



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLOGIA DA INFORMAÇÃO
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

JOSEPH DA SILVA OLIVEIRA

**ANÁLISE DE ACIDENTES DE TRÂNSITO OCORRIDOS NAS RODOVIAS
FEDERAIS DO RIO GRANDE DO NORTE: EXTRAINDO CONHECIMENTO
UTILIZANDO DADOS ABERTOS**

ANGICOS

2021

JOSEPH DA SILVA OLIVEIRA

**ANÁLISE DE ACIDENTES DE TRÂNSITO OCORRIDOS NAS RODOVIAS
FEDERAIS DO RIO GRANDE DO NORTE: EXTRAINDO CONHECIMENTO
UTILIZANDO DADOS ABERTOS**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal Rural do Semi-Árido
como requisito para obtenção do título de
Bacharel em Sistemas de Informação.

Orientador:
Prof. Me. Serafim do Nascimento Junior

ANGICOS

2021

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

d048a da Silva Oliveira, Joseph.
Análise de Acidentes de Trânsito Ocorridos nas
Rodovias Federais do Rio Grande do Norte:
Extraíndo Conhecimento Utilizando Dados Abertos /
Joseph da Silva Oliveira. - 2021.
42 f. : il.

Orientador: Serafim do Nascimento Junior.
Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Sistemas de
Informação, 2021.

1. Acidentes de trânsito. 2. Década de ação para
segurança no trânsito. 3. Análise de dados da PRF.
I. do Nascimento Junior, Serafim, orient. II.
Título.

JOSEPH DA SILVA OLIVEIRA

**ANÁLISE DE ACIDENTES DE TRÂNSITO OCORRIDOS NAS RODOVIAS
FEDERAIS DO RIO GRANDE DO NORTE: EXTRAINDO CONHECIMENTO
UTILIZANDO DADOS ABERTOS**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal Rural do Semi-Árido
como requisito para obtenção do título de
Bacharel em Sistemas de Informação.

Orientador:
Prof. Me. Serafim do Nascimento Junior

Defendida em: **01 / 06 / 2021.**

BANCA EXAMINADORA

SERAFIM DO
NASCIMENTO  Assinado de forma digital por
SERAFIM DO NASCIMENTO
JUNIOR:09511149482
JUNIOR:09511149482 Dados: 2021.06.01 16:46:04 -03'00'

Prof. Serafim do Nascimento Junior (UFERSA)
Mestre em Ciência da Computação
Presidente e Orientador

ARAKEN DE MEDEIROS  Assinado digitalmente por ARAKEN DE
MEDEIROS SANTOS:00739803409
SANTOS:00739803409 Data: 2021-06-02 17:44:39

Prof. Araken de Medeiros Santos (UFERSA)
Doutor em Ciência da Computação
Membro Examinador Interno

SARA GUIMARAES
NEGREIROS: 11051711401  Digitalmente assinado por: SARA
GUIMARAES NEGREIROS:11051711401
E-mail: SGUIMARAES@GMAIL.COM
Data: 2021/06/01 18:11:57 -03'00'

Profa. Sara Guimarães Negreiros (UFERSA)
Bacharel em Engenharia de Computação
Membro Examinador Interno

AGRADECIMENTOS

Primeiramente a Deus por ser minha força e meu escudo nos momentos e tempos difíceis durante esta longa caminhada.

Aos meus pais, Marco Antônio Barbosa de Oliveira e Jeane Cosme da Silva Oliveira, pela vida, pela força e apoio em todas as minhas decisões. A minha irmã Jade Hannah da Silva Oliveira e a minha segunda mãe Maria Auxiliadora, a mulher mais doce e bondosa que já conheci e que tenho a honra de chamar de vó. A vocês, que são o motivo de acordar e dar o meu melhor todos os dias, todo meu amor e gratidão.

A minha namorada e futura esposa Vanessa, por todo o companheirismo e compreensão.

Aos mestres que contribuíram com seus ensinamentos durante essa caminhada, José Gildo, Francisco de Assis, Wellington Barbosa, Adriana Mara, Araken de Medeiros, Enai Taveira, Joemia Leilane, Jackney Luan, Kleber Tavares, Patrício de Alencar, Sairo Raioni, Samuel Oliveira, Andrezza Cristina, Ítalo Rebouças e tantos outros que contribuíram com minha formação.

Um agradecimento especial para o professor Serafim do Nascimento Junior que aceitou essa missão de me orientar de maneira tão solícita e empática.

Para todos os colegas e amigos que estiveram no campo de batalha do saber dividindo a mesma trincheira comigo: Jerônimo de Castro, Edson Cândido, Cleyton Rage, Anderson Cirilo, Bárbara Lorena, Francisco das Chagas, Glendo Victor, Ariel Lucas, Vanderson David, Pedro Henrique, John Lennon Moraes e tantos outros que ficaram pelo caminho.

“Combati o bom combate,
terminei a corrida, guardei a fé.”
(2 Timóteo 4:7)

“A persistência é o caminho do êxito.”
Charles Chaplin.

RESUMO

As estatísticas levantadas pela Polícia Rodoviária Federal (PRF) disponibilizados para atender a Lei de Acesso à Informação (LAI) (12.527/2011) indicam que a expansão das cidades e as oportunidades de aquisição de veículos facilitadas, estão diretamente relacionadas com o aumento constante no número de ocorrência de acidentes em rodovias federais brasileiras. O relatório de status global sobre segurança no trânsito (*Global status report on road safety 2018*) aponta que no ano de 2016, 1,35 milhões de pessoas perderam suas vidas em acidentes de trânsito ao redor do mundo, tornando o tema de relevância global. Neste sentido, a Organização Mundial de Saúde (OMS) juntamente com a Organização das Nações Unidas (ONU) dá início a década de ação para segurança no trânsito 2011-2020, que tem como meta a redução de 50% de acidentes ocasionados por veículos. O presente trabalho realizou uma análise dos dados disponibilizados pela PRF utilizando conceitos de mineração de dados para verificar se o estado do Rio Grande do Norte alcançou a meta proposta, assim como classificar os acidentes ocorridos nos municípios.

Palavras-chave: Acidentes de trânsito. Década de ação para segurança no trânsito. Análise de dados da PRF.

ABSTRACT

The statistics collected by the Federal Highway Police (PRF) made available to comply with the Access to Information Law (LAI) (12,527 / 2011) indicate that the expansion of cities and the opportunities for the acquisition of facilitated vehicles are directly related to the constant increase in number of accidents occurring on Brazilian federal highways. The global status report on road safety 2018 (Global status report on road safety 2018) points out that in 2016, 1.35 million people lost their lives in traffic accidents around the world, making the topic of global relevance. In this sense, the World Health Organization (WHO) together with the United Nations (UN) begins the decade of action for road safety 2011-2020, which aims to reduce 50% of accidents caused by vehicles. The present work carried out an analysis of the data made available by the PRF using data mining concepts to verify whether the state of Rio Grande do Norte reached the proposed goal, as well as classify the accidents that occurred in the municipalities.

Keywords: Traffic accidents, Decade of action for traffic safety, PRF data analysis.

RESUMEN

Las estadísticas recopiladas por la Policía Rodoviaria Federal (PRF) puestas a disposición para dar cumplimiento a la Ley de Acceso a la Información (LAI) (12.527 / 2011) indican que la expansión de las ciudades y las oportunidades para la adquisición de vehículos facilitados están directamente relacionadas con el constante aumento del número de accidentes ocurridos en las carreteras federales brasileñas. El informe de situación global de la seguridad vial 2018 (Informe de situación global de la seguridad vial 2018 señala que en 2016, 1,35 millones de personas perdieron la vida en accidentes de tráfico en todo el mundo, lo que hace que el tema sea de relevancia mundial. En este sentido, la Organización Mundial de la Salud (OMS) junto con las Naciones Unidas (ONU) inicia la década de acción por la seguridad vial 2011-2020, que tiene como objetivo reducir en un 50% los accidentes provocados por vehículos. El presente trabajo realizó un análisis de los datos puestos a disposición por la PRF utilizando conceptos de minería de datos para verificar si el estado de Rio Grande do Norte alcanzó la meta propuesta, así como para clasificar los accidentes ocurridos en los municipios.

Palabras clave: Accidentes de tráfico, Década de acción para la seguridad vial, Análisis de datos PRF.

LISTA DE FIGURAS

Figura 1 - Número e taxa de mortes no trânsito por 100.000 habitantes: 2000–2016	17
Figura 2 - Colisão traseira	18
Figura 3 - Colisão lateral	19
Figura 4 - Colisão transversal	19
Figura 5 - Colisão frontal	19
Figura 6 - Choque contra objeto	20
Figura 7 - Visualização de Dados	22
Figura 8 - Número de ocorrências por estado do Brasil	23
Figura 9 - Ilustração de vacinação por município.	25
Figura 10 - Linha Temporal	26
Figura 11 - Ocorrências da PRF do ano de 2019	28
Figura 12 - Dados sendo carregados na biblioteca Pandas	28
Figura 13 - Merge dos dados	29
Figura 14 - Função para filtrar o RN	29
Figura 15 - Soma de dados faltantes	30
Figura 16 - Exemplo de dados fora do padrão	31
Figura 17 - Mapa de calor de 2017 a 2020	32
Figura 18 - Classificação de acidentes no município de Natal	33
Figura 19 - Número de ocorrências e óbitos dos últimos 10 anos no estado do RN	35
Figura 20 - Meses do ano em que ocorrem o maior número de acidentes	36
Figura 21 - Horários em que ocorrem o maior número de acidentes	37

LISTA DE ABREVIATURAS E SIGLAS

AT	Acidentes de Trânsito
CTB	Código de Trânsito Brasileiro
HPV	Papilomavírus Humano
KDD	<i>Knowledge-Discovery in Databases</i>
NBR	Norma Técnica
OMS	Organização Mundial de Saúde
ONU	Organização das Nações Unidas
PDA	Plano de Dados Abertos
PRF	Polícia Rodoviária Federal
SDG	<i>Sustainable Development Goal</i>
SIM	Sistema de Informação sobre Mortalidade
LAI	Lei de Acesso à Informação
UnB	Universidade de Brasília

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Justificativa	13
1.2	Objetivos	15
1.3	Organização do Trabalho	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Acidentes de trânsito	16
2.2	Classificação dos Acidentes	17
2.3	Knowledge Discovery in Databases, mineração e visualização de dados..	20
3	TRABALHOS RELACIONADOS	23
4	METODOLOGIA	26
4.1	Coleta dos dados	27
4.2	Pré-processamento dos dados	27
5	RESULTADOS E DISCUSSÃO	30
5.1	Acidentes no município de Natal nos últimos 10 anos	32
6	CONSIDERAÇÕES FINAIS	34
	REFERÊNCIAS	38
	ANEXO A – SITE (DADOS.GOV.BR)	41
	ANEXO B – SITE (PORTAL.PRF.GOV.BR)	43

1 INTRODUÇÃO

Segundo o Art.1º da Lei 9.503, de 23 de setembro de 1997 que institui o Código de Trânsito Brasileiro (CTB) define o significado de trânsito como sendo o uso das vias por pessoas, automóveis, animais (isolados ou em grupos), conduzidos ou não, para objetivo de circulação, parada, estacionamento e operação de carga ou descarga. O código afirma que é direito de todos os cidadãos brasileiros, obrigação dos órgãos e entidades responsáveis do Sistema Nacional de Trânsito (CTB), oferecer condições seguras de trânsito e adotar precauções para garantir esse direito seja exercido (BRASIL, 1997).

