

UNIVERSIDADE FEDERAL DO PARANÁ

OSIAS RAFAEL DO AMARAL NETO CESAR

APLICATIVO PARA AUXILIAR O CONSUMIDOR NA ESCOLHA DE
UM AUTOMÓVEL ZERO QUILOMETRO

CURITIBA

2017

OSIAS RAFAEL DO AMARAL NETO CESAR

APLICATIVO PARA AUXILIAR O CONSUMIDOR NA ESCOLHA DE
UM AUTOMÓVEL ZERO QUILOMETRO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal do Paraná, como requisito parcial para a obtenção do título de Mestre em Engenharia.

Orientadora: Prof^a. Dr^a. Sonia Isoldi Marty Gama Müller

CURITIBA

2017

C421a

Cesar, Osias Rafael do Amaral Neto

Aplicativo para auxiliar o consumidor na escolha de um automóvel zero quilômetro / Osias Rafael do Amaral Neto Cesar. – Curitiba, 2017.

129 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia de Produção, 2017.

Orientador: Sonia Isoldi Marty Gama Müller.

1. Aplicativo Web. 2. Estatística multivariada. 3. Extreme programming.
I. Universidade Federal do Paraná. II. Müller, Sonia Isoldi Marty Gama. III. Título.

CDD: 658.5

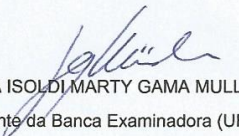


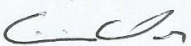
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
Setor TECNOLOGIA
Programa de Pós-Graduação ENGENHARIA DE PRODUÇÃO

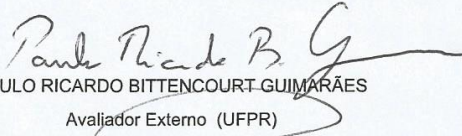
TERMO DE APROVAÇÃO

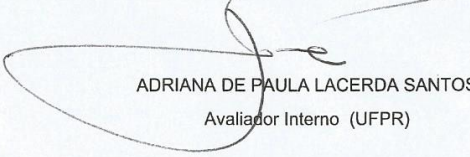
Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE PRODUÇÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **OSIAS RAFAEL DO AMARAL NETO CESAR** intitulada: **APLICATIVO PARA AUXILIAR O CONSUMIDOR NA ESCOLHA DE UM AUTOMOVEL ZERO QUILOMETRO**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua Aprovação.

Curitiba, 23 de Fevereiro de 2017.


SONIA ISOLDI MARTY GAMA MULLER
Presidente da Banca Examinadora (UFPR)


ARINEI CARLOS LINDBECK DA SILVA
Avaliador Interno (UFPR)


PAULO RICARDO BITTENCOURT GUIMARÃES
Avaliador Externo (UFPR)


ADRIANA DE PAULA LACERDA SANTOS
Avaliador Interno (UFPR)

“A leitura após certa idade distrai excessivamente o espírito humano das suas reflexões criadoras. Todo o homem que lê demais e usa o cérebro de menos adquire a preguiça de pensar. ”

Albert Einstein.

RESUMO

Com tamanha variedade de fabricantes e de modelos de automóveis oferecidos no mercado brasileiro, muitos consumidores utilizam ferramentas online para obter auxílio quando se deseja adquirir um novo automóvel. Dentro deste contexto, este trabalho tem por objetivo apresentar o desenvolvimento de um aplicativo que serve para auxiliar o consumidor na escolha de um automóvel zero quilômetro, utilizando técnicas da estatística multivariada. Para tal, foram levantadas por meio de pesquisa documental dimensões de veículos e seus itens de conforto e segurança. Em seguida, realizou-se uma Análise Fatorial com a finalidade de remover as variáveis de baixa comunalidade. Com base nas similaridades das variáveis selecionadas, utilizou-se a Análise de Agrupamentos para formarem-se grupos de automóveis. Em seguida, por meio de Análise Discriminante, encontraram-se as Funções Discriminantes Lineares de Fisher, as quais serão utilizadas para designar um dos grupos de automóveis ao usuário do aplicativo. Com base nesta análise foi desenvolvida uma ferramenta onde o usuário preenche um breve formulário e com aplicação de reconhecimento de padrões lhe é indicado um grupo de automóveis que possa atender suas expectativas. No mais, este trabalho demonstra uma aplicação prática de análise multivariada em uma abordagem inédita e o desenvolvimento de software com base na metodologia Extreme Programming.

Palavras-chave: aplicativo web, estatística multivariada, extreme programming.

ABSTRACT

With such a wide range of manufacturers and car models offered in the Brazilian market, many consumers use online tools to get help when they want to purchase a new car. Within this context, this paper aims to present the development of an application which serves to assist the consumer in choosing a zero-km automobile, using multivariate statistical techniques. For such, vehicle dimensions data and its safety and comfort items included were collected through documental research. Then, a Factorial Analysis was performed with the purpose of removing the variables of low commonality. Based on the similarities of the selected variables, a Cluster Analysis was made to form groups of automobiles. Then, through Discriminant Analysis, the Fisher Linear Discriminant Functions were found, which will be used to designate one of the car groups to the user of the application. Based on this analysis, a tool was developed which the user fills a brief form and, with pattern recognition application, it is indicated a group of cars that can meet their expectations. In addition, this work demonstrates a practical application of multivariate analysis in an unprecedented approach and the development of software based on the methodology Extreme Programming.

Keywords: web application, multivariate statistics, extreme programming.

LISTA DE FIGURAS

Figura 1 - ESTRUTURA DO TRABALHO	18
Figura 2 - AGRUPAMENTO DE MARCAS DE CARROS MODELO SEDAN	20
Figura 3 - PÁGINA INICIAL SÓ CARRÃO	25
Figura 4 - SITE WEBMOTORS	26
Figura 5 – OPÇÕES DE BUSCA WEBMOTORS	27
Figura 6 - RANKING DE CARROS CARROS NA WEB	28
Figura 7 - FATORES DE SUCESSO NO DESENVOLVIMENTO DE PRODUTOS	35
Figura 8 - METODOLOGIA DE DESENVOLVIMENTO EM CASCATA	37
Figura 9 - PRÁTICAS DA EXTREME PROGRAMMING	41
Figura 10 - ILUSTRAÇÃO DE UMA ANÁLISE FATORIAL	47
Figura 11 - EXEMPLO DE CRITÉRIO DE SELEÇÃO DE GRÁFICO DE DECLIVE	51
Figura 12 - EXEMPLO DE AGRUPAMENTO HIERÁRQUICO	56
Figura13 - PROTOCOLO DE PEQUISA	66
Figura 14 - EXEMPLO DE MUDANÇA PARA VARIÁVEIS DICOTÔMICAS	69
Figura 15 - ETAPAS DA ANÁLISE MULTIVARIADA	69
Figura 16 - MÉTODOS E MÉTRICAS UTILIZADOS NA ANÁLISE DE AGRUPAMENTOS	72
Figura 17 - DENDROGRAMA DE AGRUPAMENTOS (MÉTODO DE WARD)	72
Figura 18 – PLANEJAMENTO DE DESENVOLVIMENTO DE SOFTWARE	75
Figura 19 - MONTADORAS CONSIDERADAS NA PESQUISA	79
Figura 20 - VARIÁVEIS COLETADAS	80
Figura 21 – AGRUPAMENTO VIZINHO MAIS PRÓXIMO	82
Figura 22 – DENDROGRAMA VIZINHO MAIS DISTANTE	84
Figura 23 – DENDROGRAMA MÉTODO DE WARD	85
Figura 24 – GRUPOS FORMADOS	88
Figura 25 – FLUXOGRAMA DE FUNCIONAMENTO DO APLICATIVO	96
Figura 26 - TELA INICIAL DO APLICATIVO EXIBIDA EM UMA TELA DE COMPUTADOR	98

Figura 27 - TELA INICIAL DO APLICATIVO EXIBIDA EM UMA TELA DE SMARTPHONE	99
Figura 28 – FERRAMENTA DE DEPURAÇÃO DO NAVEGADOR	100
Figura 29 - TELA DE SELEÇÃO DE ATRIBUTOS	104
Figura 30 - TELA DE FILTRAGEM DE RESULTADOS	105
Figura 31 - TELA DE APRESENTAÇÃO DE RESULTADOS	106
Figura 32 - TELA DE PESQUISA DE OPINIÃO	107

LISTA DE QUADROS

Quadro 1 - ATRIBUTOS DE AUTOMÓVEIS QUE SÃO CONSIDERADOS IMPORTANTES PARA O CONUMIDOR, COM BASE NA LITERATURA	23
Quadro 2 - PRINCIPAIS WEBSITES A RESPEITO DE CARROS NO BRASIL	24
Quadro 3 - FERRAMENTAS DOS WEBSITES	29
Quadro 4 - CATEGORIAS DE PRODUTOS DIGITAIS	31
Quadro 5 - PRODUTOS DIGITAIS COMPARADOS COM PRODUTOS E SERVIÇOS.....	32
Quadro 6 - VARIÁVEIS ELIMINADAS DA ANÁLISE FATORIAL	81
Quadro 7 - GRUPO 1 UTILIZANDO O MÉTODO DE WARD.....	86
Quadro 8 - ACERTOS DE CLASSIFICAÇÃO SEGUNDO O MÉTODO DE FISHER.....	89

LISTA DE TABELAS

Tabela 1 - RESULTADOS DA ANÁLISE DISCRIMINANTE.....	90
Tabela 2 - COEFICIENTES DAS FUNÇÕES DISCRIMINANTES	91
Tabela 3 - CENTROIDES DOS GRUPOS NAS FUNÇÕES.....	92
Tabela 4 - TABELA DE CLASSIFICAÇÃO.....	93

SUMÁRIO

1	INTRODUÇÃO	13
1.1	PROBLEMA DE PESQUISA	14
1.2	OBJETIVOS	15
1.2.1	OBJETIVO GERAL.....	15
1.2.2	OBJETIVOS ESPECÍFICOS	15
1.3	JUSTIFICATIVA.....	16
1.4	ESTRUTURA DO TRABALHO	17
2	REVISÃO DE LITERATURA.....	19
2.1	PREFERÊNCIA DOS CONSUMIDORES DE AUTOMÓVEIS	19
2.2	ATRIBUTOS RELEVANTES PARA O CONSUMIDOR DE AUTOMÓVEIS	21
2.3	FERRAMENTAS PARA AUXILIAR O CONSUMIDOR NA ESCOLHA DE UM AUTOMÓVEL	24
2.4	PRODUTOS DIGITAIS	30
2.5	DESENVOLVIMENTO DE PRODUTOS.....	33
2.6	DESENVOLVIMENTO DE SOFTWARE.....	36
2.6.1	METODOLOGIA TRADICIONAL DE DESENVOLVIMENTO DE SOFTWARE	37
2.6.2	METODOLOGIA ÁGIL DE DESENVOLVIMENTO DE SOFTWARE ...	39
2.7	ANÁLISE MULTIVARIADA	46
2.8	ANÁLISE FATORIAL	47
2.8.1	MODELO FATORIAL ORTOGONAL.....	49
2.8.2	DETERMINAÇÃO DE FATORES	50
2.8.3	MÉTODO DAS COMPONENTES PRINCIPAIS	52
2.9	ANÁLISE DE AGRUPAMENTOS	54

2.9.1	ALGORITMOS DE AGRUPAMENTO HIERÁRQUICO.....	55
2.9.2	MÉTODO DE AGRUPAMENTO NÃO-HIERÁRQUICO.....	57
2.9.3	MÉTODOS DE LIGAÇÕES	57
2.9.4	LIGAÇÕES SIMPLES.....	58
2.9.5	LIGAÇÕES COMPLETAS	59
2.9.6	MÉTODO DAS MÉDIAS DAS DISTÂNCIAS	59
2.9.7	MÉTODO DE WARD	59
2.10	ANÁLISE DISCRIMINANTE	60
2.10.1	MÉTODO DISCRIMINANTE DE FISHER.....	62
3	MATERIAL E MÉTODOS.....	65
3.1	CLASSIFICAÇÃO DA PESQUISA.....	65
3.2	PROTOCOLO DE PESQUISA	66
3.3	DELIMITAÇÃO DA PESQUISA	67
3.4	PROCEDIMENTOS TÉCNICOS.....	68
3.4.1	COLETA DE DADOS.....	68
3.4.2	ANÁLISE DE DADOS.....	68
3.4.3	EXCLUSÃO DE VARIÁVEIS POR MEIO DE ANÁLISE FATORIAL....	70
3.4.4	APLICAÇÃO DE ANÁLISE DE AGRUPAMENTOS.....	70
3.4.5	APLICAÇÃO DE ANÁLISE DISCRIMINANTE	73
3.4.6	MÉTODO DE DESENVOLVIMENTO DA FERRAMENTA.....	74
4	RESULTADOS E DISCUSSÕES	78
4.1	BASE DE DADOS	78
4.2	ELIMINAÇÃO DE VARIÁVEIS.....	80
4.3	FORMAÇÃO DE GRUPOS DE AUTOMÓVEIS.....	82
4.3.1	CRITÉRIO VISUAL.....	82
4.3.2	CRITÉRIO DE OBJETOS CONTIDOS NOS GRUPOS.....	85

4.3.3	CRITÉRIO DE TAXA DE ACERTOS EM ANÁLISE DISCRIMINANTE	89
4.4	DISCRIMINAÇÃO DE GRUPOS DE AUTOMÓVEIS.....	90
4.5	DESENVOLVIMENTO DA APLICAÇÃO	94
4.5.1	PLANEJAMENTO.....	94
4.5.2	PROGRAMAÇÃO.....	96
4.5.3	LIBERAÇÃO DE VERSÃO E TESTES	102
4.5.4	REFATORAÇÃO DE CÓDIGO E ANÁLISE DE MELHORIAS	106
5	CONSIDERAÇÕES FINAIS	110
	REFERÊNCIAS.....	112
	APÊNDICE 1 – RESULTADOS DA ANÁLISE FATORIAL.....	118
	APÊNDICE 2 – GRUPOS DE AUTOMÓVEIS FORMADOS PELA ANÁLISE DE AGRUPAMENTOS.....	120

1 INTRODUÇÃO

Com uma representação de 5,2% do PIB brasileiro (MDIC, 2016), a indústria automotiva apresenta grande importância para a economia nacional, seja na geração de empregos diretos e indiretos, arrecadação de impostos, ou em medidas como as anunciadas pelo governo federal em janeiro de 2013, que visavam alavancar a produção de peças e direcionar uma parte das receitas geradas para as áreas de Pesquisa e Desenvolvimento.

Quanto ao setor de vendas, a rede de concessionárias em território brasileiro representa um importante papel no cenário produtivo do setor. Estes canais de distribuição, segundo Stern *et al.* (1996) requerem uma estrutura de uma rede pela qual as empresas alcançam seu público-alvo, mantém contato com clientes e outras atividades até a entrega física de seus produtos. Segundo dados do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC, 2015), o mercado automotivo brasileiro conta com 31 fabricantes de veículos divididos em 64 unidades industriais e 5.533 concessionárias.

Com tamanha variedade de fabricantes, e portanto, de modelos de automóveis oferecidos, muitos consumidores utilizam ferramentas online para obter auxílio quando se deseja adquirir um novo modelo de automóvel. Segundo dados da SimilarWeb (2016), os cinco websites mais populares do Brasil na categoria autos e veículos, totalizaram em conjunto 25,641 milhões de acessos no mês de março de 2016. No entanto, entre as principais ferramentas disponíveis, nota-se a ausência de abordagens estatísticas quanto ao agrupamento de veículos, assim como poucas ferramentas possibilitam ao consumidor a busca principalmente por atributos desejados e não por um modelo em específico.

Diante deste cenário, a proposta deste trabalho é desenvolver uma ferramenta para auxiliar o consumidor na escolha de um modelo de carro zero quilômetro. A proposta é criar um aplicativo para a web baseado nos atributos dos diversos modelos de automóveis existentes à venda no Brasil e utilizar técnicas estatísticas multivariadas para possibilitar que seja indicado um grupo de automóveis ao consumidor.

Serão testados diversos agrupamentos de modelos de automóveis com intuito de formar grupos com base em suas similaridades de dimensões e itens inclusos e, com o uso de técnicas de reconhecimento de padrões, será apresentado ao usuário um grupo de automóveis com maiores chances de atender as suas expectativas como consumidor. Como requisito para alcançar tal objetivo, será demonstrada a aplicação de uma metodologia de desenvolvimento de software, assim como conceitos de engenharia de software relacionados. Com isto, espera-se que esta ferramenta auxilie assim o processo de decisão do consumidor, apresentando opções relevantes, baseados em metodologia científica, e também, que a base de dados gerada na elaboração e utilização do aplicativo, possa ser utilizada para extrair informações a respeito do mercado de automóveis nacional.

1.1 PROBLEMA DE PESQUISA

Segundo dados dos sites Icarros (2016) e Autoesporte (2016), há mais de 40 marcas de automóveis que são comercializadas no Brasil, entre veículos fabricados em solo nacional e importados através de lojas especializadas, totalizando mais de 300 modelos à disposição do consumidor. Diante de tal variedade de modelos, escolher um novo automóvel pode ser considerada uma decisão complexa, pois trata-se de um produto com alto valor agregado, grande variedade, e diferentes produtos voltados para diferentes públicos. Existe, portanto, uma diversidade de fontes de informações e ferramentas que têm como propósito auxiliar o consumidor na decisão de compra de um novo automóvel.

Pode-se citar como exemplos de ferramentas oferecidas em sites da web para auxiliar o consumidor de automóveis, o Catálogo 0km Icarros (2016), onde pode-se escolher ou comparar modelos de automóveis com o auxílio de filtragem de resultados por preço, itens oferecidos, ano, etc. e o Comparador Autoesporte (Revista Auto Esporte, 2016), onde, de forma semelhante resultados são filtrados por preço, tipo de carroceria e marca, selecionando até três modelos para que

seja possível compará-los lado a lado. Em ambos os casos, a abordagem é que o consumidor deseja um ou mais modelos em específico e deseja informações sobre os mesmos. A ferramenta é usada então como forma de apoiar a decisão do consumidor.

No entanto, há a possibilidade de o consumidor não estar buscando um modelo de automóvel em específico, mas sim, conhecer quais veículos oferecem itens que ele deseja em um automóvel. Esta abordagem tem por base o princípio de que, se o consumidor deseja adquirir um automóvel mas não sabe qual modelo adquirir, ele pode começar identificando o que ele busca em um automóvel, e depois, analisar as opções oferecidas.

Neste contexto, o presente trabalho propõe-se a responder a seguinte pergunta: como auxiliar o consumidor na escolha de um modelo de automóvel zero quilômetro?

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

Desenvolver um aplicativo para auxiliar a compra de um carro zero quilômetro, por meio do uso de técnicas da estatística multivariada.

1.2.2 OBJETIVOS ESPECÍFICOS

- a) Identificar características e atributos que possam classificar estatisticamente modelos de automóveis zero quilômetro por meio de pesquisa documental;
- b) Agrupar modelos de automóveis com o uso da técnica de análise de agrupamentos;

- c) Realizar reconhecimento de padrão com o intuito de discriminar um grupo e automóveis ao consumidor;
- d) Desenvolver ferramenta para auxiliar na escolha de modelos de automóveis zero quilômetro;

1.3 JUSTIFICATIVA

Ferramentas disponíveis para auxiliar o consumidor de automóveis são encontradas na forma de comparadores de veículos ou de busca por um modelo desejado. Neste caso, o uso de uma ferramenta que tenha por base técnicas de estatística multivariada, apresenta uma abordagem nova para a resolução deste problema, podendo desta forma propiciar resultados inovadores.

A técnica de Análise de Agrupamentos (AA) permite que se formem grupos de automóveis sem partir de nenhum pressuposto quanto ao tamanho e conteúdo destes grupos. Tendo como base para a formação desses grupos unicamente a similaridade entre os diversos atributos que um automóvel possui, é possível aplicar técnicas de reconhecimento de padrão, para que se enquadre uma nova observação a um grupo utilizando as mesmas variáveis que se utilizaram na AA.

Partindo desse pressuposto, da mesma forma que um novo modelo de automóvel pode ser discriminado a um grupo, um consumidor que escolha os atributos que deseja em um automóvel também pode ser discriminado a um grupo. Se o consumidor escolhe estes atributos com base no que deseja, pode-se dizer que o grupo a que este foi discriminado possui grande afinidade com o mesmo e, portanto, apresenta veículos que podem atender suas expectativas. O uso de tais técnicas em uma ferramenta voltada ao consumidor apresenta uma abordagem inédita no que se refere a questão da escolha de um modelo de automóvel.

Utilizando os resultados de uma análise pode-se então desenvolver uma ferramenta adequada para auxiliar o consumidor de automóveis. No entanto, para que seja possível realizar esta análise deve-se identificar quais atributos

devem ser incluídos na pesquisa. Estes atributos devem servir tanto para diferenciar e agrupar os modelos de automóveis, como devem ser relevantes ao consumidor. Como exemplos de variáveis que podem ser utilizadas para tal abordagem, tem-se:

- Dimensões físicas do veículo, como tamanho, capacidade do tanque e do porta-malas;
- Atributos mensuráveis, como velocidade máxima, consumo de combustível e tempo de zero a cem km/h;
- Itens de segurança e de conforto oferecidos, como airbag para passageiros, bancos com regulagem de altura e teto solar.

Com a grande variedade de modelos de automóveis, surge a questão da grande variedade de atributos identificáveis em um automóvel, assim como a necessidade de selecionar os atributos que forem mais relevantes para a pesquisa.

A aplicação prática de tal ferramenta, também possibilita a coleta de dados a respeito do mercado consumidor. Se o usuário desta ferramenta escolher atributos desejáveis em um modelo de automóvel, estes dados podem ser armazenados e posteriormente convertidos em informação de mercado.

Por fim, conforme classificação disponibilizada pela Associação Brasileira de Engenharia de Produção (ABEPRO), o tema desta pesquisa possui aderência à Engenharia de Produção, especificamente, na ramificação de Processo de Desenvolvimento de Produto, da área de Engenharia de Produto.

1.4 ESTRUTURA DO TRABALHO

Com a finalidade de alcançar os objetivos delineados na estruturação desta pesquisa, este trabalho encontra-se dividido em cinco capítulos.

O primeiro capítulo aborda a introdução ao tema, problema de pesquisa, objetivos e justificativa.

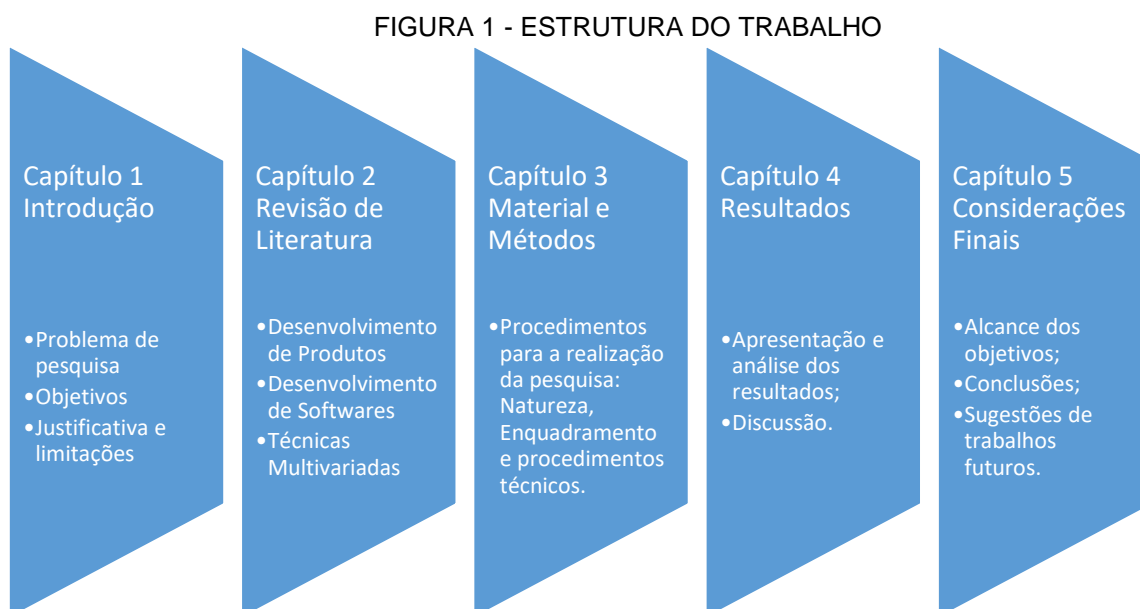
No segundo capítulo são apresentados os temas pertinentes ao desenvolvimento do trabalho com base na literatura, entre os quais abrangem uma contextualização a respeito do desenvolvimento de produtos e softwares, técnicas multivariadas empregadas para formar agrupamentos e discriminar variáveis, assim como trabalhos relevantes em torno dos temas apresentados e trabalhos correlatos a esta pesquisa onde são identificados alguns atributos importantes para este mercado consumidor.

O terceiro capítulo apresenta os aspectos metodológicos referentes ao desenvolvimento do trabalho, tais como natureza e enquadramento da pesquisa e os procedimentos adotados para a coleta de dados e desenvolvimento da ferramenta proposta, além do detalhamento sobre a aplicação das técnicas multivariadas utilizadas para alcançar os objetivos propostos.

O quarto Capítulo apresenta a análise e interpretação dos resultados alcançados assim como a demonstração da ferramenta obtida

No quinto capítulo são apresentadas as considerações finais em relação aos resultados apresentados no capítulo anterior.

A Figura 1 ilustra a estrutura deste trabalho conforme aqui apresentado.



FONTE: O autor (2016).

