



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE HUMANIDADES
UNIDADE ACADÊMICA DE ADMINISTRAÇÃO E CONTABILIDADE
CURSO DE BACHARELADO EM ADMINISTRAÇÃO**

HELLEN KAROLINE ALVES DE SOUSA

**ANÁLISE EXPLORATÓRIA DE DADOS NO *E-COMMERCE* DO
VAREJO SUPERMERCADISTA**

**CAMPINA GRANDE - PB
2021**

HELLEN KAROLINE ALVES DE SOUSA

**ANÁLISE EXPLORATÓRIA DE DADOS NO *E-COMMERCE* DO
VAREJO SUPERMERCADISTA**

**Trabalho de Conclusão de Curso
(Artigo) apresentado ao Curso de
Administração do Centro de
Humanidades da Universidade Federal
de Campina Grande, como requisito
parcial para obtenção do título de
Bacharela em Administração.**

Orientador: Professor Dr. Edvan Cruz Aguiar.

CAMPINA GRANDE - PB

2021



S725a Sousa, Hellen Karoline Alves de.
Análise exploratória de dados no e-commerce do varejo supermercadista. / Hellen Karoline Alves de Sousa. - 2021.

25 f.

Orientador: Professor Dr. Edvan Cruz Aguiar.
Trabalho de Conclusão de Curso (Artigo) - Universidade Federal de Campina Grande; Centro de Humanidades; Curso de Bacharelado em Administração.

1. Varejo supermercadista. 2. E-commerce. 3. Análise exploratória de dados. 4. Estatística descritiva e inferencial. I. Aguiar, Edvan Cruz. II. Título.

CDU: 658(045)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

HELLEN KAROLINE ALVES DE SOUSA

**ANÁLISE EXPLORATÓRIA DE DADOS NO *E-COMMERCE* DO
VAREJO SUPERMERCADISTA**

**Trabalho de Conclusão de Curso
(Artigo) apresentado ao Curso de
Administração do Centro de
Humanidades da Universidade Federal
de Campina Grande, como requisito
parcial para obtenção do título de
Bacharela em Administração.**

BANCA EXAMINADORA:

**Professor Dr. Edvan Cruz Aguiar
Orientador – UAAC/CH/UFCG**

**Professora Dra. Ana Cecília Feitosa Vasconcelos
Examinadora I – UAAC/CH/UFCG**

**Mestra Manoela Costa Policarpo
Examinadora II**

Trabalho aprovado em: 2021.

CAMPINA GRANDE - PB

Análise exploratória de dados no *e-commerce* do varejo supermercadista

Hellen Karoline Alves de Sousa¹
Edvan Cruz Aguiar, Dr.²

RESUMO

O presente artigo teve por objetivo analisar os registros de compras produzidos no *e-commerce* de um varejista supermercadista implementado durante a pandemia, a fim de aprofundar e compreender o perfil dos clientes. A partir da revisão de literatura envolvendo a temática de inovações no varejo, foi realizado um estudo de abordagem quantitativa com caráter descritivo. O método adotado consistiu na extração de dados de fontes de dados do *e-commerce*. Enquanto método de análise, fez-se uso de estatística descritivas e inferenciais a partir da abordagem exploratória de dados foram conduzidos procedimentos de *descriptive analytics* gerando *insights* a partir dos dados obtidos entre abril de 2020 e setembro de 2021. As análises foram feitas com auxílio da ferramenta *Microsoft Power BI* e evidenciaram o potencial de crescimento se adotado as preferências dos consumidores e destaca-las, seja a forma de pagamento, bairro ou classe social. Esta pesquisa contribui ao passo em que compreender o perfil do cliente, permite uma tomada de decisão mais assertiva para um melhor atendimento e prestação de serviço.

Palavras-chave: Varejo; Análise exploratória de dados; *E-commerce*.

ABSTRACT

The aim of this article was to analyze the purchase records obtained in the e-commerce of a supermarket retailer implemented during a pandemic, in order to deepen and understand the profile of customers. From the literature review involving the theme of innovations in retail, a study with a quantitative approach with a descriptive character was carried out. The method adopted consists of extracting data from e-commerce data sources. As a method of analysis, descriptive and inferential statistics were used from the exploratory data approach, descriptive analysis procedures were conducted to generate insights from the data obtained between April 2020 and September 2021. The analyzes were carried out with the aid of Microsoft Power BI tool and show the potential for growth if adopted as a consumer preference and highlights them, whether in the form of payment, neighborhood or social class. This research contributes to the understanding of the customer's profile, allowing for a more assertive decision-making for a better attendance and service provision.

Keywords: Retail; Exploratory data analysis; E-commerce.

¹ Graduanda em Administração, UFCG. Email: hellenkarolne.adm@gmail.com

² Doutor em Administração, UFCG. Email: edvan.ed@gmail.com

1. Introdução

O varejo supermercadista é um dos setores da economia que sofreu, relativamente, menos com a pandemia causada pelo COVID -19. De acordo com os dados da Associação Brasileira de Supermercados (2021) o ano de 2020 foi positivo para o setor supermercadista, que acumulou alta real de 9,36% de janeiro a dezembro em relação ao mesmo período no ano de 2019. Com base estudos é projetado um crescimento de 4,5% para o ano de 2021.

O setor tem sido impelido a inovar nos seus processos e serviços tendo em vista que houve uma mudança no comportamento do consumidor, em função do cenário pandêmico que propiciou um crescimento no comércio eletrônico no país. Assim o faturamento com vendas online cresceu cerca 4,7% no primeiro semestre de 2020, no qual cerca de 7,4 milhões de brasileiros realizaram a primeira compra online (INGIZZA, 2020).

Com o desenvolvimento do *e-commerce* e de novas formas de consumo, se faz necessário entender o cliente, pois a satisfação afeta a lucratividade. Os varejistas reconhecem que os dados orientam a tomada de decisão (KOWALCZK; BUXMANN, 2015). Com isso, a exploração de dados gerados pelos clientes deve ser analisada, extraindo informações que facilitam a tomada de decisão e geram *insights*.

O objetivo da pesquisa é analisar os registros de compras produzidos no *e-commerce* de um varejista supermercadista implementado durante a pandemia, a fim de aprofundar e compreender o perfil dos clientes.

Para o alcance do objetivo foi realizada a integração dos dados que envolve a combinação de dados que residem em diferentes fontes (LENZERINI, 2002). Em seguida foi realizada a limpeza dos dados para correção ou remoção de erros e inconsistências (RAHM, 2000) para então executar a relação entre as variáveis e por fim a validação dos dados.

A relevância da pesquisa se concentra em, por meio da análise dos dados produzidos no *e-commerce* implementado durante a pandemia por um varejista supermercadista, ter uma visualização clara da atuação do setor e sua performance, identificando áreas a serem desenvolvidas, possibilitando a implementação de práticas a fim de gerar vendas e alavancar seus resultados,

2. Referencial Teórico

2.1 Inovação no Varejo

O varejo pode ser compreendido como uma ação de comercializar um determinado produto para clientes a fim de facilitar a compra, sendo intermediada por um fornecedor que disponibiliza uma variedade de produtos. Uma clássica definição do varejo corresponde na comercialização de bens ou serviços direcionada aos consumidores finais, para uso pessoal e não comercial (PARENTE; BARKI, 2014).

O varejo sofreu alterações e evoluções na sua definição e relação com o cliente, com o crescimento do uso da internet sua influência comercial tornou-se altamente visível e em destaque no setor varejista por Zhang et al. (2014). A era da internet possibilitou diversas escolhas, onde os varejistas estão online em qualquer lugar do mundo, aumentando as alternativas para os consumidores (MUKHERJEE; NATH, 2007).