Os acidentes de trânsito são responsáveis por grande número de óbitos. No Brasil os números atingem mais de 150 mil mortos e feridos anualmente, colocando-o na quinta posição entre os países com maior mortalidade por lesões e sequelas decorrentes de acidentes de trânsito (COSTA et al., 2017).

Por meio de indicadores obtidos no registro de ocorrências da Polícia Rodoviária Federal (PRF), disponibilizados para atender a Lei de Acesso à Informação (12.527/2011) destaca como um dos principais fatores que mais causam acidentes, o comportamento inadequado dos condutores no trânsito, que por sua vez é agravado pelas más condições das vias, como a falta de sinalização e a falta de atenção.

Ainda de acordo com a Lei de Acesso à Informação, em seu Capítulo I, Art. 3º, vê-se que os procedimentos previstos na lei destinam-se a assegurar o direito fundamental de acesso à informação, devendo ser executados em conformidade com os princípios básicos da administração pública e com as seguintes diretrizes: I - observância da publicidade como preceito geral e do sigilo como exceção; II - divulgação de informações de interesse público, independentemente de solicitações; III - utilização de meios de comunicação viabilizados pela tecnologia da informação; IV - fomento ao desenvolvimento da cultura de transparência na administração pública; V - desenvolvimento do controle social da administração pública (BRASIL, 2011).

Os dados abertos são publicados em um formato que é legível por máquina, e podem ser livremente utilizados por qualquer pessoa. O usuário dos dados abertos sujeita-se a citar a fonte original e a compartilhar as licenças em que as informações foram inicialmente apresentadas (ISOTANI; BITTENCOURT, 2015). Apesar da existência da Lei 12.527, verificam-se dificuldades na utilização dos dados disponibilizados, tais como dados faltantes, dados implícitos (o significado do dado não pode ser inferido, ou não há referência a esse

dado no dicionário de dados), ou ainda lentidão no processamento devido à grande quantidade de dados. Assim sendo, a utilização de técnicas de mineração de dados pode permitir a extração de informações implícitas, previamente desconhecidas e potencialmente úteis (AGGARWAL et al, 2015).

1.1 Justificativa

De acordo com o exposto anteriormente, e embora a política de dados abertos busque franquear aos cidadãos o acesso de forma livre aos dados produzidos ou acumulados pelo Poder Público, tais dados são, atualmente, de difícil acesso e visualização. Este trabalho se justifica uma vez que, os dados abertos da Polícia PRF são disponibilizados em arquivos no formato .CSV¹, em texto simples sem formatação, que pode ser de difícil manipulação, leitura e extração de informações úteis pela maioria da população.

O portal brasileiro de dados abertos², é o meio pelo qual o governo brasileiro disponibiliza o acesso aos dados abertos e a informações públicas. Como é mostrado no Anexo A, os dados da PRF contidos no portal brasileiro de dados abertos não são atualizados desde 2017. No mesmo portal há informação para que, caso existam dados faltantes, o interessado em utilizar os dados abertos deve verificar o cronograma estabelecido no Plano de Dados Abertos (PDA) do órgão em questão. No capítulo IV do decreto nº 8.777/2016 verifica-se que a pesquisa científica de base empírica sobre a gestão pública deve ser estimulada. Ressalta-se, entretanto, que o portal da PRF destoa do portal de dados abertos do governo. O portal da PRF apresenta informações e periodicidade de atualização de dados mais frequente e com uma melhor organização, conforme pode ser visualizado no Anexo B.

Assim sendo, o trabalho proposto justifica-se por considerar importantes: (a) a dificuldade de manipulação e visualização dos dados abertos disponibilizados pelo governo brasileiro, em específico os dados disponibilizados pela PRF; (b) a utilização de técnicas de mineração de dados para análise de dados abertos disponibilizados pela PRF; (c) a importância global do tema da segurança no trânsito, estabelecida pela Organização das Nações Unidas (ONU) com meta de redução de 50% do número de mortes no trânsito no período compreendido entre os anos de 2011 a 2020.

¹ CSV é um formato de arquivo que significa “comma-separated-values” (valores separados por vírgulas). <https://support.office.com/pt-br/article/criar-ou-editar-arquivos-csv-para-importar%C3%A7%C3%A3o-para-o-outlook-4518d70d-8fe9-46ad-94fa-1494247193c7>. Acesso em: 25 jan. 2020.

² <http://dados.gov.br/pagina/sobre>. Acesso em: 28 jan. 2020.

1.2 Objetivos

Os objetivos que nortearam o desenvolvimento deste trabalho são apresentados a seguir:

- **Objetivo Geral**

Analisar os dados abertos da PRF por meio de mineração de dados, a fim de quantificar e qualificar acidentes nas rodovias federais do Rio Grande do Norte, na tentativa de analisar o avanço do estado na década de ação para segurança no trânsito.

- **Objetivos específicos**

- Coletar os dados no site da PRF.
- Realizar o pré-processamento e limpeza dos dados.
- Minerar dados para extrair padrões e relacionamentos.
- Extrair conhecimento a respeito dos dados analisados.
- Aplicar técnicas de visualização de dados.

1.3 Organização do Trabalho

Este trabalho se encontra estruturado da seguinte forma: o Capítulo 1 apresenta a introdução do trabalho; o Capítulo 2 denota a fundamentação teórica necessária para o entendimento do problema e os conceitos necessários para auxiliar no desenvolvimento do trabalho; o Capítulo 3 apresenta alguns trabalhos relacionados ao proposto neste TCC; o Capítulo 4 foca na metodologia adotada para o trabalho; o Capítulo 5 detalha os resultados e algumas discussões sobre o trabalho; o Capítulo 6 descreve as considerações finais e os trabalhos futuros; e, por fim, o capítulo que lista todas as referências utilizadas para auxiliar no processo de elaboração deste trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo é destinado à apresentação da fundamentação teórica do trabalho proposto, e foi dividida em duas subseções. Na Seção 2.1 é abordado o conteúdo referente à acidentes de trânsito, a Seção 2.2 aborda sobre a classificação e os tipos de acidente, em conjunção com uma análise dos acidentes de trânsito relatados pela Organização Mundial de Saúde (OMS). Na segunda seção, é apresentada a fundamentação teórica de aspectos da mineração de dados e da visualização de dados.

2.1 Acidentes de trânsito

O trânsito é importante para o desenvolvimento, pois auxilia no desenvolvimento econômico e na mobilidade urbana, mas a falta de atenção com esse setor fez com que os sistemas de trânsito fossem se desenvolvendo de maneira desregrada, causando uma perda de vidas, saúde e recursos (OMS, 2012). De acordo com a Norma Técnica (NBR) 10697:2018 que define acidentes de trânsito como “todo evento não premeditado e que resulte em dano em veículo ou na sua carga e/ou lesões em pessoas e/ou animais, onde pelo menos uma das partes está em movimento nas vias terrestres ou em áreas abertas ao público” (ABNT, 2018).

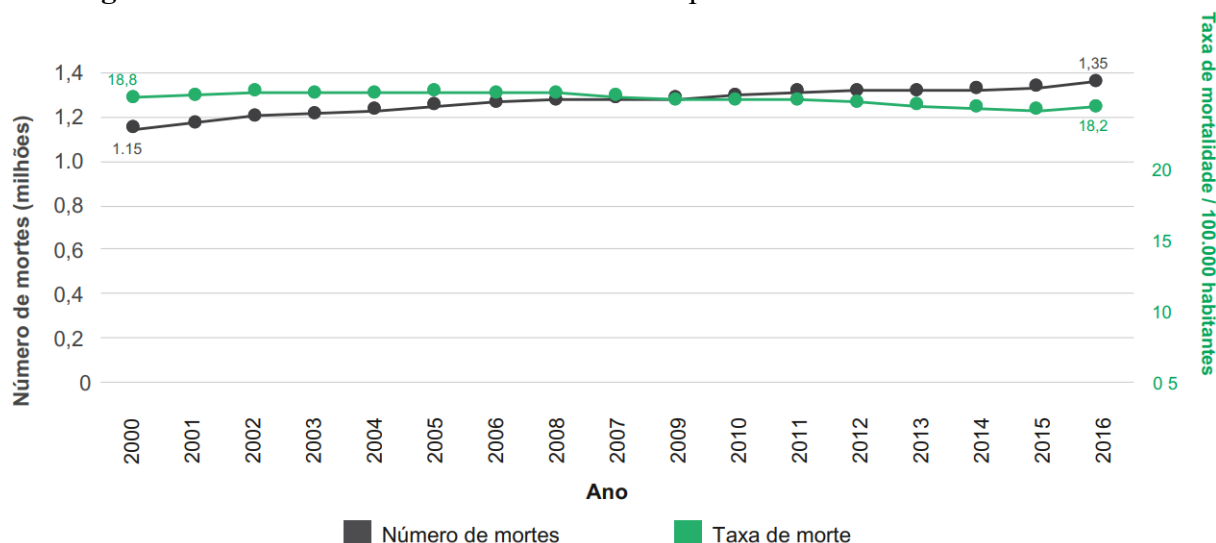
Sob o ponto de vista social, o autor Vasconcelos (2017) define o trânsito como “uma disputa pelo espaço físico, que reflete em uma disputa pelo tempo e pelo acesso aos equipamentos urbanos; é uma negociação permanente do espaço, coletiva e conflituosa. E essa negociação, dadas as características de nossa sociedade, não se dá entre pessoas iguais: a disputa pelo espaço tem uma base ideológica e política, depende de como as pessoas se veem na sociedade e de seu acesso real ao poder.

