

# ✓ Análise de Correlação entre Variáveis

## 1. Cálculo dos coeficientes de correlação

- Foi calculada a **matriz de correlação** entre todas as variáveis numéricas do `df_vendas` utilizando o **coeficiente de Spearman**.
- O Spearman é indicado por ser não-paramétrico, ou seja, adequado para identificar **associações monotônicas** mesmo que não sejam lineares.

```
corr_spearman = df_analise.corr(method='spearman')
```

---

## 2. Visualização da matriz de correlação

### ✓ Heatmap estático:

- Usado `seaborn.heatmap` com escala de cores de -1 a +1.
- Mostra claramente quais variáveis se associam fortemente (positiva ou negativamente).

```
plt.figure(figsize=(10, 8))
sns.heatmap(corr_spearman, annot=True, cmap='YlGnBu', vmin=-1, vmax=1)
plt.title('Matriz de Correlação - Spearman')
plt.show()
```

### ✓ Heatmap interativo:

- Usado `plotly.express.imshow` para interação dinâmica, facilitando o zoom e a seleção de pontos de interesse.

```
fig = px.imshow(
    corr_spearman,
    text_auto=True,
    color_continuous_scale='RdBu_r',
    title='Matriz de Correlação Interativa - Spearman'
)
fig.show()
```

### Resultado:

- A matriz foi visualizada de forma clara e interpretável, conforme o critério de aceitação.
-

### 3. Identificação das correlações mais relevantes

#### ✓ Correlações fortes ( $|\text{correlação}| \geq 0.7$ ):

```
corr_fortes = corr_pairs[corr_pairs['abs_corr'] >=
0.7].sort_values(by='abs_corr', ascending=False)
print("\n☑ Pares com correlação forte ( $>|0.7|$ ):\n", corr_fortes)
```

- Estes pares indicam **relações significativas** que podem ter implicações importantes na análise do negócio.
- Exemplo de interpretação (ajuste conforme os dados reais):
  - quantidade\_vendida vs total\_venda: **Correlação forte e positiva**, o que é esperado, pois quanto maior a quantidade, maior o total vendido.
  - desconto vs lucro: **Correlação negativa**, indicando que aumentos no desconto podem estar associados a redução no lucro.

#### ✓ Correlações fracas ( $|\text{correlação}| \leq 0.3$ ):

```
corr_fracas = corr_pairs[corr_pairs['abs_corr'] <=
0.3].sort_values(by='abs_corr')
print("\n☒ Pares com correlação fraca ( $\leq|0.3|$ ):\n", corr_fracas)
```

- Estes pares mostram **associações muito fracas ou inexistentes**, sugerindo que as variáveis não influenciam diretamente uma à outra.

---

### 4. Interpretação das correlações

#### ✓ Correlações esperadas:

- Variáveis como quantidade\_vendida e total\_venda devem apresentar **correlação positiva forte** — quanto mais se vende, maior o total.

#### ✓ Correlações inesperadas:

- Caso variáveis como desconto e quantidade\_vendida apresentem correlação **baixa ou negativa**, isso pode indicar que as promoções não impulsionam significativamente as vendas.


#### ✓ Ausência de correlação:

- Pares com  $|\text{correlação}|$  abaixo de 0.3 indicam que não há relação linear ou monotônica aparente, como por exemplo entre custo\_frete e quantidade\_vendida (hipótese a ser confirmada nos dados).
-

## 5. Levantamento de hipóteses





Com base nas correlações identificadas, podemos levantar as seguintes hipóteses:

Correlação Observada	Hipótese
Correlação forte positiva entre <code>quantidade_vendida</code> e <code>total_venda</code>	O volume de vendas impacta diretamente o faturamento.
Correlação negativa entre <code>desconto</code> e <code>lucro</code>	Altos descontos podem estar prejudicando a margem de lucro.
Correlação fraca entre <code>custo_frete</code> e <code>quantidade_vendida</code>	O frete pode não ser um fator decisivo no volume de vendas.
Correlação nula entre <code>categoria_id</code> e variáveis financeiras	A categoria do produto pode não influenciar diretamente valores financeiros.

 Estas hipóteses podem ser testadas em análises mais profundas, como regressão ou segmentação de clientes.

---

## 6. Critérios de aceitação — VERIFICAÇÃO

Critério	Status
Matriz de correlação calculada e visualizada com clareza	
Relações mais relevantes destacadas e explicadas	
Correlações fracas, fortes e inexistentes interpretadas corretamente	
Gráficos e observações documentadas para uso no relatório	

---

## Conclusão:

A análise de correlação foi realizada com sucesso, utilizando **Spearman** para identificar **associações monotônicas**.

Foram gerados gráficos claros, destacadas as relações mais relevantes, interpretadas corretamente e levantadas hipóteses úteis para futuras análises.