



Centro Federal de Educação Tecnológica de Minas Gerais
Campus Nepomuceno

Tratamento de Dados no Power BI

Profa. Rosana Massahud

Colaboração:

Brian Gabriel Silva Assis

Charles Silva

2025

Sumário

1. Introdução	2
2. Apresentação do Projeto e Importação dos Dados	2
2.1. Material do curso	2
2.2. Instalação do Power BI	2
3. Conexão com os datasets	4
3.1. Conexão ao Excel	4
3.2. Conexão ao CSV	5
3.3. Conexão ao MySQL	5
3.4. Conexão ao JSON	6
4. Power Query	7
4.1. Extrair textos	8
4.2. Tradução dos textos	10
4.2.1. Tipos de junção na mesclagem do Power BI	12
4.3. Transformações	13
4.3.1. Substituição direta	13
4.3.2. Adição de coluna	14
4.3.3. Tratamento da coluna com a linguagem M	16
4.4. Tabela originária do arquivo JSON	17
4.5. Gerenciar parâmetros	19
4.6. Etapa de alteração de Tipos de Dados	20
4.7. Excluir colunas	22
4.8. Renomear colunas	23
4.9. Editor Avançado	23
4.10. Consulta nula	24
4.11. Desafio	25
Referências	26
Apêndice	27
Apêndice A – Importação de dados da Web	27
Apêndice B – Bases de Dados	30
Apêndice C – Qualidade da Coluna	31
Apêndice D – Múltiplos pequenos	35
Apêndice E – Empilhada vs Clusterizada	37
Apêndice F – Como criar conta no Power BI	38
Apêndice G – Preparação do Ambiente MySQL Workbench	47
Apêndice H – Criação do banco de Dados no MySQL	54

1. Introdução

O foco desse curso é entender como podemos nos conectar a algumas bases de dados, dado o nosso projeto de conectar conjuntos de dados da olist. Vamos nos conectar a extensões do tipo excel, .csv, .json, e a um banco de dados MySQL.

A partir desse momento, começaremos a realizar todos os nossos tratamentos, conhecendo os recursos que o editor do Power Query, dentro do Power BI Desktop, possui para fazermos as transformações.

Aprenderemos a substituir e tratar valores, além de renomear e excluir colunas e tabelas, e trabalhar com parâmetros. Vamos aprender a utilizar o editor avançado, para entender como funcionam todas as etapas que aplicaremos ao projeto.

Também será abordado como transpor tabelas, de modo que elas tragam uma disposição que faça mais sentido para o nosso resultado final.

Por fim, seguiremos com a refatoração de todas as etapas aplicadas, visando usar todas as melhores práticas para obter um resultado mais coerente com o mercado, facilitando a análise de outras pessoas.

Esse é um curso totalmente focado em **tratamento de dados**, em vez de um olhar sobre a construção de dashboard e a utilização da linguagem DAX. Não abordaremos nenhum desses conceitos, então é importante que você tenha o Power BI instalado na sua máquina!

2. Apresentação do Projeto e Importação dos Dados

2.1. Material do curso

Durante o curso, iremos **tratar dados** utilizando uma base de dados contendo informações sobre um **e-commerce**. Você pode encontrar o material através deste [link](#).

2.2. Instalação do Power BI

1 - Acesse a [página de download](#) do Power BI.

2 - Na página de download, você encontrará diversas opções. Procure pela opção **Microsoft Power BI Desktop** e clique em **Fazer download**:



Microsoft Power BI Desktop

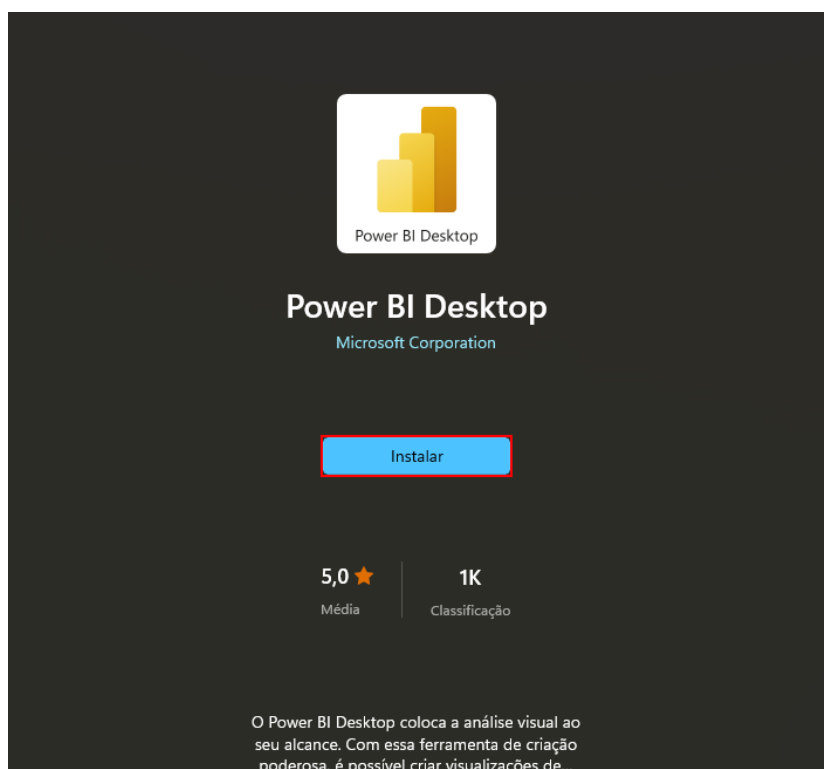
Com o Power BI Desktop, você pode explorar visualmente seus dados por meio de uma tela do tipo "arrastar e soltar" de formato livre, uma ampla variedade de visualizações de dados modernas e uma experiência de criação de relatórios fácil de usar.

Fazer download >

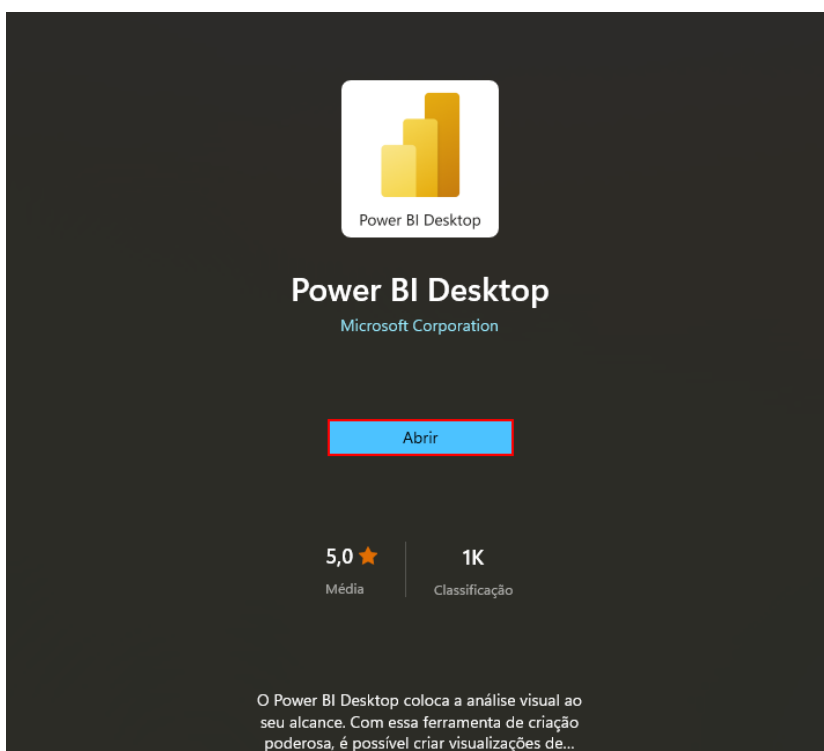
Opções de download avançadas >

3 - Após essa ação, você será redirecionado para uma página em branco, onde será solicitada a abertura da loja da Microsoft. Clique em **Abrir Microsoft Store**.

4 - Com a página inicial do Power BI Desktop na loja da Microsoft aberta, você pode clicar em **Instalar**:



5 - Nessa etapa, é necessário aguardar a instalação ser finalizada. Após a instalação ser concluída, você pode clicar em **Iniciar**:



6 - Pronto! Finalizamos a instalação do Power BI Desktop. Agora você já pode utilizá-lo e dar sequência às atividades do curso.

3. Conexão com os datasets

3.1. Conexão ao Excel

Com o Power BI aberto, vamos acessar a opção "Obter dados", que aparece para nós na janela central ao inicializar o programa. Feito isso, será aberta uma nova janela, que possui uma gama de possibilidades de conexão. Uma delas é justamente a "Pasta de Trabalho do Excel".

Vamos dar um duplo clique sobre essa opção. Será aberta uma janela onde vamos procurar o arquivo desejado. Acessaremos o arquivo olist_pedidos.xlsx.

Automaticamente, o Power BI fará a conexão com a base, visualizando o que existe nessa base de dados. Isso é feito por meio da janela "Navegador". Na lateral esquerda dessa janela, teremos duas seções:

1. O arquivo olist_orders_dataset;
2. E o arquivo olist_orders_dataset1.

Se clicarmos uma única vez sobre a primeira opção, por exemplo, teremos uma pré-visualização à direita de como está a tabela, sendo possível visualizar todas as colunas que a compõem.

Exibição das 2 primeiras linhas da tabela. Para visualizá-la na íntegra, acesse o arquivo na sua máquina.

olist_orders_dataset

orders list	Column2	Column3	Column4	Column5	Column6	Column7	Column8
date: 09/2016 - 09/2018		null	null	null	null	null	null
order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_estimated_delivery_date
e481f51cbdc54678b7cc49136f2d6af7	9ef432eb6251297304e76186b10a928d	#delivered	02/10/2017 10:56:33	02/10/2017 11:07:15	04/10/2017 19:55:00	10/10/2017 21:25:13	18/10/2017 00:00:00
53cdb2fc8bc7dce0b6741e2150273451	b0830fb4747a6cd20dea0b8c802d7ef	#delivered	24/07/2018 20:41:37	26/07/2018 03:24:27	26/07/2018 14:31:00	07/08/2018 15:27:45	13/08/2018 00:00:00

Clicando sobre a segunda opção, teremos a pré-visualização da mesma tabela, porém com uma organização diferente. Nesse cenário, qual seria a diferença entre elas?

Exibição das 2 primeiras linhas da tabela.

olist_orders_dataset1

order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_estimated_delivery_date
e481f51cbdc54678b7cc49136f2d6af7	9ef432eb6251297304e76186b10a928d	#delivered	02/10/2017 10:56:33	02/10/2017 11:07:15	04/10/2017 19:55:00	10/10/2017 21:25:13	18/10/2017 00:00:00
53cdb2fc8bc7dce0b6741e2150273451	b0830fb4747a6cd20dea0b8c802d7ef	#delivered	24/07/2018 20:41:37	26/07/2018 03:24:27	26/07/2018 14:31:00	07/08/2018 15:27:45	13/08/2018 00:00:00

Na primeira opção, temos um detalhe logo nas primeiras linhas: as colunas estão sem nome, indicadas como Column2, Column3, e assim por diante; e existem informações a mais na parte superior da tabela, logo nas primeiras linhas.

Já na segunda opção, aparentemente, temos o cabeçalho das colunas nomeados da maneira esperada, além de haver um tratamento das linhas.

Porém, como não sabemos como funcionam os pormenores desses tratamentos, vamos selecionar a primeira opção, para depois fazermos os tratamentos por conta própria.

Selecionada a opção desejada (olist_orders_dataset), teremos 3 opções no canto inferior direito da janela "Navegador": "Carregar", "Transformar Dados", e "Cancelar". Nesse caso, vamos selecionar a opção "Carregar".

Feito isso, será aberta uma nova janela que fará a carga dos dados, realizando finalmente a conexão à base de dados.

Finalizada a carga, teremos o dataset importado na seção "Dados", localizada na lateral direita da interface. A partir desse momento, temos a nossa primeira base.

Na sequência, faremos a conexão à segunda base, de extensão .csv.

3.2. Conexão ao CSV

Começaremos acessando novamente a opção "Obter dados", localizada na seção "Dados", à esquerda na guia superior da ferramenta. Ao clicar sobre ela, será aberto um menu suspenso contendo a opção "Texto/CSV". Vamos selecionar essa opção e, em seguida, dar um duplo clique sobre o arquivo `olist_itens_pedidos.csv`.

Assim como no arquivo Excel, teremos acesso a uma pré-visualização da tabela, porém com configurações um pouco diferentes.

Exibição das 3 primeiras linhas da tabela.

`olist_itens_pedidos.csv`

Origem do Arquivo		Delimitador	Detecção de Tipo de Dados			
1252: Europeu Ocidental (Windows) ▾		Vírgula ▾	Com base nas primeiras 200 linhas ▾			
order_id	order_item_id	product_id	seller_id	shipping_limit_date	price	freight_value
00010242fe8c5a6d1ba2dd792cb16214	1	4244733e06e7ecb4970a6e2683c13e61	48436dade18ac8b2bce089ec2a041202	19/09/2017 09:45:35	589	1329
000229ec398224ef6ca0657da4fc703e	1	c777355d18b72b67abbef9df44fd0fd	5b51032eddd242adc84c38acab88f23d	18/01/2018 14:48:30	1990	1787
00024acbcdf0a6daa1e931b038114c75	1	7634da152a4610f1595efa32f14722fc	9d7a1d34a5052409006425275ba1c2b4	15/08/2018 10:10:18	1299	1279

Por ser um arquivo de extensão .csv, ele é separado por **vírgula** e precisa de um **delimitador**. Na parte superior da janela de carregamento, temos 3 seções que compõem as configurações específicas:

- "Origem do Arquivo", para saber a codificação, pensando em caracteres especiais;
- "Delimitador", que é justamente a vírgula que separa as colunas;
- E "Detecção de Tipo de Dados". Nesse caso, é feita uma detecção dos tipos de cada coluna e uma formatação baseada em como os dados se comportam.
 - **Por exemplo:** se forem identificados apenas números em determinada coluna, ela será definida como numérica.

Podemos realizar a carga clicando no botão "Carregar".

Feito isso, teremos nossas duas bases importadas com êxito.

3.3. Conexão ao MySQL

Para executar seção é necessário fazer os procedimentos descritos nos Apêndices G e H dessa apostila.

Para isso, vamos clicar novamente sobre a opção "Obter dados" na barra de menu superior, e dessa vez selecionar a opção "Mais..." para visualizar todas as fontes. Será aberta a janela "Obter Dados", onde vamos explorar a seção "Banco de Dados", terceira opção na lateral esquerda.

A princípio, identificamos várias formas de fazer a conexão. Por exemplo:

- Banco de dados SQL Server;
- Banco de dados MySQL;
- Banco de dados PostgreSQL...

Nesse caso, iremos acessar o **banco de dados MySQL**. Ao clicar duas vezes sobre a opção, seremos redirecionados para uma nova janela, onde precisaremos definir o servidor e o banco de dados disponível, que é justamente a tabela que queremos importar.

Preencheremos os campos da seguinte maneira:

- **Servidor:** localhost
- **Banco de Dados:** powerquery

É muito importante que você faça essa estrutura, colocando a base e organizando todo o banco de dados do MySQL, para então conseguir fazer a conexão.

Além desses dois campos, existe uma seção de opções avançadas. Para pessoas que desejem fazer uma query, isto é, uma consulta mais específica dentro do banco de dados, é possível utilizar essa seção, que contém o campo "Instrução SQL", onde conseguimos incluir todos os comandos para fazer a consulta.

No nosso caso, vamos nos conectar a toda a tabela disponível no banco. Então, podemos clicar em "OK" para confirmar a ação e fazer a conexão.

Finalizado o processo, será aberta a janela "Navegador", trazendo o arquivo powerquery.olist_pagamentos na pasta "localhost: powerquery". Da mesma forma que fizemos anteriormente, podemos ter uma pré-visualização da tabela:

powerquery.tb_olist_pagamentos
Visualização baixada em terça-feira, 11 de junho de 2024

order_id	payment_installments	payment_sequential	pay
b81ef226f3fe1789b1e8b2acac839d17	8	1	1
a9810da82917af2d9aefd1278f1dcfa0	1	1	1
25e8ea4e93396b6fa0d3dd708e76c1bd	1	1	1
ba78997921bbcdc1373bb41e913ab953	8	1	1
42fdf880ba16b47b59251dd489d4441a	2	1	1
298fcd1f73eb413e4d26d01b25bc1cd	2	1	1
771ee386b001f06208a7419e4fc1bbd7	1	1	1
3d7239c394a212faae122962df514ac7	3	1	1
1f78449c87a54fa9e9e88ba1491fa9	6	1	1
0573b5e23cbd798006520e1d5b4c6714	1	1	1
d88e0d5fa41661ce03cf6c336527646	8	1	1
2480f727e869fdeb397244a21b721b67	1	1	1
616105c9352a9668c38303ad44e056cd	1	1	1
cf95215a722f3ebf29e6bbab87a29e61	5	1	1
769214176682788a92801d8907fa1b40	4	1	1
12e5cfe0e4716b59afb0e0f4a3bd6570	10	1	1
61059985a6fc0ad64e95d9944caacdad	1	1	1
79da3f5fe31ad1e454f06f95dc032ad5	1	1	1
8ac09207f415d55acff302df7d6a895c	4	1	1
b2349a3f20dfbeef62e7b31baa22f84b	3	1	1
5a1f6d22f7dfb061ef29216b9af687a1	4	1	1
4214cda550ece8ee66441f459dc33a8c	2	1	1

Por fim, vamos selecioná-la, assim como fizemos com o Excel, e clicar em "Carregar". Feito o carregamento da terceira base, vamos seguir para a última: o arquivo de extensão .json.

3.4. Conexão ao JSON

Começaremos clicando novamente em "Obter dados". Na janela que será aberta, selecionaremos a quarta opção, "JSON". Após clicar sobre ela duas vezes, vamos selecionar o arquivo olist_produtos.json.

A conexão será feita de forma um pouco diferente dos outros arquivos. Nesse caso, é aberta uma janela contendo um novo nome na parte superior: "Editor do Power Query".

Esse é o ambiente em que passaremos a maior parte do nosso tempo durante o curso, para fazer os tratamentos dos dados.

O arquivo do tipo JSON já traz isso para nós, ou seja, nos direciona para essa janela, pois sabe que precisam ser feitas algumas modificações para obter uma tabela bem-estruturada.

Precisamos conhecer essa área do Power BI, para depois seguir com os tratamentos.

4. Power Query

Estamos no Power Query, o ambiente no qual faremos todas as transformações necessárias nos dados para nosso projeto.

Vamos conhecer as principais guias que utilizaremos. Acessando a barra de menus superior, veremos as seguintes guias, da esquerda para a direita:

- Arquivo
- Página inicial
- Transformar
- Adicionar Coluna
- Exibição
- Ferramentas
- Ajuda
- Converter

Na primeira guia importante, "Página inicial", temos as seguintes opções relevantes, da esquerda para a direita:

- "Fechar e aplicar", que fecha o Editor de Consultas e aplica as modificações pendentes, levando-as para o *Power BI Desktop*
- "Nova Fonte", equivalente ao "Obter dados", onde realizamos as conexões
- "Fontes Recentes", que exibe as fontes recentemente utilizadas
- "Inserir Dados", onde podemos criar uma nova tabela personalizada
- "Atualizar visualização", que traz as atualizações que fizemos na base
- "Editor Avançado", que traz informações sobre os tratamentos realizados

O Power Query está sempre conectado à base e observando o caminho de conexão, o que permite acionarmos o botão "Atualizar visualização".

Além da guia "Página Inicial", temos outras com recursos igualmente importantes. Acessando a guia "Transformar", veremos dispostos de maneira intuitiva todos os recursos trazidos pelo Editor do Power Query que servem para transformar linhas, colunas e tabelas.

À direita de "Transformar", clicaremos na guia "Adicionar Coluna", que possui ferramentas para criar colunas no interior de tabelas que já existem, baseadas em interações com a base de dados ou em transformações desejadas.

Uma das mais importantes guias é a "Exibição". Ela nos traz recursos que facilitarão as análises das bases, como, por exemplo, a opção "Barra de Fórmulas", que possui uma caixa de seleção clicável. Vamos selecioná-la agora.

Com a seleção, surgirá uma barra no centro superior da tela, acima das colunas de dados. Ela possui dois campos: à esquerda, três botões, e à direita um campo de texto que exibe um código. Este último corresponde ao código que trabalha internamente.

À direita do botão "Barra de Fórmulas", temos três botões com caixas de seleção enfileirados verticalmente: "Espaçamento uniforme", "Mostrar espaço em branco" e "Qualidade da coluna". Este último se encontra não selecionável, pois a conexão que fizemos com o JSON não está configurada como tabela ainda.

Para configurá-la como tal, acessaremos o Editor de Consultas lateral e buscaremos o arquivo olist_orders_dataset, no qual realizaremos um clique duplo e renomearemos para olist_pedidos. Com o novo nome, poderemos nos orientar melhor.

Assim que selecionarmos este arquivo, a opção "Qualidade da coluna" se tornará selecionável. Vamos clicar nela, e com isso, veremos no centro da tela, além das colunas do olist_orders_dataset, um campo em cada coluna, logo abaixo dos títulos, que exibe três informações:

- Um círculo verde com o texto "Valores válidos"
- Um círculo vermelho com o texto "Erro"
- Um círculo preto com o texto "Vazio"

À direita de cada informação, temos um percentual que corresponde à quantidade de cada tipo de valor que existe naquela coluna. Com isso, conseguimos entender minimamente como os dados desta base se comportam.

Observando a tela inteira do Power BI, com toda a nossa interação exposta, podemos recapitular os elementos vistos:

- A barra superior com guias de funcionalidades
- O Editor de Consultas lateral, onde alteramos o nome do arquivo
- A tabela na parte central, onde visualizamos e alteramos dados

Além deles, temos a aba lateral de "Configurações de Consulta", onde temos duas seções: acima, a "Propriedades", e abaixo, a "Etapas Aplicadas".

Esta última corresponde às rotinas de tratamento que estamos configurando para a tabela através do Editor do Power Query para que ele rode sempre que fizermos a carga dos dados. Ou seja, ele se conecta à base de dados local, executa essas etapas e traz o resultado do tratamento.

Vamos conhecer melhor essas etapas durante o curso.

Após conhecermos o Editor do Power Query e aprendermos a localizar as ferramentas para interagir e trabalhar, seguiremos para o nosso primeiro tratamento.

Nesta que é nossa primeira tabela, já podemos analisar alguns detalhes, como, por exemplo, a coluna orders list que traz informações diferentes.

Em sua primeira linha, temos uma informação que não possui dados nas outras colunas. Já na segunda linha, todas as colunas possuem informações, mas parecem pertencer não ao corpo da tabela, mas ao cabeçalho de cada coluna.