Levando em consideração as referências supracitadas, o termo acidente utilizado neste trabalho deve ser interpretado como uma ocorrência em uma via de trânsito que envolva ao menos um veículo, motorizado ou não, que tenha colidido com outro veículo, ou pessoa, ou animal presente na via, e que tenha gerado danos físicos ou materiais aos envolvidos.

O número de mortes no trânsito continua a aumentar de forma constante, atingindo 1,35 milhões de pessoas em 2016. No entanto, a taxa de mortalidade a cada 100.000 habitantes em relação ao tamanho populacional mundial tem estabilizado, como é possível ver na Figura 1. Quando considerado o contexto do crescimento populacional e da rápida motorização que ocorreu durante o mesmo período, infere-se que os esforços de segurança rodoviária existentes podem ter atenuado a situação. Entretanto, o progresso para o

*Sustainable Development Goals*³ (SDG) promovido pelas 17 metas globais da Assembleia Geral das Nações Unidas que apela por uma redução de 50% no número de mortes no trânsito contando a partir de 2011 até o ano de 2020 está longe de ser alcançada.

Figura 1 - Número e taxa de mortes no trânsito por 100.000 habitantes: 2000–2016.



Fonte: OMS (2018).

Segundo dados do Sistema de Informação sobre Mortalidade (SIM)⁴. No ano de 2011 42.844 mortes por Acidentes de Trânsito (AT) foram registradas no Brasil, ou seja, para atingir o objetivo da ONU o número de acidentes teria que cair para 21,628. É importante ressaltar que esses números não levam em consideração apenas a base da PRF, mas sim o cruzamento de diversas bases de dados.

2.2 Classificação e tipos de acidentes

Segundo Cardoso (2006) para que seja possível desenvolver modelos de previsão de acidentes, que por sua vez irão auxiliar na tomada de decisão para realização de medidas preventivas é necessário realizar uma classificação em relação aos tipos de acidentes, sendo que essa situação pode ser estendida quanto a gravidade dos acidentes.

³ Tradução Livre: Objetivos do Desenvolvimento Sustentável.

⁴ <http://www.saude.gov.br/noticias/svs/28763-mortes-por-acidentes-caem-11-em-um-ano>. Acesso em: 24 jan. 2020.

De acordo com Coca et al. (2012) a classificação dos acidentes de acordo com seu tipo deve ser capaz de desassociar acidentes que apresentem diferentes características, com objetivo de identificar as causas prováveis do acontecimento de cada tipo, para que no futuro torne-se possível definir ações a serem implementadas para reduzir o número de acidentes.

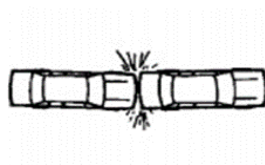
Segundo a NBR 10697:2018 (ABNT, 2018), os acidentes podem ser classificados em colisões, abalroamentos, choques, acidentes com pedestres, tombamentos, capotagens, engavetamentos, quedas e outros. Por sua vez, cada tipo de acidente pode ter sua subdivisão para tornar ainda mais claro o entendimento do ocorrido. A seguir serão citados os tipos de acidentes de acordo com a norma supracitada.

a) Colisões

São acidentes onde um ou mais veículos em movimento sofrem impacto de outros veículos que também se encontram em movimento, embora a palavra “colisão” seja comumente utilizada em casos de acidentes de trânsito (AT) com grande impacto, o uso não se limita a magnitude da força exercida. Segundo Coca et al. (2012) colisões tem como característica principal o envolvimento de pelo menos dois veículos, sendo as colisões divididas em: colisões traseiras, laterais, transversais ou frontais.

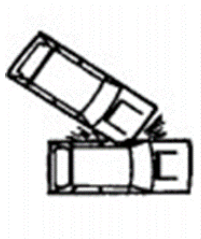
A colisão traseira acontece entre dois veículos em movimento em um mesmo sentido da via, pode acontecer por vários motivos, sendo os mais comuns o uso brusco do freio pelo motorista da frente, ou por ele estar dirigindo em uma baixa velocidade, assim impedindo que o motorista de trás tenha tempo hábil para evitar a colisão, como é ilustrado na Figura 2.

Figura 2 - Colisão traseira



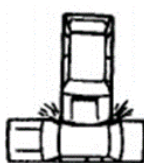
Fonte: Gold (1989, p.19)

Na colisão lateral, como é possível ver na Figura 3, os envolvidos estão trafegando em uma mesma direção, podendo o sentido ser o mesmo, ou não, esse tipo de acidente ocorre quando um dos veículos inicia uma conversão à esquerda ou à direita ou até mesmo em ultrapassagens.

Figura 3 - Colisão lateral

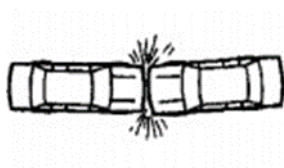
Fonte: Gold (1989, p.19)

Já as colisões transversais acontecem quando a frente de um envolvido colide com a lateral de outro formando algo parecido com ângulo de 90º graus, é comumente observada em cruzamentos viários, saídas de estacionamentos, entre outros. A Figura 5 ilustra uma colisão transversal.

Figura 4 - Colisão transversal

Fonte: Gold (1989, p.20)

A colisão frontal, na maior parte dos casos, causa danos graves tanto para os veículos, como aos seus ocupantes, tendo em vista que os veículos estão se movendo numa mesma direção, porém em sentidos opostos, como mostra a Figura 5. Geralmente ocorre em ultrapassagens quando um dos envolvidos invade a pista destinada ao tráfego do sentido oposto. Esse tipo de colisão é conhecido também como abalroamento.

Figura 5 - Colisão frontal.

Fonte: Gold (1989, p.19)

b) Choques

Diferente das colisões que envolvem pelo menos dois veículos, os choques são definidos como impacto moderado de um veículo em movimento na via de trânsito contra qualquer objeto, fixo ou móvel, mas que não esteja em movimento, podendo esse ser um objeto qualquer ou um veículo, defesa metálica, barreira de concreto, postes, árvores entre outros. A Figura 6 tem o objetivo de ilustrar um choque contra um objeto.

Figura 6 - Choque contra objeto



Fonte: Gold (1998, p.20)

2.3 Knowledge Discovery in Databases, mineração de dados e visualização de dados.

O processo de descoberta do conhecimento *Knowledge-Discovery in Databases* (KDD) é o processo que envolve 5 etapas, sendo elas: coleta dos dados de destino, pré-processamento dos dados, limpeza e execução da mineração de dados para extrair padrões e relacionamentos e em seguida, interpretação e avaliação das estruturas descobertas (HAND, 2001).

Já a mineração de dados é a análise de conjuntos de dados, para encontrar relacionamentos de dados de maneiras novas, que sejam compreensíveis e úteis para o proprietário dos dados. Os relacionamentos derivados de um processo de mineração de dados geralmente são utilizados na concepção de modelos e compreensão de padrões. Exemplos incluem: equações lineares, clusters, gráficos, estruturas em árvore e padrões recorrentes em séries temporais (HAND, 2001).

Sendo assim, o KDD está relacionado ao processo de descoberta de conhecimento de dados e a mineração de dados é a aplicação de técnicas e algoritmos para extrair modelos dos dados estudados.

Segundo Freitas (2000), o conhecimento a ser descoberto deve satisfazer a três propriedades: deve estar o mais correto possível, deve ser compreensível por usuários

humanos e deve ser algo que seja interessante, útil ou novo. O método de descoberta do conhecimento deve apresentar as seguintes três características: deve ser acurado, aplicável a vários tipos de dados e facilmente modificável. Após concluir o processo de obtenção do conhecimento faz-se necessário apresentar essas informações.

De toda forma, a complexidade no processo de KDD está na difícil tarefa de perceber e saber interpretar adequadamente os inúmeros fatos observáveis durante a execução do processo e de julgar corretamente quais ações devem ser tomadas em cada caso (GOLDSCHMIDT, 2003). Para tal tarefa, é designado um analista humano para orientar a execução do processo de KDD.

A visualização de dados emprega canais visuais para representar conjuntos de dados, transformando diferentes tipos de dados em apresentações visuais, de modo que a compreensão e análise dos dados possam ser concluídas de forma eficiente. A vantagem de visualização de dados é a incorporação de recursos humanos em uma interface visual intuitiva, combinando assim, o conhecimento armazenado na máquina com o conhecimento humano (CHEN, 2015).

Existem três campos importantes na visualização de dados, são eles: visualização científica, visualização de informações e análise visual. A visualização científica concentra-se em ilustrar estruturas e evoluções de propriedades físicas ou químicas em um determinado domínio espacial. A visualização de informações concentra-se na representação de abstratos, não estruturados e de alta dimensão de dados, incluindo dados comerciais, e dados em formato de texto em redes sociais. Na análise visual de tráfego o conjunto de dados são geralmente de alta dimensão ou espaço-temporal. Portanto, visualizar dados de tráfego emprega principalmente visualização de informações e análise visual. Essa classificação está resumida na Figura 7.

Figura 7 – Visualização de Dados.



