[Prepare Data in Power BI Desktop](https://microsoftlearning.github.io/PL-300-Microsoft-Power-BI-Data-Analyst/Instructions/01-prepare-data-with-power-query-in-power-bi-desktop.html)

# Preparar dados no Power BI Desktop

O tempo estimado para concluir o laboratório é de 45 minutos.

Este laboratório é um dos muitos em uma série de laboratórios que foi projetado como uma história completa desde a preparação de dados até a publicação como relatórios e painéis. Você pode concluir os laboratórios em qualquer ordem. No entanto, se você pretende trabalhar em vários laboratórios, sugerimos que os faça na seguinte ordem:

1. Preparar dados no Power BI Desktop
2. Carregar dados no Power BI Desktop
3. Projetar um modelo de dados no Power BI
4. Criar cálculos DAX no Power BI Desktop
5. Criar cálculos DAX avançados no Power BI Desktop
6. Projetar um relatório no Power BI Desktop
7. Aprimorar um relatório no Power BI Desktop
8. Executar análise de dados no Power BI
9. Criar um painel do Power BI
10. Aplicar segurança em nível de linha

## história do laboratório

Este laboratório foi desenvolvido para apresentar o aplicativo Power BI Desktop e como se conectar aos dados e como usar técnicas de visualização de dados para entender as características e a qualidade dos dados de origem. Os objetivos de aprendizagem são:

* Abra o Power BI Desktop
* Conectar-se aos dados de origem
* Visualizar dados de origem
* Use técnicas de visualização de dados para entender melhor os dados

## Exercício 1: Preparar dados

Neste exercício, você criará oito consultas do Power BI Desktop. Seis consultas fornecerão dados do SQL Server e duas de arquivos CSV.

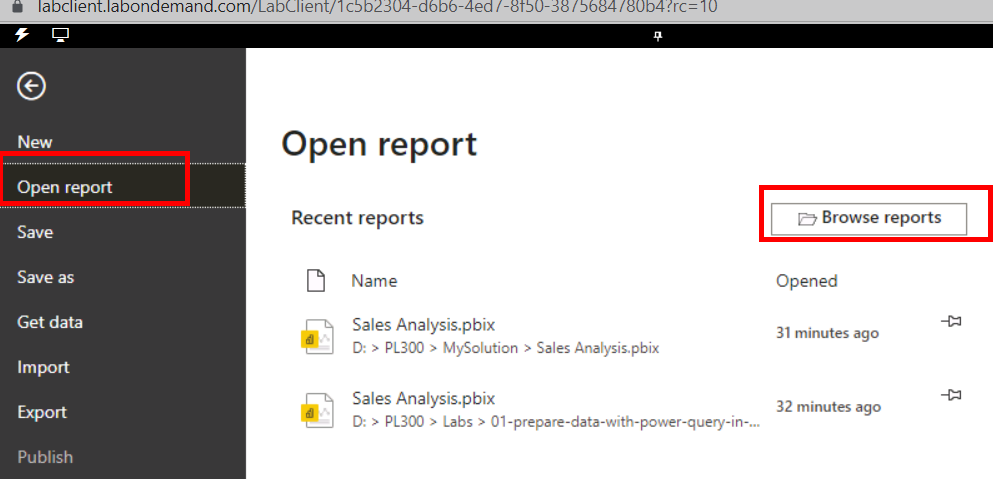
### Tarefa 1: Introdução ao Power BI Desktop

Nesta tarefa, **você começa abrindo um arquivo inicial do Power BI (.pbix)**. O arquivo inicial não contém nenhum dado, mas foi configurado especialmente para ajudá-lo a concluir o laboratório. As seguintes configurações de nível de relatório foram desativadas no arquivo inicial:

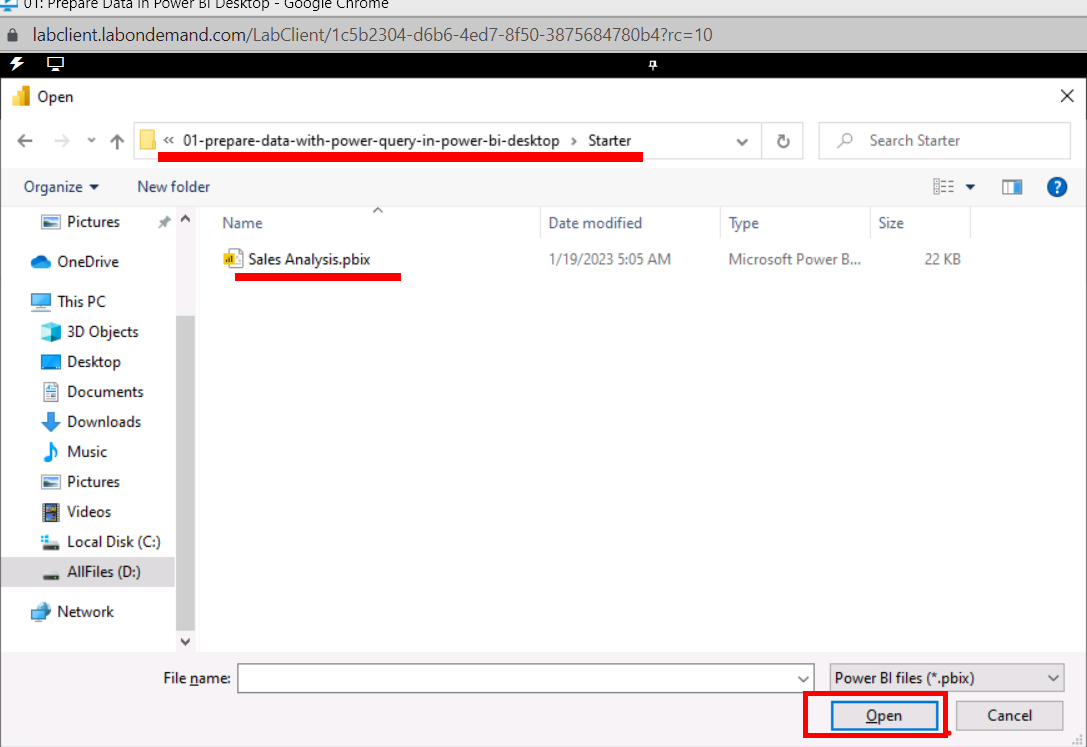
* Carregamento de dados > Relacionamentos de importação de fontes de dados no primeiro carregamento
* Carregamento de dados > Detecta automaticamente novos relacionamentos depois que os dados são carregados

*Nota: embora ter essas duas opções ativadas possa ser útil ao desenvolver um modelo de dados, você as desativou anteriormente para oferecer suporte à experiência de laboratório. Ao criar relações no laboratório Carregar dados no Power BI Desktop , você aprenderá por que está adicionando cada uma delas.*

1. Abra o Power BI Desktop.  
     
   *Dica: Por padrão, a caixa de diálogo Introdução é aberta na frente do Power BI Desktop. Você pode optar por entrar e fechar o pop-up.*
2. **Para abrir o arquivo inicial do Power BI Desktop, selecione File > Open Report > Browse Reports** .



