

# **Análise dos dados referentes à base de dados: Brazilian E-Commerce Public Dataset by Olist.zip**

Autor: Frankson Leal Silva<sup>1</sup>

## **1. Introdução**

A análise de dados de forma eficiente pode ajudar a solucionar problemas e oferecer uma série de informações importantes sobre determinados cenários. O objetivo do artigo é realizar uma ETL, ou seja, extrair, transformar e carregar os dados fornecidos para estudo e fazer uma análise exploratória sobre esses dados dentro dos tópicos de vendas, logística, satisfação do cliente, financeira e de marketing. A base de dados fornecida são referentes ao e-commerce brasileiro, contendo dados importantes sobre volume de vendas, categorias de produtos, pagamentos, e outras variáveis relacionados ao tema. A linguagem utilizada para a manipulação dos dados foi o PYTHON, dentro do ambiente de desenvolvimento VS CODE.

Na seção a seguir, serão mostrados todos os resultados obtidos sobre essa análise.

## **2. Resultados e análises de dados**

### **2.1 – Análise de performance de vendas.**

Neste tópico foram utilizadas as tabelas sobre vendas e produtos. Cada tabela está relacionada sobre a coluna “product\_id”, o que permitiu essa mesclagem de dados fazendo com que fosse alcançado o resultado esperado. A seguir, está o gráfico referente ao volume de vendas do produto que teve mais êxito, seguido do gráfico referente as volumes de vendas mensais e trimestrais.

A categoria de produtos que teve mais vendas foi a de “cama\_mesa\_banho” com 11115 vendas, tendo maiores vendas nos períodos mensais de 2017-11. No período trimestral, essa categoria vendeu mais no ano de 2018. Já a categoria com menor número de vendas foi a de “seguros\_e\_servicos”.

