

Curso de Power BI Master

Sumário

Nenhuma entrada de sumário foi encontrada.

Introdução

Linguagem M que é para tratamento de Dados e DAX.

Livros para montagem de dashboards: The Big Book of Dashboard de Wexler Shaffer e Storytelling with Data de Cole Knaflic.

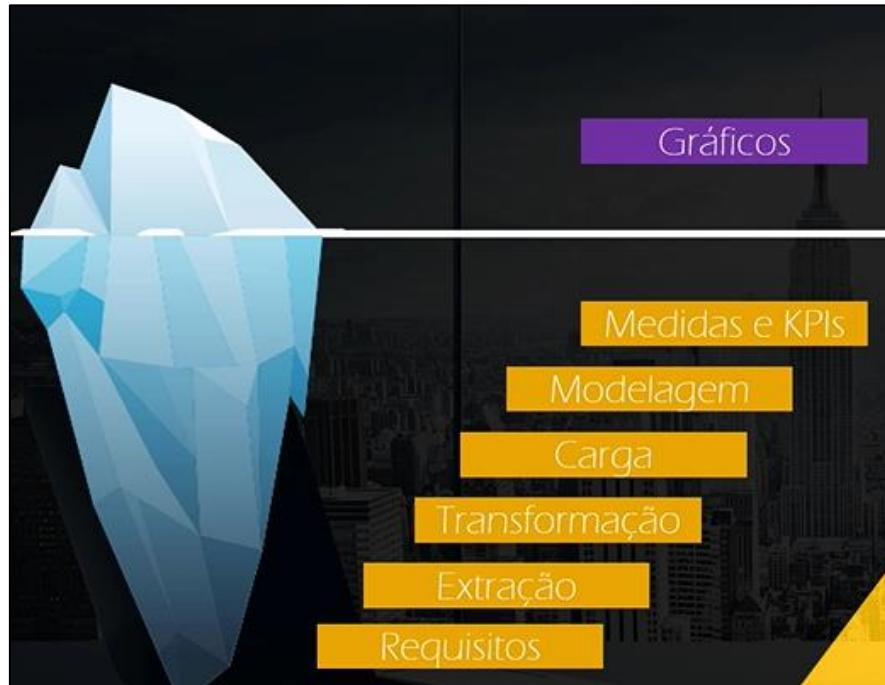
A ferramenta de **Power BI** serve para Analista(Análise de dados), para TI(aplicar governança de dados) e Devs (customização de visual).

Um analista que analisa números e tira os insights e passa um relatório para a gerência da empresa é chamado de Power User. Essa é a parte gráfica e visual.

Enquanto o dev modela os dados, carrega os dados; Carga e Modelagem de dados.

Passo a Passo

1. Levantamento dos requisitos que o usuário deseja, quais as necessidades e literalmente, os requisitos.
2. Extração dos dados, unifica e carrega para algum banco de dados, planilha, **.txt**.
3. Por fim, realiza uma modelagem dos dados, começa no ambiente OLTP (Online *Transaction Processing* ou Processamento de Transações em Tempo Real) ou seja, lá no Banco de Dados, na origem.
4. Criação de Medidas e KPI's, ou seja, os indicadores que os Power Users demandam.



Linguagem M e DAX

As linguagens presentes no **Power BI** são as M e a DAX.

A **Linguagem M** é para tratamento de dados, é utilizada para realizar ETL (*Extract, Transform, Load*), sendo uma linguagem secundária no PB, ou seja, é para otimização e não construção, é uma linguagem estática. É composta por funções.

A **Linguagem DAX**, em contrapartida, é para a parte gráfica, criação de medidas, de análise dos dados.



O que é o Power BI?

É uma ferramenta para transformar dados em gráficos visuais. É compartilhamento dos dados e dimensionar a governança da empresa.

O **Power BI** é proveniente 3 “módulos” do Excel -> **Power Query**, **Pivot** e **View** e criou uma suíte nova, que foi o **Power BI**, podendo trabalhar on-line ou on-premise(off).



Os serviços do Power BI

A manipulação dos dados é feita pelo PB Desktop, a publicação é Online, e com vários dashboards criados, se constrói um “app”. E caso queira publicar os dashboards no site da empresa, utiliza-se o “**embedded**” para a vinculação no site da empresa. E o “On Premise” é algo local, os dados não são levados para a nuvem, cria-se um servidor interno na empresa mas não estará on-line e no cloud.

Power BI x Excel

O PB trabalha com o conceito de COLUNAS. Enquanto o Excel trabalha com o conceito de CÉLULAS, que podem ser mescladas horizontal e verticalmente.

Versões do Power BI

É uma plataforma de visualização de dados, de *Business Intelligence* da **Microsoft**, essencialmente em Cloud. Pode ocorrer on-premise, que são dados “offline”

Power BI Free Version*	Power BI Pro Version **	PowerBI Premium ***
Grátis	USD 9,99 usuário/mensal	Valor Fixo
1 GB de Armazenamento	10 GB de Armazenamento	Infra única
2 Bilhões de Linhas	1 Milhão de Linhas por hora	Sem limites
Publicação Pública	Publicação Privada	

*A versão gratuita não conta com compartilhamento de painéis.

** A versão Pro, é possível informar outro usuário com licença que existe um relatório para ele visualizar ou editar.

** É possível utilizar de colaboração na elaboração dos dashboards.

** A empresa que utiliza a versão Pro tem “uma máquina” no servidor que uma outra empresa aleatória também tem, ou seja, o servidor é disponibilizado para ser trabalhado em Nuvem, porém, esse servidor é compartilhado com outras empresas.

*** Na versão Premium, é um servidor totalmente dedicado a empresa, sem compartilhamento do servidor com demais empresas. A infraestrutura não será compartilhada.

Outras versões

Power BI para celular – É a versão de visualização em mobile.

Power BI Embedded – É uma versão onde, por exemplo, o site da empresa se utiliza de um servidor da Microsoft para publicar algumas informações no site. Não está alocada nos serviços. O Embedded é através de links das Nuvens (ou seja, dos serviços na nuvem) e aí após o licenciamento do Embedded, é possível utilizar as informações da nuvem no site da empresa.

Servidor de Relatório do Power BI – On Premise

Todos os serviços das Microsoft no PB são exclusivamente em Cloud Computing, porém, o On Premise é no local, na própria empresa. São dados sensíveis. É utilizado através do **SQL Server** instalado na máquina.

Gratuito	Pro
\$0	\$9,90 / Mês
Conexão com todas as fontes de dados Publicação na Nuvem Publicamente Export de Relatórios Limite de 1GB de Dados Máximo de 2 Bilhões de Linhas Máximo de 16 Mil Colunas	Conexão com todas as fontes de dados Publicação Privada na Nuvem Export de Relatórios Análise no Excel Assinatura E-mails Workspace 8 Atualizações diárias

Para o desenvolvimento desktop nenhum email é necessário, no entanto, para a publicação on-line, apenas e-mails corporativos e educacionais são aceitos.

Você precisa de análise avançada, suporte de Big Data, relatórios no local e na nuvem?

Adicionar Power BI Premium

- Enterprise BI, análise de Big Data, relatórios na nuvem e no local
- Controles avançados de administração e implantação
- Computação de nuvem dedicada e recursos de armazenamento
- Permita que o usuário consuma o conteúdo do Power BI.

Power BI Premium

\$4,995

Preço mensal por computação de nuvem dedicada e recursos de armazenamento

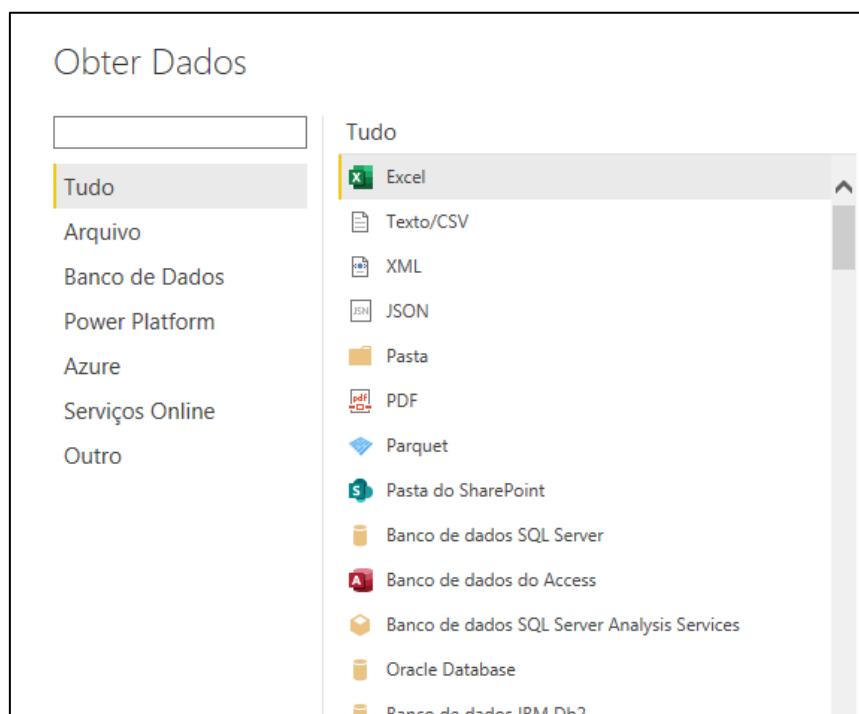
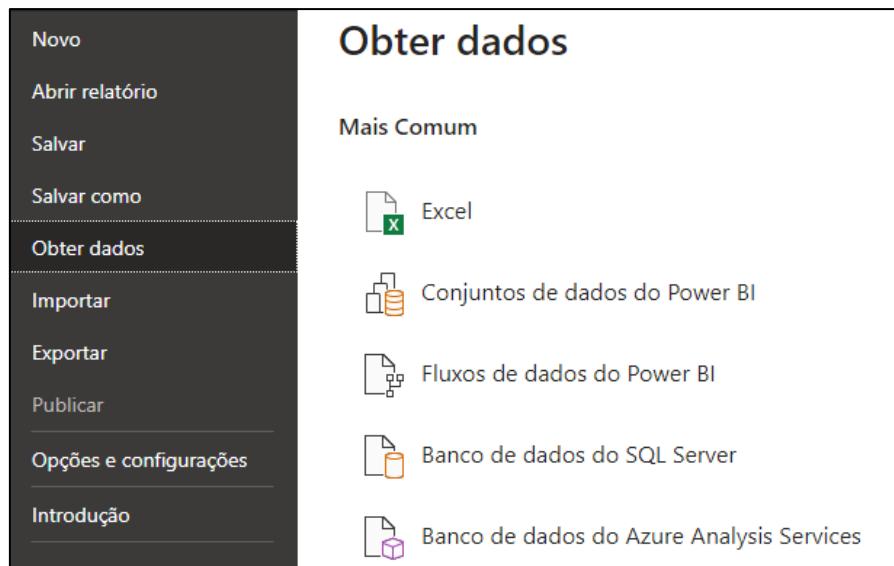
[Solicitar consulta >](#)

Interface do Power BI

Obtenção de Dados

Exemplo 01 – Planilha de Vendas

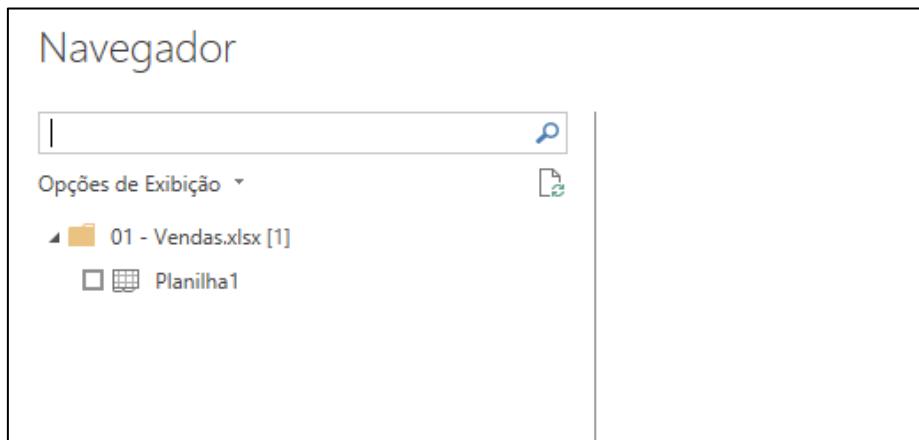
Para iniciar, vamos em Arquivo > Obter Dados > Obter dados para Começar. Existem muitos e diversos formatos para se escolher a obtenção dos dados.



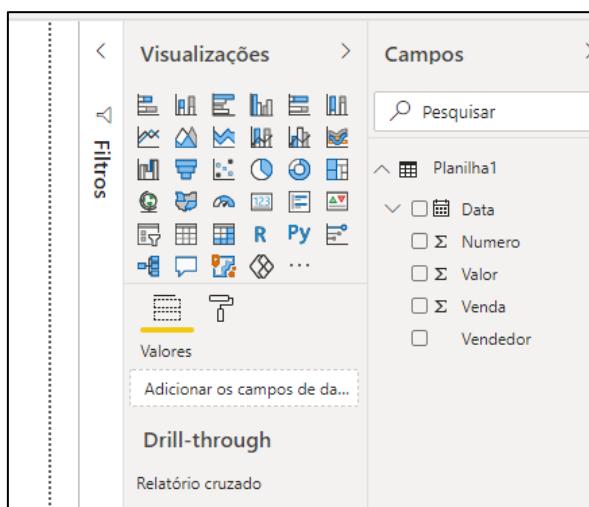
Para esse exemplo, vamos em Arquivo > Excel > Conectar > Selecionar a planilha > 01 - Vendas.xlsx

Observação durante o curso: NUNCA ABRA OS DADOS CARREGANDO. SEMPRE ABRA PRIMEIRO UTILIZANDO A TRANSFORMAÇÃO AO INVÉS DE “CARREGAR” OS DADOS.

Após estabelecer essa conexão, a seguinte tela é aberta:



Clique no flag de Planilha1 e selecione. Aparentemente nada acontece. Porém, os campos agora estão alocados a direita. Podemos agora, selecionar quaisquer uns dos gráficos nas Visualizações e a partir dessa seleção, arrastaremos os campos para o gráfico.



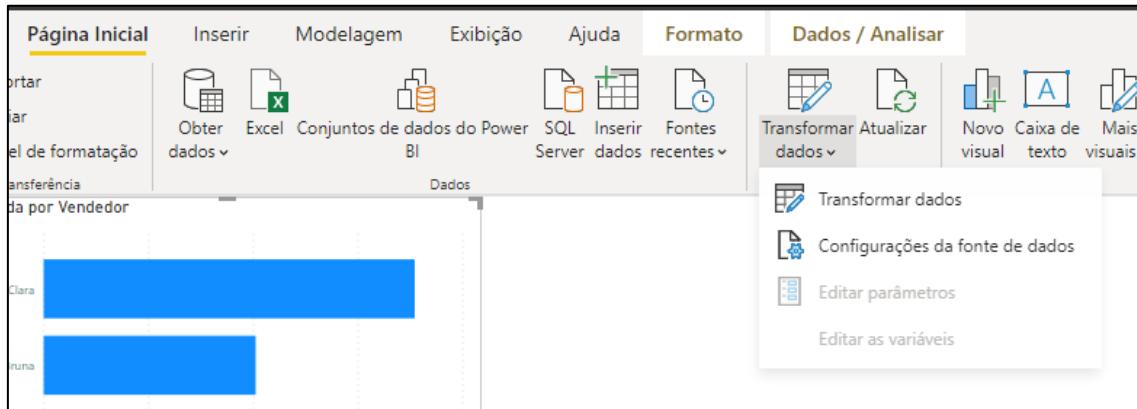
Selecionamos o primeiro gráfico. Veja, foi selecionado o campo Vendedor para o Eixo e o campo SomaVenda para os Valores,



Transformação de Dados

Caso não tenha feito a transformação antes de carregar os dados, podemos transformar os dados após carregar os dados para o PB.

Para transformar esses dados, note que por exemplo, os valores das vendas não estão formatados. Podemos transformar, vamos em Página Inicial > Transformar Dados > Transformar Dados



A janela do **Power Query Editor** deverá abrir

Vale lembrar que o **Power Query** é a camada ETL, ou seja, é onde se busca os dados e realiza o tratamento dos dados diretamente na fonte.

Podemos atualizar essa base de dados utilizando o Refresh, imagine que novos dados podem ser imputados. Basta dar um Atualizar que os dados vão ser atualizados.

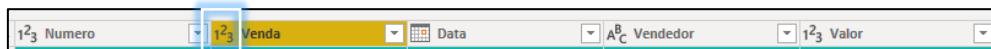
Alterar a Tipagem dos Dados

Nesse primeiro exemplo, podemos perceber que, trocamos o componente visual para Tabelas. Mas perceba que ele está “somando” as vendas? E não é para somar, nesse caso, Vendas significa qual foi o número da Venda, por exemplo, poderia ser uma Nota Fiscal.

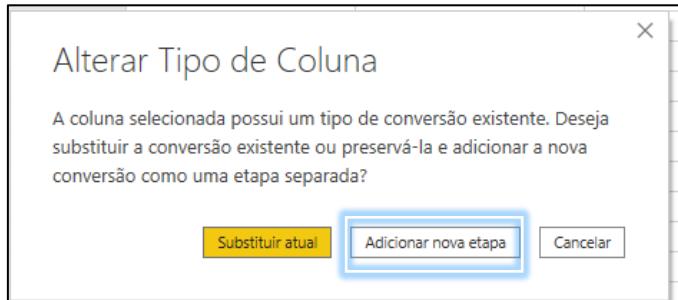
Vendedor	Venda
Ana	1021
Bruna	4052
Célia	4034
Clara	7088
João	4035
Total	20230

Para corrigir isso, precisamos arrumar a TIPAGEM dos DADOS. Essa é a primeira etapa para a solução do problema.

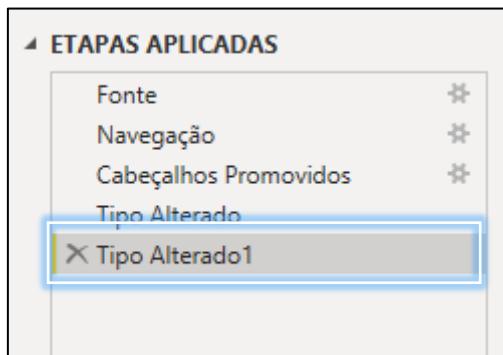
Veja que a Venda nesse caso está classificada como Número. Isso faz com que o PB some as Vendas (que não deveriam, pelo menos nesse caso).



Podemos clicar no botão onde está o “123” e substituir para String. É mensagem segue: Clicar em Adicionar Nova Etapa, pelo menos nesse exemplo.



A direita, apresenta um quadro de “Etapas Aplicadas”, nesse local, tudo que for feito como etapa, caso algo seja feito errado, basta clicar no “x” e a etapa será revertida.



Quando retornar ao PB, após sair do **Power Query (PQ)**, precisaremos aplicar novamente os campos, pois Vendas deixou de ser considerada uma somatória.

Agora, Vendas, são os Números das Notas Fiscais em Texto.

Vendedor	Venda
Ana	1021
Bruna	1002
Bruna	1012
Bruna	1018
Bruna	1020
Célia	1004
Célia	1007
Célia	1009
Célia	1014
Clara	1005
Clara	1008
Clara	1011
Clara	1013

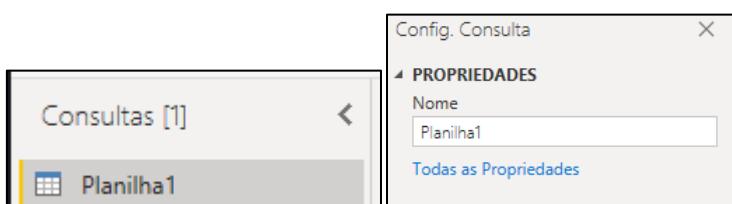
Mais uma vez, voltaremos ao PQ e alteraremos o campo Número de número para String/Texto. Note todas alterações no **TIPO** dos DADOS.



O primeiro passo é praticar a melhor TIPAGEM de dados. Futuramente vamos trabalhar com Medidas.

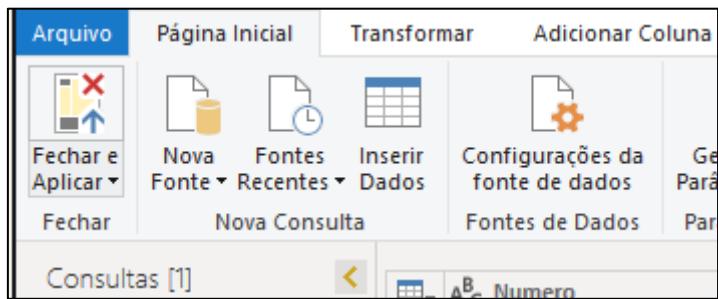
Alterando o nome da Planilha em Propriedades no PQ

Veja que o nome da planilha está como Planilha1 na direita em Propriedades e Planilha1 em Consultas.



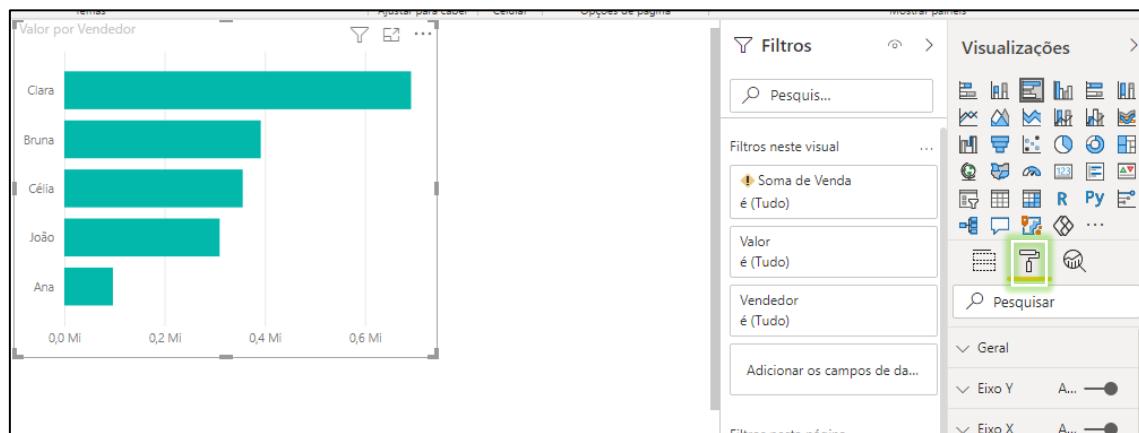
Podemos deixar “friendly” para o usuário e alterarmos o nome que automaticamente será alterado nas Consultas.

Para aplicar todas as etapas e alterações realizadas até o momento, voltaremos em *PQ > Página Inicial > Fechar e Aplicar > Aplicar ou Fechar e Aplicar*



Introdução a Painéis - Formatando Gráficos

Podemos alterar os gráficos através dos Temas, porém, a opção correta é alterar o gráfico diretamente com a formatação. Ao clicarmos no gráfico para selecioná-lo, é possível formatá-lo através do “Pincel de Formatação”. Está marcado no quadro em verde.

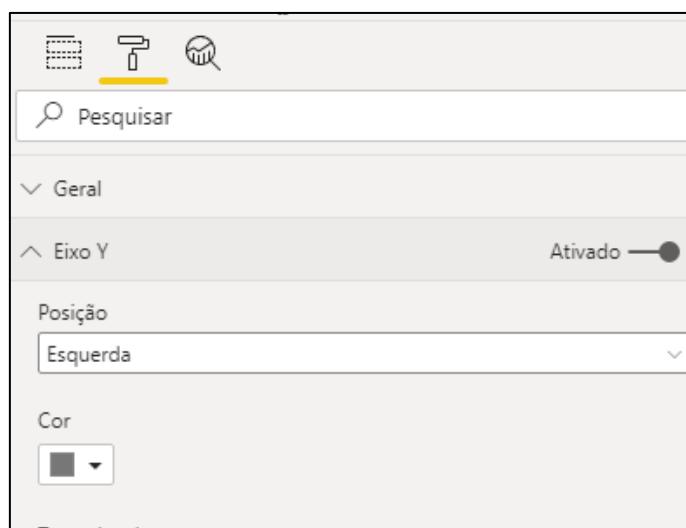


Formatando Eixos

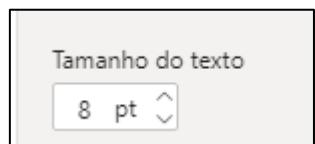
É possível alterar somente alguns elementos de cada eixo, por exemplo, os vendedores estão localizados no Eixo Vertical Y e os valores estão localizados no Eixo Horizontal X.

Formatando Eixo Vertical Y e Horizontal X

Podemos realizar diversas alterações nesses eixos, como por exemplo, alterar o tamanho da fonte dos nomes dos vendedores para melhor visualização. Toda alteração que é feita em um, também pode ser realizada no outro eixo.

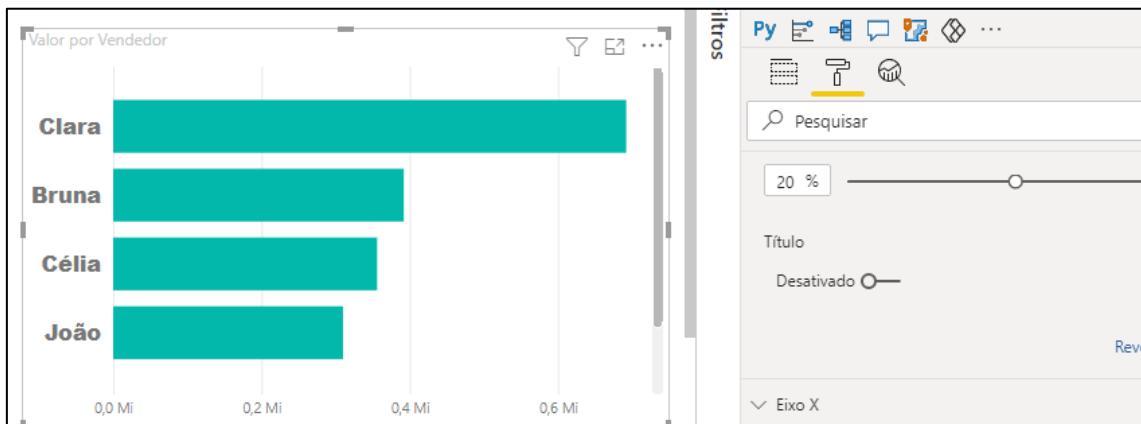


Para alterar o tamanho da fonte, e diversas outras funções.

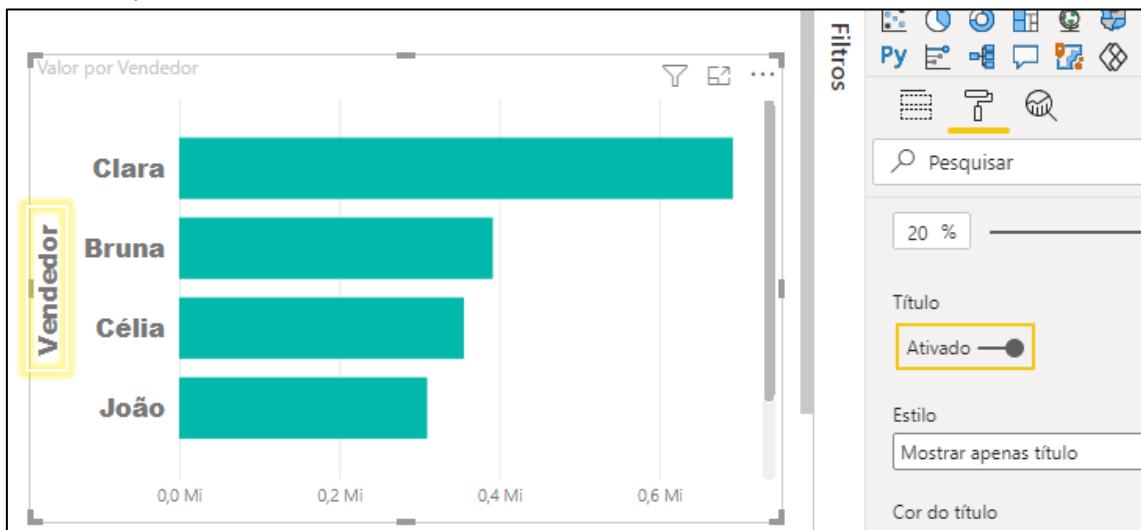


Alterando o Título do Eixo Y

Veja que não há “título” no eixo Y. O botão que ativa, está desabilitado.

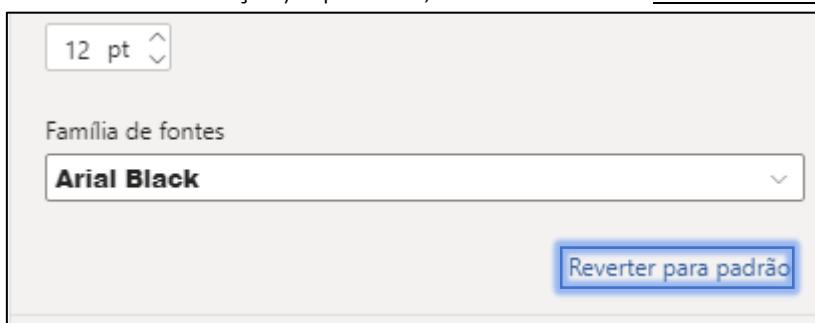


Para isso, podemos ativá-lo e o resultado será:



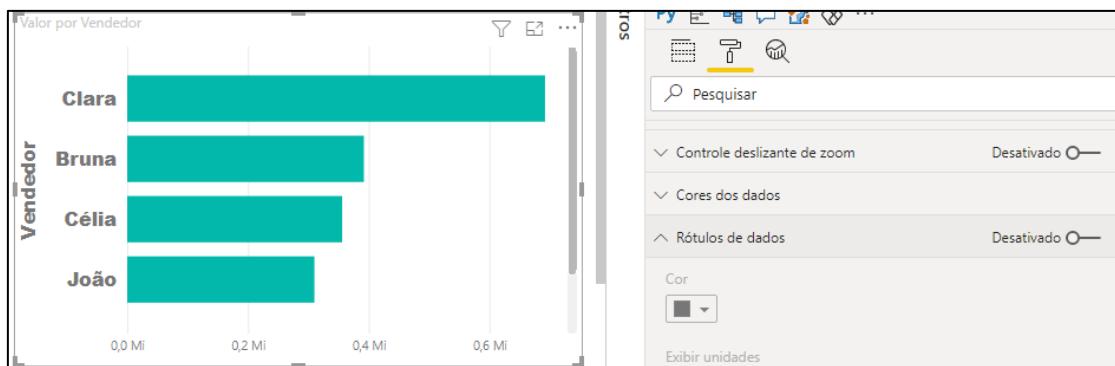
Reverter Padrão

Se por algum motivo, queira retornar tudo para o padrão anterior (o padrão básico, sem nenhuma formatação) é possível, basta clicar em “Reverter Padrão”.

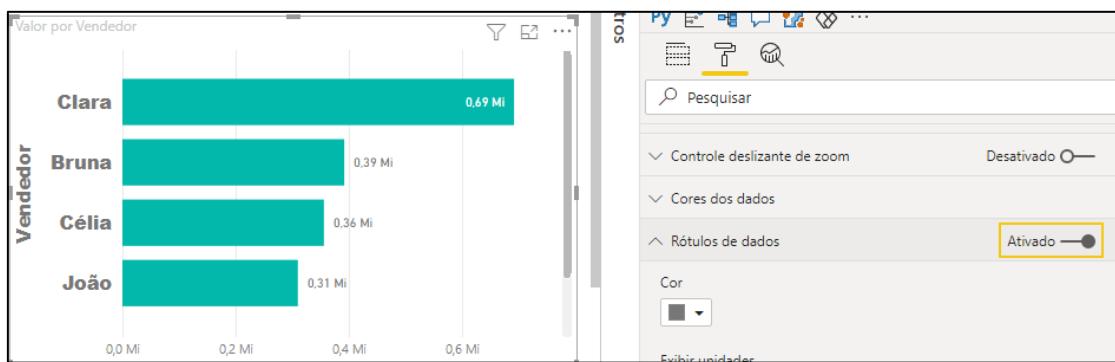


Rótulo de Dados

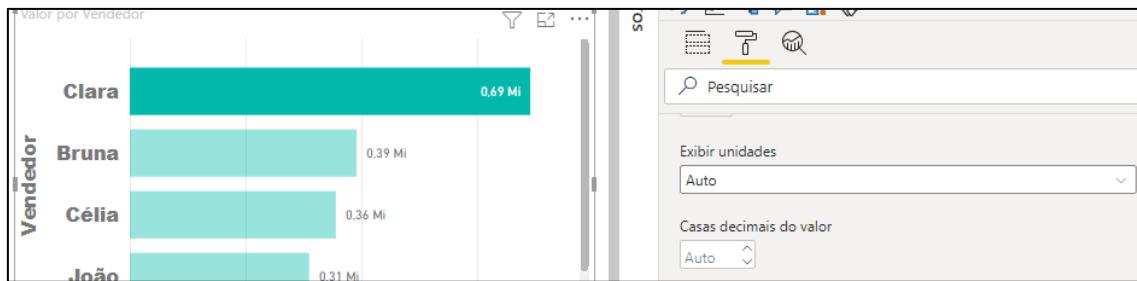
Os rótulos são importantes e por padrão, vêm desativado. Ao ativar, os valores são colocados no gráfico, veja:



Ao ativar o “Rótulo de Dados”, os valores se tornam presentes no gráfico, além do valor no eixo x.



Repare que os valores, porém são milhares, mas, pegamos a Clara que está mostrando 0.69MI. Podemos alterar os formatos pela “Exibir de Unidade”. Por padrão, está como “auto”. E por trabalharmos com números, podemos transformar.

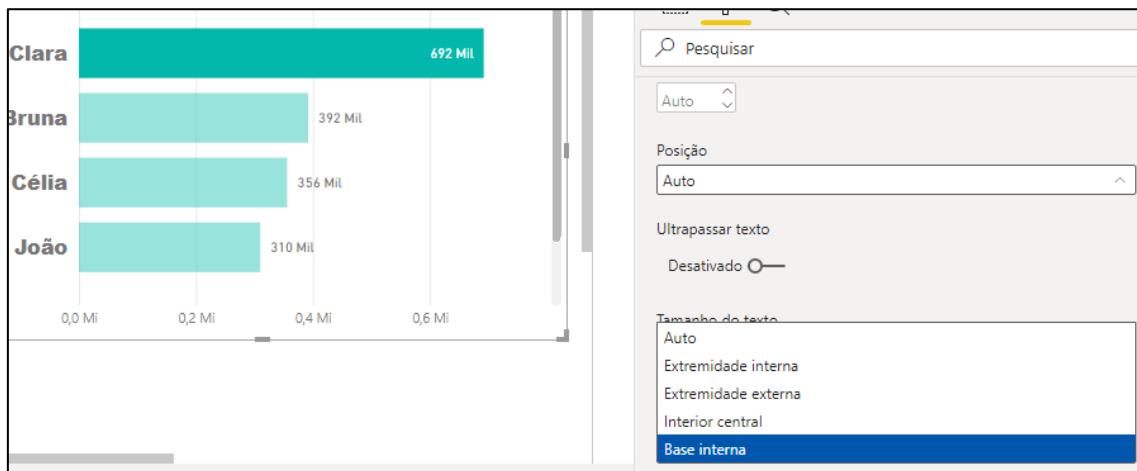


Ao alterar a posição de “auto” para “milhar”, é possível notar que os valores agora estão em formato do milhar.

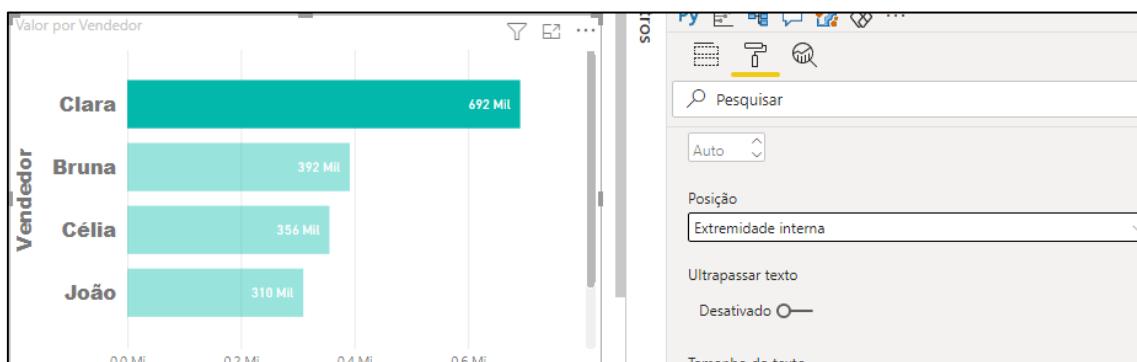


Alterando a Posição dos Rótulos dos Dados

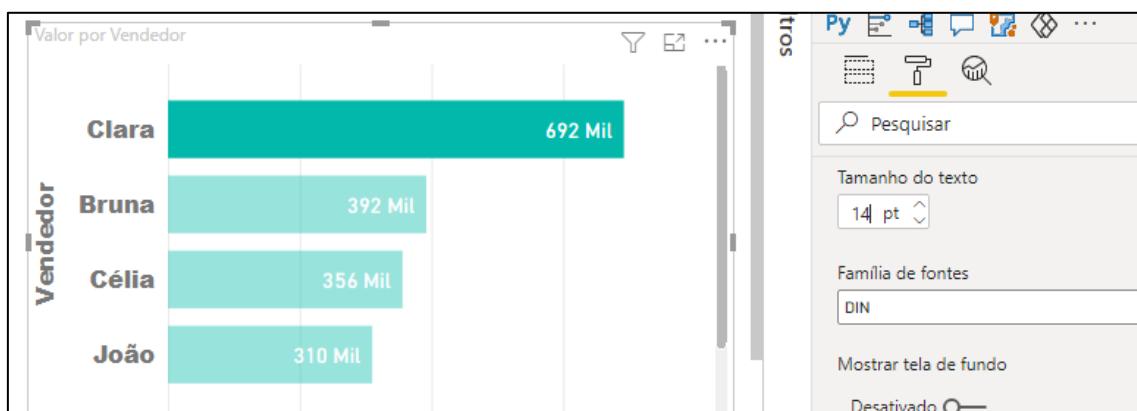
Veja que alguns dados estão “Internos”, ou seja, dentro da barra gráfica, e outros estão “Externos”, ou seja, fora da barra gráfica. O melhor seria seguir uma formatação padrão, isso é, ou todos internos ou todos externos. Para isso, podemos ir em “Rótulos de Dados > Posição”



Veja, deixando marcado a opção “Extremidade Interna”, os valores estão inclusos na barra gráfica.



Claro, os “Rótulos de Dados”, também podem ser alterados igualmente os eixos X e Y. Os tamanhos podem aumentar e diminuir. Basta alterar o “Tamanho do Texto”.

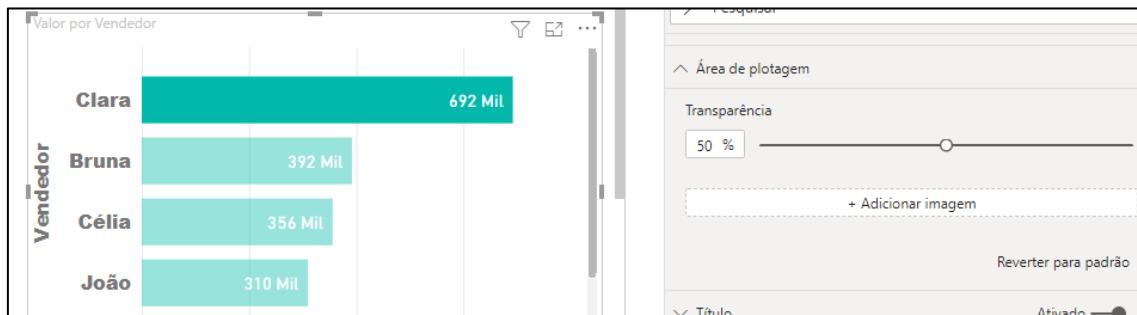


Área de Plotagem

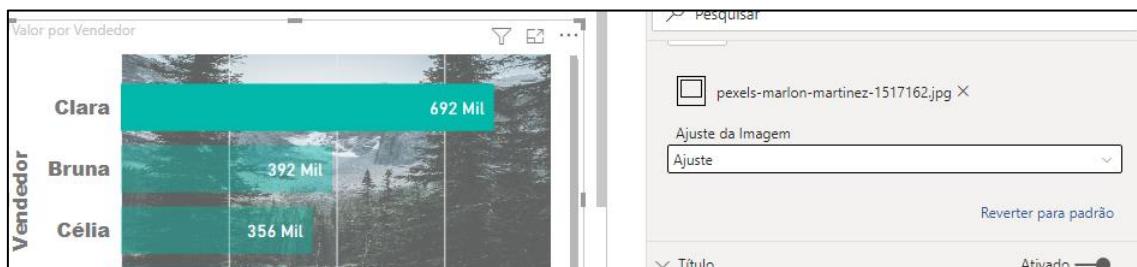
Essa área é utilizada para inserir um “Plano de Fundo” no gráfico. Nessa área, é necessário se atentar nos 3 pontos:

1. Imagem a ser inserida
2. Transparência do plano de fundo (0% significa totalmente visível)
3. Ajuste da Imagem no gráfico

Veja que o gráfico não possui nenhum plano ainda. Para isso, clicaremos em “+ Adicionar Imagens”

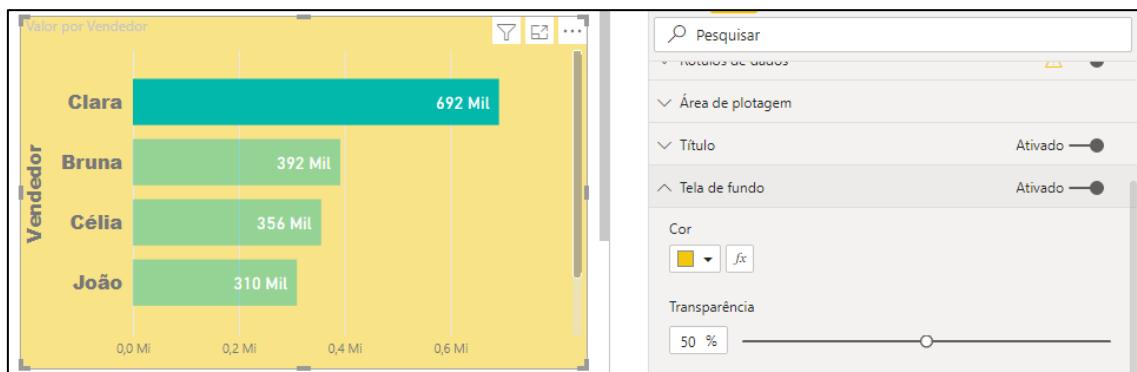


Após adicionarmos a imagem e ajustarmos para o “preencher”.



Tela de Fundo

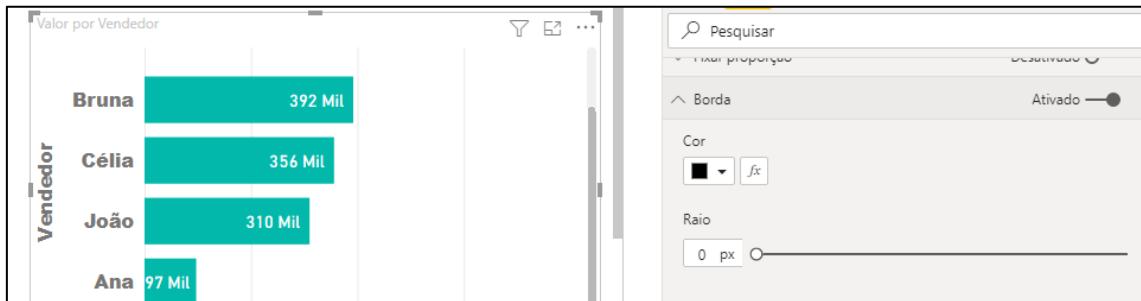
Essa área é muito semelhante a “Área de Plotagem”, porém, a diferença entre as duas é que a “Área de Plotagem” adiciona uma imagem e a “Tela de Fundo”, adiciona um **fundo sólido**. Por padrão, a “Tela de Fundo”, vem desabilitada.



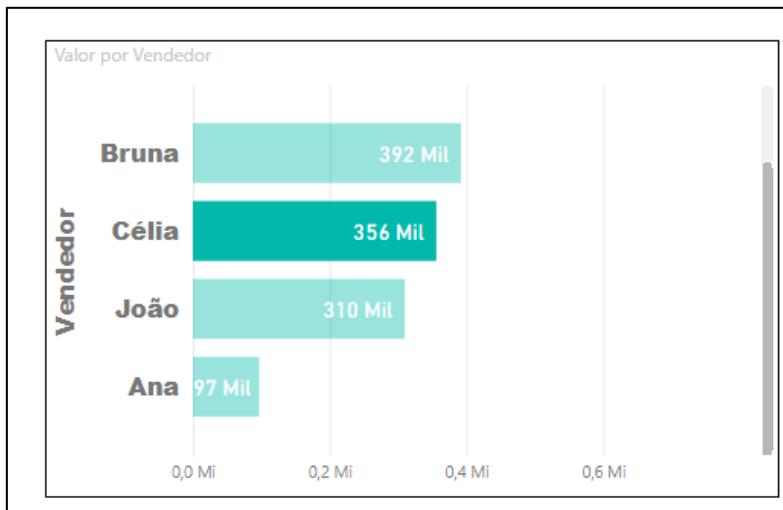
É possível deixar uma imagem e uma tela de fundo juntas.

Bordas

É possível adicionar uma borda no entorno do gráfico.



O gráfico, será apresentado da seguinte forma:

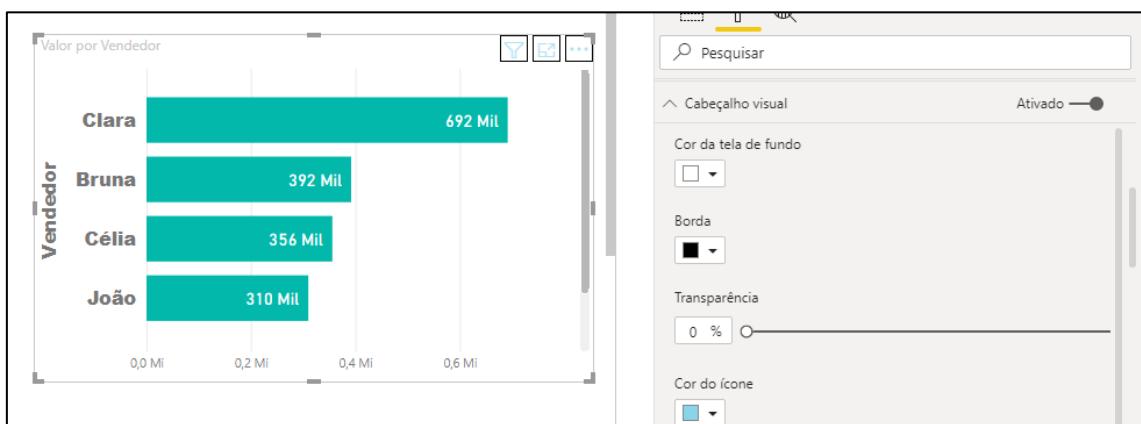


Cabeçalho Visual

Essa alteração, na verdade, altera os ícones dos gráficos, sendo eles: “Fixar/Foco, Filtro, Mais opções”

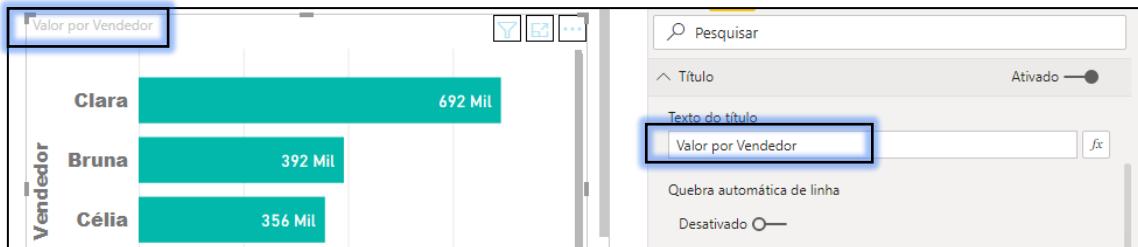


Veja no canto superior direito do gráfico como os ícones foram formatados.

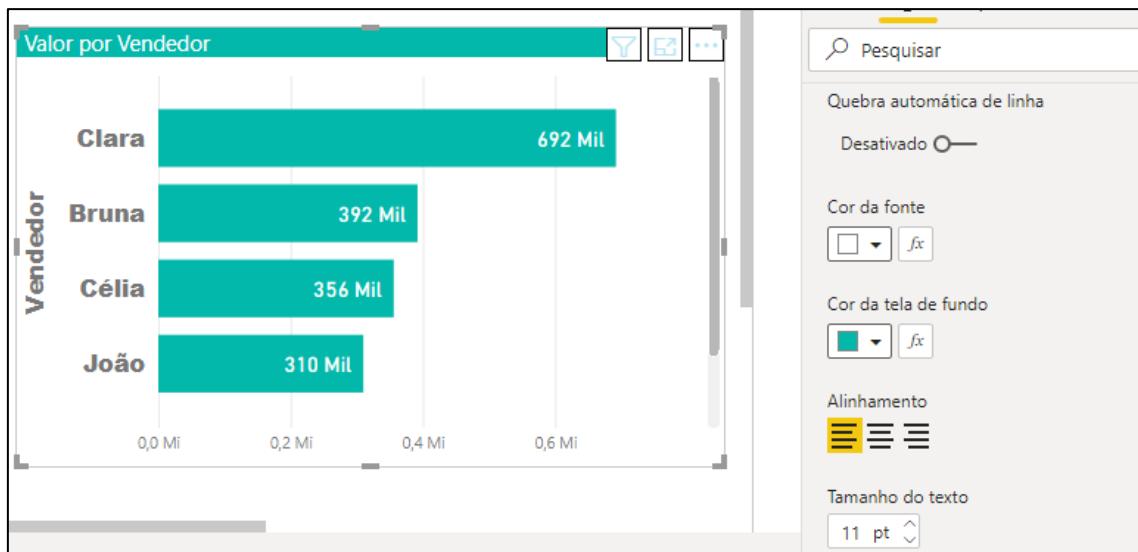


Título do Gráfico

Outro campo muito importante para a formatação do gráfico. O título não está visível no print, mas ele é proveniente automático devido a ser Vendedor x Valores Vendas nos campos.

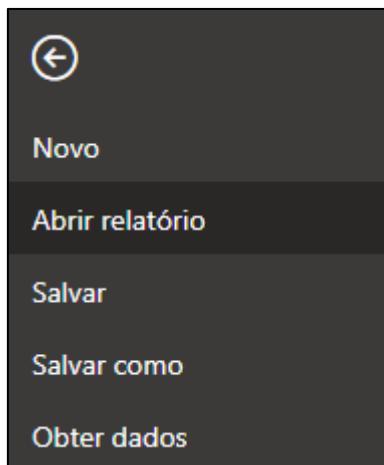


O resultado das alterações, portanto:



Salvando o Arquivo

Para salvar o arquivo que estamos trabalhando no PowerBI. É muito simples, basta ir em “Arquivo > Salvar Como > Selecionar o diretório > Nomear > Salvar”.



Vale ressaltar que a extensão do arquivo do PowerBI é a .pbix

DADOS (D:) > workspace > curso_MasterPowerBI > Arquivos PBIX - Lucas			
Nome	Data de modificação	Tipo	Tamanho
01 - Vendas Intro.pbix	01/03/2021 15:20	Microsoft.Microso...	42 KB

Primeiro Projeto

O primeiro projeto que realizaremos, terá a seguinte interface. Apenas para conhecimento prévio, os valores que estão ao lado esquerdo, como por exemplo, o “55.39 Mi”, são os “Cartões”.



Para esse caso de estudo, utilizaremos a planilha “02 – LojaDemo.xlsx” presente no caminho dos Datasets e a “Logo01” na pasta Assets.

The Excel interface shows the following details:

- File: 02 - LojaDemo.xlsx
- Tab: Página Inicial
- Table Headers:

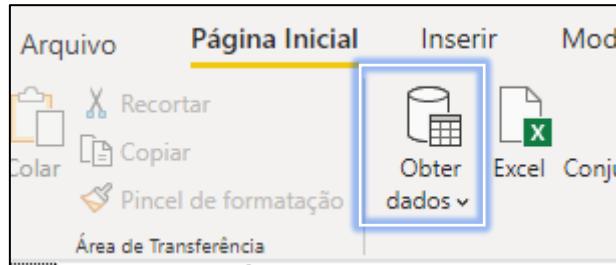
	Id Venda	Data Vend	Custo Unida	Valor	Quantidad	Custo da Venda	Valor da Venda
--	----------	-----------	-------------	-------	-----------	----------------	----------------

- Table Data (Rows 2-8):

2	7077	18/03/2013	76,09496775	304	9	684,8547097	2714
3	117	23/02/2012	7,491752798	12,99	4	29,96701119	50,14
4	7018	11/01/2012	10,12233768	159,99	9	91,10103916	1395,11
5	140	13/02/2014	0,576153263	25,69	18	10,37075874	462
6	491	17/01/2013	108,5087768	304	9	976,5789912	261
7	58	06/04/2014	77,98502883	299	24	1871,640692	705
8	7085	27/08/2014	154,987254	699	10	1549,87254	63

Lembrando que todas as planilhas são elaboradas pela Microsoft.

Para começar o projeto, vamos iniciar obtendo os dados.

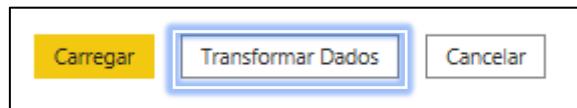


Nesse exemplo, teremos algo novo. Ao obter os dados, o PB apresenta duas opções.

Id Venda	Data Venda	Custo Unitário	Valor	Quantidade	Custo da Venda
7077	18/03/2013	76,09496775	304	9	2,23
117	23/02/2012	7,491752798	12,99	4	0,52
7018	11/01/2012	10,12233768	159,99	9	1,44
140	13/02/2014	0,576153263	25,69	18	0,46
491	17/01/2013	108,5087768	304	9	8,76
58	06/04/2014	77,98502883	299	24	7,08

Sendo “Data1” significar que os dados já foram tratados pelo PB e o “Data” são os dados sem nenhum tratamento, extraídos diretamente do excel pelo PB. Para esse exemplo, vamos selecionar o “Data1”.

Não clique em “Carregar”, utilize o “Transformar Dados”, pois é útil para verificar se os dados foram tipados da maneira correta e a que melhor atende as necessidades do dashboard. Vale lembrar que esse ato de transformação de dados, abre e utiliza o “Power Query”.



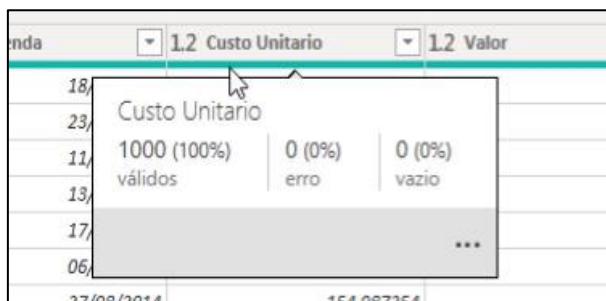
Veja que o PB se encarregou de automaticamente transformar os dados.

	Id Venda	Data Venda	Custo Unitário	Valor	Quantidade	Custo da Venda
1	7077	18/03/2013	76,09496775	304	9	2,23
2	117	23/02/2012	7,491752798	12,99	4	0,52
3	7018	11/01/2012	10,12233768	159,99	9	1,44
4	140	13/02/2014	0,576153263	25,69	18	0,46
5	491	17/01/2013	108,5087768	304	9	8,76
6	58	06/04/2014	77,98502883	299	24	7,08

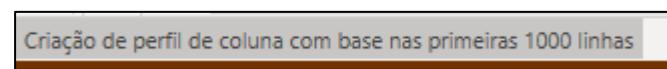
A “Data Venda” está em formato de data, valores como o “Custo Unitário” está em valor decimal, entre tantos outros.

Verificar a qualidade dos Dados

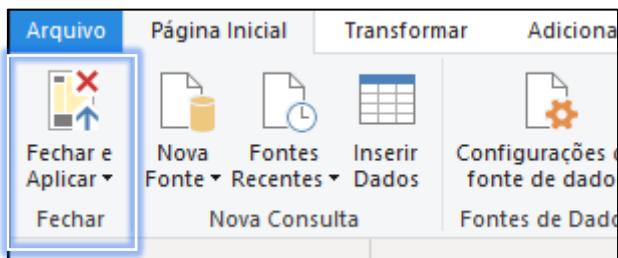
Exemplos como essa planilha, serão raros de acontecer no dia-a-dia da empresa, veja que é possível verificar se o campo coluna está apresentando alguma inconsistência como um erro ou mesmo valores vazios. Nesse caso, podemos analisar que a qualidade dos dados é perfeita, 100%.



Algo interessante ainda para salientar, é que essa planilha consta com 15.000 linhas. E o PB está apenas realizando uma “amostra” com as 1.000 primeiras ocorrências. Para que o PB aceite todas as 15mil, é necessário ir logo abaixo, próximo a barra de rolagem horizontal e alterar o tamanho.

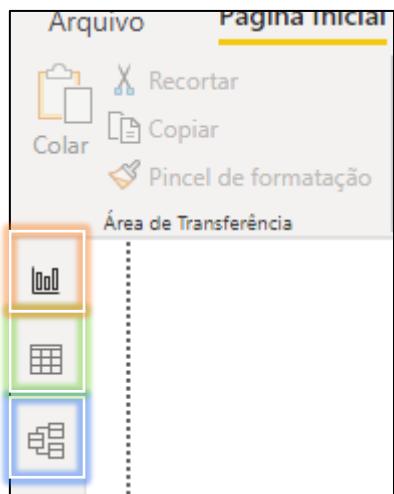


Ao trocarmos essa base das linhas, para esse caso que não alteraremos mais nada, podemos fechar e aplicar.



O Power Pivot

Após realizarmos a obtenção dos dados e a transformação deles, e ainda não iniciarmos nenhum elemento gráfico no painel, vamos nos atentar ao lado central esquerdo.



- Laranja brilhante é o “**Relatório**”, ele é onde elaboramos a parte gráfica.
- Verde brilhante são os “**Dados**”, ou seja, quais as colunas que serão utilizadas
- Azul brilhante é o “**Modelo**”, ou o **Power Pivot**, onde existe a modelagem.

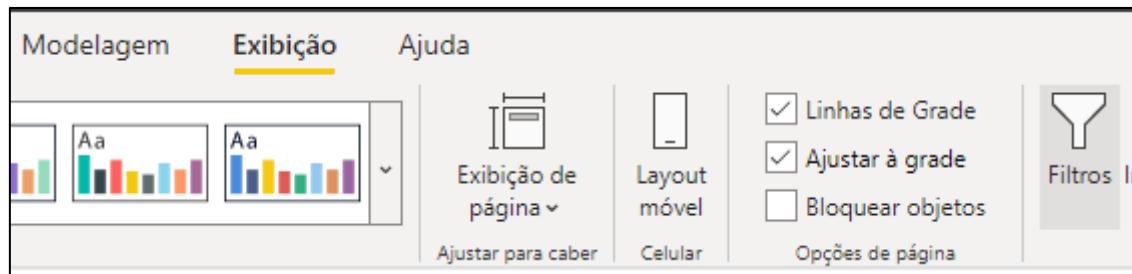
No **Power Pivot**, ou no **Modelo**, é o local onde se modela os dados, onde se criam relações entre as colunas, o que é “fato”.

No **Power Query**, ou no **Dados**, é o local onde os dados são transformados, onde consultas são geradas para a visualização dos dados.

Exibição

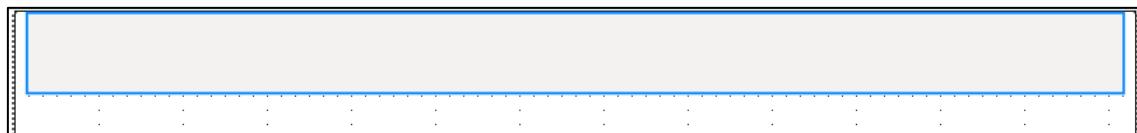
Para utilizarmos da melhor forma possível os dashboards, é altamente recomendado que antes de iniciar o projeto, o analista elabore, desenhe, crie uma forma como o seu dashboard se parecerá. Isso implica na aparência das apresentações dos elementos.

Uma forma que auxiliará é na guia de **Exibição**, fletar as duas opções de “Mostrar Linhas de Grades” e “Ajustar Objetos à Grade”.

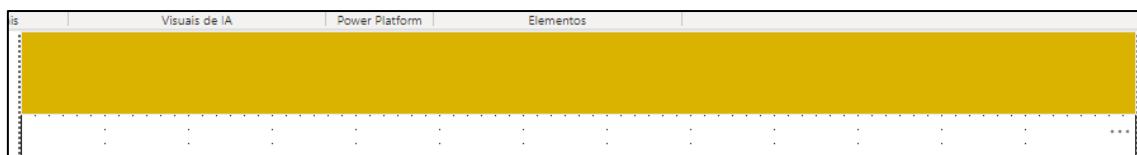


Para esse caso de estudo, utilizaremos as formas primitivas para fazer o topo do painel.

Note que há um espaço em branco, impossível de remover somente na tentativa de ampliar a forma retangular. Isso se deve pelo fato que as formas possuem a “Tela de Fundo” branca. A forma primitiva retangular é o retângulo azul.



Para corrigir isso, vamos a “Tela de Fundo” e vamos aplicar a mesma cor, precisamos remover a linha também, marcar sem nenhum ponto para a linha.



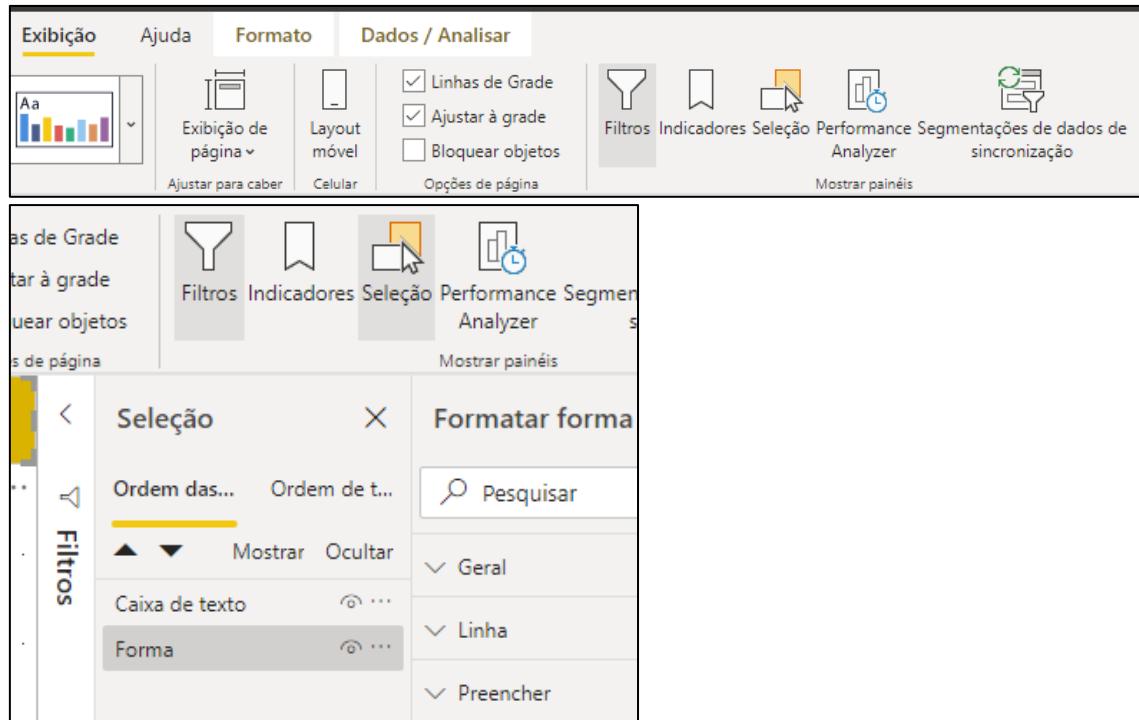
Vamos inserir uma caixa de texto nesse “cabeçalho”.



Ao jogarmos a Caixa de Texto no topo do painel em bege, a caixa de texto irá para trás. Isso acarretará um problema. Precisaremos abrir o “Painel de Seleção”.

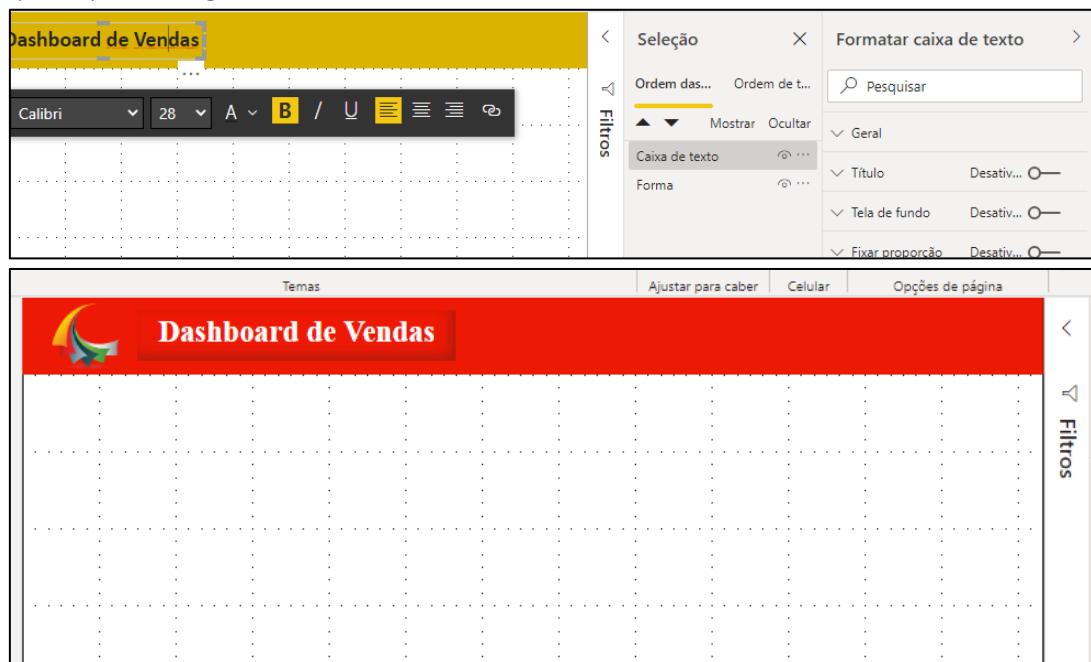
Painel de Seleção

O Painel de Seleção serve para exibir todos os elementos que estão disponíveis no painel atual. Isso é, caso alguma forma primitiva, gráfico ou qualquer outro elemento esteja sobrepondo um outro, podemos utilizar essa ferramenta para verificarmos a existência e a exibição dos elementos.



Podemos ocultar a visualização/exibição de alguns elementos, basta clicar no “olho” ao lado do nome do elemento.

Obs: A “Caixa de Texto” também tem uma “Tela de Fundo”, será preciso remover para que fique da seguinte forma.

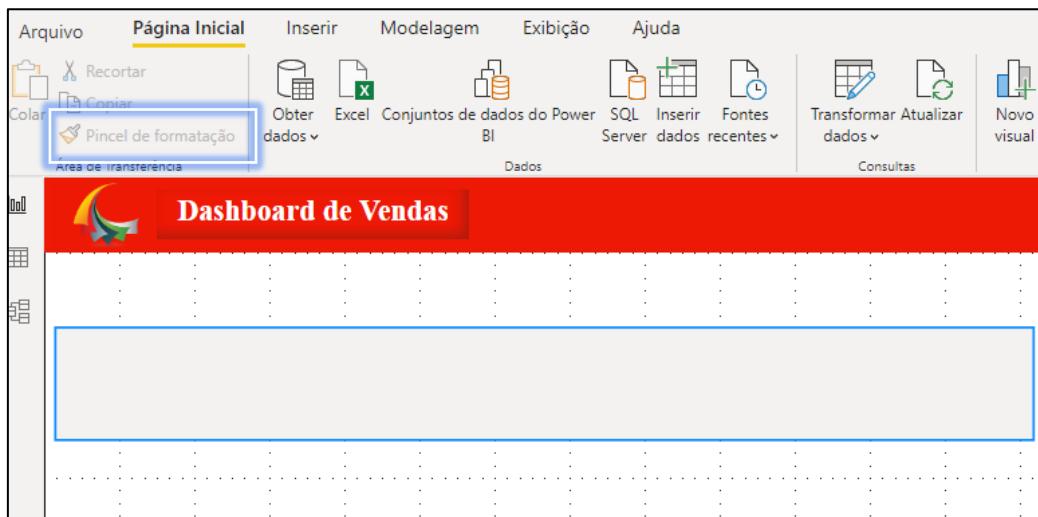


Pincel de Formatação

Continuando a elaboração do painel, realizaremos agora, a parte lateral esquerda, o quadro em azul brilhante abaixo.



Note que a lateral é uma forma primitiva também, poderíamos copiar e colar o retângulo acima. Porém, criaremos um e utilizaremos o pincel de formatação.



Clicar fora, selecionar o quadro vermelho acima e após isso, clicar diretamente no novo retângulo.



Inserindo Cartões nos Painéis

Para adicionar um cartão, é necessário ir em “Visualizações” e selecionar o “Cartão”.

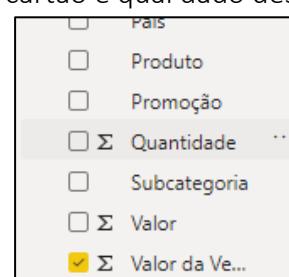


Ao clicar, um cartão “gigantesco” aparecerá na tela.



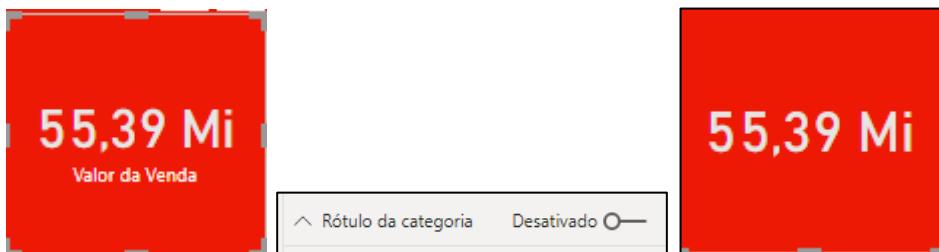
Basta diminuir até que fique no tamanho desejado.

Muito simples, após colocarmos o tamanho correto e no local desejado, selecione o cartão e qual dado deseja que apareça no cartão.



Para deixar esse cartão com um cabeçalho, vamos em “Rótulo de Dados” e podemos alterar o tamanho da fonte, a fonte, a cor.

Enquanto, existe um “texto” escrito logo abaixo do valor total da venda no cartão, é possível deixar, mas, nesse exemplo, vamos retirá-lo. Para isso, vamos em “Rótulos da Categoria”.



Para adicionar o título nesse cartão vamos em “Título”.



The screenshot shows the 'Título' (Title) configuration panel. It includes fields for 'Texto do título' (Title text) containing 'Total das Vendas', 'Cor da fonte' (Font color) with a black square, and 'Cor da tela de fundo' (Background screen color) with a black square. A preview window on the right shows the card with the title 'Total das Vendas' at the top and the value '55,39 Mi' below it.

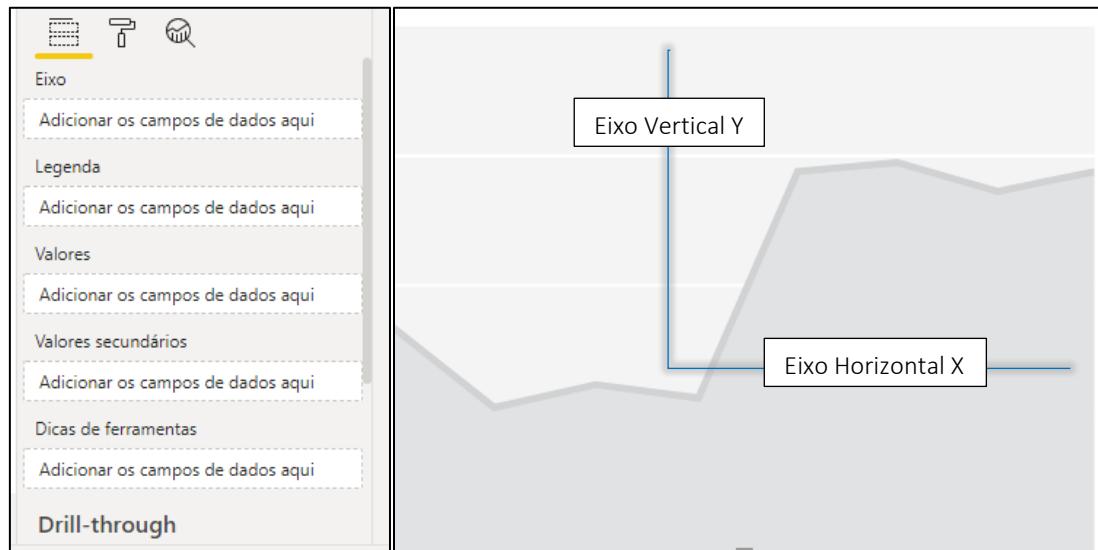
Cada um dos cartões abaixo, tiveram seus títulos alterados um a um. E cada um deles, também foram alterados conforme a necessidade do aparecimento. O primeiro cartão contou com apenas a Somatória das Vendas. O segundo, a somatória do Lucro. O terceiro, utilizamos a somatória das vendas, porém, substituimos a soma pela média. Por último, substituimos os valores das vendas pelo Produtos, nele selecionamos a “Contagem Distinta”.



Gráfico de Área

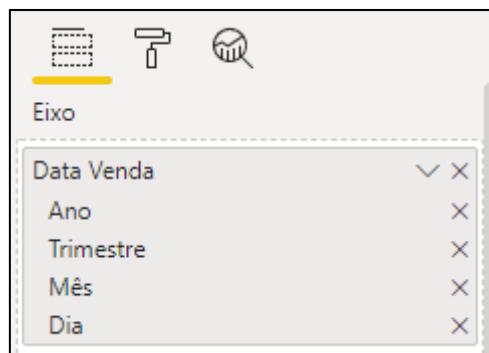
O **Gráfico De Área** é uma excelente ferramenta para analisar o “**tempo**”. Podemos utilizar o de “área comum” ou o de “área empilhado”. Para esse caso, utilizaremos o de área “simples”.

Ao inserirmos esse elemento gráfico, alguns itens são necessários para que ele seja preenchido.



Muito importante frizar que o Eixo que será alterado será o “Eixo Horizontal X”. Ele é o único eixo que pode ser configurado no gráfico de área. O “Eixo Vertical Y” somente será “*habilitado*” ao adicionarmos as “medidas” ao gráfico.

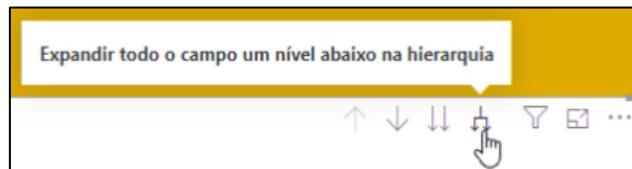
Para esse exemplo, e como vamos trabalhar com o tempo, o campo que preencherá o eixo será a “Data da Venda”.



O **Power BI** é suficientemente inteligente para identificar que nas datas, ela é composta por uma **hierarquia**, sendo, portanto o Ano, Trimestre, Mês e Dia.

Drilldown

O *drilldown* é o ato de para cada data, ir “quebrando” as datas até a última hierarquia possível. Para isso, pressionar o botão abaixo.



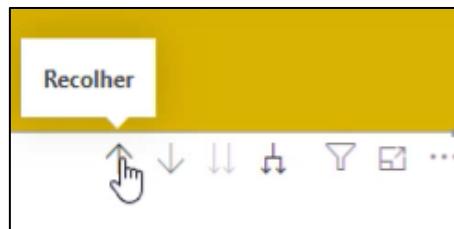
Para cada vez que esse botão for clicado, ele entrará no próximo nível da hierarquia, realizando, um *drilldown*.



O ponto importante é que ao realizar o drilldown, no primeiro nível por exemplo, realizará a **soma** e mostrará dos meses no trimestre por ano selecionado.

Drillup

É inversamente proporcional ao drilldown. Para cada nível da hierarquia, cada click no “Up”, o drillup subirá um nível da hierarquia da data. A ordem inversa, sendo, portanto, Dia, Mês, Trimestre e Ano.



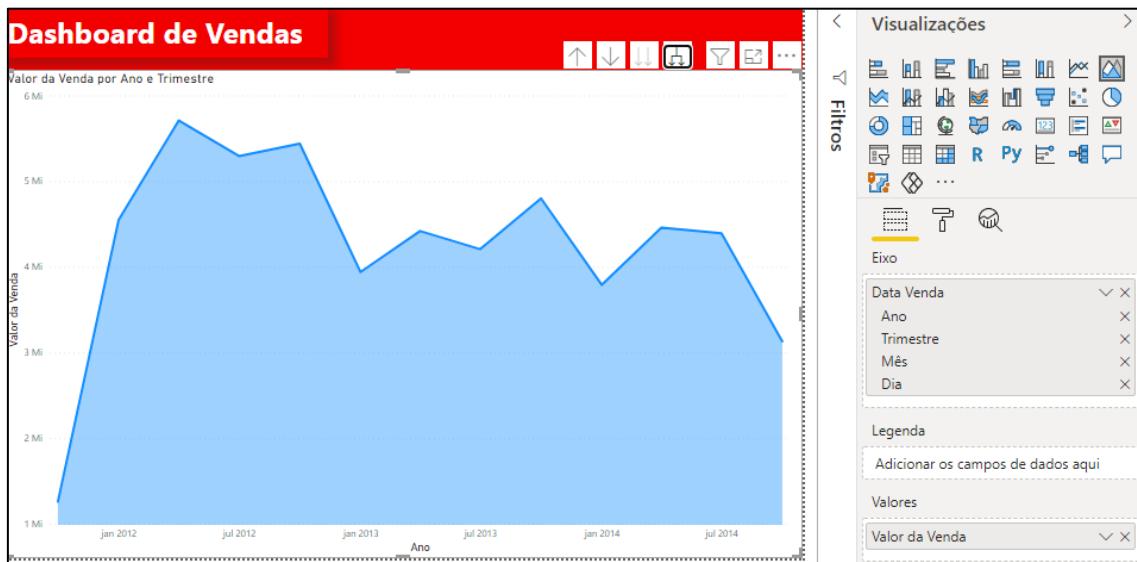
Valores no Gráfico

Voltemos aos campos a serem preenchidos no gráfico. O próximo passo será inserir os **“Valores”**.



Quais serão eles, nesse exemplo?

Os valores a serem incluídos nesse exemplo, dado as “Datas de Vendas”, serão os “Valores das Vendas”.



Veja o gráfico acima, está com o “gráfico de área”, com o *drilldown* selecionado até o nível de **Trimestre** na hierarquia, o elemento de “Data de Vendas” foi o selecionado para ocupar o “Eixo X” e os valores foram preenchidos pela somatória do “Valor de Venda”.

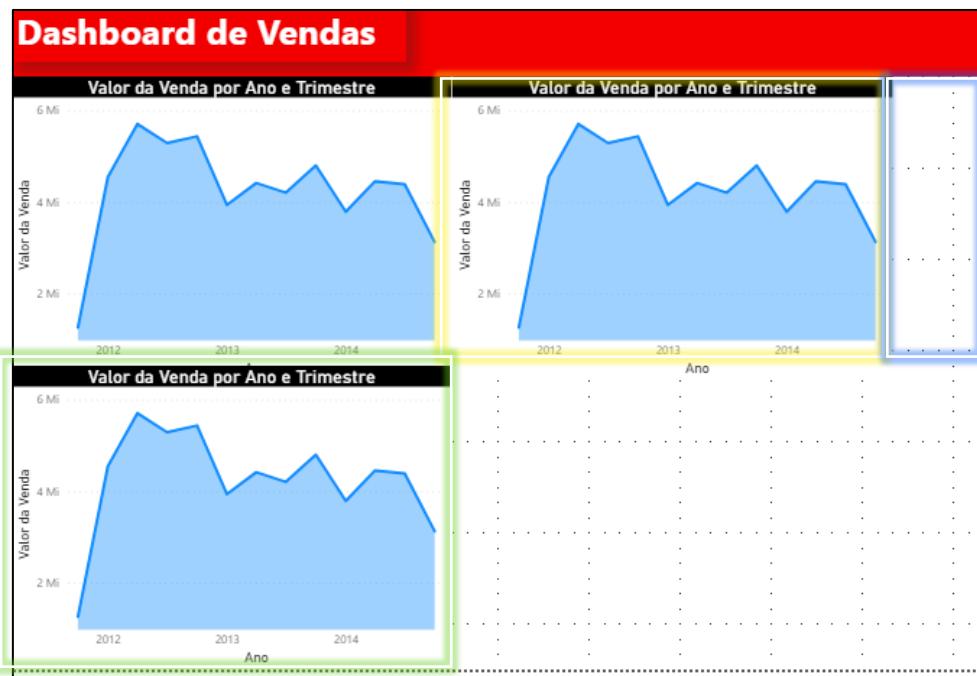
Veja que por conta do drilldown ativo, os valores foram quebrados de ano em ano, trimestre por trimestre.



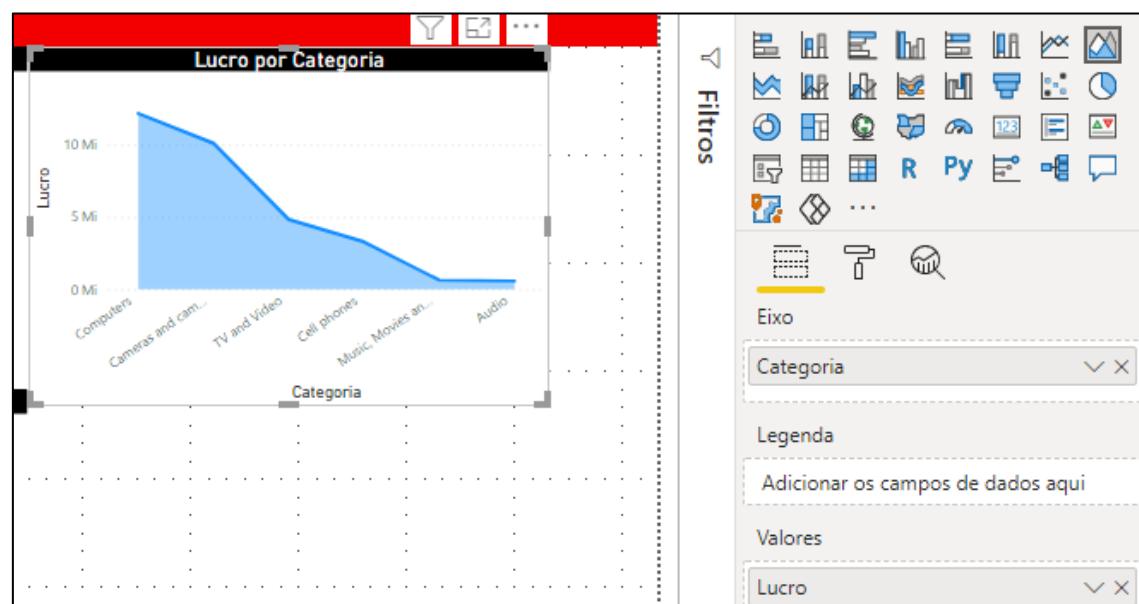
Gráfico de Barras e Rosca/Pizza

Para ganharmos tempo, e mantermos a formatação já “pronta”, vamos utilizar o “Ctrl+C” e “Ctrl+V” para copiar o gráfico e aí realizar a troca para os gráficos estabelecidos no título desse capítulo.

