



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO ACADÊMICO DO AGRESTE
NÚCLEO DE GESTÃO
CURSO DE ADMINISTRAÇÃO

ARTHUR MENDONÇA LINS CALDAS

**CIÊNCIA DE DADOS NO PROCESSO DECISÓRIO EMPRESARIAL: Um estudo de
caso sobre segmentação de mercado na Pearson Education**

Caruaru
2023

ARTHUR MENDONÇA LINS CALDAS

CIÊNCIA DE DADOS NO PROCESSO DECISÓRIO EMPRESARIAL: Um estudo de caso sobre segmentação de mercado na Pearson Education

Trabalho de Conclusão de Curso apresentado ao Curso de Administração da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de Bacharel em Administração.

Área de concentração: Administração.

Orientador: Prof^o. Dr. Elielson Oliveira Damascena.

Caruaru

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Caldas, Arthur Mendonça Lins.

Ciência de Dados no Processo Decisório Empresarial: Um estudo de caso
sobre segmentação de mercado na Pearson Education / Arthur Mendonça Lins
Caldas. - Caruaru, 2023.

42 p. : il., tab.

Orientador(a): Elielson Oliveira Damascena

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro Acadêmico do Agreste, Administração, 2023.

Inclui referências, apêndices.

1. Administração. 2. Processo decisório. 3. Ciência de Dados. 4. Segmentação
de mercado . 5. Tecnologia. I. Damascena, Elielson Oliveira . (Orientação). II.
Título.

600 CDD (22.ed.)

ARTHUR MENDONÇA LINS CALDAS

CIÊNCIA DE DADOS NO PROCESSO DECISÓRIO EMPRESARIAL: Um estudo de caso sobre segmentação de mercado na Pearson Education

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Administração da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Bacharel em Administração.

Aprovada em: 27/02/2023

BANCA EXAMINADORA

Prof^o. Dr. Elielson Oliveira Damascena (Orientador)
Universidade Federal de Pernambuco

Prof^o. Marllon Emanuel Souza Medeiros de Vasconcelos (Examinador Interno)
Universidade Federal de Pernambuco

Prof^o. Omero Galdino da Silva Júnior (Examinador Externo)
Uninassau

AGRADECIMENTOS

Agradeço primeiramente a Deus e a minha família, que me proporcionaram toda a construção do meu caráter, saúde, educação e os mais positivos estímulos para que minha jornada educacional se mantivesse perseverante e pudesse chegar ao patamar em que se encontra atualmente.

Agradeço aos meus amigos, colegas de trabalho e faculdade, por conseguirem, em meio a exaustiva rotina, trazer uma considerável leveza para o cotidiano da vida acadêmica.

Agradeço aos meus professores, desde o maternal até a faculdade, muitos dos quais adquiri conhecimento que não apenas converti em palavras que alimentam este trabalho mas também guardo comigo para a vida profissional e pessoal.

RESUMO

Na atualidade é observado cada vez mais a crescente que vive a tecnologia dentro das organizações, mais especificamente a capacidade de extrair valor e conhecimento dos dados que a organização possui. O objetivo geral da pesquisa é analisar como a Ciência de Dados apoia o processo decisório para segmentação de mercado na Pearson Education. Os dados foram coletados através de três entrevistas realizadas com funcionários da Pearson Education. A análise dos dados foi feita mediante análise de conteúdo. Os resultados apontam como o processo decisório para segmentação de mercado na organização em questão diverge nas etapas de geração e escolha de alternativas comparado a forma como estas etapas são referenciadas na literatura. Além disso, ressalta como a Ciência de Dados surge como ferramenta de apoio para a etapa de geração de alternativas dentro deste processo.

Palavras-chave: Processo decisório; Ciência de Dados; Tomada de decisão; Segmentação de mercado.

ABSTRACT

Nowadays, it has been observed the growing experience of technology within organizations, more specifically the ability to extract value and knowledge from the data that the organization possesses. The general objective of the research is to analyze how Data Science supports the decision-making process for market segmentation at Pearson Education. The data was collected through three interviews with Pearson Education employees. Data analysis was performed through content analysis. The results point out how the decision-making process for market segmentation in the organization in question differs in the steps of generation and selection of alternatives compared to how these steps are referenced in the literature. In addition, it highlights how Data Science emerges as a support tool for the alternative generation step within this process.

Keywords: Decision-making process; Data Science; Decision making; Market segmentation.

LISTA DE ABREVIATURAS E SIGLAS

DOD	Decisão orientada a dados
ELT	English Language Teaching
E1	Entrevistada 1
E2	Entrevistado 2
E3	Entrevistada 3
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
TCC	Trabalho de conclusão de curso
TI	Tecnologia da informação

SUMÁRIO

1	INTRODUÇÃO.....	10
1.1	PROBLEMA DE PESQUISA.....	11
1.2	OBJETIVOS.....	11
1.2.1	Objetivo Geral.....	11
1.2.2	Objetivos Específicos.....	11
1.3	JUSTIFICATIVA.....	11
2	REFERENCIAL TEÓRICO.....	13
2.1	PROCESSO DECISÓRIO NAS ORGANIZAÇÕES.....	13
2.1.1	Contexto histórico.....	13
2.1.2	O processo decisório e seus elementos.....	14
2.2	CIÊNCIA DE DADOS.....	16
2.2.1	Evolução e conceito.....	16
2.2.2	Tomada de decisão orientada a dados e aplicações da Ciência de Dados.....	18
2.2.3	Agrupamento em Ciência de Dados.....	21
2.3	SEGMENTAÇÃO DE MERCADO.....	21
3	PROCEDIMENTOS METODOLÓGICOS.....	25
3.1	CARACTERIZAÇÃO DA PESQUISA.....	25
3.2	O CASO ESTUDADO: PEARSON EDUCATION.....	26
3.3	PROCESSO DE ENTREVISTAS.....	27
4	ANÁLISE DOS DADOS.....	29
4.1	ENVOLVIMENTO COM O PROJETO.....	29
4.2	PROCESSO DECISÓRIO PARA SEGMENTAÇÃO DE CLIENTES.....	30
4.2.1	Identificação do problema.....	30
4.2.2	Diagnóstico.....	31
4.2.3	Geração e escolha de alternativas.....	31
4.3	APOIO DA CIÊNCIA DE DADOS NO PROCESSO DECISÓRIO.....	33
5	CONSIDERAÇÕES FINAIS.....	36

REFERÊNCIAS.....	38
APÊNDICE A - ROTEIRO DE ENTREVISTA DEMANDANTES.....	41
APÊNDICE B - ROTEIRO DE ENTREVISTA DEMANDADO.....	42

1 INTRODUÇÃO

Atualmente vemos cada vez mais a tecnologia e inovação presentes em organizações que almejam conseguir algum destaque competitivo no mercado. Logo, quando falamos de tecnologia e inovação dentro do âmbito organizacional, a Tecnologia da Informação (TI) torna-se imprescindível para ser comentada. A TI, a qual envolve conceitos mais abrangentes que os de processamento de dados, sistemas de informação e engenharia de software, evoluiu sua orientação de um suporte administrativo para um papel estratégico dentro da organização na medida em que a sua visão como arma estratégica competitiva tem sido discutida e enfatizada (LAURINDO et al., 2001).

Dentro dessa esfera tecnológica, no que se refere ao processamento de dados, as organizações têm voltado cada vez mais sua atenção para as possibilidades competitivas que os dados provêm. Por conseguinte, com o avanço tecnológico, cresce progressivamente os investimentos organizacionais em infraestrutura de negócios, consequentemente abrindo diversos aspectos do negócio como comportamento do cliente, gestão da cadeia de fornecimento, fluxo de trabalho e muito mais para a capacidade de coleta de dados (FAWCETT; PROVOST, 2018).

A abertura que o avanço tecnológico proporcionou para o mundo dos dados, ao mesmo tempo que informações sobre consumidores, concorrentes e mercado aumentaram e se tornaram cada vez mais acessíveis, criou um berço para que novos frutos e vantagens pudessem ser obtidas mediante a orientação das organizações para esse universo.

Conforme comentam Fawcett e Provost (2018, p.1): “Essa ampla disponibilidade de dados levou ao aumento do interesse em métodos para extrair informações úteis e conhecimento a partir de dados — o domínio de data science”. Dessa forma, surge a necessidade organizacional de não apenas coletar e organizar seus dados, porém, também conseguir usufruir de algo ainda mais útil com eles em mãos através da Ciência de Dados.

Nesse sentido, dentre os processos beneficiados pela extração de valor dos dados, é ressaltado o processo decisório dentro das organizações, processo o qual não apenas é de suma importância para a organização, mas também é um processo o qual a Ciência de Dados está diretamente relacionada.

