**UNIVERSIDADE PRESBITERIANA MACKENZIE**

**O CONJUNTO DE DADOS DE TRANSAÇÕES DE VENDAS DE CAFÉ**

**SÃO PAULO**

**2025**

Julia Petra T. Chaves

Leandro Henrique Pereira

Murilo Freitas Rodrigues

Thalles Rodrigues

**Conjunto de dados de transações de vendas de café**

Trabalho de projeto aplicado I, apresentado como exigência para obtenção de título em Banco de Dados da Universidade Presbiteriana Mackenzie de São Paulo / SP

**Professor:** Ismar Frango

**Disciplina:** Projeto Aplicado I

**Turma:** Turma 2025

SÃO PAULO

2025

**RESUMO**

A presente pesquisa se propõe a produzir uma dissertação sobre a análise dos dados de vendas de uma cafeteria, com vistas a compreender os padrões de consumo e propor estratégias que possam contribuir para o aumento da receita e a otimização dos processos operacionais. A investigação contemplará fatores como sazonalidade, variações mensais de faturamento, produtos mais comercializados, bem como a viabilidade de ajustes na precificação de acordo com o horário. Para além da análise descritiva, serão sugeridas ações voltadas à fidelização da clientela, incluindo promoções em dias de menor movimento e estratégias direcionadas ao fortalecimento do vínculo com os consumidores. O estudo será realizado por meio da linguagem de programação Python, utilizando a biblioteca Pandas como recurso central para a manipulação dos dados, construção de indicadores e geração de representações gráficas. Espera-se, como resultado, a proposição de insights que favoreçam a tomada de decisão da gestão, apoiando a formulação de estratégias de precificação, organização de escalas, elaboração de campanhas promocionais e melhoria da experiência do cliente.

**Palavras-chave:** Cafeteria, Análise de vendas, Sazonalidade, Precificação, Fidelização de clientes, Python, Pandas.

**ABSTRACT**

The purpose of this project is to analyze the sales data of a coffee shop, with the aim of understanding consumption patterns and proposing strategies that may contribute to revenue growth and the optimization of operational processes. The investigation will address factors such as seasonality, monthly revenue variations, best-selling products, as well as the feasibility of price adjustments according to the time of day. In addition to descriptive analysis, actions aimed at customer loyalty will be suggested, including promotions on low-demand days and strategies designed to strengthen customer relationships. The study will be conducted using the Python programming language, with the Pandas library as a central resource for data manipulation, indicator construction, and the generation of graphical representations. The expected outcome is the proposal of insights that support management decision-making, contributing to the development of pricing strategies, work schedule organization, promotional campaign planning, and the enhancement of customer experience.

**Keywords:** Coffee shop, Sales analysis, Seasonality, Pricing, Customer loyalty, Python, Pandas

# **LISTA DE TABELA E IMAGENS**

(serão acrescentadas no decorrer das entregas)

# **LISTA DE GRÁFICOS**

(serão acrescentados no decorrer das entregas)

**SUMÁRIO**

1. Introdução…………………………………………………………….................. 2

2. Objetivos……………………………………………………………….................3

3. Justificativa………………………………………………………….....................4

4. Cronograma de atividades……………………………………………………....5

5. Definição da empresa…………………………………………………………….6

6. Problema da empresa…………………………………………………………….7

7. Definição do Dataset......................................................................................9

8. Definição dos Metadados...............................................................................10

9. Links dos dados e respositório.......................................................................11

# **INTRODUÇÃO**

O café ocupa um espaço de grande relevância na cultura de consumo mundial, configurando-se não apenas como uma das bebidas mais populares, mas também como importante fonte de receita para inúmeras cafeterias. No atual cenário de crescente competitividade do setor de alimentos e bebidas, compreender os padrões de compra e o comportamento dos consumidores torna-se um diferencial estratégico para a manutenção e expansão dos negócios. A análise criteriosa dessas informações possibilita não apenas otimizar o portfólio de produtos, mas também aprimorar o planejamento de estoques, direcionar campanhas promocionais e maximizar os lucros por meio de decisões fundamentadas em dados.

Nesse contexto, a utilização de bases de dados extraídas de transações comerciais surge como um recurso fundamental para compreender a dinâmica de consumo em estabelecimentos específicos. O presente estudo parte de um conjunto de dados coletados em uma cafeteria, contendo informações detalhadas sobre vendas, formas de pagamento, horários de compra e preferências de clientes. A análise desse material oferece subsídios valiosos para a identificação de tendências, sazonalidades e padrões de comportamento ao longo do tempo, elementos indispensáveis para a gestão eficiente e para a formulação de estratégias comerciais alinhadas às demandas reais do público consumidor.

A pesquisa adota uma abordagem quantitativa e exploratória, fundamentada em técnicas de análise descritiva, visualização de dados e testes estatísticos. Por meio dessas ferramentas, busca-se não apenas interpretar o desempenho do negócio, mas também propor melhorias práticas, tais como a definição de promoções em horários de menor movimento, o ajuste do quadro de funcionários em períodos de alta demanda e a elaboração de campanhas voltadas a produtos com maior potencial de crescimento.

