



ANÁLISE DE DADOS COM POWER BI

Aula 02

- Importa dados de diferentes fontes, atendendo as necessidades da ETL
- Cria um modelo de dados, usando o editor de consultas do Power BI





[Datasets com Base de Dados para Consulta]

Na internet podemos encontrar datasets gratuitos que servirão de modelagem para nossa ideias de desenvolvimento de dashboards utilizando dados reais. Consulte e teste seu conteúdo.

Base de dados gratuitas

Kaggle, base de dados internacional com simulações fictícia e reais : <https://www.kaggle.com/datasets>

Portal governamental da transparência, reuni dados governamentais: <https://portaldatransparencia.gov.br/>

Portal de dados abertos do governo federal: <https://dados.gov.br/home>

Dataset Google: <https://datasetsearch.research.google.com>

Base de dados do Instituto de Segurança Pública: <https://www.ispdados.rj.gov.br/estatistica.html>

Base da câmara dos deputados: <https://dadosabertos.camara.leg.br/swagger/api.html?tab=staticfile#staticfile>

Base de vôo ANAC: <https://www.gov.br/anac/pt-br/assuntos/dados-e-estatisticas/historico-de-voos>

Base de análise Airbnb: <https://insideairbnb.com/get-the-data/>

Base de dados do TSE eleições 2024: <https://dadosabertos.tse.jus.br/dataset/>

[Introdução ao Editor de Consultas]

O que é o Editor de Consultas ou Power Query?

Ferramenta integrada ao Power BI que podemos conectar, transformar e carregar dados.

Permite a limpeza e transformação dos dados antes de carregá-los no modelo de dados do Power BI.

The screenshot displays the Power Query Editor window titled "Sem título - Editor do Power Query". The interface includes a ribbon with tabs: Arquivo, Página Inicial, Transformar, Adicionar Coluna, Exibição, Ferramentas, and Ajuda. The "Exibição" tab is active, showing options for data visualization and layout. The main area displays a table with columns: Data da Venda, NF de Venda, UF Venda, Vendedor, Categoria do Produto, Preço de Custo, and Quantidade. Each column has a data type and a status bar indicating the number of valid, error, and blank rows. The table contains 14 rows of data. On the left, a context menu is open, showing options like Copiar, Colar, Excluir, Renomear, and others. On the right, the "Config. Consulta" pane is visible, showing the query name "BaseDadosVendas_V01" and the applied steps: Fonte, Navegação, Cabeçalhos Promovidos, and Tipo Alterado. The status bar at the bottom indicates "27 COLUNAS, 999+ LINHAS" and "Criação de perfil de coluna com base nas primeiras 1000 linhas".

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Barra de Fórmulas | Espaço uniforme | Distribuição de colunas | Permitir sempre | Editor Avançado | Dependências de Consulta

Mostrar espaço em branco | Perfil da coluna | Ir para Coluna | Parâmetros | Dependências

Qualidade da coluna

Layout | Visualização dos Dados

Consultas [1] | BaseDadosVendas_...

Table.TransformColumnTypes(#"Cabeçalhos Promovidos",{{"Data da Venda", type date}, {"NF de Venda", type text}, {"UF Venda", type text},

Colar

Nova Consulta

Novo Parâmetro...

Novo Grupo...

