

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
ESCOLA POLITÉCNICA
GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO



DOUGLAS DE OLIVEIRA SOARES

MARKETING DIGITAL E CIÊNCIA DE DADOS
RECOMENDAÇÕES DE PRODUTOS

GOIÂNIA
2022

DOUGLAS DE OLIVEIRA SOARES

**MARKETING DIGITAL E CIÊNCIA DE DADOS
RECOMENDAÇÕES DE PRODUTOS**

Trabalho de Conclusão de Curso apresentado à Escola Politécnica, da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para a obtenção do título de Bacharel em Ciência da Computação.

Orientadora:

Prof^a. Ma. Angélica da Silva Nunes

GOIÂNIA

2022

DOUGLAS DE OLIVEIRA SOARES

**MARKETING DIGITAL E CIÊNCIA DE DADOS
RECOMENDAÇÕES DE PRODUTOS**

Trabalho de Conclusão de Curso aprovado em sua forma final pela Escola Politécnica,
da Pontifícia Universidade Católica de Goiás, para obtenção do título de Bacharel em
Ciência da Computação em: ____/____/____.

Prof^a. Ma Ludmilla Reis Pinheiro dos Santos.
Coordenadora de Trabalho de Conclusão de Curso

Banca Examinadora:

Orientadora: Prof^a. Ma. Angélica da Silva Nunes

Prof^o. Me. Fernando Gonçalves Abadia

Prof^a. Dra. Solange da Silva

GOIÂNIA

2022

AGRADECIMENTOS

Primeiramente, agradeço a Deus pela saúde e força que me concedeu durante esses anos de curso.

Aos meus pais, Lisandro e Elis Regina, por toda a paciência e auxílio em todos os momentos que precisei.

À minha irmã, Aline, que sempre me ajudou no direcionamento em momentos de tomadas de decisões importantes.

A todos os meus colegas e amigos que fizeram parte de algum momento da minha formação.

A todos meus professores, em especial minha orientadora, Ma. Angélica da Silva Nunes, a qual sempre teve muita paciência, dedicação e bons direcionamentos neste trabalho .

RESUMO

O objetivo deste trabalho é demonstrar como a ciência de dados tem um papel importante dentro de uma empresa. O trabalho é focado na ciência de dados e sua correlação com o *marketing* muito direcionado para o *marketing* digital. O *Marketing* 5.0 possui como principal pilar o direcionamento por dados, no qual as ações realizadas dentro da empresa estão embasadas na análise de dados internos. Uma das realizações da ciência de dados que podem impactar o melhor direcionamento do *marketing* de uma empresa do ramo comercial é a utilização de processos de recomendações de compra embasados nas regras de associação. Para a implementação deste caso foi criado um *site* de um mercado de pequeno porte no qual através de um algoritmo muito conhecido na área de ciência de dados chamado de *Apriori*, este foi capaz de oferecer recomendações de compra de outros itens de acordo com métodos estatísticas das correlações de conjuntos de itens. Como resultado, foi possível identificar os itens que possuem uma associação de acordo com o suporte e confiança mínimas definidas e em seguida criar esse vínculo de associações entre os produtos, podendo assim criar campanhas e estratégias de *marketing* voltadas para essas associações.

Palavras-chave: Ciência de dados, *marketing* digital, *marketing* direcionado por dados, regras de associação, *Apriori*.

ABSTRACT

The objective of this work is to demonstrate how *data science* plays an important role within a company. The work is focused on *data science* and its understanding with marketing very focused on digital *marketing*. *Marketing 5.0* has as its main pillar data-driven, in which the actions carried out within the company are based on the analysis of internal data. One of the *data science* sensations that can impact the best *marketing* targeting of a commercial company is the use of purchase recommendation processes based on association rules. For the implementation of this case, a website of a small market was created in which, through a well-known algorithm in the area of data science called *Apriori*, it was able to recommend the purchase of other items according to statistical methods of correlations of item sets. As a result, it was possible to identify the items that have an association according to the defined minimum support and confidence and then create this associations between the products, thus being able to create campaigns and *marketing* strategies met for these associations.

Keywords: *Data Science*, digital *marketing*, data driven *marketing*, rules of association, *Apriori*.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – <i>Data Science</i> : competências necessárias..... | 20 |
| Figura 2- Relação entre conceitos de dados, informações | 22 |
| Figura 3 -- Análise de Dados e suas etapas..... | 25 |
| Figura 4 – Classificações dos sistemas de recomendações | 27 |
| Figura 5 – Representação “Portugol” do algoritmo <i>Apriori</i> | 30 |
| Figura 6 – Arquitetura para automação do modelo de associação de itens proposto | 45 |
| Figura 7– Logotipo <i>Azure</i> | 46 |
| Figura 8 – Criação de conta de armazenamento <i>Azure Blob Storage</i> | 47 |
| Figura 9 – Criação de container de armazenamento na <i>Azure Blob Storage</i> | 47 |
| Figura 10 – Logotipo <i>Azure Data Factory</i> | 48 |
| Figura 11 – Logotipo <i>Azure DataBricks</i> | 48 |
| Figura 12 – Logotipo <i>Wordpress</i> | 49 |
| Figura 13 – Execução do <i>notebook mount_point</i> | 51 |
| Figura 14 – Importação das bibliotecas <i>Numpy</i> e <i>Pandas</i> no <i>Python</i> | 51 |
| Figura 15 – Variável <i>filepath</i> com o diretório da tabela de registros de compras | 51 |
| Figura 16 – Leitura da tabela de registros e sua gravação em um <i>dataframe</i> | 52 |
| Figura 17 – Método <i>removeAllOccurrencesOfValueInList</i> | 52 |
| Figura 18 – Lista que contém listas de registros de compras..... | 53 |
| Figura 19 – Instalação da biblioteca <i>Mlxtend</i> | 53 |
| Figura 20 – Utilização do método <i>TransactionEncoder</i> | 54 |
| Figura 21 – Resultado dos 10 primeiros itens acima do suporte mínimo do algoritmo | 55 |
| Figura 22 – Utilização da regra de associação confiança no algoritmo <i>Apriori</i> | 56 |
| Figura 23 – Itens frequentes com seus suportes e confiança | 56 |

| | |
|---|----|
| Figura 24 – Indicação do diretório para salvar o resultado obtido no <i>storage</i> | 57 |
| Figura 25 – Diretório “Registro” com o arquivo “Recomendados.csv” | 57 |
| Figura 26 – Inclusão no produto Arroz os itens associados | 58 |
| Figura 27 – Inclusão no produto Tomate os itens associados | 58 |
| Figura 28 – Inclusão no produto Leite UHT os itens associados..... | 59 |
| Figura 29 – Visualização do produto Leite UHT com os itens associados no <i>site</i> | 59 |
| Figura 30 – Visualização do produto Tomate com os itens associados no <i>site</i> | 60 |
| Figura 31 – Tela inicial para criação de conta na <i>Azure</i> | 68 |
| Figura 32 – Tela de validação de celular para criação de conta na <i>Azure</i> | 68 |
| Figura 33 – Tela de inserção de dados do cartão de crédito para criação de conta na <i>Azure</i> | 69 |
| Figura 34 – Tela inicial da conta criada na <i>Azure</i> | 69 |
| Figura 35 – Criar recurso <i>Azure DataBricks</i> | 70 |
| Figura 36 – Inserir os valores correspondentes para a criação do recurso <i>DataBricks</i> | 71 |
| Figura 37 – Criar um workspace do <i>Azure Databricks</i> | 71 |
| Figura 38 – Implantação do <i>workspace</i> do <i>Azure Databricks</i> | 72 |
| Figura 39 – Implantação do <i>Azure Databricks</i> finalizada | 72 |
| Figura 40 – Tela inicial do <i>Azure Databricks</i> | 73 |
| Figura 41 – Tela da criação de um cluster no <i>Azure Databricks</i> | 73 |
| Figura 42 – Tela do processo de criação de um <i>cluster</i> concluído no <i>Azure Databricks</i> | 73 |
| Figura 43 – Tela inicial no <i>Azure Databricks</i> , seleção “ <i>Create Notebook</i> ” | 74 |
| Figura 44 – Tela do <i>notebook mount-point</i> do <i>Databricks</i> | 75 |
| Figura 45 – Tela do <i>notebook mount-point</i> exibindo os conteúdos armazenados no <i>container</i> | 75 |
| Figura 46 – Tela dos arquivos armazenados no <i>container</i> | 76 |

| | |
|--|----|
| Figura 47 – Tela dos recursos que podem ser adquiridos na busca do recurso <i>Data Factory</i> | 77 |
| Figura 48 – Tela de criação do recurso <i>Data Factory</i> | 78 |
| Figura 49 – Tela de criação e parametrização do <i>Data Factory</i> | 78 |
| Figura 50 – Tela inicial do <i>Data Factory</i> criado | 79 |
| Figura 51 – Tela <i>Data Factory</i> com a criação dos <i>pipelines</i> | 79 |
| Figura 52– Tela <i>Data Factory</i> com a parametrização do <i>pipeline</i> de extração..... | 80 |
| Figura 53 – Tela <i>Data Factory</i> com a parametrização do <i>pipeline</i> de transformação | 81 |
| Figura 54 – Tela <i>Data Factory</i> do <i>pipeline</i> de transformação concluída | 81 |
| Figura 55 – Tela <i>Data Factory</i> do <i>pipeline</i> de execução construído..... | 82 |
| Figura 56 – Tela <i>Data Factory</i> parametrização da trigger semanal..... | 82 |
| Figura 57 – Tela dos recursos do <i>Azure</i> | 83 |
| Figura 58 – Tela para criação de um recurso <i>Wordpress</i> | 83 |
| Figura 59 – Tela para criação e parametrização de um recurso <i>Wordpress</i> | 84 |
| Figura 60 – Tela com os parâmetros de infraestrutura do recurso <i>Wordpress</i> | 84 |
| Figura 61 – Tela com processo de instalação do recurso <i>Wordpress</i> | 85 |
| Figura 62 – Tela para busca de acesso ao <i>site</i> | 85 |
| Figura 63 – Tela inicial direcionado para o <i>site</i> | 86 |
| Figura 64 – Tela de <i>login</i> de administrador do <i>site</i> | 87 |
| Figura 65 – Tela de instalação do tema no <i>Wordpress</i> | 88 |
| Figura 66 – Tela de instalação de um template no <i>Wordpress</i> | 89 |
| Figura 67 – Tela do <i>site</i> com o <i>template Digital Shop</i> | 89 |
| Figura 68 – Tela com a visualização de imagens de produtos..... | 90 |
| Figura 69 – Tela com modelo de importação de produtos em massa no <i>site</i> | 91 |
| Figura 70– Tela com a parametrização de importação em massa de produtos no <i>site</i> | 91 |
| Figura 71 – Tela com a seleção de colunas para importação em massa de produtos | |

| | |
|---|-----|
| no <i>site</i> | 92 |
| Figura 72– Tela com a visualização dos produtos importados e os respectivos atributos | 92 |
| Figura 73 – Tela com a visualização dos produtos no <i>site</i> | 93 |
| Figura 74 – Tela com a visualização dos produtos no <i>site</i> | 93 |
| Figura 75 – Tela com a visualização do cadastro de cliente no <i>site</i> | 94 |
| Figura 76 – Tela com a visualização de <i>login</i> de cliente no <i>site</i> | 94 |
| Figura 77 – Tela com a alteração de nome das páginas importadas do <i>template</i> | 95 |
| Figura 78 – Tela do <i>site kaggle</i> ,que contem bancos de dados gratuitos..... | 96 |
| Figura 79 – Tela dos arquivos baixados no formato csv do <i>site kaggle</i> | 96 |
| Figura 80 – Tela com o banco de dados “ <i>groceries-groceries</i> ” aberto no <i>excel</i> | 97 |
| Figura 81 – Tela com o banco de dados “ <i>groceries-groceries</i> ” em formato tabela ... | 97 |
| Figura 82 – Tela com o banco de dados “ <i>groceries-groceries</i> ” ajustado..... | 98 |
| Figura 83 – Tela com os itens contemplados no banco de dados da loja | 98 |
| Figura 84 – Tela de criação de conta na <i>Infinity Free</i> | 99 |
| Figura 85 – Tela de criação de <i>site</i> na <i>Infinity Free</i> | 100 |
| Figura 86 – Tela de finalização da criação do <i>site</i> na <i>Infinity Free</i> | 100 |
| Figura 87 – Tela de serviços disponíveis na <i>Infinity Free</i> | 101 |
| Figura 88 – Tela <i>Softaculous Apps Installer</i> disponíveis na <i>Infinity Free</i> | 101 |

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1 – Tipo de Dados | 23 |
| Quadro 2 – Classificações dos sistemas de recomendações | 26 |
| Quadro 3 – Utilidades do <i>marketing</i> direcionado por dados..... | 39 |
| Quadro 4 – Bibliotecas <i>Python</i> utilizadas no algoritmo <i>Apriori</i> | 50 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|------|---|
| AD | <i>Active Directory</i> ; em português: Diretório ativo |
| ADF | <i>Azure DataFactory</i> ; em português: sem tradução |
| BI | <i>Business Intelligence</i> ; em português: Inteligência empresarial |
| CRM | <i>Customer Relationship Managment</i> ; em português: Gestão de Relacionamento com o Cliente |
| DW | <i>Data Warehouse</i> ; em português: Armazém de Dados |
| EDW | <i>Enterprise Data Warehouse</i> ; em português: Armazém de Dados |
| ETL | <i>Extract, Transform, Load</i> ; em português: Extrair, Transformar, Carregar |
| ERP | <i>Enterprise Resource Planning</i> ; em português: Sistema de Gestão Empresarial |
| IA | <i>Artificial Intelligence</i> ; em português: Inteligência Artificial |
| IoT | <i>Internet of Things</i> ; em português: <i>Internet</i> das Coisas |
| SaaS | <i>Software as Services</i> ; em português: <i>Software</i> como serviço |
| TI | <i>Technology Information</i> ; em português: Tecnologia da Informação |

SUMÁRIO

| | |
|---|-----------|
| 1 INTRODUÇÃO | 15 |
| 1.1 Problema de Pesquisa | 16 |
| 1.2 Objetivos | 16 |
| 1.2.1 Objetivo geral | 16 |
| 1.2.2 Objetivos específico | 17 |
| 1.3 Procedimentos metodológicos | 17 |
| 1.4 Organização do trabalho | 17 |
| 2 CIÊNCIA DE DADOS | 19 |
| 2.1 <i>Business Intelligence</i> | 21 |
| 2.2 Mineração de Dados | 21 |
| 2.3 Dados | 21 |
| 2.3.1 Dados para Data Science | 22 |
| 2.3.2 Introdução ao Big Data | 23 |
| 2.3.3 Por que a maioria das organizações implementam o Big Data | 24 |
| 2.4 Análises de Dados | 25 |
| 2.5 Associações de produtos | 26 |
| 2.5.1 Regras de associação | 28 |
| 2.5.2 Algoritmo Apriori | 29 |
| 3 MARKETING | 32 |
| 3.1 As Fases do <i>marketing</i> | 33 |
| 3.2 Marketing Direcionado por dados | 36 |
| 3.2.1 Como preparar o marketing direcionado por dados | 36 |
| 3.3 Marketing Preditivo | 40 |
| 4 MODELO PROPOSTO | 45 |
| 4.1 Ambiente de Desenvolvimento | 46 |
| 4.1.1 Azure | 46 |

| | |
|---|-----------|
| 4.2 Implementação do Algoritmo de associação de produtos – <i>Apriori</i> | 49 |
| 4.2.1 Coleta de dados | 50 |
| 4.2.2 Tratamento de dados | 52 |
| 4.2.3 Transformação dos dados | 54 |
| 4.2.4 Carregamento dos resultados | 57 |
| 4.3 Atualização de itens “Você também pode gostar de” no site Mercado Goiano | 58 |
| 5 CONSIDERAÇÕES FINAIS | 61 |
| 5.1 Dificuldades encontradas | 62 |
| 5.2 Sugestão de trabalhos futuros | 62 |
| REFERÊNCIAS | 64 |
| APÊNDICE A – CRIAÇÃO DE CONTA NA AZURE | 68 |
| APÊNDICE B - INSTALAÇÃO E CONFIGURAÇÃO AZURE DATABRICKS | 70 |
| APÊNDICE C - INSTALAÇÃO E CONFIGURAÇÃO DO AZURE DATAFACTORY | 77 |
| APÊNDICE D - INSTALAÇÃO WORDPRESS NO AZURE | 83 |
| APÊNDICE E – CONFIGURAÇÃO SITE MERCADO GOIANO A PARTIR DO WORDPRESS AZURE | 88 |
| APÊNDICE F – CRIAÇÃO SIMULADA DA TABELA DE REGISTROS DE COMPRA DO MERCADO GOIANO | 96 |
| APÊNDICE G – HOSPEDAGEM NA INFINITY FREE | 99 |

1 INTRODUÇÃO

A tecnologia é muito importante e impulsionadora no percurso evolucionário da humanidade. Observa-se que a busca de melhor qualidade de vida pelos consumidores vem tomando espaço, da simples compra de produtos até a troca de objetos em melhores condições. A partir disso, os sistemas autônomos iniciaram as tomadas de decisões embasados na sua própria inteligência artificial alavancada pelo grande quantidade de dados que é gerado diariamente com o objetivo de facilitar processos rotineiros executados pelo ser humano.

A *Internet* possibilitou a coleta de informações dos clientes a partir da disponibilidades dos dados. O grande volume de dados não é assustador para as empresas, tão pouco não estão longe de garantir o controle dos mesmos até porque querem gerar valor a partir dessas informações (PROVOST; FAWCETT, 2016).

Devido a crescente competitividade do mercado, na qual tem se consumidores com uma gama de opções de seleção seja de produtos ou serviços, é necessário que as empresas criem e implementem estratégias que consigam mostrar os seus diferenciais em comparação com os seus concorrentes para que se mantenham vivos na sua segmentação de mercado e consigam prospectar clientes novos. O *marketing* é a área que cria esses tipos de estratégias. Casas (2007, p.15) afirma:

Marketing é a área do conhecimento que engloba todas as atividades concernentes às relações de troca, orientadas para a satisfação dos desejos e necessidades dos consumidores, visando alcançar determinados objetivos da organização ou indivíduo e considerando sempre o meio ambiente de atuação e o impacto que estas relações causam no bem-estar da sociedade” (CASAS, 2007 p.15).

A ciência de dados tem papel importante quando utilizado em conjunto com o *Marketing*, podendo melhorar diversas atividades do cotidiano dos ramos empresariais, sejam atividades internas ou externas. Quando é pensado na análise de mercado, ele consegue trazer previsibilidade de comportamento de consumo de clientes impulsionando as empresas na preparação de estratégias para intensificar sua representatividade no mercado.

A experiência de compra do consumidor também pode ser influenciada pela ciência de dados, podendo ser personalizado a partir de uma modelagem de dados do cliente. Com os resultados obtidos da análise dos dados se consegue entender melhor as suas necessidades, qual a melhor abordagem ao mesmo e quando essa

abordagem deve ser feita.

Busca analisar as definições conceituais que envolvem a ciência de dados como: *Business Intelligence*, Inteligência Empresarial (BI), *Big Data* e regras de associação. Além disso, procura analisar as definições sobre o *Marketing*, como também, a sua evolução e como ele vem acompanhando o surgimento de novas ferramentas tecnológicas como por exemplo o *Big Data*, solução que suporta processos de identificação de padrões de consumo de clientes facilitando as estratégias que podem ser criadas pelas empresas.

Nesse sentido, a escolha do projeto prático foi demonstrar e analisar como as bases de dados internas de uma empresa de *e-commerce* se relaciona para se obter recomendações de compra aos seus clientes de acordo com as compras feitas anteriormente na sua loja. A partir disso, utilizar um algoritmo que possa suportar uma modelagem preditiva que possa garantir a empresa um *marketing* direcional.

1.1 Problema de Pesquisa

Com a demanda crescente de *e-commerce* em vários setores comerciais e a dificuldade em se alcançar melhores e mais prováveis indicações de compras de acordo com o perfil do cliente, o uso de ciência de dados juntamente com a aplicação e uso de outros recursos tecnológicos existentes, pode-se criar um processo de compra com produtos melhores recomendados aos clientes de maneira mais assertiva. Para tanto, este trabalho possui o intuito de responder a questão a seguir: Como a ciência de dados através do algoritmo *Apriori* poderia melhorar as recomendações de compras de itens em um mercado?

1.2 Objetivos

1.2.1 Objetivo geral

Verificar de que forma a ciência de dados através de um tipo de associação de itens pode recomendar de maneira assertiva, itens para compra de acordo com compras realizadas de maneira colaborativa trazendo então a recomendação de compra.

1.2.2 Objetivos específico

Implementar a recomendação de compra a partir da ferramenta Azure, tendo como foco atingir os seguintes objetivos:

- Conexão entre o servidor/banco de dados do *site* com o Azure;
- Coletar registros de compra criando uma base de dados em *cloud*;
- Utilizar do histórico de compra em um algoritmo de recomendação para indicar os produtos relacionados entre si ;
- Criar uma rotina de automação para extrair, executar e indicar os produtos relacionados para serem recomendados dentro de um *site* de supermercado;
- Discutir se os resultados obtidos são satisfatórios.

1.3 Procedimentos metodológicos

O trabalho em questão compõe-se de trabalhos já feitos em relação aos temas embasando-se neles para explicar os assuntos abordados (WAZLAWICK, 2014, p. 21).

Em relação aos objetivos pode-se considerar uma pesquisa descritiva, pois busca a descrição dos fatos sem nenhum tipo de interferência (PRODANOV; FREITAS, 2013, p. 52)

E em relação aos procedimentos é considerado como pesquisa bibliográfica e experimental, que têm como base trabalhos já realizados na área (GIL, 2002, p. 44).

1.4 Organização do trabalho

O trabalho apresenta a divisão em em 5 capítulos. O primeiro capítulo contém a introdução sobre ciências de dados e *marketing* , também apresenta os objetivos gerais, específicos e a metodologia .

O segundo capítulo faz uma revisão sobre a área de ciência de dados, abordando os principais assuntos e conceitos do tema.

No terceiro capítulo é realizada a revisão literária sobre o *Marketing* e *Marketing* Digital nos pontos mais relevantes como o *marketing* direcionado por dados.

O quarto capítulo contém a documentação da parte prática do trabalho. Nesta

etapa foi realizada a escrita dos procedimentos utilizados para a construção do modelo proposto.

Por último o capítulo 5 apresenta as considerações finais.

2 CIÊNCIA DE DADOS

A ciência de dados é muito conhecida também pelo seu termo em inglês: *Data Science* é a ciência que tem como objetivo extrair informação, conhecimento e idéias das diversas origens e quantidade de dados, a partir da coleta ou ingestão e processamento de dados, em que esses processos podem ter apoios de automações advindos da tecnologia e também modelos matemáticos e estatísticos (COELHO, 2017).

Segundo Coelho (2017) *Data Science* ou ciência de dados é um estudo muito bem definido em relação aos dados e outras informações referentes às organizações e as visões de um tema. Portanto, é uma ciência que busca o estudo das informações, o processo de recebimento, transformação e geração para, por fim, se conseguir a análise estruturada dessas informações. A ciência de dados engloba algumas disciplinas. São elas:

- Estatística;
- Matemática;
- Computação;
- Conhecimento do negócio.

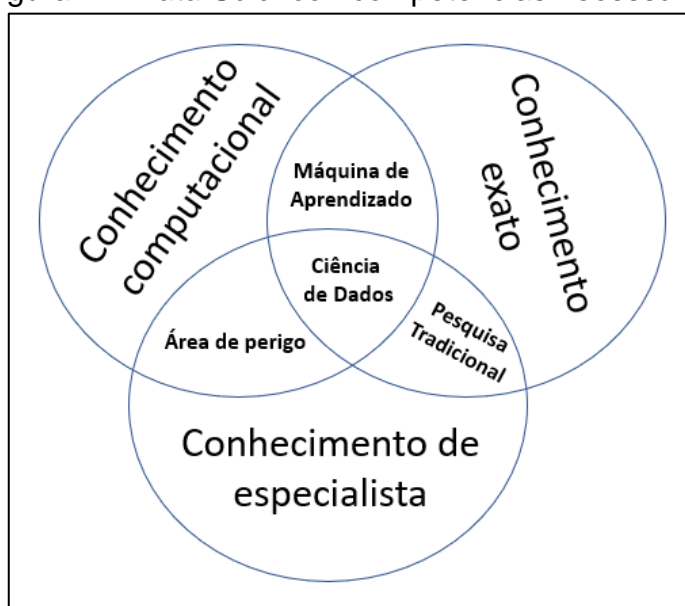
Segundo Provost e Fawcet (2016), o principal objetivo da ciência de dados é o melhoramento da tomada de decisão de uma organização sobre um determinado tema, uma vez que é sempre de interesse para as empresas.

Para Alberto Boschetti, Luca Massaron (2016) a ciência de dados é um tipo de conhecimento novo, mesmo que seus componentes já se tem estudados e pesquisas por vários anos.

Seus componentes incluem álgebra linear, modelagem estatística, visualização, linguística corporal, análise de gráficos, aprendizado de máquina, inteligência de negócios, armazenamento e recuperação de dados (MASSARON 2016, p.8).

Com a ciência de dados é fundamental a união de três competências básicas: conhecimento computacional, conhecimento exato e conhecimento de especialista, conforme ilustra a Figura 1. Segundo VANDERPLAS (2016, p. 11) *Data Science* é:

[...] o conjunto de habilidades interdisciplinares que estão se tornando cada vez mais importantes em muitas aplicações em toda a indústria e academia.”

Figura 1 – *Data Science* : competências necessárias

Fonte: Adaptado (VANDERPLAS, 2016, p. 11)

Para Provost e Fawcet (2016), a ciência de dados estuda analiticamente os dados, buscando trazer informações relevantes para soluções de problemas e para gerar conhecimento. Os autores trazem uma analogia entre a ciência de dados e uma fábrica, em que a estrutura da ciência de dados é análoga a uma linha de produção fabril, que necessita de mão-de-obra, matéria-prima e equipamentos para a produção do que se é objetivado. Na ciência de dados os próprios dados são como a matéria-prima, as modelagens matemáticas e estatísticas como os profissionais que atuam nas áreas e podem ser comparados respectivamente aos equipamentos e a mão-de-obra que conseguem modelar e transformar os dados, buscando a entrega de nova informação ou idéia que pode ser comparado como o produto final criado.

Segundo Provost e Fawcet (2016), deve-se envolver técnicas e processos na ciência de dados para assim levar a compreensão de um todo a partir da análise de dados. A ciência de dados abrange vários campos disciplinares capazes de alcançar respostas de problemas reais através da investigação das informações advindas do processo automatizado, metódico e científico de análise de dados. Para os autores, se faz necessário a atuação das organizações nessa área para terem um melhor posicionamento organizacional da Indústria 4.0.

Para se trazer o assunto de ciência de dados é necessário conceituar os itens a seguir:

2.1 Business Intelligence

As empresas possuem grande volume de dados, mas enfrentam uma grande dificuldade em extrair as informações a partir desse volume. Os sistemas presentes nas empresas, por exemplo, o *Enterprise Resource Planning*, Sistema de Gestão Empresarial (ERP), e outros complementares, não trazem as informações gerenciais de forma mais dinâmica e agradável aos gestores .

O *Business Intelligence* , Inteligência de negócios (BI) tem como objetivo a busca de maior assertividade nas tomadas de decisões das empresas. Embasadas em informações, internas ou externas que podem trazer um melhor direcionamento para o negócio em um planejamento estratégico desfocando a possibilidade de erro. Portanto, a tecnologia gera para a empresa um *status* de maior visibilidade diante o mercado e a concorrência. Sendo assim, pode-se conhecer melhor a frente de qualquer negócio quando conhece a base de dados e com suporte de tecnologias.

Com o BI, é possível ter acesso aos dados e explorar os mesmos tornando uma tomada de decisão mais coerente e precisa. É uma evolução das várias maneiras de consultas existentes nos sistemas anteriores utilizados tornando as informações flexíveis. Porém, se faz necessário uma nova forma de armazenamento dos dados, que não prejudique o desempenho da base operacional. Esta base chama-se *Data Warehouse*, Armazém de Dados (DW) (BARBIERI, 2001).

2.2 Mineração de Dados

A Mineração de Dados explora os dados em busca de padrões encontrados nos mesmos a partir de técnicas analíticas direcionadas por uma máquina. A mineração de dados faz o uso da computação em sua totalidade, na construção de seus processos de trabalho (VANDERPLAS,2016).

2.3 Dados

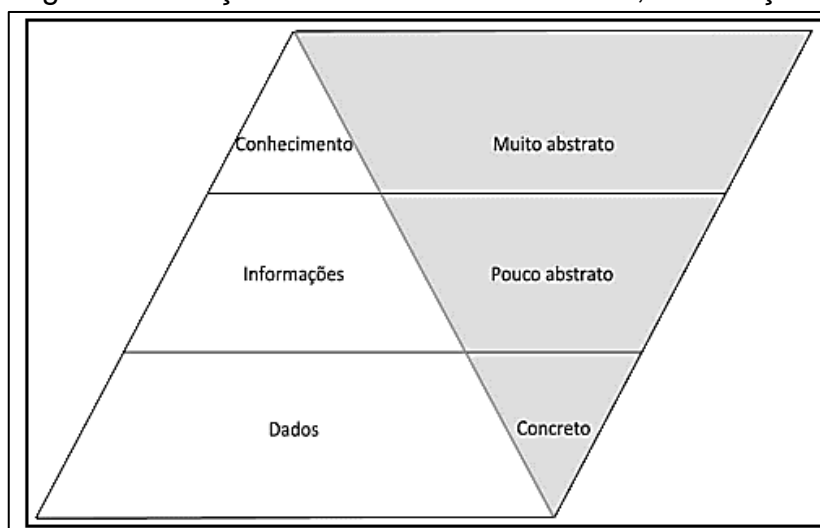
Quando se busca a origem da palavra dados ela é originária da palavra latina *datu* que corresponde no vocábulo português a a palavra “dar”. A partir disso deduz-se que os dados são fatos, pelos quais se deduz fatos adicionais (DATE, 2003, p. 13).

