

FATORES MAIS RELEVANTES PARA AGRUPAR ACIDENTES DE TRÂNSITO FATAIS.

Jesaías Kendson C. Nunes¹, Miriam Karla Rocha².

Resumo: A maior preocupação dos órgãos responsáveis pelas vias é entender e compreender o que leva a tantas mortes nas rodovias brasileiras. Neste contexto, a partir dos dados da Polícia Rodoviária Federal (PRF 2018) de acidentes de trânsito fatais em zona urbana, este presente trabalho propõe uma análise dos fatores mais relevantes nos acidentes fatais. Para tanto, os dados foram preparados e clusterizados. Esta análise demonstrou que, por exemplo, a causa do acidente, tipo da pista e fase do dia apresentaram uma melhor separabilidade dos acidentes. Conhecer quais os principais fatores que fizeram parte para a ocorrência dos acidentes de trânsito sucedem como um meio que vem ajudar na elaboração de planos e meios eficientes para a redução de acidentes de trânsito fatais.

Palavras-chave: trânsito; acidentes fatais; clusterização.

1. INTRODUÇÃO

De acordo com a Organização mundial da saúde (OMS, 2021), cerca de 1,3 milhão de pessoas morrem a cada ano em decorrência de acidentes de trânsito, onde 93% das fatalidades mundiais ocorrem em países de baixa e média renda, sendo que esses países possuem apenas 60% dos veículos do mundo. O Brasil é um dos países com o número mais acentuado de acidentes de trânsito, isso pela falta de conhecimento sobre as leis das vias, como também dos pedestres que possuem pouco conhecimento sobre o assunto. No estudo de trânsito existe um tripé que é praticamente crucial diante das causas dos acidentes, que são: motoristas, automóveis e as vias. O gerenciamento de dados é indispensável, onde, através dele, o entendimento sobre as condições mais propícias para acidentes pode ajudar no gerenciamento de especialistas de transporte e poder público. Dimensionar os dados mais importantes sobre esse tema, em nível nacional, pode colaborar para o desenvolvimento dessa área de estudo (MARÍN, 2000).

Há muitas pessoas morrendo nas rodovias brasileiras e procura-se uma saída para a diminuição desses números tão expressivos. Excesso de velocidade, ingestão de álcool, uso do celular e outras causas podem culminar em acidentes fatais. Além disso, há risco para vida do condutor, a imprudência do motorista também pode afetar muitas pessoas, dentro do veículo, como também fora do mesmo (ELAINA, 2021). A partir dos dados advindos da polícia Rodoviária Federal (PRF), seria possível distinguir as mais diversas causas, onde os fatores mais relevantes foram analisados, sendo os 3 eixos principais: vias, veículos e motoristas, mas podemos citar outros, como o horário, local do acidente, condições climáticas, como também as leis de trânsito. O desencadeamento desses fatores é primordial para aumentar a probabilidade de acontecer os acidentes.

Para destacar os principais motivos, são levantados os seguintes fatores humanos: trafegar com velocidades que estão acima dos limites máximos e seguros para as vias, realizar travessia, ultrapassagem e trafegar em locais que não permitidos, não atender às prioridades de tráfego entre as vias, não trafegar com a devida atenção necessária, seja por situação física ou psicológica. (POLITEC-MT, 2018).

Dirigir requer muita atenção do motorista, já que qualquer erro ao volante pode levar a graves acidentes. Para tentar evitar acidentes é preciso respeitar as leis de trânsito e ter muita cautela. Em caso de viagens longas, por exemplo, não dirigir de maneira irresponsável, como sob efeito de algum medicamento ou até mesmo álcool ou algo relacionado. Os motoristas não querem se acidentar, mas os acidentes acontecem de maneira rápida e na maioria das vezes a agilidade do condutor ou os recursos de segurança do carro não são suficientes (DETRAN/MS, 2016).

Os bancos de dados disponíveis nos órgãos de trânsito poderiam ser usados como ferramenta para concretizar suas ações e a divulgação dos resultados alcançados para o benefício da segurança no trânsito. Mas, a disponibilidade de dados de acidentes, conforme a regulamentação do Conselho Nacional de Trânsito (CONTRAN) é uma raridade entre a maioria dos órgãos executivos brasileiros. Para os órgãos que ainda não obtêm ou exploram adequadamente os dados de acidentes sugere-se registrar acidentes com um banco de dados próprio e em conjunto com demais órgãos do município (FARIA E FONTES, 2008).

Os acidentes de trânsito são uma série de fatores que mostram uma relação do homem com várias características do ambiente no qual circulam. Há casos que tem consequências mínimas, que são a minoria, mas quando acontece um acidente sempre vem de uma sequência de erros que se desencadeia de muitas formas.