4.1. Extraindo textos

Começaremos os tratamentos voltando ao Editor do Power Query e analisando a primeira tabela: "olist_pedidos".

Nela, temos a primeira coluna, "orders list", à esquerda, que traz a lista de pedidos. Suas duas primeiras linhas possuem informações diferentes do padrão das linhas restantes.

Se analisarmos as outras colunas à direita, veremos que nenhuma possui um cabeçalho ou nome específico, mas sim nomes genéricos como "Column2" (coluna 2). Na linha 1 de cada coluna, temos um valor nulo e na linha 2, veremos os nomes que deveriam ser os cabeçalhos.

Precisamos eliminar a primeira linha, acessar a segunda e configurar seu conteúdo como o cabeçalho.

Erros como cabeçalhos mal estruturados e valores nulos são comuns de se encontrar quando realizamos importações na base de dados. Precisamos realizar uma limpeza para ajustá-los.

Vamos começar removendo a primeira linha. Para isso, clicaremos em um botão em formato de tabela, à esquerda do cabeçalho da primeira coluna. Com esse clique, será exibido um menu suspenso com várias opções, dentre as quais selecionaremos "Remover Linhas Superiores".

Após a seleção, será aberta uma janela modal denominada "Remover Linhas Superiores" que possui um campo editável. Nele, devemos definir quantas linhas serão removidas da nossa, de cima para baixo.

Vamos remover somente a primeira, portanto digitaremos "1" e selecionaremos o botão "OK". Com isso, a primeira linha desaparecerá da nossa base, junto ao valor de data e os valores nulos presentes nela.

Vamos promover a nova linha 1 (antiga linha 2) para se tornar o cabeçalho. Para isso, clicaremos no mesmo botão com o ícone de tabela e selecionaremos no menu suspenso a opção "Usar a Primeira Linha como Cabeçalho".

Após a seleção, o conteúdo da linha 1 será visto no cabeçalho, substituindo os nomes anteriores.

Já realizamos dois tratamentos. Se lembrarmos quando exportamos o Excel, estas etapas já haviam sido realizadas como sugestões do próprio Power Query e nós as ignoramos porque queríamos entender como elas funcionam manualmente.

Agora sabemos como remover linhas e dispor o cabeçalho da melhor maneira possível.

Vamos analisar a terceira coluna da nossa base de dados, intitulada "order_status" (situação do pedido). Ela possui informações com uma tralha (#) à esquerda. Este símbolo deve ser removido, mas antes, vamos verificar se todas as linhas o possuem.

Para isso, iremos até o cabeçalho dessa coluna e clicaremos no botão à sua direita, constituído de uma caixa branca com uma seta apontando para baixo. Com o clique, será exibida uma janela com as seguintes opções de interação, de cima para baixo:

- "Classificar em ordem crescente"
- "Classificar em ordem decrescente"
- "Remover Vazio"
- "Filtros de Texto", acompanhado de um campo de pesquisa
- Um contêiner com uma lista de elementos, acompanhadas de caixas de seleção à esquerda.

Queremos verificar quais valores foram registrados nesta coluna. Para isso, acessaremos o contêiner, que nos fornece o nome de todos os *status* adicionados nesta coluna. Nessa lista, veremos que todos os dados possuem uma tralha à sua esquerda.

- #canceled
- #delivered
- #invoiced
- #processing
- #shipped
- #unavailable

Vamos fechar esta janela clicando no botão "Cancelar", no canto inferior direito.

Precisamos realizar um tratamento que verificará todas essas linhas e removerá este símbolo. Para isso, acessaremos a barra de menus superior e selecionaremos a guia "Transformar". Nela, selecionaremos a opção "Extrair", praticamente no centro dessa guia, o que abrirá um menu suspenso.

Neste, teremos sete opções:

- Duração
- Primeiros caracteres
- Últimos caracteres
- Intervalo
- Texto Antes do Delimitador
- Texto Após o Delimitador
- Texto Entre os Delimitadores

Nosso delimitador será a tralha, e queremos que seja exibido somente o que vier depois dela. Portanto, selecionaremos a opção "Texto Após o Delimitador", o que abrirá uma janela modal de mesmo nome, que possui em seu interior um campo editável chamado "Delimitador".

Neste campo, digitaremos o delimitador, ou seja, a informação a ser ignorada. Portanto, digitaremos "#" e pressionaremos o botão "OK".

Com isso, o símbolo # desaparecerá da coluna "order_status", tornando-a estruturada da forma desejada. Além disso, se clicarmos novamente no ícone de seta para baixo e visualizarmos o contêiner com os dados da coluna, veremos que nenhuma possui o símbolo.

- canceled
- delivered
- invoiced
- processing
- shipped
- unavailable

Com estas informações prontas, seguiremos para outra questão: e se quiséssemos traduzir as informações de *status* que estão em inglês para um público que fale português? Como trataremos a coluna para exibir as informações no formato desejado?

A seguir, trataremos sobre essa questão.

4.2. Tradução dos textos

Existem várias formas de traduzir a coluna "order_status":

- Substituir cada um dos valores dentro do Power BI, dado que conhecemos todos eles
- Alterar dentro da própria base

Vamos conhecer e utilizar um recurso diferente, que pode ser útil em outras soluções e em outros projetos: a **Mesclagem de Consultas**.

O primeiro passo é importar um arquivo com as informações da tradução. Para isso, acessaremos a barra de menus superior e selecionaremos a guia "Página Inicial". Em seu interior, selecionaremos "Nova Fonte".

Com isso, será exibida a janela modal "Obter Dados" em cuja lista selecionaremos a opção "Texto/CSV" para importar um arquivo do tipo CSV. Abaixo da lista, selecionaremos o botão "Conectar".

A janela "Obter Dados" será fechada e outra será aberta, exibindo o explorador do Windows. Nela, veremos um arquivo denominado "tradução", no qual faremos um clique duplo.

Com o clique, uma janela modal com o nome do arquivo será aberta, exibindo o conteúdo do arquivo em duas colunas, lado a lado:

- "Column1", onde a linha 1 possui o texto "Inglês" e as restantes exibem as informações de *status* em inglês
- "Column2", onde a linha 1 possui o texto "Português" e as restantes exibem as informações de *status* em português

No canto inferior direito dessa janela, selecionaremos o botão "OK". Com isso, o arquivo será carregado na área de visualização de dados.

Vamos promover somente a linha 1 a Cabeçalho, substituindo o anterior ("Column1" e "Column2") por "Inglês" e "Português".

Vamos realizar a mesclagem. Para isso, consultaremos a visualização do "`olist_pedidos`". Nela, selecionaremos a coluna "order_status", clicando sobre seu cabeçalho.

Dentro da guia "Página Inicial", buscaremos a opção "Mesclar consultas" e clicaremos na seta que aponta para baixo à sua direita. Com o clique, um menu suspenso será aberto, com duas opções: "Mesclar Consultas" e "Mesclar Consultas como Novas", dentre as quais selecionaremos a primeira.

A segunda opção geraria uma nova tabela com a mescla, mas neste caso, queremos aproveitar a mesma.

Clicando em "Mesclar Consultas", uma janela modal chamada "Mesclar" será aberta, onde escolheremos as duas tabelas que serão unidas.

Abaixo do título "Mesclar", um contêiner exibindo as colunas da tabela "`olist_pedidos`", pois ela já foi selecionada. Abaixo deste contêiner, temos um botão. Ao clicarmos nele, uma lista suspensa será exibida, na qual selecionaremos a tabela a ser mesclada com a "`olist_pedidos`" — ou seja, a tabela "tradução".

Com isso, será exibido um novo contêiner abaixo da opção selecionada, onde serão exibidas as colunas dessa tabela.

Para selecionar quais colunas constituirão um vínculo, basta clicar sobre elas. Faremos isso na coluna "order_status" da tabela "`olist_produtos`" e na coluna "Inglês" da tabela "tradução".

Abaixo do contêiner da tabela "tradução", serão exibidas as seções "Tipo de junção", "Opções de correspondência difusa" e abaixo delas, um ícone de check verde junto a um texto informando que a seleção de ambas as tabelas são correspondentes.

A seleção corresponde a 99441 de 99441 linhas da primeira tabela.

Isso confirma que a junção será perfeita.

Se acessarmos a seção "Tipos de junção", veremos um botão. Se clicarmos nele, serão exibidas as opções dos vários tipos de junção disponíveis:

- Externa esquerda (todas a partir da primeira, correspondência a partir da segunda)
- Externa direita (todas a partir da segunda, correspondência a partir da primeira)
- Externa Completa (todas as linhas de ambos)
- Interna (apenas linhas correspondentes)
- Anti esquerda (linhas apenas na primeira)
- Anti direita (linhas apenas na segunda)

Neste caso, escolheremos a opção "Externa esquerda" e clicaremos no botão "OK", no canto inferior direito da janela.

Com isso, a nossa tabela recebeu uma nova coluna, "tradução", à direita. Ela possui linhas com o texto "Table", em verde.

Precisamos trazer a informação traduzida. Para isso, acessaremos o cabeçalho desta coluna e clicaremos no botão quadrado à sua direita, que possui uma seta apontando para baixo.

Na janela suspensa que será aberta, teremos os botões rádio "Expandir" e "Agregar", lado a lado, abaixo dos quais haverá uma lista com duas caixas de seleção exibindo as opções "Selecionar todas as colunas", "Inglês" e "Português". Abaixo dessa lista, teremos outra caixa de seleção, exibindo o texto "Use o nome da coluna original como prefixo".

Queremos trazer somente a coluna com valores em português, portanto, marcaremos somente a opção "Português", desmarcando as outras opções da lista. Além disso, desmarcaremos a opção que utiliza o nome da coluna original como prefixo, pois senão teremos uma coluna chamada "tradução.Português" em vez de somente "Português".

Após desmarcarmos a seleção, selecionaremos o botão "OK".

Com isso, veremos que a coluna "tradução" se transformou em "Português", e exibe os dados de *status* em português. Portanto, todas as informações estão se comunicando com essa tabela.

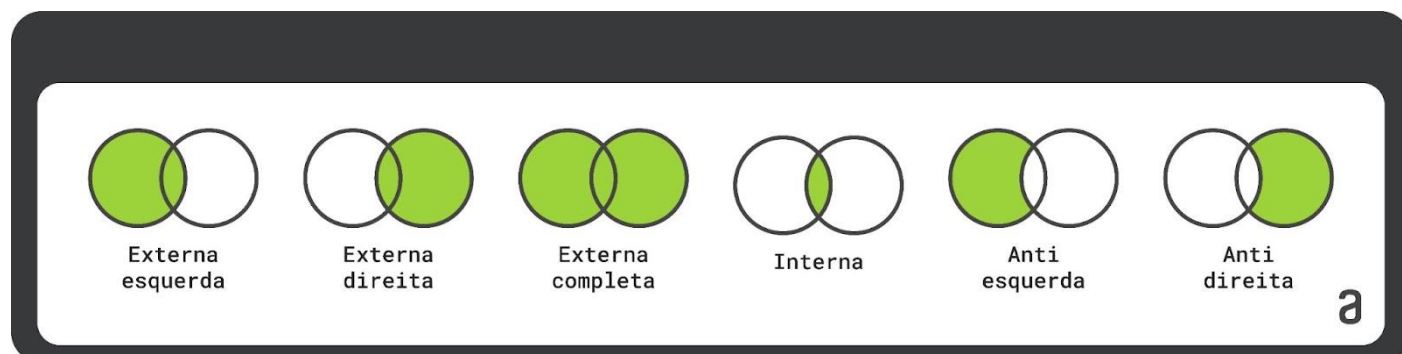
Dentro de "order_status", temos na linha 1, por exemplo, a informação "*delivered*", enquanto na coluna "Português", temos a informação "entregue".

O recurso de Mesclagem de Consultas também é interessante em situações nas quais queremos carregar juntas duas tabelas e trazê-las em uma só, além de servir para realizar interações que alteram informações de acordo com cada linha, utilizando duas tabelas.

A partir desta etapa, conseguiremos explorar outras tabelas às quais já nos conectamos, como a "olist_itens_pedidos" e a "olist_pagamentos".

4.2.1. Tipos de junção na mesclagem do Power BI

No Power BI, uma das principais funcionalidades é a capacidade de **mesclar dados de diferentes fontes** em um único conjunto de dados para análise. Essa mesclagem é possível graças ao recurso de junções, que permite combinar informações com base em critérios específicos. Existem diferentes tipos de junções disponíveis no Power BI, cada uma com suas características e finalidades. Neste texto, vamos conhecer os tipos de junções mais comuns utilizados na mesclagem de dados no Power BI.



- **Esquerda Externa (Left Outer Join):** Retorna todas as linhas da tabela esquerda e as linhas correspondentes da tabela direita com base em um critério de correspondência dos relacionamentos entre chaves primárias e estrangeiras das tabelas. Se não houver correspondência na tabela direita, os valores serão preenchidos com nulos.

- **Direita Externa (Right Outer Join):** Retorna todas as linhas da tabela direita e as linhas correspondentes da tabela esquerda com base em um critério de correspondência. Se não houver correspondência na tabela esquerda, os valores serão preenchidos com nulos.
- **Completa Externa (Full Outer Join):** Retorna todas as linhas das duas tabelas, combinando registros com base em um critério de correspondência. Se não houver correspondência em uma das tabelas, os valores correspondentes serão preenchidos com nulos.
- **Interna (Inner Join):** Retorna apenas as linhas correspondentes das duas tabelas com base em um critério de correspondência. As linhas não correspondentes são excluídas do resultado final da mesclagem.
- **Anti Esquerda (Left Anti Join):** Retorna apenas as linhas da tabela esquerda que não possuem correspondência com base em um critério de correspondência. As linhas correspondentes da tabela direita são excluídas do resultado.
- **Anti Direita (Right Anti Join):** Retorna apenas as linhas da tabela direita que não possuem correspondência com base em um critério de correspondência. As linhas correspondentes da tabela esquerda são excluídas do resultado.

Esses tipos de junções de mesclagem no Power BI são extremamente úteis para combinar e analisar dados de diferentes fontes, permitindo obter insights valiosos e tomar decisões informadas.

4.3. Transformações

Resolvemos a questão do símbolo # (tralha) dentro do *dataset* "olist_pedidos". Contudo, se clicarmos no "olist_pagamentos" por meio da aba de Consultas, veremos que uma de suas colunas possui o mesmo problema.

A coluna em questão é denominada "payment_type" e traz a forma de pagamento realizada pelas pessoas que fizeram pedidos pela Olist. Se trata da última coluna da tabela, da esquerda para a direita.

Seria possível resolver este problema de outra forma, sem mesclar consultas de uma tabela exterior? Sim.

4.3.1. Substituição direta

A primeira forma de fazer isso é realizar uma **substituição direta do valor**. Entretanto, considerando a grande quantidade de valores nesta coluna, utilizaremos um artifício para facilitar nossa tarefa.

Primeiro, clicaremos com o botão direito no título da coluna "payment_type" e, no menu suspenso que será exibido, selecionaremos a opção "Remover Duplicadas".

Com isso, a base trará somente uma ocorrência de cada valor existente naquela coluna. Podemos vê-los abaixo:

- #credit_card
- #boleto
- #voucher
- #debit_card
- #not_defined

Temos cinco tipos de valores para substituir. Vamos iniciar o processo pela "#debit_card", clicando com o botão direito em cima desse valor e selecionando, no menu suspenso, a opção "Substituir Valores".

Após o clique, será aberta a janela modal "Substituir Valores". Quando clicamos com o botão direito no valor "#credit_card", o modal já sabe que é este valor que será substituído e, portanto, preenche o campo "Valor a ser localizado" com ele.

Já no campo "Substituir por", escreveremos o valor que substituirá todas as ocorrências de "#credit_card" — ou seja, "Cartão de crédito".

Após o preenchimento, selecionaremos o botão "OK". Feito isso, a base de dados exibirá a alteração na coluna "payment_type", que ainda estará filtrada. No final do processo, retiraremos o filtro para ver todos os valores após a substituição.

Faremos o mesmo processo para os outros tipos de pagamento, efetuando as seguintes substituições:

- De "#boleto" para "Boleto"
- De "#voucher" para "Voucher"
- De "#debit_card" para "Cartão de débito"
- De "#not_defined" para "Não definido"

Após a substituição, teremos todos os valores traduzidos e sem o símbolo #.

- Cartão de crédito
- Boleto
- Voucher
- Cartão de débito
- Não definido

Para visualizar a tabela inteira sem o filtro, mas com as substituições, acessaremos a aba lateral direita, na seção "Etapas aplicadas", onde veremos todos os processos realizados na tabela. Em seu interior, excluiremos a etapa "Duplicatas Removidas", exibida acima das etapas de substituição. Este será o artifício utilizado.

Para remover a etapa, posicionaremos o cursor nela e clicaremos no ícone de "x" vermelho à sua esquerda. Com isso, uma janela modal será exibida para nos alertar sobre as consequências deste comando. Basta clicar em "Excluir" para confirmar a operação.

Com a exclusão dessa etapa, teremos a visualização completa da base de dados, com todos os valores da coluna "payment_type" substituídos da melhor maneira possível.

4.3.2. Adição de coluna

Além deste recurso, podemos efetuar a **adição de uma coluna**.

Para testá-lo, acessaremos novamente a seção "Etapas aplicadas" na aba lateral direita e removeremos os cinco processos de substituição realizados na tabela — "Valor Substituído", "Valor Substituído1", "Valor Substituído2", "Valor Substituído3" e "Valor Substituído4".

Com isso, a base de dados será exibida da forma original, com dados em inglês e que possuem tralhas.

A adição de colunas é útil quando sabemos o que queremos fazer, mas não sabemos realizar o processo. Para utilizar este recurso, acessaremos a barra superior e clicaremos na guia "Adicionar Coluna". Em seu interior, selecionaremos a primeira opção, da esquerda para a direita, denominada "Coluna de Exemplos".

Após a seleção, será exibido um contêiner cinza acima da visualização da tabela. À direita da última coluna da tabela, será exibida uma coluna cinza-claro que não possui linhas ou textos, funcionando como uma divisão. À direita desta, há uma coluna que possui o cabeçalho "Coluna1" e linhas em branco.

Este recurso seleciona todas as colunas da tabela e traz a "Coluna1" como responsável por interpretar o que queremos fazer com base no exemplo dado.

Neste momento, na base de dados, todos os cabeçalhos da tabela possuirão uma caixa de seleção à sua direita. Estes estarão selecionados, exibindo a cor verde-escuro nos cabeçalhos.

Clicaremos em cada caixa de seleção para remover esta seleção, mantendo somente a coluna "payment_type" selecionada. Com isso, registraremos na "Coluna1" a forma que gostaríamos de substituir cada tipo de ocorrência existente em "payment_type".

Na linha 1 desta coluna, temos o valor "#credit_card". Vamos acessar a linha 1 na "Coluna1" e preenchê-la com o valor que irá substituí-lo — ou seja, "Cartão de Crédito". Faremos o mesmo para a linha 2.

A partir da segunda linha preenchida, a "Coluna1" entenderá que todas as ocorrências deverão ser preenchidas com o valor "Cartão de Crédito" e sugerirá o preenchimento, mostrando o texto "Cartão de Crédito" em letras quase transparentes. A partir disso, exploraremos as exceções dessa estrutura.

Na linha 10, por exemplo, temos o valor "#boleto". Acessaremos a linha 10 da "Coluna1" e digitaremos "Boleto". Com isso, a coluna entenderá que todas as ocorrências de "#boleto" deverão ser preenchidas com o valor "Boleto" e que existem outros valores possíveis.

Para os outros tipos de valores, como, por exemplo, o "#voucher" da linha 26, ela reconhece que não será "Cartão de Crédito" e nem "Boleto", mas não sabe o valor a ser preenchido. Por isso, preenche esta célula com o valor "null".

Acessando a linha 26 da "Coluna1", vamos alterá-la de "null" para "Voucher". Com isso, serão feitos outros ajustes.

Este processo é complicado, pois temos que buscar ao menos uma ocorrência de cada valor para substituir manualmente ao longo de toda a tabela.

Para facilitar o processo, podemos utilizar o artifício de substituição dos valores, visto do método anterior. Com este, podemos remover as duplicatas antes de adicionar a nova coluna e realizar a substituição de forma mais eficiente.

Vamos cancelar o processo atual, acessando o canto inferior direito do contêiner cinza acima da tabela e clicando em "Cancelar".

Com o modo de adição encerrado, faremos o mesmo processo de antes para filtrar a coluna "payment_type" e exibir somente uma ocorrência de cada valor. Em seguida, adicionaremos novamente a coluna de exemplos, selecionando somente a coluna "payment_type".

Podemos renomear a "Coluna1" que terá os valores da substituição. Vamos clicar no cabeçalho e chamá-la de "Pagamento tratado". Em suas linhas, vamos escrever as substituições para os cinco valores exibidos.

- De "#credit_card" para "Cartão de crédito"
- De "#boleto" para "Boleto"
- De "#voucher" para "Voucher"
- De "#debit_card" para "Cartão de débito"
- De "#not_defined" para "Não definido"

Após renomear a coluna nova e definir os valores de substituição, acessaremos o contêiner cinza acima da tabela. Em seu canto inferior direito, clicaremos no botão verde "OK".

Após todo esse processo, temos uma nova coluna denominada "Pagamento tratado" com os tratamentos de tradução e remoção de tralha feitos. Vamos acessar a aba lateral direita e remover a etapa "Duplicatas removidas". Após o clique em "Excluir", veremos a base de dados completa.

Os recursos de Substituição de valores e Adição de colunas são interessantes, não somente para solucionar o problema visto na aula, como para encontrar diversas formas de tratar tabelas e dados no dia a dia. É muito importante conhecê-los.

Além deles, temos o recurso de "Coluna personalizada", visto à direita de "Coluna de Exemplos". Ele lida diretamente com a *linguagem M*, nativa do Power Query.

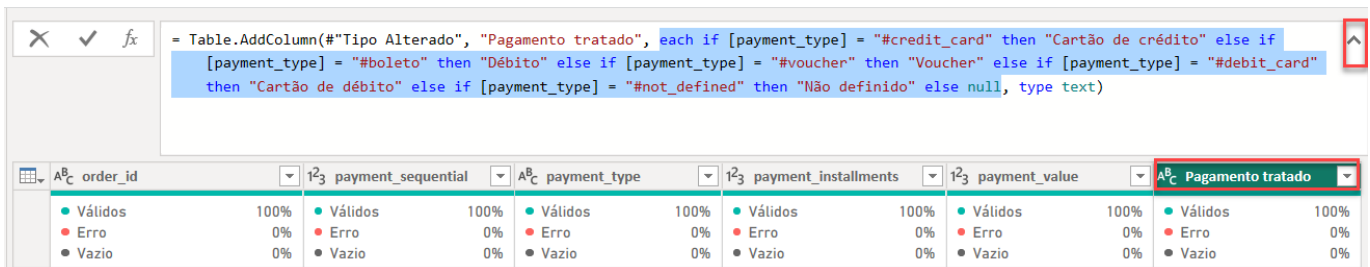
4.3.3. Tratamento da coluna com a linguagem M

Nas seções anteriores, verificamos como realizar o tratamento da coluna de Tipo de pagamentos através de uma Coluna de Exemplo, em que removemos as duplicatas, inserimos os valores desejados e voltamos com os valores duplicados novamente.

