

UNIVERSIDADE FEDERAL DO PARANÁ

RUDINEI LUIZ BOGO

ANÁLISE DE CORRELAÇÃO CANÔNICA E ANÁLISE DE AGRUPAMENTO
APLICADA AOS ACIDENTES DA BR-116 ENTRE JANEIRO DE 2007 E
NOVEMBRO DE 2009.

CURITIBA

2011

RUDINEI LUIZ BOGO

ANÁLISE DE CORRELAÇÃO CANÔNICA E ANÁLISE DE AGRUPAMENTO
APLICADA AOS ACIDENTES DA BR-116 ENTRE JANEIRO DE 2007 E
NOVEMBRO DE 2009.

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Ciências.

Orientadora: Prof^a. Dr^a. Liliana Madalena Gramani

Co-Orientador: Prof. Dr. Anselmo Chaves Neto

CURITIBA

2011

TERMO DE APROVAÇÃO

RUDINEI LUIZ BOGO

ANÁLISE DE CORRELAÇÃO CANÔNICA E ANÁLISE DE AGRUPAMENTO
APLICADA AOS ACIDENTES DA BR-116 ENTRE JANEIRO DE 2007 E
NOVEMBRO DE 2009.

Dissertação aprovada como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação em Métodos Numéricos em Engenharia, do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientadora: Prof^a. Dr^a. Liliana Madalena Gramani
Departamento de Matemática, UFPR

Co-Orientador: Prof. Dr. Anselmo Chaves Neto
Departamento de Estatística, UFPR

Prof. Dr. Jair Mendes Marques
Programa de Pós-Graduação em Métodos Numéricos
em Engenharia da UFPR,

Prof. Dr. Eloy Kaviski
Departamento de Hidráulica e Saneamento, UFPR

Prof. Dr. Leonardo Ramos Emmendorfer
Centro de Ciências Computacionais, FURG.

Curitiba, 23 de fevereiro de 2011.

AGRADECIMENTOS

Sobretudo a Deus por eu chegar até aqui.

À minha família, que sempre esteve em minha vida, me ajudando e dando força para que eu continuasse.

À orientadora Profa. Liliana M. Gramani, pela amizade, dedicação, compreensão e orientação neste trabalho.

Ao co-orientador Prof. Anselmo Chaves Neto, pela amizade, colaboração e incentivo neste trabalho. Pelos seus conhecimentos transmitidos e imprescindíveis no desenvolvimento do mesmo.

Aos demais Professores do Programa de Pós-Graduação em Métodos Numéricos em Engenharia, que de forma direta ou indireta transmitiram conhecimentos importantes para minha formação profissional e pessoal.

Aos colegas de pós-graduação, Fábio, Marina e Maiko, pela dicas e ajudas importantes na realização deste trabalho.

A todos os demais colegas do PPGMNE, pela amizade concedida nesses dois anos de trabalho.

A secretária do CESEC, Maristela Bandil, pela sua enorme alegria e amizade constantes. Pela enorme ajuda e conselhos durante todo o período do mestrado.

Ao Policial Rodoviário e Estatístico Nivaldo Aparecido Minervi, pela concessão dos dados para realização desta pesquisa.

A todas as demais pessoas que, de forma direta ou indireta, contribuíram com a realização do meu trabalho.

Determinação coragem e auto confiança são fatores decisivos para o sucesso. Se estamos possuídos por uma inabalável determinação conseguiremos superá-los. Independentemente das circunstâncias, devemos ser sempre humildes, recatados e despidos de orgulho.

Dalai Lama

RESUMO

Acidentes de trânsito têm se incorporado cada vez mais na vida das pessoas. No Brasil existe um número grande de rodovias pelas quais trafegam diariamente milhares de veículos. Em consequência disso, acidentes ocorrem com frequência tomando números preocupantes para nossa sociedade. Dentre as rodovias brasileiras, a BR-116 destaca-se por ser umas das principais rodovias brasileiras pois liga o nordeste ao sul do país. As técnicas da análise estatística multivariada não foram encontradas em trabalhos publicados com aplicações em acidentes de trânsito. O objetivo principal do presente estudo é aplicar a análise de correlação canônica e análise de agrupamentos nos dados dos acidentes registrados na rodovia BR-116, no trecho que compreende o Estado do Paraná. Foram utilizados dados obtidos da Polícia Rodoviária Federal do Paraná dos acidentes ocorridos nos anos de 2007, 2008 e 2009. A aplicação da correlação canônica foi realizada para verificar as relações existentes entre dois grupos de variáveis referentes aos acidentes: tipos de acidentes, causas de acidentes, condições meteorológicas, fases do dia e características da pista. Por outro lado, a análise de agrupamento foi feita, utilizando o método das k-médias, para agrupar os quilômetros da rodovia considerando todas as variáveis utilizadas na correlação canônica. Os resultados apontam para uma alta correlação entre os grupos considerados. Além disso, nota-se ainda uma forte correlação entre algumas variáveis e alguns quilômetros apresentaram mais problemas em relação a outros. Os grupos formados agruparam os quilômetros que apresentaram características semelhantes em relação aos acidentes. Os resultados obtidos pelas técnicas se complementam para o melhor entendimento dos acidentes que ocorrem na BR-116.

Palavras-chave: Acidentes de Trânsito, Análise de Correlação Canônica, Análise de Agrupamentos, BR-116.

ABSTRACT

Traffic accidents have become increasingly embedded in people's lives. In Brazil there is a lot of highways along which thousands of vehicles travel daily. In Consequently, accidents have occurred frequently taking hard numbers to our society. Among the roads Brazil, BR-116 is notable for being one of the main Brazilian roads linking south to northeast. The multivariate statistical analysis techniques are relatively recent published works were not found with applications in traffic accidents. The main objective of this study is to apply the canonical correlation analysis and cluster analysis on data from accidents recorded in BR-116, the stretch that includes the state of Parana. Were used data obtained by the Federal Highway Police of Parana of accidents in the years 2007, 2008 and 2009 on the highway BR-116. In the application of canonical correlation to check the relationship between two groups of variables were An analysis of some groups of variables related to accidents: types of accidents, causes of accidents, weather condition, phases of the day and the track characteristics. Moreover, the analysis clusters was made using the method of k-means, for cluster the miles of highway considering all the variables used in the canonical correlation. The results point to a high correlation between the groups. Furthermore, note still a strong correlation between some variables and some km had more problems in relation to others. The formed groups that had pooled km similar characteristics in relation to accidents. The results obtained by complementary techniques for the best understanding of the accidents that occur on the BR-116.

Key-words: Accidents, Canonical Correlation Analysis, Cluster Analysis, BR-116.

LISTA DE FIGURAS

FIGURA 2.1	– EXEMPLOS DE RODOVIAS RADIAIS	24
FIGURA 2.2	– EXEMPLOS DE RODOVIAS LONGITUDINAIS	24
FIGURA 2.3	– EXEMPLOS DE RODOVIAS TRANSVERSAIS	25
FIGURA 2.4	– EXEMPLOS DE RODOVIAS DIAGONAIS	26
FIGURA 2.5	– MORTOS E FERIDOS A CADA 1.000 ACIDENTES NAS RODOVIAS FEDERAIS EM 2004 E 2005	33
FIGURA 2.6	– MORTOS E FERIDOS NAS RODOVIAS FEDERAIS NOS ANOS DE 2004 E 2005	33
FIGURA 2.7	– RODOVIA FEDERAL BR-116	37
FIGURA 2.8	– TRECHO DA RODOVIA BR-116 EM FORTALEZA-CE	38
FIGURA 2.9	– VIA DUTRA NO TRECHO ENTRE TAUBATÉ - SP E PINDAMONHANGABA - SP	39
FIGURA 2.10	– TRECHO DA RODOVIA RÉGIS BITTENCOURT PRÓXIMO A REGISTRO - SP	40
FIGURA 2.11	– TRECHO DA RODOVIA BR-116 PRÓXIMO A CIDADE DE CURITIBA - PR	40
FIGURA 2.12	– TRECHO DA RODOVIA BR-116 EM SANTA CATARINA, PRÓXIMO À DIVISA COM O RIO GRANDE DO SUL	41
FIGURA 2.13	– RODOVIA FEDERAL BR-116 NO ESTADO DO PARANÁ	42
FIGURA 2.14	– ACIDENTES NA BR-116 (PARANÁ) POR TIPO DE PISTA	43

FIGURA 2.15	– DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 SEGUNDO AS FASES DO DIA	45
FIGURA 2.16	– DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 SEGUNDO O TRAÇADO DA VIA	45
FIGURA 3.1	– MATRIZ DICOTÔMICA DOS ACIDENTES NA BR-116	73
FIGURA 3.2	– MATRIZ DE FREQUÊNCIA DOS ACIDENTES NA BR-116 ...	73
FIGURA 3.3	– TELA DO PROGRAMA STATGRAPHICS CENTURION	77
FIGURA 4.1	– ACIDENTES NA BR-116 (JAN/07 À DEZ/07) COM DADOS DA PRF/PR	78
FIGURA 4.2	– ACIDENTES NA BR-116 (JAN/08 À DEZ/08) COM DADOS DA PRF/PR	79
FIGURA 4.3	– ACIDENTES NA BR-116 (JAN/09 À DEZ/09) COM DADOS DA PRF/PR	79
FIGURA 4.4	– GRÁFICO DAS VARIÁVEIS CANÔNICAS	82
FIGURA 4.5	– GRÁFICO DAS VARIÁVEIS CANÔNICAS	85
FIGURA 4.6	– GRÁFICO DAS VARIÁVEIS CANÔNICAS	88
FIGURA 4.7	– GRÁFICO DAS VARIÁVEIS CANÔNICAS	91
FIGURA 4.8	– KM 116 DA BR-116 (PARANÁ) - VISTA AÉREA E PAISAGEM	93
FIGURA 4.9	– KM 119 DA BR-116 (PARANÁ)- VISTA AÉREA E PAISAGEM .	93
FIGURA 4.10	– DENDOGRAMA DO AGRUPAMENTO USANDO O MÉTODO DE WARD	94
FIGURA 1.1	– ACIDENTES NA BR-116 (JAN/07 À DEZ/07) COM DADOS DA PRF/PR	101
FIGURA 1.2	– ACIDENTES NA BR-116 (JAN/08 À DEZ/08) COM DADOS DA	

PRF/PR 102

FIGURA 1.3 – ACIDENTES NA BR-116 (JAN/09 À DEZ/09) COM DADOS DA

PRF/PR 102

LISTA DE TABELAS

TABELA 2.1	– EXTENSÃO DA REDE RODOVIÁRIA FEDERAL DO BRASIL .	31
TABELA 2.2	– ESTATÍSTICAS DOS ACIDENTES DE TRÂNSITO NAS RODOVIAS FEDERAIS EM 2004 E 2005, NÚMERO DE ACIDENTES, NÚMERO DE ENVOLVIDOS POR GRAVIDADE DAS LESÕES, MORTOS, FERIDOS E ILESOS A CADA 1.000 ACIDENTES	32
TABELA 2.3	– ACIDENTES DE TRÂNSITO POR TIPO NAS RODOVIAIS FEDERAIS - 2004	35
TABELA 2.4	– ACIDENTES DE TRÂNSITO POR FATORES CONTRIBUINTES- 2004	35
TABELA 2.5	– ACIDENTES DE TRÂNSITO SEGUNDO AS CONDIÇÕES DO TEMPO - 2004	36
TABELA 2.6	– ACIDENTES DE TRÂNSITO SEGUNDO FASES DO DIA	36
TABELA 2.7	– ACIDENTES DE TRÂNSITO SEGUNDO O TRAÇADO DA VIA	36
TABELA 2.8	– ESTATÍSTICAS DOS ACIDENTES NAS TREZE RODOVIAS FEDERAIS COM MAIOR NÚMERO DE MORTES	42
TABELA 2.9	– ACIDENTES NA BR-116 (PARANÁ) POR FATOR CONTRIBUINTE	44
TABELA 2.10	– ACIDENTES DE TRÂNSITO SEGUNDO AS CONDIÇÕES DO TEMPO NA BR-116, ESTADO DO PARANÁ	46
TABELA 2.11	– ACIDENTES DE TRÂNSITO POR CAUSAS NA BR-116, ESTADO DO PARANÁ	46

TABELA 4.1	– RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CAUSAS DOS ACIDENTES	80
TABELA 4.2	– COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS	80
TABELA 4.3	– CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 2	81
TABELA 4.4	– RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CONDIÇÃO METEOROLÓGICA	83
TABELA 4.5	– COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS	83
TABELA 4.6	– CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 3	85
TABELA 4.7	– RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X FASES DO DIA	86
TABELA 4.8	– COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS	87
TABELA 4.9	– CORRELAÇÃO ENTRE VARIÁVEIS DO GRUPO 1 E GRUPO 4	88
TABELA 4.10	– RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CARACTERÍSTICAS DA PISTA	89
TABELA 4.11	– COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS	90

TABELA 4.12 – CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 5	92
TABELA 4.13 – AGRUPAMENTO DOS KM USANDO O K-MÉDIAS PARA K=5	96
TABELA 2.1 – TRECHO DETALHADO DA BR-116 NO ESTADO DO PARANÁ	110

LISTA DE SIGLAS

PRF	Polícia Rodoviária Federal
NE	Nordeste
SO	Sudoeste
SE	Sudeste
OMS	Organização Mundial da Saúde
CTB	Código de Trânsito Brasileiro
BAT	Boletim de Acidente de Trânsito
IPEA	Instituto de Pesquisa Econômica Aplicada
IML	Instituto Médico Legal
DPVAT	Danos Pessoais Causados por Veículos Automotores de Via Terrestre
ACC	Análise de Correlação Canônica
AA	Análise de Agrupamentos

SUMÁRIO

1 INTRODUÇÃO	18
1.1 O PROBLEMA	18
1.2 OBJETIVO	18
1.3 JUSTIFICATIVA	19
1.4 ESTRUTURA DO ESTUDO	20
2 REVISÃO BIBLIOGRÁFICA	22
2.1 TRÁFEGO VEICULAR	22
2.1.1 Tipos de Rodovia	22
2.2 ACIDENTES DE TRÂNSITO	27
2.2.1 Acidentes no Brasil	30
2.3 A RODOVIA BR-116	34
2.3.1 Acidentes na BR-116	41
2.4 CUSTOS DOS ACIDENTES	44
2.5 ANÁLISE DE CORRELAÇÃO CANÔNICA	49
2.5.1 Introdução	49
2.5.2 Alguns Trabalhos Relacionados	52
2.5.3 Variáveis Canônicas e Correlações Canônicas	53
2.5.4 Variáveis Canônicas Amostrais e Correlações Canônicas Amostrais	57
2.5.5 Validação da Análise de Correlação Canônica	59

2.6 ANÁLISE DE AGRUPAMENTO	60
2.6.1 Medidas de Similaridades e Dissimilaridades	60
2.6.2 Métodos de Agrupamento	61
2.6.3 Agrupamento Hierárquico	62
2.6.4 Agrupamento Não-Hierárquico	64
3 MATERIAL E MÉTODOS	66
3.1 DADOS UTILIZADOS	66
3.2 VARIÁVEIS DA ANÁLISE DE CORRELAÇÃO CANÔNICA	74
3.3 APLICAÇÃO DOS MÉTODOS PROPOSTOS - ACC E AA	76
4 RESULTADOS E DISCUSSÃO	78
4.1 RESULTADOS DA ACC	79
4.1.1 Tipos de Acidentes x Causas de Acidentes	80
4.1.2 Tipos De Acidentes x Condição Meteorológica	83
4.1.3 Tipos de acidentes x fases do dia	86
4.1.4 Tipos de Acidentes x Características da Pista	88
4.2 RESULTADOS DA ANÁLISE DE AGRUPAMENTO	93
5 CONCLUSÕES E RECOMENDAÇÕES	97
REFERÊNCIAS	99
APÊNDICE A – HISTOGRAMA DOS ACIDENTES OCORRIDOS NA BR-116 (2007, 2008 E 2009)	101
ANEXO A – PROGRAMAS UTILIZADOS PARA OBTENÇÃO DA MATRIZ DICOTÔMICA	103
ANEXO B – DETALHAMENTO DO TRECHO DA BR-116 NO ESTADO DO PARANÁ	

1 INTRODUÇÃO

1.1 O PROBLEMA

O tráfego de veículos nas rodovias federais brasileiras representa um fenômeno de grande importância sócio-econômica, principalmente, por ser, até o presente momento, o principal meio de transporte utilizado pela sociedade brasileira para se deslocar de uma cidade para outra. Também representa a principal alternativa para o transporte de produtos comercializados pela indústria.

Todavia, com o crescente aumento do número de veículos nas estradas e a deteriorização da malha asfáltica em grande parte das rodovias federais, a falta de segurança tornou-se um problema evidente, elevando assim, o número de acidentes ocorridos.

O Brasil é um país que possui um grande número de rodovias federais. São aproximadamente 73 mil¹ quilômetros de rodovias federais, nos quais a maior parte encontra-se pavimentada.

Mesmo com a sequência de privatizações de trechos de algumas rodovias, o que aumenta a qualidade de infra-estrutura e segurança, acidentes têm ocorrido com frequência e em grandes quantidades. Isso evidencia o fato de que os motoristas estão se tornando cada vez mais imprudentes colocando em risco a vida de muitas pessoas, outras vezes, a falta de condições adequadas das rodovias, conduzem a acidentes.

1.2 OBJETIVO

Objetivo Geral:

¹Fonte: Ministério dos Transportes - Plano Nacional de Viação - PNV/2004.

O presente estudo tem por objetivo geral aplicar técnicas da análise estatística multivariada para determinar os piores quilômetros da BR-116, rodovia federal brasileira, no trecho situado no estado do Paraná, utilizando dados de acidentes de trânsito ocorridos entre janeiro de 2007 à novembro de 2009.

Objetivos Específicos:

- (a)** Aplicar a análise de correlação canônica para verificar as relações existentes entre os grupos de acidentes tais como: tipos de acidentes, causas dos acidentes, condições meteorológicas, características da pista e fases do dia;
- (b)** Identificar os quilômetros da rodovia onde ocorrem mais acidentes e relacioná-las com suas possíveis causas e fatores determinantes;
- (c)** Obter com os métodos estatísticos um entendimento mais claro dos acidentes que ocorrem na BR-116 no trecho situado no estado do Paraná, no sentido de auxiliar os responsáveis pela rodovia, para que se possa fazer um planejamento mais adequado considerando os quilômetros identificados como os mais perigosos.
- (d)** Aplicar a análise de agrupamento, mais especificamente, o método das k-médias, no sentido de agrupar os quilômetros mais perigosos da rodovia.

1.3 JUSTIFICATIVA

A BR-116 é uma das principais rodovias federais brasileiras, com aproximadamente 4.594² quilômetros de malha viária, responsável por ligar o sul ao nordeste do país. De acordo com o Departamento Nacional de Trânsito (2006), no ano de 2004 foram registrados 23.135 acidentes com 1.060 mortos no total. Este fato que coloca a rodovia como uma das 5 (cinco) mais perigosas de todo o Brasil.

No Paraná, a rodovia possui uma extensão de 211 quilômetros. Deste total, 118 quilômetros da rodovia possuem pista dupla e 93 quilômetros são trechos de pista

²Fonte: Ministério dos Transportes - Plano Nacional de Viação - PNV/2010.

simples. Acidentes neste trecho ocorrem com frequência. Foram registrados pela Polícia Rodoviária Federal do Paraná (PRF), nos anos de 2007, 2008 e 2009, um total de 6.245 acidentes ao longo da rodovia. Existem alguns trechos da rodovia que possuem alguns pontos críticos, tais como: cruzamentos mal planejados e, conseqüentemente, o número de acidentes nesses trechos é maior.

A motivação principal do presente trabalho é aplicar técnicas da análise estatística multivariada nos dados de acidentes de trânsito para determinar as interações entre alguns grupo de variáveis, descritas nos acidentes, e agrupar os piores quilômetros da rodovia associados com algumas característica do fenômeno. Desta forma, pode-se fazer um planejamento adequado para realização de melhorias futuras nos trechos da rodovia que podem ser identificados com problemas.

Na literatura existente, não foram encontrados registros de trabalhos relacionados à acidentes de trânsito em rodovias, que usam alguma técnica da análise estatística multivariada para verificar relações existentes entre os tipos de acidentes com suas possíveis causas, fornecendo um estudo geral dos acidentes em uma determinada rodovia.

1.4 ESTRUTURA DO ESTUDO

A estrutura deste trabalho conta com mais quatro capítulos, além desta introdução: No capítulo 2 apresenta-se uma revisão de literatura sobre tipos de tráfego e rodovias, os conceitos mais importantes relacionados aos acidentes e alguns dados estatísticos sobre acidentes de trânsito no Brasil. Além disso, é feita uma revisão de literatura sobre as técnicas da estatística multivariada utilizadas neste trabalho: análise de correlação canônica e análise de agrupamentos. No capítulo 3 descreve-se o material utilizado para aplicação das análises, ou seja, as variáveis utilizadas, bem como os métodos aplicados. O capítulo 4 apresenta os resultados das análises feitas com os programas elaborados e, finalmente, o capítulo 5, contem a conclusão do trabalho. A dissertação ainda possui, além das referências, um apêndice e dois anexos para

melhor compreensão do trabalho.

2 REVISÃO BIBLIOGRÁFICA

2.1 TRÁFEGO VEICULAR

Santos (2008) afirma que a palavra tráfego aplicá-se de uma forma geral à movimentação de veículos: tráfego rodoviário, aeroviário, ferroviário, hidroviário; tráfego de pedestres e até a informações de dados (rede de computadores). Neste trabalho, o objetivo proposto é estudar o tráfego de veículos na rodovia federal brasileira BR-116, no trecho situado no estado do Paraná.

2.1.1 Tipos de Rodovia

As estradas pavimentadas, em geral, permitem o acesso barato e rápido de homens e mercadorias aos pontos mais remotos de uma nação. De utilidade indiscutível para indivíduos, sociedades e economias em qualquer período da história, o deslocamento por via terrestre é um dos principais meios de transporte para médias e longas distâncias do mundo atual. Uma estrada é uma via de comunicação para o deslocamento de veículos, ciclistas ou pedestres. No Império Romano, as estradas construídas pelo exército romano promoveram a circulação de mercadorias, favorecendo o comércio e aumentando o poder do Império. Já no século XIX, com a máquina a vapor, o Império Britânico construiu as estradas de ferro em todos os seus domínios e com isto contribuiu com um desenvolvimento extraordinário. Uma rodovia é uma estrada de rodagem, correspondente a uma via de transporte interurbano de alta velocidade, podendo ou não ser proibido o tráfego de pedestres ou ciclistas. A palavra rodovia é sinônimo de autoestrada ou autopista, pavimentada, de pista dupla, para veículos de alta velocidade. O termo estrada tem uso mais abrangente designando também até caminhos vicinais, veredas e trilhas sem calçamento das regiões rurais de pouco trânsito de veículos freqüentemente usados por pessoas que viajam a pé,

em carroças ou montadas em animais (SANTOS, 2008).

