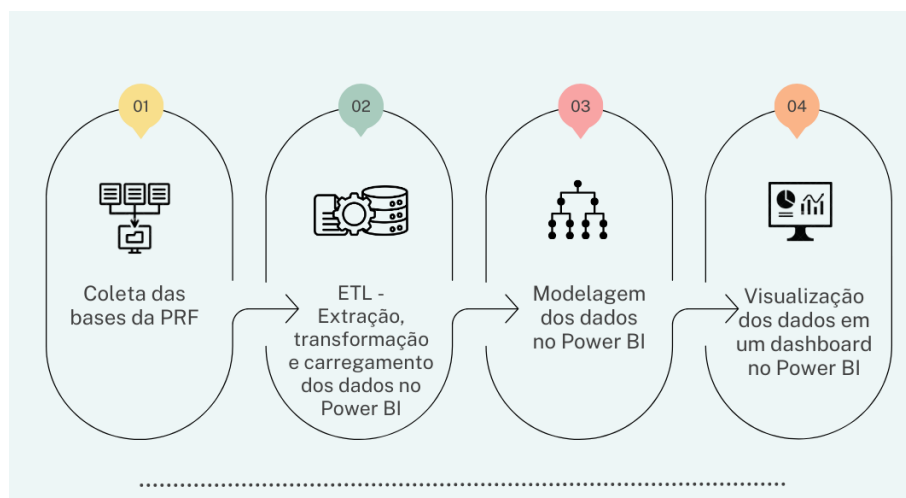


Figura 2. Demonstração da Estrutura da Metodologia

4.1. Coleta das bases

A Polícia Rodoviária Federal (PRF) disponibiliza dados abertos sobre acidentes gratuitamente em formato CSV, que podem ser usados para análises e estudos. As bases coletadas são organizadas em uma categoria principal: "Agrupados por ocorrência". Nelas é possível identificar informações sobre todas as causas e tipos de acidentes ocorridos de 2017 a 2023. Vale ressaltar que, os dados contidos nas bases são usados para analisar as causas e tipos de acidentes por ocorrência.

4.2. Extract, Transform e Load (ETL)

O processo de ETL (Extração, Transformação e Carga) é um pilar fundamental no Power BI Desktop, que permite a preparação e análise de dados de maneira eficiente e eficaz. Nesse processo, os dados são inicialmente coletados e importados a partir de arquivos CSV, um processo conhecido como extração. A extração é realizada internamente na ferramenta e, após ela, começa o processo de transformação. Inicialmente, para evitar a sobrecarga de informações desnecessárias e para melhorar a eficiência da análise de dados, foram selecionadas 19 de 30 colunas de cada base de dados "Agrupados por ocorrência", como uf, br, causa do acidente, etc. E 2 da base de dados "Feriados", como data e nome, elaborada pelos autores.

Ainda no processo de transformação, as tabelas que estavam separadas por ano são combinadas unificando-as em uma tabela. Isso é feito para facilitar a análise e a visualização dos dados ao longo do tempo. Após a combinação das tabelas, um parâmetro é definido para reduzir a quantidade de linhas a serem analisadas inicialmente, para permitir a criação do dashboard e para evitar a lentidão que pode ser causada por grandes conjuntos de dados, após isso é possível retornar com a quantidade de linhas totais para começar a análise.

Após concluídas as etapas de extração e transformação, os dados são carregados novamente para o Power BI, preparando-os para análises avançadas e visualizações significativas. Este processo integrado no Power BI Desktop destaca-se pela sua praticidade e eficácia na preparação dos dados para uma exploração mais aprofundada.