A importância deste trabalho é uma busca pela redução dos acidentes através da implementação de informações importantes a tal conjunto de dados, reconhecendo as causas e analisar os fatores mais relevantes para agrupamento

de acidentes de trânsito fatais, em nível nacional, visando colaborar para o desenvolvimento dessa área de estudo, buscando as principais estatísticas disponíveis.

O reconhecimento dos erros pode ajudar na minimização de alguma forma, podendo usar ações preventivas. A origem das causas pode ser crucial para ver onde pode ter uma melhoria. A segurança no trânsito é imprescindível, por isso não adotar normas que colocam em risco a sua vida, como também dos outros, deve ser objetivo de cada um.

O presente trabalho realiza uma análise dos fatores mais relevantes nos acidentes de trânsito fatais com base nos acidentes ocorridos no Brasil em 2018. Para viabilizar e ajustar a análise multivariada com foco no uso de clusters. Por exemplo, o agrupamento pode ser feito para verificar dados por similaridade, subgrupos em um banco de dados dos acidentes de trânsito ou para identificar áreas onde há um maior número de acidentes com os mesmos informações ou semelhantes.

2. ANÁLISE MULTIVARIADA E DE AGRUPAMENTOS (CLUSTERIZAÇÃO)

Qualquer análise de duas variáveis ou mais seria algo muito complicado há alguns anos, mas com o avanço da tecnologia, especialmente na área computacional, facilitou o processamento de dados em larga escala. Com esse avanço, a análise multivariada cresceu e deu um impulso em diversas áreas, ela que possui diversas técnicas acessíveis e de fácil acesso ajuda de forma surpreendente pesquisas, governos, além da indústria. Com o uso de dados estatísticos, onde o foco é o agrupamento de dados semelhantes, unificando-se em uma análise aquilo que antes seria necessário fazer diversas análises (HAIR, 2005).

De modo histórico, diante de muitas medidas de distância e de semelhança que remontem ao final do século XIX e início do século XX, a análise de agrupamentos, como conjunto de técnicas mais estruturado, teve origem na Antropologia, com Driver e Kroeber (1932), e na Psicologia, com Zubin (1938a e 1938b) e Tryon (1939), conforme discutem Reis (2001) e Fávero et al. (2009). Com o reconhecimento dos procedimentos de aglomeração e classificação de observações como método científico, aliado ao profundo desenvolvimento computacional, verificado principalmente após a década de 1960, a utilização da análise de agrupamentos passa a ser mais frequente após a publicação da relevante obra de Sokal e Sneath (1963), em que são realizados procedimentos para comparar as similaridades (FAVERO, 2017).

A análise de agrupamentos representa um conjunto de técnicas exploratórias muito interessantes e que podem ser usadas quando precisamos verificar os dados que podem possuir comportamentos semelhantes entre observações, como indivíduos, empresas, municípios, países, entre outros exemplos. Em relação a determinadas variáveis e o objetivo de se criarem grupos, ou clusters, em que prevaleça a homogeneidade como um todo. A análise de clusters, foca na sua função principal que é a atribuição de observações em uma quantidade parcialmente pequena de agrupamentos homogêneos internamente e heterogêneos entre si e que representem o comportamento conjunto das observações a partir de determinadas variáveis. Ou seja, as observações de determinado grupo devem ser relativamente semelhantes entre si, em relação às variáveis inseridas na análise, e consideravelmente diferentes das observações de outros grupos.

O pesquisador que realizar uma análise de agrupamentos quando tiver o objetivo de ordenar e alocar as observações em grupos e, a partir de então, estudar qual a quantidade interessante de clusters formados, ou definir a quantidade de grupos que deseja formar, embasado por determinado critério, e fazer um teste como se comportam o ordenamento e a alocação das observações naquela quantidade especificada de grupos. Além de análises que serão feitas posteriormente, com o uso de softwares adequados (FAVERO, 2017).

Ainda, é importante definir os objetivos de pesquisa, escolher determinada medida de semelhança, que servirá de base para que as observações sejam consideradas menos ou mais próximas, e determinar esquema de aglomeração, que deverá ser definido entre os métodos hierárquicos e não hierárquicos. Assim, terá circunstâncias de analisar, interpretar e comparar os resultados. É importante ressaltar que resultados obtidos por meio de esquemas de aglomeração hierárquicos e não hierárquicos podem ser comparados, e, nesse sentido, o pesquisador pode elaborar a técnica, fazendo uso de um ou outro método e se julgar necessário. Enquanto os esquemas hierárquicos permitem a identificação do ordenamento e da alocação das observações, oferecendo possibilidades para que o pesquisador estude, avalie e decida sobre a quantidade de agrupamentos formados. Já nos esquemas não hierárquicos, foca em uma quantidade conhecida de clusters e, diante disso, é feita a alocação das observações nesses clusters, com posterior avaliação da representatividade de cada variável para a formação deles (ver figura 1). Então, o resultado de um método pode agregar e ser usado como input para a realização do outro, tornando uma análise cíclica (FAVERO, 2017).