Portanto, pode-se definir os dados como valores quantitativos que em

conjunto, que se faz possível a mensuração ou adquiridos por meio de observação que foi documentada e que têm alguma finalidade de obtenção de informação posteriormente. Para Setzer (1999, p.2) os dados são facilmente armazenados e processados seja ele qualquer ambiente computacional e tecnológico que possui essa destinação e finalidade.

Para se chegar em uma obtenção de resultado a respeito do valor do dado ocorre vários processos, no qual se passa de dados a informações e de informações a conhecimento útil, assim sendo absorvidos na obtenção de *insights*. Conforme ilustra a Figura 2, quanto se utiliza um dado como informação, este passa ter uma abstração maior em relação ao seu significado e sua complexidade, sendo assim aumentando também seu nível de abstração de informação para conhecimento.

Figura 2- Relação entre conceitos de dados, informações



Fonte: (THIERAUF, 1999, p. 23)

Pode-se dizer que os computadores fazem uso de dados e os humanos trabalham com informações. Conceitos importantes para que o cientista de dados tome decisões com a intenção de gerar *insights* absorvidos de análises. Quanto maior conhecimento do que se estuda no sistema de *Data Science*, maior é a capacidade de resultados satisfatórios.

2.3.1 Dados para Data Science

É utilizado na ciência de dados *insights* como origem do conhecimento. A

partir das informações que foram processadas pode-se concluir que, se a quantidade de dados disponíveis está em larga escala, então maior são os resultados com informações assertivas no que se foi proposto.

Existe a distinção no uso de tipos de dados que são trabalhados em ciência de dados, o que permite o uso de dado estruturado, e não estruturado. O Quadro 1 apresenta a distinção entre dados estruturados e dados não estruturados segundo ELMASRI; NAVATHE (2011) .

Quadro 1 – Tipo de Dados

| Tipo de Dados | Descrição |
|-------------------------|---|
| Dados estruturados | Geralmente são organizados em forma de tabelas (linhas e colunas), assim sendo , o tipo mais utilizado e comum de encontrar em aplicações. |
| Dados semi-estruturados | Um intermediário entre dados estruturados e não estruturados. Nele os dados não possuem inteiramente um tipo e são coletados de maneira casual. Não tem um padrão estrutural para as informações e podem ter atributos em algumas entidades. |
| Dados não estruturados | Registra um grande número de informações que se deseja armazenar, mas não se preocupa com a estrutura das informações. É utilizado em larga escala para o armazenamento de vários tipos de dados como formato de textos, vídeos, imagens, entre outros. |

Fonte: Adaptado de ELMASRI; NAVATHE, p. 281 (2011)

2.3.2 Introdução ao Big Data

O termo *Big Data* é constantemente utilizado, porém o termo não tem uma definição do seu conceito totalmente alinhado, sendo assim, vários autores fazem o uso da sua definição do *Big Data* de acordo com o seu tempo.

Em 2012, por dia, se gerou em torno de 2,5 quintilhões de *bytes* de dados e 90% destes dados foram produzidos entre os anos de 2010 e 2012 . Computadores, celulares, materiais inteligentes são geralmente as origens destes dados (IBM, 2012).

Através disso, em várias situações rotineiras, toma-se decisões advindas de sistemas de inteligência artificial que foram impulsionadas por análises de um grande volume de dados com o foco na otimização da vida do ser humano, podendo então trazer uma melhor qualidade de vida.

O termo *Big Data* se faz conhecido por conseguir trazer valor ao negócio de muitas formas: definição de preço, planos e projetos de *Marketing*, entre outros. O termo trata sobre um aglomerado de registros que estão nos servidores e banco de dados das empresas. Geralmente, utiliza-se alguns métodos específicos capazes de obter um valor de dados que pode impactar significativamente os diversos negócios no cotidiano.

Grandes estratégias de *marketing* geralmente vêm de *insights* sobre o mercado. No decorrer dos anos, foi se aperfeiçoando a forma em que os profissionais de *marketing* executavam processos de pesquisa de mercado, conseguindo até, por exemplo, descobrir informações que a concorrência não conseguia obter. Levantamentos quantitativos e em conjunto de pesquisas qualitativas se tornaram direcionadores para qualquer profissional de *marketing* antes de iniciar um ciclo de planejamento de *marketing* (KOTLER, 2021).

Na publicidade e propaganda os dados possuem o papel de entregar uma comunicação de maior destaque ao público certo, trazendo ideias que norteiam as estratégias direcionadas para se obter melhor experiência para o cliente e alavancar o retorno dos investimentos.

Pode-se dizer que, a *Internet* foi e continua sendo um dos maiores fatores e mais influentes no aumento dos dados. No entanto, além dela, existe outro fator que causou grande impacto foi a ampla adoção de dispositivos móveis (MARQUESONE, 2018).

Diante disso, pesquisadores acreditam que se vivencia o início de uma nova revolução industrial, na qual os dados passam a ser elementos chaves desta mudança. Pode-se concluir, portanto, que esse é o momento ideal para se criar oportunidades a partir dos dados (MARQUESONE, 2018).

2.3.3 Por que a maioria das organizações implementam o Big Data

Para Marquesone (2016) as empresas estão à procura de melhores análises e para isso precisam analisar uma quantidade de dados maior. Geralmente a unidade de negócio *marketing* é a que dá o ponto de partida. É possível envolver os produtos em uma camada de serviços, fornecendo uma análise capaz de apoiar os clientes a usar os produtos de maneira mais eficiente e com poucas chances de erro.

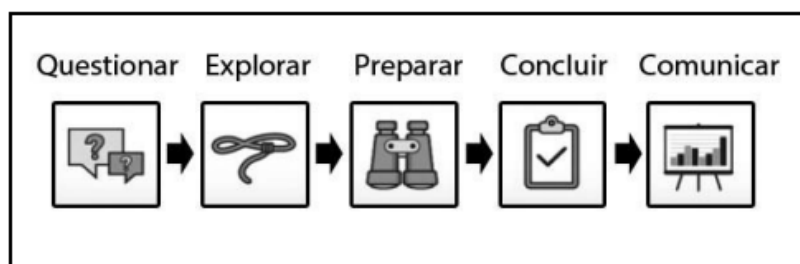
Segundo o Marquesone (2016), as empresas buscam fazer as coisas de

maneira mais rápida, melhor e mais barata usando o *Big Data* para comunicar todas as decisões relacionadas a uma unidade de negócio ou a um processo em específico. Percebem que o *Big Data* é fundamental para várias, senão para todas as unidades de negócio da organização e buscam estabelecer as bases para uma visão centralizada em dados.

2.4 Análises de Dados

Análise de Dados é o processo que procura inspecionar, limpar, transformar e modelar dados. É usada para entender como podem se relacionar os dados de uma companhia para um determinado objetivo. Como a análise de dados é constituída de diversas etapas, as suas principais fases são: questionamento, preparação, exploração, conclusão e comunicação (OLAVSRUD, 2018), conforme ilustrado na Figura 3.

Figura 3 — Análise de Dados e suas etapas



Fonte: Adaptado (LEE; Udacity, 2017)

- **Questionar:** o primeiro passo na análise de dados é questionar a respeito do que se quer solucionar e também a definição de do que é mais importante nos dados. Quando se questiona, o cenário está desenhado entre duas hipóteses. Pode-se questionar inicialmente com os dados disponíveis; ou se questiona antes para depois coletar os dados. Nos dois casos é importante chegar em um bom questionamento que seja capaz de focar-se nos aspectos de maior relevância e a direção de análises significativas (LEE; Udacity, 2017);
- **Preparar:** primeiramente é reunido os dados de maior relevância para que se responda as perguntas em questão. Após isso é feito a avaliação dos dados para que caso exista, problemas qualitativos na estrutura dos

dados. E logo em seguida os dados são tratados, substituindo e removendo dados fora do padrão para garantir um conjunto de dados bem estruturado e de qualidade (LEE; Udacity, 2017);

- **Explorar:** a análise exploratória é baseada na busca de padrões nos dados conseguindo assim extrair intuições a respeito do assunto. Após a exploração deve-se retirar os dados que estão fora dos padrões, desta forma, criando melhores recursos de dados. (LEE; Udacity, 2017);
- **Concluir:** é o último estágio antes da comunicação dos resultados, pode até conseguir trazer previsões assertivas das ocorrências. O seu caráter é uma visão preditiva e a conclusão é realizada com a implementação estatística inferenciais ou estatísticas descritivas. Esta etapa é o momento em que o processo executado consegue obter conclusões analisando padrões de maneira muito mais rápida em comparação com o ser humano (LEE; Udacity, 2017).
- **Comunicar:** o último estágio é a comunicação, que é reponsável pela justificação dos resultados encontrados através dos *insights*. Os resultados podem ser comunicados via: relatórios, *slides*, *e-mails* ou mesmo *dashboards*. (LEE; Udacity, 2017).

2.5 Associações de produtos

Esta tecnologia é presente em quase todos os aspectos da vida digital como por exemplo: com as redes sociais, o comercial e privado em *sites* de notícias e vendas.

O sistema de recomendação é uma aplicação de aprendizado de máquina para negócios que, a partir da modelagem de dados e a aplicação de algoritmos, tenta prever a preferência de um cliente com relação a um item por exemplo (PETS,2020).

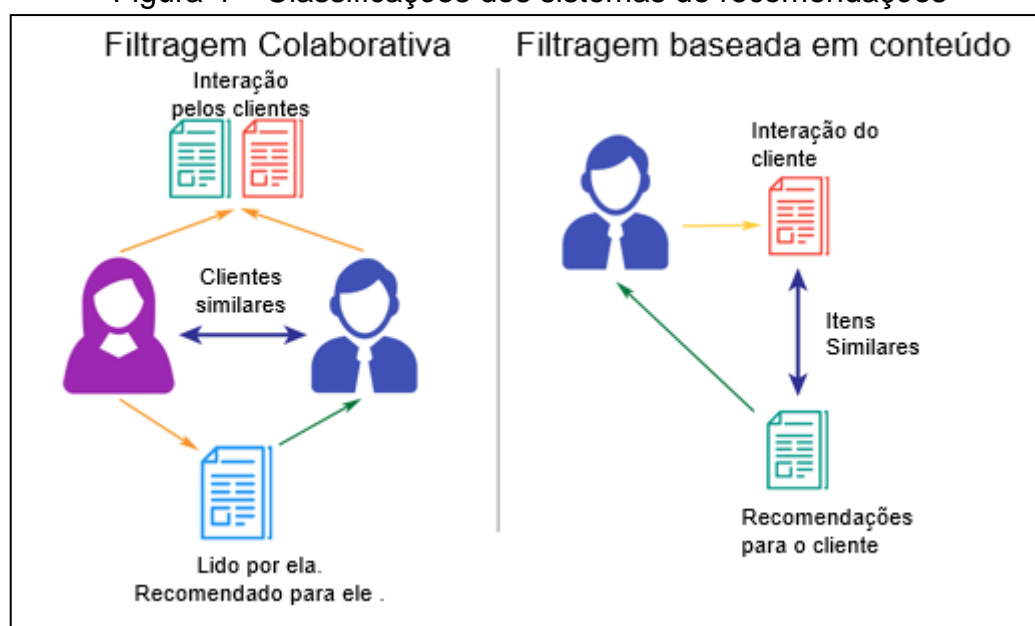
Os sistemas de recomendações de produtos são classificados destas principais formas: filtragem colaborativa, filtragem baseada em conteúdo e sistemas híbridos descritas no Quadro 2 e representadas na Figura 4.

Quadro 2 – Classificações dos sistemas de recomendações

| Classificação | Descrição |
|-------------------------------|--|
| Filtragem colaborativa | Focada na interação entre o usuário e conteúdo, recomendando itens que usuários semelhantes compraram ou interagiram. Assim, a partir dos itens comprados ou interagidos pelo cliente A, o sistema busca por usuários semelhantes que tiveram um comportamento similar e seleciona itens que o usuário A ainda não interagiu. Uma das suas desvantagens é que algoritmos de filtragem colaborativa exige uma quantidade considerável de dados. |
| Filtragem baseada em conteúdo | Focado na geração de recomendações com base de conteúdo já consumido pelo usuário de forma similar, consiste em cruzar os interesses e preferências do cliente com os atributos dos itens, para recomendar ao usuário |
| | novos itens. O usuário é colocado em uma “bolha de preferência”, tudo que se recomenda é semelhante ao que foi consumido, o que pode ser uma desvantagem, uma vez que a longo prazo, pois as preferências do cliente podem ser mudadas levando desinteresse do cliente nas recomendações. |
| Sistemas híbridos | Sistemas híbridos são algoritmos que alinham a filtragem colaborativa e filtragem baseada em conteúdo. Eles podem ser combinados de diversas formas, sendo que, aplicar os dois algoritmos e juntar seus resultados depois é o mais comum. Geralmente, sistemas híbridos possuem um resultado melhor, uma vez que as desvantagens dos tipos citados acima são amenizadas. |

Fonte: Adaptado (PETS,2020).

Figura 4 – Classificações dos sistemas de recomendações



Fonte: Imagem Adaptada (PETS,2020).

Os sistemas de recomendação são muito importantes, pois o comércio eletrônico e mídias digitais cresceram exponencialmente. Porém são levantadas também as questões de uso dos dados e a privacidade do usuário. A precisão de sistemas de recomendação depende muito de uma grande quantidade de dados (PETS,2020).

O problema se encontra na forma com que esses dados são coletados, tratados e comercializados. Até que ponto adquirir dados de consumo e preferências de clientes não entra em conflito com a privacidade do usuário? Serviços gratuitos coletam e comercializando dados de forma sutil. Tudo isso em um momento de incerteza quanto à vigência da Lei Geral de Proteção de Dados (PETS,2020).

2.5.1 Regras de associação

A principal regra para a busca de itens que se tem relação na presença de outros itens em uma mesma transação. Assim, identificar as correlações frequentes entre conjuntos de dados.

Com a tecnologia em desenvolvimento e o surgimento de várias formas de identificação de produtos as empresas conseguem processar e armazenar uma grande quantidade de dados referentes a compras feitas pelos seus clientes. A partir desse conhecimento, as companhias passam a direcionar campanhas de *marketing*, estratégias de catálogos, *layouts* de lojas físicas para favorecer compras de itens que possuem algum certo tipo de correlação. As bases que são envolvidas neste processo são complexas e extensas sendo necessário o uso de algoritmos como por exemplo o *Apriori*, um algoritmo com uma eficiência e objetividade muito bem definidas (BERRY,1997).

Para Vasconcelos e Carvalho (2004), o problema é formalizado da seguinte forma:

$$\text{Seja } I = \{I_1, I_2, I_3, \dots, I_m\} \quad (1)$$

Conjunto I de itens e seja T um banco de dados de transações, onde cada t é representada por um vetor 1 ou 0, se $t[k] = 1$ se t indica a compra do item I_k e $t[k] = 0$, caso contrário. Existe uma tupla na base de dados para cada transação. Seja X um conjunto de itens em I. É dito que a transação t satisfaz X se, para todos os itens I_k em X, $T[k] = 1$ (VANCONCELOS E CARVALHO,2004,p.7).

Para Vasconcelos e Carvalho (2004) quando se descobre as regras de

associação é verificado que o problema pode ser dividido em duas partes:

- Encontra-se os conjuntos dos itens que se relacionam e que possuam um suporte acima de um limite mínimo já determinado. O suporte para um conjunto de itens é o número das transações que contém este conjunto no todo. Os itens frequentes são os que possuem o suporte superior ao mínimo estabelecido;
- Gera-se regras de associação de acordo com conjuntos indicados pelo resultado dos itens frequentes. Deve-se selecionar apenas as regras que possuam o grau de confiança mínimo pré-estabelecido.

Sendo assim, o problema nas regras de associação está na geração das mesmas que tenham o suporte e confiança maiores que os mínimos definidos. Para Vasconcelos e Carvalho (2004), o suporte de uma regra $X \Rightarrow Y$, onde X e Y são conjuntos de itens, é dado pela Equação 2:

$$\text{Suporte} = \frac{\text{Frequência de } X \text{ e } Y}{\text{Total de } T} \quad (2)$$

A sua confiança é dada pela Equação 3:

$$\text{Confiança} = \frac{\text{Frequência de } X \text{ e } Y}{\text{Frequência de } X} \quad (3)$$

O numerador se refere ao número de transações em que X e Y ocorrem simultaneamente. O denominador se refere à quantidade de transações em que o item X ocorre. O suporte (Figura 6) pode ser descrito como a probabilidade de que uma transação qualquer satisfaça tanto X quanto Y , ao passo que a confiança (Figura 6) é a probabilidade de que uma transação satisfaça Y , dado que ela satisfaz X (Vasconcelos e Carvalho, 2004, p.9).

2.5.2 Algoritmo Apriori

O algoritmo *Apriori* emprega a busca em profundidade e gera conjuntos de itens candidatos (padrões) de k elementos a partir de conjuntos de itens de $k - 1$ elementos. Os padrões não freqüentes são eliminados. Toda a base de dados é rastreada e os conjuntos de itens freqüentes são obtidos a partir dos conjuntos de itens candidatos (VANCONCELOS E CARVALHO, 2004).

Para Vasconcelos e Carvalho (2004) o algoritmo faz uso de duas rotinas:

- **Apriori-gen** : utilizado para a geração do conjunto dos itens candidatos

mesmo que não atendam o suporte e mínimo, em seguida a eliminação dos itens que não são frequentes;

- **Subconjunto:** utilizado na extração das regras de associação. O objetivo principal é identificar relacionamento entre os itens enquanto eles são separados. E ao mesmo tempo, calcular o valor que corresponde à confiança e ao suporte.

O algoritmo trabalha sobre um banco de dados em busca de itens frequentes. Portanto, os itens que possuem suporte maior ou igual ao suporte mínimo. Desta forma, se faz necessário o fornecimentos dos valores que correspondem ao suporte mínimo e confiança mínima, além do arquivo que contem os itens e registros de compras (VANCONCELOS E CARVALHO,2004)..

O algoritmo é representado no portuol conforme a Figura 5:

Figura 5 – Representação “Portugol” do algoritmo *Apriori*

| Algoritmo 1: Algoritmo Apriori | |
|--------------------------------|---|
| | $F_1 \leftarrow \{\text{Conjuntos de itens frequentes de tamanho 1}\}$ /* Na primeira passagem $k = 1$ */ |
| 1 | para $k = 2$; $F_{k-1} \neq \text{vazio}$; $k++$ faça /* Na segunda passagem $k = 2$ */ |
| 2 | $C_k \leftarrow \text{apriori-gen}(F_{k-1})$ /* Novos candidatos */ |
| 3 | para todo <i>transação</i> $t \in T$ faça |
| 4 | $C_t \leftarrow \text{subconjunto}(C_k, t)$ /* Candidatos contidos em t */ |
| 5 | para todo <i>candidato</i> $c \in C_t$ faça |
| 6 | $c.\text{contagem}++$ |
| 7 | fim |
| 8 | $F_k \leftarrow \{c \in C_k c.\text{contagem} \geq \text{MinSup}\}$ |
| 9 | fim |
| 10 | fim |
| 11 | Resposta $F \leftarrow \text{Reunião de todos os } F_k$ |

Fonte: (VANCONCELOS E CARVALHO,2004,p.11).

Na primeira fase de iteração, e contado o suporte para cada item individual e todos aqueles que são maiores que o suporte mínimo são selecionados para agregar os itens frequentes. Na segunda iteração, os conjuntos de 2 itens são gerados pela junção dos conjuntos de apenas 1 item e seus suportes são determinados pelo processo de seleção em todo o banco de dados. Dessa forma, encontra-se o conjunto de 2 itens frequentes acima do suporte mínimo. O algoritmo prossegue iterativamente,

até que o conjunto encontrado seja um conjunto vazio (VANCONCELOS E CARVALHO,2004).

3 MARKETING

Segundo Cobra (2009), nos Estados Unidos, no início da década de 1940, o *marketing* foi impulsionado devido ao interesse em se alavancar as vendas como um todo. Foi um caminho apresentado como eficiente para o alcance de vários nichos organizacionais no mercado competitivo. Antes do surgimento do *marketing* tudo que era produzido o mercado consumia, porém foi percebido que os clientes começaram ser cada vez mais exigentes. O conceito do *marketing* teve várias definições e propósitos de acordo no decorrer do tempo.

Para Cobra (2009) a essência do *marketing* é a troca de produtos e serviços, no qual é ofertado por pessoas algum objeto ou serviço que seja útil que tem como foco a satisfação das necessidades e anseios dos consumidores e organização. Cobra (1997) afirma que, diretamente ou indiretamente o *marketing* deve ser compreendido como uma ferramenta existente no cotidiano de atividades organizacionais.

Segundo Ambler (2004), o *marketing* também surgiu no processo de troca de serviços e ou produtos, e considerou o primeiro método impulsionador comercial e ele contribui no desenvolvimento social econômico.

Chauvel (2001) afirma que no século XX a ideia de *marketing* começou a tomar forma, momento em que a expansão mercadológica no mundo ocidental estava acontecendo. Esse mercado no mundo ocidental mudou expressivamente as atividades socioeconômicas. As formas de negócio existentes entre fornecedores, clientes e trabalhadores seriam, embasadas a partir disso, nas leis de mercado e não seriam mais baseadas em controle social.

O *marketing* não é apenas propagandas e anúncios criados para impulsionar vendas de qualquer tipo. Ele está muito além do que se imagina inicialmente. As ações do *marketing* tem tamanha importância no mercado, que acompanha a evolução tecnológica e mercadológica de acordo com os objetivos traçados no decorrer dos anos; e é capaz de trazer um embasamento estratégico de qualquer organização.

Outra ponto a observar é, que geralmente a definição de publicidade e propaganda é sempre confundida com a definição de *marketing*, pois equipes comerciais ofereciam produtos e serviços para o público em geral através de uma lista telefônica e na maioria dos casos nem sabiam ao certo de qual produto se tratava.

Esse tipo de trabalho passou a ser chamado de *telemarketing*. A propaganda é um artifício que o *marketing* utiliza para trazer destaque a um produto ao serviço. Portanto, a propaganda é uma estratégia do *marketing* e não pode ser confundida com a definição do mesmo, pois o *marketing* engloba várias atividades para um objetivo final (KOTLER;KELLER,2006).

De acordo com Kotler e Keller (2006), o *marketing* utiliza do sentimento de desejo das pessoas para obtenção do que foi definido,ou seja, o oferecimento e troca de produtos de valor. É necessário gerar a identificação e satisfação dos desejos sociais, suprimindo as necessidades do consumidor.

3.1 As Fases do *marketing*

Segundo Kotler (2021), ao longo dos anos, o *marketing* evoluiu, sendo definido por cinco fases: *marketing* 1.0,*marketing* 2.0, *marketing* 3.0 , *marketing* 4.0 e, por último, o autor trouxe as definições do *marketing* 5.0. Em todas as fases é buscado o suprimento das necessidades do consumidor. Mas conforme os anos passam, a sociedade também muda as suas necessidades. Portanto é inevitável que seja necessário o desenvolvimento de novas estratégias. As fases do *marketing* representam as mudanças que vão ocorrendo ao passar dos anos.

Kotler (2017) ressalta que durante os anos 60 e 70, surgiu o *marketing* 1.0, e se definiu pela época da Revolução Industrial. Nesse período, o *marketing* tratava de comercializar somente os produtos da indústria. Portanto, Kotler (2000) afirma que o objetivo era criar um padrão a se seguir ganhando em volume de compra, reduzindo o que poderia ser reduzido em relação aos custos de produção, garantindo produtos com preço mais baixo para que pudessem ser comprados por um volume maior de clientes.

Segundo Kotler (2010) o objetivo central do *marketing* 1.0 estava ligado no processo de construção do produto, sendo trabalhado os aspectos concretos, e que não possuíam diferenças entre si e não tinham foco nas necessidades e desejos de escolha e controle do consumidor.

Kotler (2017) afirma que Henry Ford trabalhou de maneira detalhada esse tipo de estratégia, no qual foi produzido em grande escala carros na cor preta. Por causa da cor preta a pintura dos carros secava mais rápido fazendo com que a montagem dos mesmos fosse feita mais rapidamente, sendo então o objetivo empresarial da

época, baixo custo de material, produção em massa acelerada e baixo custo de mão de obra. Havia o suprimento do desejo do consumidor porém o mesmo não possuía opções de escolha.

Na década de 90, o cliente já buscava maior conhecimento dos produtos definindo melhor suas necessidades e desejos de compra, ou seja, as estratégias utilizadas anteriormente de troca e venda já não se faziam tão interessantes no momento. A sociedade foi se expandindo, foram criados vários tipos de trabalhos e as fábricas buscando estimular a eficiência na produção dos produtos começaram a utilizar tecnologias para criação de produtos (KOTLER, 2017).

Nesse período a mídia, que engloba rádio, televisão, jornal, cinema e *Internet*, consegue garantir a entrega de conteúdo para vários receptores de uma única vez, começa a disseminar na sociedade conseguindo influenciar e ter poder sobre a sociedade.

No século XX a economia evolui, trazendo consigo a era do consumidor, e a fase do *marketing* 2.0. Essa fase fundamenta-se no comportamento do consumidor. Para que seja validado a necessidade e o desejo de compra do produto ou serviço o consumidor busca obter mais informações sobre o produto e também possui a necessidade da comparação de preços entre as marcas e modelos (KOTLER, 2010).

Segundo Kotler (2010), o mercado é formado pelo agrupamento de clientes com vontades e necessidades distintas. Para se alcançar a satisfação do cliente com o produto ou serviço, as empresas precisam lidar com o agrupamento de clientes específicos para suprir as necessidades; trazendo uma experiência de compra positiva objetivando uma nova compra do produto ou até mesmo favorecendo a compra de outros produtos da marca.

Segundo Kotler (2010), ao se definir um nicho de mercado para a atuação da empresa, automaticamente é desenvolvido uma conexão empresa-consumidor mais relevante, porque, se o produto supre as necessidades e desejos do cliente, a satisfação do cliente aumenta e com isso é criado um vínculo de fidelidade com o cliente.

As campanhas de *marketing* e publicidade fazem uso de estratégias que tornam o consumidor um alvo passivo, pois o consumidor pode não possuir conhecimento necessário para compreender as “jogadas de *marketing*” e acabam sendo influenciados muita das vezes sem perceber, afetando a sua decisão de compra.

Surge então, segundo (KOTLER,2010) o *marketing* 3.0, momento conhecido também como “era do valor”. O *marketing* 3.0 surgiu com o objetivo da extração e identificação da real necessidade do consumidor. Ou seja, o que realmente deve ser utilizado nas estratégias pelas empresas é a necessidade dos consumidores. Além disso, é necessário a reafirmação constante de declarações positivas da empresa ao cliente, como por exemplo os valores e propósitos empresariais, pois os consumidores buscam por se vincular a empresas que contribuam para um mundo melhor.

O ideal do profissional de *marketing* 3.0 é buscar estratégias que tratam cada consumidor como um único indivíduo, mente e coração. Portanto, no lugar de focar apenas em vender o produto ao consumidor, voltar os olhos do profissional de *marketing* para o indivíduo, ou seja, *marketing* direcionado para os valores de cada indivíduo (KOTLER, 2010).

Para Kotler (2010) o *marketing* 3.0 tem como objetivo solucionar diversos problemas da sociedade, levando a definição de *marketing* a um lugar de valores e espiritualidade humana, fazendo com que as empresas que se baseiam no *marketing* 3.0 consigam sensibilizar os consumidores, de modo com que tragam a identificação entre produto,cliente e marca gerando também uma identificação do consumidor com os ideais da empresa.

Enquanto o *marketing* 2.0 trabalhava simplesmente com a ação comportamental do consumidor, o *marketing* 3.0 deseja garantir todos os desejos do consumidor. Desejos que muitas das vezes alavancaram pela evolução da tecnologia, do mercado e de comunicação no decorrer dos anos (KOTLER, 2010).

Kotler(2017) define o *marketing* 4.0 no ano de 2016 no qual, além de contemplar as mudanças comportamentais dos consumidores, trata também da fidelidade do consumidor podendo torná-lo até um defensor da marca e produto.

Na era do digital, o compartilhamento de experiências entre consumidores demonstra a influência dos consumidores entre si, podendo por exemplo, compartilhar críticas construtivas ou não, que o mesmo teve com o produto e com a marca e, por causa disto, se faz necessário o desenvolvimento mais assertivo de estratégias de *marketing* podendo ter os consumidores como promotores da marca, avaliando positivamente toda a sua experiência de compra (KOTLER, 2010).

O foco no *marketing* 3.0 já estava no ser humano com atenção nas suas necessidades e valores. Contudo, o *marketing* 4.0 passa também a interessar pelo

relacionamento social, porque todos os grupos sociais se auto influenciam, porque todos podem elogiar, comentar, reclamar e dar sugestões por meio das redes sociais. As empresas necessitam atuar com o real objetivo de ajudar e promover algum tipo de bem para a sociedade, pois o cliente mais munido de informação se mostra empoderado e capaz de avaliar aquilo que é prometido por elas (KOTLER, 2017).

Com isso, se definiu o *marketing* digital que está inteiramente ligado aos conceitos do *marketing* 4.0, pois é nesse momento que não se é possível definir *marketing* sem idealizar o meio digital. O *marketing* e o meio digital estão interligados e as empresas precisam estar cada vez mais atentas e presentes nesse meio, pois é onde todos os nichos de mercado estão (KOTLER, 2017).

O *marketing* 5.0 é a contínua evolução de todas as outras fases até o *marketing* 4.0 e tem como objetivo aprimorar as relações entre empresas e clientes de forma sólida e estruturada as relações apoiadas por simbiose homem-máquina. A máquina aparece nas fases de processamento de dados, coleta de dados e análise e gestão dos dados, a descoberta de padrões, uso da lógica em algoritmos e a escalabilidade de processos repetitivos e programáveis (KOTLER, 2017).

