

INTRODUÇÃO

O que você precisa saber para aproveitar esse curso

Para que você aproveite esse curso ao máximo, vamos verificar, nessa aula, alguns requisitos e recomendações. O primeiro deles é que você já tenha um bom conhecimento no Power BI e entenda a sua lógica e organização, afinal esse é um curso avançado. Caso você ainda não tenha trabalhado com Power BI, recomendo que assista o curso "Descubra o Power BI" no LinkedIn Learning. Não é obrigatório, mas recomendo que tenha conhecimento no Excel e um bom entendimento de sua lógica e suas terminologias. Também preciso que já tenha o Power BI instalado. Caso não saiba fazer isso, assista o curso "Descubra o Power BI". Nós vamos desenvolver juntos um projeto nesse curso e, para melhor aproveitamento, recomendo que tenha, no mínimo, 8 GB de memória RAM. Você pode consultar os outros requisitos do Power BI em sua documentação, nesse link. Agora que você já sabe tudo o que precisa para aproveitar o máximo desse curso, vamos começar.

1. Conceitos e objetivos do Business Intelligence

Conceito e importância do business intelligence

O Power BI é um software da Microsoft da categoria Business Intelligence. E, antes de começar a utilizá-lo, é importante sabermos o que é esse termo. Podemos traduzir Business Intelligence para o português como "Inteligência de Negócios", mas, no Brasil, é mais comum utilizar o termo em inglês, Business intelligence, que também é conhecida em sua forma abreviada como "BI". E eu devo me referenciar a essa sigla muitas vezes nesse curso. Seu principal objetivo é dar suporte à tomada de decisão por meio de informações estruturadas, atualizadas e disponíveis. É importante destacar que esse termo é utilizado de duas formas diferentes: como um processo e como um software. Ambos andam juntos, porque o software vai executar o processo. Vamos conhecer um pouco mais sobre ele. Suas etapas são muito bem definidas. Depois de estabelecidas quais informações a empresa ou o projeto precisam, o primeiro passo é obter e transformar os dados. Posteriormente, modelá-los, criando relacionamentos e cálculos. Após essa etapa, visualizá-los por meio de gráficos e tabelas. E o projeto, posteriormente, é compartilhado com as pessoas interessadas a consumirem as informações estruturadas. E tudo isso com o objetivo de melhorar a tomada de decisão. Esse foi um resumo do que é o processo de Business Intelligence. E vamos fazer, na prática, cada uma dessas etapas ao longo do projeto que desenvolveremos.

Apresentação dos arquivos e projeto do curso

Nessa aula, você vai conhecer o projeto que desenvolveremos nesse curso e os arquivos adicionais para que aprenda as variações de tratamento de dados. Começaremos com os arquivos de Excel em TXT, PDF, fontes da web e do SQL Server. Vamos explorar sobre as conexões a esses diferentes tipos de fontes de dados, que é essencial para todo início de projeto, e também seu tratamento e limpeza. A partir das aulas de relacionamento entre tabelas, utilizaremos um banco de dados no Access para criar os cálculos, utilizando a linguagem DAX e a visualização dos resultados com gráficos. Os arquivos para praticar podem ser obtidos na página desse curso. Eles estarão divididos com os nomes dos capítulos e, dentro de cada uma dessas pastas, teremos também uma numeração de acordo com a sua aula. Então, por exemplo, esse arquivo daqui está no capítulo três, referente à aula dois. Como eu já comentei, a partir do capítulo sobre relacionamento de dados, utilizaremos um mesmo arquivo, baseado num banco de dados do Access. Vou clicar duas vezes nele para visualizarmos as suas tabelas. Esse banco é de uma empresa fictícia, criada pela Microsoft, chamada AdventureWorks, e ela é amplamente utilizada para praticar e estudar softwares de BI e banco de dados. Se eu clicar nessa tabela de "Produtos Categorias", conseguimos ver que essa empresa vende bicicletas, componentes, vestuário e acessórios, portanto ela é uma empresa do ramo de aventura. Ela também vende para diversos países, em vários continentes. Agora, que você já sabe os arquivos de exercícios, vamos começar a praticar.

2. Conhecendo o ecossistema do Power BI

Compreenda as diferentes plataformas da Microsoft para análise de dados

Antes de começarmos os tratamentos, relacionamentos, cálculos e visualizações, vamos recapitular sobre as plataformas do Power BI. Se você entrar no site da Microsoft, verá que existem diversas ferramentas que

compõem o ambiente do Power BI. A primeira delas, e é que utilizaremos na maior parte do desenvolvimento desse curso, é o Power BI Desktop. Somente nele, temos todas as capacidades de tratamento, modelagem e cálculos no DAX. Depois de pronto o projeto no Power BI Desktop, faremos uma publicação para o Power BI Serviço, que pode ser acessado pelo site "app.powerbi.com", de onde compartilharemos, configurações de atualizações e gerenciamento de segurança. Os relatórios publicados no Power BI Serviço também podem ser acessados por smartphones com sistema operacional Android, iOS e Windows. Também temos outros serviços do Power BI, como o Embedded, que permite embarcar, ou seja, inserir relatórios dentro do seu site próprio, com o seu sistema de login e senha. Essa ferramenta exige a utilização de um produto da Microsoft chamado Azure e criação de um projeto no Visual Studio, que é um ambiente de desenvolvimento. Nesse curso, não falaremos sobre Power BI Embedded por abranger outros conhecimentos além do Power BI. Caso a sua empresa seja restritiva quanto a carregar informações na nuvem, que é o que faremos quando utilizarmos o Power BI Serviço, você pode também utilizar o Power BI Reporting Service, que utiliza uma instância do SQL Server e instala o ambiente do Power BI localmente na sua empresa. O ponto negativo é que você não terá acesso externo pela internet a essas informações, caso elas não sejam replicadas para fora de sua organização. Vamos recapitular também um pouco sobre o desenvolvimento no Power BI e o seu fluxo. Todo projeto começa com a obtenção de dados. Fazemos isso no Power BI Desktop. Depois dessa etapa, vamos para a modelagem, criando relacionamentos e cálculos na linguagem DAX. Posteriormente, criamos visualizações com esses resultados por meio de gráficos e tabelas. E, por fim, compartilhamos o projeto criado no Power BI Serviço, ou seja, no site do Power BI. Todas as informações publicadas e compartilhadas ficam protegidas por uma camada de segurança, que é o seu login e senha individual. Portanto, para resumir esse fluxo de desenvolvimento, primeiro, fazemos um projeto no Power BI Desktop e, depois, fazemos a sua publicação no Power BI Serviço, que é responsável por compartilhar essas informações, deixá-las disponíveis na web para acesso, por meio da camada de segurança de login e senha, e permite a interação direta dos dados conectados à fonte de dados. Também é no Power BI Serviço que fazemos a configuração da atualização dos dados e que gerenciamos quais serão os usuários que acessarão as informações publicadas. Falando, agora, um pouco sobre licenças, existem três delas para o Power BI: a gratuita, a para o e a premium. Esse terceiro tipo de licença, a premium, é direcionada para grandes empresas, acima de 499 funcionários, com acesso a relatórios ou com necessidade de processamento de dados altíssima. Seu preço inicia em 4.990,00 dólares por mês. Mas fique tranquilo! Na maior parte das empresas, o Power BI Pro é suficiente. Vamos detalhar um pouco mais essas duas primeiras licenças, que são as mais utilizadas. No Power BI Gratuito, você consegue ter acesso a todas as funcionalidades de desenvolvimento do Power BI Desktop. Portanto, caso não tenha uma licença Pro para esse curso, não se preocupe, porque você conseguirá fazer todos os desenvolvimentos, ou seja, os tratamentos, relacionamentos, cálculos e visualizações no Power BI Desktop. Com ela, conseguimos ter um limite de dados de 10 GB, 10.000 linhas por hora, em streaming, e acesso aos dados pelo Power BI Serviço, e o seu preço é gratuito. No Power BI Pro, temos outras vantagens. Além dessas iniciais, também podemos fazer oito atualizações diárias dos dados, compartilhar relatórios de forma privada, criar espaços de trabalho e assinar relatórios, e algumas outras diferenças menos relevantes. E o seu preço é 9,99 dólares, por usuário, por mês. O objetivo dessa aula foi apenas recapitular alguns conceitos importantíssimos sobre as plataformas do Power BI e seu licenciamento. Agora, podemos prosseguir com o desenvolvimento do projeto nas próximas aulas.

Conheça outros softwares de Business Intelligence

Quando trabalhamos com software, seja ele de qualquer categoria, é importante conhecer quais são os principais para comparação de funcionalidades, preço e soluções. Esse é um curso de Power BI, mas é muito bom conhecer outros para se familiarizar com o mercado e ter alternativas de funcionalidades. Nessa aula, utilizaremos como base o quadrante mágico de plataformas de Business Intelligence da Gartner, que é um instituto independente de pesquisa do ramo de tecnologia. Esse quadrante é mencionado em uma de suas pesquisas do ano de 2019 e destaca softwares de Business Intelligence de acordo com sua liderança no mercado e funcionalidades. Analisando a posição do Power BI, está entre um dos líderes. Juntamente temos outros dois que quero destacar, o Tableau e o Qlik. Apesar de trabalhar mais com Power BI e preferir suas funcionalidades e preço, também são recomendados esses dois. Não vou detalhar funcionalidades, porque estamos em um curso de Power BI, mas é importante que conheça que existem outros no mercado que podem também ser interessantes para sua empresa ou projeto. Sempre acompanhe

em pesquisas, como a da Gartner, para saber se está com uma das melhores soluções em softwares, sejam eles de bancos de dados, de BI, CRM ou qualquer outro sistema informacional.

3. Importação de dados

Conheça os diferentes tipos de dados

Eu estou com um arquivo de Power BI aberto, em branco, e vamos recapitular, nessa aula, como conectamos à fonte de dados e quais são as possibilidades dentro do Power BI. Aqui, na página inicial, nesse grupo de dados externos, podemos clicar em cima desse ícone, do "Obter Dados", para abrir a janela "Obter Dados". Aqui, temos algumas possibilidades padronizadas. Em "Arquivo", faremos uma importação de Excel, texto CSV, pasta, PDF. E, em "Banco de Dados", faremos uma conexão ao SQL Server. Também é possível utilizar outros bancos, como, por exemplo, Oracle, IBM, MySQL, Postgre, SAP, Amazon, BigQuery, entre outros. Algo interessante e muito bom do Power BI é que você não fica restrito à plataforma e produtos da Microsoft. Então, além do SQL Server, também temos alguns de seus principais concorrentes, como Oracle, aqui, disponível para ser conectado, e isso é muito bom. Além de bancos de dados, também podemos utilizar o Azure, que é um serviço da Microsoft na nuvem, e serviços online, como, por exemplo, o SharePoint, Exchange, Dynamics, Salesforce, Analytics e, até mesmo, o Facebook, para a análise de mídias sociais. É muito importante que você observe que algumas funcionalidades terão essa palavra "beta" na frente - temos aqui, e essa listagem tem muitas outras. O Power BI está em constante desenvolvimento, então, se você ver uma funcionalidade como essa, tenha cautela ao utilizar em projetos oficiais. Teste bastante e observe quais são os seus retornos para que não tenham problemas, já que é uma funcionalidade que está em teste, ou seja, é beta. Caso nenhuma dessas conexões estejam disponíveis ou tenha o produto que você utilize, também é possível vir aqui na categoria de outros e utilizar o ODBC, que é um conector universal de dados, sendo necessário ter apenas o caminho, o login e a senha. Agora que recapitulamos esse conceito importantíssimo sobre importação de dados no Power BI, vamos começar a realizar as primeiras importações.

Como importar dados do Excel e funções da linguagem M

Vamos começar essa aula fazendo uma conexão a um dos arquivos mais comuns em empresas: o Excel. Eu vou aproveitar essa conexão para analisarmos a barra de fórmulas no Power Query. Eu estou com um arquivo de Power BI em branco aberto e aqui na página inicial vou clicar nessa seta para baixo do obter dados. Teremos uma listagem dos arquivos mais comuns de importação no Power BI, entre eles o Excel. Vou clicar nessa opção, e o Power BI vai me pedir para abrir esse arquivo e eu vou escolher o "03_02_clientes", que está na pasta do capítulo 3. Vou clicar em abrir e a janela de navegação será aberta. E dois objetos são exibidos: clientes e dados clientes. Quando temos esse ícone, significa que essa é uma planilha; e esse de cima com um cabeçalho azul, significa que é uma tabela. No Excel toda tabela está dentro de uma planilha e podemos ver isso pelos dados. Aqui temos o nome do cliente Luís, Ágatha e Sophia. Na planilha clientes também temos Luís, Ágatha e Sofia, portanto se eu fizer a importação de tabela e planilha estaremos importando esses dados de forma duplicada, então eu tenho que escolher um ou outro. A minha decisão é pela tabela porque tabelas são intervalos delimitados com cabeçalhos definidos e tipos de dados também. Já a planilha não. Então com essa vantagem, aqui embaixo eu tenho uma escolha: clicar em carregar ou transformar dados; se eu clico em carregar, esses dados virão diretamente para os campos e o Power Query não será aberto automaticamente. Mas eu quero que ele seja porque eu quero transformar esses dados. Então eu vou clicar em transformar dados. No Power Query nós temos a consulta realizada com os dados importados para o Power BI, apenas recapitulando, nós estamos no Power Query que está acima do Power BI desktop. Então no Power Query fazemos os tratamentos dos dados. Eu vou maximizar novamente essa janela. Aqui na direita temos as etapas aplicadas de transformação a essa fonte de dados. E é muito importante observar que já temos três: fonte navegação e tipo alterado, que foram automaticamente criadas pelo Power BI. Os dados já estão importados mas estamos em um curso avançado e por isso não vamos simplesmente clicar em fechar é aplicar e utilizar esses dados, vamos observar antes a sua barra de fórmulas avançadas. Para exibi-la eu vou clicar na guia exibição e escolher barra de fórmulas. Ao clicar aqui teremos qual função foi utilizada para executar essa etapa daquele tipo alterado, nesse caso foi a `table.transform.ColumnTypes`. Se eu clico em fonte veremos uma outra função: `Excel.workbook`, essa função é exatamente a mesma coisa que conectar a um arquivo de Excel clicando aqui nesse botão. Na verdade dentro desse botão nós já temos essa função estabelecida com todas

as configurações para importar os dados. Então todas as vezes que clicamos em SQL; análises; texto; web, entre outros conectores, essa função vai modificar de acordo com aquele que escolhemos. Incentivo que você sempre deixa sua barra de fórmulas exibida, principalmente porque você vai aprender essas fusões e vai poder utilizá-las depois em necessidades mais avançadas, por exemplo, vamos utilizar essa função daqui em uma aula futura para importar arquivos de Excel da pasta, economizando muito tempo investido ao invés de importá-los um a um manualmente, vamos fazer isso automaticamente por meio desta função. As funções do Power Query também chamadas de linguagem M podem ser consultadas no site da Microsoft. Aqui nós temos uma documentação da referência de todas as funções, então se eu clico em Power Query M Functions, temos todas as suas categorias desde binários, de combinação, de data, de tempo, duração, erro e outras funções que podem ser utilizadas caso você queira aprender mais sobre a linguagem M você pode abrir cada uma dessas categorias e começar a estudar cada uma dessas funções. Vamos explorar muitas delas aqui dentro desse curso, mas são mais de 600 funções e recomendo que você dê uma lida em todas elas para que fique ciente sobre as funcionalidades e possibilidades dentro do Power Query. No arquivo que estávamos preparando, eu vou fechar e aplicar aqui no Power Query para fazer com que todos esses dados sejam enviados lá para o Power BI. Ao clicar aqui teremos a importação e o carregamento dos dados, que serão exibidos aqui nos campos. Podemos fechar esse arquivo, salvando antes clicando nesse disquete e inserindo o nome para ele, que será importação Excel. Vou pressionar salvar e nas próximas aulas exploraremos outros tipos de conexão sempre observando qual é a função utilizada.