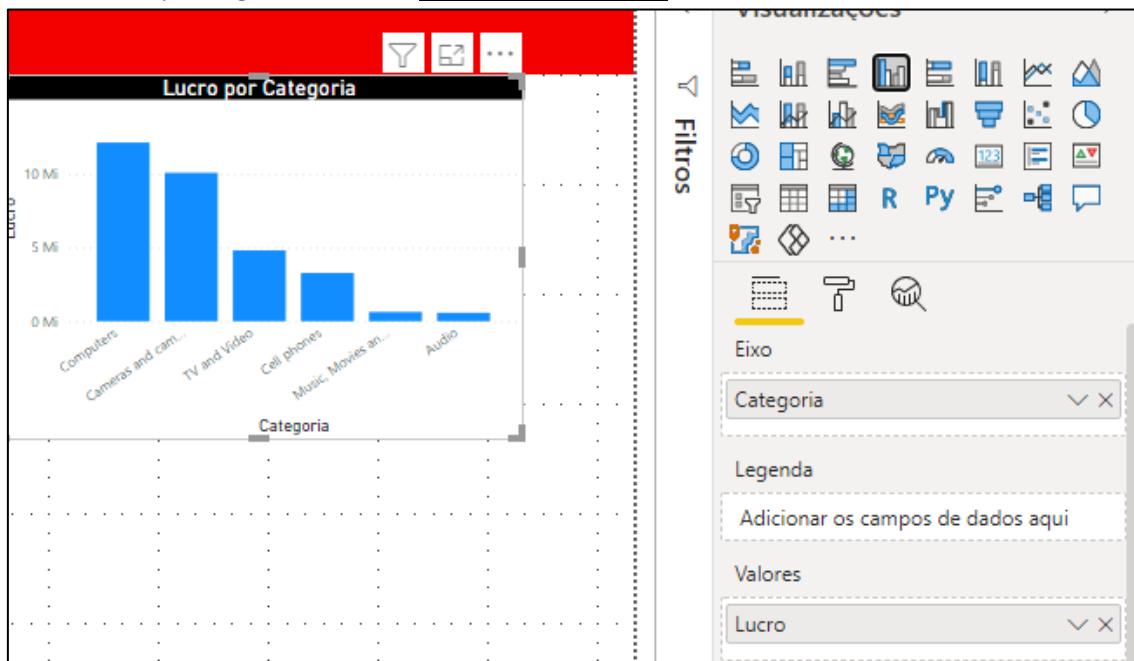
O gráfico no quadro em amarelo, será o gráfico de Barra Clusterizado. Gráfico no quadro em verde será o de Rosca/Pizza e ao lado, foi deixado um espaço para inserirmos o filtro.



Um ponto muito importante, no quadro em amarelo, desejamos saber o “*Lucro por cada Categoria*”, ou seja, precisamos colocar no “eixo X” as categorias de produtos e no campo dos “Valores”, o Lucro! Porém, esse formato de gráfico não é o mais indicado.



O formato do gráfico está para o de área, porém, não estamos tratando de eventos temporais, não estamos lidando com datas. Portanto, o gráfico que melhor se aproxima de “Valores por algo”, será o de **Barras Clusterizado**.



Observe como é muito mais legível e fácil de identificar. A categoria “computador” é a que propõem maior lucro pela categoria, enquanto a categoria de “áudio” é a que proporciona o menor lucro pela categoria. Essas informações até estavam presentes no gráfico de área, porém, **não são temporais** e esse formato apresenta uma melhor legibilidade.

Gráfico de Dispersão

O gráfico de dispersão é importantíssimo pois ele estuda a correlação entre duas variáveis, procurar uma **tendência**. Isso é, “*procura entender é como uma variável se comporta em um cenário onde outra está variando, visando identificar se existe alguma relação entre a variabilidade de ambas.*¹”. Estudar a correlação do “**Coeficiente de Correlação de Pearson**” e o “**Coeficiente de Correlação de Spearman**” no final apostila.

Para iniciar, procurar por uma “**Tabela**” na área de visualização.

The screenshot shows the Power BI visualization interface. At the top, there is a "Visualizações" (Visualizations) ribbon with various chart icons. Below it is a large workspace containing a table visual. The table has columns labeled "Subcategoria", "Valor da Venda", and "Lucro". Above the table, there are several filters: "Computers", "Cameras and cam...", "TV and Video", "Cell phones", "Music, Movies an...", and "Audio". On the right side, there is a properties pane titled "Valores" (Values). It contains a section "Adicionar os campos de dados aqui" (Add data fields here) with a dashed border. Below it is a "Drill-through" section with two options: "Relatório cruzado" (Cross-reporting) set to "Desativado" (Disabled) and "Manter todos os filtros" (Keep all filters) set to "Ativado" (Enabled). There is also a button "Adicionar os campos de drill-through..." (Add drill-through fields...).

A “Tabela” admite apenas “Valores”. Portanto, vamos adicionar os seguintes: “**Valor da Venda**”, “**Lucro**” e “**Subcategoria**”.

The screenshot shows the Power BI visualization interface with the table visual now populated with data. The table data is as follows:

Subcategoria	Valor da Venda	Lucro
Bluetooth Headphones	277.660,39	152.795,47
Camcorders	8.456.120,92	4.972.852,90
Cameras & Camcorders Accessories	321.828,87	170.688,86
Car Video	2.093.143,99	1.065.879,48
Cell phones Accessories	825.948,35	468.371,90
Computers Accessories	709.927,37	359.936,64
Desktops	3.491.787,15	1.762.032,54
Digital Cameras	2.539.792,28	1.455.482,33
Digital SLR Cameras	5.731.189,62	3.503.777,56
Home & Office Phones	319.645,11	176.717,36
Home Theater System	4.737.967,56	2.470.355,41
Laptops	6.403.174,09	3.763.974,47
Total	55.391.759,79	31.587.437,29

To the right of the table, there is a properties pane titled "Valores". It contains a section "Adicionar os campos de dados aqui" with a dashed border. Below it is a "Drill-through" section with two options: "Relatório cruzado" (Cross-reporting) set to "Desativado" (Disabled) and "Manter todos os filtros" (Keep all filters) set to "Ativado" (Enabled). There is also a button "Adicionar os campos de drill-through..." (Add drill-through fields...).

Dessa forma, e em formato de “**Tabela**”, foi descremido cada subcategoria no total de vendas e a sua participação de lucro.

¹ <https://operdata.com.br/blog/coeficientes-de-correlacao/>

Ao passar o mouse na coluna “Valor da Venda” na “Tabela”, existe uma seta direcional para baixo. Essa seta irá filtrar os valores do Maior para o Menor.

A tabela, portanto, resultará:

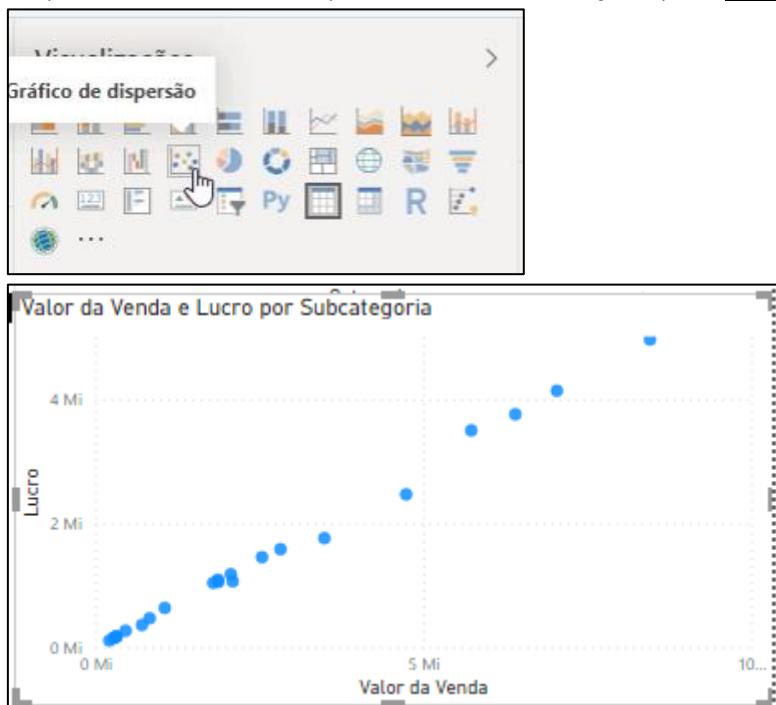
Subcategoria	Valor da Venda	Lucro
Camcorders	8.456.120,92	4.972.852,90
Projectors & Screens	7.036.174,58	4.145.050,32
Laptops	6.403.174,09	3.763.974,47
Digital SLR Cameras	5.731.189,62	3.503.777,56
Home Theater System	4.737.967,56	2.470.355,41
Desktops	3.491.787,15	1.762.032,54
Smart phones & PDAs	2.822.294,79	1.585.115,38
Digital Cameras	2.539.792,28	1.455.482,33
Car Video	2.093.143,99	1.065.879,48
Televisions	2.064.341,63	1.179.150,73
Touch Screen Phones	1.873.267,52	1.055.721,69
Monitors	1.868.935,66	1.090.990,21
Total	55.391.759,79	31.587.437,29

Note que o “Valor da Venda” foi filtrado dessa forma e note que o “Lucro” também está do Maior para o Menor, ou seja, existe uma relação (apenas de olhar, essa relação é notável), sobre a venda e o lucro.

Correlação

Podemos também, utilizar o **Gráfico De Dispersão** que irá **cruzar** cada um dos campos (Valor da Venda e Lucro) com a subcategoria de cada um. Se na plotagem desse gráfico, uma **linha** for traçada em **diagonal crescente**, isso significa estatisticamente, que existe uma relação entre “Valor de Venda” e o “Lucro”, ou seja, quanto maior o valor da venda, maior será o lucro. Indicando uma *correlação positiva*. O gráfico de dispersão é muito utilizado em *Machine Learning* para realizar “Regressão Linear”.

Clique na tabela criada e procure nas visualizações por “Gráfico de Dispersão”.



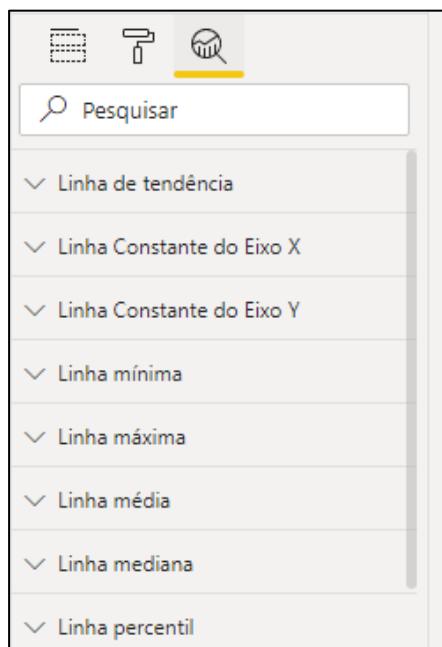
Analizando o resultado do gráfico da página 36.

Veja que cada ponto azul é a “Subcategoria”.



É claramente notado que uma está em função da outra, elas estão correlacionadas.

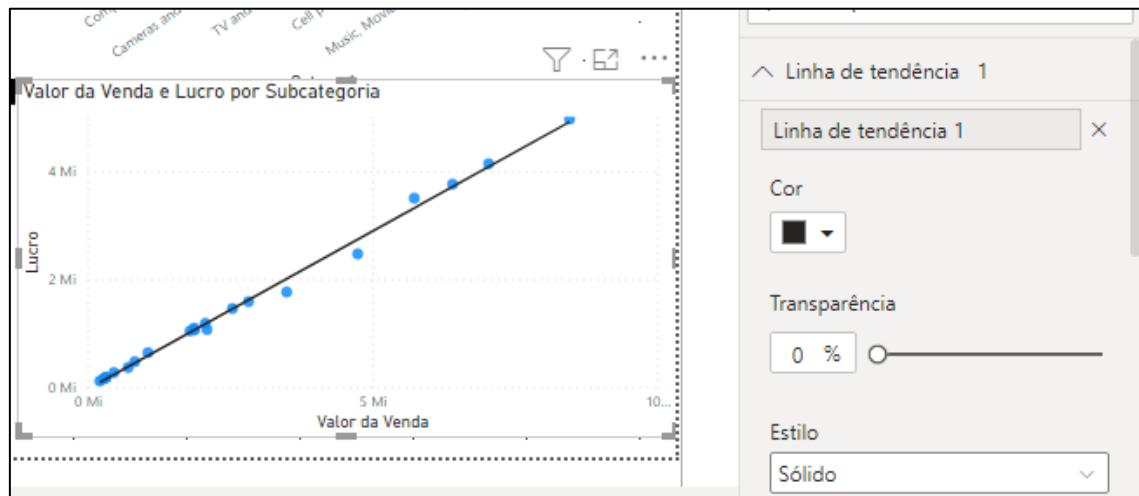
Com o gráfico de dispersão selecionado, clique em “Análise”.



As opções de análises estarão disponíveis.

Linha de Tendência

Uma ferramenta muito importante para saber a correlação, é a “Linha de Tendências”.



Quaisquer novos elementos adicionados, serão analizados para a verificação se estão dentro da “dispersão” da “Linha de Tendência”, qualquer valor fora da linha, deve ser analisado.

Filtros

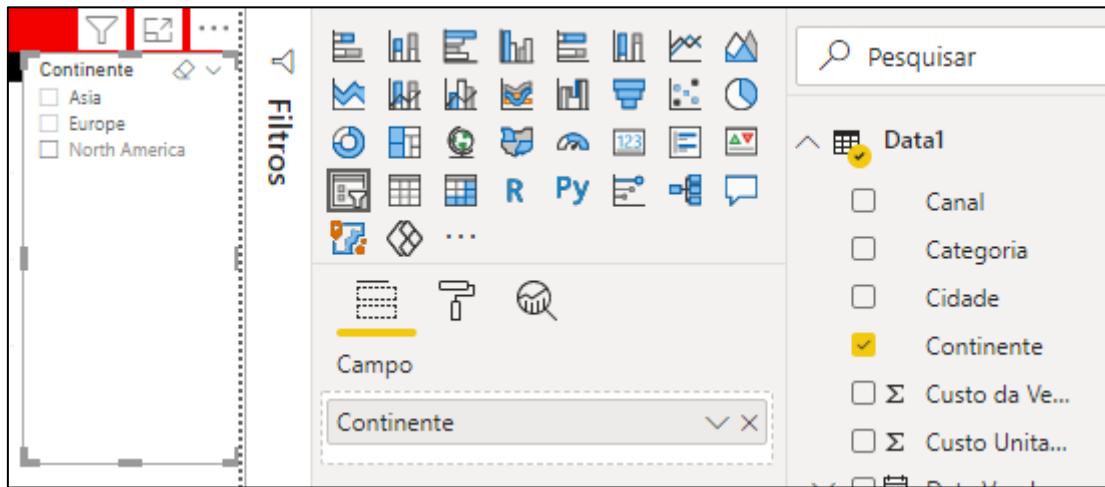
Segmentação de Dados (Filtro)

Nessa seção vamos criar um filtro, um *segmentador de dados*.

Clique em “Segmentação de Dados”.

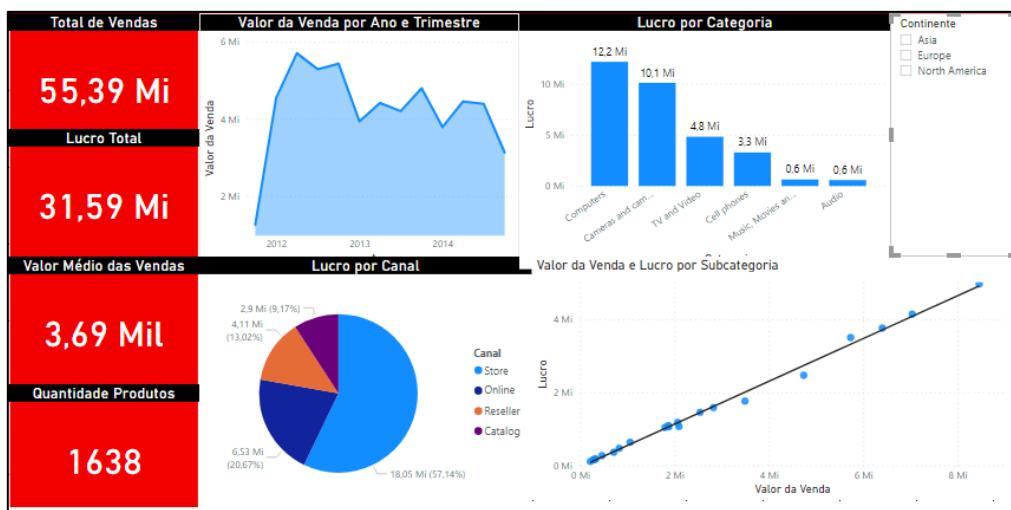
The figure shows the "Segmentação de Dados" visual. It has two data points: "0,6 Mi" and "0,6 Mi" under the category "Audio". A tooltip message "Selecione ou arraste campos para preencher este visual" is displayed above the visual area. To the right of the visual, there is a sidebar titled "Filtros".

O campo a ser preenchido no “Segmentador de Dados”, AKA “Filtro”, será o campo “Continente”, ou seja, o desejo é descobrir e quebrar os valores por continente.



Qualquer clique no Filtro, dentro desse painel, também alterará os demais gráficos inseridos no painel atual.

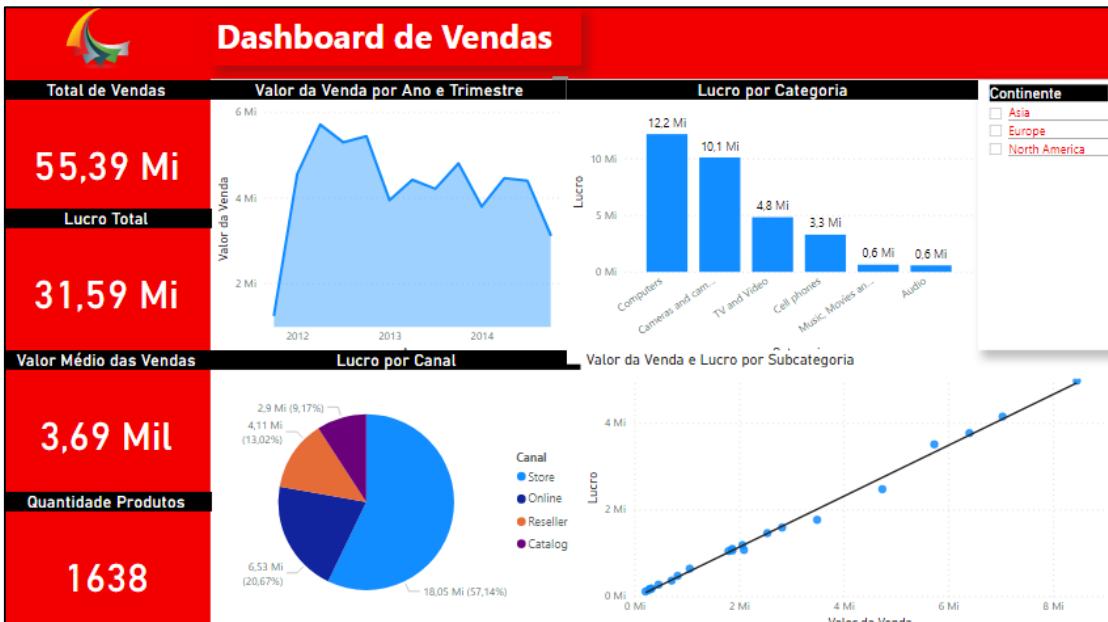
Antes do Filtro



Filtro “North America”



Término do Primeiro Projeto



Classificação de Dados

Os dados podem ser definidos de duas formas: Estruturados e Não Estruturados.

Dados	
Estruturados	Não Estruturados
Estrutura padrão – Schema	Sem estrutura padronizada
Banco de Dados Relacionais, Excel, XML, CSV	Arquivos em geral: Fotos, Músicas, Arquivos de Texto (txt) - Sem estrutura fixa

Dados Estruturados

A partir dos dados estruturados, podemos definir duas formas: Quantitativos e Qualitativos.

Dados Estruturados	
Quantitativos	Qualitativos
Expressa número, quantidade 10 Computadores, 2 pessoas, 1.5 quilos de arroz	Expressa descrição do objeto, qualidade Azul, Amarelo, Bom, Ruim, Alto, Baixo, Verão, Inverno

O **Power BI** na maioria das vezes ao identificar dados “Quantitativos” realiza uma soma ou conhecido como Somatória (Sigma - Σ)

Dados Estruturados – Quantitativos

A partir dos Dados Estruturados, o Quantitativo pode ser definido como: Contínuo e Discreto.

Dados Estruturados Quantitativos	
Contínuos	Discretos
Medidas decimais, flutuantes, reais 1.70 metros de altura, 1.3 quilos de arroz, 32.4º graus de temperatura	Valores Inteiros 10 Pessoas, 15 Canetas, 5 Gols, 50 Carros

Dados Estruturados – Qualitativos

A partir dos Dados Estruturados, o Qualitativos pode ser definido como: Nominais e Ordinais.

Dados Estruturados Qualitativos	
Nominais *	Ordinais **
Não expressam ordem ou ordem de grandeza	Expressam ordem ou ordem de grandeza
Amarelo, Verde Primavera, Outono, Verão Masculino, Feminino	1 Estrela, 2 Estrelas, 3 Estrelas, 5 Estrelas Péssimo, Ruim, Bom, Ótimo

* No Nominais não ocorre o “algo é melhor que o outro” ou “que algo é maior que o outro”.

** Por exemplo, ao realizar um “Rate” em uma plataforma de Streaming, podemos informar qual a **qualidade através da “ordem”** (nunca da quantidade), do filme visto, por exemplo: 1 estrela ou 5 estrelas.

Fluxo dos Dados



A classificação dos dados é importante para o **Power BI** realizar o tratamento e a tipagem para a melhor visualização e interpretação dos dados.

Ambientes de Banco de Dados

Existem dois ambientes, sendo eles: OLAP e OLTP.



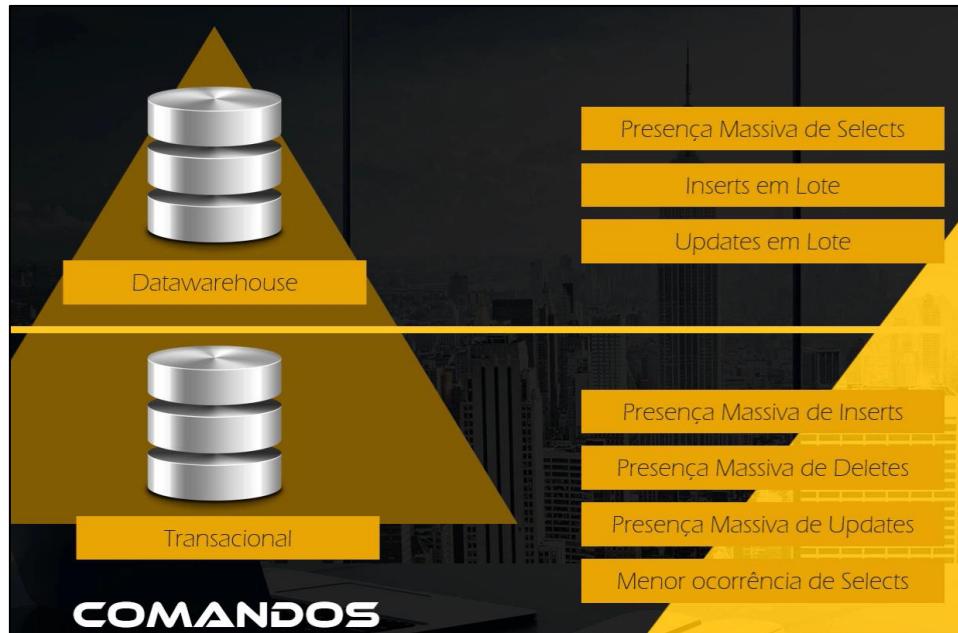
No ambiente **OLTP(On-Line Transaction Process)** é a inserção de dados, é o ambiente Operacional, por exemplo, vendedores realizando inputs de informação o tempo todo.

No ambiente **OLAP(On-Line Analytical Process)**, é o ambiente estratégico, na tomada de decisão médio e longo prazo. A análise do OLTP é analisada no ambiente OLAP.

No intervalo entre o OLTP e OLAP, existe o **Staging** que utiliza ferramentas do tipo **ETL(Extract, Transform & Load)**. O Staging extrai informações do ambiente operacional OLTP, transforma os dados e carrega para o ambiente OLAP.

Comandos do Banco de Dados

Existem dois comandos no que tange o Ambiente de Banco de Dados: os de Datawarehouse(OLAP) e o Transacionais (OLTP).



Repare que nos comandos Transacionais, são os comandos de SQL de inserção, de deletes, updates.

Enquanto em de Datawarehouse, é massivo os comandos de consultas, para tomada de decisão.

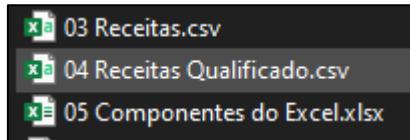
Power Query e Tratamento de Dados com Linguagem M

Arquivos .csv

O que é o arquivo CSV? É a abreviação de “Comma Separated Values”, significa que são valores separados por vírgulas.

Se o arquivo **.csv** quando chegar perfeito na mão do analista, é perfeito. Porém, se o arquivo **.csv** não estiver bem tratado, será um grande problema.

Para o próximo projeto, utilizaremos as seguintes planilhas, a “03 Receitas.csv” e a “04 Receitas Qualificado.csv”.



Lembrando que dessa vez não utilizaremos o formato do arquivo Excel e sim como “Texto/CSV”

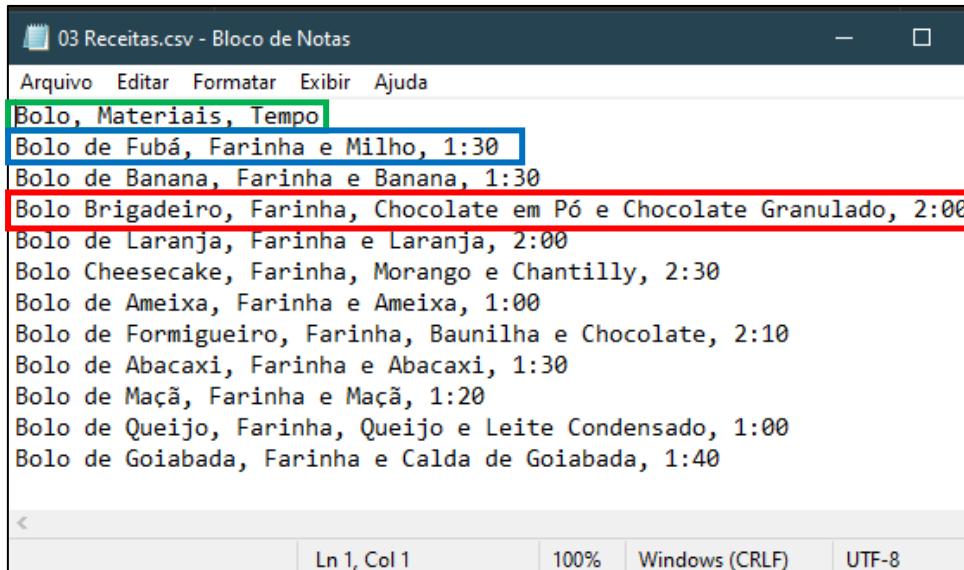


Veja como o arquivo é “complexo” de ser trabalhado, pois foge ao parâmetro comum da utilização normal dos usuários apenas em Excel.

A screenshot of the Power Query Editor. On the left, the '03 Receitas.csv' data is shown in a table with columns 'Bolo, Materiais, Tempo'. The data includes various types of cakes and their preparation times. On the right, a preview window titled '03 Receitas.csv - Bloco de Notas' shows the raw text data. The preview window has a menu bar with 'Arquivo', 'Editar', 'Formatar', 'Exibir', and 'Ajuda'. Below the menu, the text data is displayed in a monospaced font.

Interpretação do .csv

Se apenas analisarmos o que está acontecendo no **.csv**, pelo Bloco de Notas é fácil entender.



```
Bolo, Materiais, Tempo
Bolo de Fubá, Farinha e Milho, 1:30
Bolo de Banana, Farinha e Banana, 1:30
Bolo Brigadeiro, Farinha, Chocolate em Pó e Chocolate Granulado, 2:00
Bolo de Laranja, Farinha e Laranja, 2:00
Bolo Cheesecake, Farinha, Morango e Chantilly, 2:30
Bolo de Ameixa, Farinha e Ameixa, 1:00
Bolo de Formigueiro, Farinha, Baunilha e Chocolate, 2:10
Bolo de Abacaxi, Farinha e Abacaxi, 1:30
Bolo de Maçã, Farinha e Maçã, 1:20
Bolo de Queijo, Farinha, Queijo e Leite Condensado, 1:00
Bolo de Goiabada, Farinha e Calda de Goiabada, 1:40
```

Quadro em **verde**: Na primeira linha, indicando o cabeçalho, significa que os valores estão separados por vírgulas, ou seja, o *Bolo* será a primeira coluna, depois o *Material* para preparar o bolo e o *Tempo* que é necessário para o preparo do bolo.

Quadro **azul**: Na segunda linha e para baixo, note que é seguido a ordem do cabeçalho descrito acima, porém sendo os inputs, ou seja, na segunda linha por exemplo, o Bolo é o Bolo de Fubá, os Materiais é de Farinha e Milho e o Tempo é de 1:30hora.

Quadro **vermelho**: Porém, note que a primeira informação é o Bolo que é o Bolo de Brigadeiro. Na segunda informação de valor, teremos um grande problema. Note que os ingredientes estão como “Farinha, Chocolate em Pó e Chocolate Granulado” e somente depois o tempo aparece. Isso acarretará um grande problema, pois temos 4 informações separadas por valores e somente 3 de cabeçalho. E o tempo não foi devidamente informado no seu devido local.

Bolo	Material	Tempo
Bolo de Fubá	Farinha e Milho	1:30
Bolo de Brigadeiro	Farinha	Chocolate em Pó e Granulado

Resumindo, isso vai dar dor de cabeça.

Voltando ao **Power BI** na próxima página.

Correção de Charset

Veja também que devido a utilização de acentos, os caracteres não devidamente apresentados na visualização.

03 Receitas.CSV

Origem do Arquivo	Delimitador
1252: Europeu Ocidental (Windows)	Dois-pontos
Bolo, Materiais, Tempo	
Bolo de FubÃi, Farinha e Milho, 1	30
Bolo de Banana, Farinha e Banana, 1	30
Bolo Brigadeiro, Farinha, Chocolate em PÃ³ e Chocolate...	0
Bolo de Laranja, Farinha e Laranja, 2	0
Bolo Cheesecake, Farinha, Morango e Chantilly, 2	30
Bolo de Ameixa, Farinha e Ameixa, 1	0
Bolo de Formigueiro, Farinha, Baunilha e Chocolate, 2	10
Bolo de Abacaxi, Farinha e Abacaxi, 1	30
Bolo de MaÃ§Ã£o, Farinha e MaÃ§Ã£o, 1	20
Bolo de Queijo, Farinha, Queijo e Leite Condensado, 1	0
Bolo de Goiabada, Farinha e Calda de Goiabada, 1	40
	null
	null

Isso se deve por causa do charset do sistema. O charset é o conjunto de caracteres que resulta na visualização dos dados. Para alterar, a melhor solução é ado UTF-8.

03 Receitas.CSV

Origem do Arquivo	Delimitador
65001: Unicode (UTF-8)	Dois-pontos
Bolo, Materiais, Tempo	
Bolo de Fubá, Farinha e Milho, 1	30
Bolo de Banana, Farinha e Banana, 1	30
Bolo Brigadeiro, Farinha, Chocolate em Pó e Chocolate...	0
Bolo de Laranja, Farinha e Laranja, 2	0
Bolo Cheesecake, Farinha, Morango e Chantilly, 2	30
Bolo de Ameixa, Farinha e Ameixa, 1	0
Bolo de Formigueiro, Farinha, Baunilha e Chocolate, 2	10
Bolo de Abacaxi, Farinha e Abacaxi, 1	30

Delimitador de Arquivo .csv

Vamos agora, a questão do Delimitador do arquivo **.csv**.

Origem do Arquivo	Delimitador
65001: Unicode (UTF-8)	Dois-pontos
Bolo, Materiais, Tempo	
Bolo de Fubá, Farinha e Milho, 1	30
Bolo de Banana, Farinha e Banana, 1	30
Bolo Brigadeiro, Farinha, Chocolate em Pó e Chocolate...	0

Note que o delimitador está em “**dois pontos**”, mas estamos trabalhando com arquivos **.csv**. Por isso, **vamos alterar**.

Origem do Arquivo	Delimitador	
65001: Unicode (UTF-8)	Vírgula	
Column1 Column2 Column3		
Bolo	Materiais	Tempo
Bolo de Fubá	Farinha e Milho	1:30
Bolo de Banana	Farinha e Banana	1:30
Bolo Brigadeiro	Farinha	Chocolate em Pó e Chocolate Granulado
Bolo de Laranja	Farinha e Laranja	2:00
Bolo Cheesecake	Farinha	Morango e Chantilly

Porém, ainda ocorre a confusão por questão do Material e o Tempo.

Null

Um outro grande problema, existem linhas em branco, com “**Null**” pois foi apertado o botão ENTER após o término de preenchimento dos dados.

Bolo de Goiabada	Farinha e Calda de Goiabada	1:40

Após analisar, prosseguiremos com a “[Transformação dos Dados](#), nunca clique em “Carregar”

Column1	Column2	Column3
Bolo	Materiais	Tempo
Bolo de Fubá	Farinha e Milho	1:30
Bolo de Banana	Farinha e Banana	1:30
Bolo Brigadeiro	Farinha	Chocolate em Pó e Chocolate Granulado
Bolo de Laranja	Farinha e Laranja	2:00
Bolo Cheesecake	Farinha	Morango e Chantilly
Bolo de Ameixa	Farinha e Ameixa	1:00
Bolo de Formigueiro	Farinha	Baunilha e Chocolate
Bolo de Abacaxi	Farinha e Abacaxi	1:30
Bolo de Maçã	Farinha e Maçã	1:20
Bolo de Queijo	Farinha	Queijo e Leite Condensado
Bolo de Goiabada	Farinha e Calda de Goiabada	1:40

Observação Importante

Ao realizar a transformação, estava ocorrendo um erro pois não mostrava as 4 colunas, era necessário colocar uma vírgula no arquivo **.csv** depois do cabeçalho do tempo para reconhecer a 4 coluna.

Antes

	Column1	Column2	Column3
1	Bolo	Materiais	Tempo
2	Bolo de Fubá	Farinha e Milho	1:30
3	Bolo de Banana	Farinha e Banana	1:30
4	Bolo Brigadeiro	Farinha	Chocolate em Pó e Chocolate Granulado
5	Bolo de Laranja	Farinha e Laranja	2:00

Antigo

*03 Receitas.csv - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
Bolo, Materiais, Tempo

Bolo de Fubá, Farinha e Milho, 1:30

Novo

*03 Receitas.csv - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
Bolo, Materiais, Tempo,

Bolo de Fubá, Farinha e Milho, 1:30

Depois

	Bolo	Materiais	Tempo
1	Bolo de Fubá	Farinha e Milho	1:30
2	Bolo de Banana	Farinha e Banana	1:30
3	Bolo Brigadeiro	Farinha	Chocolate em Pó e Chocolate Granulado
4	Bolo de Laranja	Farinha e Laranja	2:00
5	Bolo Cheesecake	Farinha	Morango e Chantilly

Habilitando a Linguagem M

Para habilitar, vamos em “Exibição > Layout > Barra de Fórmulas”.

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

Barra de Fórmulas Espaçamento uniforme Distribuição de colunas Permitir sempre
Layout Mostrar espaço em branco Perfil da coluna Ir para Coluna Qualidade da coluna Parâmetros Editor Avançado Dependências de Consulta Visualização dos Dados Avançadas Dependências

tas [1] < = Table.TransformColumnTypes(Fonte,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}})

Por exemplo, foi identificado que na primeira coluna, é o Bolo com formato de Texto, a segunda são os Materiais com formato de Texto também.

```
table.TransformColumnTypes(#"Cabeçalhos Promovidos",{{"Bolo", type text}, {"Materiais", type text}, type text}, {"", type time}})
```

Vamos corrigir somente a última coluna, clicando no relógio e alterando para texto também.

A^B C Tempo

A	B	C
1:30		null
1:30		null
Chocolate em Pó e Chocolate Granulado		02:00:00
2:00		null
Morango e Chantilly		02:30:00
1:00		null

type text}, {"", type time}})

Essa correção além de alterar a tipagem na coluna, alterá também na “Barra de Fórmulas”, que é a Linguagem M. Podemos realizar a “Substituição” pois não queremos que seja uma nova etapa e seja possível retornar, e sim uma substituição “eterna”.

A^B C Tempo

A	B	C
1:30		null
1:30		null
Chocolate em Pó e Chocolate Granulado		02:00:00
2:00		null
Morango e Chantilly		02:30:00
1:00		null

type text}, {"", type text}})

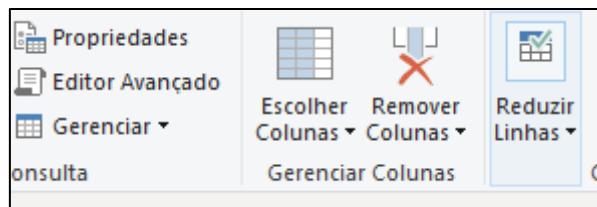
Veja que essa coluna está com nome em branco. Vamos renomear clicando duas vezes no cabeçalho e no local vazio. Vamos renomear para “Aux” que é uma abreviação para Auxiliar. Pois é somente auxiliar, passageira, essa coluna será removida após a devida alocação dos dados.

A	B	C Aux
		2:00

Removendo linhas Nulas

Lembra das linhas inferiores que estavam com valores nulos e vazios?

Para remover, procure por “Reducir/ Remover Linhas” e depois “Remover Linhas Inferiores”

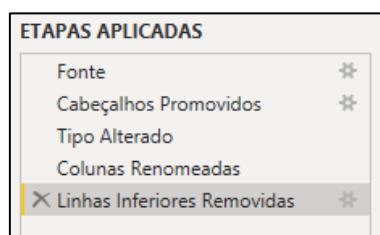


Como são 4 linhas vazias, vamos colocar as 4 e dar o Ok.



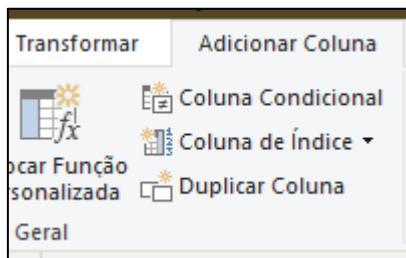
Limpeza dos Dados

Tudo isso que está sendo feito até agora, são limpezas de dados.

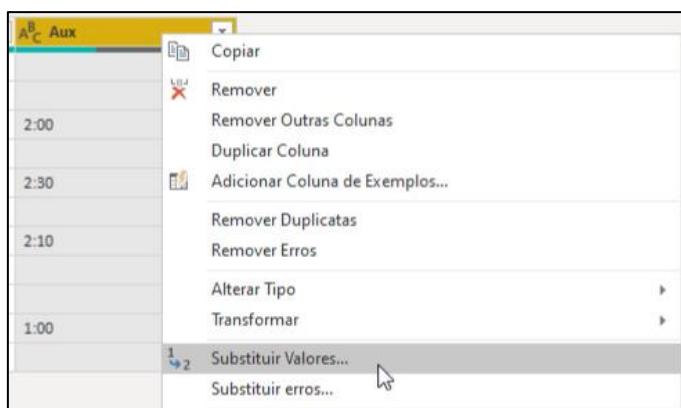


Adicionar Coluna

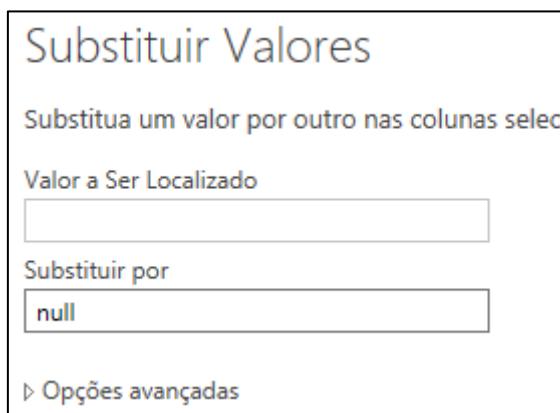
A ferramenta para adicionar “Coluna > Coluna Adicional”.



Antes de adicionar a coluna, precisamos tratar. Para adicionar, antes precisamos voltar na coluna “Aux” que possui valores brancos e vamos utilizar a opção “Substituir Valores”



Isso facilitará a trabalhar com os valores em branco dessa coluna. A substituição será, procurar por valores em branco e substituir por valores nulos.



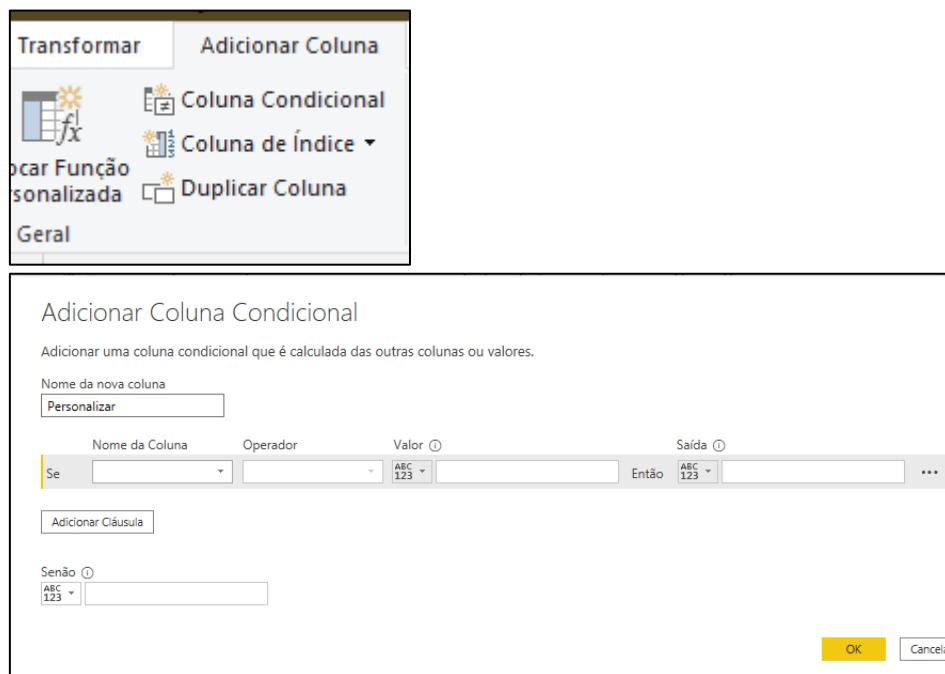
E apertar o botão “Ok”.

A	B	C	Aux
			null
			null
2:00			null
2:30			null
2:10			

Coluna Condisional

É uma coluna criada somente para atender uma condição **determinada**. É para auxiliar também a tratativa dos dados das demais colunas.

Agora com a coluna Aux tratada sem valores em branco. Podemos voltar a Coluna Adicional.



Interpretando a Condisional.

Vamos criar coluna que a chamaremos de “Aux2”. Pela condicional, Se a Coluna AUX (vai varrer linha a linha, ou seja, procurar linha por linha se na coluna Aux tem algum valor que seja “Null”) e se for verdadeiro, então substitua pelos valores da Coluna Tempo.



Continuando, se a condicional não for verdadeira, ou seja, se ao varrer linha a linha não encontrar nenhum valor nulo, pois então traga o próprio valor(que está na coluna Aux). Que se não for nulo, então é o valor do Tempo.

Isso vai gerar uma coluna Aux2, que agora está correta com os valores dos Tempos.

A ^B _C Tempo	A ^B _C Aux	A ^B _C Aux2
1:30	null	1:30
1:30	null	1:30
Chocolate em Pó e Chocolate Granulado	2:00	2:00
2:00	null	2:00
Morango e Chantilly	2:30	2:30
1:00	null	1:00
Baunilha e Chocolate	2:10	2:10
1:30	null	1:30
1:20	null	1:20
Queijo e Leite Condensado	1:00	1:00
1:40	null	1:40

E a fórmula responsável pela criação:

```
= Table.AddColumn(#"Valor Substituído", "Aux2", each if [Aux] = null then [#"Tempo"] else [Aux])  
= Table.AddColumn(#"Valor Substituído", "Aux2", each if [Aux] = null then [#"Tempo"] else [Aux])
```

Remover Colunas

Veja que agora não é mais necessário a coluna auxiliar “Aux”.

A ^B _C Tempo	A ^B _C Aux	A ^B _C Aux2
1:30	null	1:30
1:30	null	1:30
Chocolate em Pó e Chocolate Granulado	02:00	02:00
2:00	null	2:00
Morango e Chantilly	02:30	02:30
1:00	null	1:00
Baunilha e Chocolate	02:10	02:10
1:30	null	1:30
1:20	null	1:20
Queijo e Leite Condensado	01:00	01:00
1:40	null	1:40

Afina, a coluna auxiliar “Aux2” já possui os valores dos tempos que foram com a fórmula lógica na página 56. Portanto, com mouse em cima da “Aux” > “Botão Direito” > “Remover Coluna”.

A ^B _C Bolo	A ^B _C Materiais	A ^B _C Tempo	A ^B _C Aux2
1 Bolo de Fubá	Farinha e Milho	1:30	1:30
2 Bolo de Banana	Farinha e Banana	1:30	1:30
3 Bolo Brigadeiro	Farinha	Chocolate em Pó e Chocolate Granulado	02:00
4 Bolo de Laranja	Farinha e Laranja	2:00	2:00
5 Bolo Cheesecake	Farinha	Morango e Chantilly	02:30
6 Bolo de Ameixa	Farinha e Ameixa	1:00	1:00
7 Bolo de Formigueiro	Farinha	Baunilha e Chocolate	02:10

É muito importante manter o Data Set limpo!

Coluna de Exemplo

É uma *feature* do PB que permite adicionar uma **coluna totalmente em branco** e o analista pode manipular da forma que for mais conveniente.

Para adicionar a coluna exemplo:

The screenshot shows the Power BI ribbon with the 'Adicionar Coluna' tab selected. Below the ribbon, there is a table with two columns: 'A^B C Materiais' and 'A^B C Tempo'. The 'Materiais' column contains items like 'Farinha e Milho', 'Farinha e Banana', and 'Farinha'. The 'Tempo' column contains times like '1:30', '1:30', 'Chocolate em Pó e Chocolate Granulado', '2:00', 'Morango e Chantilly', and '1:00'. To the right of the table, a new column is being added, labeled 'Coluna1', which is currently empty.

Por que precisamos criar a “Coluna de Exemplo”? Porque precisamos unificar os textos que estão presentes nas colunas “Materiais” e “Tempo”, que pela criação da planilha em **.CSV** gerou o problema da separação por causa da vírgula. Ou seja, precisamos unir o “Farinha” com o “Chocolate em Pó e Chocolate Granulado”,

A ^B C Materiais	A ^B C Tempo
Farinha e Milho	1:30
Farinha e Banana	1:30
Farinha	Chocolate em Pó e Chocolate Granulado

Só que ao mesmo tempo, o PB não será inteligente o suficiente para não unir o “Farinha e Milho” e “1:30” na primeira linha. Isso é um problema para tratar também.

Importante!

Não clique no “Ok” antes de seguir os passos abaixo, somente clique no “Ok” quando terminar de editar essa coluna de Exemplo.

The screenshot shows the 'Adicionar Coluna de Exemplos' dialog box. It displays a table with four columns: 'A^B C Bolo', 'A^B C Materiais', 'A^B C Tempo', and 'A^B C Aux2'. The 'Bolo' column has 'Bolo de Fubá'. The 'Materiais' column has 'Farinha e Milho'. The 'Tempo' column has '1:30'. The 'Aux2' column has '1:30'. To the right of the table, a new column is being added, labeled 'Mesclado', which contains the combined text 'Farinha e Milho, 1:30'. At the bottom right of the dialog box, the 'OK' button is highlighted with a yellow box.

Exonerar a Engine das Colunas no Data Set

O primeiro passo para utilizar a “Coluna de Exemplo” é exonerar as colunas que não farão parte dessa coluna. Para isso, clique no botão flag ao lado do nome de cada coluna que não será utilizada, ou seja, nem a “Bolo” e nem a “Aux2”.

A ^B _C Bolo	<input checked="" type="checkbox"/> A ^B _C Materiais	<input checked="" type="checkbox"/> A ^B _C Tempo	<input checked="" type="checkbox"/> A ^B _C Aux2
Bolo de Fubá	Farinha e Milho	1:30	1:30
Bolo de Banana	Farinha e Banana	1:30	1:30
A ^B _C Bolo	<input type="checkbox"/> A ^B _C Materiais	<input checked="" type="checkbox"/> A ^B _C Tempo	<input checked="" type="checkbox"/> A ^B _C Aux2
Bolo de Fubá	Farinha e Milho	1:30	1:30
Bolo de Banana	Farinha e Banana	1:30	1:30

Por que fazer isso? Porque a coluna de exemplo “varre” todo o data set para a construção da Exemplo.

Pode parecer que não é a forma correta, mas, com o cursor do mouse, clique na linha correspondente da “C.Exemplo”, conforme abaixo.

Veja, queremos unir o que existe nessas duas colunas na Exemplo (que é a Coluna1).

A ^B _C Materiais	<input checked="" type="checkbox"/> A ^B _C Tempo	<input checked="" type="checkbox"/> A ^B _C Aux2	<input type="checkbox"/> Coluna1
Farinha e Milho	1:30	1:30	
Farinha e Banana	1:30	1:30	
Farinha	Chocolate em Pó e Chocolate Granulado	02:00	

Vamos digitar uma única vez o texto que está presente nas colunas e o **Power BI** será inteligente para identificar que queremos unir as colunas, e apertar o “Enter” após terminar de digitar, veja o resultado. O **Power BI** é inteligente para entender que queremos “mesclar” essas duas colunas.

A ^B _C Materiais	<input checked="" type="checkbox"/> A ^B _C Tempo	<input checked="" type="checkbox"/> A ^B _C Aux2	Mesclado
Farinha e Milho	1:30	1:30	Farinha e Milho, 1:30
Farinha e Banana	1:30	1:30	Farinha e Banana, 1:30
Farinha	Chocolate em Pó e Chocolate Granulado	02:00	Farinha, Chocolate em Pó e Chocol...
Farinha e Laranja	2:00	2:00	Farinha e Laranja, 2:00
Farinha	Morango e Chantilly	02:30	Farinha, Morango e Chantilly
Farinha e Ameixa	1:00	1:00	Farinha e Ameixa, 1:00

Mas, veja que além dos textos, nesse ponto, o PB não é tão inteligente. Pois ele também nos trouxe o tempo. Precisamos tratar esses valores agora.

Agora sim, podemos clicar no “Ok”.

Adicionar Coluna de Exemplos

Insira valores de exemplo para criar uma nova coluna (Ctrl+Enter para aplicar).
Transformar: Text.Combine({Text.Trim(["# Materiais"]), " ", ["# Tempo"]})

A ^B _C Bolo	<input type="checkbox"/> A ^B _C Materiais	<input checked="" type="checkbox"/> A ^B _C Tempo	<input checked="" type="checkbox"/> A ^B _C Aux2	Mesclado
Bolo de Fubá	Farinha e Milho	1:30	1:30	Farinha e Milho, 1:30

OK Cancelar

Agora voltamos ao *Data Set* e a “Coluna de Exemplo” agora é presente e está com a união de valores/texto das colunas “Materiais” e “Tempo”.

A ^B _C Materiais	A ^B _C Tempo	A ^B _C Aux2	A ^B _C Mesclado
Farinha e Milho	1:30	1:30	Farinha e Milho, 1:30
Farinha e Banana	1:30	1:30	Farinha e Banana, 1:30
Farinha	Chocolate em Pó e Chocolate Granulado	02:00	Farinha, Chocolate em Pó e Chocolate Granulado
Farinha e Laranja	2:00	2:00	Farinha e Laranja, 2:00
Farinha	Morango e Chantilly	02:30	Farinha, Morango e Chantilly

Agora, precisamos de uma “Coluna Condisional” que será nossa auxiliar de número 3.

Clique em “Coluna Condisional” e vamos criar a condição.

Vamos interpretar essa possível condição abaixo.

Note que a condição se forma como: “Se a coluna Mesclado, conter um valor de número 1, então retorne os valores presentes na Materiais.”

Nome da Coluna	Operador	Valor	Saída
Se Mesclado	contém	ABC 123	Então
		1	Materiais

Pórem, **essa não é a melhor solução**. Estamos lidando com horas, precisaríamos de 10 condicionais para atendermos esse tipo de lógica.

Portanto, vamos utilizar o dois pontos (:) para utilizar a lógica de remover horas.

Analisaremos abaixo: “Se a coluna “Mesclado” conter “:” então, retorne a coluna dos “Materiais”, se não, traga o valor do próprio valor da coluna “Mesclado”.

Nome da Coluna	Operador	Valor	Saída
Se Mesclado	contém	ABC 123	Então
		:	Materiais
Adicionar Cláusula			
Senão	Mesclado		

Após isso, podemos digitar o “Ok”. E agora é necessário remover as demais colunas do “Data Set”.

A ^B _C Bolo	A ^B _C 123 Materiais	A ^B _C Tempo
Bolo de Fubá	Farinha e Milho	1:30
Bolo de Banana	Farinha e Banana	1:30
Bolo Brigadeiro	Farinha, Chocolate em Pó e Chocolate Granulado	02:00
Bolo de Laranja	Farinha e Laranja	2:00
Bolo Cheesecake	Farinha, Morango e Chantilly	02:30
Bolo de Ameixa	Farinha e Ameixa	1:00
Bolo de Formigueiro	Farinha, Baunilha e Chocolate	02:10
Bolo de Abacaxi	Farinha e Abacaxi	1:30
Bolo de Maçã	Farinha e Maçã	1:20
Bolo de Queijo	Farinha, Queijo e Leite Condensado	01:00
Bolo de Goiabada	Farinha e Calda de Goiabada	1:40

Limpeza e Tipagem

Precisamos limpar e fazer a correta tipagem dos dados. Note que na página 60, podemos analisar que em “Bolo” e “Materiais” estão com a tipagem de acordo pois são “Textos”, mas a coluna “Tempo” precisa ser duração. Clique no “abc” ao lado do nome e selecione para “duração”.

A ^B _C Bolo	ABC 123 Materiais	Tempo
Bolo de Fubá	Farinha e Milho	0.01:30:00
Bolo de Banana	Farinha e Banana	0.01:30:00
Bolo Brigadeiro	Farinha, Chocolate em Pó e Chocolate Granulado	0.02:00:00
Bolo de Laranja	Farinha e Laranja	0.02:00:00
Bolo Cheesecake	Farinha, Morango e Chantilly	0.02:30:00

Trim

Agora sobre a limpeza, note que o texto na coluna dos “Materiais” em algumas linhas existe um espaço.

Farinha e Milho
Farinha e Banana
Farinha, Chocolate em Pó e C
Farinha e Laranja
Farinha, Morango e Chantilly
Farinha e Ameixa

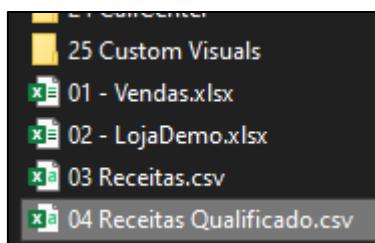
Para remover esse espaço entre a “coluna” e o texto > Clique na coluna > Transformar > Cortar. O “Cortar”, é responsável para remover esse espaço. Isso se chama “trimar”.

The screenshot shows the context menu for the 'Materiais' column. The 'Transformar' (Transform) option is selected. Underneath it, the 'Cortar' (Cut) option is highlighted with a cursor. Other options visible include 'Copiar' (Copy), 'Remover' (Delete), 'Remover Outras Colunas' (Remove Other Columns), 'Duplicar Coluna' (Duplicate Column), 'Adicionar Coluna de Exemplos...' (Add Example Column), 'Remover Duplicatas' (Remove Duplicates), 'Remover Erros' (Remove Errors), 'Alterar Tipo' (Change Type), 'Substituir Valores...' (Replace Values), 'Substituir erros...' (Replace Errors), 'Dividir Coluna' (Split Column), 'Agrupar por...' (Group by), 'Preenchimento' (Fill), 'Transformar Colunas em Linhas' (Transform Columns into Rows), and 'Transformar Outras Colunas em Linhas' (Transform Other Columns into Rows).

A ^B _C Materiais
Farinha e Milho
Farinha e Banana
Farinha, Chocolate em Pó e Chocolate Granulado
Farinha e Laranja
Farinha, Morango e Chantilly

Arquivo com Qualificadores

Para trabalharmos com esse tópico, vamos utilizar o arquivo, “04 Receitas Qualificado.csv”. E vamos abrir como arquivo texto no Bloco de Notas.

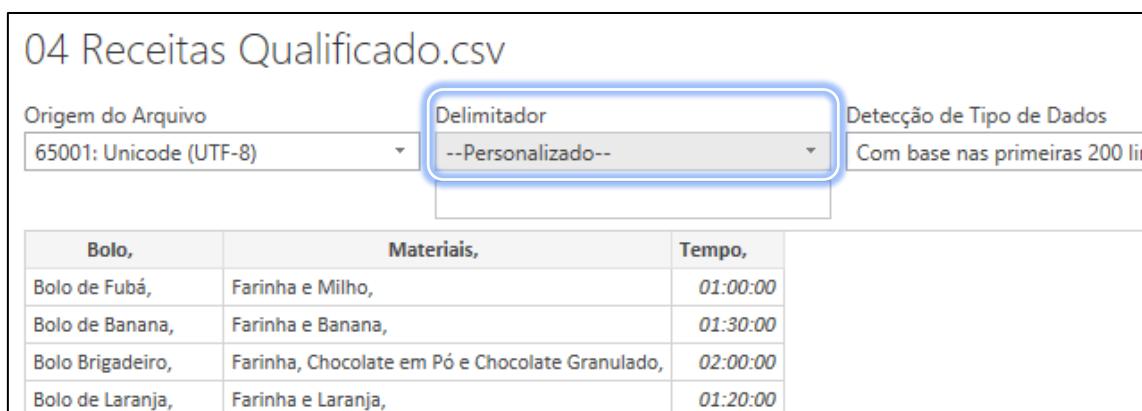


Em alguns sistemas no mundo, quando se exporta (por exemplo, um SAP quando se exporta dados para o Excel) não se escolhe o “DELIMITADOR”, como abaixo, o **delimitador** entre dados são as **vírgulas**. Os delimitadores são os agentes que delimitam as colunas, quais serão os dados. Enquanto os “QUALIFICADORES”, como abaixo, os **qualificadores** são representados pelas **aspas** (“dados”). Os qualificadores envolvem as colunas, mesmo que eu tenha uma vírgula, o qualificador envolve a vírgula e não causa problema.

04 Receitas Qualificado.csv - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
Bolo, Materiais, Tempo
"Bolo de Fubá", "Farinha e Milho", "1:00"
"Bolo de Banana", "Farinha e Banana", "1:30"
"Bolo Brigadeiro", "Farinha, Chocolate em Pó e Chocolate Granulado", "2:00"
"Bolo de Laranja", "Farinha e Laranja", "1:20"

Porém, o **Power BI** não é uma ferramenta de ETL completa. Isso vai gerar um problema.

A melhor forma, portanto, para carregar os dados iniciais, é tratar com a origem Unicode e o “Delimitador” como: “- -PERSONALIZADO- -” e sem nenhuma regra, apenas em branco.



Bolo,	Materiais,	Tempo,
Bolo de Fubá,	Farinha e Milho,	01:00:00
Bolo de Banana,	Farinha e Banana,	01:30:00
Bolo Brigadeiro,	Farinha, Chocolate em Pó e Chocolate Granulado,	02:00:00
Bolo de Laranja,	Farinha e Laranja,	01:20:00

Após isso, apertar em Transformar Dados.

Veja que dessa forma, como os dados foram qualificados pelas aspas e dentro das aspas continham vírgulas, não sofremos nenhum problema ao carregar esses dados quando informamos ao PB que o delimitador é personalizado.

	A ^B C Bolo,	A ^B C Materiais,	L Tempo,
1	Bolo de Fubá,	Farinha e Milho,	01:00:00
2	Bolo de Banana,	Farinha e Banana,	01:30:00
3	Bolo Brigadeiro,	Farinha, Chocolate em Pó e Chocolate Granulado,	02:00:00
4	Bolo de Laranja,	Farinha e Laranja,	01:20:00
5	Bolo Cheesecake,	Farinha, Morango e Chantilly,	02:30:00

A primeira movimentação será transformar o **Tempo** em **Text**.

Feito isso, perceba que nas linhas, existem vírgulas após os textos, por exemplo “Bolo de Fubá,” ou em “Farinha e Milho,“. Existem vírgulas que não que estejam erradas, mas não são mais necessárias para a produção do relatório.

Extração de Texto (Delimitador no Final)

Para remover essas vírgulas, poderíamos apenas Localizar & Substituir por nada. Porém, vamos aprofundar nas *features* e utilizar “Transformar > Extrair > Texto Antes o Delimitador”.

	A ^B C Bolo,	A ^B C Materiais,	A ^B C Tempo
1	Bolo de Fubá,	Farinha e Milho,	1:00
2	Bolo de Banana,	Farinha e Banana,	1:30
3	Bolo Brigadeiro,	Farinha, Chocolate em Pó e Chocolate Granulado,	2:00

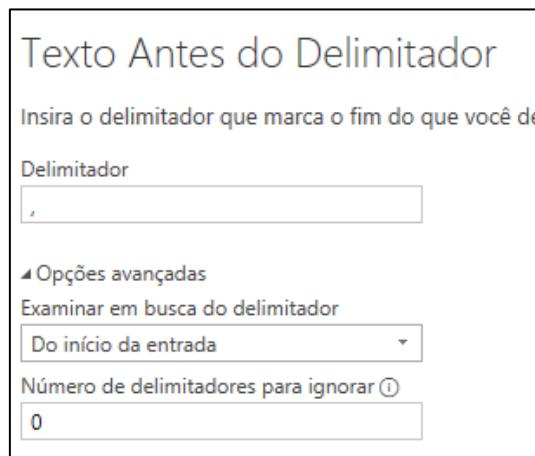
Nota Importante: Se entende que “a vírgula vai ser extraída da célula”. Porém, não é essa a interpretação correta. A extração é o que ficará na tabela, na célula. Portanto, se deseja extrair o que está ANTES do DELIMITADOR (que é a vírgula), e o resultado será apenas o “Bolo de Fubá”, todos sem vírgula no final.

Nota Importante II: **Não** podemos utilizar o mesmo método que foi utilizado para limpar a vírgula em “Bolos” para limpar a coluna dos “Materiais”, pois em algumas linhas dessa coluna, note que teremos “Farinha, Chocolate em Pó e Chocolate Granulado”, se utilizarmos o mesmo método, o resultado seria só “Farinha”, pois o PB irá recortar tudo e isso estará errado!

Extração de Texto (Delimitador no Meio)

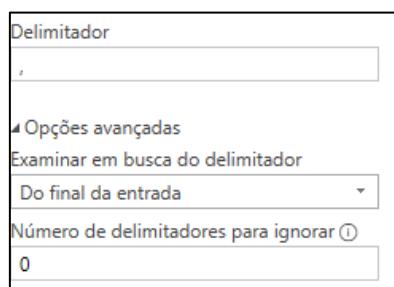
O segundo método que realizaremos a extração será quase igual ao utilizado anteriormente, porém, utilizaremos as opções avançadas dessa extração.

"Transformar > Extrair > Texto Antes o Delimitador > Opções Avançadas".



O Delimitador continuará sendo a vírgula. Porém, note que a procura, o exame é feito da esquerda para a direita, da entrada até o fim.

Podemos substituir para que seja da direita para a esquerda ou do final para o começo. E portanto, ele limpará o que estiver antes e permanecerá com o dado que encontrou após.



Resultado:

A ^B _C Bolo,	A ^B _C Materiais,	A ^B _C Tempo,
Bolo de Fubá	Farinha e Milho	1:00
Bolo de Banana	Farinha e Banana	1:30
Bolo Brigadeiro	Farinha, Chocolate em Pó e Chocolate Granulado	2:00
Bolo de Laranja	Farinha e Laranja	1:20
Bolo Cheesecake	Farinha, Morango e Chantilly	2:30

Agora as duas colunas que continham as vírgulas, estão de acordo. Apenas realizar o corte do *Trim* no início para ajustar corretamente.

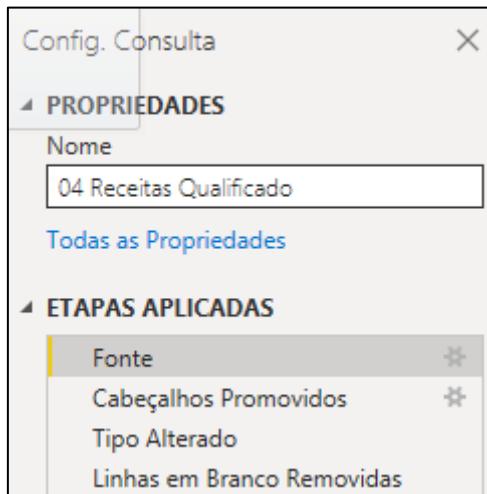
Fontes de Dados

Ainda sobre a planilha do exercício 04 – Qualificadores. Esse capítulo serve para saber qual arquivo está ligado a qual data set está ligado. Em determinados momentos, alguns arquivos estarão ligados a mais de um único data set.

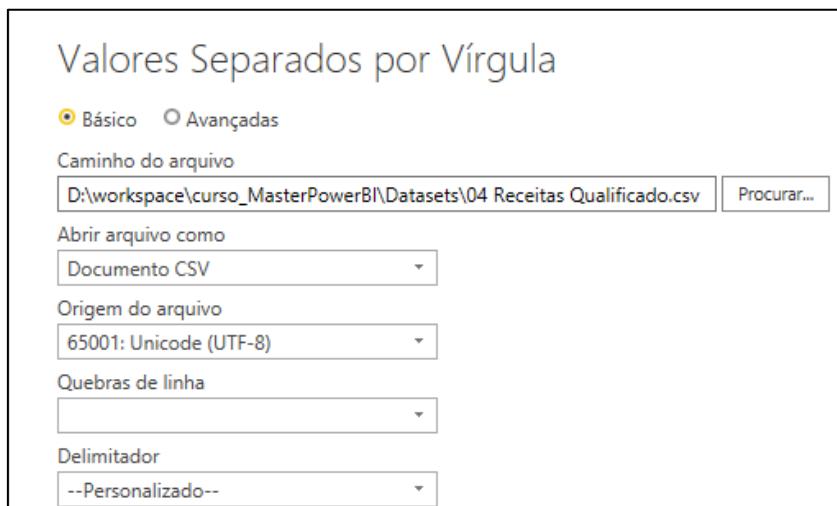
Abrir o Power Query

Vamos em “[Página Inicial > Transformar Dados](#)” para fazer as edições no **Power Query** dessa planilha.

Clique no ícone da “Engrenagem” em “Etapas Aplicadas” e em “Fontes”.



Aqui vemos qual o caminho que o arquivo está localizado.

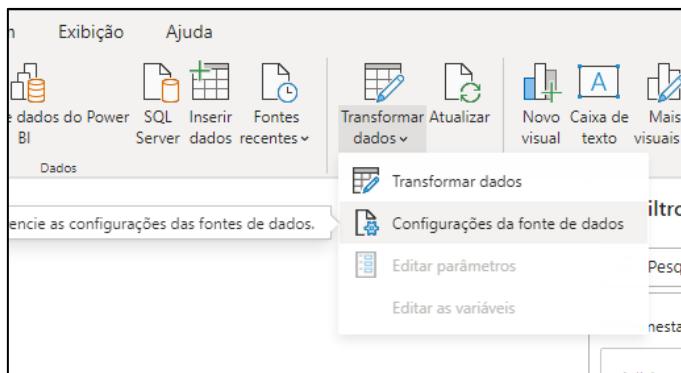


Essa é a forma mais demorada para se encontrar.

Vamos fechar sem nenhuma aplicação para voltarmos ao **Power View** (a tela inicial, ou seja, de visualização).

Acessando o Power Query por Config

Outra forma, é acessar através de Transformar Dados > Configurações da fonte de Dados



Configurações da fonte de dados

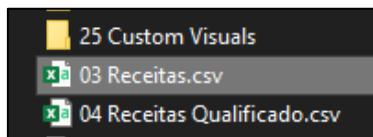
Gerenciar configurações para fontes de dados que você conectou usando o Power Query

Fontes de dados no arquivo atual Permissões globais

Pesquisar configurações da fonte de dados

d:\workspace\curso_masterpower...ts\04 receitas qualificado.csv

Por exemplo, vamos inserir mais de uma fonte de dados em um único relatório. Vamos em “Obter Dados > csv (nesse exemplo será Receitas.csv)”



E vamos carregar, sem fazer nenhum tratamento de dados, é apenas para visualização.

Veja que agora temos duas planilhas de fontes de dados diferentes em um único relatório

Campos

Pesquisar

03 Receitas

- Bolo, Mater...
- Σ Column1

04 Receitas Qualif...

- Bolo
- Materiais
- Σ Tempo

E se fizermos as consultas novamente para identificarmos quais são as fontes de dados, em de “Transformar Dados > Configurações da fonte de Dados”.

Configurações da fonte de dados

Gerenciar configurações para fontes de dados que você conectou usando o Power BI Desktop.

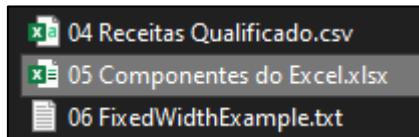
Fontes de dados no arquivo atual Permissões globais

Pesquisar configurações da fonte de dados

 d:\workspace\curso_masterpowerbi\datasets\03 receitas.csv
 d:\workspace\curso_masterpower...ts\04 receitas qualificado.csv

Componentes do Excel

Para esse capítulo, vamos utilizar a planilha “05 – Componentes do Excel”.



E vamos abrir para entender a planilha.

Numero	Nome	Sexo	Nota	Calculo	Peso	2
1	Bruna	F	9,7	19,4	As notas em questão referem-se a matéria de Matemática	
2	Renata	F	9,8	19,6		
3	André	M	8,9	17,8		
4	Leonardo	M	9	18		
5	Felipe	M	8,9	17,8	Resultado	92,6

Veja, existem números, nome, sexo, a nota e o cálculo (que é a nota vezes o Peso que está fora da tabela). Enquanto o Resultado é a soma de todos os Cálculos.

Mas veja que a planilha “cresce” para o lado(>) e não para baixo(v).

Numero	Nome	Sexo	Nota	Calculo	Peso	2
1	Bruna	F	9,7	19,4	As notas em questão referem-se a matéria de Matemática	
2	Renata	F	9,8	19,6		
3	André	M	8,9	17,8		
4	Leonardo	M	9	18		
5	Felipe	M	8,9	17,8	Resultado	92,6

Vamos tentar carregar essa planilha no Power BI.

A primeira “surpresa” é que encontramos muito mais informação que nos demais exemplos.

A screenshot of the Power BI 'View Options' pane. It shows a tree view of the file '05 Componentes do Excel.xlsx [4]'. The 'tbAlunos' table is selected. Other tables listed are 'Planilha1', 'Planilha2', and 'Peso'.

Vamos destrinchar cada um nas próximas páginas.

A primeira informação que é a “tbAlunos”, é a tabela de alunos que foi criada no Excel.

Numero	Nome	Sexo	Nota	Calculo
1	Bruna	F	9,7	19,4
2	Renata	F	9,8	19,6
3	André	M	8,9	17,8
4	Leonardo	M	9	18
5	Felipe	M	8,9	17,8

A “Peso” é apenas a fórmula do Peso. O que o “Peso” utiliza para achar o “Cálculo”.

Column1
2

Enquanto a “Planilha1” formatada com Tabelas, mas está com muito mais informação que está trazendo “lixo”, dados que não serão utilizados e que precisarão de limpeza.

iexo	Nota	Calculo	Column6	Column7	Column8
;	9,7	19,4	null	Peso	2
;	9,8	19,6	null	As notas em questão	null
;	8,9	17,8	null	referem-se a matéria	null
;	9	18	null	de Matemática	null
;	8,9	17,8	null	Resultado	92,6

A “Planilha2”, segue o mesmo, a diferença que no Excel a Planilha1 estava com tabela e a 2 não estava.

iexo	Nota	Calculo	Column6	Column7	Column8
;	9,7	19,4	null	Peso	2
;	9,8	19,6	null	As notas em questão	null
;	8,9	17,8	null	referem-se a matéria	null
;	9	18	null	de Matemática	null
;	8,9	17,8	null	Resultado	92,6

Ainda assim, todas as opções precisarão de tratamento e modelagem.

Durante a modelagem, é aconselhável trabalhar com Tabelas. Por exemplo, em Design podemos verificar o nome da Tabela, e é interessante utilizar um prefixo sempre que se utilizar tabelas, como “tbAlunos”

	A	B	C	D	E	F	G	H	I	J
1	Numero	Nome	Sexo	Nota	Calculo					
2	1	Bruna	F	9,7	19,4	Peso		2		
3	2	Renata	F	9,8	19,6	As notas em questão				
4	3	André	M	8,9	17,8	referem-se a matéria				
5	4	Leonardo	M	9	18	de Matemática				
6	5	Felipe	M	8,9	17,8	Resultado		92,6		

Veja, como é possível criar tabelas, facilmente.

Número	Linguagem
1	C#
2	Java
3	Python

Nome da Tabela:
tbLinguagens
Redimensionar Tabela
Propriedades

O nome da tabela será “tbLinguagens”.

Ao voltarmos ao Power BI, e aplicarmos o “atualizar/refresh” podemos ver essa nova tabela criada.

Número	Linguagem
1	C#
2	Java
3	Python

Opções de Exibição

- 05 Componentes do Excel.xlsx [5]
- tbAlunos
- tbLinguagens**

Porém, veja que essa não é a forma correta de se trabalhar! Veja a “bagunça” que está na visualização dos dados no PB, quando chamamos a Planilha1:

Numero	Nome	Sexo	Nota	Calculo	Column6	Column7
1	Bruna	F	9,7	19,4	null	Peso
2	Renata	F	9,8	19,6	null	As notas em
3	André	M	8,9	17,8	null	referem-se
4	Leonardo	M	9	18	null	de Matemá
5	Felipe	M	8,9	17,8	null	Resultado
null	null	null	null	null	null	
null	null	null	null	null	null	
null	null	null	null	null	null	
null	null	null	null	null	null	
null	null	null	null	null	null	
null	null	null	null	null	null	

6	Column7	Column8	Column9	Column10	Column11
null	Peso	2	null	null	null
null	As notas em questão	null	null	null	null
null	referem-se a matéria	null	null	null	null
null	de Matemática	null	null	null	null
null	Resultado	92,6	null	null	null
null		null	null	null	null
null		null	null	null	null
null		null	null	null	null
null		null	null	null	null
null		null	null	Número	Linguagem
null		null	null	1	C#
null		null	null	2	Java
null		null	null	3	Python

Realmente, fica uma confusão de dados, seria necessário muitas etapas para a limpeza.

Portanto, vamos carregar somente as tabelas, “tbAlunos” e a “tbLinguagens”, elas serão importadas corretamente, não é preciso importar a “Planilha1” inteira e vamos Transformar Dados

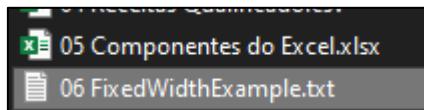
Numero	Nome	Sexo	Nota	Calculo
1	Bruna	F	9,7	19,4
2	Renata	F	9,8	19,6
3	André	M	8,9	17,8
4	Leonardo	M	9	18
5	Felipe	M	8,9	17,8

Número	Linguagem
1	C#
2	Java
3	Python

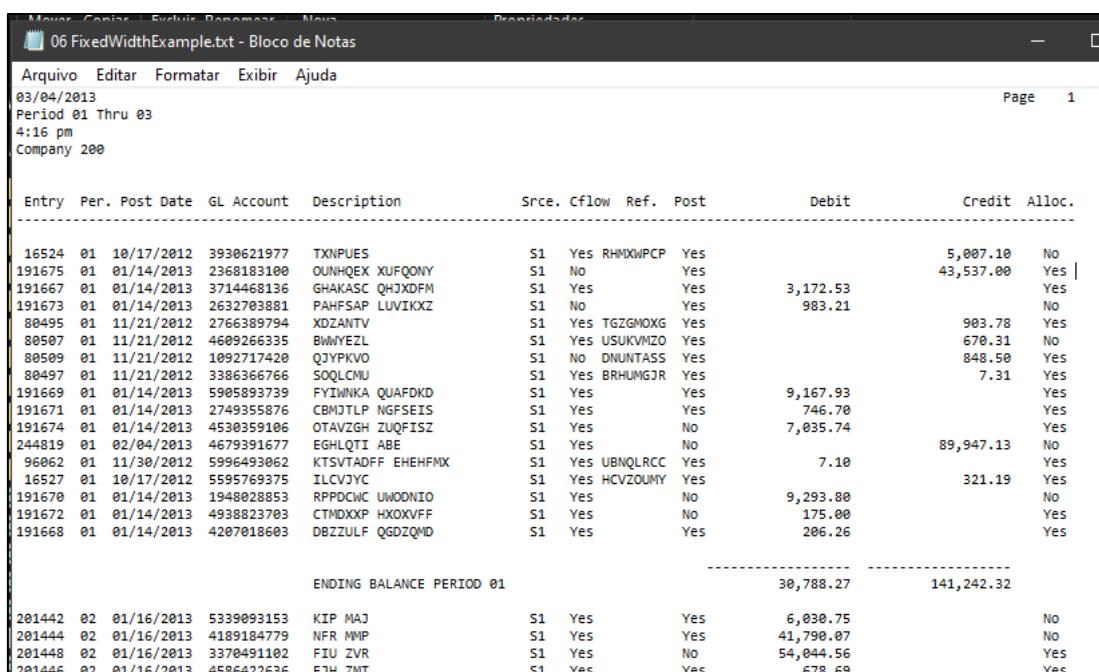
Veja, que as duas tabelas foram carregadas e serão transformadas sem seguida.

Arquivos Posicionais – Tamanho Fixo

Vamos utilizar o arquivo de número “06 FixedWidthExample”.

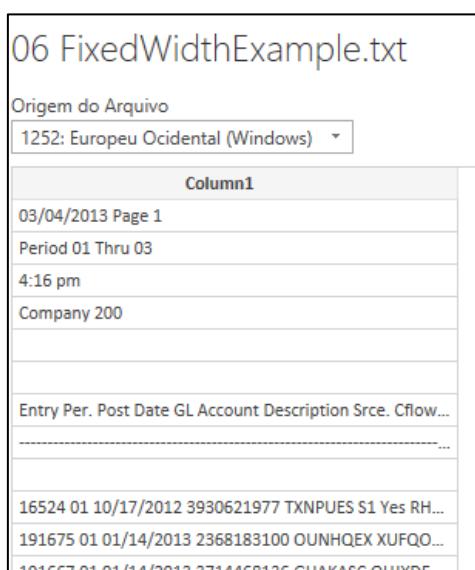


É um arquivo em formato **.txt** (texto), e que se assemelha a um “extrato bancário”, é muito comum esse tipo de arquivo em formato **.txt** e com essa formatação.