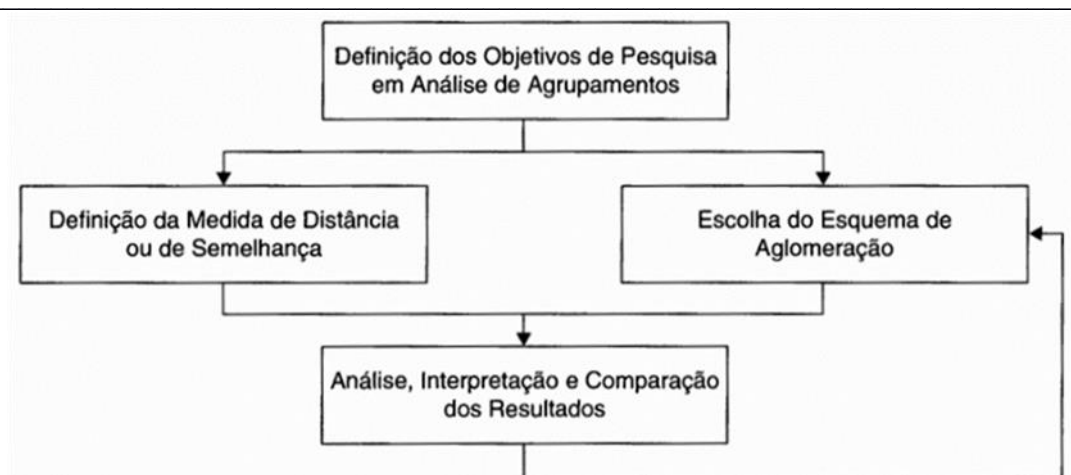


FIGURA 1 - LÓGICA PARA ELABORAÇÃO DA ANÁLISE DE AGRUPAMENTOS (HAIR, 2005)

Conforme discutem Vicini e Souza (2005) e Johnson e Wichern (2007), na análise de agrupamentos, a definição do método de aglomeração, ou grupo de aglomeração é de suma importância diante da definição da medida de semelhança, e essa escolha também precisa ser tomada tomando base tudo que o pesquisador pretende de acordo com os objetivos da pesquisa (FAVERO, 2017).

No passado, o conceito de clusterização de dados somente era conhecido entre os profissionais da ciência da computação, mas isso mudou. É o agrupamento de dados de forma que as informações, no mesmo cluster, sejam semelhantes uns aos outros do que aqueles que estão em outro grupo. Nesse caso, a classificação em clusters é feita usando critérios como distâncias menores, densidade de pontos de dados, gráficos ou várias distribuições estatísticas. Em resumo, podemos dizer que uma empresa pratica clusterização quando agrupa grandes volumes de dados e informações separando-os por similaridade. Dessa forma, é possível armazená-los, processá-los e analisá-los ordenadamente (CORTEX, 2022)

A clusterização permite que a parte funcional das operações aconteça mais rápida, ganhando agilidade e aperfeiçoando o desempenho no dia a dia. É possível utilizar máquinas simples, de fácil manutenção e que não precisam ser iguais. Também é possível utilizar a clusterização para tarefas mais difíceis ou até para algo simples. A sua capacidade pode ser vista com a adição de novos eixos, o que garante equilíbrio nos seus processos. As ferramentas de análise de dados são a base da clusterização dos dados dos acidentes, que cooperam na investigação para conseguir informações sobre o comportamento dos dados como um todo. Com todas as informações necessárias, é possível organizar de forma eficiente uma estratégia, já que o cluster é importante por diversos motivos. Para começar, a potência do processamento que é possível conseguir com esse tipo de sistema, já que ele, de maneira simples, é igual ou superior a computadores com muita potência disponíveis no mercado (MARKETING POR DADOS, 2021).

É de importante que o pesquisador defina, com base na teoria subjacente e em seus objetivos de pesquisa, bem como em sua experiência e intuição, os critérios diante dos quais as observações serão alocadas em cada um dos grupos. A experiência de muitos países tem demonstrado que são necessários dados confiáveis para a introdução de medidas para a redução de acidentes. A coleta de dados de acidentes de trânsito fornece informações essenciais para que cada país possa equacionar os problemas relacionados aos acidentes de trânsito.