1. Na janela Abrir , navegue até a pasta D:\PL300\Labs\01-prepare-data-with-power-query-in-power-bi-desktop\Starter

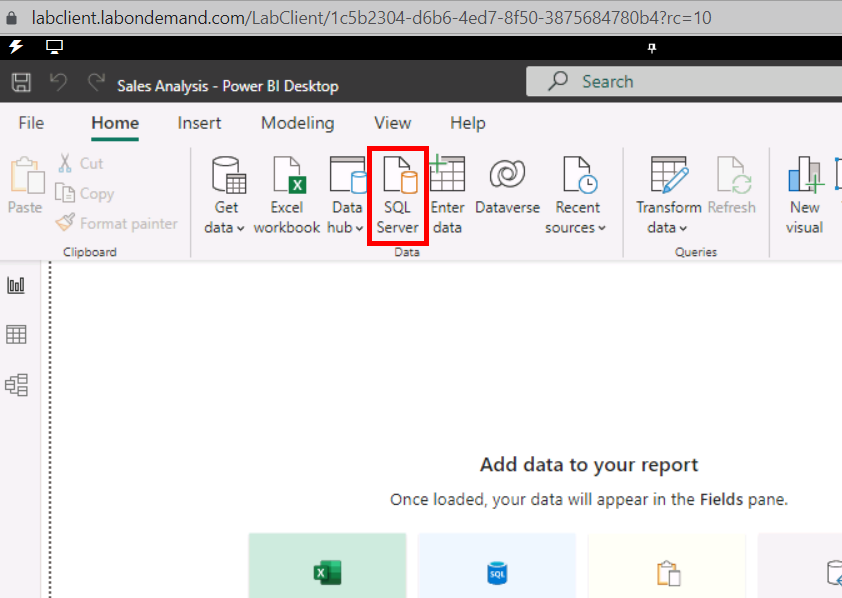


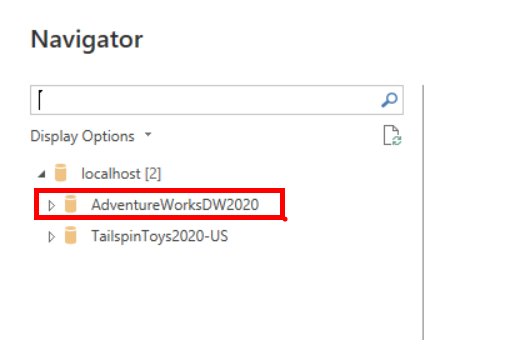
1. Selecione o arquivo de **Sales Analysis** .
2. **Salve uma cópia do arquivo com Salvar como na pasta D:\PL300\MySolution**