3.2 Marketing Direcionado por dados

O primeiro estágio do *marketing* 5.0 é a instituição do *marketing* direcionado por dados. Sendo assim, as empresas conseguem prever a probabilidade de compra dos seus clientes em potencial, a partir do uso da automação de análise de dados, embasando-se no que foi adquirido pelos clientes. Portanto, as empresas podem por exemplo, sugerir ofertas e campanhas customizadas que se adequam ao perfil do cliente. E por fim, o avanço tecnológico garante a execução destes tipos de atividades que englobam toda a segmentação mercadológica (KOTLER, 2021).

Há cerca de vinte anos os profissionais de *marketing* almejavam uma ferramenta capaz de criar um *marketing* realmente personalizado. Don Peppers e Martha Rogers foram precursores precoces do *marketing* individualizado, uma prática muito cobiçada. Os “segmentos de um” são considerados o nível mais alto da segmentação, e a implementação das tecnologias digitais no *marketing*, os tem como principal objetivo (KOTLER, 2021, p. 131).

3.2.1 Como preparar o marketing direcionado por dados

Nos anos subsequentes de 2012 foi percebido o pensamento dos profissionais de *marketing* em buscar informações em bases de dados mais robustas para trazer a ideia de uma otimização da *Customer Relationship Manangement*, gestão do relacionamento com o cliente (CRM). O *marketing* direcionado por dados foi impulsionado pelo *Big Data*. Segundo o autor, os profissionais de *Marketing* possuem a crença de que quanto maior é o volume de dados, mais chances de se encontrar ideias que possam fortalecer o *marketing* da empresa. A partir disso, eles buscam a união dos conjuntos maciços de dados, análise de dados e a pesquisa de *marketing*, em uma plataforma de gestão de dados (KOTLER, 2021).

Para Kotler (2021) é fato que o *marketing* direcionado por dados tem muita relevância no ramo empresarial, porém são poucas as empresas que foram capazes de descobrir a melhor forma de execução com foco no resultado. Segundo o autor, a grande maioria das empresas investem muito no ramo tecnológico e não conseguem a obtenção de resultados positivos no seu ambiente de dados, e isso é resultante de três motivos:

O primeiro motivo é que o projeto é tratado como responsabilidade do departamento de *Technology Information*, Tecnologia e Informação (TI). Quando se inicia o projeto, é concentrado muito foco na seleção de plataformas tecnológicas, nos recursos financeiros para a modelagem da infraestrutura e contratação de funcionários qualificados. A infraestrutura de TI precisa auxiliar a estratégia de *marketing*, e não o inverso. O *marketing* direcionado por dados deve ser um projeto elaborado especificamente pelo *marketing*. Portanto, este departamento não é apenas um auxiliador no projeto, ele se faz necessário na definição e criação de todo o processo de *marketing* direcional .

A maioria dos analistas de mercado acredita que nem sempre uma extensa gama de informações e volume de dados pode ser traduzido em *insights* melhores. É fundamental o entendimento do que se está procurando na base de dados e, portanto, é preciso ter definido qual o objetivo a ser seguido.

O segundo motivo se faz presente na existência do pensamento de que a análise de *Big Data* é o gatilho que permite alcançar os pontos principais a respeito do consumidor e consegue resolver grandes problemáticas do *marketing*. Contudo, ele não substitui nenhum método tradicional do *marketing*, como por exemplo os métodos etnográficos, pesquisa de mercado, testes de utilização ou degustações. O

Big Data deve ser um complemento da pesquisa de mercado. Portanto, um deve ser embasado pelo outro, porque para os dados serem um direcionador do *marketing*, ele precisa de métodos com ciclos formados, com focos específicos mas também precisa de uma análise em tempo real dos dados coletados através do *Big Data* garantido assim, o melhoramento contínuo do *marketing*.

O terceiro motivo é que a expectativa empresarial em conjunto com seu departamento de *marketing* consiga simplesmente inserir um extenso volume de dados em um algoritmo e que o próprio irá trazer respostas de várias de suas perguntas. Os profissionais de *marketing* acreditam que, quando é implantada a análise de *Big Data*, consegue-se deixá-la no automático. Porém, é necessário que ela tenha atuação do indivíduo e do time de *marketing*. As ferramentas tecnológicas podem detectar linhas padronizadas de resultados de maneira muito eficaz, mesmo assim, um profissional de *marketing* que tenha experiência e entendimento contextual para avaliar e interpretar os padrões se faz necessário. Criar novas ofertas e campanhas, por exemplo, também se faz importante a presença dos profissionais de *marketing* para atuar no processo. De acordo com Kotler (2021) para se preparar o *marketing* direcionado por dados é necessário seguir os seguintes procedimentos:

O primeiro é a definição dos objetivos do *marketing* direcionado por dados. O *Big Data* auxilia os profissionais de *marketing* na criação de idéias e novos produtos ou serviços para se conseguir estimulação mercadológica. É possível criar serviços e produtos personalizados a partir das análises de experiência do cliente. O *Big Data* consegue sugerir o que pode oferecer ao mercado e também auxilia no desenvolvimento da entrega do serviço ou produto. Na área de comunicação, o profissional de *marketing* utiliza o *Big Data* para definir o público-alvo, auxilia na geração de conteúdo e ajuda na seleção das mídias que serão utilizadas. Também é fácil ser visto a utilização dos dados para o auxílio na retenção do cliente e pós-venda. O *Big Data* é utilizado também na previsão de desistências de contratação de serviço e na determinação de métodos de recuperação do serviço (KOTLER, 2021).

Com a definição dos objetivos a comunicação se torna simples. Todavia colabora para diferentes unidades empresariais se alinharem em um único foco, garantindo o engajamento e a responsabilização. Os objetivos específicos instigam o profissional de *marketing* em como aumentar o desempenho e a garantir o foco nas ações. Para se obter maior sucesso é necessário que os profissionais de *marketing*

selecionem os objetivos de maiores impactos na empresa tornando a adesão do mesmo mais aceitável entre os colaboradores (KOTLER, 2021).

No Quadro 3 pode-se observar a mensuração do *marketing* direcionado por dados, que é possível somente quando se estabelecem objetivos com foco centralizado.

Quadro 3 – Utilidades do *marketing* direcionado por dados

| O que Oferecer | Como oferecer |
|---|--|
| Novas idéias e que podem gerar produtos ou serviços | Identificando o público alvo |
| Estimativa da demanda | Definindo conteúdo correto na mensagem |
| Recomendações de compra | Comunicando corretamente o mix de mídias |
| Produtos e serviços personalizados | Identificando o mix de canais |
| Experiência do consumidor | Criando a persona do cliente |
| Preço justo de produtos | Projetando níveis assistenciais do cliente |
| Precificação dinâmica | Identificando queixas e desistências de clientes |

Fonte: Kotler, 2021

O segundo procedimento é a seleção dos dados disponibilizados. Segundo Kotler (2021) no período digital a massa de dados está crescendo de uma forma acelerada. O nível de detalhe tem se aperfeiçoado e também a diversidade dos dados cresce exponencialmente. Porém, não se pode afirmar a relevância e valor dos dados. Quando as empresas focam nos objetivos, é necessário a identificação dos dados corretos que fazem sentido para serem coletados e analisados.

O *Big Data* pode ser classificado de várias maneiras. E uma delas é a classificação a partir da origem dos dados:

Dados sociais, que incluem todas as informações que os usuários de mídias sociais compartilham, como localização, perfil demográfico e interesses. Dados de mídia, que incluem medições de audiência em mídias tradicionais, como televisão, rádio, imprensa e cinema. Dados de tráfego na web, que incluem todos os logs gerados por usuários navegando na web, como visualizações, buscas e compras. Dados de pontos de venda e transações, incluindo todos os registros de transações feitas pelo consumidor, como localização, quantia, informações do cartão de crédito, compras, horários e, às vezes, número do documento do cliente. Dados da *Internet* das coisas, incluindo todas as informações coletadas por aparelhos e sensores conectados, como localização, temperatura, umidade, proximidade de outros aparelhos e indicadores de saúde. Dados de engajamento, que incluem todos os pontos de contato diretos entre empresas e clientes, como dados de call center, trocas de *e-mails* e dados de chats (KOTLER, 2021, p.140).

Para Kotler (2021) os dados de transações e engajamento são tipo de dados internos e os profissionais de dados podem acessá-los. Porém não se pode afirmar que todo dado interno é confiável para ser utilizado. Pode ser necessário a análise e limpeza dos dados de acordo com a organização dos dados e manutenção dos registros. Inclui-se também a consolidação e correção de dados duplicados, conjuntos de dados imprecisos e registros incompletos. Os dados segmentados em dados de mídia e os sociais, por exemplo, são externos e na grande maioria, são recebidos de terceiros.

O terceiro procedimento recomendado por Kotler (2021) é a criação de um ambiente integralizado de dados. Identificar um ponto comum entre as origens de todos os dados é um desafio na integralização dos dados. Integrar os dados no nível do consumidor individual se faz importante pois garante o *marketing* de “segmentos de um”. Se faz necessário uma boa qualidade na prática de registro para garantir a identidade única do cliente em cada grupo de dados capturado (KOTLER, 2021).

Identificar chaves exclusivas em fontes de dados externas é uma tarefa muito mais complexa em relação a fonte de dados internas, por exemplo, a identificação do cliente e o seu histórico de compras podem ser integrados aos dados sociais caso o cliente efetue acesso em *sites* de *e-commerce* com sua conta de redes sociais. Outro exemplo é a conexão de aplicativos de fidelização do cliente com sensores inteligentes. Por exemplo, o sensor inteligente consegue registrar o movimento do cliente nos corredores de uma loja, sendo assim, muito importante no rastreamento de jornada do cliente (KOTLER, 2021).

Cada conjunto dinâmico de dados precisa ser armazenado em uma única plataforma de gestão de dados, permitindo ao profissional de *marketing* capturar, armazenar, gerir e analisar os dados de forma abrangente. Todo projeto de *marketing* direcionado por dados com objetivos novos deve continuar a usar a mesma plataforma, o que possibilita enriquecer o ecossistema de dados, beneficiando a empresa caso decida usar a aprendizagem de máquina para automatizar a análise (KOTLER, 2021, p.142).

3.3 Marketing Preditivo

A modelagem preditiva já é utilizada e não se pode considerar como uma inovação. Os profissionais de *marketing* que se direcionam por dados criam modelos regressivos na identificação casual entre ações e resultados. Quando se faz o uso do

aprendizado de máquina, não é necessário algoritmos dos cientistas de dados que identificam padrões e modelagem de dados. Geralmente os modelos preditivos que resultam do aprendizado de máquina tendem a parecer que são modelos que vão além do entendimento humano. O profissional de *marketing* não mais se sujeita a suposições embasadas no passado na tentativa de previsão do futuro (KOTLER, 2021).

A análise preditiva permite que as empresas prevejam a movimentação mercadológica antes que ele ocorra. Anterior a fase preditiva, os profissionais de *marketing* se baseavam na estatística descritiva, que apresentava previsões baseadas no comportamento do passado, e usavam em conjunto a própria intuição para fazer previsões bem informadas sobre o futuro. Na análise preditiva, a *Artificial Intelligence*, Inteligência artificial (IA) fica com a maior parte do trabalho. Os dados passados para ela são consumidos em um motor de aprendizado de máquina para que ela consiga identificar e mostrar padrões específicos, o que é chamado de modelo preditivo. Ao abastecer de dados, o profissional de *marketing* pode prever resultados, como por exemplo quem tem mais probabilidade de comprar; que produto será vendido ou qual campanha vai trazer melhores resultados. Como o *marketing* preditivo é fortemente baseado em dados, geralmente as empresas adquirem essa funcionalidade por meio do ambiente de dados previamente implantado (KOTLER, 2021).

A partir dessas previsões, as empresas podem se organizar proativamente nos seus investimentos para o que se pode ocorrer futuramente. Conseguem, por exemplo, prever quais clientes de transações de porte pequeno têm chances de serem clientes de grande porte. Portanto, se faz o impulsionamento de investimentos no crescimento de clientes específicos.

Antes de alocar muitos recursos ao desenvolvimento de um novo produto, as empresas também podem usar a análise de dados preditiva para ajudar na filtragem de ideias. No fim das contas, a análise de dados preditiva leva a um retorno melhor do investimento em *marketing* (KOTLER, 2021, p.146).

Para montar um modelo de *marketing* preditivo é muito importante que, além dos profissionais de *marketing*, os cientistas de dados, os matemáticos e estatísticos estejam no processo de criação e desenvolvimento dos modelos. Sendo assim, os profissionais de *marketing* não necessitam de compreensão detalhada dos modelos estatísticos. Contudo, é necessário que os mesmos tenham uma compreensão,

mesmo que rasa, das ideias básicas em que os modelos preditivos se baseiam, fazendo com que consigam trazer orientações às equipes estatísticas e de dados na melhor escolha dos padrões e dados que serão utilizados no processo preditivo. Além disso, ele também ajudará na interpretação do modelo gerado, auxiliando no desdobramento das previsões em operações fatídicas (KOTLER, 2021).

Para Kotler (2021) tem-se os modelos regressivos para previsões simples, em que os profissionais de *marketing* quando utilizam a análise regressiva buscam modelos estatísticos que consigam explicar o relacionamento entre as variáveis dependentes e variáveis independentes. Portanto, ele busca entender quais ações de *marketing* têm maior relevância e quais são as melhores ações que trazem os melhores resultados para a empresa (KOTLER, 2021).

Segundo Kotler (2021) o modelo de regressão, se comparado com outras formas de modelagem, consegue observar a sua maior popularidade. A análise regressiva geralmente é muito utilizada em aplicações de *marketing*, por exemplo, na modelagem de propensão, modelagem identificadora de rejeição e na modelagem de afinidade de produtos. Normalmente, a modelagem de regressão é realizada em várias fases:

- Identificação e seleção dos dados de variáveis dependentes e variáveis independentes: Para a análise regressiva, de conjuntos de dados destes dois tipos de variáveis deve-se reunir simultaneamente todas elas tendo também uma amostragem suficiente para se analisar. O profissional de *marketing* consegue avaliar o efeito da cor de um *flyer* digital nos contadores de cliques quando coletado uma amostra suficiente o bastante de cores e as informações resultantes dos cliques;
- Definição da equação que demonstra a relação entre as variáveis: Através de um programa estatístico, o profissional de *marketing* consegue montar a equação mais adequada em relação aos dados. A equação que consegue formar uma linha reta é considerada como mais simples, chamada de linha de regressão linear. Também é comum a utilização da regressão logística, na qual utiliza funções da área logística para gerar uma variável binária dependente, como negativo/positivo ou permanece/rejeita. Portanto, a regressão logística é muito utilizada na previsão provável de determinado fim, como a probabilidade de fidelização, fidelizado/não fidelizado;

- A Interpretação da equação para identificar novas ideias e verificar a assertividade da mesma. Tem-se um exemplo na Equação 4 definida na forma:

$$Y = b + cX + dZ^2eW^3 + f \quad (4)$$

Na Equação 4, Y é a variável dependente, ou seja ela depende de outras variáveis da equação, ao passo que X , Z e W são as variáveis independentes. Tem-se o valor de Y como o reflexo de b , caso não se tenha valores nulos nas variáveis independentes. O valores de c , d e f são os coeficientes das variáveis independentes, que indicam até que ponto as variáveis independentes influenciam no valor da variável dependente. Na Equação 4, portanto, tem-se um termo que representa o erro, expresso por f . Uma modelagem regressiva precisa gerar o coeficiente de erro, porque não é sempre que as variáveis independentes conseguem explicar o valor da variável dependente. A precisão da equação está subordinado ao valor do erro. Quando o valor do erro for muito grande, a precisão da equação é menor.

- Previsão das variáveis dependentes em relação às variáveis independentes: O resultado das variáveis dependentes podem ser previstos embasados nas variáveis independentes a partir da definição da fórmula, portanto isso garante a previsão de resultado de ações de marketing.
- Seleção colaborativa para mapeamentos de recomendação: É fato que a preferência do consumidor na compra de produtos é o que se possui de semelhança com outros produtos que ele já comprou ou a produtos que outros compradores de preferências semelhantes já compraram. Para que esse mapeamento ocorra corretamente, a colaboração do cliente é de grande importância, avaliando os produtos. A seleção colaborativa não se aplica somente a produtos, também se aplica ao conteúdo, de acordo com a recomendação para o consumidor que o profissional de marketing pretende entregar (KOTLER, 2021).

Para resumir, Kotler (2021, p. 154) indica o modelo de filtragem colaborativa que é executado de acordo com sequência lógica a seguir :

1. Coleta de preferências de uma ampla base de consumidores :Para medir quantas pessoas preferem determinado produto, o profissional de marketing pode criar um sistema comunitário de notas no qual os clientes podem avaliar um produto, seja com um simples like/dislike (como no YouTube) ou uma nota de zero a cinco estrelas (como na

Amazon). Outra opção é propor ações que reflitam uma preferência, como a leitura de um texto, a visualização de um vídeo ou o acréscimo de produtos a uma lista de desejos ou um carrinho de compras. A Netflix, por exemplo, mede as preferências por meio dos filmes assistidos ao longo do tempo.

2. Agrupamento de consumidores e produtos similares: Consumidores que avaliaram tipos de produto semelhantes e que demonstraram comportamentos semelhantes podem ser colocados no mesmo cluster (grupo homogêneo). A suposição é que fazem parte dos mesmos segmentos psicográfico (com base nos likes/dislikes) e comportamental (com base em suas ações). Outra opção para o profissional de marketing é agrupar itens que tiveram avaliações semelhantes de um determinado grupo de clientes.
3. Previsão da provável avaliação que um cliente dará a um produto novo. Hoje o profissional de marketing pode prever a nota que o cliente dará a um produto que ele ainda não viu nem avaliou com base nas notas fornecidas por clientes que pensam parecido. Essa nota prevista é essencial para que o profissional ofereça os produtos corretos, que o cliente provavelmente apreciará e em relação aos quais tomará alguma atitude no futuro (KOTLER, 2021, p.154).

4 MODELO PROPOSTO

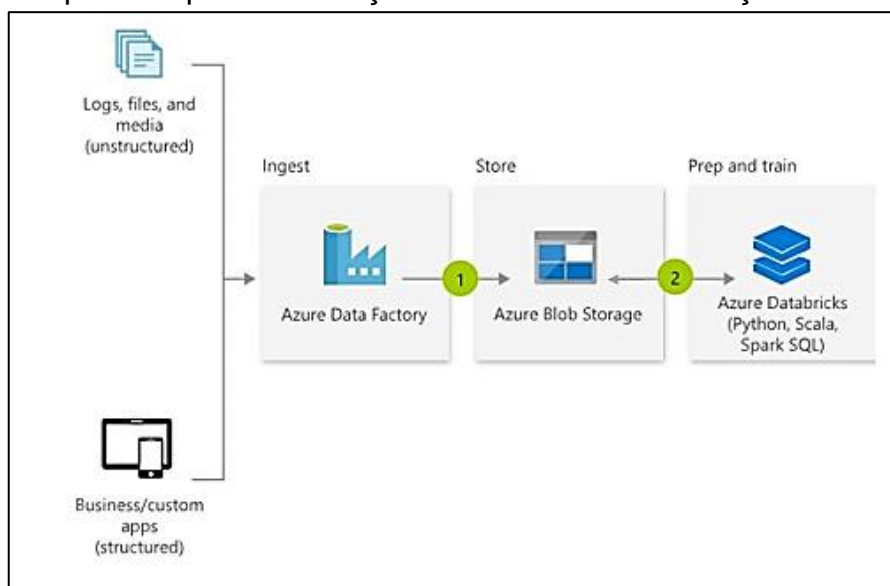
Nesta etapa foi utilizada a plataforma de nuvem *Azure* que contém cerca de 200 produtos e serviços e que são habilitados para a integração entre si. A escolha da plataforma foi feita pela familiaridade do autor deste trabalho com a mesma.

Foi utilizado os serviços e produtos fornecidos pela *Azure* a seguir:

- *Azure ActiveDirectory*;
- *Azure Blob Storage* – Containers;
- *Wordpress* como serviço de aplicação;
- *Azure DataBricks*;
- *Azure DataFactory*.

O modelo estrutural proposto para comportar o processo de execução do algoritmo para associação de itens de forma automatizada e está representado na Figura 6.

Figura 6 – Arquitetura para automação do modelo de associação de itens proposto



Fonte: Adaptado- (KRUTH,2018)

O objetivo deste trabalho está inserido na relação entre ciência de dados com o *marketing* direcionado por dados, não tendo como foco principal as ferramentas de *cloud* e os serviços disponibilizados nelas, porém se faz importante trazer conceitos dos recursos que foram utilizados neste trabalho para melhor compreensão do assunto.

4.1 Ambiente de Desenvolvimento

4.1.1 Azure

Através da abordagem *software* como serviço o usuário aluga um serviço fornecido pela *Azure*, no caso Wordpress sendo capaz então de configurar o serviço utilizando a *interface* da Azure, sem necessitar do conhecimento da infraestrutura utilizada que fornece este serviço . Essa abordagem é conhecida como *Software as a Service*, *Software* como Serviço (SaaS), no qual o cliente paga para utilizar serviços definidos. O cliente não controla nem administra o *hardware* no qual o serviço está instalado. Do mesmo modo, o cliente não tem controle sobre o sistema que executa o serviço (BRUNETTI, 2011).

Os processos de criação de uma conta e configuração do *Azure* estão contemplados conforme a descrição no Apêndice A.

Figura 7– Logotipo Azure



Fonte: Adaptado- (LOGOMARCA,2022)

O *Azure Blob Storage – Containers* é um serviço de armazenamento de dados da *Azure*, sendo eles estruturados ou não e que podem ser acessado por diversas ferramentas e serviços disponíveis na plataforma através de *linked service*. Foi criado um conta de armazenameto conforme as Figuras 8 e 9.

Figura 8 – Criação de conta de armazenamento *Azure Blob Storage*

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Figura 9 – Criação de contêiner de armazenamento na *Azure Blob Storage*

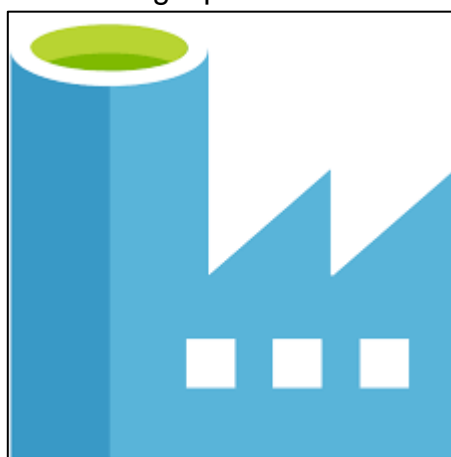
| Nome | Última modificação |
|----------------------------------|---------------------|
| <input type="checkbox"/> \$logs | 19/10/2022 02:52:47 |
| <input type="checkbox"/> load | 26/10/2022 14:40:57 |
| <input type="checkbox"/> trusted | 19/10/2022 02:53:35 |

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

O *Azure Data Factory (ADF)* , cuja logo está exibida na Figura 10, foi definido segundo a Microsoft (2022) como:

Azure Data Factory é o serviço ETL na nuvem do Azure para integração e transformação de dados sem servidor em expansão. Ele oferece uma interface do usuário livre de código para criação intuitiva e gerenciamento e monitoramento em painel único. Também é possível migrar pacotes SSIS existentes por lift-and-shift para o Azure e executá-los com total compatibilidade no ADF (Azure Data Factory). O Azure-SSIS Integration Runtime oferece um serviço totalmente gerenciado, de modo que não é necessário se preocupar com o gerenciamento da infraestrutura.

Figura 10 – Logotipo Azure Data Factory



Fonte: Adaptado- (VECTA,2022)

No ADF o primeiro passo é criar uma atividade no serviço de cópia de dados entre conexão do servidor e banco de dados do *Wordpress*. Sendo assim, capaz de extrair a tabela de registro de compras e trazer para o serviço de *container* em formato “.csv” . Em seguida, é feita a conexão entre o *Databricks* no qual é feito a chamada do *notebook* que irá executar o algoritmo de recomendação e trazer o resultado obtido através do algoritmo em um *container*.

Os processos de instalação e configuração do ADF estão contemplados conforme a descrição no Apêndice C.

O *Azure Databricks*, cuja logo está exibida na Figura 11, foi definido segundo a *Databricks* (2020) como:

O *Azure Databricks* é uma poderosa solução unificada de *Big Data* e de processamento escalável, que oferece inúmeros ganhos em relação a escalabilidade e performance. Ele oferece também a possibilidade de escrever em qualquer linguagem, como R, Python, Scala e SQL.

Figura 11 – Logotipo Azure DataBricks



Fonte: Adaptado- (ICON-ICONS, 2022)

Após a criação do serviço do *DataBricks*, é necessário criar um *cluster* que nada mais é que o processador responsável por executar todas as tarefas de um *notebook* por exemplo.

Foi criado um *notebook* que faz a conexão entre o *databricks* com o serviço de *container* da Azure, podendo assim carregar o arquivo que é utilizado nesse trabalho, no caso, um relatório de compras com todos os itens por compras. Após essa conexão o arquivo é carregado para se conseguir o objetivo através do algoritmo *Apriori*, um algoritmo de recomendações de compras de acordo com o histórico. O resultado é carregado dentro de outro diretório no *container* em arquivo CSV no qual será possível identificar as interseções mais relevantes.

O *Wordpress*, cujo logotipo é exibido na Figura 32, é um gerenciador de *sites*, *blogs*, lojas virtuais, portais de notícia, áreas de membros, que tem como objetivo facilitar a criação e edição de conteúdos em um *site* sem a necessidade do uso de linguagem de programação, democratizando o desenvolvimento de grandes *sites* para quem tem uma empresa ou demais idéias.

Figura 12 – Logotipo *Wordpress*



Fonte: Adaptado- (LOGOMARCAS,2022)

O *site* criado nesse trabalho a partir da ferramenta *Wordpress* é :
<https://mercadogioiano.azurewebsites.net/>

4.2 Implementação do Algoritmo de associação de produtos – *Apriori*

O algoritmo *Apriori* foi implementado em linguagem *Python*, que é muito utilizada na área de dados e ela não é exclusiva para análise de dados, o que favorece o uso em companhias como um todo, pois, as maiorias das equipes de

desenvolvimentos já possuem um certo conhecimento em *Python* facilitando a sua implantação (LOPES E ALMEIDA,2019).

As características mais relevantes do *Python* são a sua capacidade de integração com outras linguagens e seu sistema de bibliotecas robusto. Algumas bibliotecas do *Python* são fortemente analíticas e oferecem ferramentas de ciência de dados composta de funções altamente otimizadas para trabalhar, configuração ideal de memória, pronta para realizar operações de *script* com desempenho ideal (LOPES E ALMEIDA,2019).

Algumas bibliotecas do *Python* foram usadas para a realização deste trabalho conforme o Quadro 4:

Quadro 4 – Bibliotecas *Python* utilizadas no algoritmo *Apriori*

| Biblioteca | Descrição |
|-------------------|--|
| <i>Pandas</i> | A biblioteca tem a licença em código aberto, fornecendo arquiteturas de dados em alto desempenho e fácil utilização, direcionada a análise de dados. |
| <i>Numpy</i> | A biblioteca é utilizada geralmente na realização de cálculos em vetores e matrizes, além disso, no auxílio em se executar cálculos numéricos. |
| <i>Mlxtend</i> | A biblioteca foi desenvolvida por Sebastian Raschka, professor na Universidade de Wisconsin, ela apresenta muitas funções importantes para análise de dados e rotinas de aprendizado de máquina. |

Fonte: Lopes e Almeida, 2019

Após a conclusão dos procedimentos descritos dos Apêndices A e B desse trabalho, foi possível utilizar o serviço *DataBricks* para alcançar o objetivo deste trabalho.

4.2.1 Coleta de dados

Toda a construção da base de dados de registros de compras utilizada nesse trabalho foi simulada de acordo com o Apêndice F.

Inicialmente é executado o *notebook* responsável para a integração entre o *storage* e o *DataBricks* conforme ilustra a Figura 13.

Figura 13 – Execução do *notebook mount_point*

The screenshot shows a notebook titled "Itens Recomendados - Apriori" in Python. The command executed is `%run /Shared/mount_point`. The output indicates that the command took 41.30 seconds and that the `/mnt/registro` file system has been unmounted. The output also shows `Out[29]: True` and `Out[30]: True`.

```

Itens Recomendados - Apriori Python
File Edit View Run Help Last edit was 9 days ago
Cmd 1
1 %run /Shared/mount_point
Command took 41.30 seconds -- by cienciaDados@outlook.com
/mnt/registro has been unmounted.
Out[29]: True
Out[30]: True

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Em seguida, é feita a importação de duas bibliotecas que serão utilizadas no tratamento, limpeza e transformação dos *dataframes* e tabelas geradas conforme Figura 14.

Figura 14 – Importação das bibliotecas *Numpy* e *Pandas* no *Python*

The screenshot shows a notebook titled "Cmd 2" with two lines of Python code: `import numpy as np` and `import pandas as pd`. The output indicates that the command took 0.02 seconds.

```

Cmd 2
1 import numpy as np
2 import pandas as pd
Command took 0.02 seconds -- by cienciaDados@outlook.com at 14/11/2022 01:55:52 on Ciencia Dados's Cluster

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

No bloco seguinte conforme a Figura 15, foi criado uma variável que contém o banco de registros de compras simulado criado a partir do procedimento descrito no Apêndice F. O arquivo é nomeado como: "Registros_compra.csv".