3. DESCRIÇÃO DO BANCO DE DADOS

Dados e causas para acidentes de trânsito são essenciais para contribuir na análise e consequentemente trazer melhorias para a segurança viária. Entretanto, o registro de fatores contribuintes requer uma disponibilidade de dados que são essenciais para tal metodologia. A enumeração de fatores contribui e deve ser abrangente, como também, o mais enxuta possível para facilitar o trabalho de registro e análise destas informações. A seleção do método correto para desenvolvimento de uma pesquisa depende do objetivo e posteriormente dos objetivos que o pesquisador deseja atingir

O banco de dados utilizado para análise é proveniente da Polícia Rodoviária Federal (PRF). Estes dados constituem 12 variáveis, que são: região, causa do acidente, condição meteorológica, fase do dia, final de semana, tipo de pista, mês, entre o horário das 22h às 8h, tipo do acidente, traçado da via, pessoas e veículos. Eles foram averiguados e colocados em intervalos corretos de numeração. Este banco de dados conta com informações

referentes a 2018(janeiro a dezembro). Este período é caracterizado por ter variações nas condições meteorológicas, horários e locais do acidente (ver quadro 1).

QUADRO 1 - DESCRIÇÃO DAS VARIÁVEIS

#	CÓDIGO	DESCRIÇÃO
1	Mês	Mês do ano que ocorreu o acidente.
2	Final de semana	Se o acidente ocorreu no final de semana. 1, se sim; 0, caso contrário.
3	Entre o horário das 22h às 8h	Se o acidente ocorreu entre 22h e 8h. 1, se sim; 0, caso contrário.
4	Região	Região brasileira que ocorreu o acidente. Ex.: Norte, sul, sudeste.
5	Causa do acidente	Identificação da causa presumível do acidente. Ex.: Falta de atenção, Velocidade incompatível.
6	Tipo do acidente	Identificação do tipo de acidente. Ex.: Colisão frontal, Saída de pista.
7	Fase do dia	Fase do dia no momento do acidente. Ex. Amanhecer, Pleno dia.
8	Condição meteorológica	Condição meteorológica no momento do acidente: Céu claro, chuva, vento.
9	Tipo pista	Tipo da pista considerando a quantidade de faixas: Dupla, simples ou múltipla.
10	Traçado via	Descrição do traçado da via: reta, curva ou cruzamento.
11	Pessoas	Total de pessoas envolvidas na ocorrência.
12	Veículos	Total de veículos envolvidos na ocorrência.

De acordo com análise utilizada, as doze variáveis serão subdivididas, sendo as variáveis para a análise de clusterização: região e causa, região, causa e e final de semana, tipo de pista e condição meteorológica e tipo de pista e fase do dia.

4. MÉTODO

O banco de dados é composto por 1288 observações de acidentes fatais na área urbana do Brasil em 2018. Como os dados possuem tipos diferentes de variáveis (ver quadro 1), e variáveis com escalas de diferentes formas e magnitudes (km/h, veículos/hora), os dados foram transformados e convertidos de variáveis nominais em variáveis numéricas utilizando a técnica optimal scaling (MEULMAN e HEISER, 2005). Diante disso, foi aplicada a clusterização baseada no método hierárquico, e assim será identificado o número de grupos que serão agrupados por similaridade e a consequente análise de como eles serão distribuídos em cada cluster. (ver figura 2)

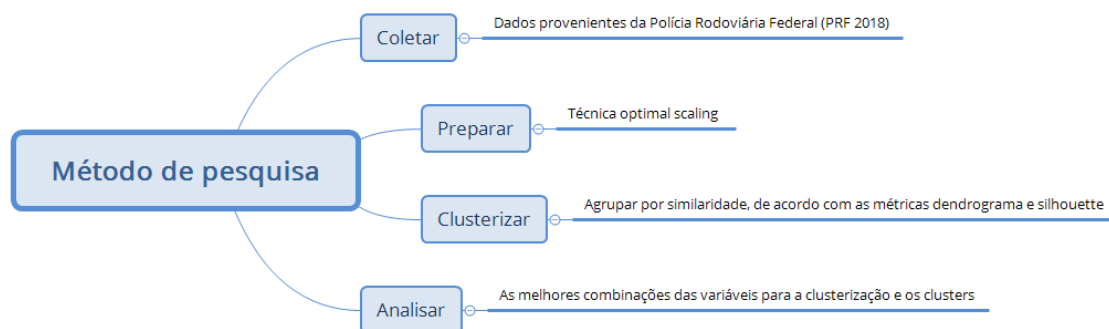


FIGURA 2 - MÉTODO DE PESQUISA

Além disso, foi realizado uma análise das melhores combinações das variáveis possíveis e cada grupo(cluster). Para, no entanto, ter uma ideia de como os dados se agrupam em um espaço de recursos de alta dimensão, Van der Maaten e Hinton (2008) desenvolveram o algoritmo de incorporação de vizinhos estocásticos t-distribuídos. Este algoritmo é usado para produzir um “t-SNE cluster plot”, que pode ser encontrado em “Plots”, que é a identificação gráfica de cada cluster, mostrando sua localização e como ficou distribuído após as análises e processamentos. Além disso, também foi usado o dendrograma, é um diagrama de árvore que mostra os grupos formados por agrupamento de observações em cada cluster e em seus níveis de similaridade. A análise é composta por testes, que foram feitos utilizando os softwares SPSS e JASP e os resultados são observados a partir da figura três em diante.