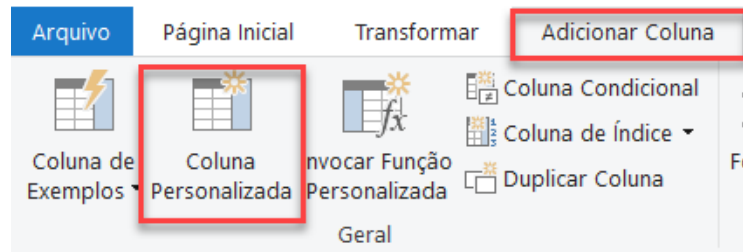
Caso desejemos realizar esse processo sem remover as duplicatas, podemos realizar esse tratamento através da **linguagem M**.

A linguagem M é uma linguagem de programação usada no Power BI para realizar transformações e manipulações de dados no processo de preparação de dados. Ela permite às pessoas usuárias escrever instruções sequenciais para filtrar, agrupar, unir tabelas e executar outras operações de transformação de dados. A linguagem M facilita a criação de processos automatizados e consistentes para preparar dados para análise no Power BI.

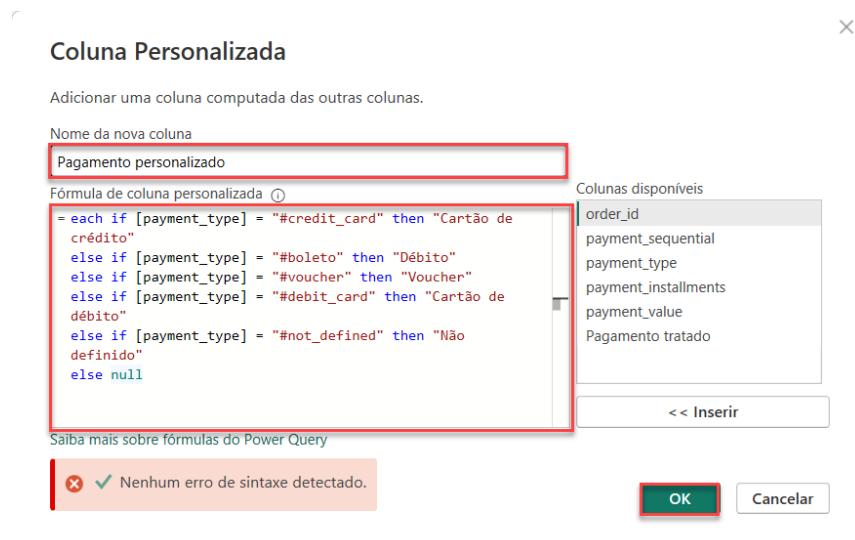
Para realizar essas alterações usando a linguagem M, vamos utilizar o próprio código gerado ao criar a coluna Pagamento tratado. Vamos copiar o código da caixa de texto acima da coluna criada, expandindo essa caixa clicando na seta à direita do campo. O código que vamos copiar será a partir do comando “each” até o “null”, que está marcado em azul:



Agora, vamos criar uma coluna personalizada, acessando o botão no canto superior direito, na aba de Adicionar coluna:



Ao clicar no botão, uma janela será aberta, onde poderemos modificar o nome da coluna e adicionar o código que copiamos:



Após adicionarmos o código, basta clicarmos em OK e a nova coluna será exibida:

```
= Table.AddColumn("#Coluna Condicional Adicionada", "Pagamento personalizado", each if [payment_type] = "#credit_card" then "Cartão de crédito"
else if [payment_type] = "#boleto" then "Débito"
else if [payment_type] = "#voucher" then "Voucher"
else if [payment_type] = "#debit_card" then "Cartão de débito"
```

quantal	AB_C payment_type	123 payment_installments	123 payment_value	AB_C Pagamento tratado	ABC 123 Pagamento personalizado
100% 0% 0%	● Válidos ● Erro ● Vazio	100% 0% 0%	● Válidos ● Erro ● Vazio	100% 0% 0%	● Válidos ● Erro ● Vazio
1	#credit_card	8	9933	Cartão de crédito	Cartão de crédito
1	#credit_card	1	2439	Cartão de crédito	Cartão de crédito
1	#credit_card	1	6571	Cartão de crédito	Cartão de crédito
1	#credit_card	8	10778	Cartão de crédito	Cartão de crédito
1	#credit_card	2	12845	Cartão de crédito	Cartão de crédito
1	#credit_card	2	9612	Cartão de crédito	Cartão de crédito
1	#credit_card	1	8116	Cartão de crédito	Cartão de crédito
1	#credit_card	3	5184	Cartão de crédito	Cartão de crédito
1	#credit_card	6	34109	Cartão de crédito	Cartão de crédito
1	#boleto	1	5195	Débito	Débito
1	#credit_card	8	18873	Cartão de crédito	Cartão de crédito
1	#credit_card	1	14190	Cartão de crédito	Cartão de crédito
1	#credit_card	1	7578	Cartão de crédito	Cartão de crédito
1	#credit_card	5	10266	Cartão de crédito	Cartão de crédito
1	#credit_card	4	10528	Cartão de crédito	Cartão de crédito
1	#credit_card	10	15745	Cartão de crédito	Cartão de crédito
1	#credit_card	1	13204	Cartão de crédito	Cartão de crédito
1	#credit_card	1	9894	Cartão de crédito	Cartão de crédito
1	#credit_card	4	24415	Cartão de crédito	Cartão de crédito
1	#credit_card	3	13671	Cartão de crédito	Cartão de crédito
1	#credit_card	4	4769	Cartão de crédito	Cartão de crédito
1	#credit_card	2	17057	Cartão de crédito	Cartão de crédito
1	#credit_card	4	6750	Cartão de crédito	Cartão de crédito

Com isso, temos mais uma opção para trabalharmos o tratamento da coluna de pagamentos, através da linguagem M.

4.4. Tabela originária do arquivo JSON

Estamos no Editor do Power Query. Quando selecionamos a tabela `olist_produtos`, a partir da visualização, conseguimos perceber que existem atributos importantes, como o ID do produto e a categoria.

```
Json.Document(File.Contents("C:\DADOS\CEFET\2024\cursoExtensaoPBI2024\curso3\dados\olist_produtos.json"))
```

product_id	Record
product_category_name	Record
product_name_lenght	Record
product_description_lenght	Record
product_photos_qty	Record
product_weight_g	Record
product_length_cm	Record
product_height_cm	Record
product_width_cm	Record

Na parte superior da tela, há uma guia chamada "Converter", porque o Power Query sabe que precisamos converter a estrutura de JSON para uma tabela. No canto superior esquerdo, encontramos a opção "Na Tabela". Apertando este botão, o Power Query traz uma estrutura semelhante a uma tabela, com duas colunas: "Nome" e "Valor".

Vamos selecionar a coluna "Value" ("Valor"). No canto direito da coluna, há um ícone de duas setas. Apertando o ícone, visualizaremos uma expansão dos valores. Para expandi-los, apertaremos "Ok". Ao fazermos isso, receberemos uma série de informações dos registros. Esse processo pode demorar um pouco, porque são várias informações.

As informações foram carregadas e podemos observar como está a estrutura: ela é bastante parecida com uma tabela, mas não está totalmente preenchida, isto é, não temos todos os valores.

Por exemplo, se analisarmos a primeira coluna, ela é composta por linhas com nomes que deveriam ser colunas, na verdade.

Conferindo os valores horizontalmente, por exemplo, da linha `product_id`, notamos que cada registro, da esquerda para a direita, são IDs. Significa que a disposição da coluna é horizontal. Ela está crescendo de forma horizontal e não vertical, que é o ideal quando trabalhamos com tabelas.

Precisamos fazer uma transposição da tabela. Vamos reorganizá-la para que ela deixe de crescer horizontalmente e, sim, verticalmente. Para isso, vamos selecionar todas as colunas. Basta selecionar a primeira e usar o atalho "Ctrl + Shift + A".

Em seguida, acessaremos a guia "Transformar" na parte superior da tela. Nela, selecionaremos a opção "Transpor". Assim, conseguiremos obter um resultado que faz mais sentido para o que queremos.

Na primeira coluna, temos o `product_id` como primeiro resultado e, logo abaixo, todos os valores de IDs, dos produtos. O mesmo vale para a segunda coluna, onde o primeiro valor é a categoria, seu nome e todos os valores de categoria abaixo.

Agora, vamos fazer uma "promoção" do cabeçalho. No canto esquerdo, vamos selecionar a opção "Usar a Primeira Linha como Cabeçalho". Aparentemente, todas as informações numéricas estão fazendo sentido.

Em uma coluna, especificamente, vale a pena realizar alguns tratamentos para deixá-la esteticamente mais apresentável, caso seja necessário apresentar o relatório para alguém.

A coluna de "Categoria", `product_category_name`, possui algumas particularidades. Uma delas é que, os espaços entre uma palavra e outra estão representados por *underline*, por exemplo, "coisas_legais", "moveis_decoracao", "cama_mesa_banho", dentre outros.

Podemos remover esse *underline* e substituí-lo por espaço. Também podemos definir que as primeiras letras de cada palavra sejam maiúsculas. Vamos começar removendo o *underline*. Para isso, apertaremos com o botão direito do mouse sob o valor e selecionar a opção "Substituir Valores".

Aparecerá uma janela com os campos "Valor a ser localizado" e "Substituir". Vamos apagar o valor a ser localizado sugerido pelo Power Query: "coisas_legais". Nossa intenção é localizar somente o *underline*, então, escreveremos `_` no primeiro campo para localizarmos todos os registros com *underline*.

No segundo campo, adicionaremos espaço (apertando tecla de espaço), assim, tudo que for *underline*, será substituído por espaço. Por fim, apertaremos "Ok" para que todos os tratamentos sejam feitos. Feito isso, todos os nomes das categorias estão sem *underline* e com espaço.

O próximo tratamento é tornar a primeira letra maiúscula. Vamos apertar a coluna de categoria com o botão direito do mouse e acessar a opção "Transformar".

Nela, encontraremos diversas opções: deixar uma palavra inteira em letras minúsculas ou maiúsculas; colocar cada palavra em maiúscula; cortar (retirar espaços em branco); limpar; identificar o comprimento; dentre outras. Vamos escolher "Colocar cada palavra em maiúscula".

As alterações foram feitas em todos os resultados da nossa coluna. Com isso, conseguimos tratar a base da melhor maneira possível, facilitando análises futuras.

Finalizamos o arquivo `olist_produtos`, mas, ainda podemos utilizar alguns recursos que facilitariam o acesso de outras pessoas, caso queiram utilizar este arquivo.

4.5. Gerenciar parâmetros

Quando pensamos em gerenciar arquivos conectados ao Power Query ou em compartilhar nosso arquivo PBIX, permitindo o acesso a outras pessoas e que elas analisem os tratamentos que estamos fazendo, uma das questões que precisamos lidar para facilitar esse acesso é o caminho do arquivo.

Por exemplo, em um processo seletivo, quando disponibilizamos um arquivo, é importante que ele seja totalmente visível.

Se analisarmos o nosso projeto, temos três arquivos com conexão local: o arquivo de JSON, `olist_produtos`; o arquivo CSV, `olist_itens_pedidos`; e o arquivo Excel, `olist_pedidos`.

Vamos acessar o `olist_pedidos` e, nas etapas aplicadas, visualizar a primeira etapa, que é "Fonte". O Power Query trará na fórmula justamente o caminho:

```
"C:\DADOS\CEFET\2024\cursoExtensaoPBI2024\curso3\dados\olist_pedidos.xlsx".
```

Se tivéssemos que alterar o local dos quatro arquivos, o Power Query não os identificaria e teríamos um problema.

Então, se trocamos o local, precisamos mudar também o caminho de cada arquivo? Existe uma forma mais prática de se fazer isso? Sim, podemos usar o **gerenciador de parâmetros**.

A ideia é criar um parâmetro incorporado em cada um dos arquivos. Cada vez que alteramos o parâmetro, ele altera os quatro arquivos ao mesmo tempo. Vamos copiar o caminho

```
"C:\DADOS\CEFET\2024\cursoExtensaoPBI2024\curso3\dados\ ".
```

Em seguida, na guia "Página Inicial", vamos selecionar a opção "Gerenciar Parâmetros". Uma janela será aberta, onde apertaremos a opção "Novo". Nela, ativamos algumas configurações que podem ser realizadas no parâmetro. Vamos começar definindo o "Nome" do parâmetro, que será "caminhoPasta".

A "Descrição" será "Diretório dos arquivos". Essa descrição serve para que qualquer pessoa que acesse o projeto saiba do que se trata o parâmetro.

Existem outras configurações logo após da descrição: se o parâmetro é obrigatório; qual é o tipo, no nosso caso, o parâmetro será textual, pois será passado em um campo de texto; quais os valores sugeridos, "qualquer valor", "lista de valores" e "consulta".

Dentre as três opções, vamos escolher "Qualquer valor" e passaremos um texto específico, o do caminho. Mas, os outros recursos, "Lista de valores" e "Consulta", são opções interessantes para explorarmos os parâmetros. Nesta aula, você encontrará uma atividade em que poderá conhecer melhor essas outras duas possibilidades.

Em "Valor Atual", vamos colar o caminho. Por fim, apertaremos "Ok". No canto esquerdo, área de consultas, surgiu o parâmetro "CaminhoPasta" que traz para a área central o "Valor Atual", que podemos alterar em qualquer momento.

Neste momento, se alterarmos o parâmetro, não teremos nenhuma consequência nos arquivos, pois não criamos um vínculo, isto é, não adicionamos esse parâmetro no arquivo. Faremos isso agora!

Vamos acessar a `olist_pedidos`, depois, no canto direito, "Etapas Aplicadas > Fonte". Dentro da barra de fórmulas que está na parte superior da tela, quando selecionamos a fonte, aparece justamente a conexão com o caminho. Vamos remover todo o caminho até o nome do arquivo:

```
"C:\DADOS\CEFET\2024\cursoExtensaoPBI2024\curso3\dados\".
```

Ficou apenas o nome do arquivo, "olist_pedidos.xlsx". Antes das aspas, vamos adicionar nosso parâmetro: caminhoPasta. O resultado é:

```
= Excel.Workbook(File.Contents(caminhoPasta"olist_pedidos.xlsx"), null, true)
```

Neste momento, o Power Query identifica que existe um parâmetro e sabe que existe o nome do arquivo, mas ainda não sabe que o caminho precisa se unir ao arquivo. Utilizaremos um recurso de concatenação da linguagem M que é o "&" ('e' comercial)".

```
= Excel.Workbook(File.Contents(caminhoPasta & "olist_pedidos.xlsx"), null, true)
```

Adicionado o &, podemos apertar "Enter". A nossa estrutura é a mesma, mas, se removêssemos o campo de &, teríamos um erro de sintaxe. Com o &, temos a garantia de que a conexão está sendo executada de forma correta.

Agora, vamos fazer o mesmo procedimento com os outros dois arquivos. Copiaremos a estrutura de caminhoPasta & e adicionaremos no olist_itens_pedidos. Acessando "Etapas Aplicadas > Fonte", vamos remover o caminho e colocar o parâmetro na frente. Depois, repetiremos o processo para o olist_produtos e para a tradução.

Os quatro arquivos estão configurados com o parâmetro. Significa que, todas as vezes em que alterarmos o caminho da pasta, os quatro arquivos serão automaticamente alterados.

Vamos fazer um teste! Abriremos a pasta com os quatro arquivos e criaremos uma nova pasta (Botão direito do mouse, "Novo > Pasta") que se chamará "Olist". Em seguida, arrastaremos os quatro arquivos para dentro dessa pasta. Vamos analisar o que vai acontecer no editor do Power Query.

Então, de volta ao Power Query, vamos selecionar uma das bases para verificarmos o que aconteceu. Também precisamos apertar "Atualizar". Apareceu um erro. Se selecionarmos os outros arquivos com a conexão e atualizarmos, eles também apresentarão erros.

Como sabemos que este problema está acontecendo porque adicionamos uma nova pasta, basta acessar o nosso parâmetro e inserir Olist, que é o nome da nossa nova pasta. Ou seja, o campo "Valor Atual" ficará:

```
C:\DADOS\CEFET\2024\cursoExtensaoPBI2024\curso3\dados\Olist\
```

Pressionando "Enter", todos os arquivos no canto esquerdo serão atualizados automaticamente. A informação de exclamação que alertava para os erros inclusive sumiu.

Este recurso é muito útil para auxiliar-nos, não só na gestão de arquivos que estão em determinado local e precisam ser remanejados, mas também para quem deseja acessar o Power Query e não tem a fonte de dados no mesmo local que a sua máquina. Basta acessar o parâmetro e inserir o diretório dos arquivos. Com isso, conseguirá acessar todos os arquivos facilmente.

Conseguimos ajustar os parâmetros, falta explorar alguns detalhes na tabela que ainda não verificamos, porque todas as vezes em que fizemos alterações nas etapas aplicadas, sempre vem uma etapa de "tipo alterado". Precisamos entender que etapa é essa.

4.6. Etapa de alteração de Tipos de Dados

Para avançar no projeto e ter um entregável que faça mais sentido, é super comum **alterar o tipo dos dados** de cada coluna.

Analisamos que nas etapas aplicadas, fizemos algumas alterações e logo em seguida surgiu uma etapa chamada "Tipo alterado". O que está acontecendo internamente?

Automaticamente, o Power BI gera essas páginas. Ele interpreta que os dados têm determinado comportamento e traz as páginas. Será que essa é a melhor solução? Vamos acompanhar um exemplo que nos ajudará a entender que nem sempre essa é a melhor tipagem.

Por exemplo, na tabela `olist_itens_pedidos`, se avaliarmos as duas últimas colunas, `price` e `freight_value`, encontraremos resultados que estão um pouco fora do normal. Por exemplo, um dos valores de frete (`freight`) é R\$1329,00 (mil trezentos e vinte e nove reais), e isso não faz o menor sentido.

Podemos antecipar uma etapa, já que os dados já foram tipados. Na lateral esquerda, em "Etapas aplicadas", vamos selecionar "Cabeçalhos Promovidos" e perceberemos que os resultados estão diferentes.

A ^B C price	A ^B C freight_value
58.9	13.29
199.0	17.87
12.99	12.79
199.9	18.14
21.9	12.69
19.9	11.85

O resultado de frete que era R1329,00, por exemplo, na verdade é de R\$13,29 (treze reais e vinte e nove centavos). Como os números estão com pontos, o Power Query os interpretou como uma casa de milhar e não como uma divisão decimal. Podemos fazer uma transformação específica para os nossos dados.

Então, selecionaremos as duas colunas com a tecla "Ctrl", apertaremos o botão direito do mouse e acessaremos "Alterar Tipo". Dentro dessa opção, conseguimos identificar todos os tipos existentes no Power Query. Nós acessaremos o "Alterar Tipo > Usando a Localidade".

A ideia é utilizar esse recurso para ensinar ao Power Query que o ponto representa uma casa decimal. Na janela que apareceu na tela, vamos apertar "Inserir". Surgirá uma nova janela, onde informaremos qual tipo de dado vamos inserir e a localidade.

O tipo de dado será "Número Decimal". Sobre a localidade, no Brasil, a vírgula é o separador decimal, mas, estamos usando o ponto. Por isso, acessaremos a localização dos Estados Unidos. Sendo assim, em "Localidade", buscaremos por "Inglês (Estados Unidos)".

Ainda na janela, aparece um exemplo dos valores de entrada, onde conseguimos conferir que a separação se dará por ponto. Por fim, vamos apertar "Ok". O Power Query definirá as duas colunas como números decimais, mas com a vírgula localizada no lugar correto.

Agora as informações estão corretas! Inclusive, se quisermos, podemos alterar de número decimal para valor monetário clicando em cima da coluna e selecionando a opção "Número decimal fixo > Inserir". Assim, o Power Query saberá que estamos lidando com valores monetários e os apresenta como dinheiro.

\$ price	\$ freight_value
58,90	13,29
199,00	17,87
12,99	12,79
199,90	18,14
21,90	12,69
19,90	11,85

O "Tipo Alterado" ao final ("Etapas aplicadas", na lateral direita) não faz mais sentido. Podemos removê-lo. O próximo passo é analisar se é necessário fazer mais ajustes na tabela.

É sempre importante ter atenção nesta parte, porque o tipo que fica determinado na base de dados pode influenciar no relatório final e nos resultados. Para ter mais controle dos tipos de dados, é possível configurar o editor do Power Query para que ele não defina os tipos de dados assim que efetuamos a conexão.

Vamos acessar "Arquivo > Opções e configurações". Na janela que será aberta, nas opções de configuração global, existe a opção de "Carregamento de Dados", que está selecionada, assim como a "Detecção de tipo".

Em "Detecção de tipo", existem três opções que podemos marcar:

- Sempre detectar tipos de coluna e cabeçalhos de fontes não estruturadas;
- Detectar tipos de coluna e cabeçalhos de fontes não estruturadas de acordo com a configuração de cada arquivo (esta opção está selecionada);
- Nunca detectar tipos de colunas e de cabeçalhos de fontes estruturadas.

A terceira opção permite que, ao conectarmos e importarmos a base de dados, nada seja definido como tipo. Por isso, vamos selecioná-la. Por fim, apertaremos "Ok" para salvar a operação.

Com isso, todas as vezes em que conectarmos qualquer base, ela já não terá o tipo dos dados. É muito importante que agora você analise os quatro arquivos, considerando se os tipos de dados fazem sentido para o projeto.

Entendemos como lidar com os tipos de dados no Power Query, mas, pensando em todo o processo que fizemos até agora, de inclusão de novas colunas e outras mudanças, podemos excluir alguns elementos, por exemplo, algumas colunas.

Inclusive, todas as nossas tabelas estão nomeadas em inglês. Já que estamos tentando manter o padrão de uma única língua, seria interessante renomear as colunas. Como podemos fazer isso dentro da estrutura do Power Query?

Agora, que já fizemos todos os tratamentos que precisávamos dentro das bases, alguns ajustes são necessários para fazermos as cargas da melhor maneira possível.

Um desses ajustes é excluir colunas que não fazem sentido para nosso entregável final.

O que precisamos analisar?