Figura 15 – Variável *filepath* com o diretório da tabela de registros de compras

The screenshot shows a notebook titled "Cmd 3" with one line of Python code: `filepath = r'/dbfs/mnt/extracao/Registros_compra.csv'`. The output indicates that the command took 0.02 seconds.

```

Cmd 3
1 filepath = r'/dbfs/mnt/extracao/Registros_compra.csv'
Command took 0.02 seconds -- by cienciaDados@outlook.com at 14/11/2022 02:09:44 on Ciencia Dados's Cluster

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

No próximo bloco de código é criado um *dataframe* que recebe o arquivo que está neste diretório, conforme Figura 16.

Figura 16 – Leitura da tabela de registros e sua gravação em um *dataframe*

```
Cmd 5
1 registros = pd.read_csv(filepath,sep=';',header=None)
2 registros.replace(np.nan, 0,inplace=True)
3

Command took 0.61 seconds -- by ciencdados@outlook.com at 14/11/2022 02:09:48 on Ciencia Dados's Cluster
Cmd 6
```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

4.2.2 Tratamento de dados

Nessa etapa foi criado o método que remove todos os registros nulos conforme Figura 17.

Figura 17 – Método *removeAllOcorrenciasOfValueInList*

```
Cmd 7
1 def removeAllOcorrenciasOfValueInList(_list, value):
2     return list(filter(lambda x: x != value, _list))

Command took 0.03 seconds -- by ciencdados@outlook.com at 14/11/2022 02:09:55 on Ciencia Dados's Cluster
Cmd 8
```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Foi criado uma lista, conforme Figura 18, no qual cada campo da lista representa um registro/transações de compra única e fazendo o uso do método criado no bloco já exibido na Figura 17 que retira os produtos considerados como nulo.

Figura 20 – Utilização do método *TransactionEncoder*

```

Cmd 19
1  from mlxtend.preprocessing import TransactionEncoder
2
3  te = TransactionEncoder()
4  #http://rasbt.github.io/mlxtend/user_guide/preprocessing/TransactionEncoder/
5  te_ary = te.fit(list_all_transactions).transform(list_all_transactions)
6
7  te_ary
8  df = pd.DataFrame(te_ary, columns=te.columns_)
9
10 df

```

Out[37]:

| | ALFACE | ARROZ | AÇÚCAR | BALINHA | BATATA | BISCOITO DE QUEIJO | CAFÉ | CAFÉ INSTANTÂNEO | CAJU | CARNE BOVINA |
|------|--------|-------|--------|---------|--------|--------------------------|-------|---------------------|-------|-----------------|
| 0 | False | False | False | False | False | False | False | False | False | False |
| 1 | False | False | False | False | False | False | False | False | False | False |
| 2 | False | False | False | False | False | False | False | False | False | False |
| 3 | False | False | False | False | False | False | False | False | False | False |
| 4 | False | False | False | False | False | False | False | False | False | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 3912 | False | False | False | False | False | False | False | False | False | False |
| 3913 | False | False | False | False | False | False | False | False | False | False |
| 3914 | False | False | False | False | False | False | False | False | False | False |
| 3915 | False | False | False | False | False | False | False | False | False | False |
| 3916 | False | False | False | False | False | False | False | False | False | False |

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

4.2.3 Transformação dos dados

Utilizando da *Mlxtend* e de suas facilidades para uso dos recursos da ciência de dados é feita a importação do Apriori e as regras de associação.

A variável *frequent_itemsets* recebe o *dataframe* resultante do algoritmo *Apriori* de acordo com o *min_suport* que pode ser considerado com a porcentagem que o item e a associação de itens está representado no banco de dados. É interessante perceber que quanto mais itens presente no banco de dados, menor deve ser considerado o valor do suporte mínimo pois o registro de cada item em todo o banco têm sua representatividade menor.

Na Figura 21, pode-se identificar que os 10 itens e associações mais presentes no banco são:

- Leite UHT : Em torno de 48,9 % das transações de compra contém o item;
- Tomate : Em torno de 48 % das transações de compra contém o item;
- Arroz : Em torno de 25 % das transações de compra contém o item;
- Água : Em torno de 22,5 % das transações de compra contém o item;
- Pão Baguete : Em torno de 16 % das transações de compra contém o item;
- Refrigerante: Em torno de 13 % das transações de compra contém o item;
- Arroz e Leite UHT : Em torno de 13 % das transações de compra contém;
- Iogurte : Em torno de 12,8 % das transações de compra contém o item;
- Leite UHT e Tomate : Em torno de 12,4 % das transações de compra contém;
- Arroz e Tomate : Em torno de 11 % das transações de compra contém o item.

Figura 21 – Resultado dos 10 primeiros itens acima do suporte mínimo do algoritmo

```

Cmd 11
1 from mlxtend.frequent_patterns import apriori, association_rules
2
3 frequent_itemsets = apriori(df, min_support = 0.1, use_colnames = True)
4 frequent_itemsets.sort_values(by=['support'], ascending = False).head(10)

```

Out[38]:

| | support | itemsets |
|----|----------|---------------------|
| 2 | 0.488895 | (LEITE UHT) |
| 6 | 0.481491 | (TOMATE) |
| 0 | 0.253255 | (ARROZ) |
| 7 | 0.224862 | (ÁGUA ENGARRAFADA) |
| 4 | 0.159561 | (PÃO BAQUETE) |
| 5 | 0.139903 | (REFRIGERANTE) |
| 8 | 0.132244 | (ARROZ, LEITE UHT) |
| 1 | 0.128870 | (IOGURTE) |
| 10 | 0.124075 | (LEITE UHT, TOMATE) |
| 9 | 0.118458 | (ARROZ, TOMATE) |

Command took 0.04 seconds -- by cienciaDados@outlook.com at 14/11/2022 02:10:20 on Ciência Dados's Cluster

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

As associações de compras presentes nesse banco de dados, são apenas conjunto entre 2 itens:

- Leite UHT e Tomate;
- Arroz e Tomate;

- Arroz e Leite.

É utilizado conforme a Figura 22 a métrica da *association_rules* chamada de *confidence* que verifica qual a confiança, ou seja, dentro dos itens frequentes identificados de acordo com o suporte mínimo quais associações tem a representatividade estatística mais relevante de acordo com o *min_threshold* definido.

Figura 22 – Utilização da regra de associação confiança no algoritmo *Apriori*

```

Cmd 12
1  from mlxtend.frequent_patterns import association_rules
2
3  rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=0.0)
4  rules=rules.sort_values(by=['lift'], ascending = False)
5  # .drop(['antecedent support', 'consequent support', 'leverage', 'conviction'], axis=1)

Command took 0.03 seconds -- by cienciaDados@outlook.com at 14/11/2022 02:10:33 on Ciencia Dados's Cluster

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

O resultado dos itens frequentes é representado na Figura 23 :

Figura 23 – Itens frequentes com seus suportes e confiança

| A | B | C | D | E |
|---|--------------------------|--------------------------|---------------------|---------------------|
| | antecedents | consequents | support | confidence |
| 0 | frozenset({'ARROZ'}) | frozenset({'LEITE UHT'}) | 0.13224406433495023 | 0.5221774193548387 |
| 1 | frozenset({'LEITE UHT'}) | frozenset({'ARROZ'}) | 0.13224406433495023 | 0.27049608355091387 |
| 2 | frozenset({'ARROZ'}) | frozenset({'TOMATE'}) | 0.1184580035741639 | 0.46774193548387094 |
| 3 | frozenset({'TOMATE'}) | frozenset({'ARROZ'}) | 0.1184580035741639 | 0.24602332979851538 |
| 4 | frozenset({'LEITE UHT'}) | frozenset({'TOMATE'}) | 0.12407454684707685 | 0.25378590078328983 |
| 5 | frozenset({'TOMATE'}) | frozenset({'LEITE UHT'}) | 0.12407454684707685 | 0.2576882290562036 |

Fonte: Tela de captura do *Excel* com conteúdo desenvolvido pelo autor deste trabalho

Pode-se identificar que :

- Arroz e Leite UHT: seja os itens antecessores entre si que estão presentes em cerca de 13,22% e dentro dessa associação o nível de confiança é cerca de 52,21% ou 27% caso o antecessor seja o Leite UHT;
- Arroz e Tomate: seja os itens antecessores entre si que estão presentes em cerca de 11,8% das compras em geral e dentro dessa associação o nível de confiança é cerca de 52,21% ou 24,6% caso o antecessor seja o Tomate;

- Leite UHT e Tomate : seja os itens antecessores entre si estão que presentes em cerca de 12,4% das compras em geral e dentro dessa associação o nível de confiança é cerca de 25,3% ou 25,7% caso o antecessor seja o Tomate.

4.2.4 Carregamento dos resultados

Na última etapa de execução do *notebook* é feito a conversão do *dataframe* que contém o resultado do algoritmo em um arquivo “.csv” e gerado o arquivo dentro do *storage* no *container* “*extracao/Recomendados.csv*”, conforme mostra as Figuras 24 e 25.

Figura 24 – Indicação do diretório para salvar o resultado obtido no *storage*

```

Cmd 15
1 filepath2=r"/dbfs/mnt/extracao/Recomendados.csv"

Command took 0.02 seconds -- by ciencdados@outlook.com at 14/11/2022 02:10:47 on Ciencia Dados's Cluster

Cmd 16
1 resultado.to_csv(filepath2, encoding = 'utf-8')

Command took 0.10 seconds -- by ciencdados@outlook.com at 14/11/2022 02:10:50 on Ciencia Dados's Cluster

Shift+Enter to run

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Figura 25 – Diretório “Registro” com o arquivo “Recomendados.csv”

Página inicial > Iakemercado | Contêineres >

registro ...
Contêiner

Pesquisar << Carregar + Adicionar Diretório Atualizar Renomear Excluir Alterar a camada Adquirir concessão Interromper concessão

Visão geral

Diagnostico e resolver problemas

IAM (Controle de Acesso)

Método de autenticação: Tecla de acesso (Alternar para a Conta de Usuário do Azure AD)

Local: registro

Pesquisar blobs por prefixo (diferenciar maiúsculas de minúsculas) Mostrar os objetos que foram excluídos

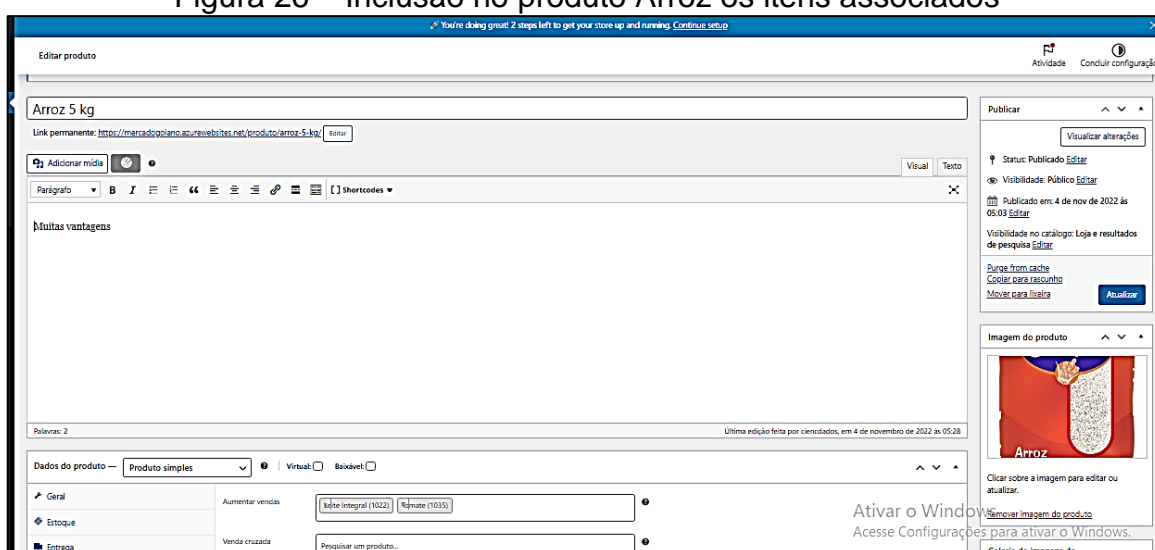
| Configurações | Nome | Modificado | Camada de acesso | Status do arquivo | Tipo de blob | Tamanho | Estado de concessão |
|-------------------------------------|--------------------------------|---------------------|-----------------------|-------------------|----------------|------------|---------------------|
| <input type="checkbox"/> | EXPORTACAO_GYMARKET_geral2.csv | 04/11/2022 12:51:22 | Principal (Inferidos) | | Blob de blocos | 12.57 KiB | Disponível |
| <input type="checkbox"/> | novo1.txt | 05/11/2022 02:45:33 | Principal (Inferidos) | | Blob de blocos | 65.73 KiB | Disponível |
| <input checked="" type="checkbox"/> | Recomendados.csv | 14/11/2022 01:49:17 | Principal (Inferidos) | | Blob de blocos | 1.19 KiB | Disponível |
| <input type="checkbox"/> | Registros_csv | 05/11/2022 01:14:48 | Principal (Inferidos) | | Blob de blocos | 359.85 KiB | Disponível |
| <input type="checkbox"/> | Registros_txt | 05/11/2022 01:31:13 | Principal (Inferidos) | | Blob de blocos | 359.85 KiB | Disponível |
| <input type="checkbox"/> | Registros_ajustado.csv | 05/11/2022 03:31:30 | Principal (Inferidos) | | Blob de blocos | 308.84 KiB | Disponível |
| <input type="checkbox"/> | Registros_ajustado2.csv | 05/11/2022 03:46:09 | Principal (Inferidos) | | Blob de blocos | 242.18 KiB | Disponível |
| <input type="checkbox"/> | Registros_ajustado3.csv | 05/11/2022 04:04:09 | Principal (Inferidos) | | Blob de blocos | 145.24 KiB | Disponível |
| <input type="checkbox"/> | Registros.csv | 04/11/2022 23:40:28 | Principal (Inferidos) | | Blob de blocos | 420.11 KiB | Disponível |
| <input type="checkbox"/> | Registros1.csv | 05/11/2022 01:09:15 | Principal (Inferidos) | | Blob de blocos | 420.05 KiB | Disponível |

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

4.3 Atualização de itens “Você também pode gostar de” no site Mercado Goiano

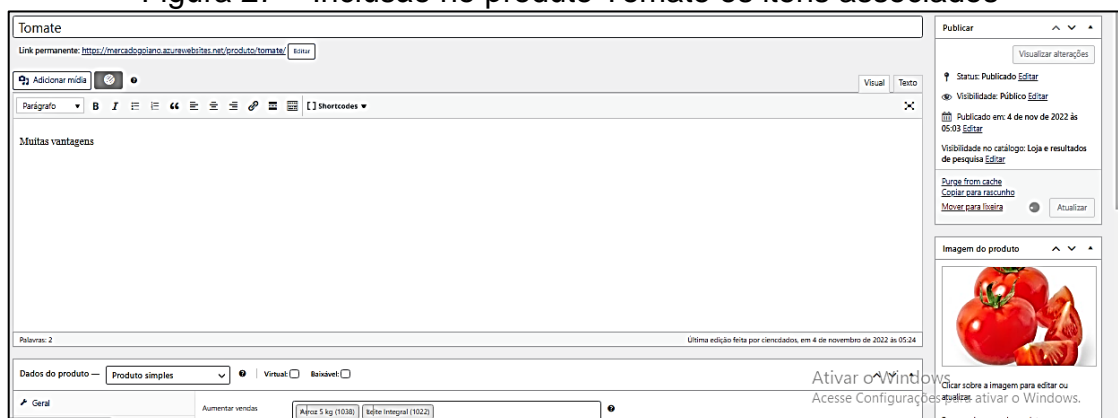
A partir desse resultado de associações de compra e feito o acesso ao *site* “mercadogoiano” com o perfil de administrador e selecionado as opções: “Produto->Todos Produtos” para selecionar os itens que foram indicados no resultado, Arroz, Tomate e Leite UHT. Ao selecionar cada item indicado e editado o campo de “Aumentar Vendas” em cada item é inserido as associações resultantes conforme ilustrado nas Figuras 26, 27 e 28 .

Figura 26 – Inclusão no produto Arroz os itens associados



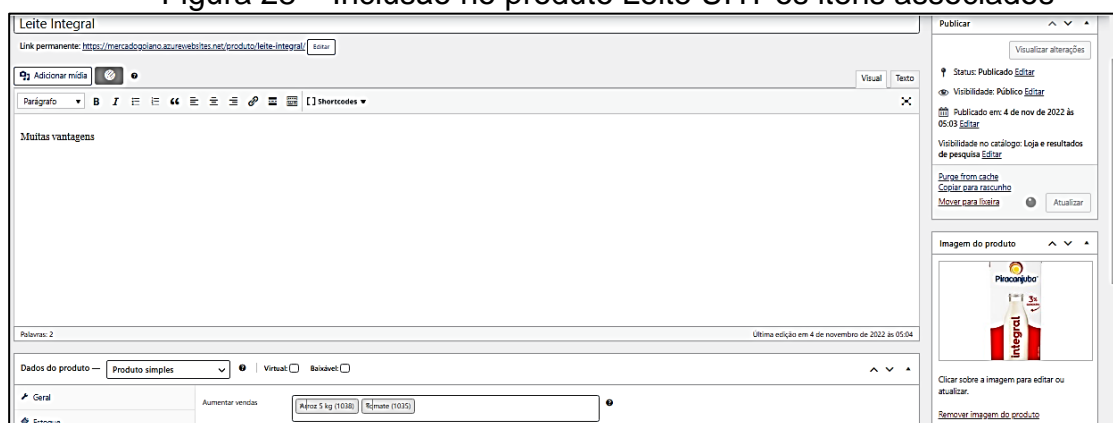
Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

Figura 27 – Inclusão no produto Tomate os itens associados



Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

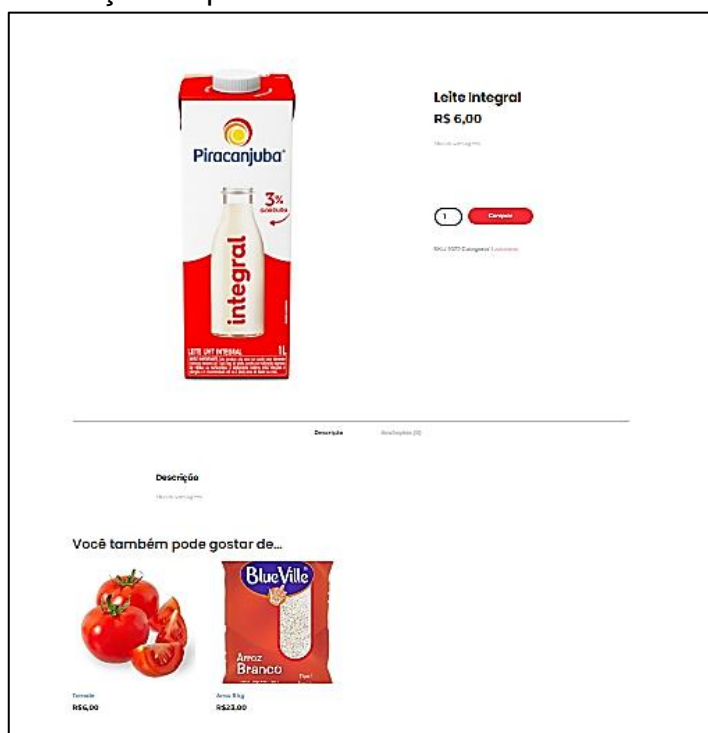
Figura 28 – Inclusão no produto Leite UHT os itens associados



Fonte:

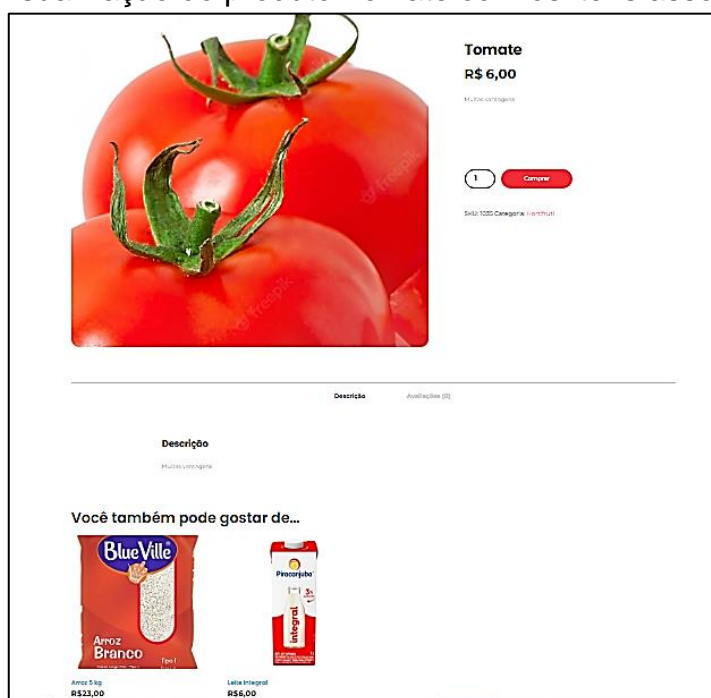
Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

O campo “Aumentar Vendas” preenchido e atualizado conforme os resultados das associações é refletido no *site* Mercado Goiano. Quando é selecionado um dos produtos informados anteriormente a página “Mercado” apresenta os itens relacionados com o título: “Você também pode gostar de” conforme ilustrado nas Figuras 29 e 30.

Figura 29 – Visualização do produto Leite UHT com os itens associados no *site*

Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

Figura 30 – Visualização do produto Tomate com os itens associados no site



Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

5 CONSIDERAÇÕES FINAIS

Com a evolução da análise de dados inicialmente alavancada em BI se teve um maior contato das companhias com uma orientação analítica originada pelos dados da própria empresa. A partir disto, o contato se tornou crescente junto com a evolução tecnológica surgindo assim conceitos como: *Big Data*, *Cloud*, Mineração de dados, *Machine Learning* e *Data Science*.

Para se utilizar o *Data Science* é necessário obter várias competências de diversas áreas como: programação, matemática, estatística, entre outras. O foco principal deste trabalho é a relação entre sistemas de recomendações de compra e o *marketing*.

O sistema de recomendações busca identificar a partir de métodos estatísticos e matemáticos a preferência de um cliente com relação a um item seja ela por vários tipos de abordagem como: colaborativa, baseada em conteúdos ou um modelo híbrido.

Pode-se fazer uso dos dados históricos que geralmente são inutilizáveis dentro de uma companhia para identificar compras de itens associados, impulsionando “vendas casadas”, ajuste de *layout* físico da loja para melhor adequar nos produtos selecionados juntos em uma compra por exemplo.

Esses sistemas estão cada vez mais presentes e são aplicações interessantes para o entendimento e compreensão. Portanto são sistemas que impulsionam uma experiência entre o cliente e a empresa mais interessante, entrando em contato com os conteúdos que mais se relacionam com os clientes e que têm mais chance de serem assistidos ou consumidos.

O *marketing* direcionado por dados é o principal estágio do *marketing* 5.0, no qual todo o processo de evolução do *marketing* muito alinhado com a tecnologia e o momento que a sociedade está inserida. Esse estágio, o *marketing* direcionado por dados, tem muita importância dentro de uma companhia, pois se têm uma extenso ambiente de dados de diversos tipos que muita das vezes são inutilizáveis por serem considerados dados defasados, mas podem ser capazes de direcionar alguma ação de *marketing* ou não e até trazer um indicativo futuros de resultados.

Portanto, conclui-se que através do estudo que é possível utilizar o algoritmo *Apriori*, um algoritmo muito utilizado em *Data Science* na definição de associações de produtos junto com suas regras de associação identificar a partir de um banco de

dados de registro de compra de itens de uma companhia para se recomendar possíveis compras de outros itens associados de acordo com a sua representatividade no todo, utilizando-se de definições de suporte e confiança estabelecidas pela companhia ou que fazem sentido para a análise.

Deve-se frisar que o ambiente deste trabalho é no âmbito de desenvolvimento, portanto fictício, pois não foi possível o acesso e utilização de uma base de dados real de um mercado. Foram considerados apenas 55 itens e menos de 10 mil registros de compras, valores muito distante do que se têm na realidade, trazendo então associações questionáveis para serem analisadas.

5.1 Dificuldades encontradas

Ocorreu dificuldades no decorrer deste trabalho. Uma delas foi que, inicialmente o provedor utilizado para o *site* chamado de *InfinityFree* apresentou uma alta latência para executar e apresentar as imagens e também no carregamento do próprio *site*, portanto para a solução do problema foi migrado todo a criação do *site* diretamente no *Wordpress* como serviço na plataforma *Azure*.

Um outro fator que ocasionou dificuldades foi que para o serviço de DataBricks conseguir acessar o serviço de armazenamento da *Azure* era necessário a implementação do *Azure AD*, porém, a conta disponibilizada gratuitamente pelo *Azure* não contempla o serviço do *Azure AD*. Foi migrado a conta *Azure* de gratuita para conta paga de acordo com o uso para sanar esta dificuldade.

Encontrar uma base de dados com registros de compras foi uma dificuldade encontrada e que não foi resolvida, visto que, foi necessário utilizar uma base fictícia de registros de compra relativamente pequena com possíveis itens oferecidos em um mercado de pequeno porte.

5.2 Sugestão de trabalhos futuros

Como sugestão de trabalhos futuros, propõe-se:

- Implementação de uma ação capaz de carregar o resultado obtido dos itens associados e inserir de maneira automática no produto do *site*;
- Criação de campanhas de *marketing* a partir dos resultados obtidos em um período de tempo pré definidos;

- Implementação do algoritmo no ambiente de produção para análise dos resultados.

REFERÊNCIAS

BARBIERI, Carlos. **BI-Business Intelligence**: modelagem & tecnologia. Rio de janeiro: Axcel Books, 2001.

BERRY, M. J; LINOFF, G. **Data Mining Techniques**: For *Marketing*, Sales, and Customer Support. John Wiley & Sons, Inc., New York, NY, USA, 1997.

BOSCHETTI, Alberto; MASSARON, Luca. **Python Data Science Essentials**, 2. ed. Birmingham – UK: Editora Packt, 2016

BRUNETTI, Roberto. **Windows Azure Step by Step**. Estados Unidos: On ne Tra n ng So ut ons, Inc, 2011. 345 p.

CASAS, A. L. L., **Marketing de serviços**. 5. Ed. São Paulo: Atlas, 2007.

COELHO, Lucas. **Ciência de dados**: o que é, conceito e definição. Disponível em 2017 < <https://www.cetax.com.br/blog/data-science-ou-ciencia-dedados/> > acesso em 13 mai. 2022.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4 ed. São Paulo: Atlas, 2002.

ICON. **DataBricks Logo**. Disponível em:< <https://icon-icons.com/pt/icone/databricks-logo/169299>> acesso em: 08 nov. 2022.

KOTLER, Philip. **Marketing 3.0**: as forças que estão definindo o novo *marketing* centrado no ser humano. 1ª. Ed. Rio de janeiro: elsevier, 2010.

KOTLER, Philip. **Marketing 4.0**: do tradicional ao digital. 1. Ed. Rio de janeiro: Sextante, 2017.

KOTLER, Philip. **Marketing 5.0**: tecnologia para a humanidade. 1ª. Ed. Rio de janeiro: Sextante, 2021.

KOTLER, Philip; **Princípios de marketing**. Rio de Janeiro: Prentice-Hall, 1993.

KRUTH. **Using Azure Databricks with Azure Data Factory**. Disponível em: <<https://www.kruth.fi/uncategorized/using-azure-databricks-with-azure-data-factory/>> acesso em: 08 nov. 2022.

LEE, Juno; Udacity. **Análise de Dados**. Disponível em <https://www.youtube.com/watch?v=qdV4sifMmWI>. Acesso em: 23 de maio de 2018.

LOGOMARCAS. **Microsoft Azure Logo**. Disponível em:<<https://logosmarcas.net/microsoft-azure-logo/>> acesso em: 08 nov. 2022.

LOPES,ALMEIDA. Delbem and, C. F. M. Toledo. 2019. “**Introdução à Análise Exploratória de Dados com Python**”. In: Escola Regional de Computação Aplicada à Saúde (ERCAS 2019).

MARQUESONE, Rosangela. **Big Data** : Técnicas e tecnologias para extração de valor de dados. São Paulo: Casa do Código, 2016.

MICROSOFT, Learn. **Data Factory**. Disponível em: <<https://learn.microsoft.com/pt-br/azure/data-factory/> 26/10> acesso em : 30 out. 2022.

MICROSOFT, Learn. **Wordpress**. Disponível em: <<https://learn.microsoft.com/pt-br/azure/app-service/quickstart-wordpress>> acesso em : 30 out. 2022.

OLAVSRUD, Thor. **Afinal o que é Ciência de Dados e o que isso tem a ver com a profissão do futuro**. Disponível em <http://idgnow.com.br/carreira/2018/07/05/afinal-o-que-e-ciencia-de-dados-e-o-que-isso-tem-a-ver-com-as-profissoes-do-futuro/> Acesso em: 10 de janeiro de 2018.

PETSI. **Algoritmos de recomendação: mais presente no dia a dia do que você pensa**. Disponível em < <http://www.each.usp.br/petsi/jornal/?p=2684>> . Acesso em: 12 nov. 2022.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2. ed. Rio Grande do Sul: Editora Feevale, 2013.

PROVOST, Foster; Fawcett, Tom. **Data Science para negócios: o que você precisa saber sobre mineração de dados e pensamento analítico de dados**. Rio de Janeiro: alta books, 2016.

ROCKET. **Infinity Free**. Disponível em <<https://rockcontent.com/br/blog/hospedagem-gratis/#:~:text=A%20InfinityFree%20%C3%A9%20uma%20hospedagem,nem%20na%20largura%20de%20banda>> acesso em: 22 ago. 2022.

SETZER, Waldemar W. **Dado, Informação, Conhecimento e Competência, Revista de Ciência da Informação**, n.0, dezembro, 1999, artigo 1. Disponível em:<<https://www.ime.usp.br/~vwsetzer/datagrama.html>> Acesso em: 26 ago 2022.

STORM. **Azure Databricks**. Disponível em:<<https://storm.kumulus.com.br/azure-databricks/#:~:text=O%20Azure%20Databricks%20%C3%A9%20uma,%2C%20Python%2C%20Scala%20e%20SQL>> acesso em: 10 out. 2022.

VANDERPLAS, Jake. **Python Data Science Handbook**, 1. ed. Sebastopol – Estados Unidos: Editora O’Reilly Media, 2016.

VASCONCELOS E CARVALHO. **Aplicação de Regras de Associação para Mineração de Dados na Web**. Disponível em: <https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_004-04.pdf> acesso em: 10 nov. 2022.

VECTA. **Data Factory Logo**. Disponível em:< <https://vecta.io/symbols/28/microsoft-azure-color/27/data-factory>> acesso em: 08 nov. 2022.

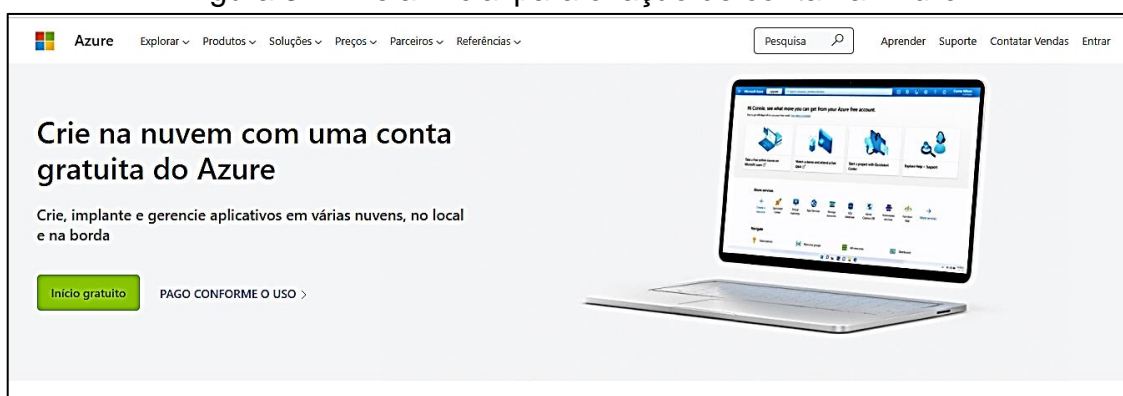
WAZLAWICK, Raul Sidnei. **Metodologia de Pesquisa para Ciência da**

Computação. 2. ed. Rio de Janeiro: Elsevier, 2014.

APÊNDICE A – CRIAÇÃO DE CONTA NA AZURE

Primeira passo é acessar o *site*: <https://azure.microsoft.com/pt-br/free/> e selecionar a opção ‘PAGO CONFORME O USO’ conforme a Figura 31, pois a versão gratuita não permite a criação de um recurso do *Azure* chamado de *Azure Active Directory* que é um dos pré-requisitos para se criar o recurso de *Azure DataBricks* que será responsável pela execução do algoritmo *Apriori*.

Figura 31 – Tela inicial para criação de conta na *Azure*



Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa apresenta-se uma tela com formulário que solicita um telefone para verificação de conta conforme a Figura 32:

Figura 32 – Tela de validação de celular para criação de conta na *Azure*

A imagem mostra a tela de validação de celular para a criação de conta na Azure. No topo, há uma barra de status com o texto 'Êxito! Vamos continuar!'. Abaixo, há um formulário com o título 'CRIAR CONTA' e o subtítulo 'Verificar sua conta'. O formulário solicita o número de celular para a verificação, com o texto 'Insira seu número de celular a seguir, para que possamos enviar um código de verificação por chamada telefônica ou mensagem de texto. Podem ser aplicadas tarifas padrão. Política de privacidade'. Há duas opções de envio: 'Enviar mensagem de texto' (selecionada) e 'Ligue-me'. O país é definido como 'Brasil (+55)'. Há campos para o número de celular e um botão 'Enviar mensagem de texto'. Um botão 'Verificar código' está disponível. Na base da tela, há uma barra de navegação com ícones de setas e um link 'Precisa de ajuda para verificar sua conta?'.

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Concluindo esse processo foi apresentado um novo formulário para o cadastro de cobrança de fatura no cartão de crédito conforme Figura 33.

Figura 33 – Tela de inserção de dados do cartão de crédito para criação de conta na Azure

Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

E conforme a Figura 34, redirecionado para a página a partir da qual é possível ter os acessos aos diversos recursos que compõem a plataforma Azure.

Figura 34 – Tela inicial da conta criada na Azure

| Nome | Tipo | Última visualização |
|---|---|---------------------|
| lakermercado | Conta de armazenamento | 12 minutos atrás |
| mercadogol-3c7b6435600840cb1730478c5885294-0bs-server | Servidor flexível do Banco de Dados do Azure para MySQL | 37 minutos atrás |
| mercadogol-9d31e63908d614209cd-endpoint [mercadogol-9d31e63908d614209cd-endpoint] | Ponto de extremidade | 37 minutos atrás |
| mercadogol-5d741bb689-private-link.mysql.database.azure.com | Zona DNS privado | 38 minutos atrás |
| mercadogoliano | Serviço de Aplicativo | 48 minutos atrás |
| dataakermercado | Conta de armazenamento | 2 horas atrás |
| gymmarket | Grupo de recursos | 2 horas atrás |
| market-gyn | Data factory (V2) | 3 dias atrás |
| DataBricks_mercado | Serviço do Azure Databricks | 3 dias atrás |
| Assinatura do Azure 1 | Assinatura | 3 dias atrás |
| mercadogol-5d741bb689-vnet | Rede virtual | 4 dias atrás |
| fire | Endereço IP público | uma semana atrás |

Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

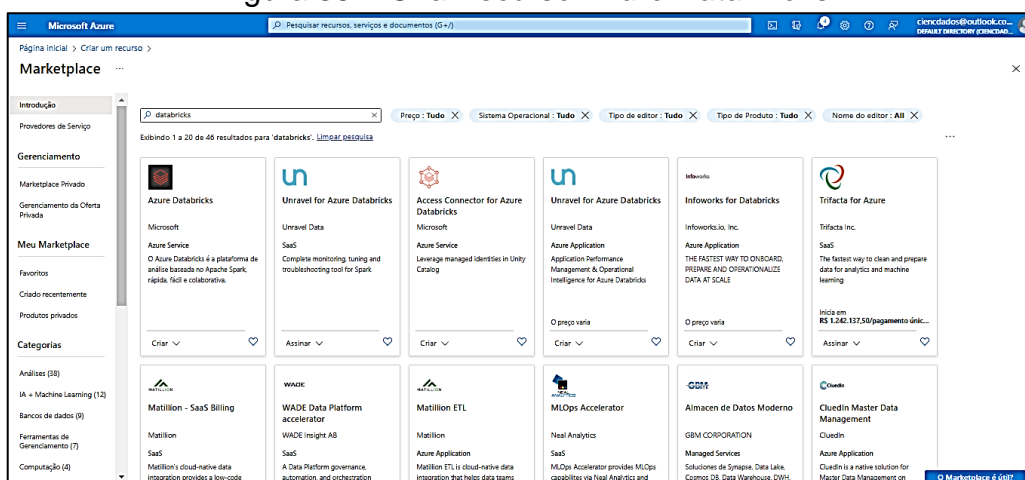
APÊNDICE B - INSTALAÇÃO E CONFIGURAÇÃO AZURE DATABRICKS

Pré-requisitos

- Subscrição *Azure*;
- Proprietário ou Colaborador *Azure*;

No Portal do *Azure*, selecione “Criar um recurso *Azure Databricks*”, conforme Figura 35.

Figura 35 – Criar recurso *Azure Databricks*



Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Em seguida conforme Figura 36 selecione o Serviço do *Azure Databricks*, forneça os valores para criar um *workspace* do *Databricks* de acordo com a subscrição.

Figura 36 – Inserir os valores correspondentes para a criação do recurso *DataBricks*

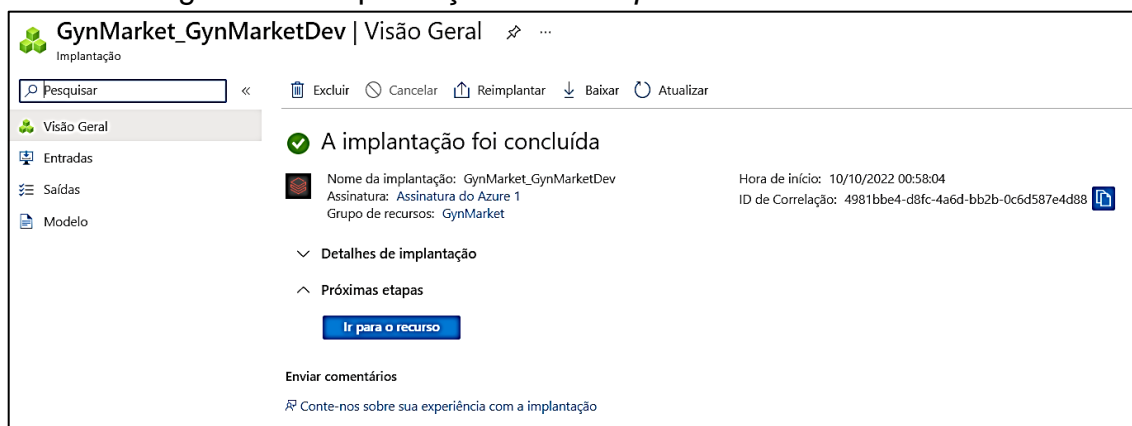
Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa, selecione “Revisar + Criar” e depois “Criar”. A criação do workspace leva um período. Quando esse processo for concluído, sua conta de usuário será adicionada automaticamente como um usuário administrador no workspace, conforme ilustrado nas Figuras 37 e 38.

Figura 37 – Criar um workspace do *Azure Databricks*

Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

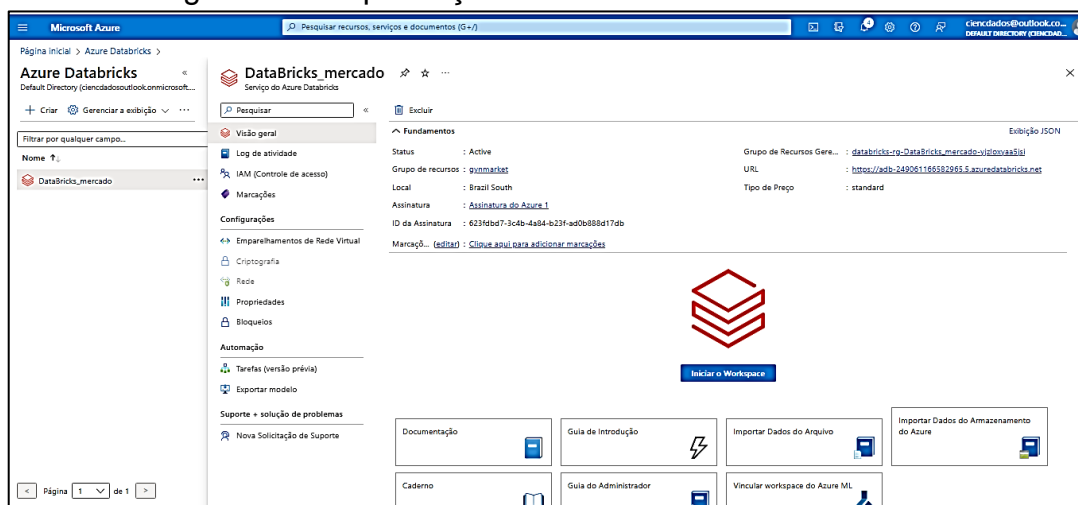
Figura 38 – Implantação do *workspace* do *Azure Databricks*



Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

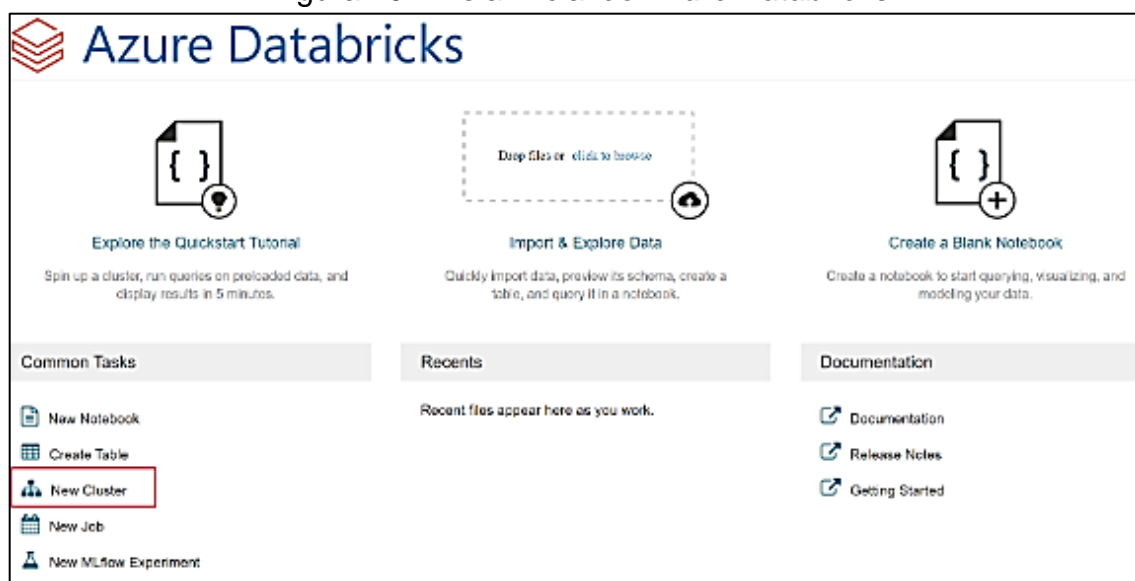
Na etapa seguinte é selecionado ir para o workspace do *Databricks* criado e clique em *Inicializar Workspace* conforme Figura 39.

Figura 39 – Implantação do *Azure Databricks* finalizada



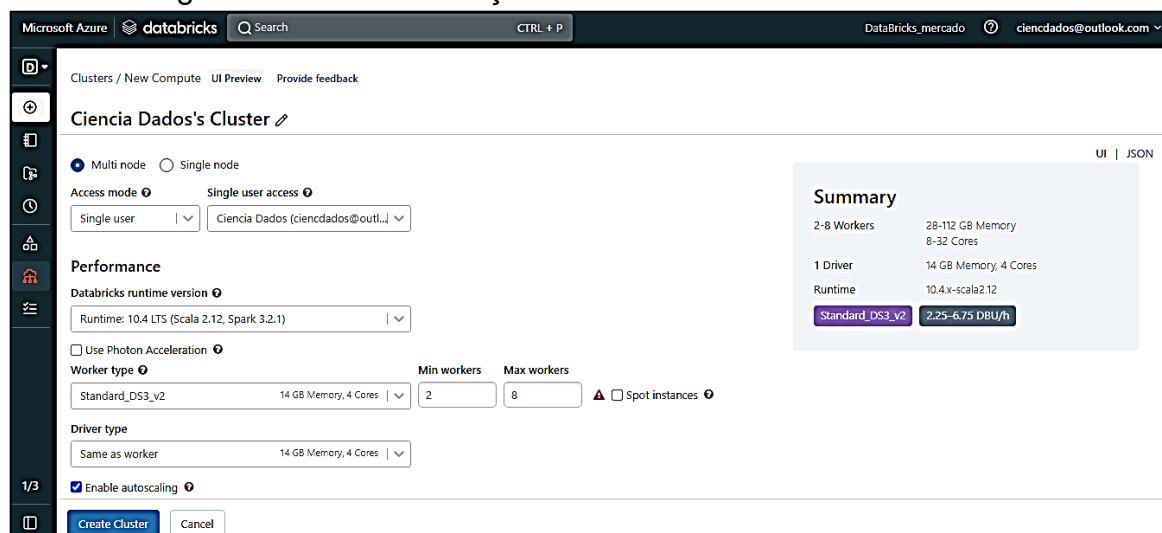
Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Nesta etapa é redirecionado para o portal do *Azure Databricks*. No portal, deve-se clicar em *Novo Cluster* conforme Figura 40.

Figura 40 – Tela inicial do *Azure Databricks*

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

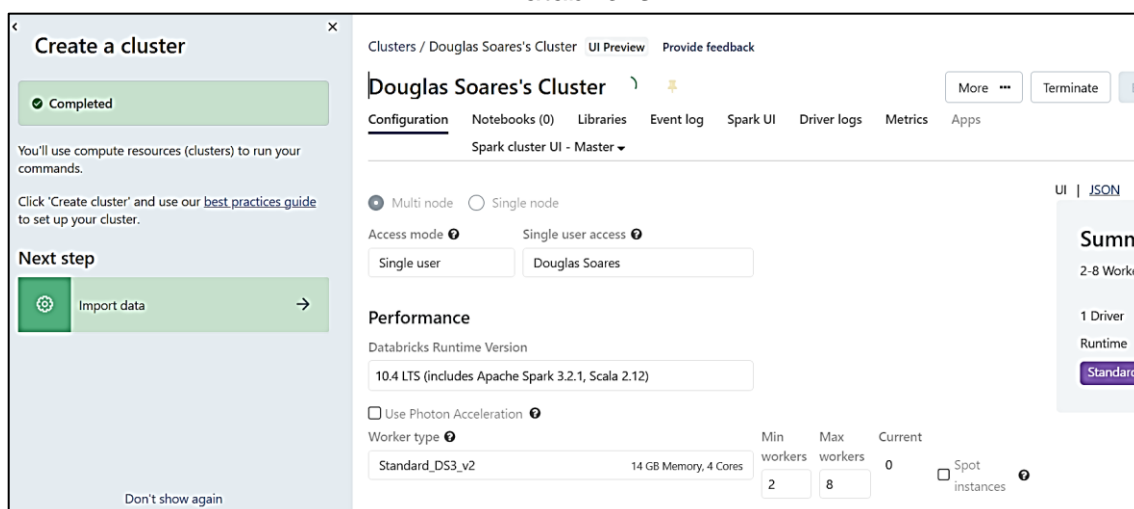
Na página Novo *cluster*, forneça os valores para criar um *cluster*, aceite todos os valores padrões, após a conclusão da criação do *cluster* o mesmo é inicializado podendo ser utilizado conforme Figura 41.

Figura 41 – Tela da criação de um *cluster* no *Azure Databricks*

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Figura 42 – Tela do processo de criação de um *cluster* concluído no *Azure*

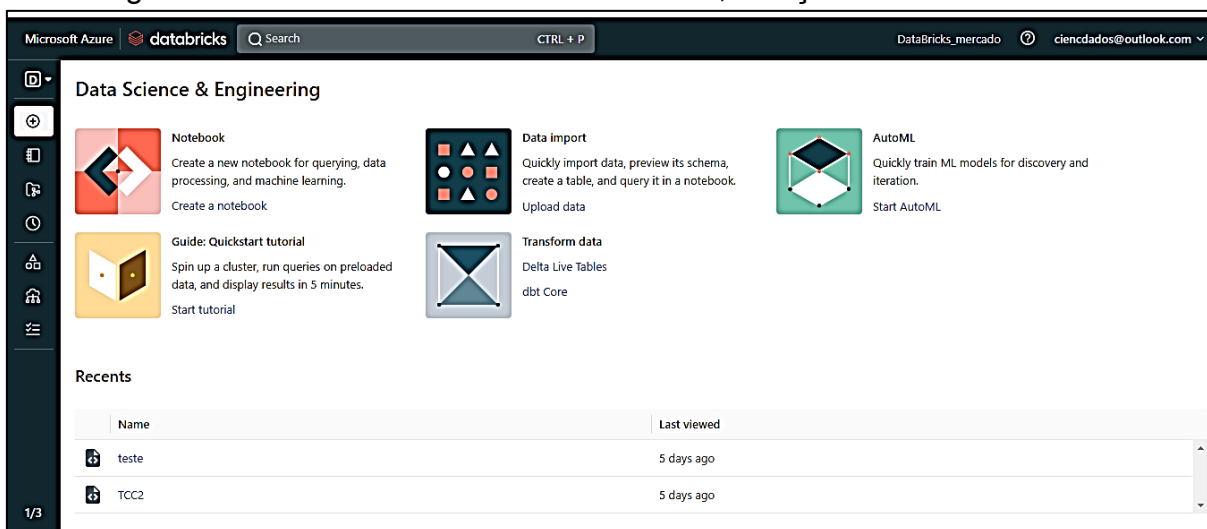
Databricks



Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Para criar um *notebook* que será executado para trazer o banco de dados de registros de compra e fazer as transformação necessárias é necessário criar um *notebook* de configurações de acesso ao container que será carregado o arquivo, sendo assim, necessário clicar em “*Create Notebook*” conforme Figura 43.

Figura 43 – Tela inicial no *Azure Databricks*, seleção “*Create Notebook*”



Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Na caixa de diálogo “*Criar Notebook*”, é inserido um nome, feito a escolha *Python* como a linguagem e escolha-se o *cluster Spark* criado anteriormente e o *notebook* criado com o nome de ‘*mount-point*’, que é através dos parâmetros incluídos neste *notebook* com o ID do *container*, e *key* e *id_key* foi gerado acessos

capazes de sincronizar o *DataBricks* com o *container*, obtendo todos os arquivos presentes no *container*, conforme Figura 44.

Figura 44 – Tela do *notebook mount-point* do *Databricks*