5. RESULTADOS E DISCUSSÕES

A análise é composta por testes, que foram feitos utilizando os softwares SPSS e JASP. Primeiro teste foi realizado usando a seleção de todas as variáveis disponíveis dos acidentes urbanos fatais, mas o com o acentuado número de variáveis, a análise ficou inviável, de acordo com a métrica silhouette (ver figura 3). Onde em um intervalo entre -1 e 1, quanto mais próximo de 1, seria bem inserida a observação está ao seu cluster de destino, então foi visto que, para adequação, teria que subdividir as variáveis em duplas ou trios, para viabilizar a verificação. Desse modo, foram feitos outros testes, onde foi definido, de acordo com a métrica anterior, que as variáveis seriam subdivididas e que serão: região e causa do acidente, região, causa e final de semana, tipo de pista e condição meteorológica e por último tipo de pista e fase do dia, que trariam uma análise mais satisfatória para o estudo (ver figura 3)

Diante das variáveis que tiveram maior rendimento no silhouette, constatou-se que cada uma delas possui um detalhamento, com uma identificação, nos dados disponibilizados pela Polícia Rodoviária Federal (PRF-2018). A variável região, que são cinco possibilidades: norte, nordeste, sudeste, sul e centro-oeste. Já o tipo de pista de pista são três: simples, dupla e múltipla; a variável final de semana possui dois intervalos: durante a semana e durante o final de semana; fase do dia possui quatro possibilidades: plena noite, pleno dia, amanhecer e anoitecer; a causa do acidente possui vinte e quatro tipos diferentes, como defeito na via, animais na pista, entre outros. Por fim, a condição meteorológica, que são nove possibilidades, como: nublado, granizo, entre outros.

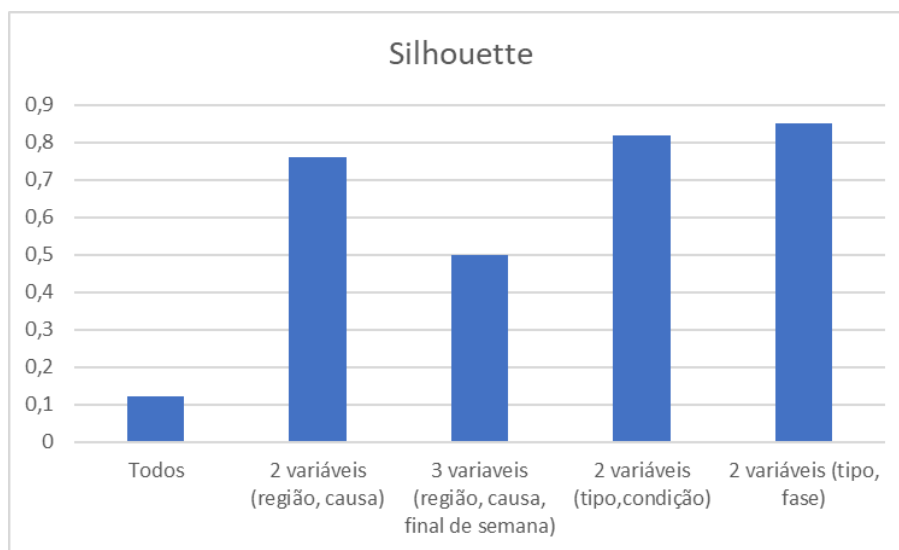


FIGURA 3 - ANÁLISE SILHOUETTE

Ao selecionar as variáveis de acordo com o silhouette, a primeira dupla analisada foi o tipo de pista e fase do dia, que foi gerado quatro clusters fixos (ver figura 4), de acordo com parâmetros de treinamento, e foi observado que os mesmos ficaram bem inseridos como mostra os clusters (ver figura 4). Dentre os quatro clusters gerados, o maior é o em cor verde, onde foi observado que os acidentes 2 (dois), 7 (sete), e 70 (setenta), possuem similaridades, já que mostram que ocorreram em pistas simples, dentre os três tipos possíveis e em horário noturno, entre os quatro horários analisados.