Entry	Per.	Post Date	GL Account	Description	Srce.	Cflow	Ref.	Post	Debit	Credit	Alloc.
16524	01	10/17/2012	3930621977	TXNPUES	S1	Yes	RHMXWPCP	Yes		5,007.10	No
191675	01	01/14/2013	2368183100	OUNHQEX XUFQONY	S1	No		Yes		43,537.00	Yes
191667	01	01/14/2013	3714468136	GHAKASC QHJXDFM	S1	Yes		Yes	3,172.53		Yes
191673	01	01/14/2013	2632703881	PAHFSAP LUVIKXZ	S1	No		Yes	983.21		No
80495	01	11/21/2012	2766389794	XDZANTV	S1	Yes	TGZGMXOG	Yes		903.78	Yes
80507	01	11/21/2012	4609266335	BMYEZL	S1	Yes	USUKVMZO	Yes		670.31	No
80509	01	11/21/2012	1892717420	QJYVKVO	S1	No	DNUNTASS	Yes		848.50	Yes
80497	01	11/21/2012	3386366766	SQLCMU	S1	Yes	BRHUMGJR	Yes		7.31	Yes
191669	01	01/14/2013	5905893739	FIWINKA QUAFDKD	S1	Yes		Yes	9,167.93		Yes
191671	01	01/14/2013	2749355876	CBMJTPL NGFSEIS	S1	Yes		Yes	746.70		Yes
191674	01	01/14/2013	4530359106	OTAVZGH ZUQFISZ	S1	Yes		No	7,035.74		Yes
244819	01	02/04/2013	4679391677	EGLHQTJ ABE	S1	Yes		No		89,947.13	No
96062	01	11/30/2012	5996493062	KTSVTADFF EHEHFMX	S1	Yes	UBNQLRCC	Yes	7.10		Yes
16527	01	18/17/2012	5595769375	ILCVJYC	S1	Yes	HCVZQUMY	Yes		321.19	Yes
191670	01	01/14/2013	1948028853	RPPDCWC UWORDNIO	S1	Yes		No	9,293.80		No
191672	01	01/14/2013	4938823703	CTMDXXP HXOXVFF	S1	Yes		No	175.00		Yes
191668	01	01/14/2013	4207018603	DBZZULF QGDZQMD	S1	Yes		Yes	206.26		Yes
-----								ENDING BALANCE PERIOD 01	30,788.27	141,242.32	
201442	02	01/16/2013	5339093153	KIP MAJ	S1	Yes		Yes	6,030.75		No
201444	02	01/16/2013	4189184779	NFR MMP	S1	Yes		Yes	41,790.07		No
201448	02	01/16/2013	3378491182	FIU ZVR	S1	Yes		No	54,044.56		Yes
201446	02	01/16/2013	4586422636	EH4 TMT	S1	Yes		Yes	678.69		Yes

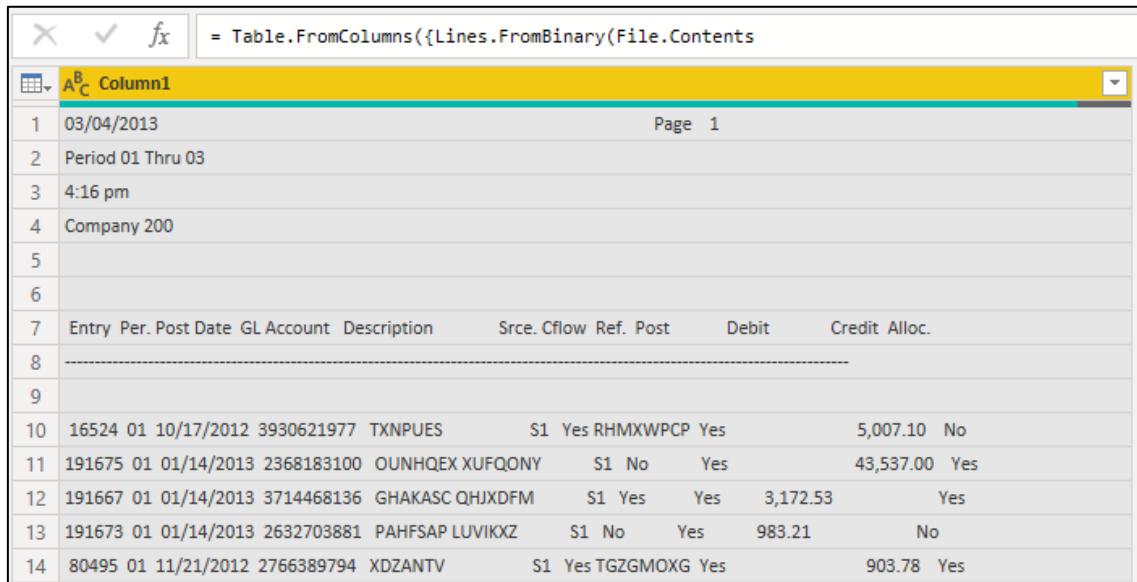
Vamos para a importação desse documento e carregar os dados. Veja quando a importação dos dados no PB.



Repare que não é disponibilizado o Delimitador para esse tipo de arquivo.

O arquivo posicional

Ao carregarmos o arquivo na transformação de dados para o **Power Query**, podemos identificar que estará uma “bagunça”. Isso acontece porque o arquivo dessa forma é um “Arquivo Posicional”.

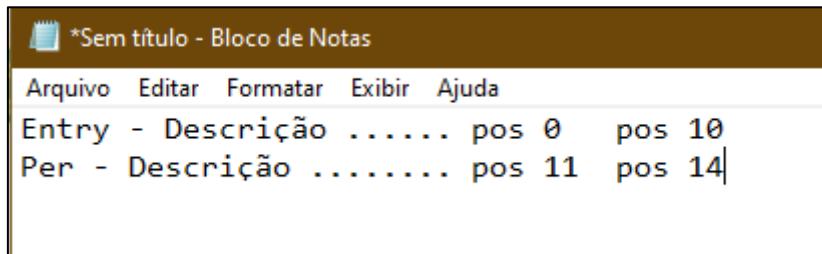


The screenshot shows the Power Query Editor interface. At the top, there's a formula bar with the code: `= Table.FromColumns({Lines.FromBinary(File.Contents`. Below the formula bar is a header row labeled "Column1". The data starts with row 1 containing "03/04/2013" and "Page 1". Rows 2 through 6 are blank. Row 7 contains column headers: "Entry Per. Post Date GL Account Description" and "Srce. Cflow Ref. Post Debit Credit Alloc.". Rows 8 through 14 are transaction data, each with multiple columns of numerical and textual values.

1	03/04/2013										Page 1		
2	Period 01 Thru 03												
3	4:16 pm												
4	Company 200												
5													
6													
7	Entry	Per.	Post	Date	GL Account	Description	Srce.	Cflow	Ref.	Post	Debit	Credit	Alloc.
8													
9													
10	16524	01	10/17/2012	3930621977	TXNPUES		S1	Yes	RHMXWPCP	Yes	5,007.10	No	
11	191675	01	01/14/2013	2368183100	OUNHQEX XUFQONY		S1	No		Yes	43,537.00	Yes	
12	191667	01	01/14/2013	3714468136	GHAKASC QHJXDFM		S1	Yes		Yes	3,172.53	Yes	
13	191673	01	01/14/2013	2632703881	PAHFSAP LUVIKXZ		S1	No		Yes	983.21	No	
14	80495	01	11/21/2012	2766389794	XDZANTV		S1	Yes	TGZGMOXG	Yes	903.78	Yes	

Arquivo Schema

É comum quando recebemos esse tipo de arquivo **.txt**, de recebermos em conjunto um arquivo chamado de “*Schema*”. Esse arquivo *schema* é responsável por indicar as descrições de cada campo e a posição deles. Para esse exemplo, não foi disponibilizado o arquivo *schema*, mas veja um exemplo feito à mão abaixo.



The screenshot shows a Notepad window with the title "*Sem título - Bloco de Notas". The content of the window is a schema definition for a .txt file. It includes a header row with column names and descriptions, followed by two data rows. The schema is as follows:

Entry	-	Descrição	pos	0	pos	10
Per	-	Descrição	pos	11	pos	14

O schema vai ser responsável por indicar uma descrição e a posição inicial e a posição final dos campos.

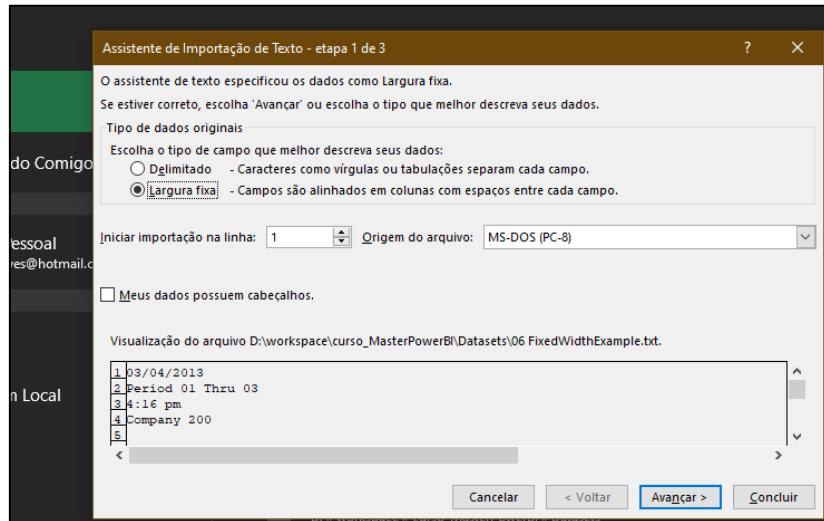
Como carregar esse arquivo para o **Power BI**? Mesmo sem receber o arquivo Schema? E mesmo com o arquivo Schema também.

Podemos seguir da forma que será apresentada na página 74, tanto para se recebermos o schema, quanto se não o recebermos.

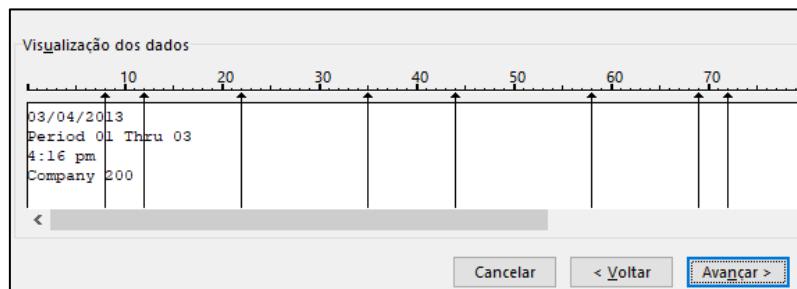
Carregando o arquivo Schema

Para isso, precisamos abrir esse arquivo **.txt** pelo Excel. Abra o Excel e procure pelo arquivo “06 FixedWidthExample”.

Veja que o Excel apresenta um “Assistente de Importação de Texto”. O próprio Excel já realizou a identificação que se trata de um arquivo de **tamanho fixo**. Clique em “Avançar”.



Agora, perceba que existe uma visualização dos dados. Precisamos saber o que são cabeçalhos ou dados inúteis e identificarmos quais são os dados que realmente importam. Por exemplo, abaixo, é uma informação inútil, esses dados são cabeçalhos que não serão usados. Vamos continuar rolando até acharmos dados utilizáveis.



Por exemplo, esses são os dados que realmente importam.

Entry	Per.	Post Date	GL Account	Description	Srce.	Cflow	Ref.	Post
16524	01	10/17/2012	3930621977	TXNPUES	S1	Yes	RHMXWPCP	Yes
191675	01	01/14/2013	2368183100	OUNHQEX XUFQONY	S1	No		Yes

Essas linhas com setas, são **importantíssimas**. Elas são os delimitadores. Devemos ajustar as linhas conforme os dados. Elas precisam estar exatamente no local do dado.

Por exemplo, na imagem acima, a seta da “Description” está alocada erroneamente. É necessário “puxar” essa seta para o início dos dados e remover a seta que está na metade da “Description” que está identificando como dois dados, mas é apenas um. Para adicionar, clique duas vezes, para remover, clique apenas uma.

Antes do Ajuste:

Visualização dos dados									
Entry	Per.	Post Date	GL Account	Description	Srce.	Cflow	Ref.	Post	
16524	01	10/17/2012	3930621977	TXNPUES	S1	Yes	RHMXWPCP	Yes	
191675	01	01/14/2013	2368103100	OUNHQEX XUFQONY	S1	No		Yes	

[Cancelar](#) [**< Voltar**](#) [**Avançar >**](#) [Concluir](#)

Depois do Ajuste:

Visualização dos dados									
Entry	Per.	Post Date	GL Account	Description	Srce.	Cflow	Ref.	Post	
16524	01	10/17/2012	3930621977	TXNPUES	S1	Yes	RHMXWPCP	Yes	
191675	01	01/14/2013	2368103100	OUNHQEX XUFQONY	S1	No		Yes	

[Cancelar](#) [**< Voltar**](#) [**Avançar >**](#) [Concluir](#)

É preciso muito, mas muito cuidado, se o arquivo for imensamente grande, precisa ser analisado coluna a coluna, linha a linha, para que nenhum valor que seja de uma linha, esteja em outra. É muito comum que os números por questão de variação da dezena-centena-milhar, por exemplo:

Veja como existe um espaço entre a linha e os números, mas ele deve acompanhar o número maior que é o início desses dados.

70	80	90	100	110	120
S1	Yes		No	87,077.68	
S1	Yes		Yes	637.36	

11,647,847.35 163,422,325.50

Criando o Próprio Schema com o Bloco de Notas

Agora, é preciso criar o nosso próprio schema. Vamos abrir o Bloco de Notas e vamos começar a anotar as posições que o Excel está nos fornecendo.

Cada espaço são 10 posições, entre os 10, o meio (um risco maior) significa 5 posições. É como se fosse um campo de futebol americano.

Visualização dos dados				
10	20	30		
201659	02	01/16/2013	4205768057	A1
201657	02	01/16/2013	1236840287	B1

Por exemplo, o primeiro campo, que é o “Entry”, acima representado pelo “201659” que é o primeiro dado, que vai até aquela linha que é representado pelo “02” que é a coluna do “Per”. É um trabalho muito complexo e cauteloso que deve ser feito agora.

Visualização dos dados				
10	20	30		
Entry	Per.	Post Date	GL Accoun	
*Sem título - Bloco de Notas				
Arquivo	Editar	Formatar	Exibir	Ajuda
0, 8, 12				

Veja que o “Entry” vai da posição 0 a 8.

A “Per” vai do 8 ao 12. E por aí em diante.

Feito todas as posições, clique em Salvar Como > Na pasta do Data Set > Nomear.

schemafor06.txt - Bloco de Notas	
Arquivo	Editar
0, 8, 12, 24, 37, 64, 69, 73, 83, 91, 110, 128	Ajuda

Essa é a primeira forma de identificação dos arquivos posicionais.

Schema pelo Arquivo Posicional

Ao invés de utilizar o método de calcular o Schema pelo “Assistente de Importação de Texto do Excel”, abra novamente o arquivo em formato **.txt**² no bloco de notas e veja:

O cursos está na segunda linha de dados, em “191675”, veja que no rodapé é possível verificarmos as Linhas e Colunas. Porém, isso está incorreto para o **Power BI**, uma vez que o PB é indexado a zero, precisamos reduzir o número por -1. Ou seja, a Coluna inicial é a coluna 0, como é demonstrando na página acima.

The screenshot shows a Windows Notepad window titled "06 FixedWidthExample.txt - Bloco de Notas". The window contains the following text:

```
Arquivo Editar Formatar Exibir Ajuda
03/04/2013
Period 01 Thru 03
4:16 pm
Company 200

Entry Per. Post Date GL Account Description Srce. Cflow Ref. Post Debit
-----
16524 01 10/17/2012 3930621977 TXNPUES S1 Yes RHMXWPCP Yes
191675 01 01/14/2013 2368183100 OUNHQEX XUFQONY S1 No Yes
191667 01 01/14/2013 3714468136 GHAKASC QHJXDFM S1 Yes Yes 3,172.53
191673 01 01/14/2013 2632703881 PAHFSAP LUVIKXZ S1 No Yes 983.21
80495 01 11/21/2012 2766389794 XDZANTV S1 Yes TGZGMOXG Yes
80507 01 11/21/2012 4609266335 BMVYEZL S1 Yes USUKVMZO Yes
80509 01 11/21/2012 1092717420 QJYPKVO S1 No DNUNTASS Yes
80497 01 11/21/2012 3386366766 SQQLCMU S1 Yes BRHUMGJR Yes
191669 01 01/14/2013 5905893739 FYIWNKA OUAFDKD S1 Yes Yes 9.167.93
< Ln 11, Col 1 100% Unix (LF) UTF-8
```

Mas, o formato que foi escrito pelo próprio schema na página 76, é o formato que o **Power BI** reconhece.

Então, esse é o correto a ser inserido para o PB.

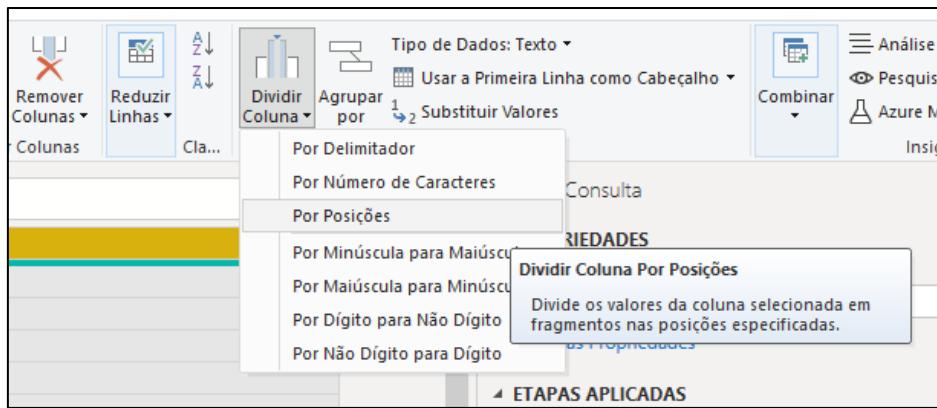
The screenshot shows a Windows Notepad window titled "schemafor06.txt - Bloco de Notas". The window contains the following text:

```
Arquivo Editar Formatar Exibir Ajuda
0, 8, 12, 24, 37, 64, 69, 73, 83, 91, 110, 128|
```

Agora, retornemos ao PB e vamos ajustar e dividir as colunas conforme o nosso schema.

“Página Inicial > Dividir Coluna > Por Posições”.

² Flat File – Arquivo de Texto



Divisão das Colunas

Veja como o **Power BI** entendeu como as colunas estão divididas, ele tentou identificar as posições.

Dividir Coluna pelas Posições

Especifique as posições nas quais dividir a coluna de texto.

Posições

Mas na verdade, essas são as posições corretas.

A schemafor06.txt - Bloco de Notas

Arquivo Editar Formatar Exibir Ajuda

```
0, 8, 12, 24, 37, 64, 69, 73, 83, 91, 110, 128|
```

Portanto, vamos pegar essas posições do Schema e substituir as posições oferecidas pelo **Power BI**.

Dividir Coluna pelas Posições

Especifique as posições nas quais dividir a coluna de texto.

Posições

► Opções avançadas

Após pressionarmos o “ok”.

= Table.TransformColumnTypes(#"Dividir Coluna pelas Posições", {"Column1.1", type text},

	A ^B _C Column1.1	A ^B _C Column1.2	A ^B _C Column1.3	A ^B _C Column1.4	A ^B _C Column1.5
1	03/04/20	13			
2	Period 0	1 Th	ru 03		
3	4:16 pm				
4	Company	200			
5					
6					
7	Entry	Per.	Post Date	GL Account	Description
8	-----	---	-----	-----	-----
9					
10	16524	01	10/17/2012	3930621977	TXNPUES
11	191675	01	01/14/2013	2368183100	OUNHQEX XUFQO
12	191667	01	01/14/2013	3714468136	GHAKASC QHJXDF
13	191673	01	01/14/2013	2632703881	PAHFSAP LUVIKZ

O Power BI já estipula as colunas e as divisões.

Realizando um “trabalho de formiguinha” e com muito cuidado, todos os dados estarão devidamente no seu lugar, devidamente em suas colunas, sem erros.

Mas nas próximas páginas, provocaremos um erro, para que se imagine que o arquivo é imensamente grande e precise de correções

Provocando Erro para Tratamento

Por exemplo, atualmente está correto, nenhum número está sobrepondo outra coluna do qual não pertence.

	5.19
67,155.69	
967,961.55	
	80,220.78
261,211.04	
655.65	
	930,183.41
986.57	
	158,267.02
27.00	

Mas provocaremos o erro para que fique dessa forma, como na segunda, perceba que não foi identificado corretamente e causará um erro.

1º Linha totalmente identificada corretamente

2º Linha identificada erroneamente, causando erro

Arquivo	Editar	Formatar	Exibir	Ajuda
0, 8, 12, 24, 37, 64, 69, 73, 83, 91, 110, 128				
0, 8, 12, 24, 37, 64, 69, 73, 83, 95, 115				

Veja a bagunça que está devido a má identificação ou por um arquivo grande sem identificação ou sem schema enviado pelo seu remetente.

3,172.53	Yes
983.21	No
	903.78 Yes
	670.31 No
	848.50 Yes
	7.31 Yes
9,167.93	Yes
746.70	Yes
7,035.74	Yes
	89,947.13 No
7.10	Yes
	321.19 Yes
9,293.80	No
175.00	Yes
206.26	Yes
-----	-----
30,788.27	1
	41,242.32

Para começarmos, é necessário excluímos e limparmos as linhas, porém, como o PB trabalha de forma “colunar”, não é possível excluímos apenas as linhas em branco, e mesmo se fosse, existem colunas que tem linhas em branco, mas em algumas outras essas linhas estão preenchidas.

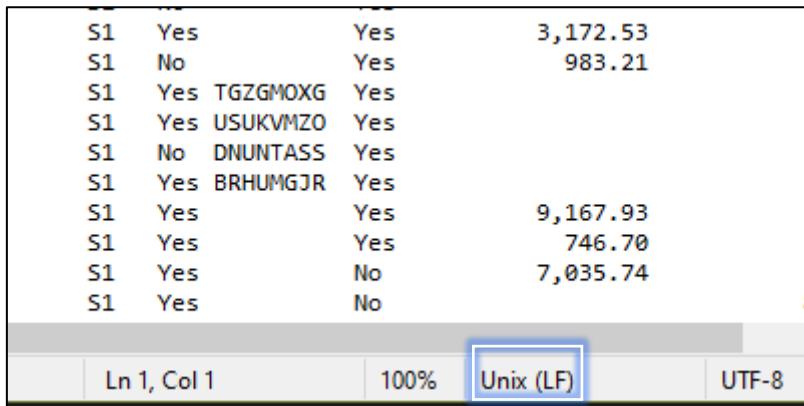
Por exemplo, poderíamos excluir as “linhas” que existem nas colunas do meio, mas veja como existe dado na primeira e na última coluna, isso me impediria.

A ^B _C Column1.5	A ^B _C Column1.6	A ^B _C Column1.7	A ^B _C Column1.8	A ^B _C Column1.9	A ^B _C Column1.10	A ^B _C Column1.11
OTAVZGH ZUQFISZ	S1	Yes		No	7,035.74	Yes
EGHLQT1 ABE	S1	Yes		No		89,947.13 No
KTSVTADFF EHEHFMX	S1	Yes	UBNQLRCC	Yes	7.10	Yes
ILCVIYC	S1	Yes	HCVZOUAMY	Yes		321.19 Yes
RPPDCWC UWODNIO	S1	Yes		No	9,293.80	No
CTMDXXP HXOXVFF	S1	Yes		No	175.00	Yes
DBZZULF QGDZQMD	S1	Yes		Yes	206.26	Yes
			-----	-----	-----	-----
ENDING BALANCE PERIOD 01					30,788.27	1
KIP MAJ	S1	Yes		Yes	6,030.75	No
NFR MMP	S1	Yes		Yes	41,790.07	No

Porém, vale lembrar que na verdade, esses valores, são os TOTAIS das somas. Ou seja, eu deveria excluir sim, justamente porque o **Power BI** já proporciona a Somatória dos valores e isso ocasionaria duplicidade na somatória.

Instalação do Notepad++

A remoção das linhas vazias não funciona porque devido a origem dos dados no bloco de notas, ser diferente de cada sistema operacional (Windows trabalha diferente de MacOS e de Linux) e portanto, o “ENTER” dando para cada linha no bloco pode ser o diferencial e isso causa erro. O UNIX (LF) é a quebra da linha pelo ENTER.

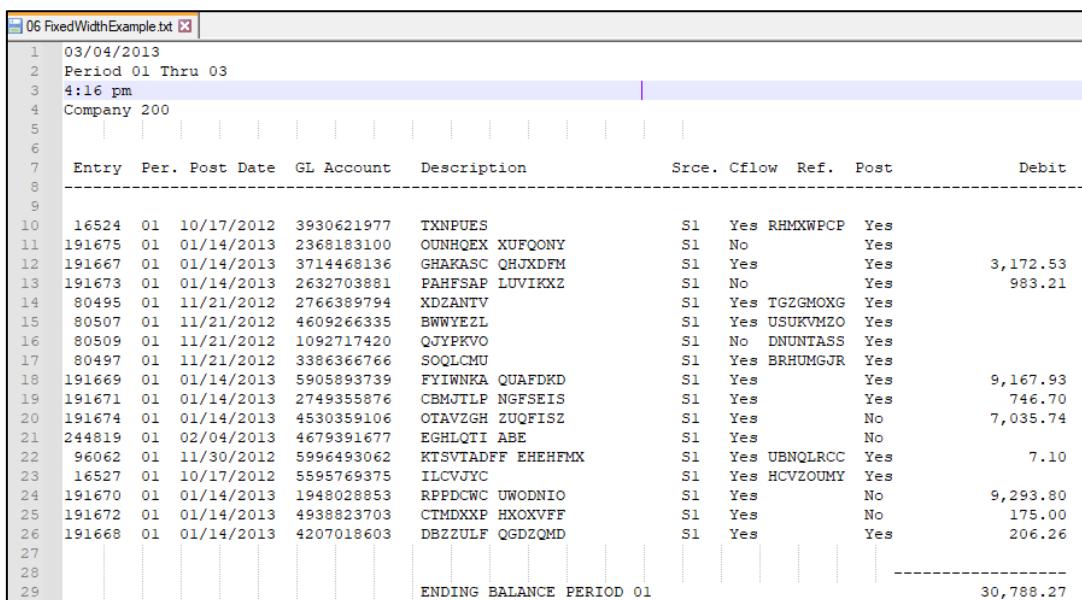


The screenshot shows a Notepad++ window displaying a table of data. The table has columns for S1, Yes, Yes, and three numerical values (3,172.53, 983.21, 9,167.93, 746.70, 7,035.74). The bottom status bar indicates "Ln 1, Col 1", "100%", "Unix (LF)", and "UTF-8". The "Unix (LF)" option is highlighted with a blue border.

S1	Yes	Yes	3,172.53
S1	No	Yes	983.21
S1	Yes	TGZGMOXG	Yes
S1	Yes	USUKVMZO	Yes
S1	No	DNUNTASS	Yes
S1	Yes	BRHUMGJR	Yes
S1	Yes	Yes	9,167.93
S1	Yes	Yes	746.70
S1	Yes	No	7,035.74
S1	Yes	No	

Para isso, vamos instalar o Notepad++, que é um Notepad turbinado. A instalação do notepad++ se deve ao poder de manipulação da formatação dos arquivos.

Veja que o Notepad++, nos permite visualizar com mais informações que o bloco de notas convencional. Podemos visualizar os espaços dados entre as informações.



The screenshot shows a Notepad++ window displaying a fixed-width text file named "06 FixedWidthExample.txt". The file contains a header with dates and times, followed by a table with columns for Entry, Per., Post Date, GL Account, Description, Srce, Cflow, Ref., Post, and Debit. The data rows show various transactions with their corresponding details and amounts. The bottom status bar indicates "Ln 1, Col 1", "100%", "Unix (LF)", and "UTF-8". The "Unix (LF)" option is highlighted with a blue border.

Entry	Per.	Post Date	GL Account	Description	Srce	Cflow	Ref.	Post	Debit
10	16524	01 10/17/2012	3930621977	TXNPUES	S1	Yes	RHMXWPCP	Yes	
11	191675	01 01/14/2013	2368183100	OUNHQEX XUFQONY	S1	No		Yes	
12	191667	01 01/14/2013	3714468136	GHAKASC QHJXDFM	S1	Yes		Yes	
13	191673	01 01/14/2013	2632703881	PAHFSAP LUVIKXZ	S1	No		Yes	983.21
14	80495	01 11/21/2012	2766389794	KDZANTV	S1	Yes	TGZGMOXG	Yes	
15	80507	01 11/21/2012	4609266335	BWWYEZL	S1	Yes	USUKVMZO	Yes	
16	80509	01 11/21/2012	1092717420	QJYPKVO	S1	No	DNUNTASS	Yes	
17	80497	01 11/21/2012	3386366766	SOQLCMU	S1	Yes	BRHUMGJR	Yes	
18	191669	01 01/14/2013	5905893739	FYIWNKA QUAFDKD	S1	Yes		Yes	9,167.93
19	191671	01 01/14/2013	2749355876	CBMJTLF NGFSEIS	S1	Yes		Yes	746.70
20	191674	01 01/14/2013	4530359106	OTAVZGH ZUQFISZ	S1	Yes		No	7,035.74
21	244819	01 02/04/2013	4679391677	EGHLQT1 ABE	S1	Yes		No	
22	96062	01 11/30/2012	5996493062	KTSVTADFF EHEHFMX	S1	Yes	UBNQLRCC	Yes	7.10
23	16527	01 10/17/2012	5595769375	ILCVJYC	S1	Yes	HCVZOUHY	Yes	
24	191670	01 01/14/2013	1948028853	RPPDCWC UWODNIO	S1	Yes		No	9,293.80
25	191672	01 01/14/2013	4938823703	CTMDXXP HXOXVFF	S1	Yes		No	175.00
26	191668	01 01/14/2013	4207018603	DBZZULF QGDZQMD	S1	Yes		Yes	206.26
27				ENDING BALANCE PERIOD 01					30,788.27

O texto foi criado em uma máquina LINUX, ou seja, os espaços não foram simplesmente ENTERS, e sim TABS ou ESPAÇOS com ENTER (esse é o formato para quebra de linha LF do Linux. E não adianta tentar transformar esse arquivo formato para o Windows (CRLF³), pois esse arquivo já foi criado como o (LF). Ou seja, ficar removendo vazio no PB desse arquivo, não adianta por causa da criação LF do Linux, se fosse criado pelo S.O do Windows, funcionava.

³ CRLF: Quebra de linhas do Windows que é através de espaços e gera nulo.

Filtro Condicional

Voltando ao PB, vamos fazer um filtro condicional. Veja que ao selecionarmos o filtro simples na coluna 1.1 que pode ser interpretada como “Número de pedido, de *nf*, ou do que for”, e perceba pela lógica que todos eles contêm números.

The screenshot shows the Power BI ribbon with the '06 FixedWidthExample' tab selected. In the 'Column1.1' dropdown, the 'Filtros de Texto' (Text Filters) option is expanded. Under 'Filtros de Texto', the 'Contém...' (Contains...) option is checked. A list of values is shown, all of which contain numbers: '(Selecionar Todo)', '(em branco)', '03/04/2', '16524', '16527', '68944', '80495', '80497', '80507', '80509', and '96062'. The 'Pesquisar' (Search) field is empty.

Então, cliquemos em “Filtro de Texto > Contém” e vamos filtrar o que queremos deixar visível.

The screenshot shows the Power BI ribbon with the '06 FixedWidthExample' tab selected. In the 'Column1.1' dropdown, the 'Filtros de Texto' (Text Filters) option is expanded. Under 'Filtros de Texto', the 'Contém...' (Contains...) option is checked. A dropdown menu is open, showing various filter criteria: 'Igual a...', 'Diferente de...', 'Começa com...', 'Não começa com...', 'Termina com...', 'Não Termina Com...', 'Contém...', and 'Não Contém...'. The 'Contém...' option is highlighted. The 'Pesquisar' (Search) field is empty.

Para esse exemplo, estamos usando uma lógica para filtrarmos o que desejamos.

Para outras planilhas, ou no dia-a-dia, deve ser usar a logica para atingir um resultado também.

Portanto, queremos que mantenha todos os números que estejam nessa coluna de 1 a 0 e OK.

Filtrar Linhas

Aplique uma ou mais condições de filtro às linhas nesta tabela.

Básico Avançadas

Manter linhas em que

E/Ou	Coluna	Operador	Valor
	Column1.1	contém	1
Ou	Column1.1	contém	2
Ou	Column1.1	contém	3
Ou	Column1.1	contém	4
Ou	Column1.1	contém	5
Ou	Column1.1	contém	6

[Adicionar Cláusula](#)

O filtro foi aplicado, permanecendo somente os valores com dados. Os vazios, da primeira coluna, não estão mais presentes, nesse filtro.

	Column1.1	Column1.2	Column1.3	Column1.4
1	03/04/20	13		
2	Period 0	1 Th	ru 03	
3	4:16 pm			
4	16524	01	10/17/2012	3930621977
5	191675	01	01/14/2013	2368183100
6	191667	01	01/14/2013	3714468136
7	191673	01	01/14/2013	2632703881
8	80495	01	11/21/2012	2766389794
9	80507	01	11/21/2012	4609266335
10	80509	01	11/21/2012	1092717420
11	80497	01	11/21/2012	3386366766
12	191669	01	01/14/2013	5905893739

Mas, perceba que ainda é necessário trimar essa coluna.

Perceba, também, que é necessário tirar as horas, as datas, os períodos que continuam aparecendo na coluna. E perceba, que esses dados são um padrão também. Nas demais linhas abaixo, eles se repetem, somente alterando as informações.

Agora, precisamos filtrar esses valores, removê-los da visualização. Precisamos encontrar algo em comum entre as linhas e as colunas, por exemplo, a data, período e hora, “compartilham” a coluna 1.4 como vazia.

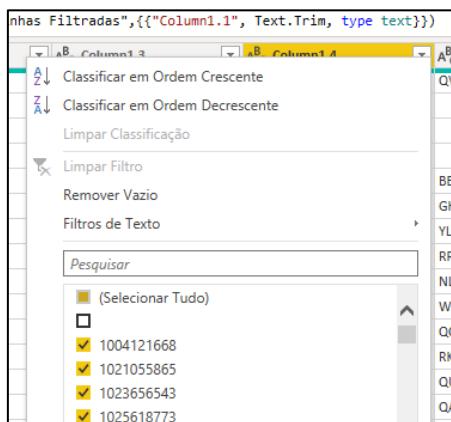
	A ^B _C Column1.1	A ^B _C Column1.2	A ^B _C Column1.3	A ^B _C Column1.4	A ^B _C
1	03/04/20	13			
2	Period 0	1 Th	ru 03		
3	4:16 pm				
4	16524	01	10/17/2012	3930621977	TXNR
5	191675	01	01/14/2013	2368183100	OUN

Veja abaixo, como os mesmos dados se repetem e na coluna 1.4 são valores vazios.

	A ^B _C Column1.1	A ^B _C Column1.2	A ^B _C Column1.3	A ^B _C Column1.4	A ^B _C Colu
47	201488	02	01/16/2013	4070447707	QVY WVI
48	03/04/2	013			
49	Period 0	1 Th	ru 03		
50	4:16 pm				
51	201492	02	01/16/2013	5901784648	BEM TNG
52	202906	02	01/17/2013	4705202419	GHC EZC

Filtro Simples

Portanto, vamos utilizar um filtro simples para isso. Clique na coluna 1.4 e desmarque a opção vazia. Tire o flag do quadro do filtro



Pronto! Aquele cabeçalho da data, hora e o período não estão mais na visualização.

	A ^B _C Column1.1	A ^B _C Column1.2	A ^B _C Column1.3	A ^B _C Column1.4	A ^B _C
1	16524	01	10/17/2012	3930621977	TXI
2	191675	01	01/14/2013	2368183100	OUE
3	191667	01	01/14/2013	3714468136	GH
4	191673	01	01/14/2013	2632703881	PA
5	80495	01	11/21/2012	2766389794	XD
6	80507	01	11/21/2012	4609266335	BW
7	80509	01	11/21/2012	1092717420	QVY
8	80497	01	11/21/2012	3386366766	SO
9	191669	01	01/14/2013	5905893739	FYI

Porém, precisamos arrumar e ajustar as posições.

Veja como ainda precisamos corrigir as posições que não foram devidamente identificadas, lá atrás, muitas páginas atrás. Veja como ainda tudo está bagunçado e desorganizado. Mesmo sem as linhas em branco e sem dados inúteis,

A ^B _C Column1.9	A ^B _C Column1.10	A ^B _C Column1.11
Yes		30,278.28 Yes
Yes 6	66,195.93	Yes
Yes 8	20,262.25	Yes
Yes	25,478.89	No
Yes 7	88,684.67	Yes
Yes	71,058.47	No
Yes	16,947.59	Yes
Yes	81.53	Yes
Yes	68,968.66	No
Yes 3	66,362.63	Yes
Yes 3	41,992.68	Yes
Yes	6,766.96	Yes
No	88.4	69,487.43 Yes

Isso não é sujeira, não é limpeza. É ajuste. Veja que o número “6” ao lado do Yes, **não** pertence a coluna 1.9 e sim a coluna 1.10, faz parte do número.

Dica: Evitar esse erro a qualquer custo. Pois a solução é ter que ficar procurando o erro na Linguagem M (Na barra de fórmulas). E isso gera muito trabalho, porque precisa ficar encontrando e indo a última etapa e verificar se a coluna foi espaçada corretamente.

Uma observação importante também, é que após dividir todas as colunas pelas posições corretas, lembre-se de usar o Trim para trimar os dados.

A ^B _C Column1.10	A ^B _C Column1.11
	5,007.10
	43,537.00
3,172.53	
983.21	
	903.78
	670.31
	848.50
	7.31
9,167.93	
746.70	

A ^B _C Column1.10	A ^B _C Column1.11
	5,007.10
	43,537.00
3,172.53	
983.21	
	903.78
	670.31
	848.50
	7.31
9,167.93	
746.70	

Múltiplos Arquivos – Arquivos Recorrentes

Quando estamos no escritório, podemos trabalhar com planilhas que são diárias, muitas vezes, que são enviadas de hora em hora, e para evitarmos o retrabalho no tratamento, sempre que uma planilha nova é recebida, o **Power BI** fornece uma conexão especial quando isso acontece.

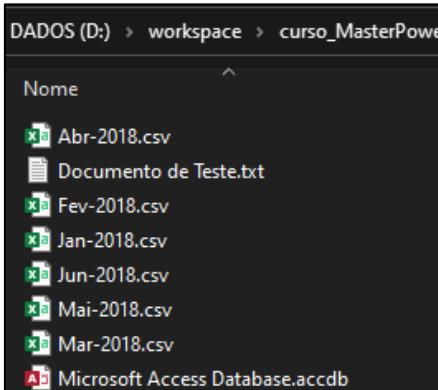
Vamos explicar melhor.

Suponha que hoje, dia 16/03/2021 recebemos uma planilha chamada "vendas160321.xlsx", nós a trataremos, limparemos e toda a parte de tipagem e gráfica. No dia seguinte, recebemos a mesma planilha, porém com dados atualizados e novos, sendo que essa planilha será a "vendas170321.xlsx". Para que não façamos o tratamento diversas vezes, o PB fará uma conexão especial.

Outro exemplo, que esse será o exemplo prático desse capítulo.

Nome	Data de modificação	Tipo
07 - Semestre 01	23/02/2021 02:47	Pasta de arquivos
07 - Semestre 02	23/02/2021 02:47	Pasta de arquivos

Nisso, selecionaremos o Semestre 01. Veja que a planilha é enviada mensalmente.



Entretanto, no **Power BI**, vamos abrir a Pasta ao Obtermos os Dados.

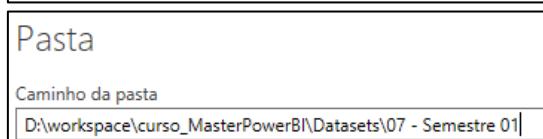
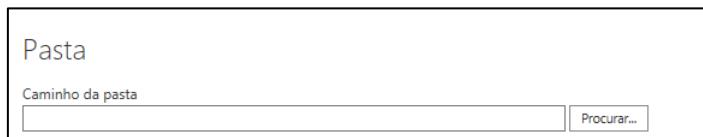
Abrindo Pastas na Obtenção de Dados

Nos exemplos anteriores, a obtenção de dados era em Excel, após tratamos .csv e extensões .txt. Nesse momento, vamos trabalhar com pastas.

The screenshot shows the 'Obter Dados' (Get Data) interface in Power BI. On the left, there is a sidebar with options: 'Pesquisar' (Search), 'Tudo' (All), 'Arquivo' (File), 'Banco de Dados' (Database), 'Power Platform', and 'Azure'. On the right, there is a main panel titled 'Tudo' with icons for 'Excel', 'Texto/CSV', 'XML', 'JSON', and 'Pasta' (Folder). The 'Pasta' option is highlighted with a gray bar at the bottom.

A conexão será na pasta.

Desse modo, não importa quais serão os documentos contidos na pasta. Pois carregaremos a pasta e todos os arquivos presentes nela. Clicar em “conectar”, para que o PB selecione o caminho da rede e qual pasta será selecionada.

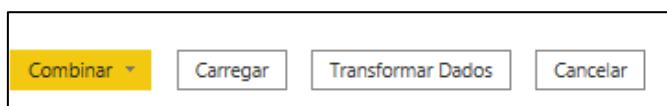


Após o PB encontrar o caminho na rede, ele apresentará quais são os documentos presentes:

D:\workspace\curso_MasterPowerBI\Datasets\07 - Semestre 01								
Content	Name	Extension	Date accessed	Date modified	Date created	Attributes		
Binary	Abr-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:45:02	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Documento de Teste.txt	.txt	08/03/2021 11:18:45	24/10/2019 06:58:20	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Fev-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:25:06	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Jan-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:08:02	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Jun-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:44:12	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Mai-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:44:20	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Mar-2018.csv	.csv	08/03/2021 11:18:45	23/10/2019 20:58:28	23/02/2021 02:47:18	Record	D:\workspace\cu	
Binary	Microsoft Access Database.accdb	.accdb	08/03/2021 11:18:45	23/10/2019 21:16:46	23/02/2021 02:47:18	Record	D:\workspace\cu	

Transformação de Dados das Pastas

Veja que existem as planilhas de Abril, um documento de texto(.txt), uma série de planilhas mensais(em .csv) e uma Database(.accdb). E as seguintes opções abaixo:



Como existe “lixo” dentro da pasta, como o arquivo .txt e o .accdb, não podemos nem “Combinar” e nem “Carregar”. Por quê? Porque o Power BI tentará combinar todos os dados presentes na pasta e não é o que deve ocorrer. Portanto, quando houver arquivos que não são do suporte padrão, precisaremos “Transformar os Dados”.

Quando carregamos dessa forma, o Power Query carrega os arquivos de forma binária. Ou seja, ele está realizando alterações na características dos arquivos e não nos dados.

Content	Name	Extension
Binary	Abr-2018.csv	.csv
Binary	Documento de Teste.txt	.txt
Binary	Fev-2018.csv	.csv
Binary	Jan-2018.csv	.csv

Filtro de Texto

Porém, o programa não é mágico para identificar sem nenhuma interação do usuário em definir quais são os arquivos a serem carregados. Veja que existe um padrão sobre as planilhas a serem filtradas, todas elas começam com o nome do mês. Ou seja, podemos realizar um “Filtro de Texto” para filtrarmos somente os arquivos de Excel e não os de **.txt** e **.db**. Sempre que qualquer arquivo for alocado e inserido nessa pasta, e estiver nessas condições do filtro, ele será carregado automaticamente.

Clique na seta para baixo ↓ e filtre os valores que “comecem com...”

Content	Name	Extension
Binary	Abr-2018.csv	.CSV
Binary	Documento de Teste.txt	.txt
Binary	Fev-2018.csv	.CSV
Binary	Jan-2018.csv	.CSV

O resultado de uma possível solução.

Content	Name	Extension
Binary	Abr-2018.csv	.CSV
Binary	Fev-2018.csv	.CSV
Binary	Jan-2018.csv	.CSV
Binary	Jun-2018.csv	.CSV
Binary	Mai-2018.csv	.CSV
Binary	Mar-2018.csv	.CSV

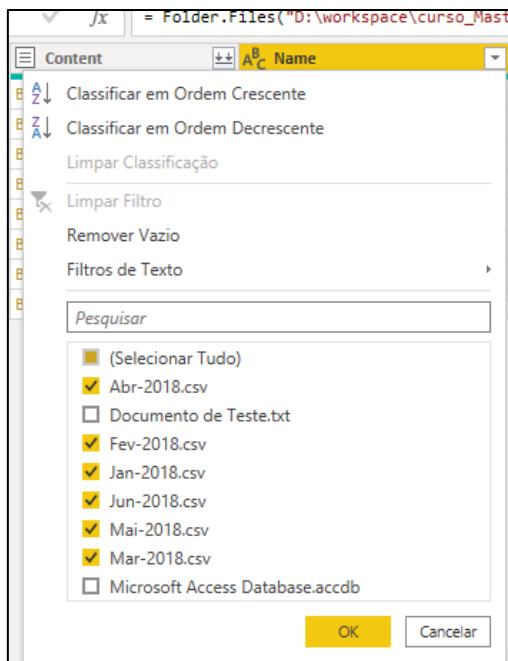
Outra possível solução também seria o filtro por extensão (seja por **.xlsx** ou **.csv**)

The screenshot shows a Windows File Explorer window with several CSV files listed. A context menu is open, and the 'Filtrar Linhas' (Filter Rows) option is selected. A modal dialog box titled 'Filtrar Linhas' appears, prompting the user to apply filters to the rows. The 'Básico' (Basic) radio button is selected. Under 'Manter linhas em que 'Name'' (Keep rows where 'Name'), there are two dropdown menus: the first shows 'termina com' (ends with) and '.csv' selected; the second shows 'E' (And) and 'Ou' (Or), with another dropdown menu below it showing 'termina com' and '.xlsx' selected. This configuration will filter the list to show only files ending in '.csv' and '.xlsx'.

Um ponto de atenção abaixo, pois deveremos realizar esse tipo de filtro quando trabalharmos com planilhas que serão diárias, semanais, mensais ou conforme for, porque é necessário sempre adicioná-las na pasta do caminho.

Cuidado para não realizar apenas o Filtro, porque assim, se novas planilhas forem adicionadas, elas não serão tratadas, por isso precisa realizar esse tipo de Filtro Simples.

O filtro aplicado abaixo mostrará apenas as planilhas já selecionadas nesse filtro, por exemplo, se uma planilha de “Dez-18” for incluída na pasta, ela não será mostrada se utilizarmos o filtro abaixo. Da forma abaixo, o filtro seria incompleto, qualquer outro arquivo com a extensão seria adicionado.



Próximo passo é realizar a limpeza de colunas desnecessárias.

Removendo Colunas das Pastas

Veja que existem colunas que são tratadas como “lixo” e podem ser descartadas.

Date accessed	Date modified	Date created	Attributes	Folder Path
08/03/2021 11:18:45	23/10/2019 20:45:02	23/02/2021 02:47:18	Record	D:\workspace\curso_MasterPowerBI\
08/03/2021 11:18:45	24/10/2019 06:58:20	23/02/2021 02:47:18	Record	D:\workspace\curso_MasterPowerBI\
08/03/2021 11:18:45	23/10/2019 20:25:06	23/02/2021 02:47:18	Record	D:\workspace\curso_MasterPowerBI\
08/03/2021 11:18:45	23/10/2019 20:08:02	23/02/2021 02:47:18	Record	D:\workspace\curso_MasterPowerBI\
08/03/2021 11:18:45	23/10/2019 20:44:12	23/02/2021 02:47:18	Record	D:\workspace\curso_MasterPowerBI\

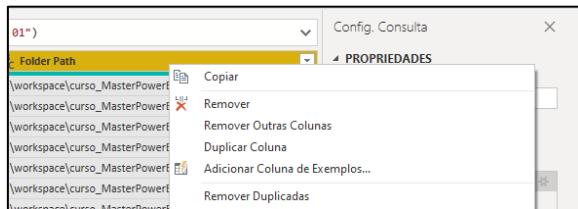
Apenas uma observação antes de seguirmos com o exemplo, existem alguns arquivos que não possuem data dentro da planilha, não é o caso desses, pois eles contêm datas.

Nota, Lucro, Unidades, Data
38894,12276.29,921,4/1/2018 0:00
38101,14071.96,936,4/2/2018 0:00

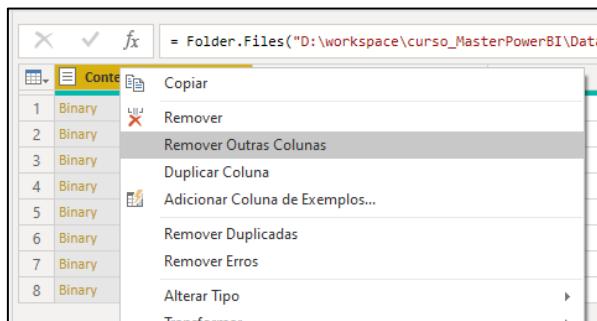
Caso os arquivos que serão carregados não contenham datas, podemos utilizar algumas das colunas que ficaram, como a “Data Accessed, Modified ou Created”.

Porém, nesses casos, podemos excluir todas que não são necessárias, pois os arquivos já possuem datas em seus dados.

Poderíamos excluir coluna por coluna através do botão direito na coluna a ser removida e Remover.



Porém, como precisamos apenas da primeira coluna, que são os arquivos binários, vamos na primeira coluna e com o **“Botão Direito > Remover Outras Colunas”**.



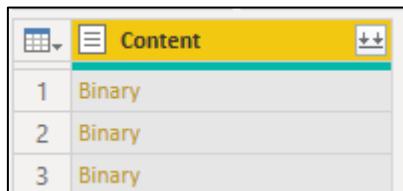
E apenas trabalharemos nessa coluna agora.

Somente após realizar todos os outros filtros anteriores que podemos excluir as colunas e mantermos somente o arquivo binário. Isso é muito importante!

Combinando Arquivos Binários

O próximo passo agora é combinar os arquivos.

Clique nas setas duplas e combine os arquivos.



O PB irá combinar, como é um arquivo **.csv** e correto uso da vírgula, ele será carregado sem nenhum problema (veja que o arquivo está perfeitamente qualificado).

Combinar Arquivos

Especifique as configurações para cada arquivo. [Saiba mais](#)

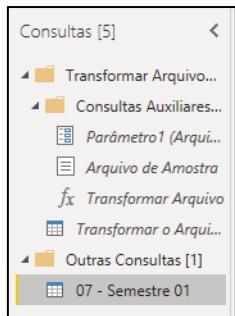
Arquivo de Amostra:

Primeiro arquivo

Origem do Arquivo		Delimitador	Detecção de Tipo de Dados
65001: Unicode (UTF-8)		Vírgula	Com base nas primeiras 200 linhas
Nota	Lucro	Unidades	Data
38894	12276.29	921	4/1/2018 0:00
38101	14071.96	936	4/2/2018 0:00
37388	13457.3	882	4/3/2018 0:00

Agora o **Power Query** nos apresentará mais passos, mais consultas a serem realizadas.

No canto esquerdo do **Power Query** podemos visualizar essas “Consultas Auxiliares”.



Muito cuidado nesse próximo passo. É necessário clicar no último arquivo da consulta, ou seja, no quadro acima, no “07 – Semestre 01”. E todas as futuras alterações serão realizadas em todos os demais arquivos contidos na pasta que foram filtrados.

Obs: [Não siga esse passo nos trabalhos do dia-a-dia, utilize apenas nessa seção, pois no próximo Capítulo – “Modelagem dos Dados – Arquivo 08”, utilizaremos a consulta de “Transformar O Arquivo Exemplo”, pois é a forma correta.]

Por exemplo, perceba que na coluna “Data” o tipo é “Texto” e vamos *tipar* como “Data e Hora”.

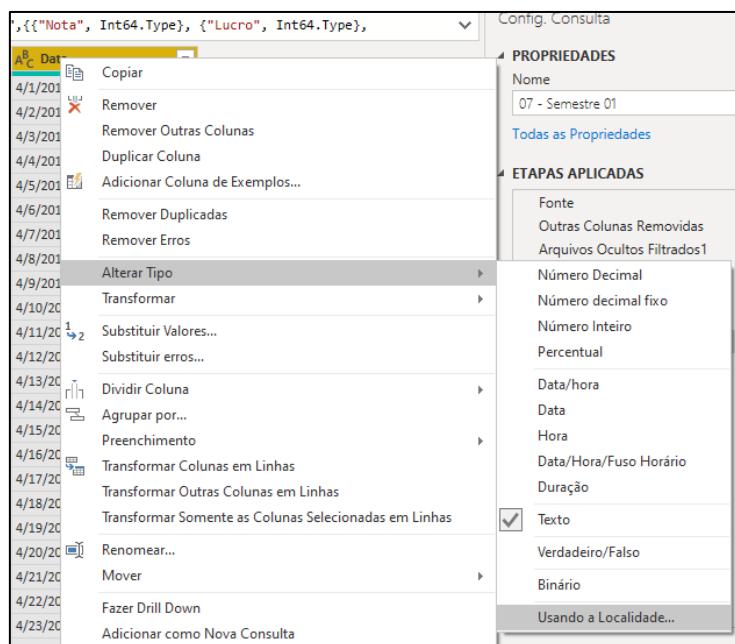
	1	2	3	Unidades	A	B	C	Data
9				921	4/1/2018	0:00		
6				936	4/2/2018	0:00		
3				882	4/3/2018	0:00		
8				909	4/4/2018	0:00		

	Data
1	04/01/2018 00:00:00
5	04/02/2018 00:00:00
2	04/03/2018 00:00:00
9	04/04/2018 00:00:00
5	04/05/2018 00:00:00
7	04/06/2018 00:00:00
2	04/07/2018 00:00:00
2	04/08/2018 00:00:00
7	04/09/2018 00:00:00
5	04/10/2018 00:00:00
3	04/11/2018 00:00:00
7	04/12/2018 00:00:00
7	Error
1	Error

Alterando Tipo pela Localidade

Algumas colunas deram “**Error**”, ou seja, essa não foi a correta tipagem dos dados. Portanto, vamos apagar essa etapa alterada e tentar identificar o que aconteceu.

O erro que ocorreu, se deve ao fato que o PB tenta converter essa data em nosso formato brasileiro de data. Porém, a data foi criada no formato americano de data. Portanto, vamos em: “**Botão Direito do Mouse na Coluna > Alterar Tipo > Usando Localidade**”.



Procure pela localidade que é a origem das datas dos dados, ou seja, os Estados Unidos.

Precisamos selecionar qual será o tipo de dado que precisaremos converter, ou seja, antes estava como “Text” e agora será “**Data/Hora**”, seguindo o padrão de EUA.

The screenshot shows a modal dialog titled "Alterar Tipo com Localidade". It has two dropdown menus: "Tipo de Dados" set to "Data/hora" and "Localidade" set to "Inglês (Estados Unidos)". Below the dropdowns, there's a section titled "Exemplos de valores de entrada:" with four examples listed:

- Tuesday, March 29, 2016 2:45 PM
- Tuesday, March 29, 2016 2:45:33 PM
- 3/29/2016 2:45 PM
- 3/29/2016 2:45:33 PM

Veja que nenhum erro mais será encontrando, porque agora estará conforme.

Por que está, então, na nossa data agora? Porque nós informamos que essa origem de dados é Americana. O Power BI converterá as datas para a nossa língua, para a data que está ajustada em nosso sistema operacional, que é o padrão de “dd/mm/aaaa hh:mm:ss”

	Data
1	01/04/2018 00:00:00
5	02/04/2018 00:00:00
2	03/04/2018 00:00:00
9	04/04/2018 00:00:00
5	05/04/2018 00:00:00
7	06/04/2018 00:00:00
2	07/04/2018 00:00:00
2	08/04/2018 00:00:00
7	09/04/2018 00:00:00
5	10/04/2018 00:00:00
3	11/04/2018 00:00:00
7	12/04/2018 00:00:00
1	13/04/2018 00:00:00
1	14/04/2018 00:00:00
2	15/04/2018 00:00:00
3	16/04/2018 00:00:00
5	17/04/2018 00:00:00
4	18/04/2018 00:00:00

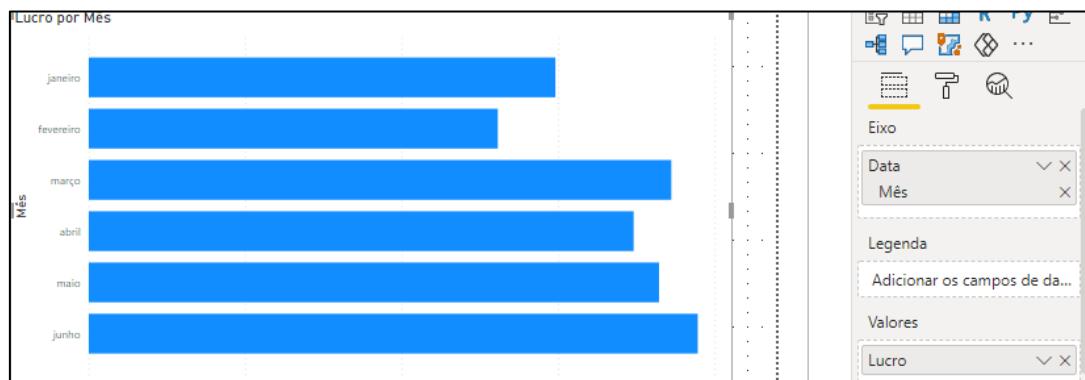
Em qualquer erro de data, verifique a formatação da localidade da data dos dados x a data do sistema operacional.

Vamos fechar e aplicar as modificações do Power Query e vamos ao Power BI.

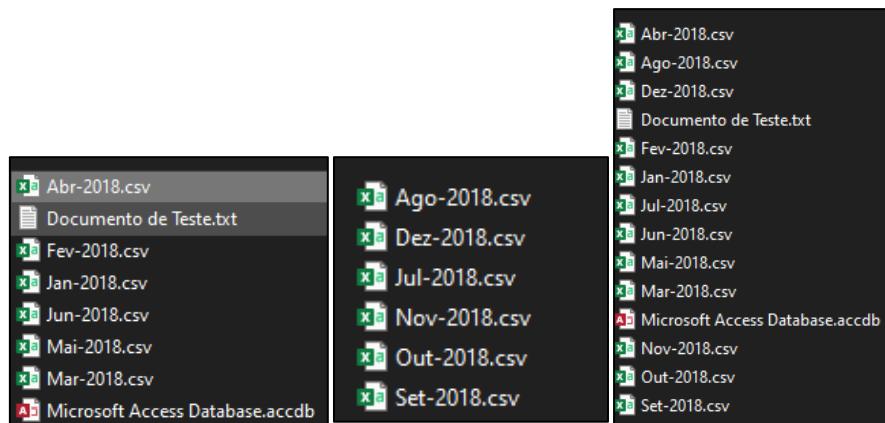
Os campos foram criados

The screenshot shows the 'Campos' (Fields) section of the Power Query editor. At the top, there is a search bar labeled 'Pesquisar'. Below it, a tree view shows a folder named '07 - Semestre 01' which contains four fields: 'Data', 'Σ Lucro', 'Σ Nota', and 'Σ Unidades'. The 'Data' field is expanded, showing its current type as a date.

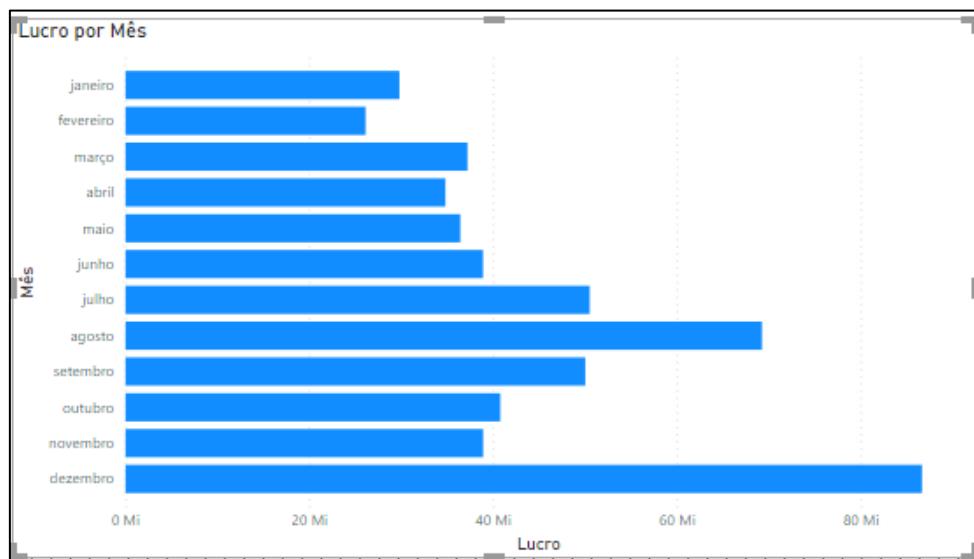
Vamos adicionar um gráfico de barra para podermos seguir com o próximo exercício.



Ótimo, agora lembra que foram criadas duas pastas? Uma que era o primeiro semestre e outra que continha o segundo semestre? Pois bem, vamos pegar os documentos que estão na pasta do segundo semestre e vamos extraí-los e alocá-los na pasta do primeiro semestre.



A primeira é o 1º semestre antes, no meio a do 2º semestre e a 3 imagem é o 1º semestre com os arquivos do 2º semestre. Após isso, basta Atualizar o **Power BI** que ele conterá todos os dados das pastas. Todos os demais agora aparecerão.



Modelagem dos Dados

Arquivo 08 Movimentos com colunas mescladas

Para esse tópico, utilizaremos principalmente as planilhas contidas na pasta:

D:\workspace\curso_MasterPowerBI\Datasets\08 – Movimento, veja um exemplo.

Localidade		Modalidade	Semestre 01				
Região	Estado	Cidade	Fato	Janeiro 2017	Fevereiro 2017	Março 2017	Abril 2017
Sudeste	Rio de Janeiro	Rio de Janeiro	Devoluções	1345	1456	5335	2143
		Rio de Janeiro	Vendas	R\$ 345.645,00	R\$ 46.575,00	R\$ 756.876,00	R\$ 67.867,00
	São Paulo	Devoluções		3543	5464	4565	5677
	São Paulo	Vendas		R\$ 43.565,00	R\$ 65.774,00	R\$ 54.643,00	R\$ 25.456,00

Essa planilha contém, células mescladas, colunas, linhas, e principalmente, ela “cresce” para o lado, os valores de dados estão “Indo para a direita”.

Vamos começar obtendo os dados para o **Power BI**. E vamos selecionar toda a pasta.

Porém, uma coisa interessante vai acontecer, veja que na pasta existem apenas duas planilhas. Enquanto no **Power BI**, foram encontradas três.

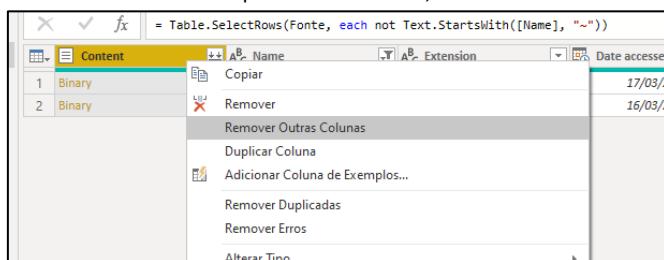
The screenshot shows the Power BI Data view on the left, displaying two files: "Movimento 2017.xlsx" and "Movimento 2018.xlsx". To the right, a file browser window shows three files: "Content", "Name", and "Extension". The "Name" column lists "Movimento 2017.xlsx", "Movimento 2018.xlsx", and "~\$Movimento 2017.xlsx".

Isso acontece porque a pasta está aberta no notebook. E ela é identificada como uma “cópia/arquivo temporário” para o Office e o **Power BI** reconhece como ativa. Identificamos através do “~” na nomenclatura do arquivo.

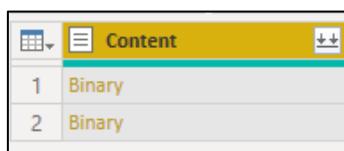
Para evitarmos que esses erros aconteçam, utilizaremos o Filtro de Texto para deixarmos somente os arquivos que **não começam** com o “~”.

The screenshot shows the Power BI Filter dialog. The "Name" column is selected, and a dropdown menu is open, showing filter options. The "Não começa com..." option is selected. Below the dropdown, the filtered results show only "Movimento 2017.xlsx" and "Movimento 2018.xlsx".

Agora é preciso remover todas as outras colunas que estão à direita porque somente utilizaremos os arquivos Binários, como nos outros exemplos anteriores.



E lembre-se de clicar nas setas ao lado do “Content” para combinar os arquivos!

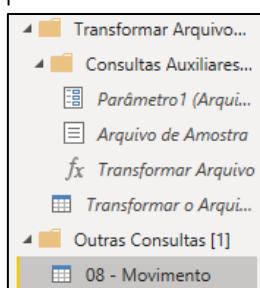


Transformar o Arquivo Exemplo

Algo diferente do que fora feito no exercício anteriores, vamos utilizar um método diferente. Veja que existem dois semestre, sendo o de 2017 e 2018.

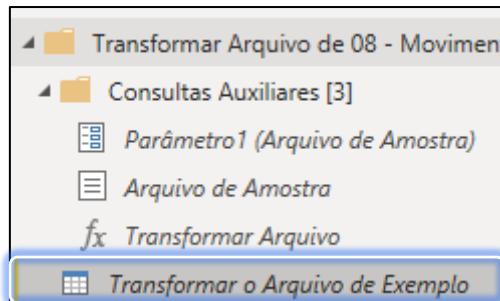
Localidade	Column2	Column3	Modalidade	Semestre 01	Semestre 02
Região	Estado	Cidade	Fato	Janeiro 2017	Fev
Sudeste	Rio de Janeiro	Rio de Janeiro	Devolução	1345	
	null	null	Vendas	345645	
	São Paulo	São Paulo	Devolução	3543	
	null	São Paulo	Vendas	43565	
	null	Piracicaba	Devolução	3566	
	null	Piracicaba	Vendas	43001	
	null	Ribeirão Preto	Devolução	2999	
	null	Ribeirão Preto	Vendas	43801	
Sul	Rio Grande do Sul	Poa	Devolução	1579	
	null	null	Vendas	344445	
	null	Porto Alegre	Devolução	1976	
	null	Porto Alegre	Vendas	42345	
	null	Santa Catarina	Devolução	3230	
	null	null	Vendas	343291	
	null	Florianópolis	Devolução	3107	
	null	null	Vendas	218291	
Total Faturado		null	null	null	1424384
Região	Estado	Cidade	Fato	Janeiro 2018	Fev
Sudeste	Rio de Janeiro	Rio de Janeiro	Devolução	1345	

Se seguíssemos com o mesmo método do exercício anterior, nós trabalharíamos nessa parte da Consultas. Porém, como foi dito, essa não é a forma correta!



E todas as alterações seriam feitas uma a uma. Teríamos que eliminar a linha de 2017, outra de 2018. Parece que é fácil, sim, pois temos somente 2 anos. E quando trabalhamos com mais de 10 mil linhas e essas mudanças se tornarão enviáveis.

Portanto, vamos utilizar o “[Transformar o Arquivo Exemplo](#)”.



Tudo que for feito no exemplo impactará todos os demais arquivos que vierem em seguida.

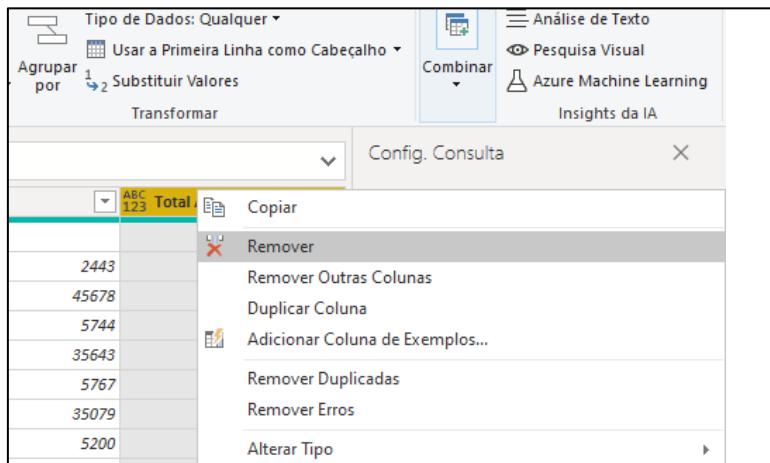
ABC 123 Localidade	ABC 123 Column2	ABC 123 Column3	ABC 123 Modalidade	ABC 123 Semestre 01
Região	Estado	Cidade	Fato	Janeiro 2017
Sudeste	Rio de Janeiro	Rio de Janeiro	Devoluçãoes	
	null	null	Vendas	345
	null	São Paulo	Devoluçãoes	3
	null	null	Vendas	43
	null	null	Devoluçãoes	3
	null	null	Vendas	43
	null	null	Devoluçãoes	2
	null	null	Vendas	43
Sul	Rio Grande do Sul	Poa	Devoluçãoes	1
	null	null	Vendas	344
	null	null	Devoluçãoes	1
	null	null	Vendas	42
	null	Santa Catarina	Devoluçãoes	3
	null	null	Vendas	343
	null	null	Devoluçãoes	3
	null	null	Vendas	218
Total Faturado		null	null	null
				1424

Veja que se manteve apenas o de 2017.

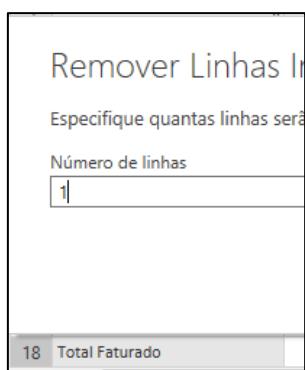
Alguns pontos para limpeza desse arquivo:

- Não precisamos deixar “Total Faturado”, não se trabalha com valores agregados, o PB se encarregará de realizar a somatória.

- Lembre-se de puxar a barra de rolagem para o lado, encontrará um “Total Anual”, também deve ser eliminado.



- Precisaremos eliminar a última linha que contém o Total Faturado.



Descartaremos somente a última linha inferior.

Todos os outros arquivos sofrerão essas mudanças pois todos os arquivos têm a mesma estrutura!

Usando Primeira Linha como Cabeçalho

Precisamos manter a primeira linha como o cabeçalho do arquivo, pois é necessário.

Veja que pela questão de células mescladas no Excel, temos o problema que em algumas linhas existe um “cabeçalho”, porém, em outras, uma “Column” sem definição. As células mescladas são um “probleminha” a serem corrigidos.

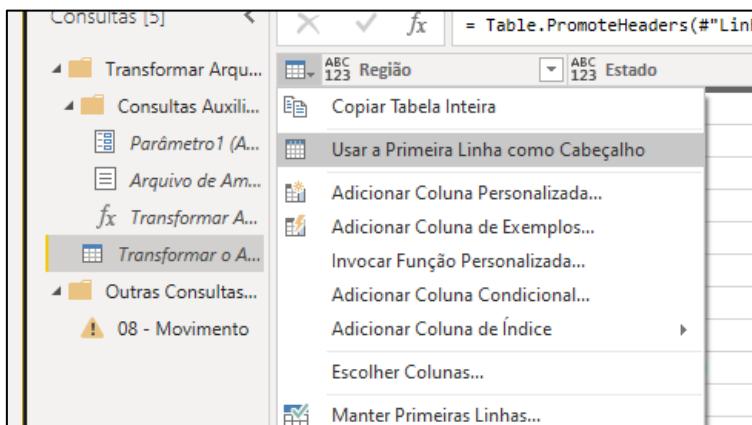
Excel:

Localidade			Modalidade	Semestre 01				
Região	Estado	Cidade	Fato	Janeiro 2017	Fevereiro 2017	Março 2017	Abril 2017	Mai
	Rio de Janeiro	Rio de Janeiro	Devoluçãoes	1345	1456	5335	2143	
		Rio de Janeiro	Vendas	R\$ 245.645,00	R\$ 16.575,00	R\$ 756.276,00	R\$ 67.967,00	R\$ 6

Power BI:

ABC 123 Localidade	ABC 123 Column2	ABC 123 Column3	ABC 123 Modalidade	ABC 123 Semestre 01	ABC 123 Column6
Região	Estado	Cidade	Fato	Janeiro 2017	Fevereiro 2017
Sudeste	Rio de Janeiro	Rio de Janeiro	Devoluçãoes		1345

Para transformar a “Primeira Linha como Cabeçalho”, clique no botão ao lado de Região, que representa a Tabela, e selecione a opção.



Observe o resultado de antes, como o após:

ABC Região	ABC Estado	ABC Cidade	ABC Fato	123 Janeiro 2017	123 Fevereiro 2017
1 Sudeste	Rio de Janeiro	Rio de Janeiro	Devoluçãoes		
	null	null	Vendas	1345	345645

Repare, que após definirmos o cabeçalho, no Excel, a célula mesclada do Sudeste significa que é “Sudeste” até chegar em “Sul”.

Localidade		
Região	Estado	Cidade
Sudeste	Rio de Janeiro	Rio de Janeiro
		Rio de Janeiro
	São Paulo	São Paulo
		São Paulo
		Piracicaba
		Piracicaba
		Ribeirão Preto
		Ribeirão Preto

No Power BI, entretanto, ele trata esses valores, como “Nulo”. Pois ele não identifica valores com a célula mesclada.

	A ^B C Região	A ^B C Estado	A ^B C Cidade
1	Sudeste	Rio de Janeiro	Rio de Janeiro
2	null	null	Rio de Janeiro
3	null	São Paulo	São Paulo
4	null	null	São Paulo
5	null	null	Piracicaba
6	null	null	Piracicaba
7	null	null	Ribeirão Preto
8	null	null	Ribeirão Preto

Como fazer que o “Sudeste” seja escrito todo para baixo, até que chegue em “Sul” e depois o “Sul” seja escrito para baixo, sem um sobreescriver o outro?

Preenchimento de Linhas

Existe uma função chamada “Preenchimento para baixo” que se encarregará de preencher as informações.

The screenshot shows the Power BI Data Editor interface. In the top ribbon, the 'Transform' tab is selected. On the left, there's a table with columns 'Região' and 'Fato'. The 'Região' column has rows for 'Sudeste' (row 1) and 'Sul' (row 9). A context menu is open over the 'Região' column at row 11, with 'Preenchimento' selected. Under 'Preenchimento', 'Para Baixo' is highlighted. The 'Fato' column contains data for 'Devoluções' and 'Vendas' repeated across the rows.

Veja o resultado, é como se fosse o “Ctrl + D” do Excel.

	A ^B _C Região
1	Sudeste
2	Sudeste
3	Sudeste
4	Sudeste
5	Sudeste
6	Sudeste
7	Sudeste
8	Sudeste
9	Sul
10	Sul
11	Sul
12	Sul
13	Sul
14	Sul
15	Sul
16	Sul

Também agimos com a coluna ao lado, pois no Excel, também existiam colunas mescladas. Agora está corretamente aplicado.

	A ^B _C Região	A ^B _C Estado	A ^B _C Cidade
1	Sudeste	Rio de Janeiro	Rio de Janeiro
2	Sudeste	Rio de Janeiro	Rio de Janeiro
3	Sudeste	São Paulo	São Paulo
4	Sudeste	São Paulo	São Paulo
5	Sudeste	São Paulo	Piracicaba
6	Sudeste	São Paulo	Piracicaba
7	Sudeste	São Paulo	Ribeirão Preto
8	Sudeste	São Paulo	Ribeirão Preto
9	Sul	Rio Grande do Sul	Poa
10	Sul	Rio Grande do Sul	Poa

Agora enfrentaremos um grande desafio. Veja que a planilha cresce para o lado. Ou seja, por exemplo, em Janeiro de 2017, houve devoluções de 1.345 e Vendas de R\$ 345.645,00. Correto?

Modalidade	Semestre 01				
	Fato	Janeiro 2017	Fevereiro 2017	Março 2017	Abril 2017
Devoluçãoẽ		1345	1456	5335	2143
Vendas	R\$ 345.645,00	R\$ 46.575,00	R\$ 756.876,00	R\$ 67.867,00	
Devoluçãoẽ		3543	5464	4565	5677
Vendas	R\$ 43.565,00	R\$ 65.774,00	R\$ 54.643,00	R\$ 25.456,00	
Devoluçãoẽ		3566	5487	4588	5700
Vendas	R\$ 43.001,00	R\$ 65.210,00	R\$ 54.079,00	R\$ 24.892,00	
Devoluçãoẽ		2999	4920	4021	5133
Vendas	R\$ 43.801,00	R\$ 66.010,00	R\$ 54.879,00	R\$ 25.692,00	
Devoluçãoẽ		1579	1690	5569	2377
Vendas	R\$ 344.445,00	R\$ 45.375,00	R\$ 755.676,00	R\$ 66.667,00	
Devoluçãoẽ		1976	3897	2998	4110
Vendas	R\$ 42.345,00	R\$ 64.554,00	R\$ 53.423,00	R\$ 24.236,00	
Devoluçãoẽ		3230	5151	4252	5364
Vendas	R\$ 343.291,00	R\$ 44.221,00	R\$ 754.522,00	R\$ 65.513,00	
Devoluçãoẽ		3107	5028	4129	5241
Vendas	R\$ 218.291,00	R\$ 31.721,00	R\$ 629.522,00	R\$ 53.013,00	
	R\$ 1.424.384,00	R\$ 429.440,00	R\$ 3.113.620,00	R\$ 353.336,00	

Enquanto o **Power BI**, também se mantém.

A ^B _C Fato	1 ² ₃ Janeiro 2017	1 ² ₃ Fevereiro 2017
Devolução	1345	1456
Vendas	345645	46575
Devolução	3543	5464
Vendas	43565	65774

Transformando Colunas em Linhas

Sabe-se que na verdade, os valores de 1345 e 345645 não são “Janeiro” e nem 1456 e 46575 não são “Fevereiro”, na verdade, são os valores dos meses.

A ^B _C Fato	1 ² ₃ Janeiro 2017	1 ² ₃ Fevereiro 2017
Devolução	1345	1456
Vendas	345645	46575
Devolução	3543	5464
Vendas	43565	65774

O **Power BI** não comprehende que Janeiro 2017 houve 1345 Devoluções, pois existem valores em Linhas e Colunas, dessa forma causa a informação cruzada.

Precisaremos, portanto, trabalhar passando a Coluna de Janeiro como se fossem linhas.

Por exemplo, a modelagem seria dessa forma, crescendo para baixo.

Devolução	Janeiro	1345
Vendas	Janeiro	345645
Devolução	Fevereiro	3543
Vendas	Fevereiro	46575

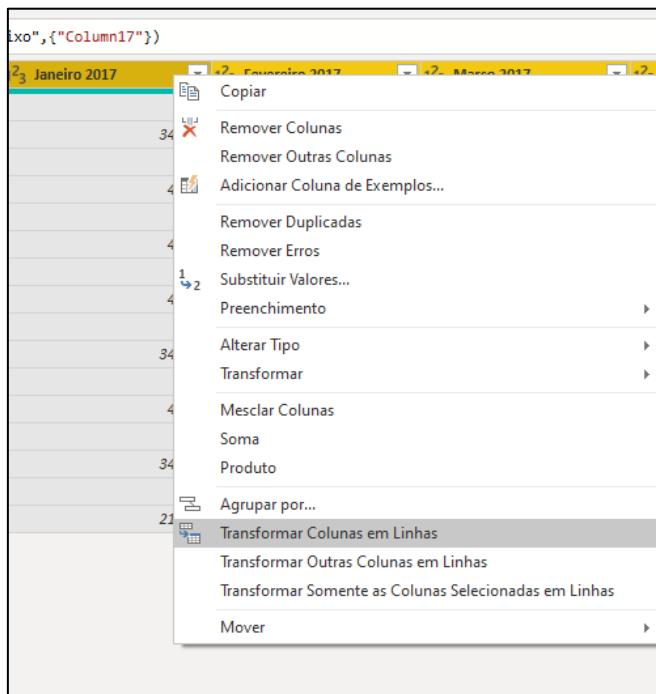
Todos os meses que seguirem, não crescerão para o lado, e sim para baixo. A planilha pode crescer para o lado, desde que seja um atributo novo, ou seja, algum novo que não seja um mês.

Clique em “Janeiro” e pressionando a tecla “SHIFT”, vá e clique no último mês ao lado, em “Dezembro”.

1 ² ₃ Janeiro 2017	Novembro 2017	1 ² ₃ Dezembro 2017
1345	2576	244
345645	87978	4567
3543	5467	574
	78876	3564

Todo o intervalo entre Janeiro a Dezembro será selecionado dessa forma.

Clique logo na primeira coluna em Janeiro, e com o botão direito, procure pela opção de “**Botão Direito > Transformar Colunas em Linhas**”.



Essa ação transformará todos os valores para baixo.

A ^B _C Fato	A ^B _C Atributo	1 ² ₃ Valor
Devolução	Janeiro 2017	1345
Devolução	Fevereiro 2017	1456
Devolução	Março 2017	5335
Devolução	Abril 2017	2143
Devolução	Maio 2017	1432
Devolução	Junho 2017	1564
Devolução	Julho 2017	1223
Devolução	Agosto 2017	1446
Devolução	Setembro 2017	6433
Devolução	Outubro 2017	5463
Devolução	Novembro 2017	2576
Devolução	Dezembro 2017	2443
Vendas	Janeiro 2017	345645
Vendas	Fevereiro 2017	46575
Vendas	Março 2017	756876
Vendas	Abril 2017	67867
Vendas	Maio 2017	678978
Vendas	Junho 2017	987978
Vendas	Julho 2017	909898
Vendas	Agosto 2017	98098

Porém, veja que no exemplo acima, existem muitas informações repetidas, por exemplo, diversas “Vendas”, diversos “Janeiros, Fevereiros, Marços”... Isso impactará na performance. Isso será visto mais à frente.

Porém, precisamos notar que existe um problema muito maior que a “performance” nesse momento.

Confusão de Fatos

O **Power BI** está cruzando os FATOS, ou seja, existem Devoluções em Janeiro, Fevereiro e outros meses. E, também existem Vendas em Janeiro, Fevereiro e outros meses. E nesse caso, esses dois valores estão coexistindo na mesma coluna. Isso é um problema, pois em Devoluções são quantidade e Vendas são valores decimais, em dinheiro. Isso precisa ser corrigido.

Transformando Linhas em Colunas

Para corrigir isso, vamos trabalhar exatamente ao contrário do anterior, vamos fazer Linhas virarem Colunas.

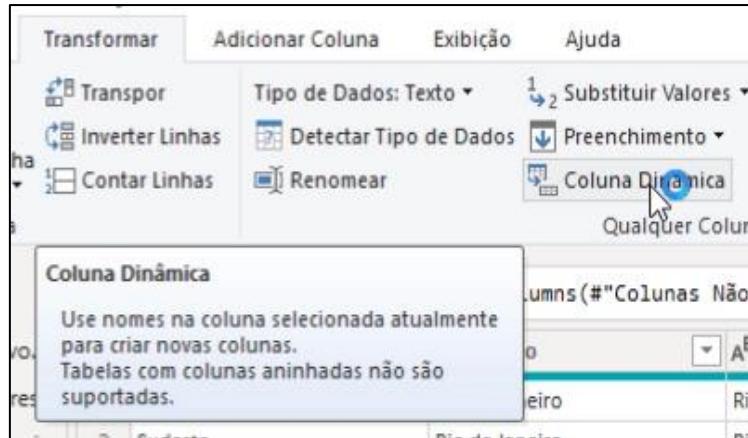
Selecione a coluna que deseja transforma, nesse exemplo, é a coluna FATO.

A	B	C	Fato
			Devolução
			Venda
			Venda

Coluna Dinâmica

O Power BI não apresenta a opção “Transformar Linhas em Colunas”, ao invés, ele se apresenta como “Coluna Dinâmica”.

Vá em “Transformar > Coluna Dinâmica”.

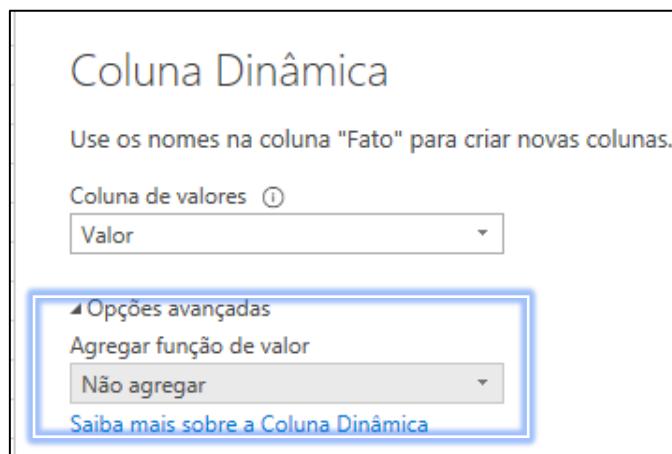


Após, o PB perguntará quem/qual valor/qual coluna que será alocada logo abaixo dessa nova coluna.