```

Microsoft Azure databricks Search CTRL + P DataBricks_mercado ciencdados@outlook.com
mount_point Python
File Edit View Run Help Last edit was 46 minutes ago Give feedback
▶ Run all • Ciencia Dados's Cluster Schedule Share
Cmd 1
1 dbutils.fs.unmount('/mnt/registro')
Python ▶ ▾ - ×
/mnt/registro has been unmounted.
Out[2]: True
Command took 20.96 seconds -- by ciencdados@outlook.com at 14/11/2022 00:57:34 on Ciencia Dados's Cluster

Cmd 2
1 configs = {"fs.azure.account.auth.type": "OAuth",
2           "fs.azure.account.oauth.provider.type": "org.apache.hadoop.fs.azurebfs.oauth2.ClientCredsTokenProvider",
3           "fs.azure.account.oauth2.client.id": "3f283653-00b7-45c5-90a3-3af8f322ebd0",
4           "fs.azure.account.oauth2.client.secret": "it00Q-6B0RpKHoTv3nmMBnuFFKd5DPu.wW6iZbVW",
5           "fs.azure.account.oauth2.client.endpoint": "https://login.microsoftonline.com/07e5e641-8bde-4b5a-bec2-70f7b466776a/oauth2/token",
6           "fs.azure.createRemoteFileSystemDuringInitialization": "true"}
7
8 dbutils.fs.mount(
9   source = "abfss://extracao@lakemercado.dfs.core.windows.net/",
10  mount_point = "/mnt/extracao",
11  extra_configs = configs)
Python ▶ ▾ - ×
Out[11]: True
Command took 20.91 seconds -- by ciencdados@outlook.com at 14/11/2022 02:03:22 on Ciencia Dados's Cluster
Ativar o Windows
Acesse Configurações para ativar o Windows.

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Verificado após a execução do *notebook mount-point* os arquivos presentes no caminho do *container* selecionado estão presentes para acesso no *DataBricks* conforme Figuras 45 e 46.

Figura 45 – Tela do *notebook mount-point* exibindo os conteúdos armazenados no *container*

```

Cmd 3
1 filepath=r'/mnt/extracao/'
Python ▶ ▾ - ×
Command took 0.02 seconds -- by ciencdados@outlook.com at 14/11/2022 02:54:33 on Ciencia Dados's Cluster

Cmd 4
1 dbutils.fs.ls(filepath)
Python ▶ ▾ - ×
Out[19]: [FileInfo(path='dbfs:/mnt/extracao/Recomendados.csv', name='Recomendados.csv', size=1215, modificationTime=1668402650000),
FileInfo(path='dbfs:/mnt/extracao/Registros_compra.csv', name='Registros_compra.csv', size=148721, modificationTime=1668401686000)]
Command took 0.62 seconds -- by ciencdados@outlook.com at 14/11/2022 02:54:35 on Ciencia Dados's Cluster
Shift+Enter to run

```

Fonte: Tela de captura do *DataBricks* com conteúdo desenvolvido pelo autor deste trabalho

Figura 46 – Tela dos arquivos armazenados no *container*

The screenshot displays the Microsoft Azure portal interface for a container named 'extracao'. The top navigation bar shows 'Microsoft Azure' and a search bar. The main content area includes a search bar, navigation icons (Carregar, Adicionar Diretório, Atualizar, Renomear, Excluir, Alterar a camada, Adquirir concessão, Interromper concessão), and a 'Visão geral' tab. Below the navigation, there are options for 'Método de autenticação' (Tela de acesso) and 'Local: extracao'. A search bar for blobs is present, along with a toggle for 'Mostrar os objetos que foram excluídos'. The main area contains a table of blobs with the following data:

| Nome | Modificado | Camada de acesso | Status do arquivo | Tipo de blob | Tamanho | Estado de concessão |
|----------------------|---------------------|-----------------------|-------------------|----------------|------------|---------------------|
| Recomendados.csv | 14/11/2022 02:10:50 | Principal (inferidos) | | Blob de blocos | 1.19 KIB | Disponível |
| registros_compra.csv | 14/11/2022 01:54:46 | Principal (inferidos) | | Blob de blocos | 145.24 KIB | Disponível |

Fonte: Tela de captura do *Azure container* com conteúdo desenvolvido pelo autor deste trabalho

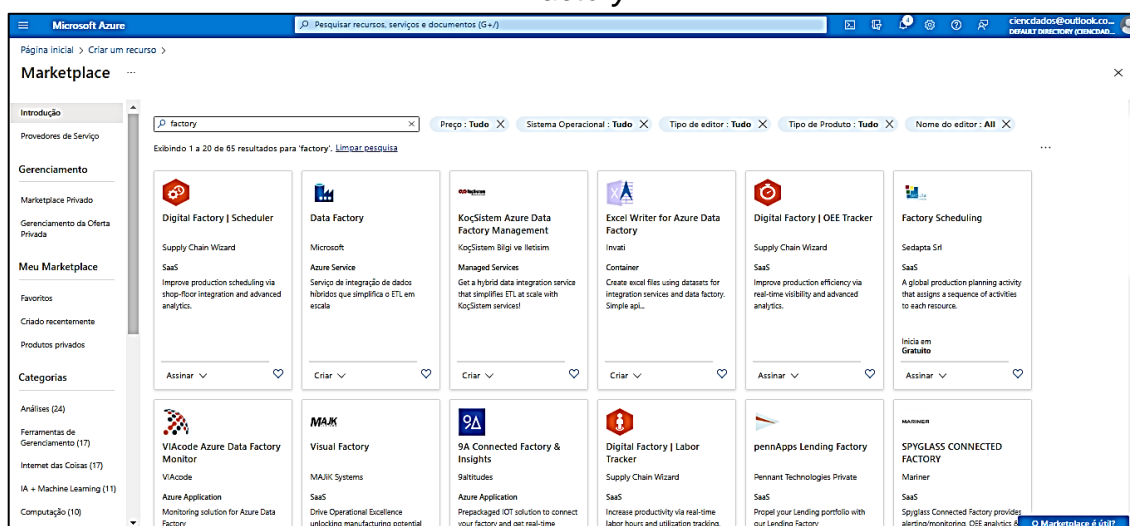
APÊNDICE C - INSTALAÇÃO E CONFIGURAÇÃO DO AZURE DATAFACTORY

Pré-requisitos:

- Subscrição *Azure* ativa;
- Proprietário ou Colaborador *Azure*;
- Uma conta de Armazenamento do Azure;
- Um aplicativo de *Azure Active Directory*.

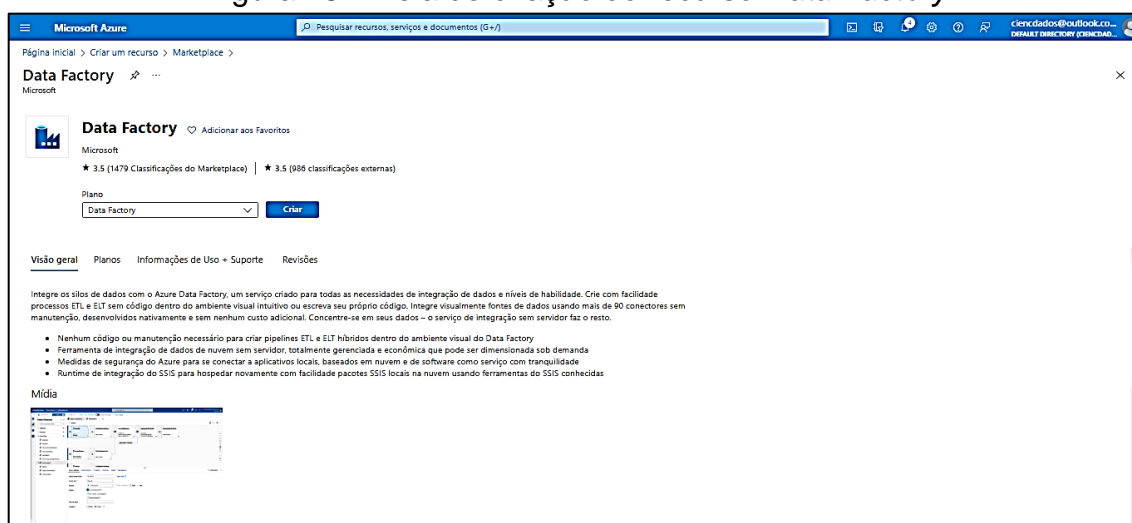
No Portal do *Azure*, é selecionado “Criar um recurso *Data Factory*”, conforme Figura 47.