As rodovias, em relação à largura e circulação de automóveis, podem ser de pista simples ou múltipla. As rodovias de pistas simples são aquelas em que possuem um pavimento¹ asfáltico, que é compartilhado por veículos nos dois sentidos de circulação separados por uma faixa (mão dupla). No Brasil, considerando este tipo de rodovia, os veículos devem circular sempre no lado direito (em relação a si) e somente podem utilizar o outro sentido da via para ultrapassagem quando permitido. Por outro lado, as rodovias de pistas múltiplas possuem um pavimento com barreiras centrais que impedem conversões ou retornos. Nesses casos, cada sentido de circulação possui uma pista própria com duas ou mais faixas, permitindo assim, o desenvolvimento de uma maior velocidade e maior segurança, já que diminui o risco de colisão frontal² entre dois veículos em alta velocidade.

O Plano Nacional de Viação (2010), no Brasil, as rodovias são conhecidas ou identificadas a partir de um código formado por três dígitos. As rodovias federais são chamadas de BR-XXX (onde XXX é o código de cada rodovia) e as estaduais chamadas de YY-XX(X) (onde YY representa a sigla de cada estado brasileiro e XXX ou XX o código de cada rodovia). As rodovias federais mais comumente chamadas de BR's estão organizadas da seguinte maneira:

- **Rodovias Radiais:** BR-0XX - São rodovias que partem da capital federal em direção aos extremos do país. A numeração pode variar de 05 a 95, segundo a razão numérica 05 e no sentido horário; exemplo: BR-020, BR-040 e BR-070 (mapa da figura 2.1).
- **Rodovias Longitudinais:** BR-1XX - São rodovias que cortam o país na direção norte-sul. A numeração dessas rodovias varia de 00, no extremo leste do País, a 50, na Capital Federal e de 50 a 99, no extremo oeste. O número de uma rodovia

¹Pavimento é uma base horizontal constituída por um ou mais materiais que se coloca sobre o terreno natural ou terraplanado, para aumentar sua resistência e servir para a circulação de pessoas ou veículos.

²No capítulo 3 é definido cada tipo de acidente apresentado aqui.



FIGURA 2.1: EXEMPLOS DE RODOVIAS RADIAIS

FONTE: Plano Nacional de Viação - Ministérios dos Transportes (2010)

longitudinal é obtido por interpolação entre 00 e 50, se a rodovia estiver a leste de Brasília e, entre 50 e 99 se estiver a oeste, em função da distância da rodovia ao meridiano da Capital Federal; exemplos: BR-101, BR-153 e BR-174 (mapa da figura 2.2).



FIGURA 2.2: EXEMPLOS DE RODOVIAS LONGITUDINAIS

FONTE: Plano Nacional de Viação - Ministérios dos Transportes (2010)

- **Rodovias Transversais:** BR-2XX - São rodovias que cortam o país na direção leste-oeste. A numeração dessas rodovias varia de 00, no extremo norte do país,

a 50 na Capital Federal, e de 50 a 99 no extremo sul. O número de uma rodovia transversal é obtido por interpolação entre 00 e 50, se a rodovia estiver ao norte da Capital, e entre 50 a 99, se estiver ao sul, em função da distância da rodovia ao paralelo de Brasília; exemplos: BR-230, BR-262 e BR-290 (mapa da figura 2.3).

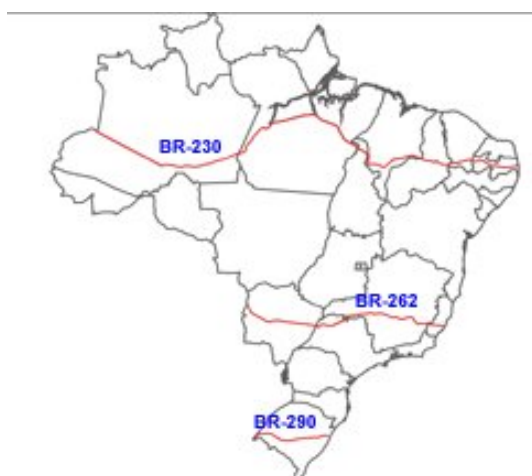


FIGURA 2.3: EXEMPLOS DE RODOVIAS TRANSVERSAIS

FONTE: Plano Nacional de Viação - Ministérios dos Transportes (2010)

- **Rodovias Diagonais:** BR-3XX - São rodovias que cortam o país em dois sentidos: noroeste-sudeste ou nordeste-sudoeste. A numeração dessas rodovias obedece o seguinte critério:
 1. Diagonais orientadas na direção geral Noroeste-Sudeste. A numeração varia, segundo números pares, de 00, no extremo Nordeste (NE) do país, a 50, em Brasília, e de 50 a 98, no extremo Sudoeste (SO). Obtém-se o número da rodovia mediante interpolação entre os limites consignados, em função da distância da rodovia a uma linha com a direção Noroeste-Sudeste, passando pela Capital Federal; exemplos: BR-304, BR-324 e BR-364 (ver Figura 2.4).
 2. Diagonais orientadas na direção geral NE-SO. A numeração varia, segundo números ímpares, de 01, no extremo NO do país, a 51, em Brasília, e de 51

a 99, no extremo Sudeste (SE). Obtém-se o número aproximado da rodovia mediante interpolação entre os limites consignados, em função da distância da rodovia a uma linha com a direção NE-SO, passando pela Capital Federal; exemplos: BR-319, BR-365, BR-381 (mapa da figura 2.4).



FIGURA 2.4: EXEMPLOS DE RODOVIAS DIAGONAIS

FONTE: Plano Nacional de Viação - Ministérios dos Transportes (2010)

- **Rodovias de Ligação:** BR-4XX - São rodovias que se apresentam em qualquer situação. A numeração dessas rodovias varia entre 00 e 50, se a rodovia estiver ao norte do paralelo da Capital Federal, e entre 50 e 99, se estiver ao sul desta referência; exemplos: BR-401, BR-407 e BR-488.

A quilometragem das rodovias federais brasileiras não é cumulativa de uma Unidade de Federação para a outro, isto é, cada vez que uma rodovia inicia dentro de uma nova UF sua quilometragem reinicia novamente a partir de zero.

O sentido da quilometragem de cada tipo de rodovia é descrito a seguir:

- **RODOVIAS RADIAIS** - o sentido de quilometragem inicia do Anel Rodoviário de Brasília em direção aos extremos do país, e tendo o quilômetro zero de cada estado no ponto da rodovia mais próxima a capital federal.

- RODOVIAS LONGITUDINAIS - o sentido de quilometragem inicia do norte para o sul. As únicas exceções deste caso são a BR-163 e BR-174, que tem o sentido de quilometragem do sul para o norte.
- RODOVIAS TRANSVERSAIS - o sentido de quilometragem inicia do leste para o oeste.
- RODOVIAS DIAGONAIS - a quilometragem se inicia no ponto mais ao norte da rodovia indo em direção ao ponto mais ao sul.
- RODOVIAS DE LIGAÇÃO - geralmente a contagem da quilometragem segue do ponto mais ao norte da rodovia para o ponto mais ao sul. No caso de ligação entre duas rodovias federais, a quilometragem começa na rodovia de maior importância.

2.2 ACIDENTES DE TRÂNSITO

O problema "Acidentes de Trânsito" tem se incorporado de forma lenta e assustadora no cotidiano das pessoas. Cada vez mais pessoas têm perdido suas vidas devido a imprudência, desrespeito dos motoristas e outros fatores que agregam a questão "trânsito".

A Organização Mundial da Saúde (2009) (OMS) afirma que, em 2002, quase 1,2 milhão de pessoas no mundo morreram em consequência de acidente de trânsito em rodovia. Isso representa mundialmente, uma média diária de 3.242 mortes. Além disso, a OMS estima entre 20 milhões e 50 milhões de pessoas no mundo saem feridas ou com algum tipo de deficiência em decorrências de acidentes nas rodovias, a cada ano.

No entanto, esses dados não refletem os números reais de envolvidos nos acidentes de trânsito. Considerando que, de maneira indireta, familiares, policiais rodoviários, agentes de controle de trânsito nas estradas, pessoal do socorro e emergências aproximam-se do problema, o número de impacto dos acidentes alcançam

uma dimensão epidêmica.

Branco, A. M. (1999), afirma que acidentes rodoviários são, desde longa data, uma grande preocupação mundial. Os problemas dos acidentes tem repercussões sociais graves, assim como também de natureza econômica. Estima-se que, nos Estados Unidos da América (EUA), o custo anual causado pelos acidentes de trânsito chega à casa dos 150 bilhões de dólares.

O Código de Trânsito Brasileiro (CTB), coloca que a segurança e prevenção de acidentes de trânsito em rodovias federais são obrigações das autoridades gestoras e operadoras de trânsito e transportes, como pode-se ver no quadro 2.1.

Ministérios das Cidades	Departamento Nacional de Trânsito - DENATRAN
Ministérios dos Transportes	Departamento Nacional de Infra-estrutura de Transportes - DNIT
Ministério da Justiça	Polícia Rodoviária Federal - PRF
Governos Estaduais	Departamento de Estradas e Rodagens (DERs) e Departamentos Estaduais de Trânsito (Detrans)

QUADRO 2.1 - AUTORIDADES GESTORAS E OPERADORAS DE TRÂNSITO E TRANSPORTES

FONTE: Retirado do IPEA, DENATRAN (2006)

A OMS adotou conceitos e definições relativos aos acidentes de trânsito na "Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - Décima Revisão"(CID-10), onde assume-se que:

- **Acidente** é um evento independente do desejo do homem, causado por uma força externa, alheia, que atua subitamente (de forma inesperada) e deixa ferimentos no corpos e na mente. Alternativamente, pode-se considerar um acidente um evento não intencional que produz ferimentos ou danos.

- **Acidente de transporte** é todo acidente que envolve um veículo destinado ou usado no momento do acidente, principalmente, para o transporte de pessoas ou de mercadorias de um lugar para outro.

Observa-se que a OMS classifica os acidentes segundo o meio de transporte utilizado, destacando os pedestres e os veículos utilizados pelas vítimas, para depois, subclassificar as vítimas segundo seu papel no acidente (condutor, passageiro, etc). No entanto, considera-se o veículo como o fator mais importante que deve ser identificado para fins de prevenção.

- **Via pública** (via de trânsito, estrada, rodovia ou rua) é a largura total entre dois limites de propriedade de todo o terreno ou caminho aberto ao público, para circulação de pessoas ou bens de um lugar para outro. Pista ou leito da via é considerada parte da via pública que, eventualmente, é utilizada para o trânsito de veículos.
- **Acidente de trânsito** é todo acidente com veículo ocorrido na via pública.
- **Pedestre** é toda pessoa que participa de um acidente, mas que no momento em que este ocorreu não estava no interior ou sobre um veículo a motor, como por exemplo: trem em via férrea, bonde, veículo de tração animal ou outro veículo, ou sobre bicicleta ou animal.
- **Condutor** é o ocupante de um veículo que manobra (dirige/guia) ou tem a intenção de manobrá-lo.
- **Passageiro** é todo ocupante de um veículo que não seja o condutor. Neste caso, exclui-se a pessoa viajando no exterior do veículo ou que não esteja ocupando local reservado normalmente ao condutor ou aos passageiros ou o local reservado para o transporte de cargas.

Dentre as diversas causas de acidentes, Branco, A. M. (1999) destaca que a maioria dos acidentes é causada por falha humana. Segundo o autor, nos Estados Unidos, atribuem-se a falhas humanas 85% dos acidentes rodoviários, e que, no

Brasil, os números são bem próximos. Por outro lado, uma rodovia bem planejada, bem operada e sinalizada pode reduzir consideravelmente o número de acidentes, fazendo com que os motoristas cometam menos erros devido a orientação adequada.

Uma estrada segura pode reduzir a gravidade dos acidentes, isto é, uma rodovia bem planejada e sinalizada, além de reduzir o número de acidentes, diminui a gravidade dos acidentes que venham a ocorrer.

As falhas humanas estão ligadas a diversos fatores muitas vezes mensuráveis (bebida, "rebite", drogas, imprudência, etc); mas também por fatores não mensuráveis (distração, cansaço, preocupação). Além do fator humano, são responsáveis por acidentes, o veículo, a via, o meio ambiente e fatores institucionais e sociais.

- **Veículo:** Qualidade do veículo e conservação de seus componentes, como pneus, freios, amortecedores, sistema de direção, etc.
- **Via/Meio-Ambiente:** Buracos, pavimentos, sinalização, dispositivos de proteção, inclinação da curva, iluminação em excesso ou insuficiente, vegetação e condições climáticas. Neste fator, ainda, considera-se o projeto da rodovia, como curvas fechadas, rampas excessivas, postes e árvores próximos da pista, etc; fatores de manutenção e fatores de ordem natural (chuva).
- **Fatores Institucionais/Sociais:** Os fatores institucionais são aqueles induzem o usuário a um comportamento adequado: informação correta e em tempo, sinalização adequada, atitudes coerentes por parte dos operadores da rodovia (policiamento, entre outros). Fatores sociais são aqueles que decorrem do contexto cultural e que se manifestam desde o meio em que o cidadão vive e cresce até a educação formal que recebeu, inclusive para o trânsito.

2.2.1 Acidentes no Brasil

O relatório do Departamento Nacional de Trânsito (2006) mostra que a malha rodoviária federal tem uma extensão oficial de 72.710 km, dos quais 79,7% são pavi-

mentadas, 20,3% das rodovias não pavimentadas e apenas 4,8% possuem pista dupla.

TABELA 2.1: EXTENSÃO DA REDE RODOVIÁRIA FEDERAL DO BRASIL

Características da via	Km	%	%
Total	72.710	100	-
Total pavimentada	57.933,1	79,7	100
Pista simples	53.632,0	73,8	92,6
Pista dupla	3.477,9	4,8	6,0
Em obras de duplicação	823,2	1,1	1,4
Total não-pavimentada	14.776,9	20,3	100
Leito natural	1.543,3	2,1	10,4
Em obras de implantação	306,2	0,4	2,1
Implantada	8.392,7	11,5	56,8
Em obras de pavimentação	4.534,7	6,2	30,7

FONTE: Retirado do IPEA, DENATRAN (2006)

Analisando os acidentes em números absolutos, no ano de 2004, ocorreu uma média de 307 acidentes por dia (aproximadamente 13 acidentes por hora) nas rodovias federais. Os acidentes envolveram em média 1,7 veículos e 4,07 pessoas. Em 2005 houve uma redução de 2,4% em relação a 2004, no entanto, o número de mortos não acompanhou essa redução, houve um acréscimo de 3,8%, mostrando que os acidentes apresentaram uma gravidade maior, embora o número de pessoas envolvidas em acidentes tenha decrescido 9,3%.

A tabela 2.2, mostra que as pessoas envolvidas em acidentes e classificadas como feridas pela PRF, reduziu 0,1% no ano de 2005 em relação a 2004. Os dados ainda revelam que 6,7% dos envolvidos classificados como ileso se revelaram feridos posteriormente. Do mesmo modo, 6,2% dos feridos no local foram a óbito. Ou seja, houve um acréscimo em torno de 66% no número de mortos declarados no Datatran³. Assim, haveria 4.067 mortos a mais em 2004 e 4.064 mortos a mais em 2005.

Analisando a tabela 2.2 ainda, pode-se observar que o número de mortos a cada 1.000 acidentes aumentou de 90,6 para 94,9, entre 2004 e 2005, incluindo

³O Datatran é um subsistema do Sistema BR-BRASIL, do Núcleo de Estatística do Departamento de Polícia Rodoviária Federal, Ministério da Justiça

TABELA 2.2: ESTATÍSTICAS DOS ACIDENTES DE TRÂNSITO NAS RODOVIAS FEDERAIS EM 2004 E 2005, NÚMERO DE ACIDENTES, NÚMERO DE ENVOLVIDOS POR GRAVIDADE DAS LESÕES, MORTOS, FERIDOS E ILESOS A CADA 1.000 ACIDENTES

Item	2004	2005	Varição
Total de acidentes	112.457	109.745	-2,4%
nº de veículos envolvidos	190.731	184.458	-3,3%
nº de pessoas envolvidas	457.409	414.663	-9,3%
nº de mortes local	6.119	6.352	3,8%
nº de mortes pós-acidente	4.067	4.064	-0,1%
Total mortes	10.186	10.416	2,3%
Mortes local/1.000 acidentes	54,4	57,9	3,5
Mortes pós-acidente	36,2	37,0	0,9
Mortes total/1.000 acidentes	90,6	94,9	4,3
nº acidentes/morte	11,0	10,5	-0,5
nº acidentes/morte local	18,4	17,3	-1,1
nº feridos local	66.117	66.066	-0,1%
nº feridos pós-acidente	25.678	22.816	-11,1%
Total de feridos	87.728	84.818	-3,3%
Feridos local/1.000 acidentes	587,9	602,0	14,1
Feridos pós-acidente/1.000 acidentes	228,3	207,9	-20,4
Feridos total/1.000 acidentes	780,1	772,9	-7,2
nº acidentes/ferido	1,3	1,3	0
nº ileso local	385.173	342.245	-11,1%
Total ileso	359.495	319.429	-11,1%
Ilesos local/1.000 acidentes	3.425	3.119	-306,5
Ilesos local/1.000 acidentes	3.197	2.911	-286,1
nº acidentes/ileso	0,3	0,3	0

FONTE: Retirado do IPEA, DENATRAN (2006)

aqueles que foram registrados como feridos e que vieram a óbito no hospital. Por outro lado, o índice de feridos a cada 1.000 acidentes diminuiu de 780,1, em 2004, para 772,9, em 2005, incluindo as pessoas que foram registradas como ileso e que depois foram classificadas como feridos.

Nos Gráficos das figuras 2.5 e 2.6, esses índices estão dispostos, de forma a comparar mortos e feridos a cada 1.000 acidentes em 2004 e 2005 e mortos e feridos nos acidentes de trânsito ocorridos nas rodovias federais nos anos de 2004 e 2005, respectivamente.

Dentre os fatores que influenciam nos acidentes, existem alguns que ocorrem mais comumente em relação a outros. O Datatran fez os registros desses fatores,

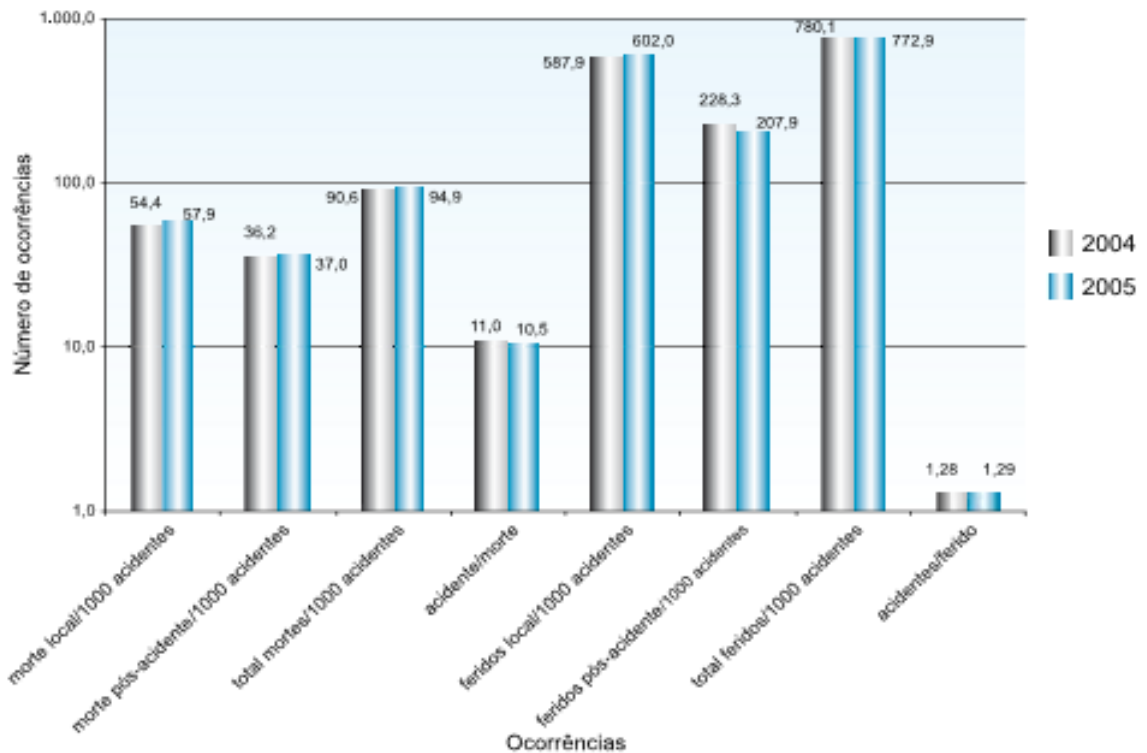


FIGURA 2.5: MORTOS E FERIDOS A CADA 1.000 ACIDENTES NAS RODOVIAS FEDERAIS EM 2004 E 2005

FONTE: Retirado do IPEA, DENATRAN (2006)

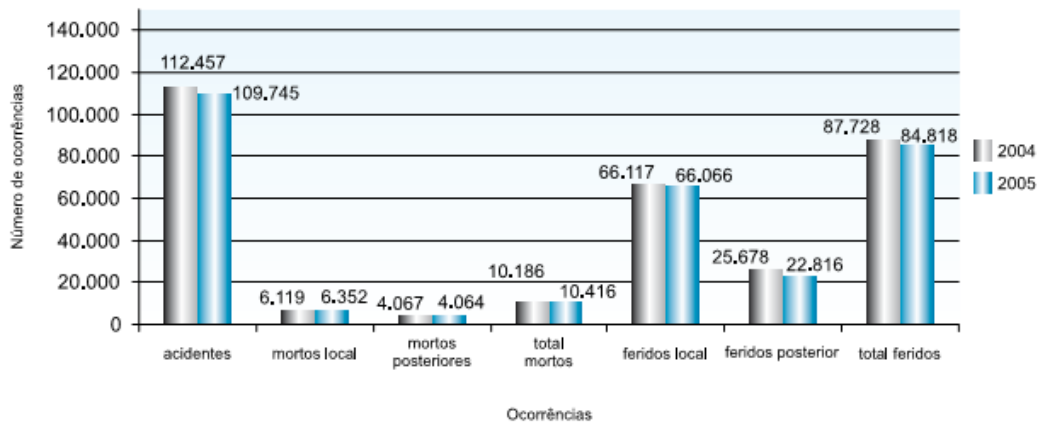


FIGURA 2.6: MORTOS E FERIDOS NOS ACIDENTES DE TRÂNSITO OCORRIDOS NAS RODOVIAS FEDERAIS NOS ANOS DE 2004 E 2005.

FONTE: Retirado do IPEA, DENATRAN (2006)

considerando os acidentes ocorridos apenas no ano de 2004.

- **Quanto ao tipo de acidente:**

A colisão traseira foi o tipo de acidente com maior número de frequência (24%),

seguido da colisão lateral (17%), saída de pista (13%) e colisão com objeto fixo (9%), totalizando 71.246 acidentes dos 112.457 ocorridos em 2004 e, que representa 63,4% desse total.

- **Quanto aos fatores contribuintes:**

Os fatores humanos/comportamentais são os que aparecem em primeiro lugar, totalizando 65.608 acidentes, 58,3% do total. O veículo aparece na segunda posição com 3.756 acidentes, o que representa apenas 3,3% do total. Por fim, a via foi responsável por 3.879 acidentes, 3,4% do total.