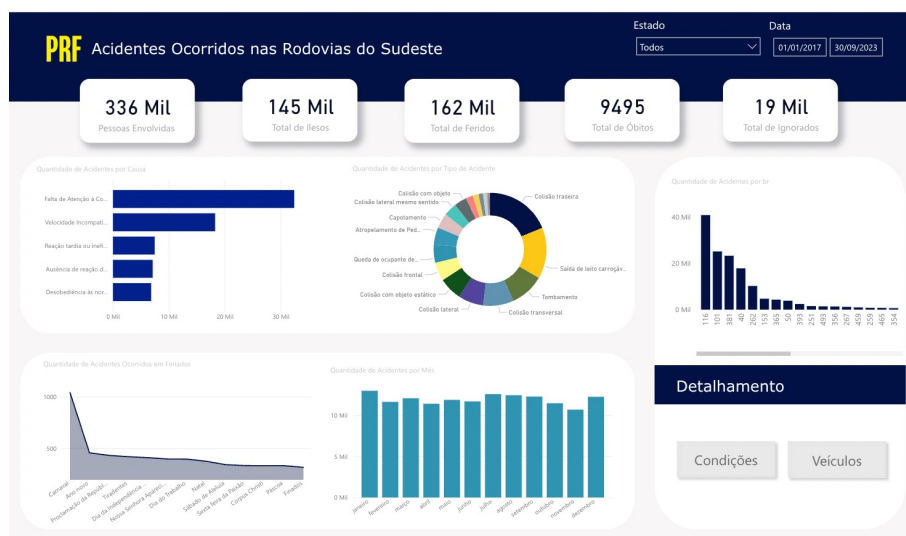


Figura 3. Demonstração do Painel Central do Dashboard

4.3. Visualização das informações

A fase final do estudo envolve a visualização dos dados, que é realizada através da criação de um painel de controle no Power BI. Este painel incorpora as informações tratadas ao longo do estudo em formato de gráficos, permitindo uma análise mais aprofundada dos acidentes ocorridos nas rodovias federais da região sudeste

Um recurso importante do painel é a inclusão de filtros anuais. Essa funcionalidade permite uma análise temporal, possibilitando a comparação e o acompanhamento dos indicadores ao longo dos anos. Ao utilizar filtros anuais, os usuários podem realizar uma investigação mais minuciosa, identificando padrões e tendências ao longo do tempo. Além disso, também foi pensado em um filtro para os estados da região sudeste, a fim de realizar uma análise mais específica da região.

5. Resultados

Esta seção se dedica aos resultados obtidos a partir das análises realizadas através do painel de controle desenvolvido durante o estudo. Com o objetivo de responder à questão proposta no presente estudo, focando na análise qualitativa, foram analisados 143.373 acidentes de trânsito ocorridos nas rodovias federais da região sudeste do Brasil entre 2017 e 2023.

A Figura 3 ilustra o painel central, que apresenta algumas consultas específicas relacionadas aos acidentes de trânsito. Para ter uma análise dinâmica, foram escolhidos dois filtros principais: Estado e Data. O filtro de Estado permite ao usuário selecionar múltiplos estados da região Sudeste, como Minas Gerais, Espírito Santo, Rio de Janeiro e São Paulo. Este recurso é útil para restringir os dados exibidos para um conjunto específico de estados, facilitando a análise de acidentes de trânsito em regiões específicas. O filtro de data permite ao usuário selecionar datas entre os anos de 2017 até setembro de 2023. Isso permite aos usuários analisar os padrões de acidentes ocorridos ao longo do tempo, identificando possíveis tendências ou mudanças nas taxas de acidentes.

Além disso, há um menu com redirecionamento para outras visões específicas no

dashboard, sendo de condições temporais e em relação ao envolvimento de veículos no acidente. Em todas as telas, são apresentados dados estáticos como a quantidade de pessoas envolvidas, o número de vítimas ilesas, vítimas feridas, vítimas fatais e acidentes ignorados. A seguir, é apresentado algumas consultas realizadas no dashboard que auxiliam na análise de padrões, fatores e tendências que contribuem para a ocorrência desses acidentes

i. Quais são as áreas de maior incidência de acidentes nas rodovias federais da região do sudeste e em cada estado da região?