Figura 47 – Tela dos recursos que podem ser adquiridos na busca do recurso *Data Factory*



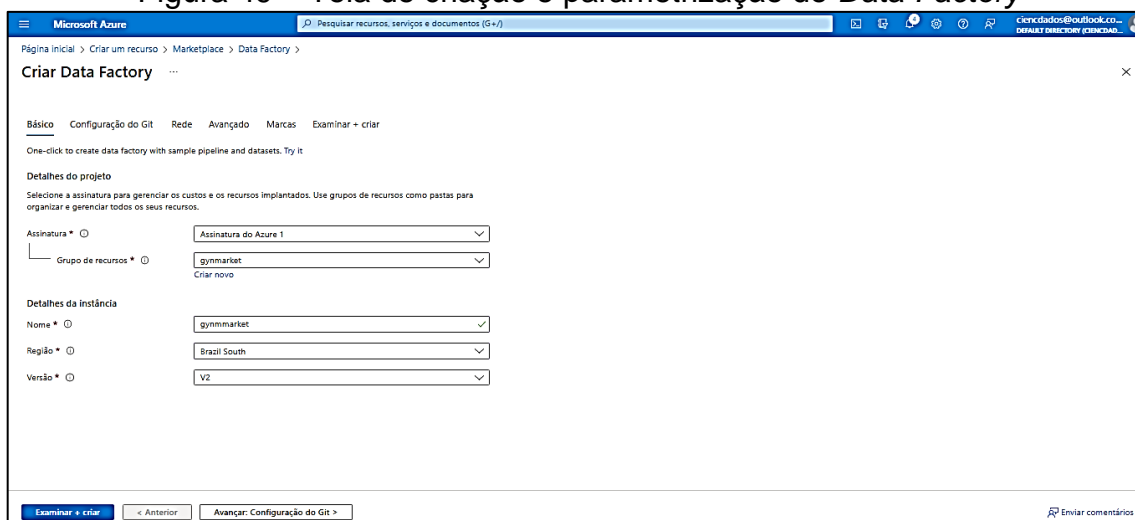
Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Em seguida, conforme a Figura 48, é acessado o recurso “*Azure Data Factory Studio*” e escolhido o botão de opção “*Criar um novo data factory*”.

Figura 48 – Tela de criação do recurso *Data Factory*

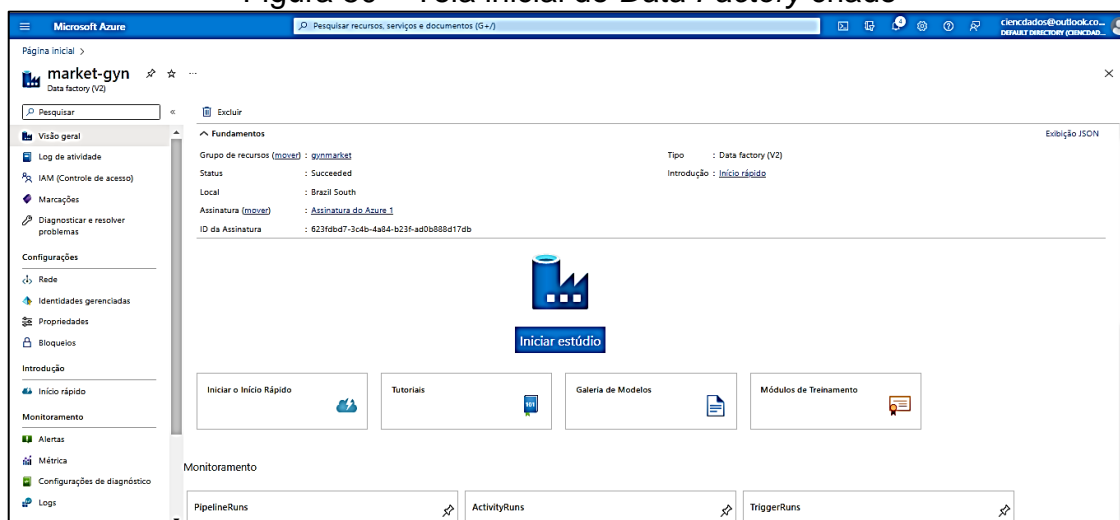
Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Em seguida conforme a Figura 49 é utilizado valores padrão para criar diretamente ou inserir um nome exclusivo e escolher assim um local e uma assinatura preferenciais que serão usados ao criar o novo *Data Factory* e Selecionado “Revisar + Criar” e em seguida “Criar”.

Figura 49 – Tela de criação e parametrização do *Data Factory*

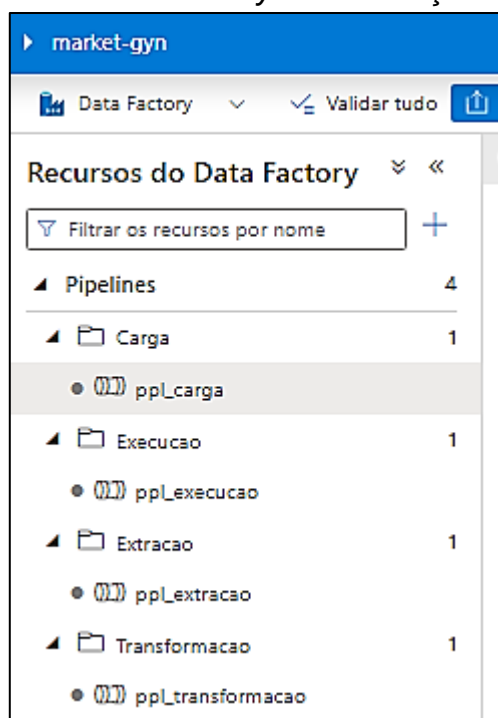
Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Após a criação, é acessado diretamente na home page do *Azure Data Factory Studio*.

Figura 50 – Tela inicial do *Data Factory* criado

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Selecionando a opção “Iniciar estúdio” é direcionado ao serviço do ADF “merket-gyn” em seguida selecionado a opção “*Author -> Pipelines -> Nova Pasta*” e cria-se pastas para cada processo de ETL que será utilizado para a automação e em cada pasta é feita a criação de um *pipeline* com o nome da atividade executada, conforme Figura 51:

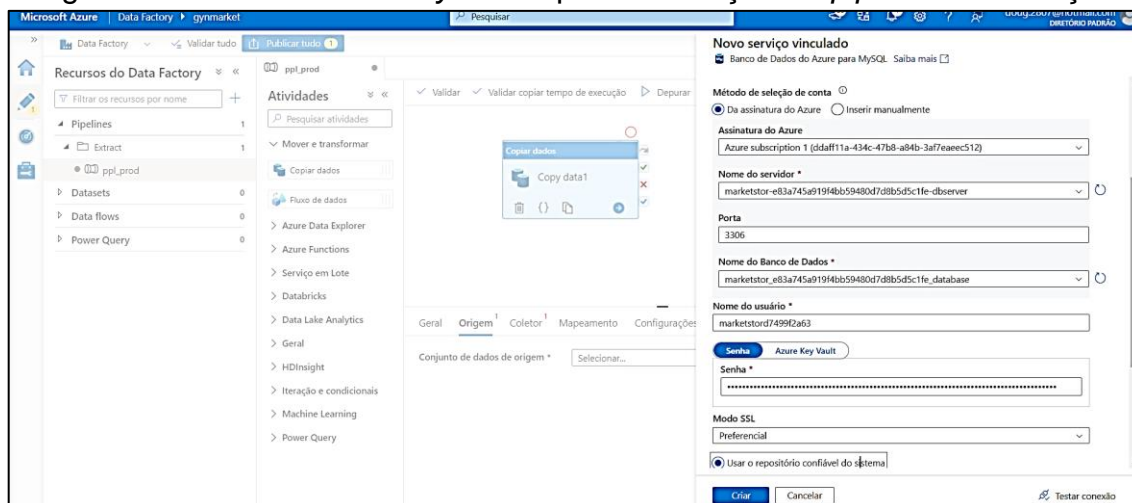
Figura 51 – Tela *Data Factory* com a criação dos *pipelines*

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Os processos que se executa em cada *pipeline* são descritos a seguir:

- Extração : extrair a tabela de registro de compras do banco de dados do *site* <https://mercadogoiano.azurewebsites.net/> no formato 'csv' e carregar a base no *container* 'gynMarket22' na pasta chamada de *extracao*, conforme Figura 52.

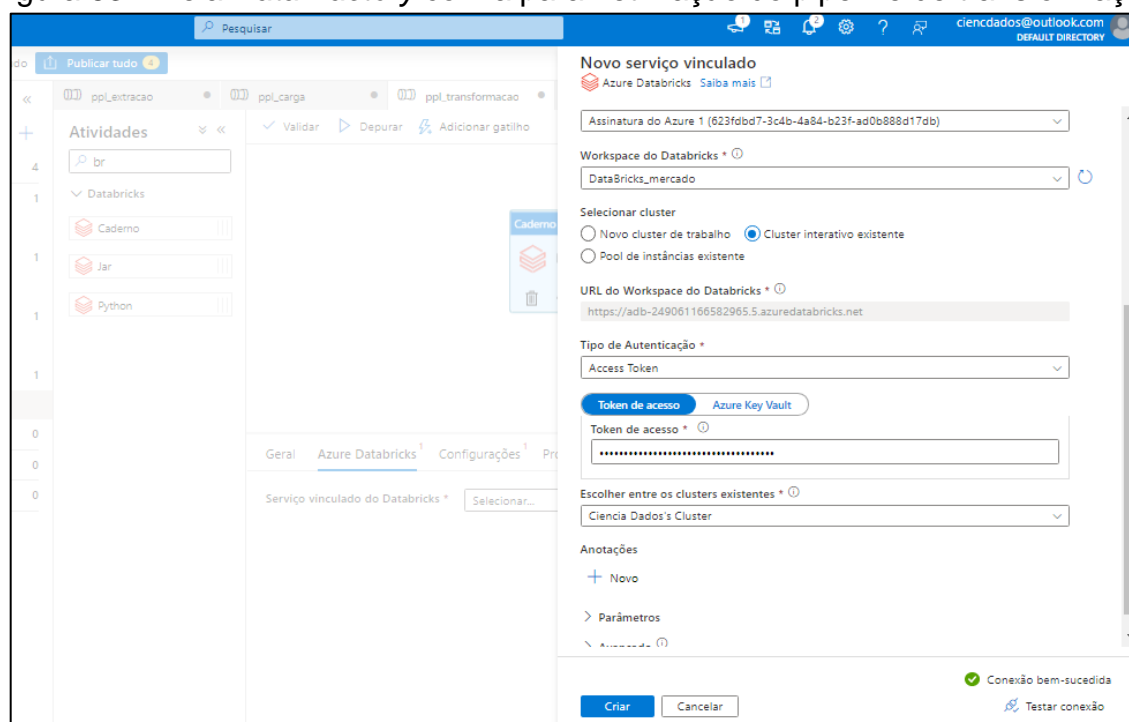
Figura 52– Tela *Data Factory* com a parametrização do *pipeline* de extração



Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

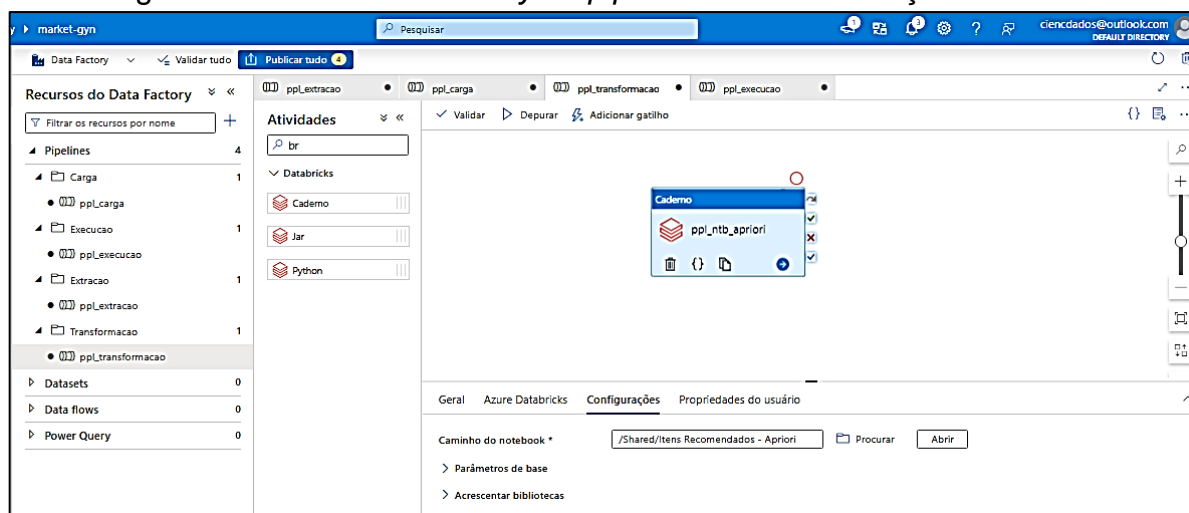
- Transformação : acessar e direcionar a tabela em formato 'csv que foi extraída no processo anterior para o ambiente do *Databricks* chamado de *DataBricks_mercado* para utilizarmos o método *Apriori* para encontrarmos os itens recomendados de acordo com a base extraída de registros de compras, conforme Figure 53.

Figura 53 – Tela *Data Factory* com a parametrização do pipeline de transformação



Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Figura 54 – Tela *Data Factory* do *pipeline* de transformação concluída

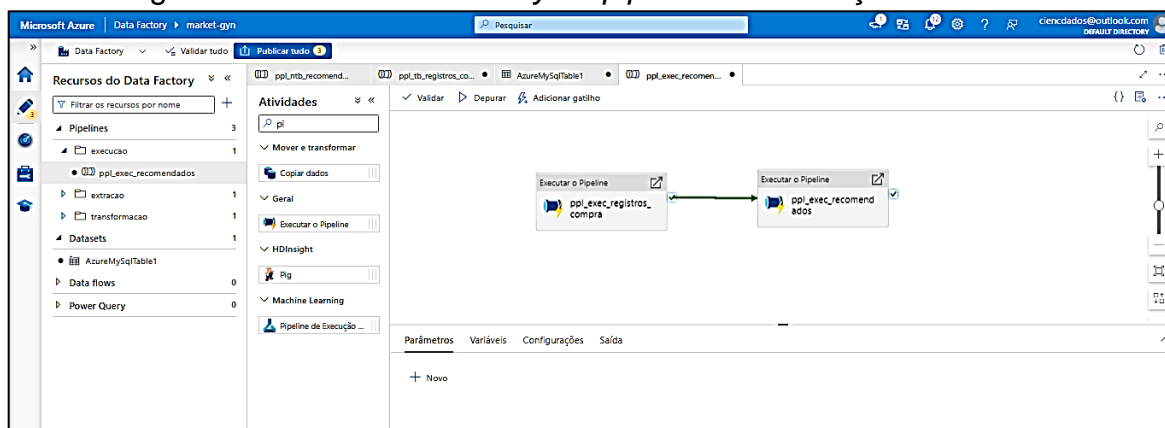


Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

- **Carga:** após se ter o resultado de recomendações feita no processo de transformação, é carregado o resultado no container '*gynMarket22->extracao*' pelo mesmo *notebook* de transformação.
- **Execução:** após feitos os passos anteriores para configurar os *pipelines* de extração e transformação, é criado um *pipeline* chamado de '*execucao*' para

executar os *pipelines* anteriores na ordem correta no qual o *pipeline* de extracao é executado e somente se ele for executado com o status de 'sucesso' o *pipeline* de transformacao é executado, conforme Figura 55.

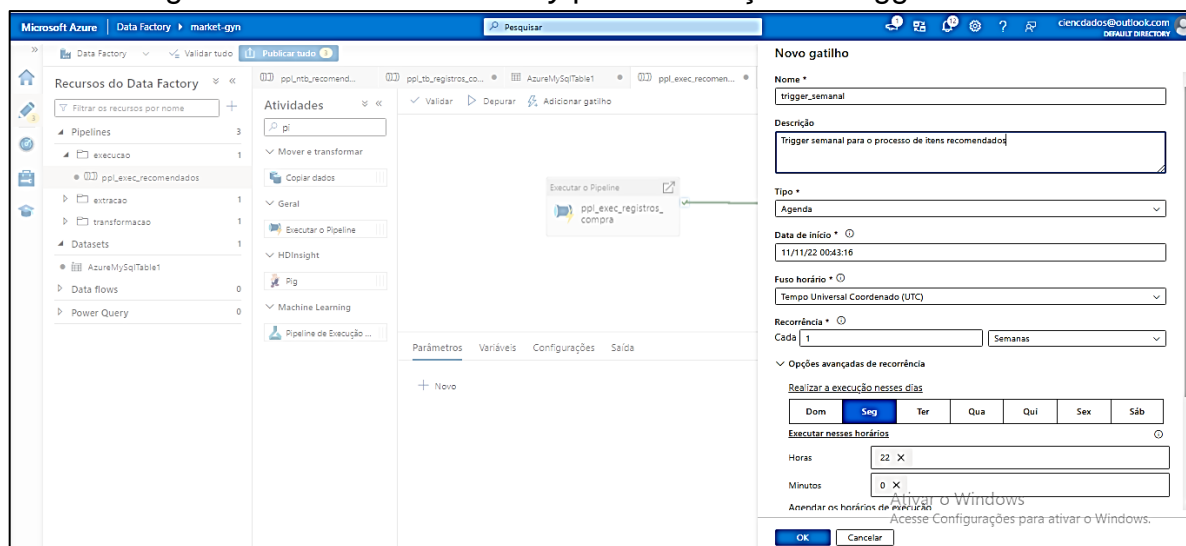
Figura 55 – Tela *Data Factory* do *pipeline* de execução construído



Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

Em seguida conforme a Figura 56 , é adicionado uma '*trigger*' para que o *pipeline* de execução seja executado semanalmente.

Figura 56 – Tela *Data Factory* parametrização da trigger semanal

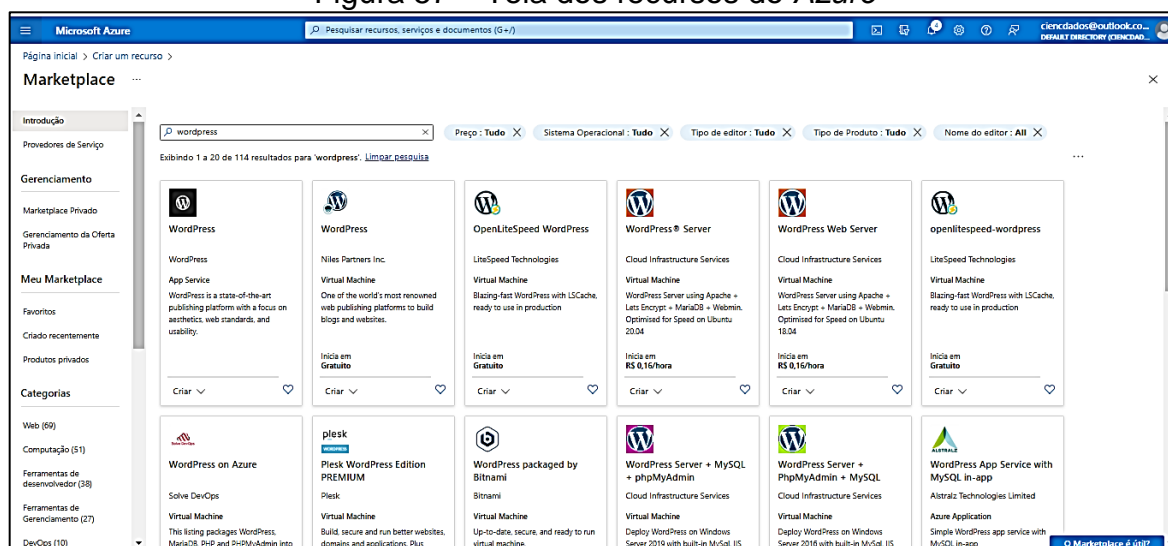


Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

APÊNDICE D - INSTALAÇÃO WORDPRESS NO AZURE

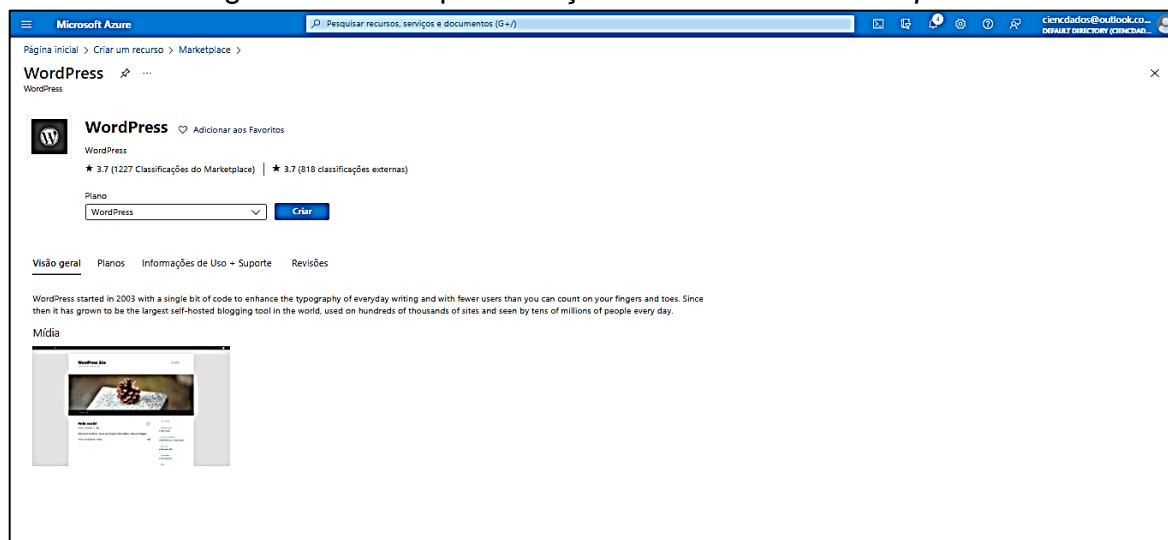
No Portal do Azure, selecione “Criar um recurso *Wordpress*”, conforme Figura 57 em seguida conforme Figura 58 selecione “Criar”.

Figura 57 – Tela dos recursos do Azure



Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

Figura 58 – Tela para criação de um recurso *Wordpress*



Fonte: Tela de captura do Azure com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa é necessário inserir os valores correspondentes a subscrição e grupo de recursos que deseja utilizar e em seguida defino o nome para o Aplicativo *Web* conforme Figura 59.

Figura 59 – Tela para criação e parametrização de um recurso *Wordpress*

Página inicial >

Criar WordPress no Serviço de Aplicativo ...

Básico Avançado Rótulos Revisar + criar

Os recursos do Serviço de Aplicativo do Azure foram aproveitados para fornecer a você uma solução de hospedagem totalmente gerenciada, de alto desempenho, segura e escalável para WordPress. [Saiba Mais](#)

Detalhes do projeto

Selecione uma assinatura para gerenciar custos e recursos implantados. Use grupos de recursos como pastas para organizar e gerenciar todos os recursos.

Assinatura * Assinatura do Azure 1

Grupo de recursos * (Novo) storepremiumgyn_group
 [Criar novo](#)

Detalhes da hospedagem

Selecione a região do servidor e forneça um nome para o Aplicativo Web. Domínios personalizados podem ser adicionados mais tarde.

Região * Brazil South

Nome * storepremiumgyn.azurewebsites.net

Sistema operacional * Linux Windows

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Nas próximas etapas referidas a tela de criação do recurso, foi utilizado os valores padrões, selecione sempre a opção “Avançar” e por fim a opção “Revisar e criar”, a etapa é executada durante alguns minutos e em seguida conforme Figura 60 é apresentado a descrição do banco de dados assim que a aplicação é instalada com sucesso.

Figura 60 – Tela com os parâmetros de infraestrutura do recurso *Wordpress*

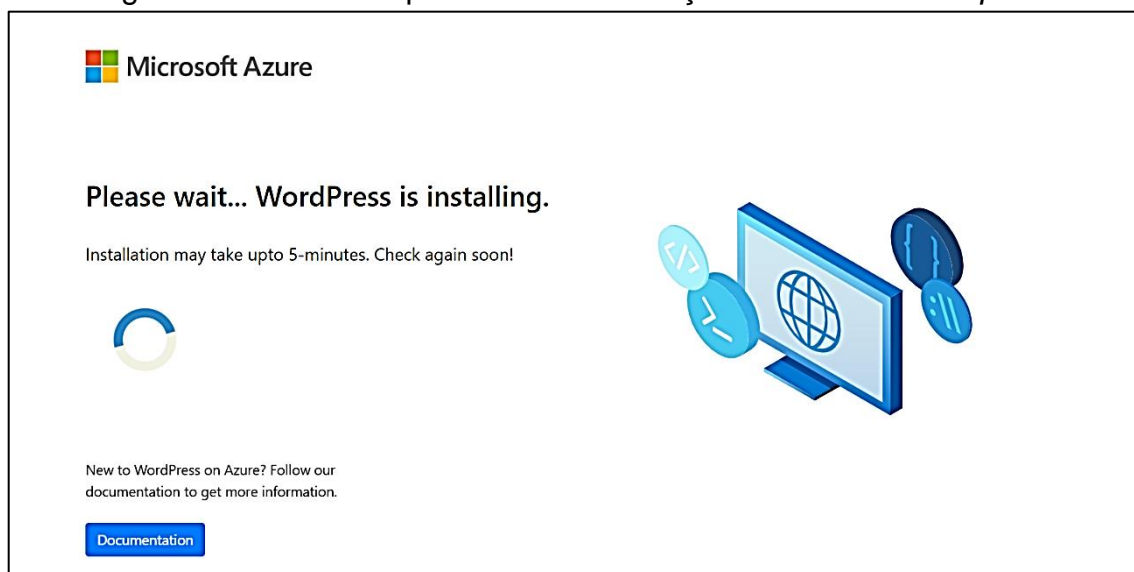
Banco de dados (Novo)

i O nome de usuário e a senha do novo banco de dados são gerados automaticamente. Para recuperar esses valores após a implantação, acesse a seção Configurações do Aplicativo da página Configuração.

| | |
|--------------------------------|--|
| Mecanismo | MySQL - Single Server (5.7) |
| Tamanho e camada de computação | Basic, Gen5 (1 vCores, 2 GiB RAM, 32 GB storage) |
| Região | Brazil South |
| Nome do servidor | storepremi-c70d3be1ec49453396a45295d6ba5d21-dbserver |
| Nome do Banco de Dados | storepremi_c70d3be1ec49453396a45295d6ba5d21_database |
| Nome de usuário | bphbewxxip |
| Senha | 473E2C707Q4VBW68\$ |

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

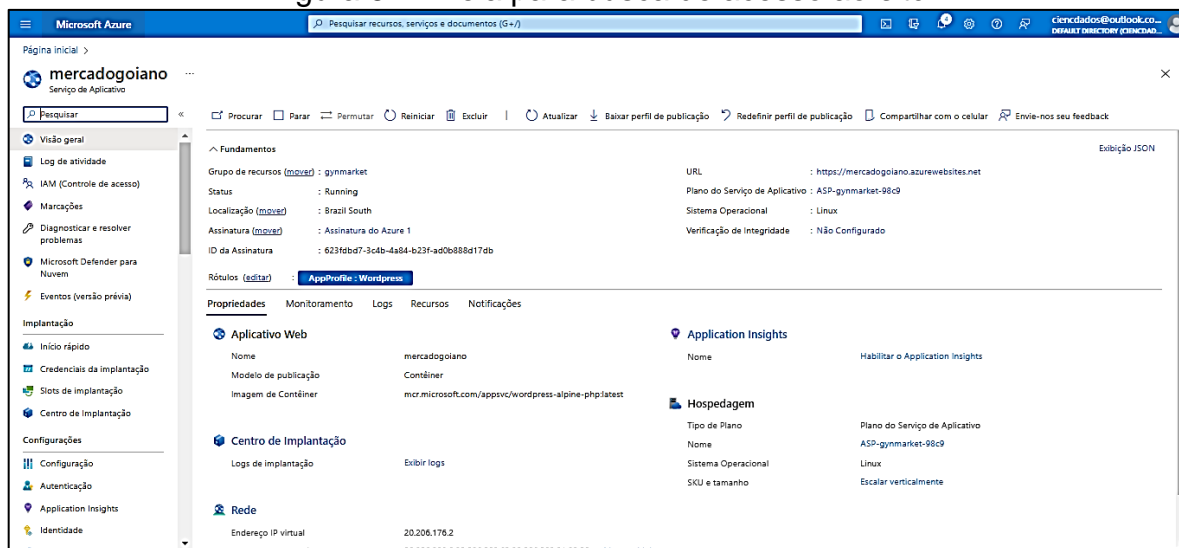
Figura 61 – Tela com processo de instalação do recurso *WordPress*



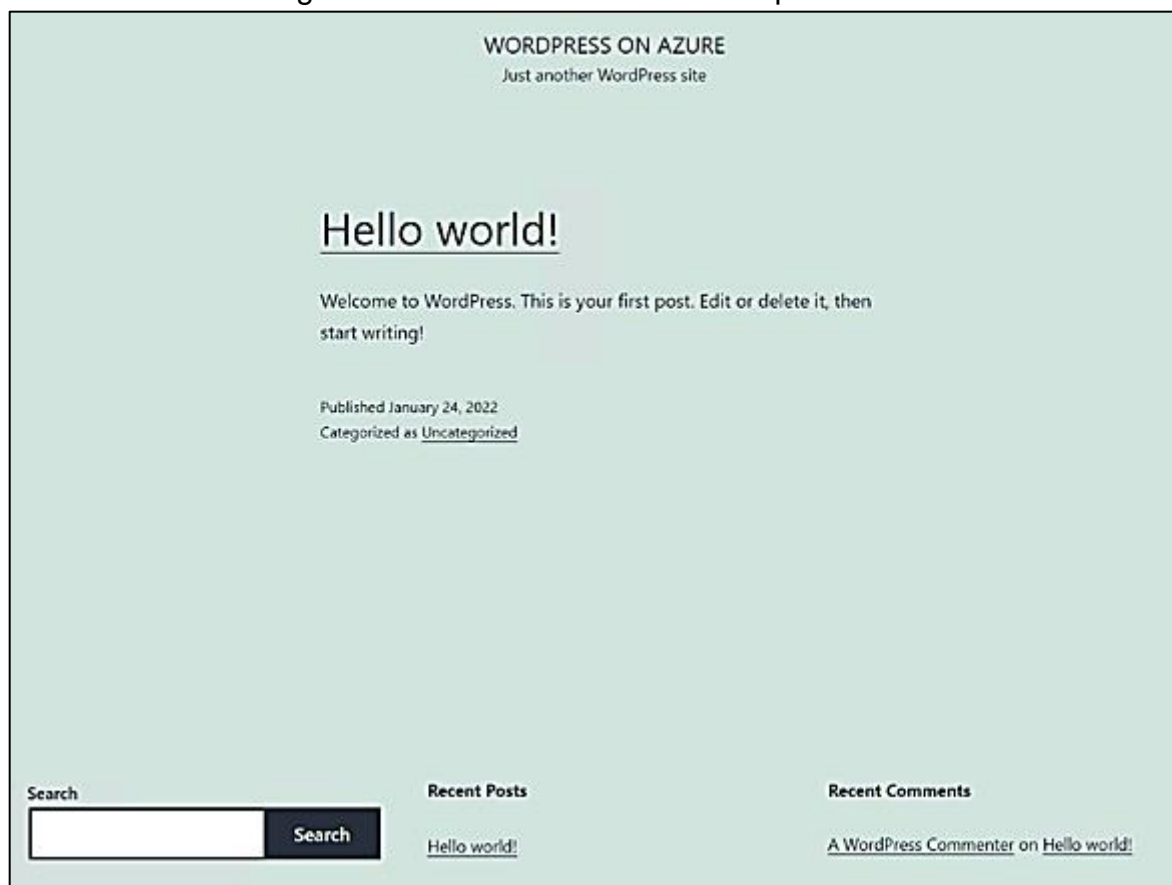
Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa, foi navegado até onde se encontra a URL do *site* e verificado se o aplicativo está sendo executado, seleciona a opção “Procurar” conforme Figura 62 e em seguida em outra aba é direcionado ao *site* conforme Figura 63.