2 REVISÃO DE LITERATURA

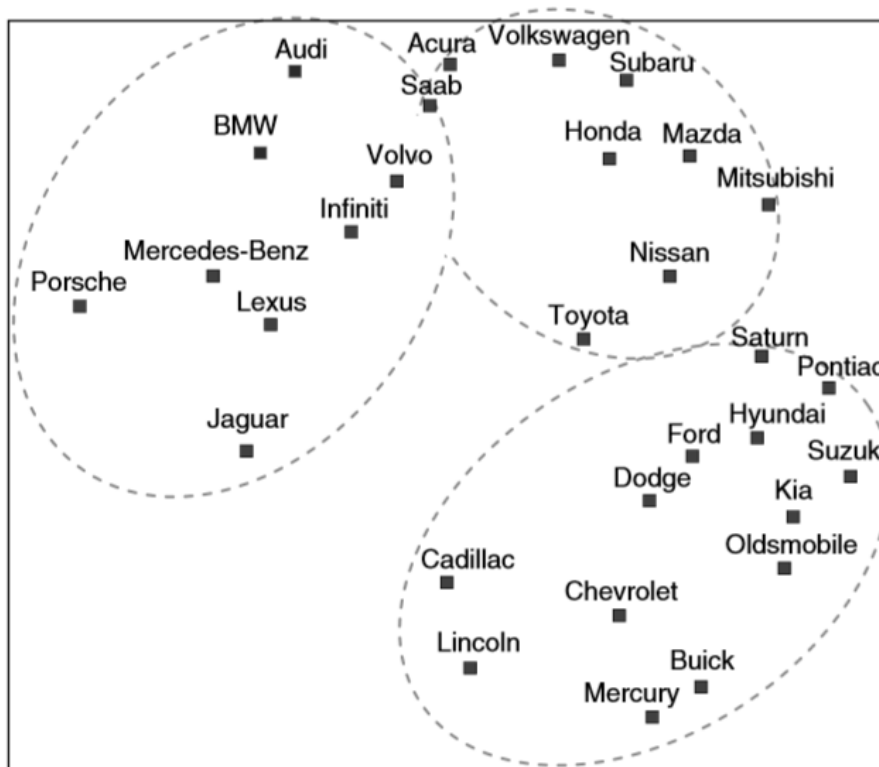
Para a revisão de literatura realizada neste trabalho, foram utilizadas como principais fontes de informação as bases de dados da CAPES/MEC, Google Acadêmico e Web Of Science™. Para que fossem encontrados trabalhos correlatos relevantes a este trabalho apresentado, utilizou-se a busca por tópicos e filtros, em conjunto com a ferramenta gratuita de mineração de texto Sobek, desenvolvida por alunos da Universidade Federal do Rio Grande do Sul. Tal ferramenta mostrou-se útil para identificar as palavras-chave e termos de título mais relevantes nos resultados das pesquisas, assim como a forma que estes termos se relacionam semanticamente, o que possibilitou a leitura e análise dos artigos pesquisados de forma mais eficiente (Sobek, 2016).

2.1 PREFERÊNCIA DOS CONSUMIDORES DE AUTOMÓVEIS

No trabalho de Netzer, Feldman et al. (2012), é discorrido a respeito de como a web 2.0, onde o usuário é quem é responsável pelo conteúdo publicado e consumido, pode ser utilizada para captar quantidades colossais de dados a respeito de pensamentos, crenças, experiências e interações de consumidores. Neste estudo, os autores propõem uma abordagem para que o conteúdo gerado pelos consumidores na web como, opiniões sobre seus produtos e produtos concorrentes, seja considerado pelas empresas. O objetivo do trabalho apresentado é, portanto, converter o conteúdo gerado por usuários em conhecimento de cenário competitivo e estrutura de mercado. Para tal, é utilizada uma abordagem de mineração de texto combinada com ferramentas de análise de rede de semânticas. Para demonstração dos resultados, é apresentado um caso de opinião a respeito de carros modelo sedan, onde são gerados mapas de percepção sem a utilização de entrevistas de consumidores, mas apenas dados de um fórum de discussão online, a respeito carros modelo sedan. Entre os dados obtidos estão: quais as marcas e modelos mais citados, os modelos

melhores avaliados, e uma proposta de agrupamento de modelos por proximidade de percepção dos usuários do fórum. Os diversos modelos e marcas de carros sedan puderam então ser agrupados por similaridade percebida conforme demonstrado na Figura 2, e os resultados foram comparados aos métodos considerados mais tradicionais de estudo de mercado, baseados em dados de vendas e questionários, validando assim o estudo.

FIGURA 2 - AGRUPAMENTO DE MARCAS DE CARROS MODELO SEDAN



FONTE: Netzer *et al.* (2012).

Faerber, Leder *et al.* (2010) e Bos, Dijksterhuis e Baaren (2011), estudam o papel do inconsciente na decisão de compra. No primeiro estudo os autores concluem que o estímulo visual influencia fortemente a percepção estética, e que a projeção de conceitos semânticos afeta a dinâmica dessa percepção. Foi possível assim, através de um experimento realizado com 144 estudantes, medir e realizar inferências estatísticas a respeito das variáveis estéticas mais significativas, sendo considerado o grau de atratividade do design como a variável-chave da apreciação estética em decisões de compra

(FAERBER, LEDER, *et al.*, 2010). O segundo estudo, através de um experimento a respeito da preferência de dois grupos de carros, testou a hipótese que o pensamento inconsciente leva a um processo de ponderação automático, onde atributos considerados importantes recebem um peso maior na decisão, enquanto atributos sem importância recebem peso menor. O experimento evidencia como tomadores de decisão imediata realizam um processo de ponderação inconsciente, decidindo assim pelo grupo de carros que tem alguns atributos considerados importantes, mas que tem muitos atributos negativos sem importância (BOS, DIJKSTERHUIS e VAN BAAREN, 2011).

Landwehr, Labroo e Herrmann (2011), visam demonstrar a importância de se levar em conta nos modelos de previsões de vendas, a qualidade na concepção de carros. Com uma base de dados de seis meses de vendas de carros na Alemanha, os autores testam então um modelo de previsão onde quantificam a complexidade do design e utilizam o valor obtido, melhorando a previsão em até 19% em relação a modelos tradicionais, onde são considerados apenas publicidade, marca, ciclo de vida, posição, preço de varejo e sofisticação tecnológica.

2.2 ATRIBUTOS RELEVANTES PARA O CONSUMIDOR DE AUTOMÓVEIS

Com o intuito de identificar na literatura os atributos de automóveis considerados relevantes para os consumidores deste mercado, foi realizada uma pesquisa em busca de trabalhos correlatos a este tema. Neste contexto, serão apresentados nessa seção alguns trabalhos que tratam de atributos relevantes para os consumidores de automóveis.

Pires e Marchetti (2000), buscam analisar os fatores que influenciam compradores de automóveis da cidade de Maringá, Paraná, na escolha de fonte de informações, identificando como atributos importantes em meio a seus entrevistados, a economia de combustível e manutenção, preço e condições de pagamento, espaço interno e conforto, design e estilo, motorização e potência e valor de mercado do carro usado.

Yamaguti (2005), também busca estudar o processo de decisão de compra de automóveis zero quilômetro pelo consumidor brasileiro. É realizada uma revisão da literatura onde são identificadas as etapas do processo de compra em um panorama geral, e posteriormente, por meio de entrevistas e aplicação de questionários, são abordados as motivações e o papel da família no processo de compra de automóveis novos. Entre os resultados, os consumidores das marcas Toyota e Honda apontaram os cinco atributos considerados mais importantes, respectivamente: preço, confiança na marca, conforto, qualidade e desempenho para os consumidores da Toyota, e preço, qualidade, confiança na marca, conforto e durabilidade para os consumidores da Honda.

Hsieh e Lindridge (2005), buscaram identificar se a importância atribuída a marca na escolha de um automóvel varia de acordo com o país. Entrevistas foram realizadas abordando características como, aspecto luxuoso, prestígio, divertido de dirigir, aceleração e velocidade, boa economia de combustível, bom valor de mercado, esportividade, robustez, durabilidade, seguro em acidentes, últimas tecnologias e estilo. Como resultados, após a aplicação de análise fatorial, a única variável retirada foi a alta qualidade. Foi identificado também que os graus de importância atribuídos às variáveis diferem de país para país.

No trabalho de Vrkljan e Anaby (2011), um estudo foi realizado com o objetivo de examinar a importância designada a atributos-chave que são levados em consideração na compra de um veículo, e qual a associação dessa importância com o gênero do consumidor. Foram entrevistados 2002 motoristas canadenses com mais de 18 anos, com o intuito de ranquear em ordem de importância, oito atributos considerados atributos-chave caso eles fossem comprar um automóvel: armazenamento, quilometragem, segurança, preço, conforto, performance, design e confiabilidade. Como resultado, segurança e confiabilidade foram os atributos melhores ranqueados, enquanto performance e design apresentaram menor importância. Ficou também evidente que não só o gênero, mas a idade dos entrevistados tem papel significativo ao explicar a importância designada a alguns atributos.

Já com o principal objetivo de identificar e classificar atributos importantes na compra de automóveis, Grohmann, Schoedler e Battistella

(2012), realizaram uma pesquisa quantitativa com 400 condutores de automóveis. Os atributos considerados mais importantes identificados nesta pesquisa foram: resistência, dirigibilidade, atendimento pós-venda e itens de conforto. Os considerados menos importantes, por sua vez, foram: tamanho do porta-malas, modelo, marca e status. Os atributos foram então agrupados em sete fatores, por meio do uso de análise fatorial. Concluiu-se que os atributos possuem forte tendência de serem agrupados em fatores que possuem a mesma funcionalidade.

Os atributos identificados são apresentados de forma resumida no Quadro 1.

QUADRO 1 - ATRIBUTOS DE AUTOMÓVEIS QUE SÃO CONSIDERADOS IMPORTANTES PARA O CONSUMIDOR, COM BASE NA LITERATURA

Atributos	Autores
- Economia, valor de manutenção, preço, condições de pagamento, espaço interno, conforto, design, motorização e potência, valor de revenda;	Pires e Marchetti (2000)
- Preço, confiança na marca, conforto, qualidade, desempenho e durabilidade.	Yamaguti (2005)
- Aspecto luxuoso, prestígio, divertido de dirigir, aceleração e velocidade, economia de combustível, valor de mercado, esportividade, robustez, durabilidade, segurança em acidentes, últimas tecnologias e estilo;	Hsieh e Lindridge (2005)
- Armazenamento, quilometragem rodada, segurança, preço, conforto, performance, design e confiabilidade;	Vrkljan e Anaby (2011)
- Resistência, dirigibilidade, atendimento pós-venda e itens de conforto.	Grohmann, Schoedler e Battistella (2012).

FONTE: Adaptado de Pires e Marchetti (2000), Yamaguti (2005), Hsieh e Lindridge (2005), Hsieh e Lindridge (2005), Grohmann, Schoedler e Battistella (2012), Vrkljan e Anaby (2011).

Os trabalhos aqui apresentados, serviram para se obter um melhor entendimento a respeito do mercado consumidor de automóveis. Nota-se a relevância da internet como ferramenta para compreender este mercado, a forma como os consumidores tem buscado fóruns de opinião entre outras ferramentas para obter informações a respeito de um produto. Em outros trabalhos contemplou-se um panorama de o que o consumidor busca em um automóvel. O próximo passo, é conhecer um pouco as ferramentas disponíveis ao consumidor buscando entender suas abordagens.

2.3 FERRAMENTAS PARA AUXILIAR O CONSUMIDOR NA ESCOLHA DE UM AUTOMÓVEL

Segundo dados da SimilarWeb (2016), os cinco websites mais populares do Brasil da categoria autos e veículos, totalizaram em conjunto 25,641 milhões de acessos no mês de março de 2016 conforme o Quadro 2.

QUADRO 2 - PRINCIPAIS WEBSITES A RESPEITO DE CARROS NO BRASIL

Site	Acessos em março de 2016 (milhões)
Webmotors	10,8
iCarros	8,3
UOL Carros	3,9
Carros na Web	1,8
Só Carrão	0,841
Total	25,641

FONTE: SimilarWeb (2016).

Importante ressaltar que dentre os websites listados no Quadro 2, apenas o UOL Carros não tem como foco principal a venda de automóveis, se tratando de um site de notícias. Os restantes possuem ferramentas que visam

auxiliar o consumidor a escolher um modelo de automóvel. Visando as ferramentas que são focadas na escolha de carros zero quilômetro, serão explanados a seguir cada uma dessas ferramentas, ressaltando seus pontos em comum, assim como peculiaridades em relação a opções disponíveis e abordagem.

Entre os sites analisados nessa pesquisa, o Só Carrão é um dos websites destinado a compra e venda de automóveis zero quilômetro ou usados. A interface da página inicial consiste em uma barra de pesquisa com abas para busca, compra, venda, serviços e revendas conforme a Figura 3.

FIGURA 3 - PÁGINA INICIAL SÓ CARRÃO



FONTE: Só Carrão (2016).

Em relação a escolha de um veículo, o site oferece uma busca simples e uma avançada. Na busca simples, o usuário seleciona a faixa do ano de fabricação do veículo, a região do Brasil desejada e entra diretamente com o nome do modelo de carro que deseja. Na busca avançada, além da opção de pesquisa citada, uma série de filtros de resultados são apresentados ao usuário, que seleciona então aquilo que deseja clicando nos botões correspondentes, como cor, número de portas, tipos de carrocerias, combustível, etc. Em ambos os tipos de busca, os resultados apresentados são veículos disponíveis para venda, com foto, descrição e contato para a compra.

De forma semelhante, o Webmotors oferece uma busca simples e uma avançada, diferenciando-se pelo fato de que o modelo e marca são selecionados por meio de botões na interface, e há um campo extra para o usuário digitar palavras-chave, para buscar uma característica específica como motorização ou tipo de combustível conforme demonstrado na Figura 4. Os resultados também são exibidos conforme há disponibilidade de compra de veículos com as características buscadas.

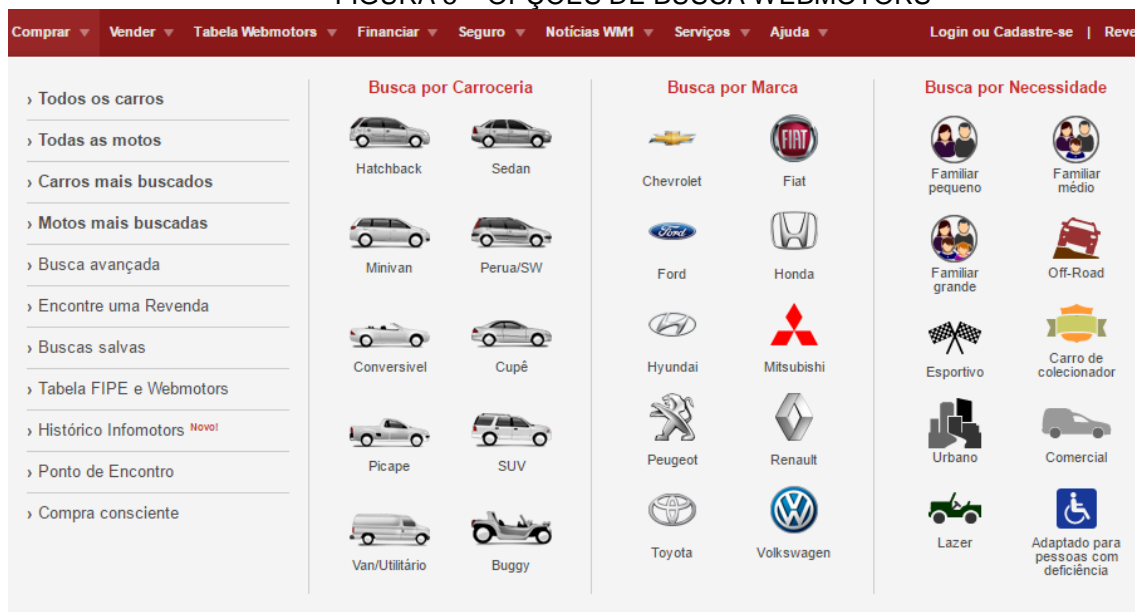
FIGURA 4 - SITE WEBMOTORS

The screenshot displays the search interface on the Webmotors website. At the top left, it says "Comprar carros" with a link to "ou motos". On the top right, there is a button labeled "Anunciar meu car" with a dollar sign icon. Below this, there are two checked checkboxes for "Novos" and "Usados". To the right of these is a link for "Busca Avançada". The main search area contains several filters: "Marca: Todas" and "Modelo: Todos" are dropdown menus; "Ano de: Min" and "Ano até: Máx" are range selectors; "Palavra-chave: 1.4, 8V, flex, turbo, automático..." is a text input field with a link "ou selecione uma versão" below it; and "Preço de: Min" and "Preço até: Máx" are range selectors. Below the filters, it indicates "Você está buscando carros em: Todo o Brasil (alterar)". A prominent red button labeled "VER OFERTAS" is positioned to the right. At the bottom, there is a section for "Catálogo de Carros 0KM" featuring a red car image, the text "Conheça nosso catálogo e descubra tudo sobre seu próximo carro.", and a button labeled "Busque seu carro".

FONTE: Webmotors (2016).

O site oferece outras opções de busca, por carroceria, por marca, por necessidade além de uma lista de mais buscados conforme a Figura 5. Há ainda um catálogo de automóveis zero quilômetro, onde é possível selecionar a marca e o modelo desejado, para que sejam exibidas algumas informações descritivas a respeito.

FIGURA 5 – OPÇÕES DE BUSCA WEBMOTORS



FONTE: Webmotors (2016).

O iCarros, possui funcionalidades e interface bem parecidas com o Webmotors, porém apresenta um catálogo de automóveis zero quilômetro mais completo, onde é possível visualizar as diferentes versões do veículo e sua ficha técnica, com informações completas a respeito dos itens e dimensões de cada veículo. Há também um comparador de veículos onde são selecionados quantos veículos se deseje e então é possível visualizar as informações do catálogo lado a lado.

O Carros na Web embora possua a interface e busca mais simples, possui uma grande diversidade de ferramentas, como catálogo com informações técnicas, galeria de fotos, notícias e uma seção onde são listados os automóveis de acordo com diversos rankings, desde os mais vendidos aos com maior número de roubos, conforme ilustra a Figura 6. O website também oferece a opção de comparar até dois modelos de automóveis, visualizando suas informações técnicas. Os resultados das buscas, no entanto, apontam para vendas de usados a serem negociados direto com o atuais proprietários, assim como um jornal de classificados. Quanto as demais ferramentas deste website, os resultados são apresentados na forma de uma listagem das informações técnicas dos veículos em questão.

FIGURA 6 - RANKING DE CARROS CARROS NA WEB

Vendas

[Carros mais vendidos](#)
[Marcas mais vendidas](#)

Concessionárias

[Melhor atendimento vendas](#)
[Melhor atendimento pós-vendas](#)

Opinião do dono

[Carros com donos mais satisfeitos](#)
[Carros com donos menos satisfeitos](#)

Roubos

[Carros mais roubados](#)

Recall

[Carros mais chamados para recall](#)

IPVA

[Alíquota por Estado \(%\)](#)

Eficiência

[Mais econômicos cidade](#)
[Mais econômicos estrada](#)
[Menos econômicos cidade](#)
[Menos econômicos estrada](#)
[Menor autonomia cidade](#)
[Maior autonomia cidade](#)
[Menor autonomia estrada](#)
[Maior autonomia estrada](#)
[Melhor coeficiente aerodinâmico \(Cx\)](#)

Pior relação peso/torque

[Carros com maior cilindrada](#)
[Carros com menor cilindrada](#)

Dimensões

[Carros mais compridos](#)
[Carros mais curtos](#)
[Carros mais largos](#)
[Carros mais estreitos](#)
[Carros mais altos](#)
[Carros mais baixos](#)
[Maior entre-eixos](#)
[Menor entre-eixos](#)
[Carros mais pesados](#)
[Carros mais leves](#)
[Maior diâmetro de giro](#)
[Menor diâmetro de giro](#)

Capacidades

[Maior porta-malas](#)
[Menor porta-malas](#)
[Maior tanque de combustível](#)
[Menor tanque de combustível](#)
[Maior carga útil](#)
[Menor carga útil](#)
[Maior tração de reboque sem freio](#)
[Maior tração de reboque com freio](#)

Fora-de-estrada

[Maior ângulo de entrada](#)
[Maior ângulo de saída](#)

FONTE: Carros na Web (2016).

O Quadro 3 apresenta de forma sucinta os principais pontos levantados a respeito dos websites pesquisados.

QUADRO 3 - FERRAMENTAS DOS WEBSITES

Site	Ferramentas	Resultados
Só Carrão	Busca por modelo.	Carros a venda na região escolhida.
Webmotors	Busca por modelo, categoria ou necessidade; Catálogo com informações descritivas.	Carros a venda na região escolhida.
iCarros	Busca por modelo categoria ou estilo; Catálogo com informações técnicas; Comparador de modelos.	Carros a venda na região escolhida; Informações técnicas.
Carros na Web	Busca por modelo; Catálogo com informações técnicas; notícias e fotos; Rankings diversos; Comparador dois modelos.	Carros anunciados em classificados; Informações técnicas.

FONTE: O autor (2016).

Entre os principais pontos levantados em relação as ferramentas estudadas, pode-se atentar a questão de que as ferramentas de busca partem do princípio que o usuário busca um modelo ou uma categoria em específico. São exceções o Webmotors e o iCarros, que além desta abordagem possibilitam a busca por estilo ou necessidade, o que pode ajudar alguém que não esteja buscando um modelo específico. Além das ferramentas de busca, há os catálogos com informações técnicas, onde é possível ver os detalhes de cada modelo em específico. O iCarros e webmotors possibilitam comparar modelos selecionados pelo usuário, o que facilita a visualização de quais itens os modelos possuem ou não, assim como suas dimensões.

Desta forma, esta etapa da pesquisa permitiu conhecer as abordagens e limitações existente no mercado, servindo como ponto de partida para o desenvolvimento de um novo produto. Na seção seguinte será apresentado um contexto a respeito do desenvolvimento de produtos, abordando um histórico de

alguns métodos amplamente utilizados no desenvolvimento de softwares e produtos digitais.

2.4 PRODUTOS DIGITAIS

Produtos digitais podem ser definidos como bens convertidos em formato digital ou objetos baseados em bits, que podem ser distribuídos através de canais eletrônicos, podendo esses se referirem a redes com ou sem fio (TURBAN e KING, 2004) (KOISSO-KANTTILA, 2004). São encontradas na literatura referências a produtos de conhecimento (QUAH, 1998) (QUAH, 1999), referindo-se à mesma ideia de produtos digitais. Quah (1999) os definem como produtos cujas características físicas parecem-se com aquelas do conhecimento, independentemente se esses contêm quantidade significativa de conhecimento, na forma como esse é tradicionalmente conhecido.

Uma vez que um pedaço de conhecimento pode ser codificado, ele pode ter uma existência própria e não mais precisa estar incorporado a outros agentes como o ser humano (HOWELL, 2004). É necessário enfatizar, entretanto, que mais do que somente informação, vários dos produtos e processos físicos que conhecemos podem ser digitalizados (CHOI, STAHL e WHINSTON, 1997). Alguns exemplos de produtos digitais são softwares, imagens, músicas, bases de dados, videogames, receitas, sequências de DNA, mensagens codificadas, e ideias e conhecimento. Os produtos digitais podem ser classificados através de categorias ou características (HUI e CHAU, 2002). O Quadro 4 mostra essas categorias, com alguns exemplos de produtos que compõem cada uma delas.

QUADRO 4 - CATEGORIAS DE PRODUTOS DIGITAIS

Categoria	O que são	Exemplos
Ferramentas e utilidades	Softwares que ajudam os usuários a desempenharem funções específicas ou agem como utilidades suplementares para atingir outros objetivos	Softwares em geral possíveis de se baixar na internet, como editores de texto ou mp3 players.
Produtos baseados em conteúdo	Produtos em que seu valor está no conteúdo informativo	Jornais eletrônicos, música digital, bases de dados.
Serviços online	Serviços que oferecem acesso a recursos úteis, como conexão a servidores, e utilidades on-line que assistem aos usuários na conclusão de determinadas tarefas	Serviço de comunicação por texto e telefonia VoIP, serviços de busca.

FONTE: Adaptado de Hui e Chau (2002)

Três características funcionais dos produtos são apresentadas como critérios para classificação: modo de entrega, granulosidade e capacidade de experimentação. A primeira refere-se ao mecanismo de entrega do produto pela internet, podendo esse ser entregue como um todo de uma vez através de download ou interativamente de forma contínua. A distinção, dessa forma, é feita pela possibilidade de se fazer um download. A granulosidade se refere à possibilidade de se desmembrar o produto para, assim, formar outros a partir do primeiro. A capacidade de experimentação distingue os produtos que podem ser testados pelo usuário antes de adquiri-los daqueles que não podem.

Freiden *et al.* (1998) e Koisso-Kantilla (2004) comparam as características dos produtos digitais com características de produtos e serviços para mostrar que naqueles estão presentes elementos convencionalmente associados tanto a serviços quanto a produtos. Koisso-Kantilla (2004) conclui, dessa forma, que produtos digitais são produtos que também se comportam como serviços. Foram utilizadas quatro dimensões nessa comparação, ilustrada

no Quadro 5: intangibilidade, deterioração, heterogeneidade e produção e consumo simultâneo.

QUADRO 5 - PRODUTOS DIGITAIS COMPARADOS COM PRODUTOS E SERVIÇOS

Característica de serviço	Posição dos Produtos Digitais	Característica de Produto
Intangível	X	Forma física definida
Heterogêneo		X Padronização possível
Produção e consumo simultâneos		X Produção e consumo separados
Impossibilidade de estocar/salvar		X Possibilidade de estocar/salvar

FONTE: Adaptado de Koiso-Kanttila (2004)

(QUAH, 2003), (CHOI, STAHL e WHINSTON, 1997) e (CABRAL e YONEYAMA, 2001) apresentam algumas características físicas dos produtos digitais, algumas delas específicas do meio digital, que determinam seu comportamento nos ambientes de produção e de consumo. Essas características são listadas e explicadas a seguir.

- a) Indestrutibilidade – quando o produto digital é utilizado por um usuário, sua utilidade não é degradada para nenhum outro;
- b) Reprodutibilidade – a quantidade de um produto digital pode ser feita arbitrariamente grande de forma rápida e sem custos;
- c) Transmutabilidade – produtos digitais são cumulativos e emergentes, isto é, novos produtos surgem facilmente a partir dos existentes, com características ausentes nos originais;
- d) Descontinuidade – produtos digitais são indivisíveis e cópias sempre vêm em quantidades inteiras;
- e) Onipresença – produtos digitais estão “em lugar algum e em todo lugar ao mesmo tempo” (QUAH, 2003), ou seja, eles são livremente disseminados pelo globo uma vez que sua comunicação significa seu transporte e sua distribuição.

Por fim, uma característica percebida no processo de geração dos produtos digitais é a dificuldade de distinção entre o que é desenvolvimento e o que é produção de um produto desse tipo (COTA JÚNIOR, 2005). Teorias tanto de gestão de desenvolvimento de produtos quanto de gestão da produção são utilizadas nesses produtos em diferentes situações, principalmente nos processos de software, e são discutidas nas seções seguintes.