Expandir Tudo

Recolher Tudo

Config. Consulta

PROPRIEDADES

Nome: BaseDadosVendas_V01

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Navegação

Cabeçalhos Promovidos

X Tipo Alterado

27 COLUNAS, 999+ LINHAS | Criação de perfil de coluna com base nas primeiras 1000 linhas

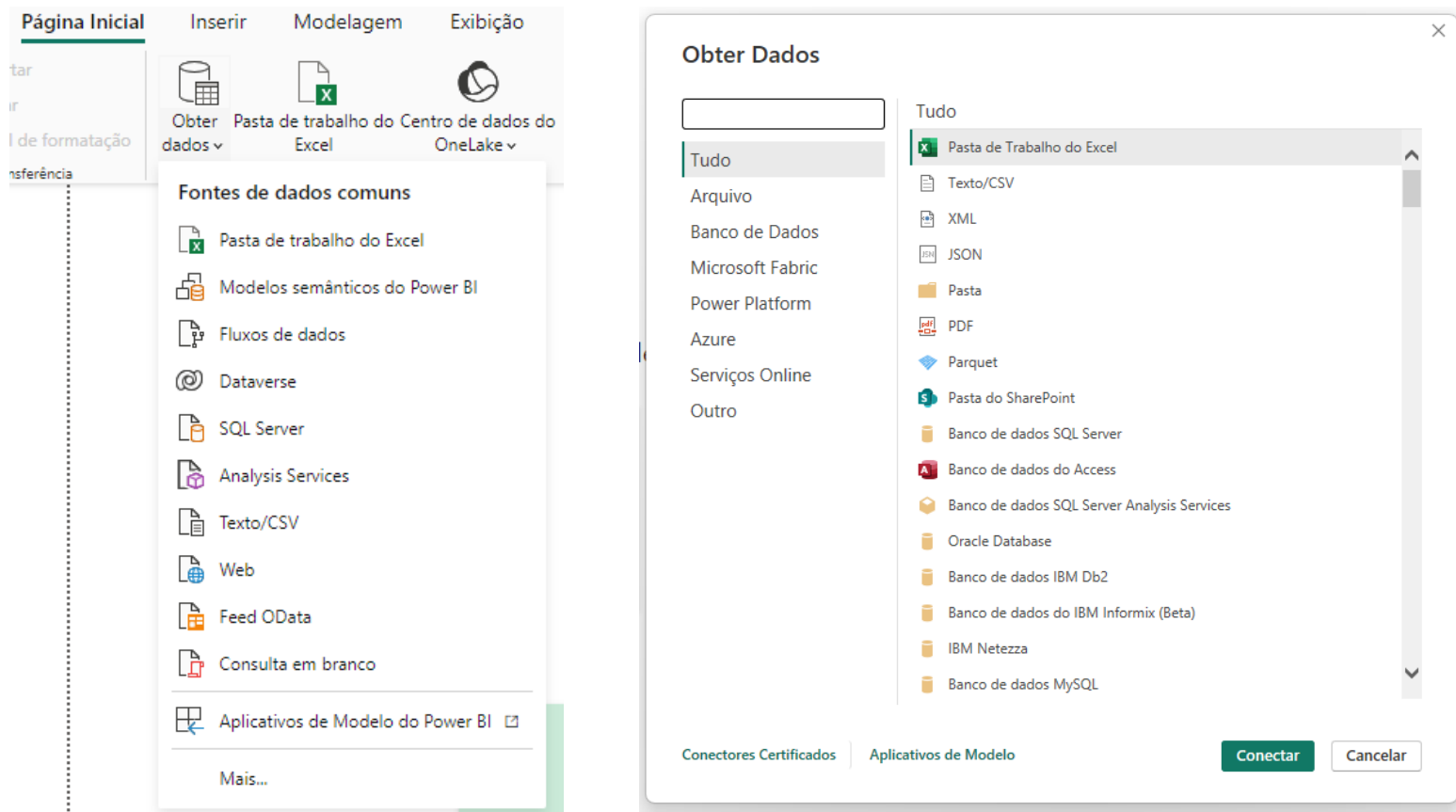
VISUALIZAÇÃO BAIXADA À(S) 01:21

[Introdução ao Editor de Consultas]

Passos para Criar um Modelo de Dados

Conectar-se a uma Fonte de Dados

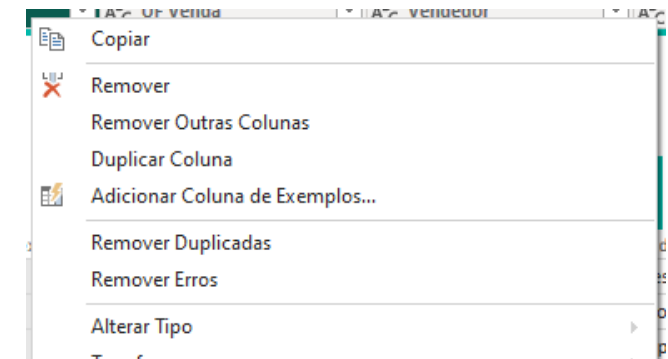
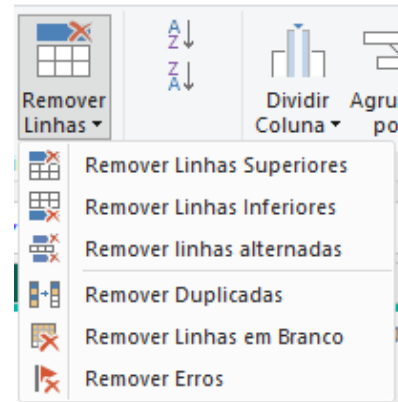
O Power BI permite conectar-se a diferentes fontes de dados como Excel, SQL Server, Web, JSON, TXT, CSV como outras fontes.