A identificação das rodovias federais (BRs) com maior incidência de acidentes em um estado específico demanda a aplicação de filtros nos conjuntos de dados disponíveis, a fim de selecionar o estado de interesse. A análise desses dados pode ser conduzida tanto considerando a totalidade dos estados de forma agregada quanto realizando análises específicas para cada estado individualmente.

A Tabela 1 ilustra um ranking das cinco principais rodovias que apresentam a maior quantidade e proporção de acidentes nos estados de São Paulo (SP), Minas Gerais (MG), Espírito Santo (ES) e Rio de Janeiro (RJ). Os dados revelam que o estado de São Paulo registra a maior proporção de acidentes, especificamente na BR-116, que representa 15,49% dos acidentes totais (143.373). Isso indica que a BR-116 no estado de SP tem a maior incidência de acidentes em relação às rodovias federais da região Sudeste.

Para aprofundar a análise nas áreas com maior índice de acidentes, foi realizado uma consulta que relaciona o quilômetro com a rodovia e o estado que apresentam a maior quantidade de acidentes ocorridos nele. A Tabela 2 exibe o total de acidentes em cada quilômetro, de acordo com as cinco principais BRs que demonstraram a maior proporção. Ao analisá-la, percebe-se que a BR-116 no estado de SP, que registra a maior quantidade de acidentes, apresenta o km 219 com a maior quantidade em relação aos outros estados e KMs da mesma rodovia.

Estado	BR	Total de acidentes	% Acidentes
SP	116	22214	15.49%
MG	381	19059	13.29%
ES	101	12674	8.84%
MG	40	12337	8.60%
RJ	101	11314	7.89%

Tabela 1. Acidentes nas rodovias federais do sudeste.

ii. Quais são as áreas de maior incidência de acidentes com vítimas fatais nas rodovias federais da região do sudeste e em cada estado da região?

Ao considerar a quantidade de acidentes que resultaram em vítimas, os dados demonstram que aproximadamente 336 mil indivíduos foram envolvidos. Dessa quantidade, 145 mil saíram ilesos, 162 mil apresentaram ferimentos, 9945 óbitos e 19 mil foram ignorados. Além disso, foi calculada a proporção de vítimas fatais em relação ao número de acidentes por BR (Tabela 3).

Embora a BR-116 tenha um número total e proporção maior de acidentes, a BR-381 no estado de Minas Gerais tem um total de óbitos mais elevado. Além disso, em

BR - Estado	KM	Acidentes
116 - SP	219	327
101 - ES	270	164
381 - MG	481	123
101 - RJ	322	193

Tabela 2. Acidentes nos km por rodovia e estado

Minas Gerais, a BR-40 tem a maior proporção de acidentes, totalizando 3,03%. Isso sugere que, embora a BR-116 possa ter mais acidentes, a BR-381 pode ser mais perigosa em termos de mortalidade. Além disso, a BR-40, apesar de ter menos acidentes, tem a maior proporção de acidentes em Minas Gerais, o que indica que essa rodovia pode ser particularmente perigosa.

Esses dados destacam a necessidade de medidas de segurança mais rigorosas em todas as rodovias, especialmente em Minas Gerais nas BR-381 e na BR-40, para reduzir o número de acidentes e mortes.

Estado	BR	Total Acidentes	Pessoas envolvidas	Óbitos	Vítimas fatais
MG	40	12337	29929	907	3.03
Rj	101	11314	27375	712	2.60
MG	381	19059	45127	1145	2.54
ES	101	12674	29440	734	2.49
SP	116	22214	50960	1114	2.19

Tabela 3. Taxa de mortalidade nas rodovias federais

iii. Quais são as principais causas de acidentes em relação ao tipo de acidente em um determinado estado ou em todos os estados da região sudeste?