O varejo 4.0 possui diversas tecnologias com variadas funções capazes de proporcionar maior agilidade ou até substituir funções operacionais realizadas por pessoas. *Big Data*, *Omnichannel*, *Customer relationship management* (CRM) são algumas das principais tecnologias. Essas tecnologias que possuem rastreamento permitem que as empresas realizem análises de dados e possibilitam maior eficácia (Dekimpe e Hanssens 2000). Uma empresa frequentemente coleta dados estruturados e não estruturados dos clientes de várias fontes, neste sentido o gerenciamento de dados do cliente é importante para que possuam qualidade e gerem informações relevantes. (Nam et al. , 2019).

2.2 Análise de Marketing

No mundo contemporâneo o varejo é desafiado a oferecer experiências memoráveis e únicas para os consumidores (ROCHA, 2018). Aperfeiçoando as formas de entregar seus produtos aos clientes melhorando a sua jornada de consumo e agregando valor ao atendimento com a finalidade de garantir a satisfação dos clientes.

De acordo Cervieri Junior et al. (2015) a tecnologia será possível interpretar os desejos da demanda e suas necessidades permitindo às empresas varejistas uma atuação mais assertiva e estratégica. No marketing o comércio móvel é considerado uma função básica, uma alternativa de vendas, essa ferramenta possibilita uma conexão entre o consumidor e a marca (MISHRA, 2017).

Os varejistas armazenam dados sobre seus clientes de diversos tipos, como demográficos ou dados que indicam se os clientes mudaram das lojas físicas para web, quais produtos colocam em suas cestas ou experimentam nos provedores, quais produtos compram e etc. Sendo dados como volume, variedade e velocidade que ultrapassam significativamente a capacidade do homem de análise (CHANG; KAUFFMAN; KWON, 2014). Para auxiliar nas análises preditivas a big data é importante referente a captura, armazenagem, transferência, compartilhamento de dados, pesquisa e análise de dados (CHEN; CHIANG E STOREY, 2012; DUAN; XIONG, 2015)

2.3 Segmentação

Podemos definir a segmentação de cliente como o processo de separação entre clientes heterogêneos e homogêneos que possuem algo em comum, esses grupos facilitam para empresa lidar com a diversidade de clientes e suas diversas preferências de maneira mais eficiente (HONG; KIM, 2012).

Existe um esforço para identificar diferentes grupos de clientes que visitam as lojas, compreendendo as necessidades específicas e preferenciais dos segmentos para disponibilizar serviços adequados e satisfazê-los como também realizar adaptações no marketing (BOONE; ROEHM, 2002).

A análise de cesta de compras é uma associação entre os itens e produtos que um determinado cliente adquire no período que está realizando a visita, esse processo é um método de mineração de dados que analisa a transferência por meio dos bancos de dados observando quais os itens são constantemente adquiridos em conjunto de acordo com Agrawal et al. (1993), Srikant e Agrawal (1995). Para a realização da análise de cestas de compras são utilizados dados de vendas e mix de categorias.

Os pesquisadores utilizam dados da venda no nível do cliente a fim de segmentar os compradores e examinar seu comportamento de compra, como exemplo examinam volumes de vendas, frequência de visitas, mix de produtos ou categorias de produtos com base em todo o histórico de compras, seja na loja física ou web por Aeron et al. (2012), Chen et al. (2009), Cheng e Chen (2009) e Khaj- VAND et al. (2011). O estudo de Aeron et al. (2012) utilizou a base de clientes para estimar o valor da vida útil do cliente (CLV), sendo útil para personalização de produtos e serviços.

Grandes varejistas como Tesco, Metro e Walmart identificaram a necessidade de tomada de decisão baseada em dados. Usam ferramentas analíticas para obter vantagem

competitiva em áreas como marketing com venda cruzada, cliente segmentação, análise de comportamento na loja e experiência multicanal por Comissão Europeia Sion (2014).

3. Métodos de Pesquisa

O capítulo apresenta os procedimentos metodológicos adotados para o alcance da realização da pesquisa. Sendo fracionados nas seguintes seções: Natureza da pesquisa, procedimentos da coleta, procedimentos da análise de dados.

3.1 Natureza da pesquisa

O presente trabalho assume uma abordagem quantitativa com caráter descritivo, uma vez que tem como objetivo compreender e analisar o registro de compras de um e-commerce implementado durante a pandemia por um varejo no segmento supermercadista. De acordo com (CRESWELL, 2007) a pesquisa quantitativa refere-se a coleta de dados com público alvo assim como sua análise e interpretação.

A pesquisa quantitativa é caracterizada por buscar a validação das hipóteses mediante a utilização de dados estruturados e analisados por meios estatísticos (MATTAR, 2001). O estudo é descritivo, pois possui como principal função descrever característica de um determinado fenômeno ou população (MALHOTRA, 2005). A pesquisa pretende reconhecer o perfil do consumidor do e-commerce do varejo supermercadista.

Para realização da pesquisa a fonte de dados é o banco de dados do e-commerce de um varejo supermercadista. Com o objetivo de produzir dados estatísticos, informações quantitativas ou numéricas de aspectos encontrados no grupo estudado (FOWLER; COSENZA, 2009).

3.2 Procedimentos de coleta

O público-alvo da pesquisa são consumidores do e-commerce de um varejista supermercadista. O processo de KDD (*Knowledge Discovery in Databases*), ou descoberta de conhecimento em base de dados, tem por objetivo extrair e transformar as informações em uma estrutura compreensível (MÁRQUEZ-VERA et al., 2016). Desse modo, técnicas e abordagens como mineração de dados *machine learning* fizeram parte deste processo que inclui áreas como estatística, análise e visualização de dados.

O modelo proposto de *machine learning* analisa o comportamento dos consumidores com base na fonte de dados adotada. Assumindo uma abordagem quantitativa com caráter preditivo (SHARDA; DELEN; TURBAN, 2019). Neste sentido, em termos metodológicos, este projeto fará uso do Método CRISP-DM (*Cross Industry Standard Process for Data Mining*), um dos processos mais utilizados para a Mineração de Dados (DRLIK; MUNK, 2018).

Para realização da pesquisa e alcance do objetivo foram realizados alguns passos sendo um dos principais a construção da base de dados, que foram extraídos através da permissão de uma varejista supermercadista ao banco de dados dos registros de compras do setor do *e-commerce* junto ao setor do T.I (tecnologia da informação) da empresa foi realizada a extração das informações relevantes para pesquisa, criando tabelas no programa excel, seguido da limpeza dos dados e aplicação na ferramenta *microsoft power BI*.

3.3 Procedimentos de análise dos dados

Através da criação de modelos preditivos foi possível a descoberta de conhecimento alcançando por meio desse método atingir o objetivo da pesquisa. Desse modo, o passo seguinte se configura na construção da base de dados.

As fontes de dados fornecidas possuem tabelas que são referentes a pedidos, clientes, itens pedidos e produtos com linhas e colunas com diversas informações. A tabela pedidos possui os seguintes registros: bairro, e-mail, forma de pagamento, modalidade de entrega, número de pedido, valor faturado, valor total, valor de venda, data de entrega, data de pedido, status do pedido e quantidade de pagamento. A tabela clientes possuem os seguintes registros: e-mail, código do cliente, e-mail e sexo. A tabela itens pedidos possui os seguintes registros: código do produto, número do preço, quantidade pedida e valor. A tabela produtos possui os seguintes registros: código do produto, categoria do produto, nome do produto, status do produto e valor.