[Transformando Dados]

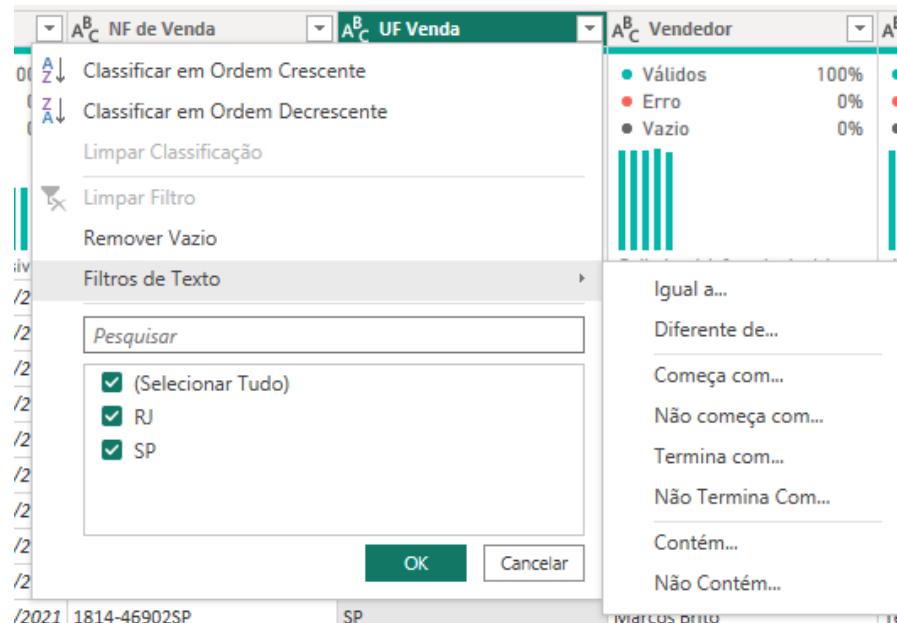
Remover Colunas Desnecessárias

Processo de remoção de colunas que não são necessárias para a análise.



Filtrar Linhas

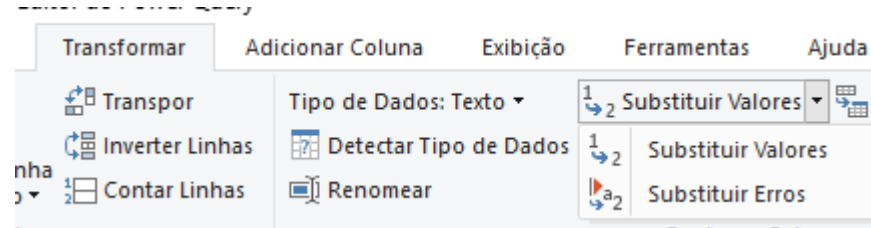
Aplicação de filtrar dados para incluir apenas as informações relevantes.



[Transformando Dados]

Substituir Valores

No Power Query podemos substituir valores nulos ou incorretos sem afetar a consistência de dados na base de origem.