Figura 62 – Tela para busca de acesso ao *site*

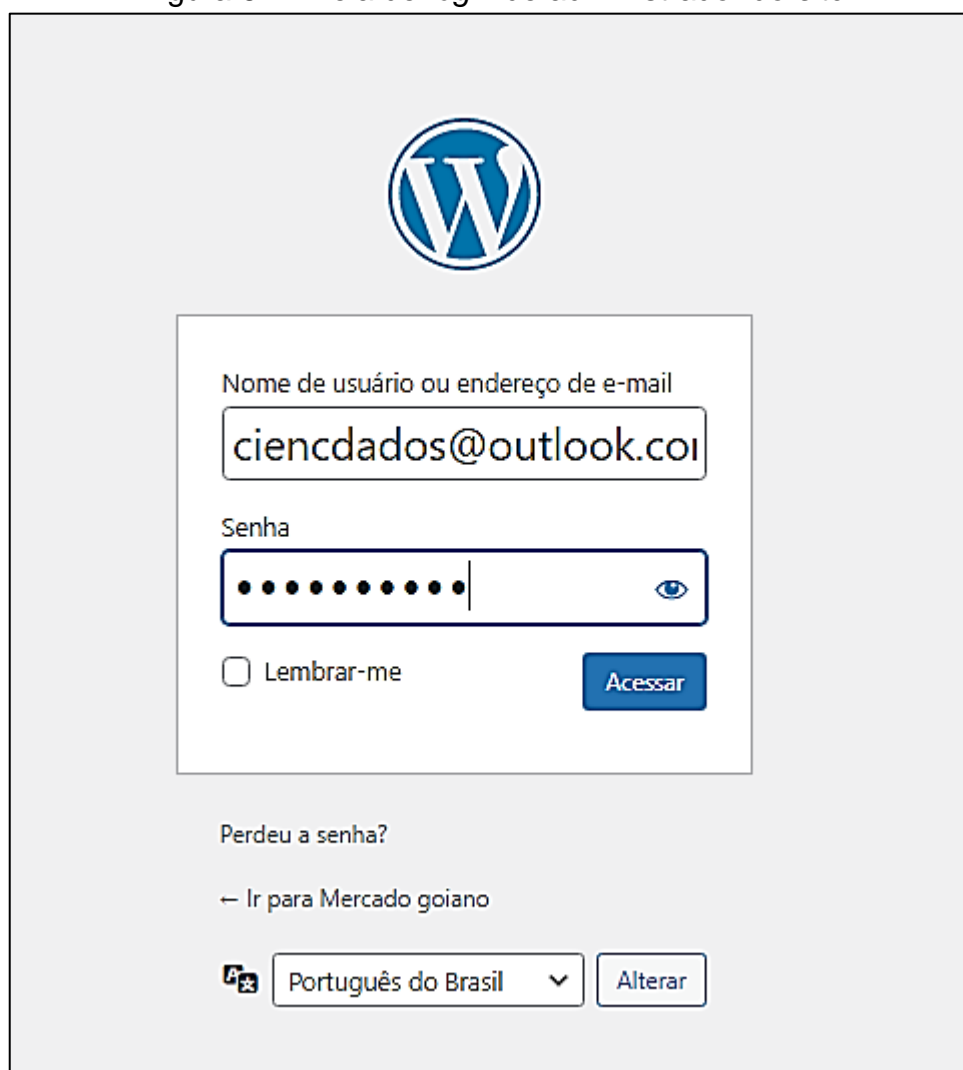


Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Figura 63 – Tela inicial direcionado para o *site*

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa, para acessar a página de administrador do *WordPress*, navegue até <https://mercadogiano.azurewebsites.net/wp-admin> e use as credenciais criadas no processo de instalação do recurso *Wordpress*, o *e-mail* cadastrado e senha do usuário da *Azure*, conforme Figura 64.

Figura 64 – Tela de *login* de administrador do *site*

Nome de usuário ou endereço de e-mail


ciencdados@outlook.coi

Senha

Lembrar-me

Perdeu a senha?

[← Ir para Mercado goiano](#)

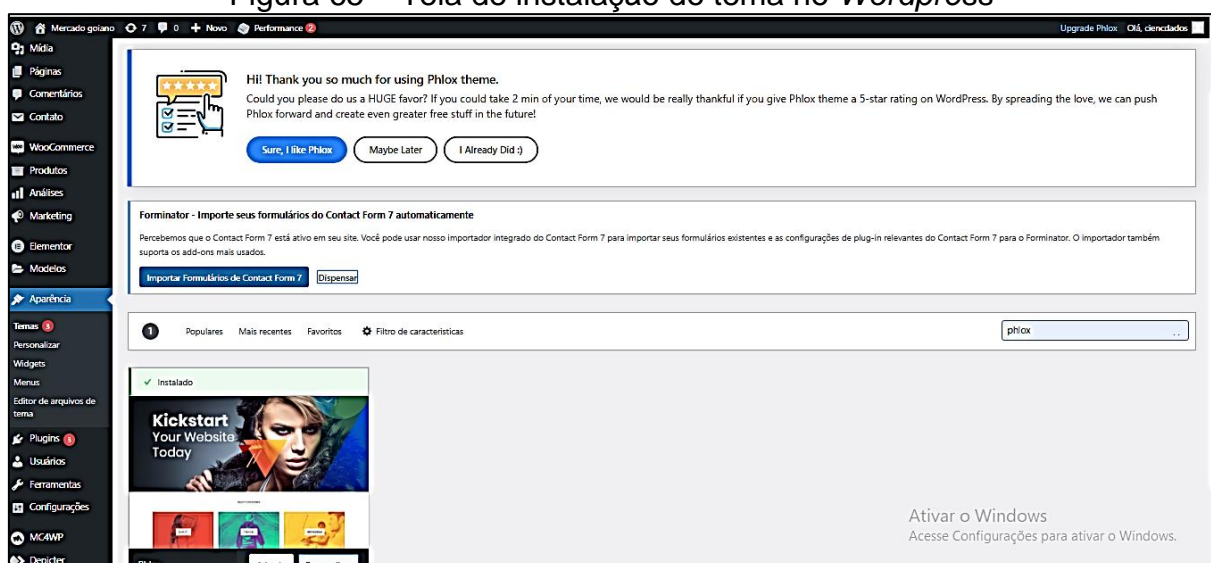
 Português do Brasil

Fonte: Tela de captura do *Azure* com conteúdo desenvolvido pelo autor deste trabalho

APÊNDICE E – CONFIGURAÇÃO *SITE* MERCADO GOIANO A PARTIR DO *WORDPRESS AZURE*

Após a conclusão dos passos de criação do *site* demonstrados no Apêndice D é iniciado o processo de aparência e configurações do *site*. Conforme Figura 65 é selecionado as seguintes opções Aparência > Temas no painel do *WordPress* e clique em Adicionar novo para inserir o tema utilizado “*PHLOX*” após a busca encontrar o tema Selecione a opção “Instalar” e “Ativar”.

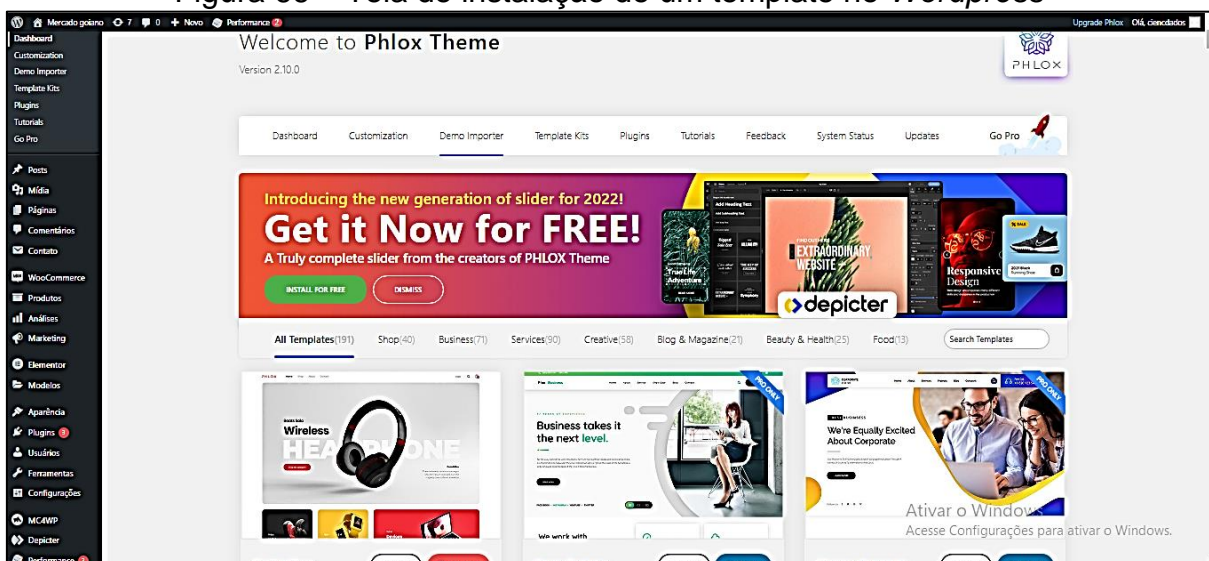
Figura 65 – Tela de instalação do tema no *Wordpress*



Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Após a conclusão da etapa anterior irá surgir uma nova opção no menu com o nome do tema *PHLOX*, dessa forma selecione a opção *PHLOX -> Demo Importer*, a aparência já pré construída de um *e-commerce*, conforme Figura 66.

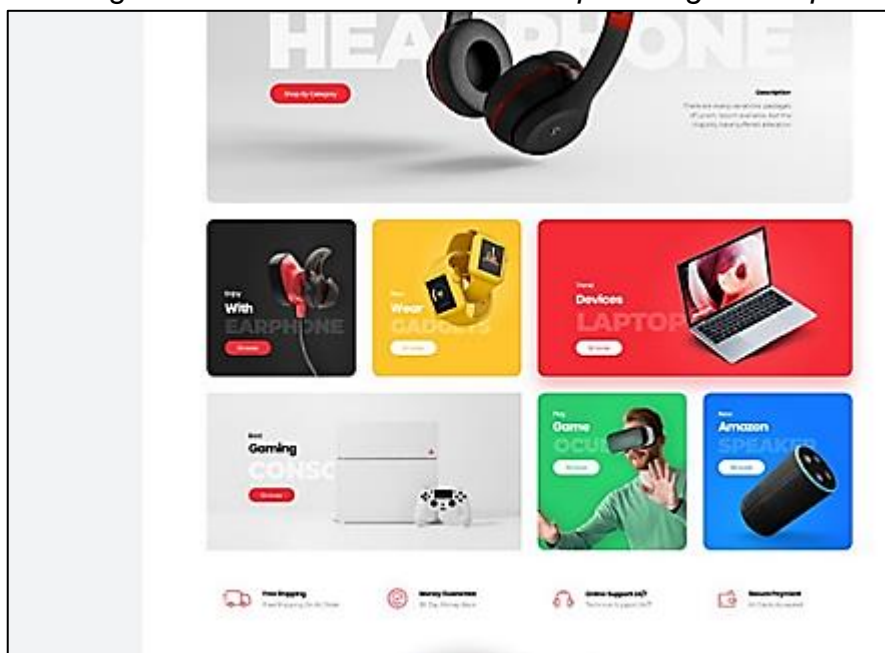
Figura 66 – Tela de instalação de um template no *Wordpress*



Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Na etapa seguinte é selecionado o template utilizado “*Digital Shop*” para a construção do *site* utilizado, instalando e ativando os plugins necessários e *layouts* e base de mídias do próprio *template*. Após a instalação do *template* a aparência do *site* inicial está representado na Figura 67.

Figura 67 – Tela do *site* com o *template Digital Shop*

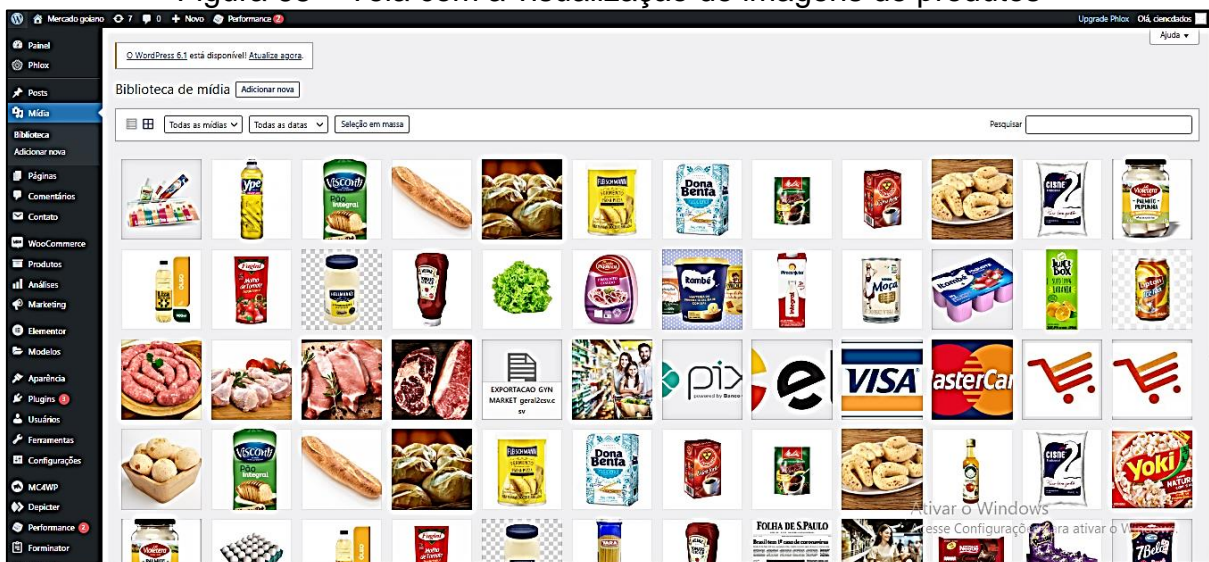


Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa foi selecionado todas imagens dos produtos que o

supermercado oferece a partir de busca no *Google* Imagens e guardadas no armazenamento local e sendo inseridas no *site* selecionado a opção Mídia->Adicionar Nova. Após a conclusão do carregamento a Biblioteca de Mídia está representado na Figura 68.

Figura 68 – Tela com a visualização de imagens de produtos



Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Na próxima etapa foi feito a exclusão de todos os produtos carregados a partir do template "*Digital Shop*", selecionando a opção Produtos -> Todos os produtos->Marcar todos->Mover para Lixeira -> Aplicar.

Na etapa seguinte foi inserindo os novos produtos a partir do carregamento em massa de um arquivo de *excel* armazenado localmente e separado por vírgulas (.csv) que contém todas as características necessárias para se criar um produto conforme Figura 69.

Figura 69 – Tela com modelo de importação de produtos em massa no *site*

| ID | Tipo | SKU | Nome | Publicado | Em destaque | Visibilidade | Status do produto | Classe de estoque | Em estoque | Estoque | Quantidade | São permitidos | Vendido | i | Peso (kg) | Comprimento | Largura (cm) | Altura (cm) | Permitir a venda | Nota da categoria | Preço |
|----|------|--------|------|------------|-------------|--------------|-------------------|-------------------|------------|---------|------------|----------------|---------|---|-----------|-------------|--------------|-------------|------------------|-------------------|-------|
| 2 | 2010 | simple | 1010 | Carne Bov | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 3 | 2011 | simple | 1011 | Carne suín | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 4 | 2012 | simple | 1012 | Frango | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 5 | 2013 | simple | 1013 | Salsicha | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 6 | 2014 | simple | 1014 | Linguica | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 7 | 2015 | simple | 1015 | Coca cola | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 8 | 2016 | simple | 1016 | Chá | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 9 | 2017 | simple | 1017 | Cerveja | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 10 | 2018 | simple | 1018 | Água | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 11 | 2019 | simple | 1019 | Suco | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 12 | 2020 | simple | 1020 | Iogurte | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 13 | 2021 | simple | 1021 | Leite Cond | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 14 | 2022 | simple | 1022 | Leite Inte | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 15 | 2023 | simple | 1023 | Manteiga | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 16 | 2024 | simple | 1024 | Margarina | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 17 | 2025 | simple | 1025 | Presunto | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 18 | 2026 | simple | 1026 | Alfafa | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 19 | 2027 | simple | 1027 | Batata | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 20 | 2028 | simple | 1028 | Caju | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |
| 21 | 2029 | simple | 1029 | Cebola | 1 | 0 | visible | taxable | - | 1 | 100 | - | 0 | 0 | | | | | | | 0 |

Fonte: Tela de captura da base de dados aberta no excel desenvolvido pelo autor deste trabalho

Na próxima etapa é selecionado a opção Produtos->Importar, escolhe o arquivo correspondente a Figura e seleccione a opção “continuar” conforme a Figura 70.

Figura 70– Tela com a parametrização de importação em massa de produtos no *site*

Importar produtos de um arquivo CSV

Esta ferramenta lhe permite importar (ou fundir) dados de produtos para a sua loja a partir de um arquivo CSV ou TXT.

Escolha um arquivo CSV do seu computador: EXPORTACA_ET_geral1.csv
Tamanho máximo: 50 MB

Atualizar produtos existentes Produtos existentes que correspondam o ID ou ao SKU serão atualizados. Produtos que não existem serão ignorados.

Alternativamente, digite o caminho para um arquivo CSV no seu servidor:

Delimitador do CSV:

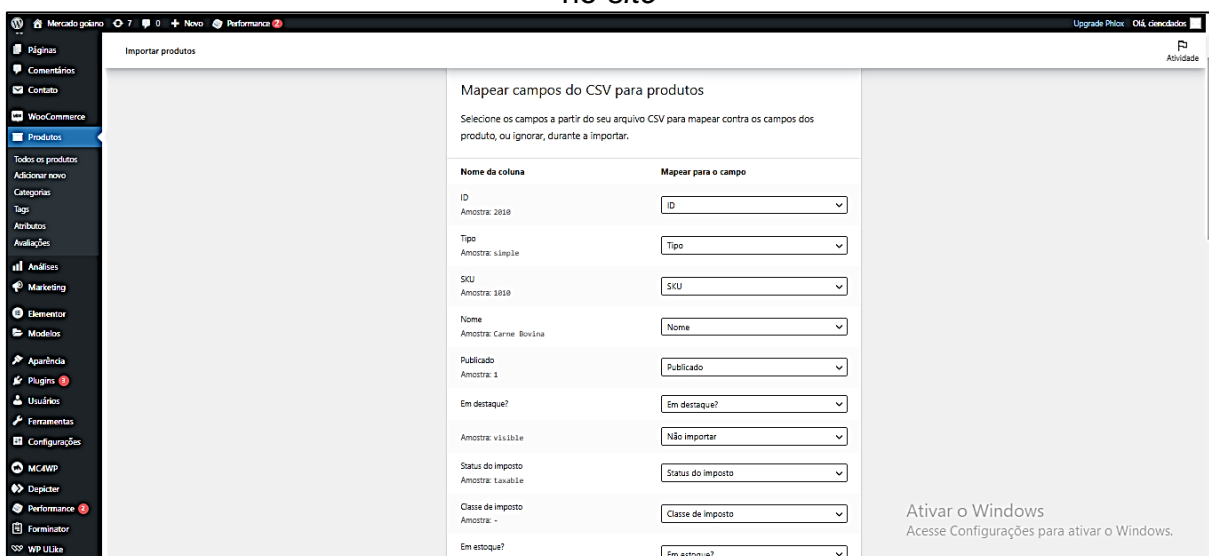
Usar as preferências salvas de mapeamento de colunas?

[Esconder opções avançadas](#)

Fonte: Tela de captura do *WordPress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Na etapa seguinte é apresentado o mapeamento do arquivo correspondendo cada campo necessário para a criação de produto com a correspondência no arquivo que foi importado conforme Figura 71, e em seguida é selecionado a opção “Executar importador”, o processo é executado em alguns minutos.

Figura 71 – Tela com a seleção de colunas para importação em massa de produtos no site



Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

Após a finalização da etapa anterior, é verificado os produtos importados selecionando a opção Produtos->Todos Produtos conforme Figura 72

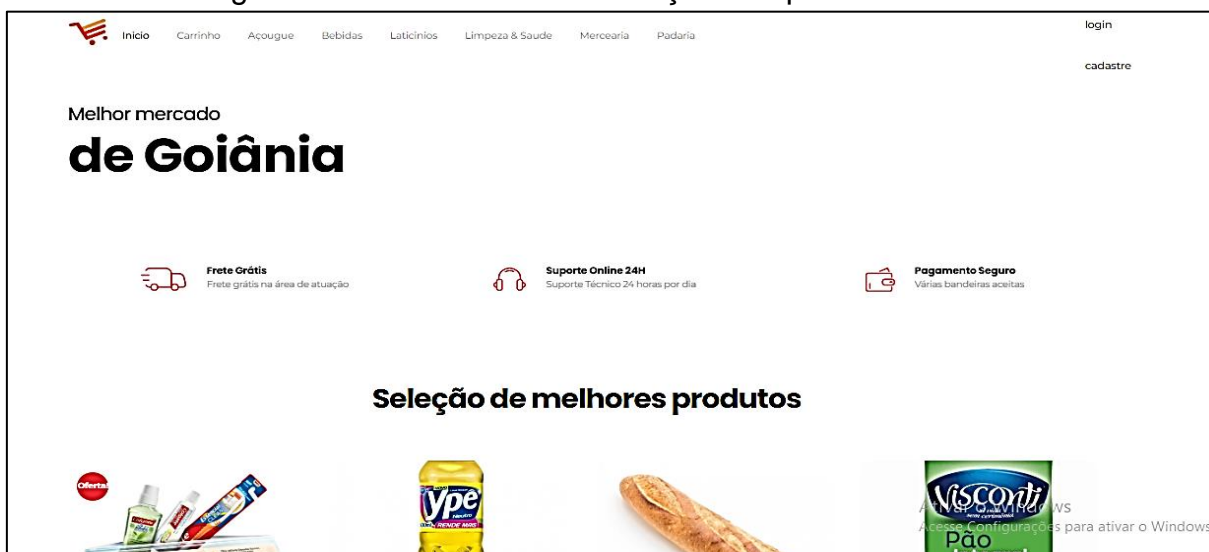
Figura 72– Tela com a visualização dos produtos importados e os respectivos atributos

| Nome | SKU | Estoque | Preço | Categorias | Tags | Data |
|-------------------|------|------------|------------------------------|------------------|--------------|-------------------------------|
| Sabão-de-quada | 1065 | Em estoque | R\$4,00 | Limpieza e Saúde | Limpieza | Publicado 04/11/2022 às 09:36 |
| Kit Saúde Bucal | 1066 | Em estoque | R\$44,99 R\$50,00 | Limpieza e Saúde | shop digital | Publicado 04/11/2022 às 09:36 |
| Kit saúde da pele | 1064 | Em estoque | R\$67,00 | Limpieza e Saúde | Limpieza | Publicado 04/11/2022 às 09:36 |
| Detergente | 1062 | Em estoque | R\$4,00 | Limpieza e Saúde | shop digital | Publicado 04/11/2022 às 09:36 |
| Pão Integral | 1061 | Em estoque | R\$11,00 | Padaria | Padaria | Publicado 04/11/2022 às 09:36 |
| Pão baguete | 1059 | Em estoque | R\$14,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Pão de queijo | 1060 | Em estoque | R\$10,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Pão francês | 1058 | Em estoque | R\$12,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Fermento | 1057 | Em estoque | R\$5,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Farinha | 1056 | Em estoque | R\$44,99 R\$50,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Café instantâneo | 1055 | Em estoque | R\$4,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |
| Café | 1054 | Em estoque | R\$4,00 | Padaria | shop digital | Publicado 04/11/2022 às 09:36 |

Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

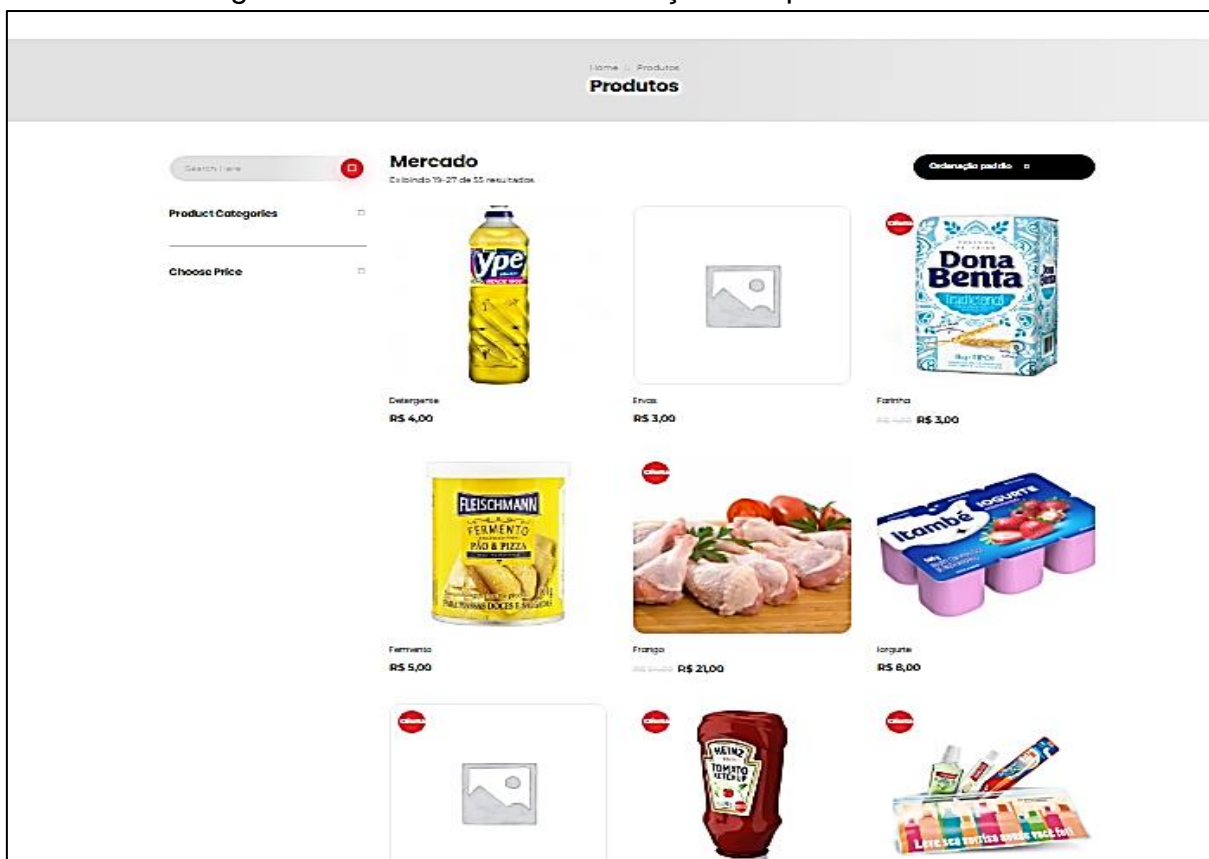
Na próxima etapa é verificado no próprio *site* os novos produtos e ajustes feitos nas etapas anteriores como as páginas de produtos, login, cadastro. Conforme ilustrado nas Figuras 73, 74, 75 e 76.