### 

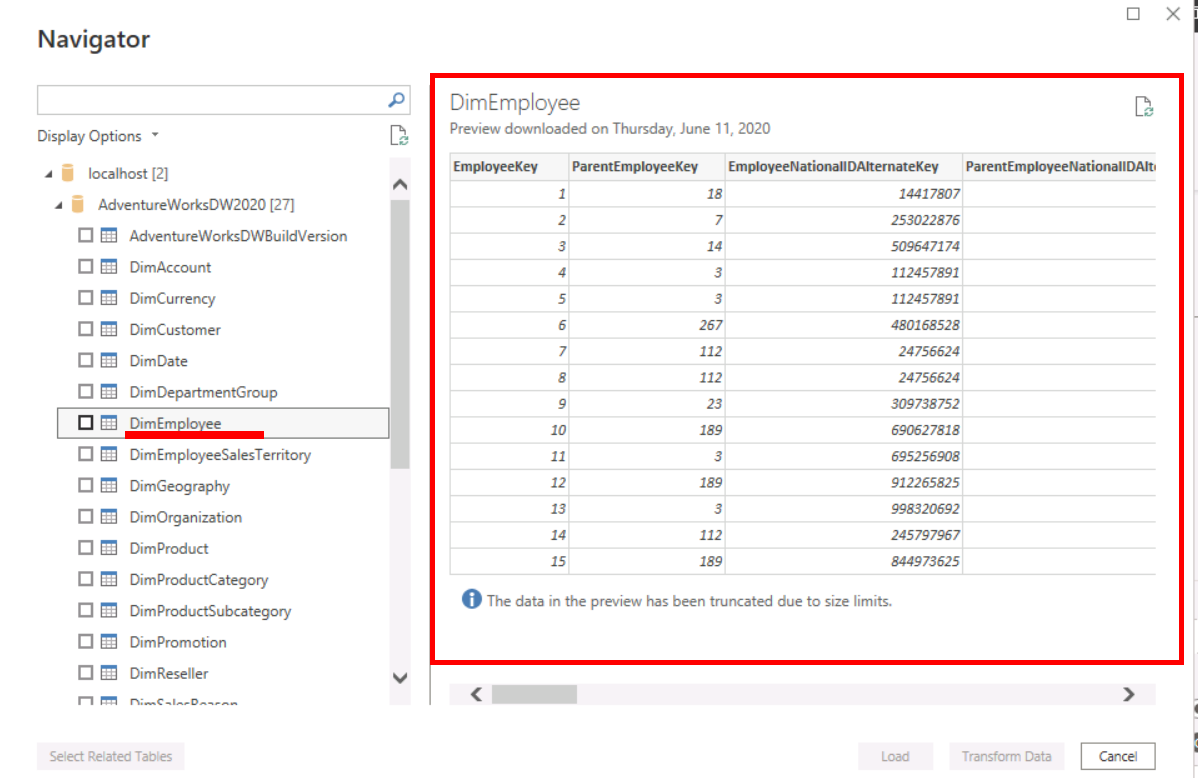
### Tarefa 2: obter dados do SQL Server

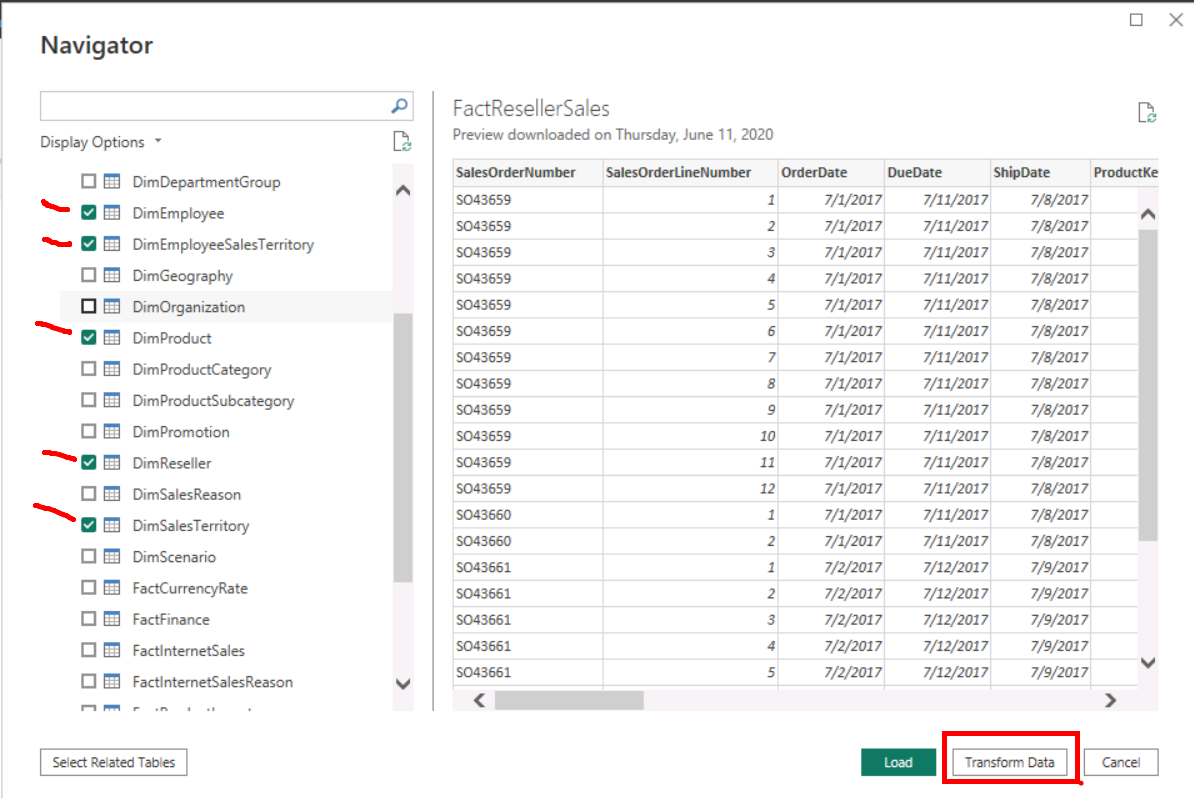
**Esta tarefa ensina como conectar-se a um banco de dados SQL Server e importar tabelas**, que criam consultas no Power Query.

1. Na guia da faixa de opções Início , dentro do grupo Dados , selecione SQL Server .  
   
2. Na janela **SQL Server Database** , na caixa Server , **digite localhost e selecione OK** .
3. Se as credenciais forem solicitadas, na janela Banco de dados do SQL Server , selecione **Use my current credentials** atuais e, em seguida, Conectar .
4. Na janela Navegador , à esquerda, expanda o banco de dados **AdventureWorksDW2020** .



*Observação: o banco de dados AdventureWorksDW2020 é baseado no banco de dados de exemplo AdventureWorksDW2017 . Ele foi modificado para apoiar os objetivos de aprendizagem dos laboratórios do curso.*

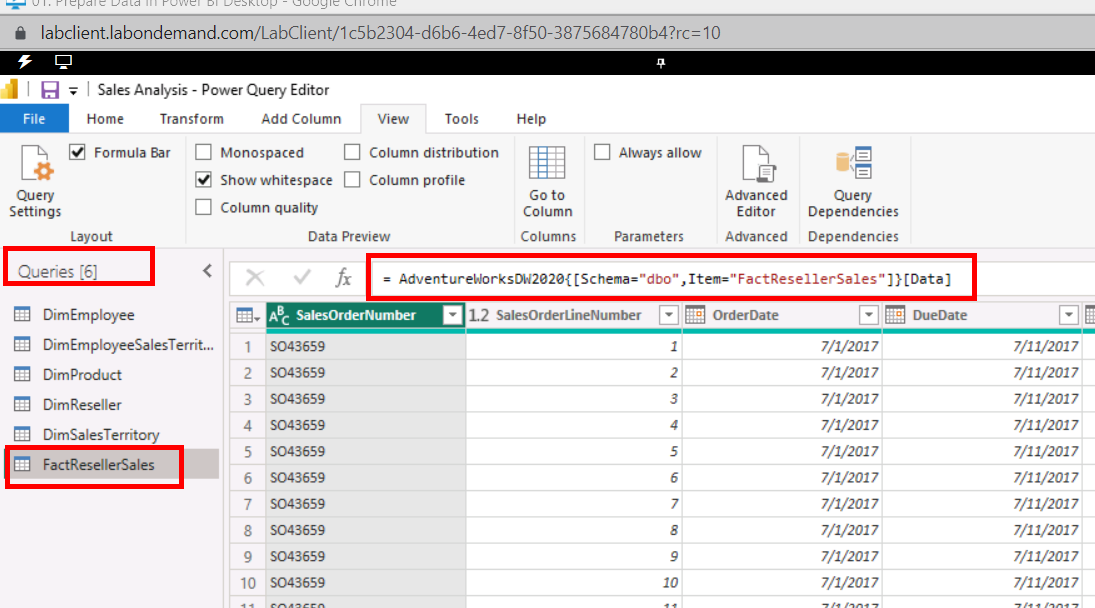
1. Selecione - **mas não marque - a tabela DimEmployee**  
   
2. No painel direito, observe uma visualização dos dados da tabela. **Os dados de visualização permitem que você veja as colunas e uma amostra de linhas**.
3. Para criar consultas, marque a caixa de seleção ao lado das **seis tabelas a seguir**:
   1. DimEmployee
   2. DimEmployeeSalesTerritory
   3. DimProduct
   4. DimReseller
   5. DimSalesTerritory
   6. FactResellerSales
4. Conclua esta tarefa clicando em Transform Data , que abrirá o Editor do Power Query.