Obviamente queremos que a coluna Valor (que contém os valores de Devolução e Vendas) sejam alocadas embaixo dessa nova coluna.

Com a Coluna que queremos que se torne linha, por exemplo a “Valor”, precisamos selecionar a coluna e aí sim, selecionarmos a “Coluna Dinâmica” e decidirmos qual será a Coluna que receberá o “benefício” da transformação.

Por exemplo: *Queremos que as linhas dos Fatos virem duas colunas. Então selecionaremos essa coluna e aplique a qual. Nesse exemplo, precisamos que aplique na coluna de Valores.*



Veja que removendo o Tipo Alterado da própria consulta do Transformação do Arquivo Exemplo, o Warning das consultas sumirá.

Não esquecer de remover o “Tipo Alterado” no momento exato que for criada! Tanto no Exemplo quanto na Consulta do Grupo!

Consultas [5]

= Table.ExpandTableColumn(#"0")

	ABC 123 Região	ABC 123 Estado
1	Sudeste	Rio de Janeiro
2		null
3		null São Paulo
4		null
5		null
6		null
7		null
8		null
9	Sul	Rio Grande do Sul

Agora, vamos seguindo com os passos normalmente, até chegarmos na hora de **pivotarmos a Coluna Dinâmica**.

Um passo diferente do que foi colocado na primeira vez, que utilizaremos as Opções Avançadas para “Não Agregar” os valores.

O resultado será uma coluna de Devoluções e outra coluna de Vendas.

ABC Cidade	ABC Mês	123 Devoluções	123 Vendas
Rio de Janeiro	Abril 2017	2143	678
Rio de Janeiro	Agosto 2017	1446	980
Rio de Janeiro	Dezembro 2017	2443	456
Rio de Janeiro	Fevereiro 2017	1456	465
Rio de Janeiro	Janeiro 2017	1345	3456
Rio de Janeiro	Julho 2017	1223	9098
Rio de Janeiro	Junho 2017	1564	9875
Rio de Janeiro	Maio 2017	1432	6785
Rio de Janeiro	Marco 2017	5335	7565

Trabalhando com Datas

Vamos antes, alterar as datas. Precisamos alterar a tipagem para “Data”. Repare que as datas estão seguindo primeiro dia de cada mês.

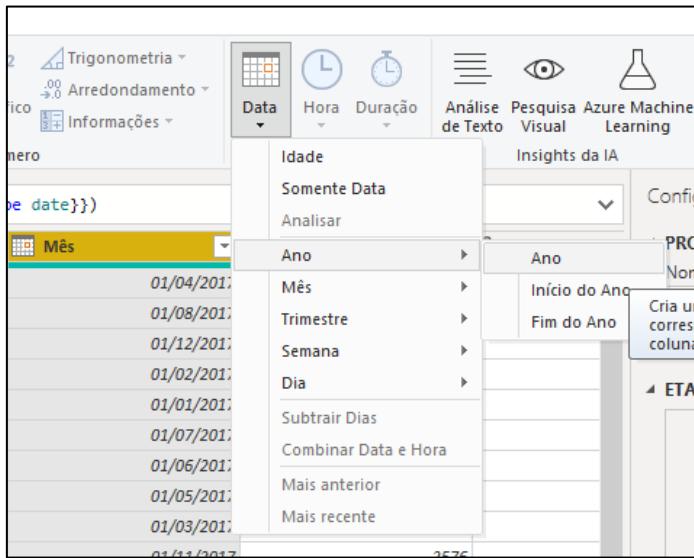
Mês

01/04/2017
01/08/2017
01/12/2017
01/02/2017
01/01/2017
01/07/2017
01/06/2017
01/05/2017
01/03/2017

A partir dessa tipagem para “Data”, vamos em “Adicionar Coluna > Data”



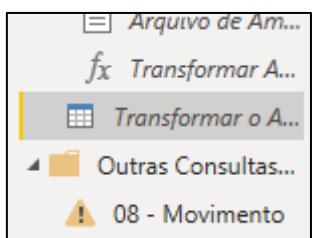
E a partir dessa seleção, podemos escolher e definir o que desejamos extrair dessa data.



Veja que a partir dessa data *tipada* corretamente, podemos extrair várias dados.

Mês	Devoluções	Vendas	Ano	Nome do Mês
01/04/2017	2143	67867	2017	abril
01/08/2017	1446	98098	2017	agosto
01/12/2017	2443	45678	2017	dezembro
01/02/2017	1456	46575	2017	fevereiro
01/01/2017	1345	345645	2017	janeiro
01/07/2017	1223	909898	2017	julho
01/06/2017	1564	987978	2017	junho
01/05/2017	1432	678978	2017	maio
01/03/2017	5335	756876	2017	março

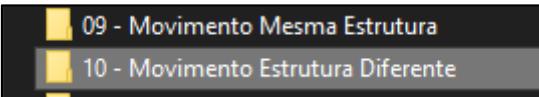
Porém, veja que existe um *Warning*. Esse Warning ativo, não deixa que após fecharmos e aplicarmos, as mudanças sejam feitas. Ele precisa ser tratado.



Essa explicação do *Warning* será demonstrada nas páginas 108 a 115. Esse exercício será retomado na página 116.

Explicando alguns erros

Vamos utilizar os arquivos abaixo para entendermos alguns erros



Abrindo a pasta do “09 – Movimentos Mesma Estrutura”, análise o formato e estrutura dos dados

Mes	Valor	Mes	Valor	Mes	Valor
Janeiro	10	Fevereiro	15	Março	60
Janeiro	20	Fevereiro	27	Março	67
Janeiro	30	Fevereiro	39	Março	74
Janeiro	40	Fevereiro	51	Março	81
Janeiro	50	Fevereiro	63	Março	88
Janeiro	60	Fevereiro	75	Março	95
Janeiro	70	Fevereiro	87	Março	102
Janeiro	80	Fevereiro	99	Março	109
Janeiro	90	Fevereiro	111	Março	116

Enquanto abrindo a pasta de “10 – Movimentos Estrutura Diferente”, veja que a coluna de Valor é diferente entre as 3 planilhas.

Mes	Valor Janeiro	Mes	Valor Fevereiro	Mes	Valor Março
Janeiro	10	Fevereiro	15	Março	60
Janeiro	20	Fevereiro	27	Março	67
Janeiro	30	Fevereiro	39	Março	74
Janeiro	40	Fevereiro	51	Março	81
Janeiro	50	Fevereiro	63	Março	88
Janeiro	60	Fevereiro	75	Março	95
Janeiro	70	Fevereiro	87	Março	102
Janeiro	80	Fevereiro	99	Março	109
Janeiro	90	Fevereiro	111	Março	116

Veja que a palavra Valor é acrescida do mês que é dito pela planilha. Isso causará um erro ao tentarmos trabalhar com os cabeçalhos e unir as planilhas.

Vamos regressar a pasta de número 09 que possuem as mesmas estruturas e vejamos o resultado delas no Power BI.

Faremos tudo que fizemos anteriormente, limpar todas as colunas, deixar somente o arquivo Binário, e selecionarmos a opção de “Transformar o Arquivo Exemplo”.

Nenhuma alteração muito significativa, apenas alteramos o tipo do Valor para Inteiro e o Mês para somente Texto.

The screenshot shows the Power Query Editor interface. On the left, the 'Consultas [5]' pane lists several items, with 'Transformar o Arquivo Exemplo' selected. On the right, a table is displayed with columns 'Mes' and 'Valor'. The table data is as follows:

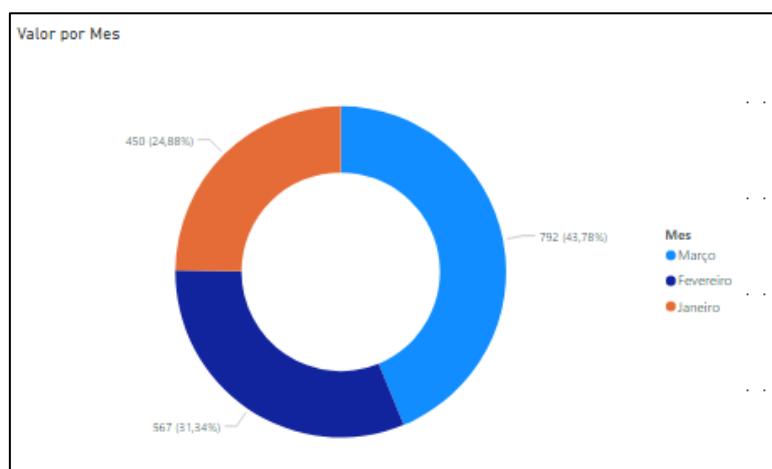
	Mes	Valor
1	Fevereiro	15
2	Fevereiro	27
3	Fevereiro	39
4	Fevereiro	51
5	Fevereiro	63
6	Fevereiro	75
7	Fevereiro	87
8	Fevereiro	99
9	Fevereiro	111

Veja que está sendo realizada no Exemplo, e não no 09 – movimento.

Vamos fechar e aplicar as mudanças.

The screenshot shows the 'Campos' (Fields) pane in Power BI. It contains a search bar and a list of fields under the '09 - Movimento ...' table. The visible fields are 'Mes' and 'Σ Valor'.

Realizando um gráfico para consultar se os valores estão ok, e veja as 3 planilhas perfeitamente alocadas.

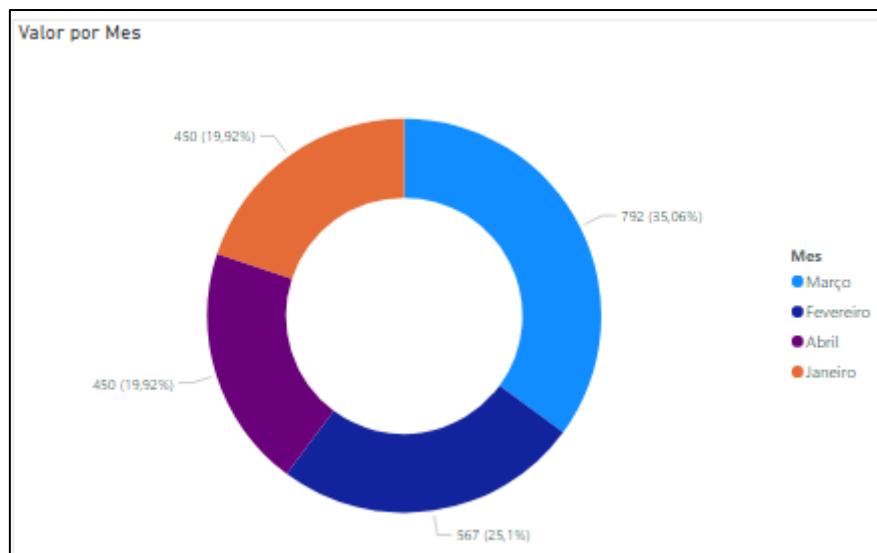


Podemos criar uma planilha de Abril na mão, desde que seja na mesma estrutura das anteriores e inseri-la na pasta.

Mes	Valor
Abril	10
Abril	20
Abril	30
Abril	40
Abril	50
Abril	60
Abril	70
Abril	80
Abril	90

DADOS (D:) > workspace > curso_MasterPowerBI > Datasets > 09 - Movimento Mensal		
Nome	Data de modificação	Tipo
Fevereiro.xlsx	25/10/2019 07:06	Planilha do Microsoft Excel
Janeiro.xlsx	25/10/2019 07:06	Planilha do Microsoft Excel
Marco.xlsx	25/10/2019 07:06	Planilha do Microsoft Excel
Abril.xlsx	17/03/2021 20:48	Planilha do Microsoft Excel

Ao retornarmos ao **Power BI** e atualizarmos o painel, com o Refresh, veja o resultado.



A planilha de Abril sofreu todas as alterações e portanto, foi adicionada corretamente e atualizada.

Vamos agora, para a outra pasta, que dessa vez, a estrutura está diferente.

Vamos conectar com a pasta da estrutura diferente.

Arquivos com Estruturas Diferentes

Após conectar, vamos remover as demais colunas, combinar os arquivos e utilizar a transformação do arquivo exemplo.

The screenshot shows the Power BI Data Editor interface. On the left, the 'Consultas [5]' pane lists several queries, with '10 - Movimento...' highlighted. On the right, a table view displays data with two columns: 'Mes' and 'Valor Fevereiro'. The table has 9 rows, each containing the value 'Fevereiro' in the first column and a numerical value in the second column: 15, 27, 39, 51, 63, 75, 87, 99, and 111. A status bar at the top indicates the formula: '= Table.PromoteHeaders(Planilha1_Sh...)'.

É sabido que a coluna de valores não poderia se chamar “**Valor Fevereiro**”, ela deveria se chamar apenas “**Valor**”.

Mas um erro muito comum é pensar: “Opa, vou substituir o nome de Valor Fevereiro para somente Valor”. Isso será um grande erro, pois nas demais planilhas não existe “Valor Fevereiro”, isso gerará um *Warning*.

Mesmo que não de um erro/*Warning* no exemplo, dará no grupo.

This screenshot shows the same Power BI Data Editor interface as the previous one. The '10 - Movimento...' query is now highlighted with a green border around its entry in the 'Consultas [5]' pane. The rest of the interface remains the same, including the table view on the right.

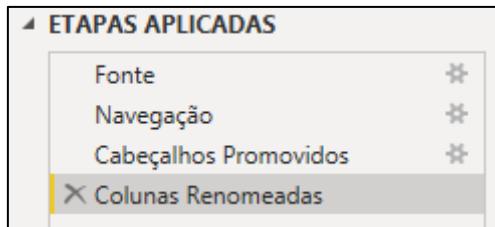
Para isso, vamos em: “[Página Inicial > Editor Avançado](#)”.

The screenshot shows the Power BI ribbon with the 'Editor Avançado' button highlighted. The ribbon tabs include 'Página Inicial', 'Transformar', 'Adicionar Coluna', 'Exibição', 'Ferramentas', and 'Ajuda'. Below the ribbon, there are several icons: 'Nova', 'Fontes', 'Inserir', 'Configurações da', 'Gerenciar', 'Atualizar', 'Propriedades', and 'Editor Avançado'.

Para verificar o erro. Significa que ele está tentando achar Fevereiro nos demais meses.

Para corrigir isso, vamos precisar deletar algumas etapas automáticas que o Power BI já inclui na iniciação do projeto.

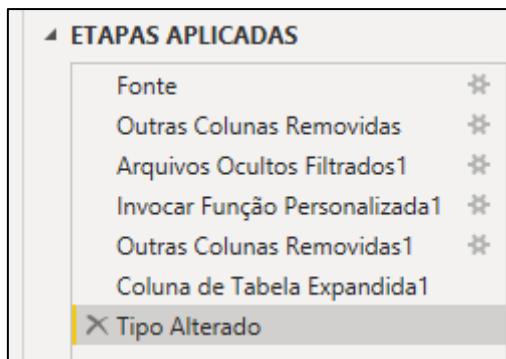
No Exemplo, vamos retirar as seguintes opções: **COLUNA RENOMEADA** [remover] e após **CABEÇALHOS PROMOVIDOS** [remover também].



Até que o Exemplo fique dessa forma

	ABC 123 Column1	ABC 123 Column2
1	Mes	Valor Fevereiro
2	Fevereiro	15
3	Fevereiro	27
4	Fevereiro	39
5	Fevereiro	51
6	Fevereiro	63
7	Fevereiro	75
8	Fevereiro	87
9	Fevereiro	99
10	Fevereiro	111

Enquanto no grupo, vamos precisar remover o **TIPO ALTERADO**[remover]



Até que fique conforme abaixo

	ABC Column1	ABC Column2
1	Mes	Valor Fevereiro
2	Fevereiro	15
3	Fevereiro	27
4	Fevereiro	39
5	Fevereiro	51
6	Fevereiro	63
7	Fevereiro	75
8	Fevereiro	87
9	Fevereiro	99
10	Fevereiro	111
11	Mes	Valor Janeiro
12	Janeiro	10
13	Janeiro	20
14	Janeiro	30
15	Janeiro	40

Agora que todos os arquivos estão “com o mesmo pé”, podemos retornar ao Exemplo e vamos remover a primeira linha. Todos têm uma primeira linha que contém o “Mês e o Valor nome_do_mês”.

	ABC Column1	ABC Column2
1	Mes	Valor Fevereiro
2	Fevereiro	15
3	Fevereiro	27

Como desejamos eliminar apenas a primeira linha de cada arquivo.

Remover Linhas Principais

Especifique quantas linhas serão removidas da parte superior.

Número de linhas

1

Se voltarmos ao grupo, podemos verificar que nenhuma planilha tem os cabeçalhos dessa forma.

Consultas [5]

	ABC 123 Column1	ABC 123 Column2
1	Fevereiro	15
2	Fevereiro	27
3	Fevereiro	39
4	Fevereiro	51
5	Fevereiro	63
6	Fevereiro	75
7	Fevereiro	87
8	Fevereiro	99
9	Fevereiro	111
10	Janeiro	10
11	Janeiro	20
12	Janeiro	30
13	Janeiro	40
14	Janeiro	50
15	Janeiro	60

Agora, retornaremos para o Exemplo, e vamos alterar e renomear as colunas, transformando-as em “Variáveis Globais”.

PROPRIEDADES

Nome
Transformar o Arquivo de Exemplo

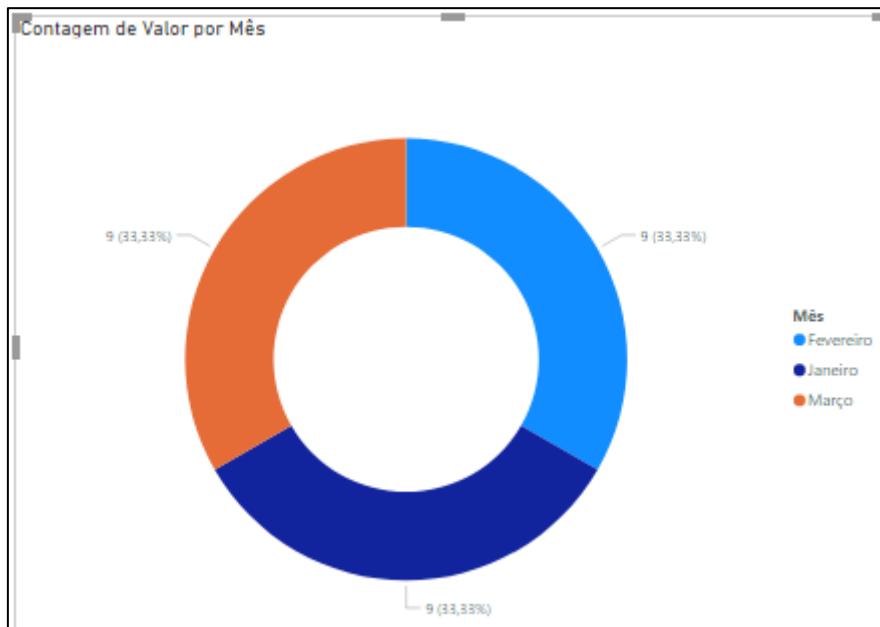
Todas as Propriedades

ETAPAS APPLICADAS

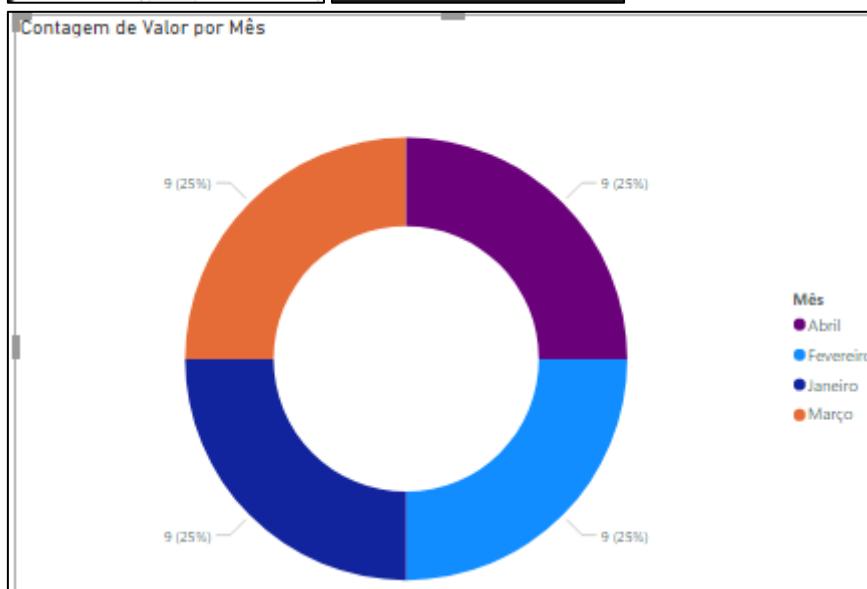
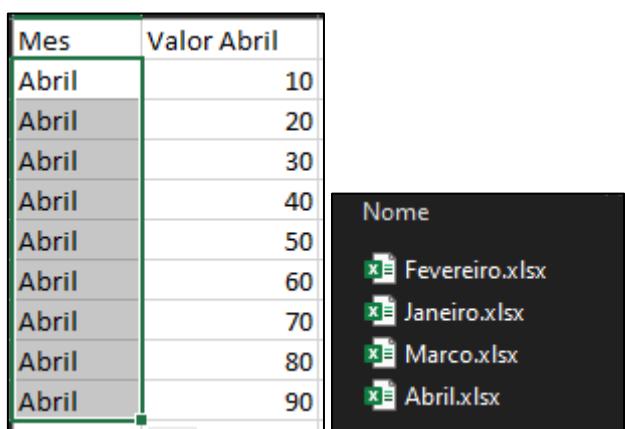
Fonte
Navegação
Linhas Principais Removidas
Colunas Renomeadas

Mês	Valor
1	15
2	27
3	39
4	51
5	63
6	75
7	87
8	99
9	111

Podemos fechar e aplicar que agora será possível.



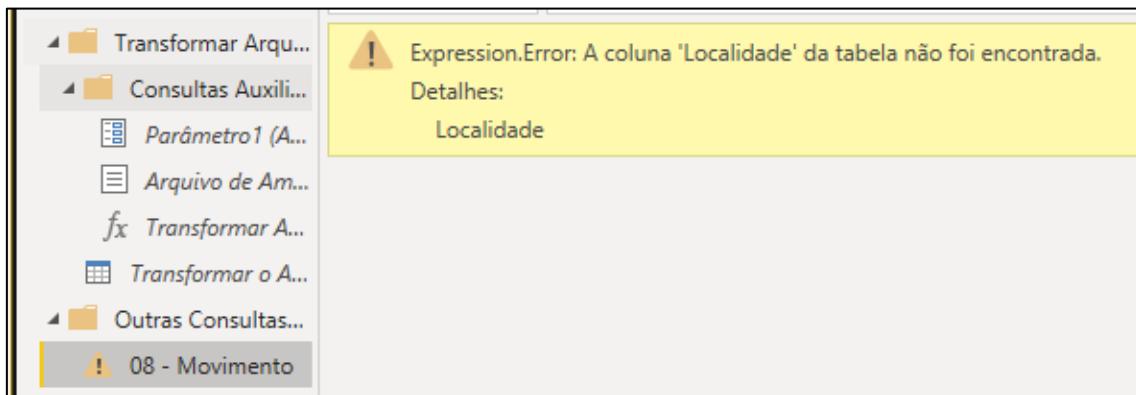
Podemos fazer aquele mesmo exemplo de adicionar a planilha de “Abril” e verificar se está funcionando e adicionando os valores de Abril.



Trabalhando com Datas – Continuação do Exercício 08

Agora com esse erro explicado, podemos regressar ao arquivo de “08 – Movimentos” e reparar o *Warning* que foi gerado.

Veja que já logo quando voltamos a mexer e tratar com as exclusões, subir a primeira linha como cabeçalho... um erro já é mostrado.



Isso acontece porque o **Power BI** começa a tipar as colunas do Exemplo automaticamente, e quando ele tipa, ele “coloca” o nome da coluna na Linguagem M. E isso faz com que não se encontre os nomes dessas colunas no grupo de planilhas.

Para corrigir isso, vamos voltar no grupo e remover o Tipo Alterado.

This screenshot shows the Power BI Data Flow interface with the 'ETAPAS APLICADAS' (Applied Steps) pane open. The pane lists several steps: Fonte, Linhas Filtradas, Outras Colunas Removidas, Arquivos Ocultos Filtrados1, Invocar Função Personalizada1, Outras Colunas Removidas1, Coluna de Tabela Expandida1, and Tipo Alterado. The 'Tipo Alterado' step is highlighted with a yellow background, indicating it is the cause of the current error.

Até que o primeiro erro desapareça, embora ainda um erro persista.

This screenshot shows the Power BI Data Flow interface with a table preview on the right. The table has two columns: 'Região' and 'Estado'. Row 1 contains 'Sudeste' and 'Rio de Janeiro'. Rows 2 through 8 all contain 'null' in both columns. A blue oval highlights the header row, and a red oval highlights the 'Região' column in the second row. The table has a red border around its preview area.

Voltando ao Exemplo, vamos remover também o “TIPO ALTERADO” e esse erro sumirá.

The screenshot shows the Power BI Data Editor interface. On the left, there's a navigation pane with several items under 'Transformar Arquivo'. In the center, a table view displays two columns: 'Região' and 'Estado'. The first row has 'Sudeste' in 'Região' and 'Rio de Janeiro' in 'Estado'. Subsequent rows from 2 to 8 have 'null' in both columns. To the right of the table is a panel titled 'ETAPAS APLICADAS' (Applied Steps) which lists several steps: 'Fonte', 'Navegação', 'Cabeçalhos Promovidos', 'Linhas Inferiores Removidas', 'Colunas Removidas', 'Cabeçalhos Promovidos1', and 'Tipo Alterado' (which is highlighted with a yellow border). Below the table, a status bar shows '1 row(s) 1 column(s)'.

Veja que removendo o Tipo Alterado, o erro sumirá.

This screenshot shows the same Power BI Data Editor interface after the 'Tipo Alterado' step has been removed. The 'ETAPAS APLICADAS' panel no longer contains the 'Tipo Alterado' step. The table now correctly displays the data: Row 1 has 'Sudeste' in 'Região' and 'Rio de Janeiro' in 'Estado'; Rows 2 through 8 have 'null' in both columns; and Row 9 has 'Sul' in 'Região' and 'Rio Grande do Sul' in 'Estado'. The status bar at the bottom indicates '9 row(s) 2 column(s)'.

Agora, vamos seguindo com os passos normalmente, até chegarmos na hora de **pivotarmos a Coluna Dinâmica**.

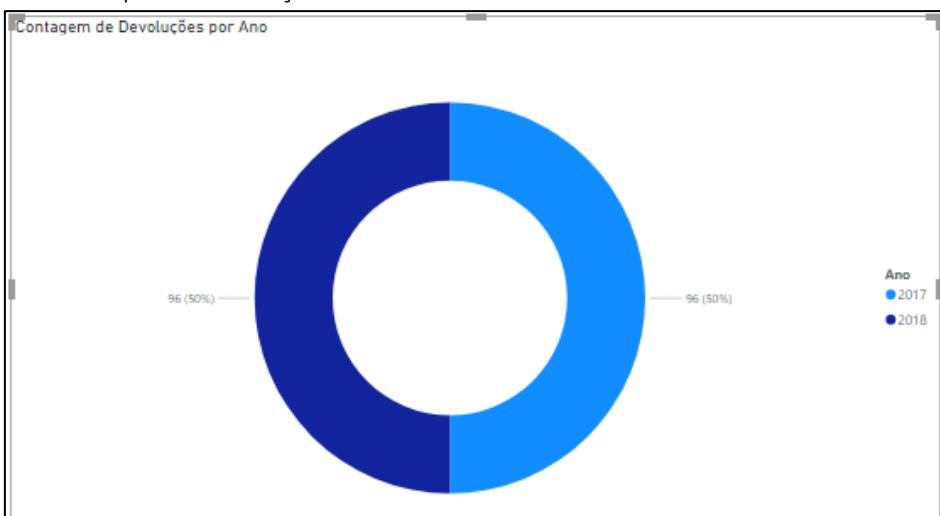
Um passo diferente do que foi colocado na primeira vez, que utilizaremos as Opções Avançadas para “Não Agregar” os valores.

This screenshot shows the 'Coluna Dinâmica' configuration dialog. At the top, it says 'Coluna Dinâmica' and 'Use os nomes na coluna "Fato" para criar novas colunas.' Below this, there's a section for 'Coluna de valores' (Value Column) with a dropdown menu set to 'Valor'. Underneath are two options: 'Opções avançadas' (Advanced Options) and 'Aregar função de valor' (Add Value Function). The dropdown menu for 'Aregar função de valor' is set to 'Não agrregar' (Do not aggregate). At the bottom, there's a link 'Saiba mais sobre a Coluna Dinâmica' (Learn more about the Dynamic Column).

Agora está perfeito.

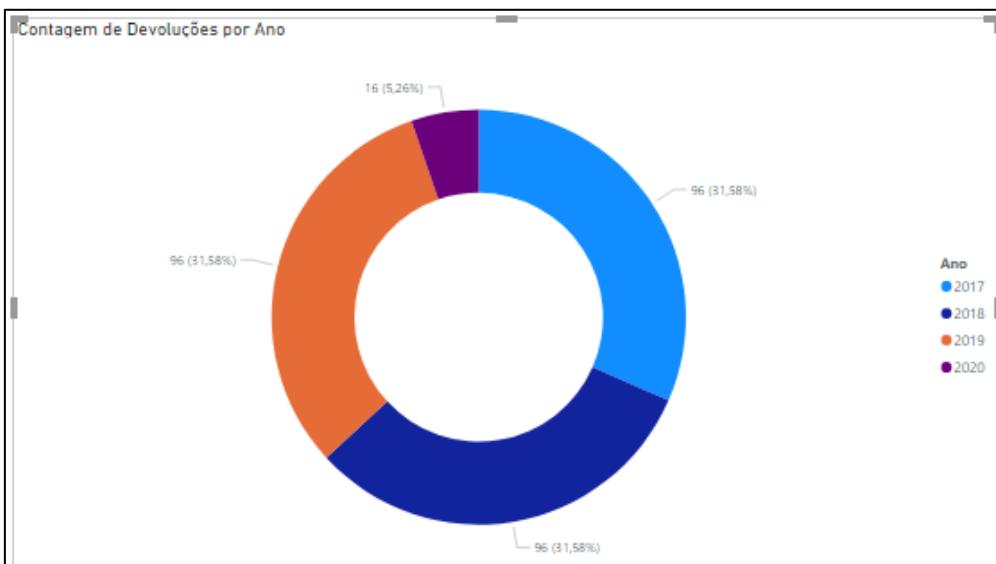
	ABC 123 Estado	ABC 123 Cidade	A ^B C Atributo	A ^B C Devoluções	A ^B C Vendas
1	Rio de Janeiro	Rio de Janeiro	Abri 2017	2143	678
2	Rio de Janeiro	Rio de Janeiro	Agosto 2017	1446	980
3	Rio de Janeiro	Rio de Janeiro	Dezembro 2017	2443	456
4	Rio de Janeiro	Rio de Janeiro	Fevereiro 2017	1456	465
5	Rio de Janeiro	Rio de Janeiro	Janeiro 2017	1345	3456
6	Rio de Janeiro	Rio de Janeiro	Julho 2017	1223	9098
7	Rio de Janeiro	Rio de Janeiro	Junho 2017	1564	9879
8	Rio de Janeiro	Rio de Janeiro	Maio 2017	1432	6789
9	Rio de Janeiro	Rio de Janeiro	Março 2017	5335	7568
10	Rio de Janeiro	Rio de Janeiro	Novembro 2017	2576	879
11	Rio de Janeiro	Rio de Janeiro	Outubro 2017	5463	346
12	Rio de Janeiro	Rio de Janeiro	Setembro 2017	6433	345
13	São Paulo	Piracicaba	Abri 2017	5700	248

Vamos agora tipar a Data e fazer os pedidos de ano e mês. E vamos tipar todos os cabeçalhos corretamente. Fecharemos e aplicaremos e faremos o gráfico de rosca somente para verificação.



Vamos aplicar as outras duas planilhas e atualizaremos para verificar se está ok.

Dessa forma, já atualizará automaticamente sempre que realizarmos adições.



Relacionamento

Case World Cup

Vamos conectar o nosso Power BI a uma tabela de internet, pela Wikipedia.

Conectaremos a uma tabela totalmente desconfigurada igual a essa

Por seleções				
Seleção	Títulos	Vice	Terceiro	Quarto
Brasil	5 (1958, 1962, 1970, 1994 e 2002)	2 (1950 e 1998)	2 (1938 e 1978)	2 (1974 e 2014)
Alemanha ^[nota 3]	4 (1954, 1974, 1990 e 2014)	4 (1966, 1982, 1986 e 2002)	4 (1934, 1970, 2006 e 2010)	1 (1958)
Itália	4 (1934, 1938, 1982 e 2006)	2 (1970 e 1994)	1 (1990)	1 (1978)
Argentina	2 (1978 e 1986)	3 (1930, 1990 e 2014)	–	–
França	2 (1998 e 2018)	1 (2006)	2 (1958 e 1986)	1 (1982)
Uruguai	2 (1930 e 1950)	–	–	3 (1954, 1970 e 2010)
Inglaterra	1 (1966)	–	–	2 (1990 e 2018)
Espanha	1 (2010)	–	–	1 (1950)
Países Baixos	–	3 (1974, 1978 e 2010)	1 (2014)	1 (1998)
Hungria	–	2 (1938 e 1954)	–	–
Tchecoslováquia	–	2 (1934 e 1962)	–	–
Suécia	–	1 (1958)	2 (1950 e 1994)	1 (1938)
Croácia	–	1 (2018)	1 (1998)	–

Para isso, vamos abrir o PB e estabelecer a conexão com os dados da seguinte forma.

Obter Dados

X

Tudo

Serviços Online

Outro

Tudo

Web

W Webtrends Analytics (Beta)

Selecionaremos o “Web” e conectar, e alocar a url da wikipedia.

Da Web

Básico Avançadas

URL

Não são todos os sites que permitem a varredura de tabelas.

Vamos fazer trabalho de formiguinha e tentaremos descobrir qual a tabela que precisamos e desejamos selecionar.

Opções de Exibição

- Tabelas HTML [27]
 - Tabela 1
 - Tabela 2
 - Tabela 3
 - Por edições[editar código-fonte]
 - Por seleções[editar código-fonte] **(selected)**
 - Por confederações[editar código-fonte]
 - 10 maiores goleadas[editar | editar có...]
 - Tabela 4
 - Tabela 5
 - Público nos estádios[editar | editar cód...]
 - Tabela 6
 - Tabela 7
 - Preços dos ingressos da Copa do Mun...
 - Valores dos direitos televisivos[190]
 - Ligações externas[editar | editar códig...]
 - Tabela 8
 - Tabela 9
 - Tabela 10

Pois veja que o Power BI procurou e encontrou diversas *tags <table>* ao estabelecer a conexão com a Wikipedia. Utilizaremos a “Por Seleções”.

Column1	Column2	Column3
Seleção	Títulos	Vice
Brasil	5 (1958, 1962, 1970, 1994 e 2002)	2 (1950 e 1998)
Alemanha[nota 3]	4 (1954, 1974, 1990 e 2014)	4 (1966, 1982, 1986 e 2006)
Itália	4 (1934, 1938, 1982 e 2006)	2 (1970 e 1994)
Argentina	2 (1978 e 1986)	3 (1930, 1990 e 2014)
França	2 (1998 e 2018)	1 (2006)
Uruguai	2 (1930 e 1950)	–
Inglaterra	1 (1966)	–
Espanha	1 (2010)	–
Países Baixos	–	3 (1974, 1978 e 2010)
Hungria	–	2 (1938 e 1954)
Tchecoslováquia	–	2 (1934 e 1962)

Vamos analisar essa tabela.

A ^B _C Column1	A ^B _C Column2	A ^B _C Column3	A ^B _C Column4
1 Seleção	Títulos	Vice	Terceiro
2 Brasil	5 (1958, 1962, 1970, 1994 e 2002)	2 (1950 e 1998)	2 (1938 e 1978)
3 Alemanha[nota 3]	4 (1954, 1974, 1990 e 2014)	4 (1966, 1982, 1986 e 2002)	4 (1934, 1970, 2006 e 2010)
4 Itália	4 (1934, 1938, 1982 e 2006)	2 (1970 e 1994)	1 (1990)
5 Argentina	2 (1978 e 1986)	3 (1930, 1990 e 2014)	-
6 França	2 (1998 e 2018)	1 (2006)	2 (1958 e 1986)
7 Uruguai	2 (1930 e 1950)	-	-
8 Inglaterra	1 (1966)	-	-

O que temos de dados? Temos **nomes de países** e **anos** que venceram ou foram vice, terceiros... Ou seja, vamos precisar transformar e fazer essa tabela crescer para baixo. Fora que percebam que o **Power BI** também não lê vários dados que estejam na mesma linha, veja o caso do Brasil, que possui 5 títulos na mesma linha.

País	Ano
Brasil	1958
Brasil	1962
Brasil	1970
Brasil	1994
Brasil	2002

Essa deveria ser a “formatação” que buscamos.

Vamos renomear a consulta da tabela para “Raw(cru)”, que são os dados sem nenhum tratamento.

The screenshot shows the Power BI Data View interface. On the left, there is a list of tables under 'Consultas [1]'. One table is highlighted with a yellow background and has its name 'Raw' visible. On the right, there is a detailed view of the 'Raw' table, showing its columns and data rows. The first column is labeled 'A^B_C Column1' and contains the value '1 Seleção'. Below it, other rows show data for Brazil and other countries.

Clicando com o botão direito na Raw, procure por “Duplicar”. Vamos renomear para Original após duplicar.

The screenshot shows the Power BI Data View interface again. Now, there are two tables listed under 'Consultas [2]': 'Raw' and 'Raw (2)'. The 'Raw (2)' table is highlighted with a yellow background and has its name 'Original' visible. This indicates that the original table has been successfully duplicated.

Agora dentro e selecionado a “Original”, vamos clicar na coluna de Títulos e procure por “Transformar > Extrair > Entre os delimitadores”.

Por que “entre os delimitadores”? Porque se perceber, os anos estão entre (parênteses). Precisamos extrair os anos dos títulos.

Extração de Texto Entre Delimitadores

A screenshot of the Power Query Editor interface. The 'Transformar' tab is selected in the ribbon. A context menu is open over the 'Títulos' column, specifically over the value '5 (1958, 1962, 1970, 1994 e 2002)'. The 'Extrair' option is highlighted in the menu. A dropdown menu for 'Text Between Delimiters' is open, showing options like 'Duração', 'Primeiros caracteres', 'Últimos caracteres', 'Intervalo', 'Texto Antes do Delimitador', 'Texto Após o Delimitador', and 'Texto Entre os Delimitadores'.

E vamos selecionar “Transformar > Extrair > Texto Entre os Delimitadores”.

A screenshot of the 'Text Between Delimiters' configuration dialog. It has two input fields: 'Delimitador de início' containing '(' and 'Delimitador de fim' containing ')'. At the bottom, there is a link labeled 'Opções avançadas'.

Veja o resultado. Agora precisaremos aplicar para todas as colunas.

	Column1	Column2	Column3
1	Seleção		Vice
2	Brasil	1958, 1962, 1970, 1994 e 2002	2 (1950 e 1998)
3	Alemanha[nota 3]	1954, 1974, 1990 e 2014	4 (1966, 1982, 1986 e 2002)
4	Itália	1934, 1938, 1982 e 2006	2 (1970 e 1994)
5	Argentina	1978 e 1986	3 (1930, 1990 e 2014)
6	França	1998 e 2018	1 (2006)
7	Uruguai	1930 e 1950	-
8	Inglaterra	1966	-
9	Espanha	2010	-
10	Países Baixos		3 (1974, 1978 e 2010)
11	Ucrânia		2 (1920 e 1954)

Lembrando do atalho do SHIFT, selecionando a coluna e então Transformar > Extrair > Entre os delimitadores

Agora teremos outros problemas na limpeza. Veja que ainda possuímos diversos valores em uma única célula separados por vírgula (,) e o uso do (e).

1958, 1962, 1970, 1994 e 2002	1950 e 1998
1954, 1974, 1990 e 2014	1966, 1982, 1986 e 2002
1934, 1938, 1982 e 2006	1970 e 1994
1978 e 1986	1930, 1990 e 2014
1998 e 2018	2006

Vamos “Padronizar” pelo menos, substituindo o (e) por somente (,). Clicando com o botão direito e “Substituir valores”.

1958, 1962, 1970, 1994 e 2002	1950 e 1998	1938 e 1978
1954, 1974, 1990 e 2014	1966, 1982, 1986 e 2002	1934, 1970, 2006 e 2014
1934, 1938, 1982 e 2006	1970 e 1994	1990
1978 e 1986	1930, 1990 e 2014	

Substituir Valores

Substitua um valor por outro nas colunas selecionadas.

Valor a Ser Localizado
e

Substituir por
,

Opções avançadas

Veja o resultado abaixo.

A ^B _C Column2	A ^B _C Column3	A ^B _C Column4
1958, 1962, 1970, 1994 , 2002	1950 , 1998	1938 , 1978
1954, 1974, 1990 , 2014	1966, 1982, 1986 , 2002	1934, 1970, 2006 , 2014
1934, 1938, 1982 , 2006	1970 , 1994	1990
1978 , 1986	1930, 1990 , 2014	
1998 , 2018	2006	1958 , 1986
1930 , 1950		
1966		
2010		

Não está no formato ok, mas pelo menos os dados estão padronizados.

Perceba que os valores ainda estão em colunas.

Nós não podemos relacionar o campeão e o vice, com terceiro, quarto.. Se fizermos isso, o Power BI vai começar a duplicar e assimilar todos os títulos várias vezes.

Agora podemos trabalhar na coluna Original.

Vamos estabelecer as relações entre os Países e seus respectivos Títulos.

A ideia aqui é o relacionamento entre uma tabela contendo País sem nenhuma outra informação e após relacionar com os títulos.

The screenshot shows the Power BI Data Editor interface. On the left, the 'Consultas [3]' pane lists 'Raw', 'Original', and 'País'. The main area displays a table with three columns: 'Seleção', 'Títulos', and 'Vice'. The 'Títulos' column contains country names followed by their titles. A context menu is open over the 'Títulos' column, with 'Remover Outras Colunas' highlighted. Other options in the menu include 'Copiar', 'Remover', 'Duplicar Coluna', 'Adicionar Coluna de Exemplos...', 'Remover Duplicadas', 'Remover Erros', 'Alterar Tipo', and 'Transformar'.

Agora queremos uma tabela que contenha os países e títulos.

Nessa Original (2), vamos apagar todas as demais outras colunas.

This screenshot shows the Power BI Data Editor with the 'Original (2)' query selected in the 'Consultas [4]' pane. The table structure has changed, with the 'Vice' and 'Terceiro' columns removed, leaving only 'Seleção' and 'Títulos'. The context menu for the 'Títulos' column is open again, with 'Remover Outras Colunas' selected. The menu includes options like 'Copiar', 'Remover', 'Duplicar Coluna', 'Adicionar Coluna de Exemplos...', 'Remover Duplicadas', 'Remover Erros', 'Alterar Tipo', and 'Transformar'.

E agora sim, podemos dividir as Colunas em Linhas.

The final screenshot shows the 'Original (2)' query selected. The table now has two columns: 'Seleção' and 'Títulos'. The 'Títulos' column contains a single cell per row, which is a list of titles separated by commas. A context menu is open over the 'Títulos' column, with 'Remover Outras Colunas' selected. The menu options are identical to the previous ones.

E agora podemos dividir a Coluna por Delimitador, “[Transformar > Dividir Coluna > Por Delimitador](#)”.

The screenshot shows the Power BI 'Transform' ribbon with the 'Dividir Coluna' (Divide Column) option highlighted. A dropdown menu is open, listing various delimiter-based division options: 'Por Delimitador', 'Por Número de Caracteres', 'Por Posições', 'Por Minúscula para Maiúscula', 'Por Maiúscula para Minúscula', 'Por Dígitos para Não Dígitos', and 'Por Não Dígitos para Dígitos'. Below the ribbon, a table is displayed with two columns: 'Seleção' and 'Títulos'. The 'Títulos' column contains values like '1958, 1962, 1970, 1994, 2002'.

Dividir Coluna por Delimitador

Especifique o delimitador usado para dividir a coluna de texto.

Selecione ou insira o delimitador

Vírgula

Dividir em

- Delimitador da extremidade esquerda
- Delimitador da extremidade direita
- Cada ocorrência do delimitador

▲ Opções avançadas

Dividir em

- Colunas
- Linhas

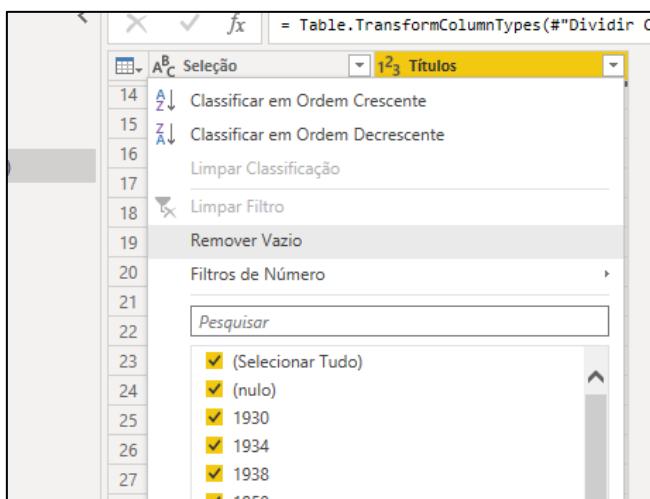
E nessa divisão agora, teremos os países duplicados conforme a quantidade de seus títulos.

	A ^B Seleção	1 ² 3 Títulos
1	Brasil	1958
2	Brasil	1962
3	Brasil	1970
4	Brasil	1994
5	Brasil	2002
6	Alemanha[nota 3]	1954
7	Alemanha[nota 3]	1974
8	Alemanha[nota 3]	1990
9	Alemanha[nota 3]	2014
10	Itália	1934
11	Itália	1938
12	Itália	1982

Alguns países não possuem títulos e estão com valores nulos, podemos remover os vazios.

	A ^B C Seleção	Títulos
14	Argentina	1978
15	Argentina	1986
16	França	1998
17	França	2018
18	Uruguai	1930
19	Uruguai	1950
20	Inglaterra	1966
21	Espanha	2010
22	Países Baixos	null
23	Hungria	null
24	Tchecoslováquia	null
25	Suécia	null
26	Croácia	null
27	Polónia	null
28	Áustria	null
29	Portugal	null
30	Bélgica	null

Com o botão direito na coluna de títulos, vamos em “Remover Vazio”.



O nosso resultado será a tabela entre os países e seus títulos de campeões.

	A ^B C Seleção	1 ² 3 Títulos	▼
1	Brasil		1958
2	Brasil		1962
3	Brasil		1970
4	Brasil		1994
5	Brasil		2002
6	Alemanha[nota 3]		1954
7	Alemanha[nota 3]		1974
8	Alemanha[nota 3]		1990
9	Alemanha[nota 3]		2014
10	Itália		1934
11	Itália		1938
12	Itália		1982

Agora precisamos fazer mais uma tabela que contenha o País e o Vice.

The screenshot shows the Power Query Editor interface. On the left, there's a navigation pane titled "Consultas [5]" containing items: Raw, Original, País, Campeão, and Vice. The "Vice" item is currently selected. The main area displays a table with two columns: "Seleção" and "Vice". The table has 6 rows:

	A ^B C Seleção	A ^B C Vice
1	Brasil	1950 , 1998
2	Alemanha[nota 3]	1966, 1982, 1986 , 2002
3	Itália	1970 , 1994
4	Argentina	1930, 1990 , 2014
5	França	2006
6	Uruguai	

E aplicaremos o mesmo passo. “Transformar > Dividir Coluna > Por Delimitador”

The screenshot shows the Power Query Editor after applying the "Divide Column by Delimiter" step. The table now has three columns: "Seleção", "Vice", and another "Vice" column. The table has 16 rows:

	A ^B C Seleção	1 ² ₃ Vice	Vice
1	Brasil		1950
2	Brasil		1998
3	Alemanha[nota 3]		1966
4	Alemanha[nota 3]		1982
5	Alemanha[nota 3]		1986
6	Alemanha[nota 3]		2002
7	Itália		1970
8	Itália		1994
9	Argentina		1930
10	Argentina		1990
11	Argentina		2014
12	França		2006
13	Uruguai		null
14	Inglaterra		null
15	Espanha		null
16	Países Baixos		1974

Lembrando de novamente, vamos remover os vazios.

E agora precisaremos realizar esse processo para o Terceiro e o Quarto lugar.

Relacionando Consultas

Após salvarmos todas as edições acima e aplicarmos para voltarmos ao Power View, o PB vai apresentar uma tela de carregamento que pode levar um tempo para relacionar cada uma das tabelas.



Se ocorrer algum problema, tente fechar e aplicar novamente as alterações.

The screenshot shows the 'Campos' (Fields) pane. On the left, there's a sidebar with tables: 'Campeão', 'Original' (selected), and 'País'. The main area shows the columns of the 'Original' table: 'Quarto', 'Seleção', 'Terceiro', 'Títulos', and 'Vice'. Below these, under the 'Raw' table, are the columns: 'Column1', 'Column2', 'Column3', 'Column4', and 'Column5'. Further down are the tables 'Terceiro' and 'Vice'.

Veja que agora todas as tabelas foram importadas.

Mas nem todas nós vamos precisar, como são os casos da Original e da Raw.

Elas estão cruas, estão sem a devida formatação.

Habilitar e Desabilitar Carga

Veja, se abrirmos essas duas tabelas cruas, estamos novamente com aquele problema.

The screenshot shows the Power BI Data View interface. On the left, there are two tables: 'Seleção' and 'Títulos'. The 'Seleção' table has columns 'Seleção' and 'Títulos', with data like 'Austrália', 'Bélgica', etc. The 'Títulos' table has columns 'Column1' and 'Column2', with data like 'Austrália', 'Bélgica', etc. To the right of the tables are several filter panes:

- Filtros neste visual**:
 - Seleção é (Tudo)
 - Títulos é (Tudo)
 - Adicionar os campos de da...
- Filtros nesta página**:
 - Adicionar os campos de da...
- Valores**:
 - Seleção
 - Títulos
- Drill-through**:
 - Relatório cruzado
 - Desativado
- Motor de todos os Filtros**

Ou seja, tanto a Original quanto a Raw, não as utilizaremos mais. Para isso, voltaremos ao **Power Query** somente para Desabilitarmos a Carga e assim não será mais carregado no View.

The screenshot shows the Power Query ribbon with 'Consultas [7]' selected. In the center, there is a table with columns 'Column1' and 'Column2'. A context menu is open over the 'Raw' table, listing options: Copiar, Colar, Excluir, Renomear, Habilitar carga (with a checked checkbox), Incluir na atualização do relatório (with a checked checkbox), Duplicar, Referência, Mover Para o Grupo, Mover para Cima, and Mover para Baixo.

Façamos isso tanto para a Raw, quanto para a Original.

Por que devemos desabilitar? Porque assim não carregará os dados adiante, mas caso precisássemos novamente dessas tabelas, elas estariam à disposição, sem causar nenhuma interrupção na performance dos dados.

The screenshot shows the Power BI Fields pane titled 'Campos'. It lists several groups:

- Campeão**: Contains 'Seleção' (checked) and 'Σ Títulos' (checked).
- País**
- Quarto**
- Terceiro**
- Vice**

Enquanto a tabela que criamos de “Campeão” nós podemos utilizá-la.

Só apenas uma observação, troque a soma dos títulos para “Não Resumir”.

The screenshot shows the Power Pivot ribbon with the 'Filtros' tab selected. On the left, the 'Seleção' table is shown with the 'Títulos' column highlighted. In the center, the 'Valores' section displays the formula Σ Títulos. On the right, the 'Pesquisar' search bar is visible.

Modelagem dos Dados no Power Pivot

Continuando o anterior, vamos relacionar as tabelas no **Power Pivot**. O acesso ao PP se dará ao botão no canto esquerdo.



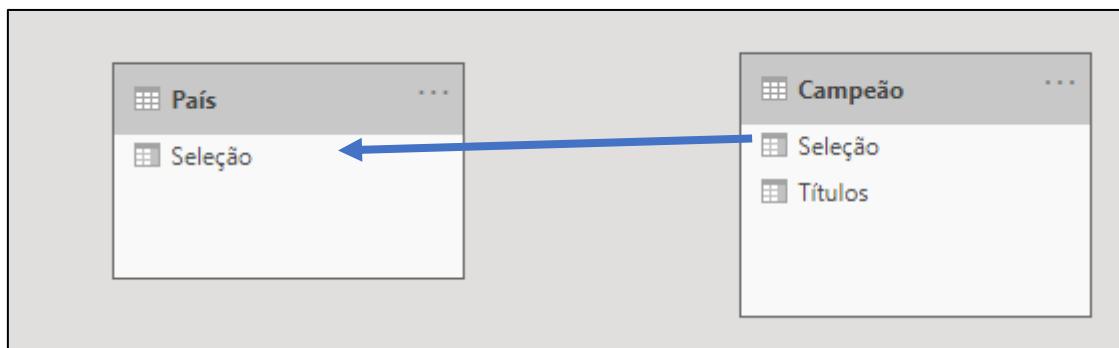
A seguinte tela será disponibilizada. Esse é o **Power Pivot**, ou, o local que se relacionam as tabelas.

The screenshot shows the Power Pivot workspace with five tables: 'Campeão', 'Quarto', 'Terceiro', 'País', and 'Vice'. Each table has a 'Seleção' column and either a 'Títulos' or 'Vice' column. A message at the top says: "Atualize para o novo modo de exibição de modelo para um design aprimorado que facilite a identificação de fontes de dados e o gerenciamento de relacionamentos. Tenha em mente que essa alteração é permanente."

Relacionamento de Um para Muitos (1 – N)

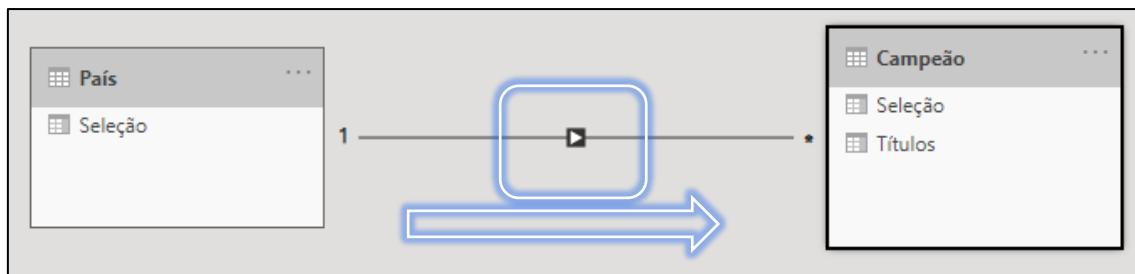
Entende-se que o País – Seleção é quem deve filtrar os títulos de campeão, vice, terceiro e quarto. Portanto, ao estabelecermos as relações, precisamos estabelecer que **UM** país pode possuir **UM ou VÁRIOS** títulos de campeão (**Relacionamento de Um para Muitos**).

Vamos puxar a “Seleção” do Campeão, ao País, arrastando o campo da “Seleção” para dentro da tabela País.



O que isso significa? Que o Filtro da Seleção na tabela “País”, agora é relacionado ao filtro na tabela de “Campeão”.

Isso significa que um país (1) pode ter UM ou VÁRIOS Títulos de Campeão (*). Isso acontece porque na tabela “País”, o país aparecerá apenas uma única vez. Enquanto na tabela “Campeão”, o país aparecerá uma ou várias vezes.

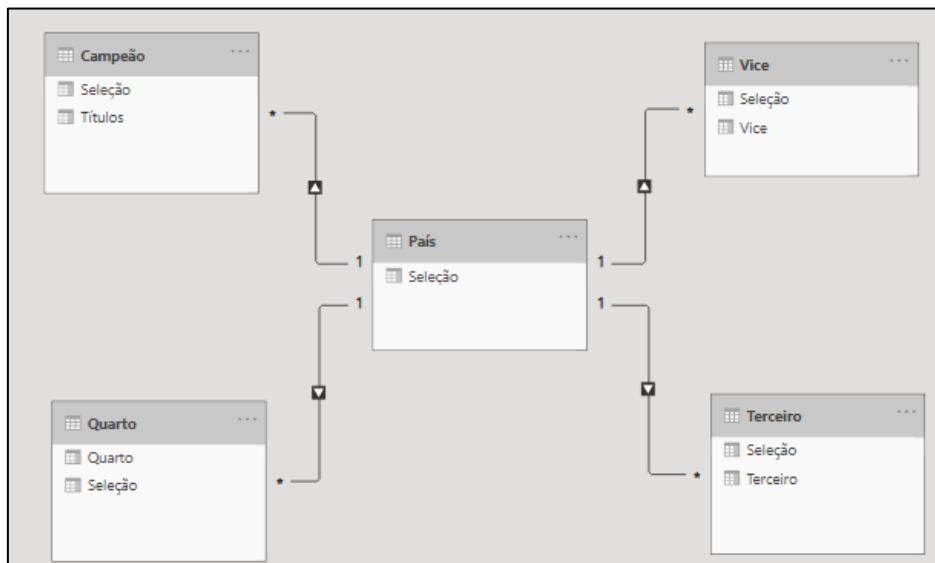


Acompanhe também que o ícone , apresenta o sentido do Filtro, ou seja, é a partir de “País” que o filtro será condicionado.

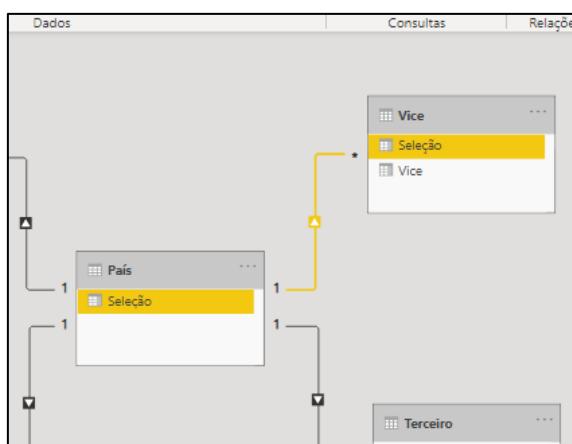
Deveremos puxar todos os fatos(Campeão, Vice, Terceiro, Quarto) para a única dimensão existente (Seleção).

Vejamos as conexões na página 132.

Lembre-se de sempre arrastar a “Seleção” de cada Fato para a “Seleção” da Dimensão.

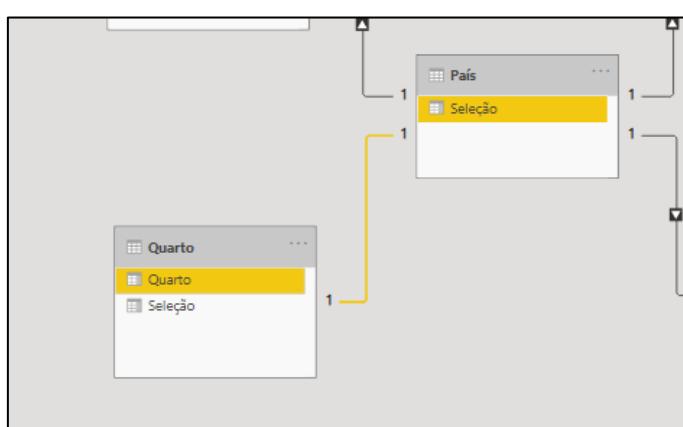


Em caso de dúvida, arraste o mouse em cima e verifique se a relação entre o fato e a dimensão foi feita corretamente.

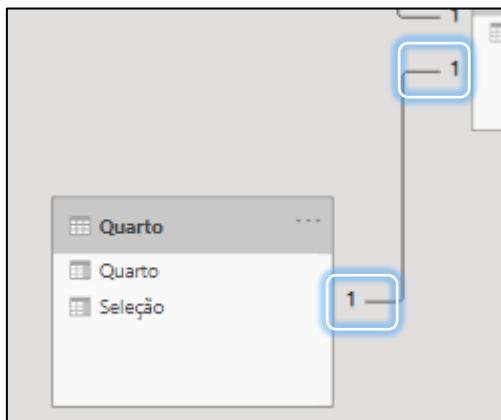


Erro de Cross-Filter

Durante o relacionamento, podemos clicar errado no fato e ocasionará um erro de “Cross Filter”. Veja que abaixo, a relação não foi feita corretamente e o ícone não foi relacionado.



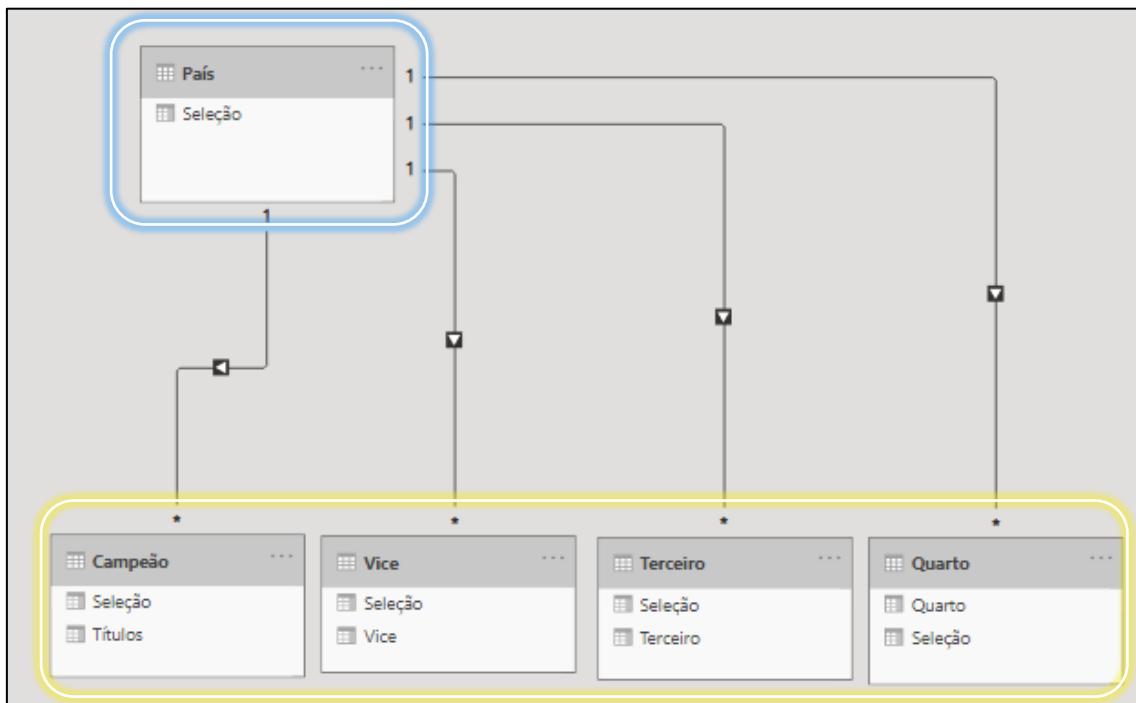
Veja que a relação está de um para um (1 – 1), isso causará erro.



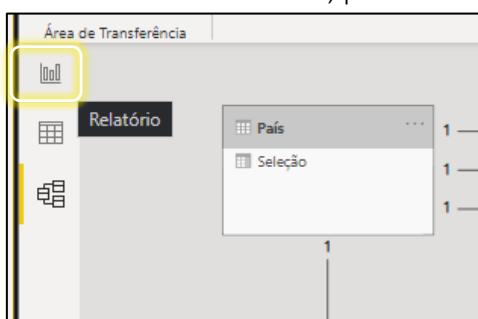
Para corrigir, é simples, basta clicar em cima da linha de relacionamento e Deletar > Excluir.

Correta Posição da Dimensão e dos Fatos no P.P

O mercado e os analistas estipularam um padrão de posicionamento para a Dimensão e os Fatos. A dimensão deve ficar no topo, enquanto os fatos logo abaixo.



Feito os relacionamentos, podemos retornar ao parte gráfica – Relatório.



Segmentação de Dados

Vamos criar um filtro de países, através de uma segmentação de dados.



Adicionaremos a Segmentação de Dados, primeiramente o país.

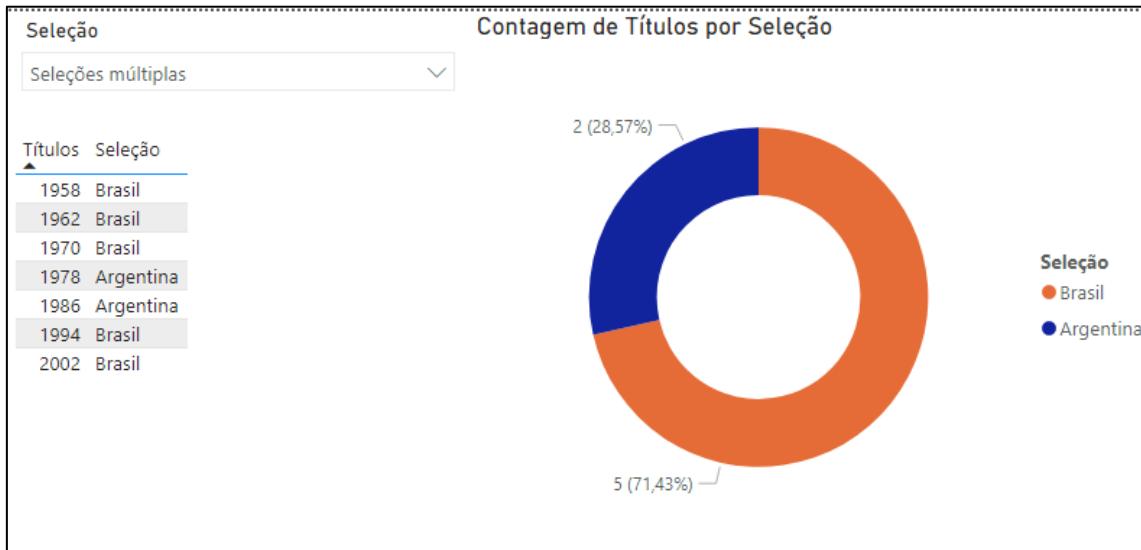
The screenshot shows the Power BI interface with the 'Filtros' (Filters) pane open. On the left, a dropdown menu titled 'Seleção' is open, displaying a list of countries: Alemanha[nota 3], Argentina, Áustria, Bélgica, Brasil, Bulgária, Chile, Coreia do Sul, Croácia, and Espanha. On the right, the 'Campos' (Fields) pane shows a hierarchy under 'Campeão': Seleção, Títulos, País, Quarto, Terceiro, and Vice. The 'Seleção' node under 'País' is highlighted with a yellow selection bar.

Após clique na seta direcional para baixo e selecione “Suspenso” para que a lista fique suspensa.

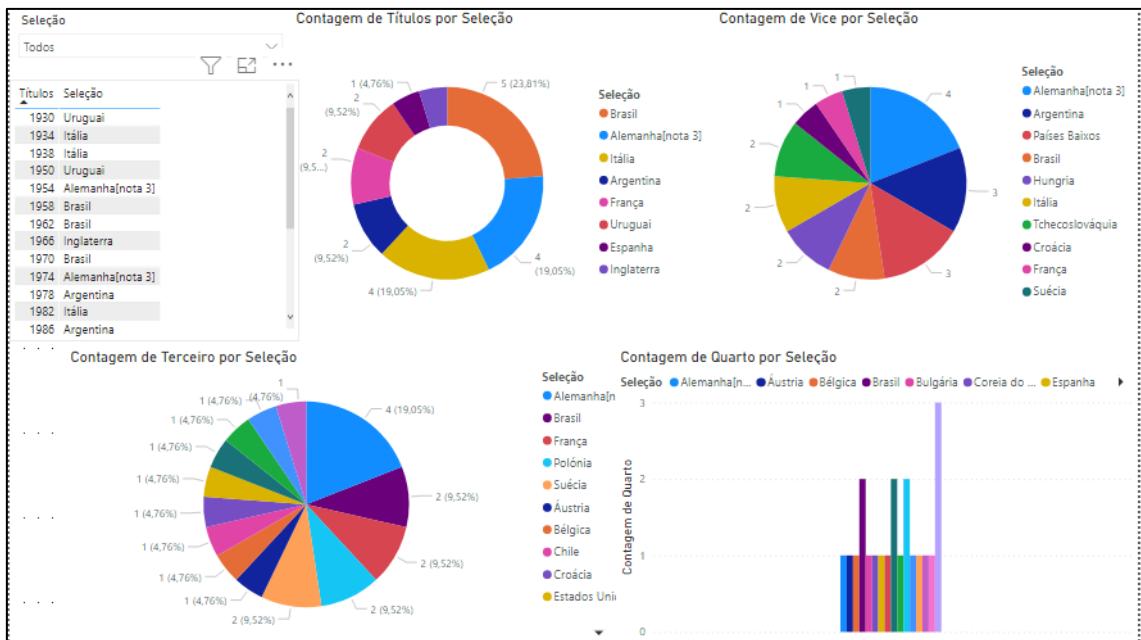


Seleções Múltiplas na Segmentação de Dados

Obs, isso é muito importante, ao criar uma segmentação de dados, para efeitos comparativos, clique e mantenha pressionado a tecla “CTRL” para habilitar a “Seleção Múltipla” e fazer vários comparativos.



O painel de relatório, portanto, ficou dessa forma.



Vamos salvar o projeto. E continuar na próxima aula, evoluindo.

Dados da Web – Estrutura NoSQL⁴

Vamos estabelecer novamente mais uma conexão de dados com a Web. No exemplo anterior, utilizamos o site da Wikipedia que já era provia tabelas em seu site, o que gerava uma certa tranquilidade para extraímos os dados dessa tabela.

Mas, e quando precisamos extraír dados de uma página Web que não possui tabelas prontas (puxando pela Tag <Table> do Html)? Por exemplo, o site de celulares da Amazon. O site da Amazon, utiliza um banco de dados NoSQL e os dados são apresentados e anunciados sem estrutura padronizada.

The screenshot shows a product listing on the Amazon website. At the top, there is a smartphone listed with a price of R\$699,00 and a rating of 4 stars from 60 reviews. Below it, another smartphone is listed with a price of R\$1,638,00 and a rating of 4 stars from 38 reviews. Both phones are shown from different angles, and there are options for delivery and payment methods like Prime.

A melhor forma de se tratar dados quando precisamos obter informações de sites que trabalham com banco de dados não relacionais (NoSQL), é identificar quais são as informações em comum. No site acima, os atributos semelhantes são os “nomes” e os “valores”.

Vamos estabelecer a conexão Web com o site da Amazon.

(https://www.amazon.com.br/s?k=celulares&__mk_pt_BR=%C3%85M%C3%85%C5%BD%C3%95%C3%91&ref=nb_sb_noss_1)

The screenshot shows a web scraping tool window titled "Da Web". It has a radio button for "Básico" selected. Below it, there is a "URL" input field containing the Amazon search URL from the previous step: "https://www.amazon.com.br/s?k=celulares&__mk_pt_BR=%C3%85M%C3%85%C5%BD%C3%95%C3%91&ref=nb_sb_noss_1".

⁴ NoSQ: Banco de Dados não relacionais

Às vezes, a melhor solução nesse caso, é utilizar a “Exibição da Web” e contar com a inteligência artificial do **Power BI**.

Pois se buscarmos em cada uma das tabelas que o **Power BI** encontrou, podemos esbarrar em alguns problemas.

The screenshot shows the Power BI interface with a 'Exibição de Tabela' (Table View) and 'Exibição da Web' (Web View) tab selected. On the left, there's a sidebar titled 'Opções de Exibição' (View Options) with a 'Tabelas Sugeridas [6]' section containing six table suggestions. On the right, a web browser window displays the Amazon website with the search term 'celulares'. The results page shows various mobile phone models and accessories. A sidebar on the right of the browser window provides details about the products.

Vamos clicar em “Adicionar a Tabela Usando Exemplos”.

This screenshot shows a modal dialog box titled 'Adicionar a Tabela Usando Exemplos'. Inside the dialog, there is a list of brand names: 'Xiaomi', 'LG', and 'SAMSUNG', each preceded by a selection checkbox. Below the list is a 'Selecionar' (Select) button.

Na forma abaixo, o Power vai mostrar os exemplos, e vamos poder selecionar do jeito que nós quisermos.

This screenshot shows the 'Adicionar a Tabela Usando Exemplos' dialog box again, but now the 'Xiaomi' checkbox is selected. The right side of the dialog shows a preview of the Amazon search results for 'celulares' with a focus on the 'Xiaomi' brand. At the bottom, there is a table with five rows labeled 1 through 5, corresponding to the selected example.

Vamos procurar primeiramente pelos dois modelos digitando o nome do “Xiaomi Redmi Note 9...” e caso de algum problema na visualização utilizar o “Xiaomi \$...” e após conseguir digitar, apague o (\$)

Às vezes pode ser que aconteçam alguns erros, tente várias e várias vezes até que a Inteligência artificial do **Power BI** reconheça os dados.

Coluna 1	Coluna 2
1 Celular Xiaomi Note 8 64GB Rom ...	R\$ 1.310,00
2 Celular Xiaomi Mi 10T 128GB 6GB ...	R\$ 3.549,97
3 Celular Xiaomi Redmi 9A Versão G...	
4 Xiaomi Redmi Note 9 128GB 4GB ...	R\$ 1.348,00
5 Celular Xiaomi Redmi 9i 128Gb 4G...	R\$ 1.126,00
6 Smartphone Xiaomi Redmi Note 9 ...	R\$ 1.539,99
7 Celular Xiaomi Redmi 9 64GB/4GB...	R\$ 1.064,89

Após isso, vamos em transformados dados.

Coluna 1	Coluna 2
1 Celular Xiaomi Note 8 64GB Rom 4GB Ram Dual Versión Global Space ...	R\$ 1.310,00
2 Celular Xiaomi Mi 10T 128GB 6GB Ram 5G Dual Sim Tela 6,67" Câmeras ...	R\$ 3.549,97
3 Celular Xiaomi Redmi 9A Versão Global 32gb / 2gb Ram/Tela 6.53" - Gr...	R\$ 809,98
4 Xiaomi Redmi Note 9 128GB 4GB RAM - Versión Global - Midnight Grey	R\$ 1.348,00
5 Celular Xiaomi Redmi 9i 128Gb 4Gb Ram Versão global - Sea Blue	R\$ 1.126,00
6 Smartphone Xiaomi Redmi Note 9S - 4GB + 64GB - Versão Global - Aur...	R\$ 1.539,99
7 Celular Xiaomi Redmi 9 64GB/4GB Dual Chip - Ocean Green	R\$ 1.064,89
8 Celular Xiaomi Mi 10T 128GB 6GB Ram 5G Dual Sim Tela 6,67" Câmera...	R\$ 3.576,00

Vamos Fechar e Aplicar.

Podemos aplicar mais dados também, por exemplo, selecionar notebooks de uma outra página de internet.

Veja que é possível continuar obtendo os dados.

Campos >

Pesquisar

Celulares

Coluna 1
Coluna 2

Notebooks

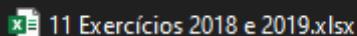
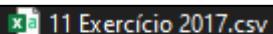
Coluna 1
Coluna 2

Coluna 1	Coluna 2
Celular Xiaomi Mi 10T 128GB 6GB Ram 5G Dual Sim Tela 6,67" Câmeras 64MP+13MP+5MP e 20MP - Black	R\$ 3.549,97
Celular Xiaomi Mi 10T 128GB 6GB Ram 5G Dual Sim Tela 6,67" Câmeras 64MP+13MP+5MP e 20MP - Lunar Silver	R\$ 3.529,90
Celular Xiaomi Note 8 128GB Rom 4GB Ram Dual - Versão Global - Space Black	R\$ 1.499,99
Celular Xiaomi Note 8 64GB Rom 4GB Ram Dual Versión Global Space Black	R\$ 1.310,00
Celular Xiaomi Poco X3 6GB/128GB NFC - Shadow Grey	R\$ 2.085,00
Celular Xiaomi Redmi 9 64GB/4GB Dual Chip - Carbon Grey	R\$ 1.043,73
Celular Xiaomi Redmi 9 64GB/4GB Dual Chip - Ocean Green	R\$ 1.064,89
Celular Xiaomi Redmi 9 Versão Global 64gb / 4gb Ram/Tela 6.53" - Cinza	R\$ 1.079,90
Celular Xiaomi Redmi 9A Versão Global 32gb / 2gb Ram/Tela 6.53" - Green	R\$ 809,98
Celular Xiaomi Redmi 9A Versão Global 32gb / 2gb Ram/Tela 6.53" - Preto	R\$ 791,30
Celular Xiaomi Redmi 9i 128Gb 4Gb Ram Versão global - Sea Blue	R\$ 1.126,00
Celular Xiaomi Redmi 9i 64Gb 4Gb Ram Versão global - Sea Blue	R\$ 1.029,90
CELULAR XIAOMI REDMI NOTE 8 DUAL 128GB NEPTUNE BLUE	R\$ 1.499,99
Celular Xiaomi Redmi Note 8 Pro 128 GB 6 GB Global Mineral Grey	R\$ 1.705,00
Celular Xiaomi Redmi Note 8 Pro Versão Global 128gb / 6gb Ram/Tela 6.53" - Verde	R\$ 1.768,00
Coluna 1	Coluna 2
HIDevolution Gigabyte Aero 17 KC-8US6150SH, 17,3" FHD 300Hz, 2,2 GHz i7-10870H, RTX 3060, 64 GB 3200MHz RAM, 1 TB PCIe SSD	R\$ 21.365,54
Notebook 2 em 1 Dell Inspiron i14-5491-A30S 10ª Geração Intel Core i7 8GB 256GB SSD Placa vídeo NVIDIA Full HD 14"	R\$ 6.693,64

Este exercício é apenas uma parte básica para o tratamento de dados. Pois é possível buscar cotação de dólar, e várias outros dados on-line.

Transposição e Mesclagem

Vamos utilizar os seguintes exemplos para tornar o exercício possível.



Note que vamos utilizar dados que tem a extensão diferentes. Um arquivo possui o formato .csv e os outros dois estão com a extensão .xlsx.

Veja o exercício 2017 .csv pelo bloco de notas:

```
Arquivo Editar Formatar Exibir Ajuda
2017;;
;;
Produtos;Ana Clara;João Nunes;Bruna Almeida
Poupança;R$345.114,21;R$224.985,63;R$223.985,50
Conta-Corrente;R$13.214.958,43;R$33.214.748,34;R$214.845,50
Produtos;R$454.214,29;R$465.214,22;R$954.214,23
```

Ele no Excel:

2017			
Produtos	Ana Clara	João Nunes	Bruna Almeida
Poupança	R\$ 345.114,21	R\$ 224.985,63	R\$ 223.985,50
Conta-Corrente	R\$ 13.214.958,43	R\$ 33.214.748,34	R\$ 214.845,50
Produtos	R\$ 454.214,29	R\$ 465.214,22	R\$ 954.214,23

O outro arquivo somente na extensão .xlsx:

Desempenho	Exercício 2018				Exercício 2019			
	Ana Clara	João Nunes	Bruna Almeida	Total 2018	Ana Clara	João Nunes	Bruna Almeida	Total 2019
Poupança	R\$552.148,85	R\$854.125,58	R\$558.456,54	R\$1.964.730,97	R\$551.478,36	R\$7.948.840,70	R\$7.385.632,60	R\$15.885.951,66
Conta-Corrente	R\$488.524,86	R\$441.575,63	R\$211.475,66	R\$1.141.576,15	R\$921.458,36	R\$11.562.251,84	R\$314.785,36	R\$12.798.495,56
Produtos	R\$55.254,87	R\$14.185,36	R\$48.726,22	R\$118.166,45	R\$33.698,25	R\$55.482,36	R\$195.412,36	R\$284.592,97
Total	R\$1.095.928,58	R\$1.309.886,57	R\$818.658,42	R\$3.224.473,57	R\$1.506.634,97	R\$19.566.574,90	R\$7.895.830,32	R\$28.969.040,19

E como juntar os arquivos .csv com os arquivos .xlsx?

Vamos primeiramente anexar o arquivo .csv. Após clicar em transformar dados, vamos para a parte de ETL (**Power Query**).

The screenshot shows the Power Query Editor interface. On the left, there's a 'Consultas [1]' pane with a single item named '11 Exercício 2017'. The main area displays a table with four columns: Column1, Column2, Column3, and Column4. The data consists of six rows: Row 1 contains '2017'; Rows 2 through 6 contain grouped data for 'Produtos', 'Poupança', 'Conta-Corrente', and 'Produtos' respectively, with values like 'R\$345.114,21' and 'R\$ 224.985,63'. The 'Column1' header is highlighted in yellow, indicating it's the current column being edited.

Primeira coisa, precisamos remover as duas linhas superiores (a que contém o ano e a vazia). Lembre-se de ir em “Remover Linhas > Remover Linhas Principais”.

Sem “Lixo” agora.

A ^B C Column1	A ^B C Column2	A ^B C Column3	A ^B C Column4
1 Produtos	Ana Clara	João Nunes	Bruna Almeida
2 Poupança	R\$345.114,21	R\$224.985,63	R\$223.985,50
3 Conta-Corrente	R\$13.214.958,43	R\$33.214.748,34	R\$214.845,50
4 Produtos	R\$454.214,29	R\$465.214,22	R\$954.214,23

Vamos promover a primeira linha como cabeçalho.

A ^B C Produtos	\$ Ana Clara	\$ João Nunes	\$ Bruna Almeida
1 Poupança	345.114,21	224.985,63	223.985,50
2 Conta-Corrente	13.214.958,43	33.214.748,34	214.845,50
3 Produtos	454.214,29	465.214,22	954.214,23

Agora precisamos transformar esses vendedores de Colunas para Linhas, certo?

Selecione todas as três colunas acima e clique com o [“Botão Direito > Transformar Colunas em Linhas”](#).

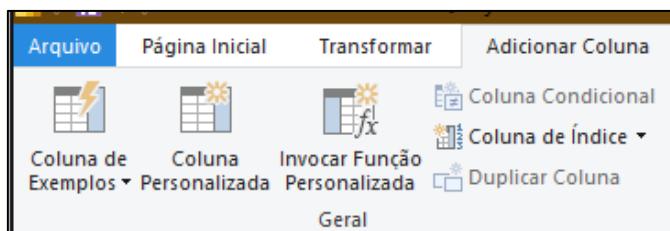
Veja o resultado abaixo.

A ^B C Produtos	A ^B C Atributo	\$ Valor
1 Poupança	Ana Clara	345.114,21
1 Poupança	João Nunes	224.985,63
1 Poupança	Bruna Almeida	223.985,50
2 Conta-Corrente	Ana Clara	13.214.958,43
2 Conta-Corrente	João Nunes	33.214.748,34
2 Conta-Corrente	Bruna Almeida	214.845,50
3 Produtos	Ana Clara	454.214,29
3 Produtos	João Nunes	465.214,22
3 Produtos	Bruna Almeida	954.214,23

Na próxima página, vamos adicionar uma coluna personalizada.