Ancorada no Método CRISP-DM (*Cross Industry Standard Process for Data Mining*), foi realizada a construção da base de dados nas seguintes etapas 1 (*Business Understanding*), 2 (*Data Understanding*) e 3 (*Data Prepatation*) do CRISP-DM. Para as etapas de modelagem (*Modeling*) e avaliação dos modelos de *machine learning* (*Evaluation*), será feito uso da Linguagem de Programação *Pyhton* (*version 3.8.5*) com apoio do *scikit-learn*, biblioteca de aprendizado de máquina de código aberto (McKINNEY, 2018; HARRISON, 2019; GÉRON, 2019).

Para Jonh W. Tukey (1997) a análise exploratória de dados (EDA) é um trabalho de detetive, na qual o pesquisador examina os dados sem qualquer pré-concepção de ideias para descobrir o que os dados dizem sobre os fenômenos em estudo. Faz uma comparação com análise de confirmação de dados (CDA) que se preocupa com a estática, teste de hipótese, intervalos, intervalo de confiança etc.

4 Análise e discussão dos resultados

O capítulo apresenta e discute os resultados obtidos por meio da fonte de dados gerados no *e-commerce* de um varejo supermercadista. Sendo ele dividido nas seguintes seções: Caracterização geral da base de dados e análise da exploratória dos dados.

4.1 Caracterização geral da base de dados

No período compreendido entre 14 de abril de 2020 a 22 de setembro de 2021 totalizando 17 meses, foram registrados 8.135 pedidos (vendas pelo *e-commerce* da empresa), gerando um faturamento total de R\$2.821.080,04 (Dois milhões, oitocentos e vinte e um mil, oitenta reais e quatro centavos.).

Estes valores foram gerados por um total de 3.350 clientes cadastrados na plataforma da empresa ao longo do período supracitado, sendo 64,93% (2.175) Mulheres, 34,78% (1.162) homens e 0,3% (10) não informaram o Gênero.

Com relação aos produtos adquiridos, das 8.135 vendas concretizadas (faturadas), possuem 19.461 produtos diferentes. Como o estudo foi realizado no varejo supermercadista que possui um vasto mix de produtos, com categorias e marcas diferentes esse valor é considerado adequado. A participação do comércio eletrônico nas vendas do varejo brasileiro antes da pandemia era de 6%, passou para 11% e o nível atual é de 7% sendo previsto uma mudança permanente de aproximadamente 1,3% (E-COMMERCE BRASIL, 2021). Sendo assim é uma tendência que promove o aumento da venda de produtos, pois oferece inúmeras vantagens como pagamento online e entrega a domicílio facilitando a jornada do consumidor e possibilitando maior comodidade e segurança.

Desse modo, ao menos uma compra foi realizada pelo site da empresa dentre os 57 (cinquenta e sete) bairros e zona rural de Campina Grande-PB. Destes, destacam-se os bairros Catolé (14,05%), Jardim Tavares (8,49%), Centro (8,31%), Alto Branco (7,64%), Mirante (7,54%), Universitário (5,92%) e Prata (5,02%) enquanto aqueles que mais contribuíram para

o faturamento neste período. Juntos eles representam 56,97% do faturamento. Ressalta ainda que 3,8% é Zona Rural (Distritos e Cidades vizinhas, como Lagoa Seca). Os cinco bairros citados anteriormente possuem maior evidência devido ao alto poder aquisitivo dos moradores fazendo parte do perfil dos consumidores, como também maior proximidade com o supermercado fazendo com que receba as compras de maneira mais rápida.

No que diz respeito às modalidades de entrega e formas de pagamento, o sistema de gerenciamento de bancos de dados registrou as seguintes informações:

Tabela 1: Modalidade de entrega, total de pedidos e faturamento total

Modalidade de Entrega	Total de Pedidos	Faturamento Total
Entrega Agendada	3.147	R\$ 587.071,72
Convencional Grátis	1.946	R\$ 981.242,03
Entrega Agendada Grátis	1.880	R\$ 1.000.970,33
Retirada na Loja	868	R\$ 209.676,40
Convencional	294	R\$ 42.119,56
Total	8.135	R\$ 2.821.080,04

Fonte: Elaborado pela autora (2021).

Na tabela 1 é demonstrado as modalidades de entrega, uma das opções é a entrega agendada no qual o cliente agenda no site a data e horário da entrega, podendo ser de forma gratuita ou cobrada, dependendo do valor total do pedido. A entrega convencional é escolhida apenas a data e o turno da entrega, sendo manhã ou tarde, podendo ser gratuita ou cobrada dependendo do valor total do pedido. Por fim, a retirada em loja é uma maneira que o cliente vai até a loja pegar o seu pedido, podendo realizar o pagamento antes ou no ato da retirada.

Também foi apresentado a quantidade total de pedidos de cada tipo de modalidade de entrega e seu faturamento. A entrega agendada possui maior número de pedidos, já a opção entrega agendada grátis possui maior faturamento, contribuindo para que os clientes comprem mais. O pagamento do pedido pode ser realizado de diversas formas, porém as mais utilizadas são PIX e cartão de crédito.

Tabela 2: Forma de pagamento, total de pedidos e faturamento total

Forma de Pagamento	Total de Pedidos	Faturamento Total
PIX	4.179	R\$ 1.552.123,58

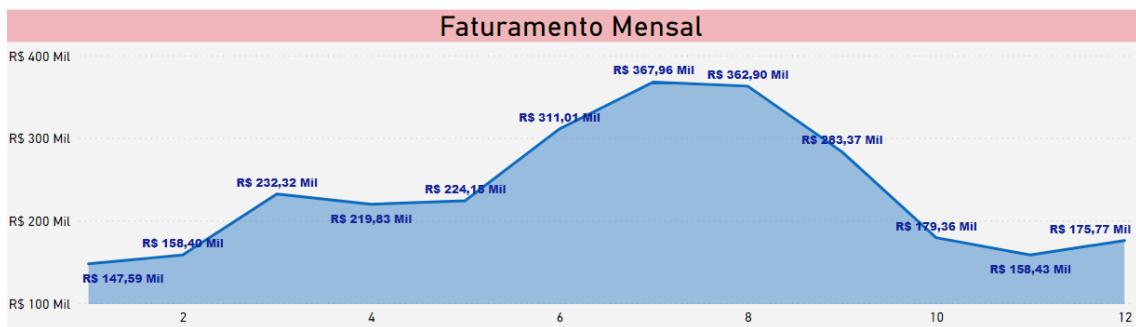
Cartão de Crédito	3.952	R\$ 1.268.956,46
Total	8.135	R\$ 2.821.080,04

Fonte: Elaborado pela autora (2021).

A tabela 2 apresenta as formas de pagamento utilizadas, sendo PIX e Cartão de crédito e respectivamente a quantidade de pedidos e faturamento de cada forma, destacando o PIX com maior quantidade de pedidos e faturamento. Para melhor percepção dessas informações foi realizada a análise exploratória de dados identificando *insights* e tendências.