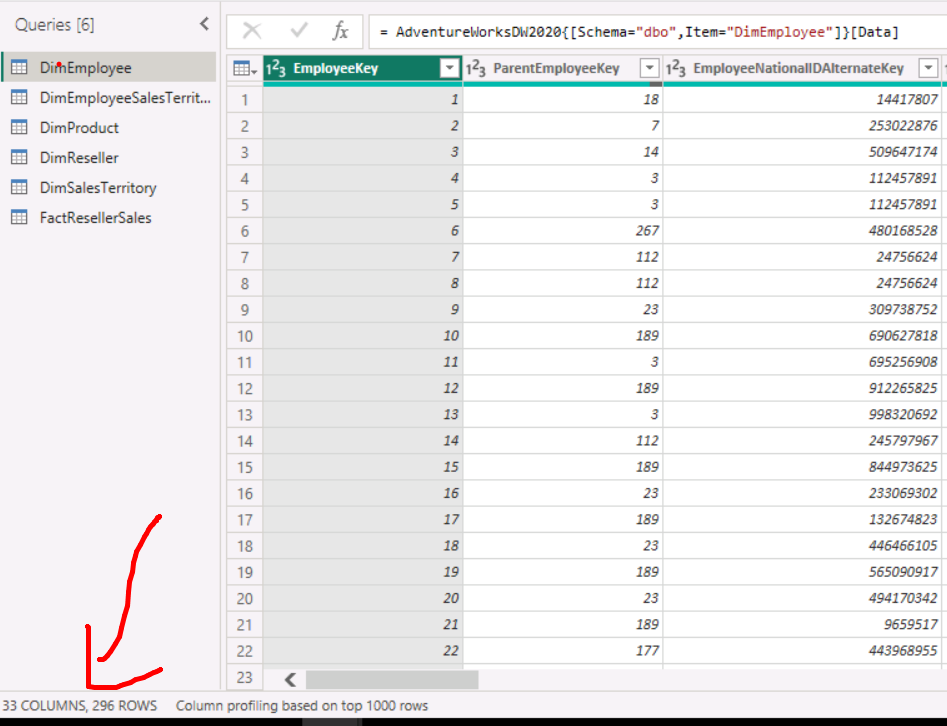


* 1. *Este laboratório destina-se apenas a conectar e criar o perfil dos dados, mas não transformar dados .*

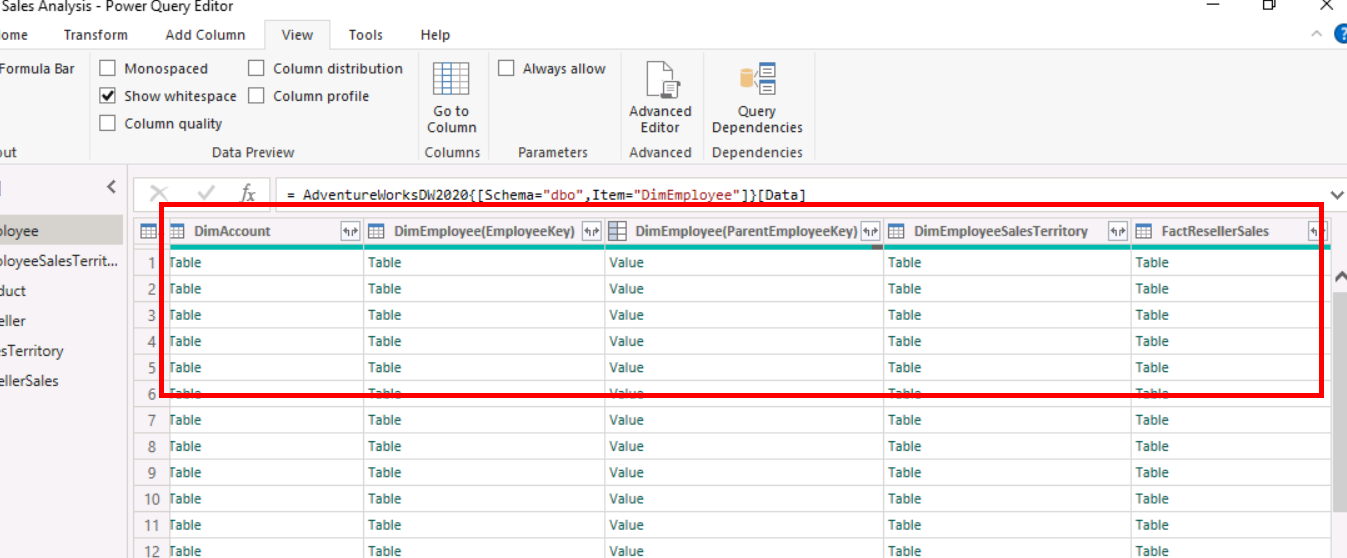
### Tarefa 3: visualizar dados no editor do Power Query

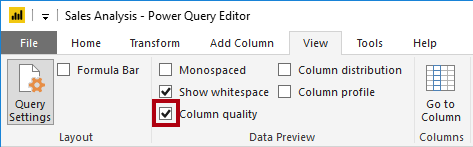
**Esta tarefa apresenta o Editor do Power Query e permite revisar e criar o perfil dos dados**. Isso ajuda a determinar como limpar e transformar os dados posteriormente.

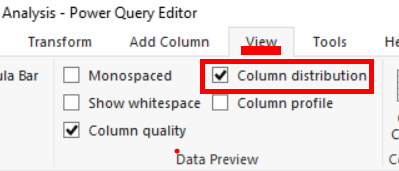
1. Na janela do Editor do Power Query , à esquerda, observe o painel Consultas . **O painel Consultas contém uma consulta para cada tabela verificada**.  
   
2. Selecione a primeira consulta — DimEmployee .  
   **A tabela DimEmployee no banco de dados do SQL Server armazena uma linha para cada funcionário.**
3. *Um subconjunto das linhas dessa tabela representa os vendedores, que serão relevantes para o modelo que você desenvolverá.*
4. **No canto inferior esquerdo da barra de status, algumas estatísticas da tabela são fornecidas** - a tabela tem 33 colunas e 296 linhas.



