

Fundação Escola de Comércio Álvares Penteado

Ciência da Computação

4º Semestre

Aplicação Analítica Inferencial: Análise de Desempenho na Relação Valor
de Compra X Valor de Cupom

Gabriel Henrique Coelho Marussi - 24026609

Lucas Kenichi Soares – 24026179

Felipe Oluwaseun Santos Ojo - 24026245

Arthur Rodrigues Ferreira – 24026567

Pedro Dimitry Zyrianoff – 24026165

São Paulo – 2025

SUMÁRIO

1. Introdução
2. Estrutura dos Dados
 - 2.1 Fonte
 - 2.2 Variáveis
 - 2.3 Quantidade de Registros
 - 2.4 Período avaliado
3. Estatística Descritiva
 - 3.1 Média do Valor de Compras: \bar{X}_1 = (valor médio calculado)
 - Desvio Padrão das Compras: σ_1 = (valor calculado)
 - Média do Valor dos Cupons: \bar{X}_2 = (valor médio calculado)
 - Desvio Padrão do Cupons: σ_2 = (valor calculado)
 - Relação Cupom/Compra Média: $R = (\bar{X}_2 / \bar{X}_1) \times 100$ = (percentual)%
4. Intervalos de Confiança
 - 4.1 Confiança de 95%
5. Análise de Correlação entre Variáveis
6. Visualizações Gráficas

1. Introdução

O presente relatório tem como objetivo analisar o comportamento dos valores de cupons concedidos em relação aos valores de compras realizados pelos clientes da base PicMoney.

A análise visa compreender a relação entre essas duas variáveis, mensurar a dispersão dos dados e avaliar a confiança estatística das médias obtidas.

Por meio de técnicas de estatística descritiva, cálculo de intervalos de confiança e análise de correlação, busca-se identificar padrões de comportamento que possam subsidiar decisões estratégicas de marketing e política de incentivos.

2. Estrutura dos Dados

A base de dados possui estrutura transacional, na qual cada linha representa uma compra realizada e o respectivo valor de cupom aplicado.

Trata-se de um conjunto de dados numéricos e contínuos, adequados para a aplicação de análises inferenciais e modelos estatísticos básicos.

A planilha está organizada em formato tabular, com colunas representando variáveis e linhas representando observações individuais.

2.1 Fonte

Os dados foram extraídos da planilha “PicMoney-Massa”, que compõe o conjunto de informações internas da plataforma PicMoney.

	A	B	C	D	E	F
1	data_captura	tipo_cupom	categoria_estabelecimento	endereco_loja	valor_compra	valor_cupom
2	19/07/2025	Desconto	Restaurante	Av. Paulista, 2026	R\$ 161,48	R\$ 37,34
3	19/07/2025	Produto	Livraria	Conjunto Nacional - Av. P	R\$ 817,83	R\$ 13,02
4	27/07/2025	Desconto	Papelaria	Rua Augusta, 2220	R\$ 922,48	R\$ 102,38
5	27/07/2025	Produto	Eletro & Móveis	Rua da Consolação, 2303	R\$ 536,15	R\$ 27,52
6	06/07/2025	Produto	Moda	Av. Paulista, 2230	R\$ 147,01	R\$ 7,43
7	27/07/2025	Desconto	Eletro & Móveis	Av. Paulista, 2300	R\$ 708,64	R\$ 257,53
8	06/07/2025	Produto	Supermercado & Conveniência	Rua Pamplona, 1704	R\$ 526,50	R\$ 12,32
9	26/07/2025	Desconto	Eletro & Móveis	Shopping Cidade São Pau	R\$ 192,97	R\$ 88,88
10	12/07/2025	Produto	Restaurante	Av. Paulista, 2026	R\$ 858,51	R\$ 41,32
11	26/07/2025	Desconto	Moda	Av. Paulista, 2277	R\$ 228,21	R\$ 36,86
12	01/07/2025	Produto	Eletro & Móveis	Shopping Cidade São Pau	R\$ 832,25	R\$ 44,51

