

# DataLAB para amazon Sales

Análise realizada por:

- Aline Dionizio
- Giulia Braga
- Taiza Ferreira



# Objetivo

A DataLAB, consultoria especializada em análise de dados, foi contratada para o Projeto Amazon Sales, com o objetivo de preparar e analisar dados de produtos e avaliações da Amazon, buscando insights sobre a relação entre categorias de produtos e as classificações dos usuários.

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

## Perguntas de Negócio

- Quanto maior o desconto, melhor será a pontuação.
- Quanto maior o número de pessoas que avaliaram o produto, melhor será a classificação.
- Produtos com preços reais mais altos (sem desconto aplicado) tendem a ter uma avaliação maior?
- Produtos com um preço real mais alto tendem a ter descontos absolutos maiores?



# CALCULO DA PROPORÇÃO (%) DE CADA CATEGORIA PRINCIPAL NO DATASET



**36.52%** → **32.92%** → **27.76%** → **2.28%**

ELECTRONICS                    HOME&KITCHEN                    COMPUTERS&ACCESSORIES                    OFFICEPRODUCTS



# Dados Gerais Observados



## Representatividade:

Electronics e Home&Kitchen lideram em quantidade de produtos, seguidos por OfficeProducts e Computers&Accessories.



## Preços e Descontos:

Electronics e Home&Kitchen têm os maiores descontos em reais.



## Percentual de Desconto:

O percentual médio de desconto varia entre categorias, refletindo estratégias de precificação.



## Avaliação Média (Rating):

A média de avaliações por produto varia entre categorias, refletindo popularidade ou volume de vendas.



## Contagem de Avaliações:

A média de avaliações difere entre categorias, indicando popularidade e volume de vendas.



# Técnicas e Testes



## Visualização e Estatísticas Descritivas:

Uso de boxplots, histogramas, mediana e desvio padrão para entender distribuições, dispersão e identificar outliers.

## Análise de Correlação:

Correlação de Pearson e Spearman para investigar relações entre variáveis numéricas.

## ANOVA e Tukey HSD:

Análise de diferenças de preços entre categorias e identificação de quais grupos diferem.

## Detecção de Outliers:

Aplicação de IQR e Z-Score para identificar valores atípicos relevantes à análise.

## Análise de Variáveis Categóricas:

Criação de variáveis com contagem de frequência e distribuição de produtos por categoria.