A Tabela 4 apresenta as cinco categorias principais de causas de acidentes de trânsito na região sudeste do Brasil. Ao examinar as três primeiras causas, a principal causa de acidentes de trânsito mais frequente nos quatro estados é a falta de atenção à condução, que resultou em 32.423 mil incidentes. Em segundo lugar, a velocidade incompatível teve 18.280 acidentes. Por fim, a reação tardia ou ineficiente do condutor foi responsável por 7.475 acidentes

Ao examinar os tipos de acidentes de trânsito mais comuns em relação à falta de atenção do condutor, nota-se que as colisões são o tipo mais frequente. Em particular, a colisão traseira é a mais comum, mantendo a primeira posição em todos os estados, com um total de 14.671 acidentes (Tabela 5).

A Figura 4 representa de maneira gráfica no dashboard, mostrando dois gráficos: um que ilustra a causa dos acidentes e outro que mostra o tipo de acidente. Estes dois gráficos estão interligados, proporcionando uma visão clara e concisa das principais causas e tipos de acidentes de trânsito na região sudeste do Brasil. Portanto, conclui-se que, a falta de atenção à condução é a principal causa de acidentes de trânsito na região sudeste, com colisões, e em particular colisões traseiras, sendo o tipo de acidente mais comum.

iv. Em quais feriados mais ocorrem acidentes na região do Sudeste?

Causa	ES	MG	RJ	SP
Falta de Atenção à Condução	4543	12391	7944	7545
Velocidade Incompatível	1432	11635	2521	2692
Reação tardia ou ineficiente do condutor	755	2613	1824	2283
Ausência de reação do condutor	692	2905	1518	2003
Desobediência às normas de trânsito pelo condutor	1625	2651	1543	1022

Tabela 4. Principais causas de acidentes na região sudeste

Tipo	ES	MG	RJ	SP	Total
Colisão traseira	1008	2689	1920	2321	7938
Colisão lateral	887	1785	1403	1479	5554
Colisão transversal	947	1881	993	474	4295
Saída de leito carroçável	236	2032	546	553	3367
Colisão com objeto estático	236	909	1006	929	3080

Tabela 5. Tipos de acidentes em relação à causa de acidente mais recorrente

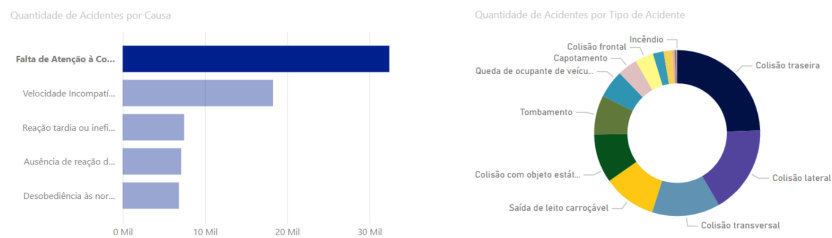


Figura 4. Tipos de acidentes em relação à causa de acidente mais recorrente no Dashboard

O dashboard apresenta uma análise detalhada dos feriados e seus impactos na incidência de acidentes de trânsito em rodovias específicas. O gráfico da Figura 5 destaca que o feriado com a maior quantidade de acidentes é o Carnaval, com um total de 1.041 registrados. Conforme a CNT 2018, o Carnaval é reconhecido como o feriado mais perigoso para as rodovias federais do Brasil.

A Tabela 6 evidencia que os dois estados com o maior número de acidentes durante o Carnaval entre os anos de 2017 e 2023 foram Minas Gerais que registrou o maior número de acidentes, com um total de 463 ocorrências, e São Paulo com 234 acidentes.

A alta incidência de acidentes durante o Carnaval pode ser atribuída a uma combinação de fatores, incluindo o aumento do consumo de álcool, o desrespeito aos limites de velocidade e a redução da quantidade de veículos nas ruas. Portanto, esses dados sublinham a importância de implementar medidas preventivas e de segurança du-