2.5 DESENVOLVIMENTO DE PRODUTOS

A inovação é considerada como um ingrediente vital para o sucesso dos negócios. No entanto, a pressão inovadora no mercado globalizado vem crescendo, enquanto a vida média de produtos diminui. De maneira geral, a cada dez ideias de novos produtos, três serão desenvolvidas e apenas um será lucrativo, ou seja, 10% das ideias inovadoras alcançam seu objetivo. O desenvolvimento bem-sucedido de um produto por si só, não garante o seu sucesso, para tal deve-se estar atento aos obstáculos, ter um destino bem definido e manter um ritmo constante no percurso ao objetivo. Isso é chamado de processo de projeto de produto, cujo ponto central é a eficiente gerência de risco, que deve prestar atenção especial as situações de:

- a) Estabelecimento de metas;
- b) Eliminação do produto, assim que se comprove que as metas não serão atingidas.

(BAXTER, 2000).

Takahashi e Takahashi (2007) enfatizam que o desenvolvimento de novos produtos pode ser visto como uma passagem do abstrato, do intangível, que contempla as ideias ainda subjetivas e não muito claras; para o concreto, o tangível, o resultado. Este processo busca transformar objetivos, intenções e ideias em algo concreto, ou seja, em produtos ou soluções pelas quais os consumidores pagarão para satisfazer suas necessidades.

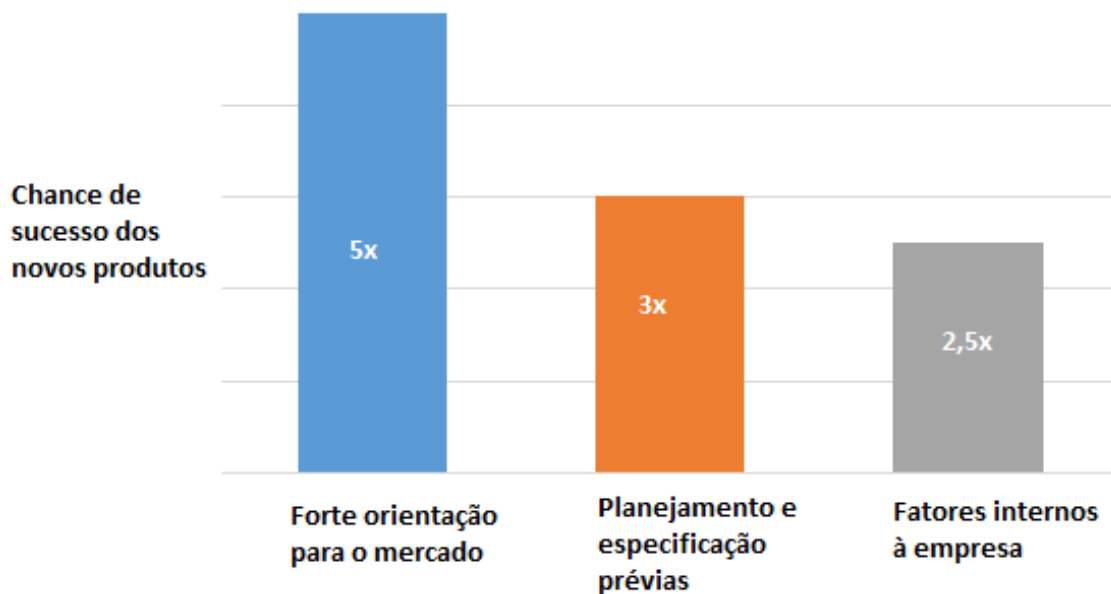
Slack, Chambers e Johnston (2007) ressaltam três aspectos dos produtos e serviços: conceito, pacote e processo. O conceito refere-se ao conjunto de benefícios esperados, isto é, a intenção global do produto/serviço quando percebido da perspectiva do consumidor. Já o pacote de produtos é o conjunto de componentes que proporcionam os benefícios definidos no conceito. Alguns desses componentes são essenciais à compra e não podem ser removidos sem destruir a natureza do pacote, outros servem para melhorar a parte essencial, sendo denominados bens e serviços de apoio. Já o processo define a relação entre os componentes, possibilitando que determinados produtos ou serviços desempenhem sua função e realizem o conceito original.

O desenvolvimento de produtos é considerado portanto uma atividade arriscada, pois envolve diversos interesses e habilidades em relação aos consumidores, vendedores, engenheiros, designers e empresários. No entanto, os fatores de sucesso e fracasso no lançamento de novos produtos podem ser classificados em três grupos principais:

- a) Orientação para o mercado – Forte diferenciação em relação aos concorrentes, apresentar características valorizadas pelos consumidores.
- b) Planejamento e especificação – Estudo de viabilidade técnica e econômica antes de começar o projeto do produto.
- c) Fatores internos – Alta qualidade nas atividades técnicas, equipe talhada para as necessidades de desenvolvimento.

Segundo Baxter (2000), Produtos que eram vistos pelos consumidores como sendo de melhor qualidade em relação aos concorrentes tinham 5,3x mais chances de sucesso dos aqueles considerados marginalmente diferentes. Produtos submetidos a cuidadosos estudos de viabilidade técnica e econômica antes do desenvolvimento, e que possuíam especificações bem definidas tinha 3,3x mais chances de sucesso. Além disso, mantendo-se a alta qualidade nas atividades técnicas relacionadas ao desenvolvimento, as chances de sucesso são 2,5x maiores, conforme ilustrado na Figura 7.

FIGURA 7 - FATORES DE SUCESSO NO DESENVOLVIMENTO DE PRODUTOS



FONTE: (BAXTER, 2000)

A abordagem apresentada por Baxter (2000) no desenvolvimento de produtos pode ser descrita em quatro etapas:

- Identificação de oportunidades. Trata-se do período de pesquisa, onde devem ser estabelecidos os objetivos e o planejamento do produto. Propõe a identificação das necessidades e restrições do consumidor, expressa de forma simples e concisa. Assim como em projetos de pesquisa, o empreendedor deve responder nesta etapa da pesquisa, qual problema que está tentando resolver, por que esse problema existe, e se este problema é uma parte específica de um problema maior ou mais amplo;
- Análise de concorrentes e produtos concorrentes. Esta etapa serve para monitorar as empresas concorrentes e seus produtos, determinar o sucesso e fracasso das mesmas e direcionar desta forma os resultados desejados;
- Configuração do projeto. Começa com o conceito escolhido e termina com o protótipo desenvolvido e testado. O conceito do produto deve começar com a geração de ideias, esquecendo das restrições

práticas e procurando ideias fora do domínio normal do problema. Para tal, é sugerido o uso da ferramenta de *brainstorming*.

- Especificação do projeto. É onde o detalhamento do projeto é apresentado de fato, com base nos resultados obtidos das configurações e avaliações estabelecidas, determinando desta forma como o produto será produzido.

(BAXTER, 2000).

No entanto, no que diz respeito ao desenvolvimento de produtos que são softwares, a literatura dispõe de uma grande diversidade de metodologias que tem por finalidade aumentar as chances de sucesso no produto final. Assim sendo, na próxima seção serão abordados alguns dos métodos de desenvolvimento de software mais conhecidos.

2.6 DESENVOLVIMENTO DE SOFTWARE

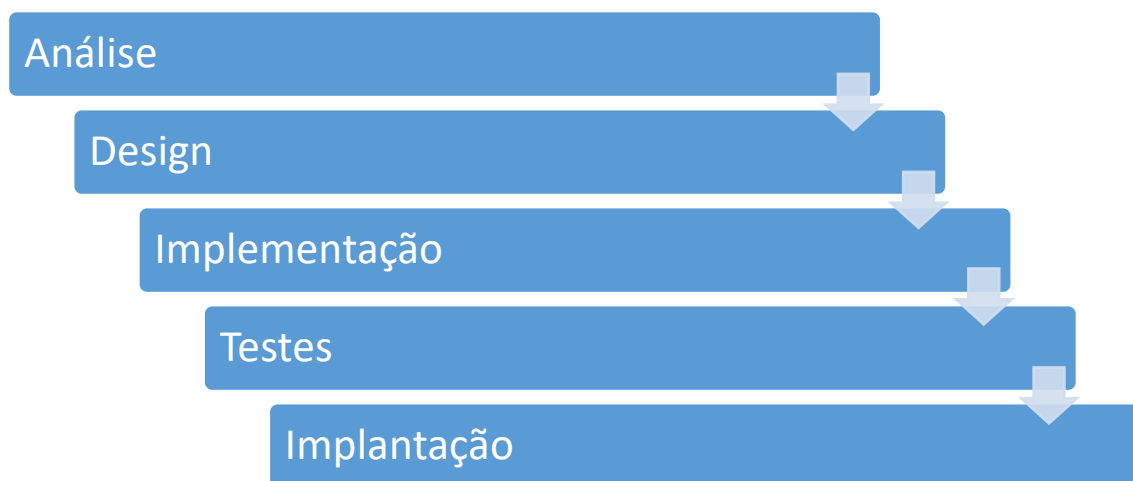
Desenvolver software é uma atividade difícil e arriscada. Entre os riscos, pode se citar gastos que superam o orçamento, consumo de tempo que supera o cronograma, funcionalidades que não resolvem os problemas dos usuários, baixa qualidade dos sistemas desenvolvidos e cancelamento do projeto por inviabilidade. Diante deste cenário, é possível afirmar que o uso de uma metodologia adequada no desenvolvimento de softwares pode reduzir ou mesmo eliminar estes riscos. As seções seguintes visam apresentar algumas premissas das metodologias tradicional e ágil no desenvolvimento de software, assim como suas principais diferenças.

2.6.1 METODOLOGIA TRADICIONAL DE DESENVOLVIMENTO DE SOFTWARE

No conceito chamado de tradicional, metodologia de desenvolvimento de software diz respeito a um conjunto de atividades e seus resultados associados que auxiliam na produção de software. Todas estas metodologias, entre outras funções, tentam reduzir o alto risco associado no desenvolvimento (PRESSMAN, 2011).

Estas metodologias têm por base de alguma forma o chamado desenvolvimento em cascata, ilustrado na Figura 8.

FIGURA 8 - METODOLOGIA DE DESENVOLVIMENTO EM CASCATA



FONTE: Adaptado de Teles (2004).

Estes modelos de desenvolvimento de software apresentam certas premissas básicas que são derivadas do modelo industrial e que influenciam profundamente o desenvolvimento tradicional. São estas:

- Linearidade;
- Determinismo;
- Especialização;
- Foco na execução;
- Crescimento exponencial do custo de alteração.

A linearidade desta abordagem fica evidente no modelo em cascata, onde os processos ocorrem de forma iterativa e incremental. O processo de desenvolvimento em cascata é conhecido por gerar determinados resultados com base nas especificações e deverá gerar resultados semelhantes quando suas regras forem rigidamente seguidas. Pode se dizer, portanto, que o desenvolvimento tradicional, assim como a indústria, persegue o determinismo por acreditar ser uma forma de reduzir erros e perdas de tempo.

A especialização no desenvolvimento tradicional aparece claramente na rígida divisão de papéis entre membros de uma equipe, como analistas, projetistas, programadores, testadores e implantadores. Desta forma, o processo de desenvolvimento pode ser desmembrado em atividades especializadas que serão executadas de independentemente e terão seus resultados integrados na fase final para formar o produto. A especialização torna as tarefas mais simples e como consequência facilita seu determinismo. (TELES, 2004).

Quando a especialização existente é suficiente para que as tarefas sejam simples e determinísticas, não existe a necessidade de pensar sobre o que se está fazendo por parte de quem executa o trabalho. Ou seja, basta que as pessoas envolvidas no processo de desenvolvimento executem corretamente as tarefas que lhe cabem, conforme estipulado, para que a especificação seja transformada corretamente em software. Isto gera uma grande valorização dos processos e uma conseqüente desvalorização das pessoas, no desenvolvimento de um software (DEMARCO e LISTER, 1987) (TELES, 2004).

Com este modelo de desenvolvimento tradicional, o custo de uma alteração cresce exponencialmente à medida que o processo de desenvolvimento avança. Aceitar esta premissa tem como consequência uma busca natural por processos determinísticos, que prometem menos alterações e maior previsibilidade. Isto também reflete e justifica de certa forma, a busca por metodologia lineares, determinísticas, especializadas e focadas na execução, próximas dos processos industriais. Estas premissas não são um problema se o design do projeto é correto e as implementações puderem ser corrigidas durante os testes, no entanto estas afirmações não funcionam bem na prática, sendo necessário passar pelas etapas do processo cascata mais de uma vez,

assim como é necessário testar o software em etapas, e não apenas no final do desenvolvimento (TELES, 2004).

As metodologias tradicionais foram muito utilizadas no passado em um contexto de desenvolvimento de software muito diferente do atual, baseado apenas em um mainframe e terminais burros. No passado, o custo de fazer alterações e correções era muito alto, uma vez que o acesso aos computadores era limitado a apenas um mainframe e terminais burros, além de não existirem ferramentas de apoio, como depuradores e analisadores de código. Por isso o software era todo planejado e documentado antes de ser implementado (SOARES, 2004).

Isto posto, metodologias ágeis têm sido apontadas como uma alternativa para o desenvolvimento de software. As metodologias tradicionais podem ainda ser aplicadas em situações em que os requisitos do sistema são estáveis e requisitos futuros são previsíveis. Entretanto, em projetos em que os requisitos são passíveis de alterações, as equipes são pequenas, as datas de entrega do software são curtas e o desenvolvimento rápido é fundamental, não pode haver requisitos estáticos, necessitando então se utilizar de metodologias ágeis.

2.6.2 METODOLOGIA ÁGIL DE DESENVOLVIMENTO DE SOFTWARE

O termo “Metodologias Ágeis” tornou-se popular quando em 2001, dezessete especialistas em processos de desenvolvimento de software representando os métodos Scrum (SCHWABER e BEEDLE, 2002), Extreme Programming (XP) (BECK, 1999) e outros, estabeleceram princípios compartilhados por todos esses métodos. Foi então criada a Aliança Ágil e o estabelecimento do Manifesto Ágil (2001). Os conceitos chave do Manifesto Ágil são:

- Indivíduos e interação entre eles mais que processos e ferramentas;
- Software em funcionamento mais que documentação abrangente;
- Colaboração com o cliente mais que negociação de contratos;
- Responder a mudanças mais que seguir um plano.

Este manifesto não rejeita os conceitos de desenvolvimento tradicional de processos, ferramentas, documentação, negociação de contratos ou planejamento, mas mostra que eles têm importância secundária se comparado aos indivíduos e interações, a capacidade do software de estar executável, colaboração do cliente e respostas rápidas a mudanças e alterações (SOARES, 2004).

Desta forma, no desenvolvimento ágil, os projetos adotam o modelo iterativo e em espiral. Neste processo, todas as fases descritas no modelo em cascata são executadas diversas vezes ao longo do projeto, produzindo ciclos curtos que se repetem ao longo de todo o desenvolvimento, sendo que, ao final de cada ciclo, sempre se tem um software funcional, testado e aprovado. Os ciclos são chamados de iterações e crescem em funcionalidades a cada repetição, sendo que, no último ciclo, todas as funcionalidades desejadas estarão implementadas, testadas e aprovadas (TELES, 2004).

Na seção seguinte serão brevemente apresentados dois dos métodos mais utilizados de desenvolvimento ágil de software.

2.6.2.1 EXTREME PROGRAMMING

A *eXtreme Programming* (XP) é uma abordagem deliberada e disciplinada para o desenvolvimento de software. A XP vem sendo bastante utilizada e vem tomando espaço que antes pertencia a metodologias tradicionais (BECK, 1999). Essa metodologia, carrega consigo todas as características de desenvolvimento ágil, desde o foco na satisfação do cliente até assumir que sempre irão ocorrer mudanças nos requisitos e especificações. Para delinear o processo de desenvolvimento, a XP foi estruturada segundo algumas práticas que permitem aumentar a qualidade do software gerado. Essas regras e práticas se dividem de forma simplificada em quatro categorias:

- Planejamento;
- Design;
- Codificação;
- Testes.

(BECK, 1999).

Desta forma a XP tem por base 12 práticas que se enquadram nas categorias anteriores, ilustradas na Figura 9 e explicadas em seguida.

FIGURA 9 - PRÁTICAS DA EXTREME PROGRAMMING



FONTE: Adaptado de Soares, 2004.

- Planejamento: visto que a XP tem por base requisitos atuais para desenvolvimento, o planejamento consiste em decidir o que é necessário ser feito e o que pode ser adiado no projeto. Além disso, o planejamento deve evitar problema entre a área de negócios e de desenvolvimento. As duas áreas devem cooperar para o sucesso do projeto, e cada uma deve focar em partes específicas do projeto. Desta forma, cabe a área de negócios decidir sobre o escopo, composição das versões e datas de entrega, e cabe aos desenvolvedores decidir sobre estimativas de prazo, processo de desenvolvimento e cronograma.

- Entregas frequentes: visa à construção de um software simples, e conforme os requisitos surgem, o software é atualizado. Cada versão entregue deve ter o menor tamanho possível, contendo os requisitos de maior valor para o negócio. Entregas com prazos que possibilitem o feedback rápido do cliente evitam surpresas caso o software seja entregue após muito tempo e melhora as avaliações do cliente, aumentando a probabilidade de o software final estar de acordo com os requisitos.
- Metáfora: são as descrições de um software sem a utilização de termos técnicos, com o intuito de guiar o seu desenvolvimento.
- Projeto simples: o programa desenvolvido pelo método XP deve ser o mais simples possível e satisfazer os requisitos atuais, sem se preocupar com os requisitos futuros. Os requisitos devem ser adicionados assim que eles realmente existirem.
- Testes: a XP focaliza a validação do projeto durante todo o processo de desenvolvimento. O software é desenvolvido criando primeiramente os testes.
- Refatoração: o aperfeiçoamento do projeto do software está presente durante todo o desenvolvimento. A refatoração deve ser feita apenas quando necessário, ou seja, quando é percebido que é possível simplificar o módulo atual do software sem perder nenhuma funcionalidade.
- Programação em pares: a implementação do código é feita em dupla, ou seja, dois desenvolvedores trabalham em um único computador. Enquanto o desenvolvedor que está com o controle do teclado e do mouse implementa o código, o outro observa continuamente, procurando identificar erros sintáticos e semânticos e pensando em como melhorar o código que está sendo programado. Esses papéis devem ser alterados continuamente. Uma grande vantagem desta abordagem é a possibilidade de os desenvolvedores aprenderem continuamente um com o outro.

- Propriedade coletiva: o código desenvolvido pertence a todos os membros da equipe. Qualquer pessoa que percebe que pode adicionar valor a um código pode fazê-lo, desde que faça a bateria de testes necessária. Isto é possível porque na XP todos são responsáveis pelo software final. A vantagem desta prática é que, caso um membro da equipe deixe o projeto antes do fim, o restante da equipe deve conseguir continuar o projeto sem grandes dificuldades, pois todos conhecem as partes do software.
- Integração contínua: interagir e construir o sistema de software várias vezes por dia, mantendo os programadores em sintonia, além de possibilitar processos rápidos. Integrar apenas um conjunto de modificações de cada vez é uma prática que funciona bem porque fica óbvio quem deve fazer as correções quando os testes falham: a última equipe que integrou código novo ao software. Esta prática é facilitada com o uso de apenas uma máquina de integração, que deve ter livre acesso a todos os membros da equipe.
- 40 horas de trabalho semanal: não se deve fazer horas extras constantemente. Caso seja necessário trabalhar mais de 40 horas por mais de uma semana consecutiva, entende-se que existe um problema sério no projeto e que deve ser resolvido não com horas extras, mas com melhor planejamento, por exemplo. Esta prática ratifica o foco da metodologia nas pessoas e não em processos e planejamentos. Altera-se o planejamento ao invés de sobrecarregar os indivíduos.
- Cliente presente: a participação do cliente durante o desenvolvimento do projeto é fundamental. O cliente deve estar disponível para resolver questões relacionadas aos requisitos, evitando atrasos e construções errôneas. O cliente pode inclusive ser integrado à equipe de desenvolvimento.
- Código padrão: a padronização na arquitetura do código é importante para que este possa ser compartilhado entre todos os programadores.

(BECK, 1999)

2.6.2.2 SCRUM

O Scrum é uma metodologia ágil que apresenta uma grande comunidade de usuários. Seu objetivo é fornecer um processo adequado para projeto e desenvolvimento orientado a objeto (SCHWABER e BEEDLE, 2002).

A abordagem da Scrum é empírica, onde aplica algumas ideias da teoria de controle de processos industriais, tendo por iniciativas essenciais:

- Comunicação;
- Trabalho em equipe;
- Flexibilidade;
- Entregas de produtos de software funcionando.

A ideia central da Scrum é que o desenvolvimento de softwares envolve muitas variáveis técnicas e do ambiente, como requisitos, recursos e tecnologia, e que as mesmas podem mudar durante o processo. Isto requer flexibilidade para acompanhar as mudanças. As coisas devem estar visíveis a todos os envolvidos no desenvolvimento, a inspeção e as ações para adaptação do produto de software devem ser realizadas frequentemente e o resultado do processo deve ser um software que é realmente útil para o cliente (SCHWABER e BEEDLE, 2002).

A metodologia Scrum é baseada em princípios semelhantes aos da XP: equipes pequenas, requisitos pouco estáveis e iterações curtas. No entanto, as dimensões diferem. O desenvolvimento é dividido em iterações (*sprints*) de trinta dias. Equipes pequenas, são formadas por projetistas, programadores, engenheiros e gerentes de qualidade. Estas equipes trabalham nas funcionalidades definidas no início de cada *sprint*.

Reuniões de acompanhamento diárias devem ser realizadas, preferencialmente de curta duração, onde são discutidos pontos como o que foi feito desde a última reunião e o que precisa ser feito até a próxima. As dificuldades e os fatores de impedimento encontrados são identificados e resolvidos (SCHWABER, 1995).

O ciclo de vida da Scrum é baseado em três fases principais:

- Pré-planejamento: os requisitos são descritos em um documento chamado *backlog*. Posteriormente estes requisitos são priorizados e são realizadas estimativas de esforço para o seu desenvolvimento. O planejamento inclui, entre outras atividades, a definição da equipe de desenvolvimento, as ferramentas a serem usadas, os riscos do projeto e as necessidades de treinamento. Por fim é proposta uma arquitetura de desenvolvimento e eventuais alterações nos requisitos são identificadas, assim como seus possíveis riscos.
 - Desenvolvimento: as variáveis técnicas e do ambiente identificadas previamente são observadas e controladas durante a fase de desenvolvimento. Na Scrum o controle é feito continuamente levando em consideração estas variáveis, ao contrário de metodologias tradicionais. Isto aumenta a flexibilidade para acompanhar as mudanças necessárias. O software é desenvolvido em ciclos (*sprints*) em que novas funcionalidades são adicionadas. Cada um desses ciclos é desenvolvido de forma tradicional, onde primeiramente é feita a análise, em seguida o projeto, implementação e testes. Cada um desses ciclos é planejado para durar de uma semana a um mês.
 - Pós-planejamento: após a fase de desenvolvimento são realizadas reuniões para analisar o progresso do projeto e demonstrar o software atual para os clientes. Nesta fase são feitas as etapas de integração, testes finais e documentação.
- (SCHWABER e BEEDLE, 2002).

Até o momento, foram apresentados trabalhos da literatura que dizem respeito ao tema escolhido para esta pesquisa, assim como a literatura a respeito

do desenvolvimento de produtos e de software. Isto posto, serão introduzidos a seguir os conceitos de estatística multivariada que serão empregados na metodologia escolhida para possibilitar a ferramenta que será desenvolvida nesta pesquisa, com o intuito de cumprir o objetivo principal proposto.

2.7 ANÁLISE MULTIVARIADA

A análise multivariada é um conjunto de técnicas estatísticas utilizadas em situações em que muitas variáveis são estudadas em um único banco de dados, originando múltiplos dados para um mesmo indivíduo ou objeto em investigação. Quanto maior o número de variáveis a serem utilizadas no estudo, mais complexas tornam-se as análises pelos métodos de estatística univariada. Este conjunto de técnicas tem sido muito utilizado para resolver diversos problemas nas áreas de educação, geologia, química, engenharia entre outros, devido à grande disseminação de pacotes e softwares estatísticos (MINGOTI, 2005).

Hair, Anderson *et al.* (2005) afirmam que técnicas de análise multivariada permitem que organizações criem conhecimento, melhorando assim suas tomadas de decisões. O termo multivariada pode ser utilizado para se referir ao exame das relações entre mais de duas variáveis.

Neste trabalho serão apresentadas algumas técnicas que serão utilizadas para:

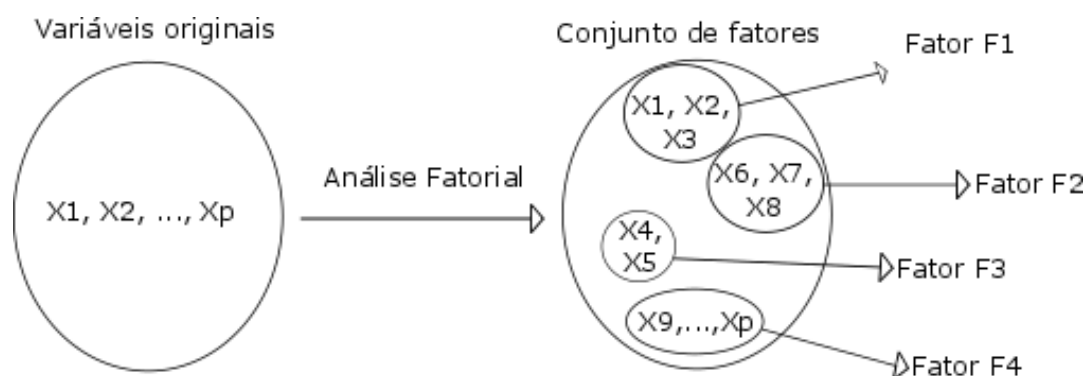
- a) Reduzir o número de variáveis que irão compor a solução final do problema;
- b) Realizar o agrupamento dos objetos (observações) de estudo de acordo com a similaridade de suas variáveis;
- c) Discriminar novos objetos a grupos previamente conhecidos, classificando assim novas observações.