## Risco Relativo e Qui-quadrado:

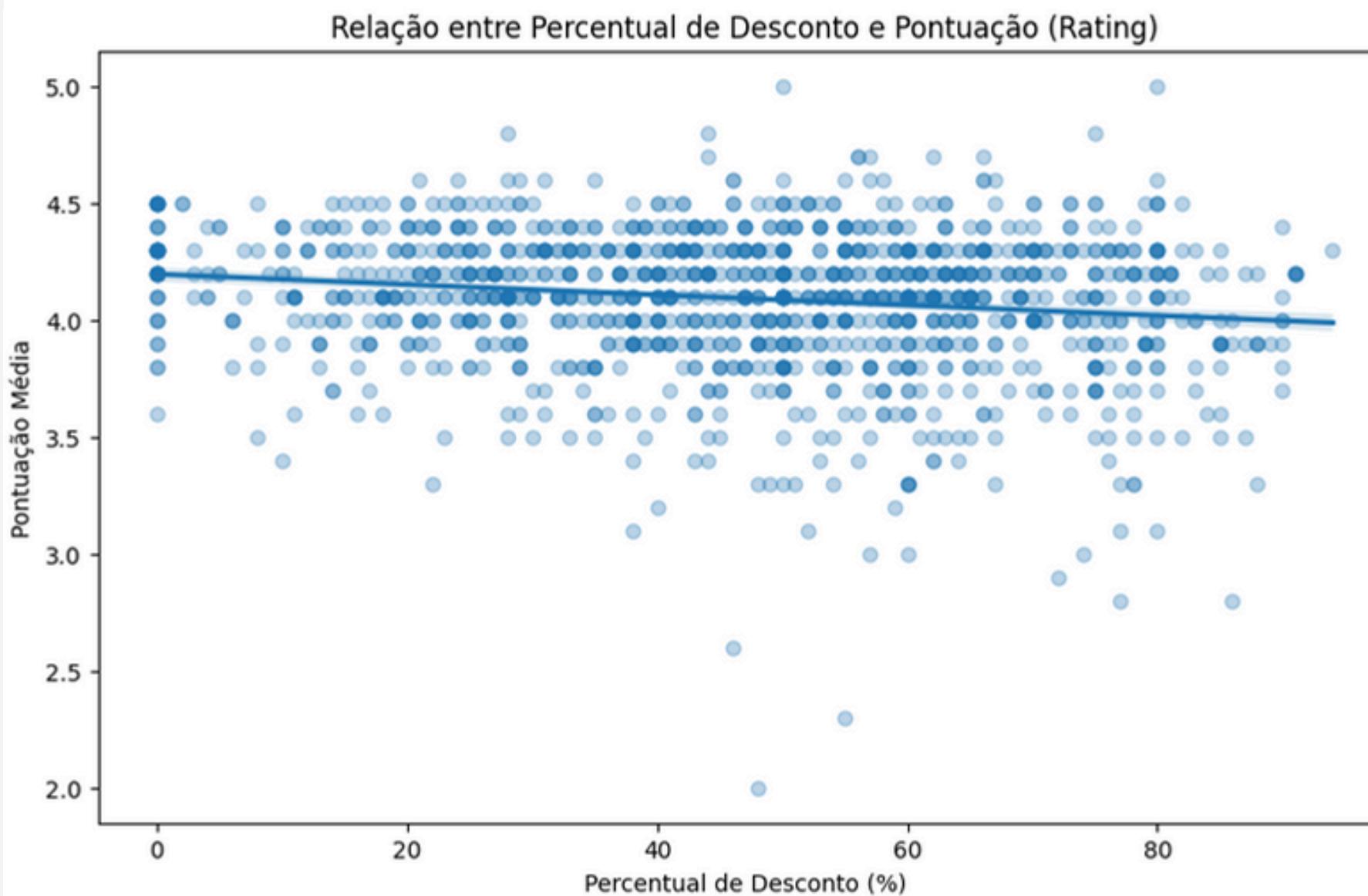
Avaliação entre categorias de produtos e probabilidade de alta avaliação.



## Hipótese 1: Desconto vs. Classificação

Correlação de Spearman entre Desconto (%) e Rating:  
-0.1512, P-valor: 0.0000.