- **Quanto às condições do tempo, às fases do dia e ao traçado da via**

Os dados mostram que a maioria dos acidentes ocorrem com o tempo bom, 65,5%. Em relação às fases do dia, o número de acidentes concentra-se em pleno dia, 53,9% e, o traçado pista com o maior número de acidentes, ocorre nos trechos em linha reta, 71,3%.

Os dados apresentados pela Polícia Rodoviária Federal reforçam a afirmação de Branco, A. M. (1999) que a maioria dos acidentes ocorrem por falha humana. Ainda, mostram que a falta de atenção é a principal causa dos acidentes nas rodovias federais, e que, provavelmente, está ligada a distração, cansaço, stress, entre outros.

As tabelas 2.5, 2.6 e 2.7 mostram que a maioria os acidentes ocorrem em tempo bom e em pleno dia. Estes ainda, concentram-se nos trechos em linha reta. O fato de ocorrerem mais acidentes em pleno dia, deve-se ao grande número de veículos que transitam nas rodovias federais.

2.3 A RODOVIA BR-116

A Rodovia Federal BR-116, como pode ser observado no mapa da figura 2.7, é uma rodovia longitudinal que tem início em Fortaleza, estado do Ceará, e término na cidade de Jaguarão, estado do Rio Grande do Sul. A extensão total é de aproximadamente 4.385 quilômetros passando por dez estados brasileiros: Ceará, Paraíba,

TABELA 2.3: ACIDENTES DE TRÂNSITO POR TIPO NAS RODOVIAIS FEDERAIS - 2004

Tipo de acidente	Número de acidentes	%	% acumulada
COLISÃO TRASEIRA	27.124	24,1	24,1
COLISÃO LATERAL	19.118	17,0	41,1
SAÍDA DE PISTA	14.574	13,0	54,1
COLISÃO COM OBJETO FIXO	10.430	9,3	63,4
Colisão transversal	7.982	7,1	70,5
Capotamento	7.353	6,5	77,0
Tombamento	6.418	5,7	82,7
Outras	6.311	5,6	88,3
Colisão frontal	4.538	4,0	92,3
Atropelamento pedestre	3.996	3,6	95,9
Atropelamento animal	3.479	3,1	99,0
Engavetamento	1.134	1,0	100,0
Total	112.457	100,0	

FONTE: Retirado do IPEA, DENATRAN (2006)

TABELA 2.4: ACIDENTES DE TRÂNSITO POR FATORES CONTRIBUENTES-2004

Fatores contribuintes	Número de acidentes	%
Componentes humanos/comportamentais	65.608	58,3
Falta de atenção	31.736	28,2
Velocidade incompatível	12.439	11,1
Distância de segmento	9.774	8,7
Desobediência à sinalização	4.971	4,4
Ultrapassagem indevida	3.862	3,4
Dormindo	1.923	1,7
Ingestão de álcool	903	0,8
Veículo	3.756	3,3
Defeito mecânico em veículo	3.756	3,3
Via	3.879	3,4
Buraco na via	2.313	2,1
Defeito na via	1.566	1,4
Outras causas	39.214	34,9
Total	112.457	100,0

FONTE: Retirado do IPEA, DENATRAN (2006)

Pernambuco, Bahia, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul⁴.

No estado do Ceará, a rodovia BR-116 começa em Fortaleza e passa por cidades como Itaitinga, Horizonte, Pacajus, Chorozinho, Limoeiro do Norte, entre ou-

⁴Texto retirado do wikipédia. Disponível em:<<http://pt.wikipedia.org/wiki/BR-116>>.

TABELA 2.5: ACIDENTES DE TRÂNSITO SEGUNDO AS CONDIÇÕES DO TEMPO - 2004

Condição do tempo	Número de acidentes	%
Bom	73.667	65,5
Nublado	11.629	10,3
Neblina	1.686	1,5
Chuva	22.466	20,0
Outros	3.009	2,7
Total	112.457	100,0

FONTE: Retirado do IPEA, DENATRAN (2006)

TABELA 2.6: ACIDENTES DE TRÂNSITO SEGUNDO FASES DO DIA

Fases do dia	Número de Acidentes	%
Amanhecer	7.753	6,9
Pleno dia	60.597	53,9
Anoitecer	7.377	6,6
Noite	36.730	32,7
Total	112.457	100,0

FONTE: Retirado do IPEA, DENATRAN (2006)

TABELA 2.7: ACIDENTES DE TRÂNSITO SEGUNDO O TRAÇADO DA VIA

Traçado da via	Número de acidentes	%
Reta	80.140	71,3
Curva	24.346	21,6
Cruzamento	7.971	7,1
Total	112.457	100,0

FONTE: Retirado do IPEA, DENATRAN (2006)

tros, e termina na cidade de Penaforte. São aproximadamente 544,5 quilômetros de malha asfáltica sendo a saída de Fortaleza duplicada até o km 50 no município de Pacajus e o restante do trecho com duas vias em sentido duplo com acostamento em ambos os lados até a divisa do estado do Ceará com o estado de Pernambuco. Um pequeno trecho, de pouco mais de 8 quilômetros, passa pelo estado da Paraíba, na cidade de Cachoeira dos Índios, devido a topografia da região. Em seguida, a rodovia continua pelo estado do Ceará até a divisa com Pernambuco.

No estado de Pernambuco, a rodovia passa pelos municípios de Salgueiro e Cabrobó, num total de 91,6 quilômetros. Neste trecho, há apenas pista simples com



FIGURA 2.7: RODOVIA FEDERAL BR-116

FONTE: Ministério dos Transportes (2010)

sentido duplo.

Seguindo pela rodovia, o próximo trajeto encontra-se no estado Bahia. O trecho neste estado tem aproximadamente 950 quilômetros de malha asfáltica e passa por cidades como: Abaré, Chorrochó, Euclides da Cunha, Feira de Santana, Poções, Vitória da Conquista e termina na cidade de Cândido Sales. No geral, a pista encontra-se em boas condições e bem sinalizada, exigindo atenção ao motorista em relação ao desnível de acostamento e cuidados com a fiscalização ao longo da rodovia. No



FIGURA 2.8: TRECHO DA RODOVIA BR-116 EM FORTALEZA-CE

FONTE: Wikipédia (2010)

município de Jaguaquara encontra-se a Serra do Mutum, ladeira bastante perigosa onde ocorrem muitos acidentes.

No estado de Minas Gerais, a BR-116 passa pelos municípios de Divisa Alegre, Águas Vermelhas, Governador Valadares, Engenheiro Caldas, Santa Bárbara do Leste, São João do Manhuaçu, Miradouro, Laranjal, entre outros. No geral a pista encontra-se em boas condições e sinalização razoável para aceitável. Entretanto, exige paciência e atenção do motorista, por apresentar pista estreita, muito sinuosa, falta de acostamento e pontos de ultrapassagem com deficiências de engenharia e operacionais, com pontos onde a distância de liberação de ultrapassagem somente permite ultrapassagens entre dois veículos pequenos.

No estado do Rio de Janeiro, a partir do km 0 (zero) até km 144,5 a BR 116 recebe o nome de Rodovia Santos-Dumont⁵. A rodovia passa pelos municípios de Teresópolis, Guapimirim, Magé, Duque de Caxias (onde termina o trecho denominado Rodovia Santos-Dumont), Rio de Janeiro (onde se inicia o trecho denominado Rodovia Presidente Dutra), São João de Meriti, Belford Roxo, Mesquita (servindo de divisa en-

⁵O nome da rodovia foi escolhido em homenagem ao grande aviador brasileiro Alberto Santos Dumont (1873-1932) que estudou na cidade de Campinas e que é considerado o *Pai da aviação*.

tre estes dois municípios), Nova Iguaçu, Queimados, Seropédica, Itaguaí, Paracambi (servindo de divisa entre estes dois municípios), Piraí, Pinheiral, Volta Redonda, Barra Mansa, Porto Real, Resende, Itatiaia e novamente Resende, chegando na divisa com o estado de São Paulo.

Em São Paulo, entre o km zero e a marginal Tietê é a continuação da Rodovia Presidente Dutra. A partir do município de São Paulo, em direção à Curitiba, representa a rodovia Régis Bittencourt, onde passa por grandes trechos de mata Atlântica preservada.



FIGURA 2.9: VIA DUTRA NO TRECHO ENTRE TAUBATÉ - SP E PINDAMONHANGABA - SP
FONTE: Wikipédia (2010)

No Paraná, a rodovia também é conhecida como rodovia Régis Bittencourt⁶. Neste estado, a rodovia tem início na divisa com o estado de São Paulo atravessando toda a cidade Curitiba (capital paranaense) e terminando na cidade de Rio Negro. A extensão total no estado do Paraná é de aproximadamente 211 quilômetros. Por ser um dos meios de ligação norte-sul mais importante do país, o aumento do número de veículos e, como consequência, o aumento dos acidentes atinge números preocupantes.

⁶Edmundo Régis Bittencourt foi um engenheiro civil brasileiro, empreiteiro e ex-presidente do Departamento Nacional de Estradas de Rodagem (DNER). Em sua homenagem, a rodovia BR-116 no trecho entre São Paulo e Curitiba, foi designada Régis Bittencourt.



FIGURA 2.10: TRECHO DA RODOVIA RÉGIS BITTENCOURT PRÓXIMO A REGISTRO - SP
FONTE: Wikipédia (2010)



FIGURA 2.11: TRECHO DA RODOVIA BR-116 PRÓXIMO A CIDADE DE CURITIBA - PR
FONTE: Wikipédia (2010)

Em Santa Catarina, a rodovia atravessa a região serrana do estado. Inicia em Mafra (divisa com o Paraná), entrocando-se ainda nesta com a BR-280, com a BR-470 em São Cristóvão do Sul, cortando o município de Lages, e neste entroncando-se com a BR-282.



FIGURA 2.12: TRECHO DA RODOVIA BR-116 EM SANTA CATARINA, PRÓXIMO À DIVISA COM O RIO GRANDE DO SUL

FONTE: Wikipédia (2010)

No Rio Grande do Sul, a rodovia faz entroncamento com a BR-285 no km 37, próximo a cidade de Vacaria. Atravessa a cidade de Caxias do Sul. No km 237 inicia o trecho duplicado. O trecho entre Dois Irmãos e Porto Alegre é considerado um dos mais movimentados do país, com tráfego diário em torno de 120 mil veículos. A rodovia ainda passa por importantes cidades gaúchas como: Novo Hamburgo, Canoas e Pelotas. O trecho duplicado da rodovia encerra na cidade de Eldorado do Sul. A rodovia termina na cidade de Jaguarão.

2.3.1 Acidentes na BR-116

A BR-116 é uma rodovia de grande extensão, é uma das rodovias onde ocorrem mais acidentes, em todo o Brasil. Conforme mostra a tabela 2.8, na BR-116 ocorreram 23.135 acidentes no ano de 2004. Isto representa quase 20,6% do total de acidentes registrados pela PRF nas rodovias federais de todo o Brasil.

O número de mortes, nesta rodovia, também impressiona, são 1.030 mortes no total de todos os acidentes ocorridos no ano de 2004. A BR-116 somente perde para a BR-101 nesta tragédia do trânsito.

No *ranking* das rodovias por gravidade (acidente/quilômetro) e periculosidade (acidente com fatalidades) a BR-116 está em quinto lugar, considerando os dados registrados no ano de 2004. Os trechos mais perigosos da BR-116, ou seja, com maior índice de periculosidade e gravidade encontram-se nos estados de São Paulo, do km 200 ao km 250, e Rio Grande do Sul, do km 250 ao km 300.

TABELA 2.8: ESTATÍSTICAS DOS ACIDENTES NAS TREZE RODOVIAS FEDERAIS COM MAIOR NÚMERO DE MORTES

BR	UF	Extensão (km)	Acidentes	Mortos	Acidentes/km	Mortos/km
BR-101	RN-PB-PE-SE-BA-ES-RJ-SP-PR-SC-RS	4.620	19.996	1.060	4,32	0,23
BR-116	CE-PB-PE-BA-MG-RJ-SP-PR-SC-RS	4.594	23.135	1.030	5,03	0,22
BR-040	DF-GO-MG-RJ	1.180	6.219	302	5,27	0,25
BR-381	ES-MG-SP	1.255	7.420	297	5,91	0,23
BR-153	PA-TO-GO-MG-SP-PR-SC-RS	4.355	3.645	237	0,83	0,05
BR-364	SP-MG-GO-MT-RO-AC	4.409	3.123	219	0,71	0,05
BR-316	PA-MA-PI-PE-AL	2.069	2.691	198	1,3	0,1
BR-262	ES-MG-SP-MS	2.299	3.661	169	1,6	0,07
BR-163	RS-SC-PR-MS-MT-PA	4.507	1.887	157	0,42	0,03
BR-277	PR	795	2.847	125	3,58	0,15
BR-230	PB-CE-PI-MA-TO-PA-AM	4.982	1.877	124	0,37	0,02
BR-222	CE-PI-MA-PA	1.841	964	123	0,52	0,06
BR-060	DF-GO-MS	1.442	1.650	108	1,14	0,07

FONTE: Retirado do IPEA, DENATRAN (2006)



FIGURA 2.13: RODOVIA FEDERAL BR-116 NO ESTADO DO PARANÁ

FONTE: O autor (2010)

A BR-116, no estado do Paraná (mapa da figura 2.13), foram registrados no ano de 2007, 1.515 acidentes, no ano de 2008, 2.393 acidentes e no ano de 2009,

2.337 acidentes, totalizando 6.245 acidentes. Os registros foram feitos pelo Boletim de Acidente de Trânsito (BAT) da PRF/PR.

A rodovia, no estado do Paraná, possui aproximadamente 118⁷ quilômetros de estrada com pista dupla e o restante em pista simples. Analisando os acidentes com relação ao tipo de pista, conforme o gráfico da figura 2.14, a maioria dos acidentes ocorreram em pista dupla (71%). Isto é evidente pela maior intensidade de tráfego e aumento de velocidade, dada as melhores condições da rodovia.

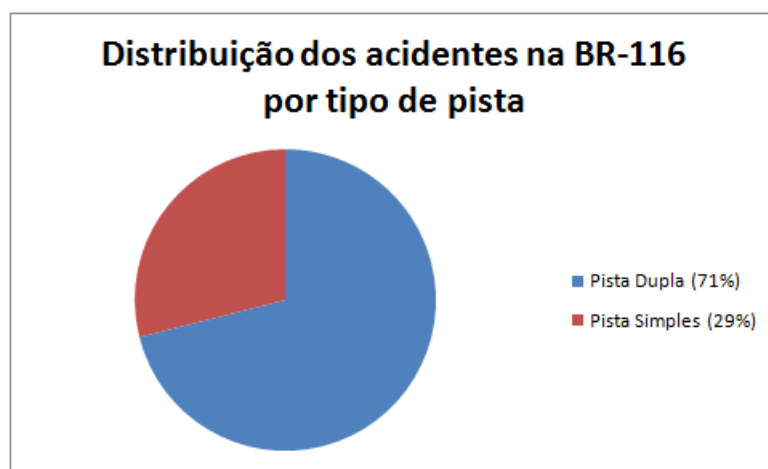


FIGURA 2.14: ACIDENTES NA BR-116 (PARANÁ) POR TIPO DE PISTA

FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

Os tipos de acidentes que mais ocorreram foram, saída de pista (24,2%), seguido por colisão traseira (21,7%), colisão lateral (12,6%), colisão com objeto fixo (9%) e colisão transversal (8,8%), perfazendo 76,3% do total de acidentes. A tabela 2.9 apresenta os resultados referentes aos tipos de acidentes ocorridos no trecho do estado do Paraná, da BR-116, de 2007 à 2009.

Quanto as fases do dia em que mais acontecem acidentes, pode-se destacar: pleno dia (55,7%) e plena noite (33,8%) como os períodos em que mais ocorrem acidentes. Somados os dois períodos, representa 89,5% do total de acidentes registrados. Em termos absolutos, foram registrados 5.587 acidentes nestes dois períodos do dia. Os acidentes restantes dividem-se nos períodos do anoitecer e amanhecer,

⁷Fonte: Banco de Informações e Mapas dos Transportes - Ministério dos Transportes. Disponível em: <<http://www2.transportes.gov.br/bit/trodo/br-116/gbr-116.htm>>

conforme o gráfico da figura 2.15.

TABELA 2.9: ACIDENTES NA BR-116 (PARANÁ) POR FATOR CONTRIBUINTE

Tipo de Acidente	Quantidade	%	% acumulada
Saída De Pista	1514	24,2	24,20
Colisão Traseira	1355	21,7	45,90
Colisão Lateral	788	12,6	58,50
Colisão Com Objeto Fixo	559	9,0	67,50
Colisão Transversal	549	8,8	76,30
Capotamento	322	5,2	81,50
Tombamento	268	4,3	85,80
Atropelamento De Pessoa	244	3,9	89,70
Queda De Motocicleta / Bicicleta	168	2,7	92,40
Colisão Frontal	142	2,3	94,70
Colisão Com Objeto Móvel	78	1,2	95,90
Derramamento De Carga	66	1,1	97,00
Colisão Com Bicicleta	64	1,0	98,00
Atropelamento De Animal	55	0,9	98,90
Danos Eventuais	52	0,8	99,70
Incêndio	21	0,3	100
Total	6245	100	

FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

Os dados analisados ainda mostram que a grande maioria dos acidentes na BR-116, no Paraná, ocorreram em trechos em linha reta, cerca de 66%. Os trechos em curva registram 31% dos acidentes, conforme o gráfico da figura 2.16.

As condições do tempo (TABELA 2.10) estabelecem que 32,1% dos acidentes ocorreram em dias em que o céu estava claro, 31,4% com chuva, 24,1% com o tempo o nublado, perfazendo 87,6 % do total registrado.

As principais causas (TABELA 2.11) de todos os acidentes ocorridos na BR-116 foram: falta de atenção (30%) e velocidade incompatível (19,3%). Outras causas não detalhadas somam 31,7% dos acidentes.

2.4 CUSTOS DOS ACIDENTES

O Instituto de Pesquisa Econômica Aplicada (IPEA) estruturou a função básica dos acidentes de trânsito definida para estimar os impactos econômicos dos acidentes

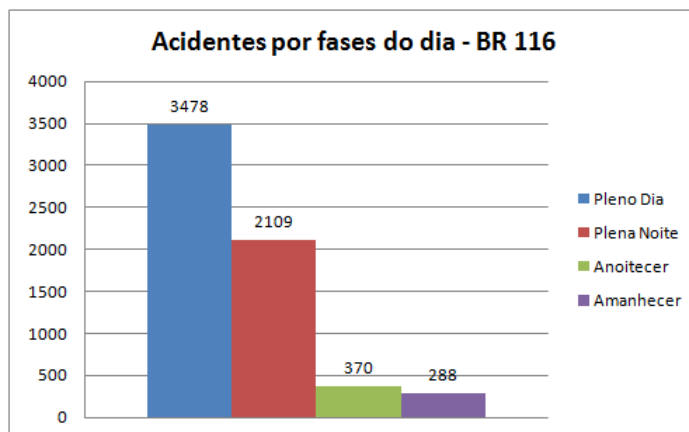


FIGURA 2.15: DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 SEGUNDO AS FASES DO DIA
FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

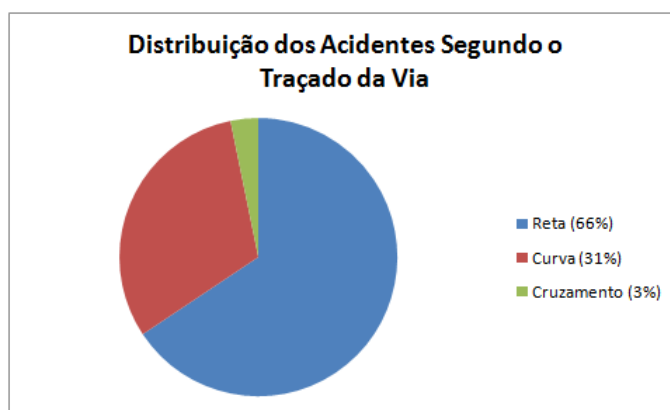


FIGURA 2.16: DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 SEGUNDO O TRAÇADO DA VIA

FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

TABELA 2.10: ACIDENTES DE TRÂNSITO SEGUNDO AS CONDIÇÕES DO TEMPO NA BR-116, ESTADO DO PARANÁ

Condição Meteorológica	Quantidade	%
Ceu Claro	2003	32,1
Chuva	1964	31,4
Nublado	1504	24,1
Sol	533	8,5
Nevoeiro/neblina	164	2,6
Ignorada	67	1,1
Granizo	10	0,2
Total	6245	

FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

TABELA 2.11: ACIDENTES DE TRÂNSITO POR CAUSAS NA BR-116, ESTADO DO PARANÁ

Causas de Acidentes	Quantidade	%
Outras	1981	31,7
Falta De Atenção	1872	30,0
Velocidade Incompatível	1207	19,3
Não Guardar Distância De Segurança	291	4,7
Ingestão De Álcool	237	3,8
Defeito Mecânico Em Veículo	204	3,3
Desobediência À Sinalização	175	2,8
Dormindo	109	1,7
Ultrapassagem Indevida	77	1,2
Animais Na Pista	67	1,1
Defeito Na Via	25	0,4
Total	6245	

FONTE: Polícia Rodoviária Federal do Paraná - BAT, 2007 à 2009

nas rodovias brasileiras (Departamento Nacional de Trânsito, 2006). A função ficou composta por quatro grupos de componentes de custos relativos:

- às pessoas;
- aos veículos;
- à via e ao ambiente onde ocorre o acidente;
- às instituições públicas.

A função global que mede os custos dos acidentes de trânsito nas rodovias brasileiras é definida da seguinte maneira:

$$C_{\text{acidente}} = C_{\text{pessoas}} + C_{\text{veículos}} + C_{\text{via/ambiente}} + C_{\text{institucionais}}$$

onde, C_{pessoas} é composto por três parcelas, ou seja,

$$C_{\text{pessoas}} = C_{\text{cuidados em saúde}} + C_{\text{perda de produção}} + C_{\text{remoção/translado}}$$

$C_{\text{veículos}}$ é composto por quatro parcelas a seguir detalhadas:

$$C_{\text{veículos}} = C_{\text{danos materiais}} + C_{\text{perda de carga}} + C_{\text{remoção/guincho}} +$$

$C_{\text{reposição}}$

$$C_{\text{via/ambiente}} = C_{\text{propriedade pública}} + C_{\text{propriedade privada}}$$

$$C_{\text{institucionais}} = C_{\text{atendimento}} + C_{\text{judiciais}}$$

Os conceitos de cada componente dos quatro grupos da função custo global dos acidentes são definidos por:

- **Custos associado às pessoas:**

1. Custo do atendimento pré-hospitalar: atendimento da vítima por unidades dotadas de equipamentos especiais, com veículos e profissionais especializados (ambulâncias, bombeiros, médicos, etc.).
2. Custo do atendimento hospitalar: soma dos custos do atendimento médico hospitalar do paciente não internado e do paciente internado na unidade de terapia intensiva e/ou enfermaria.
3. Custo pós-hospitalar: a soma dos custos com reabilitação, para os casos de seqüela temporária ou definitiva, com procedimentos, medicamentos, transporte, equipamentos e outros.
4. Custo da perda de produção: é o custo correspondente às perdas econômicas das vítimas de acidente que, em decorrência da interrupção das suas atividades produtivas, deixam de gerar renda e produção ao sistema econômico.