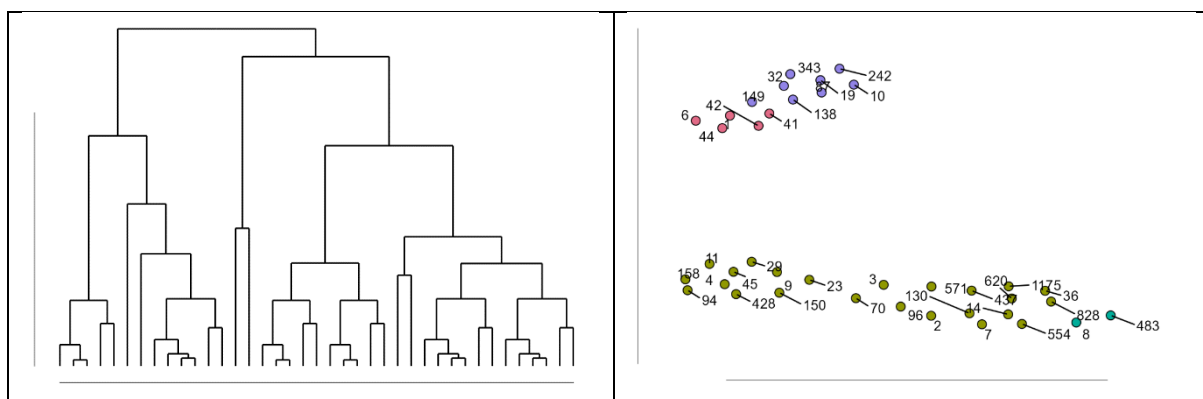


FIGURA 4 – DENDROGRAMA (A) E CLUSTERS (B): TIPO DE PISTA E FASE DO DIA

Na análise seguinte, a dupla de variáveis tipo de pista e condição meteorológica, viu-se que foi gerado quatro clusters fixos, mostrado pelo dendrograma (ver figura 5), de acordo com a métrica do silhouette. Nesse caso, foi observado que os clusters eram de pequena escala e ficaram dispersos, onde o cluster que mais se destacou foi o em cor roxo (ver figura 4), os acidentes 29 (vinte e nove) e 139 (cento e trinta e nove), por exemplo, ocorreram no mesmo tipo de pista dupla e na mesma condição meteorológica, que era nublado, dentre as nove possibilidades.

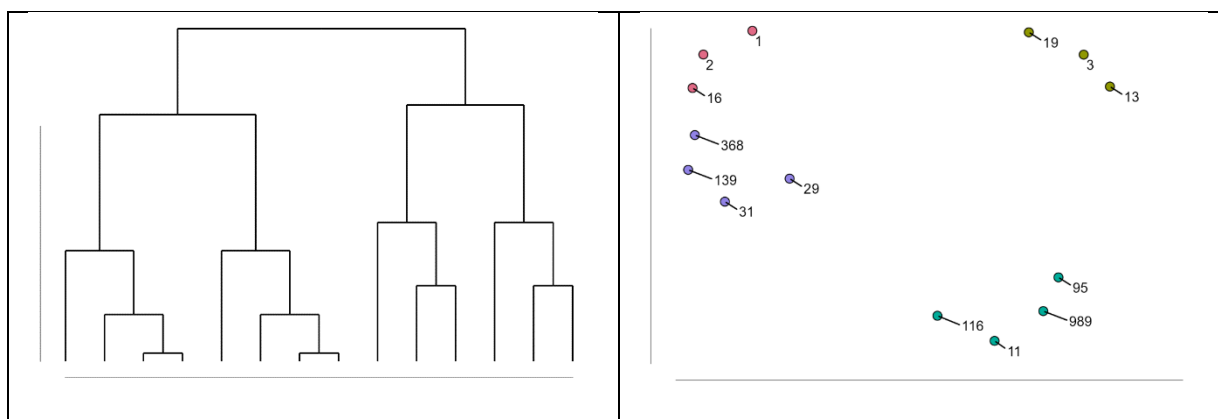


FIGURA 5 – DENDROGRAMA(A) E CLUSTERS (B): TIPO DE PISTA E CONDIÇÃO METEOLÓGICA

As variáveis região e causa do acidente tiveram uma análise bem sucinta e direta, onde mais uma vez foram formados quatro agrupamentos, como podemos ver no dendrograma (ver figura 6). Pode-se ver como os clusters formados estão uniformes e bem localizados com a função PLOT (ver figura 6). Além disso, foi observado no cluster roxo que os acidentes 4 (quatro) e 54 (cinquenta e quatro) foram similares e ocorreram na região nordeste, dentre as cinco regiões possíveis, e causa do acidente, em ambos, foi defeito na via.