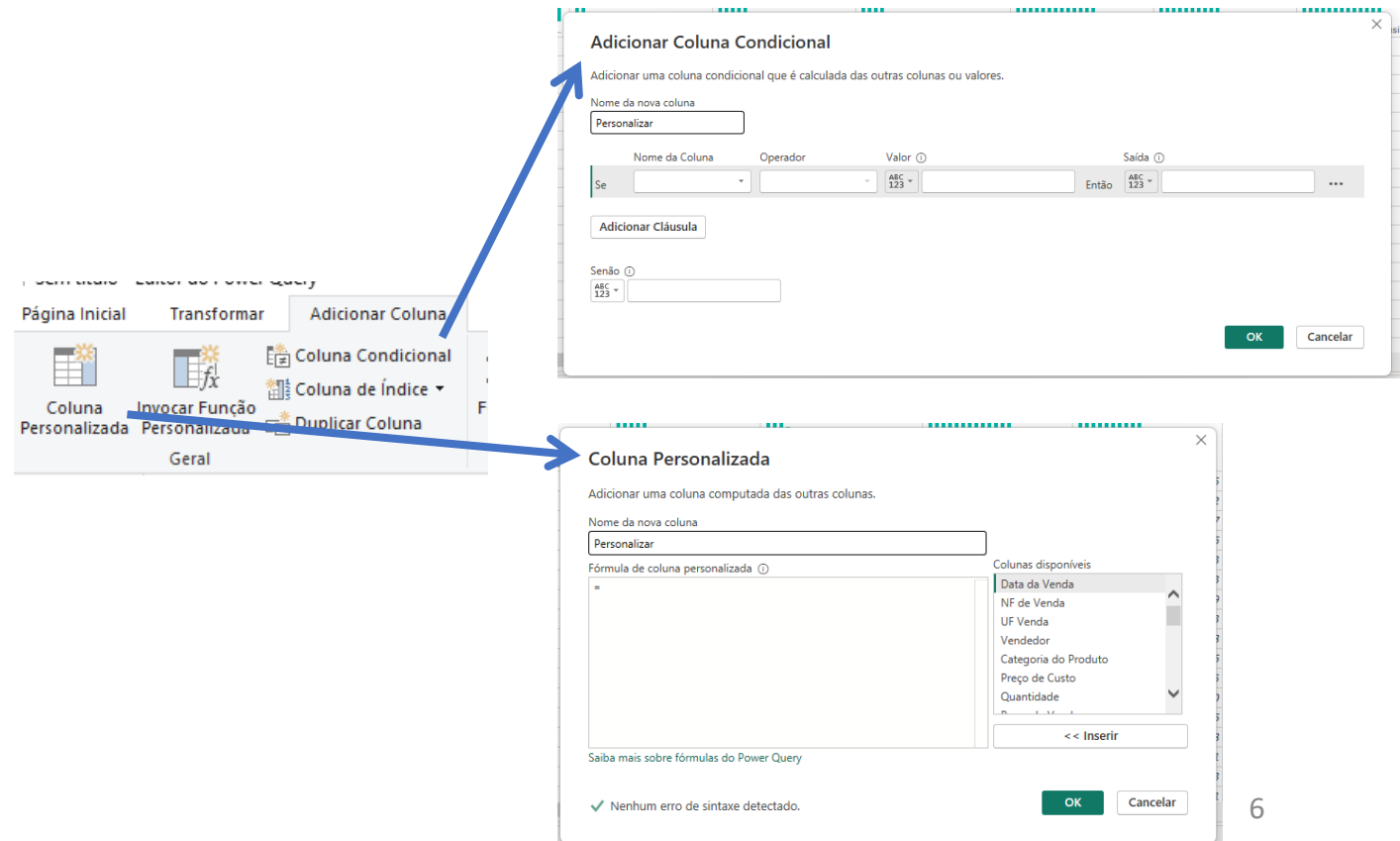


Adicionar Colunas Calculadas

Use o Editor de Consultas para criar colunas personalizadas com base em fórmulas ou condições específicas.

Podemos renomear, combinar ou criar novas colunas com base em cálculos específicos.

As colunas personalizadas são adicionadas como etapas aplicadas à consulta e podem ser modificadas a qualquer momento.

