

Exercício: Avaliando Dados para Transformação (Adventure Works)

Introdução

Imagine que você acabou de entrar na sede da Adventure Works como analista de dados. Sua gerente solicita que você avalie amostras de diversos blocos de dados usando o **Microsoft Excel**. Esses dados serão posteriormente carregados no **Power BI**. Este exercício ajudará você a entender como avaliar e transformar dados na fonte, lidando com valores ausentes, formatos inconsistentes e erros para deixar os dados prontos para análise.

Instruções Gerais

1. Baixe os arquivos CSV: **Inventory** (Estoque), **Sales** (Vendas) e **Customer Feedback** (Feedback do Cliente).
 2. Abra os arquivos no Excel (Botão direito > Abrir com > Excel).
-

1. Conjunto de Dados de Vendas (Sales)

Pergunta de análise: "Qual é o valor total de vendas por produto?"

- **Passo 1: Tratando TransactionIDs ausentes**
 1. Selecione a coluna `TransactionID`.
 2. Na guia **Dados**, clique no botão **Filtro**.
 3. Clique na seta do cabeçalho e desmarque a caixa **Vazios** (Blanks).
 4. Defina o tipo de dado da coluna como **Número**.
 - **Passo 2: Lidando com anomalias no SalesAmount**
 1. Selecione a coluna `SalesAmount`.
 2. Identifique e substitua manualmente valores de texto por números (ex: onde está `four-hundred`, digite 400).
 3. Substitua células vazias por **0**: Pressione **CTRL + H**, deixe "Localizar" vazio e em "Substituir por" digite 0. Clique em **Substituir Tudo**.
-

2. Conjunto de Dados de Estoque (Inventory)

Pergunta de análise: "Qual é o nível de estoque para cada categoria de bicicleta?"

- **Passo 1: Tratando ProductName ausente**

1. Selecione a coluna `ProductName`.
 2. Ative o **Filtro** e desmarque a opção **Vazios** (Blanks). Clique em OK.
- **Passo 2: Lidando com Categoria ausente**
 1. Selecione a coluna `Category`.
 2. Pressione `CTRL + H`. Em "Localizar", deixe vazio. Em "Substituir por", digite: `No Category Provided`.
 - **Passo 3: Padronizando RestockingFrequency**
 1. Selecione a coluna `RestockingFrequency`.
 2. Pressione `CTRL + H`.
 3. Em "Localizar", digite `30 d`. Em "Substituir por", digite `30 days`. Marque "Coincidir conteúdo da célula inteira" nas opções e clique em **Substituir Tudo**.
-

3. Conjunto de Dados de Feedback do Cliente

Pergunta de análise: "Qual é a pontuação média de feedback para cada produto?"

- **Passo 1: Tratando FeedbackID ausente**
 1. Selecione a coluna `FeedbackID`.
 2. Ative o **Filtro** e desmarque a opção **Vazios** (Blanks).
- **Passo 2: Lidando com FeedbackScore inconsistente**
 1. Defina o tipo de dado da coluna como **Número**.
 2. Substitua manualmente valores de texto (ex: onde está `four-comma-seven`, digite `4.7` ou `4,7`).
- **Passo 3: Datas inválidas na coluna FeedbackDate**
 1. Selecione a coluna `FeedbackDate`.
 2. Defina o tipo de dado como **Data** na faixa de opções.
 3. Para ajustar o formato regional: Vá em "Mais Formatos de Número" > Personalizado > Digite `mm/dd/yyyy`.
 4. Corrija manualmente valores de texto (ex: onde está `five May 2023`, digite `05/05/2023`).