5. Custo de remoção/translado: é a soma dos custos de remoção da vítima fatal ao Instituto Médico Legal (IML), e também, do traslado (terrestre ou aéreo) da vítima fatal do IML/hospital ao local do funeral.
6. Custo Previdenciário: é a soma dos custos relacionados: i) à empresa, que paga valores à previdência em um período de 15 dias de afastamento do trabalho em decorrência de um acidente de trânsito; ii) à previdência social, em virtude do afastamento, temporário ou definitivo, do trabalhador em decorrência de um acidente de trânsito; iii) As seguradoras - seguro DPVAT (DPVAT).

- **Custos associados aos veículos:**

1. Custos dos danos materiais aos veículos: custo de recuperação dos veículos danificados em acidentes de trânsito;
2. Custo de perda de carga: o custo de avaria da carga que estava no veículo envolvido no acidente;
3. Custo de remoção/pátio de veículos: custo de remoção do veículo e diárias de pátio de armazenamento.
4. Custo de reposição: despesa incorrida pela substituição do veículo, no período em que ele ficou sem condições de uso.

- **Custos institucionais**

1. Custo do atendimento policial: soma dos custos do tempo dos policiais rodoviários, da utilização de veículos para atendimento no local do acidente e do deslocamento para hospital ou delegacia.
2. Custo de processos judiciais: custo do funcionamento da estrutura judicial em função do atendimento às questões referentes aos acidentes de trânsito.

- **Custos associados à via e ao ambiente do local de acidente:**

1. Custo dos danos à propriedade pública: custo de reposição/recuperação de imobiliário ou equipamentos danificados ou destruídos em função de acidentes nas rodovias.
2. Custo dos danos à propriedade privada: custo de recuperação de propriedades particulares danificadas em função de acidentes de trânsito.

Existem outros custos "não-valorados" ou sem tradução monetária dos acidentes de trânsito, mas que quando calculados superam os demais. São custos decorrentes de perda de vidas ou de lesões permanentes que impossibilitam uma vida normal, que atingem tanto as pessoas envolvidas em acidentes de trânsito quanto sobre as pessoas de suas relações.

2.5 ANÁLISE DE CORRELAÇÃO CANÔNICA

2.5.1 Introdução

A Análise de correlação canônica é uma técnica estatística multivariada que tem como objetivo principal expressar a quantidade de relações existentes entre dois conjuntos de variáveis, ou seja, é uma técnica para a identificação e quantificação da associação entre dois grupos de variáveis (RENCHEER, 2002). Conforme Johnson e Wichern (2002), a ideia desta técnica é determinar os pares de combinações lineares $U_i = \underline{a}'\underline{X}$ e $V_i = \underline{b}'\underline{Y}$, $i = 1, 2, \dots, p$, tais que tenham a maior correlação possível.

Segundo Chaves Chaves Neto (2002), a análise de correlação canônica é considerada uma extensão da análise de regressão múltipla. Na análise de regressão múltipla, as variáveis formam o conjunto das variáveis explicativas $[x_1, x_2, \dots, x_{p-1}]$ componentes do vetor \underline{X} ($p - 1$ variáveis independentes) e a variável resposta \underline{Y} (variável única dependente) forma outro conjunto unitário. Assim, o problema na análise de regressão consiste em buscar a combinação linear $\underline{\beta}'\underline{X}$ que é altamente correlacionada com a variável \underline{Y} . Por outro lado, na análise de correlação canônica tem-se o conjunto $[y_1, y_2, \dots, y_q]$ com $q \geq 1$ variáveis, que compõem o vetor \underline{Y} . Portanto, têm-se dois

vetores aleatórios \underline{X} e \underline{Y} , devendo-se achar os vetores \underline{a}' e \underline{b}' , que formam as combinações lineares, para os quais a correlação entre $U = \underline{a}'\underline{X}$ e $V = \underline{b}'\underline{Y}$ seja máxima.

Segundo Hair e Anderson (2002), a correlação canônica está organizada em alguns estágios, descritos a seguir:

(i) Objetivos da análise de correlação canônica: Os dados apropriados à análise de correlação canônica são dois conjuntos de variáveis. Esses conjuntos podem ser definidos como: um dependente e outro independente. Assim, a correlação canônica pode ter vários objetivos. Estes objetivos podem ser:

1. Determinar se dois conjuntos de variáveis (medidas feitas sobre os mesmos objetos) são independentes um do outro ou, conseqüentemente, determinar a magnitude das relações que possam existir entre os dois conjuntos.
2. Determinar um conjunto de pesos para cada conjunto de variáveis dependentes e independentes de modo que as combinações lineares obtidas em cada conjunto possuam máxima correlação. Combinações lineares adicionais que maximizam a correlação permanente não são independentes dos conjuntos anteriores de combinações lineares.
3. Verificar a existência e explicar quaisquer relações entre os dois conjuntos de variáveis analisados, pela medida da contribuição relativa de cada variável às funções canônicas (relações) extraídas.

(ii) Planejamento de uma análise de correlação canônica: fatores importantes na aplicação da correlação canônica são o tamanho da amostra e a necessidade de um número suficiente de observações por variável. Amostras relativamente pequenas não irão representar bem as correlações, pois podem esconder relações significativas existentes. Já amostras muito grandes, podem não identificar a significância estatística na prática. Sendo assim, o número de observações recomendado para cada variável é de 10 observações.

A classificação dos conjuntos de variáveis, em dependentes ou independentes, é

de pouca importância para a estimação da função canônica, já que a correlação canônica pondera ambos os grupos de variáveis para maximizar a correlação entre os mesmos e não prioriza nenhuma variável estatística em particular. No entanto, os conjuntos de variáveis devem ser ligados conceitualmente antes de se aplicar a análise de correlação canônica.

- (iii)** Hipóteses na correlação canônica: dois aspectos influenciam nos resultados da correlação canônica - o primeiro, o coeficiente de correlação entre duas variáveis quaisquer é baseado em uma relação linear. Se a relação é não-linear, então uma ou ambas as variáveis devem ser transformadas, se possível - o segundo, a correlação canônica é a relação linear entre as variáveis estatísticas. Se essa relação não for linear, a correlação canônica não ocorrerá. Ainda, todas as variáveis devem ser testadas quanto à normalidade, homocedasticidade e multicolinearidade para explicar uma maior correlação entre as variáveis.
- (iv)** Determinação das funções canônicas e avaliação geral: o primeiro passo da análise de correlação canônica é derivar uma ou mais funções canônicas. Cada função canônica consiste em um par de variáveis, uma representando as variáveis independentes e a outra, as dependentes. O número máximo de funções canônicas que podem ser obtidas dos grupos de variáveis é igual ao número de variáveis no menor conjunto de dados, independentes ou dependentes. Para saber quais funções canônicas devem ser interpretadas, deve-se analisar o nível de significância (p), que deve ser menor que 0,05 (índice normalmente usado), na magnitude da correlação canônica e no índice de redundância.
- (v)** Interpretação da variável estatística canônica: se a correlação canônica é estatisticamente significativa e as magnitudes da raiz canônica e do índice de redundância são aceitáveis, ainda há o que se interpretar dos resultados. Isso envolve determinar a importância relativa de cada variável original nas relações canônicas. Três métodos são mais utilizados: pesos canônicos, cargas canônicas e cargas cruzadas canônicas. Para mais detalhes sobre os 3 (três) métodos apre-

sentados anteriormente, o leitor pode encontrar em (HAIR; ANDERSON, 2002), página 369.

São exemplos citados por Marques (2010) em que se pode aplicar a análise de correlação canônica:

- Um pesquisador educacional deseja determinar a relação entre três medidas de habilidade escolar com de cinco medidas de sucesso na escola.
- Um sociólogo deseja investigar a relação entre dois preditores de mobilidade social baseado em entrevistas, com a mobilidade social atual medida por quatro diferentes indicadores.
- Um médico deseja estudar a relação entre vários fatores de risco para o desenvolvimento de um grupo de sintomas.
- Um pesquisador da área de comportamento deseja estudar a relação entre a natureza do trabalho e a satisfação do trabalhador em um grande grupo de executivos.

2.5.2 Alguns Trabalhos Relacionados

Trugilho, Lima e Mori (2003) realizaram um estudo para identificar as características químicas, físicas e dimensionais da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. Foram avaliados sete tipos de clones de *Eucalyptus grandis* e três tipos de clones de *Eucalyptus saligna* considerando as seguintes características: densidade básica, densidade seca, teor de extrativos totais, teor de lignina solúvel e insolúvel, teor de cinzas, comprimento da fibra, espessura da parede celular, largura da fibra, contração radial, tangencial e volumétrica. As conclusões obtidas pela aplicação do método permitiram observar que os grupos considerados possuem uma forte correlação considerando um nível de significância igual a 1%. As associações entre os grupos considerados também foram estabelecidas, como por exemplo: Madeira

de fibras largas, de parede mais espessa e de pequeno comprimento está associada com a madeira de alta contração radial, tangencial e volumétrica e de baixa densidade básica.

Basniak (2009) utilizou a Teoria de Rough Sets e Análise de Correlação Canônica para avaliar a correlação existente entre a disponibilidade de acesso aos recursos tecnológicos nas escolas e colégios estaduais do Núcleo Regional de Educação (NRE) de União da Vitória, estado do Paraná, e o desempenho dos alunos e a prática docente dos professores.

Costa (2008) realizou um trabalho cujo objetivo foi analisar, usando a correlação canônica, como os princípios da qualidade total ligados à valorização do lado humano nas empresas são influenciados pelo comportamento gerencial nas mesmas pela busca da qualidade total.

Rigão (2009) utilizou a correlação canônica como uma técnica estatística multivariada capaz de auxiliar na seleção precoce de clones de batata, identificando os caracteres, altamente relacionados, entre tubérculos plantados e sua produção.

2.5.3 Variáveis Canônicas e Correlações Canônicas

Segundo Chaves Chaves Neto (2002), muitas vezes há interesse em medir a associação entre 2 (dois) grupos de variáveis. Seja o primeiro grupo, composto por p variáveis, representado pelo vetor aleatório $\underline{X}(p \times 1)$ e, o segundo grupo, composto por q variáveis, representado pelo vetor aleatório $\underline{Y}(q \times 1)$, considerando $p \leq q$.

Para os vetores aleatórios \underline{X}_1 e \underline{Y}_2 , pode-se definir:

$$\begin{aligned} E(\underline{X}_1) &= \underline{\mu}_1, & V(\underline{X}_1) &= \Sigma_{11}, \\ E(\underline{Y}_2) &= \underline{\mu}_2, & V(\underline{Y}_2) &= \Sigma_{22}, \\ \text{Cov}(\underline{X}_1, \underline{Y}_2) &= \Sigma_{12} = \Sigma'_{21}. \end{aligned} \quad (2.1)$$

Considerando \underline{X}_1 e \underline{Y}_2 conjuntamente, tem-se o vetor aleatório \underline{X} de dimensão

$p + q$:

$$\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \dots \\ \underline{Y}_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \\ \dots \\ Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{bmatrix}, \quad (2.2)$$

com o vetor de médias

$$\underline{\mu} = E(\underline{X}) = \begin{bmatrix} E(\underline{X}_1) \\ \dots \\ E(\underline{Y}_2) \end{bmatrix} = \begin{bmatrix} \underline{\mu}_1 \\ \dots \\ \underline{\mu}_2 \end{bmatrix}, \quad (2.3)$$

e matriz de covariância de dimensão $p + q$

$$\begin{aligned} \Sigma &= E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})' = \\ &= \begin{bmatrix} E(\underline{X}_1 - \underline{\mu}_1)(\underline{X}_1 - \underline{\mu}_1)' & E(\underline{X}_1 - \underline{\mu}_1)(\underline{Y}_2 - \underline{\mu}_2)' \\ E(\underline{Y}_2 - \underline{\mu}_2)(\underline{X}_1 - \underline{\mu}_1)' & E(\underline{Y}_2 - \underline{\mu}_2)(\underline{Y}_2 - \underline{\mu}_2)' \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}. \end{aligned} \quad (2.4)$$

Assim, as covariâncias entre os pares de variáveis dos diferentes conjuntos - uma variável de \underline{X} e outra de \underline{Y} - estão contidas em Σ_{12} ou, equivalentemente, em Σ_{21} . Isto é, os pq elementos medem a associação entre os dois conjuntos de variáveis. Quando p e q são relativamente grandes, a interpretação dos elementos de Σ_{12} torna-se impraticável. Em consequência disso, as combinações lineares são introduzidas

para tornar possível a interpretação desejada.

Sejam as combinações lineares:

$$\begin{aligned} U &= \underline{a}'\underline{X}_1, \\ V &= \underline{b}'\underline{Y}_2, \end{aligned} \quad (2.5)$$

então

$$\text{Corr}(U, V) = \frac{\text{Cov}(U, V)}{\sqrt{V(U)V(V)}} = \frac{E[(U - \bar{U})(V - \bar{V})]}{\sqrt{V(U)V(V)}}, \quad (2.6)$$

onde:

$$\begin{aligned} V(U) &= \underline{a}'\text{Cov}(\underline{X}_1)\underline{a} = \underline{a}'\underline{\Sigma}_{11}\underline{a}, \\ V(V) &= \underline{b}'\text{Cov}(\underline{Y}_2)\underline{b} = \underline{b}'\underline{\Sigma}_{22}\underline{b}, \\ \text{Cov}(U, V) &= \underline{a}'\text{Cov}(\underline{X}_1, \underline{Y}_2)\underline{b} = \underline{a}'\underline{\Sigma}_{12}\underline{b}. \end{aligned} \quad (2.7)$$

Portanto, deve-se encontrar os vetores coeficientes \underline{a} e \underline{b} tais que:

$$\text{Corr}(U, V) = \frac{\underline{a}'\underline{\Sigma}_{12}\underline{b}}{\sqrt{\underline{a}'\underline{\Sigma}_{11}\underline{a}}\sqrt{\underline{b}'\underline{\Sigma}_{22}\underline{b}}}, \quad (2.8)$$

seja a maior possível.

O primeiro par de variáveis canônicas são as combinações lineares U_1 e V_1 , com variância unitária, e que maximiza a correlação 2.8.

O segundo par de variáveis canônicas são as combinações lineares U_2 e V_2 , com variância unitária, e que maximiza a correlação 2.8 dentre aquelas que não são correlacionadas com o primeiro par de variáveis canônicas.

Na k -ésima etapa, o k -ésimo par de variáveis canônicas são as combinações lineares U_k e V_k , com variância unitária, e que maximiza a correlação 2.8 entre aquelas que não são correlacionadas com os $k - 1$ primeiros pares de variáveis canônicas definidos.

A correlação entre as variáveis canônicas do k -ésimo par é denominada k -

ésima correlação canônica.

O resultado a seguir, de acordo com Johnson e Wichern (2002), fornece detalhes de como as variáveis canônicas e suas correlações são obtidas.

Resultado 1. Suponha que $p \leq q$ e sejam os vetores aleatórios de \underline{X} e \underline{Y} com $\text{Cov}(\underline{X}) = \Sigma_{11}$, $\text{Cov}(\underline{Y}) = \Sigma_{22}$, onde $\text{Cov}(\underline{X}, \underline{Y}) = \Sigma_{12}$. Considere as combinações lineares $U = \underline{a}'\underline{X}$ e $V = \underline{b}'\underline{Y}$. Então

$$\max_{\underline{a}, \underline{b}} \text{Corr}(U, V) = \rho_1^*, \quad (2.9)$$

é obtida pelas combinações lineares (primeiro par de variável canônica)

$$U_1 = \underline{e}'_1 \Sigma_{11}^{-\frac{1}{2}} \underline{X} \text{ e } V_1 = \underline{f}'_1 \Sigma_{22}^{-\frac{1}{2}} \underline{Y}. \quad (2.10)$$

O k -ésimo par de variáveis canônicas, $k = 1, 2, \dots, p$,

$$U_k = \underline{e}'_k \Sigma_{11}^{-\frac{1}{2}} \underline{X} \text{ e } V_k = \underline{f}'_k \Sigma_{22}^{-\frac{1}{2}} \underline{Y}, \quad (2.11)$$

maximiza

$$\text{Corr}(U_k, V_k) = \rho_k^*, \quad (2.12)$$

entre as combinações lineares não correlacionadas com as precedentes $1, 2, \dots, k-1$, variáveis canônicas.

Considera-se aqui $\rho_1^{*2} \geq \rho_2^{*2} \geq \dots \geq \rho_p^{*2}$ como sendo os autovalores de

$$\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}}$$

e $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_p$ são os $(p \times 1)$ autovetores associados. Os valores de $\rho_1^{*2}, \rho_2^{*2}, \dots, \rho_p^{*2}$ são também os p autovalores da matriz $\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}}$ com os $(p \times 1)$ autovetores correspondentes $\underline{f}_1, \underline{f}_2, \dots, \underline{f}_p$. Cada vetor \underline{f}_i é proporcional à matriz $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}} \underline{e}_i$. Logo, a

correlação entre U_k e V_k é dada por:

$$\text{Corr}(U_k, V_k) = \rho_k^* = \sqrt{\rho_k^{*2}}. \quad (2.13)$$

As variáveis canônicas têm as seguintes propriedades:

- $V(U_k) = V(V_k) = 1$;
- $\text{Cov}(U_k, U_l) = \text{Corr}(U_k, U_l) = 0, k \neq l$;
- $\text{Cov}(V_k, V_l) = \text{Corr}(V_k, V_l) = 0, k \neq l$;
- $\text{Cov}(U_k, V_l) = \text{Corr}(U_k, V_l) = 0, k \neq l$ para todo $k, l = 1, 2, \dots, p$.

Segundo Lira (2004), quando se utiliza apenas amostras da população, as matrizes de covariância Σ e de correlação ρ são substituídas pela matriz de covariância amostral S e de correlação R ou $\hat{\rho}$, respectivamente. A seguir detalha-se o procedimento para encontrar as variáveis e correlações canônicas considerando apenas amostras da população.

2.5.4 Variáveis Canônicas Amostrais e Correlações Canônicas Amostrais

Considerando uma amostra aleatória composta por n observações dos vetores \underline{X} , p -dimensional e o \underline{Y} , q -dimensional, tem-se a matriz de dados de dimensão $n \times (p + q)$ representada por:

$$Z = [\underline{X} \mid \underline{Y}] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} & \vdots & y_{11} & y_{12} & \dots & y_{1q} \\ x_{21} & x_{22} & \dots & x_{2p} & \vdots & y_{21} & y_{22} & \dots & y_{2q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} & \vdots & y_{n1} & y_{n2} & \dots & y_{nq} \end{bmatrix}, \quad (2.14)$$

e o vetor de médias com dimensão $p + q$:

$$\underline{\bar{Z}} = \begin{bmatrix} \underline{\bar{X}} \\ \dots \\ \underline{\bar{Y}} \end{bmatrix}, \quad (2.15)$$

onde

$$\underline{\bar{X}} = \frac{1}{n} \sum_{n=1}^n x_n \quad \text{e} \quad \underline{\bar{Y}} = \frac{1}{n} \sum_{n=1}^n y_n . \quad (2.16)$$

A matriz de covariância amostral pode ser representada por:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, \quad (2.17)$$

e as combinações lineares

$$\begin{aligned} \hat{U} &= \underline{\hat{a}}' \underline{X}, \\ \hat{V} &= \underline{\hat{b}}' \underline{Y}, \end{aligned} \quad (2.18)$$

têm as seguintes correlações amostrais

$$r_{\hat{U}, \hat{V}} = \frac{\underline{\hat{a}}' S_{12} \underline{\hat{b}}}{\sqrt{\underline{\hat{a}}' S_{11} \underline{\hat{a}}} \sqrt{\underline{\hat{b}}' S_{22} \underline{\hat{b}}}}. \quad (2.19)$$

O primeiro par de variáveis canônicas amostrais é formado pelas combinações lineares \hat{U}_1 e \hat{V}_1 , com variância amostral unitária, e que maximiza a correlação 2.19.

O segundo par de variáveis canônicas amostrais são as combinações lineares \hat{U}_2 e \hat{V}_2 , com variância amostral unitária, e que maximiza a correlação 2.19 dentre aquelas que não são correlacionadas com o primeiro par de variáveis canônicas.

Na k-ésima etapa, o k-ésimo par de variáveis canônicas amostrais são as combinações lineares \hat{U}_k e \hat{V}_k , com variância amostral unitária, e que maximiza a correlação 2.19 entre aquelas que não são correlacionadas com os k-1 primeiros pares

de variáveis canônicas definidos.

O resultado que mostra de forma mais detalhada a obtenção das variáveis canônicas amostrais e correlações canônicas amostrais é análogo ao **Resultado 1**, apresentado na sub-seção 2.5.3. Basta substituir a matriz de covariância Σ pela matriz de covariância amostral S .

2.5.5 Validação da Análise de Correlação Canônica

Para verificar a aplicabilidade da análise de correlação canônica ao conjunto de dados do problema, deve-se verificar se os vetores \underline{X} e \underline{Y} são independentes entre si, ou não correlacionados (MINGOTI, 2007). Caso isso aconteça, a aplicação do método será inútil, pois $\underline{a}'\underline{X}$ e $\underline{b}'\underline{Y}$ terão correlação zero para qualquer escolha de \underline{a}' e \underline{b}' . Portanto, para validar a análise de correlação canônica se faz necessário uma análise da matriz de covariância ou de correlações a fim de determinar se elas são próximas ou não da matriz nula.

Assim vamos testar as seguintes hipóteses:

$$H_0 : \Sigma_{XY} = 0_{p \times q},$$

contra

$$H_A : \Sigma_{XY} \neq 0_{p \times q},$$

ou, equivalentemente, para $\Sigma = R$ a estatística do teste é definida por:

$$-2\ln(\Lambda) = - \left(n - 1 - \frac{1}{2}(p + q + 1) \right) \ln \left(\prod_{i=1}^p (1 - \lambda_i) \right), \quad (2.20)$$

onde n é o tamanho da amostra. A estatística do teste está associada a uma distribuição Qui-quadrado χ_v^2 com pq graus de liberdade. Se o valor obtido em (2.20) for maior ou igual ao respectivo valor da distribuição qui-quadrado, com o nível de significância igual a 5%, rejeita-se a hipótese nula.

2.6 ANÁLISE DE AGRUPAMENTO

A análise de agrupamentos é uma técnica da estatística multivariada com a finalidade de agregar grupos com base nas características que eles possuem. Além disso, é possível classificar cada objeto de modo que seja semelhante a outro objeto no agrupamento em relação ao critério de seleção adotado (HAIR; ANDERSON, 2002).

Classificar diz respeito a um número de grupos conhecidos e o objetivo principal é fixar uma nova observação em um destes grupos. Agrupar é a técnica mais primitiva no sentido de que nenhuma suposição é feita quanto ao número de grupos ou sobre a técnica de agrupamento utilizada (JOHNSON; WICHERN, 2002).

O agrupamento é feito com bases nas medidas de similaridades ou dissimilaridades com o objetivo básico de descobrir agrupamentos naturais dos itens ou variáveis.