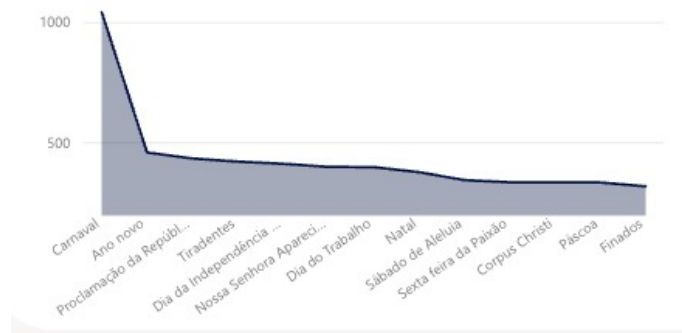


Figura 5. Total de acidentes por feriado

Estado	Acidentes	Feriado
SP	234	Carnaval
ES	117	Carnaval
MG	463	Carnaval
RJ	227	Carnaval

Tabela 6. Acidentes por feriado - Carnaval

rante o período do Carnaval, com o objetivo de reduzir esses números significativos de ocorrências.

v. Quais são os meses que mais ocorrem acidentes na região do Sudeste e qual a relação entre os feriados ocorridos neles?

Os meses de janeiro e dezembro são os que apresentam a maior quantidade de acidentes na região Sudeste (Figura 6), com um total de 12.988 e 12.251 acidentes, respectivamente. Durante o mês de janeiro no feriado de Ano novo, ocorreram 457 acidentes, o que representa uma proporção de aproximadamente 3,51% em relação ao total de acidentes no mês. Em dezembro, durante o feriado de Natal, ocorreram 375 acidentes, o que representa uma proporção de aproximadamente 3,06% em relação ao total de acidentes no mês.

Essas estatísticas ressaltam que os feriados nacionais, particularmente o Natal e o Ano Novo, podem contribuir para o aumento do número de acidentes. Esse fato pode ser atribuído a uma série de fatores, incluindo o aumento do tráfego rodoviário e o consumo de álcool durante as celebrações, por exemplo.

vi. Qual é a proporção de vítimas fatais em relação a fase do dia, condição meteorológica e dia da semana com o total de acidentes?

Os resultados apresentados indicam que a fase do dia, as condições meteorológicas e o dia da semana têm uma influência significativa na proporção de vítimas fatais em relação ao total de acidentes.

Em relação a fase do dia que apresenta a maior quantidade de acidente é pleno dia, com o total de 78.682 acidentes. No entanto, a proporção de vítimas fatais nesta fase é apenas 2.04%, sendo a fase com a menor taxa de mortalidade entre os quatro estados. Em contraste, a fase do dia mais propensa a acidentes fatais é a plena noite, com um total

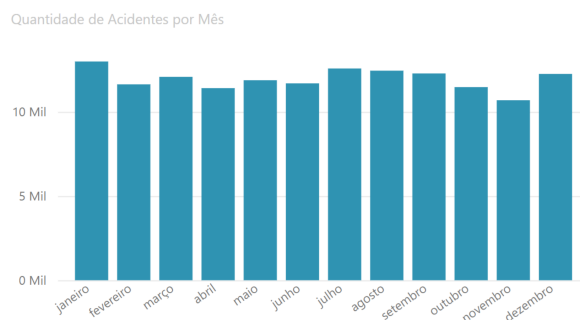


Figura 6. Meses que mais ocorrem acidentes na região Sudeste

de 49.329 acidentes e uma proporção de vítimas fatais de 4.01%, o que sugere que a noite é o período mais propenso a acidentes fatais (Tabela 7).

Quanto às condições meteorológicas, os dias mais propensos a acidentes são aqueles com céu claro totalizando 79.242, nublado com 26.827 e chuva tendo 19.681 acidentes. No entanto, a condição que apresenta a maior proporção de vítimas fatais é a nevoeiro/neblina, com 4,95% (Tabela 8), totalizando 1019 acidentes.

Por fim, em relação ao dia da semana, os dias com maior quantidade de acidentes são o domingo totalizando 23.947 acidentes, o sábado com 23.758 e sexta-feira com 22.253. O domingo apresenta a maior quantidade de acidentes e a maior proporção de vítimas fatais, com 3,20%. O sábado permanece em segundo lugar com 3,09%, enquanto sexta-feira, apesar de estar em terceiro lugar com a quantidade de acidentes, representa uma proporção de vítimas fatais de apenas com 2,55%, ficando em último lugar em relação aos outros dias da semana (Tabela 9). Conclui-se que o final de semana é o mais perigoso em relação a acidentes e mortes.