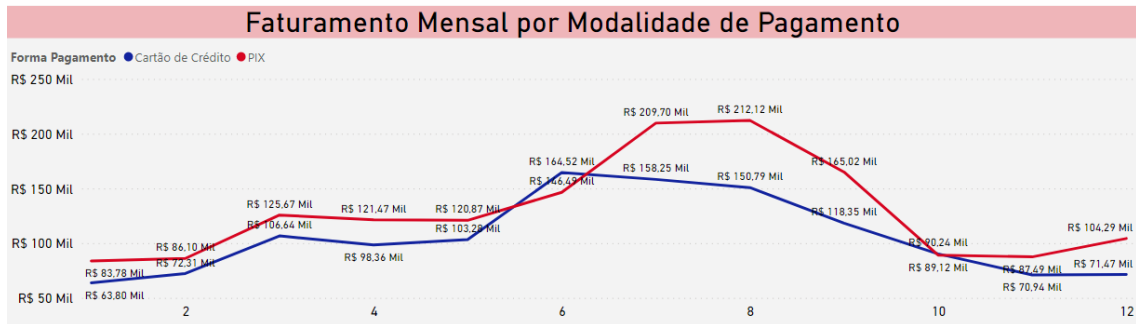
4.2 Análise exploratória dos dados

Gráfico 1: Somatório do Faturamento Mensal relativo aos anos de 2020 e 2021



Fonte: Elaborado pela autora (2021).

No gráfico 1 podemos observar a evolução das vendas por mês no *e-commerce*, podendo destacar os três primeiros meses com um crescimento gradativo. Logo após uma estabilidade no faturamento. Em seguida uma crescente significativa no mês de junho se estabilizando nos meses de julho e agosto, posterior uma queda. Esses fenômenos de alta podem acontecer por diversos motivos, sendo um deles a execução de ações de marketing voltadas ao *e-commerce* da loja, períodos sazonais ou até mesmo a pandemia com o número de óbitos causados pelo vírus da COVID-19. Os dados das transações são muito mais baratos, mais eficazes, menos demorados e mais poderosos (LIAN; YUNJIE; ZHANG, 2019). Retratando assim, a importância de analisar as informações e o contexto que estão inseridas.

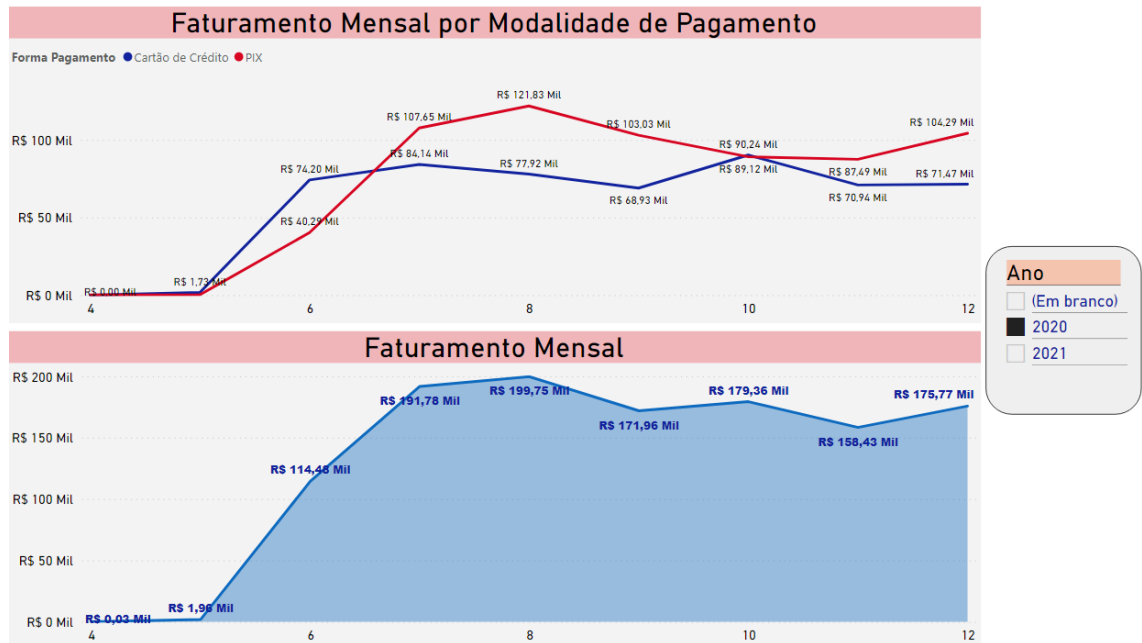
Gráfico 2: Faturamento Mensal por Modalidade de Pagamento

Fonte: Elaborado pela autora (2021).

O gráfico 2 representa o faturamento mensal em relação à modalidade de pagamento escolhida pelo cliente. No primeiro momento pode-se observar uma diferença mínima no valor do faturamento entre as modalidades analisadas, PIX e Cartão de crédito. Dado um certo momento entre os meses de junho, julho e agosto é possível observar o aumento da modalidade de pagamento PIX, impactando no faturamento cerca de R\$61.933,00 (sessenta e um mil, novecentos e trinta e três reais).

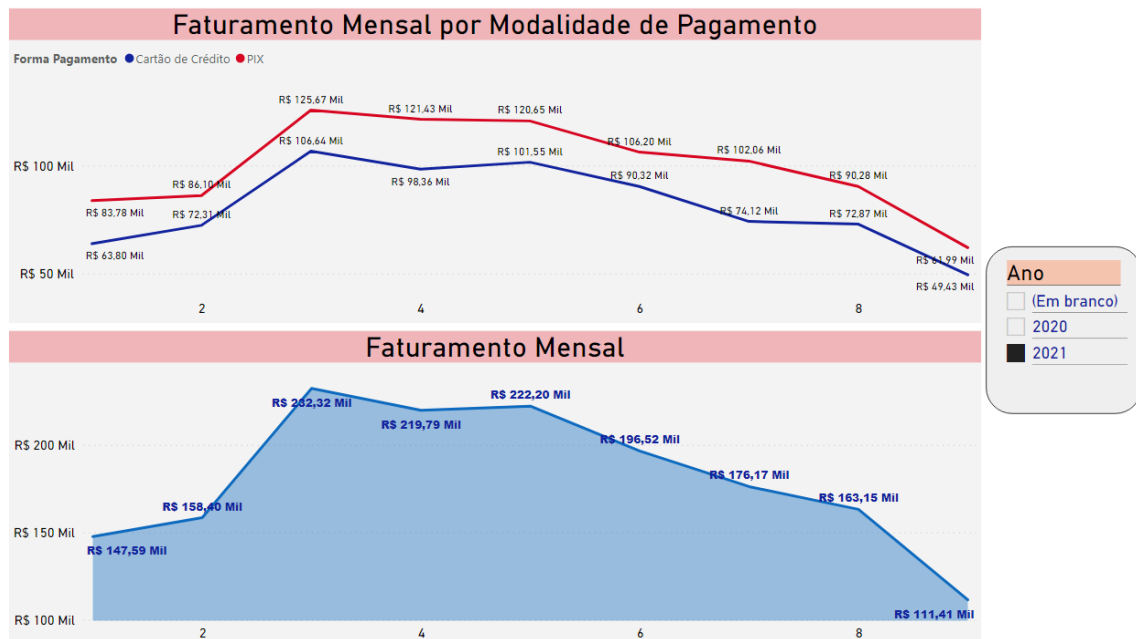
É importante compreender a forma preferencial de pagamento dos clientes, pois a partir dessa informação é possível acompanhar questões de rentabilidade e prazo de recebimento junto aos clientes, como também para compreender o perfil de consumo do público e como a empresa pode utilizar dessa informação para promoções e ações. De acordo com Kumar e George (2007) e Lee et al. (2014) o valor que uma empresa gera a partir de um cliente não se limita à transição atual, mas é o cumulativo do lucro que uma empresa obtém do cliente em todo o relacionamento ao longo do tempo.