Qualquer processo decisório organizacional diz respeito à descoberta e seleção de alternativas satisfatórias, e para se chegar a uma decisão adequada, é necessário uma análise do sistema considerando de maneira correta todas as variáveis que afetam essa decisão. A relevância do processo decisório nas organizações é tamanha que nada acontece na

organização até o momento em que as pessoas tomem decisões e ajam, além disso, a tomada de decisão encontra-se presente em todas as funções do administrador, desde a fixação de metas, divisão do trabalho, formação, condução e motivação da equipe até o controle e tomada de medidas corretivas (LACOMBE; HEILBORN, 2008).

1.1 PROBLEMA DE PESQUISA

Diante do exposto, torna-se evidente a capacidade de envolvimento da Ciência de Dados com o processo decisório dentro das organizações. Dentro deste cenário, busca-se uma maior compreensão relacionado a como a Ciência de Dados apoia o processo decisório empresarial mediante estudo de uma aplicação prática do tema em um projeto de segmentação de mercado da Pearson Education, uma empresa do ramo educacional. Desta forma, a pergunta de pesquisa deste Trabalho de Conclusão de Curso (TCC) é expresso pelo seguinte questionamento:

Como a Ciência de Dados apoia o processo decisório para segmentação de mercado na Pearson Education?

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Analisar como a Ciência de Dados apoia o processo decisório para segmentação de mercado na Pearson Education.

1.2.2 Objetivos Específicos

I - Descrever como ocorre o processo decisório da Pearson Education em um projeto de segmentação de mercado.

II - Compreender em quais etapas do processo decisório da Pearson Education a Ciência de Dados é aplicada em um projeto de segmentação de mercado.

1.3 JUSTIFICATIVA

A pesquisa torna-se relevante para organizações que possuem grandes volumes de dados e intencionam aprimorar sua tomada de decisão, propondo a elas possíveis novas ferramentas para alavancar seu processo decisório. Por conseguinte, o presente estudo também serve de referência para as organizações que já utilizam a Ciência de Dados dentro de seu processo decisório, os quais podem se inspirar e se basear nos conceitos e aplicações sugeridos neste trabalho a fim de aprimorar suas respectivas implementações já existentes.

A pesquisa também constitui-se como fonte de valor para a própria empresa em que foi realizado o estudo, a qual através das aplicações teóricas referenciadas e do estudo de caso realizado, pode refletir as concepções atuais do projeto discutido e de possíveis futuros projetos, identificando pontos de melhorias e novas aplicações da área na organização.

O TCC mostra-se relevante, ainda, para a academia, por ter buscado reunir referências teóricas que discutem um tema que é de suma importância para os administradores conforme comentam Lacombe & Heilborn (2008, p. 438) “Provavelmente, a maior responsabilidade dos administradores é tomar decisões” ao mesmo tempo em que estuda como essa responsabilidade pode ser auxiliada.

Outro aspecto que corrobora com a relevância deste TCC é o estudo de caso situado no ramo educacional, tendo em vista a expressividade deste ramo para diversos aspectos econômicos e sociais no país. Essa importância é ressaltada nos estudos de Barros e Mendonça (1997) os quais afirmam que o atraso educacional no Brasil afeta não apenas o crescimento econômico do país, mas também o crescimento populacional, a mortalidade e o desempenho educacional futuro.

Ao mesmo tempo, o trabalho em questão auxilia a compreensão de como uma empresa líder global no ramo de educação compõe seu processo decisório para segmentação de mercado. Dessa forma, apoiando a compreensão de como esta grande atuante do mercado educacional avalia suas possibilidades de consumidores e utiliza-se de tecnologia para alavancar sua tomada de decisão e seu negócio.

Através do capítulo seguinte, será abordado a referência teórica que embasa este TCC, onde foram explicitados diversos conteúdos conceituados por vários autores ao longo das décadas no que se refere aos respectivos assuntos que envolvem este estudo.

2 REFERENCIAL TEÓRICO

O referencial teórico se incube de expor conceitos encontrados na literatura que respaldam os temas tratados pelo presente estudo. Ao longo desta seção será discorrido sobre o processo decisório, seu contexto histórico e alguns de seus principais conceitos; a Ciência de Dados, sua evolução, tomada de decisão orientada à dados, conceitos e principais aplicações; e, por fim, a segmentação de mercado e seus principais conceitos.

2.1 PROCESSO DECISÓRIO NAS ORGANIZAÇÕES

2.1.1 Contexto histórico

O processo decisório é contextualizado historicamente desde os primórdios da civilização, em que o homem buscava de diversas formas auxiliar seu processo de tomada de decisão, seja através de divindades ou supostos poderes místicos. Nesse contexto, pela tomada de decisão ser dependente de algo místico e divino, as decisões tomadas pelos indivíduos eram consideradas sábias e livres de erros, e caso as decisões não dessem os esperados frutos, a culpa era atribuída às supostas divindades responsáveis, evidenciando o quanto a tomada de decisão na época era influenciada por crenças em seres sobrenaturais. Por conseguinte, ao longo da história é possível observar como líderes religiosos como Jesus, Maomé e Buda, por exemplo, passaram a exercer influência direta na vida das pessoas e na tomada de decisão delas, e como, atualmente, ainda que exista influência religiosa ou de divindades específicas na tomada de decisão das pessoas, há outros fatores que em conjunto auxiliam o processo decisório (BISPO; CAZARINI, 1998).

Ainda de acordo com os autores, no início do século persistia no ambiente organizacional a ideia de que os trabalhadores eram seres passivos e despreparados, sem capacidade para atuar nas decisões. Devido ao ambiente ser considerado estável, relativamente simples e com restrição de informações, a tomada de decisão era centrada em geral no proprietário, que em tese detinha conhecimento bastante amplo de todas as alternativas e de suas consequências. Posteriormente, com o advento de diversas teorias como a teoria das relações humanas, o indivíduo passou a não ser mais visto como mera mão de obra, mas sim como capaz de pensar, de decidir e de ser motivado.

Outra teoria que ressalta essa transição da ótica de como o indivíduo era visto também é ressaltada por Moritz e Pereira (2015), conforme comentam sobre como a Teoria das

Decisões nasceu de Hebert Simon. Os autores explicam a obra de Simon como concebendo a organização como um sistema de decisões em que cada pessoa participa racional e conscientemente, tomando decisões individuais a respeito de alternativas racionais de comportamento.

No que se refere ao estudo do processo decisório, os autores comentam o quanto evoluiu o estudo do tema desde a década de 40, na medida em que novas técnicas administrativas foram desenvolvidas, bem como com a absorção de novos procedimentos quantitativos oriundos da Matemática e da Pesquisa Operacional. Nesse sentido, os autores destacam como as decisões hoje são muito mais direcionadas para o futuro, sendo isso fruto das idealizações as quais as decisões são baseadas.

2.1.2 O processo decisório e seus elementos

No processo decisório, o tomador de decisão está inserido em uma situação e segue estratégias em busca de alcançar algum objetivo específico. Sendo assim, a decisão envolve uma opção, e o processo decisório envolve uma seleção de alternativas. Todo o curso da seleção de alternativas segue um processo racional em busca de atingir o objetivo delimitado pelo tomador de decisão (CHIAVENATO, 2014).

Para Chiavenato (p. 340, 2014) a decisão seria “[...] o processo de análise e escolha entre as alternativas disponíveis de cursos de ação que a pessoa deverá seguir”. Dessa forma, o autor em questão destrincha a decisão em elementos e o processo decisório em etapas as quais a Teoria das Decisões é parte intrínseca. Nesse sentido, para Chiavenato (p. 340-341, 2014) a decisão seria composta de seis elementos:

1. Tomador de decisão: é a pessoa que faz uma escolha ou opção entre várias alternativas futuras de ação.
2. Objetivos: são os objetivos que o tomador de decisão pretende alcançar com suas ações.
3. Preferências: são os critérios que o tomador de decisão usa para fazer sua escolha.
4. Estratégia: é o curso de ação que o tomador de decisão escolhe para atingir seus objetivos. O curso de ação é o caminho escolhido e depende dos recursos de que pode dispor.
5. Situação: são os aspectos do ambiente que envolvem o tomador de decisão, alguns deles fora do seu controle, conhecimento ou compreensão e que afetam sua escolha.
6. Resultado: é a consequência ou resultado de uma dada estratégia.

Quanto às etapas do processo decisório, o autor as define em sete partes:

1. Percepção da situação que envolve algum problema.
2. Análise e definição do problema.
3. Definição dos objetivos.
4. Procura de alternativas de solução ou de cursos de ação.