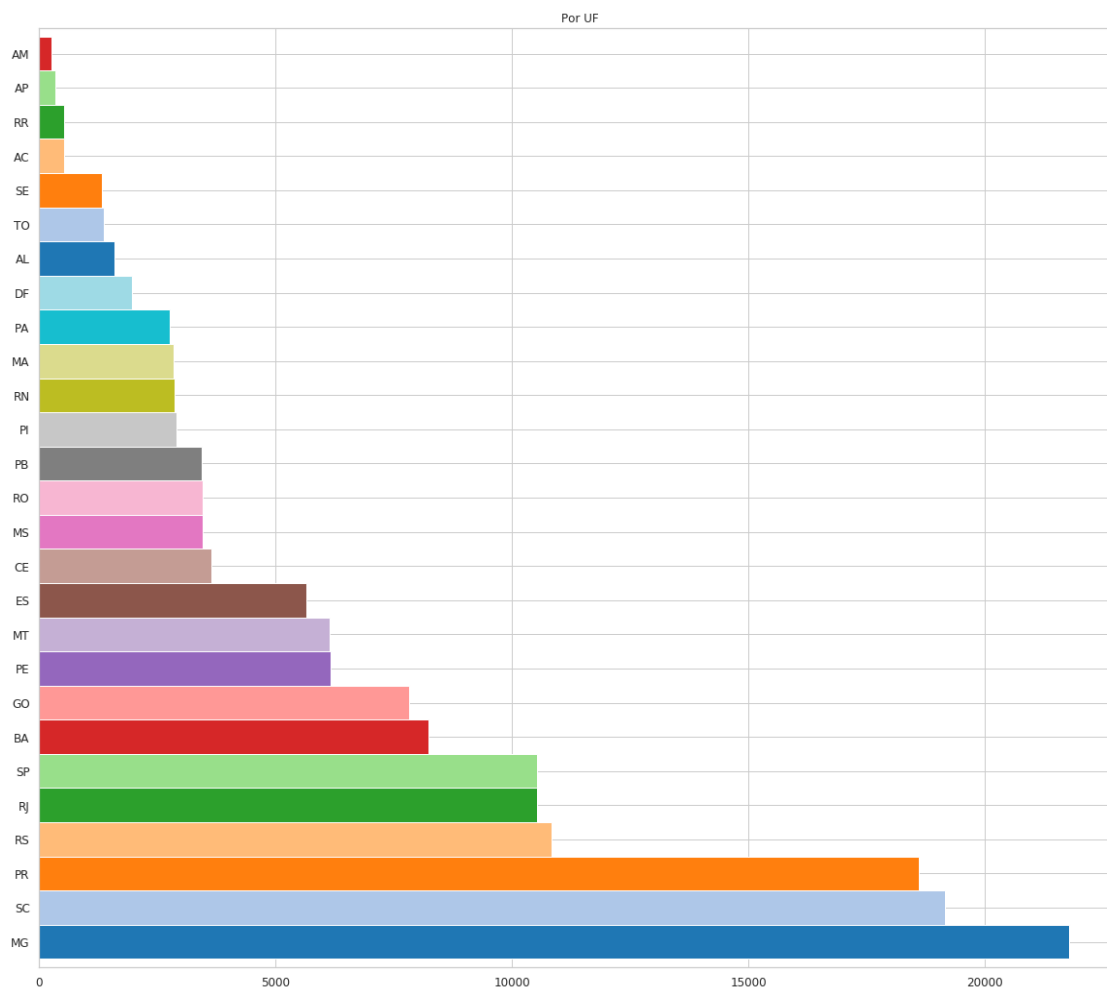
Fonte: Própria (2020).

Em geral, um sistema analítico consiste de quatro componentes principais: coleta de dados, pré-processamento de dados, consulta e análise de dados. Cada componente requer técnicas especializadas de visualização. Outros processos como agregação e cluster também podem ser aprimorados com uma interface visual (CHEN, 2015).

3 TRABALHOS RELACIONADOS

Nessa seção são apresentados dois trabalhos que se relacionam ao objetivo geral proposto. Luís Fred (2019) realizou uma análise exploratória dos dados de ocorrências da PRF no período compreendido entre 2017 e 2018. Entretanto, não foi encontrada uma publicação revisada por pares dos resultados obtidos, ou quaisquer referências acadêmicas da análise feita pelo autor citado. Luís Fred (2019) concluiu que é possível notar pelo gráfico de contagem apresentado na Figura 8, que o estado de Minas Gerais registra o maior número de ocorrências. Com a análise dos dados o autor concluiu que três dos 10 trechos mais perigosos de estradas federais brasileiras atualmente estão em Minas Gerais.

Figura 8 - Número de ocorrências por estado do Brasil.



Fonte: Luís Fred (2019)⁵.

⁵ <<https://www.kaggle.com/luisfredgs/prf-an-lise-explorat-ria-classificador/notebook>>

O trabalho foi disponibilizado na plataforma Kaggle⁶ e está dividido em duas partes, que são elas: análise exploratória e desenvolvimento de um modelo capaz de classificar acidentes. O autor usou técnicas de pré-processamento, limpeza de dados e mineração de dados com o uso do Jupyter Notebook e também fez uso de bibliotecas como: matplotlib, pandas, numpy, scikit-learn entre outras.

O segundo trabalho relacionado trata-se de um desafio de dados vencido pela equipe intitulada “A sociedade dos dados”, promovido pela ferramenta Tableau que apresenta um breve histórico sobre o Papilomavirus Humano (HPV) em sua introdução e dois Dashboards⁷, um para tratar sobre a vacinação por município e outro relatando os desafios para ampliação da vacina.

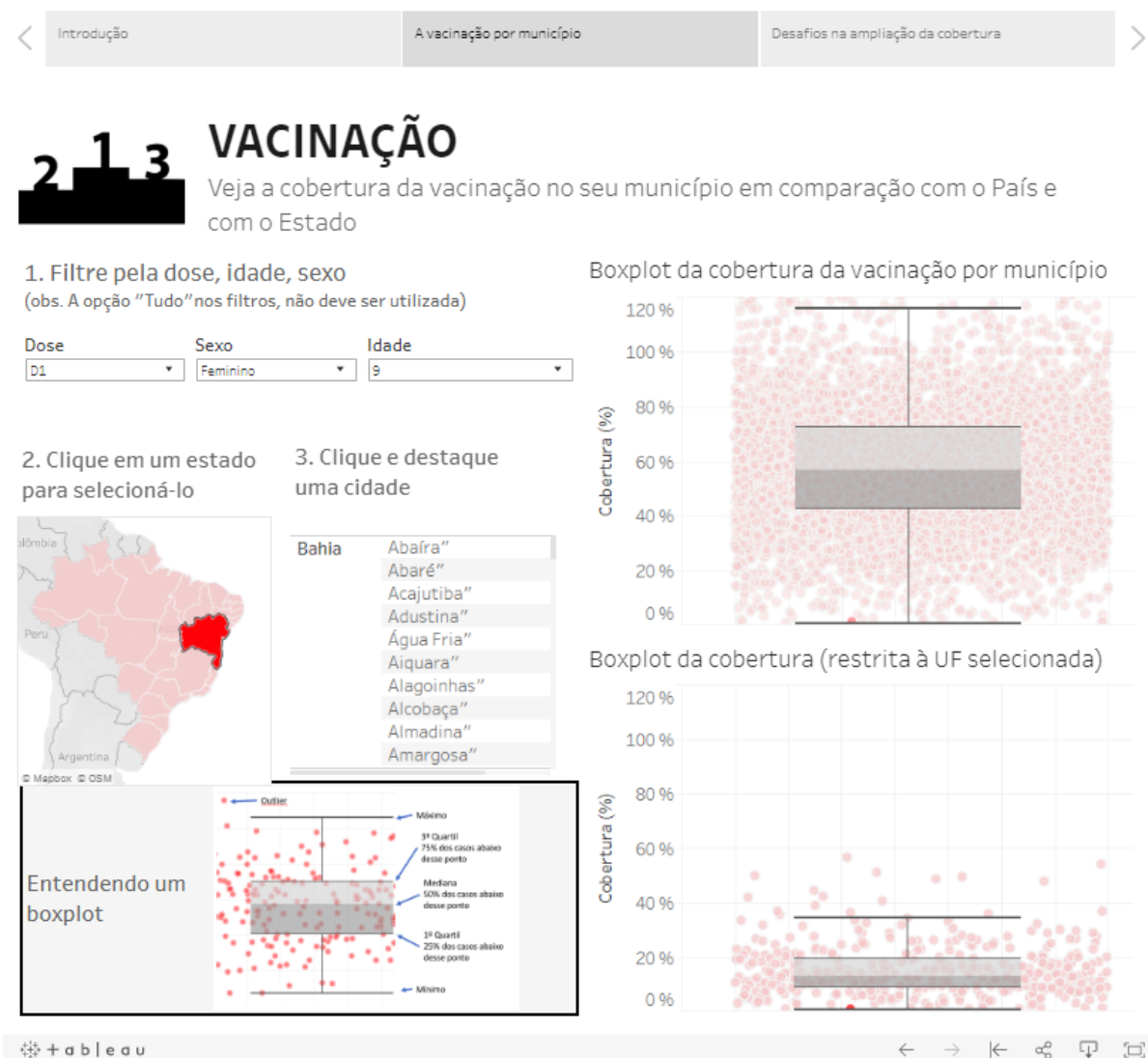
A participante da equipe Indyara Moraes, doutoranda em saúde coletiva pela Universidade de Brasília (UnB) conta que “a motivação do time ao escolher o HPV como tema da análise era fornecer ao gestor do Sistema Único de Saúde (SUS) uma ferramenta com informações atualizadas, concisas e que possam ajudar no planejamento e na prevenção não apenas do vírus, mas de doenças associadas ao HPV. De acordo com Indyara, o vírus, transmitido por via sexual, está relacionado a 99% dos casos de câncer de colo do útero (CAIXETA, 2018).

⁶ Kaggle é uma comunidade on-line de cientistas de dados e aprendizes de linguagem de máquina.

⁷

Figura 9 – Ilustração de vacinação por município

A luta contra o HPV



Fonte: Tableau (2018)⁸

⁸<<https://public.tableau.com/profile/desafio.de.dados5408#!/vizhome/AsociedadedadosdadosAlutacontraoHPV/Histria1>>. Acesso em: 31 jan. 2020.

4 METODOLOGIA

A metodologia divide-se em duas etapas, uma abordagem é a pesquisa bibliográfica apresentada anteriormente. E a outra é o levantamento dos dados de ocorrências coletados pela PRF para realização do processo KDD, com o intuito de transformar os dados brutos em conhecimento de fácil acesso por meio da visualização de dados. Para isso, fez-se necessário o uso de algoritmos de mineração de dados para filtrar e corrigir os dados disponibilizados nas planilhas da PRF.

A pesquisa considerou todas as estradas federais (BR) no território do Rio Grande do Norte, mais especificamente divididos em dois grupos, o primeiro entre os anos de 2007 e 2016 e o segundo, entre 2017 e 2020. Tal divisão ocorreu, pois houve troca de sistemas da PRF: o sistema utilizado para a coleta dos dados entre 2007 e 2016 era o BR-Brasil, substituído por um novo a partir de janeiro de 2017.

A linha temporal da mudança pode ser vista na Figura 10. Tal divisão é importante para o trabalho proposto pois observou-se mudanças nos dados coletados quando houve a mudança de sistema. Como este trabalho tem o foco na década de ações para segurança no trânsito, os dados analisados estão entre os anos de 2011 a 2020. Os dados de 2017 a 2020 estão com a latitude e longitude, por esse motivo, eles foram separados dos demais dados para que fosse feito um mapa de calor das ocorrências de acidentes registradas no Rio Grande do Norte.

Figura 10 - Linha Temporal

BR-Brasil										Novo Sistema			
2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020

Fonte: Própria (2020).

A análise preliminar dos dados coletados foi realizada com um ferramental comumente utilizado na área de mineração de dados. A seguir, detalham-se várias dessas ferramentas, que juntas compõem uma pilha (stack) de ferramentas.

4.1 Coleta dos dados

Os dados abertos de acidentes estão disponíveis no site da PRF⁹ e estão divididos conforme o Anexo B. Esse trabalho foca nos acidentes agrupados por ocorrência, por possuírem dados do ano de 2011 ao ano de 2020.