Gráfico 3: Faturamento Mensal por Modalidade de Pagamento relativo ao ano de 2020



É possível observar no gráfico 3 a evolução nas vendas, ou seja, o total de faturamento por mês durante o ano de 2020 e sua relação com a modalidade de pagamento. Sendo os meses de maior faturamento os seguintes: julho, agosto e dezembro relacionando esse período com modalidade de pagamento é notória a preferência pelo PIX nesse mesmo período, contribuindo cerca de R\$121.000,00 (cento e vinte e um mil reais) para o faturamento. Em destaque para os meses com maior faturamento mensal sendo eles: julho, agosto e outubro de 2020.

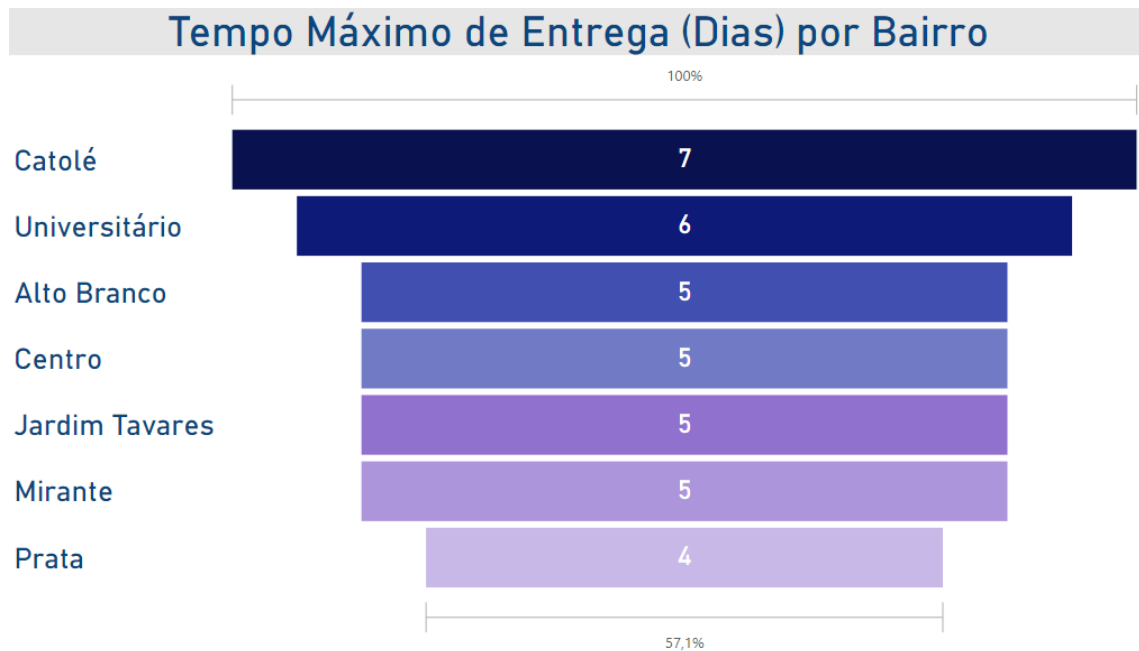
Gráfico 4: Faturamento Mensal por Modalidade de Pagamento relativo ao ano de 2021



Fonte: Elaborado pela autora (2021).

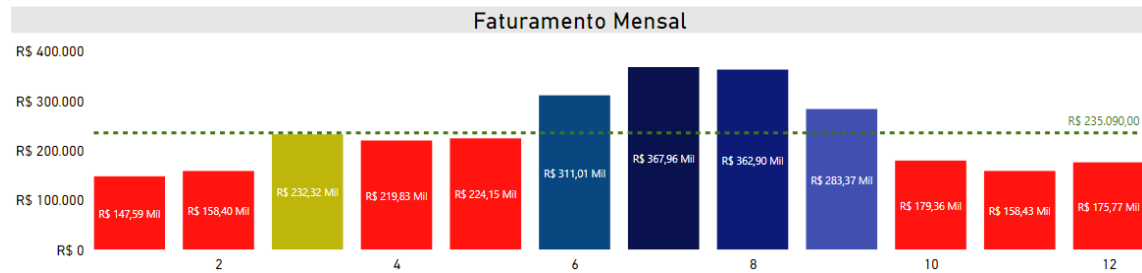
Foi analisado no gráfico 4 a evolução nas vendas, ou seja, o total de faturamento por mês durante o ano de 2021 e sua relação com a modalidade de pagamento. Sendo o maior faturamento da modalidade PIX durante todo o ano, atingindo no pico cerca de R\$125.000,00 (cento e vinte e cinco mil reais) e a modalidade cartão de crédito obteve seu maior faturamento cerca de R\$106.000,00 (cento e seis mil reais) ambos durante o mês de março. Em destaque para os meses com maior faturamento mensal sendo eles: março, abril e maio de 2021. No mês agosto inicia-se um comportamento de queda, podendo ocorrer devido ao retorno dos clientes para loja física, não mais realizando suas compras por meio da plataforma do e-commerce.

Fazendo um comparativo entre o ano de 2020 e o ano de 2021 é possível verificar a aderência da forma de pagamento PIX. De acordo com uma pesquisa realizada pela FGV (2020), foi revelado que as classes C, D, e E tem menos conhecimento e confiança no sistema PIX do que as classes A e B. Trazendo maiores percepções sobre o perfil dos consumidores e sua classe social.

Gráfico 5: Tempo Máximo de Entrega em dias por Bairro

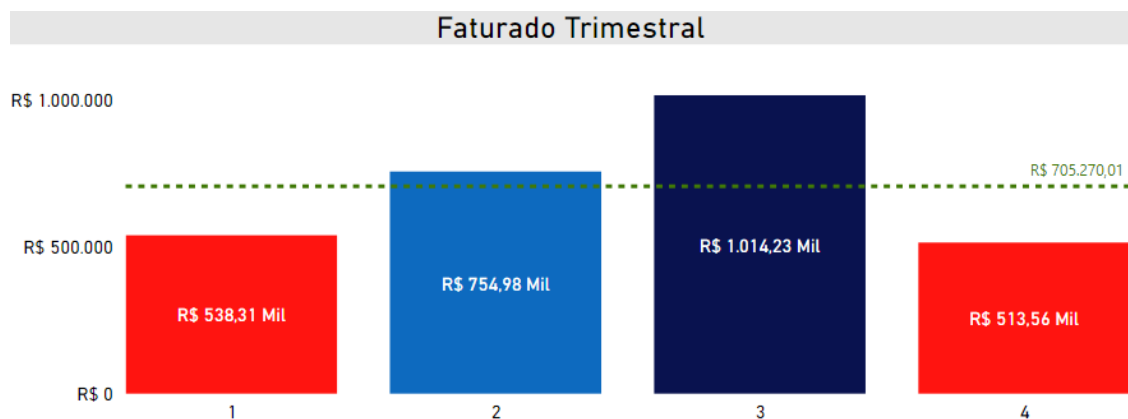
Fonte: Elaborado pela autora (2021).

O gráfico 5 representa o tempo máximo de entrega em dias por bairro, ou seja, os bairros que realizam compras com frequência e consequentemente os que mais contribuem para o faturamento. De acordo com a análise foram identificados os principais bairros sendo eles: Catolé, Universitário, Alto Branco, Centro, Jardim Tavares, Mirante e Prata. Pode-se observar que o bairro do Catolé foi identificado como o que mais realiza compras no supermercado, em contrapartida possui o tempo máximo de entrega, totalizando 7 dias. Investigar esse evento é interessante para a empresa que deseja fidelizar os clientes que proporcionam um considerável retorno financeiro. Essa visualização pode servir como ponto de partida para revisar os processos, detectando gargalos que fazem com que o tempo de entrega aumente. Importante destacar que com base os dados analisados, houve um pedido registrado com tempo de entrega de 18 dias, podendo ter influenciado na média apresentada.