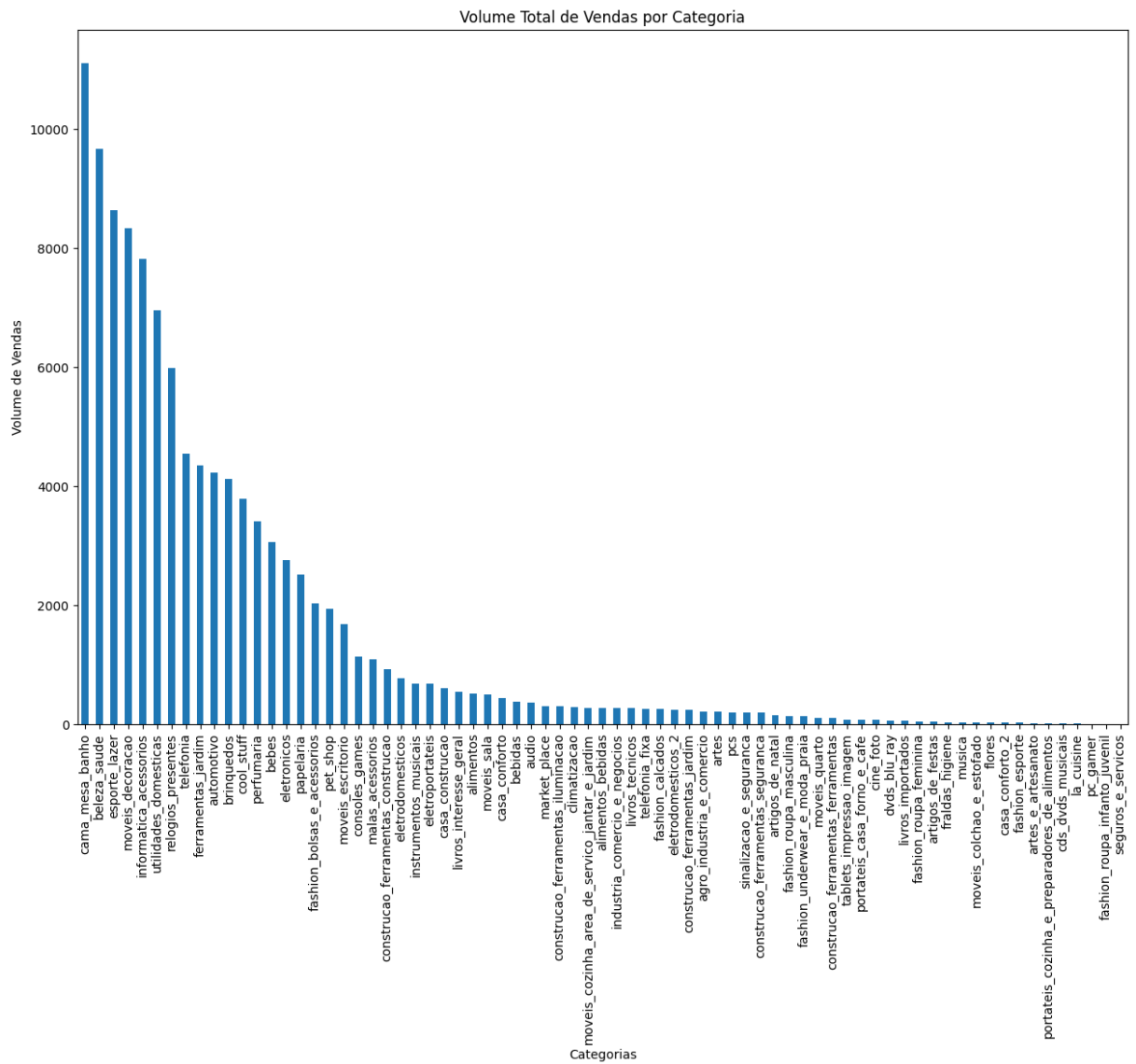


Figura 1 – Gráfico de vendas por categoria

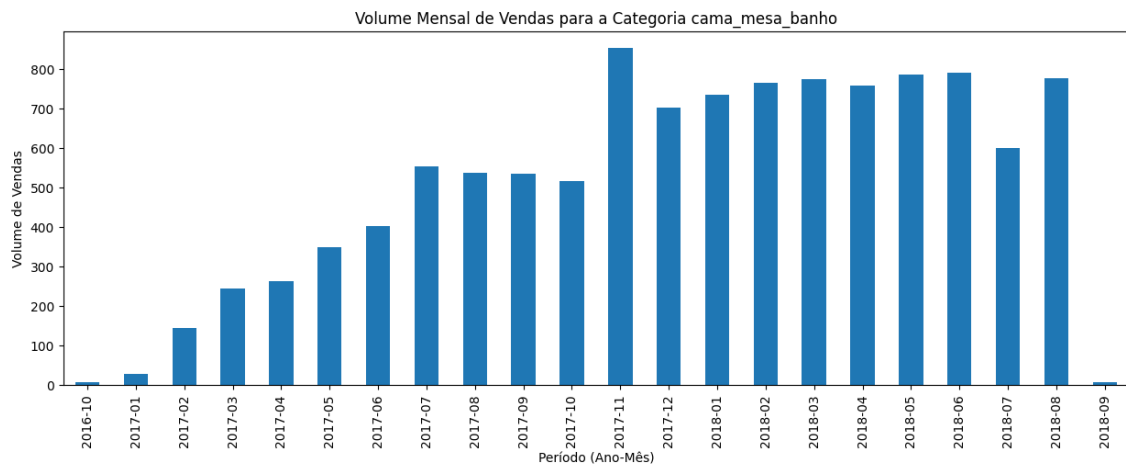


Figura 2 – Gráfico de vendas mensais

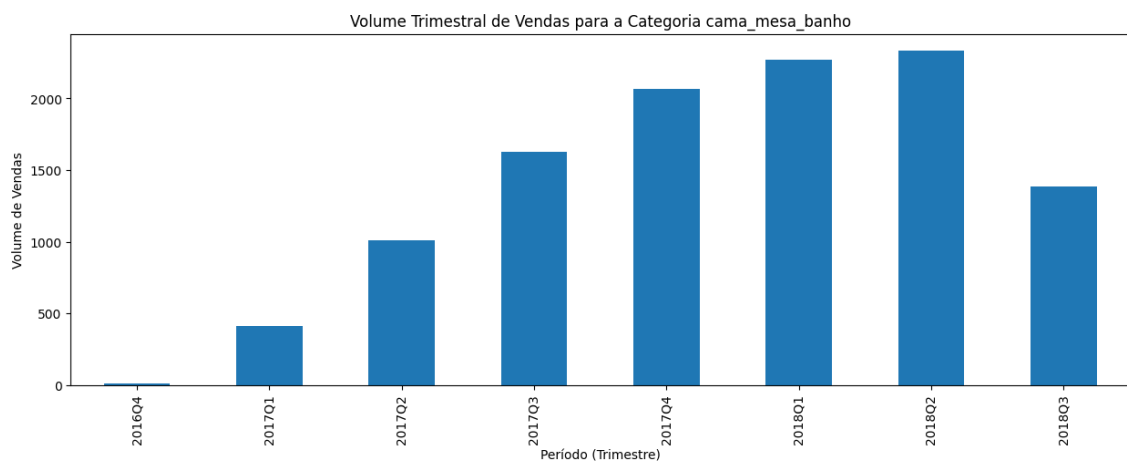


Figura 3 – Gráfico de vendas trimestralmente

## 2.2 – Análise de logística

Nesta seção, foi escolhido o tópico sobre os prazos de entregas e os fatores que influenciaram atrasos. De acordo com os resultados, o tempo médio de entrega é de 12 dias, o que para um cliente é um excelente tempo. Sobre o levantamento de informações sobre os possíveis atrasos, podemos citar a questão das outras variáveis que contém na coluna status, mostrando que existem situações de cancelado e

indisponível que resultariam assim em atrasos das entregas e consequentemente na insatisfação do cliente.

Abaixo estão os gráficos sobre essa informações.