2.8 ANÁLISE FATORIAL

A Análise Fatorial (AF) trata-se de uma abordagem estatística que pode ser usada para analisar inter-relações entre um grande número de variáveis e explicá-las em termos de suas dimensões comuns, chamadas fatores. O objetivo então é condensar a informação que um grande número de variáveis apresenta originalmente, em conjunto menor de fatores com o mínimo de perda de informação possível (HAIR, ANDERSON, *et al.*, 2005).

De forma semelhante Johnson e Wichern (2007), afirmam que o propósito da AF é, se possível, descrever as correlações entre um grande número de variáveis em termos de fatores, que seriam um conjunto de poucas variáveis aleatórias não-observáveis. A Figura 10 exemplifica, de forma visual, a realização de uma análise fatorial, onde temos a esquerda o conjunto de variáveis originais, e a direita o conjunto de fatores resultantes da análise, sendo dentro de um grupo, as variáveis são altamente correlacionadas entre si, enquanto de um grupo para outro as correlações são baixas.

FIGURA 10 - ILUSTRAÇÃO DE UMA ANÁLISE FATORIAL



FONTE: O autor (2016).

Os dois principais usos da AF são o resumo e a redução de dados. Ao resumir os dados, a AF obtém dimensões que, uma vez entendidas e interpretadas, descrevem os dados em número muito menor de conceitos. A

redução da quantidade de dados é então alcançada, calculando escores para cada dimensão obtida, e substituindo as variáveis originais por estes escores.

A principal vantagem em utilizar técnicas multivariadas, é a habilidade de acomodar múltiplas variáveis em uma tentativa de compreender as complexas relações, que não seria possível com o uso de métodos univariados e bivariados. Porém, ao aumentar o número de variáveis, também se aumenta a possibilidade de que nem todas estas variáveis são não-correlacionadas e são representativas de conceitos distintos. A AF permite que, ao invés disso, grupos de variáveis sejam inter-relacionados de forma que representem um conceito de forma mais geral. Importante ressaltar, que a experiência do pesquisador conta em relação a necessidade de conhecer como as variáveis estão correlacionadas, a fim de uma melhor interpretação dos resultados. Ainda, se o número de variáveis é muito grande, ou faz-se necessário uma melhor representação de um número menor de conceitos ao invés de representar cada uma das diversas facetas, a AF pode ser utilizada de forma a auxiliar a seleção de um subconjunto representativo de variáveis ou então na criação de novas variáveis que substituam as variáveis originais, mas ainda mantendo seu caráter original. Muitos pesquisadores consideram esta técnica apenas exploratória, útil na busca da compreensão da estrutura em um conjunto de variáveis ou então como um método de redução de dados (HAIR, ANDERSON, *et al.*, 2005).

A AF não diz respeito a uma única técnica estatística, mas sim a um conjunto de técnicas relacionadas, com o intuito de tornar os dados iniciais mais claros para a interpretação. Um dos métodos mais conhecidos para a extração dos fatores, é realizado por meio da análise de componentes principais, onde parte-se do pressuposto que se pode definir X vetores estatisticamente não correlacionados, ou ortogonais, que são as componentes principais, resultantes das combinações lineares do conjunto inicial de variáveis. Assim como a análise de componentes principais, a AF é uma técnica que pode ser aplicada a um conjunto de variáveis visando descobrir quais dessas são mais relevantes na composição de cada fator, de forma que os fatores sejam independentes um dos outros (REIS, 1997).

d) Os erros são independentes, ou seja:

$$V(\underline{\varepsilon}) = E(\underline{\varepsilon}') = \psi_{pp} = \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_p \end{bmatrix}$$

e) \underline{F}_{m1} e $\underline{\varepsilon}_{p1}$ são independentes, ou seja, $cov(\underline{\varepsilon}_{p1}, \underline{F}_{m1}) = \underline{0}$

Desta forma, a estrutura da covariância Σ deste modelo pode ser representada por:

$$\Sigma = cov(\underline{X}) = LL' + \psi \quad (3)$$

Ou:

$$V(X_i) = l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2 + \psi_i \quad (4)$$

Onde:

$l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2 =$ São as comunalidades, a porção da variância $V(X_i)$ que é distribuída por m fatores comuns;

$\psi_i =$ Variância específica.

Caso seja usada a matriz de correlação, então l_{ij} será o coeficiente de correlação. A composição da matriz covariância é $h_i^2 = V(X_i) - \psi_i$. Há a exigência de que $m \leq p$ para que a estrutura dos fatores comuns não seja mais complicada que aquela para os fatores observados (JOHNSON e WICHERN, 2007).

2.8.2 DETERMINAÇÃO DE FATORES

Uma vez especificadas as variáveis, e preparada a matriz de correlação, o pesquisador pode aplicar a análise fatorial para identificar a estrutura latente de relações, assim, devem ser tomadas as decisões em relação ao método de

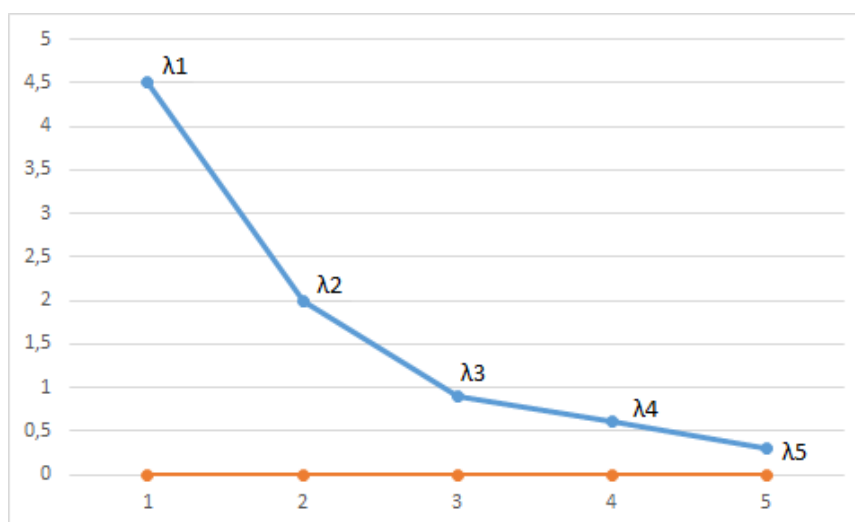
extração de fatores, e ao número de fatores que serão selecionados para representar a estrutura original dos dados (HAIR, ANDERSON, *et al.*, 2005).

Para a estimação do número de fatores é comum, entre outros critérios, a utilização do critério de Kaiser, o critério do gráfico de declive ou então a porcentagem da variância explicada.

O critério de Kaiser, também conhecido como critério do autovalor, assume que apenas fatores com autovalores maiores ou iguais a um ($\lambda_i \geq 1$) são considerados para prosseguir com a análise. Utiliza-se este critério quando estiver trabalhando com a matriz de correlação, pois trabalha-se com variáveis padronizadas, ou seja, que tem média igual a zero e variância igual a um. Quando se obtém um autovalor menor que um ($\lambda_i < 1$) em algum fator, significa que este valor é menos significativo do que na variável original (CORRAR, PAULO e FILHO, 2009; MAROCO, 2007).

De maneira similar, o critério do gráfico de declive parte do princípio que a maior parte da variância é explicada pelos primeiros fatores. Desta forma, ordenam-se os autovalores em ordem decrescente e quando a diferença entre dois autovalores consecutivos se torna pequena, uma linha de corte é traçada entre esses dois autovalores, e selecionam-se todos os anteriores a essa linha (CORRAR, PAULO e FILHO, 2009). Este critério pode ser aplicado de forma visual como exemplificado na Figura 11.

FIGURA 11 - EXEMPLO DE CRITÉRIO DE SELEÇÃO DE GRÁFICO DE DECLIVE



FONTE: O autor (2016).

Percebe-se neste exemplo que a partir do autovalor 3, o gráfico começa a formar uma tendência de reta, então de acordo com este critério seriam selecionados os autovalores λ_1, λ_2 e λ_3 (Figura 11).

Quanto ao critério da porcentagem da variância explicada, a experiência do pesquisador é levada em consideração para escolher uma quantidade de fatores que represente uma porcentagem da variância que, não representa uma grande perda de informação em relação aos dados originais, mas ao mesmo tempo represente uma simplificação que justifique o emprego da técnica.

Qualquer um destes três critérios apresentados pode ser utilizado para a seleção de número de fatores extraídos, sendo que esta quantidade pode mudar entre um critério e outro. Neste caso, dependerá dos conhecimentos teóricos do pesquisador a respeito do problema a se resolver, assim como o tipo de pesquisa que será realizada (CORRAR, PAULO e FILHO, 2009).

2.8.3 MÉTODO DAS COMPONENTES PRINCIPAIS

O método das componentes principais é um procedimento utilizado para estimar os carregamentos fatoriais l_{ij} e as variâncias específicas ψ_i através da decomposição espectral dos autovalores λ_i e os respectivos autovetores e_i provenientes da matriz covariância Σ , sendo $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. A matriz de covariância pode ser escrita da seguinte forma (JOHNSON e WICHERN, 2007):

$$\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \dots + \lambda_p e_p e_p' \quad (5)$$

Ou, se trabalhando com a covariância amostral:

$$S = \hat{\lambda}_1 \hat{e}_1 \hat{e}_1' + \hat{\lambda}_2 \hat{e}_2 \hat{e}_2' + \dots + \hat{\lambda}_p \hat{e}_p \hat{e}_p' \quad (6)$$

Desta forma, considerando que $m < p$ é o número de fatores comuns a serem extraídos, a matriz de carregamentos \hat{L}_{pm} será composta por m fatores e é calculada utilizando a seguinte equação:

$$\hat{L}_{pm} = \sqrt{\hat{\lambda}_1} \begin{bmatrix} \hat{e}_{11} \\ \hat{e}_{21} \\ \vdots \\ \hat{e}_{p1} \end{bmatrix} \sqrt{\hat{\lambda}_2} \begin{bmatrix} \hat{e}_{12} \\ \hat{e}_{22} \\ \vdots \\ \hat{e}_{p2} \end{bmatrix} \dots \sqrt{\hat{\lambda}_m} \begin{bmatrix} \hat{e}_{1m} \\ \hat{e}_{2m} \\ \vdots \\ \hat{e}_{pm} \end{bmatrix} = \begin{bmatrix} \hat{l}_{11} & \hat{l}_{12} & \dots & \hat{l}_{1m} \\ \hat{l}_{21} & \hat{l}_{22} & \dots & \hat{l}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{l}_{p1} & \hat{l}_{p2} & \dots & \hat{l}_{pm} \end{bmatrix} \quad (7)$$

Onde:

\hat{L}_{pm} = Matriz de carregamentos fatoriais estimada;

$\hat{\lambda}_i$ = Autovalor estimado referente a variável i ;

\hat{e}_{ij} = Autovetor estimado referente a i -ésima variável j -ésimo fator;

\hat{l}_{ij} = Carregamento, ou peso, para a i -ésima variável no j -ésimo fator.

Após estimada a matriz de carregamentos fatoriais \hat{L}_{pm} , é possível calcular a matriz de variâncias específicas $\hat{\psi}_{pp}$, dada por (JOHNSON e WICHERN, 2007):

$$\hat{\psi}_{pp} = S_{pp} - \hat{L}_{pm}(\hat{L}_{pm})' = \begin{bmatrix} \hat{\psi}_1 & 0 & \dots & 0 \\ 0 & \hat{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\psi}_p \end{bmatrix} \quad (8)$$

Onde:

S_{pp} = Matriz de covariância amostral;

$\hat{\psi}_1$ = Variância específica estimada da i -ésima variável, dada por:

$$\hat{\psi}_1 = s_{ii}^2 - \sum_{j=1}^m \hat{l}_{ij}^2 \quad (9)$$

Por fim, temos as comunalidades estimadas, dadas por:

$$\hat{h}_i^2 = \sum_{j=1}^m \hat{l}_{i1}^2 + \hat{l}_{i2}^2 + \dots + \hat{l}_{im}^2 \quad (10)$$

Onde \hat{h}_i^2 representa a comunalidade estimada para a i -ésima variável.

2.9 ANÁLISE DE AGRUPAMENTOS

Uma questão que pesquisadores de diversas áreas enfrentam é como organizar dados observados em estruturas que agrupem subconjuntos semelhantes. Ao trabalhar com uma determinada população ou amostra, por vezes é necessário separar as diversas observações em grupos, com o intuito de entender melhor o fenômeno estudado ou identificar se as variáveis observadas na pesquisa representam a população de forma a se aplicar reconhecimento de padrões a esta população. Para tal, pode-se utilizar a Análise de Agrupamentos (AA). A AA difere de métodos de classificação, como Análise Discriminante (AD) ou Regressão Logística, pois na classificação parte-se de um número de grupos conhecidos e o objetivo é então alocar novas observações nestes grupos enquanto agrupamento trata-se de uma técnica mais primitiva, pois nenhuma suposição é tomada quanto ao número ou estrutura de grupos, mas sim, com base na similaridade ou distância (JOHNSON e WICHERN, 2007).

A Análise de Agrupamentos é um método que reúne objetos em grupos, tais que os objetos em um mesmo grupo são mais similares entre si do que com os de outros grupos. Entretanto, deve-se atentar a observações atípicas, que distorcem a verdadeira estrutura dos grupos, tornando-os assim não representativos da verdadeira estrutura da população. Isto posto, uma verificação preliminar das observações é necessária e é sugerido, portanto, a preparação de um diagrama de perfil gráfico, onde as variáveis são listadas ao longo do eixo horizontal, enquanto os valores são listados ao longo do eixo vertical. Desta forma, cada ponto do gráfico representa o valor da variável

correspondente, sendo assim conectados para facilitar a interpretação visual (HAIR, ANDERSON, *et al.*, 2005).

Em relação às técnicas de agrupamento, são admitidas abordagens hierárquicas e abordagens não-hierárquicas (ou de partição). As primeiras consideram que de início cada indivíduo se encontra isolado e o processo segue aproximando-os de acordo com suas similaridades até que se atinja uma estabilidade relativa, que irá variar em função dos objetivos do trabalho. As técnicas de partição operam em sentido contrário, particionando um agrupamento único inicial.

É comum quando se trabalha um grande número de variáveis, tentar reduzir a dimensão deste conjunto de variáveis relacionando então a AA com outras técnicas multivariadas já conhecidas, como a análise fatorial, canônica ou de componentes principais. Além disso, para verificar se a partição obtida com essa análise é adequada, quando já são conhecidos os grupos e os seus componentes, é comum a utilização da análise discriminante.

Para se encontrar o cluster natural dos itens ou variáveis, objetivo da AA, deve-se desenvolver uma escala com o intuito de medir o grau de associação dos objetos, ou similaridades. Quando deseja-se agrupar itens, geralmente é usado algum tipo de distância, como a euclidiana, para indicar a proximidade. No entanto para agrupar variáveis, baseia-se em seus coeficientes de correlação. Entre algumas métricas de distância mais populares estão a distância euclidiana, o quadrado da distância euclidiana, a distância city-block (Manhattan), a métrica de Mahalanobis e a métrica de Minkowski (JOHNSON e WICHERN, 2007).

2.9.1 ALGORITMOS DE AGRUPAMENTO HIERÁRQUICO

Tratando-se de uma grande população em estudo, raramente pode-se examinar todas as possibilidades de grupos que podem ser formados de forma aleatório e devido a este problema, uma variedade de algoritmos de agrupamentos surgiu com a finalidade de formar grupos sem ser necessário

avaliar todas as possíveis configurações. Johnson e Wichern (2007) apontam a existência de ampla variedade de algoritmos de agrupamentos. Nos métodos de agrupamento hierárquico, inicialmente tem-se tantos objetos quanto grupos (cada objeto é um grupo). Os objetos mais similares são agrupados primeiro e estes grupos iniciais vão se fundindo até que se forme um grupo único com todos os objetos conforme o seguinte procedimento:

a) Inicialmente tem-se N grupos, cada grupo contendo um dos N objetos que se pretende agrupar e então é calculada a matriz simétrica de distâncias $N \times N$, $D = (d_{ij})$, d_{ij} sendo a distância entre o objeto "i" e o objeto "j";

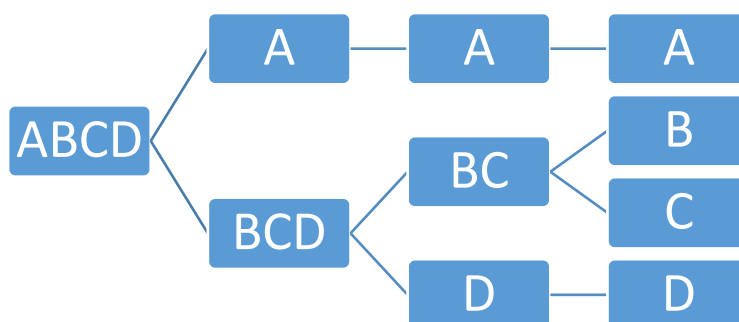
b) Na matriz D, encontra-se o par de grupos de menor distância (menor d_{ij}) e junta-se estes dois grupos em um novo grupo;

c) É denominado o novo grupo como AB por exemplo, se este é formado a partir do grupo A e do grupo B. É comum também denominar o novo grupo como N+1 se os grupos são numerados, por exemplo, os grupos iniciais são numerados de 1 a 15, então o novo grupo formado é denominado 16;

d) São repetidos os passos b e c por N-1 vezes, denominando cada grupo e anotando quais grupos deram origem a este novo, até ser formado um grupo único.

Este algoritmo é ilustrado na Figura 12.

FIGURA 12 - EXEMPLO DE AGRUPAMENTO HIERÁRQUICO



FONTE: O autor (2016).

O algoritmo separa um grupo inicial com todas as observações, em diversos agrupamento até que se tenha uma observação em cada grupo, cabendo ao pesquisador decidir quantos grupos serão formados.

Os resultados podem ser apresentados na forma de um diagrama bidimensional conhecido como dendrograma. Por meio desta ferramenta gráfica e de conhecimento prévio a respeito dos dados, determina-se uma distância de corte para definir então os grupos (JOHNSON e WICHERN, 2007).

2.9.2 MÉTODO DE AGRUPAMENTO NÃO-HIERÁRQUICO

Usa-se o agrupamento não-hierárquico quando se deseja formar k grupos de itens ou objetos, sendo o método mais conhecido o algoritmo das k-médias. O método utilizado para o agrupamento usando o método k-médias é o seguinte (JOHNSON e WICHERN, 2007):

- a) Dividem-se os objetos em K grupos arbitrariamente;
- b) Realoca-se cada objeto no grupo cujo centroide (média) esteja mais próximo. Usualmente, a distância é calculada usando a distância euclidiana de cada objeto, com dados padronizados ou não, e a média de cada grupo. Recalcula-se a média para os grupos que perderam ou ganharam objetos (houve realocações);
- c) Repete-se o passo b até não haver mais realocações;

2.9.3 MÉTODOS DE LIGAÇÕES

As seções anteriores discorreram sobre algoritmos de agrupamento hierárquicos e não-hierárquicos. Os agrupamentos hierárquicos, são realizados por meio de ligações e segundo Johnson e Wichern (2007), os tipos de ligações mais comum são as Ligações Simples, Ligações Completas, Método das Médias

das Distâncias e Método de Ward. Em seguida estes métodos serão apresentados.

2.9.4 LIGAÇÕES SIMPLES

Nas ligações simples, os grupos são formados a partir dos dois grupos com menor distância (vizinho mais próximo). Em primeiro lugar, encontra-se a menor distância da matriz D, forma-se o grupo dos objetos correspondentes formando por exemplo, o grupo AB. No passo 3 do agrupamento hierárquico, a distância entre o grupo AB e qualquer outro grupo, como por exemplo, C, é calculado segundo a equação (JOHNSON e WICHERN, 2007):

$$d_{(AB)C} = \min\{d_{AC}, d_{BC}\} \quad (11)$$

Onde:

$d_{(AB)C}$: É a métrica de distância calculada entre o grupo AB e o grupo C;

d_{AC} : É a distância entre o grupo A e o grupo C;

d_{BC} : É a distância entre o grupo B e o grupo C.

2.9.5 LIGAÇÕES COMPLETAS

Na ligação completa, o processo é semelhante ao de ligação simples, com a única diferença que, usando o exemplo anterior, a distância entre o grupo AB e os outros grupos, é de acordo com a máxima distância de cada grupo, de forma análoga a equação (1), conforme demonstrado na equação:

$$d_{(AB)C} = \max\{d_{AC}, d_{BC}\} \quad (12)$$

2.9.6 MÉTODO DAS MÉDIAS DAS DISTÂNCIAS

Na ligação das médias das distâncias, a distância entre dois grupos é dada pela distância média entre os pares de objetos pertencentes a cada grupo. Como nos outros métodos, encontra-se a menor distância entre dois grupos, forma-se o novo grupo a partir destes dois, e então determina-se a distância entre este grupo formado e os outros grupos usando o cálculo exemplificado na equação (JOHNSON e WICHERN, 2007):

$$d_{(AB)C} = \frac{\sum_i \sum_j d_{ij}}{N_{(AB)} N_C} \quad (13)$$

Onde:

d_{ij} : É a distância entre o objeto i do grupo AB e o objeto j no grupo C;

N_k : É o número de itens do grupo k .

2.9.7 MÉTODO DE WARD

Outro método de agrupamento hierárquico é o método de Ward, que tem a premissa de minimizar a perda de informação no agrupamento. Usualmente, a perda de informação é dada como soma dos erros quadráticos (ESS – *Error Sum of Squares*). Para um grupo k , calcula-se a soma dos erros quadráticos de cada

item em relação ao centroide do grupo (média), ESS_k . Soma-se então os ESS_k dos grupos, sendo $ESS = ESS_1 + ESS_2 + \dots + ESS_k$. Em cada passo é considerada a junção de todos os possíveis pares de grupos, e é refeita a soma dos ESS_k . É aceito então o agrupamento do par que resultar no menor ESS , e é realizado o próximo passo, substituindo o par de grupos pelo novo grupo formado por ambos. Quando todos os objetos forem finalmente unidos em um grupo único com N itens, o valor de ESS é dado pela equação (JOHNSON e WICHERN, 2007):

$$ESS = \sum_{j=1}^N (x_j - \bar{x})' * (x_j - \bar{x}) \quad (14)$$

Onde:

ESS : É a soma dos erros quadráticos;

x_j : É a medida multivariada associada do j -ésimo item;

\bar{x} : É a média de todos os itens.

A Análise de Agrupamentos proporciona um método de se obter grupos com base em suas similaridades de forma estatística. Da mesma forma que os itens são utilizados para formação de agrupamentos, os mesmo podem ser utilizados para empregar reconhecimento de padrões, alocando novas observações nestes grupos. Em seguida, será apresentada a técnica de Análise Discriminante.

2.10 ANÁLISE DISCRIMINANTE

A técnica multivariada conhecida como Análise Discriminante (AD), trata dos problemas de alocação de novos objetos (ou observações, objetos) em grupos previamente definidos. Esta técnica é de uma natureza preferivelmente exploratória, pois, como um procedimento de separação, é frequentemente empregada em bases de dados com o propósito de investigar as diferenças

observadas quando as relações causais entre as variáveis das observações não são bem compreendidas (JOHNSON e WICHERN, 2007).

Para Hair, Anderson *et al.* (2005), a AD é a técnica apropriada para testar a hipótese de que as médias de um conjunto de variáveis independentes para dois ou mais grupos são iguais. Para Mingoti (2005), o objetivo da AD é encontrar uma combinação linear das variáveis independentes, que minimize a probabilidade de classificação incorreta das observações.

Diz-se então que esta técnica possui dois objetivos de imediato:

- a) Dados ao menos dois grupos, ou mais, determinar quais variáveis diferenciam (discriminam) esses grupos (através da ANOVA/MANOVA);
- b) Determinadas as variáveis que melhor discriminam os grupos, utilizá-las para criar funções discriminantes que serão utilizadas para alocar novos indivíduos, objetos ou observações no grupo mais adequado (a função discriminante otimiza a alocação).

O problema da discriminação entre grupos visando uma posterior classificação foi abordado inicialmente por Fisher (1936). Uma combinação linear das características observadas era obtida de forma a apresentar o maior poder de discriminação entre os grupos (JOHNSON e WICHERN, 2007).

Desta forma, Reis (1997) afirma que as hipóteses básicas do método de AD são:

- a) O grupo de variáveis discriminantes tem distribuição normal multivariada;
- b) As matrizes de covariâncias dos grupos são iguais;
- c) Os grupos diferem quanto as médias.

2.10.1 MÉTODO DISCRIMINANTE DE FISHER

Dados dois grupos distintos de populações, nominadas como π_1 e π_2 , a ideia de Fisher (1936) foi transformar as observações multivariadas \underline{X} nas observações univariadas Y tal que os Y 's nas populações π_1 e π_2 fossem separadas tanto o quanto possível. Assim, através de combinações lineares das observações \underline{X} , são criadas as observações Y 's, de cálculo discriminante mais fácil. Seja então μ_{1Y} a média dos Y ' obtidos pertencentes a população π_1 e μ_{2Y} a média dos Y 's pertencentes a π_2 , Fisher selecionou a combinação linear que maximiza a distância quadrática entre μ_{1Y} e μ_{2Y} , ou seja, a melhor combinação linear vem da razão entre o quadrado da distância entre as médias e a variância de Y , dado por:

$$Y_0 = \underline{C}' \underline{X}_0 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{X}_0 \quad (15)$$

Que é conhecida com função discriminante linear de Fisher (FDL), onde:

Y_0 = Valor da observação univariada, obtido da observação \underline{X}_0 que deseja se discriminar;

\underline{C}' = Diferença das distâncias entre a matriz de médias das populações;

\underline{X}_0 = Matriz de dados da observação que deseja se discriminar;

$\underline{\mu}_1$ = Matriz de médias da população 1;

$\underline{\mu}_2$ = Matriz de médias da população 2;

Σ^{-1} = Matriz inversa da covariância populacional entre os dados \underline{X}_0 e as médias $\underline{\mu}$.

A FDL de Fisher transforma as populações multivariadas em populações univariadas, tais que suas médias são separadas o quanto possível relativo a variância populacional, que é considerada comum. Consideramos então o ponto médio entre duas populações univariadas:

$$m = \frac{1}{2} \underline{C}' (\underline{\mu}_1 + \underline{\mu}_2) = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2) \quad (16)$$

Assim sendo, se X_0 pertence a π_1 , espera-se que $Y_0 \geq m$ e se X_0 pertence a π_2 , espera-se que $Y_0 < m$.