Adicionando uma Coluna Personalizada

Para adicionar uma coluna personalizada, vamos em “[Adicionar Coluna > Coluna Personalizada](#)”.



Essa coluna é adicionada baseada em uma fórmula que será escrita no quadro abaixo.

Coluna Personalizada

Adicionar uma coluna computada das outras colunas.

Nome da nova coluna
Personalizar

Fórmula de coluna personalizada ⓘ
=

Colunas disponíveis
Produtos
Atributo
Valor

No lugar de “Personalizar”, vamos substituir esse campo do nome para “Ano” e a coluna conterá somente o ano de “2017” (Para esse exemplo, tratamos desse jeito! Mas essa coluna pode ser tratada de diversas formas e ter inúmeras possibilidades).

Coluna Personalizada

Adicionar uma coluna computada das outras colunas.

Nome da nova coluna
Ano

Fórmula de coluna personalizada ⓘ
= 2017

Veja que ao clicarmos em Ok, uma nova coluna foi gerada.

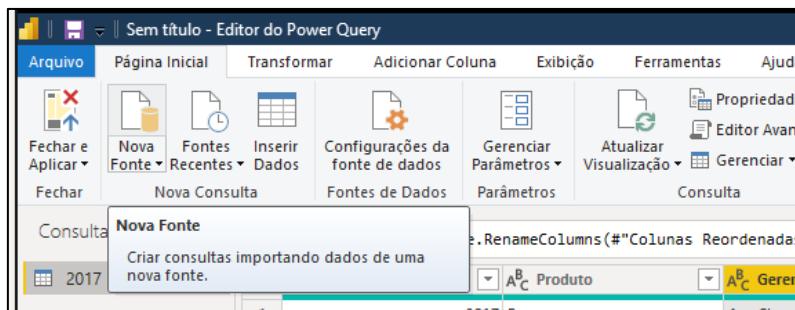
Atributo	Valor	Ano
Ana Clara	345.114,21	2017
João Nunes	224.985,63	2017
Bruna Almeida	223.985,50	2017

O próximo passo é renomear as colunas de acordo e arrastar a coluna de “Ano” para o primeiro lugar das colunas.

E se lembra que tratávamos do arquivo **.csv**, certo? Agora vamos importar as demais planilhas de **.xlsx**.

Importando mais dados dentro do Power Query

Não é preciso sair do **Power Query** para realizar mais importações de dados. Podemos obter mais dados ainda dentro do PQ, através de: “Página Inicial > Nova Fonte”.



Para a continuação desse exercício será o “11 – Exercícios 2018 e 2019.xlsx”. Repare a quantidade de “lixo” que apresenta.

Consultas [2]						
	X	V	fx	= Table.TransformColumnTypes(#"Cabeçalhos Promovidos",{{"Desempenho", type text}, {"Column2", type any}, {"Column3", type any}, {"Column4", type any}, {"Column5", type any}, {"Column6", type any}, {"Column7", type any}})		
2017				A ^B Desempenho	ABC 123 Column2	ABC 123 Column3
Exercícios 2018 e 2019					ABC 123 Column4	ABC 123 Column5
					ABC 123 Column6	ABC 123 Column7
1				null	null	null
2	Produtos / Gerentes			Exercício 2018		null
3				null	Ana Clara	João Nunes
4	Poupança				552148,85	854125,58
5	Conta-Corrente				488524,86	441575,63
6	Produtos				55254,87	14185,36
7	Total				1095928,58	1309886,57
8						
9						
10						
11						
12						
13						
14						

Precisamos deixar essa consulta exatamente igual a consulta que realizamos para o ano de 2017.

Cuidado no passo abaixo!

Para limpar esses valores nulos, é comum irmos na primeira coluna e “Remover os vazios”. Porém, esse ato irá sumir com os cabeçalhos dos gerentes que precisamos. Veja o quadro azul brilhante, não poderíamos perder esses dados, e se por acaso removêssemos os vazios logo na primeira coluna, nós os perderíamos.

A ^B Desempenho	ABC 123 Column2	ABC 123 Column3
null	null	null
Produtos / Gerentes	Exercício 2018	null
	Ana Clara	João Nunes
Poupança	552148,85	854125,58
Conta-Corrente	488524,86	441575,63
Produtos	55254,87	14185,36
Total	1095928,58	1309886,57
	null	null
	null	null

Veja o que acontece se removêssemos os vazios na primeira coluna, o erro.

A ^B C Desempenho	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4
Produtos / Gerentes	Exercício 2018	null	null
Poupança	552148,85	854125,58	558456,54
Conta-Corrente	488524,86	441575,63	211475,66
Produtos	55254,87	14185,36	48726,22
Total	1095928,58	1309886,57	818658,42

Onde estão os gerentes? Foram perdidos. Portanto, essa não é a forma correta.

Precisamos, portanto, procurar qual coluna não existe a informação cruzada.

A coluna 2, antes dessa remoção é a ideal para poder realizar a remoção dos vazios.

A ^B C Desempenho	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4
1 Produtos / Gerentes	Exercício 2018	null	null
2 null	Ana Clara	João Nunes	Bruna Almeida
3 Poupança	552148,85	854125,58	558456,54
4 Conta-Corrente	488524,86	441575,63	211475,66
5 Produtos	55254,87	14185,36	48726,22
6 Total	1095928,58	1309886,57	818658,42

Lembre-se de remover os valores totais, não são necessários! E, também verifique se existem colunas com valores nulos. E, também a linha inferior que apresenta o “Total”.

ABC 123 Column5	ABC 123 Column9	ABC 123 Column10
1 null	null	null
2 Total 2018		
3 1964730,97	15885951,66	null
4 1141576,15	12798495,56	null
5 118166,45	284592,97	null
6 3224473,57	28969040,19	null
6 Total	1309886,57	818658,42

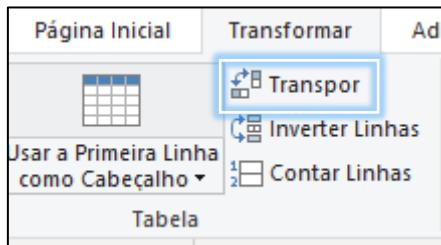
Transposição de Tabela

Podemos “Inverter” a tabela através da “transposição” da tabela, invertê-la.

Veja que queremos que os gerentes vão para baixo, um a um. Mas não podemos utilizar a “Transformar colunas em linhas” nesse caso, pois se tornará uma bagunça.

A B C Desempenho	ABC 123 Column3	ABC 123 Column4	ABC 123 Column6	A 1
Produtos / Gerentes	null	null	Exercício 2019	
	João Nunes	Bruna Almeida	Ana Clara	J
Poupança	854125,58	558456,54	551478,36	
Conta-Corrente	441575,63	211475,66	921458,36	
Produtos	14185,36	48726,22	33698,25	

Vamos transpor essa tabela.



Veja como a tabela está mais “amigável”.

A B C Column1	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4	ABC 123 Column5
Produtos / Gerentes	null	Poupança	Conta-Corrente	Produtos
	João Nunes	854125,58	441575,63	
	Bruna Almeida	558456,54	211475,66	
Exercício 2019	Ana Clara	551478,36	921458,36	
	João Nunes	7948840,7	11562251,84	
	Bruna Almeida	7385632,6	314785,36	

Vamos promover os cabeçalhos agora e extrair os anos corretamente (extraíndo pelo delimitador). Esse é um passo importante, “Transformar > Extrair > Após o Delimitador > Dê um espaço” (vamos deixar permanecer o que está após).

The screenshot shows the Power BI ribbon with the 'Transformar' tab selected. In the 'Extrair' (Extract) section of the ribbon, the 'Após o Delimitador' (After Delimiter) dialog is open. The 'Delimitador' (Delimiter) field contains a space character (' '). The main area shows a table with columns labeled 'A B C Produtos / Gerentes' and rows containing data like 'Exercício 2018', 'null', 'null', etc.

Valores Nulo x Valores Branco

Para o Power BI existe uma diferença entre Nulo e Branco.

	A	B	C	Ano
1				2018
2				
3				
4				2019
5				
6				

O correto seria utilizarmos o “**Preenchimento para baixo**”, correto? Sim e Não!! Porque mesmo realizando esse passo, o Power BI está reconhecendo essas linhas abaixo de cada ano como vazias, o que é diferente de Nulo!

Para corrigirmos isso, primeiro precisamos “Substituir os valores” de vazios para “**Null**”.

Substituir Valores

Substitua um valor por outro nas colunas selecionadas

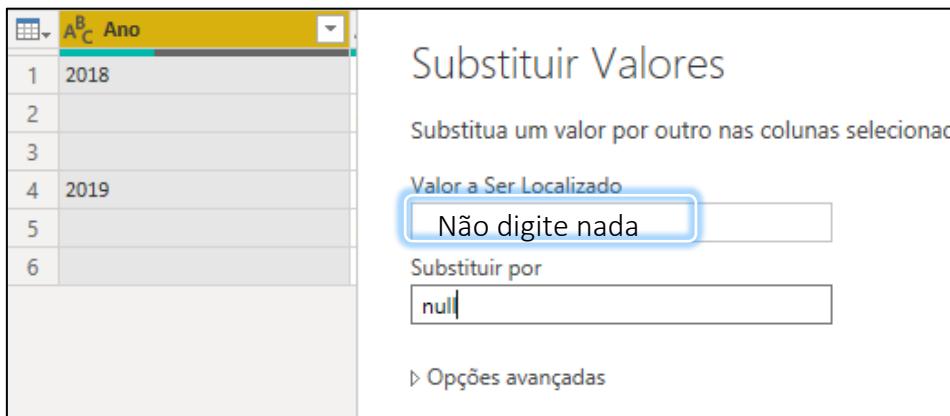
Valor a Ser Localizado

Não digite nada

Substituir por

null

Opções avançadas



	A	B	C	Ano
1				2018
2				null
3				null
4				2019
5				null
6				null

Aplicando o “Preenchimento Para Baixo”.

	A	B	C	Ano
1				2018
2				2018
3				2018
4				2019
5				2019
6				2019

Agora ficou fácil! Selecionaremos as 3 colunas que desejamos transformar em linhas.

1.2 Poupança	1.2 Conta-Corrente	1.2 Produtos
552148,85	488524,86	55254,87
854125,58	441575,63	14185,36
558456,54	211475,66	48726,22
551478,36	921458,36	33698,25
7948840,7	11562251,84	55482,36
7385632,6	314785,36	195412,36

"Shift > de Poupança até Produtos > Botão Direito > Transformar Colunas em Linhas"

	A ^B _C Ano	A ^B _C Gerente	A ^B _C Atributo	1.2 Valor
1	2018	Ana Clara	Poupança	552148,85
2	2018	Ana Clara	Conta-Corrente	488524,86
3	2018	Ana Clara	Produtos	55254,87
4	2018	João Nunes	Poupança	854125,58
5	2018	João Nunes	Conta-Corrente	441575,63
6	2018	João Nunes	Produtos	14185,36
7	2018	Bruna Almeida	Poupança	558456,54

Precisa estar no mesmo formato da de 2017, ou seja, precisamos alterar o nome de "Atributo" para "Produto". E checar se a tipagem está correta!!!!

2017:

	1 ² ₃ Ano	A ^B _C Gerente	A ^B _C Produto	1.2 Valor
1	2017	Ana Clara	Poupança	345114,21
2	2017	João Nunes	Poupança	224985,63

2018 e 2019:

	1 ² ₃ Ano	A ^B _C Gerente	A ^B _C Produto	1.2 Valor
1	2018	Ana Clara	Poupança	552148,85
2	2018	Ana Clara	Conta-Corrente	488524,86

Mas, precisamos juntar as duas consultas. Como fazer isso?

Consultas [2]

- 2017
- Exercícios 2018 e 2...

Acrescentar Consultas

Precisamos “unir” essas duas tabelas em uma única. Como faremos isso?

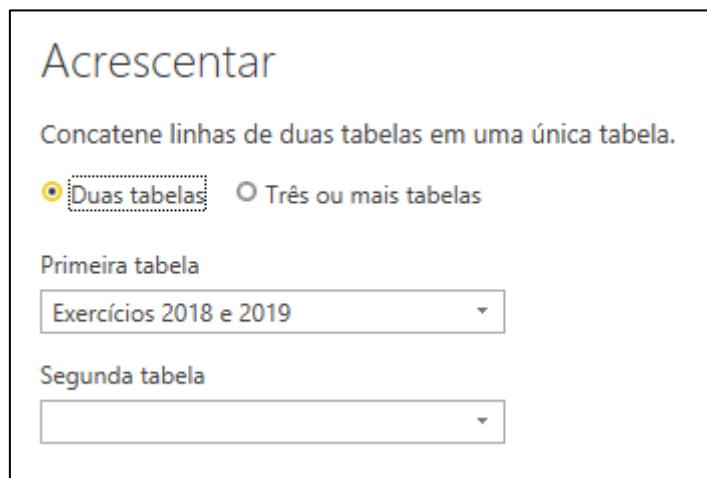
Na “[Página Inicial > Acrescentar Consultas > Acrescentar Consultas como Novas](#)”.



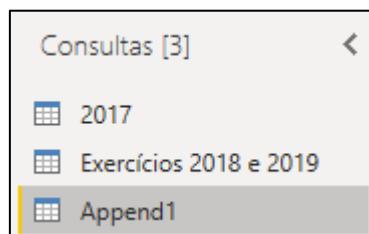
Ao clicar, o PB perguntará quantas tabelas desejamos acrescentar.

Append de Colunas

Esse próximo passo cai em prova de certificação até. Veja que o PB já está selecionando a consulta atual como a primeira tabela, agora precisamos realizar o “Append” para a consulta de 2017.



Note que gerou uma consulta nova com o nome de “Append1” ou outro nome se for em português.



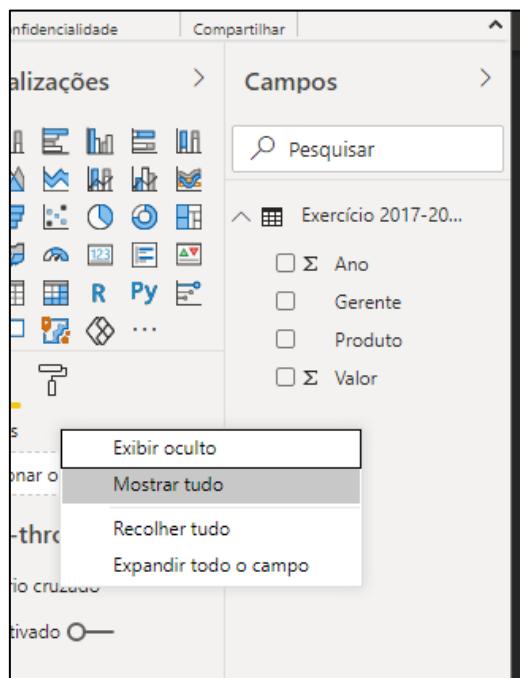
Claro, podemos renomear essa consulta. [Não esquecer de TIRAR A CARGA](#) ! Porque só vamos carregar a “Append1”, a de 2017 e a de 2018/2019, não serão mais necessárias!!

Feche e Aplique.

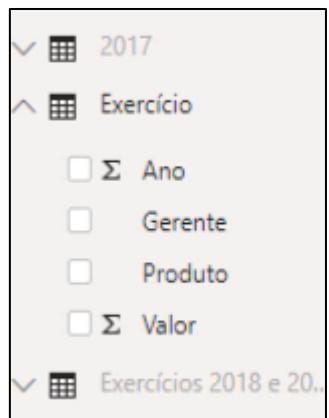
Mostrando Oculto

Há muitos analistas que não retiram a carga das consultas. E quando voltam ao Power View, eles apenas as escondem. Isso gera comprometimento da performance. Por isso, caso a performance esteja baixa, verifique se elas estão ocultas no View.

Clique com o botão direito em qualquer lugar dos Campos > Mostrar Tudo.



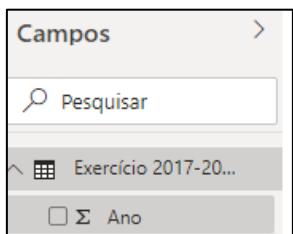
Se elas estiverem dessa forma, é porque elas estão com as cargas habilitadas, mas, ocultas.



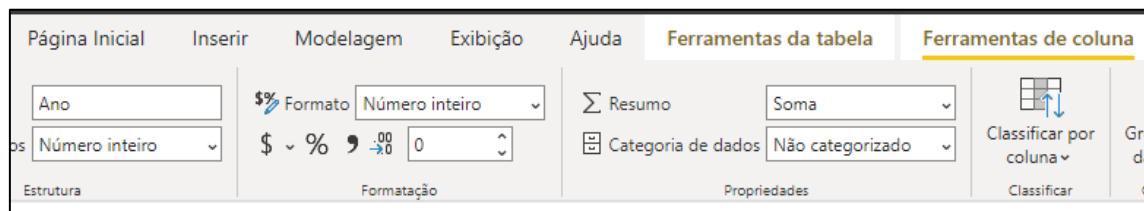
Análise Visual com Tempo

Vamos continuar o exemplo anterior.

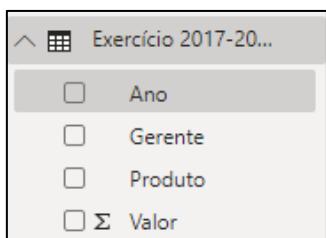
Apenas para começar, veja que o campo do “Ano” apresenta o sigma de somatória. E ano não deveria realizar uma somatória. Mas deveria continuar como número. Para corrigir isso, clique e deixe marcado em cinza.



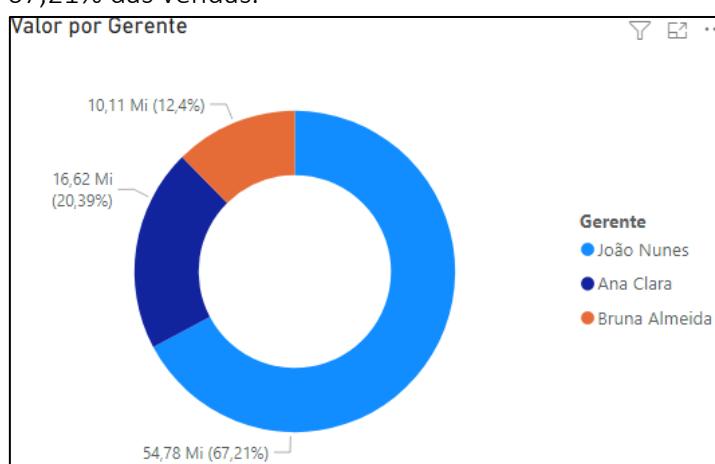
Procure por “ Σ ” e substitua o resumo de “Soma” para “Não Resumir”.



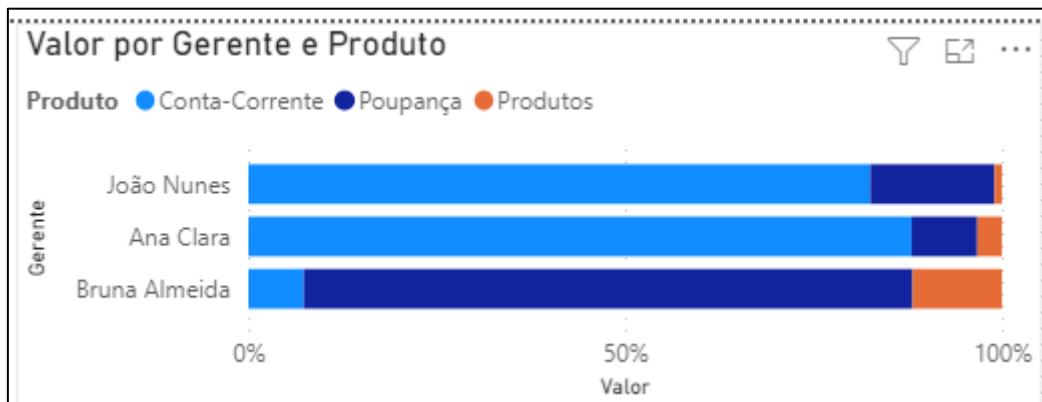
Veja que agora o ano está como número mas em formatação de texto, enquanto os valores continuam como números.



Podemos adicionar um gráfico de rosca agora e procurar saber qual valor de venda que cada gerente fez e “qual foi o melhor”. Veja que o melhor vendedor foi o “João” com 67,21% das vendas.



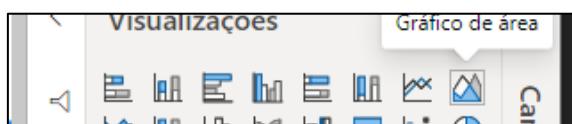
Mas e se quiséssemos visualizar quais foram os produtos que o João mais vendeu? Com um gráfico de barras.



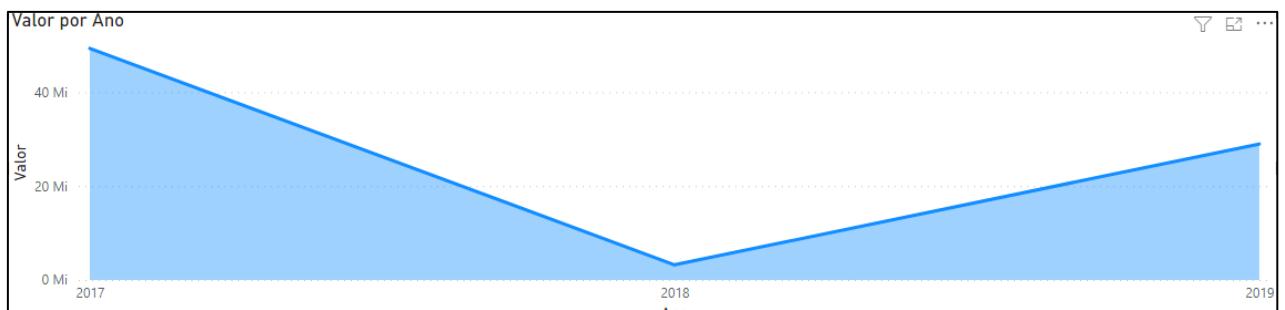
Mas, se quisermos ver ao longo do tempo, é sempre com o Gráfico de Área. São sinônimos.

Gráfico de Área – Visual com Tempo

Sempre que se for trabalhar com o tempo, se utiliza o **Gráfico De Área** ou de Linha.



Analizando o gráfico, podemos entender que em 2017 foi o melhor ano, e que em 2018, houve uma queda imensa e em 2019, os valores voltaram a subir.

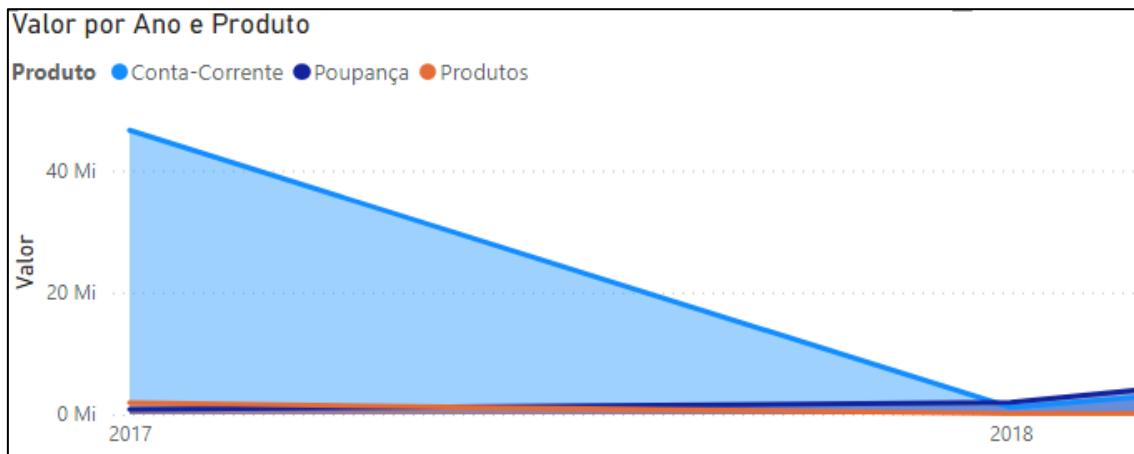


Não é possível realizar a arealização do *drilldown* pois os dados são apenas em Anos!

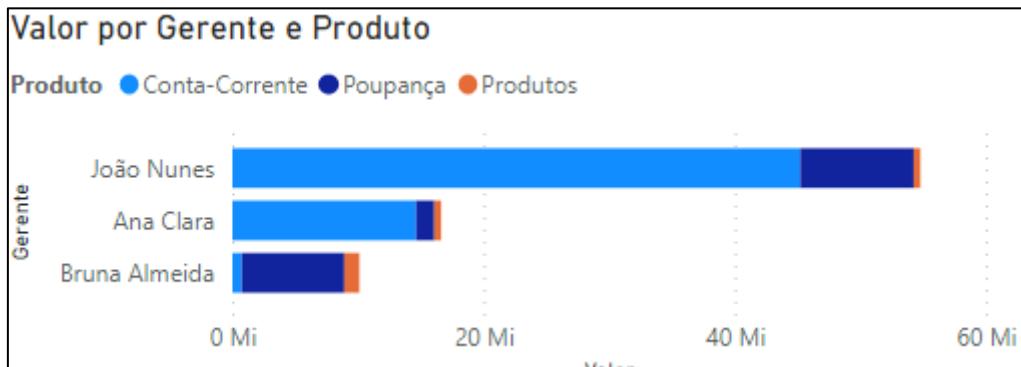
Na página 152, vamos analisar os vendedores e as suas vendas.

Podemos analisar o seguinte.

Produtos é o pior “ativo financeiro/produto” vendido, é o que menos rende dentre os três.



E podemos analisar que a Bruna é a pior vendedora entre os três.

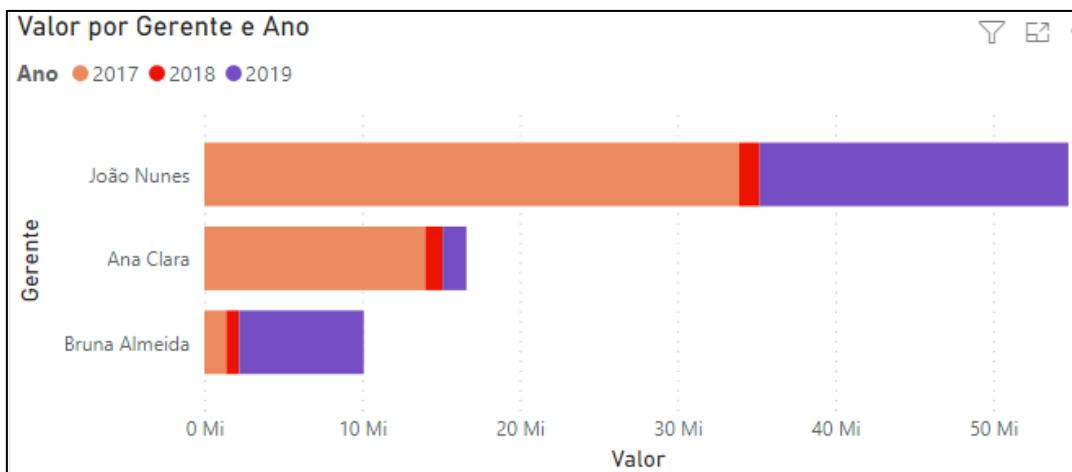


Mas, veja que dentre as suas vendas, ela é a que mais vende ou a que melhor vende o ativo de “Produtos”. Nenhum vendedor foi melhor que ela nesse ponto de vendas.

E se quiséssemos saber se os vendedores estão crescendo? Se eles estão melhorando? Poderíamos utilizar o gráfico de Barras.

Gráfico de Barras

Gráfico de barras é sinônimo de entendimento de aumento ou queda de lucratividade, rentabilidade, crescimento ou desenvolvimento e seus antônimos.



Veja que a Bruna, dada como a pior vendedora por outros gráficos, em 2019 ela vem crescendo e desenvolvendo muito mais e muito melhor que a Ana Clara, todos os vendedores em 2018 tiveram perdas e prejuízos. Porém, a Ana Clara não conseguiu se recuperar desse impacto de 2018. Enquanto o João Nunes e a Bruna Almeida, conseguiram reverter o quadro e a Bruna Almeida está recuperando.

NUNCA JULGUE POR UM ÚNICO GRÁFICO, ESTUDE, CRIE VÁRIOS GRÁFICOS, FAÇA ESTUDOS E COMPARE RESULTADOS.

Camada Binária e Camada de Dados

Existem duas camadas entre o servidor de dados e o servidor Power BI.

A camada de **Infraestrutura** – Nome/Tamanho/Compt/Extensão do arquivo (Camada **Binária**)

A camada de **Dados** – Linhas/ Colunas (Camada de **Dados**).

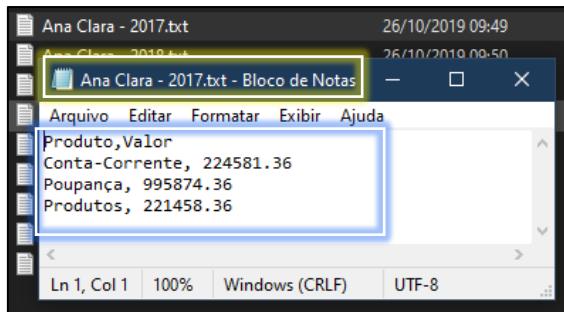
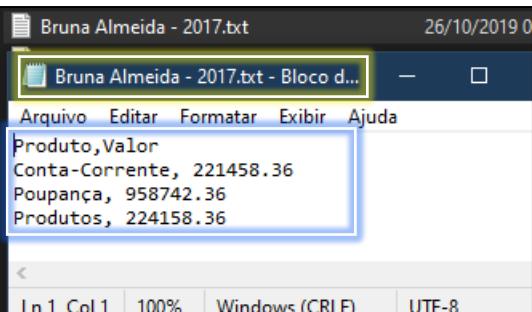
Em alguns momentos, podemos não receber nenhuma data nos arquivos, mas podemos utilizar os dados da camada de Infraestrutura (Binária) se transformando em uma Tabela na camada de Dados.

Utilizaremos a pasta 12 – Movimento

Nome	Data de modificação	Tipo
Ana Clara - 2017.txt	26/10/2019 09:49	Documento de Te...
Ana Clara - 2018.txt	26/10/2019 09:50	Documento de Te...
Ana Clara - 2019.txt	26/10/2019 09:50	Documento de Te...
Bruna Almeida - 2017.txt	26/10/2019 09:50	Documento de Te...
Bruna Almeida - 2018.txt	26/10/2019 09:51	Documento de Te...
Bruna Almeida - 2019.txt	26/10/2019 09:51	Documento de Te...
Joao Nunes - 2017.txt	26/10/2019 09:47	Documento de Te...
Joao Nunes - 2018.txt	26/10/2019 09:51	Documento de Te...
Joao Nunes - 2019.txt	26/10/2019 09:51	Documento de Te...

São os gerentes do exemplo anterior.

Essas são as informações que temos sobre a Ana Clara e Bruna Almeida.

	
---	--

O problema aqui é, teremos os dados e as informações. Mas, repare que o NOME de cada Gerente e a DATA do ano, estão no nome do arquivo, ou seja, na camada de **INFRAESTRUTURA (Na Camada Binária)**. Ao conectarmos esses dados somente pela camada de dados, nós perderemos os dados e não saberemos de quem são os valores e nem quais os anos que os fatos aconteceram.

Em Amarelo → Camada INFRAESTRUTURA (Binária)

Em Azul → Camada de DADOS

Precisamos utilizar as informações da Binária nos Dados.

Vamos conectar a pasta do Movimentos (contendo as **.txt** de 2017 e 2018, removemos a de 2019 para fazermos o teste do *update*).

D:\workspace\curso_MasterPowerBI\Datasets\12 - Movimento						
Content	Name	Extension	Date accessed	Date modified	Date created	Attributes
Binary	Ana Clara - 2017.txt	.txt	20/03/2021 00:38:44	26/10/2019 09:49:38	23/02/2021 02:47:18	Record
Binary	Ana Clara - 2018.txt	.txt	08/03/2021 11:18:47	26/10/2019 09:50:48	23/02/2021 02:47:18	Record
Binary	Bruna Almeida - 2017.txt	.txt	20/03/2021 00:40:04	26/10/2019 09:50:00	23/02/2021 02:47:18	Record
Binary	Bruna Almeida - 2018.txt	.txt	08/03/2021 11:18:47	26/10/2019 09:51:06	23/02/2021 02:47:18	Record
Binary	Joao Nunes - 2017.txt	.txt	08/03/2021 11:18:47	26/10/2019 09:47:08	23/02/2021 02:47:18	Record
Binary	Joao Nunes - 2018.txt	.txt	08/03/2021 11:18:47	26/10/2019 09:51:30	23/02/2021 02:47:18	Record

Note que estamos claramente evidenciados na camada Binária. Clique em “Transformar Dados”.

Clique na célula do conteúdo e veja que são os arquivos.

Content
1 Binary
2 Binary
3 Binary
4 Binary
5 Binary
6 Binary

Content
1 Binary
2 Binary
3 Binary
4 Binary
5 Binary
6 Binary

Content
1 Binary
2 Binary
3 Binary
4 Binary
5 Binary
6 Binary

Ana Clara - 2017.txt 85 bytes
Bruna Almeida - 2018.txt 85 bytes
Joao Nunes - 2018.txt 85 bytes

Porém, além da coluna do conteúdo (**Content**), também precisaremos da coluna ao lado, a coluna de “**Name**”.

D:\workspace\curso_Master		
Content	Name	Extension
Binary	Ana Clara - 2017.txt	.txt
Binary	Ana Clara - 2018.txt	.txt
Binary	Bruna Almeida - 2017.txt	.txt
Binary	Bruna Almeida - 2018.txt	.txt
Binary	Joao Nunes - 2017.txt	.txt
Binary	Joao Nunes - 2018.txt	.txt

Selecione as duas colunas (Content e Name), remova todas as outras colunas.

Content	Name
1 Binary	Ana Clara - 2017.txt
2 Binary	Ana Clara - 2018.txt

Vamos duplicar essa consulta e chamar de Auxiliar.

Para preservar a conexão do Power BI com a pasta, vamos duplicar a consulta. A consulta duplicada. É nessa Auxiliar que vamos trabalhar.

	Content	Name
1	Binary	Ana Clara - 2017.txt
2	Binary	Ana Clara - 2018.txt
3	Binary	Bruna Almeida - 2017.txt
4	Binary	Bruna Almeida - 2018.txt
5	Binary	Joao Nunes - 2017.txt
6	Binary	Joao Nunes - 2018.txt

Normalmente, clicaríamos diretamente nas duas setas direcionais para baixo para combinarmos os arquivos. Mas como já vimos que se fizéssemos isso, perderíamos as informações dos gerentes e os anos desses dados, pois precisamos dessas informações.

Clique exatamente no “Binary” da primeira linha para entrar no arquivo binário.

	Produto	Valor
1	Conta-Corrente	22458136
2	PoupanÃ§a	99587436
3	Produtos	22145836

O Power BI foi da camada de Infra para a camada de Dados, o PB vai entrar nos dados da Auxiliar.

Clicando na consulta Auxiliar, procure ao lado direito as “Etapas Aplicadas”. Clique em “CSV Importado”. Esse código da Linguagem M é o que viemos buscar.

Veja que buscamos o código da linguagem M que está na barra de fórmulas nesse exato ponto/momento da Auxiliar em *CSV Importado*.

```
= Csv.Document(Content,[Delimiter=",", Columns=2, Encoding=1252, QuoteStyle=QuoteStyle.None])
```

Note que o “Encoding=1252”, deverá ser substituído pelo “Encoding=65001”, que é o charset da utf-8, não na auxiliar e sim na Movimento.

Vamos copiar esse código da Linguagem M e vamos voltar a consulta de Movimentos.

```
= Csv.Document(Content,[Delimiter=",", Columns=2, Encoding=1252,  
QuoteStyle=QuoteStyle.None])
```

Agora precisamos da tabela de dados.

Vamos em “Adicionar Coluna > Coluna Personalizada”.



No quadro que permite que adicionemos a fórmula, vamos escrever o código que foi copiado da consulta auxiliar.

Coluna Personalizada

Adicionar uma coluna computada das outras colunas.

Nome da nova coluna
Personalizar

Fórmula de coluna personalizada ⓘ
=

Pontos de Atenção

Nome da nova coluna
Personalizar

Fórmula de coluna personalizada ⓘ
= Csv.Document([Content],[Delimiter=",", Columns=2,
Encoding=1252, QuoteStyle=QuoteStyle.None])

Colunas disponíveis
Content
Name

É preciso alterar o “Content”(que faz referência ao Content da consulta auxiliar) que está escrito pela “Content” das Colunas disponíveis apagando o “Content” e clicando duas vezes ao lado. E alterar o “Encoding=1252”, para a collection correta, que é a utf-8 (65001).

Nome da nova coluna
Personalizar

Fórmula de coluna personalizada ⓘ
= Csv.Document([Content],[Delimiter=",", Columns=2,
Encoding=65001, QuoteStyle=QuoteStyle.None])

Colunas disponíveis
Content
Name

Esses passos devem ser seguidos à risca.

E portanto, o **Power Query** acaba de criar a Coluna Table. Que são os dados que estavam contidos na Camada de Dados e foram buscados até a Camada de Infra.

	Content	Name	Personalizar
1	Binary	Ana Clara - 2017.txt	Table
2	Binary	Ana Clara - 2018.txt	Table
3	Binary	Bruna Almeida - 2017.txt	Table
4	Binary	Bruna Almeida - 2018.txt	Table
5	Binary	Joao Nunes - 2017.txt	Table
6	Binary	Joao Nunes - 2018.txt	Table

Não apague a coluna dos “Names”. Mas, apague a coluna do “Content”, porque não será mais necessário combinar os arquivos binários, uma vez que eles já foram convertidos em uma tabela.

	Content	Name	Personalizar
1	Binary	Ana Clara - 2017.txt	Table
2	Binary	Ana Clara - 2018.txt	Table
3	Binary	Bruna Almeida - 2017.txt	Table
4	Binary	Bruna Almeida - 2018.txt	Table
5	Binary	Joao Nunes - 2017.txt	Table
6	Binary	Joao Nunes - 2018.txt	Table

Column1	Column2
Produto	Valor
Conta-Corrente	224581.36
Poupança	995874.36
Produtos	221458.36

Veja que eu ao clicarmos na célula da primeira linha de **Table**, ela nos mostrará os dados que estavam na Camada de Dados. Não sendo mais necessário entrar e combinar os arquivos (nesse exemplo), podemos, portanto, excluir a primeira coluna dos arquivos binários agora. E clique ao lado do Personalizar, onde existem duas setas direcionais para os lados.

	Name	Personalizar
1	Ana Clara - 2017.txt	Table

	Name	Column1	Column2
1	Ana Clara - 2017.txt	Produto	Valor
2	Ana Clara - 2017.txt	Conta-Corrente	224581.36
3	Ana Clara - 2017.txt	Poupança	995874.36
4	Ana Clara - 2017.txt	Produtos	221458.36

Veja que todos os dados estão disponíveis para o manuseio e tratamento.

Agora todas as linhas estão expandidas.

Veja com muito cuidado que existem linhas e informações que são “lixos”.

Como por exemplo as linhas que contêm “Produto e Valor”.

	Name	Column1	Column2
1	Ana Clara - 2017.txt	Produto	Valor
2	Ana Clara - 2017.txt	Conta-Corrente	224581.36
3	Ana Clara - 2017.txt	Poupança	995874.36
4	Ana Clara - 2017.txt	Produtos	221458.36
5	Ana Clara - 2018.txt	Produto	Valor
6	Ana Clara - 2018.txt	Conta-Corrente	554125.36
7	Ana Clara - 2018.txt	Poupança	336524.35
8	Ana Clara - 2018.txt	Produtos	221485.36
9	Bruna Almeida - 2017.txt	Produto	Valor
10	Bruna Almeida - 2017.txt	Conta-Corrente	221458.36
11	Bruna Almeida - 2017.txt	Poupança	958742.36
12	Bruna Almeida - 2017.txt	Produtos	224158.36
13	Bruna Almeida - 2018.txt	Produto	Valor

Como existe um padrão, podemos aplicar um “Filtro de Texto” e remover. Mas, com cuidado, pois note que existe um “Produto\$” que esse é necessário pois contém valores, é um “ativo financeiro de venda”.

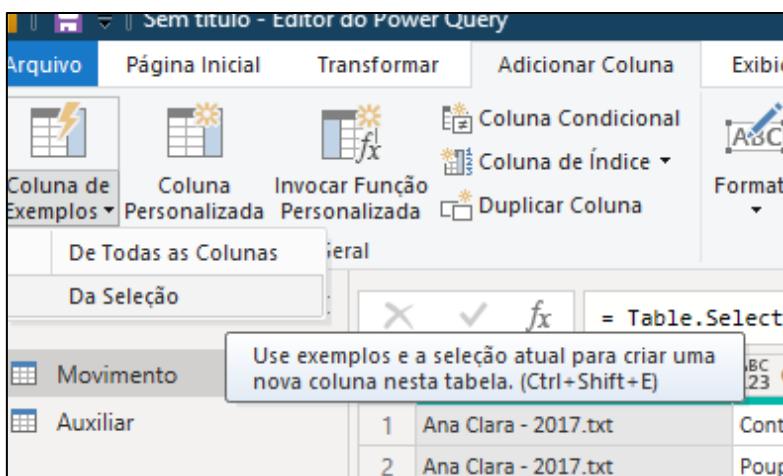
The screenshot shows the 'Filter' dialog for the 'Column1' column. The left pane lists filter options: 'Classificar em Ordem Crescente', 'Classificar em Ordem Decrescente', 'Limpar Classificação', 'Limpar Filtro', 'Remover Vazio', and 'Filtros de Texto'. The 'Filtros de Texto' section contains a 'Pesquisar' input field and a list of checked items: '(Selecionar Tudo)', 'Conta-Corrente', 'Poupança', 'Produto', and 'Produtos'. Below the dialog is a table preview showing rows 17 and 18. Row 17 is 'Joao Nunes - 2017.txt' with 'Produto' in Column1 and 'Valor' in Column2. Row 18 is 'Joao Nunes - 2017.txt' with 'Conta-Corrente' in Column1 and '124765.00' in Column2.

	Name	Column1	Column2
17	Joao Nunes - 2017.txt	Produto	Valor
18	Joao Nunes - 2017.txt	Conta-Corrente	124765.00

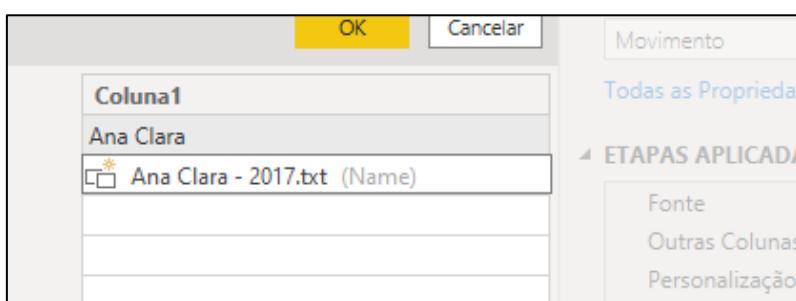
Agora a Consulta está sem “Lixo”, por enquanto.

	Name	Column1	Column2
1	Ana Clara - 2017.txt	Conta-Corrente	224581.36
2	Ana Clara - 2017.txt	Poupança	995874.36
3	Ana Clara - 2017.txt	Produtos	221458.36
4	Ana Clara - 2018.txt	Conta-Corrente	554125.36
5	Ana Clara - 2018.txt	Poupança	336524.35
6	Ana Clara - 2018.txt	Produtos	221485.36
7	Bruna Almeida - 2017.txt	Conta-Corrente	221458.36
8	Bruna Almeida - 2017.txt	Poupança	958742.36
9	Bruna Almeida - 2017.txt	Produtos	224158.36
10	Bruna Almeida - 2018.txt	Conta-Corrente	663248.36

Clique na coluna de “Name” e vamos em “Adicionar Coluna > Coluna de Exemplos > Da Seleção”. Para criamos uma Coluna de Exemplos a partir dos dados da coluna de nomes.



Basta digitar o nome da primeira ocorrência e a inteligência artificial do PB se encarregará do resto e aperte um Enter.



E finalize com o “Ok”.

Precisaremos também das datas, correto? Então realizaremos o mesmo passo anterior só que digitando os dados dos anos.

ID	Texto Entre os Delimitadores
	2017
	2017

Veja o resultado abaixo.

	Name	ABC Column1	ABC Column2	A ^B _C Texto Antes do Delimitador	A ^B _C Texto Entre os Delimitadores
1	a Clara - 2017.txt	Conta-Corrente	224581.36	Ana Clara	2017
2	a Clara - 2017.txt	Poupança	995874.36	Ana Clara	2017
3	a Clara - 2017.txt	Produtos	221458.36	Ana Clara	2017
4	a Clara - 2018.txt	Conta-Corrente	554125.36	Ana Clara	2018
5	a Clara - 2018.txt	Poupança	336524.35	Ana Clara	2018
6	a Clara - 2018.txt	Produtos	221485.36	Ana Clara	2018
7	Jana Almeida - 2017.txt	Conta-Corrente	221458.36	Bruna Almeida	2017

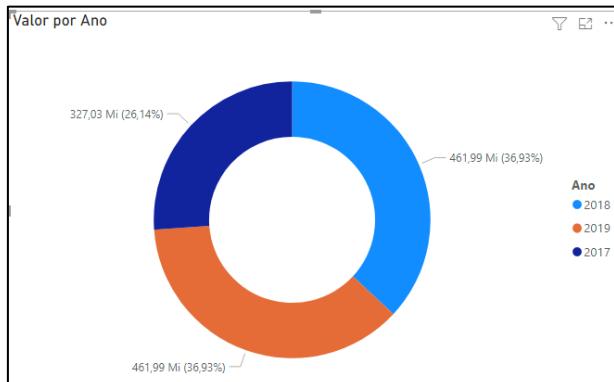
Não será mais necessária aquela primeira coluna de “Name”. Podemos remover e ajustar as colunas ao nosso prazer.

Lembre-se de deletar/excluir a Consulta Auxiliar, ela foi necessária apenas para buscarmos o código da Linguagem M.

Podemos Fechar e Aplicar as alterações.

Lembre-se que ao voltarmos no Power View, nós precisamos alterar o “Resumo” do ano para “**Não Resumir**”.

Agora com os dados totalmente modelados, tipados, as consultas “ok”. Podemos inserir novamente os arquivos de 2019 e verificarmos se as atualizações estão ocorrendo sem erro.

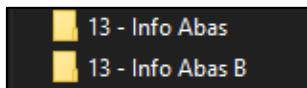


Podemos verificar se todos os dados foram incluídos corretamente no lado esquerdo em “Dados”. E verifique se os dados de 2017, 2018 e 2019 estão inseridos.

	X	✓		
	Ano	Gerente	Produto	Valor
	2017	Ana Clara	Conta-Corrente	22458136
	2017	Ana Clara	Poupança	99587436
	2017	Ana Clara	Produtos	22145836

Trabalhando com Abas

Utilizaremos o arquivo: "13 – Info Abas"



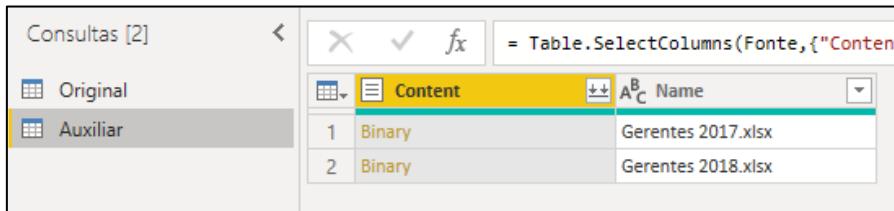
Percebe-se que vamos conectar a pasta. Mas, ao abrirmos um arquivo, perceba que semelhante ao exemplo anterior, os dados estão disponíveis, mas precisaremos novamente de informações que estão na Camada Binária (veja que os nomes dos gerentes estão nas Abas do arquivo em Excel).

Produtos	Conta-Cor	Poupança
955478,4	214586,4	214587,4

Produtos	Conta-Corrente	Poupança
885412,7	887569,32	248569,32

◀ ▶ | Bruna | João | Ana | ▶ + |

Todo aquele procedimento, transformar os dados, remover as outras colunas, duplicar a consulta original para que seja feita somente nela...



Consultas [2]

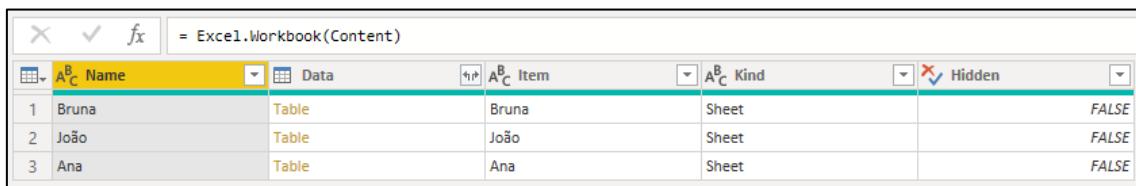
Original

Auxiliar

= Table.SelectColumns(Fonte, {"Content"})

	Content	Name
1	Binary	Gerentes 2017.xlsx
2	Binary	Gerentes 2018.xlsx

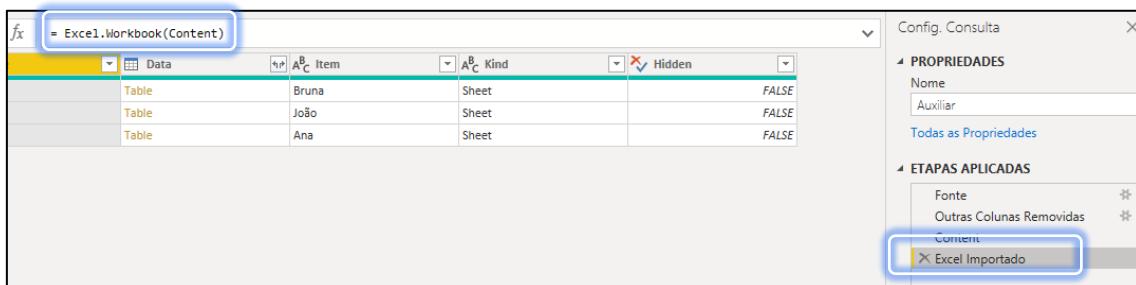
Ao clicarmos no Binário, vejamos o que acontece.



= Excel.Workbook(Content)

Name	Data	Item	Kind	Hidden
1 Bruna	Table	Bruna	Sheet	FALSE
2 João	Table	João	Sheet	FALSE
3 Ana	Table	Ana	Sheet	FALSE

Dentro do Binário, precisamos da fórmula que está presente na Etapa de "Excel Importado".



= Excel.Workbook(Content)

Data	Item	Kind	Hidden
Table	Bruna	Sheet	FALSE
Table	João	Sheet	FALSE
Table	Ana	Sheet	FALSE

Config. Consulta

PROPRIEDADES

Nome: Auxiliar

Todas as Propriedades

ETAPAS APPLICADAS

Fonte: Outras Colunas Removidas

Content

Excel Importado

Para que possamos adicionar uma nova coluna personalizada com essa fórmula.

"= Excel.Workbook(Content)"

Agora na Consulta Original, vamos adicionar uma nova coluna personalizada.

The screenshot shows the Power Query ribbon with the 'Coluna Personalizada' button highlighted under the 'Adicionar Coluna' tab. A tooltip explains that it creates a new column based on a personalized formula. Below the ribbon, a table preview shows two rows: 'Original' and 'Auxiliar'. The 'Original' row has one column labeled 'Binary'. The 'Auxiliar' row has three columns: '1' (number), 'Binary' (text), and 'Gerentes 2017.xlsx' (text).

E adicionaremos a fórmula que obtivemos na outra consulta. E lembre-se de substituir o “Content” anterior para o “Content” do arquivo atual! (Dica, o content atual virá entre “[]”).

The dialog box is titled 'Coluna Personalizada' and contains the following fields:
 - Nome da nova coluna: 'Personalizar'
 - Fórmula de coluna personalizada: '= Excel.Workbook([Content])'
 - Colunas disponíveis: 'Content' and 'Name' (with 'Content' currently selected).
 The main area shows a preview of the table with a new column 'Personalizar' added.

Vamos expandir a tabela.

The table editor shows the expanded table with the new column 'Personalizar' added. The first row contains the formula '= Table.AddColumn(#"Outras Colunas Removidas", "Personalizar", each [Content])'. The table has three columns: 'Content' (containing 'Binary'), 'Name' (containing 'Gerentes 2017.xlsx' and 'Gerentes 2018.xlsx'), and 'Personalizar' (containing 'Table').

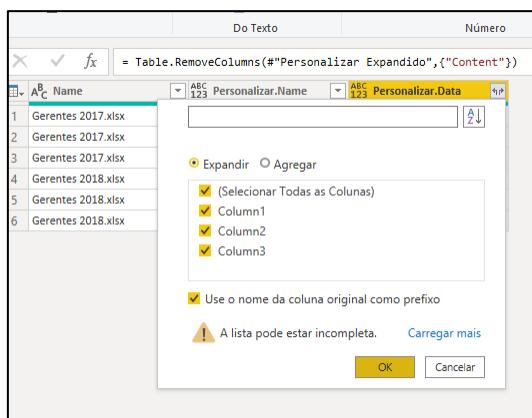
E com a coluna personalizada dessa forma, podemos obter todos os dados abaixo.

The dialog box shows the following settings:
 - 'Do Texto' and 'Número' buttons are visible.
 - Formula: '= Table.AddColumn(#"Outras Colunas Removidas", "Personalizar", each [Content])'
 - Column settings: 'Name' is selected, while 'Content' and 'Personalizar' are not.
 - 'Expandir' radio button is selected.
 - 'Seleccionar Todas as Colunas' checkbox is checked.
 - 'Use o nome da coluna original como prefixo' checkbox is checked.
 - 'OK' and 'Cancelar' buttons are at the bottom.

Vamos selecionar apenas os “Nomes” e as “Datas” (que são os dados em si no Excel). Não é mais necessário continuarmos com a coluna do Binário e nem com o arquivo anterior que é o Auxiliar.

Vamos entender as colunas. A coluna em azul brilhante, significam os arquivos, ou seja, são dois arquivos (Gerentes 2017 e Gerentes 2018). A coluna com os nomes, portanto, são as abas e a coluna com a “Table” são os dados respectivos de cada aba em seus respectivos arquivos.

Não podemos entrar na tabela, nós precisamos continuar expandindo e selecionando todos os filtros lá dentro.



Nós precisamos transformar aquelas colunas em linhas, veja, existem os dados “Produtos” e embaixo, o valor de “885474,36”. Veja que o Nome da Coluna, é exatamente o mesmo que está nos campos, isso é importante para aplicarmos a redução!

Agora, é necessário renomear o cabeçalho para que esse fique exatamente igual ao nome das informações que precisamos.

ABC 123 Produtos	ABC 123 Conta-Corrente	ABC 123 Poupança
------------------	------------------------	------------------

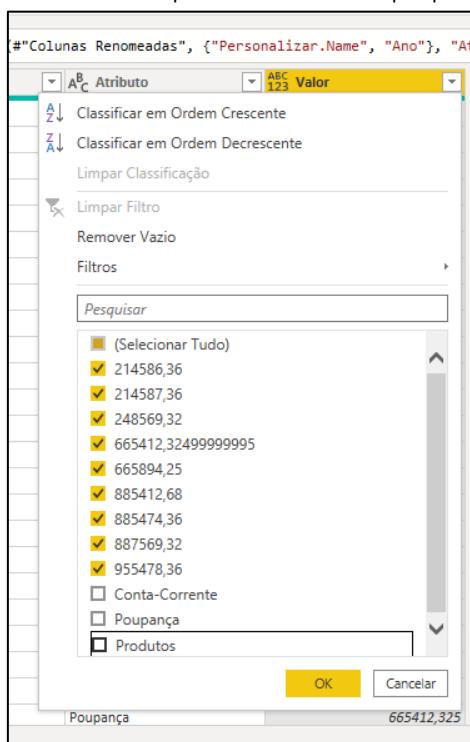
Vamos selecionar as 3 colunas e “Botão Direito > Transformar Colunas em Linhas”.

ABC 123 Nome	ABC Atributo	ABC 123 Valor
Bruna	Produtos	Produtos
Bruna	Conta-Corrente	Conta-Corrente
Bruna	Poupança	Poupança
Bruna	Produtos	885474,36
Bruna	Conta-Corrente	665894,25
Bruna	Poupança	665412,325

Pois agora, conseguiremos realizar um “Filtro Simples” de texto e remover o “Produto, Conta-Corrente e Poupança” de Valor.

ABC 123 Nome	ABC Atributo	ABC 123 Valor
Bruna	Produtos	885474,36
Bruna	Conta-Corrente	665894,25
Bruna	Poupança	665412,325
João	Produtos	885412,68
João	Conta-Corrente	887569,32
João	Poupança	248569,32

Vá até o campo de “Valor” e aplique um filtro simples removendo os 3.



Mesclagem

Vamos utilizar o arquivo na pasta 14 – Diversos.

DADOS (D:) > workspace > curso_MasterPowerBI > Datasets > 14 - Diversos		
Nome	Data de modificação	Tipo
Porcentagem.csv	26/10/2019 15:06	Arquiv
Produtos.txt	26/10/2019 14:59	Docum
Produtos.xlsx	26/10/2019 15:00	Planilh

Vejamos os arquivos (Produtos.txt – Produtos.xlsx e Porcentagem.csv)

The screenshot displays three files side-by-side:

- Produtos.txt - Bloco de Notas**: A text file containing product codes and names, separated by commas. The data includes: cod,produto, 1,cadeira, 2,geladeira, 3,fogão, 4,mesa.
- Produtos.xlsx**: An Excel spreadsheet with two columns: A (cod) and B (produto). The data is identical to the text file.
- Porcentagem.csv**: A CSV file with two columns: A (cod,porcentagem) and B (porcentagem). The data is: 1,10, 2,5, 3,14, 4,3.

Como trabalhar com esses três arquivos sendo que a estrutura de cada um é diferente?
Veja que além da extensão, a estrutura é diferente.

Para começarmos, vamos conectar inicialmente com o arquivo em Excel(o Produtos.xlsx).

Uma vez que essa consulta foi realizada, no Power Query que nós vamos obter novos dados. Clique em “Nova Fonte” e conecte o arquivo .txt.

The screenshot shows the Power Query ribbon and the query editor interface:

- Ribbon:** Arquivo, Página Inicial, Transformar, Adicionar Coluna, Exibição, Ferramentas.
- Consulta:** Consultas [1] (Sheet1 selected).
- Power Query Editor:** Shows the data from the Produtos.xlsx file. The first few rows are: 1 cod,produto, 2 celular, 3 estante, 4 cama, 5 travesseiro, 6 lençol, 7 secador, 8 creme.

O segundo arquivo em formato **.txt** foi conectado. E por último, o arquivo porcentagem.**.csv**.

The screenshot shows the Power BI Data Editor interface with three tables connected:

- Produtos Excel**: Contains columns "cod" and "porcentagem".
- Produtos txt**: Contains columns "cod" and "porcentagem".
- Porcentagem Excel**: Contains columns "cod" and "porcentagem".

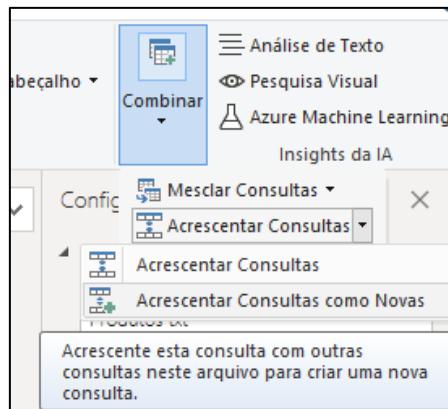
	cod	porcentagem
1	1	10
2	2	5
3	3	14
4	4	3
5	5	5

Pronto, agora os 3 arquivos estão conectados.

Combinando Extensões Diferentes

Agora precisamos combinar esses arquivos em um.

Para isso, com o arquivo Produtos.**.txt** selecionado, vamos em “[Página Inicial > Combinar > Acrescentar Consultas > Acrescentar Consultas como Novas](#)”.



Isso significa que eu não vou mexer nas consultas que já existem, eu vou criar uma nova.

Que já será a Produtos.**.txt** e sem seguida com a Produtos.xlsx (Uma complementa a outra mas foram enviadas em extensão diferente).

The screenshot shows the 'Acrescentar' (Add) dialog box with the following settings:

- Concatene linhas de duas tabelas em uma única tabela.**
- Duas tabelas** (radio button selected)
- Primeira tabela**: Produtos.txt
- Segunda tabela**: Produtos Excel

Agora possuímos uma tabela única que uniu dois arquivos de extensões diferentes que se complementavam.

The screenshot shows the Power BI Data Editor interface. On the left, there's a sidebar titled 'Consultas [4]' containing four items: 'Produtos Excel', 'Produtos txt', 'Porcentagem Excel', and 'Produtos'. The 'Produtos' item is selected. The main area displays a table with two columns: 'cod' (numbered 1 to 12) and 'produto' (product names). The formula bar at the top shows the DAX code: '= Table.Combine({#"Produtos txt"}, #'. The table data is as follows:

	cod	produto
1		cadeira
2		geladeira
3		fogão
4		mesa
5		TV
6		celular
7		estante
8		cama
9		travesseiro
10		lençol
11		secador
12		creme

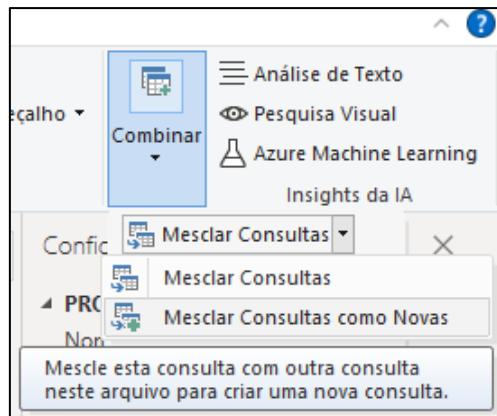
E podemos DESABILITAR A CARGA dos dois Produtos (.xlsx e o .txt).

Agora, possuímos 2 consultas. Uma de Produtos e outra de Porcentagem que evidencia a quantidade de desconto que será dada em cada produto.

Mesclar Consultas

Essa feature, é importante pois agora que temos duas consultas que devem ser unir.

Com a consulta de Produtos selecionada, vamos em "Página Inicial > Combinar > Mesclar Consultas". Não vamos adicionar uma nova pois a consulta de Produtos já está muito bem arrumada.



Na próxima página, vamos aprender a “Linkar” as colunas das consultas.

Seleção Correspondente de Consultas Mescladas

Após solicitarmos a mesclagem, precisamos definir qual será a consulta que será mesclada com a atual. E selecionarmos as colunas.

cod	produto
1	cadeira
2	geladeira
3	fogão
4	mesa
5	TV

cod	porcentagem
1	10
2	5
3	14
4	3
5	5

Logo após selecionarmos, veja que o PB reconhece as correspondências.

✓ A seleção corresponde a 12 de 12 linhas da primeira tabela.

Veja o resultado

cod	produto	Porcentagem Excel
1	cadeira	Table
2	geladeira	Table
3	fogão	Table
4	mesa	Table
5		
6		
7		
8		

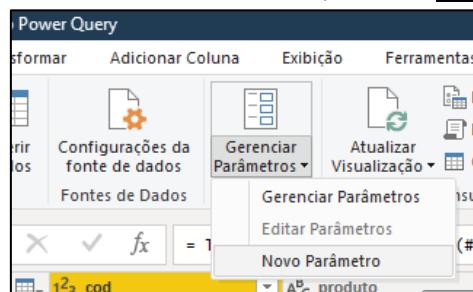
Clique nas setas direcionais da coluna, é expanda, deixando apenas a Porcentagem, uma vez que não precisaremos do código que já está na primeira coluna.

cod	produto	porcentagem
1	cadeira	10
2	geladeira	5
3	fogão	14
4	mesa	3
5		
6		
7		
8		

Parâmetros

O Parâmetro serve para quando se trocar o local do arquivo, que se altere apenas o caminho da rede apenas no Path, e que não precise alterar em todas as consultas todas as vezes.

Para criar um Parâmetro, vá em “[Página Inicial > Parâmetro > Novo Parâmetro](#)”.



Abrindo a seguinte janela. Nessa janela indicaremos onde os arquivos estão. Sempre que precisarmos alterar o local do arquivo, precisaremos alterar no Path!

The dialog box has the following fields:

- Nome:** Path
- Descrição:** (empty)
- Obrigatório:** checked
- Tipo:** Texto
- Valores Sugeridos:** Qualquer valor
- Valor Atual:** D:\workspace\curso_MasterPowerBI\Datasets\14 - Diversos

Perceba que o parâmetro foi criado

The 'Consultas [5]' list shows the following items:

- Produtos Excel
- Produtos txt
- Porcentagem Excel
- Produtos
- Path (D:\workspace\curso...

The 'Path' item is highlighted with a gray box.

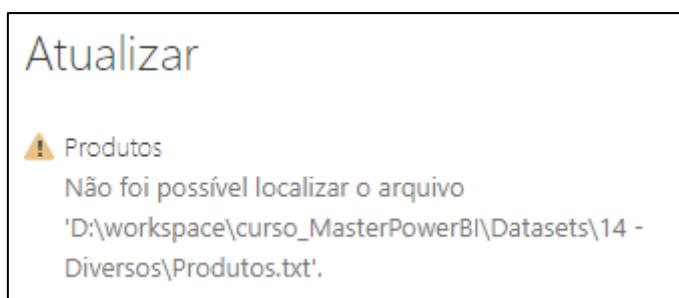
Agora, na pasta da rede, vamos cortar os arquivos da pasta anterior, daquela “14 – Diversos”. Essa pasta continha os arquivos que fizemos os parâmetros na página anterior. Veja que então, alteramos o local onde os arquivos estavam e consequentemente todas as consultas gerarão um erro.

DADOS (D:) > workspace > curso_MasterPowerBI > Datasets > 14 - Diversos		
Nome	Data de modificação	Tipo

E vamos criar uma pasta e despejarmos esses arquivos nela.

DADOS (D:) > workspace > curso_MasterPowerBI > Datasets > 14 - Novo		
Nome	Data de modificação	Tipo
Porcentagem.csv	26/10/2019 15:06	Arq
Produtos.txt	26/10/2019 14:59	Dado
Produtos.xlsx	26/10/2019 15:00	Pla

Se tentarmos voltar ao **Power BI** e atualizarmos, seremos recebidos com erros.



Para isso, vamos corrigir TODAS as consultas que foram selecionadas.

É necessário voltar no PQ, na consulta e na Etapas Aplicadas, procure pela “Fonte”.

Por exemplo, na consulta de “Produtos Excel”, vá na Etapas Aplicadas > Fonte

Veja que na barra de fórmulas, o caminho do arquivo na rede:

Isso significa que se remover o arquivo de onde ele está na Fonte, o PB não o encontrará.

Path &

Por isso criamos o Path. Para que qualquer alteração do local do arquivo, somente alteraremos o local no próprio Path.

Porém, precisaremos alterar primeiramente (antes do Path), em cada consulta, conforme abaixo:

```
= Excel.Workbook(File.Contents(Path & "\Produtos.xlsx"), null, true)
```

Apague todo o caminho do arquivo, deixando somente com “\nomeDoArquivo.extensão”. E antes, veja que o antecedente tem que ser o nome do parâmetro criado (Path &) Esse “&” serve para **concatenação** entre o “Path” e o “nome do arquivo”. E sempre que eu quiser alterar o local do arquivo, podemos jogar o arquivo de um lado para o outro na rede, e somente precisaremos alterar o local do Path!

Veja, estava com erro em Porcentagem Excel.

The screenshot shows the Power BI Data View interface. On the left, there is a list of queries: 'Produtos Excel', 'Produtos txt', 'Porcentagem Excel', and 'Produtos'. The 'Porcentagem Excel' query is selected. Its formula bar contains the code: `= Csv.Document(File.Contents("D:\workspace\curso_MasterPowerBI\Datasets\14 - Diversos\Porcentagem.csv"))`. The preview pane shows a table with two columns: 'Column1' and 'Column2'. The data rows are: 1 cod | porcentagem, 2 1 | 10, and 3 2 | 5.

Pois o caminho da rede estava errado na Fonte.

Alterando conforme o Path, ele corrigirá.

This screenshot shows the same Power BI Data View interface as the previous one, but with a corrected formula. The 'Porcentagem Excel' query now has the formula: `= Csv.Document(File.Contents(Path & "\Porcentagem.csv"))`. The preview pane shows the same table data as before: 1 cod | porcentagem, 2 1 | 10, and 3 2 | 5.

Sempre que quiser alterar o local dos arquivos na rede, basta alterarmos o parâmetro no Path.

This screenshot shows the Power BI Parameter Manager. On the left, there is a list of parameters: 'Produtos Excel', 'Produtos txt', 'Porcentagem Excel', 'Produtos', and 'Path (D:\workspace\curso...'. The 'Path' parameter is selected. The right pane shows the 'Valor Atual' (Current Value) as 'D:\workspace\curso_MasterPowerBI\Datasets\14 - Novo' and a 'Gerenciar Parâmetro' (Manage Parameter) button.

Aí é só adicionar o novo local dos arquivos e todas as consultas serão atualizadas automaticamente.

Sim, é necessário fazer apenas uma vez a alteração do Path em cada consulta gerada, mas se tornará muito mais fácil ao mudar o local do arquivo.

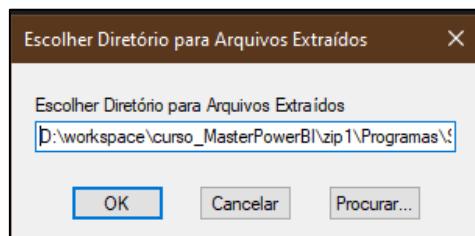
Infraestrutura – SQL Server

Vamos utilizar um servidor SQL nesse capítulo e será necessário o download do “SQL Express”.

DADOS (D:) > workspace > curso_MasterPowerBI > zip1 > Programas > SQL Server Express 64			
Nome	Data de modificação	Tipo	Ação
SQLEXPRESS_x64_PTB.exe	28/09/2019 08:36	Aplicativo	

Após concluir a Instalação, é necessário baixar também o SMSS – SQL Server Management Studio, que será disponibilizado logo após o término da instalação.

O SQL Express indicará que precisará de um diretório para Extrair e Alocar os dados.



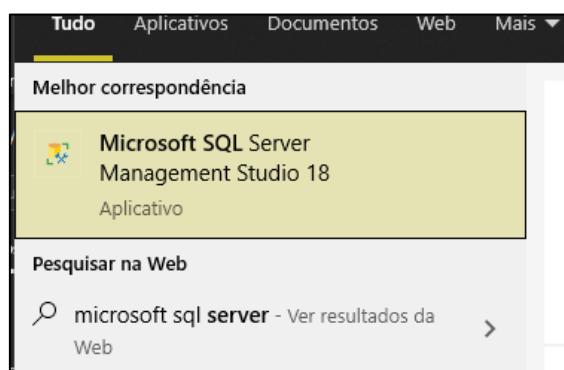
Instalar conforme a aula 65. Instalando o SQL Express até 9:35 minutos.

Verificando serviços

Verifique pelo “Windows > Serviços” se o SQL Server está em execução.

Serviços					
Arquivo	Ação	Exibir	Ajuda		
Serviços (local)		Serviços (local)			
				Nome	Descrição
				SQL Server (SQLEXPRESS)	Fornec... Em Execu...
				SQL Server Agent (SQLEXP...)	Executa tr...
				SQL Server Browser	Fornec... in...
					Tipo de Inicialização
					Fazer Logon como
					NT Service\MSQL...
					Serviço de rede
					Serviço local

Abrir o “Microsoft SQL Server Management Studio 18”



Vamos utilizar o “AdventureWork” para os estudos.

DADOS (D:) > workspace > curso_MasterPowerBI > Datasets > 15 - AdventureWorks 2017		
Nome	Data de modificação	Tipo
AdventureWorks2017.bak	20/10/2019 17:40	Arquivo BAK
AdventureWorksDW2017.bak	20/10/2019 17:40	Arquivo BAK
Restore AdventureWorks.docx	20/10/2019 17:52	Documento do Mi...

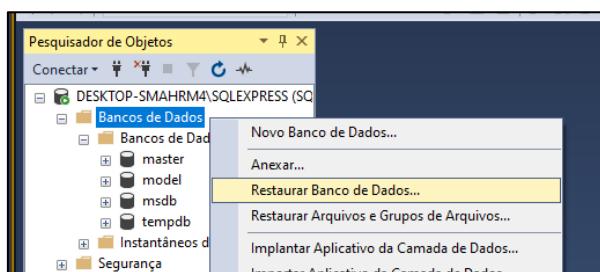
São Bancos de Dados, cada um modelado conforme a necessidade.

Obs: “AdventureWork2017.bak” é o Ambiente OLTP

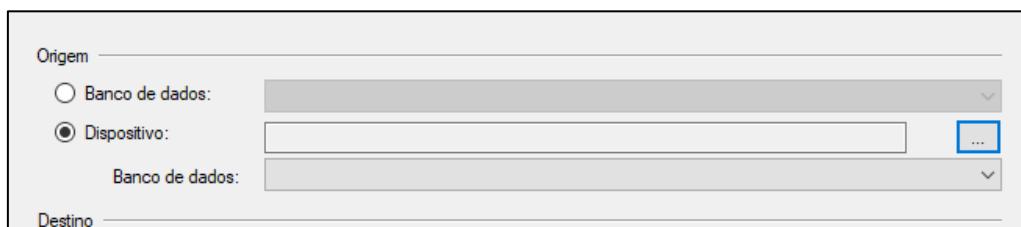
Obs[2]: “AdventureWorkDW2017.bak” é o Ambiente OLAP , para *Data Warehouse*.

Os dois arquivos “.bak” são arquivos de *Backup*. Ou seja, serão restaurados.

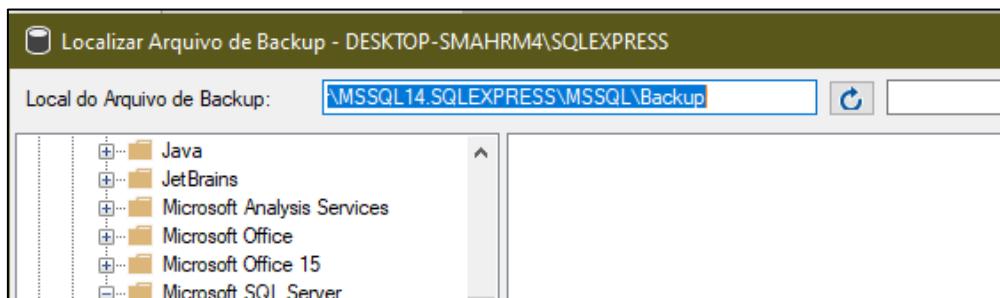
No **MSSMS**, clique em “Banco de Dados > Restaurar Banco de Dados”.



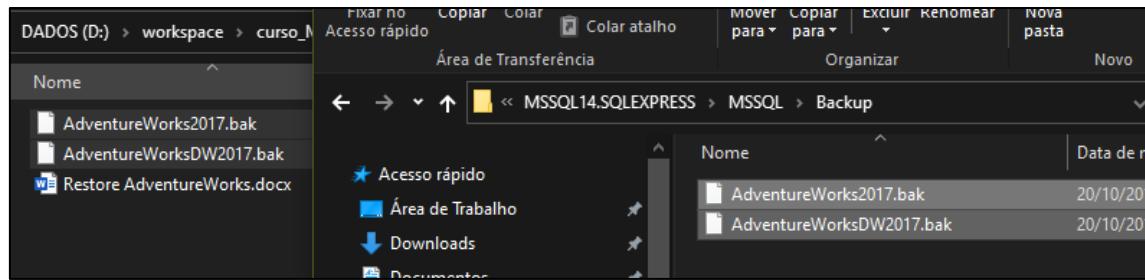
Agora clique em “Dispositivos” e na caixa com “...”. E depois “Adicionar”.



Será necessário colocar os dois arquivos dos bancos de dados exatamente dentro dessa pasta que o SQL Server criou como *Backup*..

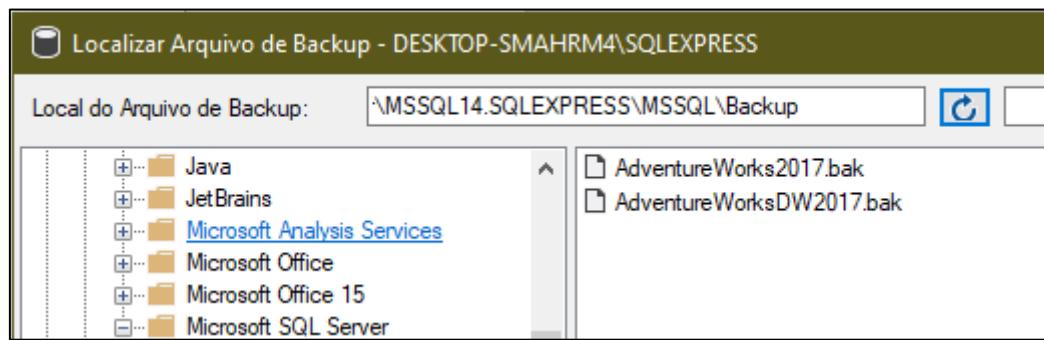


Está vazio, basta copiar o caminho, e então abrir num novo “Explorador de Arquivo” e de lá, copiar os dois arquivos “AdventureWork2017.bak” para dentro dessa página



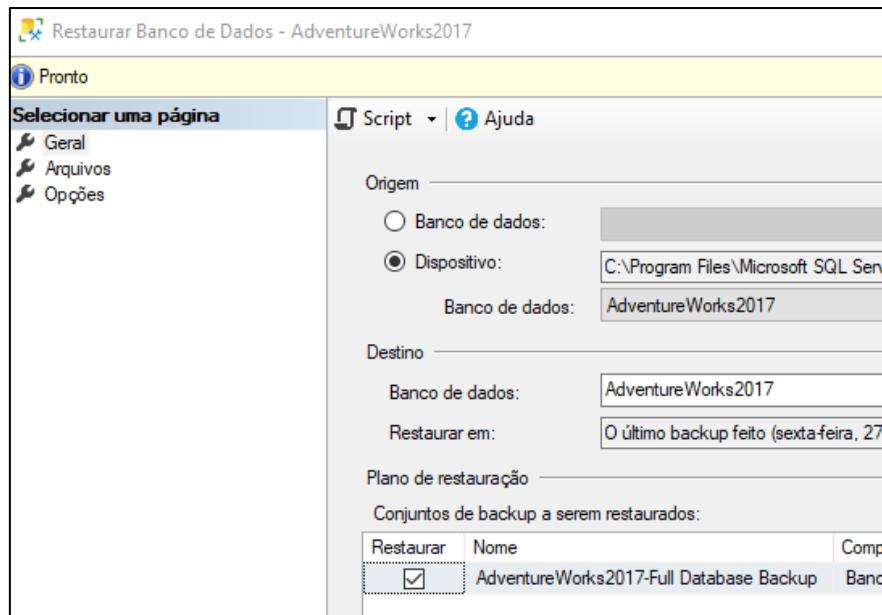
Agora estará dentro da pasta de “Backup” do SQL Server.

Ao atualizarmos o caminho no **SSMS**, veja que agora os arquivos estarão lá.

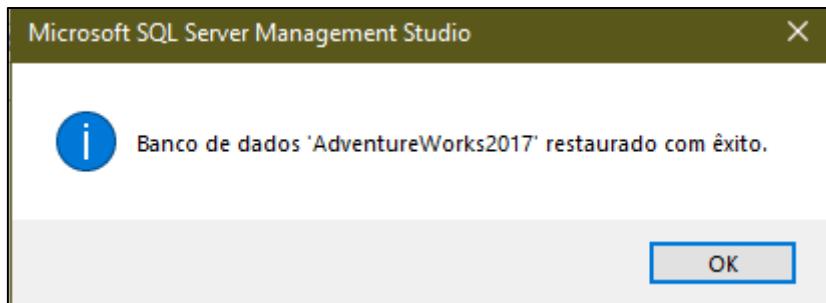


Restaurando backups

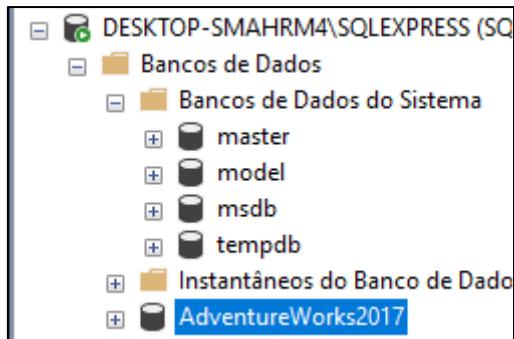
Para restaurar o banco de dados da OLTP, basta clicar duas vezes e ir clicando no “Ok”. Quando a mensagem de “Pronto” aparecer, a restauração foi concluída.



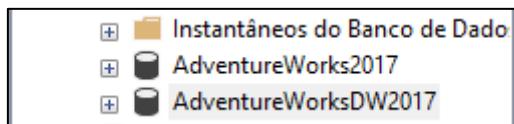
Digite “ok” novamente.



Veja que o ambiente foi restaurado com sucesso, a esquerda.



Vamos restaurar o DW agora, que é o ambiente OLAP. Seguindo o mesmo caminho anteriormente.



Pronto, os dois Bancos de Dados foram reestabelecidos.

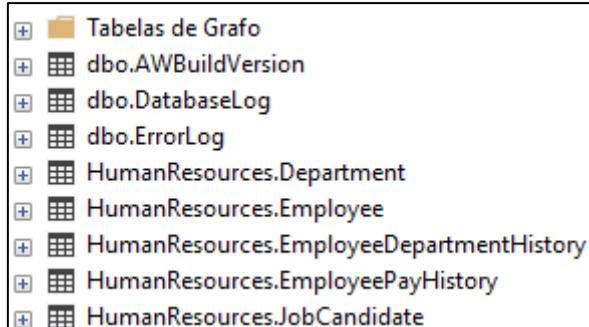
Para entendermos finalmente a diferença entre um Banco de Dados Relacional x Data Warehouse (DW). Vejamos os próximos subcapítulos (Ambiente OLTP e OLAP).

O Banco de Dados que acabamos de importar, são um estudo de caso de uma loja de aventura, de camping.

Nesse primeiro Banco de Dados Relacional.

