

Bootcamp: Analista de Dados com Ênfase no Mercado financeiro

Trabalho Prático

Módulo 2: Fundamentos em análise e ciência de dados

Objetivos de Ensino

Exercitar os seguintes conceitos trabalhados no Módulo:

1. Coleta de dados.
2. Analisar e realizar tratamento de dados.
3. Analisar resultados obtidos.
4. Conhecimento teórico ministrado nas videoaulas.

Enunciado

Como cientista de dados, você foi incumbido de analisar os dados de vendas de um e-commerce nacional, especializado em produtos vendidos por vendedores autônomos. Nesse cenário, a empresa oferece uma ampla variedade de itens para consumidores em todo o Brasil. Os dados disponíveis são ricos em informações, incluindo o nome do vendedor, produto vendido, quantidade unitária, valor do produto, valor total da venda, data da compra e localização dos clientes por estado brasileiro.

Uma característica fundamental desse e-commerce é a dinâmica dos preços dos produtos, os quais podem variar diariamente devido à demanda do mercado e às flutuações na cotação da moeda. Isso implica que o mesmo produto pode apresentar preços diferentes em momentos e regiões distintas.

Seu objetivo como cientista de dados é explorar esses dados meticulosamente, buscando insights valiosos para a empresa. A intenção é investigar os padrões de consumo em diferentes estados brasileiros, gerando informações que serão cruciais para a tomada de decisões estratégicas.

Através dessa análise profunda, espera-se fornecer à empresa informações e recomendações precisas. Essas informações serão essenciais para otimizar as estratégias de vendas, compreender as preferências dos clientes em diversas regiões e adaptar os preços dos produtos de forma inteligente, visando atender melhor às demandas do mercado e impulsionar o crescimento do negócio.

ATENÇÃO PARA TRATAMENTO DE DADOS

Avalie se será necessário realizar tratamento de dados ausentes no datasets disponibilizado.

Instruções para correção de dados ausentes

1. Para dados os dados categóricos, utilize a estratégia de exclusão dos dados.
2. Para os dados numéricos, utilize:
 - a. Regra do negócio para corrigi-los.
 - b. Atenção para a correção baseada na regra.
 - i. Ex.: valor total da venda = Quantidade unitária * Preço do produto.
 - ii. Utilize os próprios dados para fazer a correção.

Atividades

Para esta atividade, os alunos deverão realizar análise de dados utilizando as bibliotecas pandas.

1. Criar uma virtualenv e instalar as bibliotecas necessárias.
2. Coletar e inserir o arquivo dados_vendas_produtos.csv na plataforma.
3. Analisar os dados coletados.
4. Avaliar a necessidade de tratamentos de dados ausentes.
5. Manipulação e visualização dos resultados.
6. Responder às questões teóricas e práticas do trabalho.

Dicas do professor

1. Analise com cuidado os dados.
2. Antes de enviar as respostas, verifique se o gabarito está correto.
3. Eliminar dados duplicados (se houver).
4. Realize todas as manipulações e junções dos dados antes de responder às questões do desafio.
5. Atenção no momento de filtrar e corrigir dados (se necessário).
6. Tenha atenção no que pede cada questão.
7. Os dados disponibilizados no dataset são fictícios. Ou seja, não têm relação com o mundo real.
8. Siga fielmente todos os passos contidos no enunciado das questões.
9. Realize os tratamentos de dados antes de responder as questões.

10.O dataset utilizado no trabalho pode ser obtido no link:

<https://leandrolessa.com.br/datasets>

11.Abaixo segue as versões utilizadas das bibliotecas neste trabalho:

```
Bibliotecas Utilizadas:  
PANDAS: 2.0.1
```