4.7. Excluir colunas

Dentro da tabela `olist_itens_pedidos` temos uma coluna chamada `seller_id`, que é o id do vendedor. Essa coluna não faz muito sentido para o nosso negócio, pois não temos uma tabela que traga informações sobre os vendedores, Só temos esse id, então não teremos como relacionar. Por isso, vamos excluir essa coluna.

Para excluir a coluna `seller_id`, basta clicar com o botão direito em cima da coluna e selecionar a opção "Remover".

A partir desse momento, essa coluna deixa de existir na nossa base. Também temos uma etapa, no painel à direita, que evidencia essa exclusão, chamada "Colunas Removidas".

Na tabela `olist_pedidos` também temos colunas que não fazem sentido serem mantidas. Por exemplo, a coluna `order_status` é uma coluna da qual fizemos a tradução e criamos uma nova coluna chamada "Português".

E, ao lado esquerdo da `order_status`, temos a coluna `customer_id`, também não temos uma tabela que traga informações sobre isso.

Então, vamos excluir as colunas `order_status` e `customer_id`. Podemos selecionar as duas, clicar com o botão direito e selecionar "Remover".

Na tabela `olist_pagamentos` excluiríamos a coluna `payment_type`, pois já fizemos a tradução dela na coluna Pagamento tratado.

Outra forma de remover colunas é selecionar a coluna que desejamos remover e pressionar a tecla "Delete".

Além disso, podemos também selecionar as colunas que queremos manter e excluir as demais que não estiverem selecionadas. É o que faremos na tabela `olist_produtos`.

Para o nosso negócio, o que importa na tabela `olist_produtos` é trazer informações somente das duas primeiras colunas, `product_id` e `product_category_name`. Vamos selecionar essas duas colunas, clicar com o botão direito e selecionar "Remover outras colunas".

Aprendemos três formas de excluir colunas e manter o que realmente importa para nosso projeto. Agora, vamos fazer alguns ajustes.

4.8. Renomear colunas

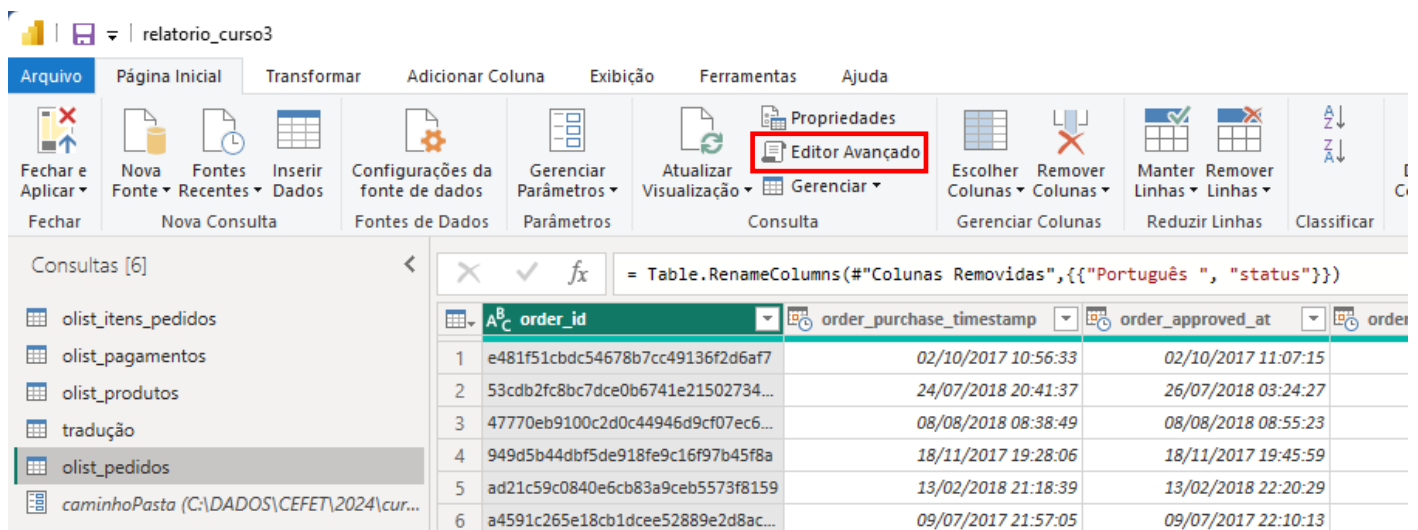
Na tabela `olist_pedidos` vamos renomear a coluna Português para status. Dessa forma ela fica com um nome que faz mais sentido ao projeto.

Para isso, podemos clicar na coluna Português com o botão direito e selecionar a opção "Renomear" ou podemos dar um duplo clique no cabeçalho da coluna, vamos renomear para status.

Agora, vamos analisar as etapas aplicadas de uma forma diferente.

4.9. Editor Avançado

Na tabela `olist_pedidos`, para acessar as etapas aplicadas como um todo, podemos acessar a guia "Página inicial" e selecionar a opção "Editor avançado".



The screenshot shows the Power BI Desktop interface with the 'Editor Avançado' (Advanced Editor) window open. The window displays a formula bar with the text '= Table.RenameColumns("#Colunas Removidas",{"Português", "status"})'. Below the formula bar, a table view shows the structure of the 'olist_pedidos' table with columns: order_id, order_purchase_timestamp, order_approved_at, and order. The table contains 6 rows of data.

	order_id	order_purchase_timestamp	order_approved_at	order
1	e481f51cbdc54678b7cc49136f2d6af7	02/10/2017 10:56:33	02/10/2017 11:07:15	
2	53cdb2fc8bc7dce0b6741e21502734...	24/07/2018 20:41:37	26/07/2018 03:24:27	
3	47770eb9100c2d0c44946d9cf07ec6...	08/08/2018 08:38:49	08/08/2018 08:55:23	
4	949d5b44dbf5de918fe9c16f97b45f8a	18/11/2017 19:28:06	18/11/2017 19:45:59	
5	ad21c59c0840e6cb83a9ceb5573f8159	13/02/2018 21:18:39	13/02/2018 22:20:29	
6	a4591c265e18cb1dcee52889e2d8ac...	09/07/2017 21:57:05	09/07/2017 22:10:13	

Ao clicar em "Editor Avançado" será aberta uma janela que lista todas as alterações que existem dentro da tabela `olist_pedidos`. A estrutura do editor avançado utiliza a **linguagem M**.

Ele tem dois blocos, o bloco `let` e o bloco `in`. O `let` engloba tudo que está sendo executado (tratamento, transformação, modificação, conexão, etc). E o `in` é o retorno, a última transformação que fizemos. Podemos ver que o `in` está da seguinte forma: "#Colunas Renomeadas".

olist_pedidos

```
let
    Fonte = Excel.Workbook(File.Contents(caminhoPasta & "olist_pedidos.xlsx"), null, true),
    olist_orders_dataset_Sheet = Fonte[[Item="olist_orders_dataset",Kind="Sheet"]][Data],
    #"Linhas Superiores Removidas" = Table.Skip(olist_orders_dataset_Sheet,2),
    #"Cabeçalhos Promovidos" = Table.PromoteHeaders(#"Linhas Superiores Removidas", [PromoteAllScalars=true]),
    #"Tipo Alterado" = Table.TransformColumnTypes(#"Cabeçalhos Promovidos",{{"order_id", type text}, {"customer_id", type
    #"Texto Extraído Após o Delimitador" = Table.TransformColumns(#"Tipo Alterado", {{"order_status", each Text.AfterDeli
    #"Consultas Mescladas" = Table.NestedJoin(#"Texto Extraído Após o Delimitador", {"order_status"}, tradução, {"Inglês"
    #"tradução Expandido" = Table.ExpandTableColumn(#"Consultas Mescladas", "tradução", {"Português "}, {"Português "}),
    #"Colunas Removidas" = Table.RemoveColumns(#"tradução Expandido",{"customer_id", "order_status"}),
    #"Colunas Renomeadas" = Table.RenameColumns(#"Colunas Removidas",{{"Português ", "status"}})
in
    #"Colunas Renomeadas"
```

Note que na linha 2 temos a fonte, onde estamos nos conectando a esse dado. Logo depois, ele pega essa fonte e identifica como vai transformar essa fonte para colocá-la como uma tabela.

Depois, ele começa a fazer todos os tratamentos, como cabeçalhos promovidos, tipos alterados, linhas superiores removidas, etc.

Essa estrutura pode facilitar muito quando precisamos fazer as mesmas alterações, às vezes na mesma pasta, em projetos diferentes. Então, podemos reaproveitar esse código. É exatamente o que faremos.

Podemos copiar toda a estrutura do código do Editor Avançado da tabela `olist_pedidos`. Em seguida, faremos novamente a importação da tabela `olist_pedidos`.

Vamos clicar em "Nova Fonte > Pasta de Trabalho do Excel" e selecionar o arquivo `olist_pedidos`. Na janela de Navegador, vamos selecionar o arquivo da forma como importamos originalmente. Agora não tem tratamento nenhum, tem apenas a fonte e a navegação.

Agora, vamos acessar o "Editor Avançado" e substituir o texto que está nele pelo código que copiamos anteriormente. Em seguida, podemos clicar no botão "Concluído" e teremos toda a tabela já tratada sem que precisemos fazer as alterações uma de cada vez.

É uma forma que facilita o compartilhamento de tratamentos e transformações.

4.10. Consulta nula

Outro recurso interessante é a consulta nula. Podemos utilizar a consulta nula para trazer essas informações também.

Para isso, podemos clicar em "Nova Fonte > Consulta Nula". Ela traz tudo tudo do zero. Em seguida, basta clicar em "Editor Avançado" e colar as informações que copiamos. Ao clicar em "Concluído", vai aparecer a base de dados com todos os tratamentos e conexões estruturados.

Essas são duas opções que podemos utilizar dentro do editor do Power Query para reproduzir rotinas que já fizemos dentro dele.

Agora, que já demonstramos o exemplo, vamos excluir essas duas tabelas que criamos, a `Consulta1` e a `olist_orders_dataset`.

4.11. Desafio



A Olist é uma plataforma de e-commerce que conecta pequenas empresas a marketplaces, permitindo que elas alcancem um público maior e expandam suas vendas. Com o objetivo de aprimorar seus processos e otimizar a experiência dos clientes, a Olist coletou uma ampla gama de dados ao longo do tempo.

E, como pessoa analista de dados, é seu papel aproveitar ao máximo esses dados, desvendando os valiosos insights por trás dessas informações. Para isso, você vai realizar uma série de tarefas para desenvolver um painel de controle utilizando a ferramenta Power BI, a fim de fornecer visões estratégicas para o negócio.

Para criar um painel de controle eficaz, usando o Microsoft Power BI, você deve executar as seguintes tarefas:

Tarefa 1: Mostrar o total de pedidos da base de dados

Uma das primeiras tarefas que você enfrentará como analista de dados nesse projeto é revelar o total de pedidos registrados na base de dados da Olist. Usando as suas habilidades no Power BI, você poderá criar uma visualização clara e concisa que mostrará o número de pedidos, fornecendo uma visão geral inicial abrangente da base de dados.

Tarefa 2: Criar uma visualização para mostrar a evolução dos pedidos no tempo

Ao mergulhar fundo nos dados da Olist, você descobrirá que a história dos pedidos ao longo do tempo é extremamente valiosa. Essa análise temporal permitirá que a equipe da empresa tome decisões estratégicas mais embasadas e antecipe as demandas futuras. E por meio do uso de visuais no Power BI, será possível obter uma visualização dinâmica e interativa que ilustra a evolução dos pedidos, destacando possíveis sazonalidades, tendências e padrões ao longo do tempo.

Tarefa 3: Criar uma visualização para mostrar os percentuais de preferência por tipo de pagamento

Para compreender melhor o comportamento dos clientes da Olist, é essencial analisar os diferentes tipos de pagamento utilizados em suas transações. Utilizando o Power BI, você poderá criar uma visualização clara e informativa que demonstrará os percentuais de preferência por cada tipo de pagamento. Essa visualização revelará quais métodos de pagamento são mais populares entre os clientes da Olist, fornecendo insights valiosos para direcionar estratégias de marketing e aprimorar a experiência do cliente.

Tarefa 4: Criar uma visualização que mostra a quantidade de produtos por categoria, observando apenas um ranking dos valores mais altos

Outro aspecto importante para entender o negócio da Olist é analisar as categorias de produtos vendidos. Com o PowerBI, você poderá criar uma visualização impactante que destacará as categorias com a maior quantidade de vendas, apresentando um ranking dos valores mais altos. Essa visualização permitirá que a equipe da empresa identifique quais categorias de produtos são as mais populares entre os consumidores e direcione esforços para otimizar o mix de produtos e a estratégia de precificação.

Tarefa 5: Criar um filtro para segmentar os registros por ano

Para facilitar a análise dos dados da Olist, será necessário criar um filtro que permita segmentar os registros por ano. Utilizando o Power BI, você poderá desenvolver um filtro interativo que permitirá à equipe explorar

os dados de forma mais precisa e detalhada, isolando informações específicas de cada ano. Esse filtro será uma ferramenta valiosa para identificar tendências sazonais, avaliar o crescimento ao longo do tempo e fazer comparações entre diferentes períodos. Além disso, é importante criar um filtro que desconsidere pedidos sem data de aprovação (filtro para valores em branco nas datas).

Referências

Microsoft Build. **Power BI documentation**. Disponível em: <https://learn.microsoft.com/en-us/power-bi/> Acesso em: abril/2024.

Microsoft. Mapas de árvore no Power BI. In: **PowerBI**. Disponível em: <https://learn.microsoft.com/pt-br/power-bi/visuals/power-bi-visualization-treemaps?tabs=powerbi-desktop> Acesso em: Jan 2024.

Apêndice

Apêndice A – Importação de dados da Web

Quando se trata de importar uma planilha do Google no Power BI, existem duas opções disponíveis: a conexão web e a opção específica chamada Planilhas Google. Embora ambas as opções permitam importar dados de planilhas do Google para o Power BI, há algumas diferenças notáveis entre elas, com destaque para a opção Planilhas Google.

Conexão Web:

A opção de conexão web é uma maneira de importar dados de planilhas do Google no Power BI usando o link compartilhado da planilha. Ao optar por essa opção, você precisa fornecer o URL da planilha para estabelecer a conexão com o Power BI. Aqui estão algumas considerações sobre a conexão web:

- **Requer um link compartilhado:** Para usar a conexão web, é necessário obter o link compartilhado da planilha no Google Drive. Isso significa que a planilha deve ser compartilhada e acessível a partir do link fornecido.
- **Necessidade de atualização manual:** Quando os dados na planilha são atualizados, você precisa atualizar manualmente a conexão no Power BI para obter os dados mais recentes. Isso pode ser feito clicando com o botão direito no conjunto de dados no Power BI e selecionando "Atualizar agora".
- **Limitações de segurança:** Ao compartilhar um link da planilha, é importante considerar as implicações de segurança. Se alguém tiver acesso ao link compartilhado, poderá visualizar a planilha, o que pode ser indesejado em certas situações.

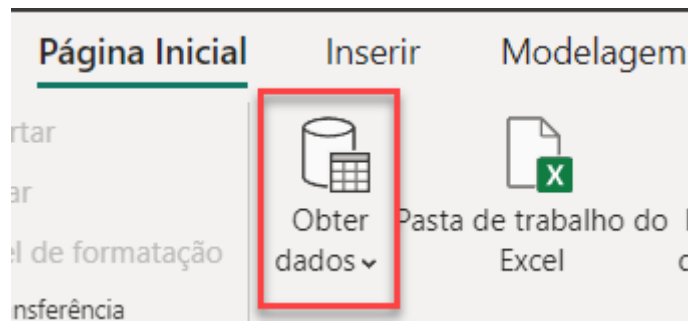
Planilhas Google:

A opção Planilhas Google é uma fonte de dados específica fornecida pelo Power BI para importar planilhas do Google diretamente, sem a necessidade de compartilhar um link. Aqui estão algumas características-chave dessa opção:

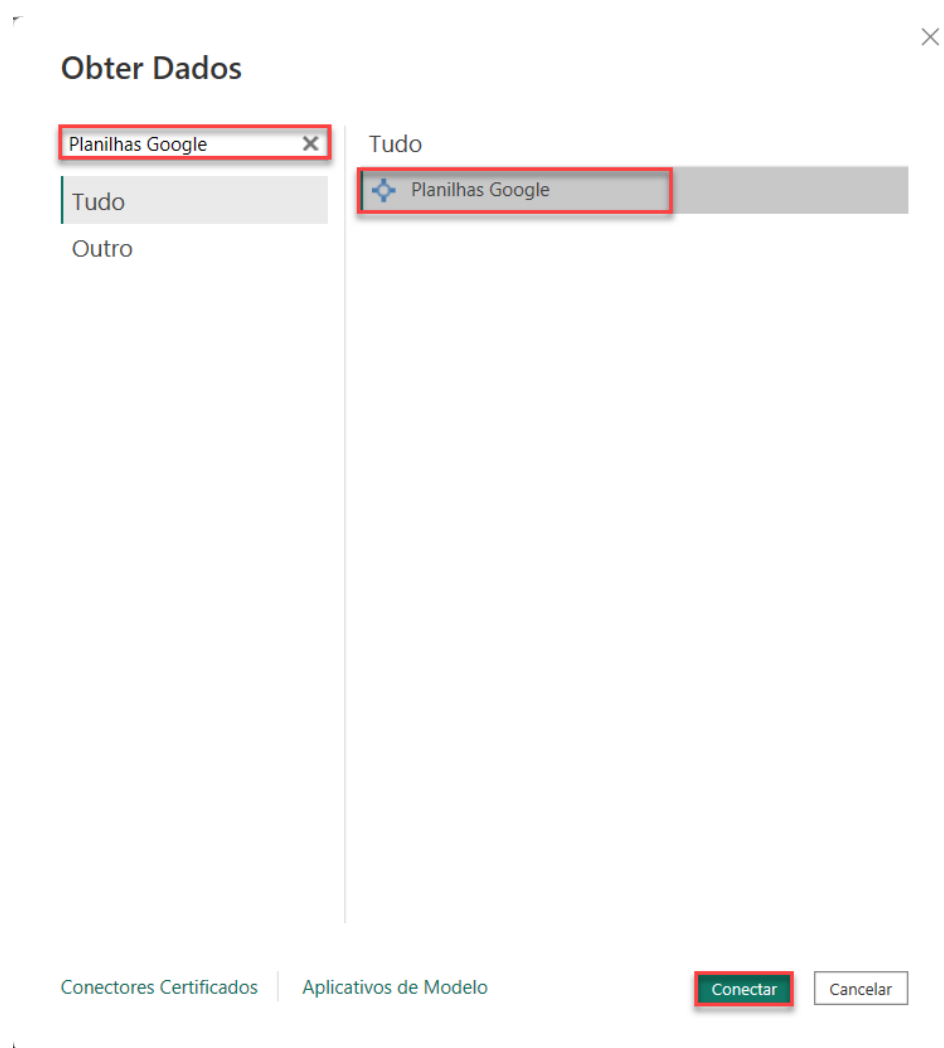
- **Conexão direta:** Com a opção Planilhas Google, você pode se conectar diretamente à sua conta do Google e selecionar a planilha que deseja importar. Isso elimina a necessidade de compartilhar um link ou depender da conexão web.
- **Segurança aprimorada:** Com a opção Planilhas Google, você pode garantir uma camada adicional de segurança, pois não precisa compartilhar um link da planilha. Os dados são importados diretamente de sua conta do Google, mantendo a confidencialidade dos dados.

Caso você queira utilizar essa opção, você pode seguir o passo a passo abaixo:

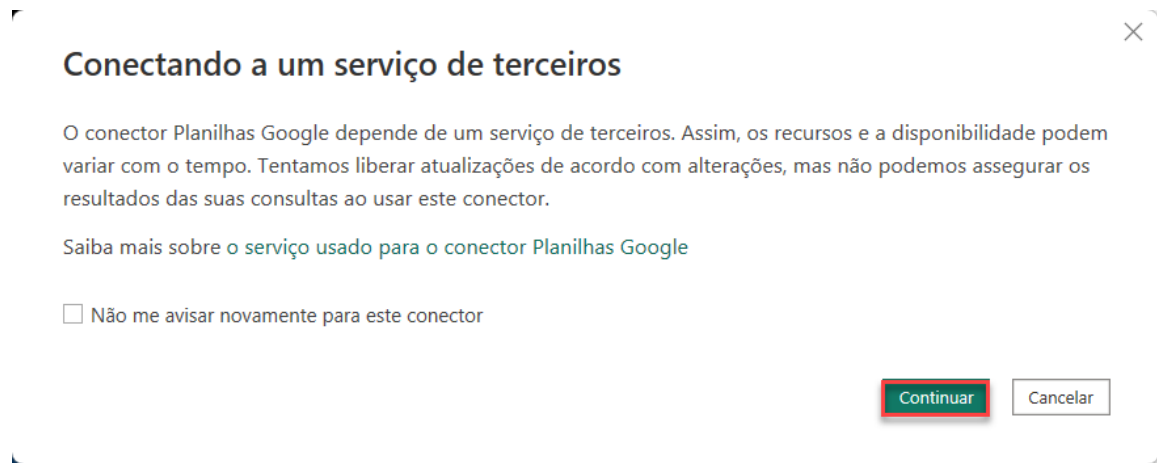
Primeiramente, clique no botão Obter dados:



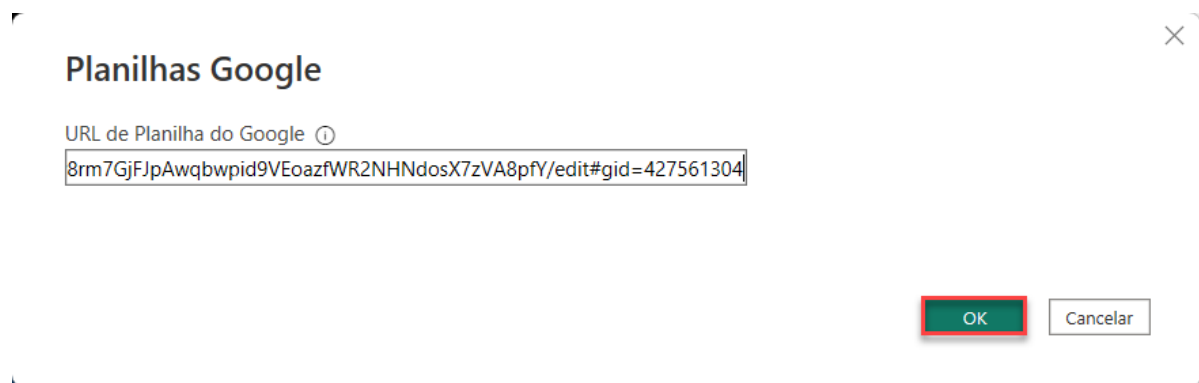
Após a abertura da janela para escolher a fonte de dados, pesquise por Planilhas Google, clique na opção e depois em Conectar:



Quando você clicar em Conectar, uma janela informativa irá aparecer, e basta clicar em Continuar:



Em seguida, basta copiar a URL da sua planilha, da barra do seu navegador mesmo, colar no campo de URL na próxima janela que irá aparecer no Power BI e clicar em OK:



Agora é só seguir com os tratamentos.