5. Escolha (seleção) da alternativa mais adequada ao alcance dos objetivos.
6. Avaliação e comparação das alternativas.
7. Implementação da alternativa escolhida.

Ainda no que se refere às etapas, o autor ressalta como nem sempre elas são seguidas à risca, podendo serem suprimidas, abreviadas, ampliadas e estendidas mediante a adequação e necessidade de cada situação.

Já pelas palavras de Maximiano (p. 139, 2000), as decisões seriam “[...] escolhas que as pessoas fazem para enfrentar problemas e aproveitar oportunidades”. Para o autor, entre diversas ações que exigem escolhas, tomar decisões para enfrentar problemas não apenas faz parte de grande parte do trabalho dos gerentes, mas faz parte do trabalho de administrar.

No que se refere ao processo decisório, Maximiano (p. 141, 2000) o define como sendo “[...] a sequência de etapas que vai desde a identificação da situação até a escolha e colocação em prática a ação ou solução. Quando a decisão é colocada em prática, o ciclo se fecha”. Por outro lado, diferentemente do que mencionava Chiavenato, Maximiano (2000) resume o processo decisório em apenas 4 etapas:

1. Constatação ou identificação do problema ou oportunidade: Inicia-se com uma frustração, interesse, desafio ou até irritação. Esta é a fase que se inicia o processo de tomada de decisão com a percepção de que existe um obstáculo entre o tomador de decisão e seu objetivo, e que será necessário tomar uma decisão.
2. Diagnóstico: O tomador de decisão irá analisar o problema ou a situação em questão e irá entender o que está acontecendo ou o que pode acontecer, analisando as causas do problema e todas as informações relevantes. Caso seja um problema comum, basta aplicar a solução padronizada, caso contrário, o tomador de decisão deve analisar todas as variáveis que envolvem o problema, algo que muitas vezes é condicionado pela urgência do problema.
3. Geração de alternativas: A fase de geração de alternativas trata-se do levantamento de possíveis opções para solucionar o problema diagnosticado na fase anterior.
4. Avaliação ou escolha de uma alternativa: Nesta fase as alternativas geradas são comparadas e julgadas por meio de critérios. É selecionada a alternativa que apresenta mais vantagens para solucionar o problema. Ainda que esta seja a fase final do processo decisório, o ciclo de resolução de problemas apenas se encerra quando a decisão é implementada.

Diante dos conceitos expostos, tendo em vista que as principais compreensões que norteiam o processo decisório foram apresentadas, a próxima seção deste capítulo abordará sobre a Ciência de Dados.

2.2 CIÊNCIA DE DADOS

2.2.1 Evolução e conceito

Em tempos anteriores, as organizações recorriam a estatísticos, modeladores e analistas que eram responsáveis por realizarem um trabalho árduo na tentativa de extrair rendimento ao máximo dos dados que eram coletados, porém, o volume desses dados era demasiado para que houvesse capacidades manuais de modelagem. Nos tempos atuais, com o avanço da tecnologia, computadores, e algoritmos, essa tarefa tornou-se muito mais simples e capaz de ser realizada, conseqüentemente, a convergência desses fenômenos permitiu de maneira cada vez mais difundida realizar a aplicação de técnicas de Ciência de Dados nos negócios (FAWCETT; PROVOST, 2018).

Por conseguinte, devido à recência da evolução tecnológica, a Ciência de Dados acaba sendo considerada uma ciência relativamente nova, ainda que a expressão “Data Science” seja datada da década de 60 (AMARAL, 2016). Como resultado, existem muitas compreensões acerca do que pode ser definido como Ciência de Dados ou *Data Science*. Essa confusão pode ser entendida devido a popularidade que norteia o termo “Ciência de Dados” e termos semelhantes que acabam sendo confundidos com este, como “Data Mining”. Este último, por exemplo, acabou desenvolvendo vida própria, visto que muitas organizações e indivíduos buscam usufruir do atual alarde que cerca o tema (FAWCETT; PROVOST, 2018).

No que se refere a confusão do conceito da temática em questão, Amaral (2016, p.4) comenta:

Normalmente, a ciência de dados é associada de forma equivocada apenas aos processos de análise dos dados, onde com o uso de estatística, aprendizado de máquina ou a simples aplicação de um filtro se produz informação e conhecimento. Nessa visão “miope”, a ciência de dados passa a ser vista apenas como um nome mais elegante para a estatística.

Dessa maneira, Amaral (2016) ressalta a diferença entre a estatística e a Ciência de Dados descrevendo primeiramente o ciclo de vida do dado. De acordo com o autor, o dado passa por 5 etapas, as quais seriam a produção, armazenamento, transformação, análise e o descarte do dado. Nesse sentido, segundo o autor, a principal divergência entre estatística e Ciência de Dados pode ser explicada pelo fato da estatística, dentro do ciclo de vida do dado, estar associada apenas à fase de análise, enquanto que a Ciência de Dados englobaria todo o ciclo de vida do dado. Mediante o exposto, Amaral (2016, p.6) define a Ciência de Dados

como sendo “[...] os processos, modelos e tecnologias que estudam os dados durante todo o seu ciclo de vida”.

Por outro lado, Fawcett e Provost (2018) ressaltam que, dentro do processamento de dados, a Ciência de Dados não está inteiramente relacionada a todas as suas etapas, conseqüentemente, algumas tecnologias, habilidades e sistemas do processamento de dados são erroneamente atribuídos como Ciência de Dados. Ainda de acordo com os autores, diversas etapas do processamento de dados, por mais que sirvam de suporte à Ciência de Dados, servem para muito mais que isso. Assim, os autores definem a Ciência de Dados como sendo “[...] um conjunto de princípios fundamentais que norteiam a extração de conhecimento a partir de dados” (FAWCETT; PROVOST, 2018, p.2).

Por sua vez, ainda que a Ciência de Dados possua elementos da estatística em sua estrutura, outras competências integram o perfil de um cientista de dados, de forma que a área da Ciência de Dados pode ser entendida como uma área interdisciplinar. Uma abordagem visual para compreender os elementos da Ciência de Dados pode ser visualizada com o Diagrama de Venn, construído por Conway (2010):

Figura 1: Diagrama de Venn



Fonte: Conway (2010, tradução nossa)

No Diagrama de Venn, é possível visualizar três áreas que englobam a Ciência de Dados, sendo elas, de acordo com o autor, Habilidades de Hacking (conhecimentos relacionados a programação e ciência da computação), Conhecimento em Matemática e Estatística e Expertise Substantiva (conhecimento suficiente para trazer hipóteses e questões acerca do mundo que possam se transformar em dados e testados com métodos estatísticos). Além disso, a figura exhibe intersecções entre as áreas demonstradas:

- Zona de perigo: De acordo com o autor, esta seria a área em que estão situadas pessoas capazes de utilizar as Habilidades de Hacking e que possuem Expertise Substantiva, porém, carecem de conhecimento de Matemática e Estatística, então não devem saber ao certo o significado das aplicações que estão fazendo.
- Aprendizado de Máquina: É a convergência entre Matemática e Estatística com Habilidades de Hacking, em que o cientista de dados consegue utilizar os dotes de programação para aplicar os conhecimentos técnicos.
- Pesquisa Tradicional: A convergência entre Matemática e Estatística com Expertise Substantiva, resultando no que o autor atribui como sendo a pesquisa tradicional.
- Data Science: A interseção que engloba as três áreas referenciadas.

Seguindo a mesma lógica do Diagrama de Venn anteriormente exibido, Rautenberg e Do Carmo (2019, p.59) comentam da interdisciplinaridade da Ciência de Dados: “Aliado ao surgimento do Big Data, tem-se o advento da Ciência de Dados como um campo de atuação de competências interdisciplinares em ascensão”. De acordo com os autores, a Ciência de Dados se utilizaria de três áreas de conhecimento para suas aplicações de forma muito semelhante à forma exibida por Conway no Diagrama de Venn, sendo elas a Programação de Computadores; Estatística e Matemática; e Domínio do Conhecimento.

Dessa forma, fica evidente a importância da relação entre negócios e Ciência de Dados no que é referenciado como Domínio do Conhecimento pelos autores:

Para o efetivo sucesso das soluções de Ciência de Dados, o Conhecimento do Domínio do problema deve ser disponível e amplamente utilizado no Processo de Tomada de Decisão. Neste sentido, as soluções de Ciência de Dados são voltadas para a formulação de hipóteses e a aquisição de informação aderente como insumo no processo decisório (RAUTENBERG; DO CARMO, 2019, p.59).