Fase do dia	ES	MG	RJ	SP	Total
Plena Noite	3.53%	4.42%	3.91%	3.60%	4.01%
Amanhecer	3.66%	4.59%	3.72%	3.08%	3.96%
Anoitecer	2.20%	3.35%	1.85%	1.06%	2.34%
Pleno dia	2.02%	2.55%	1.86%	1.12%	2.04%

Tabela 7. Taxa de vítimas fatais por fase do dia em relação a quantidade de acidentes

Condição Meteorológica	ES	MG	RJ	SP	Total
Nevoeiro/Neblina	5.16 %	5.19%	4.71%	4.68%	4.95%
Vento	0.00%	4.79%	1.80%	3.19%	3.62%
Céu Claro	2.57%	3.45%	2.76%	2.27%	2.94%
Nublado	2.92%	3.39%	2.77%	2.24%	2.90%
Chuva	2.77%	2.87%	2.38%	1.80%	2.56%

Tabela 8. Taxa de vítimas fatais por condição meteorológica em relação a quantidade de acidentes

Dia da Semana	ES	MG	RJ	SP	Total
domingo	2.86%	3.84%	2.88%	2.35%	3.20%
sábado	3.26%	3.37%	2.93%	2.55%	3.09%
quinta-feira	2.84%	3.13%	2.70%	2.00%	2.75%
segunda-feira	2.17%	3.22%	2.62%	2.02%	2.70%
terça-feira	2.21%	3.14%	2.62%	2.21%	2.70%
quarta-feira	2.42%	3.22%	2.36%	1.87%	2.63%
sexta-feira	2.27%	3.04%	2.53%	1.74%	2.55%

Tabela 9. Taxa de vítimas fatais por dia da semana em relação a quantidade de acidentes

6. Conclusões

Este estudo teve como principal objetivo realizar uma análise geoespacial detalhada das ocorrências de acidentes nas rodovias federais da região Sudeste, utilizando um dashboard desenvolvido com técnicas de Business Intelligence. Os dados utilizados para esta análise foram fornecidos pela Polícia Rodoviária Federal (PRF) e abrangem os anos de 2017 a 2023.

Para atingir o objetivo proposto, um dashboard foi desenvolvido no Power BI, a fim de identificar áreas de alto risco onde os acidentes de trânsito são mais frequentes ou mais graves. Além disso, ele foi projetado para identificar padrões temporais, como o dia da semana, a condição meteorológica e os feriados nacionais, que podem influenciar no número de acidentes. O dashboard também permite a classificação e análise de categorias de acidentes, como a causa e o tipo de acidente mais comuns. Através dessa análise, foram identificados insights significativos que podem ser utilizados para informar estratégias de prevenção e mitigação de acidentes de trânsito no futuro.

Conclui-se que os dados analisados indicam que o estado de São Paulo, especificamente na BR-116 no km 219, tem a maior incidência de acidentes de trânsito. No entanto, em termos de risco, Minas Gerais é o estado mais perigoso, com a BR-40 apresentando a maior proporção de vítimas fatais e a BR-381 registrando o maior número de óbitos. É importante destacar que a falta de atenção à condução é a principal causa de acidentes de trânsito na região sudeste do Brasil. Colisões traseiras são o tipo de acidente mais comum. Além disso, os dados mostram que ocorrem mais acidentes durante o dia (pleno dia), mas a taxa de mortalidade é maior durante a noite. A condição meteorológica mais perigosa para vítimas fatais é o nevoeiro/neblina e, por fim, o final de semana é o período mais perigoso em termos de acidentes e mortes.

Para trabalhos futuros, sugere-se a implementação de um mapa com as coordenadas geográficas, como latitude e longitude, dos dados disponibilizados da PRF. A razão para isso é que a geoespacialização dos dados pode proporcionar uma representação visual mais precisa e informativa dos acidentes nas rodovias.