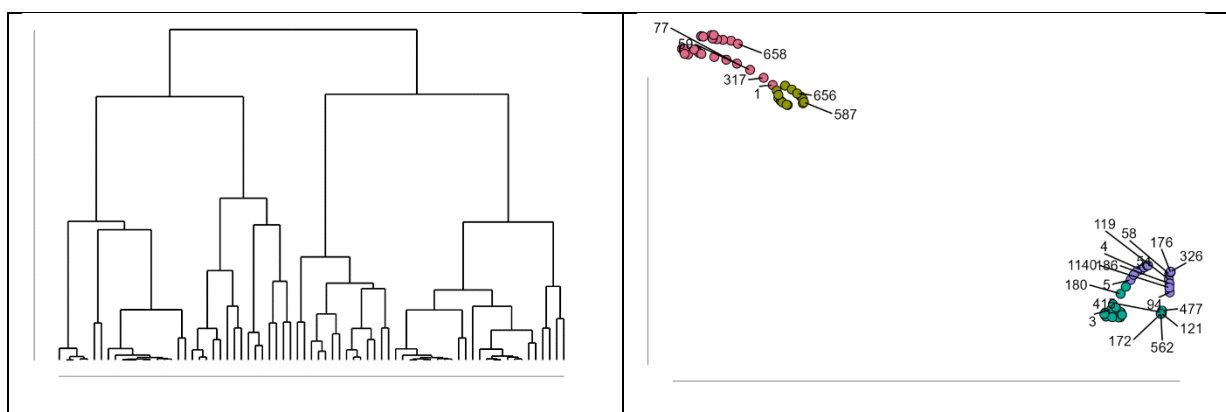


FIGURA 6 – DENDROGRAMA (A) E CLUSTERS (B) : REGIÃO E CAUSA DO ACIDENTE

Observando que número de variáveis selecionadas foram: região, causa do acidente e final de semana, foi verificado que dentre os grupos analisados, esse foi o de menor rendimento na métrica silhouette, comprovando

que os clusters não foram bem dimensionados. A partir disso, o dendrograma (ver figura 7) já mostra que a quantidade de itens foi em grande escala, o que desfavorece a análise. Com a função PLOT (ver figura 7), foi verificado os quatro clusters, onde o cluster verde comprova que a maioria dos acidentes ocorreram durante a semana (segunda a sexta). Os acidentes com identificação 347 (trezentos e quarenta e sete) e 314 (trezentos e quatorze) são similares, já que ocorreram na região centro-oeste, por falta de atenção do motorista.

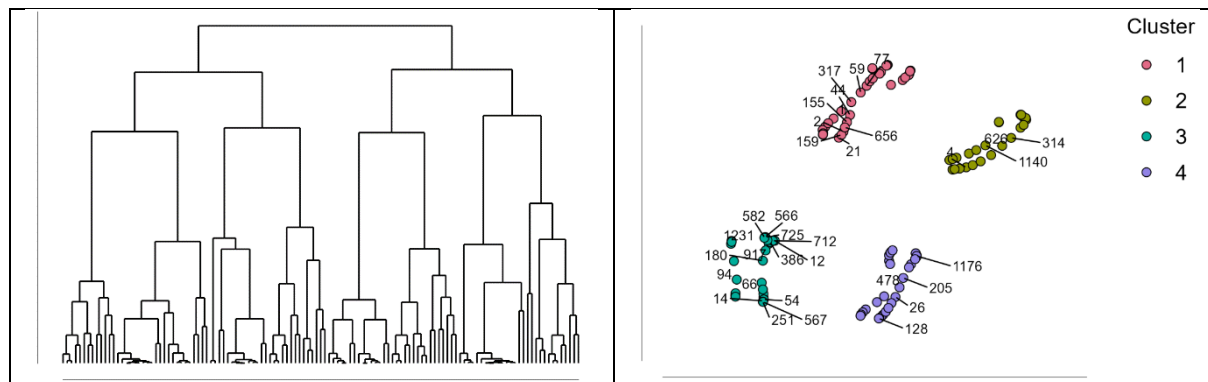


FIGURA 7– DENDROGRAMA (A) E CLUSTERS (B): REGIÃO, CAUSA DO ACIDENTE E FINAL DE SEMANA

6. CONCLUSÕES

Os dados dos acidentes de trânsito fatais urbanos são importantes para a viabilização de quais fatores são relevantes, para que se possa compreender a magnitude das estatísticas disponíveis. Os registros convencionais permitem identificar onde, quando, como os acidentes ocorrem e quem está envolvido diretamente e indiretamente. Diante do grande número de variáveis disponíveis, foi observado que somente após a separação das mesmas, usando a métrica para avaliação do clusters, silhouette, a separação dos acidentes ficou mais clara. Em vista disso, foi definido os padrões guiados pelas variáveis (causa do acidente, tipo da pista, fase do dia, entre outros), deixando a caracterização de forma uniforme e essencial.