2.2.2 Tomada de decisão orientada a dados e aplicações da Ciência de Dados

O contexto e as suposições que guiam a tomada de decisão representam aspectos fora do controle de qualquer tomador de decisão, porém, as premissas e o conhecimento de negócio depende dos dados, tendo em vista que estes são parte da organização como um sistema, sendo assim, a tomada de decisão orientada por dados poderia ser definida como sendo a prática de basear decisões em dados ao invés de puramente pela intuição (DIVÁN, 2017).

Nesse sentido, o objetivo primordial da Ciência de Dados é o aprimoramento da tomada de decisão. Os benefícios proporcionados pela tomada de decisão orientada para dados (DOD) são conclusivos no sentido de que se relacionam não apenas a um aumento de produtividade nas empresas, mas também com um maior retorno sobre seus respectivos ativos, patrimônio líquido e valor de mercado no geral (FAWCETT; PROVOST, 2018).

Ainda de acordo com os autores, é tarefa do cientista de dados, em conjunto com os investidores da empresa ou responsáveis pelo problema de negócio, decompor o problema de negócio em subtarefas de forma que cada uma dessas sub tarefas sejam elementos tratáveis pela Ciência de Dados Nesse sentido, uma das técnicas de mineração de dados comentada é o uso de algoritmos para solucionar problemas de negócio específicos.

Os autores ainda afirmam que existem incontáveis algoritmos que podem ser utilizados por um cientista de dados durante o processo de mineração dos dados, porém, ainda existe uma quantidade específica de tarefas de negócio fundamentalmente diferentes tratadas por esses algoritmos. Segundo os autores, essas tarefas são:

1. Classificação e estimativa de probabilidade: No caso da classificação, o algoritmo tenta prever para qual classe um indivíduo de uma população pertence. Um exemplo de classificação são clientes que são ou não são suscetíveis a responder a uma oferta de marketing. Enquanto a classificação atribui à classe ao qual o indivíduo pertenceria, a estimativa de probabilidade atribui um percentual que representa a probabilidade do indivíduo pertencer ou não a uma das classes determinadas.
2. Regressão: Tenta prever um valor número de um determinado indivíduo de uma população. Um exemplo de uma tarefa de regressão é determinar quanto em valor um determinado cliente usará de um serviço específico.
3. Combinação por similaridade: Identifica indivíduos semelhantes com base nos dados disponíveis sobre eles. Um exemplo de uma tarefa de combinação por similaridade é

uma empresa buscar encontrar clientes semelhantes a seus melhores clientes comerciais.

4. Agrupamento: Identifica indivíduos semelhantes dentro de uma população, porém, diferentemente da combinação por similaridade, não possui nenhum propósito inicial específico. Um exemplo de agrupamento seria uma empresa tentar entender se seus clientes formam grupos ou segmentos naturais.
5. Agrupamento de coocorrência: Busca encontrar associações entre entidades de acordo com transações. Um exemplo de agrupamento por coocorrência seria uma empresa tentar entender quais itens são comprados juntos, dessa forma, sugerindo um dos itens para o consumidor no momento da compra do outro.
6. Perfilamento: É a caracterização de um comportamento típico de um indivíduo, grupo ou até mesmo uma população. Um exemplo de perfilamento seria o de uma empresa que busca entender qual o uso típico de celular em um determinado tipo de cliente.
7. Previsão de vínculo: Busca identificar ligações entre itens de dados ou sugerir uma possível ligação que poderia existir, até mesmo estimando a força da ligação. Um exemplo de previsão de vínculo são as redes sociais sugerirem conexões novas para um determinado usuário com base em conexões semelhantes que estes compartilham.
8. Redução de dados: Busca reduzir um conjunto de dados para um conjunto de dados menor que o original, contendo grandes partes das informações importantes do conjunto de dados original. Um exemplo de redução de dados é a transformação de um conjunto de dados sobre preferências de filmes de consumidores para um conjunto menor que revela os gostos mais evidentes dos consumidores.
9. Modelagem causal: Busca a compreensão de quais acontecimentos ou ações influenciam em determinado resultado. Um exemplo de modelagem causal é uma empresa que direciona anúncios para consumidores e descobre que os consumidores alvo têm uma maior taxa de compra após terem se tornado alvos.

2.2.3 Agrupamento em Ciência de Dados

Conforme comentado anteriormente, uma das técnicas de Ciência de Dados utilizadas para tratar problemas de negócio é a técnica de agrupamento ou *clustering*. Esta técnica se encaixa no que é conhecido como método não supervisionado, que ao contrário dos métodos supervisionados, as classes (grupos) que este procedimento deseja encontrar não são conhecidas previamente, nesta situação cabe ao próprio processo de agrupamento gerar os diferentes grupos que os dados se encaixam. Dessa forma, o principal objetivo desse processo é maximizar as distâncias entre os diferentes grupos formados e minimizar a distância dentro dos grupos, de forma que os elementos de um grupo sejam parecidos entre si, mas diferente dos elementos presentes nos demais grupos (SILVA; FERRARI, 2016).

Nesse sentido, Fawcett e Provost (2018) esclarecem como podemos querer encontrar grupos de clientes em uma determinada base de dados que não são impulsionados por uma característica pré-determinada. Posto isto, a técnica de agrupamento pode ser aplicada numa tentativa de entender se os clientes de uma determinada empresa se enquadram naturalmente em grupos diferentes, desenvolvendo assim uma nova perspectiva desses clientes, gerando, por exemplo, novos produtos, campanhas de marketing ou novos métodos de atender ao grupo de melhores clientes que foram encontrados.

Diante do exposto, torna-se ideal comentar sobre alguns dos principais conceitos relacionados ao produto final proveniente do agrupamento da Ciência de Dados: a segmentação de mercado. A seção seguinte discorre alguns destes conceitos que estão presentes na literatura.

2.3 SEGMENTAÇÃO DE MERCADO

O início do trabalho de profissionais de marketing, principalmente no que se refere ao entendimento do consumidor, se dá através da segmentação de mercado, em que os profissionais identificam, através de características específicas, grupos distintos de consumidores com desejos de consumo diferentes. Afinal, uma empresa não pode atender a todos os seus clientes em mercados amplos e diversificados, porém, é capaz de segmentar esses clientes em grupos específicos, generalizando suas necessidades e desejos, assim, identificando qual desses segmentos ela pode atender com maior eficácia (KOTLER, 2012). Nesse sentido, Ferreira (p. 2, 2000) comenta: "Segmentar um mercado significa escolher um

grupo de consumidores, com necessidades homogêneas, para o qual a empresa poderá fazer uma oferta mercadológica”.

Para isso, é necessário identificar características que permitam mapear o comportamento de consumo dos clientes da organização, sendo assim, é necessário que o segmento seja identificável, mensurável, acessível, rentável e estável. Através disso, as empresas podem construir produtos sob medida para segmentos selecionados, os atendendo com mais atenção e eficiência que seus concorrentes (FERREIRA, 2000).

Ainda segundo Ferreira (2000), após estimar o tamanho do mercado, identificar os concorrentes e o potencial de crescimento do mercado, os consumidores são separados em seus respectivos grupos em que cada um dos integrantes do grupo seja semelhante entre si e diferente dos demais grupos, podendo essa segmentação ser baseada em critérios geográficos, demográficos, psicográficos ou comportamentais. Após isso, os grupos são avaliados e por fim escolhidos para serem atendidos. Desta forma, ao invés de gastar seus esforços com grupos de consumidores que talvez não tragam o retorno esperado, as empresas podem, através da segmentação de consumidores, focar nos clientes que realmente valem a pena. Nesse sentido, Kotler (p. 228 , 2012) comenta:

Para competir com mais eficácia, muitas empresas têm adotado o marketing de mercado-alvo. Em vez de dispersar seus esforços de marketing, elas concentram sua atenção em consumidores a quem poderiam efetivamente satisfazer. Um marketing de mercado-alvo eficaz exige que os gerentes de marketing:

1. Identifiquem e tracem o perfil de grupos distintos de compradores que diferem em suas necessidades e preferências (segmentação de mercado).
2. Seleccionem um ou mais segmentos nos quais a empresa tenha interesse em ingressar (mercados-alvo).
3. Para cada segmento, determinem e divulguem os principais benefícios que diferenciam os produtos da empresa (posicionamento de mercado).

Nesse contexto, para Mcdonald e Dunbar (2013), o processo de segmentação de mercado pode ser dividido em duas fases e quatro estágios. A primeira fase é composta por três estágios que devem ser seguidos para segmentar o mercado, incluindo nesse processo todo o mercado que a organização é capaz de atuar:

1. O primeiro estágio estabelece o escopo ao segmentar a área do projeto geograficamente e entendendo pela perspectiva do consumidor as intenções ou propósito dele de adquirir o produto que a organização ou os concorrentes dela podem ofertar. Através disso é feito um fluxograma que mapeia como o mercado é coberto

pelo seu projeto e como o capital dos consumidores flui entre a organização, fornecedores e concorrentes, bem como o fluxo dos serviços dos fornecedores de volta para o consumidor.