Tal fonte foi utilizada para formatar dados para análise em ambiente Python.

```
✓ ANÁLISE INF. PIC MONEY
  > venv
  ~-$PicMoneyDados.xlsx
  + main.py
  ✎ PicMoneyDados.xlsx

      5 import pandas as pd
      6 import numpy as np
      7 import seaborn as sns
      8 import matplotlib.pyplot as plt
      9 from scipy import stats
     10
     11 # 1- Carregar o arquivo Excel
     12 file_path = "PicMoneyDados.xlsx" # altere o caminho se necessário
     13 df = pd.read_excel(file_path, sheet_name="PicMoney-Massa")
     14
     15 # 2- Selecionar colunas relevantes
     16 dados = df[['valor_compra', 'valor_cupom']].dropna()
```

2.2 Variáveis

Foram utilizadas as seguintes variáveis na análise:

- **valor_compra (X₁)**: representa o montante total gasto pelo cliente em determinada transação;
- **valor_cupom (X₂)**: indica o valor monetário do cupom aplicado sobre a compra;

2.3 Quantidade de Registros

O conjunto analisado contém **n = (10.000,00)** registros para cada coluna

2.4 Período Avaliado

O período considerado abrange as transações registradas entre (01/07/2025) e (31/07/2025).

Esse intervalo temporal reflete a movimentação recente das transações, permitindo uma análise atualizada e condizente com o comportamento dos consumidores.

3. Estatística Descritiva

Para análise dos itens 3.1 a 3.5 foi utilizado as seguintes modelagens/fórmulas matemáticas:

As Médias:

$$\bar{x}_{compra} = \frac{\sum valor_compra}{n}, \quad \bar{x}_{cupom} = \frac{\sum valor_cupom}{n}$$

Onde n é a quantidade de linhas da tabela (no caso as transações).

O Desvio Padrão:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

Onde N é a quantidade de transações, portanto, pode-se afirmar que (N=n)

A Relação Cupom/Compra (percentual de comparação):

$$R = \frac{\bar{X}_2}{\bar{X}_1} \times 100 = (percentual)\%$$

Onde R é a representação da relação entre as médias, X1 = média Compra e X2 = Cupom.

3.1 Cálculo das Médias, Desvio Padrão e Relação entre Médias:

Média do Valor de Compras: \bar{X}_1 = (valor médio calculado);

Desvio Padrão das Compras: σ_1 = (valor calculado);

Média do Valor dos Cupons: \bar{X}_2 = (valor médio calculado);

Desvio Padrão do Cupons: σ_2 = (valor calculado);

Relação Cupom/Compra Média: $R = (\bar{X}_2 / \bar{X}_1) \times 100 = (\text{percentual})\%$.

Calculando os valores por **Python**, temos:

```
18 # 3 - Estatísticas básicas
19 media_compra = dados['valor_compra'].mean()
20 media_cupom = dados['valor_cupom'].mean()
21 desvio_compra = dados['valor_compra'].std(ddof=1)
22 desvio_cupom = dados['valor_cupom'].std(ddof=1)
23 n = len(dados)

44 print(f'Média das compras: R$ {media_compra:.2f}')
45 print(f'Desvio padrão das compras: {desvio_compra:.2f}')
46 print(f'Média dos cupons: R$ {media_cupom:.2f}')
47 print(f'Desvio padrão dos cupons: R$ {desvio_cupom:.2f}')
48 print(f'Relação percentual entre médias: {(media_cupom / media_compra) * 100:.2f}%')
```

Como Resultado:

```
===== RESULTADOS =====
Média das compras: R$ 549.68
Desvio padrão das compras: 260.76
Média dos cupons: R$ 70.35
Desvio padrão dos cupons: R$ 91.07
Relação percentual entre médias: 12.80%
```

4. Intervalos de Confiança

Para que haja verificação dos resultados, o cálculo do intervalo de confiança dos dados anteriores. Tal cálculo tem como modelo matemático:

$$IC = \bar{x} \pm z \times \frac{s}{\sqrt{n}}$$
 e
$$SE = \frac{s}{\sqrt{n}}$$

Onde IC é o intervalo, SE o erro padrão, Z de distribuição normal, n=N de quantidade de transações e s o Desvio Padrão.