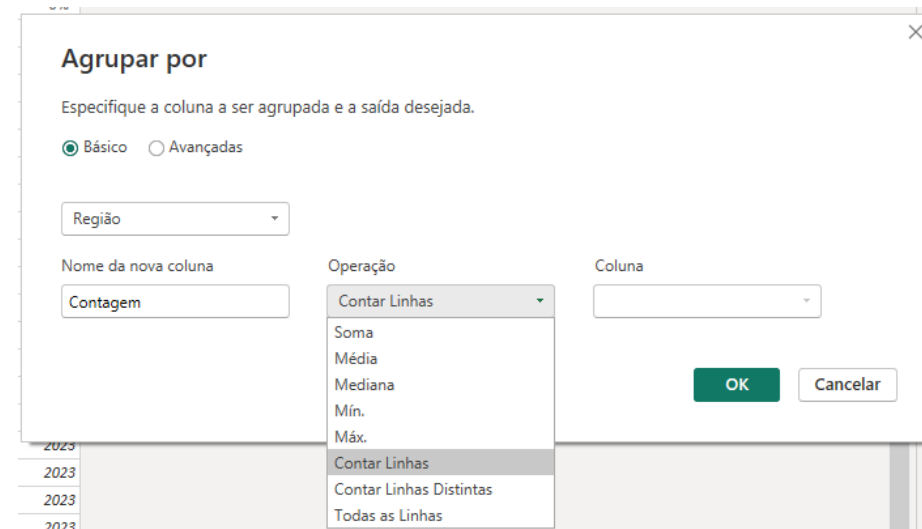
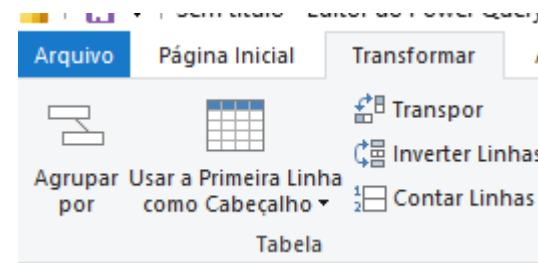


[Agrupamento de Linha]

No Power Query, o agrupamento de linhas é uma funcionalidade que permite consolidar dados de várias linhas em um único valor, com base nos valores de uma ou mais colunas. Isso é útil para resumir dados e criar relatórios mais concisos e informativos.

Como Funciona o Agrupamento de Linhas

- 1. Selecionar Colunas para Agrupar:** Você escolhe uma ou mais colunas cujos valores serão usados para agrupar as linhas.
- 2. Escolher Função de Agregação:** Você pode aplicar funções de agregação, como soma, média, contagem, mínimo, máximo, entre outras, para consolidar os dados das linhas agrupadas.
- 3. Gerar Nova Tabela:** O resultado é uma nova tabela onde as linhas são agrupadas de acordo com os critérios definidos, e os valores são agregados conforme a função escolhida.



[Agrupamento de Linha]

Benefícios do Agrupamento de Linhas

Simplificação de Dados

Reduz a quantidade de dados exibidos, facilitando a análise.

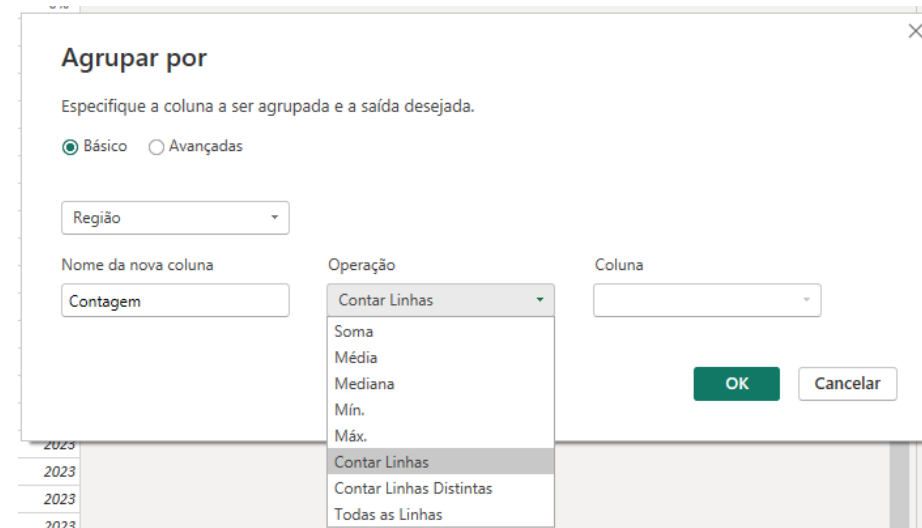
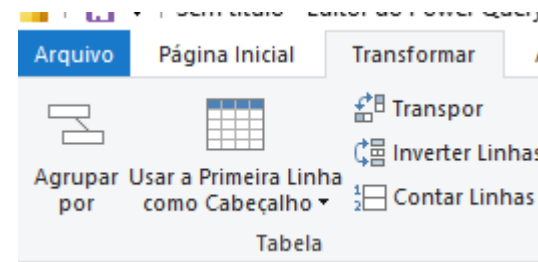
Resumo de Informações

Fornece uma visão geral dos dados, destacando tendências e padrões.