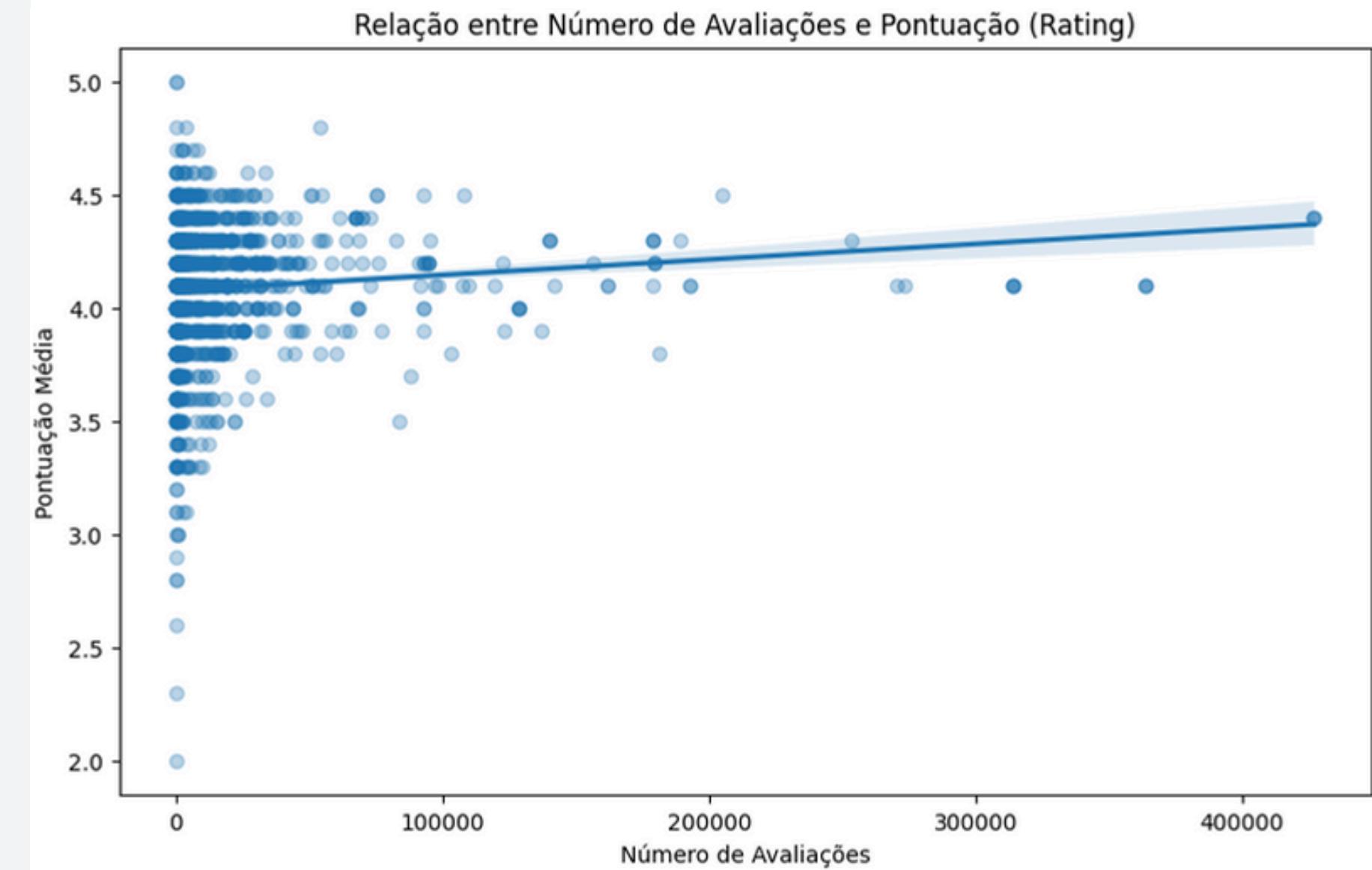
Embora haja significância estatística, a **relação é fraca**. O desconto não é um bom preditor de alta avaliação.



## Hipótese 2: Avaliações vs. Classificação

Correlação entre Desconto (%) e Rating: Pearson: 0.0982,  
P-valor: 0.0003 | Spearman: 0.1902, P-valor: 0.0000