Como importar dados de CSV e TXT

Na aula anterior fizemos a importação de um arquivo de Excel para explorar e introduzir o assunto sobre funções do Power query, também conhecidas como linguagem M. Também habilitamos a barra de fórmulas avançada. Nessa aula vamos continuar essa exploração. Vamos conectar a um arquivo de texto. Então eu estou com o Power BI em branco aberto e vou clicar nessa seta para baixo no obter dados e selecionar "texto barra CSV". Depois vou selecionar "03_03_taxa Selic.TXT", que é um arquivo que está na pasta de exercícios do capítulo três. Vou clicar em abrir. E nessa janela temos uma pré visualização dos dados e consigo ver algumas coisas erradas. Dados diários está com um caractere estranho, operações também e isso por causa da origem do arquivo. A 1252 não identifica corretamente os assentos. O Power BI tenta fazer essas alterações automáticas, mas às vezes nem sempre dá certo. Então vamos selecionar a origem do arquivo. E lá embaixo, na "65.001 Unicode utf 8" vamos selecioná-lo. Agora os assentos estão corretos, portanto, na origem do arquivo quando os seus dados tiverem assento sempre utilize a 65.001. A próxima configuração é o delimitador, ou seja, ponto e vírgula, a vírgula, sinal de igualdade, entre outros. Power BI identificou automaticamente o ponto e vírgula. É sempre bom verificarmos se realmente é esse. Portanto eu vou abrir esse arquivo na pasta e verificar. É ponto e virgula, então o Power BI identificou corretamente. Vou fechar; fechar a pasta. E agora essa última detecção de dados é uma tentativa do Power BI de alterar automaticamente esse tipo de dado. Nesse caso temos datas misturadas com texto, temos números também com textos e assim por diante. Portanto, essa detecção de tipos de dados não será correta porque identificará todas as colunas como texto. Não tem problema porque essa é uma parte do tratamento de dados, precisamos eliminar essa linha, promover os cabeçalhos e posteriormente alterar os tipos de dados. Por enquanto vamos apenas conectar a esse arquivo.

Como importar dados da pasta

Na pasta de exercícios desse capítulo, nós temos uma pasta chamada "03_04_Pedidos", que precisa ser importada para o Power BI. Nós temos cinco arquivos de Excel dentro dela, que contêm o mesmo formato, ou seja, as mesmas colunas e cabeçalhos. Vamos para o Power BI para importá-los. Eu estou com um arquivo em branco e, para importar da pasta, vou clicar na seta para baixo e, posteriormente, em "Mais...". Na janela de "Obter Dados", vou escolher "Arquivo" e selecionar "Pasta". Posteriormente, clicar em "Conectar". Observe que, apesar de serem arquivos de Excel, não vou escolher arquivo de Excel, porque, na verdade, vamos importar e fazer uma conexão da pasta e, posteriormente, juntá-los nessa conexão. Então vou voltar aqui e clicar em "Conectar". No caminho da pasta, vou procurar por essa pasta, que está em "Arquivos de Exercícios", "Capítulo 3", "03_04_Pedidos", "OK", "OK" novamente e o Power BI me mostra algumas informações desses arquivos no Windows, fala que o conteúdo dele é binário, a extensão é XLSX, a última data de acesso de modificação de criação. Eu poderia clicar em "Combinar" e juntá-los todos rapidamente, mas estamos fazendo um curso avançado, então eu vou te ensinar a utilizar uma função para

fazer a importação desses arquivos de uma forma muito mais direta e sem utilizar outros recursos do Power Query que prejudicariam a performance dessa conexão. Portanto, vou clicar em "Transformar Dados". O primeiro fato que você precisa entender é que todo arquivo importado para Power BI é binário inicialmente, antes de ser convertido por sua função específica. As funções são: "Excel.workbook", "CSV.documents", "pdf.tables", entre outras, e, nesse caso, precisamos converter esses valores para valores que são legíveis. Para isso, eu vou clicar em cima da coluna "Content" com o botão direito e escolher "Remover Outras Colunas", porque essas não me importam, eu não quero saber qual que é a data de acesso, data de modificação, qual que é o caminho desse arquivo; não preciso, então vou clicar com o botão direito na "Content" e, posteriormente, escolher "Remover Outras Colunas". Vou ficar apenas com essa no meu modelo. Então, "Remover Outras Colunas". Estou apenas com essa. Agora, eu vou adicionar uma nova coluna aqui do lado que interprete esses dados binários de acordo com o que eu quero, que é traduzi-los de uma forma legível. Então, aqui na guia "Adicionar Coluna", vou escolher "Coluna Personalizada". Nessa coluna personalizada, ela se chamará "Personalizar" mesmo e, aqui embaixo, escreveremos uma função que traduzirá o binário, que será a "Excel.workbook". À medida que eu vou escrevendo, o Power Query já indica essa função aqui e eu posso simplesmente clicar em cima dela ou pressionar "tab" no teclado para selecioná-la automaticamente e autopreencher. Vou abrir um parênteses dessa função e, posteriormente, informar o seu primeiro argumento, que é um workbook. E ele precisa estar no formato binário, e é exatamente isso que temos nessa coluna. Esse daqui é um workbook, ou seja, uma pasta de trabalho, e está no formato binário. Então, eu vou colocar essa coluna, que se chama "Content", dentro desse parênteses do argumento da "Excel.workbook". Então vou clicar duas vezes em "Content", as colunas dentro do Power Query são referenciadas em colchetes; então vou fechar agora a função "Excel.workbook", porque toda função tem um parênteses que abre e que fecha; e, posteriormente, vou clicar em "OK". Teremos, numa segunda coluna, um objeto de table, ou seja, uma tabela. Então, em cada uma dessas linhas, nós temos tabelas. Eu não preciso mais desse conteúdo aqui do lado, essas informações, elas não são úteis mais para mim, elas foram aqui para fazer essa conversão, então eu posso deletá-las do meu modelo. Vou clicar com o botão direito e escolher "Remover" essa coluna selecionada. Agora estamos apenas com a table. Não tem problema realizar essa remoção, porque ela ainda permanece nas outras etapas aqui. Então, na "Personalizar Adicionada", nós ainda temos a "Content", que foi utilizada para criar essa table. Portanto, nessa última etapa, nós estamos com essa table e, se você não entendeu direito esse processo que eu expliquei das etapas aplicadas, recomendo que reveja o curso "Descubro Power BI", disponível aqui no LinkedIn Learning. Para prosseguir, nessa table, precisamos expandi-la, ou seja, eu quero ver essas informações aqui embaixo diretamente dentro dos meus dados. Então vou clicar nessas duas setas, retirar esse "use o nome original da coluna como prefixo"; do contrário, ele usará o nome da coluna, que é "Personalizar", como prefixo para a criação de todas essas outras colunas, então ficaria "Personalizar+Name", "Personalizar+Data", e assim por diante. Não quero isso, então vou retirar, pressionar "OK" e agora temos o acesso às planilhas e tabelas dentro desses arquivos. Como você aprendeu na aula sobre importação do Excel, se eu importar planilha e tabela, estarei fazendo uma importação duplicada, porque, dentro da tabela, eu tenho também o conteúdo da planilha, e vice versa. Então, tenho que escolher apenas um. E quando temos esses dois elementos, sempre dou preferência para a tabela, pelos motivos já explicados na aula da importação do Excel. Então, aqui no "Kind", que é "tipo" em inglês, vou remover a seleção do "Sheet", que é "planilha" em inglês, e vou deixar a seleção do "Table", que é "tabela". Vou clicar em "OK" e temos um filtro que foi aplicado removendo a "Sheet" e permanecendo apenas com a "Table". Agora eu vou clicar em "Data", nessa coluna, que na verdade não é "data" em português, é "data" em inglês, que significa "dados", e somente essas informações que são interessantes para mim; eu não preciso saber qual é o nome dessa tabela ou o seu tipo ou se ela está oculta ou não. Então eu vou clicar em cima da coluna "Data" com o botão direito e "Remover Outras Colunas". Quero ficar apenas com essas. É uma prerrogativa, em todo o projeto de dados, que você sempre esteja apenas com as colunas que você quer trabalhar, então aquelas que são inúteis devem ser removidas do seu modelo. Agora, vou clicar nessas duas setas para expandir as informações das tables. Se eu clico nessa parte cinza da nossa célula, consigo ver aqui em baixo que temos os dados, então eu tenho o pedido, o código de cliente, o faturamento e algumas datas. Aqui embaixo, o mesmo processo. E assim por diante. Cada uma dessas tables são os arquivos do Excel. Havia cinco arquivos na pasta que importamos e agora temos cinco tables porque navegamos para dentro desses arquivos e traduzimos as suas informações binárias para legíveis. Então, agora, quando eu clicar nessas duas setas para expandir essas informações, teremos a junção de todos esses valores em uma

única consulta, ou seja, juntamos todos os cinco arquivos. Então clico nessas duas setas; posteriormente, verifico se essa opção está desmarcada; quero expandir todas essas colunas; clico em "OK" e temos todas as importações, com os arquivos juntos. Portanto, nessa aula, você aprendeu a utilizar uma função chamada "Excel.workbook", que antes foi utilizada apenas para importar um arquivo, mas agora importamos cinco de uma vez, fazendo uma conexão diretamente com a pasta do Windows, e não apenas com um arquivo de Excel. Caso você estivesse arquivos CSV, TXT ou outro tipo de fonte de dados, basta alterar aquela função "Excel.workbook" para a função apropriada. Na aula anterior, vimos que a função para importação de texto "ssv.document", então bastaria trocar essa função e inserir seus argumentos da forma correta.

Como importar dados do SQL Server

SQL Server é um dos bancos de dados mais utilizados no mundo, não é a toa que ele está na listagem das conexões mais comuns do Power BI. Banco de dados armazenam dados. Imagine um sistema de vendas que tem uma interface e alguém preenche algumas informações nos campos. Elas precisam ser armazenadas em algum local, normalmente é em um banco de dados, que pode ser um SQL Server. Então vamos aprender a conectar em um. Eu vou escolher SQL Server. E nessa janela, eu preciso informar um caminho, pode ser um IP, ou então um nome. Nesse caso, o meu banco de dados está em um caminho que se chama "marte.database.windows.net". Ao clicar em "OK", ele pedirá as informações de credenciais. Eu posso escolher o Windows, caso meu banco de dados aceite o meu usuário do Windows atual para o login, mas esse não é o caso; posso escolher "banco de dados", que é o próprio login e senha dele; ou uma conta da Microsoft. Nesse caso, são as credenciais do banco de dados e eu vou colocar as minhas aqui, com a senha e pressionar "Conectar". Caso a conexão seja realizada com sucesso, ele vai me mostrar quais são os bancos que tem dentro desse servidor, e, nesse caso, é apenas um, que se chama "sistemavendas". E quando eu abro nessa setinha, consigo ver todas as suas tabelas. Para importar, basta selecionar essas tabelas e, posteriormente, clicar em "Carregar" ou "Transformar Dados". Eu vou pressionar "Cancelar", apenas para te mostrar uma outra forma de fazer esse mesmo processo. Vou clicar em "obter dados", "SQL Server", vou inserir novamente o mesmo caminho e, aqui embaixo, em "Banco de Dados", vou colocar "sistemavendas", que é o nome do meu banco de dados presente nesse servidor. Em "Opções Avançadas", eu posso conectar a esse banco de dados já fazendo um filtro nele, utilizando uma instrução SQL. Eu vou colar uma rapidamente aqui, que vai fazer o seguinte: vai selecionar todos os dados dessa tabela que está presente nesse banco de dados quando a coluna "Order Date", ou seja, a data do pedido estiver entre essa data e essa data. Eu estou utilizando aqui uma linguagem SQL para fazer essa consulta, ou seja, esse filtro. Portanto, o Power BI permite que você faça um filtro antes de importar as informações. Vou clicar em "OK" e ele me mostra a tabela que eu selecionei por meio daquela linguagem SQL que eu te mostrei com a Order Date filtrada. Então, vai mostrar apenas aquelas duas datas iniciais e finais que coloquei. Posso transformar dados para abrir o Power Query e tenho a "Consulta1", com aquelas informações, conectadas em um SQL, por meio da função "SQL.database" nesse caminho, nesse banco de dados e com essa Query, ou seja, com essa consulta, que foi aquele Select + tudo do banco de dados atual que tem essa tabela quando a data do pedido, que é essa coluna, estiver entre essa data inicial e essa data final. Portanto, uma conexão a um banco de dados do SQL é bastante simples, só precisa ter as credenciais de acesso e, caso saiba SQL, fica muito mais fácil de você filtrar as informações antes de trazê-las para o Power BI.

Como importar dados de fontes da Web utilizando tabelas

Dados provenientes da internet são um dos mais utilizados em projetos, e é por isso que, nessa aula, vamos aprender a importá-los. Eu vou abrir o site da Wikipédia, onde eu tenho uma listagem de população por países. E eu quero importar essa tabela de uma forma estruturada, ou seja, eu quero que todas essas colunas venham organizadas e também atualizadas. Se esses valores mudarem, quando eu clicar em "atualizar" lá no Power BI, eu quero que eles sejam atualizados lá também. Então eu vou clicar nessa URL, pressionar CTRL+C no teclado e voltar no Power BI e agora clicar em "obter dados web". Nessa janela, eu vou colar a URL aqui com o CTRL+V no teclado e pressionar "OK". No navegador, teremos a relação de todos os objetos que tem dentro dessa página da internet. Toda página terá um "document", que é onde está armazenado o HTML, e também, se ela possuir, tabelas. Aqui em baixo temos uma, chamada "Lista", e outra chamada "Table 1". Eu quero a "Lista", porque eu vi que aqui tem as informações que eu preciso: o país, a população, a posição, a data, e assim por diante. Então eu vou selecioná-la, posteriormente clicar em "transformar dados". O Power Query vai abrir, com a lista aqui em "consultas" e as informações aqui dentro. Se eu clico

em "fonte", temos uma conexão utilizando a função "web.page", com essa URL em seu argumento. Vou voltar na "Tipo Alterado", que é a última etapa aplicada. Posso pressionar, fechar e aplicar e as informações estão prontas para serem utilizadas. Vou fazer um gráfico rápido aqui, apenas para mostrar que todas elas estão posicionados aqui dentro do Power BI e prontas para serem utilizadas. Então, aqui, eu tenho um gráfico de população por país ordenada do maior para o menor. Aqui nós temos a China, Índia, Estados Unidos, Indonésia, Brasil, e assim por diante. Poderia fazer outras configurações aqui neste gráfico, como aumentar essa letra, colocar os rótulos de dados nessas barras, mas, por enquanto, não vamos fazer esse tipo de alteração; teremos um capítulo exclusivo sobre visualizações.

Como importar dados de fontes da Web utilizando exemplos

Na aula anterior, fizemos uma importação de tabela de uma página da web, mas nem todos os sites terão tabelas. Vamos verificar o que podemos fazer nesse tipo de situação. Eu vou abrir o site do IBGE e aqui nós temos várias informações mas a que me interessa são esses preços aqui. Eu quero todos esses percentuais lá dentro do Power BI. Para isso, eu vou selecionar a URL, pressionar CTRL+C no teclado para copiar, voltar no Power BI, "Obter Dados", "Web" e colar a URL com CTRL+V. Vou pressionar "Ok" e no navegador temos apenas o "Document", que é o HTML desse site, não temos nenhuma tabela. Nesses casos, precisamos utilizar "Adicionar a Tabela Usando Exemplos", esse botão localizado aqui na parte inferior esquerda. Vou clicar. O Power BI vai abrir uma visualização do site acima e uma tabela em branco, abaixo. Então, aqui, na visualização, eu vou buscar por aquela tabela, que aparentemente é uma tabela mas não é de fato, e eu vou usar esses valores aqui dentro dessa coluna para dar exemplos ao Power BI para ele buscar essas informações automaticamente, por meio de alguma padronização na estrutura do site. Então, eu tenho que escrever: "IPCA - mensal" e "0.11%". Então aqui, na coluna 1, eu vou colocar "IPCA - mensal", e aqui do lado, "0.11%". Vou dar um nome para essa coluna de preços e aqui, na coluna 2, de percentuais. E vou dar outros exemplos; o próximo será o INPC e do lado será. 012. Quando eu dou esses dois exemplos, o Power BI já consegue identificar outros. Ou seja, esses valores têm a mesma estrutura que esses daqui de cima, lá no site. Mas ele pulou um desses valores que, nesse caso, foi o "Preços" do produtor. Então, acima dele, eu tenho o IPCA 15, que está aqui, e, abaixo, o IPCA 12 meses, que está aqui. Aqui no meio, eu preciso colocar esses valores. "IPP" e, do lado, será "-1.24", que é o que tem aqui do lado do IPP. Quando eu coloco essas informações, o Power BI já consegue identificar um padrão e puxa todos os preços. O último que eu tenho é o Sinapi, que é exatamente esse. Então, eu posso pressionar "Ok" e, agora, temos uma tabela aqui embaixo, criada por esses exemplos. Ela já está selecionada, então, vou clicar em "Transformar Dados" e eles estão disponíveis aqui, dentro de uma tabela. E essa fonte de dados utilizou uma "Web.BrowserContents", que outra função diferente da anterior que a gente utilizou para importar dados da web. Então, se eu clico na etapa "Tipo Alterado", eu tenho as informações prontas para serem utilizadas. Agora, basta ir na página inicial, "Fechar e Aplicar" e elas estarão disponíveis aqui. Para esse arquivo, vou simplesmente pressionar o "X", porque eu não vou precisar utilizá-lo futuramente.