No entanto, percebe-se que este método é indicado para a discriminação de uma observação entre duas populações. Para a discriminação entre diversas populações esse método pode ser estendido, assumindo então que as matrizes de covariâncias populacionais Σ 's sejam iguais.

Sendo então $\underline{\bar{\mu}}$ o vetor médio das diversas populações dado por:

$$\underline{\bar{\mu}} = \frac{1}{g} \sum_{i=1}^g \underline{\mu}_i \quad (17)$$

Onde:

g = número de populações.

E seja B_0 a matriz de variação entre grupos populacionais dada por:

$$B_0 = \sum_{i=1}^g (\underline{\mu}_i - \underline{\bar{\mu}})(\underline{\mu}_i - \underline{\bar{\mu}})' \quad (18)$$

Temos então uma média global dada por:

$$\bar{\mu}_Y = \frac{1}{g} \sum_{i=1}^g \mu_{iY} = \underline{C}' \underline{\bar{\mu}} \quad (19)$$

Como consequência, a razão entre a soma dos quadrados das distâncias das populações para média global de Y , e a variância de Y' é:

$$\frac{\underline{C}' B_0 \underline{C}}{\underline{C}' \underline{\Sigma} \underline{C}} \quad (20)$$

Esta é uma generalização para grupos múltiplos do caso de duas populações, que mede a variabilidade entre os grupos de valores Y em relação à variabilidade entre os grupos. De forma análoga ao problema com duas populações, selecionamos então o \underline{C} que maximiza esta razão.

Ainda, sendo $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s > 0$ Os autovalores não-nulos de $\Sigma^{-1}B_0$, e $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_s$ os autovetores escalonados correspondentes tal que dado que $\underline{e}'\Sigma\underline{e} = 1$, dado que:

$$s \leq \min(g - 1, p) \quad (21)$$

Onde:

g = número de populações;

p = número de variáveis.

Então, o vetor que maximiza a razão (6) é dado por $\underline{C}_1 = \underline{e}_1$.

A combinação $\underline{C}_1'X = \underline{e}_1'X$ é chamada primeira discriminante, a combinação $\underline{C}_2'X = \underline{e}_2'X$ é chamada segunda discriminante e assim obtemos os vários discriminantes sucessivamente (JOHNSON e WICHERN, 2007; HAIR, ANDERSON, *et al.*, 2005).

Para testar a validade das funções discriminantes, entre outras técnicas, é possível classificar as observações originais e comparar os grupos obtidos com os grupos pré-definidos e em seguida estimar a porcentagem de casos classificados corretamente, a partir das variáveis utilizadas (HAIR, ANDERSON, *et al.*, 2005).

As análises referentes a este trabalho, serão então realizadas com base nas técnicas aqui apresentadas e seus resultados serão utilizados na elaboração do aplicativo proposto.

3 MATERIAL E MÉTODOS

Neste capítulo serão apresentadas as etapas definidas para a realização desta pesquisa, evidenciando assim método, abordagem, objetivos e procedimentos técnicos adotados.

3.1 CLASSIFICAÇÃO DA PESQUISA

A natureza desta pesquisa é aplicada, segundo Silva (2004), pois utilizando os conhecimentos gerados pela pesquisa básica, será demonstrado o desenvolvimento de um produto voltado ao consumidor de automóveis, que possui a finalidade de auxiliá-lo na escolha de um automóvel zero quilômetro.

A abordagem do problema é quantitativa em relação aos métodos utilizados em análises estatísticas das variáveis selecionadas dos modelos automóveis. Será qualitativa no que diz respeito a seleção das variáveis obtidas dos automóveis assim como nos métodos de desenvolvimento de produtos e softwares empregados.

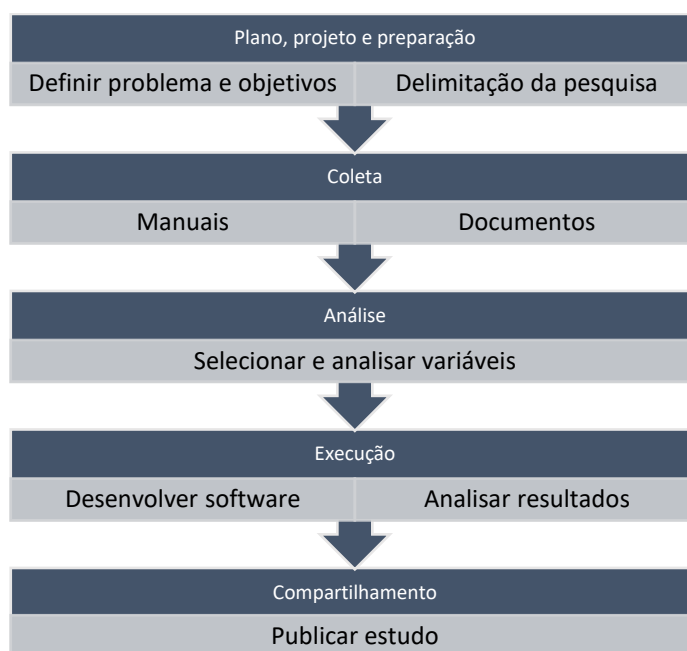
A classificação da pesquisa em relação aos seus objetivos é descritiva, segundo Gil (2002), pois visa a descrição das características dos automóveis, formando agrupamentos.

O método utilizado será a pesquisa documental. Serão analisados dados técnicos públicos a respeito de automóveis, que posteriormente serão utilizados na formação de agrupamentos dos diversos modelos desses automóveis, seguido da seleção e verificação dos grupos que serão utilizados na aplicação.

3.2 PROTOCOLO DE PESQUISA

O projeto de pesquisa será então realizado de acordo com os procedimentos ilustrados na Figura13.

FIGURA13 - PROTOCOLO DE PEQUISA



FONTE: O autor (2016).

Após definidos o problema e os objetivos da pesquisa, o passo seguinte é então realizar a coleta de dados sobre os diversos atributos que um automóvel possui, através de dados disponibilizados publicamente pelas montadoras. Os atributos em questão tratam-se de dimensões físicas e itens de segurança, conforto e luxo, que possam ser quantificados em variáveis numéricas.

Em seguida, serão selecionados os atributos com maior representação estatística quanto as similaridades e dissimilaridades dos veículos, com o intuito de formar grupos de automóveis com características estaticamente similares. Será aplicada uma técnica de reconhecimento de padrões sobre os grupos formados, e os mesmos que serão apresentados posteriormente ao consumidor, visando auxiliar na decisão de compra.

Concluída a análise multivariada, o próximo passo é desenvolver um software onde serão inseridos os grupos de automóveis formados e os resultados da análise de reconhecimento de padrões, com a finalidade de designar estes grupos ao consumidor com base nas respostas de um questionário incluído no aplicativo.

3.3 DELIMITAÇÃO DA PESQUISA

Neste trabalho serão levantados os atributos de automóveis relativos as marcas comercializadas no Brasil, com base em informações fornecidas por montadoras, revistas e sites especializados.

Como o intuito da utilização desses dados é realizar uma análise multivariada, variáveis que possuam dados divergentes, incompletos ou que não possam ser quantificáveis serão excluídas.

Salvo algumas exceções, cada modelo de automóvel possui pelo menos duas variações de motorização e itens inclusos de fábrica, sendo que a maioria possui ao menos quatro variações e em alguns casos chegando ao número de 12. Então, por motivos práticos, selecionou-se para esta pesquisa apenas uma variação de cada modelo, sendo a variação escolhida a considerada mais popular de acordo com a Auto Esporte (2016) e iCarros (2016).

Embora se tenha coletado os dados de todos os veículos encontrados em documentos disponíveis, fogem do escopo desta pesquisa veículos comerciais e vans.

De modo a alcançar os objetivos propostos, será realizada uma análise multivariada com base nestes dados, e então os resultados serão utilizados no desenvolvimento de uma aplicação online.

3.4 PROCEDIMENTOS TÉCNICOS

Os procedimentos aplicados no desenvolvimento deste trabalho podem ser divididos em coleta de dados, análise de dados e desenvolvimento de produto e software.

3.4.1 COLETA DE DADOS

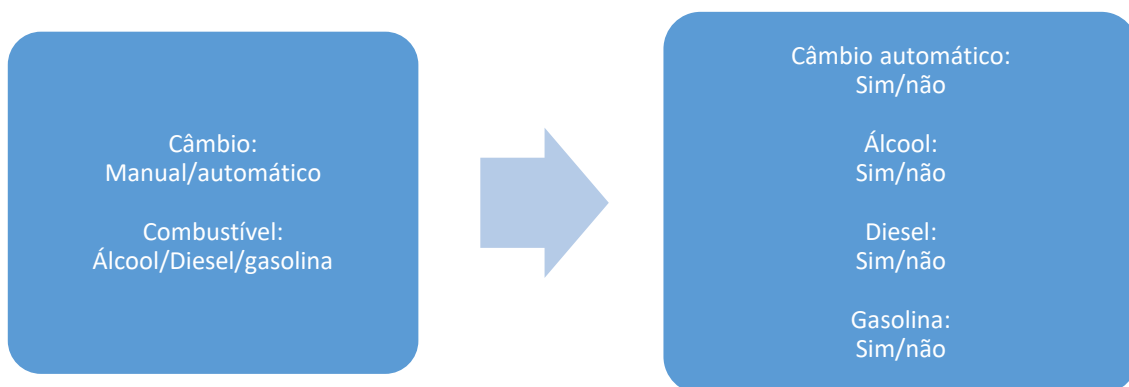
Esta pesquisa tem por base dados dos modelos de automóveis disponíveis no mercado brasileiro. Estes dados são obtidos através de pesquisa documental, obtendo-se informações técnicas disponibilizadas por montadoras de veículos e websites especializados. Foram levadas em consideração as variáveis com dados precisos e quantificáveis com o intuito de evitar viés na análise.

3.4.2 ANÁLISE DE DADOS

Para aplicação da análise multivariada, serão utilizados 40 dos 42 atributos coletados, pois optou-se por não utilizar nesta análise o preço e a motorização. O preço se trata da variável mais relevante e, portanto, será o principal filtro de resultados na ferramenta final. A motorização possui mais de uma opção para a maioria dos modelos de automóveis, no entanto, esta pesquisa utiliza apenas uma variação de cada modelo, o que pode levar a um viés nos agrupamentos se utilizada.

Para possibilitar esta análise, houve a necessidade de dividir algumas das variáveis cuja resposta não é numérica em variáveis que possibilitem obter-se respostas dicotômicas, como ilustrado na Figura 14. Desta forma, a base para análise será de um total de 44 variáveis.

FIGURA 14 - EXEMPLO DE MUDANÇA PARA VARIÁVEIS DICOTÔMICAS



FONTE: O autor (2016).

De tal maneira, os procedimentos metodológicos utilizados na análise são apresentados na Figura 15.

FIGURA 15 - ETAPAS DA ANÁLISE MULTIVARIADA



FONTE: O autor (2016).

Esta análise será realizada utilizando os conceitos expostos nas Seções 2.7 a 2.10 do Capítulo 2.

3.4.3 EXCLUSÃO DE VARIÁVEIS POR MEIO DE ANÁLISE FATORIAL

Tendo variáveis apenas numéricas para se trabalhar, é possível então realizar uma Análise Fatorial (AF), sobre os dados. Esta análise tem por finalidade reduzir o número de variáveis que o usuário do aplicativo terá de escolher ao usar a ferramenta desenvolvida, assim como eliminar variáveis que indiquem a mesma informação, por exemplo, no caso de itens comumente vendidos em conjunto como travas elétricas, vidros elétricos e alarme, utilizar as três variáveis na pesquisa ao invés de apenas uma delas pode provocar um viés no resultado.

Não se deseja com esta análise reduzir as variáveis em fatores comuns, pois estas variáveis terão de ser utilizadas posteriormente pelo usuário da ferramenta que será desenvolvida e para tal, as mesmas devem ser facilmente interpretáveis. Utilizou-se então como critério para reduzir o número de variáveis eliminar da análise as que apresentem comunalidade inferior a 0,6, pois não representam importância no reconhecimento de padrões. Para cada uma ou mais variáveis que são removidas da análise, deve-se repetir a AF com as variáveis restantes até que não haja mais variáveis com baixa comunalidade. Utilizou-se o método das componentes principais, e rotação Varimax.

3.4.4 APLICAÇÃO DE ANÁLISE DE AGRUPAMENTOS

Será realizada AA utilizando os automóveis e variáveis coletados para este estudo, buscando identificar visualmente se a técnica de AA pode ser adequada para formar agrupamentos com a população de veículos em estudo. Para tal, são utilizados os métodos de agrupamento hierárquicos e não-hierárquicos disponíveis no software Statgraphics Centurion. Os métodos hierárquicos agrupam os objetos da população em um único grupo de acordo com suas similaridades, cabendo ao pesquisador identificar e decidir quantos grupos serão formados na composição final. Já os métodos não-hierárquicos

partem de um número de grupos determinado arbitrariamente e vai se realocando as observações para o grupo com centroide mais próximo, repetindo a análise até que não haja mais realocação de objetos nos grupos (JOHNSON e WICHERN, 2007).

Serão utilizados como critérios para a seleção do método de agrupamento, a análise gráfica dos dendrogramas obtidos, a análise dos objetos contidos em cada um dos grupos formados e a precisão do reconhecimento de padrões aplicada no agrupamento.

Quanto a análise do dendrograma obtido, o agrupamento formado será considerado adequado de acordo com os seguintes critérios:

- Grupos obtidos bem definidos visualmente, ou seja, com poucos objetos órfãos ou grupos indistinguíveis entre si;
- Grupos bem divididos, onde seja possível identificar a divisão de agrupamento em pelo menos oito grupos.

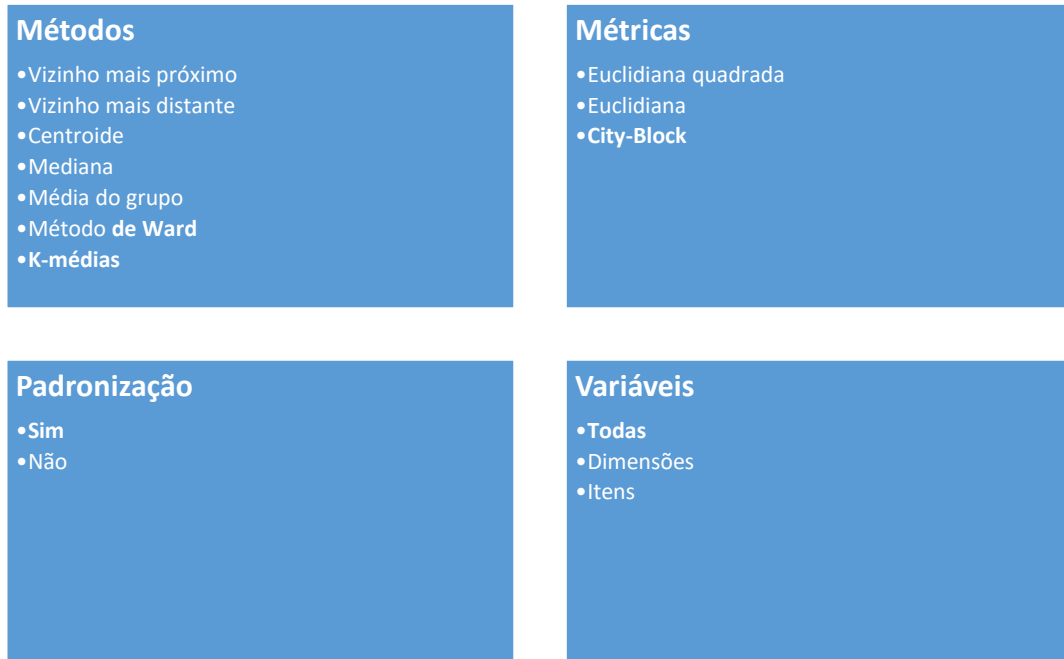
Quanto ao método de k-médias, o qual não é possível gerar um dendrograma da análise, os critérios utilizados serão os de melhor aplicação de reconhecimento de padrão e os objetos contidos nos grupos formados. Os grupos devem ser explicáveis de acordo com as variáveis utilizadas e também de acordo com a classificação já conhecida de automóveis (carroceria e estilo).

Como critério adicional, será testado também a possibilidade de agrupamentos utilizando:

- Apenas as dimensões do veículo (largura, altura, capacidade do porta-malas, etc.);
- Somente os itens que o veículo possui;
- Todas as variáveis simultaneamente.

Os métodos e métricas testados nesta etapa estão descritos na Figura 16, assim como a possibilidade de uso das variáveis.

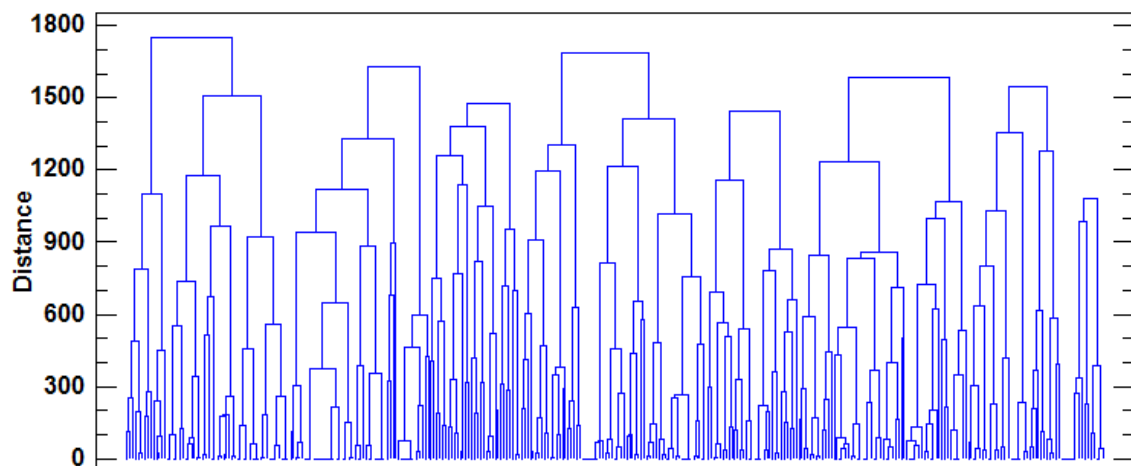
FIGURA 16 - MÉTODOS E MÉTRICAS UTILIZADOS NA ANÁLISE DE AGRUPAMENTOS



FONTE: O autor (2016).

Com isso, serão analisados 126 agrupamentos formados. Na Figura 17 é possível contemplar o dendrograma obtido com o método de Ward, como forma de ilustrar a análise de agrupamentos por métodos hierárquicos.

FIGURA 17 - DENDROGRAMA DE AGRUPAMENTOS (MÉTODO DE WARD)



FONTE: Stagraphics Centurion (2016).

As linhas horizontais representam a distância entre os itens (similaridades) enquanto as linhas verticais representam, na parte inferior os objetos originais da população em análise, e na parte central e superior as ligações (agrupamentos) formados entre estes objetos. Nota-se neste exemplo como os grupos são facilmente identificáveis.

Ao final desta etapa, os grupos de automóveis formados devem ser explicáveis e aptos para a aplicação de reconhecimento de padrões.

3.4.5 APLICAÇÃO DE ANÁLISE DISCRIMINANTE

A última etapa da análise multivariada é a realização da Análise Discriminante, que tem por objetivo nesta pesquisa alocar os novos objetos, ou neste caso, os usuários do aplicativo, nos grupos previamente definidos. A AD trata-se de uma técnica apropriada para testar a hipótese de que as médias de um conjunto de variáveis independentes para dois ou mais grupos são iguais (HAIR, ANDERSON, *et al.*, 2005). Para alocar os usuários a um grupo de automóveis, será empregado o Método Discriminante de Fisher para diversas populações, que consiste em encontrar funções lineares que calculem o centroide uma observação em relação ao centroide dos grupos, tornando possível desta forma classificar novas observações com a menor taxa de erros possível (FISHER, 1936).

Uma das técnicas para se obter o valor da taxa de erros, consiste em utilizar as funções lineares encontradas para tentar classificar os próprios objetos originais da população aos grupos identificados. A porcentagem de acertos na classificação indica se a função é adequada para se aplicar o reconhecimento de padrões a novas observações (JOHNSON e WICHERN, 2007).

Como resultado desta análise, serão obtidos os coeficientes das funções lineares, a tabela com a taxa de acertos no método de classificação e os centroides dos grupos utilizados. Com isto, estes dados podem ser utilizados no desenvolvimento da aplicação proposta nesta pesquisa.

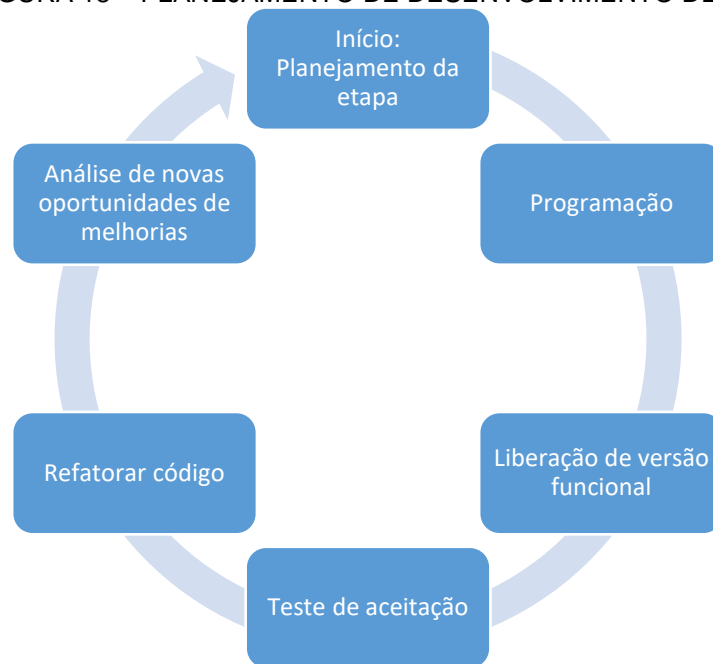
3.4.6 MÉTODO DE DESENVOLVIMENTO DA FERRAMENTA

No que diz respeito ao planejamento de um novo produto, a abordagem de Baxter (2000) atribui os aspectos de identificação de oportunidade, análise de concorrentes, configuração do projeto e especificação do projeto. Em relação a estes dois aspectos iniciais, o teste e estudo das ferramentas online disponíveis proporcionou um entendimento necessário para que seja possível desenvolver a ferramenta que é objeto de estudo deste trabalho, conforme esclarecido na Seção 2.3 do Capítulo 2. Além disso, há a questão da identificação de lacuna, pois não há ferramentas disponíveis que tenham por base os atributos dos automóveis ou que se utilizem de métodos multivariados para classificação de produtos e reconhecimento de padrão.

No que tange a configuração e especificação de projeto, buscou-se complementar o método com a abordagem metodológica de desenvolvimento ágil de software *eXtreme Programming*. Para atender tais requisitos, a ferramenta será desenvolvida com base no planejamento de produto e de software e nos resultados da análise multivariada.

Diante disto e com base nas etapas da metodologia de programação XP, o plano de desenvolvimento da ferramenta tem por base o diagrama exibido na Figura 18.

FIGURA 18 – PLANEJAMENTO DE DESENVOLVIMENTO DE SOFTWARE



FONTE: O autor (2016)

A etapa de planejamento consiste em estimar e priorizar as funcionalidades. As funcionalidades priorizadas estão descritas nos objetivos propostos nesta pesquisa, sendo prioridade fazer com que a ferramenta seja útil para se encontrar um modelo de automóvel zero quilômetro que possa atender as expectativas do usuário. Para tal, o funcionamento da ferramenta desenvolvida deve ser simples e intuitivo.

A etapa de programação, deve ser realizada de maneira simples de modo a se obter uma versão funcional. Para isto a programação será executada na seguinte ordem: layout, base de dados, funções matemáticas e coleta de dados. A programação será realizada com base em uma combinação de linguagens PHP, HTML, MySQL, Java Script e AJAX. O layout será programado com base no *framework* Bootstrap. Cada uma dessas linguagens tem uma funcionalidade específica no funcionamento da aplicação como um todo. Enquanto o HTML serve como base de toda a aplicação, os comandos PHP são utilizados para interagir com o servidor MySQL, transmitindo os comandos de forma a acessar a base de dados. JavaScript e AJAX são utilizados para executar modificações no conteúdo exibido na página sem a necessidade de gerar uma nova página inteira.

Bootstrap é a biblioteca mais popular de HTML, CSS e Java Script, utilizada para desenvolver aplicações com design responsivo, que adapta automaticamente o layout de acordo com o dispositivo utilizado pelo usuário (Bootstrap, 2016). Assim sendo, não há necessidade de programar diversas vezes a interface de usuário para os diferentes tamanhos de tela que irão acessar a ferramenta, pois o próprio *framework* se encarrega de preencher o espaço disponível de forma ajustada a tela.

Concluído o layout funcional, a base de dados original contendo os automóveis e suas variáveis correspondentes é convertida em formato csv, para que possa ser exportada para um banco de dados MySQL. Este banco de dados que permitirá o aplicativo acessar e apresentar os grupos de automóveis como resultados ao usuário. De maneira similar, são inseridos em banco de dados os centroides dos grupos e os coeficientes das funções lineares, obtidos com a Análise Discriminante.

Tendo layout funcional, e base de dados acessível pela interface desenvolvida, é possível inserir os cálculos matemáticos que serão encarregados de apresentar os resultados ao usuário.

Até esta etapa, já se obtém um produto funcional e pronto para ser distribuído, porém, é proposto ainda armazenar os dados que o usuário do aplicativo preencher nos formulários, com a finalidade de se formar um grande banco de dados que pode indicar informações relevantes quanto a preferência desses usuários. Para tal, deve-se criar uma nova base de dados em branco, que será preenchida conforme os acessos recebidos ao aplicativo.

Após concluída e testada as funcionalidades do aplicativo, o mesmo deve ser hospedado em um servidor online, acessível via navegador de internet e distribuído em primeiro lugar a um grupo de testes, verificando se há necessidade de alterações, para em seguida ser divulgado ao público.

Como teste de aceitação do produto, optou-se por inserir um breve formulário após a apresentação de resultados, como maneira de se obter *feedback* diretamente dos usuários a respeito da satisfação com os resultados e simplicidade da utilização da ferramenta.