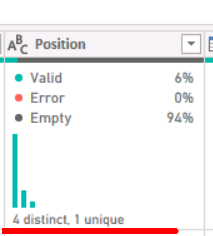
1. No painel de visualização de dados, role horizontalmente para revisar todas as colunas. **Observe que as últimas cinco colunas contêm links de Tabela ou Valor** .  
   **Essas cinco colunas representam relacionamentos com outras tabelas no banco de dados***. Eles podem ser usados ​​para unir tabelas. Você unirá tabelas no laboratório Carregar dados no Power BI Desktop .*

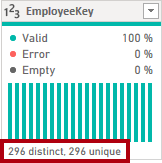
**

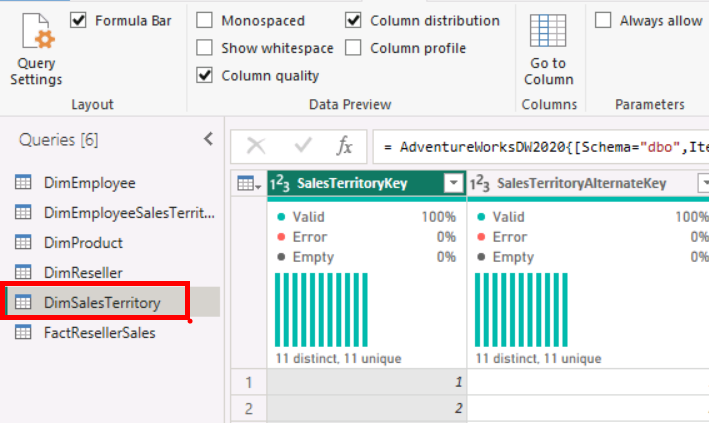
1. Para avaliar a qualidade da coluna, na **guia View** da faixa de opções, dentro do grupo **Data Preview** , marque **Column Quality** . **O recurso de qualidade da coluna permite determinar facilmente a porcentagem de valores válidos, de erro ou vazios encontrados nas colunas.**
2. Observe que a coluna **Position** tem 94% de linhas vazias (nulas).  
   
3. Para avaliar a distribuição da coluna, na **guia View** da faixa de opções, dentro do grupo **Data Preview** , marque **Column Distribution** .



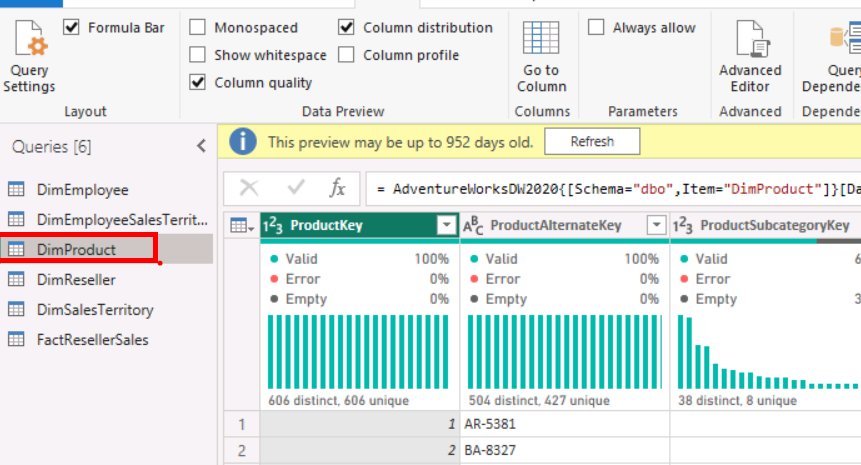
1. Revise a coluna **Position** novamente e observe que há quatro valores distintos e um valor exclusivo.



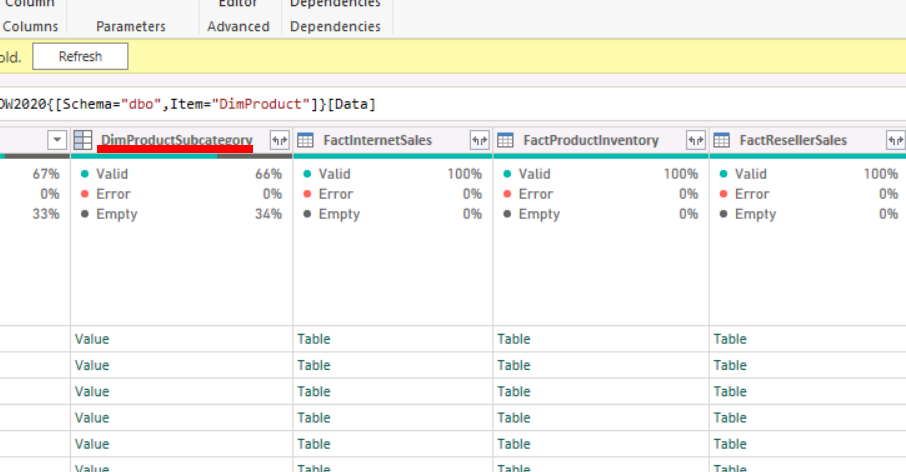
1. Revise a distribuição da coluna EmployeeKey — há 296 valores distintos e 296 valores exclusivos.
2. *Quando* **as contagens distintas e exclusivas são iguais, significa que a coluna contém valores exclusivos***. Ao modelar, é importante que algumas tabelas de modelo tenham colunas exclusivas. Essas colunas exclusivas podem ser usadas para criar relações um-para-muitos, o que você fará no laboratório Dados do modelo no Power BI Desktop .  
   *
3. No painel Consultas , **selecione a consulta DimEmployeeSalesTerritory** .  
   **A tabela DimEmployeeSalesTerritory armazena uma linha para cada funcionário e as regiões de território de vendas que eles gerenciam***. A tabela permite relacionar várias regiões a um único funcionário.* **Alguns funcionários gerenciam uma, duas ou possivelmente mais regiões***. Ao modelar esses dados, você precisará definir uma relação muitos-para-muitos.*

**

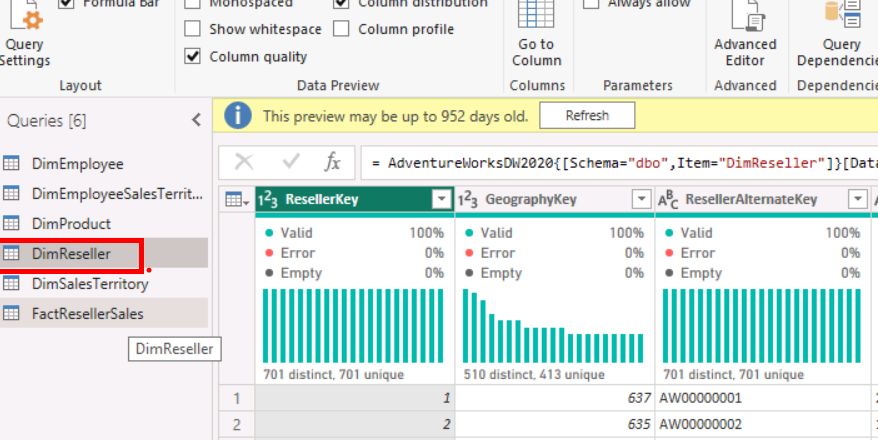
1. No painel Consultas , selecione a consulta **DimProduct** . **A tabela DimProduct contém uma linha por produto vendido pela empresa**.