4. Tratamento de dados importados

Apresentação da linguagem M

Como mencionei em aulas anteriores, onde observamos as funções de conexões à fontes de dados, o Power Query é um suplemento que está dentro do Power BI e possui sua própria estrutura lógica, formada pela Linguagem M, que podemos afirmar que toda conexão e transformação de dados no Power Query utiliza a Linguagem M, que é baseada em funções com argumentos. E ela tem várias categorias, com uma vasta biblioteca de funções de transformações de dados que podem ser consultados diretamente da documentação da Microsoft, mostrada na segunda aula do capítulo 3. E, a partir desse capítulo, começaremos a fazer transformações dos dados importados. E é importante ressaltar que as alterações realizadas não alteram a fonte de dados, ou seja, o arquivo original ou o banco de dados permanecerá intacto, alterações serão apenas dentro do Power BI ou, então, dentro do Power Query. Vamos verificar algumas características da Linguagem M. Ela é realizada por etapas, o editor avançado vai agrupar todas as funções das etapas - na próxima aula, eu vou mostrar o editor avançado -, a barra de fórmulas, que vimos nas aulas anteriores, mostrará apenas a função da etapa selecionada e as funções sempre serão no idioma inglês; seus argumentos também são separados pela vírgula. E a Linguagem M está presente apenas dentro do Power Query. Portanto, não confunda: Linguagem M é diferente de linguagem DAX. Temos ambas dentro do Power BI, mas a Linguagem M pode ser encontrada apenas dentro do Power Query, que deve ser utilizada

para conexão e tratamento de dados. Já a Linguagem DAX, deve ser utilizada para criar cálculos e medidas e só pode ser acessada dentro do Power BI, ou seja, não posso utilizar funções DAX dentro do Power Query. E a sintaxe das funções M é a seguinte: começamos com um símbolo de igual, posteriormente com o nome da função que, nesse caso, é "Excel.Workbook"; posteriormente, abrimos o parênteses para declarar os seus argumentos, que são separados por vírgula e nem todos eles serão obrigatórios; e finalizamos a função com parênteses final. Na próxima aula, eu vou detalhar mais sobre o editor de consultas que eu comentei nessa aula para que você entenda como que as etapas no Power Query são armazenadas.

Introdução ao editor de consultas avançado

No capítulo anterior, fizemos diversas conexões em fontes de dados e vimos que cada uma delas utiliza uma função da Linguagem M diferente para importá-las. Conseguimos ver as funções pela barra de fórmulas mas, também, existe um outro local, chamado "Editor Avançado", que armazena todas as funções utilizadas em uma consulta. Vamos verificar no Power BI onde ele está localizado. Eu estou com um arquivo onde foi realizado uma importação de vários arquivos de Excel da pasta e eu preciso acessar o Power Query para poder ver as etapas. Para isso, vou vim na "Página Inicial" e, depois, "Editar Consultas". Clico nessa opção, o Power Query foi aberto, temos a consulta aqui na esquerda e as etapas aplicadas na direita. Você já deve saber disso mas vou recapitular que: cada uma dessas etapas são transformações aplicadas nos dados, que utilizam algumas funções. Então, eu tenho etapa "Fonte", "Outras Colunas Removidas" e muitas outras aqui embaixo, que foram criadas ao clicar em "Ferramentas", aqui na faixa de opções. Então, eu tenho a guia "Transformar", com várias transformações que eu posso executar, a guia "Adicionar" e, também, algumas transformações realizadas em filtros e nesses botões. Cada uma das alterações e cliques que damos aqui dentro do Power Query geram funções automáticas que são armazenadas nessas etapas, que o próprio Power Query também faz. É ele que dá esse nome, "Outras Colunas Removidas", e é ele que escreve essa função daqui com os argumentos, de acordo com o clique que você dá na coluna ou na transformação que você está tentando executar. Então, já sabemos verificar qual é a função, utilizada aqui na barra de fórmulas, mas agora eu quero te mostrar onde que está o editor avançado. Aqui na guia "Exibição", vou clicar em "Editor Avançado", e teremos um grande editor, como se fosse um bloco de notas, onde podemos escrever o que queremos aqui. Então, eu poderia criar essas etapas manualmente, escrevendo-as - caso soubesse qual é a função para utilizar -, e os seus argumentos. Então, eu posso ou escrever diretamente aqui ou, então, utilizar as ferramentas que já tem aqui na faixa de opções. Vamos observar um pouco esse editor avançado: nós temos um let; um in; as etapas, entre eles; e depois uma etapa aqui que se referencia a etapa anterior. É muito importante que você entenda essa estrutura, principalmente quando avançar no conhecimento no Power BI. Dessa forma, você vai conseguir fazer modificações diretamente no editor avançado e também vai conhecer um pouquinho mais sobre a sua estrutura de formação das transformações. Então, vamos ver cada uma dessas etapas detalhadamente para aprendermos como o Power Query armazena todas essas informações. Vou clicar em "Cancelar" e voltar para a apresentação. Vamos verificar agora as características do editor avançado. A primeira delas é que as etapas são inseridas após o let. Nesse exemplo, eu vou destacar agora todas as etapas que estão após a palavra let e antes da palavra in. O Power Query foi desenvolvido assim e sempre que você abrir o editor avançado teremos um let, um in e as etapas internas. Para a segunda característica, vamos observar que os nomes das etapas, eles podem ter espaço mas, se tiverem, terão que ser precedidos por um jogo da velha e inseridos entre aspas. Observe o exemplo da palavra fonte e da personalização adicionada. Na fonte, nós não temos espaço e por isso ela não tem o jogo da velha e as aspas; já na personalização adicionada, temos um espaço entre essas duas palavras e, por isso, temos o elemento do jogo da velha e aspas. É importante ressaltar novamente que quando criamos transformações no Power Query, ou seja, utilizamos a faixa de opções e clicamos em alguma transformação, ele já automaticamente insere esses nomes em português, mas as funções permanecem em inglês. Os nomes das etapas não tem uma relação direta com a função utilizada, apesar de que o Power Query insere em português de uma forma extremamente descritiva, ou seja, se eu estou fazendo uma transformação de colunas removidas, ele vai inserir esse nome, "Colunas Removidas". Apenas para fins de referência, para facilitar se eu precisar mudar essa etapa e, por exemplo, removê-la ou alterá-la. A característica número três é que as etapas são separadas por vírgula, exceto a última antes do in. Vamos verificar esse exemplo: nós temos várias vírgulas após cada uma dessas etapas, entretanto, antes da palavra in, que foi destacado agora, nós não temos uma vírgula, destacado pelo retângulo vermelho. Você não pode confundir as vírgulas que separam os argumentos dentro das

funções, que sempre ficam entre os parênteses e após o nome das funções, com essas vírgulas que separam as etapas. Elas servem para propósitos diferentes. E a palavra in, no final do editor avançado, determina que as etapas intermediárias terminaram, por isso não temos uma vírgula antes da etapa do in. Agora, vamos verificar a característica número quatro: os dados importados do Power Query para o Power BI sempre serão provenientes da etapa informada depois do in, que normalmente é a última etapa. Então, aqui nós temos um exemplo, que a última etapa se chama "Data Expandido", e ela também foi referenciada depois do in. Ou seja, esse será o estágio dos dados que será importado para o Power BI. É sempre preciso ter em mente que todas as etapas são estágios intermediários de transformações para se chegar no estágio final dos dados, que é a última etapa. Elas são todas armazenadas dentro do editor avançado para termos total controle sobre as transformações que são realizadas e isso também é muito bom porque se precisarmos alterar alguma transformação aplicada facilmente podemos apagar a etapa necessária ou, então, alterá-la. Essa foi uma introdução sobre o editor avançado. Nas próximas aulas, faremos transformações dos dados. E você sempre deve ter em mente que cada funcionalidade que clicarmos na faixa de opções vai gerar uma etapa com uma função escolhida pelo Power Query para executar as transformações e, à medida que avançar no conhecimento do Power Query e do Power BI, você também pode livremente alterá-las ou criar manualmente dentro do editor avançado.

Limpar dados indesejados

Agora que você já sabe que conexões e tratamento de dados dentro do Power Query geram etapas que utilizam funções da linguagem M e que são armazenadas dentro do editor avançado, vamos prosseguir utilizando outros arquivos para transformar dados. Eu estou com o "0403 População 2001 a 2018" aberto, que se encontra na pasta de exercícios do capítulo 4 desse curso, e nós temos um relatório de população por estado, de 2001 a 2018, e nessa planilha daqui, de estados regiões, temos uma informação simples de estado e suas siglas. Em "Populações por ano" teremos um trabalho grande, que será tratar esses dados. Eu tenho um cabeçalho do relatório, uma linha em branco e posteriormente os cabeçalhos dos dados, que é o que me interessa para descrever esses dados aqui embaixo. Também temos alguns dados não preenchidos aqui que precisaremos tratar. Faremos tudo isso dentro do Power Query. Então, eu vou clicar no "X" do Excel para fechar, em um arquivo de Power BI em branco, na página inicial, vou clicar em "Obter dados" e escolher "Excel". Vou selecionar esse arquivo da pasta, do "Capítulo 4", e abrir. O Power BI está fazendo a importação dos dados e na janela "Navegador", nós temos a planilha "Estados e regiões" que eu quero importar, então eu vou selecionar, e a "População por ano", que também quero importar, mas ela já está bem bagunçada. Vamos tratá-la dentro do Power Query. Vou clicar no botão "Transformar dados" e ele realizará a importação desses dados. Com as duas tabelas aqui dentro, em "Consultas", vamos realizar as transformações, começando com a avaliação de que eu tenho esse cabeçalho promovido, aqui do lado nós temos uma etapa chamada "Cabeçalhos promovidos". O Power BI automaticamente promove cabeçalhos ao importar uma fonte de dados, essa é uma etapa padronizada dele. Eu posso ou não aceitar essa promoção. Você já deve saber disso, principalmente se já fez o outro curso "Descubra o Power BI" aqui no LinkedIn Learning, que cada uma dessas etapas podem ser removidas clicando nesse "X" ao seu lado. Então, se eu clico nesse "X" do "Tipo alterado", ela foi removida e depois, na "Cabeçalhos promovidos", também foi removida. Agora estamos em um estágio dos dados de navegação.

Transforme colunas em linhas

Vamos continuar transformando o mesmo arquivo da aula anterior. Se você acompanhou junto comigo, já possui o arquivo "População Estados Brasil". Caso essa seja a primeira aula que você esteja vendo desse capítulo, abra o arquivo "0404 População Estados Brasil" que contém o mesmo estágio da aula anterior. A necessidade que temos de transformação dessa aula é a seguinte: eu preciso garantir que tenham colunas com valores armazenados abaixo delas, de acordo com a sua descrição. Então, aqui eu tenho a região e as regiões, a publicação e as datas, e aqui eu tenho os anos e os seus valores. Entretanto, eu não posso ter várias colunas com valores armazenados do mesmo tipo de dado; 2001, 2, 3 e 4 são valores de anos, ou seja, eles deveriam estar armazenados todos em uma coluna, e não horizontalmente como estão aqui. Esses anos precisam se transformar em linhas e cada um desses valores precisa dinamicamente referenciar a cada um desses anos. Então eu preciso de uma nova coluna chamada "Anos", com 2001 e assim por diante, até 2018, e os seus valores logo ao lado, ou seja, eu quero ter uma tabela com apenas cinco colunas: unidades da federação, região, publicação, anos e valor. Para fazer essa transformação, utilizaremos uma

funcionalidade chamada "Transformar colunas em linhas", que está localizada aqui na guia "Transformar", nesse ícone, "Transformar colunas em linhas". Mas, antes de clicar nela, precisamos realizar a seleção dos dados que queremos transformar. Nesse caso, são todas essas colunas, de 2001 até 2018. Então vou selecionar todas elas, voltar aqui para frente e agora, na guia "Transformar", vou clicar em cima desse símbolo de transformar colunas em linhas. Não confunda com esse ícone daqui, é bem parecido, mas são funcionalidades diferentes. Então, em "Transformar colunas em linhas", depois de realizar a seleção de todas as colunas, e eu realizei clicando aqui na 2001, depois pressionando a tecla SHIFT, usando a barra de rolagem aqui para ir para a última coluna, e com o SHIFT pressionado, cliquei em 2018. Dessa forma, selecionei todas as colunas. Agora, em "Transformar", vou clicar na funcionalidade e teremos uma transformação peculiar aqui, dentro dessa tabela. Olhe que eu tenho apenas cinco colunas, que era o nosso objetivo, com os anos aqui verticalmente, escrito "Atributo" aqui em cima e os seus valores correspondentes. Ou seja, fizemos uma transformação de colunas em linhas. Agora, eu posso clicar duas vezes em cima dessa coluna de Atributo para editar o seu nome, vou colocar "Anos", e agora, em "Valor", está correto, não preciso fazer nenhuma transformação adicional aqui dentro. Ainda precisamos fazer algumas alterações nessa tabela, mas deixaremos para a próxima aula. Salve esse arquivo clicando no disquete, posicionado aqui na parte superior esquerda, e continuaremos na próxima aula.

Utilize o preenchimento para cima e para baixo

Continuando com as transformações realizadas no arquivo da aula anterior, chamado "População Estados Brasil", vamos avaliar essa seguinte informação aqui: quando eu posiciono o mouse em cima dessa pequena faixa verde, temos a informação de que tenho 108 valores válidos e 486 vazios. Essa informação é muito importante principalmente se você possuir erros dentro dessa coluna, o que não temos. Vamos avaliar esses vazios. Eu vou selecionar uma célula e usar a barra de rolagem para verificar que eu tenho vários valores nulos correspondentes à Rondônia, que é da região Norte. Aqui embaixo também, Acre é Norte e assim por diante. Se eu uso mais a barra de rolagem, teremos uma informação da região Nordeste para Maranhão, que está nulo, ou seja, eu posso preencher esses valores que estão aqui para baixo, para automaticamente já mudar a região do Maranhão de nulo para a região do Nordeste. O mesmo acontecerá com outras informações de estados aqui embaixo, região Sudeste, Minas Gerais, está nulo. Portanto, podemos preencher esses valores para baixo. Para isso, vou selecionar a coluna "Região" e aqui na guia "Transformar" vou escolher essa opção da seta para baixo e escolher "Transformar para baixo". Ao clicar nela, teremos as informações aqui dentro dessa faixa verde, todas com 594 válidos, zero erros e zero vazios, que era o que queríamos. Então, se eu utilizo a barra de rolagem, o Acre está na região Norte, Amazonas, região Norte, aqui embaixo teremos a mudança para a região Nordeste, com Maranhão também posicionado com a região correta. Então, a funcionalidade de preencher para cima ou para baixo, de acordo com a necessidade que você tem, é muito útil para alterar automaticamente esse tipo de situação quando temos uma avaliação de muitos valores vazios. Ao tratar outros dados em projetos que desenvolverá, sempre verifique a informação que tem dentro dessa barra verde e trate-as principalmente se você tiver erros sendo apontados nela.