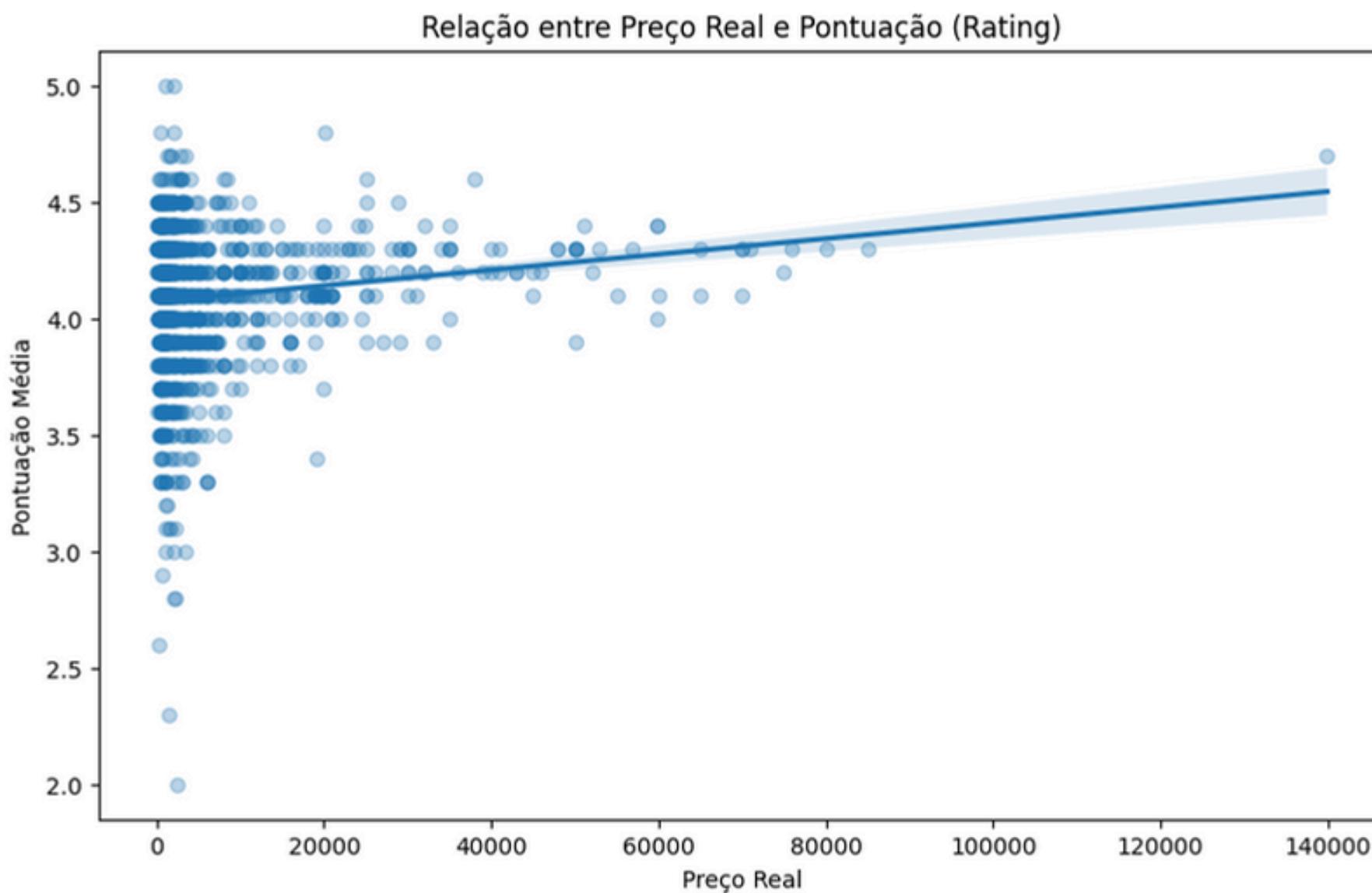
Produtos com mais avaliações tendem a ter notas maiores, mas a **relação é fraca** e não é determinante.



## Hipótese 3: Preço Real vs. Avaliação

Correlação entre Desconto (%) e Rating: Pearson: 0.1280, P-valor: 0.0000 | Spearman: 0.0325, P-valor: 0.2320.