Flexibilidade

Permite agrupar dados de diversas maneiras, aplicando diferentes funções de agregação conforme necessário.

Essa funcionalidade é essencial para criar relatórios e dashboards mais eficientes e informativos no Power BI.



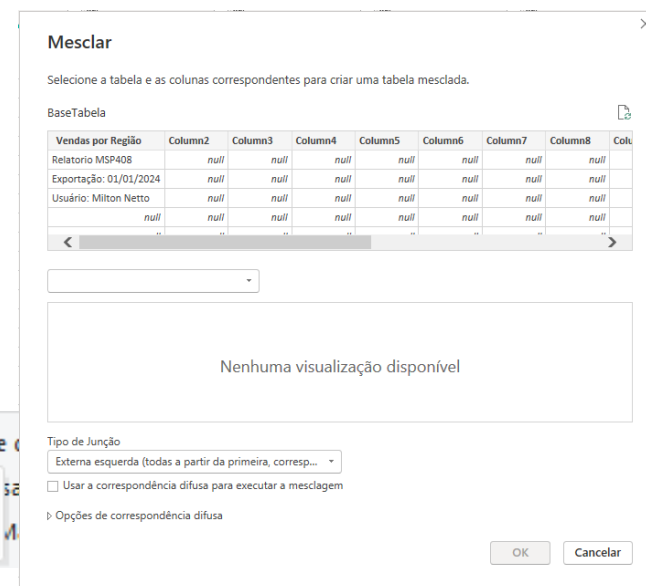
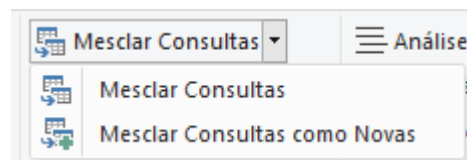
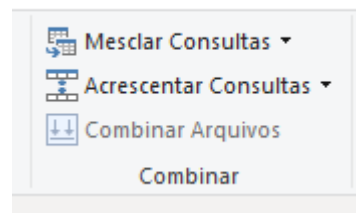
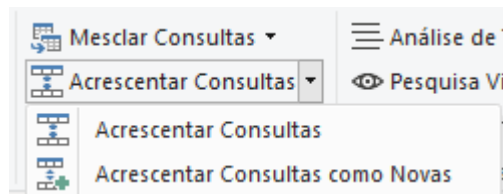
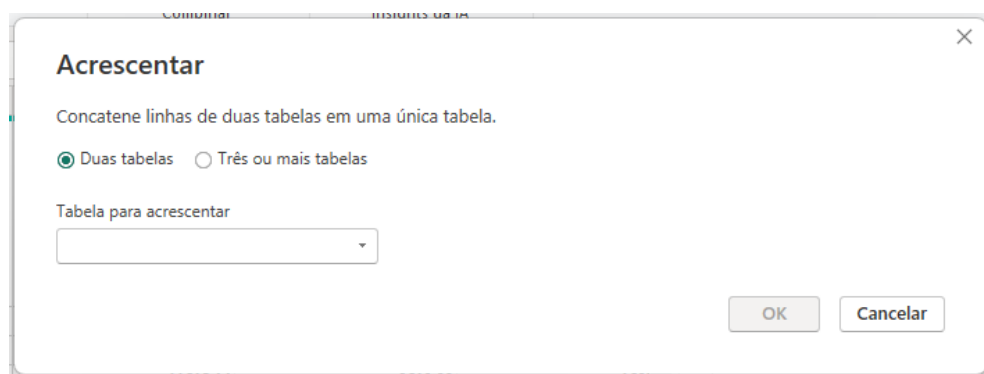
[Conceitos de Combinação de Dados]

O Power BI, o Editor de Consultas é uma ferramenta poderosa que permite transformar e combinar dados de várias fontes. Aqui estão os passos para Mesclar, Acrescentar e Combinar arquivos.

Mesclar Consultas – Mesclar consultas é útil quando você deseja combinar colunas de duas ou mais tabelas com base em uma coluna comum.

Acrescentar Consultas - Acrescentar consultas é útil quando você deseja empilhar dados de duas ou mais tabelas com a mesma estrutura.

Combinar Arquivos - Combinar arquivos é útil quando você tem vários arquivos com a mesma estrutura e deseja consolidá-los em uma única tabela.