A importância dos tipos de dados

Continuando as transformações realizadas na aula anterior, eu estou no arquivo "População Estados Brasil", e agora verificaremos a importância dos tipos de dados. Ao lado de cada nome de coluna teremos alguns símbolos: ABC no calendário, ABC 123; esses são os indicativos de qual é o tipo de dado armazenado nesta coluna. Aqui é um texto, porque se eu cliço aqui no ABC, vejo que esse símbolo é referente a um texto; aqui em "Data", o calendário, por mais óbvio que pareça, é a data. E é muito importante definir todos os tipos e, principalmente, verificar se eles estão corretos. Aqui em "Valor" nós não temos um tipo de dado definido, está ABC 123. Se eu vejo nessa listagem não existe esse tipo de dado aqui, não temos esse ícone. Isso porque essa coluna ainda não foi determinada em qual tipo de dado ela se encontra. Portanto precisamos alterar o seu tipo. Os tipos dessas colunas já foram alterados porque promovemos o cabeçalho em algum momento de transformação, que foi nesse daqui: cabeçalhos promovidos; e, automaticamente, o Power BI encerra uma etapa de tipo alterado tentando alterar esses tipos de dados. Mas como a coluna de Valor surgiu apenas na etapa de colunas não dinâmicas, o seu tipo não foi alterado porque não temos outros tipos aqui abaixo, apenas acima depois da etapa de cabeçalhos. Portanto, precisamos alterá-lo manualmente. Antes de fazer isso vamos selecionar a consulta "Estados e Regiões" e verificar que ainda não temos a promoção desses dois cabeçalhos. Aqui no cabeçalho está "Coluna 1" e "Coluna 2" em inglês. Para

promovê-los basta selecionar esse símbolo de tabela e escolher usar a primeira linha como cabeçalho. Já que a primeira linha é o Estado e Sigla, podemos fazer isso. Então cliquei no "Símbolo de Tabela", posteriormente em "Usar a Primeira Linha como Cabeçalho". Quando eu fiz isso, automaticamente o Power Query inseriu uma etapa chamada "Cabeçalhos Promovidos" e a "Tipo alterado" que é realizada automaticamente, assim como você já aprendeu anteriormente. Então, aqui em População Por Ano precisamos, ainda, alterar esse tipo de dado. Em "Valor", vou clicar em ABC 123 e realizar a sua alteração conforme o conteúdo dessa coluna, que são números inteiros. Então eu vou escolher 123, número inteiro. Ao clicar nessa opção temos um "Tipo Alterado 1" que transforma apenas essa coluna, mas temos um indicativo muito importante aqui. Você lembra dessa barra verde que tínhamos em outras colunas e que foi mencionado na aula anterior? Quando eu posiciono meu mouse em cima dela está apontando os seis erros. Se eu uso a barra de rolagem para verificar quais são, e clico na célula desse erro, aqui embaixo, Power Query está falando que não conseguiu converter em número um valor que tem um número e, posteriormente, um (*) entre parênteses. Ou seja, estamos com valores textuais e numéricos misturados em um mesmo registro. Antes de tratar desse tipo, precisamos transformar esse asterisco e parênteses em nada, ou seja, removê-los daqui. Para isso eu vou permanecer com a Tipo Alterado aqui, e vou selecionar a etapa de cima: A preenchido abaixo. Depois eu vou selecionar a coluna Valor, e vamos realizar uma alteração nesse estágio de etapa. E aí, quando fizermos a etapa do tipo alterado, ela conseguirá alterar todos sem erros, porque não teremos mais o asterisco e os parênteses. Então, com essa etapa selecionada selecione também a coluna aqui em transformar vou procurar por esse ícone 1 2 que significa substituição de valores e vou clicar nele. O Power Query vai alertar que eu estou inserindo uma etapa no meio de duas etapas. Normalmente fazemos alterações lineares sempre clicando na última etapa, mas eu estou propositalmente colocando antes da tipo alterado um para substituir esses valores. Como eu tenho certeza do que eu estou fazendo vou clicar em Inserir e agora eu vou dizer qual é o Valor a ser localizado. Vou colocar parênteses asterisco parênteses (*). E aqui embaixo vou substituir por nada, não vou colocar nada. Quando eu vou clicar no botão ok, não consigo, porque ele está desabilitado, e tem um alerta aqui do lado. Quando eu clico nele, está falando para inserir um valor Number, ou seja, ele não está permitindo alterar os valores aqui desta coluna, porque ela está sendo considerada com valores numéricos, mesmo estando formatada como ABC 123. Consigo verificar isso porque todos esses valores estão à direita. Valores textuais ficam na esquerda, e valores numéricos e de data ficam na direita. Então eu vou cancelar essa inserção e vou alterar essa coluna para Texto. Clicando no ABC 123, posteriormente, Texto, vou dizer que eu quero inserir essa etapa no meio, e agora eu vou substituir esses valores. Vou verificar se eu tenho outros tipos de tratamentos que eu posso fazer. Então um asterisco (*) com os parênteses, outro, e também tenho essa situação de dois asteriscos. Então precisamos abranger ambas. Vou selecionar a Valor na guia Transformar. Vou escolher 12, que é a etapa de substituir valores. Vou clicar em Inserir, porque eu quero colocar uma etapa aqui no meio. E o valor a ser localizado será: (*) parênteses asterisco parênteses, ou substituir por nada e clicar em ok. Fizemos a primeira substituição, vou fazer outra: transformar o 12. Para substituir os valores, vou inserir essa etapa. E o valor a ser localizado será parênteses, dois asteriscos, parênteses (**), vou substituir por nada e clicar em ok. Portanto, só podemos fazer substituições em valores textuais, mesmo se eles estiverem classificados como ABC 123, que não determina que o valor é numérico. Agora, quando eu clico em Tipo Alterado, nesta etapa vamos verificar se ainda temos alguns erros. Vou utilizar a barra de rolagem, não temos nenhum erro por enquanto. Vamos verificar se tem mais algum, porque aqui ainda está apontando com erros. Temos dois, eu posso clicar nesses três pontinhos e dizer para "manter erros" para verificar o motivo deles. Ao clicar aqui vou selecionar essa célula em branco e ainda tínhamos mais um, que tinham três asteriscos. Então, agora, essa etapa que eu criei aqui de Erros Mantidos, foi apenas para consultar os erros. Portanto, eu não posso permanecer com ela. Posso clicar no x para voltar ao estágio anterior. Vou selecionar a coluna de Valor e continuar realizando a substituição de valores, mas eu preciso fazer isso antes da Tipo Alterado para que, quando ele for alterado, não tenha nenhum valor que seja diferente de numérico. Então selecionei a Valor, Transformar, 12. Vou inserir a etapa no meio, procurar por parênteses, três asteriscos, parênteses (***), substituir por nada, clicar em ok. E aqui, em Tipo Alterado temos um valor com a avaliação se estão válidos com erros ou vazios 100% válidos. Portanto, alterar tipos de dados é bastante importante em algumas funcionalidades, como a Substituir Valores, não será possível executar caso eles não estejam apropriadamente formatados como texto, ou como número. Todas essas substituições são completamente irrelevantes na hora de fazer os cálculos utilizando a linguagem DAX. Portanto, você sempre deve observá-los antes de concluir qualquer projeto. Para prosseguirmos, clique no botão de disquete para salvar esse projeto.

Juntar dados

Nessa aula você vai aprender a juntar consultas diferentes em uma mesma consulta escolhida. Eu estou continuando a transformação dos dados realizados na aula anterior, no arquivo "População Estados Brasil". Se você está acompanhando junto comigo, pode utilizar o seu arquivo. Caso essa seja a primeira aula que você está vendo desse capítulo, utilize o arquivo "0407 População Estados Brasil" que está no mesmo estágio de alterações da aula anterior. Nessa tabela População Por Ano, nós temos os anos e suas populações de acordo com as regiões do Brasil de 2001 até 2018 apenas. Mas eu tenho um arquivo com a mesma estrutura de dados, ou seja, com cinco colunas, todas nomeadas iguais, está escrito: Unidade da Federação, Região, Publicação, Anos e Valor de 2019, e eu quero atualizá-lo e inseri-la aqui dentro. Para isso, vamos verificar o arquivo de Excel, que é esse daqui que está presente na pasta de exercícios desse curso, "0407 População 2019.xlsx". Tenho cinco colunas: Unidades, Região, Publicação, Anos e Valor, e por estarem na mesma estrutura de dados, podemos importar diretamente para aquela tabela. Então eu vou fechar esse Excel, e aqui no Power Query mesmo, vou vir em Nova Fonte e escolher Excel. Vou selecionar "0407 População 2019" e clicar em abrir. Vou selecionar essa planilha, posteriormente "ok". E teremos a sua importação, com todos os valores corretos aqui em cima. Sempre vou verificar essa barra verde e, quando eu posiciono meu mouse, está dizendo que tem 24 vazios. Vou usar a barra de rolagem, e temos vários nulos, que é o mesmo que vazios. Então eu vou retirá-los. Para isso eu vou filtrar clicando na seta para baixo, desmarcando nulo, e depois "ok". Agora temos essa barra verde informando 27 válidos. Ok, podemos prosseguir, estamos com os tipos de dados alterados. Então eu tenho ABC, ABC, data, número e número. É muito importante, antes de juntarmos esses dados de População Ano 2019 com a População Por Ano, que os tipos de dados sejam iguais. Ou seja: ABC ABC, data, 123 e 123. Vou clicar em População Por Ano e verificar se são os mesmos. ABC ABC, data, ABC. Lá na População Ano 2019, aqui estava como número, ambos precisam ser do mesmo tipo de dado. Então, tenho que escolher um dos dois para alterar. Lá em População Ano 2019, vou alterar 123 para o texto no cabeçalho de anos. Agora, na População Por Ano, quero juntar essa consulta nessa daqui de cima, ou seja, teremos do ano 2001 até 2019, e não apenas até 2018. Para juntar duas consultas nessa ativa daqui, eu vou clicar na página inicial, e nesse botão de combinar vou escolher Acrescentar Consultas. Ao clicar nessa opção posso selecionar duas tabelas, ou seja, essa atual e mais uma, ou então três ou mais tabelas. Só tenho duas, então vou selecionar a consulta que eu quero acrescentar nessa tabela atual, que é a População Ano 2019. Vou clicar em ok. E agora fizemos uma consulta acrescentada. Essa etapa daqui que gerou essa função table.combine, de combinar, em português, e eu posso fazer uma classificação, apenas para fins de validação, para verificar se o ano de 2019 está aqui. Então vou clicar nessa seta para baixo, vejo que já está aqui, 2019, mas eu posso ver os dados clicando em Classificar em Ordem Decrescente. Não faz sentido eu deixar essa etapa de linhas classificadas, porque eu fiz ela apenas para ver os dados. Então eu vou clicar no x e temos duas consultas separadas juntadas na mesma tabela. E com esses dados atualizados até o ano que eu preciso, vamos prosseguir na próxima aula com as transformações necessárias.

Mesclar dados

Continuando com o mesmo arquivo da aula anterior, População Estados Brasil, agora faremos uma mesclagem de dados para incrementar essa tabela. Nós temos os Estados aqui dentro dessa coluna: Rondônia, Acre, Amazonas, e assim por diante, e eu também quero a informação abreviada, ou seja, aparecer AC para Acre, AM para Amazonas, RR para Roraima, e assim por diante. Essas informações estão nessa consulta: Estados e Regiões, nesta coluna: Sigla. Então precisamos fazer uma mesclagem entre essas duas consultas. Na População Por Ano, como eu quero que essa coluna apareça aqui, eu vou selecioná-la e, posteriormente, aqui na página inicial, vou escolher esse botão de combinar, e selecionar "Mesclar Consultas". Vou clicar nessa opção, e agora nós temos a População por ano, que é essa consulta e tenho que selecionar a outra, que é Estados e Regiões, onde eu tenho a coluna que eu quero trazer. Posteriormente, eu preciso relacioná-las, dizer para o Power Query quais são as duas colunas que têm os mesmos dados, que é a Unidade da Federação, onde eu mostrei que tem os nomes dos Estados, com uma coluna de Estado da Estados e Regiões. Eu sei que aqui está escrito Brasil, mas é porque dentro dessa coluna temos outras informações também, como o país, que é somente o Brasil, e suas regiões, por exemplo: região Norte, Nordeste. E eu quero te mostrar o que acontecerá nesse caso, porque não temos Brasil aqui nesses estados. Vamos ver assim que fizemos a mesclagem. No tipo de junção vou deixar Externa Esquerda, dizendo que todas as linhas da primeira coluna, que é essa daqui, serão considerados ainda, serão

mantidas, e apenas as linhas da segunda tabela que serão consideradas. Ou seja, essa daqui. Embaixo, temos uma pré-visualização que a seleção corresponde a 513 linhas de 621. Isso significa que temos Brasil; Brasil Região Norte, Nordeste; que não tem aqui nesse estado porque aqui temos estados de fato, e que essa não correspondência vai gerar um tipo de valor. Não é um problema, eu vou mostrar para você o que vai acontecer. Então eu vou clicar em ok, e o resultado foi a inserção de uma table, ou seja, uma tabela, dentro dessa consulta daqui de População Por Ano. Quando eu clico nessa parte cinza ou em branco da tabela, aqui embaixo conseguimos verificar quais são as informações dessa linha. Como eu tenho Brasil aqui, o resultado foi nada. Mas se eu seleciono um estado de fato, que é aqui em Roraima, vou selecionar aqui, temos RR. Para expandir essa tabela e mostrar os seus valores, basta clicar nessas duas setinhas, e, posteriormente, escolher qual é a coluna que quer considerar. Não quero estado porque eu já tenho essa informação aqui dentro. Então eu vou exportar apenas a sigla dessa tabela. Vou clicar em ok, e o resultado para os valores que não tinham na tabela Estados de Regiões, como o Brasil, foi nulo. É esse valor de nulo. Mas quando temos um estado que corresponde nas duas tabelas, como Rondônia, temos o seu retorno correto. E se eu verificar usando a barra de rolagem, temos um retorno correto para todos esses estados. E o objetivo dessa aula foi concluído com sucesso. É importante ressaltar que eu não posso apagar esse Estados e Regiões, se eu clicar com o botão direito em cima da tabela e depois excluir, o Power Query vai reclamar, vai exibir essa informação dizendo que não pode ser excluída, porque está sendo referenciada por outra consulta, que é a População Por Ano. Isso porque, dentro da função utilizada lá dentro da População Por Ano, e em Consultas Mescladas, o nome dessa tabela Estados e Regiões é referenciado dentro da função Table.NestedJoin que é a execução da Consulta Mesclada, que é a mesma coisa que clicar aqui na página inicial, combinar, iniciar as consultas. Esse processo gerou essa etapa. Então essa consulta está referenciada aqui dentro, Estados e Regiões. Por isso não podemos apagá-la. O mesmo acontece com a População Ano 2019, que foi o que fizemos na aula anterior, juntando a População Por Ano com a População Ano 2019. Se eu clico com o botão direito em cima dela e excluo, não será possível porque ela está sendo utilizada dentro de uma função dentro dessa consulta. Portanto, nessa aula você aprendeu a fazer uma mesclagem de dados e incrementar suas tabelas. E também analisamos o motivo de não conseguir deletar essas consultas por causa da referência dentro das funções.

Crie colunas condicionais

Você pode enriquecer e adicionar novos dados em uma tabela por meio da coluna condicional, que é o recurso que aprenderemos nessa aula. Eu estou com o mesmo arquivo da aula anterior, "População Estados Brasil", e vamos adicionar dentro da consulta "População por Ano", novas informações. A primeira delas é que na coluna de "anos", quando tiver o ano de 2019, será o ano atual; e todos os outros serão passado. Então eu vou posicionar o meu mouse aqui na faixa de opções... na "Adicionar coluna", e escolher essa opção, "Coluna condicional". O nome da nova coluna será "Informação Atual". E a coluna que eu quero validar é a "Anos". Se ela é igual à "2019". Eu quero que seja inserido dentro de uma nova coluna, que se chamará "Informação Atual", o dado "Ano Atual". Caso contrário, ou seja, se não, quero que coloque "Passado", ao pressionar "ok"... teremos essa nova coluna com o "Passado" para os anos de 2001 até 2018, mas quando tivermos o ano de 2019 - e para localizar rapidamente eu vou classificá-lo em ordem decrescente - quando tiver 2019 será "Ano Atual". Vou remover essa classificação, porque foi uma etapa que eu criei apenas para visualizar os dados; estão aqui em "Linhas Classificadas" vou clicar no X", e agora vamos realizar mais um tipo de coluna condicional mas utilizando valores numéricos. Nós temos essa coluna de "Valor", com a população por cada um desses estados, regiões e o Brasil. Eu quero inserir uma nova coluna aqui do lado que vai validar se essa população daqui for maior do que 10 milhões. Se for, quero que classifique como "Maior que 10 milhões". Se não for, quero que classifique como "Menor". Para isso, aqui na "Adicionar coluna", "Coluna condicional"... Vou selecionar o nome da coluna "Valor" e quero validar se ela é maior ou igual a 10 milhões. Então é 1 seguido de sete 0. A saída será "Acima... de 10 milhões". Vou colocar abreviado "10 m". Se não, caso contrário, se for menor do que 10 milhões, quero que coloque "Abaixo... de 10 milhões". O nome da nova coluna será "Grupo de tamanho". E posteriormente vou clicar em "ok". Uma nova coluna foi criada, aqui temos "Acima de 10 milhões" porque a população do Brasil, mas vamos localizar um outro estado... que está abaixo de 10 milhões, que é o estado do Acre. E também teremos alguns estados acima de 10 milhões... que será... a Bahia. Vamos localizá-la aqui. Então aqui na "Bahia", se eu verificar, temos acima de 10 milhões. E também teremos outros estados nessa classificação. Então você aprendeu nessa aula a como enriquecer os seus dados com informações baseadas nos próprios dados dentro

da consulta. Durante projetos de BI, lembre-se dessa funcionalidade para enriquecer os filtros que poderão ser utilizados em gráficos.