Conhecer quais os principais fatores que fizeram parte para a ocorrência dos acidentes de trânsito sucedem como um meio que vem ajudar na elaboração de planos e meios eficientes para a redução de acidentes de trânsito fatais. Com as análises realizadas, pode-se ter uma ideia de como os acidentes fatais ocorrem em diferentes segmentos que compõem o tripé: humano, via e veículo. Foi observado que a clusterização mostra como a similaridade e dissimilaridade ocorrem em uma mesma variável dentro de um grupo de dados. Além de que, teve-se a identificação dos padrões distintos de comportamento das variáveis quando há ocorrência de variações nas condições meteorológicas, nos horários e locais do acidente. Os métodos de clusterização aplicados apresentaram resultados satisfatórios e importantes. Dessa forma, destaca-se a transcendência de utilizar vários métodos em análises de bancos de dados, já que esses métodos podem ser integrantes entre todas as possibilidades.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ELAINA, Jeniffer. Quais são as maiores causas de acidentes de trânsito? São Paulo: Smartia, 2021. Disponível em: <https://www.smartia.com.br/blog/maiores-causas-acidentes-transito/>. Acesso em: 5 nov. 2021.

MARÍN, Letícia; QUEIROZ, Marcos S. A atualidade dos acidentes de trânsito na era da velocidade: uma visão geral. Rio de Janeiro: Scielo, 2000. Disponível em: <https://www.scielo.br/j/csp/a/8Y6XGs7ZTK63sbHnZHdkmDF/?lang=pt>. Acesso em: 10 nov. 2021

OMS (Organização Mundial da Saúde). Lesões causadas pelo trânsito. Genebra: Desconhecido, 2021. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>. Acesso em: 10 nov. 2021.

ALMEIDA, Rosa Livia Freitas de; BEZERRA FILHO, José Gomes; BRAGA, José Uelers; MAGALHÃES, Francismeire Brasileiro; MACEDO, Marinila Calderaro Munguba; SILVA, Kellyanne Abreu. Via, homem e veículo: fatores de risco associados à gravidade dos acidentes de trânsito. Fortaleza: Scielo, 2013. Disponível em: <https://www.scielo.br/j/rsp/a/wJPcRrcLkfxJ6HBTXyr8MbP/?lang=pt>. Acesso em: 25 abr. 2022.

NASCIMENTO, Rodrigo. Clusterização: o que é e porque é tão importante para sua empresa? Desconhecido: Marketing Por Dados, 2021. Disponível em: <https://marketingpordados.com/analise-de-dados/clusterizacao/>. Acesso em: 10 mar. 2022.

FARIA, E. DE O.; FONTES, A. S. Uma contribuição ao tratamento de dados disponíveis nos órgãos executivos de trânsito para avaliar a fiscalização de velocidade. ANPET. Anais..., 2008. Disponível em: <http://www.anpet.org.br>. Acesso em: 10 mar. 2022.

TEIXEIRA, Tita Mara. Fator humano é o que mais contribui para os acidentes de trânsito, alerta Politec. Cuiabá: Politec-Mt, 2019. Disponível em: <http://www.mt.gov.br/-/11848914-fator-humano-e-o-que-mais-contribui-para-os-acidentes-de-transito-alerta-politec> . Acesso em: 07 abr. 2022.

DETRAN-MS. Veja as principais causas de acidentes nas vias e rodovias. Campo Grande: Desconhecida, 2016. Disponível em: <https://www.detrans.ms.gov.br/veja-as-principais-causas-de-acidentes-nas-vias-e-rodovias/> . Acesso em: 25 abr. 2022.

CORTEX, Analista de Conteúdo da. O que é clusterização e por que é importante para entender dados do negócio? Desconhecido: Cortex-Blog, 2022. Disponível em: <https://www.cortex-intelligence.com/blog/o-que-e-clusterizacao> . Acesso em: 10 maio 2022.

HAIR JUNIOR, Joseph F.; BLACK, William C.; TATHAM, Ronald L.; ANDERSON, Rolph E.; TATHAM, Ronald L.. Análise Multivariada de Dados. 5. ed. São Paulo: Bookman, 2005.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. Manual de Análise de Dados - Estatística e Modelagem Multivariada com Excel®, SPSS e Stata. Rio de Janeiro: Gen Ltc, 2017.

MEULMAN, Jacqueline J.; HEISER, Willem J.. SPSS Categories 14.0. Chicago: Spss Inc., 2005.