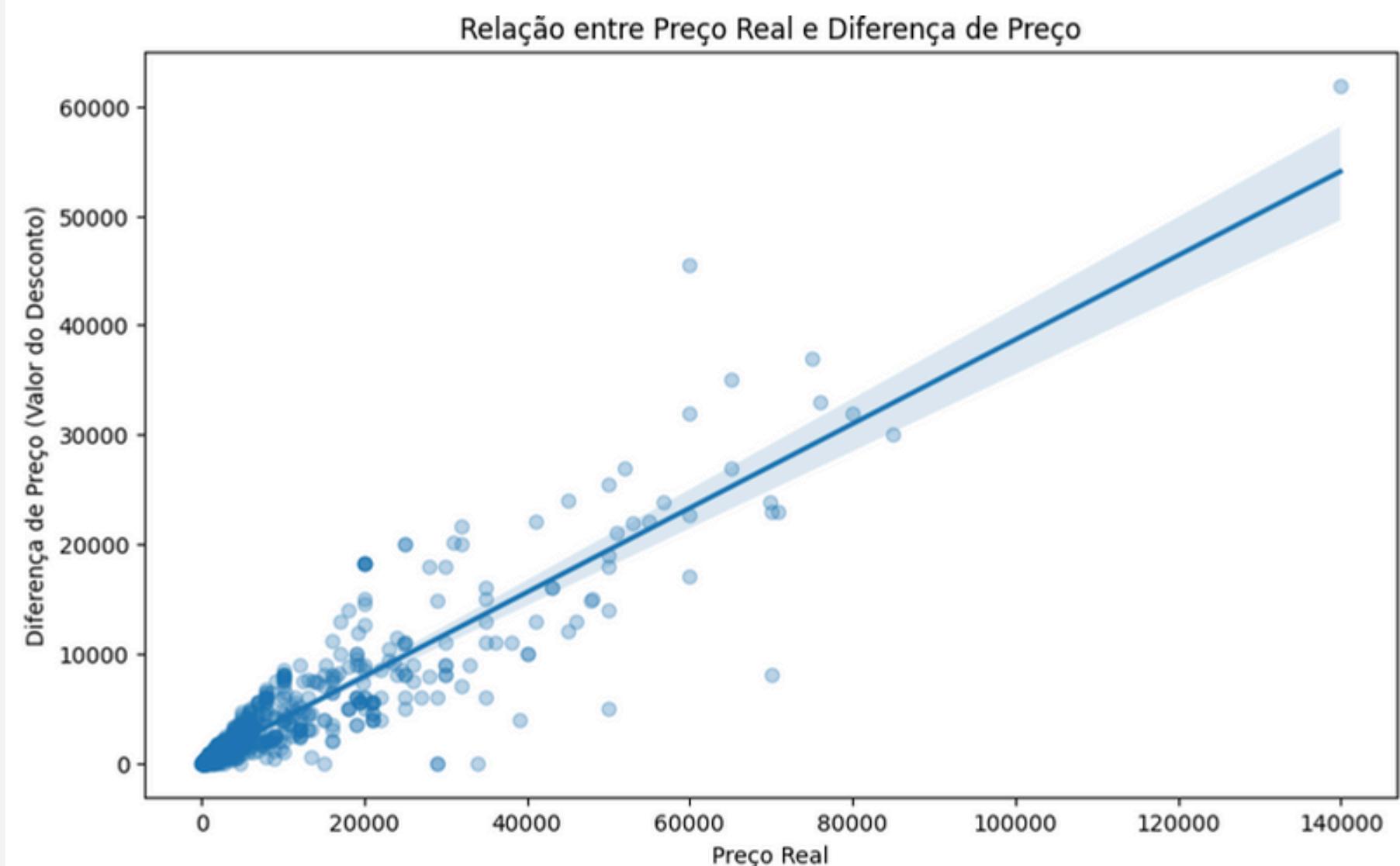
Correlação de Pearson indica **fraca relação positiva**, mas a de Spearman não confirmou, sugerindo inconsistência.



## Hipótese 4: Preço Real vs. Diferença de Preço

Correlação entre Desconto (%) e Rating: Pearson: 0.9108, P-valor: 0.0000 | Spearman: 0.8952, P-valor: 0.0000.

Produtos mais caros tendem a ter descontos absolutos maiores, com **forte relação positiva** confirmada.





