

GRÁFICO DE DISPERSÃO COM LINHA DE TENDÊNCIA

O cálculo realizado tem como objetivo analisar a relação entre o **Valor da Compra (R\$)** e o **Valor do Cupom (R\$)**, verificando se há um padrão que indique que compras maiores estão associadas a cupons de maior valor. Para isso, foi aplicada uma **análise descritiva** seguida de um **ajuste de modelo de regressão linear** para avaliar a força e direção dessa relação.

Para isso aplicamos um modelo de **Regressão Linear Simples**, utilizando a variável **Valor da Compra** como variável independente (X) e o **Valor do Cupom** como variável dependente (Y).

O algoritmo responsável por fazer esse cálculo foi feito em python no ambiente do google colab:

```
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import pearsonr

file_name = 'PicMoney-Massa_de_Teste_com_Lojas_e_Valores-10000 linhas (1).csv'

df = pd.read_csv(file_name, sep=';')

df['valor_compra'] = pd.to_numeric(df['valor_compra'], errors='coerce')
df['valor_cupom'] = pd.to_numeric(df['valor_cupom'], errors='coerce')
df_analise = df[['valor_compra', 'valor_cupom']].dropna()

r, _ = pearsonr(df_analise['valor_compra'], df_analise['valor_cupom'])
r_squared = r**2

print(f"Coefficiente de Correlação de Pearson (R): {r:.4f}")
print(f"Coefficiente de Determinação (R²): {r_squared:.4f}")

plt.figure(figsize=(10, 6))
sns.regplot(data=df_analise, x='valor_compra', y='valor_cupom',
            scatter_kws={'alpha': 0.3, 's': 15},
            line_kws={'color': 'red', 'linestyle': '--'})

plt.title('Dispersão entre Valor da Compra e Valor do Cupom', fontsize=16)
plt.xlabel('Valor da Compra (R$)', fontsize=12)
plt.ylabel('Valor do Cupom (R$)', fontsize=12)
plt.grid(True, linestyle='--', alpha=0.6)
plt.text(0.05, 0.95, f'R = {r:.4f}\nR² = {r_squared:.4f}',
        transform=plt.gca().transAxes, fontsize=12,
        verticalalignment='top', bbox=dict(boxstyle='round,pad=0.5', fc='wheat', alpha=0.5))

plt.show()
```

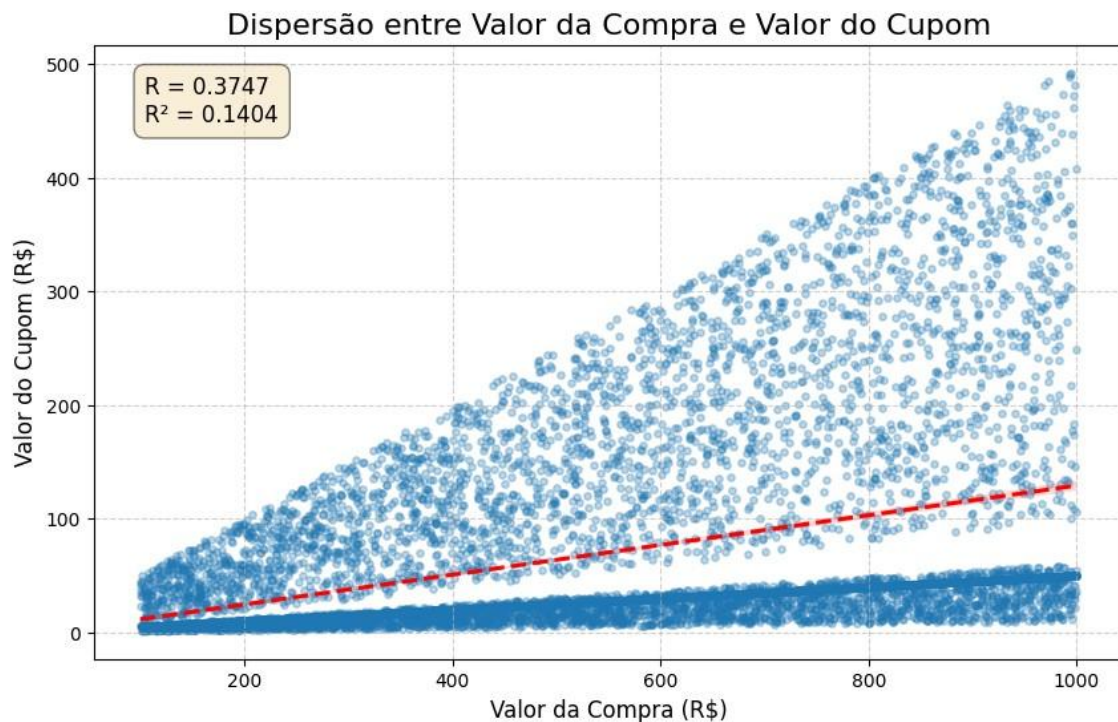
Coeficientes de Correlação e Determinação

O coeficiente de correlação de Pearson obtido foi **R = 0,3747**, indicando uma correlação positiva, porém fraca, entre as variáveis.

O coeficiente de determinação obtido foi **R² = 0,1404**, mostrando que cerca de **14% da variação no valor do cupom pode ser explicada pelo valor da compra**.

Gráfico de Dispersão

A imagem apresenta o gráfico de dispersão com a linha de regressão ajustada:



Interpretação dos Resultados

O coeficiente angular (β_1) indica que **para cada aumento de R\$1,00 no valor da compra, o valor do cupom tende a aumentar em R\$ 0,13, em média.** O intercepto (β_0) de -1.576 representa o valor esperado do cupom quando o valor da compra é zero.

A análise do R^2 sugere que **o modelo explica apenas 14% da variação do valor do cupom**, o que indica que existem outros fatores que influenciam esse valor além do valor da compra.

Conclusão

Com base nos resultados obtidos, conclui-se que:

- Há uma correlação positiva, mas fraca, entre o **Valor da Compra** e o **Valor do Cupom**.
- Apenas 14% da variabilidade do **Valor do Cupom** é explicada pelo **Valor da Compra**.