Gráfico 6: Faturamento por mês

Fonte: Elaborado pela autora (2021).

O gráfico 6 demonstra o faturamento mensal do *e-commerce* em relação ao valor médio de faturamento por mês. Sendo os meses de maior faturamento: **junho, julho, agosto e setembro** ultrapassando consideravelmente em relação à média mensal de R\$ 235.090,00 (duzentos e trinta e cinco mil e novecentos reais). Compreender porque esses meses se comportam dessa maneira, pode fazer com que tenha um melhor aproveitamento desse período, entendendo o aumento de consumo e elaborando estratégias para se beneficiar desse fenômeno. Como também executar ações para os meses em que possui um faturamento bem abaixo do valor médio mensal. Com essas análises possibilita a elaboração de um planejamento estratégico de maneira mais assertiva, revertendo períodos com resultados insatisfatórios e potencializando períodos positivos.

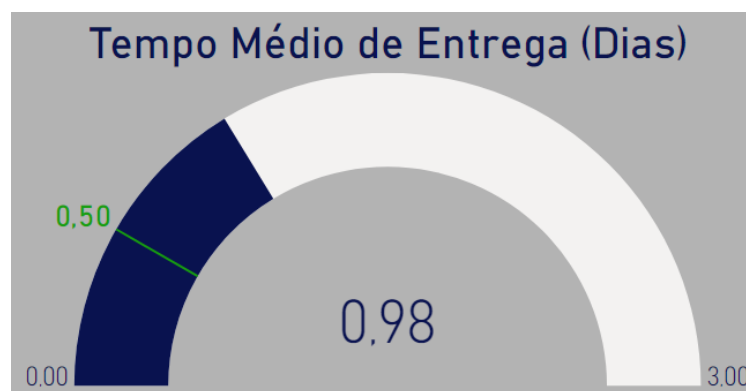
Gráfico 7: Faturamento trimestral

Fonte: Elaborado pela autora (2021).

Neste gráfico 7 foi ilustrado o faturamento do *e-commerce* por trimestre durante todo o período analisado. Observou-se que o faturamento médio por trimestre é de R\$ 705.270,01 (setecentos e cinco mil, duzentos e setenta reais e um centavo). Durante o terceiro trimestre que corresponde aos meses de julho, agosto e setembro foi registrado um aumento relevante

comparado com os demais, em relação ao valor da média. De acordo com Potdar et al. (2018), o envolvimento do cliente inclui comunicação, interação, envolvimento, satisfação, contínuo pontos de conexão, ligação e endosso. Nesse sentido, a interação com cliente deve acontecer durante todo o ano, fazendo com que se tenha um relacionamento próximo, utilizando todas as ferramentas de comunicação disponibilizadas para isso, revertendo os resultados negativos do primeiro e quarto trimestre.

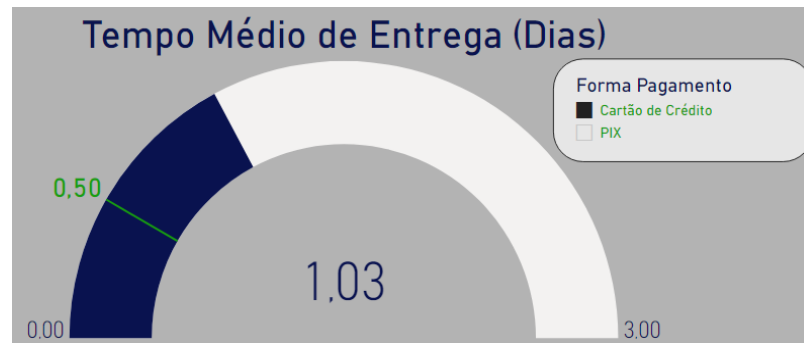
Gráfico 8: Tempo médio de Entrega em dias



Fonte: Elaborado pela autora (2021).

O tempo médio de entrega foi desenvolvido a partir da data do pedido e a data de entrega do pedido representado no gráfico 8, obtendo a média de 0,98, ou seja, o pedido é entregue em menos de 24 horas. Com isso, pode-se estabelecer metas em relação ao tempo médio de entrega, para maior satisfação dos clientes visto que o tempo é levado em consideração por clientes que realizam compras *online*. A capacidade da empresa de detectar e adaptar as necessidades dos clientes às mudanças nas condições do mercado ajuda a conectar clientes (DAY, 2011; VORHIES et al., 2011). Portanto, esse dado, pode ser transformado em um diferencial competitivo realizando entregas no tempo eficiente, atendendo as necessidades da população que priorizam agilidade na entrega para escolha do supermercado.

Gráfico 9: Tempo médio de entrega em dias em relação a forma de pagamento com cartão de crédito.



Fonte: Elaborado pela autora (2021).

O tempo médio de entrega em dias foi relacionado com a forma de pagamento de cartão de crédito no gráfico 9. Foi verificado um aumento de 1h e 12 min. no tempo de entrega, fato que pode ter como causas o prazo de confirmação das bandeiras de cartões aceitas, tendo por consequência aguardar o pagamento para iniciar o processo de entrega.

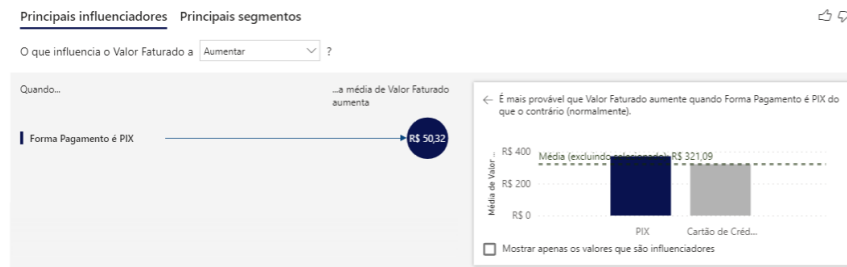
Gráfico 10: Tempo médio de entrega em dias em relação a forma de pagamento com PIX



Fonte: Elaborado pela autora (2021).

O tempo médio de entrega em dias foi relacionado com a forma de pagamento PIX no gráfico 10. Foi verificado uma diminuição de 1h e 12 min. no tempo de entrega, fato que pode ter como causas o prazo de confirmação direto dessa modalidade de pagamento, dando celeridade no processo de entrega.

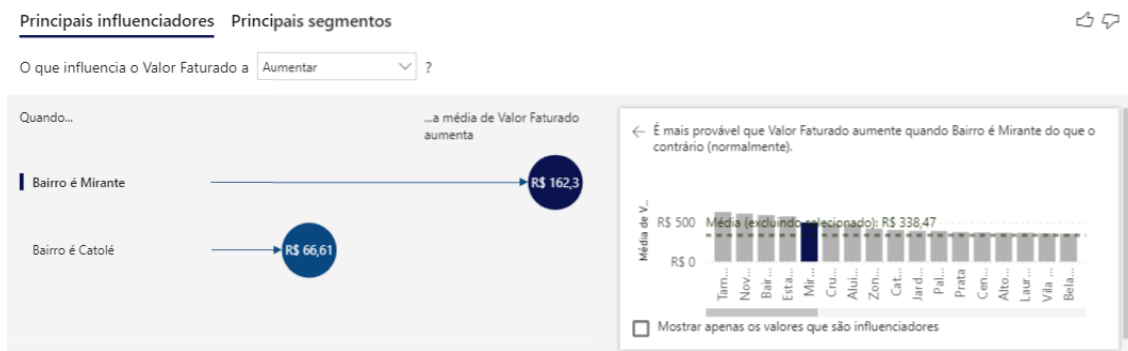
Gráfico 11: Contribuição para o faturamento em relação a forma de pagamento com PIX



Fonte: Elaborado pela autora (2021).