2.6.1 Medidas de Similaridades e Dissimilaridades

A similaridade entre duas variáveis é uma medida de relação mútua ou semelhança entre as variáveis a serem agrupadas (HAIR; ANDERSON, 2002). Quanto maior o valor observado menos parecidos (mais dissimilares) serão as variáveis analisadas. O Coeficiente de correlação é considerado como medida de similaridade. Por outro lado, a distância euclidiana é um exemplo de medida de dissimilaridades (BUSSAB, 1990).

- **Distância Euclidiana:** Também conhecida como distância geométrica no espaço multidimensional. Dados duas observações multivariadas de dimensão p

$$\underline{X} = [x_1, x_2, \dots, x_p] \text{ e } \underline{Y} = [y_1, y_2, \dots, y_p] , \quad (2.21)$$

a distância Euclidiana é da forma:

$$d = (\underline{X}, \underline{Y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (2.22)$$

- **Quadrado da Distância Euclideana:** Neste caso, tem-se a vantagem de que não é necessário extrair a raiz quadrada, o que facilita os cálculos. Sendo dada por:

$$d = (\underline{X}, \underline{Y}) = \sum_{i=1}^p (x_i - y_i)^2. \quad (2.23)$$

- **Distância de City-Block:** Também chamada de função distância absoluta ou distância de Manhathan, dada por:

$$d = (\underline{X}, \underline{Y}) = \sum_{i=1}^p |x_i - y_i|. \quad (2.24)$$

- **Distância de Mahalanobis:** Também conhecida como distância estatística, dada por:

$$d = (\underline{X}, \underline{Y}) = \sqrt{(\underline{x} - \underline{y})' \cdot S^{-1} \cdot (\underline{x} - \underline{y})}. \quad (2.25)$$

onde S é a matriz de covariância amostral.

- **Métrica de Minkowski:**

$$d = (\underline{X}, \underline{Y}) = \sqrt[n]{\sum_{i=1}^p |x_i - y_i|^n} \quad (2.26)$$

Observa-se que, quando $n = 1$ tem-se a distância City-Block.

2.6.2 Métodos de Agrupamento

Após calculada a matriz de distâncias ou similaridades entre os pares de itens ou variáveis, inicia-se o processo de agrupamento que é feito com base na matriz encontrada, pois, deseja-se encontrar informações relativas a cada par de itens ou variáveis. Todavia, existem na literatura, métodos para formar tais grupos homogêneos.

Johnson e Wichern (2002) propõem dois procedimentos dos mais utilizados na análise de agrupamentos, hierárquicos e não-hierárquicos.

2.6.3 Agrupamento Hierárquico

No método aglomerativo hierárquico, a suposição inicial é de que se tem tantos indivíduos quantos objetos. Os indivíduos são classificados em grupos em diferentes etapas, no processo hierárquico, formando-se uma árvore de classificação. O procedimento do método aglomerativo hierárquico, está descrito abaixo:

1. Inicia-se com n grupos, cada qual contendo um par de entidades, então, calcula-se a matriz de distâncias (ou similaridades) de ordem $n \times n$ com $D = \{d_{ik}\}$ a distância ou similaridade entre i e k ;
 2. Procurar na matriz de distâncias o par de grupos mais próximos (mais similares juntando-se os mesmos);
 3. Tem-se o novo grupo formado pelos pares A e B. Então, deve-se atualizar a matriz de distâncias excluindo as linhas e colunas correspondentes ao grupo (A,B) e adicionando as novas linhas e colunas dadas pela distância entre (A,B) e os grupos remanescentes;
 4. Por fim, repete-se os passos 2 e 3 de um total de $(n - 1)$ vezes. Todos os objetos formarão um único grupo no final do algoritmo. Deve-se observar a identidade dos grupos que são agrupados e os níveis (distâncias ou similaridades) nos quais ocorrem os agrupamentos. Como foi mencionado o que caracteriza os algoritmos de produzir agrupamentos é o critério usado para definir a distância entre grupos. Desta forma, pode-se fazer os agrupamentos usando alguns tipos de ligações. Os resultados obtidos com essas ligações pode ser observado graficamente na forma de um dendrograma ou diagrama em árvore.
- **Ligação Simples:** Na ligação simples ou método do vizinho mais próximo (*single linkage method*) o agrupamento é feito juntando-se dois grupos com menor

distância ou maior similaridade. Assim, dados os conjuntos de objetos X e Y , a distância entre eles será definida como:

$$d(X, Y) = \min\{d_{ij}(x, y) | i \in X \text{ e } j \in Y\}. \quad (2.27)$$

- **Ligação Completa:** Na ligação completa ou do vizinho mais longe (*complete linkage*) o agrupamento é semelhante ao da ligação simples. A única diferença é que a distância (ou similaridade) entre dois elementos de dois grupos distintos é determinada como sendo a maior distância entre eles, isto é, que são mais distantes. Assim, a ligação completa assegura que todos os itens de um grupo estão dentro de alguma distância máxima (ou mínima similaridade) de cada outro grupo.
- **Ligação Média:** Na ligação média ou método das médias das distâncias (*average linkage*) o agrupamento é feito a partir das distâncias médias entre todos os pares de itens entre dois grupos. Assim, dada a matriz de distâncias D_{ik} e o grupo (AB) formado pela distância mínima $D_{AB} = \min\{d_{ik}\}$, a distância entre (AB) e C é dada por:

$$d_{(AB)C} = \frac{\sum_i \sum_k d_{ik}}{n_{(AB)}n_C}, \quad (2.28)$$

onde d_{ik} é a distância entre o objeto i no grupo (AB) e o objeto k no grupo C, $n_{(AB)}$ e n_C são o número de itens no grupo (AB) e (C), respectivamente.

- **Método de Ward:** Neste método hierárquico, o procedimento é feito com base na minimização da "perda de informação" na junção de dois grupos. Este método é geralmente aplicado com perda de informação, tomada como um aumento no critério das somas dos erros dos quadrados. O Método de Ward é baseado na noção de que os grupos de observações multivariadas são esperados ser aproximadamente elíptico. Este método hierárquico é precursor dos métodos de agrupamento não-hierárquicos que buscam otimizar segundo algum critério para dividir os dados em determinados grupos elípticos.

- **Método da Centróide:** Dentre os critérios utilizados para produzir agrupamentos, o método da centróide é o mais direto, pois substitui cada junção de objetos num único ponto representado pelas coordenadas de seu centro. A distância entre os grupos é definida pela distância entre os centros e, em cada etapa, deve-se procurar fundir os grupos que tenham a menor distância entre si.

2.6.4 Agrupamento Não-Hierárquico

O agrupamento não-hierárquico é uma técnica utilizada quando se deseja formar k grupos de itens ou objetos. O número de grupos k , pode ser especificado com antecedência ou determinado como parte do processo de agrupamento (JOHNSON; WICHERN, 2002). Para obter a melhor partição de ordem k , deve-se empregar algum critério de qualidade de partição. É impossível, computacionalmente, obter todas as partições possíveis de ordem k e, a partir do conhecimento dessas partições, decidir qual seria a mais adequada (MINGOTI, 2007).

Os métodos não-hierárquicos diferem dos hierárquicos em vários aspectos. Primeiramente, requerem que o usuário tenha especificado previamente o número de *clusters* k desejado, ao contrário das técnicas hierárquicas aglomerativas. Em cada estágio do agrupamento, os novos grupos podem ser formados através da divisão ou fusão de grupos já combinados em passos anteriores. Isto significa que, se em algum passo do algoritmo dois elementos tiverem sido colocados num mesmo conglomerado, não necessariamente eles "estarão juntos" na partição final. Como consequência, não é mais possível a construção de dendrogramas. Geralmente, os algoritmos computacionais utilizados nos métodos não hierárquicos são do tipo iterativo e, em comparação com os métodos hierárquicos, têm uma maior capacidade de análise de conjunto de dados de maior porte, ou seja, com um grande número de observações (MINGOTI, 2007).

A técnica de agrupamento não-hierárquico mais usada é o algoritmo das K-médias. Basicamente, cada elemento amostral é alocado àquele *cluster* cujo cen-

tróide (vetor de médias amostral) é o mais próximo do vetor de valores observados para o respectivo elemento. O método é composto por quatro passos:

1. Partição arbitrária dos itens em k centróides, chamadas de "sementes" ou protótipos, para se inicializar o processo de partição;
2. Prossegue-se com a lista de itens, atribuindo um ponto para o grupo cuja média (centróide) está mais próximo. A distância normalmente utilizada é a distância Euclidiana quer com observações padronizadas ou não-padronizadas.
3. Depois de aplicar o passo 2 para cada um dos n elementos amostrais, recalcula-se os valores dos centróides para cada novo grupo formado, e repete-se o passo 2, considerando os centróides destes novos grupos;
4. Os passos 2 e 3 devem ser repetidos até que todos os elementos amostrais estejam "bem alocados" em seus grupos, isto é, até que nenhuma realocação de elementos seja necessária.

Computacionalmente, os softwares estatísticos podem utilizar formas diferentes de implementação do método das k -médias. A escolha as sementes iniciais de agrupamento influencia no agrupamento final.

3 MATERIAL E MÉTODOS

Neste capítulo são apresentados os materiais utilizados neste trabalho, que incluem os dados utilizados e também os *softwares* utilizados para aplicação dos métodos propostos no Capítulo 2, análise de correlação canônica (ACC) e análise de agrupamentos (AA).

3.1 DADOS UTILIZADOS

Os dados utilizados neste trabalho foram gentilmente cedidos pela Polícia Rodoviária Federal do Paraná, e se referem aos acidentes ocorridos no período entre 2007, 2008 e 2009 na BR-116. Os dados estão organizados em planilhas do Excel e contêm diversas informações levantadas sobre cada acidente ocorrido. No ano de 2007, foram registrados 1.515 acidentes, no ano de 2008, foram 2.393 acidentes e, em 2009, 2.337 acidentes na rodovia BR-116. Dentro de cada planilha constam as seguintes informações:

- Descrição do local onde ocorreu cada acidente, bem como, informações relativas à pista, condição meteorológica, causa do acidente, tipo do acidente, entre outros;
- Informações sobre o condutor e as demais pessoas envolvidas no acidente;
- Os veículos envolvidos no acidente;
- Dados do Policial Rodoviário Federal que fez o atendimento e registro do acidente no local.

Todas essas informações mostram de forma detalhada como ocorreu cada acidente. Ainda, revelam o número de ilesos, feridos e mortos no local. Cada acidente

ocorrido, possui um número de registro no Boletim de Ocorrência de Trânsito (BAT) da Polícia Rodoviária Federal.

A partir da examinação desses dados, a planilha foi submetida ao Microsoft Visual Basic, no qual foram implementados programas para obter a matriz de dados relativa aos acidentes na BR-116. O primeiro procedimento foi organizar a planilha com as informações dos acidentes de modo a obter as variáveis separadas em colunas.

- **Sentido da via:**

- Crescente;
- Decrescente.

- **Tipos de Acidentes:**

- Tombamento;
- Colisão traseira;
- Saída de Pista;
- Colisão lateral;
- Queda de motocicleta/bicicleta/veículo;
- Colisão transversal;
- Capotamento;
- Colisão com Objeto fixo;
- Atropelamento de Pessoa;
- Colisão com objeto móvel;
- Incêndio;
- Atropelamento de animal;
- Danos eventuais;
- Derramamento de carga;

- **Causas de Acidentes:**

- Velocidade incompatível;
- Falta de atenção;
- Ingestão de álcool;
- Desobediência à sinalização;
- Não guardar distância de segurança;
- Defeito mecânico em veículo;
- Dormindo;
- Ultrapassagem indevida;
- Defeito na via;
- Animais na pista;
- Outras.

- **Condição da Pista:**

- Pista molhada;
- Pista seca;
- Escorregadia;
- Em obra;
- Pista com buraco;
- Outra.

- **Restrição de Visibilidade:**

- Configuração do terreno;
- Veículo estacionado;
- Poeira;
- Neblina;

- Fumaça;
- Vegetação;
- Ofuscamento;
- Placas;
- Inexistente;
- Outros.

- **Sinalização:**

- Vertical;
- Horizontal;
- Manual;
- Luminosa;
- Inexistente.

- **Fases do dia:**

- Pleno dia;
- Plena noite;
- Anoitecer;
- Amanhecer.

- **Condição Meteorológica:**

- Nublado;
- Chuva;
- Céu claro;
- Sol;
- Nevoeiro/neblina;
- Granizo;

- Vento;
- Neve;
- Ignorada.

● **Tipo de Localidade:**

- Não edificada;
- Comercial;
- Industrial;
- Residencial;
- Lazer;
- Escolar.

● **Condição do Acostamento:**

- Desnível acostamento;
- Pavimentado;
- Bom;
- Com erosão;
- Ruim.

● **Obstáculo ao Cruzamento:**

- Muro;
- Cerca vegetal;
- Sarjeta;
- Meio-fio;
- Canal;
- Rio;
- Outro.

- **Conservação da Pista:**

- Bom;
- Com erosão;
- Ruim.

Algumas dessas variáveis merecem uma descrição mais detalhada. Tomando como base o Anuário Estatístico das Rodovias Federais (Departamento Nacional de Infraestrutura e Transportes, 2009) que classifica os tipos de acidentes como o primeiro evento relacionado com o veículo precursor da ocorrência como base para sua interpretação.

1. **Colisão com objeto fixo** - colisão de veículo motorizado com objeto estacionário ou fixo (exceto veículo estacionado), tais como: poste, meio-fio, mureta, baranco, entre outros, presente na área da via destinada ao trânsito de veículos.
2. **Capotamento** - ocorrência que se caracteriza pelo fato do veículo girar sobre si até ficar de rodas para cima, ou mesmo de lado, ou voltar a ficar sobre as próprias rodas.
3. **Atropelamento** - colisão de veículo motorizado com pessoa a pé ou conduzindo animal ou veículo não motorizado, na área da via destinada ao trânsito de veículos.
4. **Atropelamento de animal** - colisão de veículo motorizado com animal solto na área de influência da rodovia.
5. **Choque com veículo estacionado** - colisão de veículo motorizado com outro veículo motorizado estacionado na área da via destinada ao trânsito de veículos (usualmente, pista de rolamento e acostamento).
6. **Colisão traseira** - colisão de veículo motorizado com outro veículo motorizado que trafegue à sua frente, quando o impacto se dá com a parte traseira do veículo.

7. **Colisão lateral** - colisão de veículo motorizado com outro veículo motorizado que trafegue no mesmo sentido, quando o impacto se dá entre as laterais dos veículos envolvidos.
8. **Colisão transversal** - colisão de veículo motorizado com outro veículo motorizado que trafegue em sentido perpendicular (usualmente em cruzamento de fluxos), quando o impacto se dá com a parte da frente de um com a lateral do outro.
9. **Tombamento** - ocorrência que se caracteriza pelo fato do veículo motorizado tombar sem ter girado sobre si, ficando usualmente, de lado.
10. **Saída de Pista** - ocorrência que se caracteriza pelo fato do veículo motorizado projetar-se para fora da área destinada ao tráfego de veículos sem que tenha colidido, tombado ou capotado dentro da área referida.
11. **Atropelamento e fuga** - colisão de veículo motorizado com pessoa a pé ou conduzindo animal ou veículo não motorizado, da qual o motorista do veículo motorizado tenha escapado furtivamente do local da ocorrência.
12. **Queda de veículo** - ocorrência em que uma das vítimas (condutor ou passageiro) tenha caído do veículo em movimento, na área da via destinada ao trânsito de veículos.

Após obter a matriz de dados dos acidentes, obteve-se, novamente utilizando o programa implementado, a matriz dicotômica (FIGURA 3.1), com 6.245 linhas (acidentes) e 211 colunas (variáveis), desmembrando as variáveis originais, onde cada elemento da matriz indica a ocorrência (1) ou não (0) de uma variável em um determinado acidente. A partir da matriz dicotômica obteve-se a matriz de frequência (FIGURA 3.2) (211km × 211 variáveis) onde cada elemento da matriz é o número de acidentes em um quilômetro com relação a uma determinada variável.

Sentido Da Via_Crescente	Sentido Da Via_Decrescente	Tipo De Acidente_Tombamento
1	0	1
0	1	0
0	1	0
1	0	0
1	0	0
1	0	0
0	1	1
0	1	0
0	1	0
1	0	0
0	1	0
0	1	0

FIGURA 3.1: MATRIZ DICOTÔMICA DOS ACIDENTES NA BR-116
 FONTE: O autor (2010)

Sentido Da Via_Crescente	Sentido Da Via_Decrescente	Tipo De Acidente_Tombamento
42	70	4
56	12	0
13	16	1
10	22	2
19	23	1
17	97	0
20	50	3
26	21	1
29	19	1
22	18	1

FIGURA 3.2: MATRIZ DE FREQUÊNCIA DOS ACIDENTES NA BR-116
 FONTE: O autor (2010)

3.2 VARIÁVEIS DA ANÁLISE DE CORRELAÇÃO CANÔNICA

O primeiro grupo de variáveis refere-se às variáveis respostas e é relacionada aos tipos de acidentes. Foram considerados apenas as variáveis em que o número de ocorrências foi elevado. Algumas variáveis referentes à este grupo foram desconsideradas por ocorrerem esporadicamente. Assim, o **primeiro grupo de variáveis respostas** foi definido como:

- CT - Colisão Traseira;
- SP - Saída de Pista;
- CL - Colisão Lateral;
- CTR - Colisão Transversal;
- COF - Colisão com objeto fixo.

O segundo grupo de variáveis denominadas de variáveis de decisão, são as variáveis que correspondem aos fatores que determinaram os tipos de acidentes. Neste grupo, também, foram considerados apenas as variáveis em que o número de ocorrência mostrou-se elevado. Assim, obteve-se o segundo grupo de variáveis de decisão:

- VI - Velocidade incompatível;
- FA - Falta de Atenção;
- DES - Desobediência à sinalização;
- NDS - Não guardar distância de segurança;
- UI - Ultrapassagem indevida.

O terceiro do grupo de variáveis decisão é formado por algumas características da pista quanto da ocorrência de cada acidente. Logo, as variáveis deste grupo são:

- PM - Pista Molhada;
- PSC - Pista Seca;
- TPR - Reta;
- TPC - Curva;
- PS - Pista simples;
- PD - Pista dupla;
- FE - Faixa com erosão;
- OCM - Obstáculo ao cruzamento - Muro;
- OCS - Obstáculo ao cruzamento - Sarjeta;
- OCM - Obstáculo ao cruzamento - Meio fio.

O quarto grupo é formado por variáveis relacionadas as fases do dia em que os acidentes aconteceram. Estas fases são:

- PDia - Pleno dia;
- PN - Plena noite;
- AN - Anoitecer;
- AM - Amanhecer.

O quinto grupo de variáveis de decisão é formado por variáveis relacionadas as condições do tempo.

- NUB - Nublado;
- CH - Chuva;
- CC - Céu claro;
- SOL - Sol;
- NEV - Nevoeiro/Neblina.

3.3 APLICAÇÃO DOS MÉTODOS PROPOSTOS - ACC E AA

Neste trabalho, para a aplicação dos métodos propostos no capítulo 2, foi utilizado o *software* STATGRAPHICS Centurion XVI. O programa dispõe de processos apropriados para recolher, organizar, classificar e apresentar conjuntos de dados e oferecer estatísticas essenciais de forma amigável, com poder de análise, e facilidade nos benefícios de uso. Dentre as diversas análises de dados disponíveis pelo programa, a análise de correlação canônica e análise de agrupamentos foram utilizadas neste trabalho, como pode-se ver na figura 3.3. Na análise de correlação de canônica foi considerado apenas o primeiro par de variáveis canônicas em todas as análises feitas e na análise de agrupamento utilizou-se a distância Euclidiana e a ligação pelo método de Ward.

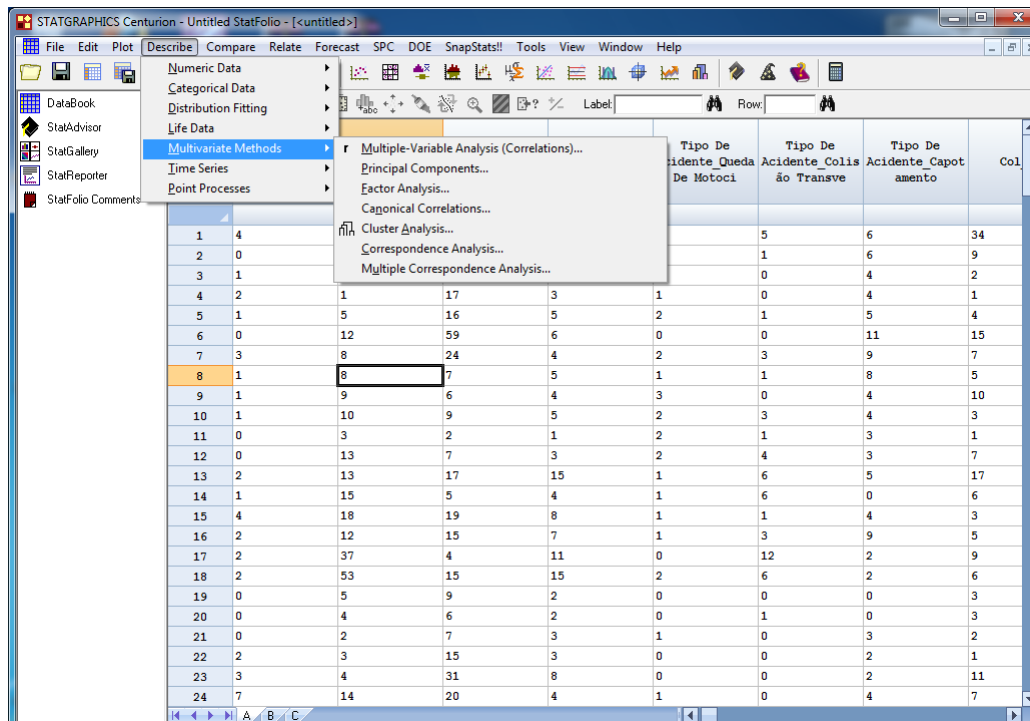


FIGURA 3.3: TELA DO PROGRAMA STATGRAPHICS CENTURION

4 RESULTADOS E DISCUSSÃO

A primeira etapa do trabalho foi encontrar os quilômetros da rodovia onde ocorrem o maior número de acidentes. No histograma (Figura 4.1), referente aos acidentes ocorridos no ano de 2007, pode-se observar uma grande concentração de acidentes nos primeiros quilômetros da rodovia (km 0 a km 20) e, também, um grande número de acidentes do quilômetro 100 ao quilômetro 125, aproximadamente. Este último trecho localizá-se entre as cidades de Curitiba e Fazenda Rio Grande.