2. O segundo estágio consiste em olhar para o mapeamento de mercado e seus pontos de tomada de decisão, construindo um modelo do mercado baseado nos diferentes consumidores encontrados e nas transações que eles fazem, requer dessa forma que as principais variáveis levantadas pelo mercado que influenciam no momento de comparação de alternativas sejam registradas.
3. O terceiro estágio consiste em olhar para as escolhas dos tomadores de decisão nos estágios anteriores e entender quais variáveis influenciam a comparação de alternativas destes decisores. Assim, a importância de cada variável é atribuída a cada célula dentro do modelo de mercado construído. Após isso, técnicas de agrupamento são aplicadas, de forma que células que se parecem, no sentido de importância das necessidades que elas procuram ser sanadas, são combinadas para formar segmentos com o melhor agrupamento possível. Após isso, cada segmento é sujeito a uma checagem de realidade baseada no tamanho, diferenciação de ofertas demandadas, viabilidade de atender aos consumidores e compatibilidade com a organização.

A segunda fase, de acordo com os autores, consiste em olhar para os segmentos compostos e entender quais deles a organização deve operar e compor as estratégias de marketing. O quarto estágio, contido nesta fase, consiste em:

4. Olhar para os segmentos e definir o critério de atratividade de um dos segmentos para a organização, criando assim uma pontuação de atratividade para cada um dos segmentos. O passo final da segmentação seria estabelecer a habilidade da organização para atender a necessidade dos segmentos comparado a habilidade dos concorrentes para atender essas necessidades.

No que se refere a segmentação de mercado, tendo em vista a natureza do estudo de caso deste TCC, torna-se essencial comentar da segmentação de organizações. Para Kotler (2012), a segmentação de mercados organizacionais assim como a segmentação de mercado de consumidores pode ser executada através de variáveis geográficas, benefícios procurados e índices de utilização, que da mesma forma que funcionam para segmentar o mercado organizacional, funcionam para segmentar o mercado de consumidores.

No entanto, ainda de acordo com o autor, é possível aplicar outras variáveis a segmentação de organizações, como porte da empresa, recursos dos clientes, tecnologia, tamanho do pedido, entre outras diversas variáveis que, por natureza, não se enquadram exclusivamente na segmentação de mercado de organizações. Ainda vale ressaltar que no cenário do mercado organizacional, as variáveis demográficas e operacionais se mostram as mais relevantes, enquanto que variáveis que refletem características pessoais acabam sendo as menos importantes.

Tendo em vista os referenciais teóricos mencionados neste capítulo, a seguir será exposto como ocorrerá os procedimentos metodológicos deste estudo.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo será discorrido a respeito dos procedimentos metodológicos que compõem o caráter da pesquisa deste TCC. Nas próximas seções será tratado da caracterização da pesquisa, descrição da empresa que faz parte do estudo de caso e sobre o processo de entrevistas.

3.1 CARACTERIZAÇÃO DA PESQUISA

Tendo em vista a natureza do estudo deste trabalho bem como os objetivos definidos previamente, optou-se por uma abordagem de pesquisa de caráter qualitativo. O método de pesquisa qualitativa pode ser compreendido como sendo a tentativa de uma compreensão detalhada de significados e características expostas por indivíduos entrevistados, diferindo do método quantitativo no sentido de que não utiliza-se de um método estatístico para a análise de um problema (RICHARDS, 1985).

Nesse sentido, Minayo (p. 21, 2011) comenta sobre a pesquisa qualitativa:

[...] trabalha com o universo dos significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes. Esse conjunto de fenômenos humanos é entendido aqui como parte da realidade social, pois o ser humano se distingue não só por agir mas por pensar sobre o que faz e por interpretar suas ações dentro e a partir da realidade vivida e partilhada com seus semelhantes.

Além do caráter qualitativo e conseqüentemente exploratório deste estudo, será usado como método de pesquisa o estudo de caso. O estudo de caso é uma investigação empírica, em que um fenômeno contemporâneo é estudado em profundidade assumindo que o entendimento obtido a partir de um caso específico estudado engloba condições contextuais relevantes do fenômeno como um todo (YIN, 2015).

Nessa perspectiva, o estudo de caso como modalidade de pesquisa é entendido como um método que visa a investigação de um caso em específico que seja bem delimitado e contextualizado para que se possa realizar uma busca circunstanciada de informações, podendo focar sua análise tanto em casos únicos (uma unidade ou indivíduo) ou em casos múltiplos em que diversos estudos são conduzidos simultaneamente (VENTURA, 2007).

Assim sendo, o presente TCC conduz sua pesquisa qualitativa através do estudo de um caso individual de uma empresa de educação, a Pearson Education. A escolha da organização em questão como objeto de estudo parte de um anseio de entender organizações em que seus

respectivos propósitos representem os valores obtidos ao longo da jornada de educação e da vida do autor deste estudo, bem como a facilidade de acesso que o autor deste trabalho possuía para com o projeto da empresa, tendo em vista que este faz parte do quadro de funcionários da companhia. Aliado a isto, enxerga-se este estudo como oportunidade de estudar a segmentação de mercado de uma empresa global que não apenas é líder de seu mercado de atuação conforme será demonstrada na seção posterior, porém, é referência em diversos ramos que contemplam o mercado educacional tanto no Brasil quanto no mundo inteiro.

Posto isto, ao longo da próxima seção deste capítulo será discorrido as informações e características da organização em questão.

3.2 O CASO ESTUDADO: PEARSON EDUCATION

A empresa escolhida para realizar o estudo de caso, conforme comentado anteriormente, é a Pearson Education, mais especificamente, a Pearson Education do Brasil, com sua sede situada na cidade de São Paulo. As informações discriminadas a seguir foram extraídas diretamente do site da Pearson Brasil (2022).

A Pearson é considerada a maior empresa de aprendizagem do mundo, contendo mais de 22,5 mil colaboradores atuando em 70 países. A organização oferece conteúdo, avaliação e serviços digitais para estudantes, instituições educacionais, empresas, governos e outros parceiros ao redor do mundo, tendo como referência seu compromisso de ajudar as pessoas a desenvolver as habilidades que elas precisam para aumentar suas chances de empregabilidade e obter sucesso dentre as mudanças no mercado de trabalho, acreditando que, onde quer que a aprendizagem prospere, as pessoas também prosperarão.

No que se refere a sua atividade no Brasil, local onde atua desde a década de 1970, a Pearson entrega soluções completas e customizadas desde a educação infantil até a idade adulta através de seu vasto portfólio de linhas de negócios voltadas para educação básica, sendo referência com os sistemas de ensino COC by Pearson, Dom Bosco by Pearson e Atitude by Pearson; idiomas com as marcas Wizard by Pearson, Yázigi e Skill; ensino superior; formação corporativa e profissional; e um catálogo variado de livros, conteúdo, recursos tecnológicos e serviços para escolas e institutos de idiomas.

Conforme mencionado, uma das linhas de produto da empresa se refere ao catálogo de livros e material didático disponibilizado pela companhia, a linha de English Language Teaching - ELT - (tradução livre: Ensino de Língua Inglesa). Estes materiais, no Brasil,

geralmente têm seu foco voltado para escolas privadas, que adquirem o material didático da companhia a fim de lecionar Inglês nas referidas instituições.

Diante do exposto, o presente estudo de caso intenciona estudar um projeto de segmentação de mercado de escolas privadas que adquirem o material de ELT da Pearson Education. Nesse sentido, a próxima seção irá tratar do processo de entrevistas realizado para auxiliar na presente pesquisa.

3.3 PROCESSO DE ENTREVISTAS

As entrevistas foram realizadas de maneira semiestruturada virtualmente entre 18 de Outubro e 9 de Novembro de 2022 com os sujeitos da pesquisa sendo constituídos por três funcionários da organização pertencentes a três diferentes áreas de atuação. Estas foram conduzidas a partir de dois roteiros de entrevista elaborados para atender as cada um dos diferentes entrevistados e que podem ser encontradas na seção de “Apêndice A” e “Apêndice B” deste TCC. As entrevistas ainda tiveram uma duração média de aproximadamente 13 minutos, foram gravadas e posteriormente transcritas para a realização da análise dos dados.