Dessa forma, espera-se que este trabalho contribua para o fortalecimento do processo de tomada de decisão na cafeteria estudada, fornecendo recomendações baseadas em evidências empíricas. Além disso, a pesquisa reforça a importância da análise de dados como instrumento de apoio à gestão empresarial, demonstrando como sua aplicação pode gerar vantagens competitivas e promover uma atuação mais eficiente e estratégica no mercado.

**Objetivos do Projeto**

**Objetivo Geral**Realizar uma análise exploratória, quantitativa e descritiva dos dados de vendas de uma cafeteria, com o intuito de identificar padrões de consumo, tendências sazonais e oportunidades de melhoria que subsidiem decisões estratégicas voltadas à operação e ao desempenho comercial.

**Objetivos Específicos**

* Examinar o conjunto de dados disponíveis, contemplando informações sobre volume de vendas, receita, formas de pagamento, horários de compra e preferências dos clientes;
* Identificar os produtos mais populares e rentáveis, por meio da análise do volume de vendas e da receita gerada por tipo de café;
* Avaliar a distribuição das vendas ao longo do dia e da semana, destacando horários de pico e períodos de menor movimento;
* Investigar a sazonalidade mensal e as variações percentuais de vendas, identificando tendências de crescimento ou queda;
* Calcular métricas de desempenho, como o ticket médio por dia e por período do dia;
* Aplicar testes estatísticos a fim de verificar a significância das diferenças entre produtos e períodos de consumo,
* Elaborar visualizações gráficas e um relatório de insights que apoiem a formulação de recomendações gerenciais, como promoções direcionadas, ajustes no quadro de funcionários e estratégias para o mix de produtos.

**Justificativa**

O setor de cafeterias tem se consolidado como um segmento dinâmico e competitivo no mercado de alimentos e bebidas, exigindo cada vez mais estratégias de gestão baseadas em informações precisas e atualizadas. A compreensão dos padrões de consumo, da sazonalidade e do comportamento dos clientes torna-se fundamental para a definição de políticas comerciais, planejamento operacional e maximização da receita.

Nesse cenário, a análise de dados se apresenta como uma ferramenta indispensável para apoiar a tomada de decisão, permitindo transformar registros de transações em informações estratégicas. A utilização de métodos quantitativos e exploratórios, aliada à aplicação de técnicas de análise descritiva, possibilita não apenas a identificação de tendências, mas também a antecipação de demandas e a proposição de ações práticas voltadas ao crescimento sustentável do negócio.

A relevância deste projeto justifica-se, portanto, pela possibilidade de oferecer subsídios concretos para a gestão de uma cafeteria, a partir da identificação de oportunidades de melhoria em aspectos como precificação, mix de produtos, alocação de funcionários e definição de promoções em períodos específicos. Além de contribuir para a eficiência operacional e para a otimização dos resultados financeiros, este estudo reforça a importância da cultura de decisões orientadas por dados no contexto empresarial contemporâneo.

**Cronograma de atividades**

|  |  |  |
| --- | --- | --- |
| **Dia** | **Etapa** | **Atividades** |
| 06/09/2025 | Planejamento e organização inicial | Definição do escopo do projeto, objetivos gerais e específicos, delimitação do problema e distribuição das responsabilidades entre os integrantes. Estabelecimento dos critérios metodológicos a serem seguidos. |
| 07/09/2025 | Levantamento e análise preliminar dos dados | Organização, verificação e anonimização dos registros de transações da cafeteria. Preparação do conjunto de dados para análise estatística e visualização. |
| 08/09/2025 | Construção da fundamentação teórica e metodológica | Revisão da literatura sobre consumo de café, análise exploratória de dados e gestão de cafeterias. Estruturação da abordagem metodológica, incluindo técnicas de análise descritiva e exploratória. |
| 09/09/2025 | Análise e sistematização dos resultados | Aplicação de técnicas quantitativas e de visualização para identificar padrões, sazonalidades e métricas de desempenho. Interpretação dos resultados à luz da teoria e das práticas gerenciais. |
| 10/09/2025 | Redação final e revisão do documento | Consolidação de todas as seções do trabalho (introdução, definição da empresa, justificativa, objetivos, metodologia, resultados e considerações finais). Revisão coletiva visando coerência, clareza e adequação às normas acadêmicas. |

**Definição da empresa**

A organização objeto deste estudo caracteriza-se como uma cafeteria pertencente ao setor de alimentos e bebidas, cuja principal atividade é a comercialização de diferentes tipos de café e produtos correlatos. Seu modelo de negócio está voltado para o atendimento direto ao consumidor, oferecendo um portfólio diversificado que contempla opções tradicionais e especiais da bebida, além de itens complementares.