4.2 Pré-processamento dos dados

Após a etapa de coleta foi dado início a etapa de pré-processamento dos dados no *Google Colab*. O Colab permite o uso da linguagem Python em nuvem. Os arquivos que foram baixados no formato .rar e .zip foram extraídos, após a extração, os arquivos ficaram no formato .csv e foram carregados em *dataframes*¹⁰ com o uso da biblioteca Pandas que vem incluso na plataforma do *Google Colab*. Um exemplo do arquivo disponibilizado pela PRF pode ser visto na Figura 11.

Figura 11 – Ocorrências da PRF do ano de 2019.

Área de Transferência

Fonte

Alimentação

Número

Condicional -

Tablas -

Célula -

Células

POSSÍVEL PERDA DE DADOS

Alguns recursos podem ser perdidos se você salvar esta pasta de trabalho no formato (.csv) delimitada por vírgulas. Para preservar estes recursos, salve-a em um formato de arquivo do Excel.

Não

A1

</

Fonte: Própria (2019).

⁹ <https://portal.prf.gov.br/dados-abertos-acidentes>. Acesso em: 26 jan. 2020.

¹⁰ DataFrames são estruturas de dados tabulares bidimensionais, mutável em tamanho e potencialmente heterôneas, com eixos rotulados (linhas e colunas). <https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html> Acesso em: 26 jan. 2020.

Logo em seguida os arquivos foram lidos pela biblioteca Pandas, dando uma prévia inicial das linhas e colunas que foram trabalhadas no pré-processamento de dados, como pode ser visto na figura 12.

Figura 12 – Dados sendo carregados na biblioteca Pandas.

```
[ ] # Esse são todos os dados agrupados por ocorrência
df_2020 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2020.csv", sep=";", encoding='cp1252', low_memory=False)
df_2019 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2019.csv", sep=";", encoding='cp1252', low_memory=False)
df_2018 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2018.csv", sep=";", encoding='cp1252', low_memory=False)
df_2017 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2017.csv", sep=";", encoding='cp1252', low_memory=False)
df_2016 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2016.csv", sep=";", encoding='cp1252', low_memory=False)
df_2015 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2015.csv", sep=";", encoding='cp1252', low_memory=False)
df_2014 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2014.csv", sep=";", encoding='cp1252', low_memory=False)
df_2013 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2013.csv", sep=";", encoding='cp1252', low_memory=False)
df_2012 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2012.csv", sep=";", encoding='cp1252', low_memory=False)
df_2011 = pd.read_csv("/content/drive/My Drive/TCC/Joseph/TCC1 - Joseph/Code/input/agrupados_por_ocorrendia/datatran2011.csv", sep=";", encoding='cp1252', low_memory=False)
```

df_2020.head()

	id	data_inversa	dia_semana	horario	uf	br	km	municipio	causa_acidente	tipo_acidente	classificacao_acidente	fase_dia	sentido_via	condicao_meteorologica	ti
0	260031.0	2020-01-01	quarta-feira	01:00:00	TO	153.0	678,1	GURUPI	Animais na Pista	Atropelamento de Animal	Com Vítimas Feridas	Plena Noite	Crescente	Nublado	
1	260036.0	2020-01-01	quarta-feira	01:00:00	RJ	116.0	178	NOVA IGUAÇU	Falta de Atenção do Pedestre	Atropelamento de Pedestre	Com Vítimas Fatais	Plena Noite	Decrescente	Céu Claro	
2	260037.0	2020-01-01	quarta-feira	01:52:00	SC	101.0	206,9	SAO JOSE	Ingestão de Alcool	Colisão lateral	Com Vítimas Feridas	Plena Noite	Crescente	Nublado	
3	260038.0	2020-01-01	quarta-feira	01:15:00	RO	364.0	236	CACOAL	Velocidade Incompatível	Saída de leito carroçável	Com Vítimas Feridas	Plena Noite	Decrescente	Nublado	
4	260039.0	2020-01-01	quarta-feira	01:30:00	PR	153.0	360	REBOUCAS	Ingestão de Alcool	Capotamento	Com Vítimas Feridas	Plena Noite	Crescente	Garoa/Chuvisco	

Fonte: Própria (2021).

Após ter uma noção das colunas que serão trabalhadas começa a etapa de limpeza dos dados, essa é a parte mais demorada do processo de extração de conhecimento. Para facilitar o processo de limpeza foi feito um merge (fusão de dados) das informações para transformar os dez Dataframes em um só, como é possível ver na figura 13.

Figura 13 – Merge dos dados.

```
#Merge dos dados agrupados por ocorrência em um novo arquivo csv
# Com todos os anos
df = pd.concat([df_2020,df_2019,df_2018,df_2017,df_2016, df_2015, df_2014,
                df_2013, df_2012, df_2011], ignore_index=True)

df = df.to_csv('/content/drive/My Drive/2011_2020.csv', sep=";", encoding='cp1252', index=False)
```

Fonte: Própria (2021).

Após o merge foi necessário eliminar os dados pertencentes aos outros estados, já que o foco do trabalho são as ocorrências no estado do Rio Grande do Norte. Para isso, foi feito

uma função na linguagem Python para que esses dados fossem filtrados, como pode ser visto na Figura 14.

Figura 14 – Função para filtrar o RN.

```
[ ] def filtra_RN(df):
    RN = df[(df['uf']=='RN')]
    return RN

[ ] filtra_RN(df).to_csv('/content/drive/MyDrive/TCC/TCC_Joseph/TCC1 - Joseph/DadosPRF/Dados do RN/rn2011_2020.csv', sep=";", encoding='cp1252', index=None)
```

Fonte: Própria (2021).

O próximo passo foi instalar a biblioteca pandas-profiling, essa biblioteca é usada para dar uma melhor visão dos dados que estão sendo trabalhados, algumas das principais informações que o Pandas Profiling fornece são o tamanho do dataset (MB,GB), quantidade de linhas duplicadas, quantidade de linhas do Dataframe, quantidade de linhas por colunas, bem como o tipo de dado de cada coluna; e quais colunas contém valores faltantes.

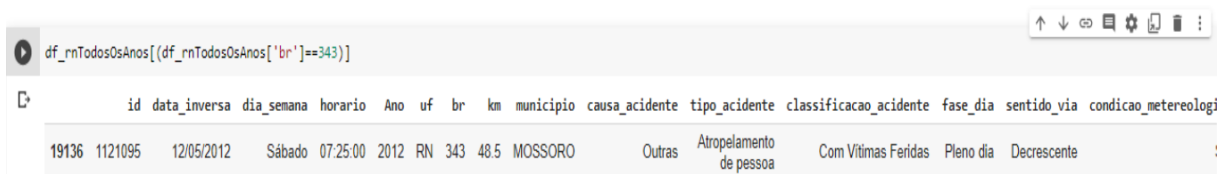
Após analisar as informações com a biblioteca, notou-se colunas com dados faltantes, conforme apresentado na Figura 15.

Como foi relatado anteriormente, houve uma mudança no sistema da PRF. O sistema BR-Brasil continha os campos, mas aparentemente por algum motivo nenhum dos dados foi coletado, por isso, existe essa grande quantidade de dados faltantes. Os dados do campo Br e Km foram adicionados por aproximação com outros dados semelhantes. Outros problemas também foram encontrados, como por exemplo na Figura 15 que mostra um dado com a Br 343, porém não existe essa BR no município de Mossoró. Todos os dados que tinha esse problema foram atualizados.

Figura 15 – Soma de dados faltantes


The screenshot shows a Jupyter Notebook interface with a search bar, expand/collapse icons, and a file explorer on the left. The main area displays the command `df_rnTodosOsAnos.isnull().sum()` and its output, which is a list of columns and their corresponding counts of missing values.

Column	Count
id	0
data_inversa	0
dia_semana	0
horario	0
Ano	0
uf	0
br	9
km	9
municipio	0
causa_acidente	0
tipo_acidente	0
classificacao_acidente	0
fase_dia	0
sentido_via	0
condicao_meteorologica	0
tipo_pista	0
tracado_via	0
uso_solo	0
peessoas	0
mortos	0
feridos_leves	0
feridos_graves	0
ilesos	0
ignorados	0
feridos	0
veiculos	0
latitude	19748
longitude	19748
regional	19748
delegacia	19748
uop	19748

Fonte: Própria (2021)**Figura 16 – Exemplo de dados fora do padrão**


The screenshot shows a Jupyter Notebook interface with a search bar, expand/collapse icons, and a file explorer on the left. The main area displays the command `df_rnTodosOsAnos[(df_rnTodosOsAnos['br']=='343')]` and its output, which is a single row of data from the dataset.

id	data_inversa	dia_semana	horario	Ano	uf	br	km	municipio	causa_acidente	tipo_acidente	classificacao_acidente	fase_dia	sentido_via	condicao_meteorologica
19136	1121095	12/05/2012	Sábado 07:25:00	2012	RN	343	48.5	MOSSORO	Outras	Atropelamento de pessoa	Com Vítimas Feridas	Pleno dia	Decrescente	

Fonte: Própria (2021)

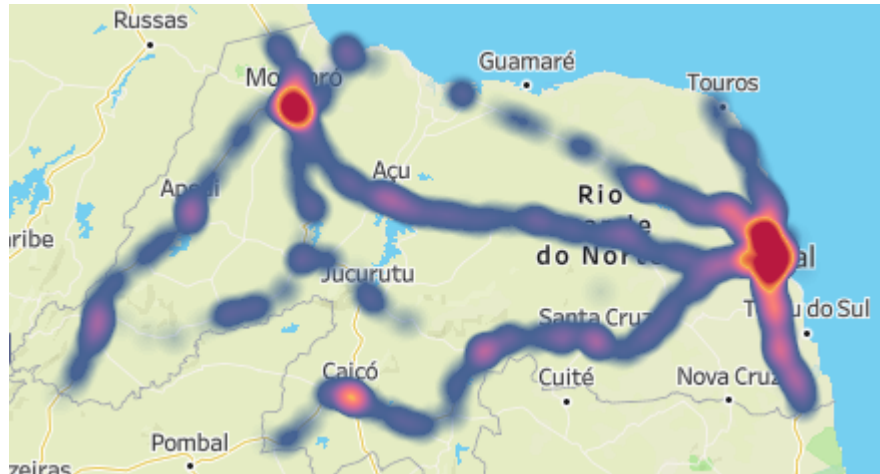
5 RESULTADOS E DISCUSSÃO

Segundo o IBGE¹¹, a população estimada do estado do Rio Grande do Norte em 2020 é de 3,534,165 milhões de pessoas e o total de veículos informados no ano de 2018 foi 1,290,903 milhões. A sua área territorial é composta por 167 municípios. Segundo (CHAGAS, 2015), em termos de áreas rodoviárias estaduais, o estado do Rio Grande do Norte conta com 2.863 km. Dessa forma, são nove os principais eixos rodoviários federais que cruzam o estado:

- **BR-101:** liga o Rio Grande do Norte à Paraíba, tendo o seu marco zero em Touros, no Litoral Norte e passando pelo porto de Natal;
- **BR-104:** liga Macau-RN até Maceió, em Alagoas;
- **BR-110:** liga Areia Branca—RN até Catu, na Bahia;
- **BR-226:** liga Natal-RN à cidade de Paraíso do Tocantins, em Tocantins;
- **BR-304:** começa em Beberibe no Ceará, passando por Natal e cruzando Zona Mossoroense;
- **BR-405:** vem da Paraíba para conectar Pau dos Ferros e Mossoró, permitindo a ligação entre o Alto Apodi e as demais zonas homogêneas do Estado;
- **BR-406:** conecta Natal a Macau, cidade relevante pela produção de sal;
- **BR-427:** tem seu início em Currais Novos, passando por Caicó para adentrar o interior da Paraíba;
- **BR-437:** inicia em Chapada do Apodi-RN e segue até o Vale do Jaguaribe, no estado do Ceará.