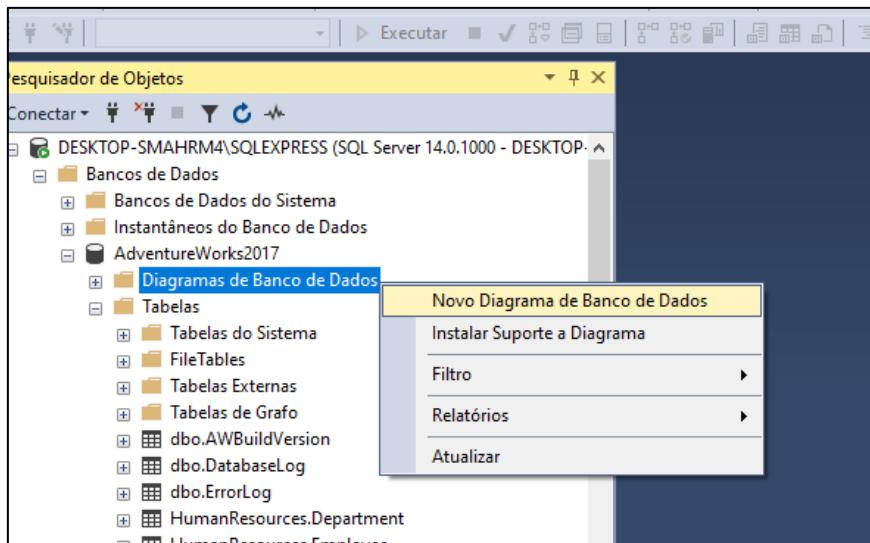
Ambiente OLTP

Veja que o pessoal que desenvolveu esse BD, fizeram por “*Schemas*”, ou seja, fizeram o “`porÁrea.nomeDaTabela`”, isso é muito prático pois facilita o entendimento, por exemplo, “`HumanResources.Employee`” ou “`Production.ProductModel`”.

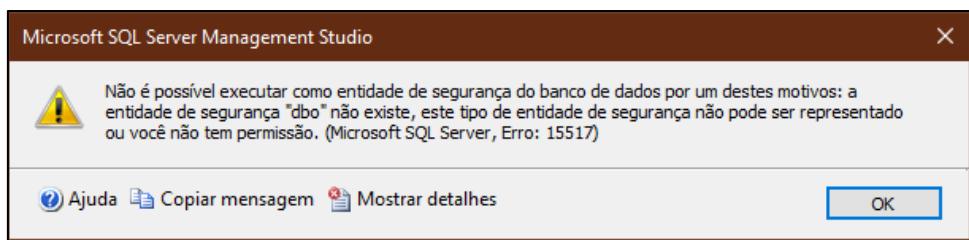


Vamos criar um DIAGRAMA dessas tabelas. E esse Diagrama é um MAPA que comprehende todas essas tabelas.

Em “[Diagrama de Banco de Dados > Botão Direito > Novo Diagrama](#)”.



Caso o erro abaixo aconteça,



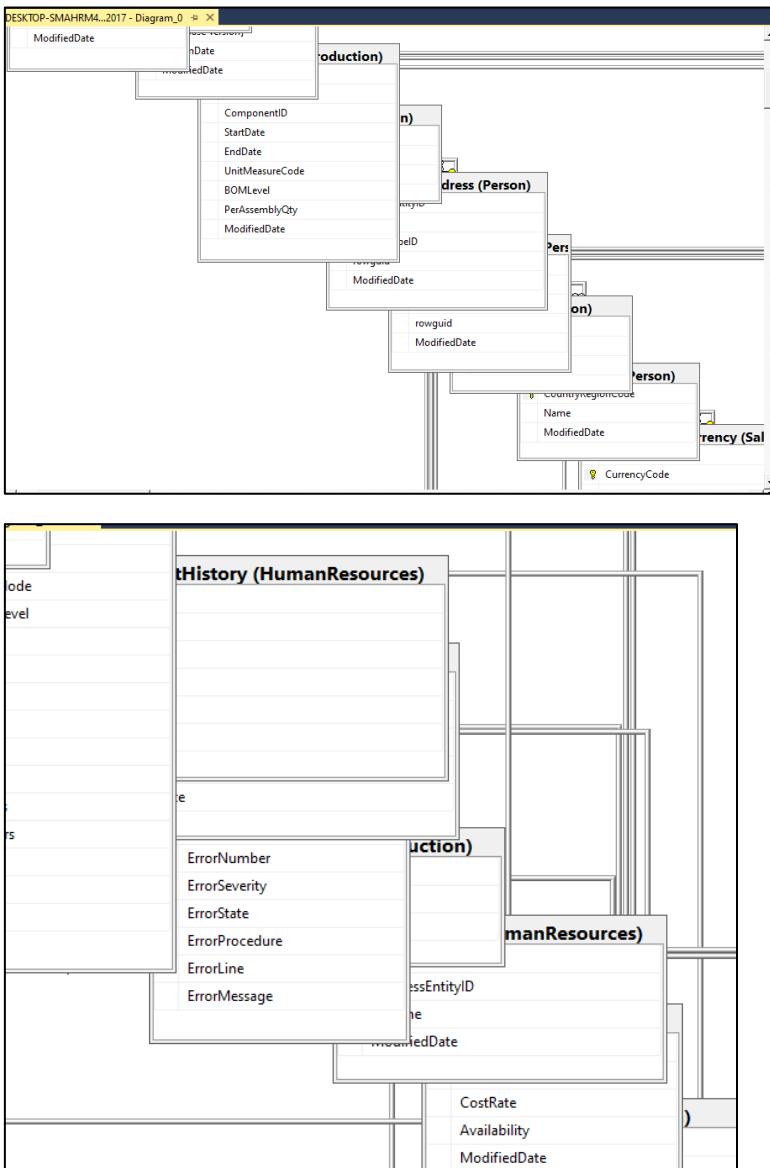
Vá em “[AdventureWorks2017 > Botão Direito > Propriedades > Arquivos > Proprietário > ... > Adicionar > Procurar > OK](#)”

Esse é o caminho para se adicionar como um proprietário do Banco.

Após isso, tente criar o Diagrama novamente.

E agora selecione todas as tabelas e clique em Adicionar. Isso deve levar algum tempo.

Se torna algo gigantesco.

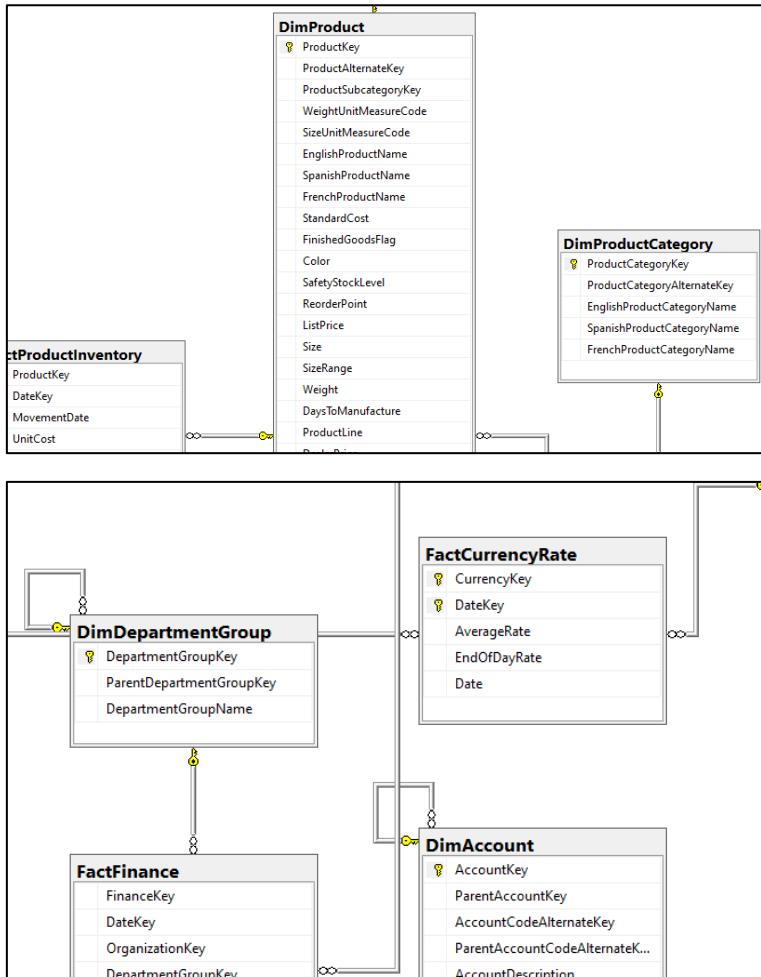


O mais ideal caso a empresa não consiga criar um ambiente DW, se ela for obrigada a trabalhar com o Relacional (caso acima), aconselha-se elaborar uma forma que transforme essas tabelas em Excel.

Ambiente OLAP

Vejamos a diferença do que acontece em um Data Warehouse (DW).

Muito menos tabelas, muito mais organização.



Modelagem de Dados e Power Pivot

A Linguagem **Dax** é utilizada para o **Power Pivot**.

Relacionamentos

Inicie o Banco de Dados do SQL Server Express.

Existe um Script inicial para a criação de um Banco de Dados.



Esse arquivo de “Queries Iniciais.sql” é um Script básico para a criação de um BD.

```
SELECT * FROM FactInternetSales
GO

SELECT FIRSTNAME, LASTNAME, YEARLYINCOME, C
FROM DimCustomer DC
INNER JOIN FactInternetSales FS
ON DC.CustomerKey = FS.CustomerKey
INNER JOIN DimCurrency DIC
ON DIC.CurrencyKey = FS.CurrencyKey
ORDER BY FIRSTNAME, LASTNAME
GO

-----
--- JOINS E RELACIONAMENTOS
-----

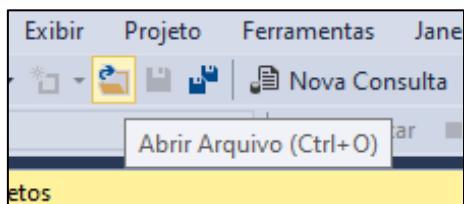
CREATE DATABASE CLINICA
GO

USE CLINICA
GO

CREATE TABLE PACIENTE(
    IDPACIENTE INT PRIMARY KEY,
    NOME VARCHAR(30),
    SEXO CHAR(1)
)
GO

CREATE TABLE ENDERECO(
    IDENDERECO INT PRIMARY KEY,
    RUA VARCHAR(30),
    BAIRRO VARCHAR(30),
    CIDADE VARCHAR(30),
    ID PACIENTE INT UNIQUE
```

Para criarmos o BD, vamos em “Abrir Arquivo (CTRL + O)”.



Agora com o Banco de Dados criado, vamos começar a entender.

O professor do curso, Felipe Mafra, que criou a query abaixo, durante as aulas, fez várias execuções das Queries para mostrar o que está acontecendo.

```
-- DIFERENÇA ENTRE FATOS E DIMENSÕES

USE AdventureWorksDW2017
GO

SELECT * FROM DimCustomer
GO

SELECT * FROM DimCurrency
GO

SELECT * FROM FactInternetSales
GO

SELECT FIRSTNAME, LASTNAME, YEARLYINCOME, CURRENCYNAME, ORDERQUANTITY, SALESAMOUNT
FROM DimCustomer DC
INNER JOIN FactInternetSales FS
ON DC.CustomerKey = FS.CustomerKey
INNER JOIN DimCurrency DIC
ON DIC.CurrencyKey = FS.CurrencyKey
ORDER BY FIRSTNAME, LASTNAME
GO
```

Em vermelho, utilizaremos o Banco de Dados em formato do DW (Data Warehouse).

Após, o roxo, é uma consulta (Select) da Dimensão dos Clientes, demonstrando todas as informações dos clientes da empresa.

CustomerKey	GeographyKey	CustomerAlternateKey	Title	FirstName	MiddleName	LastName	NameStyle	BirthDate	MaritalStatus	Suffix	Gender	EmailAddress
11000	26	AW00011000	NULL	Jon	V	Yang	0	1971-10-06	M	NULL	M	jon24@adventure-works.com
11001	37	AW00011001	NULL	Eugene	L	Huang	0	1976-05-10	S	NULL	M	eugene10@adventure-works.com
11002	31	AW00011002	NULL	Ruben	NULL	Torres	0	1971-02-09	M	NULL	M	ruben35@adventure-works.com

Em Verde, é uma consulta (Select) da Dimensão das Moedas, demonstrando todas as características das moedas.

CurrencyKey	CurrencyAlternateKey	CurrencyName
16	BRL	Brazilian Real
17	BND	Brunei Dollar
18	BGN	Bulgarian Lev
19	CAD	Canadian Dollar
20	GHC	Cedi
21	XOF	CEA Franc BCEAO

Em Amarelo/Laranja, é novamente uma consulta (Select) que demonstra os Fatos, ou seja, as vendas realizadas.

ProductKey	OrderDateKey	DueDateKey	ShipDateKey	CustomerKey	PromotionKey	CurrencyKey	SalesTerritoryKey	SalesOrderNumber	SalesOrderLineNumber
310	20101229	20110110	20110105	21768	1	19	6	SO43697	1
346	20101229	20110110	20110105	28389	1	39	7	SO43698	1
346	20101229	20110110	20110105	25863	1	100	1	SO43699	1
326	20101229	20110110	20110105	14501	1	100	4	SO43700	1

Veja que os fatos não relacionam diretamente quem os comprou , apenas uma chave de “CustomerKey”.

Em azul, é uma consulta (Select) que recebe joins, ou seja, que eu demonstro o que eu quero receber como consulta.

FIRSTNAME	LASTNAME	YEARLYINCOME	CURRENCYNAME	ORDERQUANTITY	SALESAMOUNT
Aaron	Adams	50000,00	US Dollar	1	3,99
Aaron	Adams	50000,00	US Dollar	1	24,99
Aaron	Adams	50000,00	US Dollar	1	34,99
Aaron	Adams	50000,00	US Dollar	1	53,99
Aaron	Alexander	40000,00	US Dollar	1	69,99
Aaron	Allen	10000,00	Canadian Dollar	1	3399,99
Aaron	Baker	40000,00	US Dollar	1	1700,99
Aaron	Baker	40000,00	US Dollar	1	49,99
Aaron	Bryant	70000,00	US Dollar	1	35,00
Aaron	Bryant	70000,00	US Dollar	1	4,00

Poderíamos adicionar quaisquer campos presentes nas outras consultas para realizarmos essa. Essa consulta, é a que recebemos no Excel, podemos solicitar esses dados/consultas ao **DBA** (Analistas de Database) ou nós mesmo podemos escrever o script e recebermos esses dados.

Clicando na área roxa, selecionaremos toda a consulta que foi recebida, podemos copiar os dados e criamos uma planilha em Excel e colarmos.

Resultados				Mensagens
	FIRSTNAME	LASTNAME	YEARLYINCOME	CURRENCYNAME
1	Aaron	Adams	50000,00	US Dollar
2	Aaron	Adams	50000,00	US Dollar

Ou podemos salvar essa pesquisa. A desvantagem que ela não será simultânea com a realidade.

Resultados					Mensagens
	FIRSTNAME	LASTNAME	YEARLYINCOME	CURRENCYNAME	ORDERQUANTITY
1	Aaron	Adams	50000,00	US Dollar	1
2	Aaron	Adams	50000,00	US Dollar	1
Aaron	Adams				
Aaron	Adams				
Aaron	Alexander				
Aaron	Allen				
Aaron	Baker				
Aaron	Baker				
Aaron	Bryant				
Aaron	Bryant				

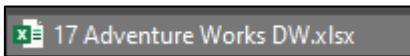
...

Consulta executada com êxito

...

Fim dessa parte do Script, apenas para entendimento.

A alternativa a utilização da conexão do Banco de Dados, podemos utilizar a planilha



Imagine que o **DBA** encaminhou os dados em formato .xlsx. E, portanto, possuímos os dados que supostamente só estariam em um Banco de Dados, pela planilha.

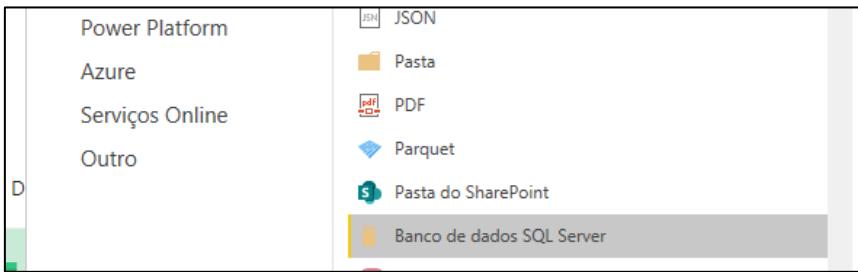
Porém, o correto seria estabelecer a conexão com o Banco de Dados, coisa que veremos logo abaixo.

Conectando o Power BI ao SQL Server

Nesse capítulo vamos aprender como estabelecer a conexão entre o **Power BI** e um SQL Server de Data Warehouse (DW).

Vamos conectar os dados que estão naquela planilha gigantesca acima, porém, não vamos conectar a planilha e sim estabelecer a conexão entre o PB x DW.

“Obter Dados > Banco de Dados SQL Server > Conectar”.



A próxima tela é muito importante.

Modo de Conectividade de Dados

Banco de dados SQL Server

Servidor ⓘ

Banco de Dados (opcional)

Modo de Conectividade de Dados ⓘ

Importar

DirectQuery

► Opções avançadas

Existe uma diferença entre **IMPORTAR** e **DIRECTQUERY**.

Através da **IMPORTAÇÃO**, o PB realizará um import direto dos dados do Banco de Dados, trazendo absolutamente todos os dados lá presente (embora a compressão de dados do PB seja muito boa, dependendo do tamanho do BD, a desvantagem será que serão muitos dados e isso impactará a performance.) E, também será necessário atualizar o PB para verificar se existem novas informações/dados do BD. Veja mais em Gateway.

Através do **DIRECTQUERY**, é uma buscar direta no Banco de Dados, ou seja, ele vai trazer apenas os dados especificados. Em compensação, a conexão com o BD será constante e repetidamente. É um método que vai impactar e muito a performance, mas é um método *LIVE* (ou seja, muito atualizado, os dados sempre estarão atualizados).

Servidor ⓘ

Sobre o acesso ao “SERVIDOR”. Como nesse exercício estamos trabalhando com o local, utilizaremos o (.) ponto que é para a Microsoft, o acesso “Local Host”. Após utilizaremos o (\) que é a instância do BD e por fim o nome dessa instância.

Se estivéssemos nos conectando a um BD na empresa, seria necessário que o DBA informasse qual é o servidor de acesso ao BD.

Servidor ⓘ

\SQLEXPRESS

Banco de Dados (opcional)

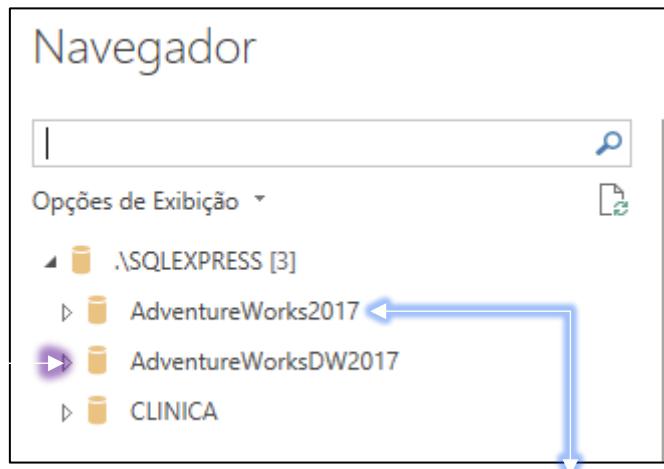
Modo de Conectividade de Dados ⓘ

Importar

DirectQuery

A próxima tela será a de segurança, caso tenha estabelecido alguma segurança no Banco de Dados, mais uma vez, o DBA deverá informar qual é o padrão da segurança.

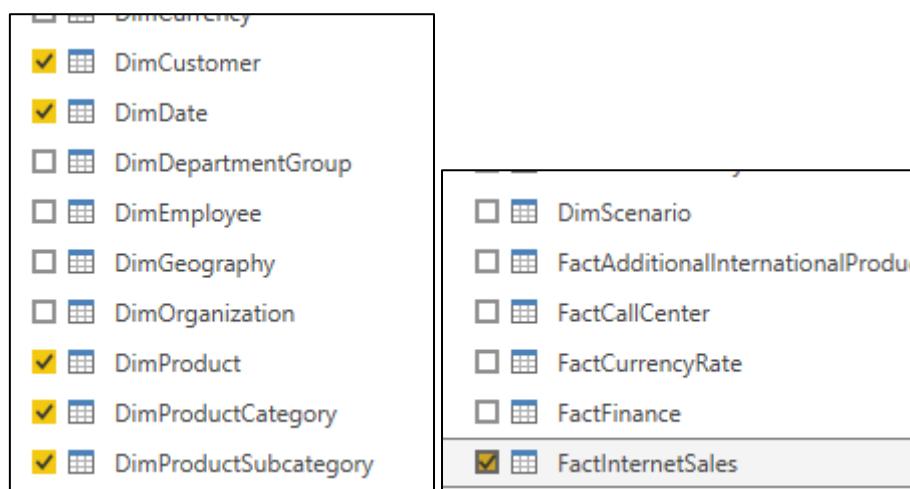
Para esse exercício, após passarmos pela segurança, vamos estabelecer a conexão com o “AdventureWorksDW2017”.



Esse Banco de Dados é o Relacional, é onde os vendedores, consultores, ou quem quer que seja, utiliza o INPUT, não podemos conectar nesse. Somente estabelecemos a conexão no DW.

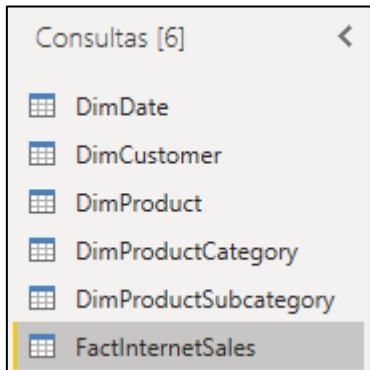
→ Esse será o nosso Banco a ser conectado, sempre no DW.

Para as **Dimensões** que buscaremos nesse exercício, serão as seguintes:
“DimCustomer”, “DimDate”, “DimProduct”, “DimProductCategory”,
“DimProductSubCategory”, e, para os **Fatos**, apenas a “FactInternetSales”.



Próximo passo, transformar os dados.

Veja que todas as consultas que foram selecionadas acima, estão conectadas.



Estamos nesse momento, na consulta da Fato de Vendas de Internet. Veja que as colunas que estão no BD foram conectadas.

1^2_3 ProductKey	1^2_3 OrderDateKey	1^2_3 DueDateKey	1^2_3 ShipDateKey
310	20101229	20110110	20110105
346	20101229	20110110	20110105
346	20101229	20110110	20110105

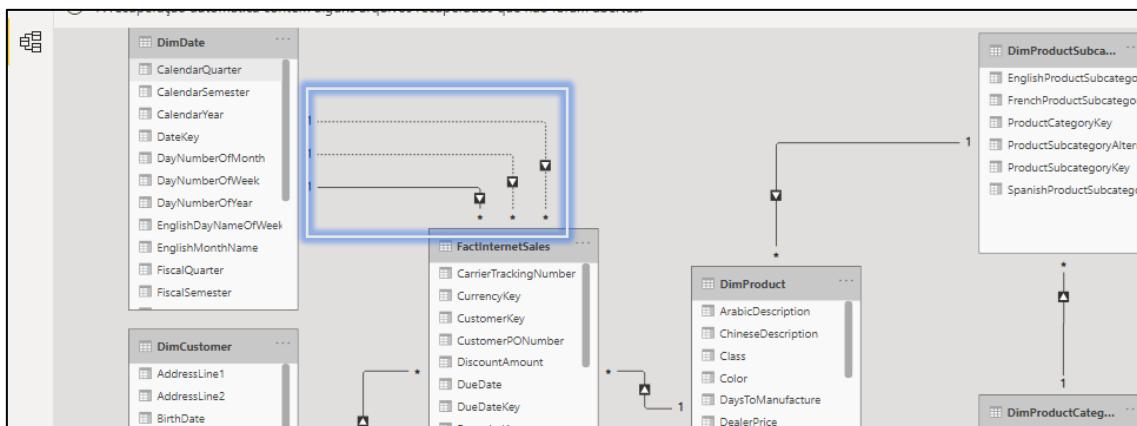
E, também, perceba que o **Power BI**, apresenta sugestões de Tabelas que estão “linkadas”, a essa de Fato de Vendas de Internet.

DimCurrency	DimCustomer	DimDate(DueDateKey)	DimDate(OrderDateKey)	Dim
Value	Value	Value	Value	Value
Value	Value	Value	Value	Value

Uma observação muito importante, é que sempre que estivermos trabalhando com DW, precisamos de uma Dimensão Date/Time (ex: DimDate).

Para o momento, vamos apenas Fechar e Aplicar, sem nenhuma transformação no Power Query. Nesse exercício nós não tratamos os dados pois eles já estão “tratados” pelo Banco de Dados. Porém, poderíamos esconder algumas colunas para reduzir... Mas para esse exercício a ideia é irmos no “Modelo” para estabelecermos as relações.

Veja que o próprio **Power BI** se encarregou de estabelecer os relacionamentos que estão presentes pelo BD.

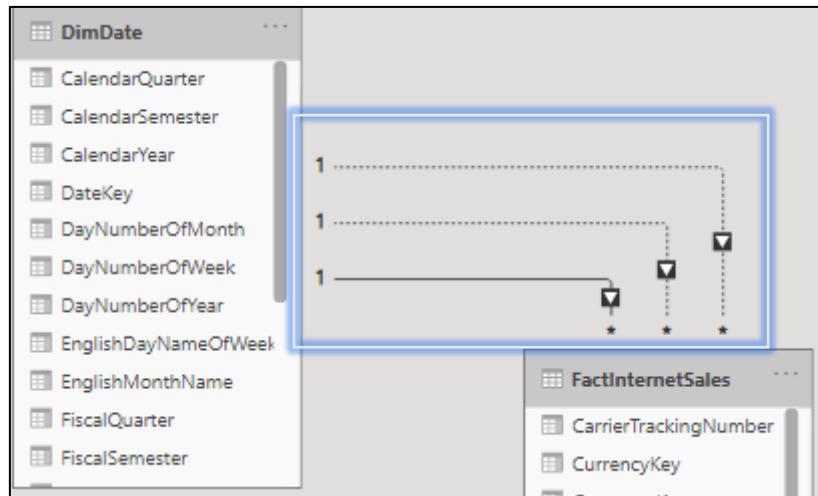


Tipos de Relacionamentos

Mas um ponto interessante, é entendermos que existem alguns tipos de relacionamentos, conforme abaixo.

Os Relacionamentos Ativos (-----)

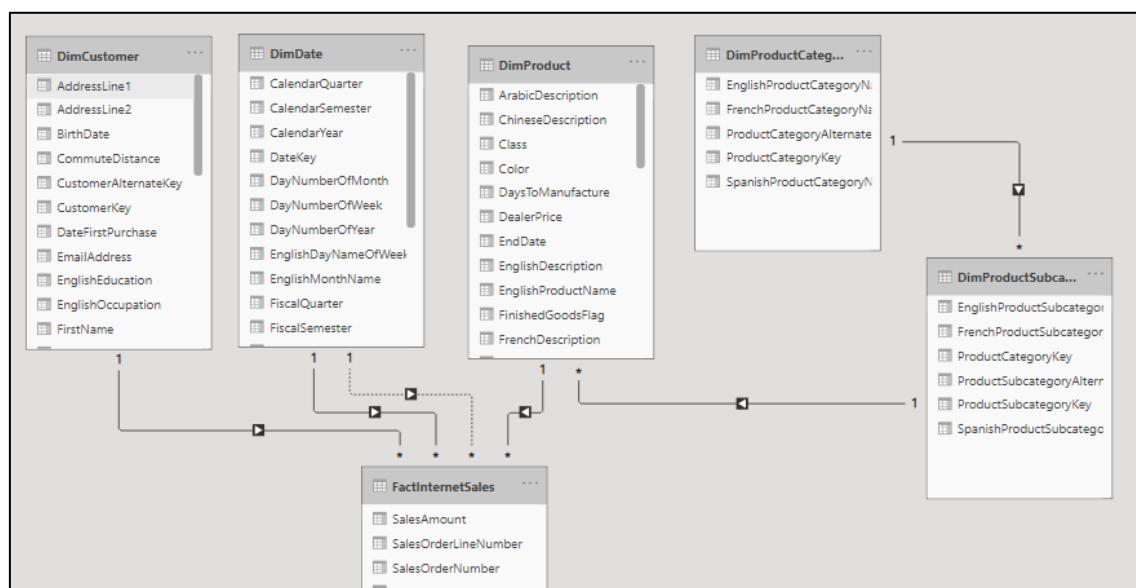
Os Relacionamentos Inativos (- - - - -)



É muito importante verificarmos se os campos que foram relacionados estão corretos, sempre verifique! Algumas vezes, pode ser que os campos não estejam relacionados corretamente!

E o **Power BI** faz isso, porque ele acredita que estão relacionados e nós também podemos estabelecer várias e várias relações da forma que nós queremos.

Lembre-se da ordem, as Dimensões ficam em cima e os Fatos, embaixo, pois Dimensões filtram Fatos.. E, também, quanto menor o nível, é filtrado do Maior para o Menor.



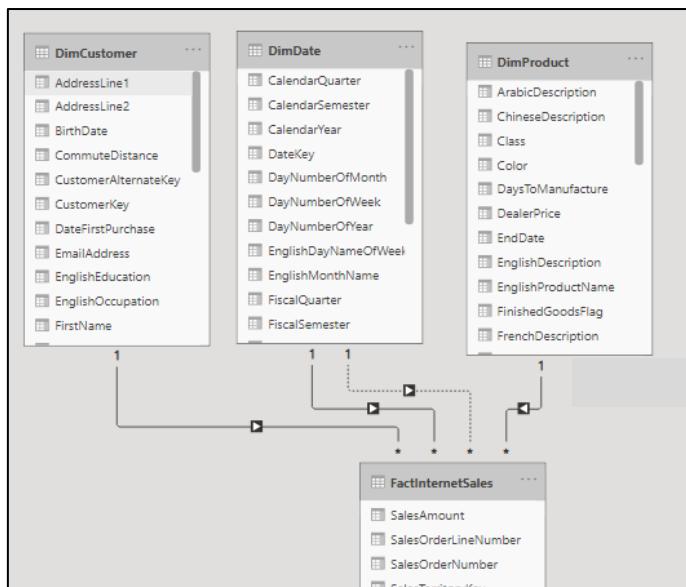
Esse modelo é conhecido como SNOWFLAKES, não é a melhor opção de trabalho, a melhor é se fosse o STAR.

O Modelo SnowFlakes, é uma tabela (DimProductSubCategory) que é filtrada pela (DimProductCategory) que por final é filtrada um Fato (FactInternetSales). Imagine que não tivéssemos as duas consultas da direita, esse seria o modelo Star.

Modelo SNOWFLAKES



Modelo STAR



Voltemos ao nosso *Dashboard* na Visualização.

Cruzando duas Tabelas Fatos

Para esse exercício, utilizaremos a planilha “18 – Setores.xlsx”.



Esse exercício não é para ETL, ou seja, não é sobre tratamento de dados.

Veja que na planilha, existem duas abas, uma de “Compras (Outcome)” e outra de “Vendas(Income)”.

1	IdSetor	Setor	Data
2	1	Operação	11/2/20
3	2	Contabilidade	11/2/20
4	3	Comercial	12/2/20
5	4	Marketing	13/2/20
6	5	Segurança	14/2/20
7	6	TI	16/2/20
8	7	RH	16/3/20
9	1	Operação	03/11/
10	2	Contabilidade	03/11/
11	3	Comercial	13/3/20
12	4	Marketing	13/3/20

1	IdSetor	Setor	Data
	1	Operação	13/2/
	3	Comercial	13/2/
	3	Comercial	14/2/
	4	Marketing	15/2/
	1	Operação	17/2/
	4	Marketing	17/2/
	1	Operação	18/2/
	1	Operação	13/2/
	3	Comercial	13/2/
	3	Comercial	14/2/
	4	Marketing	15/2/

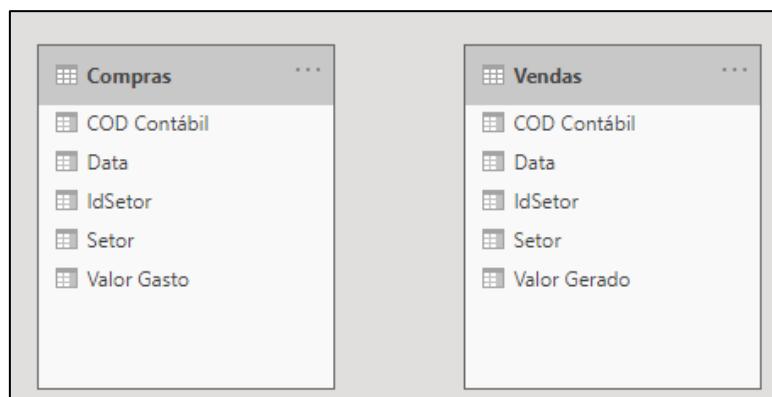
Vamos entender uma coisa também. Repare que o setor de RH tem o IdSetor de número 7. Mas existem dois códigos diferentes, um de RHSAL(Salário) e outro de RHFER(Férias). Mesmo que eles contêm como mesmo IdSetor.

IdSetor	Setor	Data	COD Contábil	Valor Gasto
7	RH	16/3/2019	RHSAL	420,10
7	RH	16/3/2019	RHSAL	420,10
7	RH	16/4/2019	RHSAL	420,10
7	RH	16/4/2019	RHFER	420,10

Isso vale tanto para venda quanto para compra.

Após carregarmos e configurarmos a tipagem correta, vamos ao **Power Pivot (Modelo)**.

Não existe nenhum relacionamento entre as tabelas até o momento. O relacionamento ainda não foi modelado.



Veja que ao tentarmos fazer uma tabela “comparativa”, encontraremos um erro.

Valor Gasto	Valor Gerado	Setor
R\$ 1.260,60	R\$ 17.048,00	Comercial
R\$ 2.881,80	R\$ 17.048,00	Contabilidade
R\$ 1.382,76	R\$ 17.048,00	Marketing
R\$ 2.701,80	R\$ 17.048,00	Operação
R\$ 1.680,40	R\$ 17.048,00	RH
R\$ 2.670,45	R\$ 17.048,00	Segurança
R\$ 2.700,30	R\$ 17.048,00	TI
R\$ 15.278,11	R\$ 17.048,00	

Valor Gasto	Valor Gerado	Setor
R\$ 15.278,11	R\$ 3.000,00	Comercial
R\$ 15.278,11	R\$ 5.000,00	Diretoria
R\$ 15.278,11	R\$ 3.000,00	Marketing
R\$ 15.278,11	R\$ 6.048,00	Operação
R\$ 15.278,11	R\$ 17.048,00	

Se seguirmos pedindo o “Setor” ou de Compras ou de Vendas, esse erro acima será apresentado.

Valores

Valor Gasto	▼ X
Valor Gerado	▼ X
Setor	▼ X

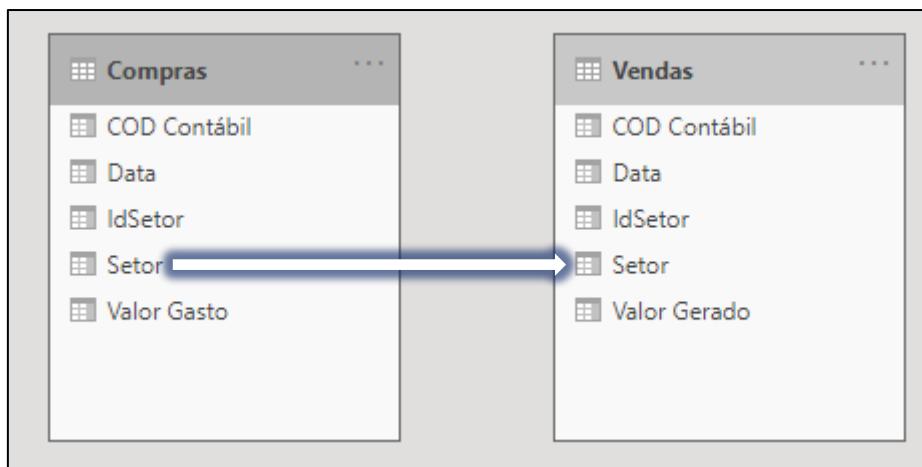
E não podemos colocar dois setores, pois não será possível exibir o visual.

Valores

Valor Gasto	▼ X
Valor Gerado	▼ X
Setor	▼ X
Setor	▼ X

X
Não é possível exibir o visual. [Ver detalhes](#)
Corrigir isto

Isso acontece porque não tem nenhuma relação, nenhum relacionamento entre as tabelas.



Na próxima página iniciaremos o relacionamento.

Precisamos, portanto, ligar o **setor** de uma tabela com o **setor** da outra.

Na hora que fizermos isso.

Criar relacionamento

Selecione tabelas e colunas relacionadas umas às outras.

Compras				
IdSetor	Setor	Data	COD Contábil	Valor Gasto
1	Operação	segunda-feira, 11 de fevereiro de 2019	OPRDAT	R\$ 900,60
2	Contabilidade	segunda-feira, 11 de fevereiro de 2019	CT09	R\$ 960,60
3	Comercial	terça-feira, 12 de fevereiro de 2019	CM102	R\$ 420,20

Vendas				
IdSetor	Setor	Data	COD Contábil	Valor Gerado
1	Operação	terça-feira, 13 de fevereiro de 2019	OPR09	R\$ 500,00
3	Comercial	quarta-feira, 13 de fevereiro de 2019	CMCB	R\$ 500,00
3	Comercial	quinto-feira, 14 de fevereiro de 2019	CM102	R\$ 250,00

Cardinalidade Direção do filtro cruzado

Muitos para muitos (*;*) Ambas

Ativar este relacionamento Aplicar filtro de segurança em ambos os sentidos

Desativar cardinalidade de referência

Note a CARDINALIDADE é de “muitos para muitos”.

Veja que isso vai dar erro! Porque não podemos cruzar fatos. Um fato é a VENDA e outro fato é a COMPRA.

Quando aplicamos o filtro, uma tabela em um determinado momento, tem preferência do filtro sobre a outra.

The screenshot displays two separate Power BI reports side-by-side, each showing a table with three columns: Valor Gasto, Valor Gerado, and Setor. The top report's filters pane shows both the Compras and Vendas tables expanded, with the Setor column checked under both. The bottom report's filters pane shows the same setup, but only the Vendas table's Setor column is checked, while the Compras table's Setor column is unchecked. This visualizes how the 'Muitos para muitos' cardinality setting can lead to different filtering behaviors across related tables.

O “erro” acima acontece porque no determinado momento do filtro, ainda mesmo com o relacionamento criado, um FATO tem predominância sobre o outro FATO, por exemplo, na primeira, o FATO da COMPRA tem predominância sobre o FATO da VENDA e na segunda imagem foi o contrário.

Modelagem dos Dados de Dois Fatos

Vamos entender o que está acontecendo com os dados, com a planilha.

Sabemos que temos dois fatos, certo? Um fato que foi a **Venda** e outro fato que foi a **Compra**. E não podemos cruzar esses dois.

1	IdSetor	Setor	Data	1	IdSetor	Setor	Data
2	1 Operação		11/2/20	2	1 Operação		13/2/
3	2 Contabilidade		11/2/20	3	3 Comercial		13/2/
4	3 Comercial		12/2/20	3	Comercial		14/2/
5	4 Marketing		13/2/20	4	Marketing		15/2/
6	5 Segurança		14/2/20	1	Operação		17/2/
7	6 TI		16/2/20	4	Marketing		17/2/
8	7 RH		16/3/20	1	Operação		18/2/
9	1 Operação		03/11/	1	Operação		13/2/
10	2 Contabilidade		03/11/	3	Comercial		13/2/
11	3 Comercial		13/3/20	3	Comercial		14/2/
12	4 Marketing		13/3/20	4	Marketing		15/2/

Outra informação que temos é sobre as dimensões.

Quais são as dimensões que temos? É o do **setor**. E por que não temos a dimensão da COD Contábil? Porque ela é apenas “uma ramificação do setor”, pois é a dimensão SETOR que determina os **FATOS**, uma dimensão SETOR realiza um fato de **VENDA** e uma dimensão de SETOR que realiza um fato de **COMPRA**.

IdSetor	Setor	Data	COD Contábil	Valor Geral
1 Operação		13/2/2018	OPR09	500,00
3 Comercial		13/2/2019	CMCB	500,00
3 Comercial		14/2/2019	CM102	250,00
4 Marketing		15/2/2019	MKTNATAL	500,00
1 Operação		17/2/2019	OPRDAT	500,00
4 Marketing		17/2/2019	MKTNATAL	250,00
1 Operação		18/2/2019	OPRDAT	512,00

Nomenclatura correta

Ou seja, até o momento, as informações que podemos extrair são:

Fatos

- 1 - **ftVendas**
- 2 - **ftCompras**

Dimensão

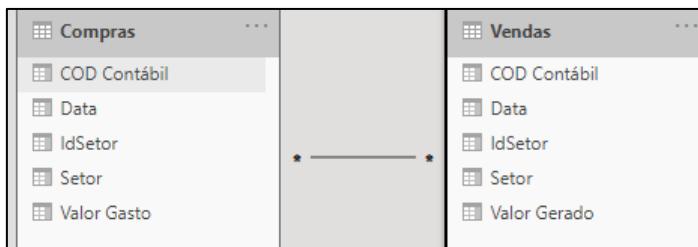
- 1 - **dimSetores**

Fatos se devem iniciar com “**ft**” e dimensões com “**dim**”.

Poderíamos criar uma dimensão de Data também, mas nesse caso, deixaremos a data que o fato ocorreu. Aliás, a dimensão de Data, geralmente são gigantescas, pois contam com o ano, mês, semanas...

Criando as Fatos e Dimensão

Voltando, portanto, ao exercício, no **Power Pivot**, vamos excluir aquela relação que foi estabelecida previamente.



Clicando com o “Botão direito na linha > Excluir”.

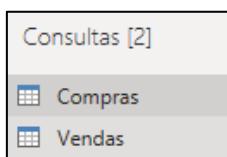
Agora, sem nenhum relacionamento entre as tabelas, sairemos do **Power Pivot** e vamos editar as consultas/transformar dados no **Power Query**.

Repare que no fato de Vendas, existem mais Setores que existem no fato de Compras.

	A	B	C	D	E
1	IdSetor	Setor	Data	COD Contábil	Valor Gerado
26	4	Marketing	15/2/2019	MKTVER	500,00
27	1	Operação	17/2/2019	OPRGER	500,00
28	4	Marketing	17/2/2019	MKTVER	250,00
29	1	Operação	18/2/2019	OPRDATA	512,00
30	8	Diretoria	25/4/2019	DIR02	1500,00
31	8	Diretoria	25/5/2019	DIR02	3500,00
32					

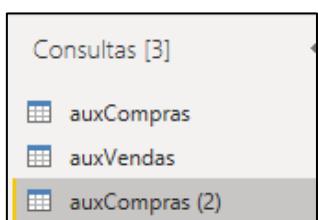
Por isso, vamos “*appendar*” a Vendas com Compras na coluna dos Setores.

E em resumo, para entender, existem duas fatos, atualmente, correto?



Então, precisaremos que saiam mais pelo menos 3 consultas, sendo duas fatos (ftVendas e FtCompras) e uma dimensão que é a dimSetores.

Vamos duplicar a Compras



Com a “auxCompras(2)” selecionada, em “Página Inicial > Combinar > Acrescentar Consultas > Acrescentar Consultas” e acrescente a consulta de auxVendas.

Se quiséssemos também pelo COD Contábil, faríamos os próximos passos abaixo, só que ao invés de realizar pelo Setor, realizar para a COD Contábil.

Selecionado o Setor e o IdSetor, vamos remover as outras colunas.

	1 ² ₃ IdSetor	A ^B _C Setor
1		1 Operação
2		2 Contabilidade
3		3 Comercial

Selecionando a coluna do IdSetor, clique com o [Botão Direito > Remover Duplicatas](#).

O que aconteceu? Criamos a nossa dimensão de Setores!

	1 ² ₃ IdSetor	A ^B _C Setor
1		1 Operação
2		2 Contabilidade
3		3 Comercial
4		4 Marketing
5		5 Segurança
6		6 TI
7		7 RH
8		8 Diretoria

Agora precisamos duplicar as duas fatos, a de Compras e Vendas e aproveitar para renomear cada uma.

Consultas [5]
auxCompras
auxVendas
dimSetores
ftCompras
ftVendas

Em cada uma das fatos, mas começando pela de **ftCompras**.

Não precisamos dos setores e nem da COD Contábil, apenas do IdSetor, Data e o Valor.

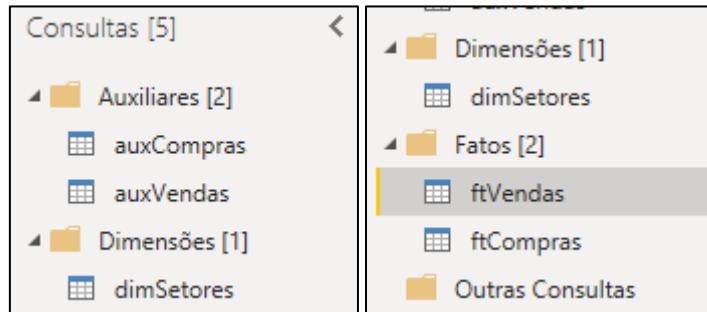
	1 ² ₃ IdSetor	Data	1.2 Valor Gasto
1		11/02/2019	900,6
2		11/02/2019	960,6

	1 ² ₃ IdSetor	Data	1.2 Valor Gerado
1		13/02/2018	500
2		13/02/2019	500

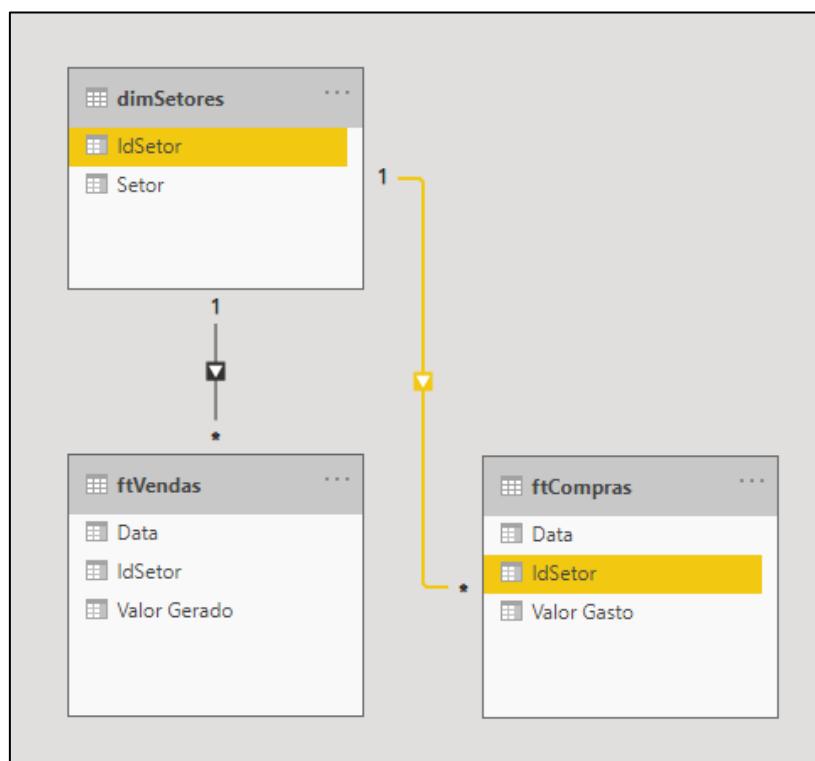
Criando Grupos

Para melhor organizar o dataset, é sempre bom criar “grupos” para cada um.

Por exemplo, em qualquer espaço em branco em Consultas > “Botão Direito > Novo Grupo”. E depois arraste as consultas para cada uma de suas respectivas pastas.



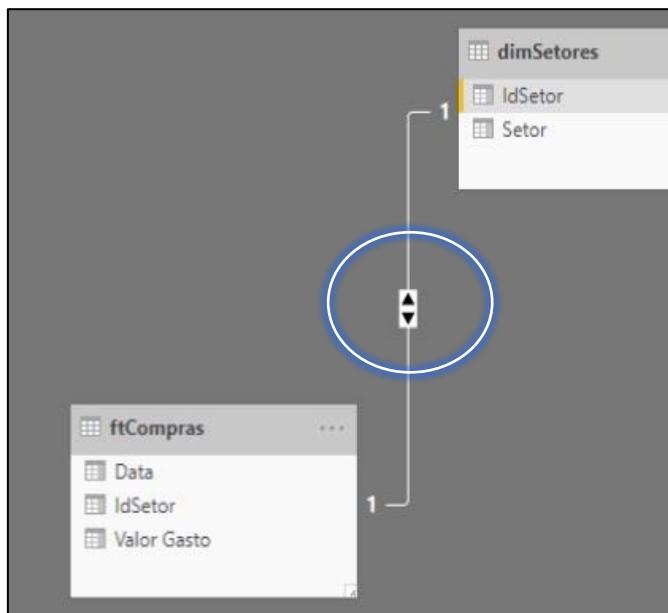
Ao clicarmos para aplicar e fechar, voltemos ao **Power Pivot** e vejamos as tabelas.



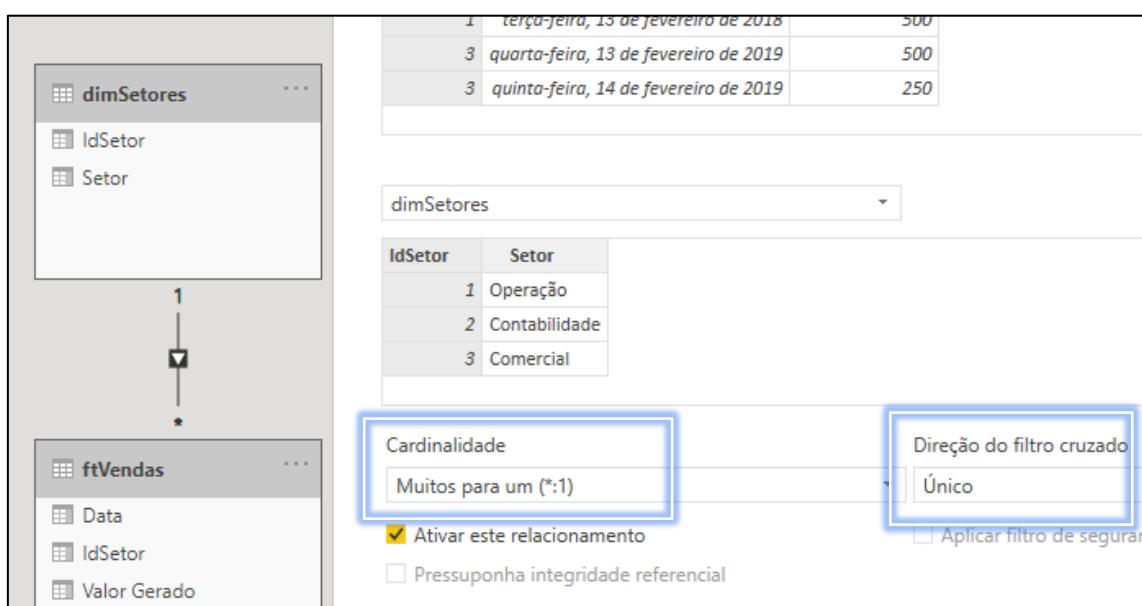
Pode ser que dependendo da atualização do **Power BI**, ele mesmo se encarregará de estabelecer o relacionamento, apenas verifique se foi estabelecido de forma correta.

Cardinalidade e Filtros

Caso, o Power BI não crie automaticamente, é preciso clicar com o botão direito na relação e Propriedades. Pois não pode ocorrer o “Cross-Filter”, que é o que acontece abaixo.



Além de que o relacionamento está de “Um para Um”. Ou seja, a ftCompras está filtrando a dimSetores. Isso é errado, a Cardinalidade correta é de “Um para Muitos”, conforme abaixo.

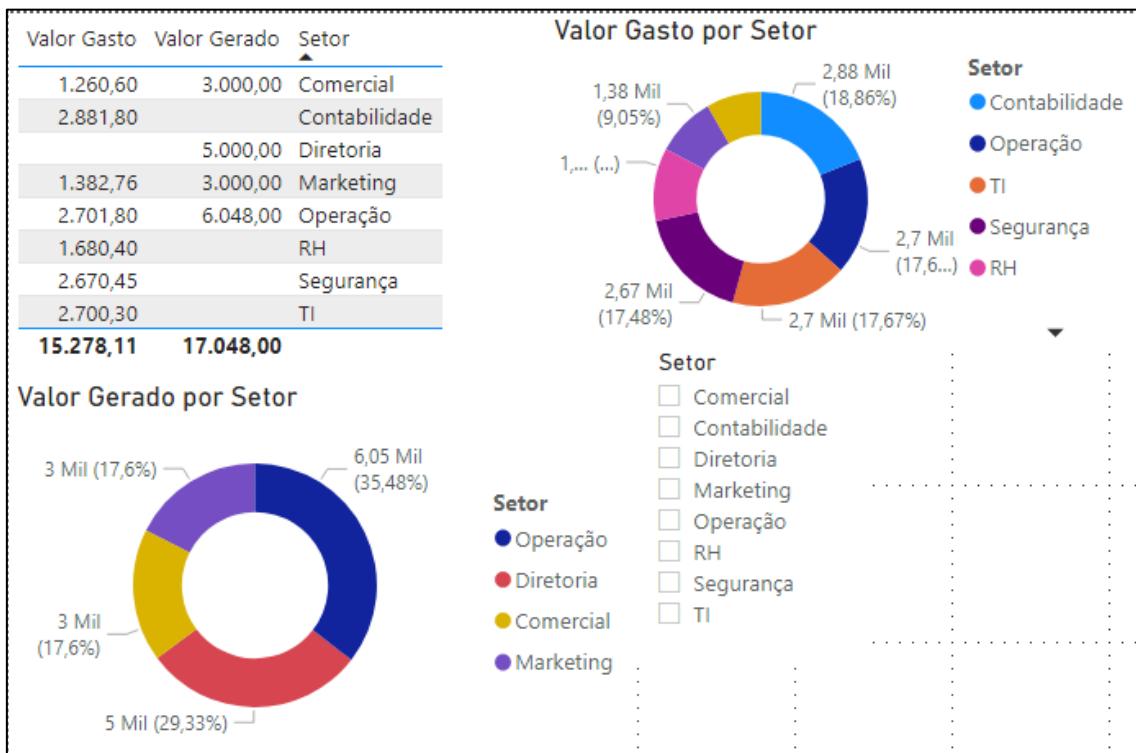


Lembre-se também que a “Direção do Filtro Cruzado” dever ser única. Somente um deve filtrar o muitos.

Veja que agora está tudo de acordo.

The screenshot shows the Power BI Data View interface. On the left is a table with three columns: 'Valor Gasto', 'Valor Gerado', and 'Setor'. The data includes rows for Comercial, Contabilidade, Diretoria, Marketing, Operação, RH, Segurança, and TI, with summary values at the bottom: 15.278,11 and 17.048,00. To the right is the data model pane, which displays three fact tables: 'dimSetores' (with columns IdSetor and Setor), 'ftCompras' (with columns Data, IdSetor, and Valor Gasto), and 'ftVendas' (with columns Data, IdSetor, and Valor Gerado). The 'Valores' section contains three dropdowns for 'Valor Gasto', 'Valor Gerado', and 'Setor'.

Poderíamos aplicar diversos gráficos.



Organizando Relacionamentos e Regra de Negócio

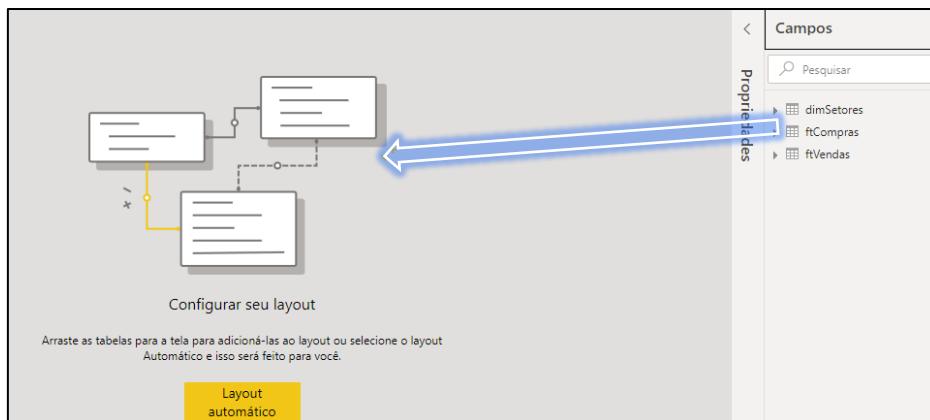
Quando se trabalha com um DW muito grande, é comum inúmeras tabelas aparecerem no **Power Pivot**, e para trabalhar e criar os relacionamentos, é um pesadelo.

Por isso, deve-se levar em consideração, que devemos trabalhar inicialmente, com cada parte do negócio isoladamente. Por exemplo, quando estamos analisando as Vendas, precisamos apenas trabalhar na análise de Vendas. Deve ser a regra de negócio da venda, criar todos os relacionamentos de Vendas. Ou seja, separe os fatos.

No **Power Pivot**, clique no “sinal de +” para abrir uma nova janela.

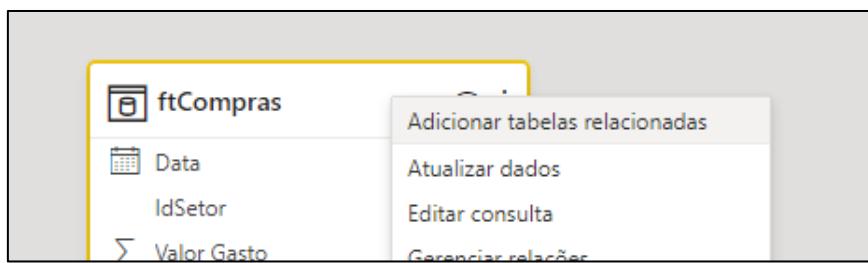


Não clique em “Layout Automático”, pois queremos configurar manualmente, conforme o nosso interesse do negócio.



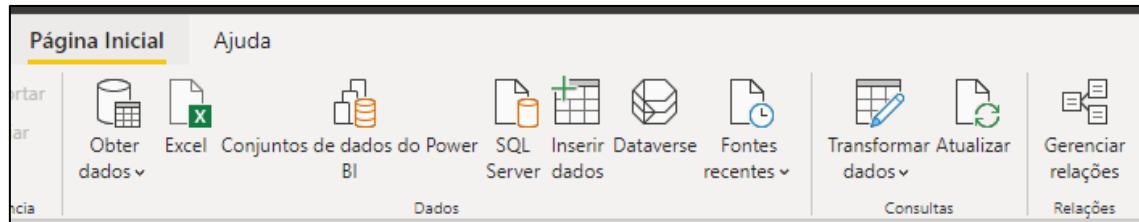
Clique e arraste o “ftCompras” para dentro do Layout.

Com o “Botão Direito > Adicionar Tabelas Relacionadas”. Esse será o resultado de apenas tabelas relacionadas ao Fato de Compras e não ao fato de vendas ou quaisquer outros que existam.



Gerenciar Relações no Geral

Para saber todas as relações ativas no Power Pivot, procure por “[Página Inicial > Relacionamentos > Gerenciar Relações](#)”.



Ao clicar, uma tela apresentará todas as relações estabelecidas dentro de Todas as Tabelas (a aba original e criada automaticamente pelo Power BI).

Gerenciar relações

Ativo	De: Tabela (Coluna)	Para: Tabela (Coluna)
<input checked="" type="checkbox"/>	ftCompras (IdSetor)	dimSetores (IdSetor)
<input checked="" type="checkbox"/>	ftVendas (IdSetor)	dimSetores (IdSetor)

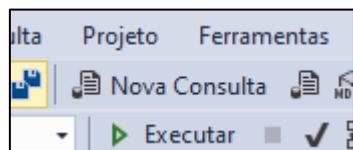
Dimensão Tempo

Em todo BI é necessário ter uma dimensão de tempo. O BI é justamente para entender o que aconteceu no passado para a tomada de decisão. Enquanto o Data Science é para entender e realizar previsões do futuro. Uma área complementa a outra. O BI é reativo e enquanto o *Machine Learning* é preditivo.

Dimensão Tempo no Banco de Dados

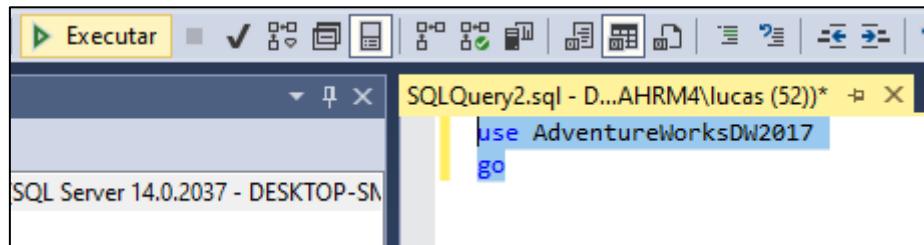
Vamos novamente, iniciar o Banco de Dados.

Clique em “Nova Consulta”.



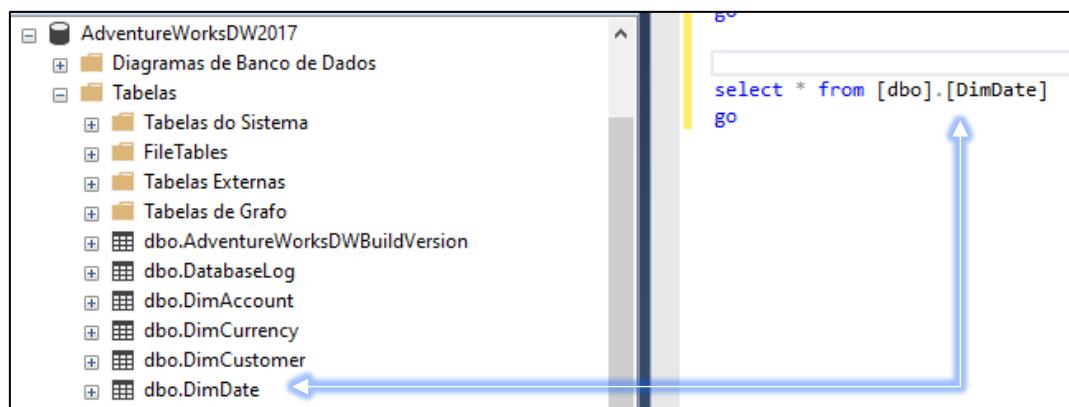
Para escrevermos um script de Consulta (Não vamos escrever um Banco de Dados! É um script para consulta).

Ao escrevermos o código abaixo e Executarmos.



Nos conectamos ao BD do “AdventureWorksDW2017” para acessar os seus dados.

Veja que possuímos uma dimensão tempo/calendário, é a “dbo.DimDate”.



Vamos selecionar a consulta. Podemos puxar e arrastar que ganharemos produtividade.

Veja que esse resultado da consulta é o calendário, é tempo, é data.

The screenshot shows a SQL query window with the following content:

```
select * from [dbo].[DimDate]
go
```

The results pane displays the following data:

	DateKey	FullDateAlternateKey	DayNumberOfWeek	EnglishDayNameOfWeek	SpanishDayNameOfWeek	FrenchDayNameOfWeek	DayNumberOfMonth	DayNur
1	20050101	2005-01-01	7	Saturday	Sábado	Samedi	1	1
2	20050102	2005-01-02	1	Sunday	Domingo	Dimanche	2	2
3	20050103	2005-01-03	2	Monday	Lunes	Lundi	3	3

Utilizaremos o “19 – Carga Dimensão PO.sql”. É um script de Banco de Dados criado pelo instrutor do curso, o Felipe Mafra.



Vamos, então, abrir esse script. Com o atalho (CTRL + O), abra o arquivo.

The screenshot shows a SQL query window titled "19 - CARGA DIMENSAO TEMPO.lucas (54)" with the following content:

```
--CRIANDO A DIMENSÃO TEMPO--
```

```
CREATE TABLE DIM_TEMPO(
    IDSK INT PRIMARY KEY IDENTITY(1,1),
    DATA DATE,
    DIA CHAR(2),
    DIASEMANA VARCHAR(10),
    MES CHAR(2),
    NOMEMES VARCHAR(10),
    QUARTO TINYINT,
    NOMEQUARTO VARCHAR(10),
```

Essa dimensão tempo pode ser criada pelo DBA. Caso ela não tenha, criaremos no Power BI em algumas páginas a frente através da Linguagem M.

Nesse script ainda não estava com a criação e a utilização do DB.

Vamos escrever o script para iniciarmos o banco de dados “calendar”.

The screenshot shows the Object Explorer and a SQL query window. The Object Explorer shows a connection to "DESKTOP-SMAHRM4\SQLEXPRESS (SQL Server 14.0.2035)". The "Bancos de Dados" node is expanded, showing "Bancos de Dados do Sistema", "Instantâneos do Banco de Dados", "AdventureWorks2017", "AdventureWorksDW2017", and "calendar". The SQL query window contains the following script:

```
create DATABASE calendar
go

use calendar
go
```

Below the query window, a dashed line indicates the continuation of the script with the text "--CRIANDO A DTMENSÃO".

Veja o resultado desse script.

Lembre-se que para consultar é preciso “ select * from DIM_TEMPO”

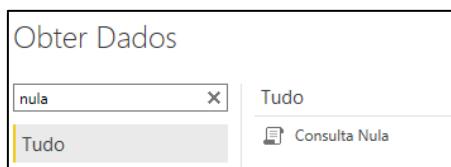
IDSK	DATA	DIA	DIASEMANA	MES	NOMEMES	QUARTO	NOMEQUARTO	ANO	ESTACAOANO	FIMSEMANA	DATACOMPLE	
1	50000	1950-01-01	01	Domingo	01	Janeiro	1	Primeiro	1950	Verão	Sim	19500101
2	50001	1950-01-02	02	Segunda	01	Janeiro	1	Primeiro	1950	Verão	Não	19500102
3	50002	1950-01-03	03	Terça	01	Janeiro	1	Primeiro	1950	Verão	Não	19500103
4	50003	1950-01-04	04	Quarta	01	Janeiro	1	Primeiro	1950	Verão	Não	19500104
5	50004	1950-01-05	05	Quinta	01	Janeiro	1	Primeiro	1950	Verão	Não	19500105
6	50005	1950-01-06	06	Sexta	01	Janeiro	1	Primeiro	1950	Verão	Não	19500106

Agora aprenderemos como criar essa dimensão tempo dentro e diretamente no Power BI. Acima, foi criado no Banco de Dados para entendimento. Abaixo criaremos uma Dimensão de Calendário/ Tempo/ Data.

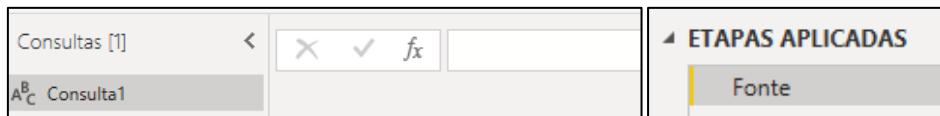
Consulta Nula

Para criamos um script em uma Linguagem M, precisaremos de uma **Consulta Nula**.

Para isso, clique em “Obter Dados > Consulta Nula”.

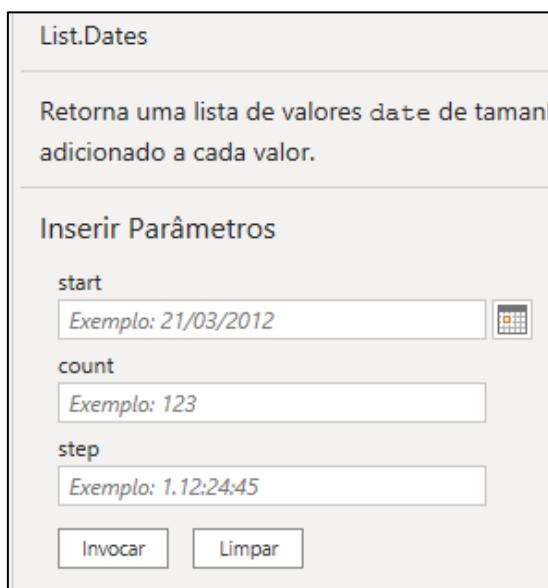


Veja que o prompt está aberto para a escrita.



Digite: “ = List.Dates ” e aperte o Enter. Não se esqueça do (=).

Agora vai aparecer uma interface para o usuário a preencher.



Vamos interpretar os Parâmetros acima.

- **Start:** O dia de início do negócio, pode ser qualquer dia, mas o melhor é o dia real e efetivo. *Formato:* dd/mm/yyyy.
- **Count:** É o contador de anos. O valor deve ser em dias. Por exemplo, se desejo dois anos, o valor será de 730, se o valor for 10 anos, será de 3.650.
- **Step:** São os pulos dos dias. Se quisermos de 1 em 1, o valor será 1.

Clique em “Invocar”, já que é uma função.

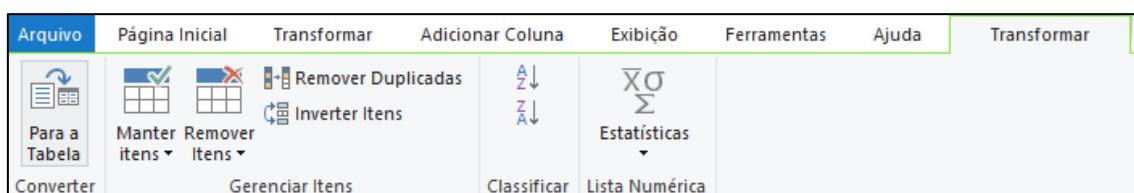
Lista	
1	01/01/2010
2	02/01/2010
3	03/01/2010
4	04/01/2010
5	05/01/2010
6	06/01/2010

Repare que nas Consultas, um novo ícone agora é conhecido.

A screenshot of the Microsoft Power BI ribbon. The 'Consultas [2]' section is selected. Within this section, the 'Função Invocada' icon (a grey square with a white grid) is highlighted, indicating it is the active function.

Para podermos trabalhar, precisaremos **transformar** a Função Invocada em uma Tabela.

Logo quando é criada, ela já abre uma aba nas guias.



Clique em “Transformar > Converter > Para a Tabela”. E o ícone já será alterado.

A screenshot of the Microsoft Power BI ribbon. The 'Consultas [2]' section is selected. The 'Função Invocada' icon is no longer highlighted, indicating the transformation process has been completed.

Não se esqueça de converter para Data e renomear a Coluna!



Agora precisamos definir um ID para essas datas, correto?

Para isso, clique e selecione a coluna da Data.

Vá em “Adicionar Coluna > Coluna de Índice > A partir de 1”.

The screenshot shows the Power BI Data Editor interface. On the left, there's a 'Transformar' pane with a 'Coluna de Índice' section open, showing options 'De 0' and 'De 1'. The 'De 1' option is selected. On the right, a table is displayed with two columns: 'idData' (containing values 1 through 10) and 'Data' (containing dates from 01/01/2010 to 10/01/2010).

1 ² 3 idData	Data
1	01/01/2010
2	02/01/2010
3	03/01/2010
4	04/01/2010
5	05/01/2010
6	06/01/2010
7	07/01/2010
8	08/01/2010
9	09/01/2010
10	10/01/2010

Agora podemos realizar aquela extração de dados, ou seja, extrair o ano, mês, nome do mês, dias...

“Adicionar Coluna > Data e Hora > Data”.

The screenshot shows the Power BI Data Editor with a table containing four columns: 'Data', 'Ano', 'Mês', and 'Nome do Mês'. All rows show the same data: '01/01/2010' in 'Data', '2010' in 'Ano', '2010' in 'Mês', and 'janeiro' in 'Nome do Mês'.

Data	Ano	Mês	Nome do Mês
01/01/2010	2010	2010	janeiro
02/01/2010	2010	2010	janeiro
03/01/2010	2010	2010	janeiro
04/01/2010	2010	2010	janeiro
05/01/2010	2010	2010	janeiro

Para uma boa dimensão de calendário, o interessante é extrair muitas variações.

The screenshot shows the Power BI Data Editor with a table containing five columns: 'Ano', 'Mês', 'Nome do Mês', 'Semana do Mês', and 'Nome do Dia'. The data shows various combinations of these dimensions, such as '2010' in 'Ano' and '1 janeiro' in 'Nome do Mês'.

Ano	Mês	Nome do Mês	Semana do Mês	Nome do Dia
2010	1	janeiro		sexta-feira
2010	1	janeiro		sábado
2010	1	janeiro	2	domingo
2010	1	janeiro	2	segunda-feira

Vamos criar uma coluna condicional com base no “Dia da Semana” onde “0 e 6” são os finais de semana.

The screenshot shows the Power BI Data Editor with a table containing two columns: 'Dia da Semana' and 'Nome do Dia'. The 'Dia da Semana' column has values 5, 6, 0, and 1, which correspond to the days 'sexta-feira', 'sábado', 'domingo', and 'segunda-feira' respectively.

Dia da Semana	Nome do Dia
5	sexta-feira
6	sábado
0	domingo
1	segunda-feira

Através da Coluna Condisional.

Adicionar Coluna Condisional

Adicionar uma coluna condicional que é calculada das outras colunas ou valores.

Nome da nova coluna
Final de Semana

Nome da Coluna	Operador	Valor ⓘ	Saída ⓘ
Se	dia da semana	igual a	ABC 123 0 Então ABC 123 Sim
Senão...	dia da semana	igual a	ABC 123 6 Então ABC 123 Sim

Adicionar Cláusula

Senão ⓘ ABC 123 Não

E o resultado.

1 2 3 Dia da Semana	A B C Nome do Dia	ABC 123 Final de Semana
5	sexta-feira	Não
6	sábado	Sim
0	domingo	Sim
1	segunda-feira	Não
2	terça-feira	Não
3	quarta-feira	Não
4	quinta-feira	Não
5	sexta-feira	Não
6	sábado	Sim

E se criarmos um Primeiro Semestre e Segundo Semestre?

ABC 123 Semestre				
Primeiro Semestre				
Primeiro Semestre				
Primeiro Semestre				
181 junho	5	3 quarta-feira	Não	Primeiro Se
182 julho	1	4 quinta-feira	Não	Segundo Se

Introdução a Linguagem M

A Linguagem M é composta basicamente de Funções. Para ver mais, consulte a documentação no site da *Microsoft – PowerQuery-m*.

Para trabalhar com a Linguagem M é preciso ter uma consulta pelo menos.

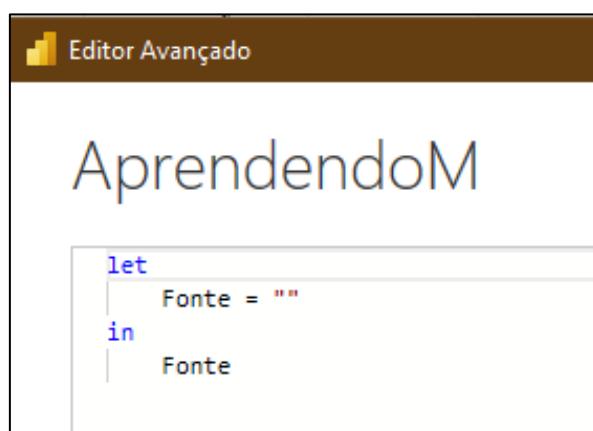
Para iniciarmos o entendimento, uma breve introdução ao menos, vamos clicar em “Obter Dados > Consulta Nula/Branco”.

Existem **duas** formas de trabalhar com a Linguagem M:

Através da **Barra de Fórmulas**.



Ou pelo **Editor Avançado**. Clicando com o “Botão Direito na Consulta selecionada > Editor Avançado”. Ou em “Exibição > Editor Avançado”.



O “Let” é onde manipulamos as variáveis.

O “In” serão as saídas. O que será mostrado.

Nota: Não poderemos mostrar nada no “In” se não foi declarado no “Let”.

Veja que abaixo, iniciamos uma variável 5.



E queremos mostrar “Fonte”. Isso vai dar um erro.

E qual será o erro? Que a fonte não foi inicializada.

A screenshot of a programming interface titled 'Consultas [1]'. In the top right, there are icons for close, save, and execute, followed by a search bar containing 'let'. Below the search bar, a message box displays an error: 'AprendendoM' with an exclamation mark, and 'Erro na consulta ". Expression.Error: O nome 'Fonte' não foi reconhecido. Certifique-se de que ele foi soletrado corretamente.'

Mas se no "In" pedimos para mostrar a variável A, não teremos erro.

A screenshot of a programming interface titled 'AprendendoM'. The code area contains:

```
let
|   A = 5
in
|   A
```

The output window shows the result of the execution.

Veja o resultado.

A screenshot of a programming interface titled 'Consultas [1]'. The code area shows '1 2 3 AprendendoM'. The output window shows the result of the execution: '5'. To the right, a panel titled 'ETAPAS APLICADAS' shows the step 'A'.

Separação por Vírgula

É importante saber que só passaremos para a linha de baixo, após separarmos por vírgula, a separação de variáveis é por vírgula.

Veja que teremos um erro se mantivermos assim como está a esquerda. O correto é o quadro a direita, note a vírgula após declararmos o A = 5.

A screenshot of a programming interface titled 'AprendendoM'. It shows two examples of code:
Left example:

```
let
|   A = 5
|   B = 10
in
|   A
```

Right example:

```
let
|   A = 5,
|   B = 10
in
|   A
```

É muito comum, na segunda linha e em diante, essa variável receber as informações contidas na linha acima, para continuar a operação. Em breve veremos isso acontecer.

Podemos imprimir variáveis como:

A screenshot of a programming interface titled 'AprendendoM'. The code area contains:

```
let
|   A = 5,
|   B = 10,
|   C = A + B
in
|   C
```

Veja que o resultado será 15. E as etapas sendo criadas.

Consultas [1] < X ✓ fx = A + B

1²3 AprendendoM 15

ETAPAS APLICADAS

A
B
C

Tipo dos Dados na Linguagem M

Números são declarados sem aspas duplas (5). Enquanto Textos, são declarados com aspas duplas ("5"). O sinal da concatenação no BI é o E Comercial (&).

AprendendoM

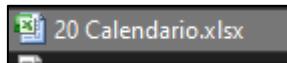
```
let
  A = "5",
  B = "10"
in
  A & " " & B
```

Consultas [1] < X ✓ fx = let

A^BC AprendendoM 5 10

Datas Dinâmicas

Para esse exercício, utilizaremos o arquivo “20 – Calendário”.



Conectamos, transformamos os dados e renomeamos para a ftEstoque.

A screenshot of the Power BI Data Editor interface. The title bar says 'Consultas [1]'. A table named 'ftEstoque' is selected. The columns are labeled 'ID', 'DataContagem', and 'Unidades'. The data shows two rows: Row 1 with ID 1, DataContagem '01/01/2018', and Unidades 3503; Row 2 with ID 2, DataContagem '02/01/2018', and Unidades 3524.

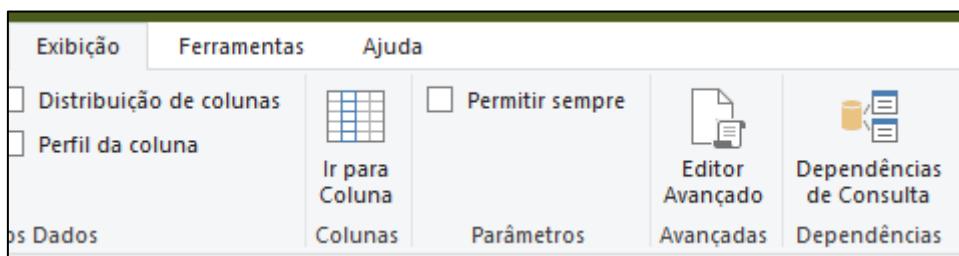
ID	DataContagem	Unidades
1	01/01/2018	3503
2	02/01/2018	3524

Dessa vez, vamos criar uma dimCalendário através da **Linguagem M**.

“Obter Dados/Nova Fonte > Consulta Nula/Em Branco”.

Após criarmos a dimensão de tempo igual ao que fizemos no capítulo anterior, vamos agora para a parte que interessa.

Em **“Exibição > Dependências > Dependências de Consulta”**.



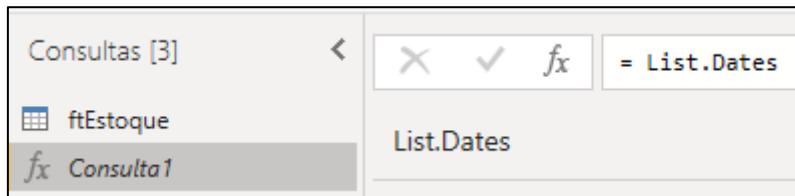
Veja que ao clicarmos, ele nos mostrará quais são as consultas que dependem uma da outra.



Nota-se que a dimCalendario, é dependente da *Consulta1* (que foi a Nula que criamos do zero para fazer essa dimensão).

Porém, a *Consulta1* não é nada mais que apenas uma consulta “inútil”, pois ela contém apenas a invocação da data de início, o final e o passo. Ela poderia ser “descartada”.

Veja que na *Consulta1*, apenas existe a fórmula do “List.Dates”



O que isso significa? Que podemos ir à dimCalendário e lá no Editor Avançado, nós podemos agora, jogar essa função diretamente no Editor e ele será o “encarregado” dessa função, e não será mais necessário a *Consulta1*.

Clicando no “Editor Avançado” na dimCalendario.

A screenshot of the Power BI Advanced Editor for the 'dimCalendario' table. The title 'dimCalendario' is at the top. The M code in the editor is as follows:

```
let
    Fonte = Consulta1(#date(2017, 1, 1), 3650, #duration(1, 0, 0, 0)),
    #"Convertido para Tabela" = Table.FromList(Fonte, Splitter.SplitByNothing()),
    #"Tipo Alterado" = Table.TransformColumnTypes(#"Convertido para Tabela", {
        {"Data", typeDate}
    }),
    #"Colunas Renomeadas" = Table.RenameColumns(#"Tipo Alterado", {
        {"Data", "Ano Inserido"}
    }),
    #"Índice Adicionado" = Table.AddIndexColumn(#"Colunas Renomeadas", "Índice"),
    #"Colunas Reordenadas" = Table.ReorderColumns(#"Índice Adicionado", {
        "Índice", "Ano Inserido"
    }),
    #"Nome do Mês Inserido" = Table.AddColumn(#"Colunas Reordenadas", "Nome do Mês Inserido", Text.From),
    #"Nome do Dia Inserido" = Table.AddColumn(#"Nome do Mês Inserido", "Nome do Dia Inserido", Text.From)
in
    #"Nome do Dia Inserido"
```

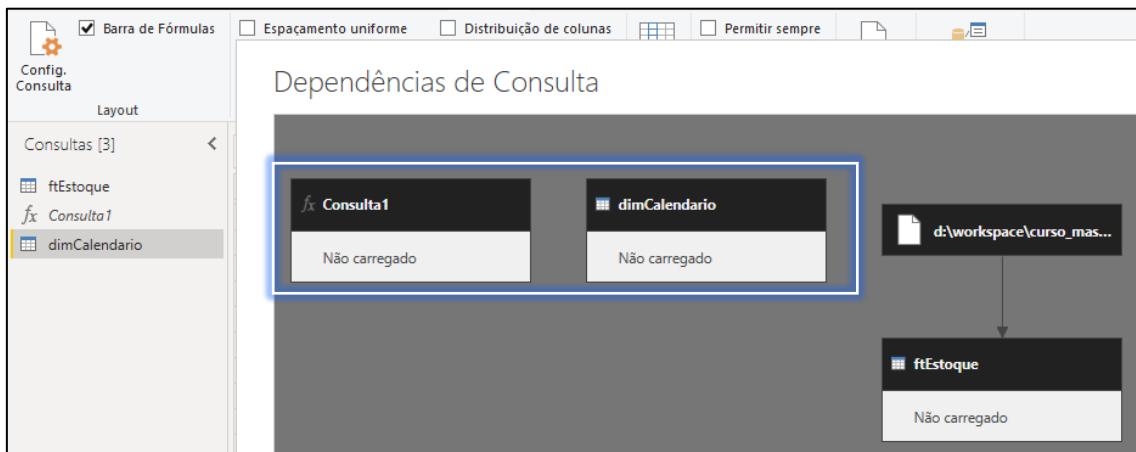
Faça, substitua o “*Consulta1*” por “*List.Dates*”, isso implicará que não é mais necessário a utilização dessa *Consulta1*.