[Etapa de consulta]

Cada vez que você aplica uma transformação nos dados, como renomear uma coluna, filtrar uma linha, dividir uma célula, etc., o editor de consulta registra essa ação como uma etapa de consulta.

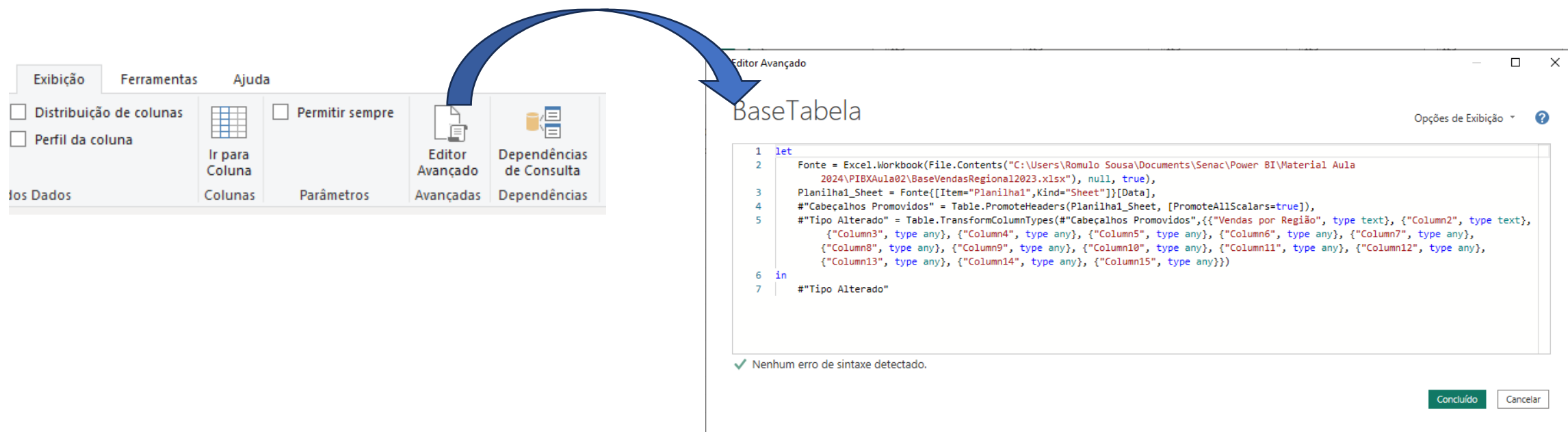
As etapas de consulta são listadas no painel Configurações de Consulta e podem ser editadas, reordenadas ou excluídas conforme necessário.



[Editor de Consulta da Linguagem M]

A linguagem do Power Query M ou Linguagem M é a estrutura de programação usada pelo editor de consulta para definir as consultas e as transformações.

Cada etapa de consulta é traduzida em uma expressão da linguagem M, que pode ser vista na barra de fórmula ou no editor avançado. É possível escrever ou modificar as expressões da linguagem M diretamente para obter resultados personalizados.





[Contexto da Atividade]

O objetivo da atividade consiste em conhecer algumas técnicas de tratamento da modelagem de dados utilizando as ferramentas de obtenção de dados do Power Query.

Assim, teremos os seguintes cenários:

Tratamento de inconsistência

Tabelas desnormalizadas

Relatório em arquivos separados

Incrementação e exclusão de dados (automático)

Mesclar fatos relacionados

Separar ou juntar elementos em colunas

Relatórios em tabelas separadas

Detalhamento de datas (extração de formas mês, Ano e dia da semana)

Classificação condicional

Conexão de dados da Web



[Atividade]

Pasta 01_FormaDemo

Técnicas: Tratamento inconsistência e Tabelas desnormalizadas

- Foi enviada uma base de dados na qual o tratamento de dados está direcionado a correção de intervalos mesclados e intervalos vazios na tabela.
- Técnica a ser aplicada:
 - Remover cabeçalho sem nexos
 - Preencher para baixo compôs sem valor
 - Usar primeira linha como cabeçalho
 - Filtrar linhas utilizando a opção do filtro NÃO CONTÉM
 - Remover colunas com valor null (caso ocorram) e colunas não contabilizadas
 - Pivotar tabela que será a transformação de colunas em linhas

Pasta 02_VendasHistorica

Técnicas: Relatório em arquivos separados, incrementar e exclusão de dados (automático)

- Técnica aplicada: **(somente para arquivos csv)**
 - Preparar arquivos e salvando em pasta
 - **Carregamento de arquivo em pasta**
 - Clique no controle duplo da coluna Content para carregar o conteúdo.
 - Visualizar o conteúdo a ser expandido utilizando o controle
 - Selecionar todos os dados e expandir.
 - Remover e incluir arquivos caso houver necessidade.
- Técnica aplicada: **(somente para arquivos EXCEL)**
 - Repetir o mesmo processo.