Por meio da investigação dos dados foi identificado a preferência dos clientes em realizar o pagamento por PIX, como demonstrado no gráfico 11. Quando o pagamento é realizado através dessa modalidade, contribui em média R\$50,32 reais para o faturamento. A crescente utilização do PIX, tem contribuído para a construção de um mercado de pagamentos de varejo mais competitivo, mais eficiente, mais inclusivo e mais seguro (GARCIA, 2021). Neste sentido, a facilidade da informação dessa modalidade de pagamento irá tornar ainda mais atrativo a realização das compras, visto que é uma preferência do consumidor.

Gráfico 12: Contribuição para o faturamento em relação aos bairros

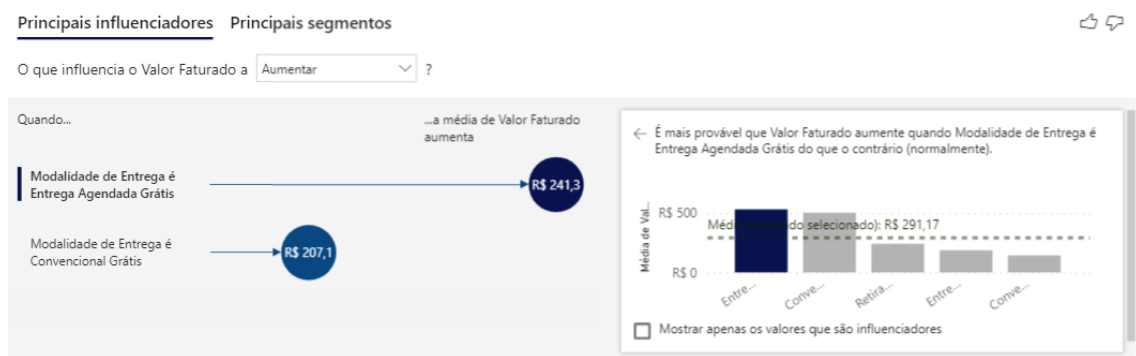


Fonte: Elaborado pela autora (2021).

Dentre os bairros dos clientes cadastrados que realizaram compras podemos destacar os o Catolé e o Mirante como os que mais contribuem para o faturamento. O bairro do Mirante é classificado em quinto lugar, em relação aos bairros que mais compram, porém, sua contribuição para o faturamento é maior dentre todos os outros. Esse comportamento se deve em relação a classe social dos ocupantes desse bairro, possuindo uma renda superior, conseqüentemente realizando compras de valores maiores. De acordo com o censo 2010 do

IBGE o bairro do Serrotão possui uma renda per capita de R \$201,24 já o bairro do Mirante possui R \$3.117,80 (MONTE, 2015). Comprovando a relação dos moradores do bairro e seu poder econômico. Essa análise permite com que o marketing da empresa possa explorar mais esse bairro, com a criação de campanhas direcionadas a esse público com o objetivo de aumentar a quantidade de compras e causando um impacto no faturamento.

Gráfico 13: Tempo médio de entrega em dias em relação a forma de pagamento com PIX



Fonte: Elaborado pela autora (2021).

Dentre as formas de entrega disponibilizadas, podemos ver no gráfico 13 o destaque as que possuem maior contribuição para o faturamento, sendo a entrega agendada grátis com média de R\$241,3 reais (duzentos e quarenta e um reais e três centavos) e entrega convencional grátis com média de R\$ 207,1 reais (duzentos e sete reais e um centavo).

Esses resultados devem servir de insumo para diversos setores da empresa como marketing, comercial, logística entre outros, que por meio dessa análise irão atuar de maneira mais assertiva, atraindo o cliente de forma personalizada e causando maior satisfação. Essas informações também permitem à empresa ter um diferencial competitivo no mercado, criando alternativas e soluções para atender as necessidades e desejos de seus clientes. O registro de compras analisados trouxe reconhecimento do perfil dos consumidores do e-commerce e quais as práticas com melhores retornos financeiros, como também da sua atuação durante esse período pandêmico.

5. Conclusões

O presente trabalho se propôs a analisar os registros de compras produzidos no e-commerce de um varejista supermercadista implementado durante a pandemia, a fim de aprofundar e compreender o perfil dos clientes. Os achados na pesquisa mostraram através da análise exploratória de dados a identificação e compreensão do perfil dos consumidores e sua relação com o faturamento, permitindo gerar insights para uma tomada de decisão estratégica, podendo assim adaptar seus serviços de acordo às necessidades e desejos dos clientes.

Sobre os achados da pesquisa pode-se destacar o reconhecimento que moradores do bairro do Mirante quando realizam compras o valor médio aumenta de maneira significativa, porém esses moradores estão em quinta colocação em relação aos bairros que mais realizam compras, os consumidores que realizam compras e pagam por PIX em média contribuem mais para o faturamento em relação ao cartão de crédito e os clientes possuem preferência em entrega gratuita entre outras.

O estudo é importante, pois contribuiu para o entendimento e percepção do perfil dos consumidores, suas preferências e seu impacto no faturamento de um varejo supermercadista. Além de servir de subsídio para os gestores enxergarem oportunidades de desenvolver melhorias nos processos para um atendimento de forma personalizada e atrativa. Também é possível estabelecer metas sobre prazos e faturamento, utilizando como base nos valores médios, permitindo ganho financeiro.

Contudo, a pesquisa apresentou algumas limitações como a extração do banco de dados inconsistente, apresentando palavras distorcidas, fazendo com que o processo de limpeza de dados se tornasse ainda mais importante e minucioso. Além de serem dados de um período considerado atípico, pois vivenciamos um cenário pandêmico, onde houve considerável crescimento do comércio eletrônico.

Como sugestão para pesquisas futuras, espera-se que esse trabalho possa contribuir para uma investigação do perfil de consumo dos clientes em um contexto mais amplo e com outros segmentos de varejo, como também segmenta-los, ou até mesmo desenvolver um sistema de *clusters*, diferente do que foi feito para este estudo, que buscou compreender as variáveis geradas no banco de dados e como se relacionam, no contexto da cidade de Campina Grande.

Referências Bibliográficas

ABRAS. **Setor supermercadista registra crescimento de 9,36% em 2020**. Disponível em: <<https://www.abras.com.br/clipping/noticias-abras/72398/setor-supermercadista-registra-crescimento-de-936-em-2020>>. Acesso em: 18. set. 2021.

AERON, H., Kumar, A., & Moorthy, J. **Data mining framework for customer lifetime value-based segmentation**. Journal of Database Marketing and Customer Strategy Management, 19(1), 17–30, 2012.

AGRAWAL, R., Imielinski, T., & Swami, A. **Mining association rules between sets of items in large databases**. In ACM SIGMOD International conference on management of data (pp. 207–216), 1993.

BOONE, D. S., & Roehm, M. **Retail segmentation using artificial neural networks**. International Journal of Research in Marketing, 19(3), 287–301, 2002.

BORNAC, G. 2015. **“The Power of Store Clustering.”** Disponível em: <<https://www.manh.com/resources/articles/2015/08/27/power-store-clustering>>. Acesso: 10 set. 2021.