Referencie e duplique tabelas

Nessa aula, nós vamos analisar a diferença entre duplicar uma consulta ou referenciá-la. Eu estou prosseguindo com os tratamentos de dados no arquivo de "População Estados Brasil" que estamos trabalhando nas últimas aulas, e eu preciso duplicar essa consulta porque utilizaremos na próxima aula. Para duplicá-la, eu tenho duas opções: se eu clicar com o botão direito em cima dela, posso realmente duplicar; ou então fazer uma referência dela. Vou escolher primeiro essa opção de "duplicar"... E nós temos a "População por Ano (2)". Agora eu vou clicar de novo em cima da "População por Ano"... E escolher "referência". Agora temos uma outra, "(3)". Vamos analisar a diferença entre as duas. Na "População por Ano (2)", que foi a duplicada, nós temos todas as etapas de transformação aqui na direita; já na "População por Ano (3)" temos apenas uma, que se chama "Fonte". Essa daqui foi uma referência... dessa consulta daqui, que se chama "População por Ano", que está aqui em "Consultas". Já a "População por Ano (2)" não fez uma referência, ela literalmente duplicou todas as etapas - inclusive a "Fonte", "Navegação", e todas aquelas que desenvolvemos ao longo das aulas anteriores. Então essa é a diferença entre duplicar e referenciar. A escolha entre essas duas opções deve ser tomada de acordo com a seguinte análise: Tudo o que você fizer dentro de "População por Ano" será automaticamente replicada para a "População por Ano (3)", que foi aquela consulta que fizemos uma referência. E isso porque está escrito aqui: igual "População por Ano", então todas as alterações aqui dentro serão automaticamente passadas para essa. Já na outra, "População por Ano (2)", que é uma duplicata da "População por Ano", nenhuma das alterações realizadas aqui em "População por Ano" será aplicada nessa consulta, porque nós temos uma do zero - ou seja, uma nova consulta aqui que referenciou a todos os passos sem fazer nenhuma referência à consulta, e sim às transformações realizadas de uma forma independente. Portanto, para decidir entre "duplicar" ou "referenciar", você tem que pensar nisso. Tudo que você fizer na sua tabela base pode vir para a tabela que está duplicada ou referenciada? Se sim, você escolhe uma referência, porque é muito melhor apenas referenciar do que fazer todas as etapas novamente. Do ponto de vista de performance, teremos várias etapas - tanto nessa primeira consulta quanto na outra - sendo executadas, duplicar uma tabela é muito mais pesado em termos de performance para o Power Query. Já a referência vai aproveitar todas as etapas da consulta referenciada, tendo uma performance muito melhor. Vou manter apenas uma dessas duas consultas, quero apenas a referência; vou clicar na "População por Ano (2)" com o botão direito e excluir... E vou renomear essa "População por Ano (3)" para "Agrupamento... Estados". E na próxima aula aplicaremos um agrupamento nesses dados.

Agrupe dados no Power Query

Na aula anterior fizemos uma referência da consulta "População por Ano" e criamos a consulta "Agrupamento Estados". Nessa aula faremos um agrupamento de todas as regiões, suas populações, do ano de 2019. Para isso, aqui na consulta "Agrupamento Estados", vou clicar no botão de "Filtro" e selecionar apenas o ano de 2019. Para isso eu vou desmarcar a opção "Selecionar Tudo" e aqui embaixo vou selecionar "2019". Vou pressionar "ok", e agora temos apenas a população desses estados no ano de 2019. A minha necessidade é saber qual é a população por cada uma dessas regiões agrupadas, sem ter o detalhamento por estado. Para isso, aqui na página inicial, vou selecionar a ferramenta "Agrupar por". Nessa janela eu vou configurar a visão que eu quero ver. Então, em "Agrupamento por" vou escolher "Região". Então selecionei esse menu, "Região", e em "Nome da nova coluna", é aquela que vamos criar fazendo uma somatória de toda a população, e por isso é importante que esses valores estejam formatados como numéricos - e estão. No "Nome da nova coluna" vou colocar "População"... E na "Operação" será uma soma. Quero fazer uma soma da coluna "Valor". Vou clicar em "ok". E temos as regiões do Brasil e as somatória das respectivas populações. Podemos fazer também outros tipos de agrupamento. Para isso, eu vou duplicar essa consulta clicando com o botão direito em "Duplicar"... e agora, no "Agrupamento Estado (2)", eu vou editar o agrupamento realizado. Aqui em "Linhas Agrupadas" nós temos essa engrenagem aqui do lado. Vou clicar nela... e ao invés de agrupar por "Região", agora eu quero agrupar por "Grupo de Tamanho", que foi a coluna condicional que fizemos. O nome da nova coluna será "Quantidade de Estados". E a operação, ao invés de soma, eu quero fazer uma "Contagem de Linhas Distintas". Também poderia escolher "Contar Linhas" porque não temos valores duplicados na coluna que estamos contando. Então escolhi "Contar Linhas". Vou pressionar "ok" e agora temos uma informação que, abaixo de 10.000.000 de população, temos

21 estados; e acima de 10 milhões, temos 6. O total dá 27, que são 26 estados mais o Distrito Federal. O que acabamos de fazer foi uma contagem de todas as linhas que tinha esse agrupamento de "Abaixo de 10 milhões", e esse "Acima de 10 milhões". Se eu clico em "Linhas Filtradas", nós temos 27 linhas aqui; e para cada um dos agrupamentos desses estados aqui nessa coluna, essa funcionalidade de agrupar contou quantos estados tem. Vou selecionar de volta a etapa "Linhas Agrupadas" para vermos o resultado, que foi uma contagem correta dos estados acima e abaixo de 10 milhões em população. Portanto conseguimos agrupar dados utilizando a ferramenta "Agrupar por", e você pode realizar vários cálculos - entre eles: a soma e a contagem, que vimos nessa aula; e outros como a média, o máximo, o mínimo e a mediana.

Modifique o endereço de uma fonte de dados

Na aula anterior, finalizamos todas as transformações previstas nesses dados. Nessa aula, vamos finalizar esse projeto clicando em "fechar e aplicar" para trazer os dados que estão aqui do Power Query para o Power BI. Então vou clicar aqui, os dados serão importados e posteriormente disponibilizados aqui nos "campos", com todas as transformações realizadas. Antes de finalizarmos o capítulo 4 sobre Power Query, é muito importante que você saiba alterar as fontes de dados. Não podemos presumir que as pastas e os caminhos serão os mesmos para sempre, portanto eu vou simular uma alteração do nome da pasta e posteriormente vamos trocar o seu caminho aqui. Antes de fazer isso, eu vou salvar esse arquivo clicando aqui no disquete. E posteriormente fechá-lo. Todos os arquivos que utilizamos neste capítulo estavam na pasta chamada "Capítulo 4". Vou clicar nela e renomear para "Capítulo 4 Renomeado". Agora vou abrir e posteriormente abrir o arquivo de Power BI dessa aula. Por enquanto temos os dados prontos aqui para serem utilizados, entretanto, se abirmos o Power Query aqui no editar consultas... Temos esse alerta dizendo "Data Source.Error: Não foi possível localizar uma parte do caminho", que é esse daqui, olha: "Capítulo 4", nós colocamos um renomeado depois. Tivemos uma aula apenas sobre o editor avançado e podemos alterar esse nome do caminho de duas formas: pelo editor ou pelas configurações da fonte de dados. Vamos ver primeiro no "editor avançado". Vou clicar em cima dele, aqui na "página inicial"... E eu posso alterar o caminho dessa consulta que está selecionado vindo aqui no caminho da fonte e colocando "renomeado" após o "capítulo 4". Vou clicar em concluído... E teremos os dados disponíveis de volta. A segunda forma de alterar o caminho é vindo em "configurações da fonte de dados". Nessa janela, posso escolher um desses caminhos, "alterar fonte"... E aqui eu posso colocar o "renomeado" após o "capítulo 4". Vou clicar em ok. Vou fazer o mesmo em todas, então cliquei agora nesse último, "alterar fonte"... E se o caminho tiver mudado muito, você também pode clicar no botão "procurar". Mas como mudou pouco, eu vou apenas renomear por aqui. Ok. Consigo ver que essas alterações já estão válidas porque automaticamente as consultas mudaram daquele símbolo de exclamação, para símbolos de tabela. Fecho essa janela e agora verificar que todos os dados estão disponíveis. Você acabou de aprender a alterar o caminho da fonte de dados. Posteriormente você pode fechar e aplicar essas alterações, clicando aqui nesse botão. Ele vai aplicar as alterações novas para o Power BI. E salvar esse projeto.

O Power Query realiza modificações apenas dentro do Power BI. Os arquivos originais ficarão intactos e não é possível alterá-los pelo Power BI e sobrescrever seu estado original.

5. Relacionando dados

Compreenda a cardinalidade dos dados

Antes de começarmos a utilizar as funções no DAX para criar cálculos e responder questões de negócios, precisamos entender de relacionamentos e como realizá-los no Power BI. Relacionamentos possuem cardinalidade. Vamos entender o que essa palavra significa. Na modelagem de dados, cardinalidade é o grau de relação entre colunas de tabelas. E esse grau pode ser explorado com os tipos de cardinalidade. Entre eles 1 : 1, ou seja, duas colunas com valores únicos relacionados; 1 : N, representado por esse "N", que são duas colunas relacionadas, uma com valores únicos e outra com valores duplicados. E N:N, ou seja, muitos para muitos, que são 2 colunas com valores duplicados relacionadas. E normalmente, esses relacionamentos de 1 para muitos, é realizado com chaves primárias e estrangeiras. Vamos verificar o que significam essas duas palavras. A chave primária, é uma coluna com valores únicos dentro de uma tabela, ou seja, que não se repetem. É uma coluna de referência. É como se fosse o CPF da coluna. No Brasil cada pessoa tem um CPF diferente e ele não se repete. Esse conceito é parecido com a chave primária. Também temos a chave

estrangeira, que é uma coluna com valores duplicados que não pertence ao assunto direto da tabela atual e é utilizada apenas para estabelecer um link entre tabelas. Por exemplo, imagine uma tabela de vendas onde temos todos os fatos que aconteceram de vendas, ou seja, todas as vendas realizadas. Nós temos uma coluna, que se repete, dos códigos de produto. Esses códigos, eles são uma chave estrangeira da nossa tabela de vendas, porque ele não pertence à venda diretamente, ele é um assunto indireto. Algo foi vendido e suas informações, normalmente, são armazenadas em outra tabela. Por exemplo, para cadastrar uma venda, precisamos que o produto já esteja cadastrado previamente em outra tabela que tem seus atributos, como o preço unitário, sua cor, ou então o modelo e outras informações. A tabela de vendas vai simplesmente usar esse código de produto para relacionar com o código de tabela de cadastro e poder trazer outras informações detalhadas, que são essas que eu acabei de comentar. E quando temos uma tabela de cadastro, podemos falar que ela é uma "Tabela de Dimensão", que possui colunas descritivas. Então pode ser um cadastro de produto, um cadastro de funcionário, com a data de admissão, seu salário, turno e outras informações; um cadastro de regiões com o estado, a sua região, o país, o continente; e outros tipos de tabela que descrevem um fato com mais de uma coluna. Já as "Tabelas Fato", são fatos que aconteceram, ou seja, vendas realizadas. Normalmente possuem uma coluna de data porque os fatos aconteceram em algum momento e essas Tabelas Fato, normalmente, estão associadas com tabelas de cadastro, ou seja, tabelas de dimensão. Por exemplo, quando uma venda é realizada, ela tem uma data que aconteceu, um produto que foi vendido, um cliente que comprou, uma região que foi atendida, um vendedor que vendeu, e todas essas colunas que eu falei aqui agora, esses nomes, se relacionam com tabelas de cadastro, que podem ter sido utilizadas durante a Tabela Fato, de uma venda que foi executada, quando foi cadastrado, que a venda utilizou o produto X, foi vendido por esse vendedor, para esse cliente que se encontra nessa região. Agora que você já sabe esses dois conceitos, podemos prosseguir com a análise de dados.

Explore os tipos de tabelas e colunas do modelo

Nessa aula eu vou te apresentar qual é o projeto que utilizaremos para explorar os relacionamentos e as funções DAX. Até o momento estávamos utilizando um tipo de arquivo para explorar melhor os tratamentos de dados no Power query, e agora utilizaremos um novo arquivo para explorar essa outra parte do curso. Na pasta de exercícios do capítulo 5 temos 2 arquivos. Um chamado "Adventure" e o outro "AdventureWorks". Essa é uma base de uma empresa fictícia criada pela Microsoft e que foi adaptada por mim para explorar melhor os conceitos aprendidos nesse curso. Na aula anterior, nós vimos sobre relacionamentos, cardinalidade e tabelas dimensão e fato. Vamos aprender melhor, na prática, o que significam todos esses conceitos. Eu vou clicar 2 vezes em "AdventureWorks" no access para abri-lo e aqui temos algumas tabelas na esquerda. Vamos explorar cada uma delas. Começando com as "Tabelas Fato". Nós temos uma tabela chamada "Pedidos Detalhes", "Pedido Geral" e meta. A "Pedidos Geral" determina que fatos aconteceram de vendas, ou seja, fizemos alguns cadastros de vendas realizadas nesta data daqui para esse cliente, foi um produto adquirido diretamente do e-commerce, por isso não tem nenhum vendedor, foi vendido para uma região, também, algum produto foi comprado, que temos nessa tabela. Portanto são fatos que aconteceram. Tem data, tem um agente de compra, um agente de vendas, tem um objeto que foi comprado, entre outras informações. É isso que determina o que é uma "Tabela Fato", mas não necessariamente precisa já ter acontecido, pode ser, também, uma previsão. Aqui na meta nós temos o ano, o mês e uma meta prevista para cada um desses anos e meses. Também podemos considerá-la como uma Tabela Fato. É um fato que está previsto de acontecer, agora vamos verificar um outro conceito de tabela que temos tanto no Power BI, quanto em qualquer análise de banco de dados. Também vimos na aula anterior que existem "Tabelas Dimensão", ou seja, aquelas que descrevem algo. Normalmente são tabelas de cadastro. Aqui no Acces, eu posso apontar como Tabelas Dimensão, essa daqui de produtos, ou então a de categorias, a de subcategorias, e as de regiões. Eu vou abri-la e aqui nós temos o ID da região, posteriormente, o país e o seu grupo. Imagine que essa é uma tabela de cadastro e observe que não temos uma repetição aqui dentro dessa coluna. Daqui a pouco vamos ver qual é o nome a característica dessa coluna. Mas observe que aqui em "Pedidos Detalhes", nós temos os fatos que aconteceram com os IDs dos produtos e também em Pedidos Geral, temos os fatos que aconteceram com outras colunas e se eu vier aqui em "ID da Região", temos repetições para essas regiões. Vamos ver o que significa isso nas características do dax. Mas antes de prosseguirmos, eu preciso que você entenda o que são Tabelas Dimensão. São simplesmente cadastros. Então eu tenho o cadastro de regiões, o cadastro de subcategorias com os seus nomes aqui, cadastro de categoria e também de produtos. Cada uma dessas

tabelas são individualmente armazenadas dentro desse banco de dados. E lá no Power BI, vamos fazer um relacionamento entre elas para poder, por exemplo, tentar fazer uma relação entre o ID do produto e sua categoria e trazer, também, qual é a região comprada de algum produto específico, entre outras relações. Mas antes de prosseguirmos, preciso que você entenda, também, sobre chave primária, que é aquela coluna que não tem repetições e normalmente são códigos. Vamos verificar, diretamente, no nosso banco de dados esse conceito. Aqui na "Tabela de Regiões" nós temos a ID da Região que não se repete tem 1, 2, 3, 4, 5 até 10. Já em "Pedidos Geral", a ID Região, se repete. Eu já tenho dois "10" aqui, tenho vários números "9", vários números "4" e assim por diante. Então, aqui dentro de regiões, nós temos uma chave primária que está diretamente relacionada com as colunas adjacentes à ela. Então, quando eu tenho a ID da região 1, significa Brasil/América do Sul. Quando eu tenho 10, significa Inglaterra/Europa. Aqui em Pedido Geral, esse 10 também vai significar Inglaterra/Europa, mas só poderemos utilizar outras colunas de outras tabelas depois de realizarmos um relacionamento entre elas. E é bastante importante você também aprender um outro conceito, que são as "Chaves Estrangeiras", ou seja, se eu tenho uma coluna que não pertence a tabela atual, mas que está lá apenas para descrever melhor algo que aconteceu dentro dessa tabela, ela pode ser considerada como uma chave estrangeira. Vamos ver lá no Access. Eu estou na tabela Pedidos Geral e temos fatos que aconteceram de vendas. E também uma coluna de ID Região. Se eu quiser saber qual foi a região que aconteceu essa venda, eu preciso consultar, diretamente, na Tabela de Dimensão, de Regiões que o número 10 é Inglaterra, como já vimos. A tabela que possui as informações de regiões é essa daqui que se chama "Regiões". Portanto, nós temos uma chave primária que pertence a essa tabela de regiões e eu também tenho essa mesma chave, só que dê uma forma estrangeira, ou seja, de uma forma apenas para complementar os dados que tem dentro dessa tabela que não é uma dimensão, é um fato. E para descrever esse fato eu, simplesmente, relacionam qual é o ID da região e não todas as suas informações. Por exemplo, não seria performático colocar aqui, ao invés de 10, escrever "Inglaterra". Para conceito de banco de dados essa prática vai consumir muito mais espaço. Apareceria os nomes dos países em cada uma dessas linhas e se eu quisesse, também, saber qual é a região ou continente que ela está, também apareceria uma coluna aqui do lado com o nome da região. Ou seja, toda vez que um cliente que estivesse na Inglaterra comprasse, estaria escrito diretamente Inglaterra no banco de dados, e isso não é uma prática executada em sistemas. Você sempre vai encontrar códigos relacionados com códigos de outras tabelas e relacionamentos devem ser feitos entre eles. É por isso que temos essa de relação entre Tabelas Dimensão, que é região e Tabelas Fato. Para não precisar escrever diretamente "Inglaterra, Inglaterra, Brasil, Colômbia" em cada um desses números porque é muito mais fácil, no computador, armazenar números do que textos explícitos escrito com todas as suas letras dessa forma, como está em regiões. Ao invés de duplicar a Inglaterra, colocamos ela em apenas uma linha e depois relacionamos com o seu código, que nesse caso é o número 10. Agora que você entendeu melhor o que são Tabelas Fato, Dimensão, Chaves Primárias e Estrangeiras, vamos abrir o Power BI na próxima aula para fazer o relacionamento entre elas.