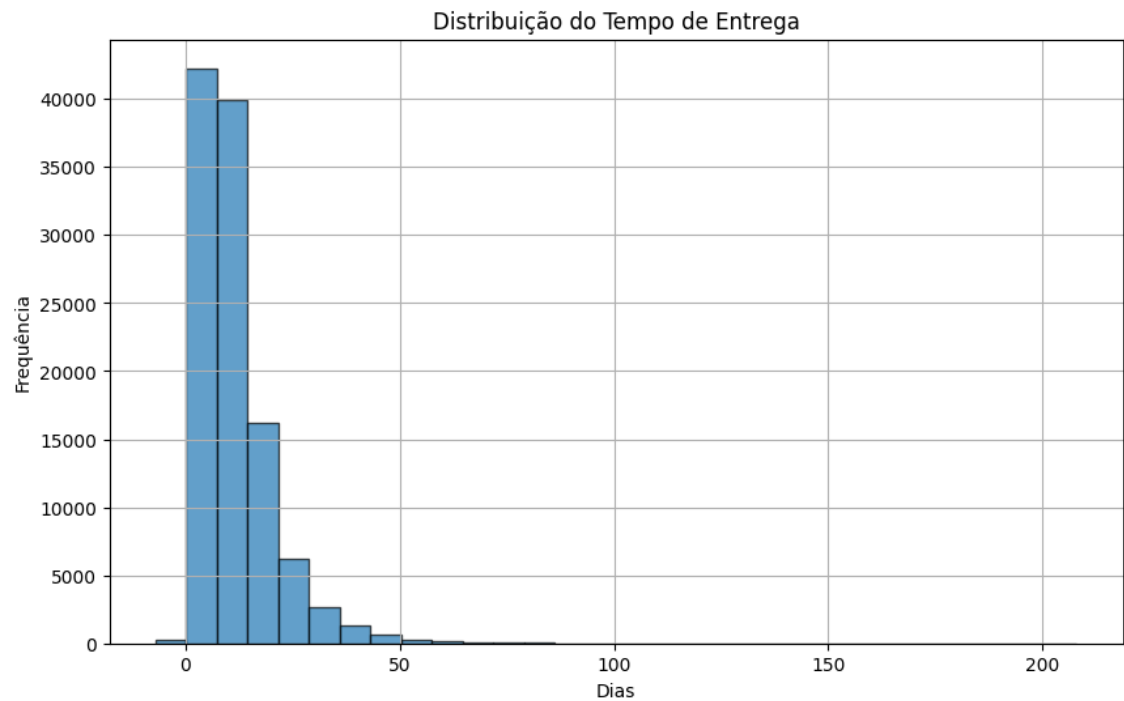


Figura 4 – Gráfico sobre a distribuição do tempo de entrega

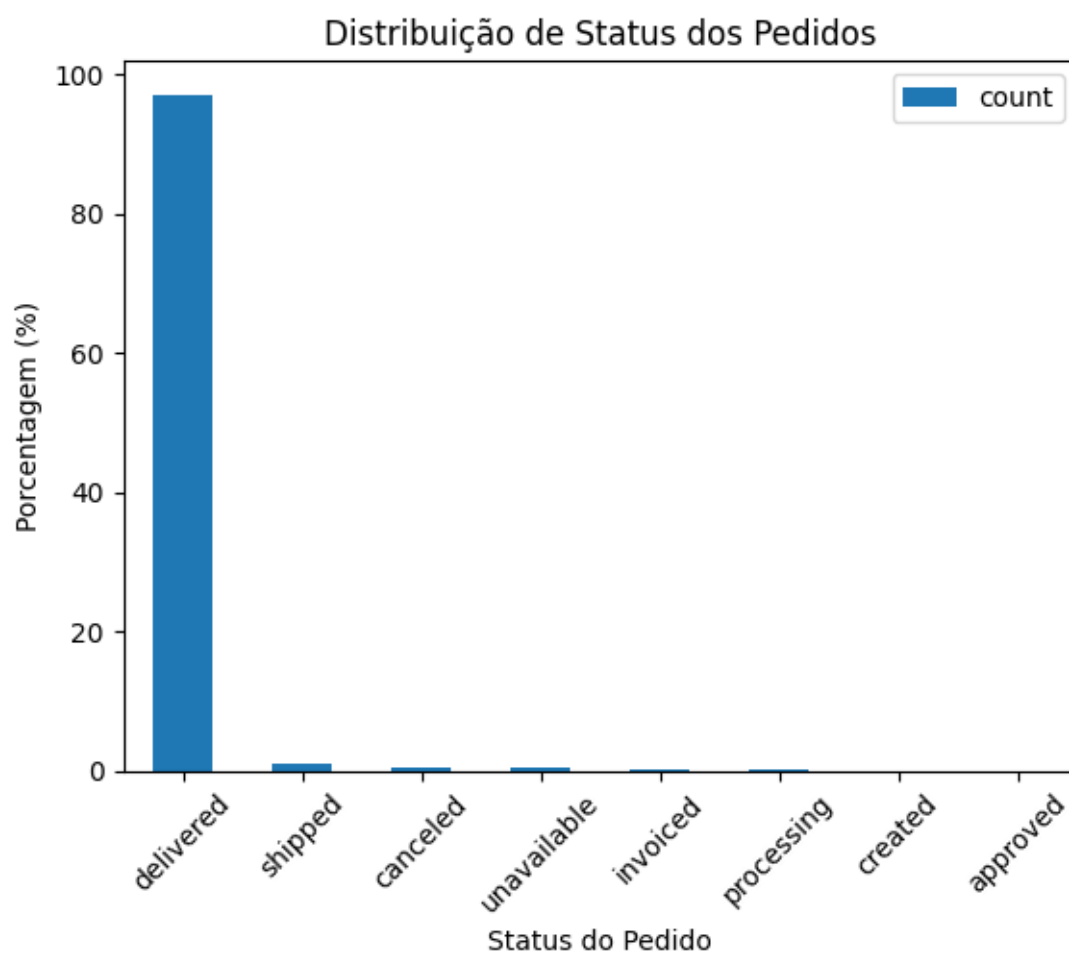


Figura 5 – Gráfico sobre os satus dos pedidos

### 2.3 – Análise da satisfação do cliente

Nesta parte da análise, foram utilizadas as tabelas referentes às vendas, revisões e produtos. Como resultados para este ponto de estudo, foi observado que a maioria dos pedidos tiveram avaliação máxima, o que é muito importante para as empresas fornecedoras. Como categoria para esta avaliação, podemos destacar: perfumaria, cama mesa e banho, acessórios de informática, brinquedos, e outros. Com as piores avaliações, com notas entre 0 e 1, temos as categorias: pet shop, beleza, móveis e decoração, entre outros. Os gráficos sobre essas informações estão logo a seguir.