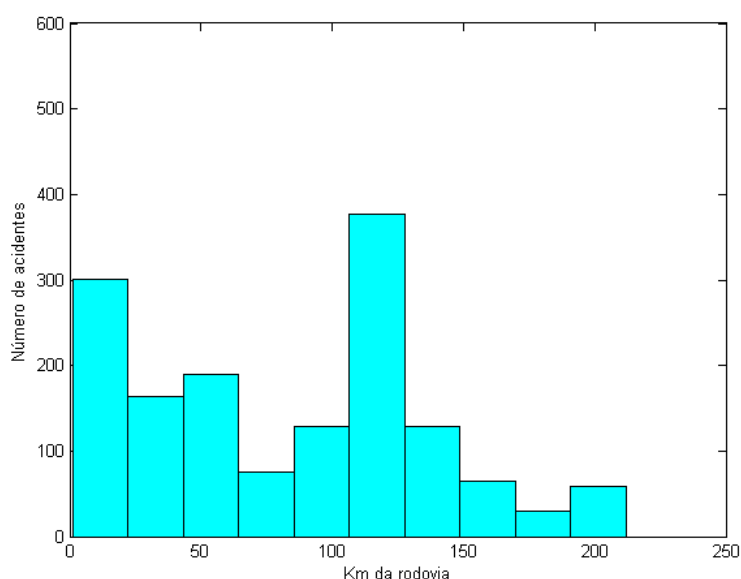


FIGURA 4.1: DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 (JAN/07 À DEZ/07)

FONTE: O autor (2010)

Tanto no ano de 2008 (Figura 4.2), como no ano de 2009 (Figura 4.3) os acidentes estão distribuídos praticamente da mesma forma. Os quilômetros onde ocorrem mais acidentes, se assemelham com os dados obtidos em 2007. Isso mostra que, nesses trechos, onde ocorrem um maior número de acidentes, o tráfego de veículos é intenso e há problemas evidentes relacionados com a rodovia, e até mesmo com os motoristas que trafegam diariamente por esses locais.

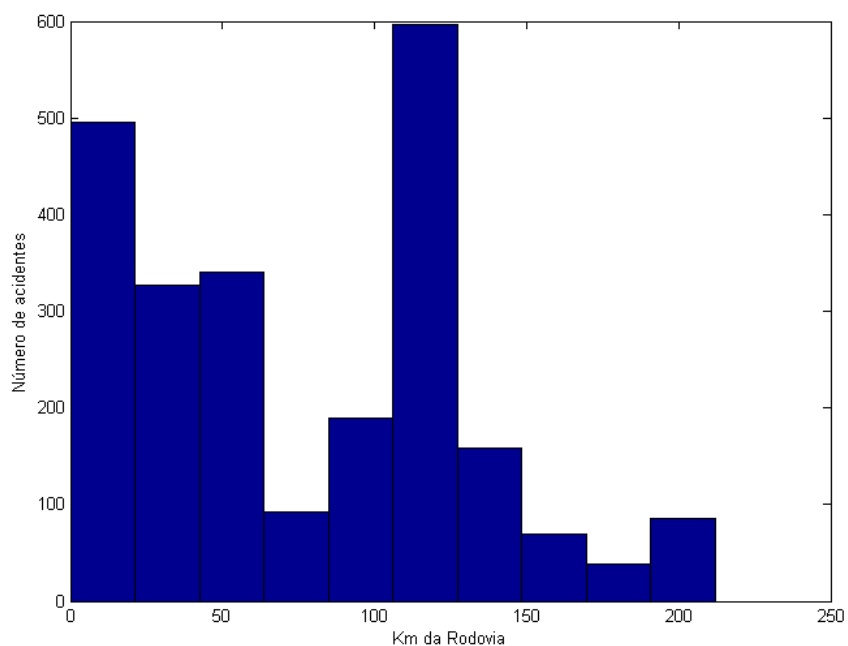


FIGURA 4.2: DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 (JAN/08 À DEZ/08)

FONTE: O autor (2010)

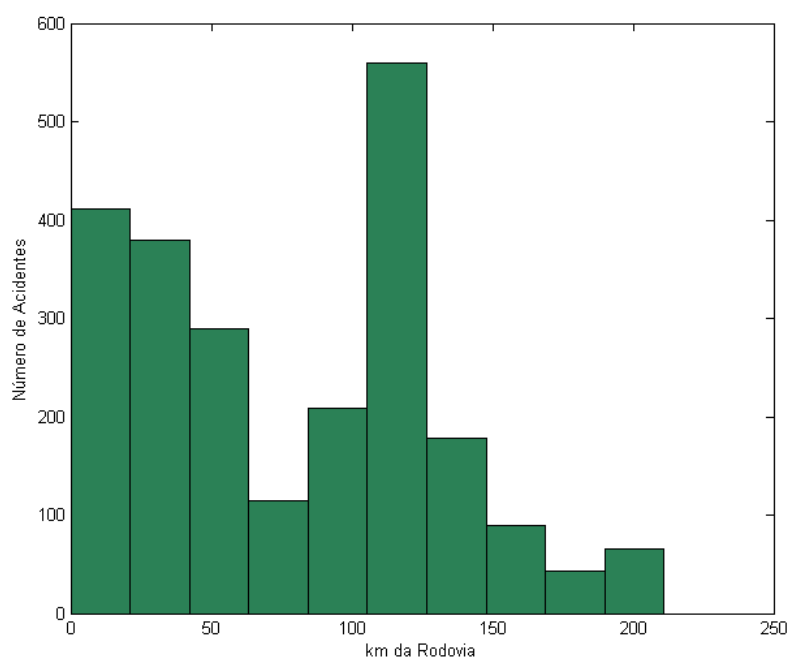


FIGURA 4.3: DISTRIBUIÇÃO DOS ACIDENTES NA BR-116 (JAN/09 À DEZ/09)

FONTE: O autor (2010)

4.1 RESULTADOS DA ACC

Nesta seção é apresentado os resultados das análises realizadas entre os grupos de acidentes utilizando o programa STATGRAPHICS.

4.1.1 Tipos de Acidentes x Causas de Acidentes

Na aplicação da análise de correlação canônica, por meio do software STAT-GRAPHICS, para verificar as relações existentes entre os tipos de acidentes e as causas dos acidentes (TABELA 4.1), obteve-se:

TABELA 4.1: RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CAUSAS DOS ACIDENTES

Número	Autovalores	Correlações Canônicas	Qui-Quadrado	Graus de Liberdade	valor-p
1	0,97001	0,98489	1165,89	25	<0,001
2	0,79085	0,88929	448,73	16	<0,001
3	0,42957	0,65542	128,75	9	<0,001
4	0,05525	0,23505	13,95	4	0,0075
5	0,01131	0,10638	2,32	1	0,1271

FONTE: Análise dos dados

Observando o valor-p, tem-se que 4 (quatro) correlações canônicas são estatisticamente importantes considerando um nível de significância igual a 5% pelo teste do qui-quadrado. Analisando ainda os autovalores, pode-se concluir que o primeiro autovalor bem como sua correlação canônica (raiz quadrada do primeiro autovalor) é mais significativa que os demais. O autovalor correspondente a primeira variável canônica é 0,97001 cuja raiz quadrada é 0,98489. Na tabela 4.2 são apresentados os coeficientes canônicos (autovetor associado).

TABELA 4.2: Coeficientes canônicos do primeiro par de variáveis canônicas

Coeficientes de U_1		Coeficientes de V_1	
VI	0,01328	CT	0,48192
FA	0,64699	SP	0,05971
DES	0,20968	CL	0,31446
NDS	0,21820	CTR	0,11460
UI	-0,01235	COF	0,01061

FONTE: Análise dos dados

Assim, a primeira variável canônica do primeiro conjunto de variáveis é representada por:

$$U_1 = 0,0132824 \cdot VI + 0,646996 \cdot FA + 0,209684 \cdot DES + 0,218201 \cdot NDS - 0,0123585 \cdot UI$$

e a primeira variável canônica do segundo conjunto de variáveis é representada por:

$$V_1 = 0,481924 \cdot CT + 0,0597115 \cdot SP + 0,314469 \cdot CL + 0,114604 \cdot CTR + 0,0106105 \cdot COF.$$

O par de variáveis canônicas U_1 e V_1 possui correlação canônica igual a 0,98489, o que indica uma forte correlação entre os grupos de variáveis analisados.

Analisando as correlações entre variáveis consideradas, para verificar quais variáveis são mais fortemente correlacionadas, com nível de significância igual a 5%, como pode-se ver na tabela 4.3, que apresenta a matriz de correlações entre as variáveis analisadas na aplicação da análise de correlação canônica entre tipos de acidentes e causas de acidentes.

TABELA 4.3: CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 2

	CT	SP	CL	CTR	COF	VI	FA	DES	NDS	UI
CT	1	0,0816	0,8309	0,7146	0,3005	0,0908	0,926	0,7248	0,882	0,3652
SP	0,0816	1	0,0781	-0,0969	0,1309	0,8851	0,1351	-0,025	0,003	-0,1082
CL	0,8309	0,0781	1	0,6999	0,3162	0,0963	0,8986	0,7132	0,82	0,2359
CTR	0,7146	-0,0969	0,6999	1	0,2094	-0,0535	0,797	0,8725	0,6372	0,3325
COF	0,3005	0,1309	0,3162	0,2094	1	0,1308	0,3442	0,1502	0,2833	0,0284
VI	0,0908	0,8851	0,0963	-0,0535	0,1308	1	0,148	0,0252	-0,0041	-0,1096
FA	0,926	0,1351	0,8986	0,797	0,3442	0,148	1	0,753	0,8212	0,3049
DES	0,7248	-0,025	0,7132	0,8725	0,1502	0,0252	0,753	1	0,6499	0,3485
NDS	0,882	0,003	0,82	0,6372	0,2833	-0,0041	0,8212	0,6499	1	0,4003
UI	0,3652	-0,1082	0,2359	0,3325	0,0284	-0,1096	0,3049	0,3485	0,4003	1

FONTE: Análise dos dados

onde as variáveis CT, SP, CL, VI, FA e NDS referem-se, respectivamente:

- CT - Colisão traseira;
- SP - Saída de Pista;
- CL - Colisão lateral;
- VI - Velocidade Incompatível;

Gráfico das Variáveis Canônicas

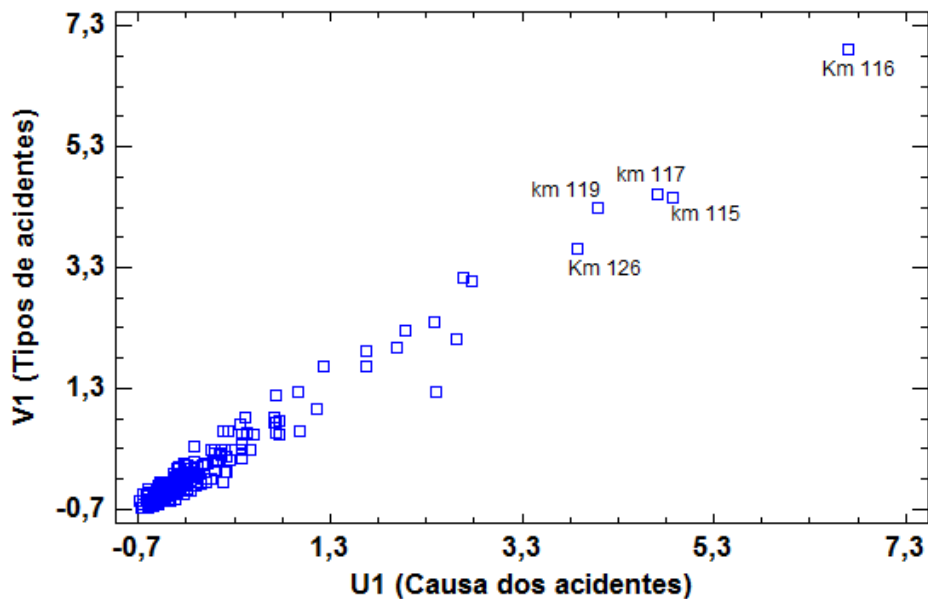


FIGURA 4.4: GRÁFICO DAS VARIÁVEIS CANÔNICAS

- FA - Falta de Atenção;
- NDS - Não guardar distância de segurança.

Analisando ainda os pares de variáveis canônicas dos dois conjuntos de variáveis, percebe-se que existem alguns pontos que se distanciam dos demais. Esses pontos referem-se, respectivamente, aos quilômetros 126, 119, 117, 115 e 116, da rodovia. Ao analisar os dados das variáveis desses quilômetros, percebe-se uma grande quantidade de acidentes do tipo: colisão traseira e lateral - e também, uma grande quantidade de acidentes causados por falta de atenção. Deste modo, pode-se classificar os quilômetros, em ordem crescente, onde ocorrem mais acidentes na rodovia. O gráfico da figura 4.4 apresenta o gráfico das relações entre os pares de variáveis canônicas dos dois grupos considerados, U_1 e V_1 . Observe que os valores de V_1 podem ser obtidos a partir dos valores de U_1 .

4.1.2 Tipos De Acidentes x Condição Meteorológica

Analisando agora os resultados obtidos pela análise de correlação canônica para verificar as relações existentes entre os tipos de acidentes com as condições meteorológicas, obteve-se:

TABELA 4.4: RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CONDIÇÃO METEOROLÓGICA

Número	Autovalores	Correlações Canônicas	Qui-Quadrado	Graus de Liberdade	valor-p
1	0,95279	0,976114	976,909	25	<0,001
2	0,80119	0,895096	352,505	16	<0,001
3	0,07261	0,269462	22,147	9	0,0084
4	0,02632	0,162258	6,732	4	0,1507
5	0,00622	0,078868	1,276	1	0,2586

FONTE: Análise dos dados

Na tabela 4.4, observa-se pelo valor-p que as 3 primeiras correlações canônicas são estatisticamente importantes considerando o nível de significância igual a 5%. Analisando ainda os autovalores, pode-se concluir que o primeiro autovalor bem como sua correlação canônica (raiz quadrada do primeiro autovalor) é mais significativa que os demais. O autovalor da primeira variável canônica é 0,952798 cuja raiz quadrada é 0,976114. Na tabela 4.5 são apresentados os coeficientes canônicos (autovetor associado).

TABELA 4.5: COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS

Coeficientes de U_1		Coeficientes de V_1	
NUB	0,26445	CT	0,44187
CH	0,36478	SP	0,45769
CC	0,47209	CL	0,24964
SOL	0,11898	CTR	0,23027
NEV	-0,01376	COF	0,05777

FONTE: Análise dos dados

Logo, pode-se definir a primeira variável canônica do primeiro conjunto de variáveis referente as condições meteorológica como sendo: $U_1 = 0,264457 \cdot \text{NUB} + 0,364788 \cdot \text{CH} + 0,472097 \cdot \text{CC} + 0,118987 \cdot \text{SOL} - 0,0137602 \cdot \text{NEV}$

e a primeira variável canônica do segundo conjunto de variáveis considerados, os tipos de acidentes, é representada por:

$$V_1 = 0,441874 \cdot CT + 0,457693 \cdot SP + 0,249644 \cdot CL + 0,23027 \cdot CTR + 0,0577702 \cdot COF$$

O primeiro par de variáveis canônicas, U_1 e V_1 , tem correlação igual a 0,976114. Isso significa que os tipos de acidentes estão fortemente correlacionados com as condições meteorológicas.

Analisando, novamente, a matriz de correlação entre as variáveis dos dois conjuntos considerados, com nível de significância igual a 5%, para verificar quais variáveis possuem forte correlação, entre tipos de acidentes e condições meteorológicas, o resultado é ilustrado na tabela 4.6.

Assim as variáveis mais fortemente correlacionadas são:

- CT - Colisão traseira;
- SP - Saída de pista;
- CL - Colisão lateral;
- NUB - Nublado;
- CH - Chuva;
- CC - Céu claro;
- SOL - Sol.

Verificando o primeiro par de variáveis canônicas dos dois conjuntos analisados, percebe-se outra vez, que existem alguns pontos que se distanciam dos demais. Esses pontos referem-se aos quilômetros 119, 117, 115 e 116, respectivamente. Nesses quilômetros, ao analisar os dados, percebe-se que a grande maioria dos acidentes ocorrem quando o tempo está nublado e, também, quando o céu está claro. Acidentes em dias de chuva ocorrem com uma frequência bem menor. Os tipos de

TABELA 4.6: CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 3

Variáveis	CT	SP	CL	CTR	COF	NUB	CH	CC	SOL	NEV
CT	1	0,0816	0,8309	0,7146	0,3005	0,8133	0,2811	0,9098	0,8756	0,2559
SP	0,0816	1	0,0781	-0,0969	0,1309	0,4008	0,8938	0,1105	0,0728	0,4558
CL	0,8309	0,0781	1	0,6999	0,3162	0,7913	0,2657	0,8746	0,817	0,3033
CTR	0,7146	-0,0969	0,6999	1	0,2094	0,6368	0,1265	0,796	0,6974	0,0935
COF	0,3005	0,1309	0,3162	0,2094	1	0,3651	0,2138	0,3335	0,3197	0,104
NUB	0,8133	0,4008	0,7913	0,6368	0,3651	1	0,5161	0,8312	0,7639	0,4133
CH	0,2811	0,8938	0,2657	0,1265	0,2138	0,5161	1	0,2807	0,2478	0,4724
CC	0,9098	0,1105	0,8746	0,796	0,3335	0,8312	0,2807	1	0,879	0,2906
SOL	0,8756	0,0728	0,817	0,6974	0,3197	0,7639	0,2478	0,879	1	0,2544
NEV	0,2559	0,4558	0,3033	0,0935	0,104	0,4133	0,4724	0,2906	0,2544	1

FONTE: Análise dos dados

Gráfico das Variáveis Canônicas

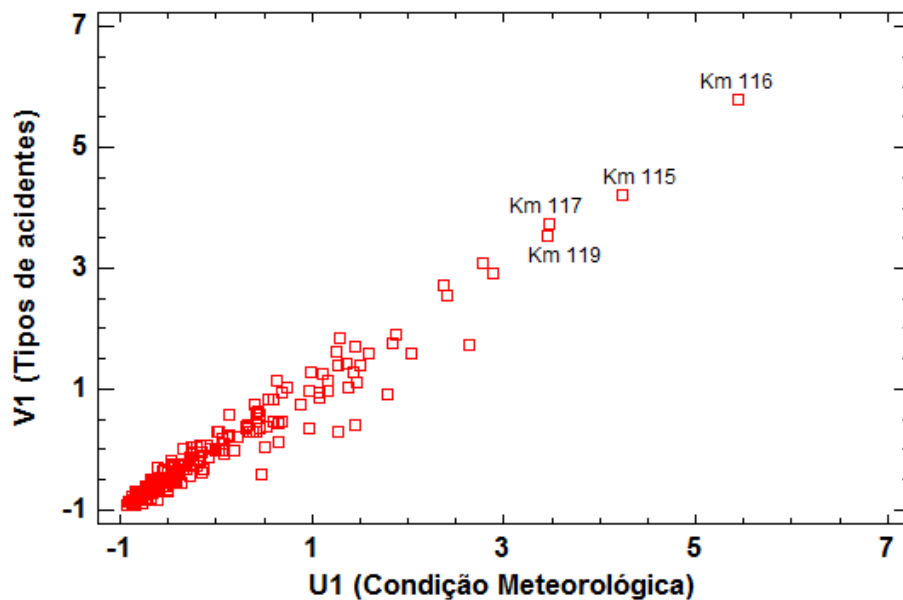


FIGURA 4.5: GRÁFICO DAS VARIÁVEIS CANÔNICAS

FONTE: Análise de dados: Programa Statgraphics

acidentes que mais ocorrem, com relação a condição meteorológica, foram: colisão traseira e lateral. Os quilômetros estão classificados em ordem crescente em relação ao risco de acidentes.

O gráfico da figura 4.5 apresenta o gráfico das variáveis canônicas entre os dois conjuntos analisados, U_1 e V_1 . Deste modo, pode-se obter os valores de V_1 a partir dos valores de U_1 .

4.1.3 Tipos de acidentes x fases do dia

Aplicou-se a análise de correlação canônica, por meio do programa STAT-GRAPHICS, para verificar o grau de associação entre os grupos de variáveis, tipos de acidentes e fases do dia, conforme mostra a tabela 4.7. Analisando o valor-p, tem-se que duas correlações canônicas são estatisticamente importantes considerando o nível de significância igual a 5%. Os autovalores obtidos indicam que somente o primeiro é mais significativo que os outros. Logo considera-se somente o primeiro par de variáveis canônicas.

Pela Tabela 4.8, que apresenta o primeiro autovetor associado ao autovalor (0,952652), define-se a primeira variável canônica do primeiro conjunto de variável, como sendo:

$$U_1 = 0,4062 \cdot CT + 0,536924 \cdot SP + 0,250317 \cdot CL + 0,21537 \cdot CTR + 0,0498286 \cdot COF$$

e a primeira variável canônica do segundo conjunto de variáveis considerados, é representada por:

$$V_1 = 0,649657 \cdot PDia + 0,258817 \cdot PN + 0,113227 \cdot AN + 0,0303759 \cdot AM.$$

TABELA 4.7: RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X FASES DO DIA

Número	Autovalores	Correlações Canônicas	Qui-Quadrado	Graus de Liberdade	valor-p
1	0,952652	0,97604	677,779	20	<0,001
2	0,178370	0,42234	52,479	12	<0,001
3	0,043492	0,20855	12,204	6	0,0576
4	0,014952	0,12228	3,0883	2	0,2135

FONTE: Análise dos dados

Assim, o primeiro par de variáveis canônicas, U_1 e V_1 , tem correlação canônica igual a 0,976039, indicando novamente uma forte correlação entre os tipos de acidentes e as fases do dia.

Calculando a correlação entre as variáveis do primeiro conjunto com as variá-

TABELA 4.8: COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS

Coeficientes de U_1		Coeficientes de V_1	
CT	0,4062	PDia	0,6496
SP	0,5369	PN	0,2588
CL	0,2503	AN	0,1132
CTR	0,2153	AM	0,0303
COF	0,0498		

FONTE: Análise dos dados

veis do segundo para verificar quais variáveis estão mais fortemente correlacionadas, com o nível de significância igual a 5%, conforme a tabela 4.9, pode-se verificar que:

- CT - Colisão traseira;
- CL - Colisão lateral;
- PDia - Pleno dia;
- PN - Plena noite;
- AN - Anoitecer.

Analisando ainda os dois conjuntos de variáveis canônicas, é possível observar alguns pontos que se distanciam dos demais. Esse pontos referem-se aos quilômetros: 126, 17, 5, 119, 117, 115 e 116. Observando os dados do km 126 da rodovia, tem-se um grande número de acidentes do tipo, colisão traseira, que ocorreram em pleno dia. O km 17 da rodovia, apresenta dados elevados para colisão traseira, e também, em pleno dia. Já o km 5, mostra um número maior de acidentes com saída de pista, com a maior parte sendo em pleno dia e plena noite. Os quilômetros 119, 117, 115 e 116, ocorrem muitas colisões traseira e lateral, respectivamente, e em pleno dia. Assim, é possível classificar os quilômetros da rodovia em crescente em relação as fases do dia e aos tipos de acidentes. O gráfico da figura 4.6 apresenta o gráfico das variáveis canônicas dos dois conjuntos, U_1 e V_1 . Assim, os valores de V_1 podem ser obtidos a partir de U_1 .