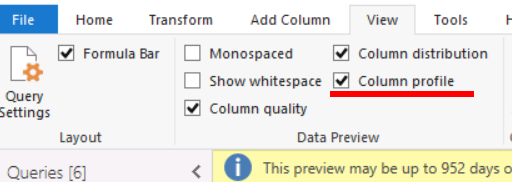
1. Role horizontalmente para revelar as últimas colunas. Observe a coluna **DimProductSubcategory** .  
   *Ao adicionar transformações a esta consulta no laboratório Carregar dados no Power BI Desktop , você usará a coluna DimProductSubcategory para unir tabelas.*

**

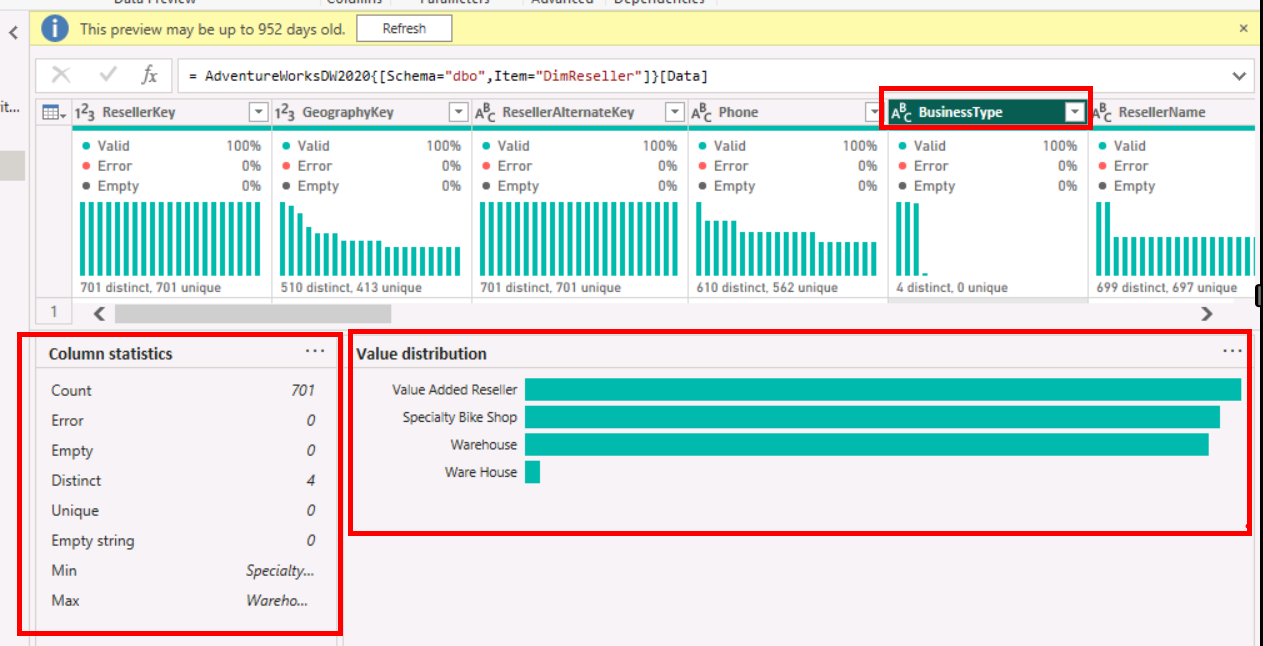
1. No painel Consultas , selecione a **consulta DimReseller** .  
   **A tabela DimReseller contém uma linha por revendedor***. Os revendedores vendem, distribuem ou agregam valor aos produtos Adventure Works.*

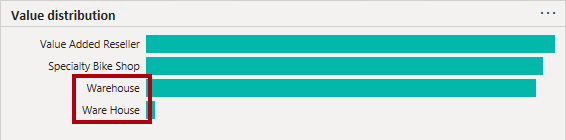
**

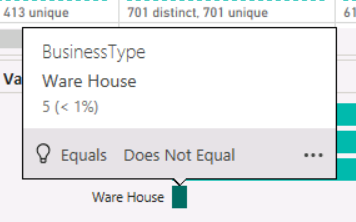
1. Para exibir os valores da coluna, na guia Exibir da faixa de opções, dentro do grupo Visualização de dados , marque Perfil da coluna .



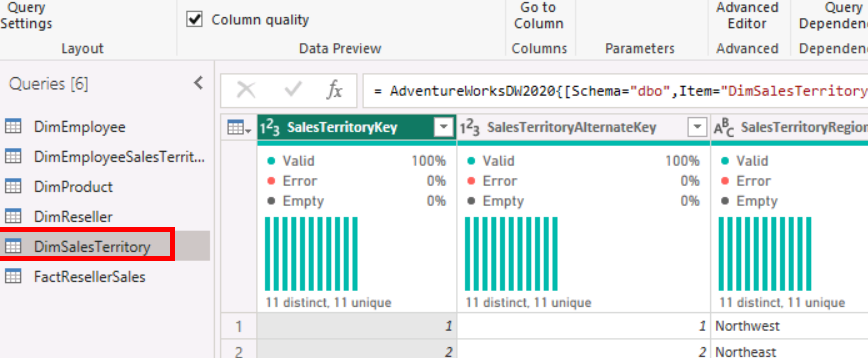
1. Selecione o cabeçalho da **coluna BusinessType e observe o novo painel abaixo do painel de visualização de dados.**

****

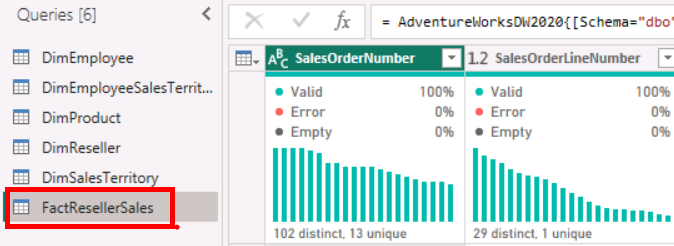
1. Revise as estatísticas da coluna e a distribuição de valores no painel de visualização de dados.  
   *Observe o problema de qualidade de dados: há dois rótulos para armazém ( Armazém e Armazém com erros ortográficos ) .  
   *
2. Passe o cursor sobre a barra Ware House e observe que há cinco linhas com esse valor.