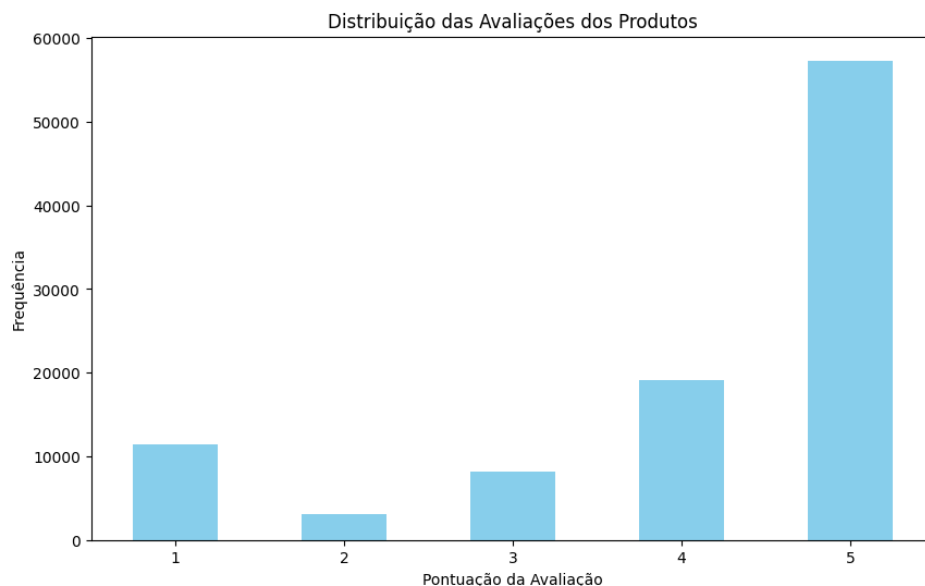


Figura 6 – Gráfico da avaliação dos produtos dos clientes

	product_id	product_category_name	review_score
0	00066f42aeeb9f3007548bb9d3f33c38	perfumaria	5.0
3	000b8f95fcb9e0096488278317764d19	utilidades_domesticas	5.0
4	000d9be29b5207b54e86aa1b1ac54872	relogios_presentes	5.0
6	00126f27c813603687e6ce486d909d01	cool_stuff	5.0
8	001b237c0e9bb435f2e54071129237e9	cama_mesa_banho	5.0
10	001c5d71ac6ad696d22315953758fa04	cama_mesa_banho	5.0
13	0021a87d4997a48b6cef1665602be0f5	informatica_acessorios	5.0
16	002959d7a0b0990fe2d69988affcbc80	moveis_decoracao	5.0
17	002af88741ba70c7b5cf4e4a0ad7ef85	cool_stuff	5.0
18	002c6dab60557c48cfd6c222ef7fd76	brinquedos	5.0