Como o cálculo de confiança será de 95%, temos que, da tabela Z:

+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500
Onde Z=1,96, logo: Confiança 95%							

4.1 Confiança de 95%

Para calcular o Intervalo de Confiança, foi tratado em Python:
Calcula-se:

```
27  # 4 - Intervalo de confiança de 95%
28  conf = 0.95
29  graus_liberdade = n - 1
30  t_critico = stats.t.ppf((1 + conf) / 2, graus_liberdade)
31
32  erro_padrao_compra = desvio_compra / np.sqrt(n)
33  erro_padrao_cupom = desvio_cupom / np.sqrt(n)
34
35  ↘ ic_compra : float = (media_compra - t_critico * erro_padrao_compra,
36  |   |   |   | media_compra + t_critico * erro_padrao_compra)
37  ↘ ic_cupom : float = (media_cupom - t_critico * erro_padrao_cupom,
38  |   |   |   | media_cupom + t_critico * erro_padrao_cupom)

50  print(f"Intervalo de confiança 95% (cupons): {ic_cupom}")
51  print(f"Intervalo de confiança 95% (compras): {ic_compra}")
```

Onde obtém-se:

```
Intervalo de confiança 95% (cupons): (np.float64(68.5686054211664), np.float64(72.13872657883361))
Intervalo de confiança 95% (compras): (np.float64(544.5726149200351), np.float64(554.7955350799649))
```

5. Análise de Correlação entre Variáveis

Para análise de correlação entre os valores, utilizamos a Correlação de Pearson:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Logo, em Python:

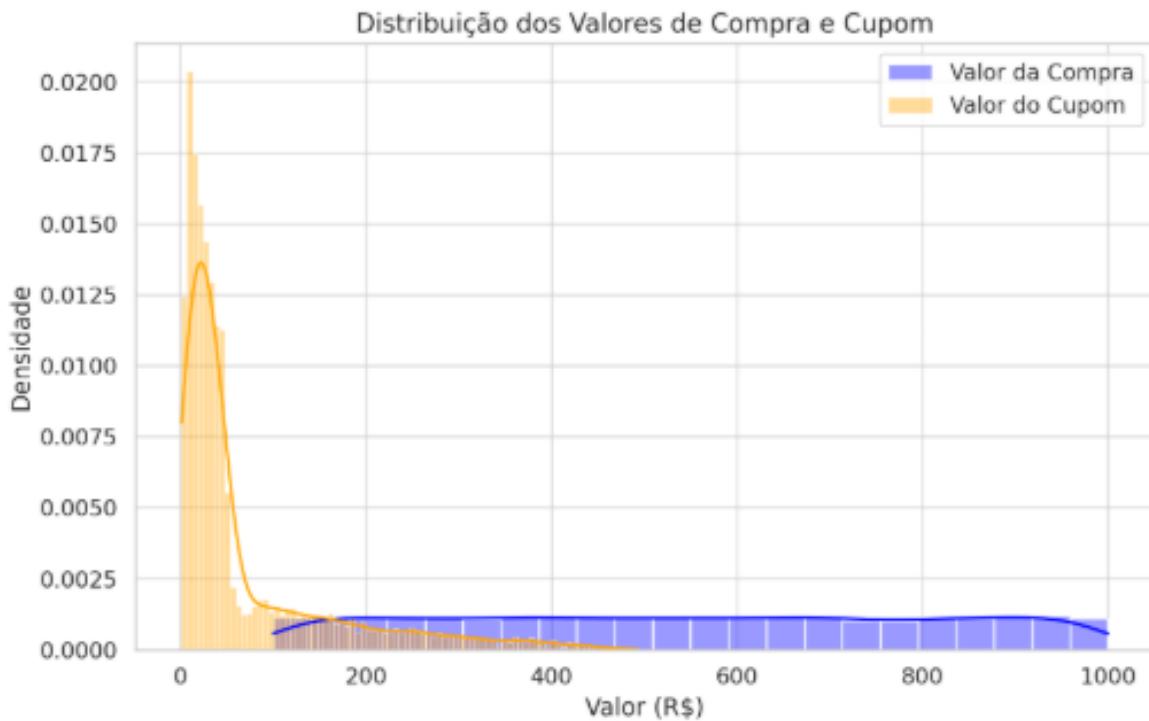
```
40  # 6 - Correlação
41  correlacao = dados['valor_compra'].corr(dados['valor_cupom'])
...
52  print(f"Correlação (compra x cupom): {correlacao:.3f}")
```

Logo o valor da correlação é aproximadamente $r = 0.86$ Correlação positiva forte.

- Significa que quanto maior o valor da compra, maior tende a ser o valor do cupom.
- Há uma relação linear bem consistente entre as duas variáveis.

6. Visualizações Gráficas

Visualização referente aos valores de cupons e de compras respectivas.



Referente a correlação das variáveis, verificando o intervalo de confiança