[Atividade]

Pasta 03_DemonsVendas

Técnicas: Mesclar fatos relacionados e Separar ou juntar elementos em colunas

- Técnica aplicada:
 - Consiste em juntar e separa conteúdo transformando em novos valores agregados
 - Prepara a visão para o relacionamento de valores
 - Identificar qual o tipo de Combinação a ser utilizada: Acrescentar e entender o que é
 - Diferenciação entre DUPLICAR e REFERÊNCIA.

Técnicas: Relatórios em tabelas separadas

- Conectar o arquivo no Power Query e visualizar seu carregamento
- Realizar carregamento em tabelas separadas
- Efetuar combinação por acréscimo
- Somente subir para o modelo de dados a base resultante da combinação.

Pasta 04_BaseAlunos

Técnica: Detalhamento de datas e Classificação condicional

- Obtenção da base de dados utilizando o método de carregamento em pasta
- Higienização dos dados e corrigir possíveis inconsistências.
- Criação de cálculo de tempo com o formato de data (identificação das referências temporais)
- Criação de coluna condicional segmentando a idade por faixa etária



[Atividade]

Técnica: Coleta de dados da Web

Cotação do dólar:

<https://www.bcb.gov.br/estabilidadefinanceira/historicocotacoes>

<https://br.investing.com/currencies/usd-brl>

<https://wise.com/br/currency-converter/dolar-hoje>

<https://www.gov.br/receitafederal/pt-br/assuntos/orientacao-tributaria/declaracoes-e-demonstrativos/ecf/taxas-de-cambio-incluindo-valor-do-dolar-para-fins-fiscais-irpj-AC-anteriores>

[Atividade Desafio]

A	B	C	D	E	F	G
Data da Venda	NF de Venda;Id CT Distribuição;Centro de Distribuição;UF Ct	Id Vendedor;Vendedor	Id Produto Categoria do Produto Descrição do Produto	Preço de Custo	Quantidade	Modo de PGTO
28/05/2023	2646-84312SP;2;São Paulo;SP	1:Marcos Brito	11 Desktop Desktop Dell I5, 8gb	R\$ 1.450,00	4	Cartão de Crédito
17/02/2022	13-95993RJ;1;Rio de Janeiro;RJ	3:Angela Martins	8 Mouse Mouse Microsoft XL29	R\$ 59,00	8	PIX
20/05/2023	3556-48326RJ;1;Rio de Janeiro;RJ	3:Angela Martins	5 Laptop Notebook Avell i7, 64gb de Ra	R\$ 4.600,00	4	Cartão de Crédito
15/05/2023	3145-83560RJ;1;Rio de Janeiro;RJ	2:Marcela Soares	12 Desktop Desktop Positivo, Dual Core,	R\$ 1.100,00	9	Cartão de Crédito

Pasta 05_Atividade_Desafio

Técnica: Analise a base e implemente a técnica

Desafio a atingir:

- Uma base em Excel foi encaminhada e seu desafio será organizar os campos. Observe que temos campos conectados em uma só coluna com: NF de Venda;Id CT Distribuição;Centro de Distribuição;UF Ct. Assim seu objetivo será separar estes campos com os seus respectivos registros.
- Com todas as colunas contendo seus respectivos registros, realize a combinação dessas bases criando uma única.
- Finalize carregando somente para o modelo de dados a base de dados a qual será realizada a análise.
- Na área de relatório aplique visuais que o usuário possa identificar resultados como Total de Valor Vendido por UF, Vendedor ou Categoria de Produto. Pode explorar sua criatividade, mas lembrando que nesse momento não há necessidade de aplicação de cálculo DAX, somente pegar, arrastar e mudar o contexto do visual.