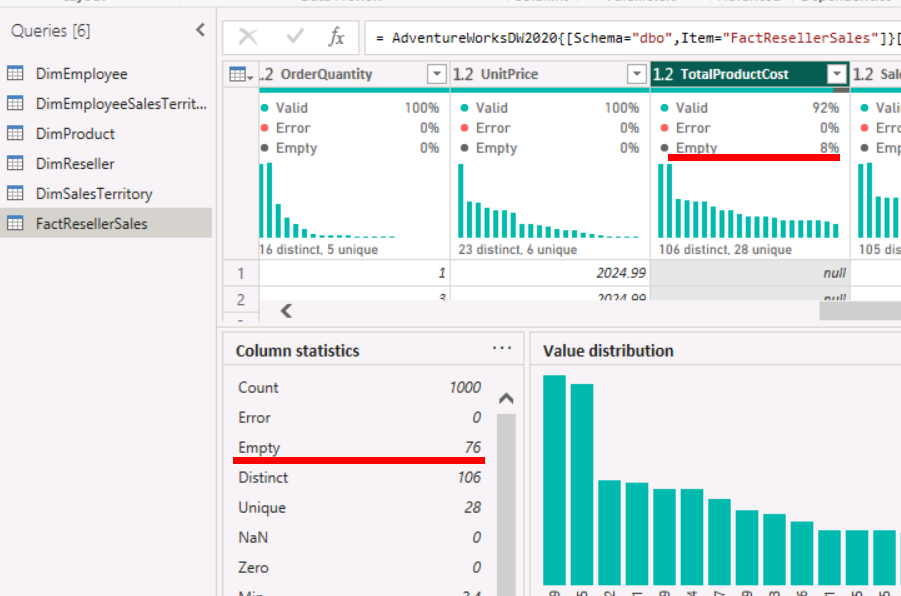
1. *Você aplicará uma transformação para renomear essas cinco linhas no laboratório Carregar dados no Power BI Desktop .*
2. No painel Consultas , **selecione a consulta DimSalesTerritory** .  
   **A tabela DimSalesTerritory contém uma linha por região de vendas***, incluindo Corporate HQ (sede). As regiões são atribuídas a um país e os países são atribuídos a grupos. No laboratório Model Data in Power BI Desktop , você criará uma hierarquia para dar suporte à análise em nível de região, país ou grupo.*

**

1. No painel Consultas , **selecione a consulta FactResellerSales** .  
   **A tabela FactResellerSales contém uma linha por linha de ordem de venda — uma ordem de venda contém um ou mais itens de linha***.*

**

1. R**evise a qualidade da coluna TotalProductCost** e observe que 8% das linhas estão vazias.  
   **A falta de valores da coluna TotalProductCost é um problema de qualidade de dados***. Para resolver o problema, no laboratório Carregar dados no Power BI Desktop , você aplicará transformações para preencher os valores ausentes usando o custo padrão do produto, armazenado na tabela DimProduct relacionada.*

**

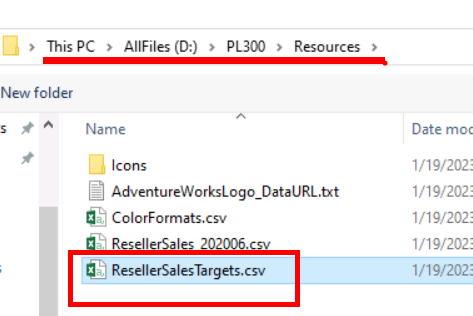
### Tarefa 4: obter dados de um arquivo CSV

Nesta tarefa, **você criará uma nova consulta baseada em arquivos CSV.**

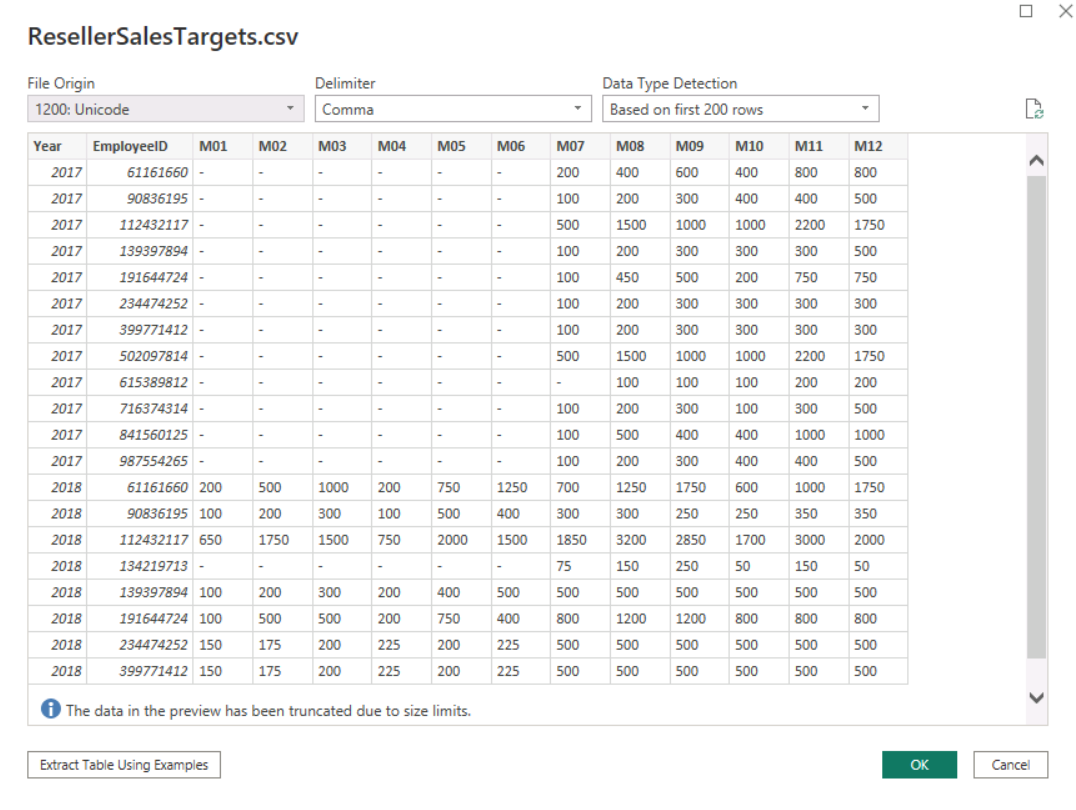
1. Para adicionar uma nova consulta, na janela no **Power Query Editor** , na guia da faixa de opções **Home** , de dentro do grupo **New Query** , selecione a seta para baixo **New Source** e, em seguida, **selecione Texto/CSV** .