Em resumo, a opção Planilhas Google oferece uma conexão direta e automática com planilhas do Google. Além disso, o usuário precisa ser proprietário da planilha, evitando a necessidade de compartilhar links e permitindo maior segurança em relação à exposição dos dados.

Apêndice B – Bases de Dados

Agora que você já aprendeu a importar arquivos para dentro do Power BI, é hora de expandir suas habilidades explorando diferentes bases de dados. Felizmente, existem vários sites que disponibilizam bases de dados gratuitas, especialmente destinadas a estudantes que desejam praticar e explorar dados reais.

Um dos sites recomendados para essa finalidade é o [Open Data NI](#). Esse site oferece uma ampla variedade de bases de dados organizadas por áreas. Embora o site esteja em inglês, ele é intuitivo e você pode usar recursos de tradução do seu navegador para facilitar a compreensão.

Outra opção interessante é o [Dataset Search](#), disponibilizado pelo Google. Através deste site, você pode pesquisar e encontrar conjuntos de dados relevantes para a sua área de interesse.

Para quem está no Brasil, o [Instituto Brasileiro de Geografia e Estatística \(IBGE\)](#) disponibiliza diversos datasets em seu site oficial. Esses dados abrangem diversos levantamentos estatísticos realizados pelo instituto, proporcionando uma ampla gama de informações para análise. Além disso, temos o [Portal Brasileiro de Dados Abertos](#), que fornece um grande conjunto de dados do governo.

Se você estiver interessado em dados relacionados ao [Exame Nacional do Ensino Médio \(ENEM\)](#), pode explorar os datasets disponível neste site. Esses dados oferecem insights valiosos sobre os alunos e são frequentemente utilizados para pesquisas e análises.

A própria Microsoft recomenda o uso do [banco de dados AdventureWorks](#). Ele pode ser utilizado em conjunto com outros softwares da área de dados, como o SQL, proporcionando várias opções de exploração e prática.

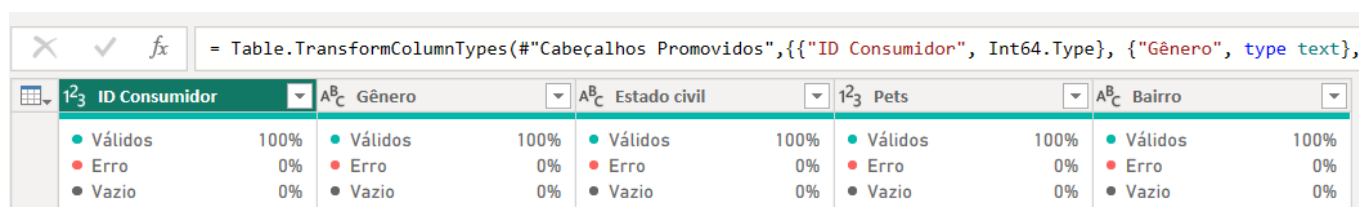
Por fim, o [Kaggle](#) é um dos sites mais conhecidos na área de Data Science. É amplamente utilizado por estudantes e profissionais para encontrar e compartilhar conjuntos de dados interessantes. Navegar pelo Kaggle pode revelar uma infinidade de possibilidades para aprimorar suas habilidades de análise de dados.

Apêndice C – Qualidade da Coluna

O recurso de Qualidade da Coluna no Power BI rotula os valores em linhas em cinco categorias, fornecendo informações sobre a qualidade dos dados em cada coluna:

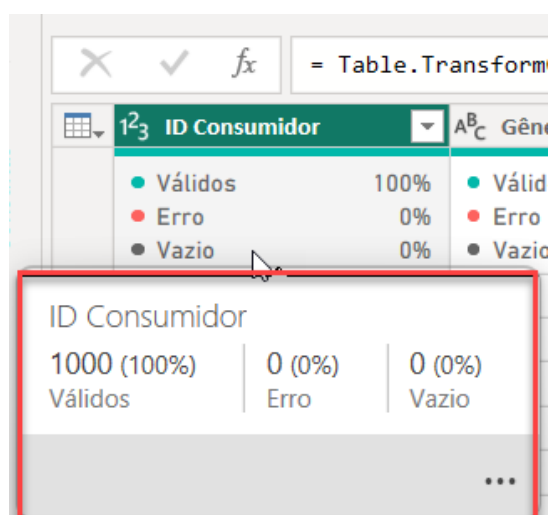
- Válido (verde): indica que os valores na coluna estão corretos e dentro dos critérios definidos.
- Erro (vermelho): sinaliza a presença de erros na coluna, indicando que os valores não estão de acordo com as regras ou critérios estabelecidos.
- Vazio (cinza escuro): representa valores ausentes ou nulos na coluna, indicando que não há dados presentes.
- Desconhecido (verde pontilhado): indica a presença de erros em uma coluna, resultando em uma qualidade de dados desconhecida para os demais valores.
- Erro inesperado (vermelho pontilhado): identifica a ocorrência de erros inesperados na coluna, que não se enquadram nas categorias anteriores.

Esses indicadores são exibidos abaixo do nome da coluna. O número de registros em cada categoria de qualidade de coluna é apresentado como uma porcentagem, como podemos verificar na imagem abaixo:



ID Consumidor	Gênero	Estado civil	Pets	Bairro
Válidos 100%	Válidos 100%	Válidos 100%	Válidos 100%	Válidos 100%
Erro 0%	Erro 0%	Erro 0%	Erro 0%	Erro 0%
Vazio 0%	Vazio 0%	Vazio 0%	Vazio 0%	Vazio 0%

Ao passar o mouse sobre qualquer uma das colunas, é possível visualizar a distribuição numérica da qualidade dos valores em toda a coluna:

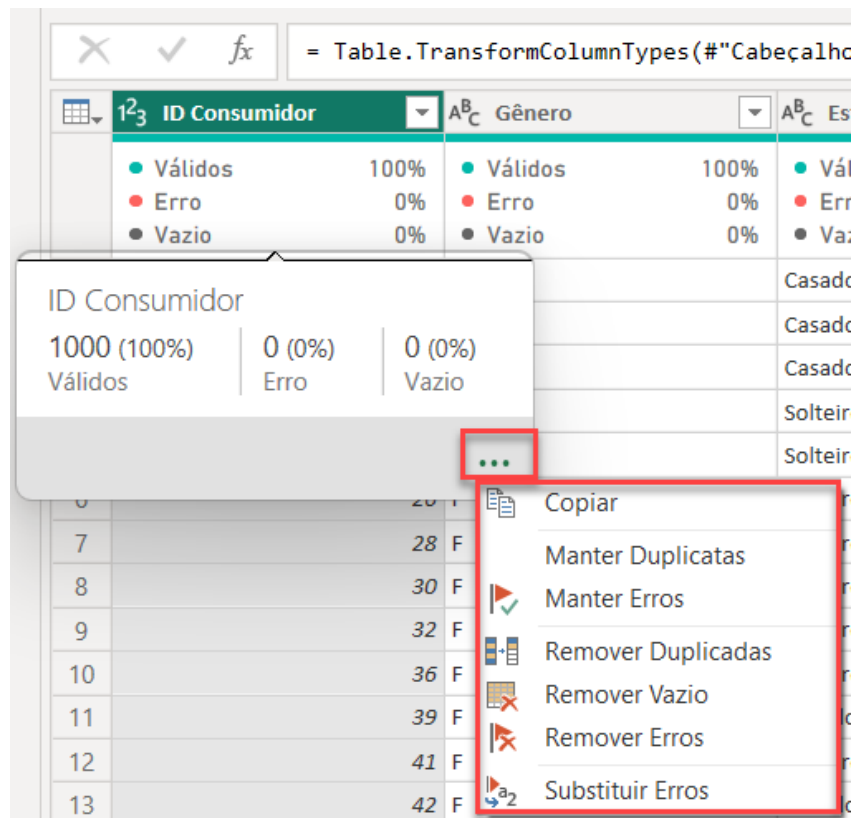


ID Consumidor	Gênero
Válidos 100%	Válidos 100%
Erro 0%	Erro 0%
Vazio 0%	Vazio 0%

ID Consumidor

1000 (100%)	0 (0%)	0 (0%)
Válidos	Erro	Vazio

Além disso, ao selecionar o botão de reticências (...), são exibidos botões de ação rápida que permitem realizar operações nos valores:

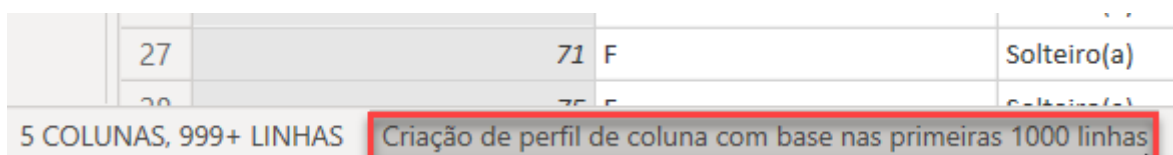


Essa funcionalidade de Qualidade da Coluna no Power BI proporciona uma visão rápida e clara sobre a qualidade dos dados em cada coluna. Entretanto, devemos estar atentos a um quesito muito importante quando se trata do Power BI, que é o fato de se tratar de uma ferramenta que tem como padrão resumir os dados, principalmente por questão de performance.

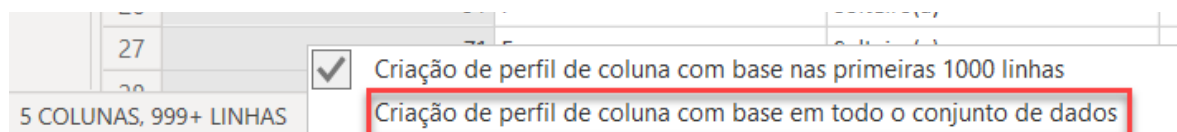
Pensando nisso, o Power BI oferece a opção de filtrar os dados das tabelas, por meio de duas opções: criação de perfil da coluna com base nas primeiras 1000 linhas, que é a opção padrão; e criação de perfil da coluna com base em todo o conjunto de dados. Essa opção de filtragem por todo o conjunto pode ser especialmente importante para garantir uma análise mais precisa e abrangente dos seus dados.

Ao ativar a opção de criação de perfil da coluna com base em todo o conjunto de dados, o Power BI analisará todas as linhas do conjunto de dados, permitindo identificar padrões, distribuições e problemas de qualidade que podem não ser detectados apenas com uma amostra limitada de linhas.

Como exemplo, vamos utilizar a tabela de Clientes, a qual já estávamos utilizando nas imagens anteriores. Assim como pôde ser observado, não foi encontrado nenhum erro, porém, não verificamos como a filtragem estava sendo definida. Para verificar isso, no canto inferior esquerdo, vamos procurar pelo botão da filtragem:



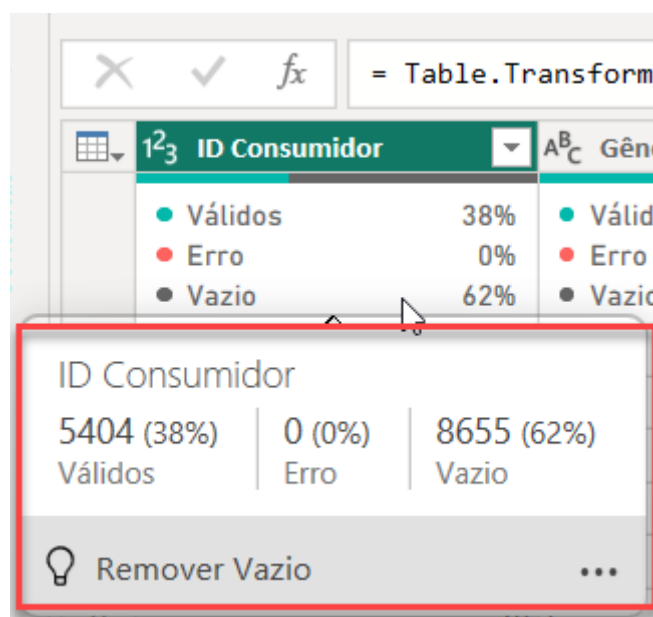
Como podemos verificar, estamos filtrando nossos dados pelas primeiras 1000 linhas. Para resolvermos isso, pode clicar nessa opção, pois se trata de um botão, e selecionar a opção para todo o conjunto de dados:



Agora, voltando para averiguar a qualidade da coluna, podemos ver o que realmente temos em mãos:

= Table.TransformColumnTypes("#Cabecalhos Promovidos",{{"ID Consumidor", Int64.Type}, {"Gênero", type text}},									
123 ID Consumidor	AB C Gênero	AB C Estado civil	123 Pets	AB C Bairro					
● Válidos 38%	● Válidos 38%	● Válidos 38%	● Válidos 38%	● Válidos 38%					
● Erro 0%	● Erro 0%	● Erro 0%	● Erro 0%	● Erro 0%					
● Vazio 62%	● Vazio 62%	● Vazio 62%	● Vazio 62%	● Vazio 62%					

Encontramos uma porcentagem significativa de valores vazios. Se colocarmos o mouse em cima do campo de Vazio, percebemos a diferença:



Além de garantir que a detecção de erros e demais ações de melhoria sejam aplicadas de maneira completa e abrangente, o tratamento desses dados vazios foi essencial para que pudéssemos realizar o relacionamento entre as tabelas de Clientes e Vendas na área de Modelagem de Dados.

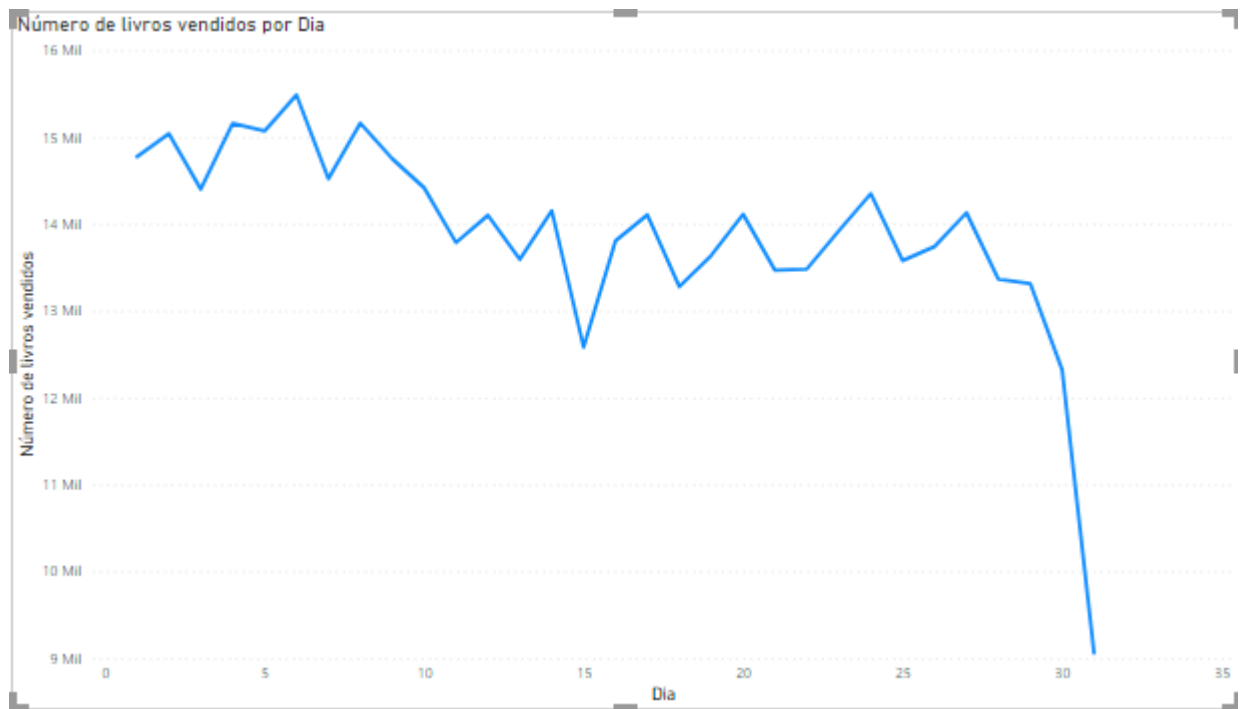
Portanto, a funcionalidade de Qualidade da Coluna no Power BI proporciona uma visão rápida e clara sobre a qualidade dos dados em cada coluna. Adicionalmente, é altamente recomendável utilizar a opção de criação de perfil da coluna com base em todo o conjunto de dados. Isso permitirá uma análise mais precisa e confiável dos dados, fornecendo informações valiosas para aprimorar a qualidade dos seus relatórios e tomada de decisões.

Apêndice D – Múltiplos pequenos

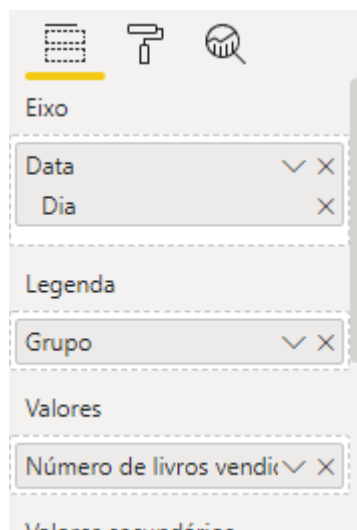
Vimos que temos alguns campos no visual do gráfico de linhas. Temos dois em particular que nem sempre são utilizados, mas podem te auxiliar em algumas análises. São eles os campos de *Eixo y secundário* e *múltiplos pequenos*.

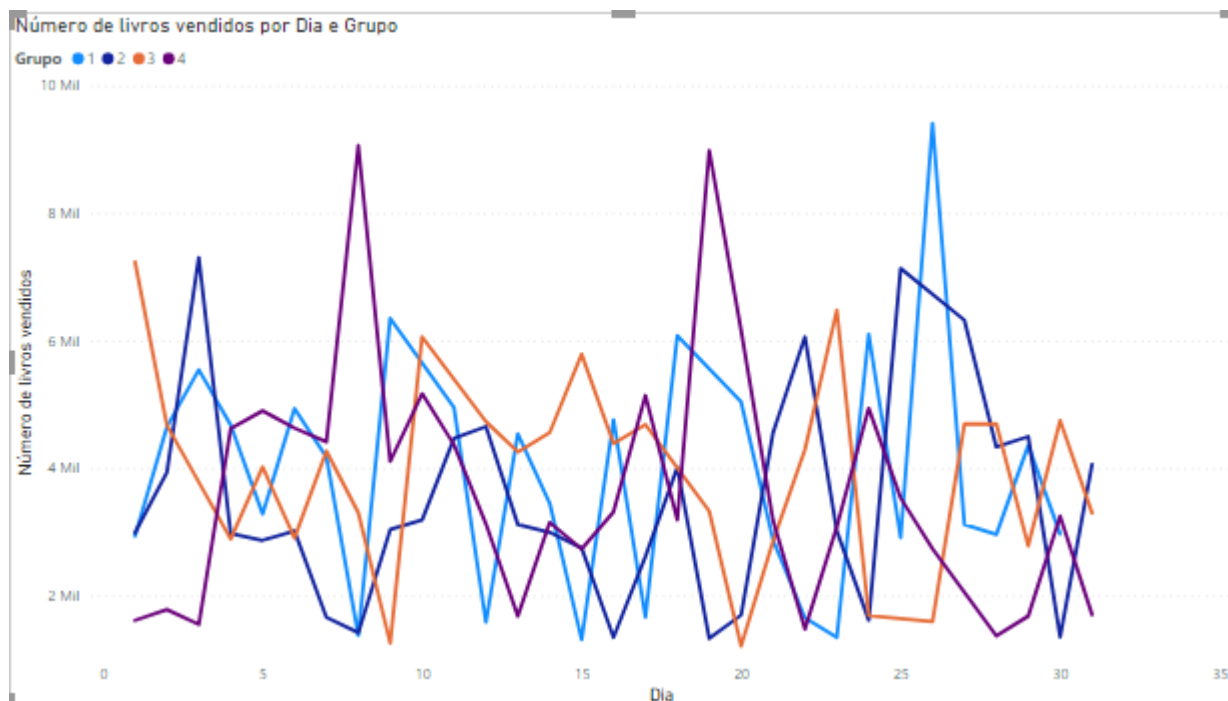
Com os múltiplos pequenos conseguimos dividir nossa análise em subgrupos para verificar o comportamento dos nossos dados em paralelo.

Vamos supor que temos uma série com valores de vendas em um período relativos a 4 grupos de vendedores diferentes:

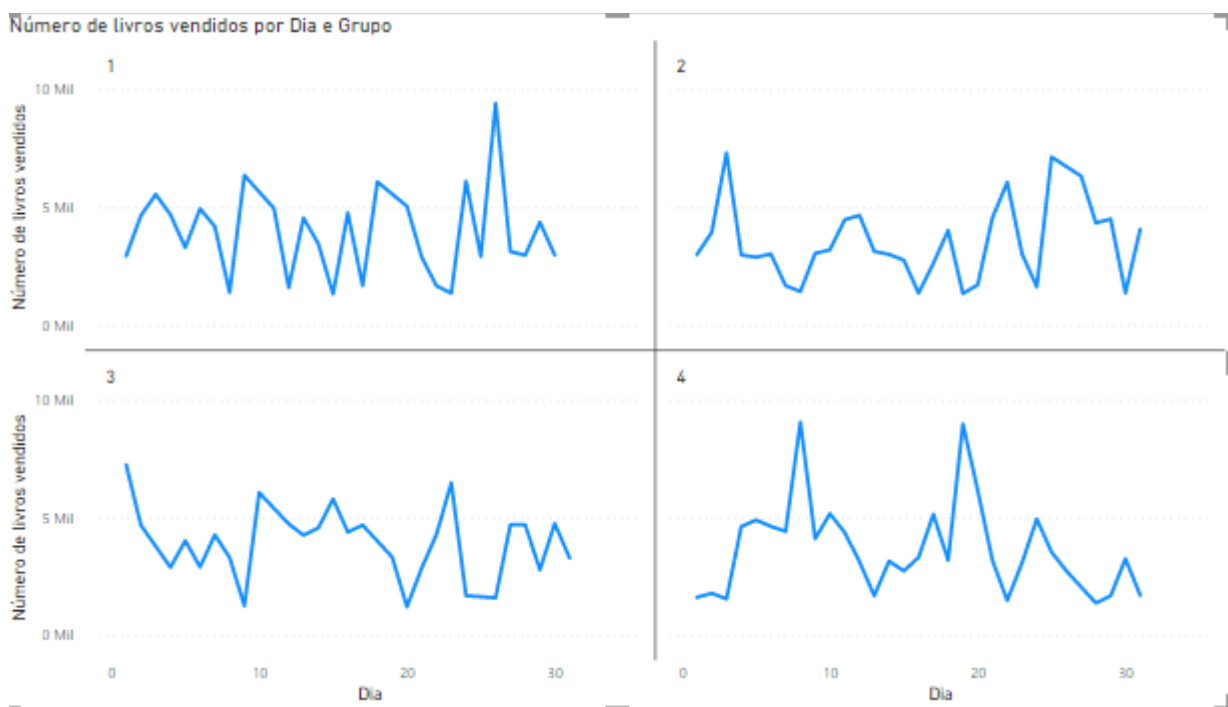


Para analisar os 4 grupos poderíamos arrastar a coluna dos grupos para a legenda, para assim visualizar cada linha de forma independente:





Porém, como dissemos antes, olha só como a análise fica muito poluída, não está claro quais informações queremos destacar. Dessa forma, uma maneira de solucionar isso é que podemos utilizar então o grupo no campo de **múltiplos pequenos**:



Os múltiplos pequenos nos ajudam a dividir a análise, deixando assim a leitura e interpretação muito mais fáceis de se fazer.