A escolha dos participantes para a realização da entrevista foi feita de forma que as áreas mais envolvidas com a construção e utilização da solução estudada pudessem estar presentes e fornecer diferentes perspectivas e informações que contribuem para o atingimento dos diferentes objetivos que este trabalho pretende atingir. Sendo assim, foram entrevistados uma coordenadora pertencente a área de Research & Insights da companhia, área que elaborou a solução estudada por este TCC; um coordenador da área de planejamento comercial, área também envolvida com o projeto estudado; e uma gerente da área de ELT privado, área demandante da solução estudada por este TCC. Respectivamente, por questão de sigilo de identidade, os participantes serão identificados como E1 (entrevistada 1), E2 (entrevistado 2) e E3 (entrevistada 3). As informações referenciadas nesta seção podem ser visualizadas de acordo com o quadro 1:

Quadro 1 - Dados dos entrevistados e dados do corpus

Participante	Área de atuação	Duração da entrevista	Palavras transcritas	Roteiro utilizado
E1	Coordenadora de Research & Insights	17 minutos e 33 segundos	1792	Apêndice B
E2	Coordenador de planejamento comercial	9 minutos e 5 segundos	771	Apêndice A
E3	Gerente de ELT privado	12 minutos e 45 segundos	969	Apêndice A

Fonte: O autor.

As análises foram feitas mediante análise de conteúdo, a qual se incube de analisar dados qualitativos através de materiais provenientes de entrevistas ou observações e os categorizando em temas a fim de compreender o que está por trás dos discursos (SILVA; FOSSÁ, 2015). Sendo assim, a análise de conteúdo deste trabalho será realizada avaliando o material transcrito das entrevistas em comparação com a referência bibliográfica que embasa o TCC, intencionando, ao longo da análise, responder aos objetivos pré estabelecidos por este trabalho.

Diante do exposto, o próximo capítulo irá realizar a análise dos dados obtidos mediante as entrevistas realizadas.

4 ANÁLISE DOS DADOS

Neste capítulo serão discutidas e analisadas todas as informações obtidas através das entrevistas descritas anteriormente a fim de atingir os objetivos previamente definidos por este trabalho.

4.1 ENVOLVIMENTO COM O PROJETO

A entrevistada E1 coordena a equipe responsável pela inteligência de mercado e pesquisa dentro da Pearson, conseqüentemente, tornou-se a área a qual a solução de segmentação de mercado foi demandada, tendo em vista que, por característica da própria área em si, esta equipe atua como suporte não apenas para a área de ELT dentro da companhia, mas diversas outras linhas de negócio que podem vir a necessitar de apoio quando se fala de dados, pesquisas e inteligência de negócio.

O entrevistado E2 atua como coordenador da equipe de planejamento comercial, sendo assim, também atua dando suporte a equipe comercial que demanda o projeto de segmentação de clientes. Nesse sentido, o entrevistado descreve que sua área é responsável por todo apoio à equipe comercial, desde precificação, planejamento e administração de vendas até identificação de oportunidades, carteirização e distribuição da equipe.

Por outro lado, a entrevistada E3 atua gerenciando a área que demanda o projeto estudado neste trabalho. Nesse sentido, a entrevistada em questão descreve dentre suas principais atividades dentro da área de ELT a gestão da equipe para metas e resultados, participação na elaboração estratégica de ELT e também a relação com clientes, administrando carteiras visando metas e objetivos.

Assim sendo, tendo compreendido a relação de cada uma das respectivas áreas dos entrevistados com o projeto estudado e, conseqüentemente, com a área de ELT, é possível observar que ambos os participantes E1 e E2 atuam servindo de suporte para a área demandante do projeto, integrada pela entrevistada E3. Especialmente, é perceptível que a área da entrevistada E1 é a responsável pela construção do projeto e a área do entrevistado E2 por apoiar tudo que se refere a planejamento para a área da entrevistada E3 mediante a solução construída.

4.2 PROCESSO DECISÓRIO PARA SEGMENTAÇÃO DE CLIENTES

Esta seção irá relacionar, de acordo com a literatura exposta no referencial teórico e as informações obtidas através das entrevistas, como cada etapa do processo decisório se caracteriza dentro da segmentação de mercado na Pearson.

4.2.1 Identificação do problema

Conforme ressalta Maximiano (2000), é na etapa de identificação do problema que será compreendido quais de fato serão os obstáculos que impedem o tomador de decisão atingir seus respectivos objetivos. Sendo assim, diversos obstáculos foram identificados como problemas que posteriormente resultaram na implementação da segmentação de clientes.

Para que o problema pudesse ser identificado, primeiro era preciso esclarecer de fato qual o objetivo da equipe responsável pela área de ELT. Ao longo das entrevistas, foi possível perceber, mediante as diferentes perspectivas de negócio dos entrevistados, que haviam dois objetivos principais a serem alcançados. O primeiro, é encontrado na fala da E1, área ao qual foi demandada a solução de segmentação de mercado:

Então foi esse o contexto, que o time comercial de ELT tinha que conseguir novos clientes, e não assim, clientes novos em geral, mas clientes de fato em potencial, clientes que realmente vão ficar com a empresa, gastar um ticket médio relevante, então foi esse desafio que nos foi proposto (E1).

Assim, é possível identificar um dos objetivos como sendo, resumidamente, a prospecção de clientes e principalmente clientes em potencial. Além disso, outro objetivo, identificado também na fala da equipe comercial e de planejamento comercial, é a necessidade de não apenas prospectar clientes novos, porém, entender o perfil dessas escolas novas e das escolas que já eram clientes da companhia.

Todos esses aspectos vão de encontro ao que é referenciado por Ferreira (2000) em termos de segmentação de mercado. Fica evidente que a empresa tinha como objetivo a segmentação de mercado desde o início ao identificar a necessidade de perfilamento dos clientes e focar nos clientes que realmente valem a pena, conforme a E1 falou, clientes que vão gastar um ticket médio relevante.

Dessa forma, diante dos objetivos mencionados, diversas frustrações foram expostas, entre elas está a dificuldade de prospecção de clientes por parte da equipe de ELT, a falta de uma maior visão de mercado, o entendimento de qual o potencial que essa equipe tem ao

prospectar determinados clientes e a perda da noção de qual o perfil dos clientes que já estão na base da companhia.

4.2.2 Diagnóstico

Na etapa posterior a de identificação, inicia-se o diagnóstico do problema, ou seja, o tomador de decisão irá analisar os problemas e entender o que está acontecendo e o que pode acontecer, atribuindo também as causas que resultaram naquele problema. Nessa situação, caso o problema fuja da rotina do tomador de decisão, será necessário analisar todas as causas e informações relevantes acerca do problema (MAXIMIANO, 2000).

Sendo assim, principalmente devido ao fato de os problemas identificados não serem parte da rotina da organização, uma análise e identificação de todas as causas e informações com relação ao problema tiveram de ser realizadas. Uma das principais causas mapeadas foi a mudança de foco da organização que antes enxergava o segmento de ELT como uma fonte de renda secundária em detrimento de outra área educacional que recebia mais foco por parte da organização. Consequentemente com a venda dos sistemas de ensino responsáveis por essa outra área educacional, o olhar da empresa se engrandeceu para o segmento de ELT. Aliado a isto, pela carência de foco em ELT que ocorria antes da venda comentada, houve um descuido da gestão da informação de mercado e dos clientes que a empresa se relacionava e poderia vir a se relacionar no segmento de ELT. Essa questão é exposta na fala do E2:

A questão é que a gente tava com todo o banco de dados, toda gestão de informação em cima do ELT muito descuidada, e com isso a gente tinha perdido um pouquinho da noção de quem são os clientes da nossa base, como a gente atua nesses clientes (E2).

Outro fator apontado era que além do descuido da gestão de informação, esta, quando obtida, não possuía um refinamento profundo dos dados a qual demandava a equipe de ELT, impossibilitando a compreensão do que de fato as escolas mapeadas representavam dentro do cenário.

4.2.3 Geração e escolha de alternativas

Esta seção compreende as duas etapas finais na compreensão de processo decisório de Maximiano (2000). A primeira delas é a fase de geração de alternativas, em que serão

levantadas as possíveis soluções para resolver o problema, e a segunda é a fase de avaliação e seleção da alternativa a ser implementada. O autor propõe que as alternativas sejam comparadas e julgadas mediante critérios avaliativos, porém, na Pearson, isso ocorre de maneira diferente.

Diante dos problemas discriminados anteriormente, diversas alternativas foram levantadas, porém, devido à falta de informações de mercado por parte da empresa e a necessidade de soluções de prontidão, não houve um critério avaliativo dessas alternativas e escolha em paralelo delas, como propunha a teoria. As diversas alternativas geradas eram implementadas de maneira simultânea, havendo necessidade em um primeiro momento de múltiplas soluções para adereçar o problema. Também foi observado que, ao longo do tempo, as alternativas anteriormente implementadas acabavam sendo substituídas ou aprimoradas, ainda assim, não havendo comparação e julgamentos com critérios predefinidos como havia proposto Maximiano.

