

# Relatório

## Análise de dados DataLab Amazon Sales

**Preparado por:**

Aline Dionizio | Giulia Braga | Taiza Ferreira

# 1. Introdução

## Objetivo:

Analisar dados de produtos e avaliações da Amazon para identificar padrões, relações e insights, orientando decisões estratégicas de precificação, posicionamento e avaliação de percepção dos clientes.

O objetivo secundário do projeto foi explorar o uso da IA para otimizar o processo de análise.

## 2. Metodologia e Ferramentas

Este relatório documenta as principais etapas e achados do processo de análise de um dataset contendo informações sobre produtos e avaliações da plataforma Amazon.

- Carregamento dos datasets (amazon\_product.csv, amazon\_review.csv) e união via product\_id.
- Limpeza de nulos e duplicados.
- Conversão de tipos e criação das variáveis:
  - diferenca\_preco (economia absoluta).
  - categoria\_principal (nível superior da hierarquia).
- Preparação do unificada\_df para análise estatística e exploratória.



### 3. Análise exploratória

#### 3.1 Medidas de Tendência Central e Dispersão

Variáveis	Média	Mediana	Moda	DP	IQR
discounted_price	3.294,57	899,00	2,99	7.157,17	1827.50
actual_price	5.687,36	1.795,00	999,00	11.191,95	3738.50
discount_percentage	46,84	49,00	50	21,65	31.75
diferenca_preco	2.392,79	803,50	0,0	4.727,66	1620.75
rating	4,09	4,10	4,1	0,29	0.40
rating_count	17805,42	4863,00	9378	42.161,54	15370.00

Há alta variabilidade em preços e contagem de avaliações, com distribuições assimétricas puxadas por produtos premium. Ratings possuem distribuição estável, sugerindo homogeneidade na percepção do consumidor.

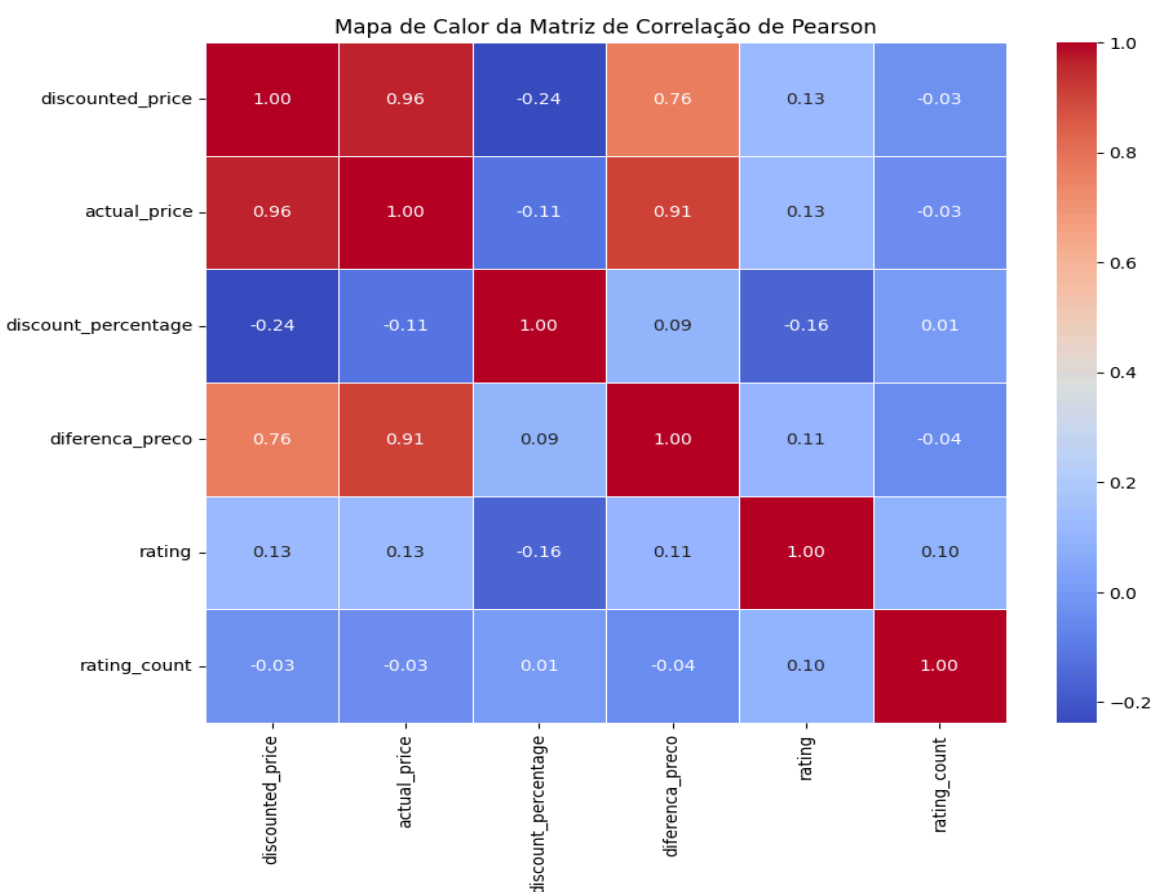
#### 3.2 Distribuição de Produtos por Categoria

Categoria	%
Electronics	36,5%
Home&Kitchen	32,9%
Computers&Accessories	27,7%
Outras	2,9%

Três categorias concentram 97% dos produtos, indicando foco estratégico para análises de preço, desconto e percepção de valor nestas categorias.

### 3.3 Correlação entre Variáveis

- **Forte correlação positiva:** actual\_price, discounted\_price, diferenca\_preco (Pearson >0,91).
- **Correlação negativa moderada:** discount\_percentage vs preços.
- **Fraca correlação:** Ratings e número de avaliações com preço ou desconto.