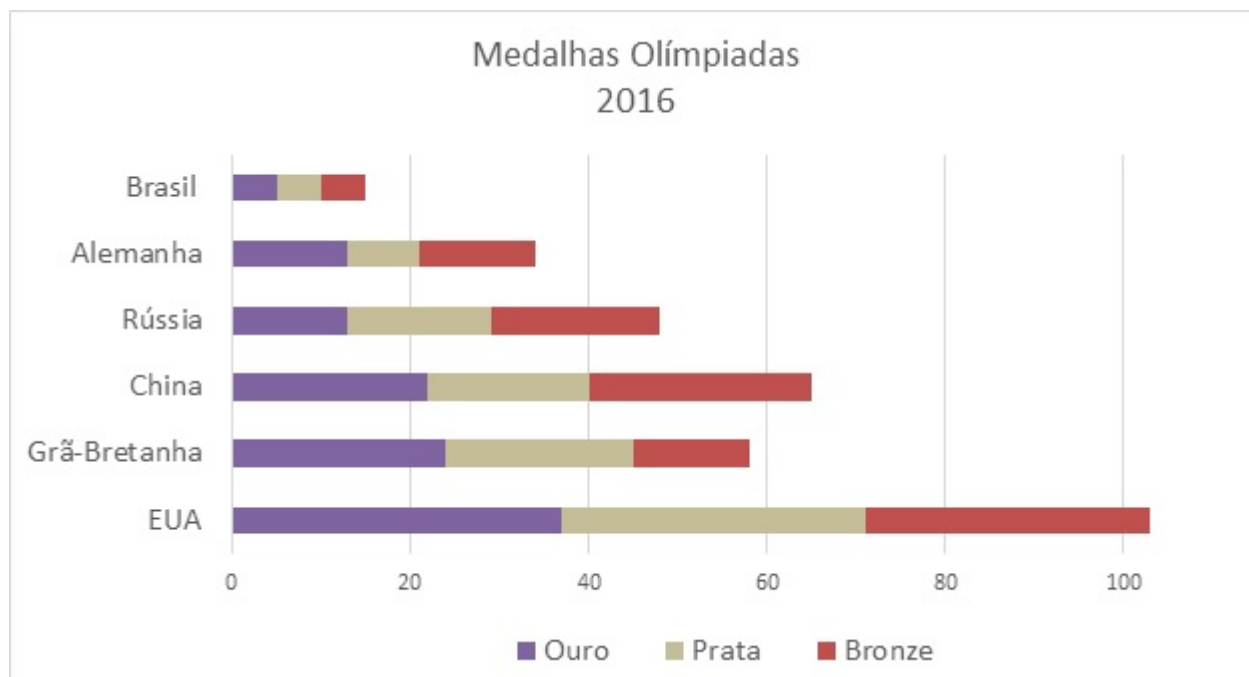
Então aproveite esse recurso do Power BI e o utilize para ajudar ainda mais a tirar insights e conclusões interessantes.

Apêndice E – Empilhada vs Clusterizada

Vimos as diferenças práticas entre os gráficos de barras e de colunas. Em resumo, suas diferenças estão na orientação das barras ou colunas, com o primeiro mostrando as categorias no eixo vertical e o segundo mostrando as categorias no eixo horizontal. A escolha entre eles depende do objetivo da análise e da melhor forma de representar visualmente os dados para facilitar a compreensão.

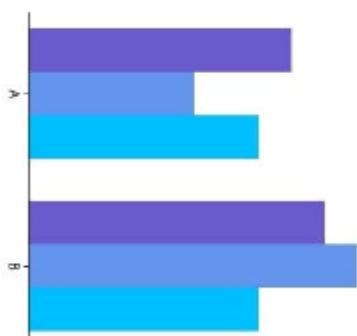
Ainda sobre gráficos apropriados para comparar categorias, podemos ter outras diferenças com gráficos empilhados ou clusterizados, seja de barra ou de colunas, vamos entender a diferença entre cada um.

1 - Gráfico de Barras Empilhadas: O gráfico de barras empilhadas é usado para comparar a contribuição de diferentes categorias em um todo. As barras são empilhadas uma sobre a outra, representando a totalidade do valor, enquanto as diferentes cores nas barras representam as categorias individuais.



2 - Gráfico de Barras Clusterizadas: O gráfico de barras clusterizadas é usado para fazer comparações lado a lado entre diferentes categorias. As barras são agrupadas em clusters, cada um representando uma categoria, e são organizadas lado a lado para facilitar a comparação entre elas.

Portanto, ao escolher por um dos tipos de gráficos aqui comentados, sempre leve em consideração quais são os objetivos das suas análises, o que você gostaria de destacar visualmente de modo que fique compreensível encontrar as informações mais relevantes para a ocasião.



Apêndice F – Como criar conta no Power BI

É importante frisar que você não precisa pagar por nenhuma assinatura para criar uma conta do Power BI, pois é um processo totalmente gratuito.

Para obter as funcionalidades de publicação e distribuição dos dashboards Power BI, é preciso possuir um e-mail comercial ou de estudante. Caso você não tenha, é possível tornar a sua conta pessoal numa conta corporativa para utilizar o Power BI acessando o Microsoft 365 Business Basic.

○

○

○

Assinatura e detalhes da conta

Detalhes da entrada

Informações de pagamento e término

Experimente grátis por um mês

Antes de iniciar sua avaliação, você precisa configurar sua assinatura adicionando até 25 usuários gratuitos.

Para quantas pessoas isso é?

1

^

v

Resumo do seu pedido

Detalhes do seu pedido após o término da avaliação em 25 de novembro de 2023. Você será cobrado mensalmente após essa data.

Microsoft 365 Business Basic

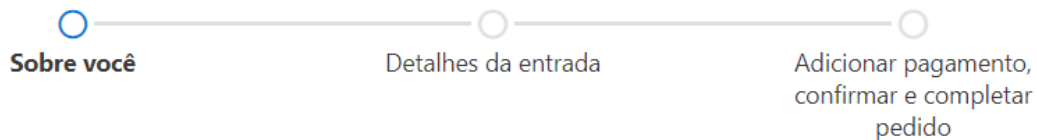
Assinatura 1 ano, Pagar R\$ 32,00 usuário/mês para 1 usuário	
Subtotal após a avaliação (imposto não incluído)	R\$ 32,00

Pagamento com vencimento hoje (imposto não incluído)	R\$ 0,00
--	----------

Após o término da avaliação, ela se tornará uma assinatura paga de 1 ano. Você não será cobrado se cancelar antes de 25 de novembro de 2023. Após essa data, você tem 7 dias para cancelar e obter um reembolso proporcional. [Saiba mais sobre cancelamento](#)

Próximo

A primeira tela é um resumo da conta que estamos pedindo. Você pode apenas clicar no botão “Próximo” e continuar com o cadastro.



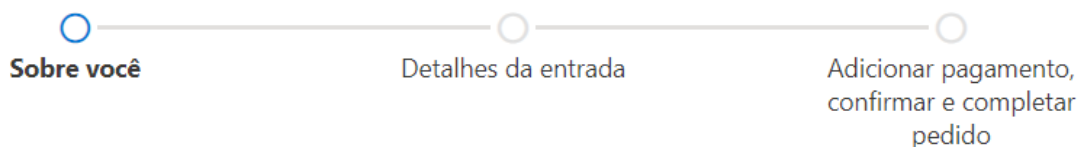
Vamos começar

Insira seu endereço de email corporativo ou de estudante. Verificaremos se você precisa criar uma conta para o Microsoft 365 Business Basic.

Email

Próximo

1 - Após inserir o e-mail pessoal, basta clicar em Configurar conta.



Vamos começar

Parece que você precisa criar uma nova conta. Vamos ajudá-lo a começar!
Continuar como **seuemail@gmail.com**.

Configurar conta

Alterar meu email

2 - Depois é só preencher os campos e clicar em Próximo.

Conte-nos sobre você

País ou Região

Brasil

Nome *

Segundo nome (Opcional)

Sobrenome *

Código do país

(+55) Brasil

Número de telefone comercial *

Nome da empresa *

Tamanho da empresa *

Selecione uma das opções abaixo



- ☐ Eu gostaria de obter informações, dicas e ofertas sobre Soluções para Empresas e Organizações e outros produtos e serviços da Microsoft. Para saber mais ou cancelar a assinatura a qualquer momento, veja a [Política de Privacidade](#).
- ☐ Gostaria que a Microsoft compartilhasse minhas informações com parceiros específicos para que eu possa receber informações sobre seus produtos e serviços. Para saber mais ou cancelar a inscrição a qualquer momento, veja a [Política de Privacidade](#).

Próximo

3 - Nessa etapa, seu telefone celular será verificado para validação dos seus dados. Escolha a opção para validação: SMS ou Ligação. Após escolher a opção e preencher com seu número, clique em Enviar código de verificação.

Conte-nos sobre você

Uma mensagem de texto ou uma chamada telefônica ajuda-nos a garantir que é você.

Insira um número que não seja VoIP ou chamada gratuita.

☒ Enviar-me uma mensagem de texto

☐ Ligar para mim

Código do país

(+55) Brasil

Número de telefone

11999999999

Não salvaremos este número de telefone nem o usaremos para qualquer outra finalidade.

Enviar código de verificação

Voltar

4 - Após a escolha da opção de validação, basta inserir o código de verificação enviado.

Sobre você

Detalhes da entrada

Adicionar pagamento,
confirmar e completar
pedido

Conte-nos sobre você

Uma mensagem de texto ou uma chamada telefônica ajuda-nos a garantir que é você.
Insira um número que não seja VoIP ou chamada gratuita.

☒ Enviar-me uma mensagem de texto

☐ Ligar para mim

Código do país

(+55) Brasil

Número de telefone

11999999999

Não salvaremos este número de telefone nem o usaremos para qualquer outra finalidade.

Digite o código de verificação

Não entendeu ou precisa de um novo código? [Tente novamente.](#)

Verificar

Alterar meu número de telefone

5 - Após a validação do seu número, basta escolher o nome do seu usuário, domínio e senha e clicar em Próximo.

Sobre você

Detalhes da entrada

Adicionar pagamento,
confirmar e completar
pedido

Como você fará o login

Esse nome de usuário é o que você usará para fazer login sempre que usar seus aplicativos. O nome de domínio é uma sugestão. Você pode alterar seu domínio agora ou mais tarde a qualquer momento com seu próprio domínio personalizado.

Nome de usuário

JoaoSilva

@

Nome do domínio

JS

.onmicrosoft.com

Salvar

Senha

.....

Confirmar senha

.....

Próximo

6 - ATENÇÃO! Esta é uma etapa muito importante. Ao clicar em “Próximo”, você será direcionado para a tela de pagamento. NÃO É NECESSÁRIO REALIZAR NENHUM PAGAMENTO. O processo de criação da conta já finalizou, você pode simplesmente fechar a aba, adicionar uma forma de pagamento é um passo opcional.

The screenshot shows a three-step progress bar at the top: 'Sobre você' (completed), 'Detalhes da entrada' (completed), and 'Adicionar pagamento, confirmar e completar pedido' (current step). Below the progress bar is the heading 'Quantidade e pagamento'. A table lists the product details:

Nome do produto	Preço (BRL)	Quantidade	Subtotal (BRL)
Microsoft 365 Apps para Pequenos e Médios negócios Compromisso anual Fatura mensal	R\$ 52,80/Licenses/mês	1	R\$ 52,80

Below the table, it states 'Subtotal antes do imposto: R\$ 52,80'. At the bottom, there is a blue button labeled 'Adicionar uma forma de pagamento'.

7 - Com a criação da conta finalizada, é só se direcionar ao Power BI. Quando você acessar esse link, irá entrar na sua conta. Após o login, surgirá uma tela contendo as etapas necessárias para ativar a conta. Na etapa 1 - Vamos começar, você terá o e-mail conectado da sua conta do Power BI. Para seguir, basta clicar em Continuar:

Você selecionou Microsoft Power BI

1

Vamos começar

Você está conectado com

Usuario [redacted]@Usuari [redacted].onmicrosoft.com

Selecione **Continuar** para começar.

Continuar

[Fazer logoff e criar outra conta](#)

2

Criar sua conta

3

Detalhes de confirmação

8 - Na etapa 2 - Criar sua conta, você precisará inserir seu número de telefone novamente, e clicar em Introdução:

Você selecionou Microsoft Power BI

1 Vamos começar

2 Criar sua conta

Oi **Usuario**,

Precisamos apenas de mais algumas informações antes de você começar

Email

UsuarioAnonimo@UsuarioWork.onmicrosoft.com

País ou Região

Brasil

Número de telefone comercial

Eu entendo que a Microsoft pode contactar-me sobre minha avaliação.

- ☐ Eu gostaria de obter informações, dicas e ofertas sobre Power BI, Soluções para Empresas e Organizações e outros produtos e serviços da Microsoft. Para saber mais ou cancelar a assinatura a qualquer momento, veja a [Política de Privacidade](#).
- ☐ Gostaria que a Microsoft compartilhasse minhas informações com parceiros específicos para que eu possa receber informações sobre seus produtos e serviços. Para saber mais ou cancelar a inscrição a qualquer momento, veja a [Política de Privacidade](#).

Ao escolher **Introdução**, você concorda com nossos [termos e condições](#) e [Política de Privacidade](#).

Introdução

3 Detalhes de confirmação

9 - Na etapa 3 - Detalhes de confirmação, você irá clicar em Introdução:

Você selecionou Microsoft Power BI

- 1 Vamos começar
- 2 Criar sua conta
- 3 Detalhes de confirmação

Obrigado por inscrever-se para Microsoft Power BI

Seu nome de usuário é

Usuario @Usuario .onmicrosoft.com

[Introdução](#)

Agora, você já pode usar a conta que você acabou de criar.

Lembrando que essa é uma possibilidade de criar uma conta corporativa, no entanto, se você possuir uma conta corporativa, é possível usar a conta da sua empresa, da escola ou faculdade para acessar o Power BI.

É importante dizer que o software continua gratuito e ainda é possível desenvolver painéis e relatórios de forma local, ou seja, na sua própria máquina. No entanto, sem a conta corporativa não é possível acessar o Power BI Serviço.

Apêndice G – Preparação do Ambiente MySQL Workbench

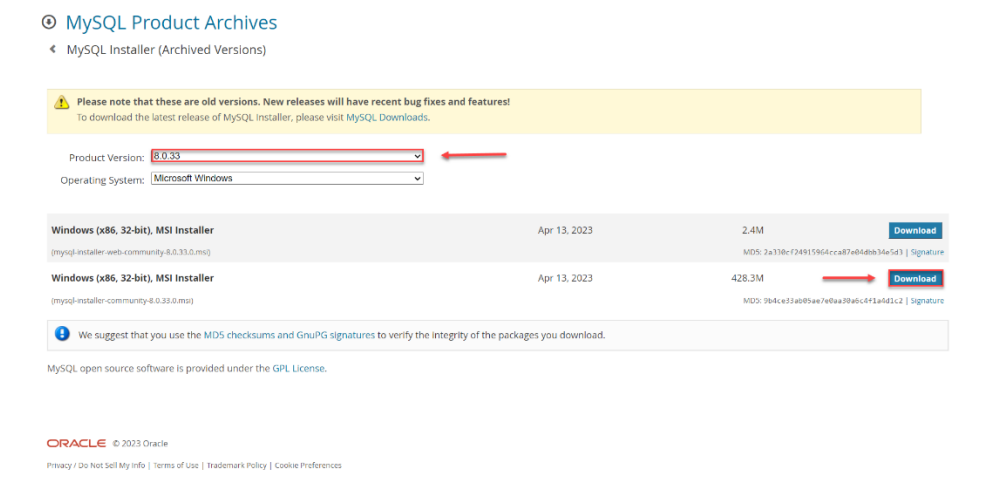
O MySQL Workbench é uma ferramenta de administração e modelagem de bancos de dados MySQL que possui uma interface gráfica amigável. É amplamente utilizado por profissionais que trabalham com banco de dados para **projetar, desenvolver e administrar bancos de dados relacionais**. Com o MySQL Workbench, é possível criar, modificar e excluir tabelas, definir relacionamentos entre elas, executar consultas SQL, importar e exportar dados, entre outras funcionalidades. Ele oferece uma visão geral do esquema do banco de dados, permitindo que as pessoas usuárias visualizem e modifiquem sua estrutura de forma intuitiva.

Para instalar o MySQL Workbench, vamos seguir as seguintes etapas:

Passo 1: Acesse o site oficial do MySQL para baixar o instalador do MySQL. **Recomendamos baixar a versão 8.0.33**. Você pode encontrar o site pesquisando "MySQL Installer download" em um mecanismo de busca, ou clicando no link abaixo que já está direcionado para as outras versões:

[Download MySQL Installer](#)

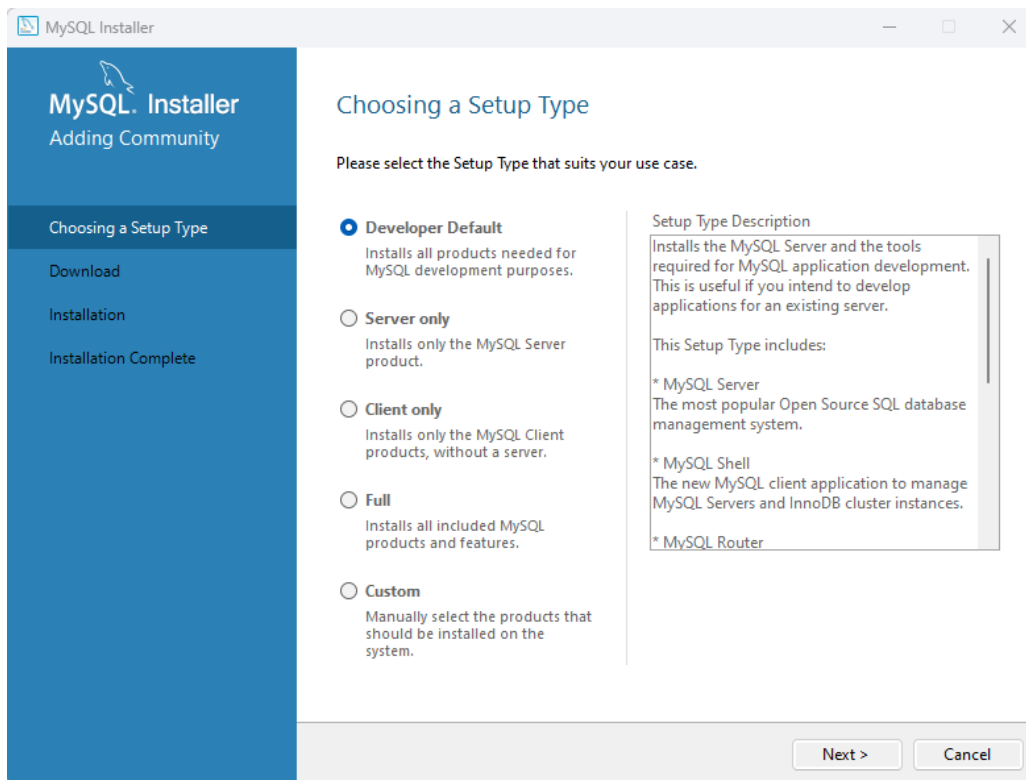
Passo 2: Na página de downloads do MySQL Installer, você encontrará diferentes versões disponíveis para diferentes sistemas operacionais. Selecione a **versão 8.0.33** e faça o download recomendado.



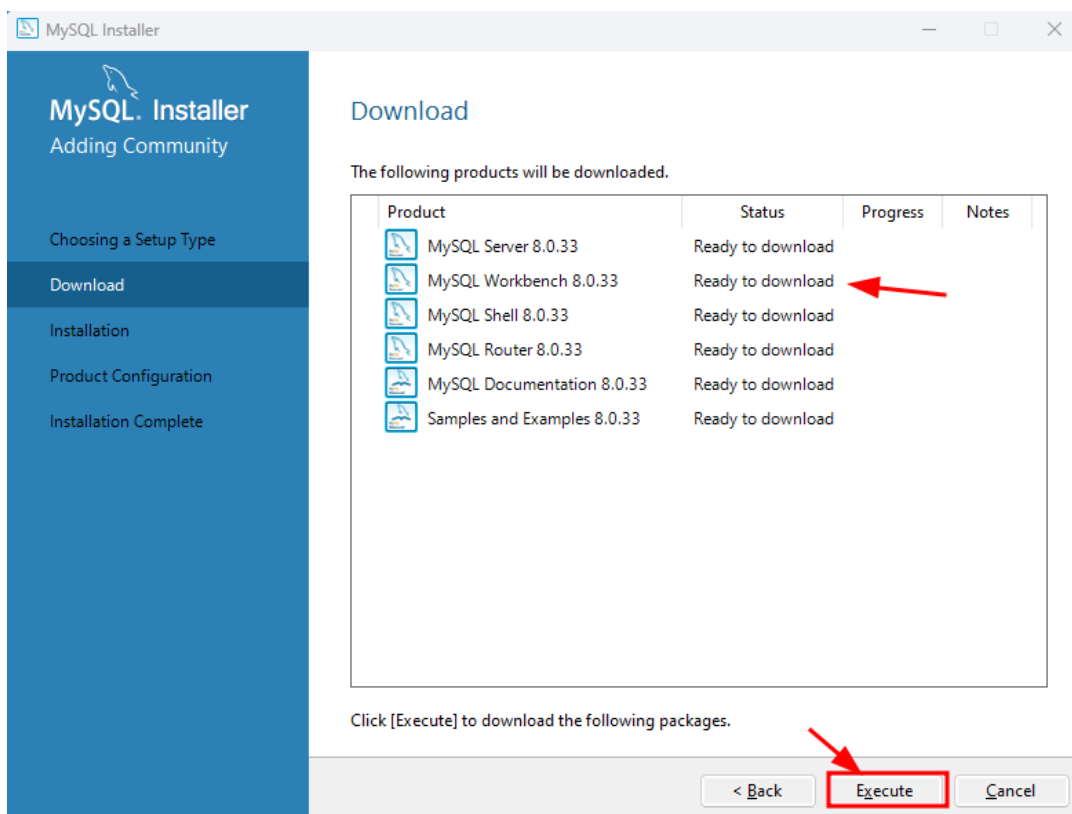
Passo 3: Após escolher o local de download do arquivo e finalizá-lo, localize o arquivo de instalação no seu computador e execute-o. Você pode ser solicitado a confirmar se deseja permitir que o programa faça alterações no seu dispositivo - nesse caso, clique em "Sim" ou "Permitir".