Inicialmente, no que se refere a aquisição de dados de mercado, algumas das alternativas ressaltadas pela E3 foram parcerias com associações educacionais, visitas de campo e experiência da equipe de vendas a fim de prover ao máximo a equipe com dados e informações sobre as escolas, porém, sem nenhum tipo de segmentação ou priorização de mercado.

Posteriormente, foi construído um banco de dados com informações do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), que era cruzado com dados de um fornecedor da empresa, dessa forma, auxiliando na obtenção de informações de mercado. A fim de priorizar os clientes, foi feita uma segmentação de mercado rasa com base em alguns critérios lógicos utilizando o banco de dados mencionado anteriormente, conforme menciona o E2:

Ambas foram implementadas, a gente teve um banco de dados de escolas do INEP, que foi utilizado junto com dados da Geofusion para montar um cenário de potencial de mercado, quais escolas existem no mercado. Em cima disso foi aplicado uma segmentação bem rasa com base em uma série de critérios, se não me engano tinha de critérios o tamanho da escola, a estrutura terminológica que ela tinha e pouco mais (E2).

Por conseguinte, é perceptível que alguns aspectos dos problemas mapeados puderam ser minimamente atendidos pelas alternativas levantadas, porém, havendo considerável margem para aprimoramento dessas alternativas. Nesse aspecto é compreensível que os problemas de prospecção e visão de mercado, através do banco de dados, conseguiram ser

parcialmente solucionados, porém, conforme a própria fala do E2, aspectos como o mapeamento do perfil e priorização dos clientes foi realizado de maneira rasa, carecendo o potencial que a organização poderia atingir com a solução.

4.3 APOIO DA CIÊNCIA DE DADOS NO PROCESSO DECISÓRIO

Tendo em vista que nem todos os problemas puderam ser solucionados da melhor maneira, partiu do time de Research & Insights a construção de uma segmentação de mercado utilizando Ciência de Dados. A solução utilizada foi o aprimoramento da segmentação de mercado existente através de uma técnica comumente referenciada pelos entrevistados como clusterização.

Conforme visto nas referências bibliográficas, esta técnica de Ciência de Dados de agrupamento ou *clustering* se responsabiliza por encontrar grupos em uma determinada base de dados de acordo com as similaridades dos dados (SILVA; FERRARI, 2016). Posto isto, no caso da Pearson, a técnica buscaria entender em quais grupos os seus clientes de ELT se enquadram naturalmente, de forma que os clientes em um grupo sejam semelhantes entre si dada determinadas características e diferentes dos clientes presentes nos demais grupos. Dessa forma, através dos grupos construídos com a técnica de Ciência de Dados, a clusterização evoluiu para a segmentação de mercado proposta por Kotler (2012), em que dado grupos de consumidores com necessidades homogêneas, a empresa pode generalizar suas necessidades e desejos, identificando em quais segmentos ela pode atuar da melhor forma.

Na perspectiva da E1, a clusterização apoia todas as dificuldades que compõem o projeto, mas principalmente entrega uma linha de raciocínio muito mais coerente para o time comercial no momento de entender as alternativas de mercado que eles podem atender:

Primeiramente, a clusterização apoiou a dificuldade como um todo, sabe? Mas eu ressaltaria que a clusterização ela ajudou sobretudo no que eu chamaria de organização da linha de raciocínio, porque como eu te falei, quando essa dificuldade ela foi apresentada a nós, foi uma questão de “onde nós podemos achar?”, “onde estão as escolas?”, “onde estão as oportunidades?”, “como nós vamos saber que uma escola pode ser nossa potencial cliente mais do que outra escola?”. Então acho que a primeira coisa que eu citaria é na organização do racional, lógica e na organização de onde nós vamos focar os nossos esforços (E1).

A entrevistada ainda resalta a capacidade de mapeamento de perfil e potencial que a segmentação proporciona. Isso evidencia a fala de Ferreira (2000) quando o autor comenta

sobre a capacidade das empresas, através da segmentação de mercado, construírem produtos sob medida para os segmentos e atendê-los com maior atenção e eficácia.

Além disso, para a entrevistada, o banco de dados existente anteriormente, com a segmentação rasa e simples que foi comentada, não era totalmente irrelevante, porém, os ganhos que a Ciência de Dados trouxe para o projeto em termos de compreensão de potencial de mercado e mapeamento do perfil das escolas que a Pearson atende e pode atender foram muito relevantes.

Outro ponto ressaltado pela E1 é a perda do viés humano no mapeamento do perfil das escolas e a capacidade de detalhamento dos grupos dessas escolas em termos de similaridade, algo que, conforme comentado por Silva e Ferrari (2016), faz parte do trabalho de agrupamento da Ciência de Dados:

E outra coisa, a clusterização é feita via Machine Learning, então é interessante a gente saber que é uma abordagem que perde o viés humano na construção, antes o que a gente tinha a planilha mas que não trazia clusters com tantos detalhes sobre os grupos das escolas, as similaridades que elas têm, e tinha esse viés humano que continham premissas que nós assumimos, e com Machine Learning parece que fica algo mais técnico, do tipo “o algoritmo que calculou”. Então vemos muito potencial desde a primeira versão que foi apresentada da clusterização (E1).

O E2 também ressalta a eficiência do embasamento que a Ciência de Dados proporciona. Para o entrevistado, por mais que as soluções anteriores tivessem auxiliado a equipe comercial em momentos em que não haviam outras alternativas, a solução de segmentação proposta anteriormente era mais embasada no “achismo”, algo que, do seu ponto de vista, ocorre de maneira oposta com o uso de Ciência de Dados.

Posto isto, é possível notar aspectos de apoio da Ciência de Dados no que é referido por Diván (2017) como tomada de decisão orientada a dados, em que a tomada de decisão se baseia em dados ao invés de puramente pela intuição, um ponto que se mostra bastante destacado pelos entrevistados.

Além disso, é possível notar na fala do E2 os problemas específicos que a solução proposta consegue atender:

Tanto gestão de carteira quanto prospecção de clientes. Por uma lado tem uma visão de tamanho de mercado por trás, que me ajuda a ter uma noção de como que eu tenho distribuído o time e ao mesmo tempo também tenho a visão de como que eu vou dar a carteira para o time e dar os objetivos para eles, ataca primeiro esse grupo ou outro, utiliza essas armas. As duas vêm sendo apoiadas com a clusterização (E2).

Assim como já referenciado anteriormente na fala da E1, o E2 ressalta novamente a algo visto na literatura de Ferreira (2000), em que a empresa pode, através da segmentação de mercado, focar nos clientes que realmente valem a pena, ou como o E2 comenta, atacar de antemão um determinado grupo.

A gerente da área demandante, a E3, também ressaltou diversos ganhos com a alternativa de Ciência de Dados:

Quando eu olhei essa versão eu vi que ela me trouxe diversos dados super relevantes e que eles me mostram o potencial da escola mas de uma forma muito mais detalhada, então assim, qual é o potencial de compra dessa escola, qual é a relevância dela dentro do cenário. Então, pra mim, quando eu vi foi algo inédito, a quantidade de salas de aula, o segmento que essa escola trabalha, a quantidade de equipamentos. Você tem diversos dados ali que me trazem muito mais do que simplesmente ela ser uma escola numerosa ou não, então eu tenho certeza que esse projeto ele norteia muito mais a forma de atuação e a própria decisão estratégica “como iremos atuar daqui pra frente? Em quem queremos ir e aonde queremos ir?” acho que ajuda demais (E3).

Conforme visto no trecho anterior, um dos principais ganhos ressaltados é o direcionamento, que, para a entrevistada, fica muito mais evidente fatores que denotam a real relevância da escola e seu potencial de compra para a Pearson, apoiando diretamente a estratégia como um todo da companhia. Nesse contexto, a entrevistada ainda ressalta um ponto que também foi levantado pelos dois entrevistados anteriores, que é a canalização de recursos e equipe, no lugar certo e na hora certa, em que a companhia pode focar nos clientes que realmente valem a pena.

5 CONSIDERAÇÕES FINAIS

O processo decisório é, sem dúvida, dentre todos os principais componentes que fazem parte do dia a dia de uma organização, o principal sistema que guia a estratégia e o curso para o atingimento de seus objetivos. Paralelo a isto, é irrefutável o papel que os dados têm e devem ter como suporte a este processo e a tomada de decisão como um todo.

A partir da fundamentação teórica e da análise das entrevistas foi possível responder ao problema de pesquisa que inspirou este TCC, expresso pela pergunta: “*Como a Ciência de Dados apoia o processo decisório para segmentação de mercado na Pearson Education?*”.