TABELA 4.9: CORRELAÇÃO ENTRE VARIÁVEIS DO GRUPO 1 E GRUPO 4

Variáveis	CT	SP	CL	CTR	COF	PDia	PN	AM	AN
CT	1	0,0816	0,8309	0,7146	0,3005	0,7826	0,7776	0,7301	0,4842
SP	0,0816	1	0,0781	-0,0969	0,1309	0,5922	0,4759	0,3564	0,4365
CL	0,8309	0,0781	1	0,6999	0,3162	0,7474	0,7549	0,7158	0,4992
CTR	0,7146	-0,0969	0,6999	1	0,2094	0,582	0,6086	0,6736	0,3953
COF	0,3005	0,1309	0,3162	0,2094	1	0,3259	0,4046	0,2999	0,2374
PDia	0,7826	0,5922	0,7474	0,582	0,3259	1	0,8794	0,8068	0,6459
PN	0,7776	0,4759	0,7549	0,6086	0,4046	0,8794	1	0,7508	0,619
AN	0,7301	0,3564	0,7158	0,6736	0,2999	0,8068	0,7508	1	0,5194
AM	0,4842	0,4365	0,4992	0,3953	0,2374	0,6459	0,619	0,5194	1

FONTE: Análise dos dados

Gráfico das Variáveis Canônicas

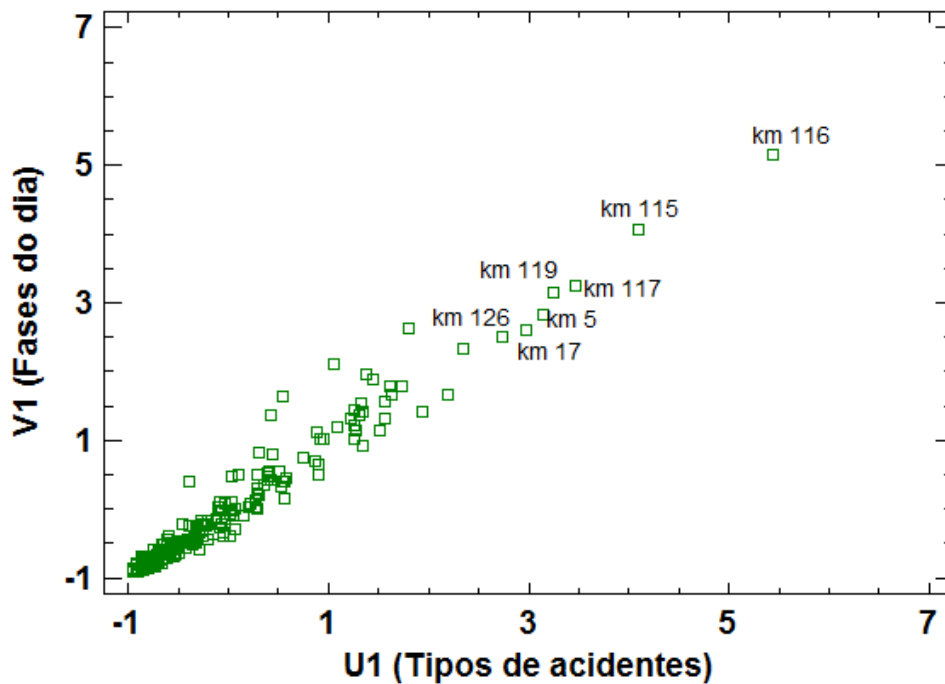


FIGURA 4.6: GRÁFICO DAS VARIÁVEIS CANÔNICAS

FONTE: Análise dos dados: Programa Statgraphics

4.1.4 Tipos de Acidentes x Características da Pista

A última análise de correlação canônica foi aplicada para verificar as relações existentes entre os tipos de acidentes e algumas características relacionadas a pista. A tabela 4.10 apresenta os resultados obtidos pelo programa STATGRAPHICS.

Observando o valor-p, nota-se que todas as correlações canônicas obtidas

TABELA 4.10: RESULTADO DA APLICAÇÃO DA ANÁLISE DE CORRELAÇÃO CANÔNICA REFERENTE AOS TIPOS DE ACIDENTES X CARACTERÍSTICAS DA PISTA

Número	Autovalores	Correlações Canônicas	Qui-Quadrado	Graus de Liberdade	valor-p
1	0,973886	0,986857	1325,860	50	<0,001
2	0,883629	0,940015	589,507	36	<0,001
3	0,409400	0,639844	155,011	24	<0,001
4	0,141156	0,375708	48,634	14	<0,001
5	0,084784	0,291178	17,896	6	0,0065

FONTE: Análise dos dados

são estatisticamente importantes considerando o nível de significância igual a 5%. Analisando os autovalores é possível concluir que o primeiro autovalor bem como sua correlação canônica (raiz quadrada do primeiro autovalor) é mais significativa que os demais. O autovalor da primeira variável canônica é 0,973886 cuja correlação canônica é 0,986857. Na tabela 4.11 são apresentados os coeficientes canônicos (autovetor associado ao primeiro autovalor) do primeiro par de variáveis canônicas.

Assim a primeira variável canônica do conjunto de variáveis relacionados as características da pista é definido por:

$$U_1 = 0,320364 \cdot PM + 0,454549 \cdot PSC + 0,00326829 \cdot OCM + 0,0301317 \cdot OCS + 0,13291 \cdot OCMF - 0,0482 \cdot FE + 0,498697 \cdot PD + 0,398577 \cdot PS - 0,163027 \cdot PR - 0,399233 \cdot PC$$

e a segunda variável canônica do conjunto de variáveis referente aos tipos de acidentes é definida por:

$$V_1 = 0,488926 \cdot CT + 0,24177 \cdot SP + 0,246269 \cdot CL + 0,293707 \cdot CTR + 0,0498594 \cdot COF$$

O primeiro par de variáveis canônicas tem correlação canônica igual a 0,986857 indicando uma forte correlação entre tipos de acidentes e as características da pista.

Fazendo as correlações entre as variáveis dos dois conjuntos, aqui analisados, para verificar quais variáveis são mais fortemente correlacionadas. Tem-se os resultados na tabela 4.12.

TABELA 4.11: COEFICIENTES CANÔNICOS DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS

Coeficientes de U_1		Coeficientes de V_1	
PM	0,3203	CT	0,4889
PSC	0,4545	SP	0,2417
OCM	0,0032	CL	0,2462
OCS	0,0301	CTR	0,2937
OCMF	0,1329	COF	0,0498
FE	-0,0482		
PD	0,4986		
PS	0,3985		
PR	-0,1630		
PC	-0,3992		

FONTE: Análise dos dados

- CT - Colisão traseira;
- SP - Saída de pista;
- CL - Colisão lateral;
- PM - Pista molhada;
- PSC - Pista seca;
- FE - Faixa com erosão;
- PD - Pista dupla;
- PS - Pista simples;
- TPR - Traçado da pista - reta;
- TPC - Traçado da pista - curva.

Analisando o primeiro par de variáveis canônicas, existem alguns pontos que se distanciam dos demais. Esses pontos referem-se aos quilômetros 17, 120, 126, 119, 115, 117 e 116, respectivamente. Observando os dados do km 17, 120 e 126, estes apresentam uma grande quantidade de colisões traseira e com pista seca. Os quilômetros 17 e 126 possuem o maior número de acidentes com pista dupla, já no

km 120 ocorrem mais acidentes em pista simples. O quilômetro 126 possui pista dupla pois localizá-se no perímetro urbano do município de Fazenda Rio Grande. Acidentes com obstáculo ao cruzamento - meio fio, ocorrem mais nos quilômetros 17 e 126. Observando agora os quilômetros 119, 117 e 115, percebe-se mais acidentes do tipo: colisão traseira, lateral e transversal, com pista seca. Os quilômetros 117 e 115 mostram acidentes onde a faixa apresenta erosão. No quilômetro 115 os acidentes ocorrem mais em pista dupla, já nos km 119 e 117, pista simples. O quilômetro 116 é o que apresenta números mais elevados para colisões traseira, lateral e transversal, pista seca e dupla, obstáculo ao cruzamento - meio fio. Todos os quilômetros analisados aqui, registraram a maior parte dos acidentes em **reta**.

Gráfico das Variáveis Canônicas

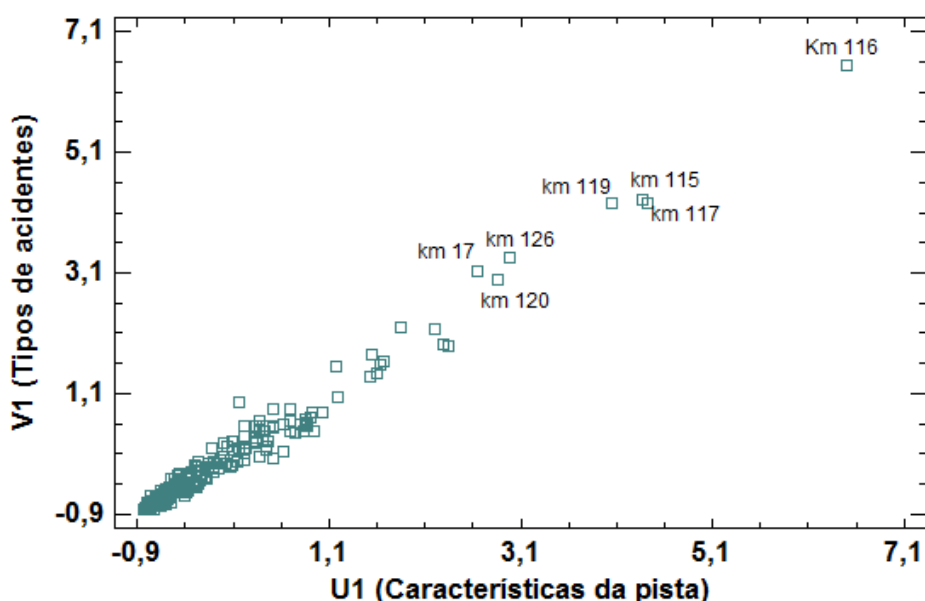


FIGURA 4.7: GRÁFICO DAS VARIÁVEIS CANÔNICAS

Assim, os quilômetros da rodovia são classificados em ordem crescente em relação as características da pista e aos tipos de acidentes. O gráfico da figura 4.7 apresenta o gráfico do primeiro par de variáveis canônicas dos dois conjuntos, U_1 e V_1 . Observe que os valores de V_1 podem ser obtidos a partir dos valores de U_1 .

As figuras 4.8 e 4.9 mostram os quilômetros 116 e 119 da rodovia BR-116, respectivamente. Estes quilômetros estão localizados no trecho entre Curitiba e Fazenda

TABELA 4.12: CORRELAÇÃO ENTRE AS VARIÁVEIS DO GRUPO 1 E GRUPO 5

Variáveis	CD	SD	CL	CTR	COF	PM	PSC	OCM	OCS	OCMF	FE	PD	PS	PR	PC
CD	1	0,0816	0,8309	0,7146	0,3005	0,3184	0,9223	0,2999	0,2254	0,6704	0,7706	0,5789	0,5095	0,9396	-0,0156
SP	0,0816	1	0,0781	-0,0969	0,1309	0,9011	0,153	0,5138	0,5566	-0,0211	0,5383	0,6785	-0,0987	0,1264	0,8793
CL	0,8309	0,0781	1	0,6999	0,3162	0,304	0,8837	0,2592	0,1864	0,7661	0,7183	0,6041	0,3887	0,8779	0,012
CTR	0,7146	-0,0969	0,6999	1	0,2094	0,1512	0,7711	0,071	-0,0627	0,6089	0,564	0,2693	0,7099	0,772	-0,1239
COF	0,3005	0,1309	0,3162	0,2094	1	0,2253	0,3682	0,2476	0,2488	0,1835	0,2758	0,3323	0,101	0,3572	0,1094
PM	0,3184	0,9011	0,304	0,1512	0,2253	1	0,361	0,5798	0,5139	0,195	0,7128	0,7936	0,0723	0,3458	0,8788
PSC	0,9223	0,153	0,8837	0,7711	0,3682	0,361	1	0,3153	0,2069	0,6657	0,8298	0,6268	0,5388	0,9473	0,0896
OCM	0,2999	0,5138	0,2592	0,071	0,2476	0,5798	0,3153	1	0,4174	0,1112	0,4812	0,6437	-0,1495	0,3242	0,4316
OCS	0,2254	0,5566	0,1864	-0,0627	0,2488	0,5139	0,2069	0,4174	1	0,0238	0,3691	0,5906	-0,2333	0,2795	0,3583
OCMF	0,6704	-0,0211	0,7661	0,6089	0,1835	0,195	0,6657	0,1112	0,0238	1	0,5468	0,5466	0,1405	0,6904	-0,0498
FE	0,7706	0,5383	0,7183	0,564	0,2758	0,7128	0,8298	0,4812	0,3691	0,5468	1	0,8089	0,3927	0,8035	0,4934
PD	0,5789	0,6785	0,6041	0,2693	0,3323	0,7936	0,6268	0,6437	0,5906	0,5466	0,8089	1	-0,1226	0,628	0,5996
PS	0,5095	-0,0987	0,3887	0,7099	0,101	0,0723	0,5388	-0,1495	-0,2333	0,1405	0,3927	-0,1226	1	0,5099	-0,0677
PR	0,9396	0,1264	0,8779	0,772	0,3572	0,3458	0,9473	0,3242	0,2795	0,6904	0,8035	0,628	0,5099	1	-0,0152
PC	-0,0156	0,8793	0,012	-0,1239	0,1094	0,8788	0,0896	0,4316	0,3583	-0,0498	0,4934	0,5996	-0,0677	-0,0152	1

FONTE: Análise dos dados

Rio Grande.



FIGURA 4.8: KM 116 DA BR-116 (PARANÁ) - VISTA AÉREA E PAISAGEM
 FONTE: O autor e Google Earth (2011)



FIGURA 4.9: KM 119 DA BR-116 (PARANÁ)- VISTA AÉREA E PAISAGEM
 FONTE: O autor e Google Earth (2011)

4.2 RESULTADOS DA ANÁLISE DE AGRUPAMENTO

Na entrada de dados do programa Statgraphics foram utilizadas todas as variáveis consideradas na aplicação da análise de correlação canônica para fazer os agrupamentos dos quilômetros da rodovia BR-116. Primeiramente, foi usado o método de ward, uma técnica de agrupamento hierárquico, juntamente com a ligação Euclidiana. Esta técnica foi aplicada para auxiliar na identificação da quantidade de grupos a serem obtidos para executar o k-médias na sequência.

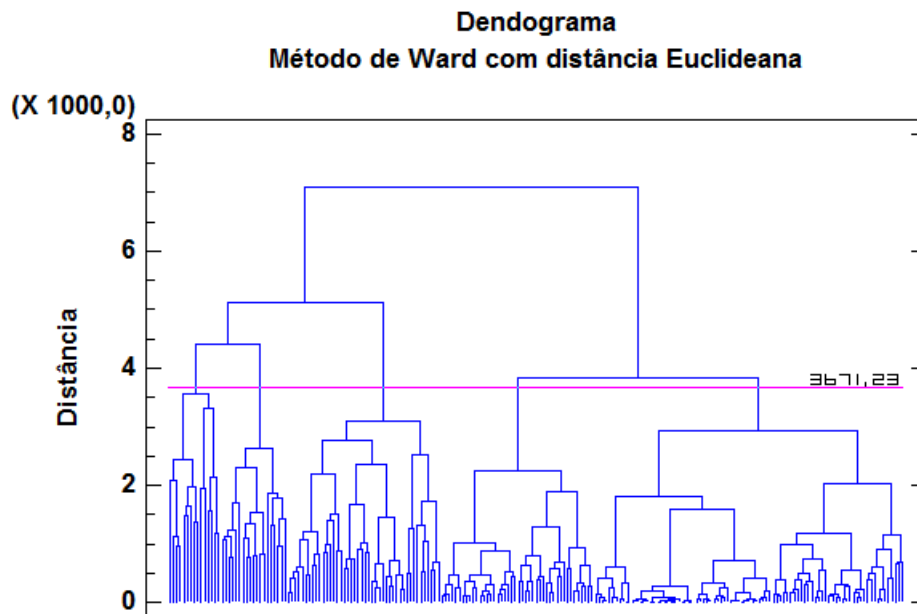


FIGURA 4.10: DENDOGRAMA DO AGRUPAMENTO USANDO O MÉTODO DE WARD
FONTE: Análise de dados: Programa Statgraphics

A partir da figura 4.10, que apresenta o dendograma dos grupos formados pelo método de ward com a distância Euclideana, pode-se observar a formação final de 2 grandes grupos. No entanto, fazendo um corte na metade do intervalo, obteve-se 5 (cincos) grupos.

Aplicando o método de agrupamento não-hierárquico, o k-médias, baseado no agrupamento feito pelo método de Ward, obteve-se 5 (cincos) grupos. Os resultados são apresentados na Tabela 4.13.

Na tabela 4.13, o primeiro grupo é formado por 11 membros, que representam os quilômetros da rodovia. Pode-se dizer que esses quilômetros são os mais perigosos da rodovia por registraram um maior número de ocorrências de acidentes. Os acidentes que mais ocorrem são **colisão traseira**, com exceção do quilômetro 122, que predomina a colisão transversal. A principal causa de acidentes nesses quilôme-

tros é a **falta de atenção**. Outras características mais presentes nesse grupo são: pista seca, pleno dia, céu claro e traçado da pista - reta.

O segundo grupo é formado por 9 membros. Todos os quilômetros pertencentes à esse grupo localizam-se no trecho antes de Curitiba. Nesses quilômetros, a grande maioria dos acidentes são **saídas de pista**, ocasionadas por **velocidade incompatível**. Além disso, pista molhada, chuva, pleno dia e traçado da pista - curva, são as características mais presentes nos acidentes registrados nesses quilômetros.

O terceiro grupo agrupa a grande maioria dos quilômetros. São 147 membros no total. No geral, nos quilômetros pertencentes a esse grupo ocorrem menos acidentes que nos demais quilômetros. Foram registrados um total de 2.451 acidentes, cerca de 39,3% do total. Por outro lado, os outros grupos com uma quantidade bem inferior de quilômetros registraram 3.794 acidentes (aproximadamente 60,7%), totalizando 6.245 acidentes registrados na rodovia BR-116.

O quarto grupo é composto por 11 membros no total. As características dos acidentes nos quilômetros deste grupo são as mesmas apresentadas no grupo 2, com exceção do km 59 em que a falta de atenção aparece como sendo a principal causa dos acidentes. Nesse grupo, a quantidade de acidentes é menor em relação ao grupo 2.

O quinto e último grupo é formado por 31 membros. Nesse grupo, evidenciou-se uma variabilidade com relação aos tipos, causas e outras características dos acidentes nos quilômetros.

TABELA 4.13: AGRUPAMENTO DOS KM USANDO O K-MÉDIAS PARA K=5

Grupos	Km	Quantidade	%
1	16, 115, 116, 117, 118, 119, 120, 121, 122, 124, 125, 126 e 127	11	6,16
2	1, 5, 28, 29, 33, 42, 43, 44, 46	9	4,27
3	8, 9, 10, 11, 13, 19, 25, 27, 31, 34, 35, 36, 37, 48, 50, 54, 55, 56, 58, 60, 62, 63, 68, 69, 70, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90 91, 93, 94, 95, 96, 97, 98, 99, 100, 101, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 123, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 154, 155, 156, 157, 159, 160, 161, 162, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210 e 211	147	69,67
4	6, 22, 23, 26, 32, 39, 40, 45, 47, 51 e 59	11	5,21
5	0, 2, 3, 4, 7, 12, 14, 15, 17, 18, 20, 21, 24, 30, 38, 41, 49, 52, 53, 57, 61, 64, 65, 66, 67, 71, 92, 102, 153, 158 e 194	31	14,69

FONTE: Análise dos dados

5 CONCLUSÕES E RECOMENDAÇÕES

Acidentes de trânsito têm sido uma preocupação constante em nossa sociedade. Cada vez mais é possível observar que a imprudência dos motoristas ligadas a outros fatores determinantes contribuem para o crescimento do número de acidentes no Brasil e no mundo.

O aumento da quantidade de veículos rodando por todas as estradas e rodovias também é um fator que contribui fortemente para o aumento dos acidentes, visto que, a infraestrutura oferecida pela rodovia não suporta uma quantidade além da que foi planejada ou muitas vezes o planejamento da rodovia foi mal elaborado.

A rodovia federal brasileira, BR-116, objeto de estudo deste trabalho, apresentou números alarmantes de acidentes considerando o trecho compreendido no estado do Paraná. Foram em média, quase 30 acidentes por quilômetro, sendo um total de 211 quilômetros.

Neste trabalho, foram utilizadas duas técnicas da análise estatística multivariada, correlação canônica e análise de agrupamentos, onde avaliou-se a correlação existente entre tipos de acidentes e outros conjuntos de atributos relacionados aos acidentes. Também foi feito o agrupamento dos quilômetros que possuem as mesmas características em relação aos acidentes, com as mesmas variáveis utilizadas na correlação canônica.

Assim, com a aplicação da análise de correlação canônica pode-se concluir que:

- Existe uma forte correlação entre os tipos de acidentes e causas de acidentes, o que implica em que a imprudência do motorista ou até mesmo a falta de experiência resultam em mais acidentes;

- As fases do dia, condição meteorológica e características da pista também apresentaram forte correlação com os tipos de acidentes. Sendo assim, pode-se dizer que acidentes também ocorrem devidos a problemas inerentes aos motoristas;
- Os piores quilômetros da rodovia BR116, no estado do Paraná, estão localizados no trecho entre Curitiba e Fazenda Rio Grande. São eles: km 115, km 116, km 117, km 119 e km 120, sendo o pior de todos o km 116, determinado pelas análises feitas. Nesse trecho, recomenda-se a duplicação da rodovia, visto que, o tráfego de veículos é intenso. Há a necessidade também, de se implantar trincheiras em lugares onde a entrada e saída de veículos é constante.

Na aplicação da análise de agrupamentos pode-se perceber que os grupos formados, usando o método das k-médias, reúne os piores quilômetros da rodovia, os quilômetros que menos ocorrem acidentes, e os quilômetros em nível intermediário de ocorrências.

Os resultados da correlação canônica e análise de agrupamentos se complementam, a primeira verificando as relações existentes nos acidentes e a segunda separando os quilômetros em grupos que se assemelham pelo número de acidentes ocorridos.

Para sugestões futuras, pode-se aplicar as técnicas usadas neste trabalho para toda a rodovia BR-116 ou aplicar em outras rodovias federais ou estaduais brasileiras. Pode-se também aplicar técnicas como a análise fatorial para resumir a grande quantidade de variáveis existentes em cada acidente em fatores.

REFERÊNCIAS

BASNIAK, M. I. **Teoria Rough Sets e Análise de Correlação Canônica: Aplicação na Avaliação de Recursos Tecnológicos do Ensino Básico**. Dissertação (Mestrado) — UFPR, 2009.

Branco, A. M. **Segurança Rodoviária**. 1999. On line; acesso em 15 junho de 2010. Disponível em: <http://www.vias-seguras.com/publicacoes/manual_de_seguranca_rodoviaria>.

BUSSAB, W. O. **Introdução à Análise de Agrupamentos**. São Paulo: [s.n.], 1990.

Chaves Neto, A. **Análise Estatística Multivariada Aplicada à Pesquisa. Notas de Aula**. Curitiba, 2002.

COSTA, G. G. O. Análise de correlação canônica entre conjuntos de princípios da qualidade total. **Cadernos de Administração**, v. 1, 2008.

Departamento Nacional de Infraestrutura e Transportes. **Anuário Estatístico das Rodovias Federais**. 2009. On line; acesso em 29 dezembro de 2010. Disponível em: <<http://www.dnit.gov.br/rodovias/operacoes-rodoviaras/estatisticas-de-acidentes/anuario-2009.pdf>>.