Figura 7 – Tabela das melhores avaliações

	product_id	product_category_name	review_score
2	0009406fd7479715e4bef61dd91f2462	cama_mesa_banho	1.0
5	0011c512eb256aa0dbbb544d8dfcf6e	automotivo	1.0
19	002d4ea7c04739c130bb74d7e7cd1694	pet_shop	1.0
32	0043c62d00db47eff6a6bc4cf6bfaeda	esporte_lazer	1.0
33	0043d1a25ef08fb6f41b8fa6f91742ab	construcao_ferramentas_construcao	1.0
40	0060b415594c5e1200324ef1a18493c4	esporte_lazer	1.0
86	00cf06147ed4880ec5fbba2adbb20e1d	livros_interesse_geral	1.0
104	00e62bcf9337ca4c5d5b4c5c8188f8d2	beleza_saude	1.0
106	00efb9a8ce0cf3b2f37892ab003edc10	esporte_lazer	1.0
118	01084e8138d03dc691daa19aa6c12471	moveis_decoracao	1.0

Figura 8 - Tabela das piores avaliações

De acordo com os dados fornecidos, os produtos especificamente não são mostrados pois não existe uma tabela que descreva o produto, e sim apenas a categoria do mesmo. Porém, já podemos ter uma visão bem ampla sobre os produtos com melhores e piores avaliações apenas avaliando a categoria.

## 2.4 – Análise financeira

Serão mostrados neste tópico referente aos resultados, informações sobre o frete, entrega e satisfação do cliente. Foi perceptível que quanto menor for o custo de frete do produto e menor o tempo de entrega, a satisfação do cliente é maior. A seguir estão os gráficos dos resultados para uma melhor visualização.

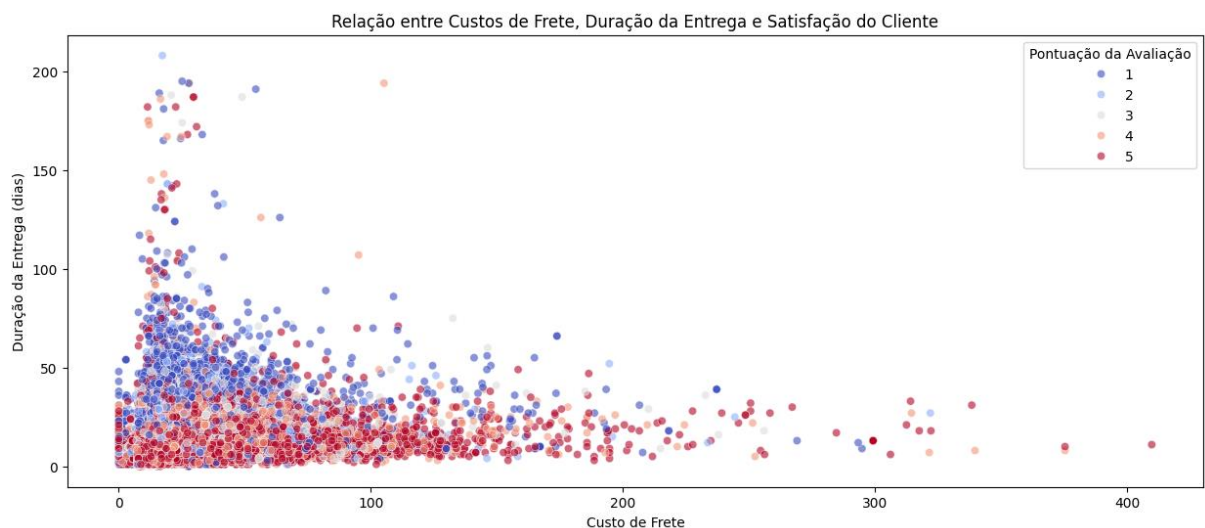


Figura 8 – Gráfico sobre relação de custo de frete e prazo de entrega

O gráfico mostra essa relação de custo de frete e entrega baseado na satisfação do cliente. É possível notar que quanto mais próximo a zero maiores são as movimentações de cores, que são representadas pela satisfação do cliente. A cor vermelha se encontra mais na região do zero devido ao custo de frete baixo e o prazo de entrega também baixo, resultado em uma melhor satisfação. Quanto mais a movimentação é afastada, pior fica a satisfação do cliente diante à compra.