Compreenda o motivo de tabelas dimensão e fato

Nessa aula vamos começar a criar os relacionamentos no Power BI, mas antes eu preciso que você entenda o motivo de utilizar tabelas separadas. Muitos alunos me perguntam: "Por que que não utilizamos um relatório sem precisar dividir em cada uma dessas tabelas e criar esses relacionamentos?" E a resposta é que relatórios são diferentes de modelos de dados. Relatórios servem para rápida consulta, onde já temos todas as descrições dos Campos lá dentro e conseguimos fazer, por exemplo, uma tabela dinâmica utilizando apenas um conjunto de dados. Já no Power BI, temos uma capacidade muito incrível de relacionar várias tabelas separadas e utilizar os seus campos de uma forma muito mais performática. Você ainda pode utilizar relatórios dentro do Power BI, mas não estará explorando todas as suas capacidades. A base de dados que vamos utilizar para criar os relacionamentos é parecida com essa tabela daqui. Nós temos IDs de produtos, IDs de regiões com códigos. Eu não sei o que significa a ID região número 10, mas com o relacionamento criado no Power BI, vamos conseguir utilizar campos de outras tabelas para descrever o que é esse ID 10. E normalmente, quando exportamos um relatório do sistema, ele vem dessa forma daqui. Ao invés de IDs, estará, por exemplo, com o nome do país. Tenho aqui Colômbia, Chile... aí consigo, já ver que nesse pedido temos o Chile como país, e o cliente que comprou os produtos relacionados desse pedido. Mas eu preciso que você entenda que quando armazenamos os valores dessa forma, estamos consumindo muito mais espaço e processamento. Ao invés de Brasil poderíamos ter armazenado apenas como "1". Dessa forma, ao

invés de ter 6 letras, estamos apenas com um número, o que deixa o relatório muito mais rápido para fins de processamento computacional. E se quiséssemos saber outras descrições, por exemplo, aqui no relatório 1, também preciso saber a dimensão de qual é o continente que esse país está localizado. No relatório 2 teríamos uma outra coluna que teria América do Sul, Europa, repetido para todos esses países, América do sul novamente para esses, e agora temos um relatório muito mais pesado em termos de processamento, ou seja, requer muito mais do computador para fazer o seu processamento, porque o arquivo, ele é maior em tamanho, em megabytes, e também em complexidade de dados. Então, ter tabelas separadas para cada um desses países e regiões, e aqui na "Pedidos Detalhes", relacionarmos apenas os seus códigos e posteriormente criar um relacionamento de fato entre essas duas tabelas, é o melhor caminho a se seguir. Não estou dizendo que você não pode utilizar no Power BI um relatório como esse, você pode, mas teremos um modelo muito mais pesado. E provavelmente, se esses dados crescerem, você tiver 1 bilhão de linhas, por exemplo, terá problemas de cálculo, eles vão demorar muito para processar. E é por isso que no Power BI, utilizaremos tabelas separadas onde temos um ID da região e, posteriormente, conectaremos com uma tabela de regiões com o ID da região, o seu país, e o grupo, ou seja, o continente que ele pertence. Então te espero na próxima aula para criarmos esses relacionamentos juntos.

Estabelecendo relacionamentos no Power BI

Eu finalizei a aula anterior te mostrando esse projeto no Power BI, que já possui todas as tabelas importadas. Esses dados vieram do Access, que eu também te mostrei em uma aula anterior. Vou abri-lo agora e aqueles dados estão aqui dentro do Access, em cada uma dessas tabelas. Eu te ensinei no capítulo 3 a importar dados de diversas fontes, então você já sabe que para importar esse arquivo lá no Power BI, eu vim em "Obter dados", posteriormente "Mais", e na janela de obter dados, cliquei em "Banco de dados", "Banco de dados do Access". Para focar nos ensinamentos de relacionamento e funções DAX, esse projeto já está preparado para utilizarmos, com todas as importações e tratamentos realizados. Eu tenho várias tabelas segmentadas nesse projeto porque elas vieram de um banco de dados, e normalmente, quando você faz importação para esse tipo de dado, você terá tabelas separadas. Vou clicar em "Cancelar" nessa janela de obter dados, porque eles já foram obtidos e vamos começar com o relacionamento. Você aprendeu em outra aula desse capítulo que relacionamentos possuem cardinalidade, e elas podem ser de um para muitos, representado pelo "1:N". Vamos verificar como que isso é executado dentro do Power BI. Começando com a tabela "Calendário", onde temos uma data e suas descrições, ou seja, uma admissão de dia, de mês, o nome do mês, o nome do dia e outras dimensões, vamos conectar a sua chave primária, ou seja, a data, com uma chave estrangeira em outra tabela. Antes, para entendermos melhor esses conceitos, vamos visualizar os seus dados. Vou clicar nesse ícone de dados, posteriormente, vou selecionar a tabela de data e verificar que nós temos uma data para cada dia do calendário, sem repetições. Então, do dia 1 ao dia 13/07/2020 temos apenas datas únicas para cada um desses dias, ou seja, essa é uma chave primária que não se repete e pertence a essa tabela. Já na tabela de vendas, que é a "Pedidos detalhes", nós temos uma data de pedido que se repete, ou seja, é uma chave estrangeira com valores duplicados, repetidos, e aqui eu já consigo ver que tenho 11/07/2020 e aqui embaixo também. Normalmente, utilizamos tabelas "Calendário", que é essa daqui, para segmentar as nossas análises. Se eu quiser analisar não apenas por data, mas também com um ano, posso escolher esse campo ou então apenas pelo mês, ou então apenas pelo nome do dia e originalmente, lá na tabela "Pedidos detalhes", não temos essas dimensões, tem apenas a data do pedido. Portanto, vamos realizar um relacionamento entre essas duas tabelas para poder realizar diversos tipos de análises utilizando essas dimensões. Para isso, eu vou selecionar a coluna de data, pressionar o botão esquerdo do mouse e soltar em cima da data do pedido. Quando eu solto, automaticamente, o Power BI vai fazer um relacionamento representado por essa linha, de um para muitos. Nós temos um, ou seja, valores únicos dentro dessa coluna, e muitos, representados por esse asterisco, porque temos valores duplicados, ou seja, repetidos dentro da data do pedido e o Power BI já identifica isso automaticamente. O mesmo acontecerá com o exemplo de regiões, temos o "ID região" aqui e podemos conectar com a coluna que se corresponde ao ID da região aqui em "Pedidos detalhes". Essa relação será de um para muitos, automaticamente identificada pelo Power BI, porque aqui dentro dessa tabela, quando vierem dados, nós temos apenas um número para cada ID da região, sem repetições. Essa é a chave primária dessa tabela e na "Pedidos detalhes", temos uma chave estrangeira de ID de região que se repete. E para conseguir avaliar quais são as regiões e países onde essas vendas foram executadas, eu preciso conectar essas duas tabelas. Então, aqui em "Modelo", vou selecionar a ID de região e colocar em cima da

"ID região", da "Pedidos detalhes". O relacionamento foi criado, assim como aquele realizado com essas duas tabelas de cima. Já conseguimos analisar as vendas de acordo com as dimensões do calendário, e também com as regiões e países onde foram realizadas. Vamos continuar com os relacionamentos, agora com produtos. Vou identificar qual é a coluna em comum entre essas duas tabelas para poder verificar quais são os produtos vendidos, as suas subcategorias, categorias e outros campos, como, por exemplo, estilo, a linha do produto, o modelo. Vou tentar identificar isso primeiro aqui na "Pedidos detalhes". Consigo ver que eu tenho uma coluna chamada "ID produto". Aqui em "Produtos", também tenho uma "ID produto", portanto eu vou relacioná-las. A "ID produto" é uma chave primária da tabela "Produtos", e a "ID produtos" da tabela "Pedidos detalhes" é uma chave estrangeira. Agora, se eu quiser relacionar "Pedidos detalhes" com "Subcategoria", preciso verificar se eu tenho uma coluna em comum entre elas, e eu não tenho, porque aqui eu tenho apenas o ID do produto e, dentro dessa tabela, eu tenho o ID da categoria e subcategoria, mas podemos aproveitar o relacionamento de outra tabela, como a de "Produtos", para propagar um relacionamento proveniente de outra tabela. Então eu vou verificar se aqui dentro de "Produtos" nós temos uma coluna que se corresponde com a subcategoria. Então eu vou procurar, aqui tem "ID subcategoria produto". Vou selecioná-la e relacionar com "ID subcategoria produto" da mesma tabela. É muito importante ter campos que são parecidos com os seus nomes para não precisar ficar procurando em cada uma dessas tabelas, vindo aqui em "Dados", e também para se certificar que está conectando com os dados certos. O mesmo eu vou fazer com a subcategoria e a categoria. Nós temos aqui o ID da categoria de produto e, também, embaixo o mesmo nome. Então, vou selecionar, relacionar e temos mais uma relação. Vamos avaliar por quê que o Power BI colocou um para essa tabela e muitos para a tabela de subcategorias. Em "Produto categoria", se eu venho em "Dados" e posteriormente abro essa tabela, teremos bicicletas, componentes, vestuários e acessórios. Dentro da tabela de subcategorias, teremos várias subcategorias que se encaixam na categoria "Bicicleta", ou seja, vou selecionar a tabela "Produtos subcategoria" e nós temos bicicletas montain, bicicletas urbanas, bicicletas touring, todas em uma só categoria, que é "Bicicletas", por isso que foi um relacionamento de um para muitos. O ID de categoria se repete dentro da nossa tabela de "Produtos subcategoria". Vamos continuar com os relacionamentos, vou voltar em "Modelo" e agora falta apenas uma tabela, a "Meta". Vamos começar identificando os campos que são semelhantes entre "Pedidos detalhes" e a "Meta", e é a data. Eu tenho data aqui e data também no pedido. Então, se eu tento relacioná-las, o Power BI vai ter outro comportamento, ele exibe essa janela reclamando aqui embaixo que essa cardinalidade é muitos para muitos, ou seja, são duas colunas com valores repetidos. Aqui na coluna de data eu tenho vários valores iguais, então, aqui em "Dados", vou selecionar "Meta" e consigo ver que eu tenho 01/02, se eu classifico em ordem crescente, tenho vários 1/01/2017. E o mesmo vai acontecer em "Pedidos detalhes" porque aqui na "Data do pedido", eu posso ter mais de um pedido realizado no mesmo dia, e no Power BI nós vamos evitar esse tipo de relacionamento. Portanto, eu pressionei "Cancelar" naquela janela que surgiu e faremos um relacionamento indireto entre a "Meta" e a "Pedidos detalhes" com uma outra tabela que possui, uma coluna parecida, que é a "Calendário". Então, eu vou selecionar a data da "Calendário", relacionar com a data da "Meta" e a relação foi feita sem o Power BI reclamar e exibir mensagens de erro. Isso porque temos uma relação de um para muitos. Nessa aula, nós fizemos todos os relacionamentos necessários entre essas tabelas e nas próximas aulas, vamos continuar explorando quais são as suas consequências e resultados nos cálculos.

6. Introdução à Linguagem DAX

Introdução ao DAX

Agora que fizemos os relacionamentos entre as tabelas, chegou a hora de utilizar funções DAX dentro do Power BI para criar cálculos e posteriormente visualizá-los em gráficos e tabelas. A linguagem DAX é um acrônimo para "Data Analysis Expressions", em português, são expressões de análise de dados, que é uma linguagem utilizada no Power BI para realizar cálculos. É muito importante que você não confunda, nós temos várias linguagens dentro do Power BI. A primeira delas é a Linguagem M, para tratamento de dados no Power Query, que vimos no capítulo anterior. Posteriormente, temos a linguagem DAX, que é utilizada para fazer cálculos e ela não é utilizada dentro do Power Query. Também temos a linguagem R e Python, para tratamento e estatística. Nesse curso, falaremos da linguagem M e DAX, que são as principais dentro do Power BI. Existem diferentes categorias de funções DAX: matemática e trigonometria; filtro; lógica; texto; estatística; hierarquia; informação; data e tempo; e as minhas preferidas, inteligência de tempo, que são funções que fazem cálculos bastante complexos para avaliar o mesmo período do ano anterior, ou então o

mês anterior, ou fazer comparações de períodos muito bem estabelecidos. Vamos utilizar algumas dessas funções aqui, nesse curso. Assim como na linguagem M, caso você queira consultar todas as funções disponíveis no Power BI de DAX, você deve ir diretamente na documentação da Microsoft nesse site, em "docs.microsoft.com", nesse caminho temos a referência de todas as expressões no DAX. Aqui temos as categorias, data e tempo, posteriormente, inteligência de tempo e todas as outras. Também falaremos sobre Syntax, Operadores, Queries, que é o mesmo que consultas, e você pode realizar o download dessa documentação baixando o seu PDF. Recomendo muito que leia toda essa documentação para que tenha um domínio muito maior sobre powered BI e sobre o DAX. Os operadores aceitos nessa linguagem são: o "igual", "maior", "menor", "maior ou igual", "menor ou igual" e "diferente". Você vai notar que o DAX é muito parecido com o Excel, mas você não pode confundir os dois softwares. No Power BI, o DAX possui uma abordagem completamente diferente utilizando contextos, que vamos ver nas próximas aulas. O Excel é bem mais simples. Vamos ver a Sintaxe do DAX. Primeiro, temos um símbolo de "igual", posteriormente o nome da função, e é importante ressaltar que não é case sensitive, ou seja, você pode colocar a função toda em letra minúscula ou maiúscula que também funcionará. Todas as funções do DAX também serão em inglês, não teremos a tradução para português, assim como no Excel. Depois do nome da função, temos a abertura dos parênteses, os argumentos divididos por vírgulas, se possuir, e depois o fechamento do parênteses. Como eu já comentei, é muito parecido com o Excel. Entretanto, tem diversos locais de aplicação. Entre eles: as medidas, as colunas calculadas e as tabelas. As medidas são como se fossem cálculos armazenados para serem utilizados em visuais; as colunas materializam um cálculo dentro da tabela; e as tabelas são tabelas virtuais criadas de dados já existentes, ou seja, daqueles já importados pelo Power Query. Só que dessa vez, essas colunas estão sendo criadas com DAX. Essa foi apenas uma introdução para você entender o cenário da aplicação da linguagem DAX. Nas próximas aulas, vamos explorar cada um deles com mais detalhe.