Conseqüentemente, em função do problema, surgiu o objetivo geral “*Analisar como a Ciência de Dados apoia o processo decisório para segmentação de mercado na Pearson Education*” o qual foi abordado através de objetivos específicos.

O primeiro objetivo específico é descrever como ocorre o processo decisório da Pearson Education em um projeto de segmentação de mercado. Posto isto, foi possível perceber que após a identificação dos potenciais problemas, devido ao problema fugir da rotina organizacional, foi necessário, dentro da etapa de diagnóstico, uma análise de todas as causas e informações relevantes acerca dos problemas. Porém, por mais que estas duas etapas tivessem sido mapeadas conforme identificado na literatura, o mesmo não podia ser dito das etapas de geração e escolha de alternativas. Nestas últimas duas etapas, não havia necessariamente comparação entre as alternativas ou algum tipo de ponderação no momento da implementação. O problema já carregava consigo alternativas que eram implementadas, e, caso surgisse alguma nova alternativa ou aprimoramento da já existente, esta era implementada ao longo do tempo, substituindo ou aperfeiçoando a implementação já existente.

O segundo objetivo específico é compreender em quais etapas do processo decisório da Pearson Education a Ciência de Dados é aplicada em um projeto de segmentação de mercado. Portanto, ficou evidente que diante de todas as etapas percorridas que compõem o processo decisório para segmentação de mercado, nota-se que a Ciência de Dados apoia diretamente na etapa de geração de alternativas, surgindo não apenas como uma alternativa para solucionar os problemas que serviam de obstáculo para os objetivos da organização, mas também como melhoria e complementação das alternativas anteriormente implementadas. Além disso, a Ciência de Dados surge como uma alternativa que traz consigo um embasamento para a tomada de decisão organizacional muito mais sólida, proporcionando guiar as decisões através de dados ao invés da intuição.

Algumas das limitações que elencaram este trabalho é o enfoque em apenas um projeto da organização referida, conseqüentemente, restringindo os diferentes aspectos e nuances que o processo decisório possa ter ao longo dos diferentes níveis de tomada de decisão dentro da organização. Por conseguinte, perde-se também a capacidade de compreender se em outros tipos de projetos a Ciência de Dados pode apoiar outras etapas do processo decisório além da etapa de geração de alternativas, conforme foi evidenciado neste trabalho.

Sendo assim, como indicações para futuras pesquisas sugere-se o aprofundamento do estudo em mais projetos e trabalhos da organização que não se restrinjam apenas a um projeto de segmentação de mercado, dessa forma, capturando diferentes formas e características que o processo decisório possa assumir e a Ciência de Dados possa apoiar, bem como explorar o processo decisório de outros tipos de organizações do mesmo mercado que a organização trabalhada neste TCC.

REFERÊNCIAS

AMARAL, Fernando. **Introdução à ciência de dados: mineração de dados e big data**. Alta Books Editora, 2016.

BARROS, Ricardo Paes de; MENDONÇA, Rosane Silva Pinto de. **Investimentos em educação e desenvolvimento econômico**. 1997.

BISPO, Carlos Alberto Ferreira; CAZARINI, Edson Walmir. **A evolução do processo decisório**. CD-ROM) In: ENCONTRO NACIONAL DA ENGENHARIA DA PRODUÇÃO, v. 18, 1998.

CHIAVENATO, Idalberto. **Introdução à teoria geral da administração**. 9. ed. Barueri - SP: Manole Ltda, 2014.

CONWAY, Drew. **The data science venn diagram** (2010). Disponível em: <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram> Acesso em: 02 de Outubro de 2022.

DIVÁN, Mario José. **Data-driven decision making**. In: 2017 international conference on Infocom technologies and unmanned systems (trends and future directions)(ICTUS). IEEE, 2017. p. 50-56.

SILVA, Leandro Nunes De Castro; FERRARI, Daniel Gomes. **Introdução a mineração de dados**. Saraiva Educação SA, 2016.

DE OLIVEIRA MORITZ, Gilberto; PEREIRA, Maurício Fernandes. **Processo decisório**. 2015.

FAWCETT, Tom; PROVOST, Foster. **Data Science para Negócios: O que você precisa saber sobre mineração de dados e pensamento analítico de dados**. Alta Books Editora, 2018.

FERREIRA, Francis Haime Giacomelli. **Segmentação de mercado**. Biblioteca temática do empreendedor, 2000.

KOTLER, Philip. Administração de marketing: análise, planejamento, implementação e controle. **Administração de marketing: análise, planejamento, implementação e controle**. 14. ed. São Paulo - SP: Pearson, 2012.

LACOMBE, Francisco; HEILBORN, Gilberto. **Administração: Princípios e Tendências**. 2. ed. São Paulo: Saraiva, 2008.

LAURINDO, Fernando José Barbin et al. **O papel da tecnologia da informação (TI) na estratégia das organizações**. Gestão & Produção, v. 8, p. 160-179, 2001.

MAXIMIANO, Antonio Cesar Amaru. **Introdução à Administração**. 5a edição. Editora Atlas, 2000.

MCDONALD, Malcolm; DUNBAR, Ian.. **Market segmentation: How to do it and how to profit from it**. 4a edição. John Wiley & Sons, 2013.

MINAYO, Maria Cecília De Souza; DESLANDES, Suely Ferreira; GOMES, Romeu. **Pesquisa social: teoria, método e criatividade**. Editora Vozes Limitada, 2011.

PEARSON. Disponível em: <<https://br.pearson.com/>>. Acesso em: 5 de Setembro de 2022.

RAUTENBERG, Sandro; DO CARMO, Paulo Ricardo Viviurka. **Big data e ciência de dados: complementariedade conceitual no processo de tomada de decisão**. Brazilian Journal of Information Science: research trends, v. 13, n. 1, p. 56-67, 2019.

RICHARDSON, Roberto Jarry et al. **Pesquisa social: métodos e técnicas**. São Paulo: atlas, 1985.

SILVA, Andressa Hennig; FOSSÁ, Maria Ivete Trevisan. **Análise de conteúdo: exemplo de aplicação da técnica para análise de dados qualitativos**. Qualitas revista eletrônica, v. 16, n. 1, 2015.

VENTURA, Magda Maria. **O estudo de caso como modalidade de pesquisa**. Revista SoCERJ, v. 20, n. 5, p. 383-386, 2007.

YIN, Robert K. **Estudo de Caso-: Planejamento e métodos**. Bookman editora, 2015.

APÊNDICE A - ROTEIRO DE ENTREVISTA DEMANDANTES

- 1 - Antes de começarmos nossa entrevista, poderia informar qual seria seu cargo e respectiva área de atuação na empresa?
- 2 - Poderia descrever a sua função e a sua relação com o segmento de ELT?
- 3 - Quais as dificuldades iniciais que você e seu time tinham que levaram a implementação do projeto de segmentação de clientes de ELT?
- 4 - Como vocês imaginavam que essa dificuldade poderia ser sanada?
- 5 - Quais as soluções embrionárias que foram tomadas dentro da Pearson para solucionar as dificuldades mencionadas?
- 6 - A partir desse ano, usando Ciência de Dados, começou a ser implementada uma clusterização para auxiliar o processo de segmentação de clientes. O que você achou da primeira versão de clusters de clientes de ELT apresentada?
- 7 - Considerando a segmentação de clientes de ELT, de que forma você diria que a clusterização trouxe apoio ao projeto?
- 8 - Mais especificamente, dentre as dificuldades descritas anteriormente qual delas você diria que a clusterização apoiou e de que forma?

APÊNDICE B - ROTEIRO DE ENTREVISTA DEMANDADO

- 1 - Antes de começarmos nossa entrevista, poderia informar qual seria seu cargo e respectiva área de atuação na empresa?
- 2 - Poderia descrever a sua função e a sua relação com o segmento de ELT?
- 3 - Quais as dificuldades iniciais que você e seu time foram informados ou foram passados que levaram a implementação do projeto de segmentação de clientes de ELT?
- 4 - Como você e sua equipe imaginavam que essa dificuldade poderia ser sanada? Vocês já possuíam as sugestões ou contribuíram para o brainstorming do tema?
- 5 - Quais as soluções embrionárias que foram tomadas pela sua área ou pela área demandante para solucionar as dificuldades mencionadas?
- 6 - A partir desse ano, usando Ciência de Dados, começou a ser implementada uma clusterização para auxiliar o processo de segmentação de clientes. O que você achou da primeira versão de clusters de clientes de ELT apresentada?
- 7 - Considerando a segmentação de clientes de ELT, de que forma você diria que a clusterização trouxe apoio ao projeto?
- 8 - Mais especificamente, dentre as dificuldades descritas anteriormente qual delas você diria que a clusterização apoiou e de que forma?