|  |  |
| --- | --- |

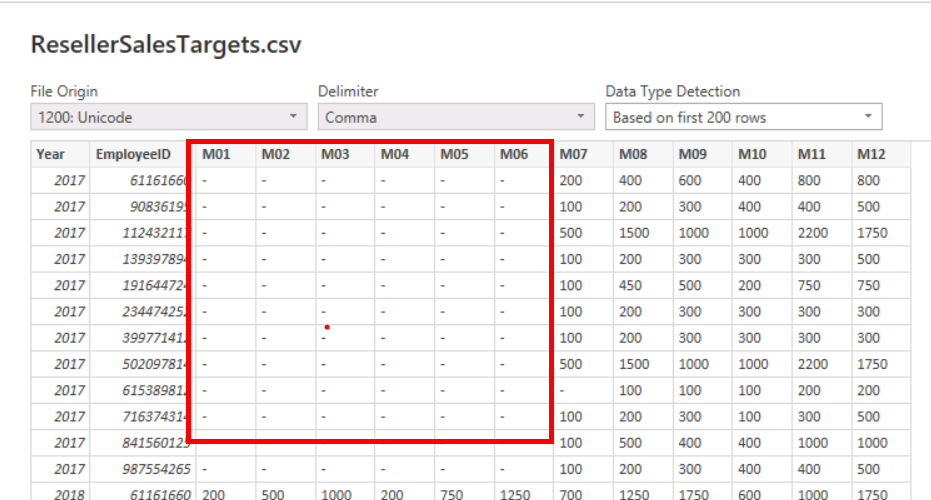
1. Na janela Abrir , **navegue até a pasta D:\PL300\Resources** e selecione o arquivo **ResellerSalesTargets.csv** . Selecione Abrir .
2. Na janela ResellerSalesTargets.csv , revise os dados de visualização. Selecione OK .

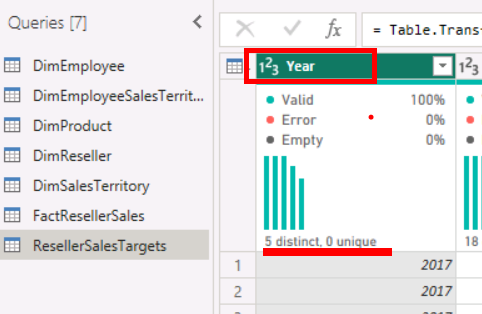


1. No painel Consultas , **observe a adição da consulta ResellerSalesTargets .  
   O arquivo CSV ResellerSalesTargets contém uma linha por vendedor, por ano***.* **Cada linha registra 12 metas de vendas mensais** *(expressas em milhares). O ano comercial da empresa Adventure Works começa em 1º de julho.*

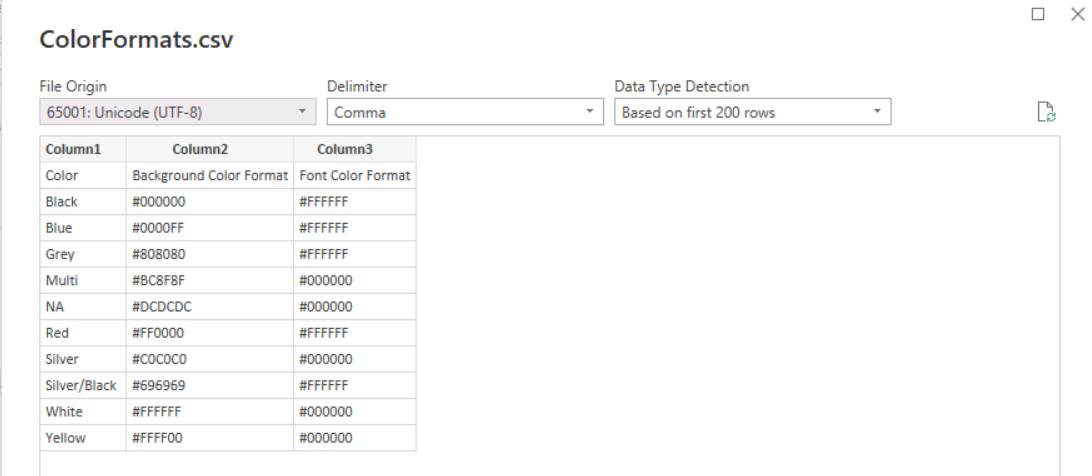
**

1. Observe que nenhuma coluna contém valores vazios. **Quando não há uma meta de vendas mensal, um caractere de hífen é armazenado**.

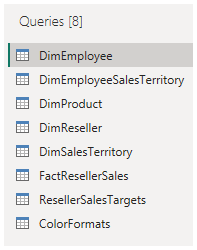


1. Revise os ícones no cabeçalho de cada coluna, à esquerda do nome da coluna. Os ícones representam o tipo de dados da coluna. 123 é um número inteiro e ABC é um texto.  
   
2. **Use as etapas da tarefa anterior para criar uma consulta com base no arquivo D:\PL300\Resources\ColorFormats.csv** .

*O arquivo ColorFormats CSV contém uma linha por cor do produto. Cada linha registra os códigos HEX para formatar o plano de fundo e as cores da fonte.*

**

*Agora você deve ter duas novas consultas, ResellerSalesTargets e ColorFormats .*

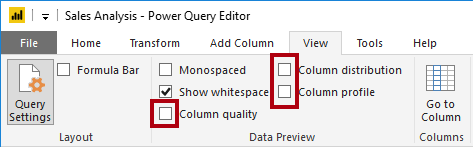
**

### Tarefa 5: Concluir

Nesta tarefa, você concluirá o laboratório.

Na guia da faixa de opções Exibir , dentro do grupo Visualização de dados , desmarque as três opções de visualização de dados que foram ativadas anteriormente neste laboratório:

* + Qualidade da coluna
  + Distribuição de colunas
  + Perfil da coluna



**Salve o arquivo do Power BI Desktop. Quando solicitado a aplicar as alterações pendentes, selecione Aplicar mais tarde** .  
*Dica: Aplicar as consultas carregará seus dados no modelo de dados.* **Você não está pronto para fazer isso, pois há muitas transformações que devem ser aplicadas primeiro.**