O mapa de calor apresentado na Figura 14 apresenta o ponto de maior intensidade no município de Natal na Rodovia Governador Mário Covas uma rodovia duplicada, próximo ao Complexo Viário Trampolim da Vitória que dá acesso a BR-304 sentido Macaíba, BR-101 sentido Natal e a BR-101 sentido Parnamirim. O segundo ponto de calor com maior intensidade fica no município de Mossoró onde três pontos se destacam, são eles: Avenida Presidente Dutra, na BR-110, Avenida Wilson Rosado na BR-304 e Rua Felipe Camarão na BR-405. O terceiro ponto de calor fica no município de Caicó, na Avenida Coronel Martiniano BR-427. E o último ponto de calor começa no município de Apodi e segue a BR-405 até o município de Pau dos Ferros.

¹¹ Instituto Brasileiro de Geografia e Estatística. **Rio Grande do Norte**. Disponível em: <https://cidades.ibge.gov.br/brasil/rn/panorama>. Acesso em: 17 maio 2021.

Figura 17 – Mapa de calor de 2017 a 2020.

Fonte: Própria (2021).

5.1 Acidentes no município de Natal nos últimos 10 anos

O maior número de ocorrências é de 5.625, que corresponde a 22,16% dos acidentes, dos quais 2.827 acidentes foram colisões traseiras e 1.400 foram colisões laterais, aconteceram na capital do estado do Rio Grande do Norte, Natal, especificamente na BR-101 conforme é apresentado na Figura 17.

Figura 18 – Classificação de acidentes no município de Natal.

Acidentes no município de Natal

	Br			
Classificacao Acidente	101	226	304	406
Com Vítimas Fatais	78	24	1	
Com Vítimas Feridas	1.802	365	6	17
Ignorado	18	2		
Sem Vítimas	3.727	248	7	13

Fonte: Própria (2021).

A BR-101 é a segunda rodovia mais longa do Brasil, com 4.658 quilômetros partilhados por 12 estados, estando diretamente conectada aos estados do Rio Grande do Norte, em torno de 194 quilômetros, tendo o seu marco zero em Touros, seguindo por Ceará-mirim, município de Parnamirim, Canguaretama, Rio do Fogo, São Gonçalo do

Amarante, São José de Mipibu, Maxaranguape, Natal e Goianinha, finalmente fazendo fronteira com a Paraíba.

Em sua maioria, os acidentes registrados em Natal, na BR-101 não possuem vítimas. No total o número de ocorrências sem vítimas feridas é de 3.727 mil. Logo em seguida, verificam-se 1.802 mil contingências com vítimas feridas. O número de vítimas fatais é de 78. Por fim temos 18 casos que foram ignorados, para os quais não há informação sobre o estado físico das pessoas envolvidas na ocorrência.

Na BR-226, por sua vez, os dados indicam que 365 acidentes resultaram em vítimas feridas. Os números de ocorrências sem vítimas foram de 248. O número de ignorados foi 2 e as vítimas fatais foram 24.

A BR-304 não apresentou nenhum acidente ignorado, houve uma vítima fatal, 6 vítimas feridas e 7 acidentes sem vítima. Por fim, a BR-406 apresenta 17 ocorrências com vítimas feridas e 13 acidentes sem vítimas. A Figura 15 ilustra o que foi citado anteriormente.

O Departamento Nacional de Estradas de Rodagem (DNER), atualmente conhecido como DNIT, apresenta algumas alternativas de custo reduzido para prevenção de acidentes. De acordo com o DNER (1998, p. 49) possíveis soluções de baixo custo indicadas são:

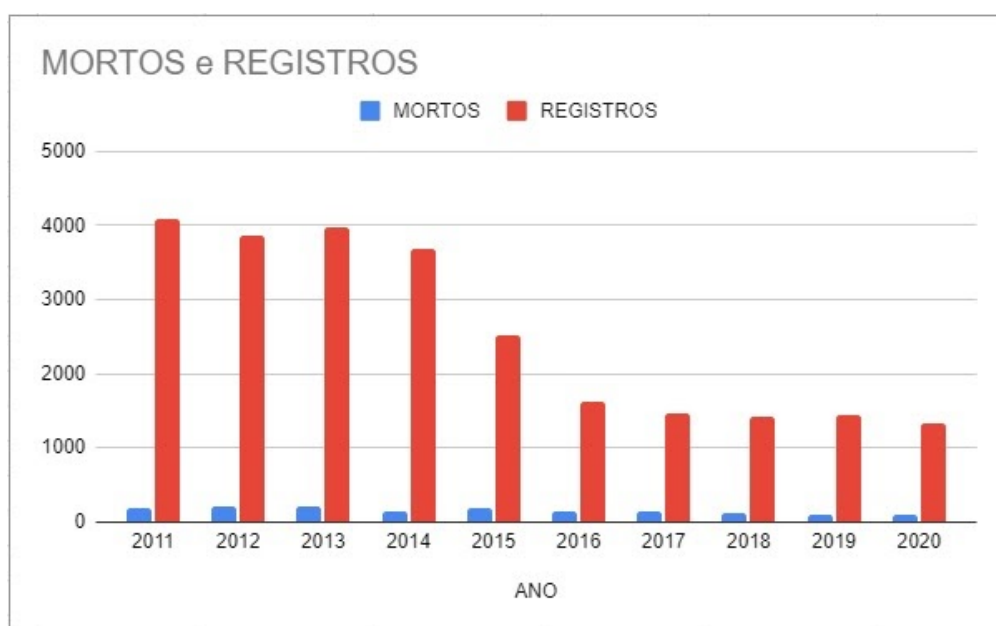
- Orientar os condutores quanto à posição adequada do veículo na pista, locais de ingresso de veículos na pista ou inadequados para ultrapassagem.
- Advertir os usuários de perigos adiantes, isto é, preveni-los da existência de curvas, pista derrapante, declives, entre outras ocorrências.
- Impedir que os veículos sofram quedas em precipícios ou choques contra barrancos.
- Suprir deficiências de motoristas inexperientes que perderam o ponto de frenagem e de veículos sem manutenção, com pneus gastos.
- Orientar o fluxo de veículos para evitar manobras perigosas e tráfego em local impróprio
- Evitar conflito entre veículos e pedestres que trafegam pelo acostamento, ou mesmo pela via, em áreas urbanas que apresentem concentração de pedestres e orientar os pedestres para uma travessia segura em local próprio e sinalizado.
- Evitar conflitos entre pedestres e veículos em áreas de concentração de pedestres (áreas urbanas).
- Dar melhores condições aos motoristas de enxergarem os pedestres na via e outros veículos.

- Reduzir a velocidade dos veículos próximos a áreas urbanas, onde aumenta o fluxo de pedestres e de veículos cruzando a pista.
- Alertar os motoristas de que se deve reduzir a velocidade ao se aproximarem de um ponto que ofereça algum tipo de perigo.
- Evitar conflito entre veículos de porte e veículos leves de propulsão humana em áreas urbanas onde há concentração de bicicletas.
- Levar as paradas de ônibus para fora da via, evitando-se, assim, entre outros acidentes, as colisões traseiras, e oferecer condições seguras de acesso e desembarque.
- Alertar o motorista de que existe, à frente, uma situação atípica em que se inclui a redução de pista e canalização do tráfego.

6 CONSIDERAÇÕES FINAIS

Os altos índices de acidentes de trânsito é um tema frequentemente debatido em agendas mundiais ao longo dos anos. Como medida preventiva, resultante desses encontros, podemos citar a década de ação para segurança de trânsito 2011-2020, que tem como meta a redução de 50% de acidentes ocasionados por veículos. O relatório de status global sobre segurança no trânsito (*Global status report on road safety 2018*) aponta que no ano de 2016, 1,35 milhões de pessoas perderam suas vidas em acidentes de trânsito ao redor do mundo, tornando este tema de relevância global. Este estudo se propôs a realizar uma análise das ocorrências da PRF nas principais rodovias federais do estado do Rio Grande do Norte para saber se com base nesses dados houve uma diminuição de 50% no número de mortes.