## 2.5 – Análise de marketing

Nesta última seção dos resultados será abordado sobre a análise de marketing. Foi uma seção mais complicada devido a falta de informações suficientes para tratar sobre as fontes de conversão de vendas pois não consta na base de dados fornecida nenhuma coluna ou tabela que tenha esses dados específicos, mas, com base nos dados de pagamentos foram criadas supostas explicações sobre as conversões de vendas.

Abaixo está o grafico de quantidade de vendas e quais os tipos.

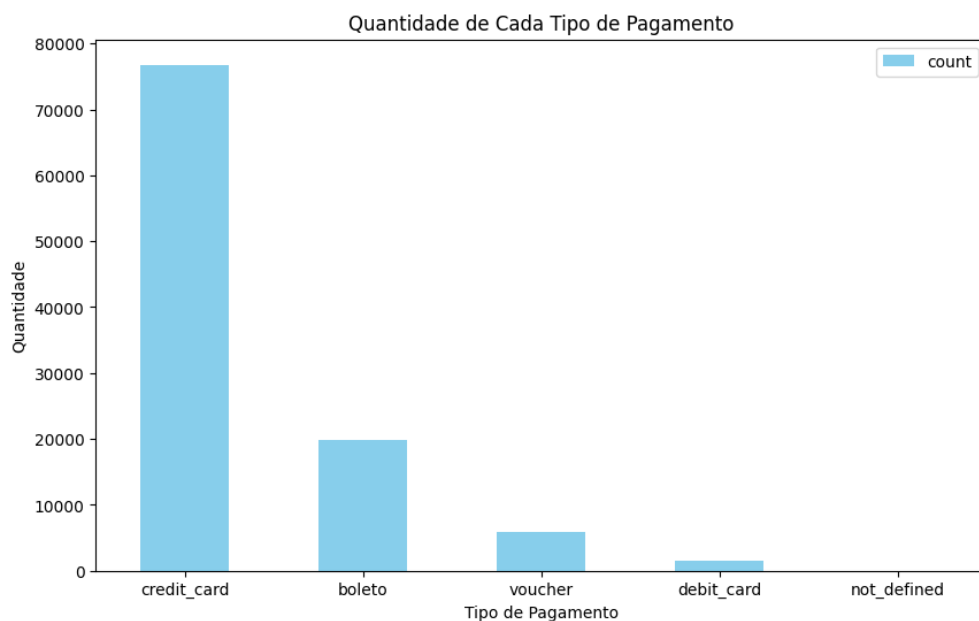


Figura 9 – Gráfico sobre os tipos de pagamentos

De acordo com os dados apresentados sobre os tipos de pagamentos, podemos dizer que as conversões de vendas foram principalmente por meio de tráfego pago devido ao alto uso do cartão de crédito. Por se tratar de dados sobre o e-commerce, as vendas ocorreram pela internet, e podemos destacar as redes sociais pagas que patrocinam suas publicações para maior visibilidade, sites e aplicativos de vendas.



### **3. Conclusão**

O presente trabalho buscou analisar os dados referentes ao e-commerce brasileiro, abordando sobre os tópicos de vendas, logística, satisfação do cliente, financeiro e marketing. Como resultados foi observado que os maiores números de vendas vieram da categoria de produtos cama, mesa e banho. A maioria dos produtos tiveram satisfação máxima pelos clientes, com mais de 80% de avaliações. A análise financeira mostrou que as taxas de frete juntamente com o prazo de entrega influenciam diretamente na satisfação do cliente.

Essa análise embora simplificada, ressalta a importância da análise dos dados referentes ao e-commerce brasileiro e pode ser utilizada como ferramenta de melhoria para as empresas, extraindo as informações apresentadas e modelando novas táticas e estratégias competitivas.