

IFPI

Curso: TDS

Disciplina: Tópicos Especiais em Programação

Exercício: Explorando o NumPy com Análise de Dados

Contexto da Atividade

Nesta atividade, você será um analista de dados que recebeu informações de uma empresa sobre o desempenho de vendas de produtos em diferentes regiões. Seu objetivo é manipular, analisar e extrair insights desses dados utilizando as funcionalidades do NumPy.

Dataset Sugerido

Criaremos um pequeno conjunto de dados fictícios sobre vendas. Aqui está a estrutura do dataset:

- **Colunas:**
 - ID: Identificação do produto (1 a 10)
 - Preço: Preço unitário do produto (em reais)
 - Vendas Região A: Quantidade de unidades vendidas na Região A
 - Vendas Região B: Quantidade de unidades vendidas na Região B
 - Custo: Custo unitário de produção (em reais)
-

Código para Gerar o Dataset

Antes de iniciar a atividade, execute o seguinte código para criar o dataset:

```
import numpy as np
```

```
# Gerar dados aleatórios
```

```
np.random.seed(42) # Para resultados reproduzíveis
```

```
ids = np.arange(1, 11) # IDs de 1 a 10
```

```
precos = np.random.randint(20, 200, size=10) # Preços entre R$20 e R$200
```

```
vendas_a = np.random.randint(50, 500, size=10) # Vendas na Região A
```

```
vendas_b = np.random.randint(30, 400, size=10) # Vendas na Região B
```

```
custos = np.random.randint(10, 150, size=10) # Custos entre R$10 e R$150
```

```
# Criar o dataset como um array
```

```
dataset = np.column_stack((ids, precos, vendas_a, vendas_b, custos))
```

```
# Exibir o dataset
```

```
print("Dataset de Vendas:\n", dataset)
```

Tarefas

1. Exploração de Dados

- Imprima o formato do array (número de linhas e colunas).
- Verifique o tipo dos dados no array.
- Encontre o valor mínimo e máximo de cada coluna.

2. Cálculos Estatísticos

- Calcule a média, mediana e desvio padrão das vendas em cada região.
- Determine a receita total de vendas para cada região (preço × quantidade vendida).
- Identifique o produto com maior margem de lucro (Preço - Custo).

3. Manipulação de Dados

- Adicione uma coluna de "Receita Total" ao dataset (soma das receitas das Regiões A e B).
- Filtre apenas os produtos com receita total acima de R\$10.000.

4. Visualização de Insights

- Ordene o dataset pela receita total (em ordem decrescente).
- Encontre a região com maior volume de vendas.