

Exercício: Preparando um Dataset

Caso de Estudo

A Adventure Works precisa de ajuda para analisar seus dados de vendas e criar um relatório Power BI que visualize esses dados de forma significativa. Antes que a análise possa começar, os dados brutos devem ser limpos e transformados para garantir que sejam precisos e consistentes. A empresa envia um arquivo CSV contendo o dataset bruto chamado SalesFile.csv. O dataset consiste em dados importantes relacionados a vendas recentes, como categorias de produtos, preços de fabricação, preços de venda, unidades vendidas e outros detalhes similares.

No entanto, o arquivo também contém erros como valores ausentes, tipos de dados incorretos e formatação inconsistente. Ajude a Adventure Works a resolver esses problemas usando o Power Query Editor para que possam produzir um relatório preciso e confiável.

Objetivos

Entender como abordar problemas comuns de dados como valores ausentes, tipos de dados incorretos e formatação inconsistente. Aprender a importar, limpar e transformar dados no Power BI Desktop para garantir análise precisa e confiável.

Instruções

Crie um novo projeto Power BI chamado Exercise – Preparing a dataset. Siga as etapas abaixo para completar o exercício.

Passo 1: Carregar o Arquivo Baixar o arquivo Excel SalesFile.xlsx e importar o arquivo no Power BI.

Passo 2: Abrir Power Query Editor Abrir o Power Query Editor para começar a editar os dados.

Passo 3: Tratar Valores Ausentes Localizar e selecionar a coluna Units Sold. Identificar todos os valores nulos e substituí-los por 0. Repetir para as colunas Sale Price, Sales e Profit.

Passo 4: Limpar Colunas Manufacturing Price e Sale Price Localizar e selecionar as colunas Manufacturing Price e Sale Price. Alterar o tipo de dados para Decimal Number. Repetir para as colunas Sales e Profit.

Passo 5: Limpar Coluna Discount Band Selecionar a coluna Discount Band. Localizar cada instância do valor 1 na coluna. Substituir cada instância por None. Alterar o tipo de dados da coluna para Text.

Passo 6: Limpar Coluna Units Sold Selecionar a coluna Units Sold. Localizar todas as instâncias do texto "six hundred". Substituir cada instância pelo valor numérico 600. Alterar o tipo de dados para Whole Number.

Passo 7: Tratar Inconsistências na Coluna Date Selecionar a coluna Date. Garantir que o tipo de dados seja Date. Substituir todos os valores nulos pela data padrão 03/03/2023. Selecionar a coluna Month Number e alterar o tipo de dados para Whole Number.

Passo 8: Remover Registros com Erros Selecionar a coluna Manufacturing Price. A coluna contém erros nas linhas 6 e 38. Usar Remove Errors para remover esses registros. Repetir para erros nas colunas Sales e Profit.

Passo 9: Remover Linhas Duplicadas Identificar linhas duplicadas no dataset. Dica: Verificar a coluna Products. Remover as linhas duplicadas.

Passo 10: Aplicar Transformações Após completar todas as etapas de limpeza, selecionar Close & Apply. Isso aplicará as transformações e atualizará o dataset.