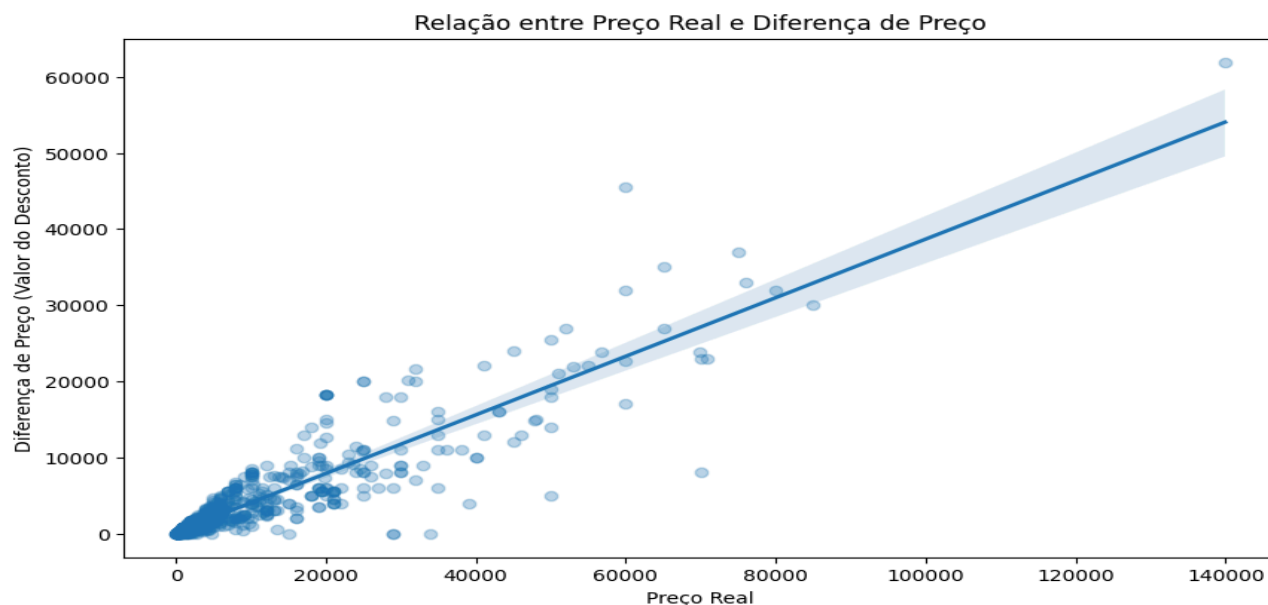


O valor absoluto do desconto depende fortemente do preço real. Ratings e contagem de avaliações não estão relacionados ao preço ou desconto de forma significativa, indicando que qualidade percebida e valor agregado podem ser mais relevantes que estratégias de desconto.

## 4. Validação de Hipóteses

Hipótese	Resultado
1- Quanto maior o desconto, melhor será a pontuação?	<b>Refutada</b> (correlação fraca negativa).
2- Quanto maior o número de pessoas que avaliaram o produto, melhor será a classificação?	<b>Confirmação parcial</b> (correlação fraca positiva).
3- Produtos com preços reais mais altos (sem desconto aplicado) tendem a ter uma avaliação maior?	<b>Refutada</b> (sem relação clara).
4- Produtos com um preço real mais alto tendem a ter descontos absolutos maiores?	<b>Confirmada</b> (correlação forte).

### Visualização Hipótese 4 Confirmada:



O preço real impacta o valor absoluto do desconto, mas não influencia diretamente a avaliação do cliente, reforçando que estratégias de desconto precisam ser analisadas em conjunto com percepção de valor e categoria do produto.

## 5. Análise de Risco Relativo por Categoria

Comparação	RR	Significância
Electronics vs Home&Kitchen	1,09	p=0,0427
Electronics vs Computers&Accessories	0,90	p=0,0050
Home&Kitchen vs Computers&Accessories	0,82	p=0,0000
Electronics vs OfficeProducts	0,74	p=0,0023

Electronics possui 9% mais chances de alta avaliação em relação a Home&Kitchen, mas tem desempenho inferior frente a Computers&Accessories e OfficeProducts, indicando a importância da categoria na percepção de valor pelo cliente.

## 6. Conclusões

A análise evidencia que a categoria do produto influencia significativamente as avaliações, com Computers&Accessories e OfficeProducts apresentando notas mais altas e desempenho superior em comparação a Electronics e Home&Kitchen.

Esses resultados ressaltam a necessidade de considerar a categoria ao interpretar métricas de avaliação e indicam potenciais vieses ao comparar segmentos distintos, recomendando-se ampliar a análise para outras categorias e realizar testes post-hoc para aprofundar a investigação.

- O preço real está fortemente relacionado ao desconto absoluto.
- Descontos não melhoram significativamente as avaliações.
- O número de avaliações impacta as notas de forma marginal.

## 7. Recomendações Estratégicas

Com base nos resultados, recomenda-se:

- **Não depender apenas de descontos** para influenciar avaliações. A percepção de qualidade e valor é mais determinante.
- **Melhorar o posicionamento e apresentação de produtos** em categorias com menor avaliação média, como *Electronics* e *Home&Kitchen*.
- **Estudar boas práticas das categorias mais bem avaliadas**, como *Computers&Accessories* e *OfficeProducts*, para replicar estratégias bem-sucedidas.
- **Personalizar ações por categoria**, adotando campanhas específicas conforme o comportamento dos consumidores.
- **Implementar testes A/B e monitoramento contínuo** de avaliações por categoria, permitindo ajustes estratégicos baseados em dados reais.