Neste momento, o código recém-programado é analisado em busca de possíveis ajustes que facilitem as manutenções futuras. Ainda, são corrigidas possíveis falhas e quebras.

Por fim, são analisadas as funcionalidades existentes, modo de funcionamento e operação da ferramenta, oportunidades de incrementar o software com novas funções e outras melhorias.

É esperado com isso que os resultados obtidos da primeira versão do software possam prover as informações necessárias para que sejam adicionadas novas funcionalidades e melhorias no funcionamento da ferramenta.

Como observação final, deve-se elaborar um planejamento a respeito da atualização da base de dados dos automóveis e variáveis.

4 RESULTADOS E DISCUSSÕES

No decorrer deste capítulo serão empregados os procedimentos necessários ao cumprimento dos objetivos delineados nesta pesquisa. Neste sentido, serão discutidos os resultados obtidos e os aspectos técnicos de acordo com o fundamento teórico apresentado ao longo deste trabalho.

Primeiramente, serão descritos os resultados da coleta de dados, seguido da utilização dos métodos de Análise Fatorial e da Análise de Agrupamentos para formar grupos de automóveis. Posteriormente, as técnicas de Análise Discriminante de Fisher são aplicadas para reconhecimento de padrões e discriminação de grupos a usuários da ferramenta. Por fim, é descrito o método de desenvolvimento empregado.

4.1 BASE DE DADOS

Como ponto inicial deste estudo, buscou-se formar uma base de dados com a população de automóveis disponíveis para venda no Brasil com o intuito de realizar uma análise estatística multivariada. Para tal é primordial a consistência desses dados, assim como é necessário que estes sejam quantificáveis de forma numérica. O primeiro passo então foi buscar, em revistas e websites especializados em automóveis, quais marcas de veículos possuem meios de venda no Brasil sem ser a exportação direta pelo consumidor. Na Figura 19 é possível observar as montadoras selecionadas para este estudo.

FIGURA 19 - MONTADORAS CONSIDERADAS NA PESQUISA

Aston Martin	Audi	Bentley	BMW	Chery	Chevrolet	Chrysler
Citroen	Dodge	DS	Ferrari	Fiat	Ford	Geely
Honda	Hyundai	JAC	Jaguar	Jeep	Kia	Lamborghini
Land Rover	Lexus	Lifan	Mahindra	Maserati	Mercedes	MINI
Mitsubishi	Nissan	Peugeot	Porsche	Renault	Rolls-Royce	Smart
SsangYong	Subaru	Suzuki	Toyota	Volkswagen	Volvo	

FONTE: O autor (2016).

Foram identificados 304 modelos de automóveis disponíveis no mercado com exceção de veículos comerciais, vans e utilitários. Buscou-se em seguida as especificações destes veículos, e assim, foram coletadas informações disponibilizadas online por fabricantes e websites especializados de forma a viabilizar o estudo proposto neste trabalho.

Algumas das especificações disponíveis a respeito dos veículos não permitem a realização de uma análise numérica, como é o caso de cores disponíveis de itens ou tipo de suspensão. Também, muitas das especificações não são disponibilizadas por todas as fabricantes, como tamanho do console central, se possui porta-copos, iluminação no porta-malas, entre outras. Como é de suma importância que as variáveis coletadas tenham os valores informados para todos os automóveis do estudo, a base será formada por 42 variáveis cujas informações foram encontradas disponíveis para todos os 304 modelos de automóveis.

Isto posto, as variáveis coletadas e que serão utilizadas nesta análise estão descritas na Figura 20.

FIGURA 20 - VARIÁVEIS COLETADAS

Preço	Motorização	Direção	Câmbio	Marchas	Combustível	Potência
Torque	Velocidade	0-100km/h	Consumo	Altura	Largura	Comprimento
Entre-eixos	Peso	Tanque	Porta-malas	Portas	Ocupantes	Airbag adicional
Alarme	Controle de tração	EBD	Sensor de farol	Ar-condicionado	Piloto automático	Travas elétricas
Regulagem volante	CD player	MP3	Entrada USB	Kit multimídia	Rádio	Ajuste elétrico
Bancos de couro	Desembaçador traseiro	Teto solar	Vidros elétricos dianteiros	Vidros elétricos traseiros	Computador de bordo	Farol de neblina

FONTE: O autor (2016).

4.2 ELIMINAÇÃO DE VARIÁVEIS

Respeitando os critérios adotados realizou-se uma análise multivariada com base nas variáveis coletadas nesta pesquisa, com o intuito de formar grupos de automóveis de acordo com suas similaridades e, posteriormente, utilizar reconhecimento de padrões para indicar um destes grupos a um consumidor interessado em escolher um novo modelo de automóvel. Para todas as etapas utilizou-se o software Statgraphics Centurion.

Iniciou-se a análise multivariada com a realização de uma Análise Fatorial com a finalidade de reduzir o número de variáveis. Para tal, foi utilizado o método de componentes principais, rotação Varimax e critério de Kaiser para extração de fatores. Foram analisados os valores das comunalidades estimadas das 42 variáveis selecionadas. Variáveis cuja comunalidade foi inferior ao valor de 0,6 foram eliminadas da análise, sendo necessário realizar nova rodada de Análise Fatorial cada vez que alguma variável foi eliminada. Os resultados da análise realizada estão expostos no Apêndice 1.

Após quatro rodadas de AF, as comunalidades das variáveis restantes não apresentaram valores abaixo de 0,6. Portanto, foram eliminadas as 15 variáveis listadas no Quadro 6.

QUADRO 6 - VARIÁVEIS ELIMINADAS DA ANÁLISE FATORIAL

Variável	Comunalidade estimada	Rodada da análise
Airbag passageiro	0,457505	1
Distribuição eletrônica de frenagem	0,397221	1
Travas elétricas	0,589262	1
Entrada USB	0,520693	1
Bancos de couro	0,469404	1
Desembaçador traseiro	0,548419	1
Teto solar	0,537737	1
Computador de bordo	0,552794	1
Farol de neblina	0,489336	1
Número de marchas	0,491273	2
Volante com regulagem de altura	0,565488	2
Ajuste elétrico de bancos	0,577975	2
Sensor de farol	0,567384	2
Ar-condicionado	0,535168	3
Piloto automático	0,589751	3

FONTE: O autor (2016).

A exclusão dessas variáveis se deve ao fato de apresentarem baixa contribuição na explicação da população, além de ser um indício de que as variáveis não são linearmente correlacionadas (JOHNSON e WICHERN, 2007).

Embora autores apontem que atributos como ar-condicionado e travas elétricas são levados em consideração pelo consumidor no momento da compra, é possível explicar estes resultados de acordo com o poder discriminatório das variáveis. Em relação a população estudada, que é a de modelo de automóveis, pode-se dizer que se o atributo está disponível em poucos modelos, ou então na maioria dos modelos, este atributo pouco contribui na explicação da população, como é o caso teto solar e do ar-condicionado respectivamente. Há também casos em que o item é oferecido quase sempre em conjunto com outros itens como o vidro-elétrico e a entrada USB. Assim sendo, faz sentido não utilizar estas variáveis na formação de grupos de automóveis e posteriormente na

discriminação destes grupos. No entanto, estas variáveis ainda podem ser incluídas no aplicativo de maneira a personalizar os resultados apresentados ao usuário.

Com as variáveis definidas, o próximo passo é realizar a Análise de Agrupamentos.

4.3 FORMAÇÃO DE GRUPOS DE AUTOMÓVEIS

Nesta etapa da análise buscou-se encontrar o melhor método de agrupamento de automóveis com base nos dados disponíveis e na técnica de Análise de Agrupamentos.

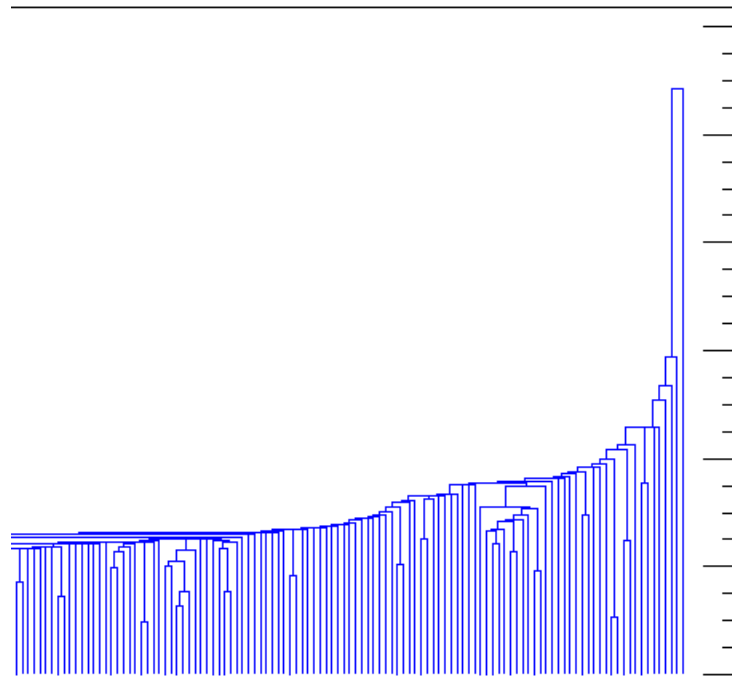
Verificando os critérios definidos na Seção 3.6.2 do Capítulo 3, 126 possíveis agrupamentos devem ser analisados.

4.3.1 CRITÉRIO VISUAL

DE MANEIRA A DIMINUIR O NÚMERO DE AGRUPAMENTOS A SEREM ANALISADOS DE ACORDO COM SEU CONTEÚDO, FOI APLICADO COMO PRIMEIRO CRITÉRIO DE ELIMINAÇÃO DOS MÉTODOS DE AGRUPAMENTO A ANÁLISE GRÁFICA DO DENDROGRAMA. NÃO É DE INTERESSE PARA ESTA APLICAÇÃO UM AGRUPAMENTO DE AUTOMÓVEIS ONDE ALGUNS GRUPOS POSSUEM UM OU DOIS VEÍCULOS E OUTROS GRUPOS POSSUEM MAIS DE CEM. UTILIZANDO O MÉTODO DE AGRUPAMENTO DO VIZINHO MAIS PRÓXIMO, DISTÂNCIA EUCLIDIANA E VARÁVEIS PADRONIZADAS, É POSSÍVEL OBSERVAR NO DENDROGRAMA PARCIAL, APRESENTADO NA

Figura 21, como os automóveis tendem a não se agruparem com nenhum outro modelo, obtendo-se diversos grupos contendo apenas um automóvel, o que inviabiliza o método para esta aplicação.

FIGURA 21 – AGRUPAMENTO VIZINHO MAIS PRÓXIMO



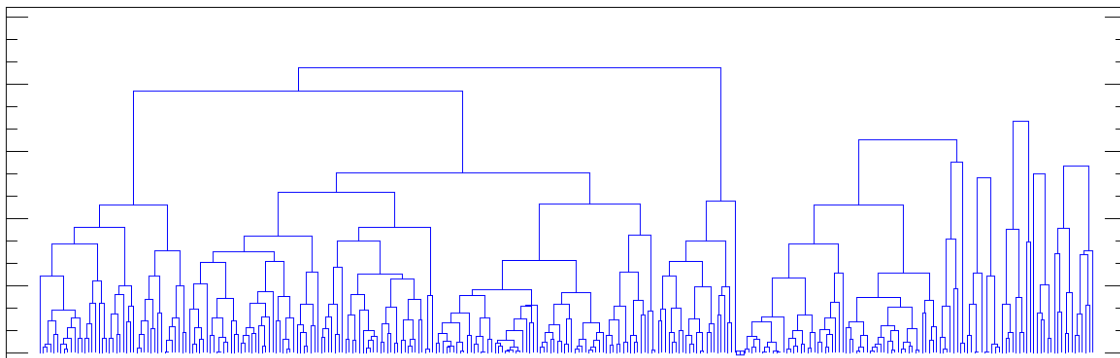
FONTE: Stagraphics Centurion (2016).

Esse comportamento é observado também utilizando como variáveis somente as dimensões ou somente os itens dos veículos. Um comportamento similar é visto utilizando também as métricas de City-Block e euclidiana quadrada para este método, assim como utilizando as variáveis não-padronizadas. O mesmo se repete para os métodos centroide, média do grupo e mediana.

Foi observado que utilizando as métricas euclidiana e euclidiana quadrada, sempre um dos grupos formados continha a grande maioria dos veículos, enquanto os outros grupos são menores. Nota-se este comportamento no dendrograma da

Figura 22, um exemplo da utilização do método do vizinho mais distante, métrica euclidiana quadrada, variáveis padronizadas.

FIGURA 22 – DENDROGRAMA VIZINHO MAIS DISTANTE



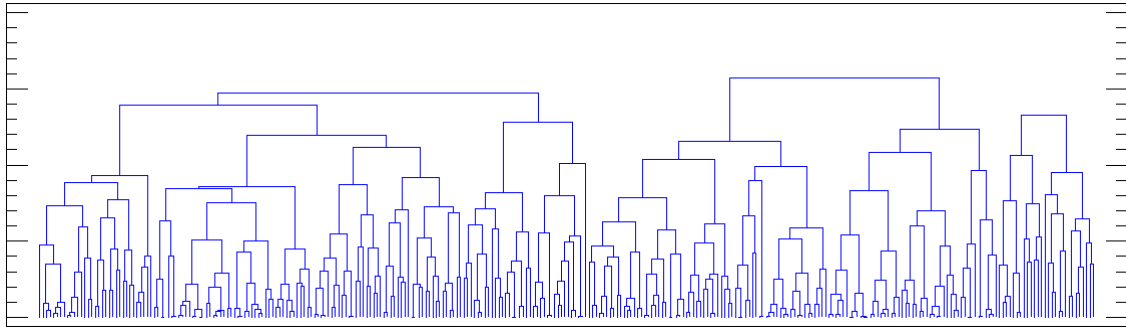
FONTE: Statgraphics Centurion (2016).

Mesmo dividindo o agrupamento em um número maior de grupos, a divisão ocorre de forma mais acentuada nos grupos que já tem um número reduzido de automóveis, não justificando assim um aumento no número de grupos formados. Este comportamento não é desejado, pois no maior grupo formado há uma grande variedade de categorias de veículos, tanto em relação a carroceria, quanto em relação a preço e tamanho, o que pode confundir o usuário e dá indícios de que o agrupamento resultante não é adequado.

COM BASE NO CRITÉRIO VISUAL E CONFORME DEMONSTRADO NA

Figura 23, pode-se dizer que o método que melhor separa os grupos é o método de Ward, métrica de City-Block e variáveis padronizadas. O comportamento se repete quando utilizando somente as dimensões e somente os itens para formar os agrupamentos. O fato do método de Ward separar bem os grupos visualmente, pode ser considerado um indicativo de que o método de k-médias também deve separar bem os grupos, de acordo com a semelhança matemática dos métodos.

FIGURA 23 – DENDROGRAMA MÉTODO DE WARD



FONTE: Stagraphics Centurion (2016).

Desta forma, resta para próxima etapa da análise os métodos:

- Método de Ward;
- K-médias.

A métrica utilizada será a de City-Block.

4.3.2 CRITÉRIO DE OBJETOS CONTIDOS NOS GRUPOS

Como o conteúdo dos grupos formados será analisado nesta etapa, primeiramente deve-se definir o número de grupos que irão compor a solução, mesmo que este número seja alterado posteriormente. É esperado que quanto maior o número de grupos, mais complexo será o reconhecimento de padrões para estes grupos. Por outro lado, quanto menor o número de grupos, mais difícil será para se obter um resultado satisfatório do ponto de vista do usuário, visto que um grupo poderá conter modelos muito distintos, e grupos contendo muitos veículos podem não ser de grande ajuda na escolha de um modelo. Desta forma, definiu-se que o agrupamento será dividido em ao menos oito grupos, que é o

número de categorias de estilo apresentado no catálogo de carros do website iCarros (2016). Este número pode aumentar de acordo com os resultados da Análise Discriminante, desde que seja possível empregar reconhecimento de padrão nos mesmos.

Analisou-se os automóveis contidos nos oito grupos obtidos, para os métodos de AA selecionados. Desta maneira, foi possível perceber que os dois métodos apresentaram grupos coerentes em relação a maneira como os automóveis se agrupam. Porém, todos os métodos apresentaram grupos que contem veículos que não parecem ser semelhantes aos demais em quesitos de desempenho, dimensões, categoria ou sequer itens oferecidos. Optou-se por adicionar a AA a variável preço e comparar os resultados. Foi possível observar que alguns grupos permaneceram os mesmos com ou sem a utilização desta variável, enquanto outros grupos foram modificados, tornando-se mais coerentes com a proposta desta análise. No Quadro 7 é possível observar, como exemplo, a diferença em um dos grupos formados com o método de Ward, com e sem a utilização da variável preço na formação do agrupamento.

QUADRO 7 - GRUPO 1 UTILIZANDO O MÉTODO DE WARD

(continua)

Sem a variável preço	Com a variável preço
Aston Martin Rapide	Aston Martin Rapide
Audi A6	Audi A6
Audi A6 Avant	Audi A6 Avant
Audi A7	Audi A7
Audi A8	Audi A8
Audi RS6 Avant	Audi RS6 Avant
Audi RS7	Audi RS7
Audi S6	Audi S6
Audi S7	Audi S7
Audi S8	Audi S8
Bentley Continental Flying Spur	Audi SQ5
Bentley Continental GT	Bentley Continental Flying Spur
Bentley Mulsanne	Bentley Continental GT
BMW M5	Bentley Mulsanne
BMW Série 6 Gran Coupe	BMW M5

(conclusão)

Sem a variável preço	Com a variável preço
BMW Série 7	BMW Série 5
Chevrolet Camaro	BMW Série 6 Gran Coupe
Jaguar XJ	BMW Série 7
Kia Quoris	Chevrolet Camaro
Land Rover Range	Chrysler 300C
Lexus LS 460	Jaguar XJ
Maserati Quattro	Kia Quoris
Mercedes Benz Classe CLS AMG	Land Rover Range Rover
Mercedes Benz Classe E AMG	Lexus ES 350
Mercedes Benz Classe G AMG	Lexus LS 460
Mercedes Benz Classe GL AMG	Maserati Quattro
Mercedes Benz Classe GLE AMG	Mercedes Benz Classe C AMG
Mercedes Benz Classe ML AMG	Mercedes Benz Classe CLS
Mercedes Benz Classe S	Mercedes Benz Classe CLS AMG
Mercedes Benz Classe S AMG	Mercedes Benz Classe E AMG
Mercedes Benz Classe SL AMG	Mercedes Benz Classe G AMG
Rolls-Royce Ghost	Mercedes Benz Classe GL AMG
Rolls-Royce Phantom	Mercedes Benz Classe GLE AMG
	Mercedes Benz Classe ML AMG
	Mercedes Benz Classe S
	Mercedes Benz Classe S AMG
	Mercedes Benz Classe SL AMG
	Porsche Macan
	Porsche Panamera
	Rolls-Royce Ghost
	Rolls-Royce Phantom
	Subaru Legacy Sedan
	Subaru Outback

Fonte: O autor (2016).

Os grupos formados com a inclusão da variável preço podem ser considerados mais coerentes e permitem um melhor entendimento de como cada grupo foi formado. A explicação mais plausível de como os automóveis se agruparam é apresentada na Figura 24.

FIGURA 24 – GRUPOS FORMADOS



FONTE: O autor (2016).

No entanto ainda existem divergências nos grupos formados que desta vez podem ser explicadas em sua maior parte pelos itens que cada veículo oferece, visto que são umas das principais variáveis deste estudo, e em alguns casos pelas dimensões físicas do veículo.

Até o momento, a análise indica que os dois métodos de AA selecionados são adequados para formar grupos de automóveis de acordo com suas características. Desta forma, o critério utilizado para selecionar o melhor método de AA para esta aplicação foi verificar a adequabilidade do agrupamento formado para a utilização de reconhecimento de padrões. A próxima etapa então é verificar qual dos métodos de agrupamentos selecionados obtém uma melhor taxa de acertos de classificação utilizando o método discriminante de Fisher para diversas populações.

4.3.3 CRITÉRIO DE TAXA DE ACERTOS EM ANÁLISE DISCRIMINANTE

Visto que os dois métodos de AA selecionados apresentam agrupamentos que podem ser utilizados na aplicação, será selecionado para compor a solução o método que apresentar melhor taxa de acertos de classificação na Análise Discriminante. Desta forma, é aplicado o método discriminante de Fisher nos grupos formados e são comparadas as taxas de acerto de classificação da AD. Os resultados são apresentados no Quadro 8.

QUADRO 8 - ACERTOS DE CLASSIFICAÇÃO SEGUNDO O MÉTODO DE FISHER

Método de agrupamento	Utilização da variável preço no agrupamento	Porcentagem de casos classificados corretamente
Ward	Não	73,03%
Ward	Sim	89,14%
K-médias	Não	77,63%
K-médias	Sim	94,08%

FONTE: O autor (2016).

Os resultados de agrupamento mais satisfatórios foram os obtidos com o uso do método de K-médias, métrica de City-Block, utilizando todas as variáveis e com as variáveis padronizadas. Os automóveis que compõe cada grupo segundo este método são descritos no Apêndice 2.

Com os resultados da Análise de Agrupamentos, e a partir da taxa de acertos de classificação, é possível realizar a Análise Discriminante, buscando as Funções Discriminantes Lineares de Fisher que deverão empregar o reconhecimento de padrões.

4.4 DISCRIMINAÇÃO DE GRUPOS DE AUTOMÓVEIS

Para realização do reconhecimento de padrão, o qual deve ser utilizado para discriminar um dos grupos de automóveis a um usuário que busque um veículo com as características daquele grupo, foi realizada uma Análise Discriminante. Para tal, adicionou-se como parâmetro da população de automóveis o grupo a qual o automóvel pertence de acordo com a AA. Foram levados em consideração para a AD apenas os parâmetros selecionados por meio da AF, totalizando 29 variáveis entre dimensões e itens dos automóveis. Como principal resultado da AD, foram obtidas sete funções discriminantes, necessárias para representar 100% da variância total explicada, conforme apresentado na Tabela 1.

TABELA 1 - RESULTADOS DA ANÁLISE DISCRIMINANTE

Função Discriminante	Autovalor	Percentual Relativo	Lambda de Wilks	Qui-quadrado
1	9,834	48,17	0,000470556	2179,7237
2	4,55997	22,34	0,00509801	1501,8486
3	2,7576	13,51	0,0283448	1013,7625
4	1,58239	7,75	0,106508	637,1466
5	0,904453	4,43	0,275046	367,2375
6	0,52275	2,56	0,523812	183,964
7	0,253706	1,24	0,797635	64,3267

Fonte: O autor (2016).

O teste de significância estatística qui-quadrado e o valor do lambda de Wilks mostram como as informações nas sucessivas funções discriminantes são retiradas. O baixo valor do lambda de Wilks na primeira função indica que a série de variáveis que estão sendo usadas possuem um considerável poder de discriminação. Cada função foi gerada segundo sua magnitude e seu autovalor associado, enquanto o valor do lambda de Wilks e do teste qui-quadrado avaliam a quantidade de informação discriminante não incorporada pelas funções

precedentes. Desta forma, ao nível de significância de 0,01 o teste qui-quadrado indicou que nenhuma função discriminante deveria ser eliminada da análise.

A Tabela 2 apresenta os coeficientes das funções discriminantes que irão compor a solução para reconhecimento de padrões e serão integradas no software com o intuito de designar um grupo de automóveis ao usuário de acordo com suas preferências. Utilizando estes coeficientes, é possível calcular os escores discriminantes de cada automóvel, permitindo assim classificá-los em um dos grupos determinados.

TABELA 2 - COEFICIENTES DAS FUNÇÕES DISCRIMINANTES
(continua)

Variável	Função Discriminante						
	1	2	3	4	5	6	7
Direção Elétrica	0,312493	-0,23882	0,591086	-0,56081	-0,77737	-0,24773	1,11387
Direção Hidráulica	0,037728	-0,19944	0,136937	-1,00719	-0,43498	-0,02308	0,187864
Câmbio Automático	1,37197	0,164966	0,661886	1,1523	1,09472	-0,63376	-0,89274
Álcool	-1,27585	0,257087	-0,95595	0,038581	0,592149	1,2812	-0,29103
Diesel	-0,25918	4,92662	4,27443	-1,39046	9,45989	-0,50756	10,9476
Gasolina	0,521656	-0,58725	6,40783	-4,391	10,366	-1,93291	13,6666
Potência(cv)	0,004873	-0,00318	-0,00643	-0,01051	-0,00449	-0,00506	-0,0036
Torque (kgf,m)	-0,01927	0,026716	-0,00069	0,005439	0,049995	0,070563	0,0406
Velocidade Máxima (km/h)	0,006196	-0,00483	0,002854	0,006328	-0,01105	-0,00666	-0,00033
Tempo 0-100 (s)	0,047834	0,027972	-0,01794	-0,05962	-0,02199	-0,03538	-0,03359
Consumo cidade (km/l)	-0,02159	0,031609	-0,02993	-0,03645	-0,04731	0,13563	0,071579
Consumo estrada (km/l)	0,046953	-0,04922	0,062069	0,015615	0,143741	-0,15246	0,153719
Altura (mm)	-0,00033	0,000849	0,002324	-1,7E-05	-0,00317	-0,00454	0,002406
Largura (mm)	0,00254	-0,00085	-0,00428	0,001654	0,00024	0,004026	-0,00256
Comprimento (mm)	-0,00067	0,000242	0,001079	7,87E-05	0,00047	-0,00073	-0,00172
Entre-eixos (mm)	-2,8E-05	0,001849	0,000502	-0,00358	-0,00216	-0,00194	0,001404
Peso (kg)	0,001618	-0,00035	0,000103	0,000168	-0,0008	-0,00053	0,002503

(conclusão)

Variável	Função Discriminante						
	1	2	3	4	5	6	7
Tanque (L)	0,003868	0,035918	-0,03365	-0,01486	0,037466	0,03401	0,010419
Porta-malas (L)	0,000445	0,001441	4,64E-06	0,001463	-0,0003	-0,00168	-0,00029
Portas	-0,18129	0,217089	0,675544	-0,43084	0,35615	0,257016	-0,3008
Ocupantes	-0,00528	0,178727	0,139898	0,131496	0,284227	0,466342	0,124969
Airbag lateral	1,58269	-0,01489	0,387756	0,775532	0,565659	0,240167	0,074123
Controle de tração	1,02436	0,179096	0,419494	0,206539	0,367178	0,371811	-0,99621
Cd player	0,057333	0,119453	0,119843	-0,10618	-0,4057	0,309707	0,87531
Cd player com MP3	0,536768	-0,20891	1,03788	0,255253	-1,58434	0,635257	-0,74777
Kit Multimídia (GPS e/ou DVD)	1,08165	0,081753	0,144498	0,425482	0,116654	0,061617	0,343278
Rádio FM/AM	-0,02666	-0,06683	-0,29944	-0,12757	-0,33099	0,262032	-0,32134
Vidros elétricos dianteiros	0,381085	-0,52798	-0,32137	0,532788	0,028815	0,275264	0,510757
Vidros elétricos traseiros	-0,27372	0,359639	1,11658	-1,08275	0,131738	-0,05961	-0,30969
CONSTANTE	-9,24498	-8,60089	-11,2994	12,9604	-4,34621	6,10313	-14,5994

FONTE: O autor (2016)

Na Tabela 3 encontram-se os centroides ou pontos médios de cada grupo, nas sete funções discriminantes.