Departamento Nacional de Trânsito. **Impactos Econômicos e Sociais dos Acidentes de Trânsito nas Rodovias Brasileiras**. 2006. On line; acesso em 10 junho de 2010. Disponível em: <<http://www.prsp.mpf.gov.br/prdc/area-de-atuacao/direitos-humanos/arquivos/RelatorioIPEA.pdf>>.

HAIR, J. F.; ANDERSON, R. E. **Applied Multivariate Statistical Analysis**. New York: [s.n.], 2002.

JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. New York: Prentice Hall, 2002.

LIRA, S. A. **Análise de Correlação: Abordagem Teórica e de Construção dos Coeficientes com Aplicações**. Dissertação (Mestrado) — UFPR, 2004.

MARQUES, J. M. **Análise Estatística Multivariada Aplicada à Pesquisa. Notas de Aula**. Curitiba, 2010.

MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada**. Belo Horizonte: Editora UFMG, 2007.

Ministério dos Transportes. **Banco de Informações e Mapas**. 2010. On line; acesso em 29 dezembro de 2010. Disponível em: <<http://www2.transportes.gov.br/bit/trodo/br-116/gbr-116.htm>>.

Organização Mundial da Saúde. **Global Status Report on Safety Road**. 2009. On line; acesso em 29 dezembro de 2010. Disponível em: <http://whqlibdoc.who.int/publications/2009/9789241563840_eng.pdf>.

Plano Nacional de Viação. **Rede Rodoviária do PNV - Divisão em Trechos**. 2010. On line; acesso em 03 janeiro de 2011. Disponível em: <<https://gestao.dnit.gov.br/plano-nacional-de-viacao/pnv-2010-relacao-de-trechos-xls/pnv-2010.pdf>>.

RENCHER, A. C. **Methods of Multivariate Analysis**. New York: [s.n.], 2002.

RIGÃO, M. H. **Correlação Canônica para Identificação de Caracteres Importantes na Seleção Precoce de Tubérculos de Batata**. Tese (Doutorado) — UFSM, 2009.

SANTOS, L. J. R. **Generalização do Modelo Computacional de Tráfego Veicular IDM (Intelligent Driver Model)**. Dissertação (Mestrado) — UFRPE, 2008.

TRUGILHO, P. F.; LIMA, J. T.; MORI, F. A. Correlação canônica das características químicas e físicas da madeira de clones de eucalyptus grandis e eucalyptus saligna. **CERNE**, v. 9, p. 66–80, 2003.

Wikipédia. **Rodovia Federal BR-116**. 2010. On line; acesso em 21 outubro de 2010. Disponível em: <<http://pt.wikipedia.org/wiki/BR-116>>.

APÊNDICE A – HISTOGRAMA DOS ACIDENTES OCORRIDOS NA BR-116 (2007, 2008 E 2009)

Acidentes ocorridos na BR-116 no estado do Paraná nos anos de 2007, 2008 e 2009.

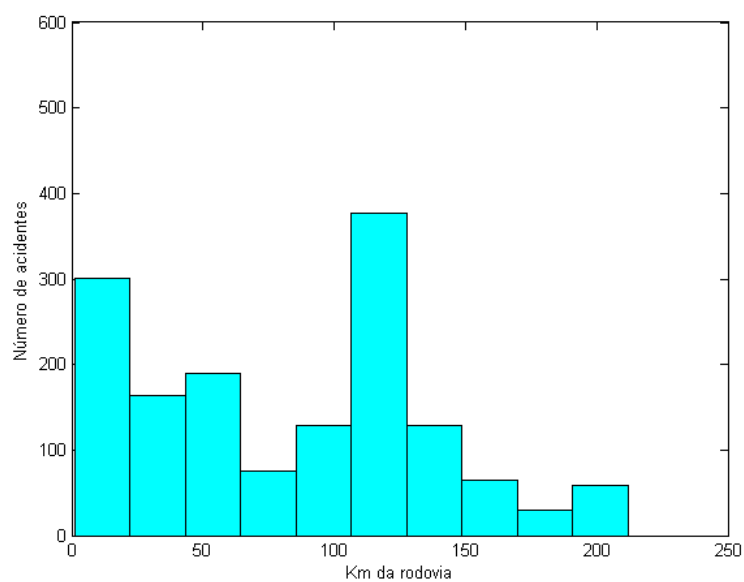


FIGURA 1.1: ACIDENTES NA BR-116 (JAN/07 À DEZ/07) COM DADOS DA PRF/PR

FONTE: O autor (2010)

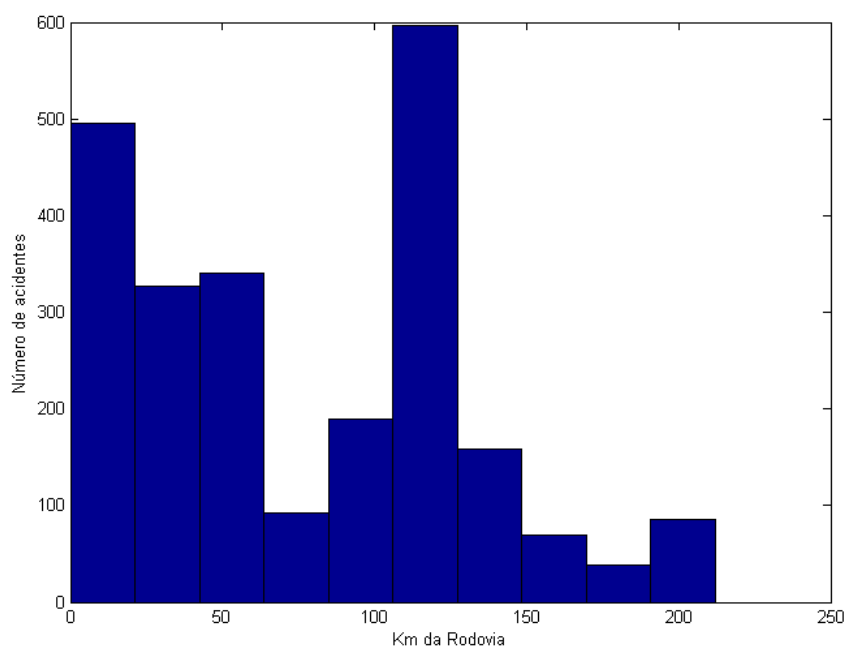


FIGURA 1.2: ACIDENTES NA BR-116 (JAN/08 À DEZ/08) COM DADOS DA PRF/PR
 FONTE: O autor (2010)

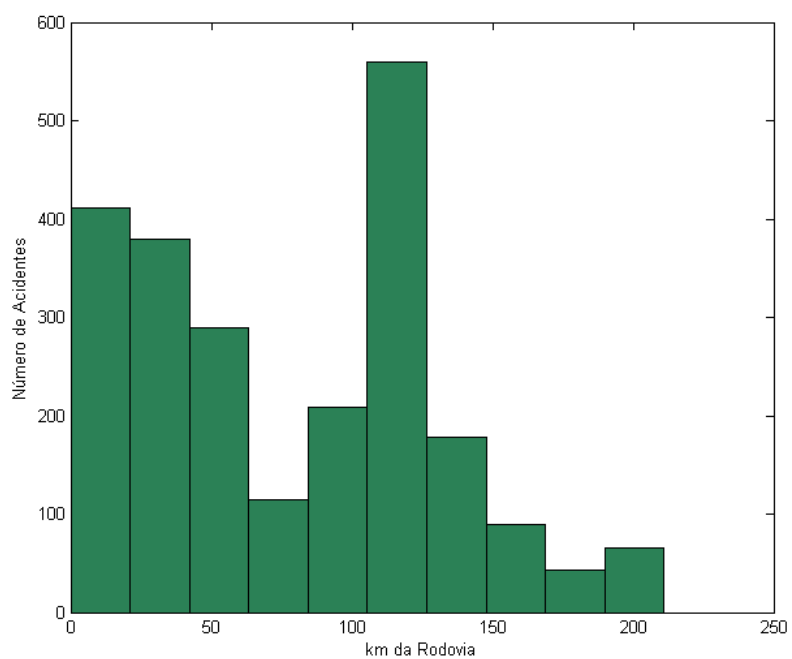


FIGURA 1.3: ACIDENTES NA BR-116 (JAN/09 À DEZ/09) COM DADOS DA PRF/PR
 FONTE: O autor (2010)

ANEXO A – PROGRAMAS UTILIZADOS PARA OBTENÇÃO DA MATRIZ DICOTÔMICA

Função programada por Fábio André Negri Balbo, aluno do PPGMNE e colega do grupo de pesquisa em Tráfego de Veículos e Pedestres utilizando a linguagem FORTRAN.

```

Sub Organizar()

'Variáveis

Dim r Dim j As Integer Dim m As Integer Dim n As Integer Dim i As
Integer Dim k As Integer Dim aux As String

'Escolha das colunas

i = 0 r = Array(15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 36, 37,
38, 39, 40, 41, 42, 44, 48, 51, 56, 57, 58, 59, 62, 63, 64, 65,
66, 68, 69, 61, 43, 67, 6, 6, 6, 12, 11)

'Seleciona planilha e conta o número de linhas do banco de dados

Sheets("Plan1").Name = "Sel_Dados"

Sheets("Acidente_Rodovia").Select While
Sheets("Acidente_Rodovia").Cells(i + 1, 1).Value <> Empty
    i = i + 1
Wend

'cria uma matriz de strings

ReDim A(i, 39) As String

'Seleciona as colunas escolhidas em r e coloca na planilha
Sel_Dados

For k = 1 To 39
    For j = 1 To i
        Sheets("Sel_Dados").Cells(j, k).Value =

```

```

        Sheets("Acidente_Rodovia")
        .Cells(j, r(k - 1))
    Next j
Next k

'Passa os valores escolhidos para a matriz A

For m = 1 To i
    For n = 1 To 39
        A(m, n) = Sheets("Sel_Dados").Cells(m, n).Value
    Next n
Next m

'Tira os espaços em branco das strings

For m = 1 To i
    For n = 1 To 31
        A(m, n) = Trim(A(m, n))
        Sheets("Sel_Dados").Cells(m, n).Value = A(m, n)
    Next n
Next m

'coloca 9999 onde não possui valores na planilha

For m = 1 To i
    For n = 1 To 31
        If Sheets("Sel_Dados").Cells(m, n).Value = Empty Then
            Sheets("Sel_Dados").Cells(m, n).Value = "9999"
        End If
    Next n
Next m

For m = 1 To i
    For n = 32 To 39
        If Sheets("Sel_Dados").Cells(m, n).Value = " " Then
            Sheets("Sel_Dados").Cells(m, n).Value = 9999
        End If
    Next n
Next m

'formata as colunas AE, AF e AG

Application.CutCopyMode = False Columns("AE:AG").Select
Selection.NumberFormat = "0.000"

'Converte em numeros as colunas 32 e 39

For k = 2 To i

```

```

Sheets("Sel_Dados").Cells(k, 32).Value = Sheets("Sel_Dados")
.Cells(k, 32).Value * 1
Sheets("Sel_Dados").Cells(k, 39).Value = Sheets("Sel_Dados")
.Cells(k, 39).Value * 1
Next k
'passa para intervalo as variaveis 33 e 34
For k = 2 To i
  If Sheets("Sel_Dados").Cells(k, 38).Value <= 30 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "0 - 30"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 60 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 30 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "30 - 60"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 90 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 60 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "60 - 90"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 120 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 90 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "90 - 120"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 150 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 120 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "120 - 150"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 180 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 150 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "150 - 180"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 210 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 180 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = "180 - 210"
  ElseIf Sheets("Sel_Dados").Cells(k, 38).Value <= 750 \\
  And Sheets("Sel_Dados").Cells(k, 38).Value > 210 Then
    Sheets("Sel_Dados").Cells(k, 38).Value = " 210 - 750"
  End If
  If Sheets("Sel_Dados").Cells(k, 33).Value = 0 Then
    Sheets("Sel_Dados").Cells(k, 33).Value = "= 0"
  ElseIf Sheets("Sel_Dados").Cells(k, 33).Value <= 3 \\
  And Sheets("Sel_Dados").Cells(k, 33).Value > 0 Then
    Sheets("Sel_Dados").Cells(k, 33).Value = "<= 3"
  ElseIf Sheets("Sel_Dados").Cells(k, 33).Value <= 20 \\
  And Sheets("Sel_Dados").Cells(k, 33).Value > 3 Then
    Sheets("Sel_Dados").Cells(k, 33).Value = ">= 3"
  ElseIf Sheets("Sel_Dados").Cells(k, 33).Value = 9999 Then
    Sheets("Sel_Dados").Cells(k, 33).Value = "9999"
  End If
  If Sheets("Sel_Dados").Cells(k, 34).Value <= 7.5 Then
    Sheets("Sel_Dados").Cells(k, 34).Value = "<= 7.5"
  ElseIf Sheets("Sel_Dados").Cells(k, 34).Value <= 15 \\
  And Sheets("Sel_Dados").Cells(k, 34).Value > 7.5 Then
    Sheets("Sel_Dados").Cells(k, 34).Value = "entre 7.5 e 15"
  ElseIf Sheets("Sel_Dados").Cells(k, 34).Value <= 21 \\

```

```

        And Sheets("Sel_Dados").Cells(k, 34).Value > 15 Then
            Sheets("Sel_Dados").Cells(k, 34).Value = "entre 15 e 21"
        ElseIf Sheets("Sel_Dados").Cells(k, 34).Value > 21 \\  

        And Sheets("Sel_Dados").Cells(k, 34).Value < 50 Then
            Sheets("Sel_Dados").Cells(k, 34).Value = "> 21"
        ElseIf Sheets("Sel_Dados").Cells(k, 34).Value = 9999 Then
            Sheets("Sel_Dados").Cells(k, 34).Value = "9999"
        End If
    Next k

'converte todas as primeiras letras das palavras em maiuscula

For m = 1 To i
    For n = 1 To 31
        Sheets("Sel_Dados").Cells(m, n).Value =  

        StrConv(Sheets("Sel_Dados").Cells(m, n).Value, vbProperCase)
    Next n
Next m

'arruma as datas

For k = 2 To i Sheets("Sel_Dados").Cells(k, 35).Value =  

Left((Sheets("Sel_Dados").Cells(k, 35).Value), 4) Next k

For k = 2 To i
    aux = Sheets("Sel_Dados").Cells(k, 36).Value
    aux = Left(aux, 7)
    aux = Right(aux, 2)
    Sheets("Sel_Dados").Cells(k, 36).Value = aux
Next k

For k = 2 To i
    aux = Sheets("Sel_Dados").Cells(k, 37).Value
    aux = Left(aux, 10)
    aux = Right(aux, 2)
    Sheets("Sel_Dados").Cells(k, 37).Value = aux
Next k

Sheets("Sel_Dados").Cells(1, 35).Value = "Ano"
Sheets("Sel_Dados").Cells(1, 36).Value = "Mês"
Sheets("Sel_Dados").Cells(1, 37).Value = "Dia"

Sheets("Sel_Dados").Select Application.CutCopyMode = False
Columns("AI:AK").Select Selection.NumberFormat = "0"

'cria intervalos para os dias

For k = 2 To i

```

```

        If Sheets("Sel_Dados").Cells(k, 37).Value > 21 Then
        Sheets("Sel_Dados").Cells(k, 37).Value = "Dia 22 - fim do mes"
        ElseIf Sheets("Sel_Dados").Cells(k, 37).Value > 14 Then
            Sheets("Sel_Dados").Cells(k, 37).Value = "Dia 15 - 21"
        ElseIf Sheets("Sel_Dados").Cells(k, 37).Value > 8 Then
            Sheets("Sel_Dados").Cells(k, 37).Value = "Dia 8 - 14"
        ElseIf Sheets("Sel_Dados").Cells(k, 37).Value > 0 Then
            Sheets("Sel_Dados").Cells(k, 37).Value = "Dia 1 - 7"
        End If
    Next k

End Sub

Sub Separar_Rodovias() Dim i As Integer Dim j As Integer Dim k As
Integer Dim n As Integer n = 0 k = 2

Sheets("Plan2").Name = "BR-116"

'contador

While Sheets("Sel_Dados").Cells(n + 1, 1).Value <> Empty
    n = n + 1
Wend

'seleciona os nomes das variaveis

For j = 1 To 38
    Sheets("BR-116").Cells(1, j).Value =
        Sheets("Sel_Dados").Cells(1, j).Value
Next j

'selecionas os acidentes da BR-116

For i = 1 To n
    If Sheets("Sel_Dados").Cells(i, 39).Value = 116 Then
        For j = 1 To 38
            Sheets("BR-116").Cells(k, j).Value = Sheets("Sel_Dados").
                Cells(i, j).Value
        Next j
        k = k + 1
    End If
Next i

End Sub

Sub Desagrupar_Variáveis() Dim i, j, n, k, p, l, m, w, s As
Integer Dim A(20, 39) As String l = 0 n = 1 k = 0 p = 1 m = 1 w =
1 s = 1

```

```

Sheets("Plan3").Name = "Dados"

'contadores

While Sheets("BR-116").Cells(n, 1).Value <> Empty
    n = n + 1
Wend

While Sheets("BR-116").Cells(1, p).Value <> Empty
    p = p + 1
Wend

'cria vetores dentro da matriz A que contem os diferentes valores
de cada variável

For q = 1 To p - 1
    For i = 2 To n - 1
        For j = 1 To 20
            If Sheets("BR-116").Cells(i, q).Value <> A(j, q)\\
                And Sheets("BR-116").Cells(i, q).Value <> "9999" Then
                    k = k + 1
                End If
            Next j
            If k = 20 Then
                l = l + 1
                A(l, q) = Sheets("BR-116").Cells(i, q).Value
            End If
            k = 0
        Next i
        l = 0
    Next q

'concatena as strings para formar os novos nomes das variáveis

For q = 1 To p - 1
    For i = 1 To 20
        If A(i, q) <> Empty Then
            Sheets("Dados").Cells(1, w).Value = Sheets("BR-116").
            Cells(1, q).Value + "_" + A(i, q)
                w = w + 1
            End If
        Next i
    Next q

' coloca 0 ou 1, dependendo se a comparação for falsa ou
verdadeira, respectivamente

```

```

For q = 1 To p - 1
  While A(m, q) <> Empty
    m = m + 1
  Wend
  For i = 1 To m - 1
    For j = 2 To n - 1
      If A(i, q) = Sheets("BR-116").Cells(j, q).Value Then
        Sheets("Dados").Cells(j, s).Value = 1
      Else
        Sheets("Dados").Cells(j, s).Value = 0
      End If
    Next j
    s = s + 1
  Next i
  m = 1
Next q
End Sub

```

PROGRAMA UTILIZADO PARA OBTER A MATRIZ DE FREQUÊNCIA DOS DADOS

```

function C= agrupar_km(A) [m,n] = size(A); s=1; l=1; C =
sortrows(A,n); vetor = C(:,n); j=1; while j<=m
  B(s,:)=zeros(1,n);
  [r,c]= find(vetor == vetor(j,1));
  k=r(length(r));
  for i= 1 : k
    B(s,:)= B(s,:)+C(i,:);
  end
  B(s,n)=B(s,n)/length(r);
  s=s+1;
  j =k+1;
  l=k+1;
end C= B(:,1:n-1);

```

ANEXO B – DETALHAMENTO DO TRECHO DA BR-116 NO ESTADO DO PARANÁ

TABELA 2.1: TRECHO DETALHADO DA BR-116 NO ESTADO DO PARANÁ

	Início	Fim	Extensão	Condição
DIV SP/PR (CAB NORTE PONTE S/ RIO PARDINHO) - INÍCIO VARIANTE DO ALPINO (PISTA DIREITA)	0,0	22,7	22,7	DUP
INÍCIO VARIANTE DO ALPINO (PISTA DIREITA) - FIM VARIANTE DO ALPINO (PISTA DIREITA)	22,7	34,4	11,7	PAV
FIM VARIANTE DO ALPINO (PISTA DIREITA) - REPRESA DO CAPIVARI	34,4	42,3	7,9	DUP
REPRESA DO CAPIVARI - ENTR PR-410 (GRACIOSA)	42,3	59,3	17,0	DUP
ENTR PR-410 - INÍC CONT. LESTE CURITIBA (QUATRO BARRAS)	59,3	71,1	11,8	DUP
INÍC CONT LESTE CURITIBA (QUATRO BARRAS)- ENTR PR-415 (P/PIRAQUARA)	71,1	86,0	14,9	DUP
ENTR PR-415 (P/PIRAQUARA) - ENTR BR-277(A)	86,0	92,9	6,9	DUP
ENTR BR-277(A) - ENTR. BR-376(A)	92,9	102,4	9,5	DUP
ENTR BR-376(A) - ENTR BR-376(B)/476/277(B) (CURITIBA SUL/PINHEIRINHO)	102,4	115,2	12,8	DUP
ENTR BR-376(B)/476/277(B) (CURITIBA SUL/PINHEIRINHO) - FIM PISTA DUPLA	115,2	118,2	3,0	DUP
FIM PISTA DUPLA - CABECEIRA NORTE PONTE RIO IGUAÇÚ	118,2	124,9	6,7	PAV
CABECEIRA NORTE PONTE RIO IGUAÇÚ - ENTR PR-510(MANDIRITUBA)	124,9	142,3	17,4	PAV
ENTR PR-510 (MANDIRITUBA) - ENTR PR-419 (AREIA BRANCA)	142,3	153,1	10,8	PAV
ENTR PR-419 (AREIA BRANCA) - ENTR PR-511 (QUITANDINHA)	153,1	167,5	14,4	PAV
ENTR PR-511 (QUITANDINHA) - ENTR PR-427 (CAMPO DO TENENTE)	167,5	191,9	24,4	PAV
ENTR PR-427 (CAMPO DO TENENTE) - DIV PR/SC (RIO NEGRO/MAFRA)	191,9	211,9	20,0	PAV
ACESSO CONT LESTE CURITIBA (ANT KM 71,1) - ENTR PR-506 (QUATRO BARRAS (ANTIGO KM 75,7))	0,0	4,6	4,6	DUP
ENTR PR-506 (QUATRO BARRAS (ANTIGO KM 75,7)) - ACESSO CONT NORTE CURITIBA (ANT KM 81,7)	0,0	6,0	6,0	DUP
ACESSO CONT NORTE CURITIBA (ANT KM 81,7) - ENTR BR-476 (CURITIBA NORTE/ATUBÁ (ANTIGO KM 88,3))	0,0	6,6	6,6	DUP
ACESSO CONT NORTE CURITIBA/VIA METROPOL - ENTR PR-415 (P/PIRAQUARA)	0,0	9,9	9,9	PLA
ENTR PR-415 (P/PIRAQUARA) - ENTR BR-277	0,0	7,4	7,4	PLA
ENTR ACES CONT NORTE CURITIBA/VIA METROP - ENTR BR-476	0,0	6,4	6,4	PLA
ENTR BR-476 - ENTR PR-417 (P/COLOMBO)	0,0	4,8	4,8	PLA
INÍCIO VARIANTE DO ALPINO (P ESQUERDA) - FIM VARIANTE DO ALPINO (PISTA ESQUERDA)	0,0	11,7	11,7	PAV

FONTE: Ministério dos Transportes (2010)