Para fins desta pesquisa, as informações foram anonimizadas, de modo a preservar a identidade da empresa, mantendo-se apenas os elementos necessários para a realização de uma análise estatística consistente. A cafeteria opera diariamente, com registro contínuo de transações que abrangem dados sobre vendas, formas de pagamento, horários de compra e preferências de consumo.

Esse conjunto de informações representa a base para a investigação proposta, permitindo a identificação de padrões de comportamento e de tendências sazonais. O objetivo central é extrair informações acionáveis que possam subsidiar decisões estratégicas relacionadas ao marketing, à gestão operacional e ao planejamento financeiro, reforçando a importância da utilização de dados como suporte à gestão empresarial contemporânea.

**Problema da pesquisa**

* **Problemas Resolvidos a Partir dos Insights**

Através da Identificação dos Produtos de Maior Rentabilidade. Antes: a empresa não sabia quais cafés geravam maior receita e poderiam ser priorizados em campanhas.

Com os insights: ficou claro que produtos como Americano with Milk e Latte são os mais vendidos e rentáveis, permitindo concentrar marketing e planejamento de estoque neles.

* **Reconhecimento de Horários de Pico e Baixa Demanda**

Antes: a equipe era escalada de forma genérica, sem considerar variações de fluxo ao longo do dia.

Com os insights: foi possível ajustar escalas para reduzir ociosidade em períodos fracos e reforçar a equipe nos horários de pico, melhorando produtividade e reduzindo custos.

* **Detecção de Sazonalidade e Variação Mensal**

Antes: as compras de insumos eram feitas sem considerar sazonalidade.

Com os insights: a empresa passou a antecipar períodos de maior demanda (ex.: meses frios) e ajustar pedidos para evitar falta de produtos ou desperdício.

* **Planejamento de Promoções Inteligentes**

Antes: promoções eram feitas de forma aleatória, sem medir impacto.

Com os insights: foram definidos dias de menor faturamento (finais de semana) e horários de menor movimento (tarde/noite) para direcionar promoções e combos, aumentando o fluxo nesses períodos.

* **Monitoramento de Crescimento e Queda de Produtos**

Antes: produtos que estavam perdendo relevância não eram percebidos rapidamente.

Com os insights: foi possível detectar quedas de vendas de determinados cafés e agir rapidamente — ajustando preço, fazendo campanhas ou revisando o cardápio.

* **Melhora na Gestão Financeira**

Antes: não havia acompanhamento estruturado do ticket médio.

Com os insights: foi identificado o ticket médio por dia e por período, possibilitando campanhas de upsell e cross-sell para elevar o valor médio gasto por cliente.

* **Decisões Embasadas em Evidências**

Antes: decisões eram tomadas com base em percepção e experiência.

Com os insights: testes estatísticos (ANOVA e Kruskal-Wallis) confirmaram que as diferenças entre produtos e meses eram significativas, garantindo que as ações fossem fundamentadas em dados confiáveis. Os insights resolveram problemas de falta de visibilidade, ineficiência operacional e estratégia genérica, resultando em decisões mais inteligentes sobre estoque, promoções, equipe e planejamento financeiro.

**Definição do Dataset**

O dataset analisado contém registros históricos de vendas de café realizadas ao longo de vários meses. Cada linha da base de dados representa uma venda individual e contém informações detalhadas como:

* Data e hora da transação – permitindo análises temporais e identificação de sazonalidades;
* Tipo de café vendido – possibilitando comparações de popularidade e rentabilidade entre produtos;
* Valor monetário da venda – utilizado para calcular receita, ticket médio e desempenho financeiro.
* No total, a base conta com milhares de registros, garantindo robustez estatística e permitindo identificar tendências com maior precisão.

|  |  |  |
| --- | --- | --- |
| **Campo** | **Descrição** | **Tipo de Dado** |
| Date | Data da transação no formato AAAA-MM-DD | Data (datetime) |
| hour\_of\_day | Hora em que a venda foi realizada (0-23) | Numérico (int) |
| weekday | Dia da semana da transação (Mon-Sun) | Categórico |
| Time\_of\_Day | Período do dia (Manhã, Tarde, Noite) | Categórico |
| coffee\_name | Nome do café vendido (Latte, Americano, etc.) | Texto (string) |
| money | Valor monetário da transação (em reais) | Numérico (float) |

**Definição dos Metadados**

As informações foram anonimizadas, mantendo apenas os elementos necessários para permitir uma análise estatística consistente. A empresa opera durante todos os dias da semana, com um portfólio diversificado de cafés, e os dados refletem sua operação de vendas ao longo de um período contínuo. O objetivo é gerar um estudo que permita extrair informações acionáveis, apoiando decisões estratégicas de marketing, operação e planejamento financeiro.

**Links dos dados e repositório**

Os dados foram consumidos do site Kaggle:  
<https://www.kaggle.com/datasets/navjotkaushal/coffee-sales-dataset>  
  
Link do repositório no GitHub:

<https://github.com/ThallesRodri/Mackenzie_Projeto_Aplicado_I/tree/main>