Figura 73 – Tela com a visualização dos produtos no site



Fonte: Tela de captura do Wordpress com conteúdo desenvolvido pelo autor deste trabalho

Figura 74 – Tela com a visualização dos produtos no site



Fonte: Tela de captura do Wordpress com conteúdo desenvolvido pelo autor deste trabalho

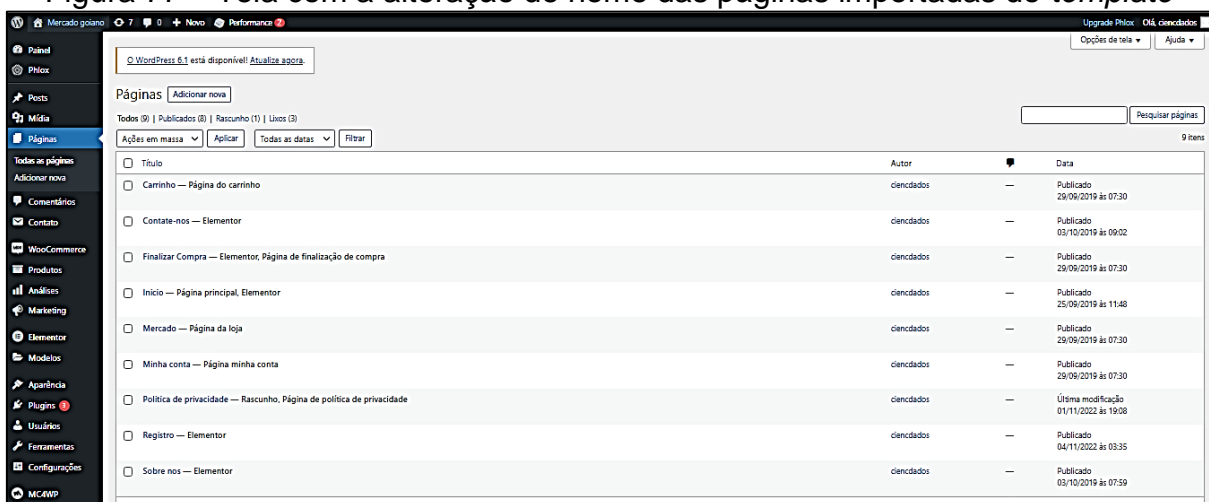
Figura 75 – Tela com a visualização do cadastro de cliente no *site*

Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

Figura 76 – Tela com a visualização de *login* de cliente no *site*

Fonte: Tela de captura do *WordPress* com conteúdo desenvolvido pelo autor deste trabalho

Na etapa seguinte é selecionado a opção Páginas->Todas as páginas para visualizar todas as páginas importadas pelo template e feito a renomeação de cada página na língua portuguesa conforme Figura 77.

Figura 77 – Tela com a alteração de nome das páginas importadas do *template*

The screenshot shows the WordPress admin interface for 'Mercado goiano'. The 'Páginas' (Pages) section is active, displaying a list of pages imported from a template. The interface includes a sidebar with navigation options like 'Painel', 'Posts', 'Mídia', and 'Páginas'. The main content area displays a list of pages with columns for 'Título', 'Autor', and 'Data'. A search bar and filter options are visible at the top of the list.

| <input type="checkbox"/> | Título | Autor | Data |
|--------------------------|---|------------|--|
| <input type="checkbox"/> | Carrinho — Página do carrinho | ciencdados | — Publicado 29/09/2019 às 07:30 |
| <input type="checkbox"/> | Contate-nos — Elementor | ciencdados | — Publicado 03/10/2019 às 09:02 |
| <input type="checkbox"/> | Finalizar Compra — Elementor, Página de finalização de compra | ciencdados | — Publicado 29/09/2019 às 07:30 |
| <input type="checkbox"/> | Início — Página principal, Elementor | ciencdados | — Publicado 25/09/2019 às 11:48 |
| <input type="checkbox"/> | Mercado — Página da loja | ciencdados | — Publicado 29/09/2019 às 07:30 |
| <input type="checkbox"/> | Minha conta — Página minha conta | ciencdados | — Publicado 29/09/2019 às 07:30 |
| <input type="checkbox"/> | Política de privacidade — Rascunho, Página de política de privacidade | ciencdados | — Última modificação 01/11/2022 às 19:08 |
| <input type="checkbox"/> | Registro — Elementor | ciencdados | — Publicado 04/11/2022 às 03:35 |
| <input type="checkbox"/> | Sobre nos — Elementor | ciencdados | — Publicado 03/10/2019 às 07:59 |

Fonte: Tela de captura do *Wordpress-admin* com conteúdo desenvolvido pelo autor deste trabalho

APÊNDICE F – CRIAÇÃO SIMULADA DA TABELA DE REGISTROS DE COMPRA DO MERCADO GOIANO

Foi acessado a plataforma *Kaggle*, conforme Figura 78, que é uma plataforma muito conhecida em *Data Science* que contém diversos datasets para serem utilizados. O *dataset* selecionado foi o *Groceries Market Basket Dataset* que contém mais de 9 mil registros de compras e 169 itens diferentes.

Figura 78 – Tela do site *kaggle*, que contém bancos de dados gratuitos

The screenshot shows the Kaggle dataset page for 'Conjunto de dados da cesta de mercado de mantimentos'. The page features a search bar at the top, a navigation menu on the left, and a main content area with the dataset title and description. The description states: 'O conjunto de dados contém 9.835 transações e 169 itens exclusivos'. There is also a 'Baixar (172 kB)' button and a 'Novo bloco de anotações' button. The page is in Portuguese and includes a sidebar with navigation options like 'Casa', 'Competições', and 'Conjuntos de dados'.

Fonte: Tela de captura do processo de criação do banco de dados com conteúdo desenvolvido pelo autor deste trabalho

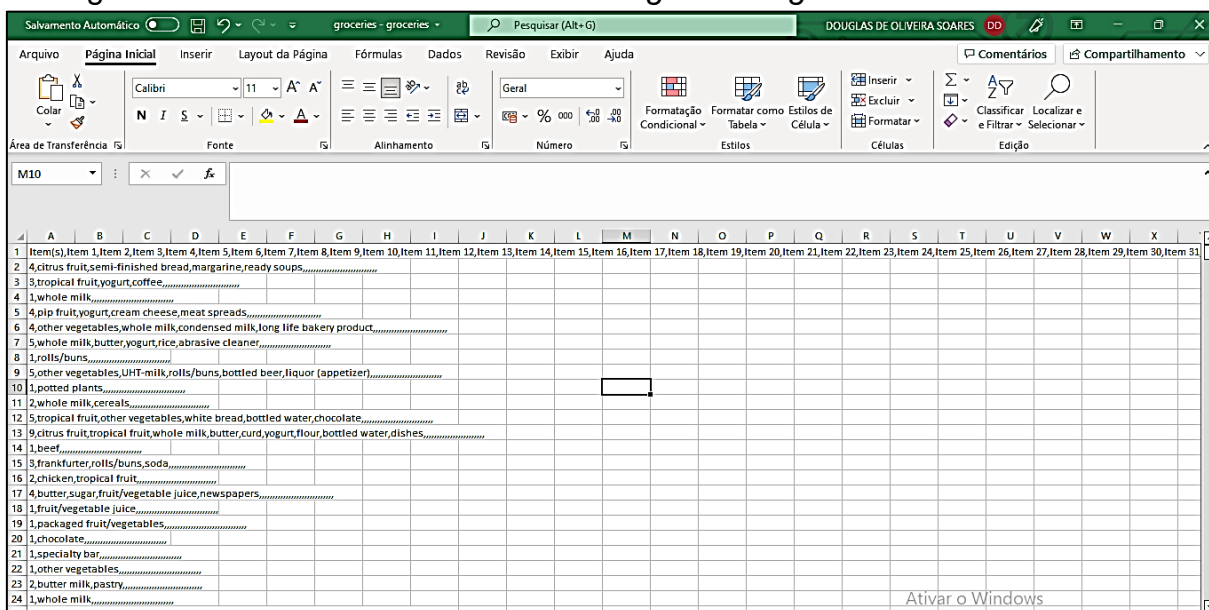
Foi utilizado o arquivo de nome “*groceries-groceries*” conforme figura 79

Figura 79 – Tela dos arquivos baixados no formato csv do site *kaggle*

| Nome | Data de modificação | Tipo | Tamanho |
|-----------------------|---------------------|----------------------|---------|
| groceries - groceries | 10/11/2022 18:23 | Arquivo de Valore... | 785 KB |
| groceries | 10/11/2022 18:23 | Arquivo de Valore... | 499 KB |

Fonte: Tela de captura do processo de criação do banco de dados com conteúdo desenvolvido pelo autor deste trabalho

Figura 80 – Tela com o banco de dados “groceries-groceries” aberto no excel



Fonte: Tela de captura do processo de criação do banco de dados com conteúdo desenvolvido pelo autor deste trabalho

Em seguida, como próximo passo foi feita a conversão do arquivo em tabela para melhor visualização e tratamento conforme Figura 81.

Figura 81 – Tela com o banco de dados “groceries-groceries” em formato tabela

| Item | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 | Item 7 | Item 8 | Item 9 | Item 10 | Item 11 | Item 12 |
|------|---------------------------|---------------------|-----------------------|--------------------------|--------------------|--------------------|---------------|---------------|---------|---------|------------------|---------|
| 1 | citrus fruit | semi-finished bread | margarine | ready soups | | | | | | | | |
| 2 | tropical fruit | yogurt | coffee | | | | | | | | | |
| 3 | whole milk | | | | | | | | | | | |
| 4 | pip fruit | yogurt | cream cheese | meat spreads | | | | | | | | |
| 5 | other vegetables | whole milk | condensed milk | long life bakery product | | | | | | | | |
| 6 | whole milk | butter | yogurt | rice | abrasive cleaner | | | | | | | |
| 7 | rolls/buns | | | | | | | | | | | |
| 8 | other vegetables | UHT-milk | rolls/buns | bottled beer | liquor (appetizer) | | | | | | | |
| 9 | potted plants | | | | | | | | | | | |
| 10 | whole milk | cereals | | | | | | | | | | |
| 11 | tropical fruit | other vegetables | white bread | bottled water | chocolate | | | | | | | |
| 12 | citrus fruit | tropical fruit | whole milk | butter | curd | yogurt | flour | bottled water | dishes | | | |
| 13 | beef | | | | | | | | | | | |
| 14 | frankfurter | rolls/buns | soda | | | | | | | | | |
| 15 | chicken | tropical fruit | | | | | | | | | | |
| 16 | butter | sugar | fruit/vegetable juice | newspapers | | | | | | | | |
| 17 | fruit/vegetable juice | | | | | | | | | | | |
| 18 | packaged fruit/vegetables | | | | | | | | | | | |
| 19 | chocolate | | | | | | | | | | | |
| 20 | specialty bar | | | | | | | | | | | |
| 21 | other vegetables | | | | | | | | | | | |
| 22 | butter milk | pastry | | | | | | | | | | |
| 23 | whole milk | cream cheese | processed cheese | detergent | newspapers | | | | | | | |
| 24 | tropical fruit | root vegetables | other vegetables | frozen dessert | rolls/buns | flour | sweet spreads | salty snack | waffles | candy | bathroom cleaner | |
| 25 | bottled water | canned beer | | | | | | | | | | |
| 26 | yogurt | | | | | | | | | | | |
| 27 | sausage | rolls/buns | soda | chocolate | | | | | | | | |
| 28 | other vegetables | | | | | | | | | | | |
| 29 | brown bread | soda | fruit/vegetable juice | canned beer | newspapers | shopping bags | | | | | | |
| 30 | yogurt | beverage | bottled water | specialty bar | | | | | | | | |
| 31 | hamburger meat | other vegetables | rolls/buns | spices | bottled water | household articles | nackins | | | | | |

Fonte: Tela de captura do processo de criação do banco de dados com conteúdo desenvolvido pelo autor deste trabalho

Cada linha do arquivo representa uma compra podendo ter até 32 itens. Na base simulada fizemos a redução para até 10 itens por compra e fazendo a tradução de linguagem para o produto correspondente em português, de acordo com a lista

criada de itens para um mercado de pequeno porte, conforme ilustrado na Figura 82.

Figura 82 – Tela com o banco de dados “groceries-groceries” ajustado

| # | A | B | C | D | E | F | G | H | I | J |
|----|------------------------|--------------------|-------------------|--------------------|--------------------|--------------------|-------------------|--------------------|-----------------|--------------------|
| 1 | ITEM 1 | ITEM 2 | ITEM 3 | ITEM 4 | ITEM 5 | ITEM 6 | ITEM 7 | ITEM 8 | ITEM 9 | ITEM 10 |
| 2 | Tomate | Kit saude da pele | Tomate | Limao | Molho de tomate | Maca | Linguica | Pao Baguete | Detergente | Pao Baguete |
| 3 | Vinagre | Tomate | | | | | | | | |
| 4 | Pao Frances | Mandioca | Chá gelado | Açucar | Jornal | Alface | Café instantaneo | Maionese | Batata | Caju |
| 5 | Balinha | Café | Cebola | | | | | | | |
| 6 | Uva | Sabão de quadra | Chá gelado | Açucar | Ervas | Pao Integral | | | | |
| 7 | Balinha | Açucar | Detergente | Sucos de caixinha | Sal | Chiclete | Café | Café | Salsicha | Ervas |
| 8 | Café instantaneo | fermento | Café | Barra de chocolate | Pao de queijo | Vinagre | Café | Leite Condensando | | |
| 9 | Salsicha | | | | | | | | | |
| 10 | Presunto | Carne Bovina | Detergente | Carne Bovina | Carne de Frango | Alface | Uva | Arroz 5kg | Maca | Arroz 5kg |
| 11 | Ervas | Biscoito de queijo | Caju | Kit saude da pele | farinha | | | | | |
| 12 | Café instantaneo | Tomate | Açucar | Presunto | Uva | Café | Coca-Cola | Pipoca | | |
| 13 | Tomate | Balinha | Manteiga | | | | | | | |
| 14 | Garrafa de água 300 ml | | | | | | | | | |
| 15 | Biscoito de queijo | Laranja | Kit Saude Bucal | Ketchup | Limao | Molho de tomate | Vinagre | | | |
| 16 | Kit saude da pele | Uva | | | | | | | | |
| 17 | Sabão de quadra | Cerveja | Presunto | Caju | Pao Baguete | Chiclete | Sal | Leite Condensando | Kit Saude Bucal | Caju |
| 18 | Sabão de quadra | Kit Saude Bucal | Salsicha | Tomate | Kit saude da pele | | | | | |
| 19 | Cerveja | Presunto | Carne de Frango | Tomate | Kit saude da pele | | | | | |
| 20 | Carne de Frango | Presunto | Palmito | | | | | | | |
| 21 | Óleo de soja | Limao | Café | | | | | | | |
| 22 | fermento | Pao Baguete | Manteiga | Óleo de soja | Sabão de quadra | Manteiga | Óleo de soja | Limao | Linguica | Maca |
| 23 | Alface | Biscoito de queijo | Chá gelado | Limao | Sal | Ketchup | Balinha | Leite Condensando | Sal | Kit Saude Bucal |
| 24 | farinha | Carne de Frango | Pao Baguete | Molho de tomate | Carne Suina | Sabão de quadra | Café | Kit saude da pele | Pao Baguete | Sucos de caixinha |
| 25 | Detergente | Chiclete | Açucar | Maionese | Pipoca | Carne de Frango | Macarrao | Carne de Frango | | |
| 26 | Pipoca | Caju | Kit saude da pele | Chiclete | Biscoito de queijo | Barra de chocolate | Laranja | Biscoito de queijo | Óleo de soja | Arroz 5kg |
| 27 | Limao | Carne Suina | Arroz 5kg | Tomate | Arroz 5kg | Pipoca | Molho de tomate | Linguica | Pao Baguete | Vinagre |
| 28 | Salsicha | Detergente | farinha | Pao Integral | Carne Bovina | Chá gelado | Tomate | lorgurte | | |
| 29 | Sal | Sabão de quadra | Molho de tomate | | | | | | | |
| 30 | Macarrao | Pao Frances | Sal | Cerveja | Ketchup | Presunto | | | | |
| 31 | Maionese | Limao | Ovo | Pao Frances | Batata | Sal | Sucos de caixinha | Macarrao | farinha | Café instantaneo |
| 32 | Manteiga | Limao | Açucar | Tomate | Chiclete | Pao Baguete | farinha | Caju | | |
| 33 | Pao Integral | Pao Baguete | Tomate | Maca | Laranja | Alface | Uva | Carne de Frango | Maca | Barra de chocolate |

Fonte: Tela de captura do processo de criação do banco de dados com conteúdo desenvolvido pelo autor deste trabalho

Conforme Figura 83, segue os itens que compõem a loja virtual.

Figura 83 – Tela com os itens contemplados no banco de dados da loja

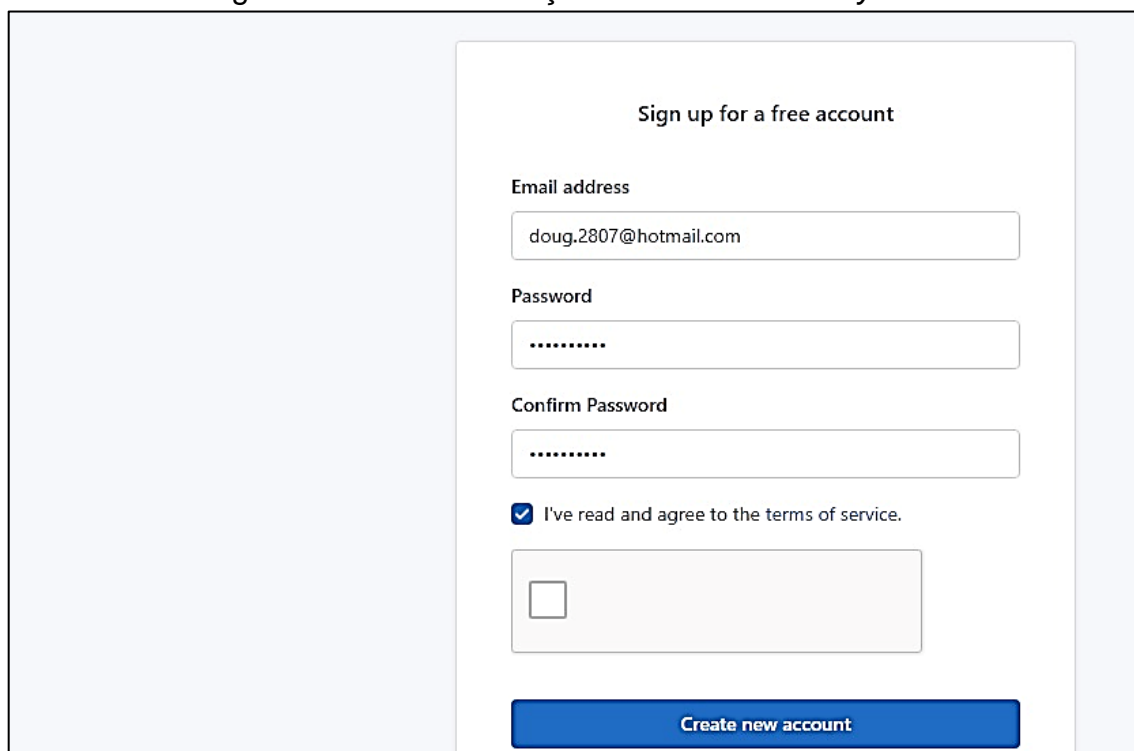
| ITENS | |
|---------------------------|----------------------|
| 1 Açucar | 30 Leite Condensando |
| 2 Alface | 31 Leite UHT |
| 3 Arroz 5kg | 32 Limao |
| 4 Balinha | 33 Linguica |
| 5 Barra de chocolate | 34 Maca |
| 6 Batata | 35 Macarrao |
| 7 Biscoito de queijo | 36 Maionese |
| 8 Café | 37 Mandioca |
| 9 Café instantaneo | 38 Manteiga |
| 10 Caju | 39 Molho de tomate |
| 11 Carne Bovina | 40 Óleo de soja |
| 12 Carne de Frango | 41 Ovo |
| 13 Carne Suina | 42 Palmito |
| 14 Cebola | 43 Pao Baguete |
| 15 Cerveja | 44 Pao de queijo |
| 16 Chá gelado | 45 Pao Frances |
| 17 Chiclete | 46 Pao Integral |
| 18 Coca-Cola | 47 Pipoca |
| 19 Detergente | 48 Presunto |
| 20 Ervas | 49 Sabão de quadra |
| 21 farinha | 50 Sal |
| 22 fermento | 51 Salsicha |
| 23 Garrafa de água 300 ml | 52 Sucos de caixinha |
| 24 lorgurte | 53 Tomate |
| 25 Jornal | 54 Uva |
| 26 Ketchup | 55 Vinagre |
| 27 Kit Saude Bucal | |
| 28 Kit saude da pele | |
| 29 Laranja | |

Fonte: Tela do processo de criação do banco de dados desenvolvido pelo autor deste trabalho

APÊNDICE G – HOSPEDAGEM NA *INFINITY FREE*

Foi feito o acesso ao site da *Infinity Free* a seguir: <https://app.infinityfree.net/login> e criado uma conta de maneira simples com os dados de nome, *email* e senha, conforme Figura 84.

Figura 84 – Tela de criação de conta na *Infinity Free*



The screenshot shows a registration form titled "Sign up for a free account". The form includes the following fields and elements:

- Email address:** A text input field containing the email address "doug.2807@hotmail.com".
- Password:** A text input field with masked characters (dots).
- Confirm Password:** A text input field with masked characters (dots).
- Terms of Service:** A checked checkbox followed by the text "I've read and agree to the terms of service."
- Image Placeholder:** A square box with a small square icon inside, likely for a profile picture.
- Create new account:** A blue button at the bottom of the form.

Fonte: Tela de captura de criação de conta na *Infinity Free* com conteúdo desenvolvido pelo autor deste trabalho

No próximo passo é solicitado o nome e domínio do *site* que deseja construir e a senha de administrador conforme Figura 85.

Figura 85 – Tela de criação de *site* na *Infinity Free*

The screenshot shows the 'New Hosting Account: Enter Additional Information' page. At the top, there's a navigation bar with 'Home', 'Profile', 'Accounts', 'Free SSL Certificates', 'Site Builders', 'Domain Checker', 'Knowledge Base', and 'Community Forum'. A 'Go Premium' button is in the top right. Below the navigation, there's a banner for Adobe Stock with the text 'Aprimore os seus projetos. Aproveite modelos grátis do Adobe Stock para email.' and a 'Baixe grátis' button. A green notification box states 'The domain storepremiumgyn.rfgd is available!'. The main form area is titled 'Step 2: Enter Additional Information' and contains three input fields: 'Account Label' (with the value 'Website for storepremiumgyn.rfgd'), 'Account Username' (with the value '(generated automatically)'), and 'Account Password' (with the value '#Brasil2807').

Fonte: Tela de captura de criação de *site* na *Infinity Free* com conteúdo desenvolvido pelo autor deste trabalho

Após a conta criada foi clicado no botão “*Open Control Panel*” para acessar os serviços e configurações disponíveis ao *site* criado, conforme Figura 86.

Figura 86 – Tela de finalização da criação do *site* na *Infinity Free*

The screenshot shows the 'New Hosting Account: Account Created' page. At the top, there's a navigation bar with 'Home', 'Profile', 'Accounts', 'Free SSL Certificates', 'Site Builders', 'Domain Checker', 'Knowledge Base', and 'Community Forum'. A 'Go Premium' button is in the top right. Below the navigation, there's a banner for HBO Max with the text 'DESCUBRA POR R\$ 14,16 /MÊS* HBO MAX ASSINE AGORA'. A green notification box states 'Your account has been created with username epiz_32497916!'. The main form area is titled 'Step 3: Account Created' and contains a message: 'Your account has been created with username epiz_32497916! Here are some things you need to know:'. Below the message is a list of instructions:

- It will take a few minutes for your account to be set up, but you can login to the control panel already.
- It can take up to 72 hours for the new domain to be visible everywhere due to DNS caching.
- Please login to the control panel once to enable all features.
- Not sure what to do next? Please see this guide for some ideas on how to get started.

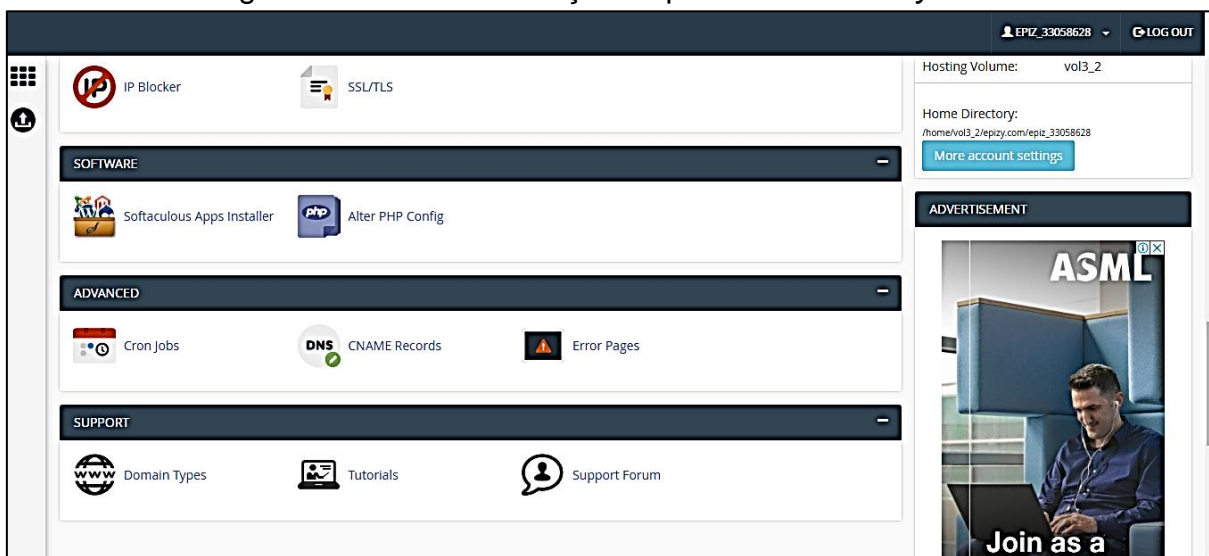
At the bottom of the form area, there is a large green button labeled 'Open Control Panel' and a blue button labeled 'Finish'. A 'Back' button is also visible at the bottom left.

Fonte: Tela de captura de criação de *site* na *Infinity Free* com conteúdo desenvolvido pelo autor deste

trabalho

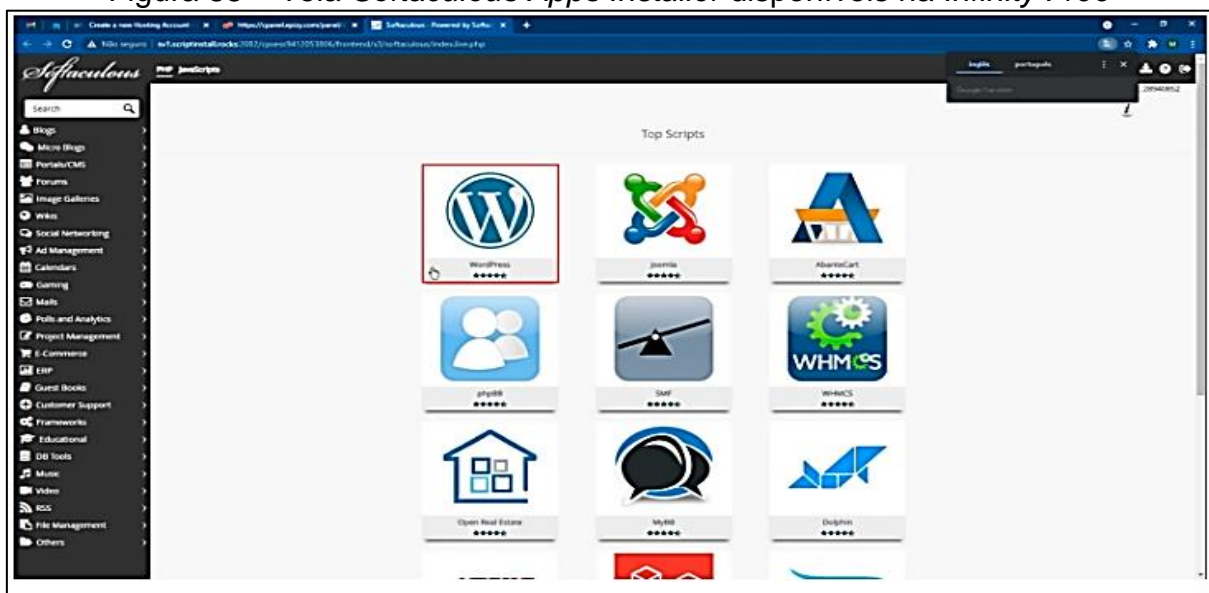
Na próxima etapa é selecionado a opção dentro do painel de controle da *Infinity Free* o serviço: “*Software->Softaculous Apps Installer->Wordpress*”, conforme Figuras 87 e 88.

Figura 87 – Tela de serviços disponíveis na *Infinity Free*



Fonte: Tela de captura de seleção de serviços na *Infinity Free* com conteúdo desenvolvido pelo autor

Figura 88 – Tela *Softaculous Apps Installer* disponíveis na *Infinity Free*



Fonte: Tela de captura de seleção do *Wordpress* na *Infinity Free* desenvolvido pelo autor

Após a criação do serviço do *Wordpress* na *Infinity Free* os passos seguintes foram feitos conforme o Apêndice E, que contém as configurações de temas e do

mercado criado para este trabalho.

RESOLUÇÃO n° 038/2020 – CEPE

ANEXO I

APÊNDICE ao TCC

Termo de autorização de publicação de produção acadêmica

O(A) estudante Douglas de Oliveira Soares
do Curso de Ciências da Computação, matrícula 20172002300267,
telefone: 981301725 e-mail soares2007@hotmail.com, na qualidade de titular dos
direitos autorais, em consonância com a Lei n° 9.610/98 (Lei dos Direitos do autor),
autoriza a Pontifícia Universidade Católica de Goiás (PUC Goiás) a disponibilizar o
Trabalho de Conclusão de Curso intitulado
Big Data e Marketing Digital,
gratuitamente, sem ressarcimento dos direitos autorais, por 5
(cinco) anos, conforme permissões do documento, em meio eletrônico, na rede mundial
de computadores, no formato especificado (Texto (PDF); Imagem (GIF ou JPEG); Som
(WAVE, MPEG, AIFF, SND); Vídeo (MPEG, MWV, AVI, QT); outros, específicos da
área; para fins de leitura e/ou impressão pela internet, a título de divulgação da
produção científica gerada nos cursos de graduação da PUC Goiás.

Goiânia, 12 de AGOSTO de 2022.

Assinatura do(s) autor(es): Douglas

Nome completo do autor: Douglas de Oliveira Soares

Assinatura do professor-orientador: Angélica

Nome completo do professor-orientador: Angélica da Silva Nunes