Passo 4: Na primeira tela, selecionamos a opção "Developer Default" e selecionamos o botão "Next >".

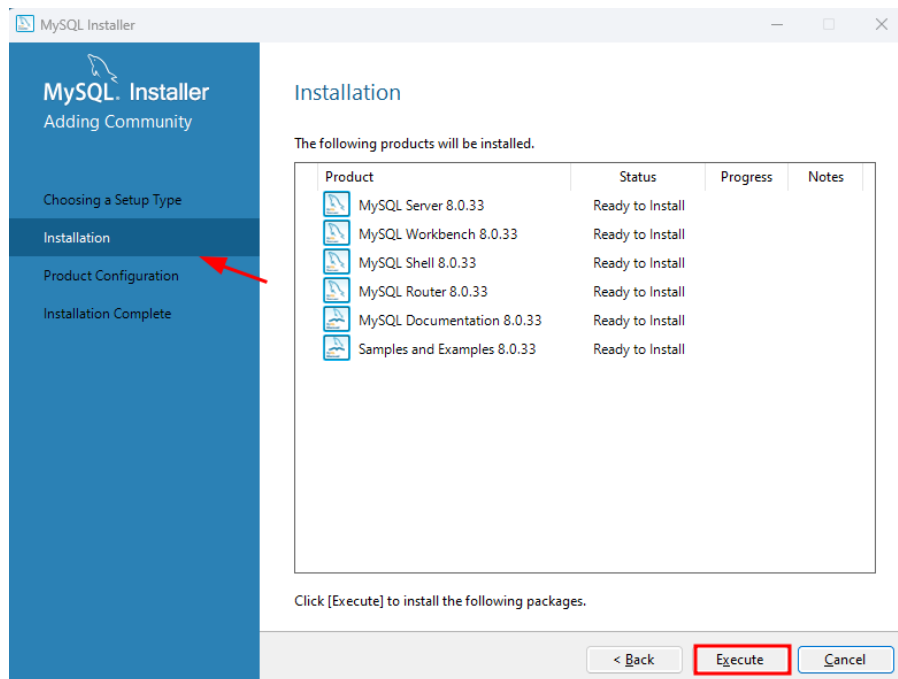
Observação: Caso a opção "Developer Default" não apareça para você, basta marcar a opção **Full** e seguir os próximos passos.



Passo 5: Na próxima tela, de Download, e durante a instalação, você será solicitado a selecionar os componentes que deseja instalar. Certifique-se de que a opção "MySQL Workbench" está na lista para que seja baixado no seu sistema. Após isso, selecione a opção "Execute".



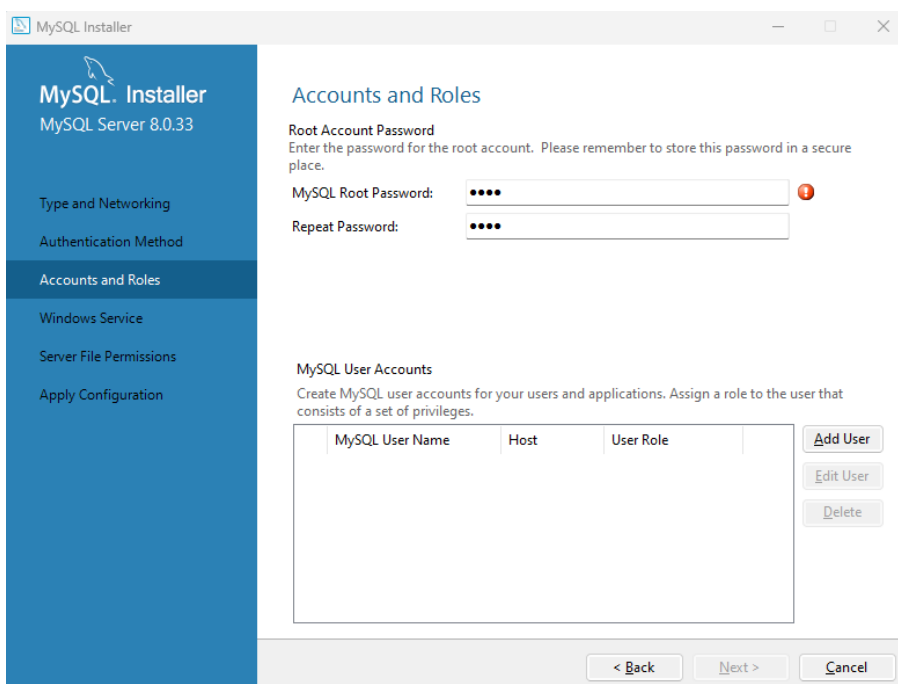
Passo 6: Devemos aguardar até que o processo de download esteja concluído. Uma vez concluído, aperte o botão "Next >" e prossiga para o processo de instalação. Na tela de instalação, vamos selecionar a opção "Execute".



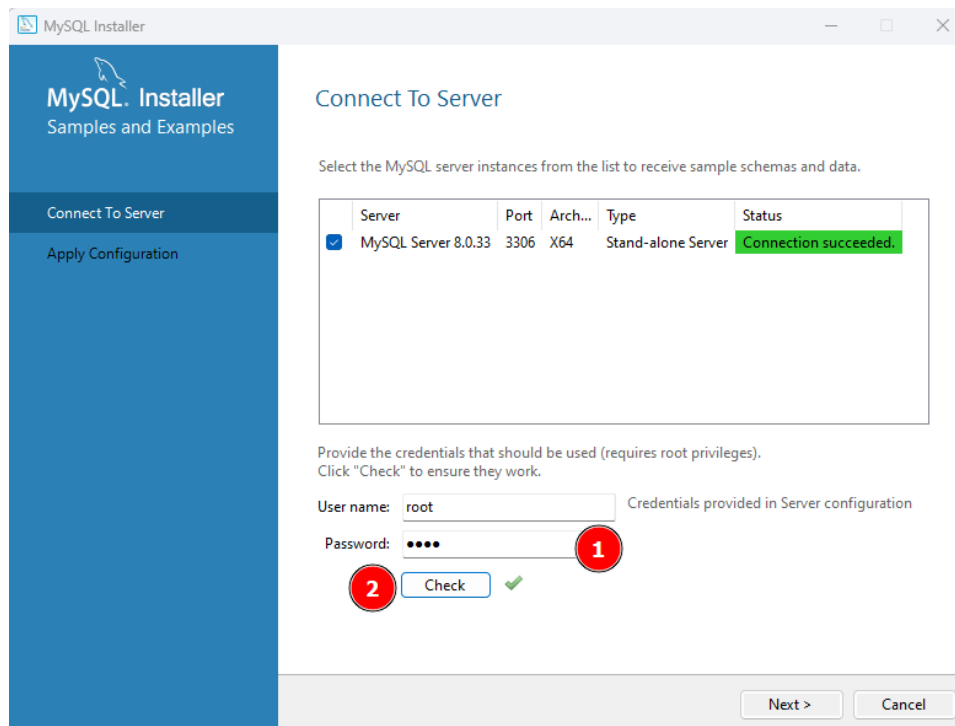
Passo 7: Devemos aguardar até que o processo de instalação esteja concluído. Uma vez concluído, aperte o botão “Next >” e prossiga até o final do assistente de instalação.

Nas próximas telas, irão aparecer configurações relativas a configuração do servidor do MySQL Server. Apenas trabalhe sobre os campos que considerar achar necessário. Para esses cursos, não iremos nos aprofundar nessas opções, apenas na criação da conta de administrador necessária (Root account), localizada na tela de “Account and Roles”.

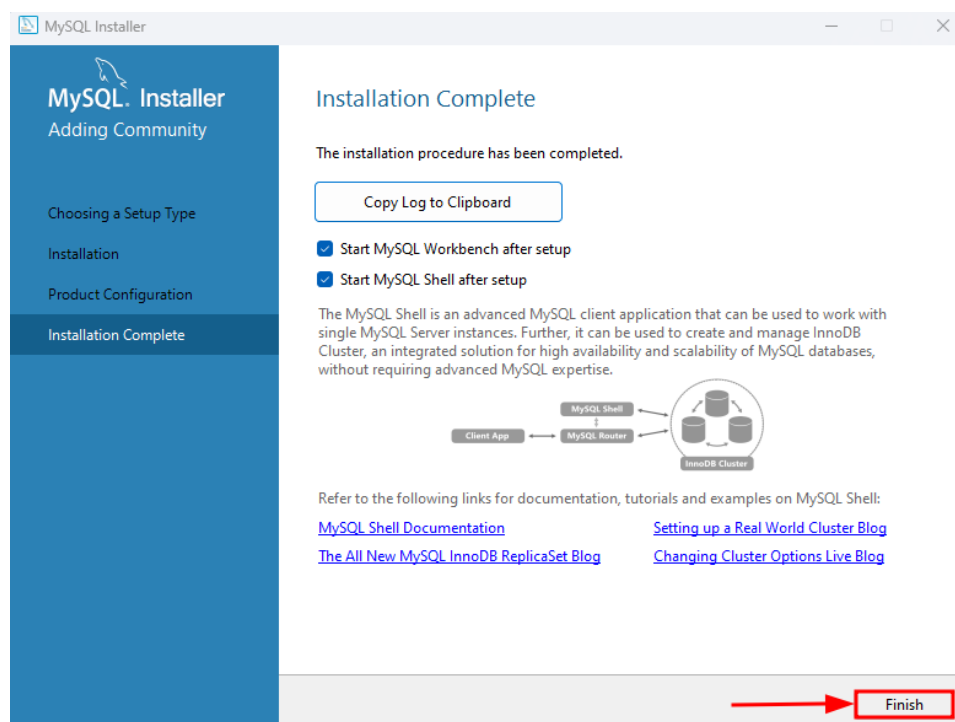
Passo 8: Na janela “Account and Roles”, precisamos criar uma senha para o administrador do MySQL. Digite a senha desejada e a salve um local de fácil acesso. Caso esse banco venha a ser utilizado em outros projetos, lembre-se de construir uma senha forte. Após digitar a senha, prossiga selecionando “Next >” para prosseguir até chegar na tela “Connect to Server”.



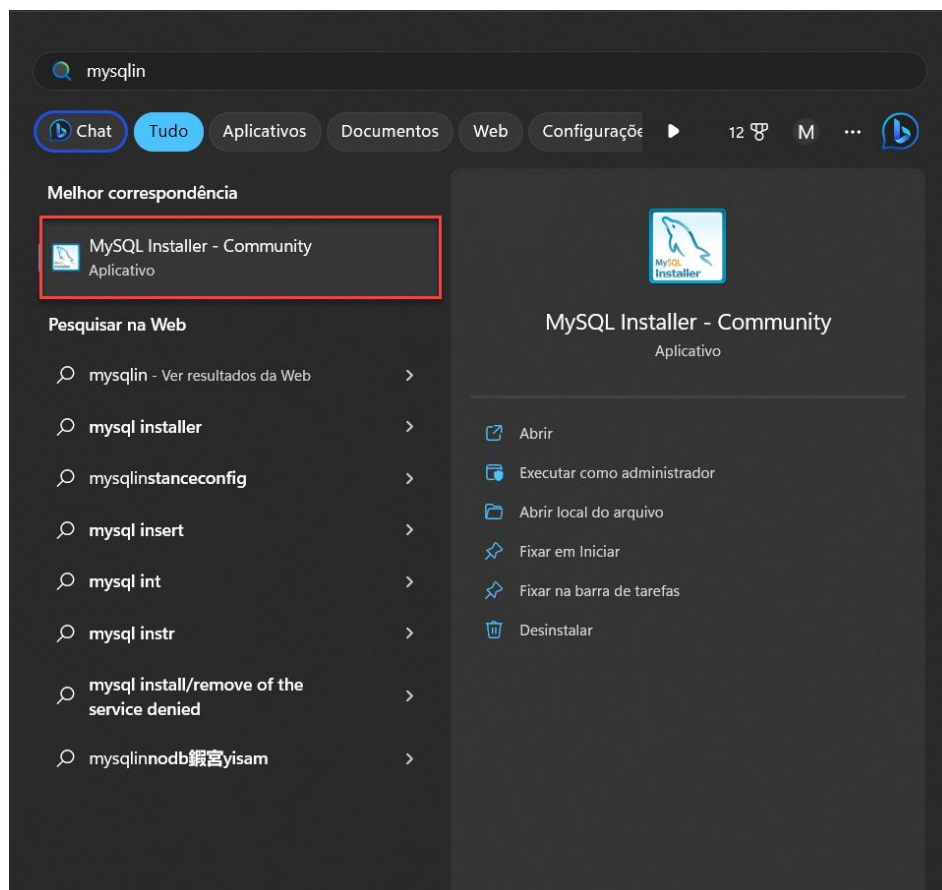
Passo 9: Agora, precisamos digitar a senha que foi criada para o usuário root no passo anterior. Após isso, clique no botão “Check” e siga prosseguindo para as próximas telas.



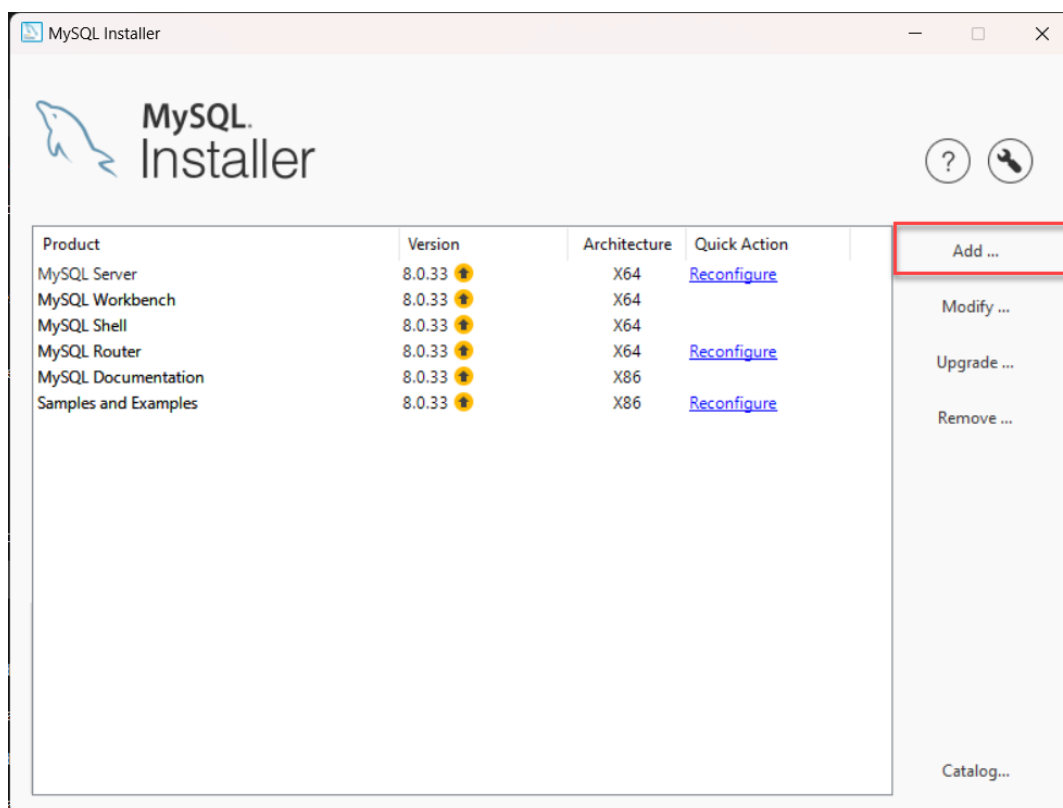
Passo 10: Quando a instalação estiver completa, você irá para a página de título “Installation Complete”. Para finalizar a instalação, clique em “Finish”.



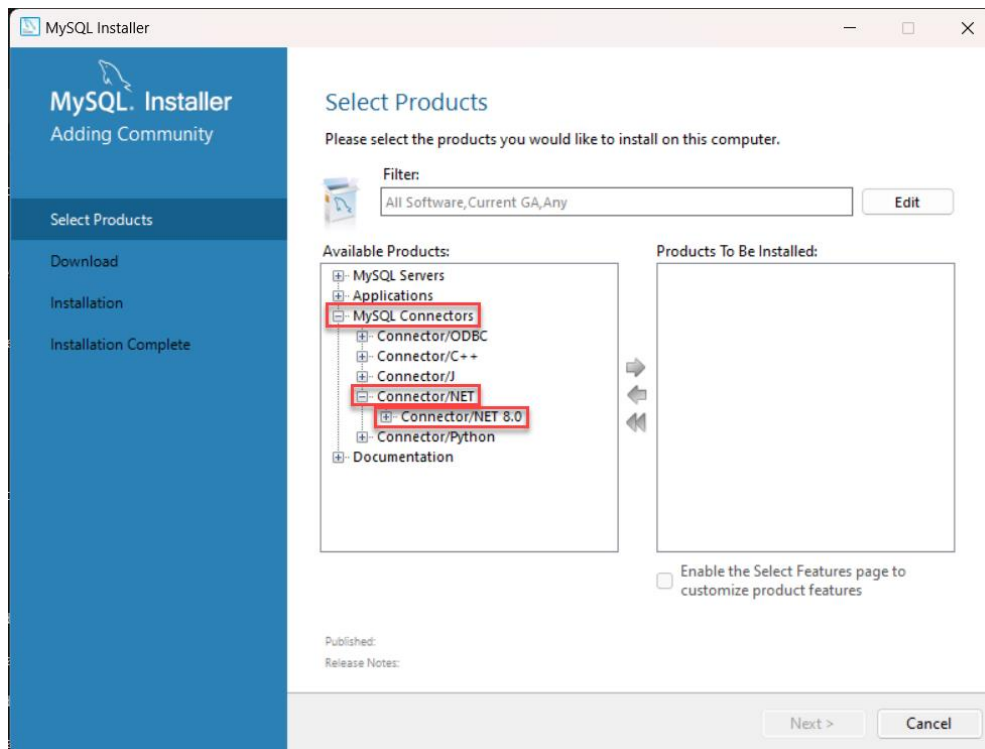
Passo 11: Após finalizar a instalação desses componentes, precisamos adicionar mais um componente: o Connector. Esse componente é necessário para que o Power BI identifique o MySQL. Para isso, pesquise pelo MySQL Installer no seu computador e o inicie:



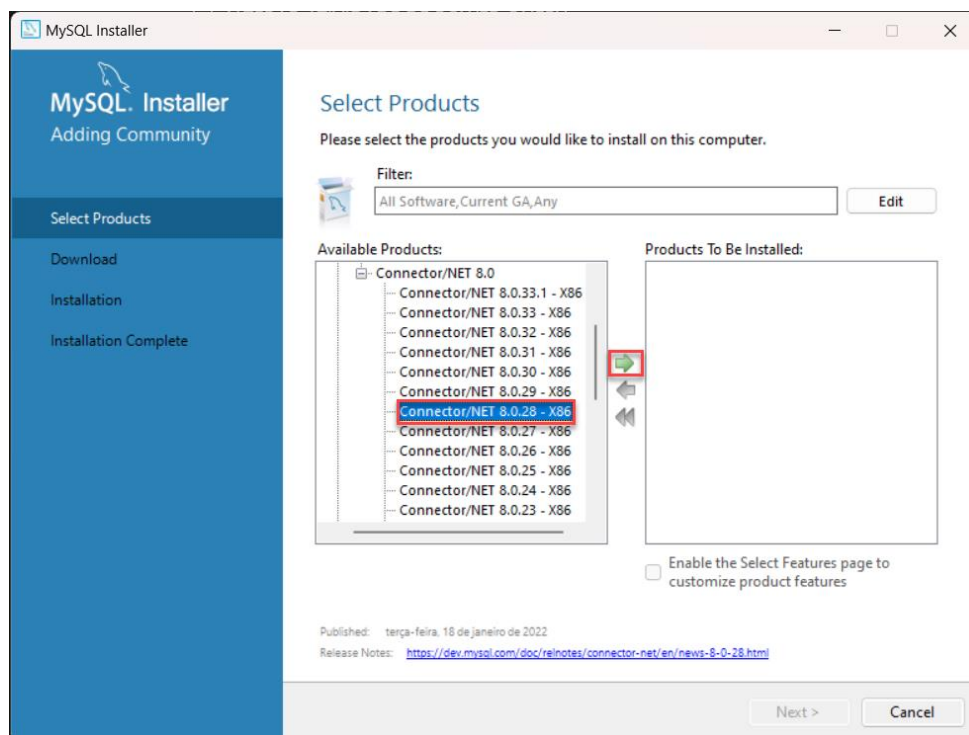
Passo 12: Na tela inicial do MySQL Installer, você irá encontrar os componentes do MySQL que estão instalados. Para adicionar o componente do Connector, clique no botão “Add” na direita:



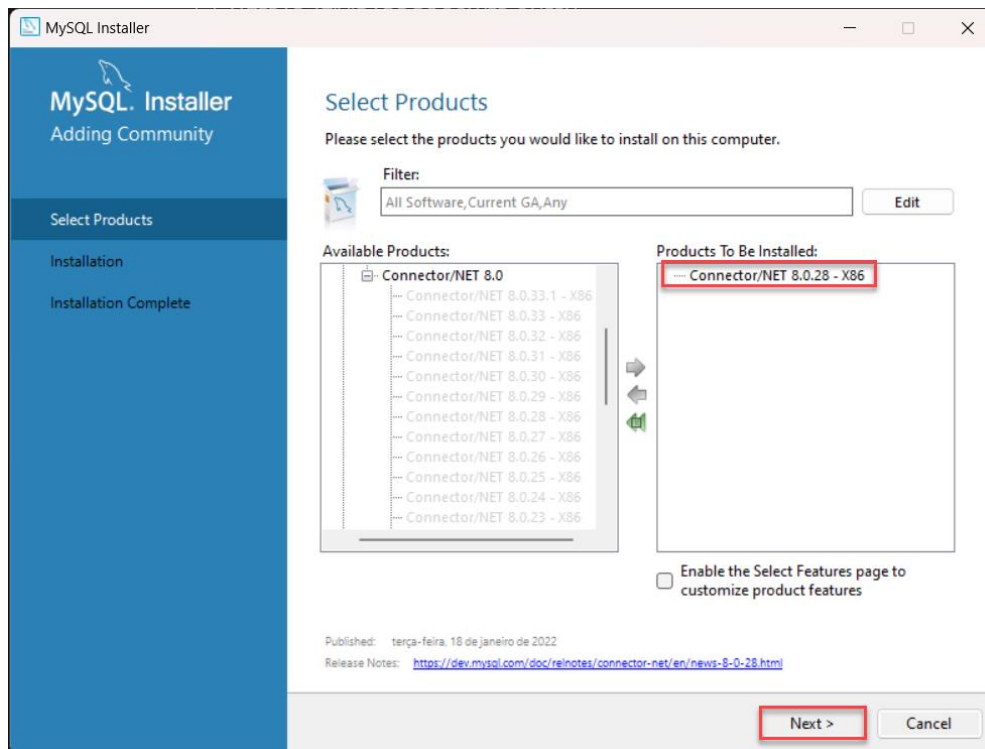
Passo 13: Na tela de Seleção de Produtos, expanda o campo “MySQL Connection”, depois o campo “Connector/NET” e, por fim, o campo “Connector/NET 8.0”.