Utilize funções de agregação

Nessa aula, vamos começar a utilizar funções no DAX para criar cálculos de soma e média. Eu estou com o arquivo "Adventure" aberto, que foi aquele que fizemos todos os relacionamentos entre essas tabelas. Ele também está disponível na pasta de exercícios desse capítulo. E aqui na visão de relatório, nesse botão aqui da esquerda, vamos criar tabelas e gráficos mostrando a soma, a média de vendas e de meta, e faremos isso utilizando funções DAX, especificamente a SUM e a AVERAGE, em português, soma e média. Na última aula, você viu que dá para fazer vários tipos de cálculos com o DAX e podemos encontrá-los aqui na guia "Modelagem", nesses três botões. Então, nós temos uma medida, uma coluna e uma tabela. Nessa aula, vamos criar medidas. Existem outras formas também de criá-las. A outra é clicando em cima de uma tabela que você quer que ela seja armazenada e escolhendo "Nova medida". Quando clicamos aqui, a barra de fórmulas vai ser exibida com o nome da medida antes do "igual", e depois do "igual", temos que colocar a função, que, nesse caso, será uma soma. Então, antes do "igual", vou escrever "Soma", "Vendas" e a escolha dos nomes das medidas precisam ser bem descritivas. Se eu vou fazer uma soma da coluna de vendas, eu vou colocar o nome de "Soma de vendas". Então, depois do "igual", vamos chamar a função SUM. Dentro do DAX nós temos um recurso chamado "Intellisense", isso significa que quando eu digito algumas letras aqui, ele já sugere outras sugestões parecidas, então aqui temos "SUM", "SUMMARIZE", "SUMMARIZECOLUMNS" e "SUMX", mas eu quero a soma. Então eu posso selecionar duas vezes nela e já abre o parênteses obedecendo à sintaxe que vimos na última aula. O argumento da função SUM é apenas um ColumnName. ColumnName em português é o nome da coluna. E nesse caso, eu quero somar a coluna que possui os valores de vendas, e no DAX sempre temos que referenciar a tabela e a coluna. Então, aqui, nesse caso, eu sei que essa coluna está dentro da tabela "Pedidos detalhes", então basta eu começar a escrever "Pedidos" que ele já relaciona todos os campos que tem para eu escolher aqui dentro. Nesse caso, eu quero escolher a Total, que é a coluna que tem os valores vendidos. Ao clicar duas vezes nessa opção, temos o autopreenchimento com a tabela e, posteriormente, com a coluna. O próximo passo é fechar essa função com parênteses no final. Ao pressionar ENTER, teremos a criação dessa medida, mas, aparentemente, nada aconteceu aqui no Power BI. Entretanto, se eu observo aqui nos campos, em "Calendário", eu tenho um cálculo novo representado por essa calculadora, chamado "Soma vendas". O que nós acabamos de fazer foi armazenar esse cálculo dentro de uma medida, e agora ela pode ser utilizada dentro de gráficos e tabelas. Para utilizá-la, posso selecionar e possivelmente arrastar aqui para dentro do relatório; quando eu solto o mouse, temos um gráfico criado com apenas uma coluna, isso porque não temos outras dimensões

filtrando essa medida, ou seja, distribuindo esses valores. Se quiser, eu posso arrastar outros campos para cá. Eu estou na tabela "Calendário", se eu quiser ver as vendas por ano, posso utilizar a dimensão de ano e arrastar para debaixo do eixo. Agora, nós temos uma visão diferente. Conseguimos ver claramente qual foi a distribuição de vendas ao longo dos anos. Em 2017, começamos a vender bem; 2019 melhorou bastante; 2020 ficou muito bem; 2021, não tanto. O gráfico criado automaticamente foi o de coluna, mas podemos escolher outros. Basta clicar em um deles, por exemplo, um gráfico de área, um gráfico de linha, ou então um gráfico de pizza e assim por diante. Estou apenas recapitulando um conhecimento que, talvez, você já tenha. Caso eu queira fazer uma comparação de meta e vendas, agora eu preciso criar uma "soma meta", ou seja, uma outra medida. Então, aqui em modelagem, eu posso pressionar "Nova medida", ou então clicar com o botão direito em "Calendário" e escolher "Nova medida". Aqui, eu vou dar o nome dela de "Soma meta" e, depois do igual, vou colocar "SUM" e somar os valores de meta. Então vou colocar primeiro o nome da tabela para ver quais são seus campos e vou selecionar "Meta (meta)", porque eu sei que aqui nós temos os valores armazenados de meta. Vou clicar duas vezes para auto preencher, fechar os parênteses da SUM e pressionar ENTER. Novamente, aparentemente nada aconteceu no nosso projeto, mas temos uma outra medida aqui do lado, que pode ser utilizada nesse gráfico. Então eu vou selecioná-la e arrastar para debaixo de "Vendas". Quando eu solto, temos uma comparação de vendas versus meta. Para visualizar esses campos aqui na direita, você precisa se certificar que o gráfico estava selecionado antes, então, se ele não estivesse selecionado, os campos não apareceriam, mas, se estivesse selecionado, eles apareceriam. Se eu quiser saber qual é a média de vendas, posso, de novo, criar uma outra medida aqui em "Calendário", vou clicar com o botão direito, "Nova medida" e ela se chamará "Média vendas", e depois do "igual", vou colocar a função AVERAGE. Na sua descrição está falando que ela vai retornar a média aritmética e todos os números de uma coluna. Vou clicar duas vezes para auto preencher, o argumento dela é uma "ColumnName", nesse caso, eu vou colocar uma coluna da "Pedidos detalhes", coloquei apenas a palavra "pedido" para ter essa listagem, vou selecionar a coluna de "Total", que é aqui que eu tenho os valores de pedidos, vou fechar com parênteses da AVERAGE, pressionar ENTER e teremos uma nova medida aqui do lado. Por enquanto, não vamos utilizá-la, só vou deixar armazenada aqui. Colocamos todas as medidas na tabela "Calendário", entretanto, ela poderia estar em qualquer tabela que funcionariam da mesma forma. Apenas para fins de organização, eu vou sempre escolher a tabela "Calendário", porque nem sempre utilizaremos Campos apenas de vendas ou apenas de meta. Então, o meu objetivo aqui é deixá-las todas armazenadas no mesmo local, apenas para rápido encontro. Criamos as primeiras medidas no DAX para você ter uma exposição sobre esse assunto. Na próxima aula, vamos explorar as formas de filtros dessas medidas. Quando colocamos essa medida em um gráfico, ela automaticamente distribui os seus valores pelos anos, ou seja, pela dimensão inserida em eixo. Vamos observar com mais detalhes esse processo.

Compreenda a lógica de cálculo de funções DAX

Nessa aula vamos explorar as lógicas dos cálculos no Power BI, e esse é um dos conceitos mais complexos que temos. Portanto, preste bastante atenção. Todo o cálculo realizado, seja ele em uma medida, uma coluna, ou em uma tabela, utiliza contextos, que é como se fossem os cenários para o cálculo realizar os seus procedimentos. E todo o cálculo é realizado por meio de filtros no cenário que é inserido. Na última aula nós fizemos uma Medida de Soma Vendas e colocamos dentro de um gráfico de coluna. Quando inserimos um ano dentro desse mesmo visual, a medida foi redistribuída, ou seja, conseguimos ver para cada um dos anos quanto que vendemos, que podemos dizer que foi um filtro aplicado dentro do visual. E dentro do Power BI, temos dois deles: o contexto de filtro do visual, que é esse que acabei de descrever; e o contexto de filtro de linha, que veremos na próxima aula, quando criarmos uma coluna calculada. Apenas para fins de introdução, dentro de medidas calculadas, elas utilizam sempre contexto de filtro do visual, que são calculadas de acordo com os Campos e Filtros inseridos em visuais ou argumentos, dentro de suas funções. É importante ressaltar, também, que ela vai calcular todos os valores da coluna sem identificar individualmente suas linhas. Ou seja, ela não consegue ver cada uma dessas linhas, porque ela não está inserida no contexto da tabela, e sim apenas no contexto de filtro do visual. Quando fizemos a Medida Soma Vendas, isso é uma meta, inserimos dentro da tabela calendário. Mas não tem nada a ver inseri-la lá, mas isso é porque a medida não precisa estar necessariamente na tabela dos Campos que ela relaciona, porque ela ainda não contém um contexto, ele é criado apenas dentro do visual. Já as colunas calculadas são diferentes, elas utilizam o contexto de Filtro de Linha, ou seja, são calculadas de acordo com cada linha da tabela, porque quando criamos uma coluna, ela é materializada dentro da tabela, ou seja, possui um contexto, que é a tabela de

suas linhas. Resumindo, as influências de contextos de Filtro de visual e medidas são os Campos dentro do visual, por exemplo, o ano que inserimos dentro do gráfico de coluna, os segmentadores de dados, que são filtros, os filtros aplicados no painel de filtros e outros visuais. Quando clicamos, por exemplo, em um visual na mesma página, ele é capaz de filtrar os outros ao seu redor, caso tenha alguma relação. Já o contexto de Filtro de linha em colunas, cada linha é individualmente calculada por causa do contexto que ela está inserida, ou seja, dentro da tabela. E os filtros de visuais e segmentadores não influenciam no cálculo dentro da tabela. Já vimos o comportamento de como uma medida é calculada e, na próxima aula, veremos como que uma coluna executa seus cálculos para entendermos, de uma vez por todas, essa prática de contextos de Filtro de linha e de Visual. Como falei no início da aula, é um dos conceitos mais complexos dentro do Power BI. No início do meu aprendizado eu demorei mais ou menos umas duzentas a trezentas funções criadas dentro do Power BI, entre os diferentes projetos de testes que eu estava realizando para compreender 100% da sua atuação. Portanto, você precisa praticar tudo o que vamos ensinar aqui e aplicar também em projetos de teste para, realmente, solidificar esse conhecimento.

Criando uma coluna calculada

Utilizando o mesmo projeto das últimas aulas, nessa vamos criar uma Coluna Calculada para compreender melhor os contextos de Filtro de visual e de linha. Eu vou clicar em dados e depois vou selecionar a Pedidos Detalhes que, na verdade, já está exibida aqui. A minha necessidade é calcular qual é o valor de desconto por cada um desses preços unitários. Por exemplo, aqui temos 10%, eu quero multiplicar com esse valor para encontrar qual foi o valor de desconto. Como aprendemos na aula anterior, as colunas calculadas conseguem identificar cada uma dessas linhas, já as medidas não conseguem porque elas não estão inseridas nesse contexto, elas simplesmente ficam armazenadas dentro dessas calculadoras, e só são utilizadas quando são inseridas dentro de visuais. Mesmo se tentássemos criar esse cálculo dentro de uma medida não conseguiríamos. Vamos tentar. Aqui em modelagem vou clicar em Nova Medida, e vou chamar de Valor Desconto. Depois do igual, vou tentar fazer uma multiplicação entre o preço unitário e a Desconto. Então vou começar a escrever "pre" de preço unitário, que é o nome da coluna, e aqui no Intellisense ela não me relaciona qual é a coluna que eu posso buscar. Quando o Intellisense não identifica automaticamente essa coluna quer dizer que tem alguma coisa errada, e nesse caso tem: é uma regra que as medidas sempre precisam ser relacionadas com uma função que faz algo com a coluna que estamos utilizando. Por exemplo: se eu coloco "sum" e depois coloca "pre", essa coluna de preço unitário é relacionada. Então essa é uma dica que eu te dou, se no Intellisense do DAX, quando você digita alguma coisa, ele não sugere aqui embaixo, é porque você está fazendo no local errado, ou utilizando o cálculo incorreto, que, nesse caso, nós estamos. Eu só conseguiria fazê-lo da forma como estava fazendo antes, somando toda essa coluna, por exemplo, chamando a preço unitário aqui, fechando parênteses, e depois multiplicando pela coluna de desconto. E eu não consigo colocá-la aqui dentro dessa função, porque a "sum" só tem um argumento, que é uma column name, ou seja, um nome de coluna. Portanto, eu teria que multiplicar externamente também. Fazer um vezes sum desconto. E aí eu posso selecionar a coluna desconto e fechar. Esse cálculo não faz o menor sentido, porque nós estamos primeiro somando toda essa coluna, posteriormente, somando toda essa coluna, e multiplicando ambas. Tudo isso porque dentro de uma medida não conseguimos chamar apenas a coluna sem cápsula-la dentro de uma função. Eu vou pressionar ENTER para finalizar essa função e armazenarmos ela aqui dentro da tabela Pedidos Detalhes, ela está aqui em formato de calculadora, mas ela não foi materializada, ou seja, ela não está presente visualmente aqui nessa tabela. Eu tenho aqui da coluna e da região, até ia de pedido, ela não está aqui, porque medidas não criam colunas, são utilizadas apenas dentro de visuais. Para solucionar esse problema temos que criar uma coluna calculada, que pode ser acessada aqui na modelagem, nova coluna; ou então clicando aqui nesses três pontinhos, na Pedidos Detalhes ou com o botão direito e escolhendo Nova Coluna. Vou selecionar essa opção, uma nova coluna foi materializada dentro dessa tabela. Essa já é uma diferença que temos das medidas e também foi criado esse símbolo de coluna aqui dentro da tabela. Eu vou modificar o seu nome ao invés de chamar de coluna será Valor Desconto 2, porque já temos uma medida criada com o nome Valor Desconto, então tô colocando diferente. E aqui depois do igual, eu vou colocar "pre" para tentar chamar o Intellisense do DAX e ele sugerir essa coluna, e nesse caso ela parece, o que não estava acontecendo na medida. Isso porque nós estamos no contexto de Filtro de linha, ou seja, estou dentro de uma tabela criando essa coluna. Vamos selecionar duas vezes para auto preencher, posteriormente vou tentar fazer uma multiplicação com a coluna desconto também escrevendo seu nome ela também relacionada aqui no

interessantes e vou clicar duas vezes, posteriormente pressionar ENTER e temos a coluna criada dentro do contexto da tabela com o valor de desconto. Então esse resultado é 119, arredondando, e se eu vou no seus valores, o preço unitário é $1192 \times 10\%$. O valor está correto. Se observarmos agora aqui dentro dessa tabela, nós temos uma medida chamada Valor Desconto, criada anteriormente, e depois uma coluna calculada identificada por esses símbolos. Se colocarmos ambas dentro de um visual, teremos resultados completamente diferentes. Eu vou fazer isso agora. Aqui é um relatório, nesse ícone. Vou criar uma nova página, clicando em mais, e depois vou selecionar a Valor Desconto, trazer aqui para o visual. Depois a Valor Desconto 2, que é a coluna calculada, trazer para o mesmo visual, e vou modificar para uma tabela, para conseguir ver os números por extenso. Como está bastante pequeno esse visual, vou aumentar um pouco a sua fonte. Vou selecioná-lo, clicar nesse rolinho de tinta que é o formato, posteriormente grade, e selecionar tamanho. Aumento um pouco, temos para o Valor Desconto, que é a medida, um valor muito grande. O Valor Desconto 2, que é a coluna calculada, o valor de 130.821,08, que é o correto porque faz uma multiplicação para cada uma das linhas que temos de preço unitário vezes o desconto. Quando colocamos no visual, é realizada uma soma desses valores. Consigo ver que é uma soma porque aqui está escrito soma. Já a valor desconto, está somando todos os valores de preço unitário, depois todos os descontos, e fazendo uma multiplicação. Por isso que fica esse valor tão alto. Portanto, essa é a diferença entre a criação de uma medida, e de uma coluna, e as suas diferenças entre o contexto de linha, de uma coluna, e ele existe porque essa coluna é criada dentro de uma tabela, portanto ela tem esse contexto daqui. Já a medida não possui contexto, porque ela só é utilizada quando colocamos dentro de visuais e, apesar de estar armazenadas dentro da tabela Pedidos Detalhes, ela não considera esse contexto. Na verdade nunca vai considerar. É importante ressaltar que conseguimos fazer com que uma medida tenha o comportamento de uma coluna calculada, só que esse é um assunto mais avançado e será tratado em outra aula desse curso.