A screenshot of the Power BI Advanced Editor for the 'dimCalendario' table. The title 'dimCalendario' is at the top. The M code in the editor is as follows:

```
let
    Fonte = List.Dates(#date(2017, 1, 1), 3650, #duration(1, 0, 0, 0)),
    #"Convertido para Tabela" = Table.FromList(Fonte, Splitter.SplitByNothing()),
    #"Tipo Alterado" = Table.TransformColumnTypes(#"Convertido para Tabela", {
        {"Data", typeDate}
    }),
    #"Colunas Renomeadas" = Table.RenameColumns(#"Tipo Alterado", {
        {"Data", "Ano Inserido"}
    }),
    #"Índice Adicionado" = Table.AddIndexColumn(#"Colunas Renomeadas", "Índice"),
    #"Colunas Reordenadas" = Table.ReorderColumns(#"Índice Adicionado", {
        "Índice", "Ano Inserido"
    }),
    #"Nome do Mês Inserido" = Table.AddColumn(#"Colunas Reordenadas", "Nome do Mês Inserido", Text.From),
    #"Nome do Dia Inserido" = Table.AddColumn(#"Nome do Mês Inserido", "Nome do Dia Inserido", Text.From)
in
    #"Nome do Dia Inserido"
```

The line 'Fonte = Consulta1(#date(2017, 1, 1), 3650, #duration(1, 0, 0, 0))' has been replaced by 'Fonte = List.Dates(#date(2017, 1, 1), 3650, #duration(1, 0, 0, 0))'.

Veja como na “Dependência de Consultas”, a *Consulta1* agora é inútil, ela não injeta mais a função em nenhum lugar.



Mas, afinal, e a data dinâmica entra onde?

Repare que nós ao invocarmos a função da `List.Dates` na *Consulta1*, informamos uma data de início em 01/01/2017. Mas no Excel, o início das “operações” foi em 01/01/2018 e o final, o último registro em 31/12/2020.

ID	DataContagem	Unidades	1095	1094	29/12/2020	6013
1	01/01/2018	3503	1096	1095	30/12/2020	6033
2	02/01/2018	3524	1097	1096	31/12/2020	6053

Mas do jeito que o nosso BI está atualmente, ele está armazenando data anterior ao início das atividades e na verdade ainda por mais 3650 dias (10 anos de algo que somente durou 3 anos).

Como deixar essa data dinâmica alocada?

Iniciaremos a variável = chama a função que retorna o valor mínimo da consulta (o nome do fato [abre a coluna e aqui coloca qual é a coluna que deve passar a lista de valores das datas]),

`DataMinima = List.Min (ftEstoque[DataContagem]),`

Veja que na primeira etapa ela foi criada corretamente.

Agora vamos aplicar a data máxima.

dimCalendario

```
let
    /* Iniciando as Variáveis */

    DataMinima = List.Min( ftEstoque[DataContagem]),
    DataMaxima = List.Max( ftEstoque[DataContagem]),
```

Veja que foi criada com sucesso:

The screenshot shows the Power BI Data Editor interface. On the left, there's a preview pane displaying a single date value: "31/12/2020". On the right, the "Config. Consulta" pane is open, showing the properties for the new dimension. Under "PROPRIEDADES", the "Nome" field is set to "dimCalendario". Under "ETAPAS APlicadas", two steps are listed: "DataMinima" and "DataMaxima", with "DataMaxima" currently selected.

Ótimo. Agora precisamos pegar o ano mínimo e máximo dos valores.

dimCalendario

```
let
    /* Iniciando as Variáveis */

    dataMinima = List.Min( ftEstoque[DataContagem]),
    dataMaxima = List.Max( ftEstoque[DataContagem]),
    anoMinimo = Date.Year( dataMinima ),
    anoMaximo = Date.Year( dataMaxima ),
```

Utilizaremos a função “Date.Year”. Essa função, seguindo a documentação, aceita um valor (ano). Por isso precisamos trazer a variável que recebia a mínima e a máxima.

The screenshot shows the Power BI Data Editor interface. On the left, there's a preview pane displaying the year value: "2018". On the right, the "Config. Consulta" pane is open, showing the properties for the new dimension. Under "PROPRIEDADES", the "Nome" field is set to "dimCalendario". Under "ETAPAS APlicadas", three steps are listed: "dataMinima", "dataMaxima", and "anoMinimo", with "anoMinimo" currently selected.

Tanto o mín quanto o máx estão corretos.

E para achar a Duração, ou seja, quantos dias efetivamente foram necessários.

dimCalendario

```
let
    /* Iniciando as Variáveis */

    dataMinima = List.Min( ftEstoque[DataContagem]),
    dataMaxima = List.Max( ftEstoque[DataContagem]),
    anoMinimo = Date.Year( dataMinima ),
    anoMaximo = Date.Year( dataMaxima ),
    dataInicial = #date(anoMinimo, 1, 1),
    dataFinal = #date(anoMaximo, 12, 31),
```

A “dataInicial = #date(anoMinimo, mês, dia), sendo que mês é sempre janeiro(1) e o dia primeiro (1) e o máximo é o oposto: #date(anoMaximo, 12, 31).

Para isso, invocaremos uma função chamada “*Duration.Days*”, porém é necessário acrescentar “+1” pois o BI é indexado ao zero.

dimCalendario

```
let
    /* Iniciando as Variáveis */

    dataMinima = List.Min( ftEstoque[DataContagem]),
    dataMaxima = List.Max( ftEstoque[DataContagem]),
    anoMinimo = Date.Year( dataMinima ),
    anoMaximo = Date.Year( dataMaxima ),
    dataInicial = #date(anoMinimo, 1, 1),
    dataFinal = #date(anoMaximo, 12, 31),
    duracao = Duration.Days(dataFinal - dataInicial) + 1,
```

Agora está tudo certo para **excluirmos** o *hardcode*!

```
/* Criando a Dimensão Calendario */
Fonte = List.Dates(#date(2017, 1, 1), 3650, #duration(1, 0, 0, 0)),
#"Convertido para Tabela" = Table.FromList(Fonte, Splitter.SplitByN
```

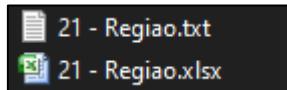
Agora na linha da Fonte, poderemos substituir o *hardcode* pelas variáveis.

```
/* Criando a Dimensão Calendario */
Fonte = List.Dates(dataMinima, duracao, #duration(1, 0, 0, 0)),
```

Internacionalização de Datas

No Brasil, o formato de datas é de dd/mm/aaaa. Em muitos países o formato é mm/dd/yyyy.

Utilizaremos os dois arquivos “21 – Região.txt e o .xlsx”



A screenshot of a table titled "21 - Regiao.txt - Bloco de Notas". The table has two columns: "Americano" and "Brasileiro". Both columns contain 10 rows of date data. The "Americano" column uses the mm/dd/yyyy format, while the "Brasileiro" column uses the dd/mm/aaaa format.

ID	Americano	Brasileiro
1	31/12/2018	31/12/2018
2	05/05/2018	05/05/2018
3	25/04/2018	25/04/2018
4	08/08/2017	08/08/2018
5	29/08/2017	29/08/2018
6	25/07/2018	25/07/2018
7	15/02/2019	15/02/2019
8	05/05/2017	05/05/2017
9	12/12/2017	12/12/2018
10	29/01/2018	29/01/2018

Repare que os dois arquivos são idênticos, apenas a extensão é diferente. Além do que nos dados, a data é diferente. No padrão americano e a no padrão brasileiro.

Nota: Acesse o Painel de Controle > Região > Result: Nossa máquina é Português.

A screenshot of the Windows Control Panel under "Painel de Controle". The left sidebar shows "Controle" and "Região". The right pane is titled "Região" and shows the "Formatos" tab selected. It displays the following information:

- Formato: Português (Brasil)
- Coincidir com idioma do Windows (recomendado)
- Preferências de idioma**
 - Formatos de data e hora
 - Data abreviada: dd/MM/aaaa
 - Data por extenso: dddd, d' de 'MMMM' de 'aaaa'
 - Hora abreviada: HH:mm
 - Hora por extenso: HH:mm:ss
 - Primeiro dia da semana: domingo
 - Exemplos
 - Data abreviada: 01/04/2021
 - Data por extenso: quinta-feira, 1 de abril de 2021
 - Hora abreviada: 01:42
 - Hora por extenso: 01:42:16

Veja que ao carregarmos o arquivo .txt e clicarmos para a transformação de dados, o Power BI já interpretou o idioma e região do Sistema Operacional automaticamente e já fez a tipagem das Datas de “Brasileiro” corretamente. Por isso, ele trata as datas no formato “Americano” como texto.

21 - Regiao	1 ² ₃ ID	A ^B _C Americano	Brasileiro
	1	12/31/2018	31/12/2018
	2	5/5/2018	05/05/2018
	3	4/25/2018	25/04/2018
	4	8/8/2017	08/08/2018

O que acontece se alterarmos o tipo de texto no Americano para Data? Erros.

1 ² ₃ ID	Americanoo	Brasileiro
1	Error	31/12/2018
2	05/05/2018	05/05/2018
3	Error	25/04/2018
4	08/08/2017	08/08/2018
5	Error	29/08/2018

Por que esses erros ocorrem? Porque o Power BI tenta converter para o formato dd/mm/aaaa. E não existe 12/31/2020. Não, ele não inverte, ele realmente tenta passar o número 31 como um mês, como isso não é possível, ele causa um erro.

Apague essa etapa acima.

Vamos com o “[Botão Direito > Alterar Tipo > Usando Localidade > US > Data](#)”

```
bole.TransformColumnTypes("#"Linhas Classificadas", [{"Americanoo": type date}], "en-US")
```

The screenshot shows the Power BI Data Editor interface. A context menu is open over a column named "Americanoo". The "Alterar Tipo" (Change Type) option is selected, and a submenu is displayed with various data types. The "Data" type is checked and highlighted. Other options in the submenu include Número Decimal, Número decimal fixo, Número Inteiro, Percentual, Data/hora, Hora, Data/Hora/Fuso Horário, Duração, Texto, Verdadeiro/Falso, Binário, and Usando a Localidade... (Using Locale...).

Isso é válido para arquivos em formato .txt.

Para os arquivos em extensão .xlsx, o método muda um pouco.

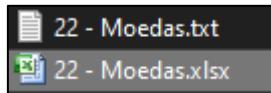
O interessante é que quando carregamos o arquivo em Excel, o Power BI se encarrega de alterar automaticamente.

ID	Americano	Brasileiro
1	31/12/2018	31/12/2018
2	05/05/2018	05/05/2018
3	25/04/2018	25/04/2018
4	08/08/2017	08/08/2018
5	29/08/2017	29/08/2018
6	25/07/2018	25/07/2018

É preciso muito cuidado e atenção nas datas.

Internacionalização de Números

Utilizaremos nesse exercício o “22 – Moedas..txt e .xlsx”



Sabemos que, o padrão de moeda no Brasil e nos US, são:

Brasil	R\$ 2.500,43
US	USD 2,500.43

Um é o inverso do outro.

Mas e quando recebemos: 2.500.00 ? Ou 2,500,00 ? Ou 250000? Como configurar?

Vamos conectar o arquivo .txt inicialmente.

	A ^B Cliente	A ^B \$Americano	A ^B Americano	\$ R\$Brasileiro	1.2 Brasileiro	1 ² ₃ Pontos
1	A	\$1,100.01	1,100.01		1.100,01	1100,01
2	A	\$1,100.01	1,100.01		1.100,01	1100,01

Usando o mesmo princípio das datas, vamos alterar a Localidade para a US. Porém, precisamos deixar como “Decimal Fixo”, porque é a unidade da tipagem monetária.

Veja que foi alterado e que a coluna do \$Brasileiro já foi reconhecida automaticamente.

\$ \$Americano	\$ Americano	\$ R\$Brasileiro
1.100,01	1.100,01	1.100,01
1.100,01	1.100,01	1.100,01

As colunas que serão trabalhosas, serão essas:

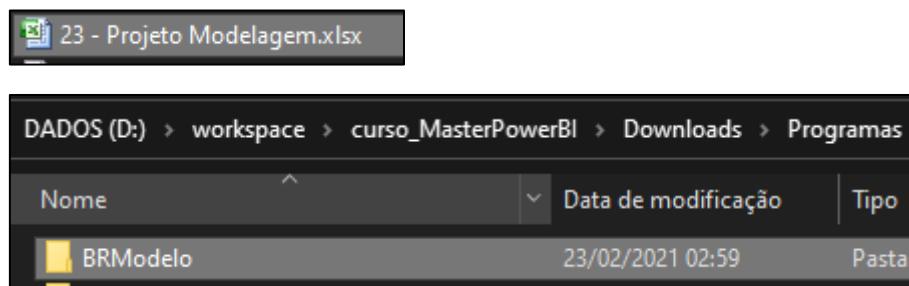
1.2 Brasileiro	1 ² ₃ Pontos	A ^B Vírgulas
1100,01	110001	1,100,01
1100,01	110001	1,100,01

Por isso, é muito mais fácil ajustar o tipo na origem dos dados, ajuste no Banco de Dados ou Excel. Caso contrário, serão necessários os cortes de cada uma das do final até o começo pela vírgula.

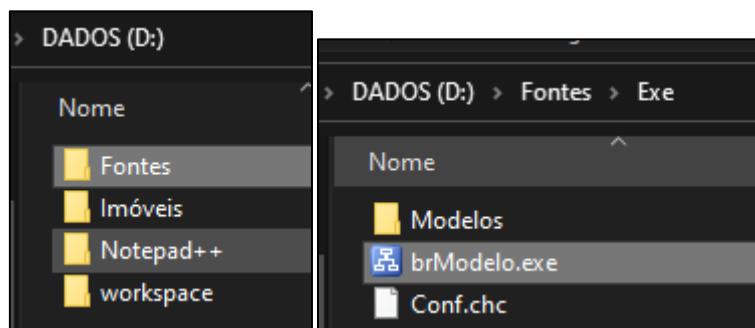
Instalando o BR Modelo

Para esse exercício, utilizaremos o “23 – Projeto Modelagem.xlsx”.

E utilizaremos e instalaremos o “Br Modelo”, que é um software *Free* para a modelagem de dados.



Ao extrairmos, pegamos a pasta “Fontes” e jogamos no Disco Local C ou no caso, no Hd Externo que está executando. O arquivo é o “brModelo.exe”, ele é um executável e não um arquivo instalável, não é necessário a instalação.



Entendendo a Regra de Negócio

Abra o arquivo .pdf de “Documento+de+Requisitos+01”.

Análise o trecho:

O projeto consiste na implantação do serviço de Business Intelligence para a empresa Ruschel Importações SA. A empresa possui dados de compra e venda armazenados em um banco de dados, que são regularmente exportados para um arquivo XLSX para que não ocorra lentidão no ambiente operacional.

Fica claro que por esse trecho, os dados serão enviados diretamente pelo depto operacional, não tem uma estrutura Data Warehouse (DW), é enviado sempre pelo Operacional em extensão .xlsx.

A empresa não possui ambiente analítico e não é uma opção a criação de um Datawarehouse em um primeiro momento.

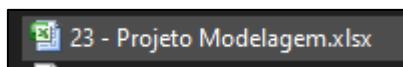
Artefatos de Entrada e Entregáveis

Em toda implementação de um serviço de Business Intelligence, ocorrem fases, que possuem Artefatos. Os **Artefatos** são as “informações, documentos, origem de dados, regra que o cliente envia” e os **Entregáveis** são literalmente, as entregas baseadas nos artefatos.

Exemplo:

Fase 01	
Artefatos: Documento de Requisitos	
Origem dos Dados	
Regras de Negócios.	
Entregáveis: Extração e tratamento dos dados	
Modelagem completa do Datawarehouse no Power BI	

Imagine que recebemos a planilha “23- Processo Modelagem.xlsx.



Essa planilha é uma muito parecida com um que fizemos em uns capítulos anteriores.

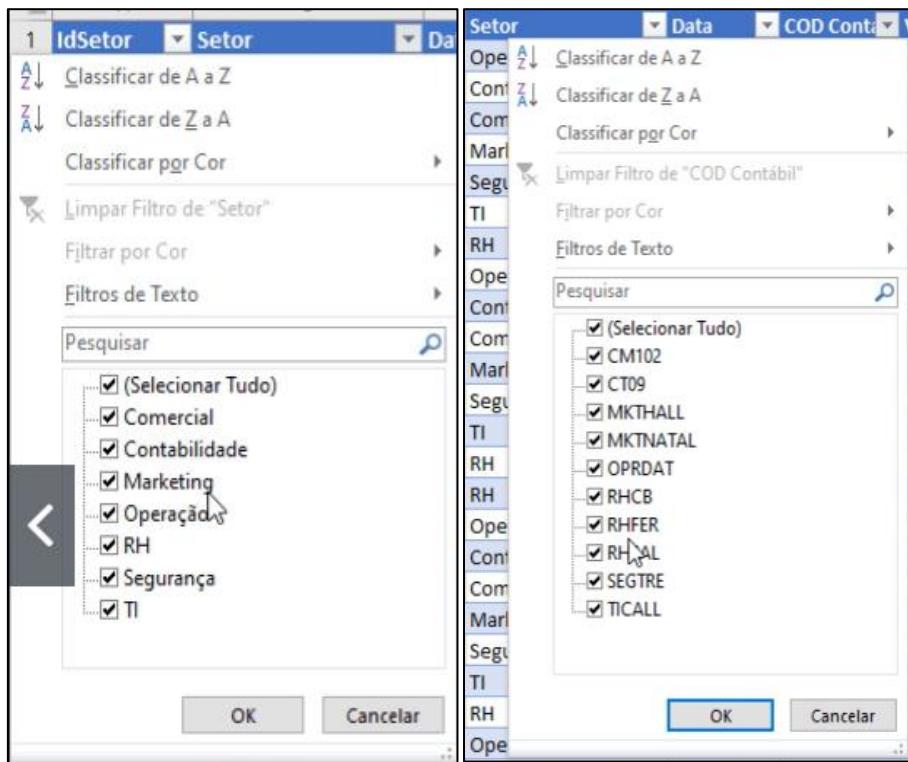
IdSetor	Setor	Data	COD Contábil	Valor Gerado
1	Operação	13/2/2018	OPR09	500,14
3	Comercial	13/2/2019	CMCB	500
3	Comercial	14/2/2019	CM102	250,98
4	Marketing	15/2/2019	MKTNATAL	500
1	Operação	17/2/2019	OPRDAT	500
4	Marketing	17/2/2019	MKTNATAL	250
1	Operação	18/2/2019	OPRDAT	512
1	Operação	13/2/2018	OPRDAT	500
3	Comercial	13/2/2019	CM102	500
3	Comercial	14/2/2019	CMGIN	250
4	Marketing	15/2/2019	MKTNATAL	500,14

IdSetor	Setor	Data	COD Conta	Valor Gasto
1	Operação	11/2/2019	OPRDAT	900,6
2	Contabilidade	11/2/2019	CT09	960,6
3	Comercial	12/2/2019	CM102	420,2
4	Marketing	13/2/2019	MKTHALL	460,92
5	Segurança	14/2/2019	SEG TRE	890,15
6	TI	16/2/2019	TICALL	900,1
7	RH	16/3/2019	RHSAL	420,1
1	Operação	03/11/19	RHCB	900,6
2	Contabilidade	03/11/19	CT09	960,6
3	Comercial	13/3/2019	CM102	420,2
4	Marketing	13/3/2019	MKTHALL	460,92

São **duas tabelas fatos**, sendo uma de **Compras** e uma de **Vendas**. Contendo um Id, o Setor, a Data, Cód Contábil e Valor Gasto/Gerado.

Analise primeiramente, a planilha.

Veja que o IdSetor é apenas 1, enquanto para um único Setor, temos um ou mais CODContábil, ou seja, a CARDINALIDADE do CODContábil é muito maior.



Por exemplo, se filtrarmos pelo RH, veja que ele possuirá um ou mais CODContábil.

IdSetor	Setor	Data	COD Contábil	Valor Gasto
7	RH	16/3/2019	RHSAL	420,1
7	RH	16/3/2019	RHSAL	420,1
7	RH	16/4/2019	RHSAL	214,36
7	RH	16/4/2019	RHFER	574,25
7	RH	16/3/2019	RHSAL	358,24
7	RH	16/3/2019	RHSAL	420,1
7	RH	16/4/2019	RHSAL	420,1

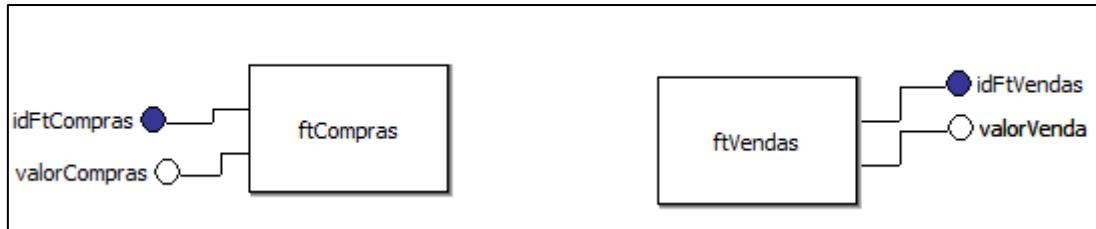
Ou seja, descobrimos uma relação já.



Então o Setor vai se repetir porque a cardinalidade da CODContábil é maior.

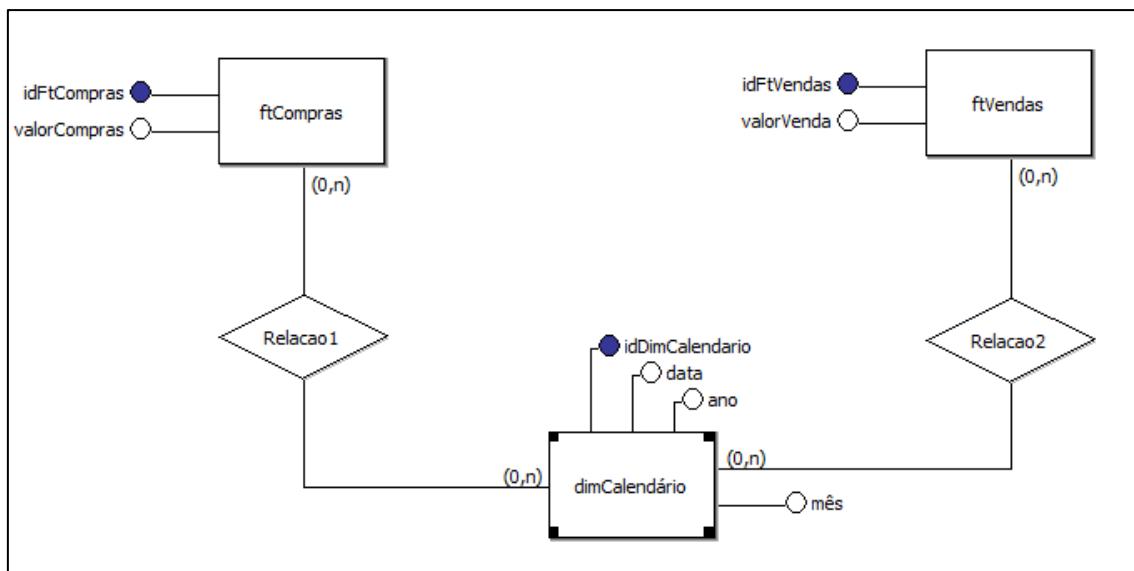
A **dimCalendário** é fácil. Não tem nenhum mistério, é a mais simples de ser feita.

Abra o “brModelo” e vamos modelar os dados.



Dica para melhor visualização: Clique na Entidade e Ctrl + O.

Entenda que as Entidades(**ftCompras** e **ftVendas**), recebem atributos. O **idFtCompras** e **idFtVendas** são as chaves primárias. Isso é, pode ser qualquer coluna, por exemplo, vamos atribuir o CODContábil (mas isso é um pouco mais para a frente). E o que mais? Cada um tem um valor.

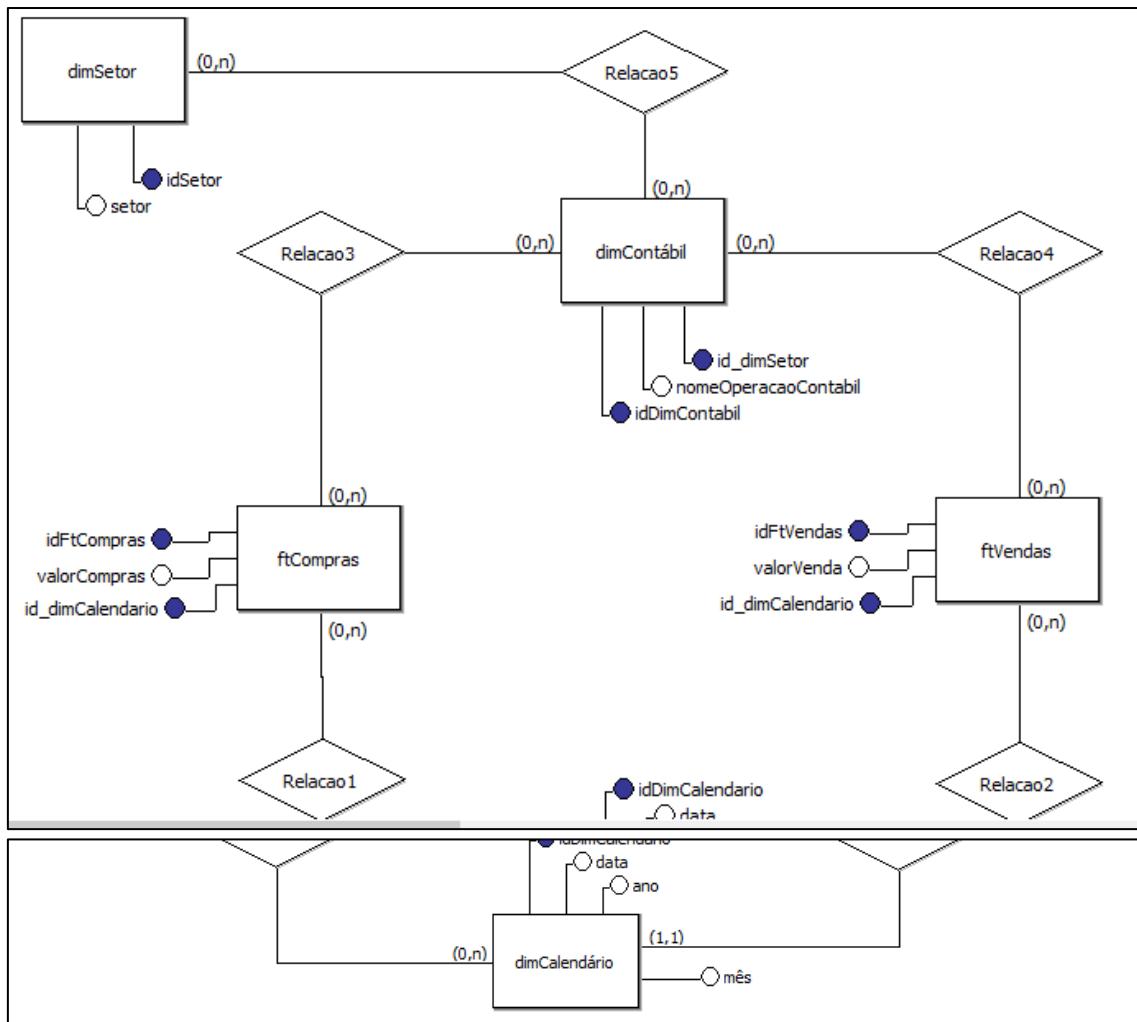


Vamos analisar até aqui.

Criamos a **dimCalendário**. Ela contém as datas nos seus atributos e um id único. E o **dimCalendário** se relaciona com as duas fatos, correto? Por isso é necessário construir a relação entre os 3.

Porém, precisamos lembrar que os Fatos se repetem todos os dias e até mais de uma vez no mesmo dia. Um cliente pode entrar e comprar um pão na padaria no dia 01/04/2021. E nesse mesmo dia, um outro cliente também entra. Então o dia também se repetirá.

Por isso também, precisamos criar uma chave estrangeira nos fatos, essa **Foreign Key** deve ser das dimensões. Ela é identificada pelo começo **id_nomeDaDimensão** (o *underline* é o que distingui como uma FK).



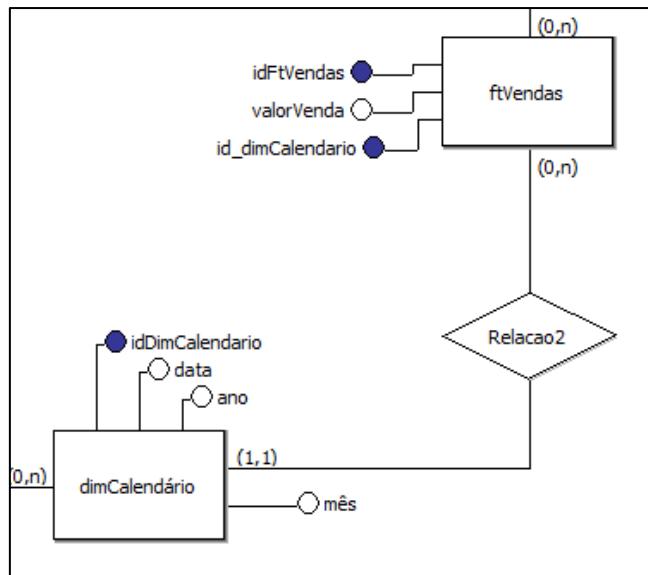
Essa abordagem está em SNOWFLAKE

Por enquanto está assim.

Vamos analisar a Cardinalidade entre cada Entidade.

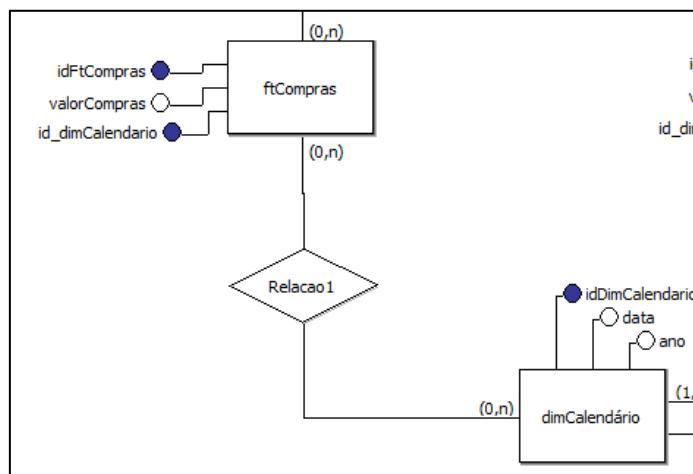
dimCalendario x ftVendas

A cardinalidade espera que exista pelo menos 1 data para 1 venda. E que na a Venda pode ocorrer nenhuma (por exemplo um final de semana, o mercado pode estar fechado) ou n vendas que aconteçam.



dimCalendario x ftCompras

A mesma premissa também é verdadeira. Se ocorrer uma data, ela pode ocorrer ao menos uma vez para uma compra. Enquanto a Compra de insumos por exemplo, pode ocorrer nenhuma vez ou várias vezes.



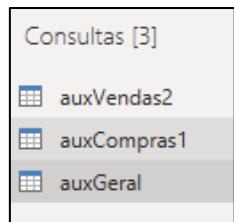
Nas próximas páginas veremos a modelagem no próprio BI

Modelagem Compra e Vendas

Abrindo o Power BI, utilizaremos o “23 – Projeto Modelagem.xlsx”.

Com base na modelagem feita no capítulo anterior.

Aquele padrão básico, duas planilhas que serão auxiliares, uma geral que conta com o acréscimo de outra.

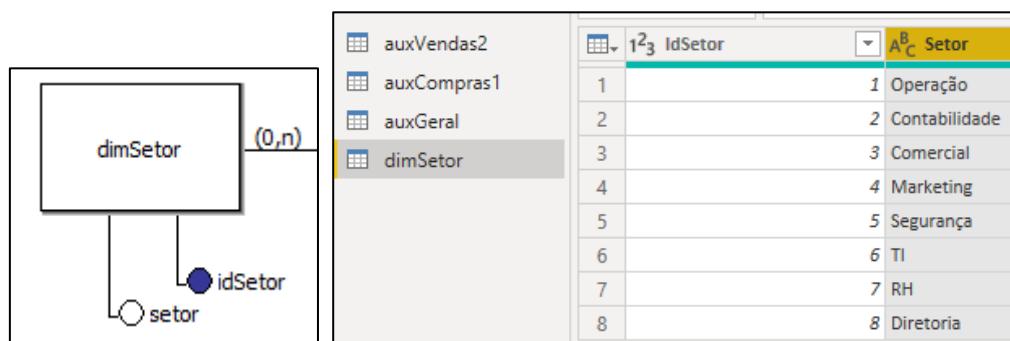


A partir da auxGeral, vamos realizar todas as dimensões, os fatos...

Criação da dimSetor

A primeira será a montagem da dimSetor.

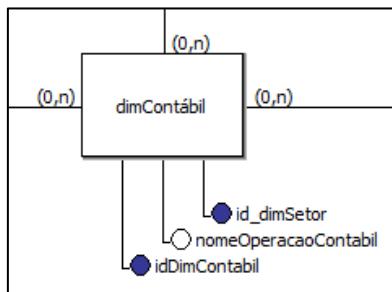
Duplique a auxGeral e mantenha apenas os atributos que foram definidos no brSetor.



Criação da dimContábil

	t^2_3 IdSetor	A^B_C Setor	Data	A^B_C COD Contábil	\$	Valor Gasto
1	1	Operação	11/02/2019	OPRDATA		900,60
2	2	Contabilidade	11/02/2019	CT09		960,60
3	3	Comercial	12/02/2019	CM102		420,20
4	4	Marketing	13/02/2019	MKTALL		460,92
5	5	Segurança	14/02/2019	SEGTRE		890,15
6	6	TI	16/02/2019	TICALL		900,10

Veja que é necessário manter o 'id_DimSetor' porque é ele que faz a conexão com a dimSetor. Ele é a chave estrangeira necessária para conectar com a dimSetor.



Portanto, removeremos as demais colunas, até que fique somente assim

	t^2_3 IdSetor	A^B_C COD Contábil
1		1 OPRDATA
2		2 CT09
3		3 CM102
4		4 MKTHALL
5		5 SEGTRE
6		6 TICALL
7		7 RHSAL
8		1 RHCB
9		2 CT09
10		3 CM102
11		4 MKTHALL
12		5 SEGTRE
13		6 TICALL
14		7 RHSAL
15		7 RHSAL
16		1 OPRDATA

Agora precisaremos remover os valores duplicados que existem na COD Contábil.

Deverá permanecer assim.

	t^2_3 IdSetor	A^B_C COD Cont		
1		1 OPRDATA	12	3 CMCB
2		2 CT09	13	3 CMGIN
3		3 CM102	14	1 OPD
4		4 MKTHALL	15	4 MKTVER
5		5 SEGTRE	16	1 OPRGER
6		6 TICALL	17	8 DIR02
7		7 RHSAL	18	4 MKTINV
8		1 RHCB	19	3 CMO98
9		4 MKTNATAL	20	4 MKTCAMP
10		7 RHFER	21	3 CMBF
11		1 OPR09	22	8 DIRBONUS

Veja que agora não teremos mais valores repetidos, além do que, por exemplo, 1 é Operacional lá no Setor, certo? E todos os COD Contábil que são de IdSetor 1, são relacionados ao Operacional.

	A ^B _C COD Contábil
1	1 OPR09
2	1 OPRDAT
3	1 OPD
4	1 RHCB
5	1 OPRGER
6	2 CT09
7	3 CMBF
8	3 CMGIN
9	3 CMCB
10	3 CM102
11	3 CMO98

Mas, acima não é o Id da CODContábil. Aquele número ao lado é a chave estrangeira da dimSetor!

Ainda precisamos criar o ID único para a CODContábil.

Vá em “Adicional Coluna > Coluna de Índice > Personalizado”.

Veja o resultado.

1 ² ₃ idDimContabil	A ^B _C NomeOpContabil	1 ² ₃ id_dimSetor
1000	OPRDAT	1
1010	OPRGER	1
1020	OPD	1
1030	RHCB	1
1040	OPR09	1
1050	CT09	2
1060	CM102	3
1070	CM098	3

A coluna “idDimContabil”, é a **Primary Key** da dimensão de CODContábil.

A segunda coluna, é o nome da operação, apenas renomeamos.

Enquanto a terceira, renomeamos para identificar que ela é a **Foreign Key** da dimSetor.

Criação da dimCalendario

NÃO DUPLIQUE A AUXGERAL PARA A CRIAÇÃO DA DIMCALENDARIO!

Para a elaboração de uma dimensão Calendário é necessário, iniciar uma consulta nula e depois torná-la dinâmica. Não acrescente o índice logo após a criação, pois a criação é feita com aquele intervalo de tempo longo que somente será alterado quando iniciarmos a manipulação com a Linguagem M.

Lembra daquele Script que foi criado num exercício anterior? Pois podemos usá-lo sem problema algum!!!

```
let
    /* Trabalhando nas variáveis */
    DataMinima = List.Min( NomeColunaData ),
    DataMaxima = List.Max( NomeColunaData ),
    AnoMinimo = Date.Year( DataMinima ),
    AnoMaximo = Date.Year( DataMaxima ),
    DataInicial = #date(AnoMinimo, 1, 1),
    DataFinal = #date(AnoMaximo, 12, 31),
    Duracao = Duration.Days( DataFinal - DataInicial ) + 1,
    /* Criando a Dimensão */
    Fonte = List.Dates(DataInicial, Duracao, #duration(1,
    0, 0, 0)),
```

Porém, precisamos lembrar que não podemos achar as datas com base nas auxiliares de Compra ou Venda pois em alguns dias ocorreram vendas e outros não e vice-versa. Por isso, utilizaremos a tabela **auxGeral** para realizar a data dinâmica.

Como adaptar essa função lá no **Editor Avançado**? Copie o código acima lá no editor.

Função Invocada

```
let
    /* Trabalhando nas variáveis */
    DataMinima = List.Min( auxGeral[Data] ),
    DataMaxima = List.Max( auxGeral[Data] ),
    AnoMinimo = Date.Year( DataMinima ),
    AnoMaximo = Date.Year( DataMaxima ),
    DataInicial = #date(AnoMinimo, 1, 1),
    DataFinal = #date(AnoMaximo, 12, 31),
    Duracao = Duration.Days( DataFinal - DataInicial ) + 1,
    /* Criando a Dimensão */
    Fonte = List.Dates(DataMinima, Duracao, #duration(1, 0, 0, 0)),
```

Agora precisamos verificar se ainda existem dependências dessa Função Invocada (que em breve se tornará a dimCalendário) e a Consulta1 que foi aquela consulta nula que foi criada.

Veja que não existem mais dependência entre as duas.



Portanto, podemos excluir a *Consulta1*.

Criando as Fatos de Compras e Vendas

Iniciaremos primeiramente, com a Fato de Compras.

Não utilizaremos a auxGeral pois ela foi feita apenas para as dimensões, porque ela compõe todas as informações dos fatos de Compras e Vendas. Para trabalhar criando a ftCompras, utilizaremos a auxCompras.

Comece duplicando a auxCompras e renomeie para ftCompras. Como não haverá data dinâmica, podemos já criarmos o idFtCompras logo de início.

Logo em seguida, exclua as duas colunas do idSetor e Setor, pois estamos em uma modelo SNOWFLAKE. Nossa identificador será com a CODContábil.

	1 ² 3 idFtCompras	1 ² 3 IdSetor	A ^B C Setor	Data	A ^B C COD Contábil
1	1		7 RH		16/04/2019 RHSAL
2	2		6 TI		16/04/2019 TICALL
3	3		7 RH		16/04/2019 RHFER
4	4		7 RH		16/04/2019 RHSAL
5	5		6 TI		16/04/2019 TICALL

Mas precisaremos do ID da Data e, também o ID do CODContábil, correto?

1 ² 3 idFtCompras	Data	A ^B C COD Contábil	\$ Valor Gasto
1	16/04/2019	RHSAL	214,36
2	16/04/2019	TICALL	958,36
3	16/04/2019	RHFER	574,25
4	16/04/2019	RHSAL	420,10

Vamos em “[Página Inicial > Combinar > Mesclar Consultas](#)”.

Selecione com cuidado, a coluna de Data na “ftCompras” e a dimensão de dimCalendario e, também selecione a coluna Data. Dessa forma o Power BI encontrará as correspondências.

Mesclar

Selecione a tabela e as colunas correspondentes para criar um ftCompras

idFtCompras	Data	COD Contábil	Valor Gasto
1	16/04/2019	RHSAL	214,36
2	16/04/2019	TICALL	958,36
3	16/04/2019	RHFER	574,25
4	16/04/2019	RHSAL	420,10
5	16/04/2019	TICALL	542,36

dimCalendario

idDimCalendario	Data
2000	13/02/2018
2001	14/02/2018
2002	15/02/2018
2003	16/02/2018
2004	17/02/2018

Tipo de Junção

Externa esquerda (todas a partir da primeira, correspo...)

Usar a correspondência difusa para executar a mesclagem

Opções de correspondência difusa

✓ A seleção corresponde a 66 de 66 linhas da primeira tabela.

Veja que o resultado dessa mesclagem, e a dimCalendario em forma da *Table* que pode ser expandida.

1^2_3 idFtCompras	Data	A ^B _C COD Contábil	\$ Valor Gasto	dimCalendario
1	16/04/2019	RHSAL	214,36	Table
2	16/04/2019	TICALL	958,36	Table
3	16/04/2019	RHFER	574,25	Table

Clique na expansão e selecione apenas o ID da dimCalendario, não é necessário a data.

E veja que datas que se repetem, ou seja, que outra COD Contábil realizou no mesmo dia que outra, o ID é mantido e se repete também.

Data	A ^B _C COD Contábil	\$ Valor Gasto	1^2_3 idDimCalendario
11/02/2019	OPRDATA	900,60	2363
11/02/2019	CT09	960,60	2363
11/02/2019	OPRDATA	698,41	2363
11/02/2019	CT09	257,36	2363
11/02/2019	OPRDATA	475,36	2363

Se eu já tenho o idDimCalendario, ou seja, eu já tenho a identificação da data, não é mais necessário que a Data esteja na ftCompras. Podemos excluir também, uma vez que já possuímos.

Agora, faremos o mesmo procedimento para a CODContábil.

1^2_3 idFtCompras	A ^B _C COD Contábil	\$ Valor Gasto	1^2_3 idDimCalendario
49	OPRDATA	900,60	2363
50	CT09	960,60	2363
51	OPRDATA	698,41	2363
52	CT09	257,36	2363
53	OPRDATA	475,36	2363
54	CT09	578,36	2363

Lembre-se de fazer a seleção que desejamos mesclar.

Mesclar

Selecione a tabela e as colunas correspondentes para criar uma ta

ftCompras

idFtCompras	COD Contábil	Valor Gasto	idDimCalendario
49	OPRDATA	900,60	2363
50	CT09	960,60	2363
51	OPRDATA	698,41	2363
52	CT09	257,36	2363
53	OPRDATA	475,36	2363

dimContabil

idDimContabil	NomeOpContabil	id_dimSetor
1000	OPRDATA	1
1010	OPRGER	1
1020	OPD	1
1030	RHCB	1
1040	OPR09	1

Tipo de Junção

Externa esquerda (todas a partir da primeira, correspo...)

Usar a correspondência difusa para executar a mescla

► Opções de correspondência difusa

A seleção corresponde a 66 de 66 linhas da primeira

O resultado da mesclagem.

A ^B _C COD Contábil	\$ Valor Gasto	1 ² ₃ idDimCalendario	dimContabil
OPRDAT	900,60	2363	Table
CT09	960,60	2363	Table
OPRDAT	698,41	2363	Table

Lembre-se de fazer uma verificação rápida com o ID da origem. Por exemplo, o OPRDAT lá na dimContabil é realmente o **ID 1040?** Se for, siga abaixo.

A única coisa que precisamos é o idDimContabil e aí precisaremos somente excluir também o CODContábil.

1 ² ₃ idFtCompras	1 ² ₃ idDimCalendario	1 ² ₃ idDimContabil	\$ Valor Gasto
49	2363	1040	900,60
51	2363	1040	698,41
53	2363	1040	475,36
50	2363	1050	960,60
52	2363	1050	257,36
54	2363	1050	578,36
46	2364	1060	420,20

Vamos analisar o que temos na ftCompras?

Seta **azul**: É o id da própria compra.

Seta **vermelha**: é o id da Data que a compra foi feita.

Seta **verde**: é o id que corresponde ao CODContábil que a compra foi feita.

Seta **amarela**: é o valor gasto da própria compra.

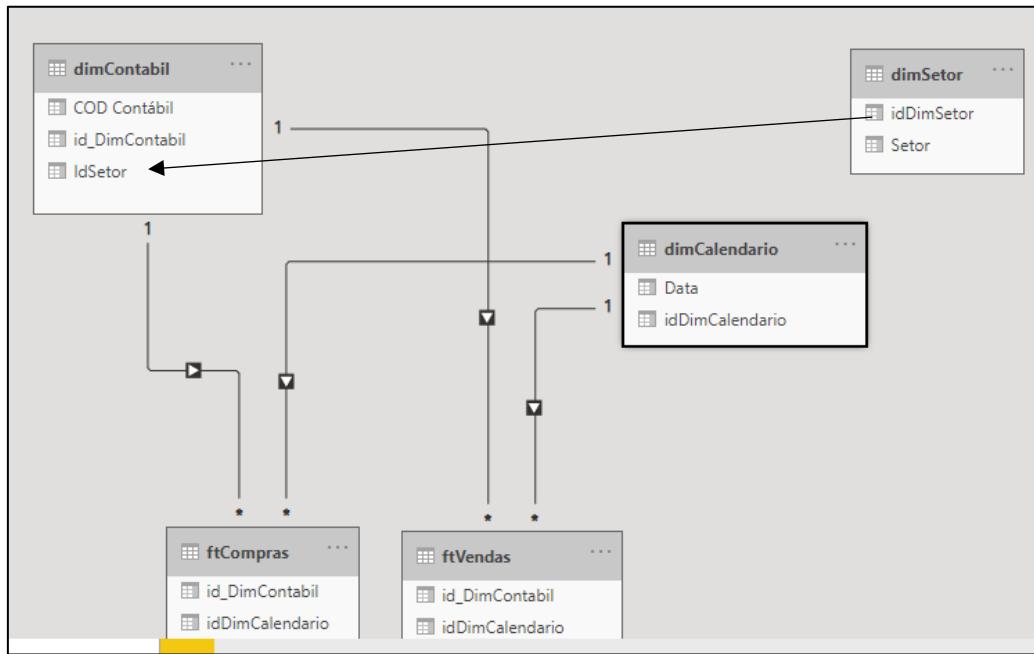
E agora é necessário fazer também para a **ftVendas**.

Nota: *Muito cuidado, pois durante o curso, o instrutor se confundiu e não arrumou corretamente na planilha de origem continha um “RH” no setor 01. Isso deu um problema na correção.* Mas seguindo os passos e arrumando diretamente na planilha de origem, tudo saiu corretamente (Obs: a planilha já está corrigida, a diferença será desse manual com o PBI e o Excel).

Nota[2]: Esse projeto deveria ter sido conectado em uma PASTA e, também deveríamos ter criado o PATH, pois novos arquivos seriam adicionados futuramente e isso seria um problema. Ou até mesmo se o arquivo fosse alterado da sua pasta original.

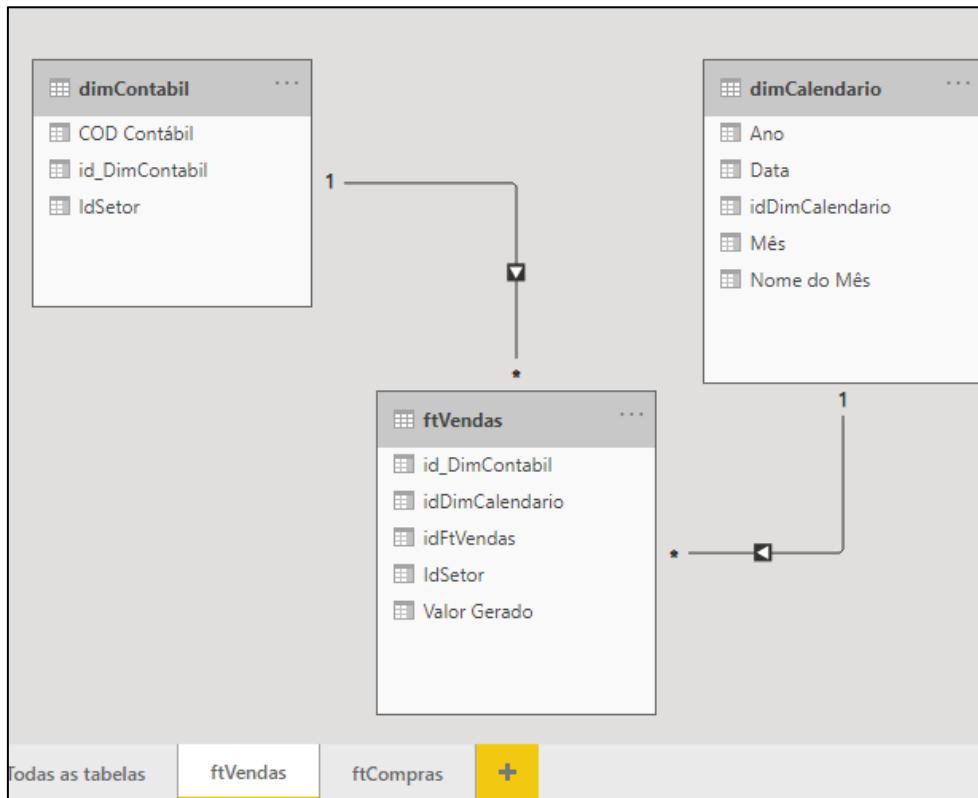
Relacionamento da Modelagem no Power Pivot

Para estabelecer o relacionamento, vamos em Modelo.



Lembre-se que o PBI pode estabelecer o relacionamento incorreto, por isso, devemos verificar sempre. Relacione também o “idDimSetor” com o “idDSetor” na dimContabil.

Próximo passo é criar as relações “Lógicas” para a Vendas e Compras.



Trabalhando com Matriz

A Matriz fica ao lado da Tabela.

O melhor posicionamento das informações na Matriz é “utilizar” em COLUNAS o que se tem de menos informação.

Por exemplo, vejamos o modo **errado** de posicionamento.

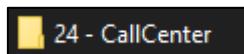
Setor	CM102	CMBF	CMCB	CMGIN	CMO98	DIR02	DIRBONUS
Comercial	R\$ 1.000,98	R\$ 1.681,98	R\$ 2.250	R\$ 1.188,79	R\$ 632,58		
Diretoria						R\$ 5.000	R\$ 1
Marketing							
Operação							
Total	R\$ 1.000,98	R\$ 1.681,98	R\$ 2.250	R\$ 1.188,79	R\$ 632,58	R\$ 5.000	R\$ 1

Abaixo, é o modo correto de posicionamento na matriz. Veja que acima, as colunas contêm muitas informações, isso gera uma matriz gigantesca, enquanto abaixo, as colunas são as que contêm menos informações.

COD Contábil	Comercial	Diretoria	Marketing	Operação	Total
CM102	R\$ 1.000,98				R\$ 1.000,98
CMBF	R\$ 1.681,98				R\$ 1.681,98
CMCB	R\$ 2.250				R\$ 2.250
CMGIN	R\$ 1.188,79				R\$ 1.188,79
CMO98	R\$ 632,58				R\$ 632,58
DIR02	R\$ 5.000				R\$ 5.000
DIRBONUS	R\$ 1.180,7				R\$ 1.180,7
MKTcamp	R\$ 124,36				R\$ 124,36
MKTHall	R\$ 731,47				R\$ 731,47
Mktinv	R\$ 810				R\$ 810
Mktnatal	R\$ 1.978,69				R\$ 1.978,69
Mktver	R\$ 2.382,47				R\$ 2.382,47
OPD	R\$ 1.063,12				R\$ 1.063,12
OPR09	R\$ 3.835,22				R\$ 3.835,22
OPRdat	R\$ 5.026,66				R\$ 5.026,66
OPRGER	R\$ 2.284,69				R\$ 2.284,69
Total	R\$ 6.754,33	R\$ 6.180,7	R\$ 6.026,99	R\$ 12.209,69	R\$ 31.171,71

Gráficos e Análises Visuais

Utilizaremos o arquivo “24 – Call Center”.



Entendendo o case, é uma empresa de Call Center que está avaliando o atendimento de um Banco X. Conforme o cliente liga para o Call Center/ SAC para algum serviço financeiro. E tem o seu tempo de espera até ser atendido/finalizado o atendimento.

IDCliente	Data	Operação	Cidade	IDPesquisa	Tempo de Espera
13	01/04/2018	Depósito Poupança	São Paulo	126	58
15	01/09/2018	Depósito Poupança	São Paulo	228	33
18	01/04/2018	Depósito Poupança	Porto Alegre	331	50
24	01/01/2018	Depósito Poupança	Porto Alegre	437	31
31	01/11/2018	Depósito Poupança	Porto Alegre	544	19
98	01/02/2018	Depósito Poupança	São Paulo	6111	46
99	01/12/2018	Depósito Poupança	São Paulo	7112	34
156	01/08/2018	Depósito Poupança	Porto Alegre	8169	2
212	01/01/2018	Depósito Poupança	Florianópolis	9225	24

Após o atendimento, é realizado uma pesquisa de satisfação.

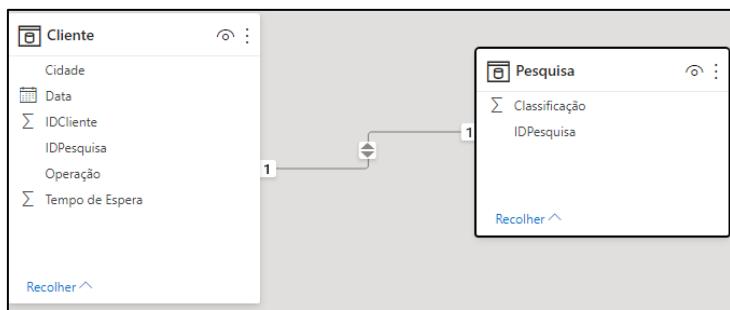
IDPesquisa	Classificação
126	4
228	4
331	5
437	2
544	5
6111	4
7112	2
8169	3
9225	4

O Banco X que contratou os serviços, quer entender a relação entre o tempo, a cidade com a pesquisa de satisfação para uma visão mais Macro.

Nesse caso não precisaremos classificar, tipar os dados. Só precisamos ver as relações.

IDCliente	Data	Operação	Cidade	IDPesquisa
13	01/04/2018	Depósito Poupança	São Paulo	126
15	01/09/2018	Depósito Poupança	São Paulo	228
18	01/04/2018	Depósito Poupança	Porto Alegre	331
24	01/01/2018	Depósito Poupança	Porto Alegre	437
31	01/11/2018	Depósito Poupança	Porto Alegre	544
98	01/02/2018	Depósito Poupança	São Paulo	6111

Nesse caso, a relação está realmente 1 para 1 (1-1). E podemos permanecer assim. Nesse caso, abordaremos as dimensões, porque possuímos um único JOIN. Veja que tem que uma está filtrando a outra.



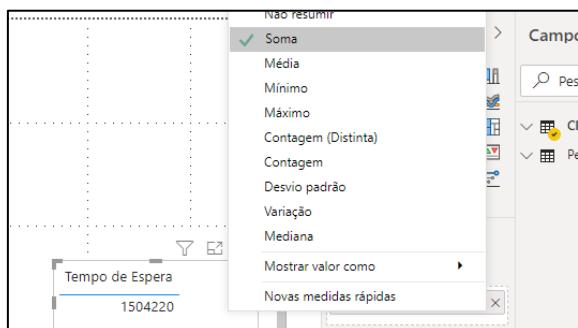
Por que então, não fazemos uma dimCalendario? Uma dimTempoEspera? Uma dimCidade? Porque isso não foi solicitado no Requisito do Cliente. Não é necessário. Caso ele peça, caso seja necessário, aí sim, nós precisaremos. Mas, ele só quer a relação entre a Cidade e a Pesquisa.

Medida em DAX

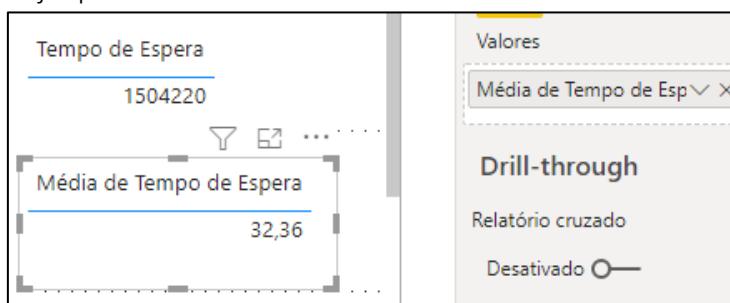
Veja que o Tempo de Espera do Cliente está somando.



Isso faz sentido, uma vez que é uma FATO. Então de fato, deveria somar. Porém, e se quisermos trabalhar com a MÉDIA do Tempo de Espera? Ou o MENOR Tempo de Espera, ou o MAIOR Tempo de Espera? Uma solução seria em cada gráfico, alterar a medida cada vez que for necessário.



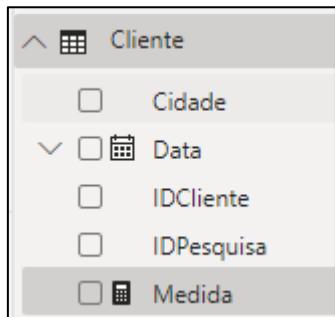
Veja que um está a soma e o outro está a média.



Mas, com a DAX não é necessário. Podemos trabalhar com a MEDIDA.

Criando uma Medida DAX

Selecionando o campo PESQUISA, com o botão direito > Nova Medida.



Veja que agora a barra de fórmula está disponível para criação da DAX.

Se trocar o nome na Barra de Fórmula, logo mudará o nome do campo.



Vamos digitar a fórmula de AVG(). Assim que começamos a digitar, o PBI já disponibiliza



The screenshot shows a Power BI report with a table containing two rows: "Tempo de Espera" (1504220) and "Tempo Médio de Espera" (32,36). To the right of the table is a context pane. In the context pane, under the "Valores" section, there is a button labeled "Adicionar os campos de da...". Below it, under "Drill-through", is the text "Relatório cruzado". On the far right, a list of fields is shown, with "Tempo Médio de Espera" being the last item in the list. The "Pesquisa" field is expanded, revealing the DAX formula: `Tempo Médio de Espera = AVERAGE([Cliente[Tempo de Espera]])`.

Para saber a quantidade de Clientes que ligaram.

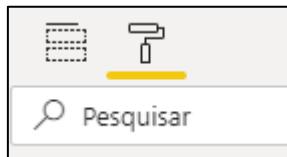


The screenshot shows a Power BI report with a table containing one row: "Contador de Clientes" (46487). To the right of the table is a context pane. The "Contador de Clientes" field is highlighted with a gray background. The context pane also lists "Cidade" and "Data".

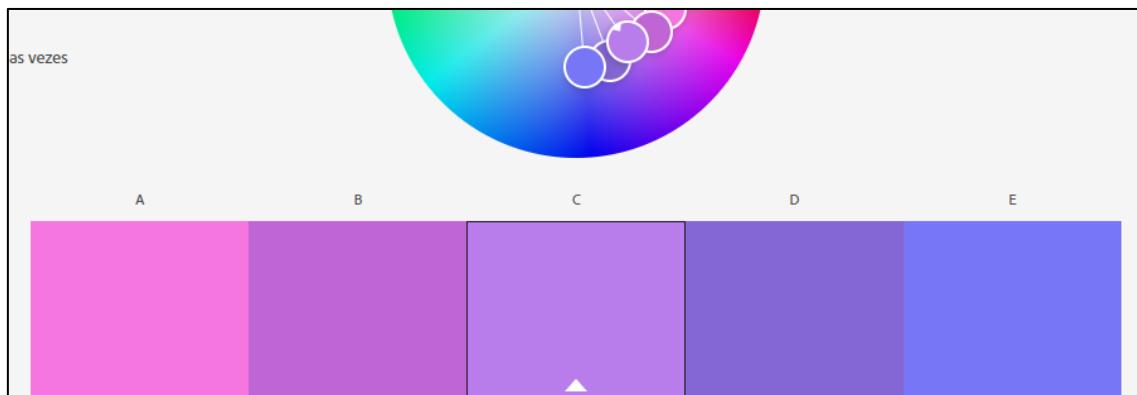
Isso poupa tempo.

Inserindo Imagem, cores no Background

Para inserir qualquer formatação no background, vá em Formato

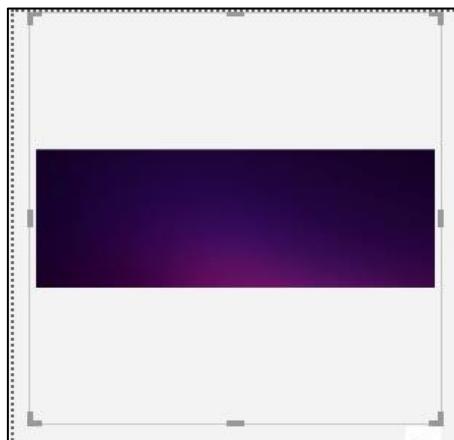


Acesse <https://color.adobe.com/pt/create/color-wheel>. Para design de cores do Background.

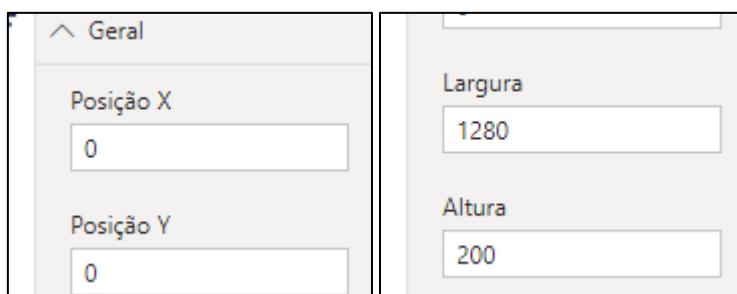


Clique em “Inserir > Inserir Imagem”.

Veja que a importação da imagem está na proporção original da imagem.



Ao invés de ficar redimensionando a imagem pelo mouse. Clique na imagem, vá em Formato > Geral > zere o eixo Y e zere o eixo X para início.



Porque deixamos a Largura em 1280? Porque é o tamanho da página.

Para descobrir o tamanho da página, clique no “Fundo > Formato > Tamanho da Página”.

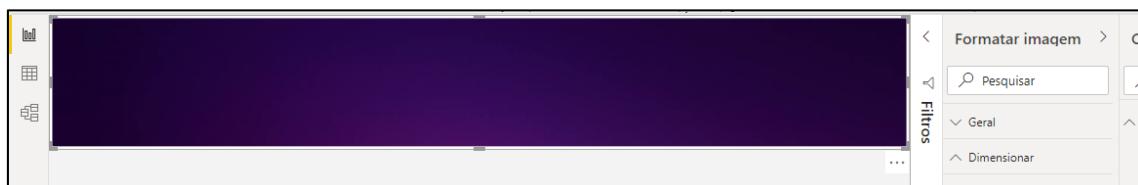


Veja como está.



Mas está “feio”.

Para corrigir, clique na “Imagem > Formato > Dimensionar > Preencher”.



Não se esqueça de enviar para trás de tudo.

Clique na “Imagem > Formato (na barra de Pesquisa) > Enviar para Trás”.

Acesse <https://www.flaticon.com/> para baixar ícones gratuitos para o dashboard.

Leitura Z

A tendência do ser humano é fazer leitura exatamente nesse sentido que está em rosa, começando pelo lado esquerdo e fazendo um “Z” no dashboard.

Ou então empilhado, como está em verde, sendo a esquerda primeiro.



Gráfico de Barras Verticais

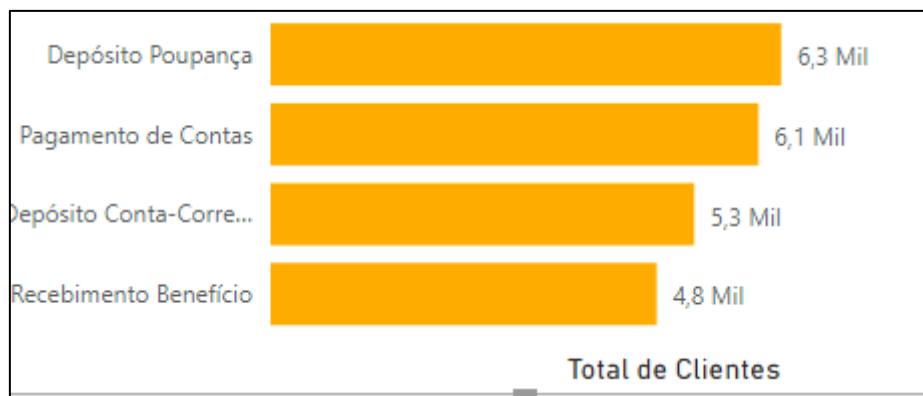
Dê preferência aos gráficos de Barras Verticais quando houver muitos nomes, dados, operações.

Por exemplo:

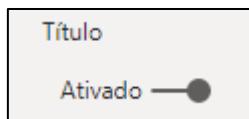
- (Selecionar Tudo)
- Depósito Conta-Corrente
- Depósito Poupança
- Gerência
- Pagamento de Contas
- Recebimento Benefício
- Saque Conta-Corrente
- Saque Poupança

Remova o Título do Eixo X

Para remover o título do eixo X, que não é algo elegante, pode ser prático e necessário, mas também pode ser descartado.



Vá no Formato > Eixo X > Título do Eixo > Desativar



Resultado:

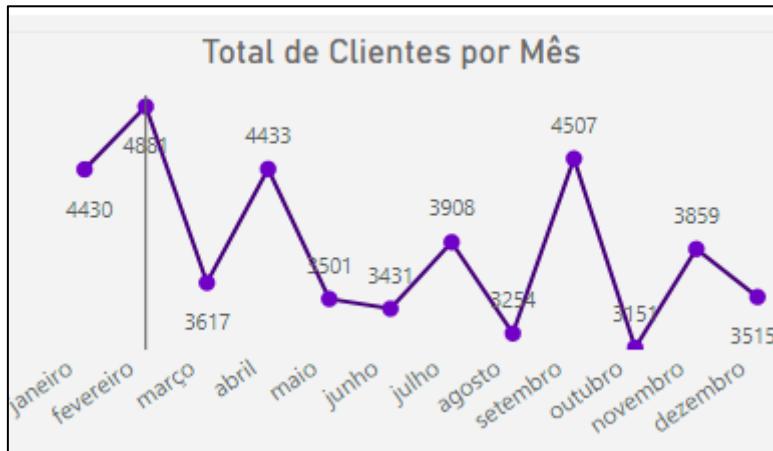


Gráfica de Área x Gráfico de Colunas e Linhas Empilhadas

O gráfico de área é indicado para se trabalhar com o tempo assim com o de colunas e linhas empilhadas. Porém, o gráfico de área é indicado para quando se compara apenas duas variáveis do mesmo tipo.

Gráfico de Área: Ao longo do tempo de Vendas do Cliente A x Vendas do Cliente B.

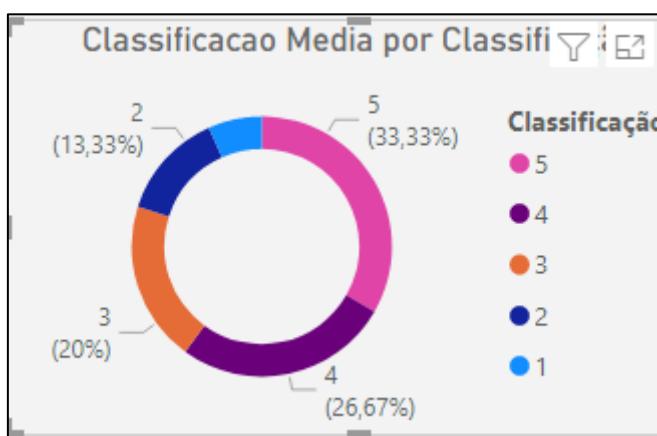
Gráfico de Linhas e Colunas: Ao longo do tempo, as *Vendas* do Cliente A ou B e cruzar com a *Meta* do Cliente A ou B.



Técnica do Gráfico de Rosca com Cartão

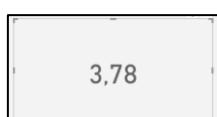
Vamos utilizar para a classificação média, a avaliação de satisfação.

Primeiro, faça um gráfico de rosca, e aumente o Raio Interno para até 75 pts. Para isso, selecione o gráfico e vá em “Formas > Raio Interno”.



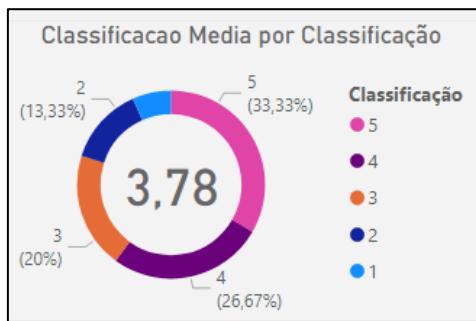
Após, crie um cartão com o campo desejado, nesse caso, queremos a classificação média.

Reduza o tamanho do cartão até que ele caiba dentro do raio do **Gráfico de Rosca**.



E por fim, **arraste** o cartão para dentro do *raio* do Gráfico de Rosca. Caso seja necessário, diminua o tamanho do Rótulo de Dados.

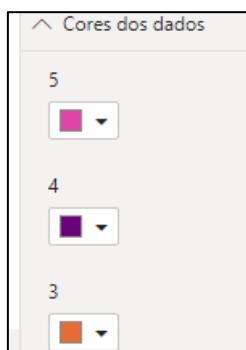
Veja o resultado.



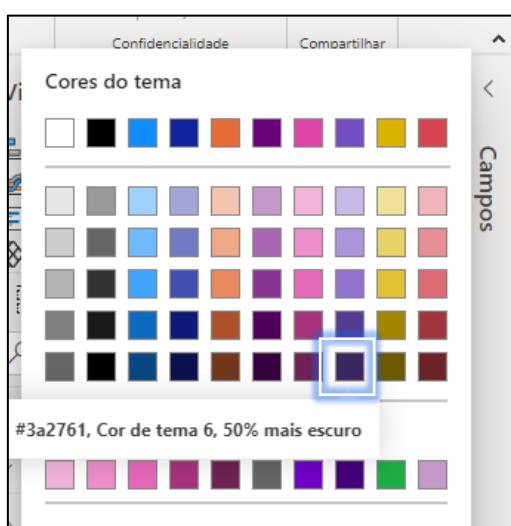
Técnica do Tom sobre Tom no Gráfico de Rosca

Veja que o gráfico está colorido, não necessariamente, “feio”. Porém, em Gráfico de Rosca, é comum a utilização do Tom sobre Tom. Vá em “Formato > Cores dos Dados”.

Veja que o padrão é o mostrado.

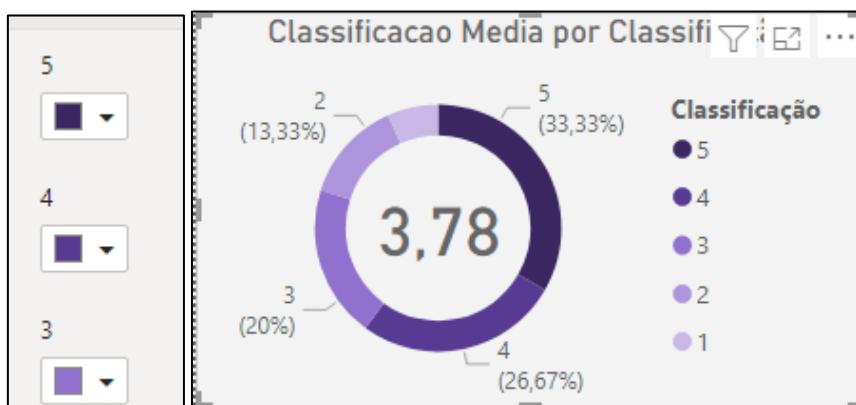


No primeiro elemento, o número 5, selecione a cor mais forte e escura que deseja da paleta de cores. Por exemplo, se deseja trabalhar com azul, escolha o mais forte e vá diminuindo na mesma cor, só que o tom abaixo.



Caso opte por essa em azul brilhante, suba cada um dos tons conforme o nível.

Veja o resultado abaixo



Tabela

Vamos inserir uma tabela, com os campos abaixo, lembre-se de ajustar o Estilo para Nenhum e o Tamanho da Grade para que as linhas não fiquem comprimidas.

Operação	Total de Clientes	Classificacao Media	Média de Tempo de Espera
Depósito Conta-Corrente	5257	3,78	32,
Depósito Poupança	6337	3,76	33,
Gerência	7631	3,74	33,
Pagamento de Contas	6054	3,46	38,
Recebimento Benefício	4792	3,75	36,
Saque Conta-Corrente	7008	3,76	34,
Saque Poupança	9408	4,08	22,
Total	46487	3,78	32,

Técnica da Barra de Dados na Tabela

Para isso, vamos em “Formatação Condisional” e ative a “**Barra de Dados**”.

Selecione o campo que deseja que a Barra de Dados apareça.

Operação	Total de Clientes	Classificacao Media	Média de Tempo de Espera
Saque Poupança	9408	4,08	22,34
Depósito Conta-Corrente	5257	3,78	32,41
Depósito Poupança	6337	3,76	33,87
Saque Conta-Corrente	7008	3,76	34,92
Recebimento Benefício	4792	3,75	36,97
Gerência	7631	3,74	38,42
Pagamento de Contas	6054	3,46	38,35
Total	46487	3,78	32,36

Pegue por exemplo o “Pagamento de Contas”, essa operação é a que demanda mais tempo na Média de Espera.

Agora vamos customizar essas Barras de Dados.

Clique em “Controles Avançados”.

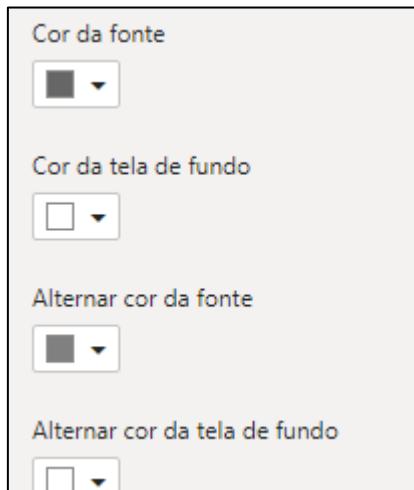


Operação	Total de Clientes	Classificacao	Média	Média de Tempo de Espera
Saque Poupança	9408	4,08		22,34
Depósito Conta-Corrente	5257	3,78		32,41
Depósito Poupança	6337	3,76		33,87
Saque Conta-Corrente	7008	3,76		34,92
Recebimento Benefício	4792	3,75		36,97
Gerência	7631	3,74		33,42
Pagamento de Contas	6054	3,46		38,35
Total	46487	3,78		32,36

De preferência a uma cor das barras bem próximas a tela de fundo do dashboard. Isso fará que essas barras sejam “Integradas”.

Alterar o Cabeçalho da Tabela

Para isso, procure por “Valores”.

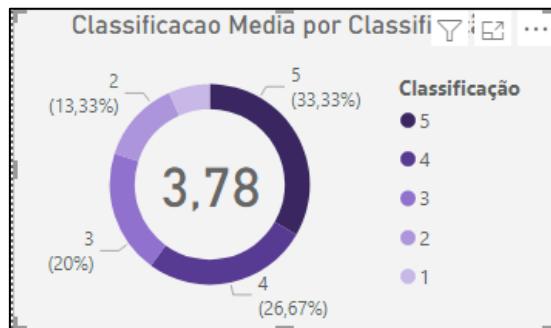


E o cabeçalho da Coluna, como abaixo.

Total de Clientes	Classificacao	Média	Média de Tempo de Espera	Cabeçalhos da coluna
9408	4,08		22,34	Cor da fonte

Painel de Seleção

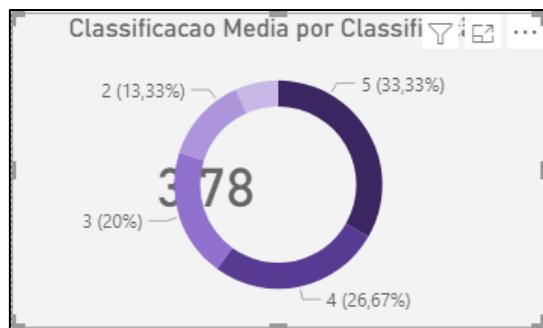
Voltando ao Gráfico de Rosca, veja como ele contém informação demais.



Não é necessário deixar a Legenda e, também a porcentagem, evite informações redundantes.

Remova a Legenda

Veja que o Cartão vai ficar intocável.



Procure em “Exibir > Painel de Seleção” para pode ajustar o Cartão para dentro do raio novamente. Desabilite a visualização temporariamente para ajustar.

The screenshot shows the 'Seleção' (Selection) panel on the left and the donut chart on the right. The selection panel has several options: 'Ordem das...', 'Mostrar' (Show), 'Ocultar' (Hide), 'Tabela' (Table), 'Cartão' (Card), and 'Classificacao Media p...'. The 'Cartão' option is highlighted with a yellow bar under 'Ordem das...'. The donut chart on the right displays the same data as the previous screenshots, with the central value being 3,78 and the segments labeled with their respective values and percentages: 5 (33,33%), 4 (26,67%), 3 (20%), and 2 (13,33%).

Use o Painel de Seleção para ajustar caso algo esteja na frente do outro.

Mapas

Primeiramente, existem dois tipos de Mapas: **Globo** e **Mapa Coroplético**. São os padrões contidos no Power BI. Lembre-se que mais visuais podem ser acionados a qualquer instante.



Mapa de Globo: Plota “bolas” de acordo com a quantidade mensurada. Eficiente para verificar desigualdades regionais (mais em um local e menos em outro). Esse mapa de globo é indicado quando se trabalha com regiões menores, como cidades.



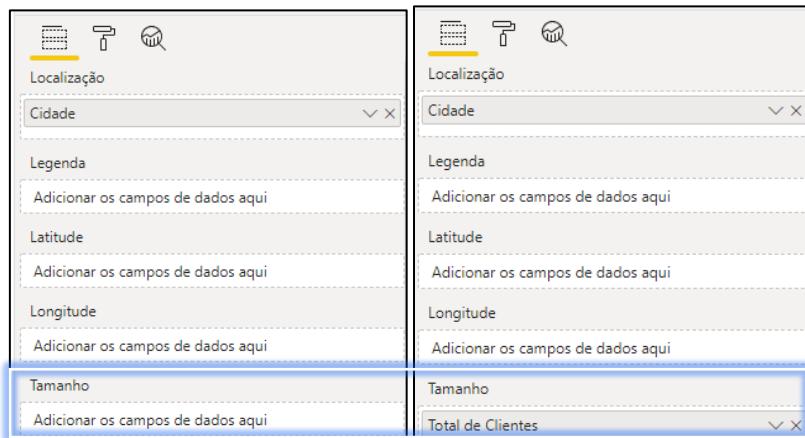
Mapa Coroplético: Apenas colore a região. Utilize esse tipo de mapa quando tratar dados extensos e de muitas regiões como estado/país/continente. Caso contrário, como nesse exemplo estamos tratando apenas de Cidades, a visualização não será de qualidade.



Veja abaixo como identificar qual região possui mais clientes que a outra.

Para descobrir qual região de cidade possui mais clientes que a outra, primeiro, aplicaremos o Mapa de Globo.

Em seguida, é necessário informar ao PBI qual será a unidade mensurável, de tamanho. Ou seja, é a QUANTIDADE/ TOTAL DE CLIENTES.

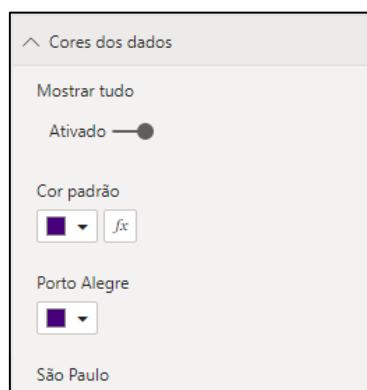


The screenshot shows two side-by-side 'Localização' cards in the Power BI Data View. Both cards have 'Cidade' selected in the 'Localização' dropdown. The left card's 'Tamanho' field is set to 'Adicionar os campos de dados aqui'. The right card's 'Tamanho' field is set to 'Total de Clientes'.

Resultado



Ainda é necessário **formatar** esse Mapa.



The screenshot shows the 'Cores dos dados' (Data Colors) settings in Power BI. The 'Ativado' (Enabled) switch is turned on. Under 'Cor padrão' (Default Color), a dark purple square is selected. Below it, there are two additional color swatches: one for 'Porto Alegre' and another for 'São Paulo'.

Em Mapas, não é uma boa prática alterar as cores dos dados dessa forma, pois, e se em uma atualização os estados mudam? Aumentam as ocorrências? Para isso vamos ao Power Query realizar os ajustes.

Hierarquia dos Mapas

Crie uma Coluna Condicional e altere o tipo para Texto.

Adicionar Coluna Condisional

Adicionar uma coluna condicional que é calculada das outras colunas ou valores.

Nome da nova coluna

Região

	Nome da Coluna	Operador	Valor ⓘ		Saída ⓘ
Se	Cidade	igual a	ABC 123	São Paulo	Então ABC 123 Sudeste
Senão...	Cidade	igual a	ABC 123	Rio de Janeiro	Então ABC 123 Sudeste
Senão...	Cidade	igual a	ABC 123	Belo Horizonte	Então ABC 123 Sudeste
Senão...	Cidade	igual a	ABC 123	Porto Alegre	Então ABC 123 Sul
Senão...	Cidade	igual a	ABC 123	Florianópolis	Então ABC 123 Sul
Senão...	Cidade	igual a	ABC 123	Recife	Então ABC 123 Nordeste

Adicionar Cláusula

Senão ⓘ

ABC 123

null

Retornando ao Relatório, adicione a Região recém criada ao campo de Legenda.