CERVIERI JUNIOR, O. et al. **Tecnologias emergentes aplicáveis ao varejo**. BNDES Setorial, Rio de Janeiro, n. 42, 2015.

CHANG, R. M., Kauffman, R. J., & Kwon, Y. **Understanding the paradigm shift to computational social science in the presence of big data**. Decision Support Systems, 63, 67–80, 2014.

CHEN, H., Chiang, R. H., & Storey, V. C. **Business intelligence and analytics: From big data to big impact**. MIS Quarterly, 36(4), 1165–1188, 2012.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 2. ed. Porto Alegre: Artmed, 2007.

DAY, G.S. Closing the marketing capabilities gap. J. Market. 75 (4), 183–195, 2011.

DEKIMPE, Marnik G. and Dominique M. Hanssens, **“Time-series Models in Marketing: Past, Present and Future,”** International Journal of Research in Marketing, 17 (2), 183–93, 2000.

DRLIK, M.; MUNK, M. **Understanding Time-Based Trends in Stakeholders' Choice of Learning Activity Type Using Predictive Models**. IEEE, v. 7, p. 3106–3121, 2018.

DUAN, L., & Xiong, Y. **Big data analytics and business analytics**. J. Manag. Anal., 2(1),1–21, 2015.

E-COMMERCE BRASIL. **Com crescimento de 75%, e-commerce brasileiro representou 11% das vendas do varejo em 2020**. Disponível em: <<https://www.ecommercebrasil.com.br/noticias/e-commerce-brasileiro-vendas-varejo-2020/>>. Acesso em: 17 de set. 2021.

EUROPEAN COMMISSION, Directorate-general for research and innovation, expert group on retail sector innovation. **Six perspectives on retail innovation**, 2014.

FGV. **Brasileiros precisam de mais informações para poder utilizar o PIX, revela pesquisa**. 2020. Disponível em: <<https://portal.fgv.br/noticias/brasileiros-precisam-mais-informacoes-poder-utilizar-pix-revela-pesquisa>>. Acesso em: 18. set. 2021.

FOWLER, F. J. Jr., & COSENZA, C. (2009). Design and evaluation of survey questions. **The SAGE handbook of applied social research methods**, p.375-412, 2009.

GARCIA, Larissa. Pix é o sistema de pagamentos instantâneos com adesão mais rápida no mundo, diz BC. **Folha de S.Paulo**, 2021. Disponível em: <<https://www1.folha.uol.com.br/mercado/2021/05/pix-e-o-sistema-de-pagamentos-instantaneos-com-adesao-mais-rapida-no-mundo-diz-bc.shtml>> Acesso em 20 set. de 2021.

GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow: Conceitos, Ferramentas e Técnicas para a Construção de Sistemas Inteligentes**. Rio de Janeiro: Alta Books, 2019.

HARRISON, M. **Machine Learning: Guia de Referência Rápida**. Rio de Janeiro: Novatec, 2019.

HONG, T., & Kim, E. Segmenting customers in online stores based on factors that affect the customer's intention to purchase. **Expert Systems with Applications**, 39(2), 2127–2131, 2012.

INGIZZA, Carolina. Vendas online no Brasil cresceram 47% no 1º semestre, maior alta em 20 anos. **Exame**, 2020. Disponível em: <<https://exame.com/pme/e-commerce-brasil-cresce-47-primeiro-semester-alta-20-anos/>> Acesso em 25 set. de 2020.

KUMAR, V., and K. Karande. “The Effect of Retail Store Environment on Retailer Performance.” **Journal of Business Research** 49 (2): 167–181, 2000.

KOWALCZYK, M., & Buxmann, P. An ambidextrous perspective on business intelligence and analytics support in decision processes: **Insights from a multiple case study**. *Decision Support Systems*, 80, 1–13, 2015.

LENZERINI, M. **Data integration: A theoretical perspective**. In Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems (pp. 233–246), 2002.

LIAN, S., X. Yunjie, and C. Zhang. “**Family Profile Mining in Retailing**.” *Decision Support Systems* 118: 102–114, 2019.

MALHOTRA, Naresh K. *Introdução à pesquisa de marketing*. São Paulo: **Pearson Prentice Hall**, 2005.

MATTAR, F. N. *Pesquisa de marketing*. 3.ed. São Paulo: **Atlas**, 2001.

MÁRQUEZ-VERA, C.; CANO, A.; ROMERO, C.; NOAMAN, A. Y. M.; FARDOUN, H. M.; VENTURA, S. **Early Dropout Prediction Using Data Mining: A Case Study with High School Students**. *Expert Systems*. v. 33, n. 1, p. 107-124, 2016.

McKINNEY, W. *Python for Data Analysis: Data Wrangling with Pandas, Numpy, and Ipython*. **O’Reilly Media**. 2ed, 2018.

MIGUÉIS, V. L., A. S. Camanho, and C. João Falcão E. “Customer Data Mining for Lifestyle Segmentation.” *Expert Systems with Applications* 39, 2012

MISHRA, P **Emerging trends in Indian retailing**. *J Retail Mark Distrib Manag* 1(1):1–6, 2017.

MUKHERJEE, A., & Nath, P. **Role of electronic trust in online retailing**. *European Journal of Marketing*. 41(9/10), 1173–1202, 2007.

MONTE, Rammom. Bairros de Campina Grande têm renda média em extremos, diz IBGE. **G1**, 2015. Disponível em: <<http://g1.globo.com/pb/paraiba/noticia/2015/10/bairros-de-campina-grande-tem-renda-media-em-extremos-diz-ibge.html>> Acesso em 18 set. de 2021

NAM, D., Lee, J., Lee, H. **Business analytics use in CRM: a nomological net from IT competence to CRM performance**. *Int. J. Inf. Manag.* 45, 233–245. 2019.

PARK, C. H., Y.-H. Park, and D. A. Schweidel. “A Multi-category Customer Base Analysis.” **International Journal of Research in Marketing** 31 (3): 266–279. 2014.

PARENTE, J.; BARKI, E. **Varejo no Brasil: gestão e estratégia**. 2. ed. São Paulo: 2014.

POTDAR, V., Joshi, S., Harish, R., Baskerville, R., Wongthongtham, P. **A Process Model for Identifying Online Customer Engagement Patterns on Facebook Brand Pages**. Information Technology & People, 2018.

ROCHA, A. B. B. V. **A experiência de compra do consumidor popular no varejo de moda**. Escola Superior de Propaganda e Marketing, 2018.

RAHM, E. **Data cleaning : Problems and current approaches**. Informatica, 1–11, 2000.

SHARDA, R.; DELEN, D.; TURBAN, E. **Business Intelligence e Análise de Dados para Gestão do Negócio**. Porto Alegre: Bookman, 2019.

SRIKANT, R., & Agrawal, R. **Mining Generalized Association Rules**. In VLDB '95 Proceedings of the 21th international conference on very large data bases (pp. 407–419), 1995.

TUKEY, J. W. **Exploratory Data Analysis**, New York: Addison-Wesley, 1977.

VORHIES, D.W., Orr, L.M., Bush, V.D. Improving customer-focused marketing capabilities and firm financial performance via marketing exploration and exploitation. **J. Acad. Market. Sci.** 39 (5), 736–756, 2011.

VOHRA, G. “**Store Clustering**.” Disponível em: <<https://analyticstraining.com/store-clustering/>> Acesso em: 20 set. 2021.

ZHANG, M., Huang, L., He, Z., & Wang, A. G. **E-service quality perceptions: an empirical analysis of the Chinese e-retailing industry**. Total Quality Management & Business Excellence, 2014.