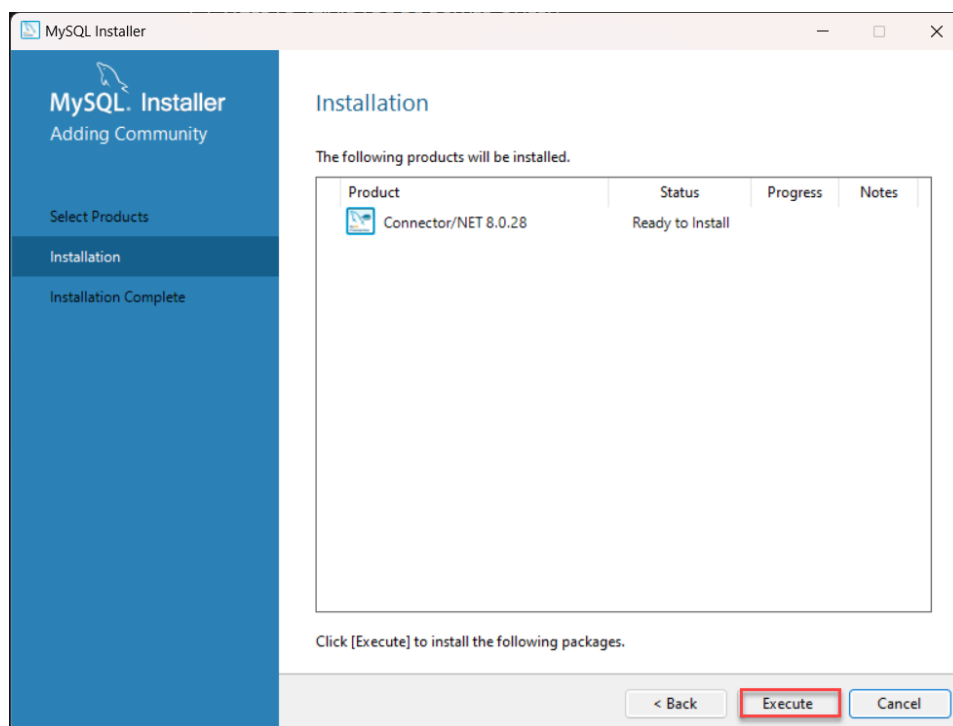
Passo 14: Ao expandir o campo “Connector/NET 8.0”, várias versões do Connector serão listadas. Para que o MySQL seja reconhecido pelo Power BI, as versões 8.0.29 em diante não funcionam. Nesse caso, vamos **utilizar a versão 8.0.28**. Para isso, clique no campo dessa versão e depois na seta verde ao meio, para que ela seja selecionada:



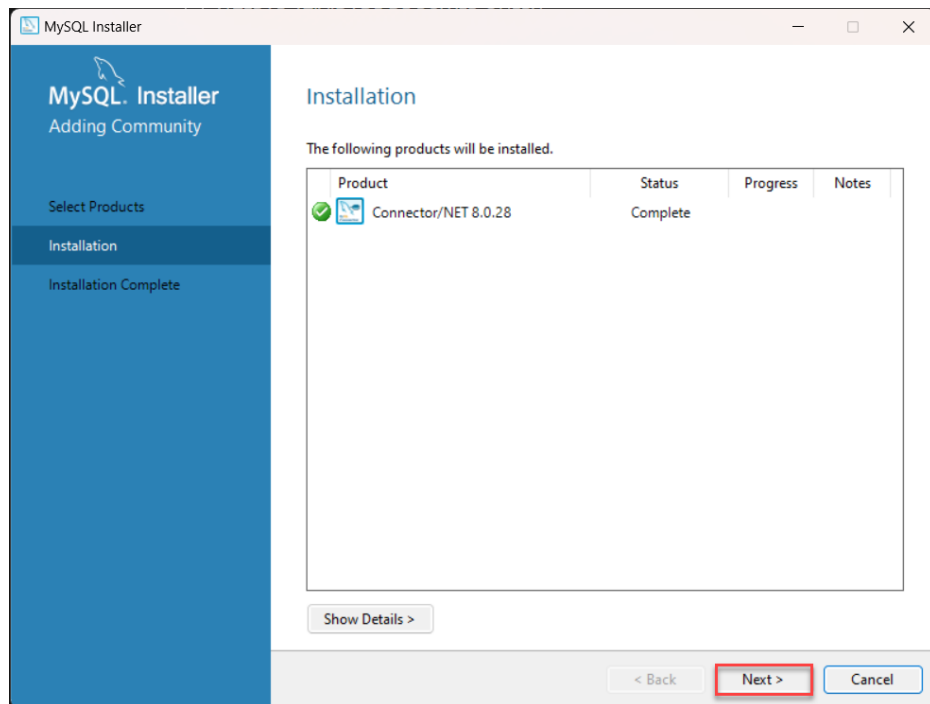
Passo 15: Após selecionar a versão adequada do conector, clique em Next:



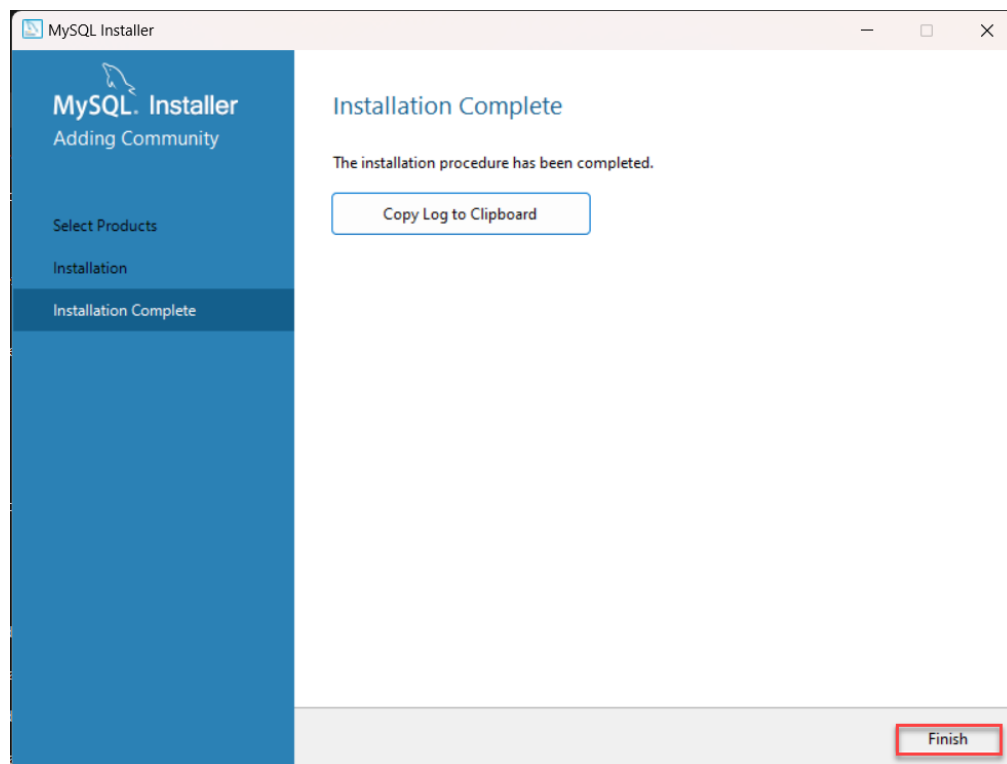
Passo 16: Na janela de Instalação, clique em Execute:



Passo 17: Assim que a execução finalizar, clique em Next:



Passo 18: Para completar o processo, clique em Finish. Com isso, finalizamos a instalação do conector do MySQL:



Pronto, temos instalado em nossa máquina o MySQL com o conector adequado

Apêndice H – Criação do banco de Dados no MySQL

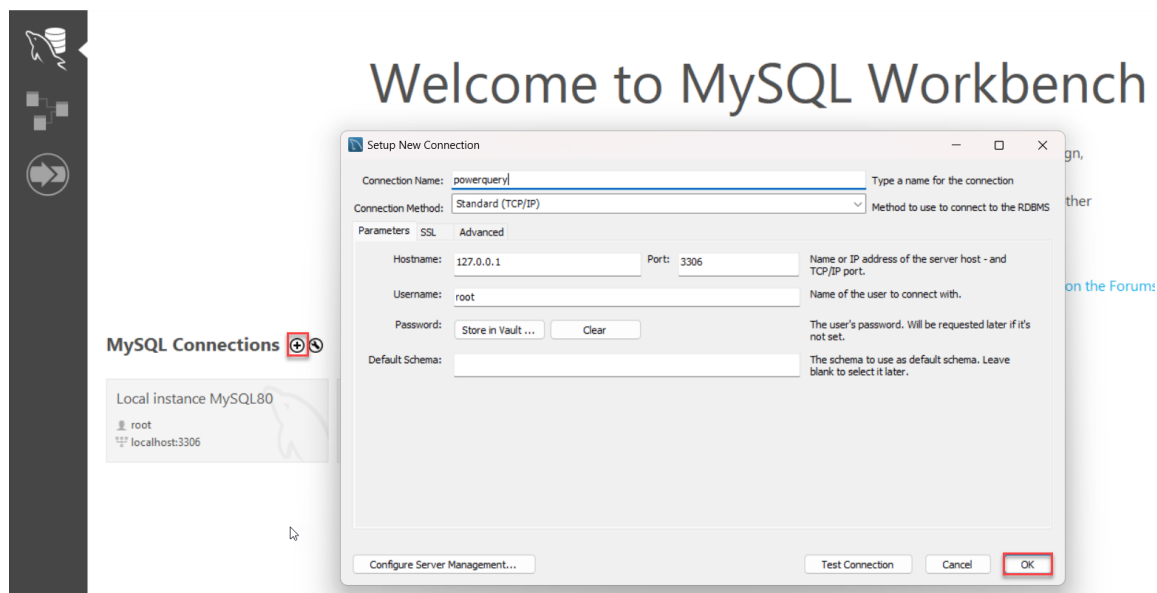
Observação: Caso ao decorrer do passo a passo abaixo encontre dificuldades na conexão com o MySQL, uma alternativa é usar os dados em formato CSV, você pode baixá-lo neste [link](#).

A criação do banco de dados no MySQL será realizada através da utilização de um arquivo de dump, que contém um registro da estrutura de tabela e ou dados de um banco de dados.

Para isso, vamos realizar algumas etapas:

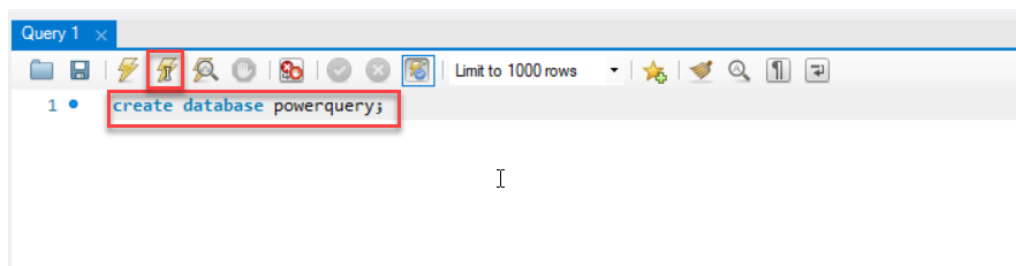
1 - Primeiramente, você terá que baixar o arquivo de dump (o arquivo de dump é um arquivo com extensão .sql. Para este projeto está disponível neste [link](#)).

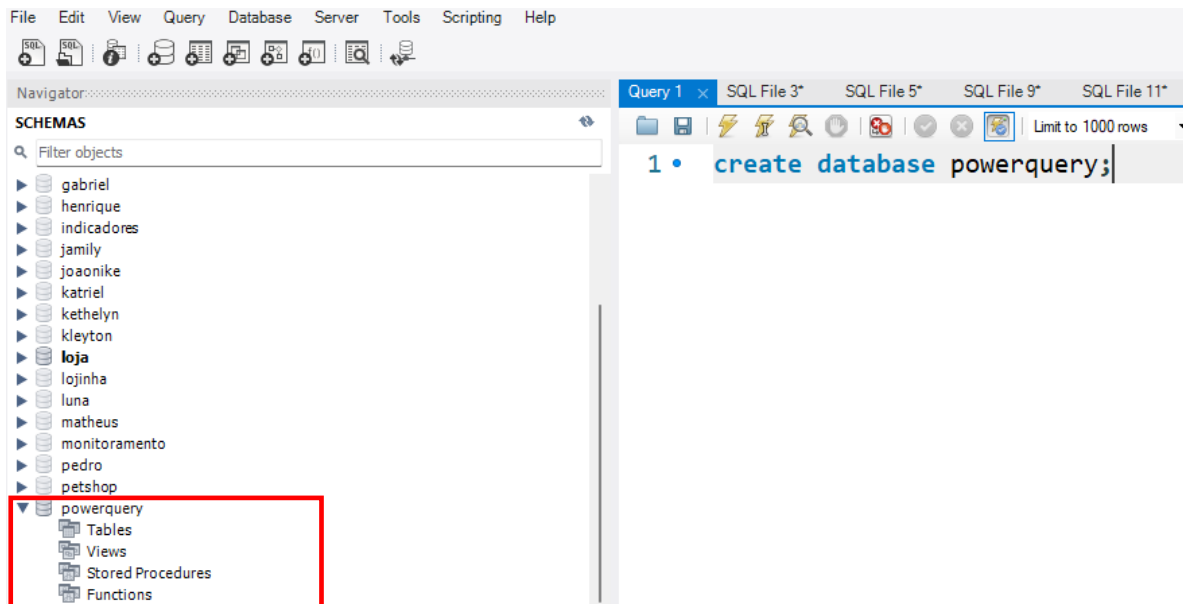
2 - Antes de mais nada, vamos criar o banco de dados no MySQL Workbench. Para isso, você irá abri-lo, e ao lado do título "MySQL Connections", clicar no botão de "+" para criar uma nova conexão. Ao fazer isso, uma janela será aberta, você irá preencher os campos necessários e, ao final, clicar em OK:



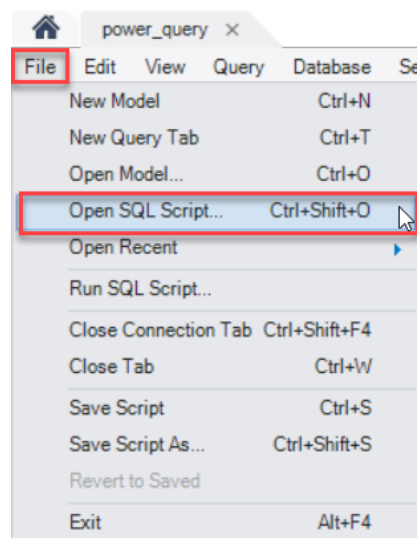
3 - A nova conexão aparecerá na tela principal, abaixo do título "My SQL Connections". Você deve acessá-la clicando nela. Ao fazer isso, vamos criar o banco de dados "powerquery", utilizando o comando `create database powerquery;`

E clicar no raio destacado no menu:

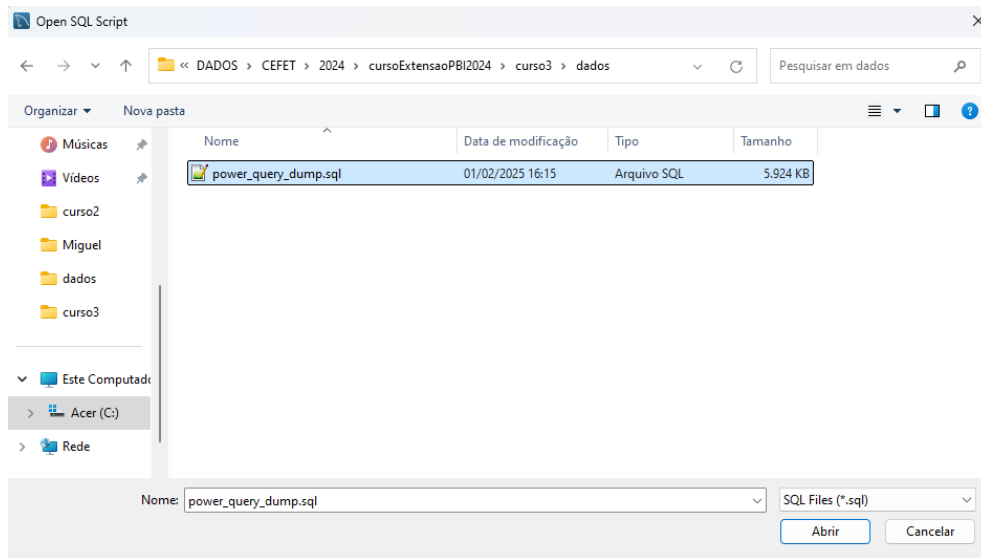




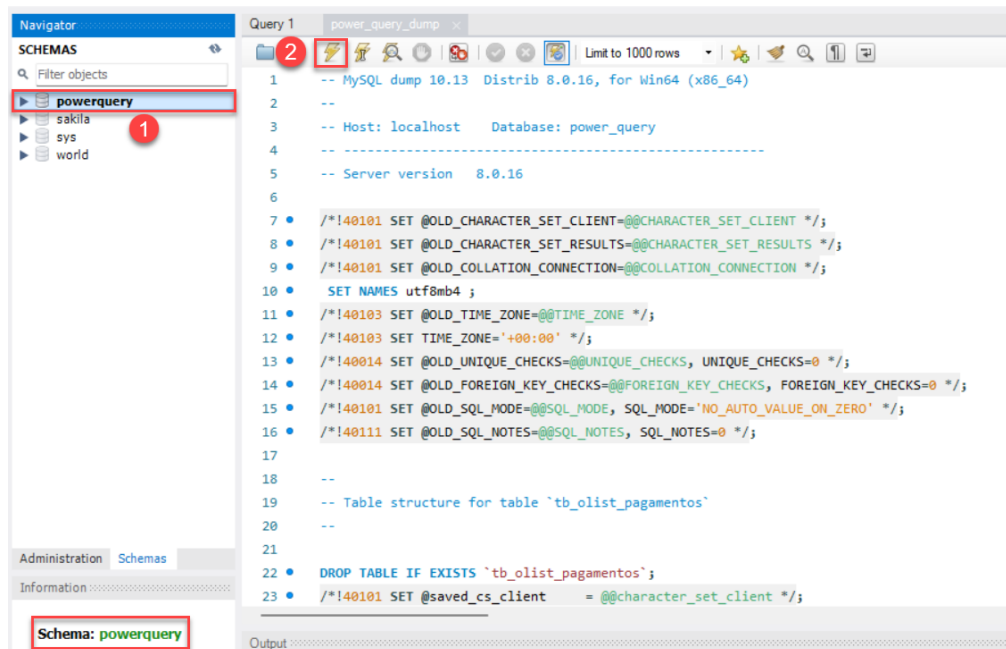
4 - Ao abrir a seção Schemas, no menu Navigator, localizado a esquerda, podemos clicar sobre a opção "Tables" e verificar que não temos nenhuma tabela criada ainda, pois ela será criada com o processo de *dump*. Agora que criamos o banco, vamos realizar o *dump*. Para isso, vamos acessar o menu de "File", no canto superior esquerdo, e clicar na opção "Open SQL Script" para procurar pelo arquivo de *dump* baixado:



5 - Vamos procurar pelo arquivo na pasta onde ele foi salvo e clicar em Abrir.



6 - Após abrir o arquivo, uma nova aba será criada com o código do *dump*. Para executar o comando, precisamos selecionar o banco no qual vamos importar os dados, que no caso é o "powerquery", clicando duas vezes no seu nome no menu lateral esquerdo (após isso, ele ficará em negrito). Com o banco de dados e o script selecionados, agora basta clicar no botão de executar no topo.



7 - Após realizar o dump, retornando à aba inicial onde criamos o banco, podemos realizar uma seleção na tabela de pagamento e verificar o resultado:

Query 1 x power_query_dump

Limit to 1000 rows

```

1 • CREATE DATABASE powerquery;
2
3 • select * from tb_olist_pagamentos;

```

Result Grid

	order_id	payment_installments	payment_sequential	payment_value	payment_type
▶	b81ef226f3fe1789b1e8b2acac839d17	8	1	99	#credit_card
	a9810da82917af2d9aefd1278f1dcfa0	1	1	24	#credit_card
	25e8ea4e93396b6fa0d3dd708e76c1bd	1	1	66	#credit_card
	ba78997921bbcdc1373bb41e913ab953	8	1	108	#credit_card
	42fd880ba16b47b59251dd489d4441a	2	1	128	#credit_card
	298fcd1f73eb413e4d26d01b25bc1cd	2	1	96	#credit_card
	771ee386b001f06208a7419e4fc1bbd7	1	1	81	#credit_card
	3d7239c394a212faae122962df514ac7	3	1	52	#credit_card
	1f78449c87a54faf9e96e88ba1491fa9	6	1	341	#credit_card
	0573b5e23cbd798006520e1d5b4c6714	1	1	52	#boleto
	d88e0d5fa41661ce03cf6cf336527646	8	1	189	#credit_card
	2480f727e869fdeb397244a21b721b67	1	1	142	#credit_card
	616105c9352a9668c38303ad44e056cd	1	1	76	#credit_card
	cf95215a722f3ebf29e6bbab87a29e61	5	1	103	#credit_card

Export: | Wrap Cell Content: | Fetch rows: |

Result Grid
Form Editor
Field Types
Query Stats

8 - Pronto, agora o seu banco de dados está configurado e pronto para conectarmos ao Power BI.