Que o PBI se encarregará de criar a relação.

A map of Brazil and surrounding countries (Argentina, Uruguay, Paraguay, Bolivia, Peru, Chile, Colombia, Suriname) showing the total number of clients per city and region. The map uses a color-coded legend where purple dots represent different regions: Nordeste (dark purple), Sudeste (medium purple), and Sul (light purple). The size of the dots corresponds to the number of clients in each city. A legend box in the top-left corner identifies the controls for 'Localização' (City, Legend, Region), and a zoom control (+/-) is located in the bottom-right corner.

Total de Clientes por Cidade e Região

Região ● Nordeste ● Sudeste ● Sul

Bing

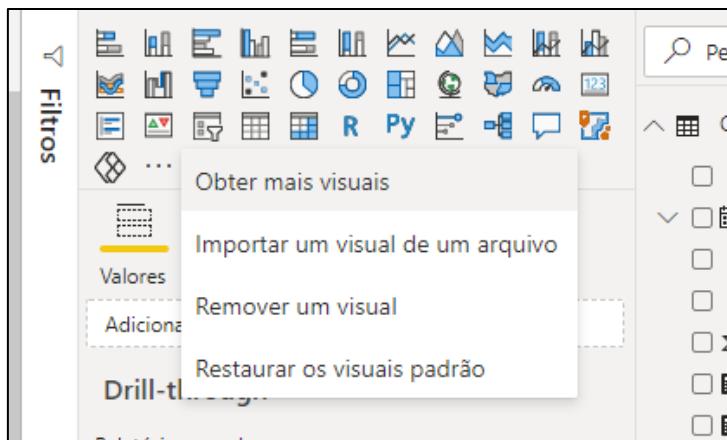
© 2021 TomTom, © 2021 Microsoft Corporation. [Terms](#)

Lembre-se de aumentar as Bolhas e alterar as cores.

Em “Controles de Mapas” adicione os “**Botões de Zoom**”, que o Mapa ficará igual acima.

Marketplace

O **marketplace** é um local onde **devs** de Power BI colocam os gráficos que eles desenvolveram. A linguagem utilizada é a **C#**.

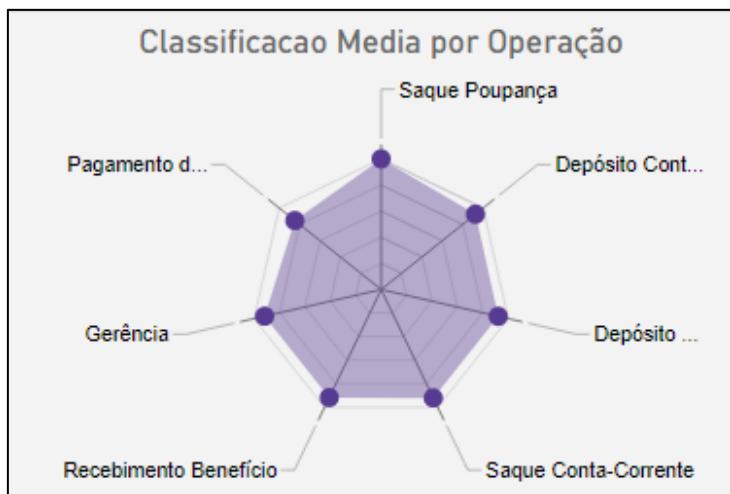


Clique em “Obter mais Visuais”.

Os visuais certificados e homologados pela Microsoft são os que possuem o sinal de “Check” em azul.



Por exemplo:



Esse gráfico foi importado.

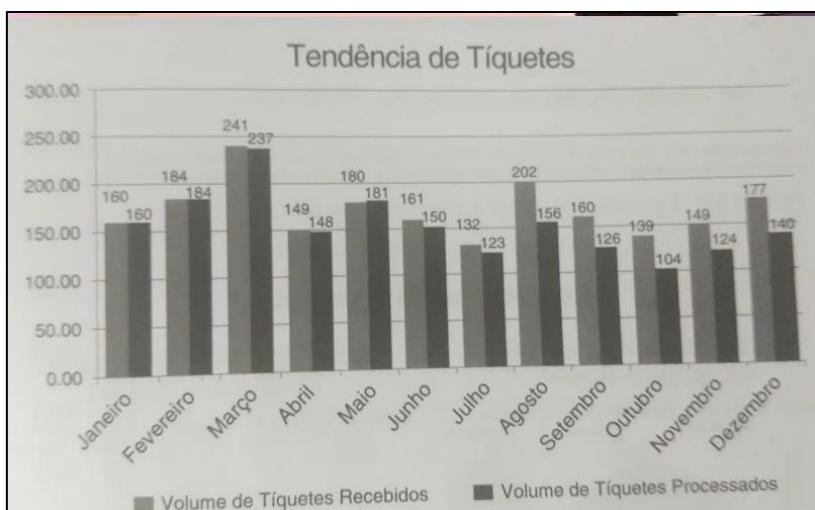
Dashboard Finalizado

Um exemplo de Dashboard finalizado.

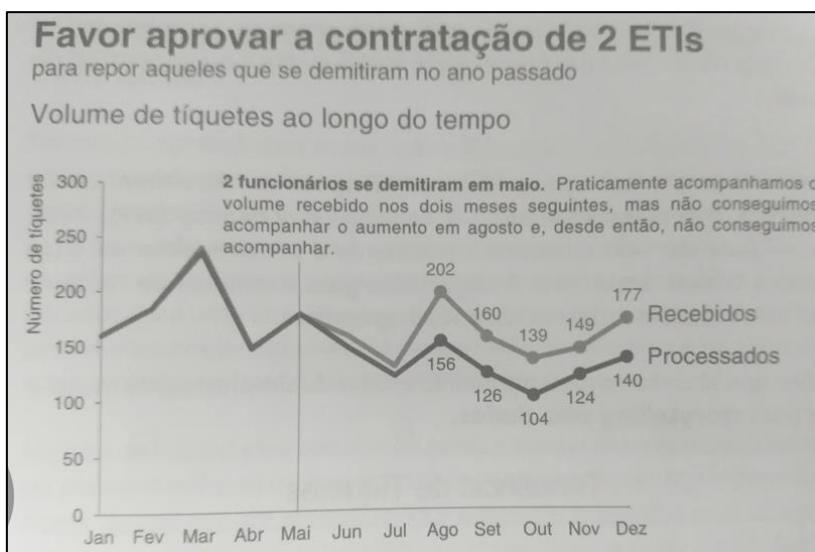


Storytelling

Storytelling é contar uma história com letras e números. Uma empresa que recebe tickets(em cinza claro) e que processa tickets(em cinza escuro).



Agora, veja abaixo. Adicionando um CONTEXTO.



Handout

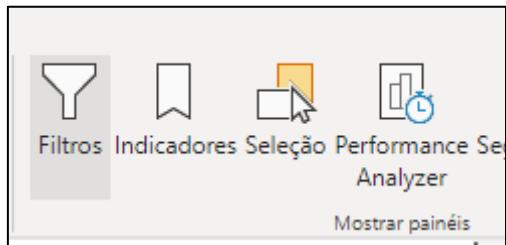
Uma vez que não é possível inserir contexto no gráfico, seja por aparência ou visualização, utilizamos o Handout, é um papel, uma informação em texto, um folder que contém uma apresentação sobre o contexto dos dados. O Handout é utilizado para enviar para pessoas que não conhecem os dados.

Bookmarks

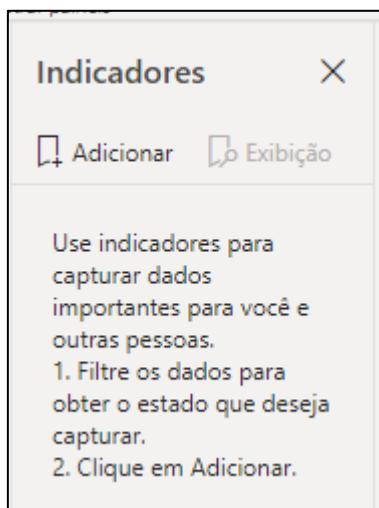
É uma forma de contar históricas através do Dashboard. É contar histórica no gráfico.

Painel de Indicadores

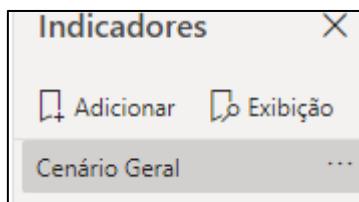
Vá em “Exibição > Painel de Indicadores”, esses são os “Bookmarks”.



O resultado será uma coluna ao lado indicado o Indicadores.



Clique acima em Adicionar e renomeie para “Cenário Geral”.

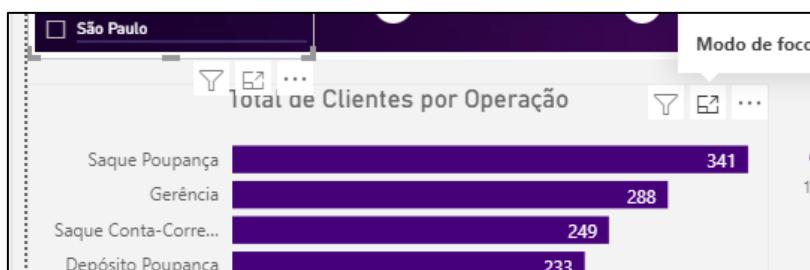


Agora, clique em Rio de Janeiro, lá na Segmentação de Dados e depois, clique em Adicionar mais um indicador.

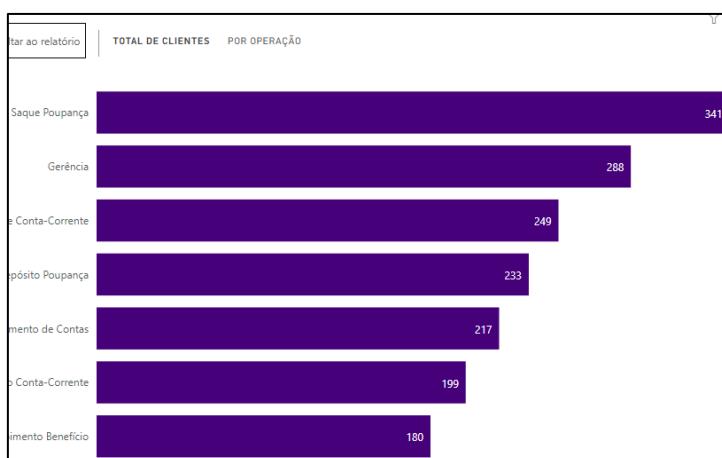


Acima, é um dashboard exclusivamente da situação do Rio de Janeiro.

Mas, e se em algum momento precisássemos evidenciar os Clientes por Operação, exclusivamente do RJ? Continue com ele selecionado, clique no PIN/Modo de Foco do Gráfico de Barras.



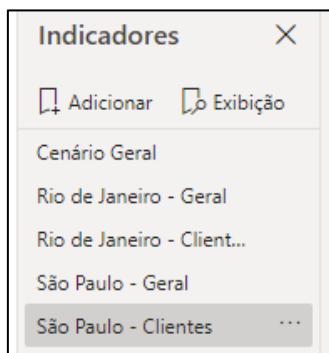
Esse será o resultado clicando no PIN



Adicione um Bookmark (indicador) mais um vez.

Volte ao relatório e repita o procedimento para cada cidade, ou para cada ação que deseja que seja mostrada.

Por exemplo, abaixo repetimos os passos de mostrar a situação geral e os clientes com o Foco para a cidade de São Paulo.



Vantagens

As Vantagens dos Bookmarks, é a velocidade de reprodução que fica armazenada em cada indicador. Durante uma apresentação, basta clicar nos indicadores e a atualização da página é realizada rapidamente, mais fácil do que clicar em cada um dos filtros separadamente.

Exibição dos Bookmarks

Ao lado de Adicionar, existe a opção de “Exibição”, é uma apresentação dos Bookmarks e contém uma barra de rolagem com setas na parte inferior do dashboard.

Operação	Total de Clientes	Classificacao Media	Média de Tempo de Espera
Depósito Conta-Corrente	5257	3,78	32,41
Depósito Poupança	6337	3,76	33,87

Cenário Geral < > X

Imagen com Filtro de Bookmark

Adicione uma imagem no dashboard.

Volte a tela inicial, sem filtro nenhum ativo, sem nenhuma movimentação , em nenhum gráfico.

Agora, clique na imagem recém adicionada, e adicione um novo bookmark.

Ainda com essa imagem selecionada, vá em “Formatar Imagem”, ao lado dos Indicadores. E procure por “Ação”. Ative.

Indicadores X Formatar imagem

Indicador

Ativado

Título Desativado

Tela de fundo Desativado

Fixar proporção Ativado

Borda Desativado

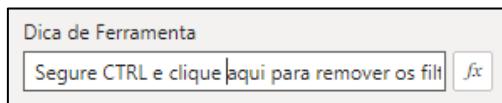
Sombra Desativado

Ação Desativado

Selecione o tipo como um Indicador/Bookmark.

E o “Indicador” deve ser algo que foi criado para ser associado a essa ação. E nesse caso, queremos associar o Bookmark que “é o cenário geral, sem filtro, sem ação”.

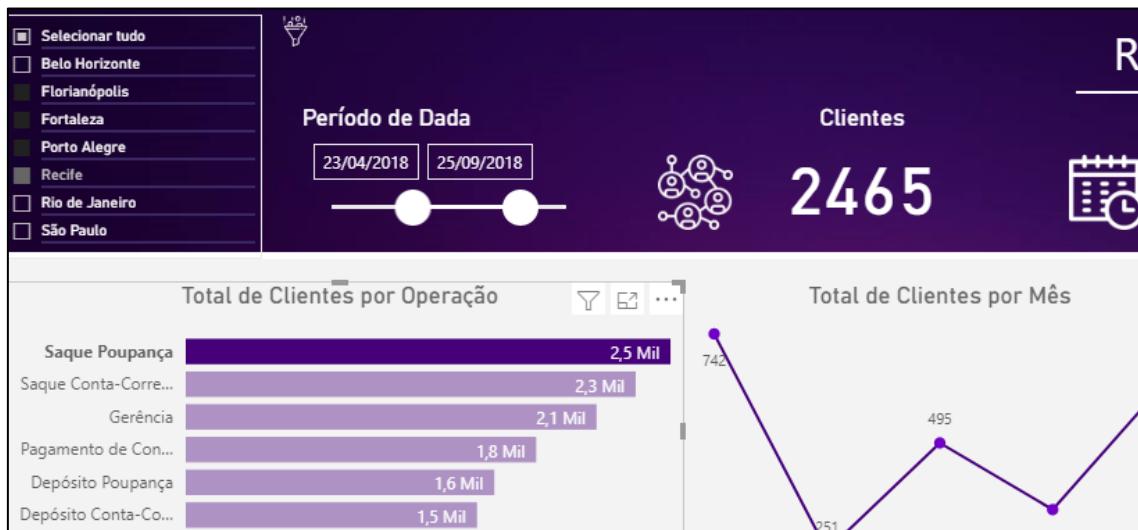
Não se esqueça de informar com uma dica a ação que será realizada.



Ok.

Agora, bagunce o dashboard.

Selecione várias cidades, altere o range de datas..

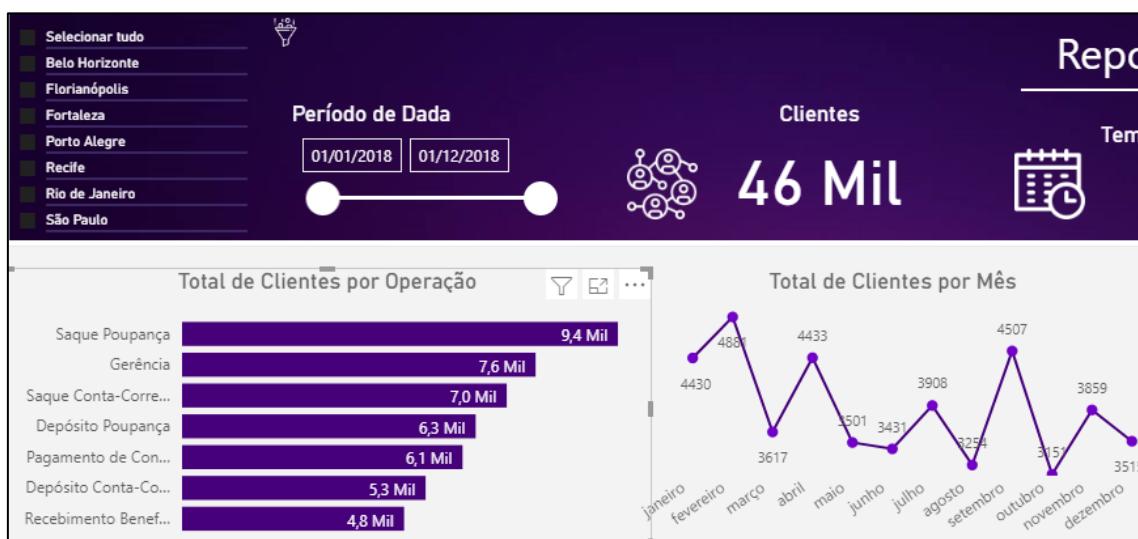


Perceba que o dashboard está uma bagunça.

Clique com o **CTRL** na imagem do filtro que o padrão será reestabelecido ao padrão normal.



Veja o padrão reestabelecido.



Desenvolvimento Extra curso

Java Run Time Environment (JRE) – Instalação

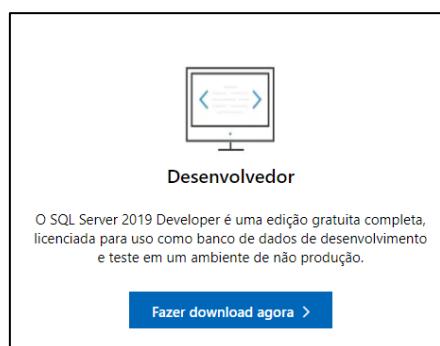
Procure no Google por “Java JRE”.

É necessário para a instalação completa do SQL Server.

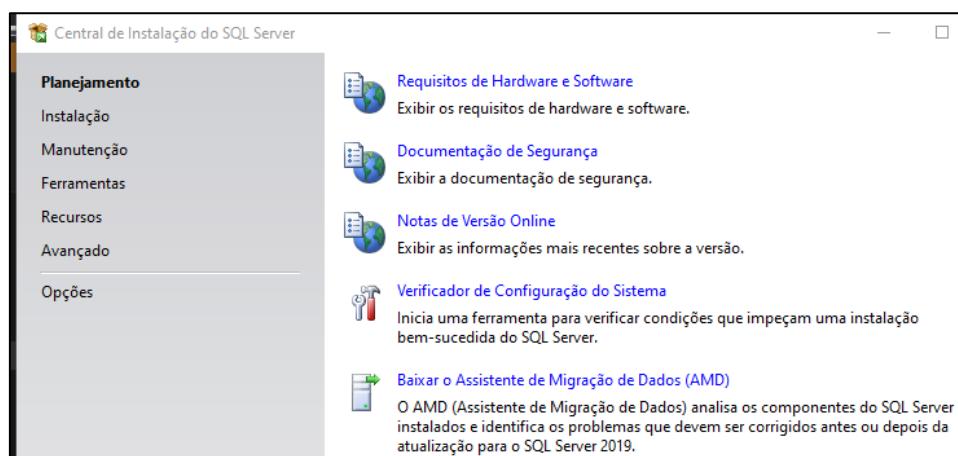
SQL Server 2019 – Instalação

Procure no Google.

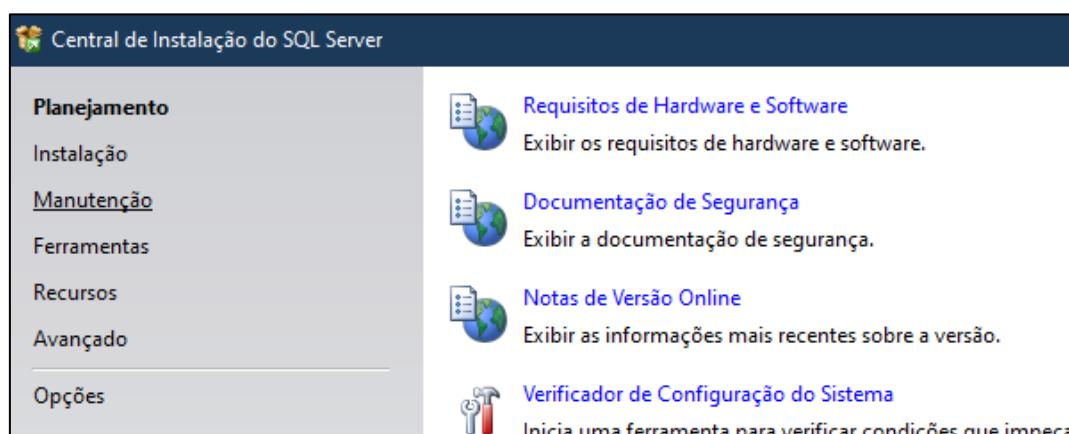
Baixe a versão **developer**.



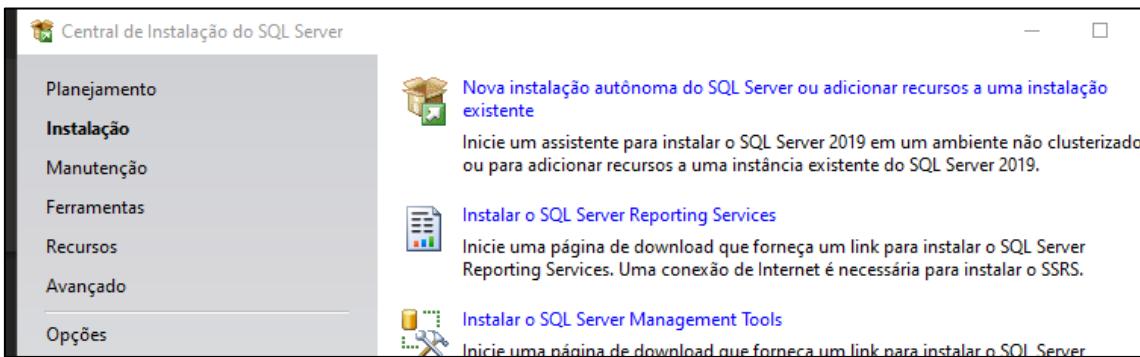
Após clicar para a instalação, uma outra janela será disponibilizada.



Clique na aba de Instalação.



Selecione a primeira opção (**Stand Alone**), que é a “**Nova Instalação do SQL Server**”

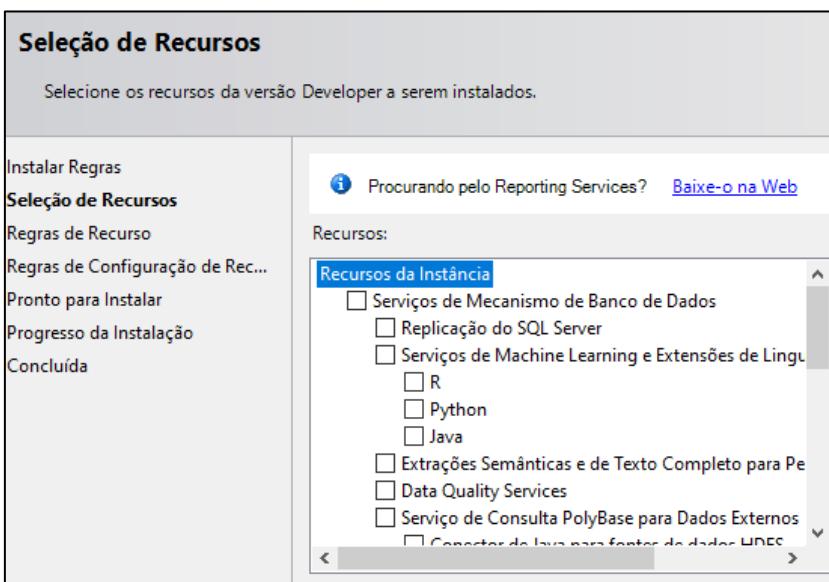


Como estamos trabalhando em um ambiente de desenvolvimento para estudos, marque a opção de “**Developer**”.



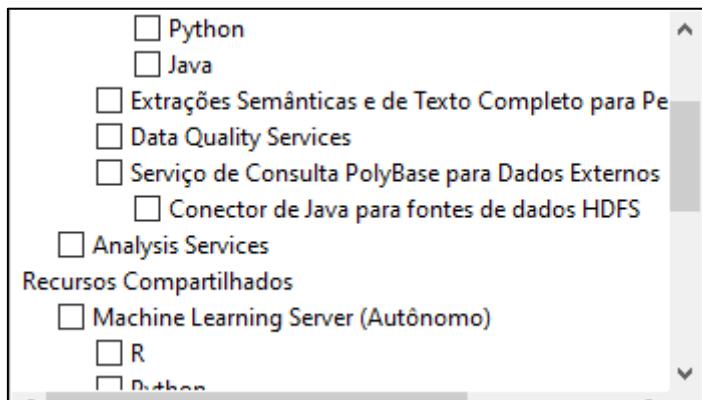
Aceite os termos e avance.

A próxima etapa é a de “**Seleção de Recursos**”.



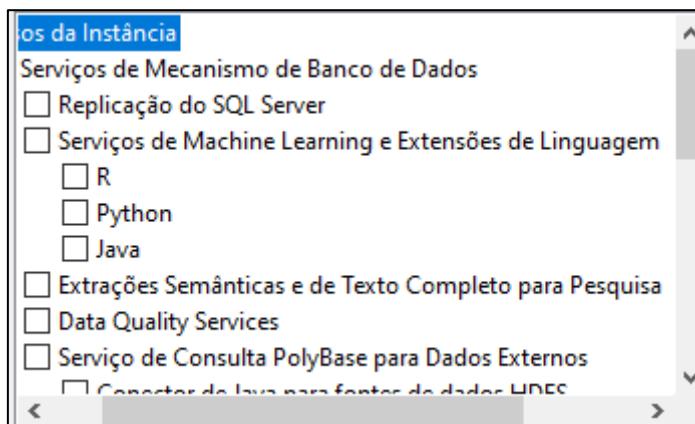
Analysis Services

Alguns elementos como o “*Analysis Services*” que faz a comunicação com o Cubo.



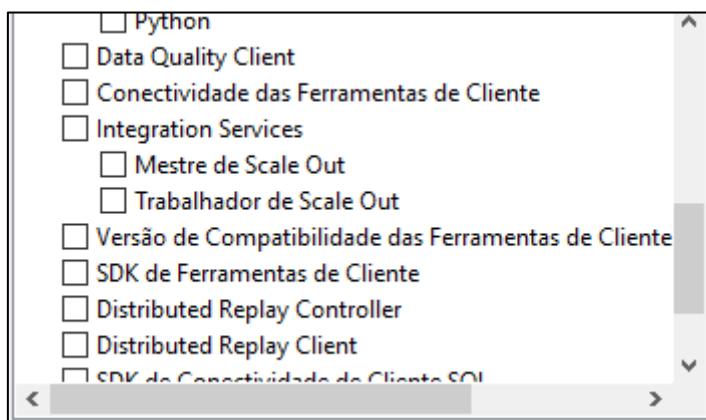
PolyBase

Ou como a “*PolyBase*” para trabalhar com Big Data e é necessário o Java instalado na máquina.



Integration Services

“*Integration Services*”



E enquanto ao “*Reporting Services*”, que é o servidor de relatório.

Reporting Services

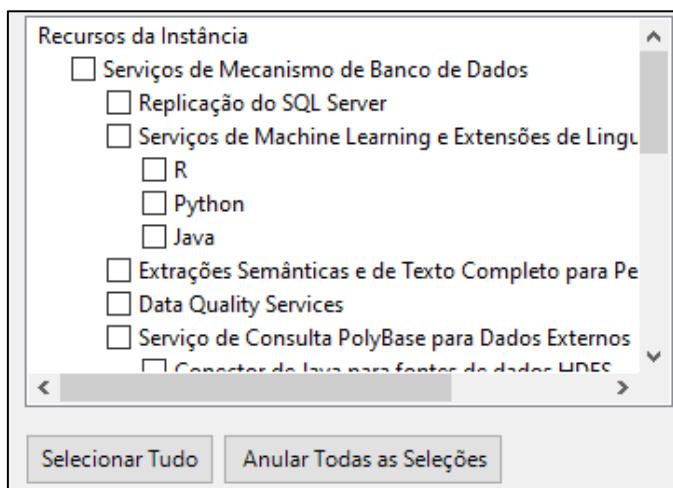
Clique acima e faça o download diretamente pelo link descrito no instalador.



A partir da versão de 2019, o instalador do Reporting Services é realizado “por fora”, através do site: <https://www.microsoft.com/pt-BR/download/details.aspx?id=55252>.

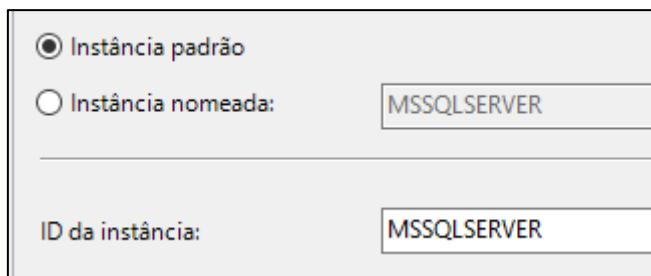
Instalando as Features do SQL Server

Clique em “Select All”/”Selecionar Todas” e avance.



Instale como a “instância Default”, pois é um ambiente própria.

Uma instância é uma instalação de um Banco de Dados, e dentro de uma única instalação, podem conter vários Bancos de Dados.



Contas de Serviços

A Microsoft delega e cria os serviços e as contas que realizaram os serviços.

Serviço	Nome da Conta	Senha	Tipo de Inicialização
SQL Server Agent	NT Service\SQLSERVER...		Manual
Mecanismo de Banco de Dados ...	NT Service\MSSQLSER...		Automática
SQL Server Analysis Services	NT Service\MISSQLServ...		Automática
SQL Server Integration Services 1...	NT Service\MsDtsServe...		Automática
Master do SQL Server Integratio...	NT Service\SSISScaleO...		Automática
Trabalho do SQL Server Integrati...	NT Service\SSISScaleO...		Automática
SQL Server Distributed Replay Cli...	NT Service\SQL Server ...		Manual
SQL Server Distributed Replay C...	NT Service\SQL Server ...		Manual
Mecanismo PolyBase do SQL Ser...	AUTORIDADE NT\SE...		Automática
Movimentação de Dados PolyBa...	AUTORIDADE NT\SERVI...		Automática
Launchpad do SQL Server	NT Service\MSSQLLau...		Automática
Iniciador do Daemon de Filtro d...	NT Service\MSSQLFDL...		Manual
SQL Server Browser	NT AUTHORITY\LOCA...		Desabilitado

Modo de Autentificação

Esse modo é a segurança do Banco de Dados.

Sempre selecione a opção “Modo Misto”, e informe a senha.

Instalar Regras	Configuração do Servidor
Seleção de Recursos	Diretórios de Dados
Regras de Recurso	TempDB
Configuração da Instância	MaxDOP
Configuração do PolyBase	Memória
Local de Instalação do Java	FILESTR
Configuração do Servidor	
Configuração do Mecanismo d...	Especifique o modo de autenticação e os administradores para o Mecanismo de Banco de Dados.
Configuração do Analysis Servi...	
Configuração do Integration Se...	Modo de Autenticação
Configuração do Integration Se...	<input checked="" type="radio"/> Modo de Autenticação do Windows
Distributed Replay Controller	<input type="radio"/> Modo Misto (autenticação do SQL Server e do Windows)
Distributed Replay Client	Especifique a senha da conta do sa (administrador do sistema) do SQL Server.
	Digite Senha:
	Confirmar senha:
	Especificar administradores do SQL Server

Isso significa que ele vai aceitar o usuário do Windows e também solicitar uma senha.

Clique em “Adicionar Usuário Atual”.

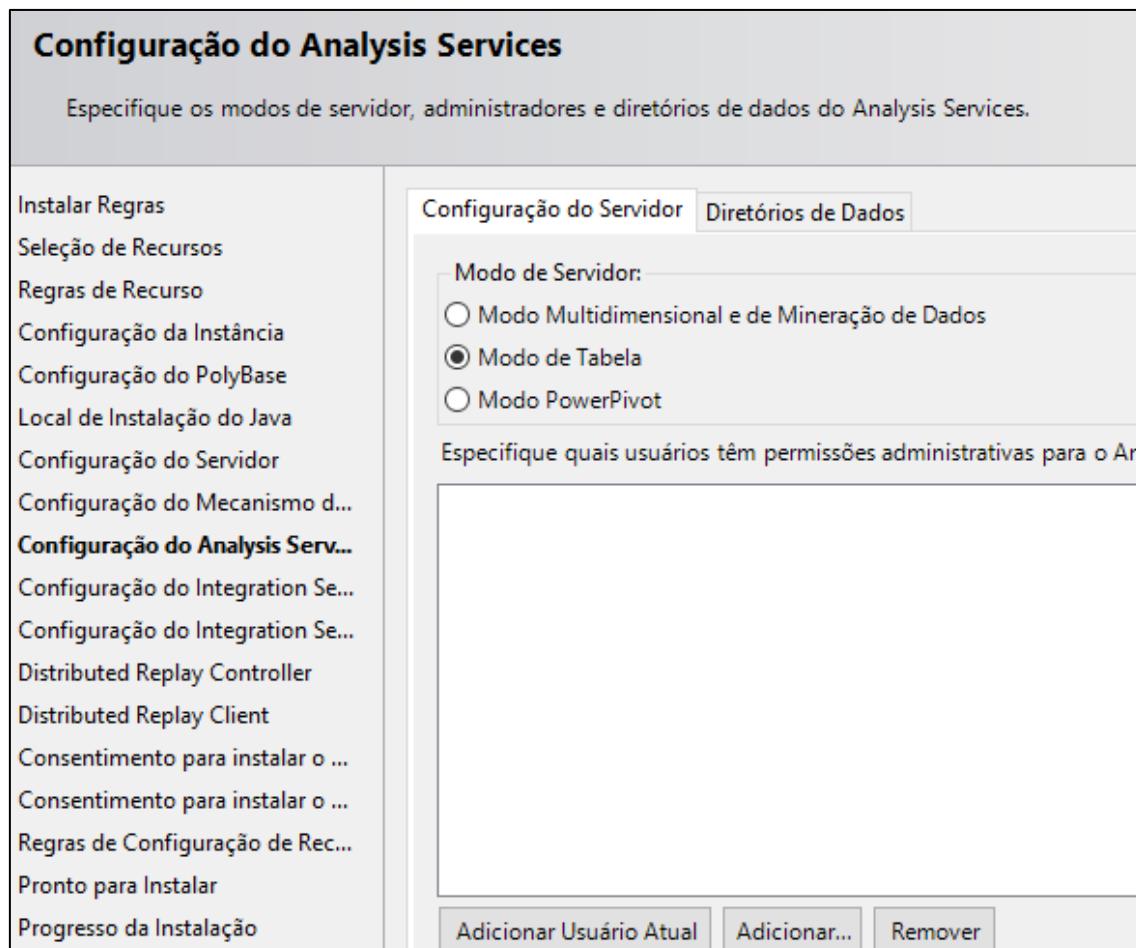
Digitar Senha:	****
Confirmar senha:	****
Especificando administradores do SQL Server	
<input type="button" value="Adicionar Usuário Atual"/> <input type="button" value="Adicionar..."/> <input type="button" value="Remover"/>	

Instalação do Analysis Services do SQL

Durante a instalação do SQL Server, a opção do “*Analysis Services*” é solicitada.

E a pergunta é sobre, “Para o seu servidor de Cubo, você quer o modo tabular ou multidimensional”. Selecione por enquanto o “Modo de Tabela/Tabular”.

Não esqueça de adicionar o usuário igual foi o método de autentificação.



Master Node – Nô Mestre

Não vamos configurar nesse curso, apenas avance.

Configuração do Integration Services Scale Out – Nô Mestre

Especifique o número da porta e o certificado de segurança para o nô mestre de scale out.

Instalar Regras Seleção de Recursos Regras de Recurso Configuração da Instância Configuração do PolyBase Local de Instalação do Java Configuração do Servidor Configuração do Mecanismo d... Configuração do Analysis Servi... Configuração do Integration S... Configuração do Integration Se... Distributed Replay Controller Distributed Replay Client Consentimento para instalar o ... Consentimento para instalar o ... Regras de Configuração de Rec... Pronto para Instalar Progresso da Instalação Concluída	Especifique o número da porta que o nô mestre usará para comunicação com os nós de trabalho na topologia de scale out. Um certificado autoassinado padrão será criado se você não escolher um novo certificado SSL. Número da Porta: <input type="text" value="8391"/>
	Selecione um certificado SSL que seja usado para a comunicação entre o nô mestre e os trabalhos na topologia de scale out. Um certificado autoassinado padrão será criado se você não escolher um novo certificado SSL. <input checked="" type="radio"/> Criar um novo certificado SSL CNs no certificado: <input type="text" value="CN=desktop-smahrm4; CN=192.168.56.1"/> <input type="radio"/> Usar um certificado SSL existente <input type="text"/> <input type="button" value="Procurar..."/>
	<input type="button" value="Avançar >"/> <input type="button" value="Voltar <"/>

Worker Node – Nô de Trabalho

Não vamos configurar nesse curso, apenas avance.

Configuração do Integration Services Scale Out – Nô de Trabalho

Especifique o ponto de extremidade do nô mestre e o certificado de segurança usados pelo nô de Trabalho de Scale Out.

Instalar Regras Seleção de Recursos Regras de Recurso Configuração da Instância Configuração do PolyBase Local de Instalação do Java Configuração do Servidor Configuração do Mecanismo d... Configuração do Analysis Servi... Configuração do Integration S... Configuração do Integration S... Distributed Replay Controller Distributed Replay Client Consentimento para instalar o ... Consentimento para instalar o ... Regras de Configuração de Rec... Pronto para Instalar Progresso da Instalação Concluída	Forneca o ponto de extremidade do nô mestre ao qual o nô de trabalho precisa se conectar: https://[NomeDoComputadorDoNôMestre]:[Porta]; <input type="text" value="https://desktop-smahrm4:8391"/>
	Configure o certificado SSL do nô mestre para que este computador seja confiável. Esse certificado precisa ser confiável neste nô de trabalho para que a conexão entre o trabalhador e o nô mestre seja estabelecida. Isso é opcional quando o certificado do nô mestre já é confiável, ou seja, quando é emitido por uma AC (autoridade de certificação) confiável ou quando o nô mestre já está no mesmo computador. Caso contrário, forneça um certificado SSL de cliente do nô mestre (certificado criado e autoassinado por você e cujo mestre não esteja no mesmo computador).
	<input type="button" value="Avançar >"/> <input type="button" value="Voltar <"/>

Reply Controller
Apenas adicione o usuário.

Distributed Replay Controller

Especifique as permissões de acesso ao serviço Distributed Replay Controller.

Instalar Regras	Especifique os usuários que têm permissões para o ser...
Seleção de Recursos	
Regras de Recurso	
Configuração da Instância	
Configuração do PolyBase	
Local de Instalação do Java	
Configuração do Servidor	
Configuração do Mecanismo d...	
Configuração do Analysis Servi...	
Configuração do Integration Se...	
Configuração do Integration Se...	
Distributed Replay Controller	
Distributed Replay Client	
Consentimento para instalar o ...	
Consentimento para instalar o ...	
Regras de Configuração de Rec...	
Pronto para Instalar	
Progresso da Instalação	
Concluída	

Adicionar Usuário Atual **Adicionar...** **Remover**

Distributed Replay Client
Não faça nada.

Distributed Replay Client

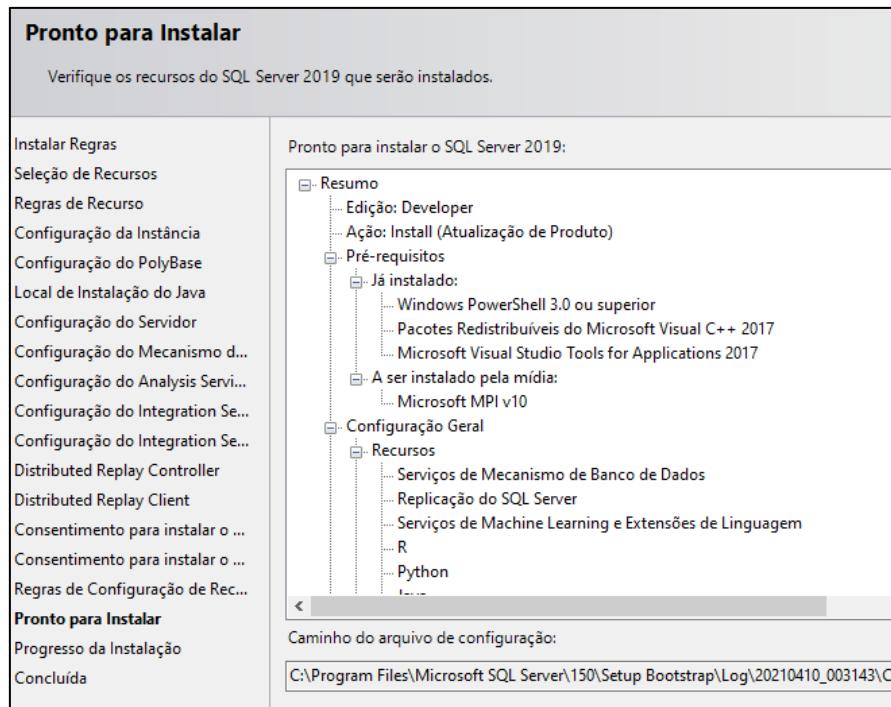
Especifique o controlador correspondente e os diretórios de dados para o Distributed Replay Client.

Instalar Regras	Especifique o nome do computador do controlador e os locais do diretório.
Seleção de Recursos	
Regras de Recurso	
Configuração da Instância	
Configuração do PolyBase	
Local de Instalação do Java	
Configuração do Servidor	
Configuração do Mecanismo d...	
Configuração do Analysis Servi...	
Configuração do Integration Se...	
Configuração do Integration Se...	
Distributed Replay Controller	
Distributed Replay Client	
Consentimento para instalar o ...	
Consentimento para instalar o ...	
Regras de Configuração de Rec...	
Pronto para Instalar	
Progresso da Instalação	
Concluída	

< Voltar **Avançar >**

Roteiro-Resumo da Instalação

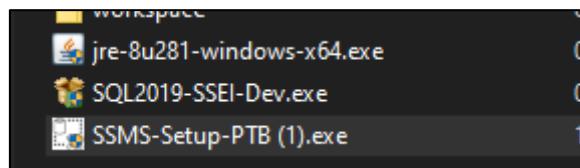
No final, após todas as configurações, e até mesmo do consentimento de instalar Python e R, o roteiro é apresentado.



SQL Management Studio Instalação

Clique novamente no instalador do SQL.

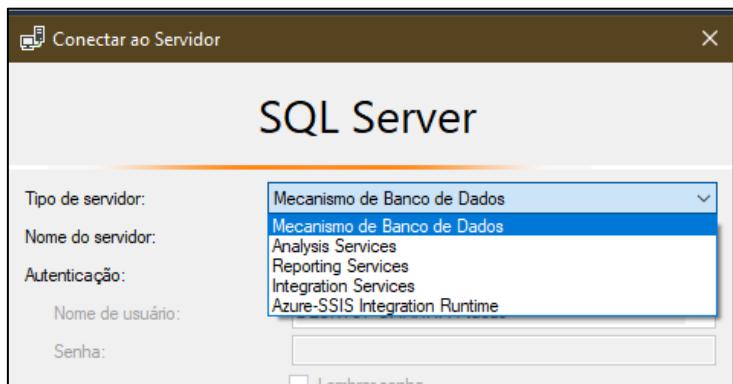
E selecione o “Install SQL Server Management Tools/Studio”, que pode ser encontrado no site: <https://docs.microsoft.com/pt-br/sql/ssms/download-sql-server-management-studio-ssms?redirectedfrom=MSDN&view=sql-server-ver15>.



Apenas instale.

Para abrir, procure no menu iniciar por “SQL Server Management Studio”.

Para entrar selecione o Mecanismo de Banco de Dados, mas é nessa tela de conexão de servidor que realizaremos o acesso das ferramentas mais pra frente



Instalação do Integration Service e Analysis Service

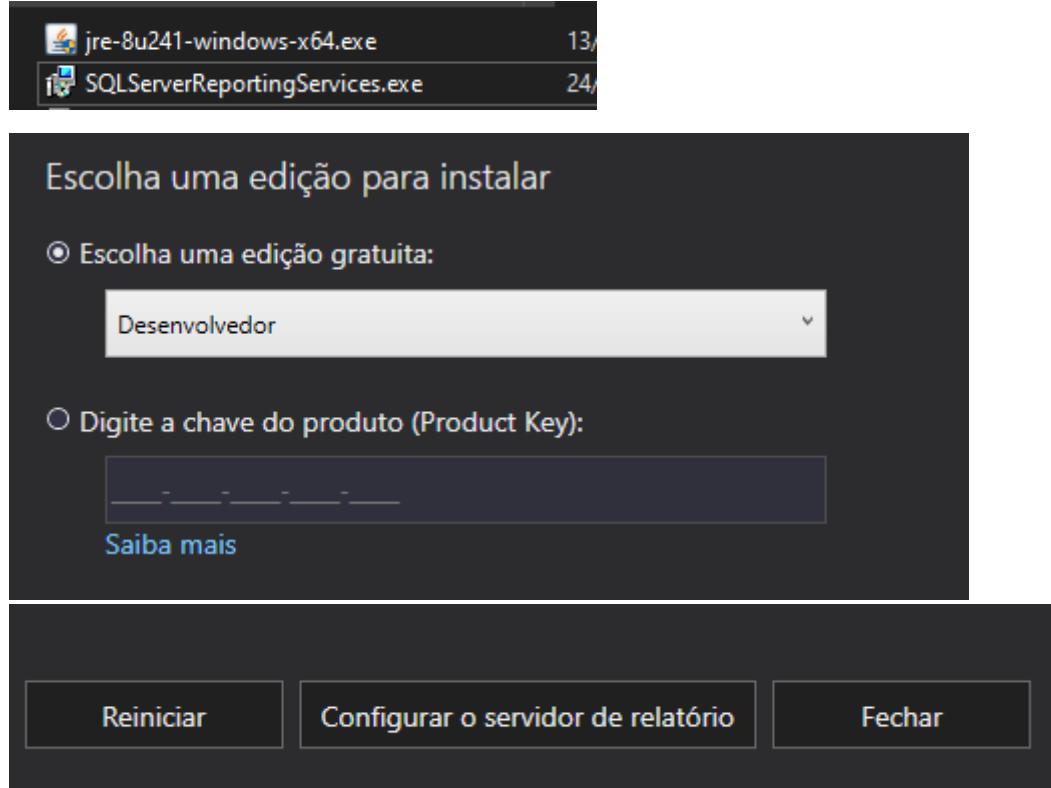
É a instalação dos Datas Tools através do Visual Studio, pelo acesso do site:
<https://visualstudio.microsoft.com/pt-br/downloads/>.

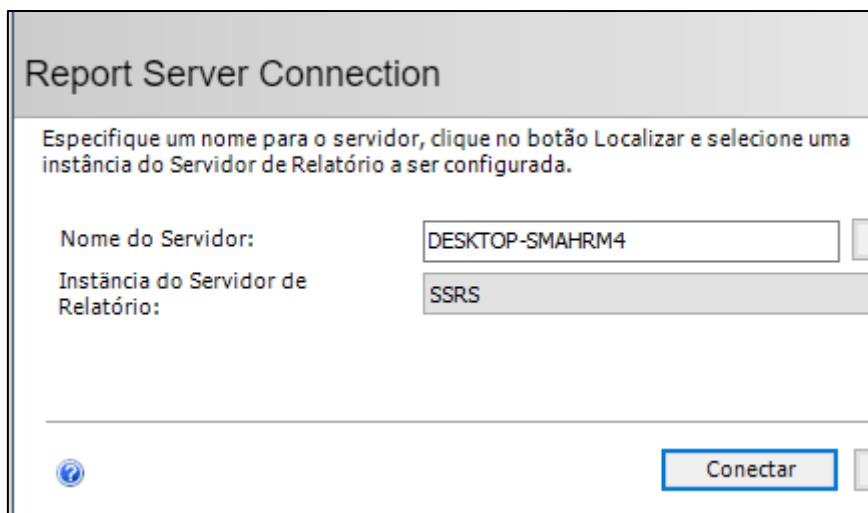
Instale o Microsoft Visual Studio e selecione a opção de Banco de Dados e partir dali, selecione as opções do SSDT (Microsoft SQL Server Data Tools).

Siga os passos da instalação pelo Visual Studio Xano.

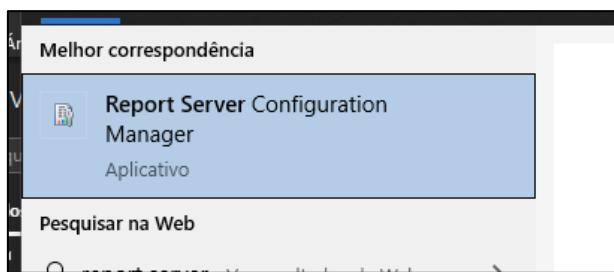
Reporting Services

Instale o Microsoft SQL Server 2017 Reporting Services

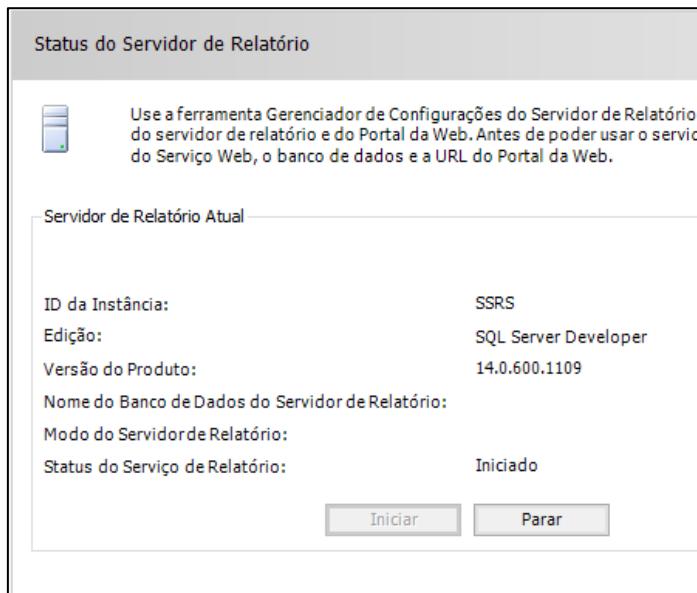




Para acessar no menu iniciar:



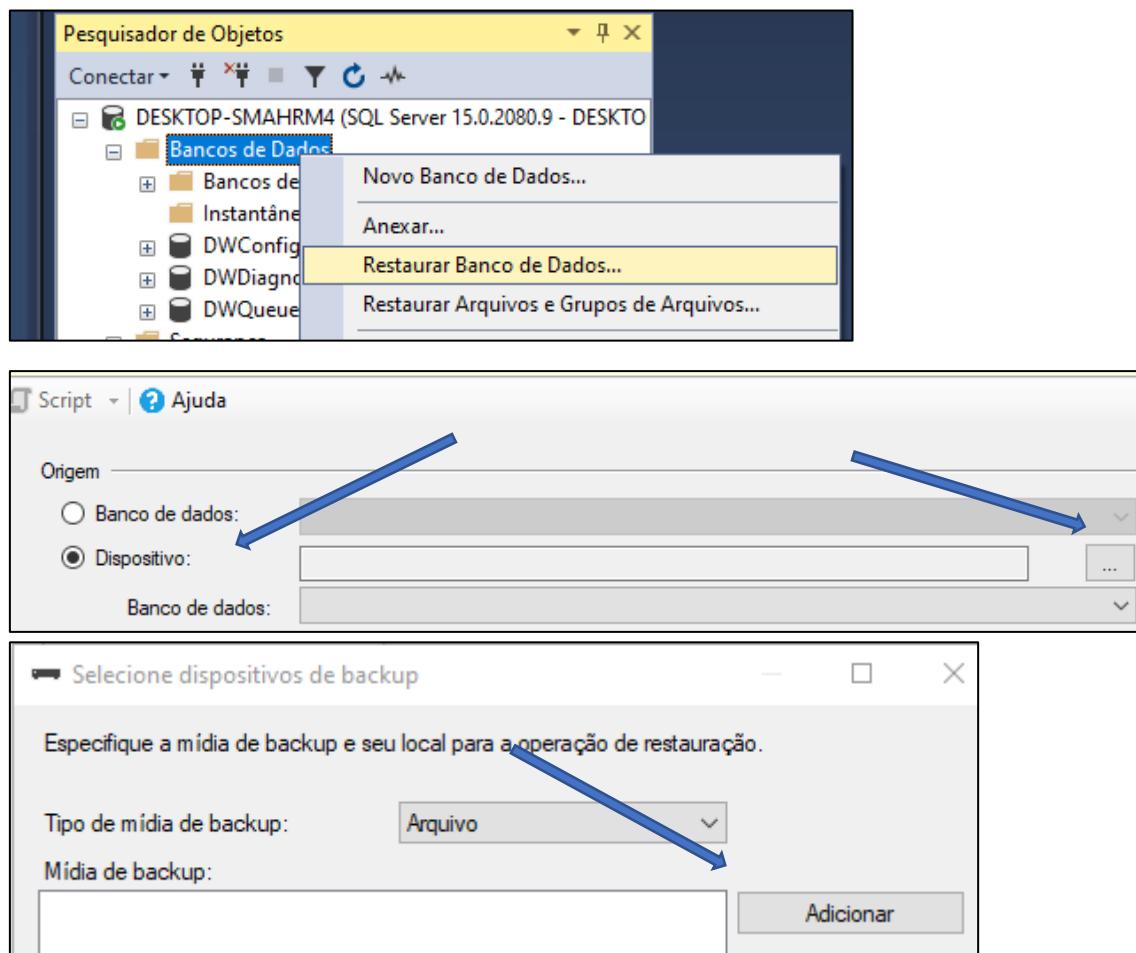
Para parar o servidor:



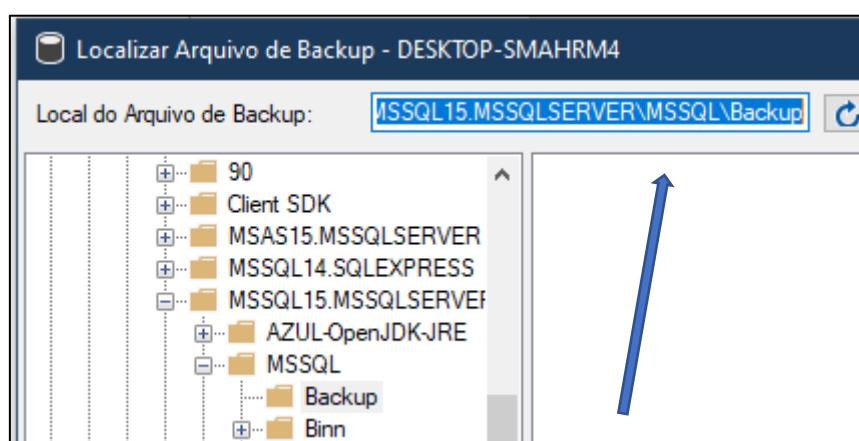
Abra o SQL Server Management Studio.

Estabeleça a conexão com o servidor, nesse exemplo, foi o “DESKTOP-SMAHRM4”.

Clique com o botão esquerdo no Banco de Dados > Restaurar Banco de Dados.

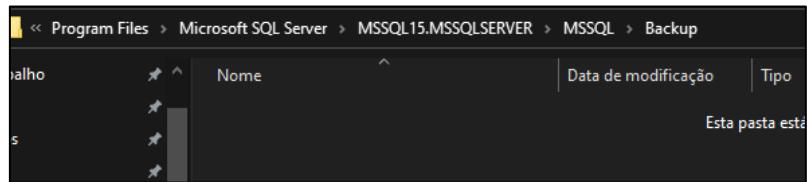


Para restaurar um Banco de Dados, é preciso que ele esteja dentro da pasta de backup do servidor de SQL. (C:\Program Files\Microsoft SQL Server\MSSQL15.MSSQLSERVER\MSSQL\Backup)

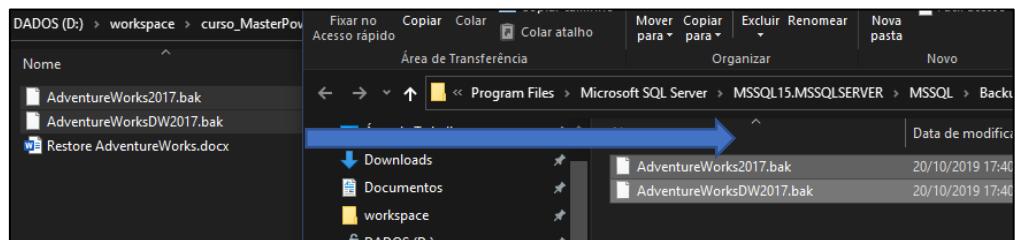


Copie o endereço e abra num explorador de arquivo.

Ele provavelmente estará vazio.



Adicione os bancos lá dentro.



Adicione os dois, selecione um de cada vez e faça o restore.

Overview do BI

SSIS – SQL Server Integration Service

O SSIS, ou é Server do Integration Service. Ele é o responsável pela integração entre o ambiente OLTP (transação, operacional) com o ambiente OLAP (de análise).

Datamart

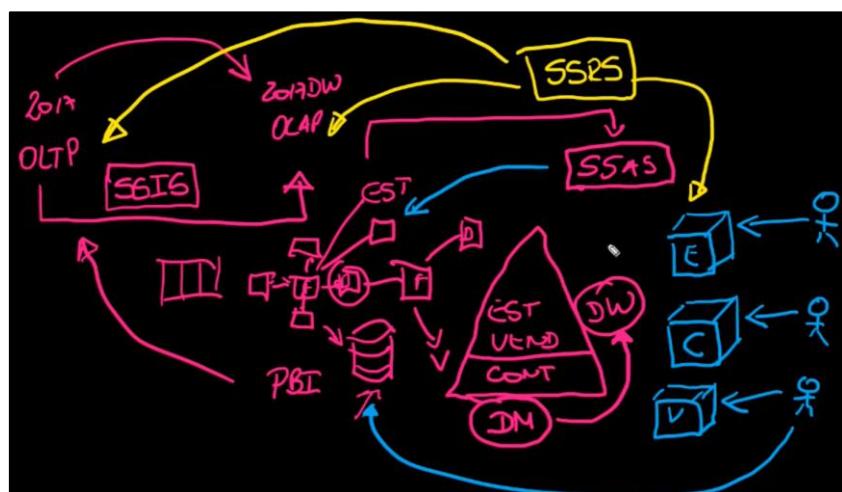
É parte de um DW, que corresponde a uma área de negócio da empresa.

SSAS – SQL Server Analysis Services

Ele transforma a tabela/física em uma abordagem lógica. Ele observa como “desenhos”.

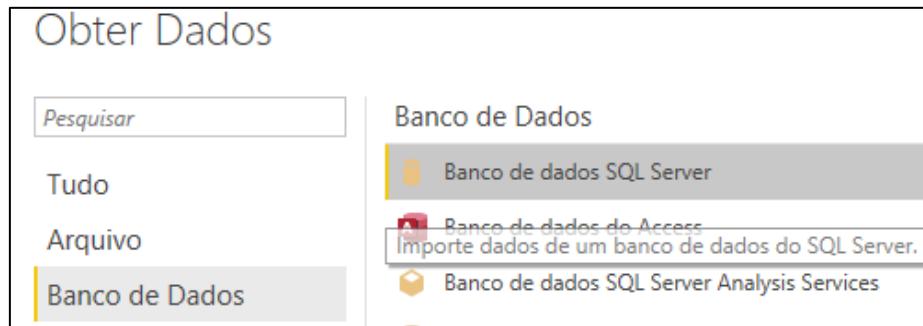
SSRS – SQL Server Reporting Service

O Reporting Service ele tem “habilidade” de buscar o dado onde não ocorreu nenhum tratamento do dado (lá no OLTP) ou no ambiente onde os dados foram tratados e estão em formato STAR ou SNOWFLAKES (que é o ambiente OLAP) ou até mesmo no SSAS.



Teste de Conexão entre PBI e BD

Acesse o Power BI e > Obter Dados > Banco de Dados > Banco de Dados SQL Server.

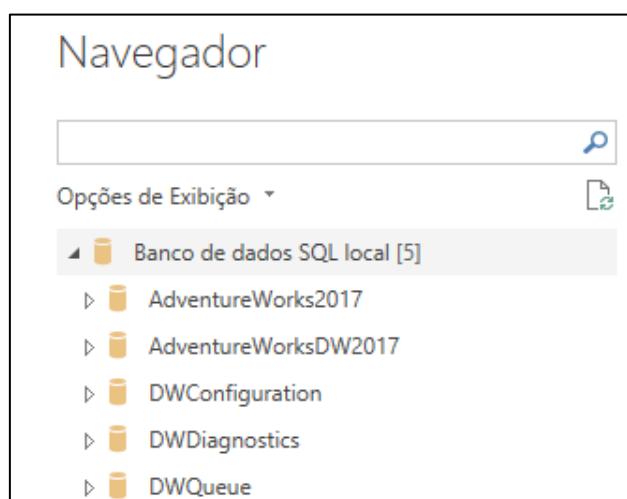


Se for “localhost”: Digite um “.”

Se for “Conexão em Rede”, o DBA/Analista de Rede precisa informar o servidor.



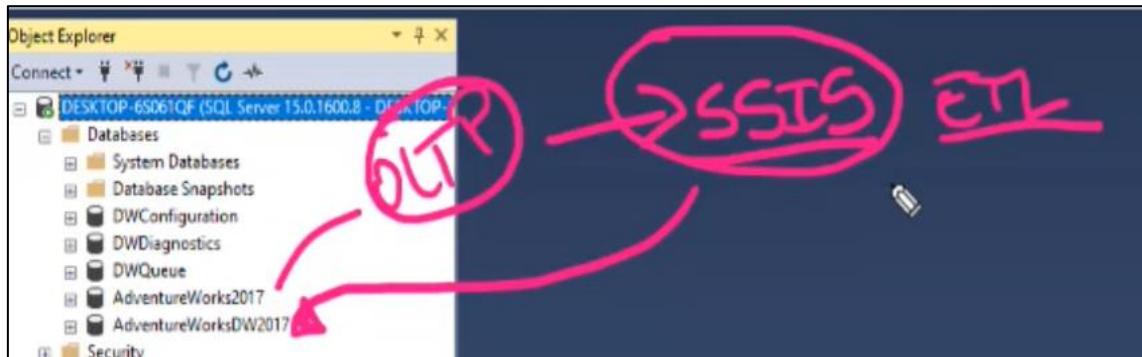
Pronto, conectado.



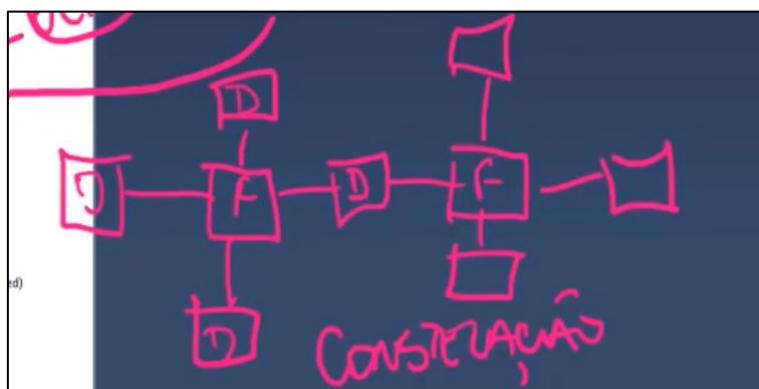
Somente teste de conexão. Cancele.

Mais um Overview

Veja que o DWAdventureWorks2017 é um OLTP, ou seja, um ambiente transacional. Ele deve ser tratado pela ferramenta de ETL que será o SQL Server Integration Service, e aí os dados serão modelados e transformados para o DW.



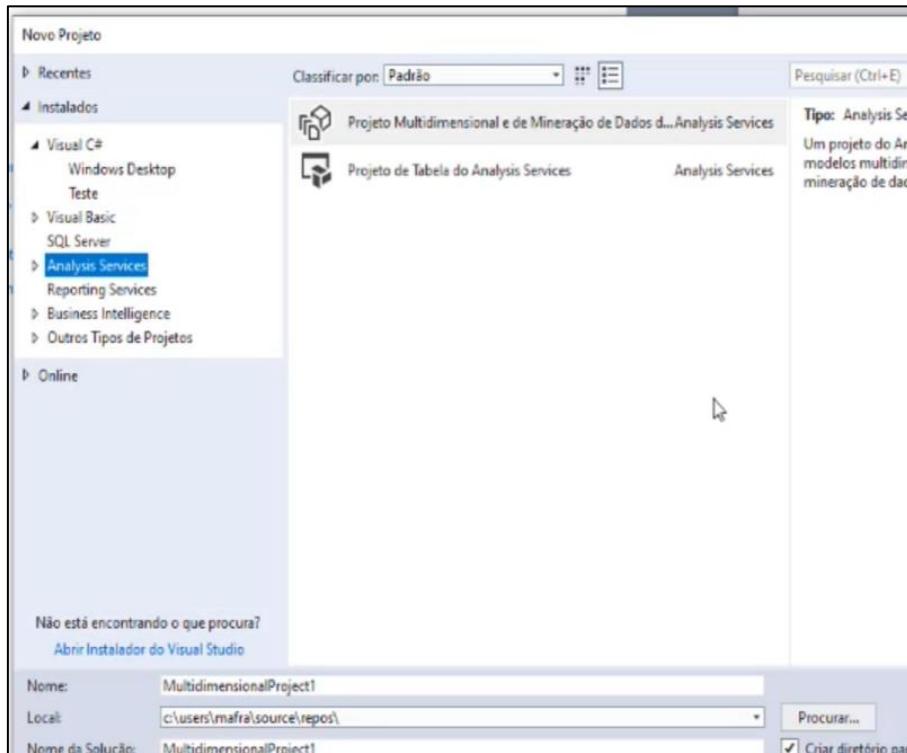
Quando esse passo ocorre, o DW é criado e é criado como um SNOWFLAKE, ou seja, fato ligado a dimensões. Quando se têm vários fatos ligados as várias dimensões, teremos uma “constelação”.



Criando o Cubo no Analysis Services

Abra o Visual Studio xAno (SSDT).

Crie um: Novo projeto > Projeto de Tabela do Analysis Services (Formato Tabular)



Se trabalha com o modelo Tabular quando os dados são mais simples.

Se trabalha como modelo do Cubo multidimensional quando os dados são mais complexos e as suas ligações também.

Assistir novamente a aula 121.

Conectando o Power BI ao Analysis Services

Se conectássemos/obtivéssemos os dados através da forma abaixo, a conexão seria na Camada Física do Banco de Dados (DW).

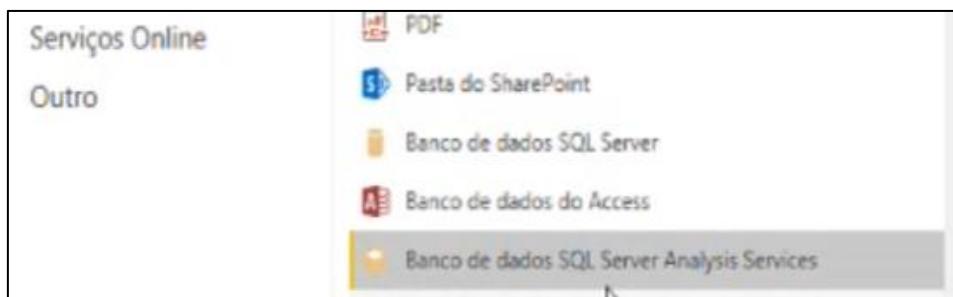
The image consists of three screenshots illustrating the connection process:

- Obter Dados (Get Data) screen:** Shows the "Banco de Dados" (Data Source) section. The "Banco de dados SQL Server" (SQL Server Database) option is selected, highlighted by a yellow box. A tooltip for "Importe dados de um banco de dados do SQL Server." (Import data from a SQL Server database) is visible.
- Banco de dados SQL Server screen:** Shows connection parameters:
 - Servidor (Server): A dropdown menu.
 - Banco de Dados (opcional) (Optional Database): A dropdown menu.
 - Modo de Conectividade de Dados (Data Connectivity Mode): Radio buttons for "Importar" (Import) (selected) and "DirectQuery".
 - A "Mais" (More) button is visible.
 - A "Opções avançadas" (Advanced Options) link is at the bottom.
- Navegador (Navigator) screen:** Shows the data model structure:
 - Opções de Exibição (Display Options): A dropdown menu.
 - Tree view of tables:
 - Banco de dados SQL local [5]
 - AdventureWorks2017
 - AdventureWorksDW2017 [39]
 - vAssocSeqLineItems
 - vAssocSeqOrders
 - vDMPrep
 - vTargetMail
 - vTimeSeries
 - AdventureWorksDWBuildVersion
 - DatabaseLog
 - DimAccount

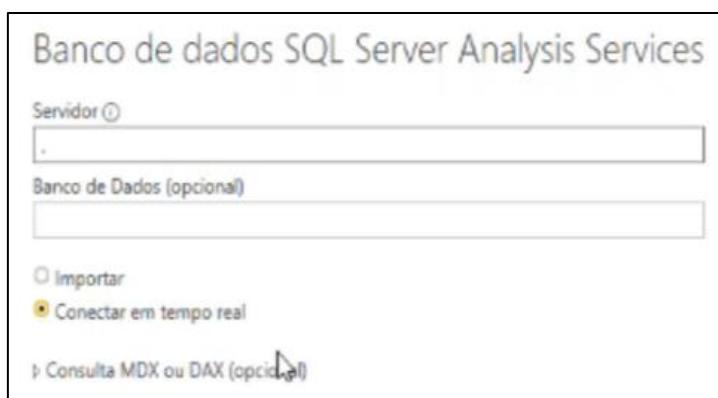
Essa conexão seria em TODAS as Tabelas do DW Adventure Works 2017. Mas, queremos a conexão com o CUBO.

Conexão na Camada Lógica (Cubo) – Analysis Services

Queremos conectar na Camada Lógica. Para estabelecer essa conexão, vá em “Obter Dados > Tudo > Banco de Dados SQL Server Analysis Service > Conectar”.



Será solicitado a conexão.



Nesse caso, estabeleceremos a conexão local (poderia ser 127.0.1) ou um ponto(.). Mas em caso de um CUBO, é necessário que o DBA informe o Servidor.

Atenção!

- **Importar:** Vai “trazer/importar” todas as informações sobre esse BD e alojar em memória. No Disco/HD. Não é uma conexão constante.
- **Conectar em Tempo Real:** É a conexão em tempo real, é em memória, é query diretamente na memória e não no HD.

Outros tópicos

Limpar o Dashboard de Campos Não Utilizados

É muito importante manter o dashboard bem limpo e organizado.

Há alguns campos que não são utilizados nesse dash, porém, não podemos deletá-los.

Para isso, vá na Visualizações de Dados.

IDCliente	Data	Operação	Cidade	IDPesquisa
569	domingo, 1 de abril de 2018	Saque Poupança	Porto Alegre	13231582
786	domingo, 1 de abril de 2018	Saque Poupança	Porto Alegre	13250799

E clique na coluna com o botão direito e “Oculte da Visualização do Relatório”.

IDCliente	Data	Operação	Cidade	IDPesquisa	Tempo de Espera
569	domingo, 1 de abril de 2018	Saque Poupança	Porto Alegre	13231582	32
786	domingo, 1 de abril de 2018	Saque Poupança	Porto Alegre	13250799	48

Entenda que essas colunas NÃO foram deletadas e suas cargas CONTINUAM CARREGADAS. A diferença será notada aqui, nos Campos. Caso necessário, ative a visualização novamente.

The screenshot shows the 'Campos' (Fields) visualization in Power BI. On the left, there's a tree view of fields under 'Cliente'. Under 'Cliente', 'Cidade' is checked. Under 'Data', 'Operação', 'Região', and 'Tempo de E...' are listed. Under 'Pesquisa', 'Classificação' and 'Classificaca...' are listed. The checkbox for 'Total de Cli...' under 'Data' is checked.

Informações sobre gráficos

Gráfico de Funil: Utilizar para método comparativo de porcentagem.



Gráfico de Rosca/ Pizza: Utilizar quando no máximo 5 unidades.

Desativar relacionamento e tipagem automática

Menu opções > Opções > Arquivo Atual > Carregamento de Dados do Arquivo Atual > Remover o “Detectar automaticamente novosrelacionamentos”.

Detecção de tipo
<input checked="" type="checkbox"/> Detectar tipos de coluna e cabeçalhos de fontes não estruturadas
Relacionamentos
<input checked="" type="checkbox"/> Importar relacionamentos de fontes de dados no primeiro carregamento ⓘ
<input type="checkbox"/> Atualize ou exclua relacionamentos ao atualizar os dados ⓘ
<input checked="" type="checkbox"/> Detectar automaticamente novos relacionamentos depois que os dados são carregados ⓘ

Linguagem M x Linguagem DAX

Sempre que realizamos esse tipo de alteração, o Query é alterado no seu *script*, que é feito na linguagem de programação M. Ou seja, toda e qualquer alteração realizada no **Power Query** é alterado diretamente o script na linguagem M.

Enquanto o **Power Pivot** sempre que alterado, será alterado conforme a linguagem DAX.

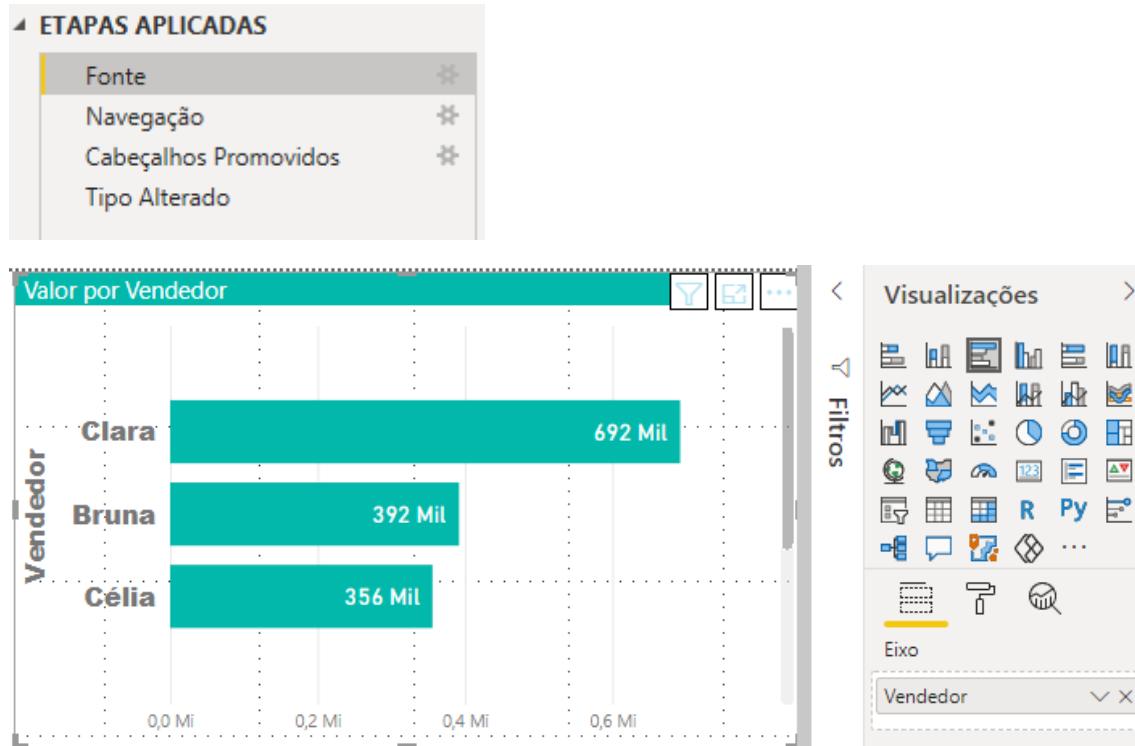
Coeficiente de Correlação de Pearson

Coeficiente de Correlação de Spearman

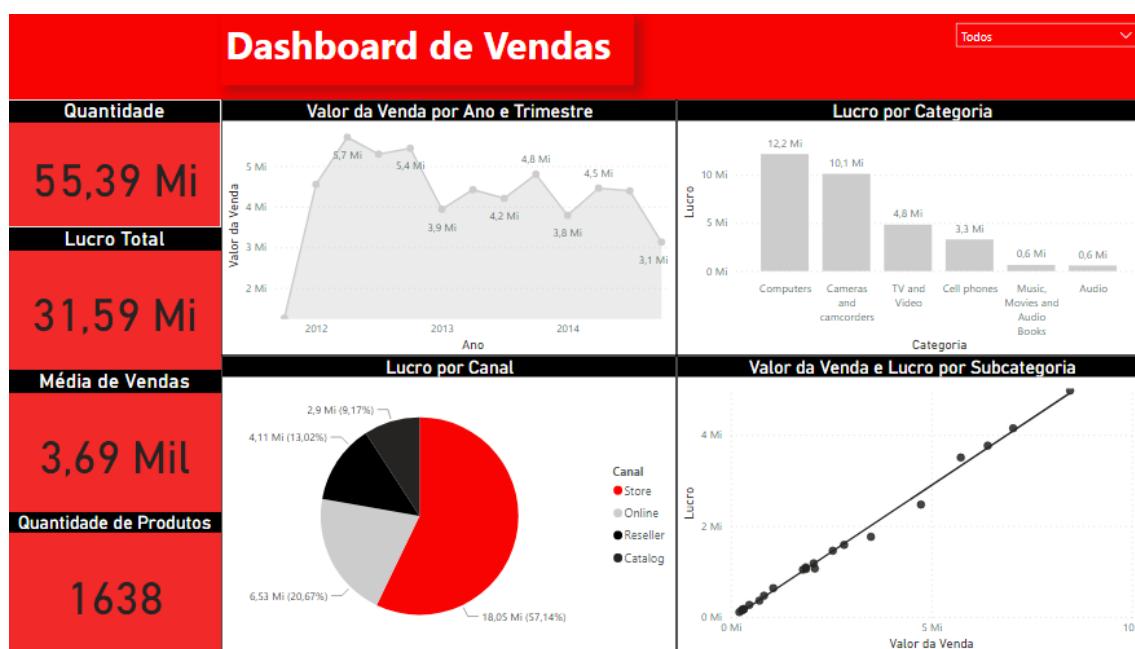
Etapas de Cada Exercício

Esse capítulo conterá todas as fotos do final de cada exercício resolvido para verificação após refazer o exercício. Verifique quais etapas foram e quais foram puladas.

Exercício 01 - Vendas Intro



Exercício 02 - Dashboard de Vendas



Exercício 03 - Tratamento CSV

▲ ETAPAS APLICADAS

- Fonte *
- Cabeçalhos Promovidos *
- Tipo Alterado
- Colunas Renomeadas
- Tipo Alterado1
- Linhas em Branco Removidas
- Valor Substituído *
- Coluna Condisional Adicionada *
- Tipo Alterado2
- Valor Substituído1 *
- Colunas Removidas
- Coluna Mesclada Inserida *
- Coluna Condisional Adiciona... *
- Colunas Removidas1
- Colunas Reordenadas
- Colunas Renomeadas1
- Tipo Alterado3

✖ Texto Aparado

Exercício 04 - Receitas Qualificado

▲ ETAPAS APLICADAS

- Fonte *
- Cabeçalhos Promovidos *
- Tipo Alterado
- Linhas em Branco Removidas
- Texto Extraído Antes do Deli... *
- Texto Extraído Antes do Deli... *
- Texto Aparado

✖ Colunas Renomeadas

Exercício 05 – FixedWidthExample

▲ ETAPAS APLICADAS

- Fonte *
- Dividir Coluna pelas Posições *
- Tipo Alterado
- Linhas Filtradas *
- Linhas Filtradas1 *
- Linhas Filtradas2 *

✖ Texto Aparado

Exercício 07- Múltiplos arquivos

▲ ETAPAS APLICADAS

- Fonte
- Outras Colunas Removidas
- Arquivos Ocultos Filtrados1
- Invocar Função Personalizada1
- Outras Colunas Removidas1
- Coluna de Tabela Expandida1
- Tipo Alterado
- ✗ Tipo Alterado com Localidade

Exercício 08 - Movimentos com Colunas Mescladas e Diferentes

▲ ETAPAS APLICADAS

- Fonte
- Linhas Filtradas
- Outras Colunas Removidas
- Arquivos Ocultos Filtrados1
- Invocar Função Personalizada1
- Outras Colunas Removidas1
- ✗ Coluna de Tabela Expandida1

▲ ETAPAS APLICADAS

- Fonte
- Navegação
- Cabeçalhos Promovidos
- Linhas Inferiores Removidas
- Colunas Removidas
- Cabeçalhos Promovidos1
- Preenchido Abaixo
- Colunas Não Dinâmicas
- Coluna em pivô
- Tipo Alterado
- Ano Inserido
- Nome do Mês Inserido
- Colunas Removidas1
- ✗ Tipo Alterado1

Transformar o Arquiv...

Exercício 09 - Movimentos Mesma Estrutura]

The image shows two separate 'ETAPAS APLICADAS' (Applied Steps) panes from a Power BI report. Both panes list the same sequence of steps:

- Fonte
- Outras Colunas Removidas
- Arquivos Ocultos Filtrados1
- Invoker Função Personalizada1
- Outras Colunas Removidas1
- Coluna de Tabela Expandida1
- Tipo Alterado

In both panes, the last step, 'Tipo Alterado', is highlighted with a yellow selection bar. Below the first pane, there is a small icon labeled 'Transformar o Arquiv...'.

Exercício 10 - Movimentos Estrutura Diferentes

The image shows two separate 'ETAPAS APLICADAS' (Applied Steps) panes from a Power BI report. The left pane lists the following steps:

- Fonte
- Outras Colunas Removidas
- Arquivos Ocultos Filtrados1
- Invoker Função Personalizada1
- Outras Colunas Removidas1
- Coluna de Tabela Expandida1

The right pane lists the following steps:

- Fonte
- Outras Colunas Removidas
- Arquivos Ocultos Filtrados1
- Invoker Função Personalizada1
- Outras Colunas Removidas1
- Coluna de Tabela Expandida1

In both panes, the last step, 'Coluna de Tabela Expandida1', is highlighted with a yellow selection bar. Below the first pane, there is a small icon labeled 'Transformar o Arquiv...'.

Exercício 11 - Dados web em Tabela

Exercício 13 - Transposição e Mesclagem

Exercício 14 - Trabalhando com Abas

Exercício 15 - Path