Figura 19 - Número de ocorrências e óbitos dos últimos 10 anos no estado do RN



Fonte: Própria (2021)

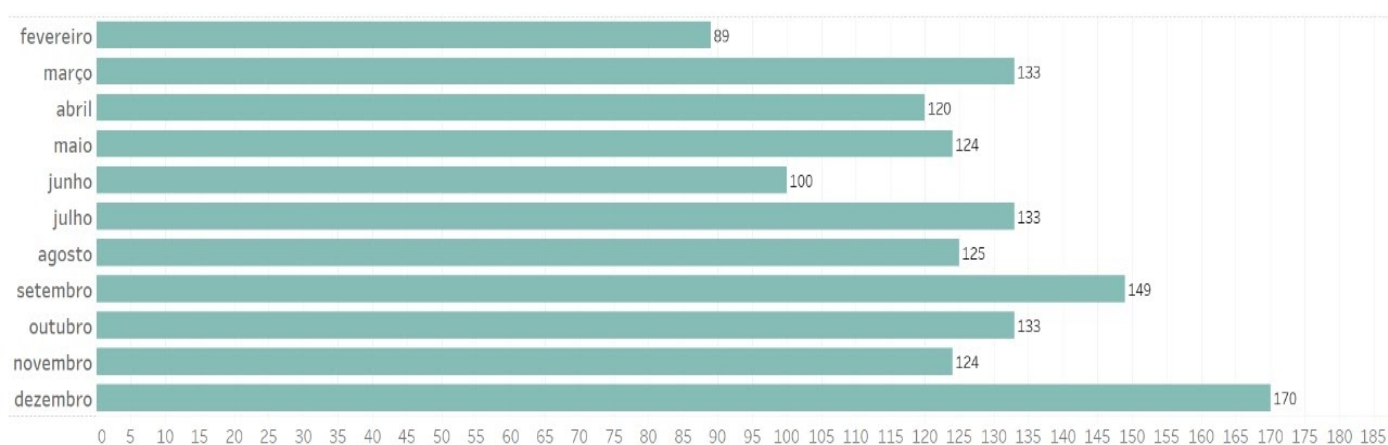
O gráfico acima apresenta os dados coletados relativos a acidentes entre os anos de 2011 a 2020. No ano de 2011, ano base dessa análise, a PRF coletou 4.083 mil registros de ocorrências e o número de mortes foi de 183 pessoas. Em 2012 o número de registros caiu para 3.868 mil e o número de mortes subiu para 207. No ano de 2013 o número de registros aumentou para 3.975 mil registros de ocorrência o número de mortes foi de 208. O ano de 2014 apresentou uma redução para 3.675 mil registros e o número de mortes caiu para 150. No ano de 2015 foi registrado uma redução nos registros de ocorrência para 2.521 mil e o

número de mortes voltou a subir tendo 187 registros. Em 2016 o número de registros caiu para 1.626 e o número de mortes caiu para 144. No ano de 2017 houve uma queda para 1.414 mil registros de ocorrência e o número de mortes caiu para 139. Em 2018 o número de registros teve uma pequena queda para 1.414 registros e o número de mortes caiu para 118. Em 2019 foi o ano que o estado do Rio Grande do Norte atingiu a meta da década de ação para segurança no trânsito, o número de mortes chegou a 90 que corresponde a uma redução de aproximadamente 50,8% comparado ao ano de 2011 que apresentou 183 mortes, o número de registros foi de 1.437. O ano de 2020 apresenta um número de 100 mortes e o menor número de registros dos 10 anos, totalizando 1.333 mil registros.

Com base na análise dos dados da PRF, é possível afirmar que o único ano onde o estado do Rio Grande do Norte esteve dentro da meta para redução de 50% no número de mortes decorrentes de acidentes de trânsito foi no ano de 2019, onde foram registradas 90 mortes, que representa uma queda de aproximadamente 50,8% no número de mortes, logo em seguida, temos o ano de 2020 com uma tendência de alta apresentando 100 mortes o que deixaria o estado com aproximadamente 45,4% de redução no número de mortes. Por outro lado, o número de registro de acidentes que estava em 4.083 ocorrências caiu para 1.333 o que corresponde uma diminuição de aproximadamente 67,4% no número de ocorrências.

A partir do levantamento e estudo dos dados disponíveis, foi possível constatar que há um crescente número de acidentes de trânsito com vítimas fatais em tempos específicos do dia e do ano, como todo, o período do ano onde esse acréscimo se torna mais evidente é o mês de dezembro, como apresentado na Figura 20 adiante, acredita-se que pelo aumento de viagens realizadas devido as festas de fim de ano.

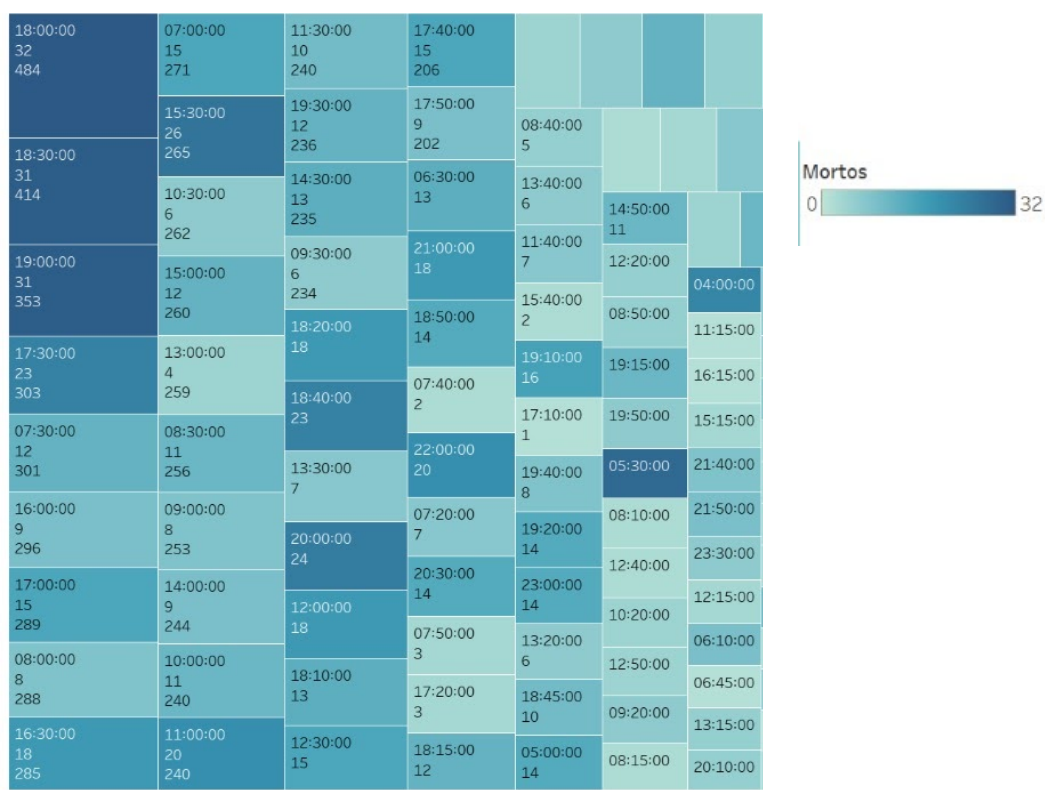
Figura 20 - Meses do ano em que ocorrem o maior número de acidentes



Fonte: Própria (2021).

Relacionado ao espaço do dia em que as ocorrências são registradas em maior número destacam-se os horários do fim da tarde, início da noite, sendo o ponto central as 18:00h, geralmente associado ao trajeto de saída dos trabalhadores do seu local de atividade profissional em direção a sua residência, conforme manifestado no gráfico a seguir:

Figura 21 - Horários em que ocorrem o maior número de acidentes



Fonte: Própria (2021).

A partir da percepção acerca do crescimento do número de ocorrências vinculadas a períodos específicos do dia e do ano. Propõe-se que haja uma maior fiscalização por parte dos poderes responsáveis, a fim de desenvolver e aplicar medidas de orientação aos motoristas, relacionadas a questões pontuais como o excesso de velocidade, do fluxo de trânsito superior, da importância de se ter uma maior atenção durante a viagem, também quanto ao uso do cinto de segurança , a não utilização do telefone celular por parte do motorista e especialmente a não ingestão de bebidas alcoólicas.

Com objetivo de melhoria futura das análises de dados disponibilizados pela PRF, são propostas as seguintes atividades: incluir a latitude e longitude do ano de 2007 ao de 2016,

período em que o sistema utilizado era o BR-Brasil. Além de cruzar os dados com bases relacionadas ao sistema DATASUS, visto que as ocorrências da PRF só registram óbitos ocorridos no local do incidente.

REFERÊNCIAS

Aggarwal, Charu C. (2015), Data Mining, Springer Depaire, B., Wets, G., Vanhoof, K., (2008). **Traffic accident segmentation by means of latent class clustering**. Accident Analysis and Prevention, 40, 1257–1266.

Associação Brasileira de Normas Técnicas (ABNT). **NBR 10697: Pesquisa de acidentes de trânsito**. Rio de Janeiro, 2018. 10 p.

BRASIL. Constituição (1997). Lei nº 9503, de 23 de setembro de 1997. **Código de Trânsito Brasileiro**. Brasília, DF.

BRASIL. Constituição (2011). Lei nº 12527, de 18 de novembro de 2011. **Acesso A Informação**. Brasília, DF.

CAIXETA, Fernando. **Análise sobre vírus HPV vence primeiro Datathon de saúde pública**. 2018. Disponível em: <https://www.metropoles.com/distrito-federal/ciencia-e-tecnologia-df/analise-sobre-virus-hpv-vence-primeiro-datathon-de-saude-publica>. Acesso em: 31 jan. 2020.

CARDOSO, G. **Modelos para previsão de acidentes de trânsito em vias arteriais urbanas**. 2006. 195 f. Tese (Doutorado em Engenharia de Produção) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

CHAGAS, Carlos Magno Rodrigues das. **Distribuição das rodovias federais no Rio Grande do Norte sob a luz do princípio dos transportes da teoria dos lugares centrais**. 2015. Trabalho de Conclusão de Curso. Universidade Federal do Rio Grande do Norte.

Chen, W. and Jovanis P. (2000). **Method for identifying factors contributing to driver-injury severity in traffic crashes**, Transportation Research Record 1717 1-9
CHEN, Wei; GUO, Fangzhou; WANG, Fei-Yue. **A survey of traffic data visualization**. IEEE Transactions on Intelligent Transportation Systems, v. 16, n. 6, p. 2970-2984, 2015.

COCA, A. C. P.; RAIA JUNIOR, A. A.; BEZERRA, B. S.; BASTOS, J. T.; SILVA, K. C. R. **Segurança Viária**. 1. ed. São Carlos: Suprema Gráfica e Editora, 2012. 322 p.

da Costa Medeiros, Wilma Maria, et al. **"Perfil epidemiológico das vítimas de acidentes de trânsito atendidas num serviço público de emergência da região metropolitana de Natal/RN."** HOLOS 7 (2017): 213-224.

DEPARTAMENTO NACIONAL DE ESTRADAS DE RODAGEM - (DNER). **Guia de redução de acidentes com base em medidas de engenharia de baixo custo**. Rio de Janeiro, 1998. Disponível em:

<http://ipr.dnit.gov.br/normasemanuais/manuais/documentos/703_guia_de_reducao_de_acidentes.pdf>. Acesso em: 21 fev. de 2018.

FREITAS, A. A. **Uma Introdução a Data Mining**. Informática Brasileira em Análise, Centro de Estudos e Sistemas Avançados do Recife (C.E.S.A.R.), Recife, Pe, ano II, n. 32, 2000.