## Hipótese 4 com Testes Complementares

A **Hipótese 4 foi confirmada**. No entanto, essa relação varia entre as categorias de produto.

Electronics e Home & Kitchen tendem a oferecer os maiores descontos absolutos médios, enquanto, como Car & Motorbike, oferecem menos.

### ANOVA: ANÁLISE POR CATEGORIA PRINCIPAL

O Resultado do ANOVA para a média da diferença\_preco entre as categorias principais foi: p-valor = 4,49e-24

→ Rejeita-se a hipótese nula, indicando que há diferenças estatisticamente significativas.

### TUKEY HSD: COMPARAÇÃO ENTRE CATEGORIAS

Computers & Accessories oferece, em média, ₹753,27 a mais de desconto do que Car & Motorbike ( $p\text{-adj} < 0,0001$ ).

Outros pares, como Electronics vs Home & Kitchen, não apresentaram significância estatística ( $p\text{-adj} = 1.0000$ ).



# Risco Relativo

## Cálculo Risco Relativo:

Produtos da categoria Electronics (RR = 1,0916) têm 1,09x mais chance de receber alta avaliação do que Home&Kitchen, indicando um aumento de 9% na probabilidade de boa avaliação.

## Teste Qui-quadrado:

O teste confirmou que a associação entre a categoria do produto e a chance de alta avaliação é significativa (Qui-quadrado = 4,11, p = 0,0427).

Esse resultado reforça a análise anterior, indicando que a diferença observada não ocorreu por acaso.



# Resultados e Insights

Conclui-se que a categoria do produto impacta mais nas avaliações, destacando Computers&Accessories e OfficeProducts com melhor desempenho em notas altas.

- Produtos com preços mais altos recebem maiores descontos monetários. As hipóteses 1, 2 e 3 não foram sustentadas pelos dados.
- A Hipótese 1 foi refutada por correlação fraca e negativa. A Hipótese 2 indicou relação positiva fraca. A Hipótese 3 foi descartada por ausência de relação clara.
- No Risco Relativo, Electronics têm mais chance de alta nota que Home&Kitchen, mas menor chance que Computers&Accessories e OfficeProducts. Teste Qui-quadrado confirmou as associações como estatisticamente significativas.



A categoria do produto impacta mais do que preço ou desconto.



# 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.



## PERSONALIZAR AÇÕES POR CATEGORIA,

adotando campanhas específicas conforme o comportamento dos consumidores.



# Responsáveis pelo Projeto:



## AlineDion - Overview

Análise de Dados | Engenharia de Dados | SQL | ETL  
| Python | Power BI - AlineDion

GitHub

[linkedin.com/in/aline-dionizio/](https://linkedin.com/in/aline-dionizio/)



## TaizaFerreira - Overview

BigQuery | Power BI | Google Sheets | Google Colab |  
Looker Studio - TaizaFerreira

GitHub

[linkedin.com/in/taiza-ferreira-dados/](https://linkedin.com/in/taiza-ferreira-dados/)



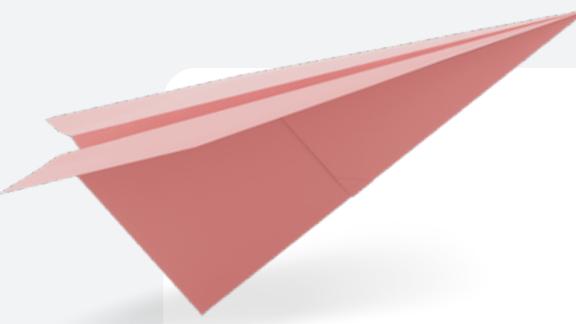
## BragaGiu - Overview

Estudante de Análise e Desenvolvimento de  
Sistemas - BragaGiu

GitHub

[linkedin.com/in/bragagiu/](https://linkedin.com/in/bragagiu/)





# Obrigada!