Além disso, embora este estudo tenha identificado as principais causas dos acidentes, uma análise mais aprofundada e detalhada dessas causas pode ajudar a desenvolver novas estratégias de prevenção mais precisas.

Por fim, sugere-se a implementação de novas fontes de dados é uma estratégia importante para aprofundar a análise dos acidentes de trânsito. Dados de veículos da

Polícia Rodoviária Federal (PRF) podem ser usados para analisar a relação entre o tipo de veículo e o risco de acidentes. Além disso, a inclusão de informações adicionais sobre os condutores e passageiros, como idade e sexo, pode proporcionar uma visão mais completa e detalhada dos acidentes.

Referências

ANDRADE, F. R. d. and ANTUNES, J. L. F. (2019). **Tendência do número de vítimas em acidentes de trânsito nas rodovias federais brasileiras antes e depois da década de ação pela segurança no trânsito.** Cadernos de Saúde Pública, 35.

BARBIERI, C. (2011). **BI2 Business Intelligence. Modelagem e Qualidade.** Elsevier, Rio de Janeiro.

BARROSO JUNIOR, G. T., BERTHO, A. C. S., and VEIGA, A. d. C. (2019). **A letalidade dos acidentes de trânsito nas rodovias federais brasileiras em 2016.** Revista Brasileira de Estudos de População, 36:e0074.

BRASIL. Agência CNT Transporte Atual. **Feriado do Carnaval é o mais perigoso do país em rodovias federais.** Brasília, DF: Confederação Nacional do Transporte, 2018. Disponível em : <https://www.cnt.org.br/agencia-cnt/carnaval-feriado-mais-perigoso-rodovias>. Acesso em: 23 nov. 2023

BRASIL. Agência CNT Transporte Atual . **Índice de acidentes nas rodovias cai, mas o número de vítimas fatais se mantém em 2020.** Brasília, DF: Confederação Nacional do Transporte, 2020. Disponível em : <https://www.cnt.org.br/indice-acidentes-cai-vitimas-fatais-se-mantem> Acesso em: 20 Ago. 2023

BRASIL. Confederação Nacional do Transporte . **Painel cnt de consultas dinâmicas dos acidentes rodoviário.** Brasília, DF: Confederação Nacional do Transporte, 2018. Disponível em : <https://www.cnt.org.br/painel-acidente>. Acesso em: 20 Ago. 2023

BRASIL. Confederação Nacional do Transporte. **Anuário do transporte.** Brasília, DF: Confederação Nacional do Transporte, 2022. Disponível em: <https://anuariodotransporte.cnt.org.br/2021/Rodoviario/1-1-/Principais-dados>. Acesso em: 20 Ago. 2023

CALDEIRA, J. (2010). **Dashboards: Comunicar eficazmente a informação de gestão.** Almedina, Lisboa.

DA COSTA JÚNIOR, E. E. and JÚNIOR, L. C. S. (2022). **Risco de morte em acidentes ocorridos nas rodovias federais brasileiras.** Refas-Revista Fatec Zona Sul, 9(2):1–22.

INMON, W. H. (1997). **Como construir o Data Warehouse.** Campus, Rio de Janeiro.

KIMBALL, R. (1998). **Data Warehouse Toolkit:Técnicas para construção de data warehouses dimensionais.** Makron Books, São Paulo.

KIMBALL, R. and ROSS, M. (2010). **The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence.** Wiley, Indianapolis.

SOUZA JÚNIOR, M., MELVILLE, M., MEDEIROS, J., and ALVES, A. (2015).

Adotando técnicas de geolocalização e business intelligence para visualização dos atendimentos móveis de urgência a partir de um dashboard.

TURBAN, E., SHARDA, R., KING, D., and ARONSON, J. E. (2008). **Business Intelligence: Um Enfoque Gerencial para a Inteligência do Negócio.** BOOKMAN, Porto Alegre.