TABELA 3 - CENTROIDES DOS GRUPOS NAS FUNÇÕES

Grupo	Coeficientes das funções						
	Função 1	Função 2	Função 3	Função 4	Função 5	Função 6	Função 7
Grupo 1	3,19437	0,838708	0,221144	-2,06287	0,276598	-0,17334	0,600253
Grupo 2	4,10329	-1,93022	-4,24728	-0,58777	-0,38789	0,592775	-0,87153
Grupo 3	-3,19893	5,01267	-2,04246	0,189788	-1,51664	-1,86359	-0,25338
Grupo 4	-3,23283	-0,84611	0,91109	-0,5334	-1,32385	0,710357	0,043973
Grupo 5	-4,84533	-0,67532	-1,09849	-0,11039	1,59081	-0,03687	0,071806
Grupo 6	2,01204	-2,99003	-1,46266	2,83633	-0,66846	-0,44944	1,18049
Grupo 7	1,5137	-0,21519	1,52994	0,630858	0,274536	-0,25134	-0,3776
Grupo 8	2,1182	8,19757	-0,81536	2,54149	0,689135	2,44266	0,341646

FONTE: O autor (2016)

É possível observar na tabela de classificação, demonstrada na Tabela 4, que as funções discriminantes obtidas utilizando o método de Fisher apresentaram uma taxa de acerto de classificação de 94,08%, índice considerado bastante satisfatório para a aplicação desejada e indicando também a precisão na técnica de agrupamento escolhida. Portanto, pode se considerar que as sete funções podem ser usadas para classificar novas observações dentro dos grupos.

TABELA 4 - TABELA DE CLASSIFICAÇÃO

Grupo real	Tamanho do grupo	Observações previstas por grupo							
		1	2	3	4	5	6	7	8
1	46	43	0	0	0	0	0	3	0
2	21	0	21	0	0	0	0	0	0
3	16	0	0	15	1	0	0	0	0
4	53	0	0	0	50	0	0	3	0
5	44	0	0	0	1	42	0	1	0
6	19	0	0	0	0	0	19	0	0
7	96	1	0	1	2	0	4	87	1
8	9	0	0	0	0	0	0	0	9

FONTE: O autor (2016)

De acordo com os critérios estabelecidos, o método de agrupamento, o número de grupos selecionados e o método de reconhecimento de padrão foram considerados adequados para empregar o aplicativo de forma a auxiliar o consumidor de automóveis a escolher um novo modelo. Contudo, nada impede que em implementações futuras do aplicativo, uma nova análise seja realizada, utilizando novas variáveis coletadas ou testando outros critérios de seleção de variáveis, agrupamentos e números de grupos. O roteiro utilizado nesta análise, pode servir para tal base.

Com base nos resultados obtidos da análise multivariada, inicia-se a programação do software conforme proposto.

4.5 DESENVOLVIMENTO DA APLICAÇÃO

Com o intuito de atingir os objetivos propostos no escopo e de acordo com as limitações definidas neste estudo, iniciou-se o desenvolvimento da ferramenta com base na análise multivariada realizada. Os resultados aqui apresentados dizem respeito ao primeiro ciclo de desenvolvimento, do planejamento inicial até a análise de oportunidades de melhorias sobre a versão desenvolvida.

4.5.1 PLANEJAMENTO

Esta etapa consiste em especificar tanto as ferramentas necessárias para garantir o sucesso no desenvolvimento do aplicativo, quanto o próprio funcionamento do aplicativo. Partindo do pressuposto que mudanças de especificação poderão ocorrer a qualquer momento durante o processo de desenvolvimento, buscou-se em primeiro lugar, garantir que ao final do ciclo o resultado obtido fosse um produto funcional em nível de usabilidade. Isto quer dizer que ao final do desenvolvimento detalhado neste estudo, o produto deve estar online, com todas as funções implementadas em funcionamento. Ainda, esta etapa deve assegurar que o processo de desenvolvimento atenda aos valores detalhados na seção 2.6.2 do Capítulo 2.

Como definido na seção 3.4.6 do Capítulo 3, o aplicativo será desenvolvido utilizando linguagem web. Para tal, deve-se em primeiro lugar configurar um ambiente para desenvolvimento onde se possa manipular banco de dados e testar códigos que irão rodar apenas no servidor que hospedará o aplicativo. Será utilizado desta forma o ambiente de desenvolvimento Wampserver, que permite criar aplicações utilizando Apache2, PHP e MySQL, além de oferecer ferramentas para gerenciar estas aplicações.

Uma vantagem de programação web é que o código gerado não requer uma ferramenta específica para compilar, podendo ser escrito, por exemplo,

diretamente no bloco de notas e testado em qualquer navegador de internet. A ferramenta utilizada para escrever o código neste estudo é o Notepad++, que nada mais é que um editor de texto com funcionalidades para facilitar a visualização e formatação de código em diversas linguagens.

Isto posto, na parte do planejamento que diz respeito ao desenvolvimento do aplicativo em si, buscou-se especificar as etapas de programação, e o funcionamento básico do aplicativo. A programação deve atender as etapas:

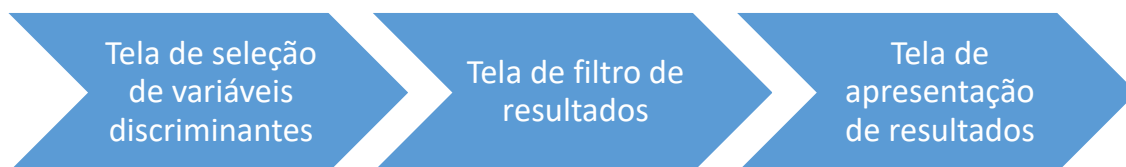
- Programação e teste de layout;
- Programação das bases de dados;
- Programação da lógica do aplicativo;
- Testes de funcionamento.

A ordem em que as tarefas de programação devem ser realizadas dependem apenas da preferência do programador. No entanto, caso existam mais de um programador na equipe, as tarefas podem ser realizadas simultaneamente. Neste trabalho optou-se por programar o código na ordem apresentada, tendo em vista que com o layout e a base de dados previamente preparados, é possível programar e testar a lógica de funcionamento de forma simultânea. O funcionamento consistirá no seguinte pressuposto:

- O usuário preenche um formulário com itens e dimensões que deseja em um veículo zero quilômetro;
- Em seguida, seleciona entre os filtros disponíveis para que os resultados sejam mais consistentes;
- O aplicativo então discrimina o usuário a um dos grupos de automóveis pré-carregados no servidor, utilizando a FDL de Fisher encontrada na análise multivariada;
- O grupo de automóveis é exibido ao usuário, excluindo os automóveis que não se enquadram no filtro aplicado.
- As opções selecionadas pelo usuário são armazenadas em uma base de dados.

De forma sucinta, o fluxograma de funcionamento do aplicativo é apresentado na Figura 25.

FIGURA 25 – FLUXOGRAMA DE FUNCIONAMENTO DO APLICATIVO



FONTE: O autor (2016).

Nota-se a simplicidade de funcionamento do aplicativo, com poucas interações do usuário e poucas telas a serem apresentadas. No entanto, a realização do planejamento e sua devida documentação são importantes para que futuras atualizações sejam implementadas de forma consistente com o estilo e as premissas do software aqui desenvolvido. Há também a importância de que sejam informados de forma clara os objetivos, premissas e descobertas relacionadas ao desenvolvimento do aplicativo. Assim, caso outro desenvolvedor necessite adicionar funcionalidades ou refatorar o código, o mesmo deve ser capaz de fazê-lo com base na documentação gerada.

O intuito deste planejamento, é que o aplicativo possa ser entregue com as funcionalidades mais básicas bem definidas e com o código funcionando adequadamente. Detalhamentos mais específicos em relação a propriedades ou particularidades do funcionamento do aplicativo serão comentadas a medida em que a etapa em questão seja apresentada.

4.5.2 PROGRAMAÇÃO

Conforme a abordagem proposta, a primeira etapa foi programar o layout do software. Levou-se em consideração para tal a simplicidade de uso, prezando principalmente a funcionalidade e uniformidade de estilo. Tendo em vista que o

aplicativo deve funcionar adequadamente em diferentes plataformas e tamanhos de tela, tem-se duas opções para tal: programar uma versão de aplicativo para cada plataforma ou programar uma única versão universal com o auxílio de alguma biblioteca. Embora uma versão específica para cada plataforma ofereça uma experiência mais consistente para o usuário daquela plataforma, o desenvolvimento de um aplicativo universal pode ser realizado de forma mais rápida e não impede que o código seja portado para diferentes linguagens de programação caso seja conveniente. Por estes motivos, o aplicativo foi escrito utilizando principalmente linguagem PHP/HTML e a biblioteca encarregada de manter a uniformidade de estilos no layout é a Bootstrap.

Como a biblioteca Bootstrap possibilita a padronização de estilos das páginas HTML, foi necessário apenas se preocupar com a funcionalidade do layout. Três telas foram programadas inicialmente, sendo uma tela onde são inseridos os itens que o usuário deseja em um automóvel, uma tela para filtrar os resultados e uma tela onde os resultados serão exibidos. Os layouts desenvolvidos, assim como todas as etapas seguintes, foram testados em uma tela de notebook de 14 polegadas e em uma tela de smartphone de 4,7 polegadas. Conforme demonstrado na Figura 26 e na Figura 27 respectivamente, o conteúdo se adaptou ao tamanho de tela como esperado, e no mais não foram detectadas anomalias ou quebras de funcionamento em relação ao esperado, portanto iniciou-se a etapa seguinte.

FIGURA 26 - TELA INICIAL DO APLICATIVO EXIBIDA EM UMA TELA DE COMPUTADOR

MultiCar

Selecione os atributos desejados em um carro zero quilômetro

Tipo de combustível
 Gasolina Álcool Flex Diesel

Tipo de direção
 Hidráulica Elétrica

Tipo de câmbio
 Automático Manual

Número de portas
 2 3 4 5

Número de ocupantes
 2 3 4 5 6 7

FONTE: O autor (2017).

FIGURA 27 - TELA INICIAL DO APLICATIVO EXIBIDA EM UMA TELA DE SMARTPHONE

MultiCar

Selecione os atributos desejados em um carro zero quilômetro

Tipo de combustível
 Gasolina Álcool Flex Diesel

Tipo de direção
 Hidráulica Elétrica

Tipo de câmbio
 Automático Manual

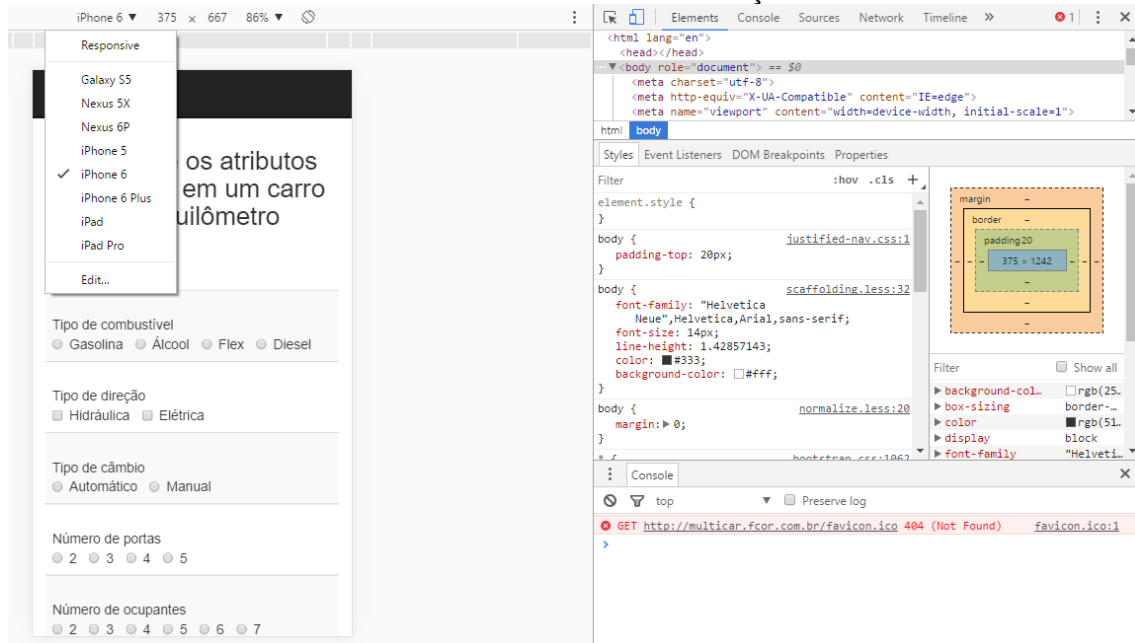
Número de portas
 2 3 4 5

Número de ocupantes
 2 3 4 5 6 7

FONTE: O autor (2017)

Pequenos testes foram realizados com o layout desenvolvido, visando verificar o comportamento do conteúdo das telas quando executadas em diferentes dispositivos. Para tal, foi usado a ferramenta de depuração do navegador de internet, conforme ilustrado na Figura 28.

FIGURA 28 – FERRAMENTA DE DEPURAÇÃO DO NAVEGADOR



FONTE: O autor (2017).

É possível observar que a ferramenta oferece tamanhos de layout pré-definidos de acordo com tamanhos de telas de aparelhos populares, além de análise de código e indicação de erros. No exemplo demonstrado, foi acusado um erro de ícone faltando, o que pode ser resolvido especificando o endereço correto onde se encontrava o arquivo. Desta forma percorreu todo o desenvolvimento, realizando pequenos testes a cada funcionalidade adicionada. Esta abordagem possibilita uma rápida identificação de falhas e eventuais necessidades de revisão de planejamento. No entanto, o layout pode ser finalizado sem maiores complicações.

Concluído o layout, passou-se para a etapa de criação dos bancos de dados. Serão necessárias ao longo do desenvolvimento, a utilização de cinco base de dados diferentes, sendo que três delas só serão atualizadas caso ocorra alterações na análise multivariada ou na base de automóveis. As outras duas servirão para armazenar dados gerados durante a utilização do aplicativo. Desta forma, foram gerados dois arquivos no formato csv contendo os resultados da Análise Discriminante, sendo um arquivo com os coeficientes das funções discriminantes e outro arquivo com os centroides dos grupos nas funções discriminantes. Um terceiro arquivo foi gerado com a base de dados de veículos

completa utilizada na análise multivariada e em seguida cada um destes arquivos foi importado para uma base online distinta com acesso em MySQL. Programou-se em linguagem PHP algumas funções de escopo, com a finalidade de realizar a interface de acesso entre a base de dados e o layout que é apresentado ao usuário do software. Os testes mostraram que o acesso a base de dados não apresenta erros e os dados requisitados são exibidos corretamente na tela.

Importante ressaltar que os resultados da Análise Discriminante poderiam ser inseridos diretamente no código do software, sem a necessidade de permanecer em uma base a parte que requer acesso cada vez que é utilizada. No entanto, esta prática não é recomendada, visto que quando necessário alguma alteração no algoritmo de reconhecimento de padrão, fica muito difícil controlar as alterações, deixando margens para erros de digitação ou atualização incorreta de variáveis. Da maneira que este problema foi tratado, assim que necessário atualizar a análise, basta substituir o arquivo na base de dados sem a necessidade de inserir novo código na programação.

Os cálculos necessários são então inseridos no código, com o uso de linguagem Javascript, que possibilita obter os coeficientes, os centroides e os dados preenchidos pelo usuário para em seguida implementar esta funcionalidade sem maiores complicações. Os testes desta etapa foram realizados utilizando dados pré-definidos e de respostas conhecidas, com a finalidade de validar o comportamento do software. Mais uma vez os resultados foram exibidos conforme esperado.

Como última etapa deste ciclo de programação, resta apenas criar uma base de dados onde serão armazenadas as respostas inseridas pelos usuários da ferramenta e um formulário de pesquisa a respeito da usabilidade do produto. Para tal, foi programado em MySQL os comandos necessários para gerar uma base de dados em branco. Como não há informações pessoais ou sensíveis como por exemplo senhas, o método de armazenamento desta informação é por meio de HTML simples, o qual é tratado com o uso de PHP. Apesar de o método permitir que as variáveis sejam manipuladas ou interceptadas por alguém mal-intencionado, isto não é uma preocupação visto que o dano potencial é mínimo. Testes com um pequeno grupo de controle demonstraram que o software se mostrou funcional e já preparado para ser posto online. Foram adicionadas três

perguntas em relação a usabilidade do software logo abaixo dos resultados exibidos. O armazenamento das respostas a essas perguntas foi tratado de forma semelhante à descrita previamente.

4.5.3 LIBERAÇÃO DE VERSÃO E TESTES

Até este ponto, todo código gerado foi testado no ambiente Wampserver, em uma máquina offline, o próximo passo então deve ser colocar os arquivos necessários no servidor dedicado online, de maneira a possibilitar o acesso ao aplicativo por meio de qualquer dispositivo.

Desta forma, todos os arquivos gerados da programação e de dados foram colocados em um servidor online utilizando as ferramentas disponibilizadas para tal pelo serviço contratado. Visto que esta ferramenta irá se apresentar em forma de um desenvolvimento contínuo, onde funcionalidades continuarão a ser implementadas após o lançamento, optou-se por utilizar um domínio genérico e o subdomínio MultiCar, proveniente da junção das palavras multivariada e carros.

A versão desenvolvida até este ponto foi disponibilizada à um grupo de testes, com o intuito de verificar falhas de funcionamento e uma opinião preliminar a respeito da ferramenta desenvolvida. Por se tratar de uma versão preliminar de um aplicativo com poucas telas e pouco código, não foi gerada nenhuma documentação quanto ao feedback do uso, no entanto, foi discutido com o grupo de testes e os seguintes pontos foram observados:

- Não houve falhas de execução no aplicativo;
- Variáveis como torque e potência são complexas para se requisitar ao usuário a escolha de um valor desejado;
- Falta de itens gráficos, como fotos;
- Por vezes não é exibido nenhum item na tela de resposta.

Quando o aplicativo não exibe os resultados, em seu lugar é exibido uma mensagem de que alguma variável de filtro eliminou todos os veículos do grupo discriminado ao usuário. A variável com maior poder para tal é o preço, pois existem, por exemplo, grupos contendo apenas automóveis com valores acima de 100.000 reais.

O aplicativo funcionou conforme esperado e com isto, pode-se dizer que foi alcançado o sucesso no desenvolvimento da ferramenta proposta. Mesmo que tenham sido identificadas falhas pontuais quanto ao layout conceito do funcionamento, estas falhas podem ser corrigidas ou trabalhadas em futuras atualizações do aplicativo. A versão desenvolvida neste trabalho consiste em duas telas que requerem preenchimento de dados pelo usuário, respectivamente demonstradas na Figura 29 e na Figura 30, e uma tela exibindo o grupo de carros resultante demonstrada na Figura 31.

FIGURA 29 - TELA DE SELEÇÃO DE ATRIBUTOS

MultiCar

Número de portas
 2 3 4 5

Número de ocupantes
 2 3 4 5 6 7

Airbag lateral
 Sim Não

Controle de tração
 Sim Não

Vidros elétricos
 Dianteiros Traseiros

Player Multimídia
 Com CD Com MP3
 Com rádio AM/FM Com GPS ou DVD

Próximo

FONTE: O autor (2017)

FIGURA 30 - TELA DE FILTRAGEM DE RESULTADOS

MultiCar

Faixa de Preço (R\$)

Min. 79000

Max.

Potência (cv)

Min.

Max.

Consumo km/L

Min. cidade

Min. estrada

Dimensões mínimas (mm)

Altura

Largura

Comprimento

Tempo 0-100 km/h (s)

Min

Ver Carros

The image shows a mobile application interface for filtering cars. At the top, there is a dark header with the text 'MultiCar'. Below this, the interface is divided into several sections, each with a title and a slider control. The first section is 'Faixa de Preço (R\$)' with a 'Min.' slider set to 79000 and an empty 'Max.' slider. The second section is 'Potência (cv)' with 'Min.' and 'Max.' sliders. The third section is 'Consumo km/L' with 'Min. cidade' and 'Min. estrada' sliders. The fourth section is 'Dimensões mínimas (mm)' with 'Altura', 'Largura', and 'Comprimento' sliders. The fifth section is 'Tempo 0-100 km/h (s)' with a 'Min' slider. At the bottom of the filtering area is a blue button labeled 'Ver Carros'.

FONTE: O autor (2017)

FIGURA 31 - TELA DE APRESENTAÇÃO DE RESULTADOS



FONTE: O autor (2017).

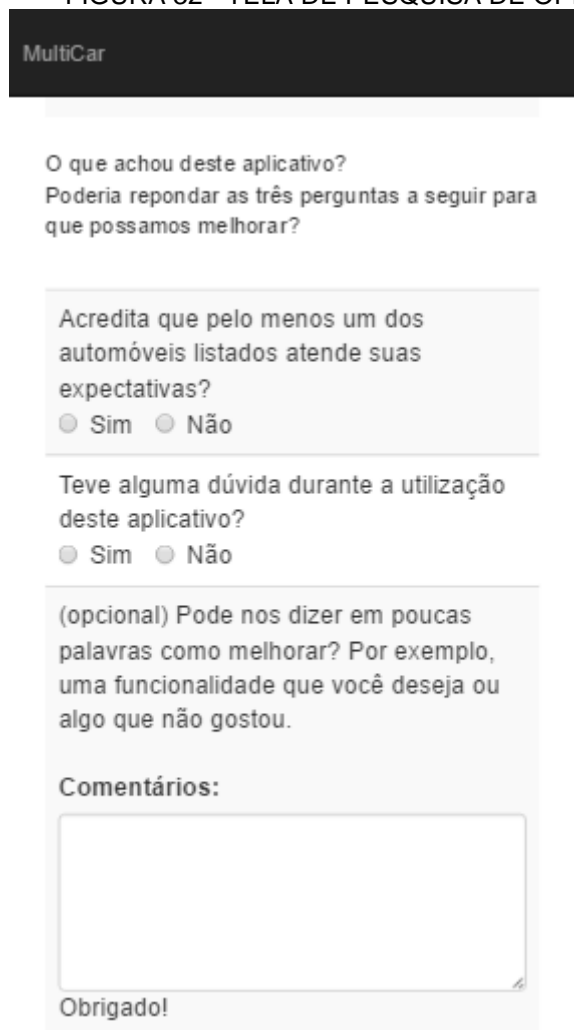
Em seguida serão discutidos os pontos de melhoria identificados e o planejamento para atualizações futuras.

4.5.4 REFATORAÇÃO DE CÓDIGO E ANÁLISE DE MELHORIAS

Embora o software esteja funcional e de acordo com os objetivos propostos, algumas preocupações carecem de planejamento e podem ser listadas em relação às futuras versões desenvolvidas.

Como primeira providência, foi idealizada a adição de uma breve pesquisa de opinião ao final da tela de resultados, conforme ilustrado na Figura 32..

FIGURA 32 - TELA DE PESQUISA DE OPINIÃO



The image shows a mobile application screen for an opinion survey. At the top, there is a black header with the text "MultiCar" in white. Below the header, the text reads: "O que achou deste aplicativo? Poderia reponder as três perguntas a seguir para que possamos melhorar?". The survey consists of three sections: 1. A question: "Acredita que pelo menos um dos automóveis listados atende suas expectativas?" with radio buttons for "Sim" and "Não". 2. A question: "Teve alguma dúvida durante a utilização deste aplicativo?" with radio buttons for "Sim" and "Não". 3. An optional question: "(opcional) Pode nos dizer em poucas palavras como melhorar? Por exemplo, uma funcionalidade que você deseja ou algo que não gostou." followed by a text input field labeled "Comentários:". At the bottom of the form, it says "Obrigado!".

FONTE: O autor (2017)

A cada atualização de versão do aplicativo, as perguntas que compõem esta pesquisa podem ser alteradas de maneira a alcançar um feedback de alguma função adicionada em específico ou possíveis falhas. Espera-se que este questionário seja fixo, alterando periodicamente as perguntas que o compõem e que não seja ultrapassado o número de três perguntas. Caso os resultados desta pesquisa sejam inconclusivos, uma pesquisa mais abrangente deve ser realizada por outros meios.

Em relação aos pontos de melhoria identificados na etapa de testes, foram enumeradas possíveis soluções que devem ser estudadas para implementação na próxima versão do aplicativo:

- Variáveis complexas: Entre as soluções disponíveis, pode ser colocado uma versão do aplicativo onde se substituam as variáveis utilizadas na AD, pelos fatores obtidos na AF. Desta forma, variáveis como altura largura e comprimento, seriam reduzidas a um fator chamado de dimensões externas e assim por diante. Outra opção seria a de se buscar novas variáveis ou alterar os critérios de formação e seleção de grupos na análise multivariada.
- Falta de elementos visuais: Na versão desenvolvida foi priorizada apenas a consistência do layout, por isto e por questões de direitos de imagem, optou-se por não se utilizar elementos visuais até que se tenha uma base consistente para tal e uma identidade visual definida.
- Necessidade de telas de ajuda e apresentação da ferramenta: Entre as prioridades da próxima versão está a elaboração de uma apresentação da ferramenta que seja inclusa antes do usuário iniciar sua utilização e também algum botão de ajuda que esteja disponível a qualquer momento.
- Divulgação do aplicativo: atingidas as metas de planejamento citadas anteriormente, a divulgação do aplicativo pode ser realizada através de plataformas amplamente utilizadas, como a do Facebook, Android e iOS.