GOLD, P. A. **Segurança de Trânsito: Aplicações de Engenharia para Reduzir Acidentes**. 1 ed. Nova York. Banco Interamericano de Desenvolvimento, 1998. 211 p.

GOLDSCHMIDT, Ronaldo. Assistência Inteligente à orientação do processo de descoberta de conhecimento em bases de dados. **Rio de Janeiro, Tese de Doutorado–Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro**, 2003.

HAND, David J.; MANNILA, Heikki; SMYTH, Padhraic. **Principles of data mining** (adaptive computation and machine learning). MIT Press, 2001.

ISOTANI, Seiji; BITTENCOURT, Ig Ibert. **Dados Abertos Conectados: Em busca da Web do Conhecimento**. Novatec Editora, 2015.

MCKINNEY, Wes. **Python for data analysis: data wrangling with Pandas, NumPy, and IPython**. 2017.

Organização Mundial da Saúde (OMS). **Sistemas de dados: um manual de segurança viária para gestores e profissionais da área**. Brasília, D. F.: Organização Pan-Americana da Saúde; 2012. Disponível em:

<http://apps.who.int/iris/bitstream/10665/44256/3/9789275717110_por.pdf?ua=1&ua=1>.

Acesso em: 28 jan 2020.

VASCONCELOS, Eduardo A. **O que é trânsito**. Brasiliense, 2017.

ANEXO A – SITE (DADOS.GOV.BR)

The screenshot displays the website dados.gov.br/dataset/acidentes-rodovias-federais. The browser's address bar shows the URL and a 'Não seguro' (Not secure) warning. The page layout includes a left sidebar with social media links (Facebook, Twitter) and a 'Licença' (License) section for Open Data Commons Open Database License (ODbL). The main content area is titled 'Dados e recursos' (Data and resources) and lists various datasets with corresponding file icons (PDF, CSV) and 'Explorar' (Explore) buttons.

Formato	Nome do Recurso	Ação
PDF	Modelo Entidade Relacionamento	Explorar
PDF	Dicionário de Dados	Explorar
CSV	Ocorrências 1º semestre de 2007	Explorar
CSV	Ocorrências 2º semestre de 2007	Explorar
CSV	Ocorrências 1º semestre de 2008	Explorar
CSV	Ocorrências 2º semestre de 2008	Explorar
CSV	Ocorrências 1º semestre de 2009	Explorar
CSV	Ocorrências 2º semestre de 2009	Explorar
CSV	Ocorrências 1º semestre de 2010	Explorar
CSV	Ocorrências 2º semestre de 2010	Explorar
CSV	Ocorrências 1º semestre de 2011	Explorar
CSV	Ocorrências 2º semestre de 2011	Explorar
CSV	Ocorrências 1º semestre de 2012	Explorar
CSV	Ocorrências 2º semestre de 2012	Explorar
CSV	Ocorrências 1º semestre de 2013	Explorar
CSV	Tabelas de domínio	Explorar
CSV	Tabelas de pessoas e veículos	Explorar

Não seguro | dados.gov.br/dataset/acidentes-rodovias-federais

Apps Assista os melhores... Fóruns - Fórum Scri... Apostilando.com... Cursos de TI Instawload: Downlo... The Huxley Faça o download d... GUJ e-Aulas da USP :: B... Escavador - Seu ass... Outros favoritos

Tabelas de domínio

Explorar

Tabelas de pessoas e veículos

Explorar

Informações Adicionais

Campo	Valor
Autor	Departamento de Polícia Rodoviária Federal
Mantenedor	Ministério da Justiça
Última Atualização	30 de Junho de 2017, 00:40 (UTC-03:00)
Criado	30 de Agosto de 2013, 17:58 (UTC-03:00)
Cobertura Temporal	2007 a 2014
Granularidade Geográfica	Municipal
Periodicidade	Semestral
VCGE	Transporte Rodoviário [http://vocab.e.gov.br/2011/03/vcge#transporte-rodoviario], Segurança Pública [http://vocab.e.gov.br/2011/03/vcge#seguranca-publica]

O Portal

- ▶ Catálogo
- ▶ Organizações
- ▶ Catalogue seus dados
- ▶ Mapa do site
- ▶ Sobre o portal
- ▶ Perguntas frequentes
- ▶ Desenvolvimento do portal
- ▶ Agradecimentos

Comunidade

- ▶ Comunidades
- ▶ Twitter
- ▶ Identificação
- ▶ Lista de discussão
- ▶ Ferramentas de colaboração
- ▶ Últimas alterações
- ▶ Sugira dados
- ▶ Agenda
- ▶ Sugestões

Manuais

- ▶ Cartilha de publicação de dados
- ▶ Arquitetura Técnica Referencial
- ▶ Manual de Catalogação
- ▶ Manual de Dados Abertos [W3C](#)
- ▶ Manual do desenvolvedor [W3C](#)
- ▶ Cartilha do desenvolvedor [W3C](#)
- ▶ Guia de Dados Abertos [OK!](#)
- ▶ Manual para Elaboração de Planos de Dados Abertos

Saiba mais

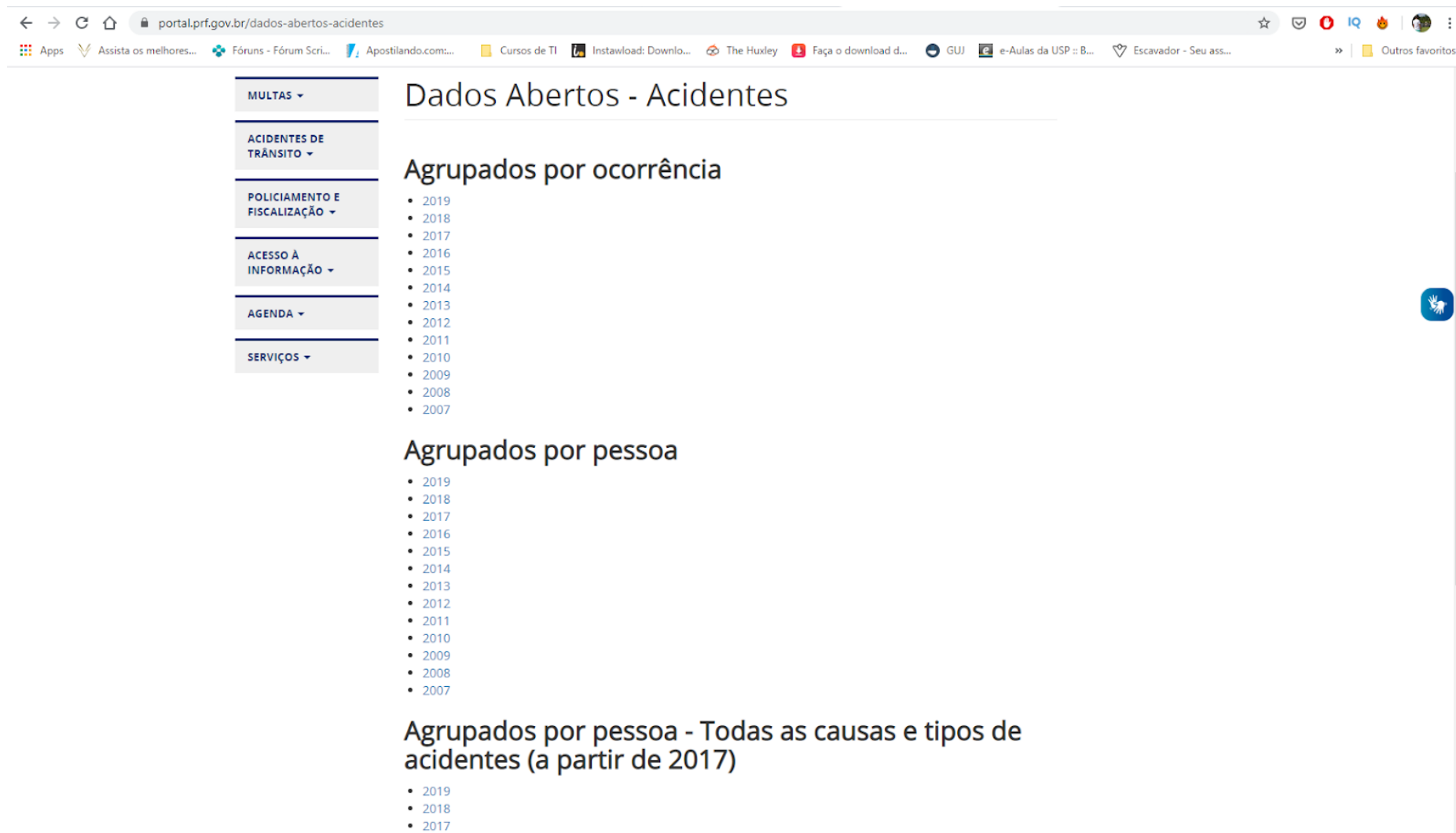
- ▶ O que são dados abertos?
- ▶ Lei de acesso a informação
- ▶ Wiki da [INDA](#)
- ▶ Instrução Normativa da [INDA](#)
- ▶ Parceria para Governo Aberto
- ▶ Outras iniciativas
- ▶ Indicadores de governo
- ▶ Painel de Monitoramento

Desenvolvido pela [Secretaria de Tecnologia da Informação, Ministério do Planejamento, Desenvolvimento e Gestão](#)

Todo conteúdo licenciado sob uma Licença [Creative Commons Attribution-ShareAlike 3.0 Unported](#)

Logotipo Licenciado sob [Licença Pública de Marca](#)

ANEXO B – SITE (PORTAL.PRF.GOV.BR)



portal.prf.gov.br/dados-abertos-acidentes

Apps Assista os melhores... Fóruns - Fórum Scri... Apostilando.com... Cursos de TI Instawload: Downlo... The Huxley Faça o download d... GUJ e-Aulas da USP :: B... Escavador - Seu ass... Outros favoritos

MULTAS ▾

ACIDENTES DE TRÂNSITO ▾

POLICIAMENTO E FISCALIZAÇÃO ▾

ACESSO À INFORMAÇÃO ▾

AGENDA ▾

SERVIÇOS ▾

Dados Abertos - Acidentes

Agrupados por ocorrência

- 2019
- 2018
- 2017
- 2016
- 2015
- 2014
- 2013
- 2012
- 2011
- 2010
- 2009
- 2008
- 2007

Agrupados por pessoa

- 2019
- 2018
- 2017
- 2016
- 2015
- 2014
- 2013
- 2012
- 2011
- 2010
- 2009
- 2008
- 2007

Agrupados por pessoa - Todas as causas e tipos de acidentes (a partir de 2017)

- 2019
- 2018
- 2017