Estas questões devem ser resolvidas na próxima versão do aplicativo. Em seguida serão apresentadas as preocupações que devem ser levadas em consideração a cada atualização da ferramenta:

- Atualização da base de dados dos automóveis: com o lançamento de novas versões e novos modelos, a base deve ser atualizada. Com o atual calendário de lançamentos de veículos no mercado brasileiro, é possível que esta atualização ocorra mensalmente, visto que esta é a periodicidade de revistas especializadas (Revista Auto Esporte, 2016).

- Novas versões: Com base nos princípios do método de programação XP, o código gerado é de fácil manutenção sendo possível implementar novas funcionalidades ou correções assim que as mesmas sejam especificadas.
- Expansão das variáveis utilizadas: este estudo foi realizado com uma grande variedade de automóveis, no entanto, na medida em que o número de usuários do aplicativo aumente, é possível analisar os dados obtidos com o intuito de remover alguns tipos de automóveis da base de dados, permitindo desta forma a inclusão de outras variáveis disponibilizadas apenas por algumas montadoras.

5 CONSIDERAÇÕES FINAIS

Este trabalho abordou uma metodologia de desenvolvimento ágil para se programar uma ferramenta computacional que tem o intuito de auxiliar o consumidor na escolha de um automóvel zero quilômetro. Para tal, foram utilizados métodos de agrupamento, seleção de variáveis e reconhecimento de padrões provenientes da estatística multivariada.

Pode-se observar o uso de métodos estatísticos na classificação de um produto (automóveis). Embora os automóveis possuam inúmeras características que podem ser relevantes ao consumidor, foi possível agrupá-los com base apenas em seus atributos e dimensões. Os grupos apresentam coerência de acordo com a proposta da ferramenta desenvolvida, de que ao invés de utilizar a classificação por tipo de carroceria ou preço, leva-se em consideração apenas o que o automóvel oferece, em relação as variáveis coletadas no estudo.

Neste trabalho foi visto que os automóveis podem ser agrupados de acordo com suas dimensões e itens de conforto, segurança e luxo oferecidos.

Quanto a base de dados que se pretende formar com informações de como a ferramenta é usada, espera-se que seja possível fornecer elucidações a respeito da importância de dispor-se de informações sobre preferências de seu mercado consumidor. A própria premissa de funcionamento da ferramenta serve também à este propósito de coletar dados.

No entanto, identificou-se durante os testes que, embora a ferramenta cumpra com os objetivos propostos neste estudo, há problemas que necessitam ser resolvidos de maneira a oferecer uma experiência mais concisa ao usuário. Além dos pontos discutidos na seção 4.5.4 do Capítulo 4, pode ser levada em consideração a aplicação de uma pesquisa de mercado, com o intuito de se identificar se há categorias de automóveis que podem ser removidas da população de estudo, de maneira a apresentar resultados mais satisfatórios ao usuário da ferramenta. A adição de uma breve pesquisa ao final da utilização da ferramenta, como foi realizado, também pode contribuir para que o aplicativo seja melhorado de forma contínua. Usando o próprio espaço da ferramenta para se identificar problemas pontuais e falhas de código é possível obter feedback de

forma rápida, agilizando também as correções necessárias, além de ser um bom ponto de partida para pesquisas mais elaboradas.

Para trabalhos futuros com base nesta pesquisa, é sugerido expandir o uso de técnicas de estatística multivariada na análise de agrupamentos para outros tipos de produtos. Um exemplo de produto que permite aplicar esta mesma abordagem de pesquisa é o aparelho celular, um produto que possui grande diversidade de modelos e especificações técnicas bem detalhadas. Desta forma, pode-se desenvolver uma nova ferramenta utilizando os métodos aqui propostos, mas voltado à produtos diversos. Há também a possibilidade de se aplicar outros métodos, estatísticos ou não, para análise de uma população de produtos com o intuito de auxiliar o consumidor na decisão de compra.

Por fim, sugere-se pesquisas no sentido de demonstrar a eficácia de ferramentas que possuem além do propósito de atender a um mercado em específico, obter informações deste mercado. Como um bom exemplo desse tipo de filosofia de produto, pode-se citar a ferramenta de busca do Google, onde os termos inseridos em uma pesquisa são armazenados e utilizados para identificar quais anúncios são mais relevantes para se exibir àquele usuário.

REFERÊNCIAS

- ABEPRO. **Associação Brasileira de Engenharia de Produção**, 2016. Disponível em: <<http://www.abepro.org.br/>>. Acesso em: 10 mar. 2016.
- AGILE Manifesto. **Agile Manifesto**, 2001. Disponível em: <<http://agilemanifesto.org>>. Acesso em: 25 Novembro 2011.
- AMA. **American Marketing Association**, 2015. Disponível em: <<https://www.ama.org/>>. Acesso em: 15 novembro 2015.
- ANFAVEA. **Associação Nacional dos Fabricantes de Veículos Automotores**, 2016. Disponível em: <<http://www.anfavea.com.br/>>. Acesso em: 20 dezembro 2015.
- ARAÚJO, C. A. Bibliometria: evolução história e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11-32, janeiro junho 2006. ISSN 1808-5245.
- BAXTER, M. R. **Projeto de produto**: guia prático para o design de novos produtos. 2ª. ed. São Paulo: Bluscher, 2000.
- BECK, K. **Programação Extrema Explicada**. Porto Alegre: Bookman, 1999.
- BOEHM, B. W. **Software Engineering Economics**. New Jersey: Prentice Hall, 1981.
- BOOTSTRAP , 2016. Disponível em: <<http://getbootstrap.com/>>. Acesso em: 6 Agosto 2016.
- BOS, M. W.; DIJKSTERHUIS, A.; VAN BAAREN, R. B. The benefits of "sleeping on things": Unconscious thought leads to automatic weighting. **JOURNAL OF CONSUMER PSYCHOLOGY**, v. 21, n. 1, p. 4-8, janeiro 2011.
- CABRAL, A. S.; YONEYAMA, T. **Economia Digital**: Uma Perspectiva Estratégica para Negócios. São Paulo: Atlas, 2001.
- CHIAVENATO, I. **Administração de produção**: uma abordagem introdutória. Rio de Janeiro: Elsevier, 2005.
- CHOI, S. Y.; STAHL, D. O.; WHINSTON, A. B. **The economics of electronic commerce**. Indianapolis: Macmillan Technical Publishing, 1997.

COBRA, M. **Administração de marketing no Brasil**. 3ª. ed. Rio de Janeiro: Elsevier, 2009.

CORRAR, L. J.; PAULO, E.; FILHO, J. M. D. **Análise multivariada**. São Paulo: Atlas, 2009.

CÔRREA, H. L.; CORRÊA, C. A. **Administração da produção e operações: manufatura e serviços: uma abordagem estratégica**. 2ª. ed. São Paulo: Atlas, 2010.

COTA JÚNIOR, M. B. G. **Aplicação dos métodos QFD, de gestão de desenvolvimento de produtos, e PCP, de gestão de produção, a produtos digitais em uma empresa brasileira de internet móvel**. Universidade Federal de Minas Gerais. Monografia (Graduação em Engenharia de Produção). 2005.

DEMARCO, T.; LISTER, T. **Peopleware: Productive Projects and Teams**. New York: Dorset House, 1987.

FAERBER, S. J. et al. Priming semantic concepts affects the dynamics of aesthetic appreciation. **ACTA PSYCHOLOGICA**, v. 135, n. 2, p. 191-200, outubro 2010.

FENABRAVE. **Anuário 2015**, 2015. Disponível em: <<https://issuu.com/fenabrave/docs/anuario2015/2?e=6659190/34656224>>. Acesso em: 20 março 2016.

FIAT. **Fiat**, 2016. Disponível em: <<http://www.fiat.com.br/>>. Acesso em: 25 março 2016.

FISHER, R. A. **The statistical utilization of multiple measurements**. Londres, p. 179-188. 1936.

FLEURY, A. **Introdução à engenharia de produção**. Rio de Janeiro: Elsevier, 2008.

FREIDEN, J. et al. Information as a product: Not goods, not services. **Marketing Intelligence and Planning**, v. 16, n. 3, p. 210-220, 1998.

GALVÃO, M. B. **Relacionamento de clientes: proposição de um modelo para empresas de serviços com base no CRM**. 2012. 87f. Tese - Universidade Federal de Pernambuco. Recife. 2012.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4ª. ed. São Paulo: Atlas, 2002.

- GOOGLE. Android Developer, 2016. Disponível em: <<https://developer.android.com/index.html>>. Acesso em: 06 Agosto 2016.
- GOVERNO do Estado de São Paulo. **Indústria Automobilística**, 2016. Disponível em: <http://www.saopaulo.sp.gov.br/conhecasp/historia_republica-industria-automobilistica>. Acesso em: 8 abril 2016.
- GROHMANN, M. Z.; SCHOEDLER, A. R.; BATTISTELLA, L. F. Atributos importantes para o consumidor de automóveis: classificação em função instrumental ou expressiva. **R. Adm. Faces Journal**, Belo Horizonte, v. 11, n. 1, p. 67-86, janeiro 2012. ISSN 1984-6975.
- HAIR, J. F. et al. **Análise multivariada de dados**. 5ª. ed. Porto Alegre: Bookman, 2005.
- HOWELL, B. **Information, Technology, and the Productivity Paradox: Incorporating Digital Goods into Productivity Measurement**. Asia-Pacific Productivity Conference. Brisbane: [s.n.]. 2004.
- HSIEH, M.; LINDRIDGE, A. Universal appeals with local specifications. **Journal of Product & Brand Management**, 14, n. 2, 2005. 14-28.
- HUI, K. L.; CHAU, P. Y. K. Classifying Digital Products. **Communications of the ACM**, v. 45, n. 6, p. 73-79, 2002.
- ICARROS. **Catalógo 0km**, 2016. Disponível em: <<http://www.icarros.com.br/catalogo/index.jsp>>. Acesso em: 25 março 2016.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6ª. ed. New Jersey: Pearson Prentice Hall, 2007.
- KOISSO-KANTTILA, N. Digital Content Marketing: A Literature Synthesis. **Journal of Marketing Management**, v. 20, p. 45-65, 2004.
- KOTLER, P.; KELLER, K. L. **Administração de marketing**. 12ª. ed. São Paulo: Pearson Prentice Hall, 2006.
- LANDWEHR, ; WENTZEL, D.; HERRMANN, A. The Tipping Point of Design: How Product Design and Brands Interact to Affect Consumers' Preferences. **PSYCHOLOGY & MARKETING**, v. 29, n. 6, p. 422-433, junho 2012. ISSN 0742-6046.
- LANDWEHR, J. R.; LABROO, A. A.; HERRMANN, A. Gut Liking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecasts. **Marketing Science**, v. 30, n. 3, p. 416-429, maio-junho 2011.

- LARICA, N. J. **LARICA, Neville Jordan. Design de transportes: arte em função da mobilidade.** Rio de Janeiro: 2AB, 2003.
- MANNERING, F. et al. Brand Loyalty and the Decline of American Automobile Firms. **Brookings Papers on Economic Activity**, Washington, DC, 1991. 67-114.
- MAROCO, J. **Análise estatística com a utilização do SPSS.** 3ª. ed. Lisboa: Sílabo, 2007.
- MDIC. **Ministério do desenvolvimento, indústria e comércio exterior**, 2016. Disponível em: <<http://www.mdic.gov.br/sitio/interna/interna.php?area=2&menu=327>>. Acesso em: 30 março 2016.
- MELCHER, C. **Proposta metodológica para avaliações otimizadas de usabilidade em websites desenvolvidos com método ágil: Um Estudo de Caso.** Dissertação de mestrado - PUC-RJ. Rio de Janeiro, p. 299. 2012.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada.** Belo Horizonte: UFMG, 2005.
- MORSE, S. **A administração de marketing.** São Paulo: McGraw-Hill, 1988.
- NETZER, O. et al. Mine Your Own Business: Market-Structure Surveillance Through Text Mining. **Marketing Science**, v. 31, n. 3, p. 521-543, maio 2012.
- PEINE, K.; WENTZEL, D.; HERRMANN, A. Getting better or getting worse? Consumer responses to decreasing, constant, and ascending multi-dimensional price profiles. **REVIEW OF MANAGERIAL SCIENCE**, v. 6, n. 1, p. 81-101, janeiro 2012. ISSN 1863-6683.
- PIRES, V. C.; MARCHETTI, R. Z. **Fatores influenciadores na escolha da fonte de informação na compra de automóveis novos: A importância da comunicação boca a boca.** Encontro Anual da Associação Nacional de Pós-Graduação em Administração. [S.l.]: Anpad. 2000. p. 304.
- PRESSMAN, R. S. **Engenharia de software: uma abordagem profissional.** 7ª. ed. Porto Alegre: AMGH, 2011.
- PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico.** 2ª. ed. Novo Hamburgo: Universidade Feevale, 2013.
- QUAH, D. A weightless economy. **UNESCO Courier**, Dezembro 1998.

QUAH, D. The weightless economy in growth. **The Business Economist**, v. 30, n. 1, p. 40-53, 1999.

QUAH, D. **Digital goods and the new economy**. Centre for Economic Policy Research. London: School of Economics and Political Science. 2003. p. Discussion paper n. 3846.

REIS, E. **Estatística multivariada aplicada**. Lisboa: Edições Sílabo, 1997.

REVISTA Auto Esporte. **Comparador Auto Esporte**, 2016. Disponível em: <<http://revistaautoesporte.globo.com/comparador.html>>. Acesso em: 27 março 2016.

REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. São Paulo: Manoele, 2003.

ROVER, D. et al. **Advantages of Agile Methodologies for Software and Product Development in a Capstone Design Project**. IEEE Frontiers in Education Conference. Madrid: [s.n.]. 2014.

SCHWABER, K. Scrum Development Process. **OOPSLA'95 Workshop on Business Object Design and Implementation**, 1995.

SCHWABER, K.; BEEDLE, M. **Agile Software Development with Scrum**. New Jersey: Prentice Hall, 2002.

SIMILARWEB , 2016. Disponível em: <<https://www.similarweb.com/>>. Acesso em: 27 mar. 2016.

SLACK, N. et al. **Administração da Produção**. 1ª reimpressão 10. ed. São Paulo: Atlas, 2006.

SOARES, M. D. S. Comparação entre Metodologias Ágeis e Tradicionais para o Desenvolvimento de Software. **INFOCOMP Journal of Computer Science**, v. 3, n. 2, p. 8-13, 2004.

SOBEK. **Minerador de Textos Sobek**, 2016. Disponível em: <<http://sobek.ufrgs.br/>>. Acesso em: 10 Julho 2016.

SOMASUNDARAM, G.; SHRIVASTAVA, A. **Armazenamento e Gerenciamento de Informações: Como armazenar, gerenciar e proteger informações digitais**. São Paulo: Bookman, 2009.

STERN, L. W.; EL-ANSARY, A.; COUGHLAN, A. T. **Marketing Channels**. 5ª. ed. New Jersey: Prentice Hall, 1996.

TEIXEIRA, U. **Marketing e Semiótica: Um estudo sobre a produção do sentido no espaço mercadológico**. 2000. 102 f. Dissertação de Mestrado - Universidade Federal do Rio de Janeiro. Rio de Janeiro. 2000.

TELES, V. M. **Extreme Programming**. São Paulo: Novatec, 2004.

TURBAN, E.; KING, D. **Comércio eletrônico: estratégia e gestão**. São Paulo: Prentice Hall, 2004.

VRKLJAN, B. H.; ANABY, D. What vehicle features are considered important when buying an automobile? An examination of driver preferences by age and gender. **JOURNAL OF SAFETY RESEARCH**, v. 42, n. 1, p. 61-65, fevereiro 2011. ISSN 0022-4375.

WOMACK, J. P.; JONES, D. T.; ROOS, D. **A máquina que mudou o mundo**. Rio de Janeiro: Campus, 1992.

WOS. **Web Of Science™**, 2016. Disponível em: <<https://apps.webofknowledge.com/>>. Acesso em: 11 março 2016.

YAMAGUTI, C. L. **O COMPORTAMENTO DO CONSUMIDOR E A INFLUÊNCIA DA FAMÍLIA NO PROCESSO DE DECISÃO DE COMPRA DE AUTOMÓVEIS NOVOS**. 2005. 172 f. Dissertação de Mestrado - Pontifícia Universidade Católica de São Paulo. São Paulo. 2005.

APÊNDICE 1 – RESULTADOS DA ANÁLISE FATORIAL

Variável	Comunalidade análise 1	Comunalidade análise 2	Comunalidade análise 3	Comunalidade análise 4
Preço	0,693141	0,614883	0,633027	0,634092
Motorização	0,836121	0,860143	0,864192	
Direção Elétrica	0,822161	0,803994	0,847047	0,861931
Direção Hidráulica	0,807293	0,722133	0,717058	0,854017
Câmbio Automático	0,682423	0,491273	0,633027	0,70826
Marchas	0,602391	0,655129	-	-
Alcool	0,669505	0,927136	0,679696	0,677209
Diesel	0,858266	0,947093	0,946527	0,950741
Gasolina	0,893567	0,855437	0,960192	0,962007
Potência(cv)	0,901849	0,864301	0,868101	0,873348
Torque (kgf,m)	0,902775	0,844738	0,869794	0,873896
Velocidade Máxima (km/h)	0,844448	0,609673	0,84626	0,848636
Tempo 0-100 (s)	0,60416	0,956567	0,615497	0,613242
Consumo cidade (km/l)	0,957134	0,944049	0,972361	0,977339
Consumo estrada (km/l)	0,947786	0,766461	0,969859	0,976082
Altura (mm)	0,770983	0,800433	0,768039	0,768858
Largura (mm)	0,800393	0,853745	0,802853	0,805823
Comprimento (mm)	0,857223	0,881009	0,866093	0,866354
Entre-eixos (mm)	0,877154	0,918871	0,887861	0,888728
Peso (kg)	0,931283	0,837304	0,924077	0,926521
Tanque (L)	0,839428	0,799427	0,837031	0,837198
Porta-malas (L)	0,859297	0,692177	0,853286	0,863676
Portas	0,733025	0,767573	0,734538	0,770241
Ocupantes	0,775169	0,726388	0,823698	0,837701
Airbag lateral	0,755699	0,614883	0,715156	0,699077
Airbag passageiro	0,457505	-	-	-
Controle de tração	0,754733	0,847471	0,865777	0,871278
Distribuição eletrônica de freia	0,397221	-	-	-
Ar-condicionado	0,675082	0,610704	0,535168	-
Piloto automático	0,696098	0,634802	0,589751	-
Travas elétricas	0,589262	-	-	-
Volante com regulagem de altura	0,666391	0,565488	-	-
Cd player	0,787967	0,81716	0,831013	0,835691
Cd player com MP3	0,808809	0,852697	0,857461	0,863133
Entrada USB	0,520693	-	-	-
Kit Multimídia (GPS e/ou DVD)	0,797736	0,842815	0,849684	0,855872
Rádio FM/AM	0,689375	0,642286	0,65423	0,649563

Ajuste elétrico	0,617688	0,577975	-	-
Bancos de couro	0,469404	-	-	-
Desemb, traseiro	0,548419	-	-	-
Teto solar	0,537737	-	-	-
Vidros elétricos dianteiros	0,767455	0,728648	0,751178	0,865532
Vidros elétricos traseiros	0,785648	0,628055	0,679265	0,778098
Computador de bordo	0,552794	-	-	-
Farol de neblina	0,489336	-	-	-
Sensor de farol	0,650493	0,567384	-	-

Mercedes Benz Cl	1
Mercedes Benz Cl	1
Mercedes Benz Cl	1
Porsche Cayenne	1
Porsche Panamera	1
Rolls-Royce Ghos	1
Rolls-Royce Phan	1
Volkswagen Touar	1
Volvo XC90	1
Aston Martin Van	2
Aston Martin Van	2
Audi R8	2
Audi RS5	2
Bentley Continen	2
BMW M4	2
BMW M6	2
Chevrolet Camaro	2
Ferrari 458 Ital	2
Ferrari 488 GTB	2
Ferrari Califór	2
Ferrari FF	2
Jaguar F-Type	2
Lamborghini Aven	2
Lamborghini Hura	2
Maserati Grancab	2
Maserati GranTur	2
Mercedes Benz AM	2
Mercedes Benz Cl	2
Mercedes Benz Cl	2
Rolls-Royce Wrai	2
Chevrolet S10 Ca	3
Chevrolet S10 Ca	3
Fiat Ducato	3
Ford Ranger (Cab	3
Ford Ranger (Cab	3
Hyundai HR	3
Mahindra Mahindr	3
Mitsubishi L200	3
Mitsubishi Pajer	3
Nissan Frontier	3
SsangYong Actyon	3
Toyota Hilux Cab	3
Toyota Hilux Cab	3

Toyota SW4	3
Troller T4	3
Volkswagen Amaro	3
Audi A1	4
Chery Face	4
Chery QQ	4
Chery Tiggo	4
Chevrolet Cruze	4
Chevrolet Montan	4
Chevrolet Tracke	4
Citroen C4 Loung	4
Fiat 500	4
Fiat Grand Siena	4
Fiat Idea	4
Fiat Linea	4
Fiat Punto	4
Fiat Strada	4
Ford Ecosport	4
Ford Focus Hatch	4
Ford New Fiesta	4
Geely GC2	4
Honda City	4
Honda Civic	4
Honda Fit	4
Honda HR-V	4
Hyundai HB20S	4
Hyundai i30	4
Hyundai ix35	4
JAC J2	4
JAC J3	4
JAC J3 Turin	4
JAC J5	4
JAC J6	4
JAC T6	4
Kia Cerato	4
Kia Picanto	4
Kia Sportage	4
Lifan X60	4
Mitsubishi ASX	4
Mitsubishi Lance	4
Mitsubishi Pajer	4
Peugeot 208	4
Peugeot 308	4

Renault Duster	4
Renault Duster O	4
Suzuki Grand Vit	4
Suzuki Jimny	4
Suzuki S Cross	4
Suzuki Swift Hat	4
Suzuki SX4	4
Volkswagen Cross	4
Volkswagen Fox	4
Volkswagen Golf	4
Volkswagen Jetta	4
Volkswagen Space	4
Volkswagen Space	4
Chery Celer	5
Chevrolet Celta	5
Chevrolet Classi	5
Chevrolet Cobalt	5
Chevrolet Onix	5
Chevrolet Prisma	5
Chevrolet Spin	5
Citroen Aircross	5
Citroen C3	5
Citroen C3 Picas	5
Fiat Bravo	5
Fiat Doblò	5
Fiat Mobi	5
Fiat Palio	5
Fiat Siena	5
Fiat Uno	5
Ford Ka Hatch	5
Ford Ka+	5
Ford New Fiesta	5
Hyundai HB20	5
Hyundai HB20X	5
Hyundai Tucson	5
JAC T5	5
Jeep Renegade	5
Kia Soul	5
Lifan 530	5
Nissan March	5
Nissan Sentra	5
Nissan Versa	5
Peugeot 2008	5

Peugeot 207	5
Peugeot 207 Seda	5
Renault Clio	5
Renault Fluence	5
Renault Logan	5
Renault Sandero	5
Renault Sandero	5
Toyota Corolla	5
Toyota Etios Hat	5
Toyota Etios Sed	5
Volkswagen Gol	5
Volkswagen Savei	5
Volkswagen Up!	5
Volkswagen Voyag	5
Audi TT	6
Audi TTS	6
BMW I3	6
BMW I8	6
DS DS 3	6
Mercedes Benz Cl	6
Mercedes Benz Cl	6
MINI Cabrio	6
MINI Cooper	6
MINI John Cooper	6
MINI Paceman	6
Peugeot RCZ	6
Porsche 911	6
Porsche 911 Cabr	6
Porsche Boxster	6
Porsche Cayman	6
Smart fortwo Cab	6
Smart fortwo Cou	6
Volkswagen Fusca	6
Audi A3	7
Audi A3 Cabriole	7
Audi A3 Sedan	7
Audi A4	7
Audi A4 Avant	7
Audi A5	7
Audi A6	7
Audi Q3	7
Audi Q5	7
Audi RS Q3	7

Audi RS3	7
Audi RS4 Avant	7
Audi S3	7
Audi S3 Sedan	7
Audi S5	7
BMW M3 Sedan	7
BMW Série 1	7
BMW Série 2	7
BMW Série 3	7
BMW Série 4	7
BMW Série 5	7
BMW X1	7
BMW X3	7
BMW X4	7
Chevrolet Captiv	7
Chevrolet Cruze	7
Citroen C4 Picas	7
Citroen Grand C4	7
Dodge Journey	7
DS DS 4	7
DS DS 5	7
Fiat Freemont	7
Fiat Toro	7
Ford Focus Sedan	7
Ford Fusion	7
Honda Accord	7
Honda CR-V	7
Hyundai Azera	7
Hyundai Elantra	7
Hyundai Santa Fe	7
Jaguar XF	7
Jeep Cherokee	7
Jeep Compass	7
Jeep Wrangler	7
Kia Cadenza	7
Kia Optima	7
Kia Sorento	7
Land Rover Disco	7
Land Rover Freel	7
Land Rover Range	7
Lexus CT 200h	7
Lexus ES 350	7
Lexus IS 250	7

Lexus NX 200t	7
Lexus RX 350	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
Mercedes Benz Cl	7
MINI Clubman	7
MINI Countryman	7
Mitsubishi Lance	7
Mitsubishi Outla	7
Mitsubishi Pajer	7
Peugeot 3008	7
Peugeot 308 CC	7
Peugeot 408	7
Porsche Macan	7
SsangYong Korand	7
Subaru Forester	7
Subaru Impreza S	7
Subaru Legacy Se	7
Subaru Outback	7
Subaru XV	7
Toyota Camry	7
Toyota Prius	7
Toyota RAV4	7
Volkswagen CC	7
Volkswagen Golf	7
Volkswagen Passa	7
Volkswagen Passa	7
Volkswagen Tigua	7
Volvo S60	7
Volvo V40	7
Volvo V60	7
Volvo XC60	7

Dodge Ram Pickup	8
Kia Mohave	8
Land Rover Disco	8
Land Rover Range	8
Mercedes Benz Cl	8
Mercedes Benz Cl	8
Mercedes Benz Cl	8
Mercedes Benz Cl	8
SsangYong Rexton	8