Realize uma divisão segura com a DIVIDE

Medidas podem ser reaproveitadas dentro de outras medidas. Em uma aula anterior nós fizemos uma medida chamada Soma Meta e outra Soma Vendas, e depois colocamos dentro desse visual. É quase imperceptível algumas diferenças, por exemplo, em 2017, 2018 e 2001, em relação às vendas e à meta, seria muito melhor dividir esses dois valores para verificá-los percentualmente. Para isso vamos criar uma medida com uma função chamada "divide". Vou clicar com o botão direito em cima de calendário e escolher nova medida. Essa nova medida se chamará "Vendas versus Meta". Depois do igual vou colocar a função divide e verificar qual é a sua descrição, ela fará uma safedivide com capacidade para tratar casos de divisão por zero. Então esse safe divide é como se fosse um Se erro. Se essa função e o seu cálculo der erro, então o valor alternativo será exibido, e ela perde três argumentos: um numerador, um denominador, e um valor alternativo. Onumerador serão as vendas, então eu vou chamar a medida "Soma Vendas". Quando eu coloco apenas soma ela já se relaciona aqui em baixo, vou clicar duas vezes em cima dela, e ela foi auto preenchida. Todas as medidas referenciadas dentro de outras medidas são exibidas nesta cor roxa. Agora eu vou colocar um ponto e vírgula para irmos para o segundo argumento, que é o denominador, que serão as metas. Então vou colocar Meta, clicar duas vezes nessa sugestão do Intellisense. Agora eu vou preencher o último argumento, que é o valor alternativo. Então vou colocar um ponto e vírgula e quero que, se tiver um caso de divisão por zero, seja de um numerador ou no denominador, que retorne zero. Poderia, também, retornar Blanc, que é essa função daqui que retorna um espaço em branco, ou seja, nada. Posso clicar duas vezes nela, ela não tem argumento, por isso abre parênteses e fecha automaticamente. Vou colocar um parênteses no final para fechar a divide, pressionar ENTER para a criação dessa medida, e temos a criação aqui. Vou selecionar esse gráfico, colocá-lo na esquerda e trazer essa medida para dentro de um novo gráfico. Um gráfico de colunas foi criado, vou selecionar o campo Ano, trazer para Eixo, e temos a distribuição da "Vendas versus Meta" por ano. Mas eu acho que ainda não está bom, não é possível ver tanta variação ao longo dos anos, eu posso colocar o rótulo de dados para ficar mais evidente, selecionando esse visual, clicando em formato e, posteriormente, rótulo de dados. Eu consigo ver que aqui está 1,01, aqui também está 1,00, e assim por diante. Eu vou aumentar um pouco esse rótulo de dados apenas para conseguirmos visualizá-los melhor. E agora eu quero fazer uma variação de Vendas versus Meta para explicitar melhor essas diferenças, ou seja, eu vou diminuir desses valores e resultados 1. Eu vou retirar 1 inteiro dessa fração, 1 inteiro também dessa, aí assim por diante. Os valores de metas que não forem atingidos ficarão como negativos. Para editar uma medida, vou selecionar a Vendas versus Meta novamente, teremos ela aqui na barra de fórmulas, e aqui no final do parênteses, vou colocar um -1 para retirar 1

inteiro. Ao pressionar "OK", teremos uma visibilidade muito maior dessa variação. Melhor do que esse gráfico, naquele anterior utilizando essa mesma medida, e agora exibindo a sua variação de fato. Para formatar uma medida, por exemplo, eu quero mostrar o percentual aqui, eu posso selecioná-la e depois vim na guia modelagem, e selecionar esse percentual. Ao clicar nessa opção teremos 1,44 - 0,17 e assim por diante, todos em percentual. Portanto, medidas podem ser reaproveitados dentro de outras medidas. Você também aprendeu nessa aula a como utilizar a divide que faz uma divisão segura caso esteja fazendo de visões por zero. Você também aprendeu como exibir melhores gráficos dentro de um visual, com um resultado muito mais representativo para facilitar a análise.

Introdução à função CALCULATE

Nessa aula, você vai aprender sobre uma das funções mais importantes no DAX: a "Calculate". Utilizando o mesmo projeto das últimas aulas, aqui nós temos uma tabela com o nome da categoria e a soma vendas, que é aquela medida que criamos. Utilizando o SUM do "Pedido detalhes total", e agora eu preciso calcular quanto que "Acessórios" representa de componentes; quanto que "Bicicletas" representa de componentes; "Vestuário", de componentes, ou seja, eu preciso fazer uma divisão desse valor daqui, 11 milhões, sobre todos esses. Mas existe um conceito muito importante que você precisa aprender no DAX. Cada uma dessas linhas são filtros individuais, então eu não consigo fazer uma relação entre esse valor e todos esses outros; preciso, necessariamente, colocar esse valor de 11 milhões ao lado de cada uma dessas linhas daqui para poder fazer essa divisão e realizar a comparação entre essas categorias. Para isso, precisamos alterar o contexto de filtro. Quando colocamos o nome da categoria no visual, ele se torna um filtro e está filtrando a medida "Soma Vendas" de acordo com essa linha; o mesmo acontece com todas as outras. Portanto, o filtro natural dessa linha é mostrar as vendas de acessórios, mas para fazer um cálculo entre um milhão e 11 milhões, precisamos colocar 11 milhares aqui do lado e é a função "Calculate" que vai nos ajudar. Então eu vou clicar com o botão direito em cima de "Calendário", "Nova medida" e aqui nessa medida, vou colocar o nome de "Vendas componentes". Depois do igual vou chamar a função "Calculate" e vamos ler a sua descrição. Ela vai avaliar uma expressão, que é o seu primeiro argumento, e podemos colocar aqui um cálculo, uma medida ou então uma função e vai fazer essa avaliação em um contexto modificado por filtros; e esse filtro, vamos explicitamente colocar dentro da função, neste segundo argumento. Então, a expressão que eu quero calcular de uma forma diferente é a "Soma Vendas", quero mostrar os valores de vendas de componentes, mas de uma forma filtrada para cada uma dessas linhas. Então vou colocar a medida aqui dentro, "Soma Vendas", posteriormente, colocar um ponto e vírgula para preencher o filtro que será o nome "Categoria" que é essa coluna daqui quando ela for igual - coloquei um símbolo de igual - a componentes; então eu vou colocar entre aspas esse texto, porque ele é um texto então "Componentes". Fecho aspas, vou fechar parenteses da "Calculate" e pressionar ENTER. Basicamente, o que essa função está fazendo é o seguinte: calcular a soma de vendas, de acordo com esse filtro daqui, independente do que tiver dentro dessas linhas. Então, ao colocá-la dentro desse visual, vou puxar a "Venda componentes" e colocar dentro da tabela. Quando eu solto, teremos 11 milhões em todas essas linhas, ou seja, alteramos o filtro natural desse visual. E agora conseguimos fazer o cálculo necessário; dividir esse valor por esse, esse por esse e assim por diante. Esse cálculo já sabemos fazer, podemos utilizar uma função chamada "Divide". Então aqui em "Nova medida" vai se chamar "Comparação Componentes" e vou chamar a função "Divide", que já fizemos na última aula, o numerador será a "Soma Vendas", que serão todos esses valores no visual; vou colocar um ponto e vírgula para preencher o denominador que será essa medida de vendas componentes. Não vou preencher o valor alternativo porque ele é opcional, está mencionando aqui nesse colchetes; então sempre que tiver um colchete, eu não preciso preencher. Coloco um parênteses no final para finalizar a função "Divide", pressionar ENTER e ela acabou de ser criada. Seleciono e arrasto para o visual, temos o resultado. É melhor formatá-la com um percentual para visualizarmos melhor, vou selecioná-la e, posteriormente, em "Modelagem", clicar no símbolo de percentual. Portanto, a única forma de fazer uma divisão entre esses valores de linhas diferentes é inserindo-os na mesma linha, que foi o que fizemos com a função "Calculate", que alterou o contexto natural dessas linhas. Essa daqui, o contexto natural era "Acessórios"; na soma de vendas, ela naturalmente mostraria um milhão 271 mil, mas, quando utilizamos a "Calculate", que é uma função que altera o contexto natural do filtro, exibiu 11 milhões, porque foi isso que colocamos em seu argumento. Não precisamos exibir essa coluna, estou exibindo apenas para fins didáticos; então posso removê-la do visual, clicando em "Vendas componentes" no "X". Temos a comparação necessária de ser realizada utilizando a "Calculate" e a

"Divide"". Você acabou de aprender sobre uma das mais importantes funções no DAX. Nas próximas aulas utilizaremos a "Calculate" o tempo inteiro.

Função ALL e ALLSELECTED

Nessa aula, você vai aprender a utilizar as funções "ALL" e "ALLSELECTED" que removem filtros realizados dentro e fora de visuais. A necessidade desse projeto, que é o mesmo que estamos utilizando nas últimas aulas, é calcular a representatividade de cada uma dessas categorias sobre o total, ou seja, preciso dividir esse valor, do valor do total e assim por diante para cada uma dessas categorias. Para isso, você aprendeu na última aula, da "Calculate", que precisamos alterar o contexto de filtro e mostrar esse total ao lado de cada uma dessas categorias para poder calcular o seu percentual. Portanto, já sabemos que vamos utilizar a função "Calculate"; agora, precisamos saber como referenciar o total que não é um campo dessa tabela, na verdade, ele é apenas um valor inserido aqui dentro do visual de tabela, mas ele não pertence a tabela de categorias. Com essa análise, vamos precisar utilizar uma função chamada "ALL". A "ALL" vai remover todos os filtros, ou seja, "Acessórios" é um filtro aplicado para essa linha; "Bicicleta" também. Ao remover todos os filtros, teremos o total da coluna que será 109 milhões, nesse caso. Vamos ver na prática todos esses conceitos. Eu vou clicar com o botão direito em cima da "Calendário", escolher a "Nova medida" e vou chamá-la de "Todas as vendas". Vou colocar um igual e, posteriormente, chamar a função "Calculate". Os argumentos da "Calculate" são: uma Expressão, que pode ser uma medida, um cálculo ou uma função; nesse caso será a "Soma Vendas" porque queremos calcular esse valor de total, mas com um filtro diferente. Então vou colocar "Soma Vendas", posteriormente, ponto e vírgula, para o segundo argumento que é o "Filtrar" e aqui vamos utilizar o "ALL". A função "ALL" vai retornar todas as linhas de uma tabela ou todos os valores de uma coluna, ignorando todos os filtros que estiverem aplicados. Esse daqui é um filtro, este aqui também; então a função "ALL" vai ignorá-los. Essa função me pede uma tabela ou uma coluna. Eu vou referenciar a tabela de "Pedidos", que é onde temos todos esses valores de vendas; então aqui dentro vou chamar a "Pedidos Detalhes" e fechar a função "ALL". Posteriormente, eu vou fechar a "Calculate", pressionar ENTER e teremos a medida criada. O que estamos pedindo para o Power BI fazer é calcular essa soma de vendas, que são esses valores aqui, retirando todos os filtros, ou seja, sem considerar cada um deles. Quando colocarmos aqui dentro desse visual, vai mostrar apenas o valor de total; então, vou selecionar essa medida, "Todas as vendas", trazer para a "Tabela" e agora temos o total em cada uma dessas linhas, e agora conseguimos fazer uma divisão entre "Soma Vendas" e "Todas as vendas" e achar o percentual de cada uma dessas categorias. Então aqui em "Calendário" vou clicar com o botão direito, "Nova medida" e agora vai se chamar "Repres", de "Representatividade", abreviado, "Categorias"; depois do igual vou chamar a função "Divide", que já aprendemos em outra aula, que faz uma divisão com o numerador e o denominador; o numerador vai ser a "Soma Vendas", ponto e vírgula para o segundo argumento, o denominador vai ser essa medida, que é a "Todas as vendas". Selecionei ambas, vou fechar com um parênteses da direita. Vou pressionar ENTER; não preciso preencher o último argumento da "Divide" porque ele alternativo tem esses colchetes então não vou preencher e agora eu vou trazer essa medida aqui para dentro do visual. Temos o resultado em decimal. Vou formatá-lo para percentual, para ter uma visualização melhor; então selecionei a medida, vou clicar em "Modelagem" e depois "Percentual" e agora sabemos que "Acessórios" representa 1,16% do total; "Bicicletas", 86% do total e assim por diante. E agora eu vou te mostrar um problema dessa função. Eu vou colocar um visual de filtro com os países aqui na esquerda, então os países estão na tabela "Regiões", vou trazer o país para cá, fora do visual de tabela; ele automaticamente criou um mapa, mas eu não quero mapa, então vou selecionar esse visual em modificar para "Segmentador de dados", nesse ícone. Vou aumentar também um pouco seu tamanho, porque está muito pequeno. Então aqui em "Geral", na verdade, aqui em "Itens", vou aumentar o tamanho do texto e se eu selecionar agora o "Brasil", o total mudou ficou, 16 milhões, mas a medida "Todas as vendas" permanecem com o mesmo valor. Isso porque a função "ALL" vai remover todos os filtros internos e externos. Esse daqui está externo ao visual e os filtros internos é cada uma dessas categorias. Se você quisesse considerar filtros realizados externamente, por exemplo, buscar esse valor daqui de 16 milhões e fazer essa comparação de percentual com ele, precisamos usar uma outra função chamada "ALL SELECTED". Para isso, eu vou selecionar essa tabela, pressionar CTRL+C e CTRL+V apenas para criar uma cópia, e vou retirar esses últimos dois campos: "representação categorias" e "todas as vendas". Agora vou criar uma outra medida que vai utilizar a "All Selected". Então aqui em "Calendário" vou clicar com o botão direito, "Nova medida", o nome da medida se chamará "Todas selecionadas", que indica que serão todas as

vendas selecionadas; vou utilizar a "Calculate" para alterar o contexto de filtro desses visuais; a expressão será a soma de vendas, ponto e vírgula. O filtrar será a "All Selected". Ao invés de utilizar a "All", vou selecionar essa daqui que vai retornar todas as linhas de uma tabela ou todos os valores de uma coluna, ignorando também os filtros aplicados dentro da consulta que são esses aqui, porém, mantendo todos os filtros que vierem de fora, que é esse segmentador aqui. Essa segunda frase é o que a "All" sozinha não faz, então precisamos usar a "All Selected". Vou colocar a tabela "Pedidos detalhes", igual eu coloquei aqui em cima, fechar a "All Selected" com seu parênteses, fechar também a função "Calculate" com o seu último parênteses e pressionar ENTER para criar essa medida. Agora vou selecioná-la e colocar dentro da tabela; temos os 16 milhões, que é exatamente o que tem aqui no total. Já na tabela de cima temos 109 milhões. Se eu seleciono agora "Canadá", ele automaticamente vai mudar o "Todas selecionadas", que essa medida, para o total das vendas, de acordo com os filtros externos aplicados. Então se fizermos uma divisão entre a "Soma Vendas" e a "Todas selecionadas", ao invés de dar 7%, igual esse que deu, dará sempre 100%. Então aqui em "Calendário", vou clicar com o botão direito, "Nova medida", essa medida se chamará Representatividade, abreviado, Categorias selecionadas e aqui no igual vou colocar um "Divide", que é a função que divide; o numerador será "Soma Vendas", o denominador será "Todas selecionadas", que é essa medida daqui. Vou fechar o parêntese da "Divide" e apertar ENTER, Selecionar a medida, aqui, na tabela "Categoria" e inserir dentro do visual. Então vou selecioná-lo, inserir no visual de tabela, temos a representatividade aqui dentro; vou diminuir um pouco essa coluna; selecionar a medida e aqui em "Modelagem" colocar como "Percentual" para vermos o mesmo formato que a de cima. Então aqui temos 100% e agora sabemos qual é a representatividade de bicicletas sobre total selecionado, inclusive, aplicando aos filtros. Nesse caso, a representatividade de bicicleta sobre esse total não está correto aqui, está 6%; quando, na verdade, é 85% sobre o total considerando esse filtro. Você acabou de aprender nessa aula a como utilizar a "Calculate" juntamente com outras duas funções, que a "All" e a "All Selected" para manipular filtros dentro e fora de tabelas. Eu exibi essas colunas aqui apenas de forma didática para você visualizar o que está acontecendo; elas não precisavam estar exibidas, então eu posso retirar a seleção dela retirar a seleção da de cima, "Todas as vendas" e agora temos apenas os resultados das representações das categorias. E aqui em baixo apenas das seleções, incluindo os filtros aqui de fora. Estamos criando todos esses cálculos e aplicando em visuais de tabela, mas eles funcionariam em qualquer outro tipo de visual aqui. O que eu quero que você entenda é apenas o contexto de alteração do filtro. Além de utilizar a "Calculate", também alteramos o filtro que cada uma dessas linhas.

7. Cálculos avançados com DAX

Introdução às funções iterantes

Na aula chamada criando uma coluna calculada vimos que não é possível utilizar uma medida para calcular valores individuais em colunas. Isso porque uma medida não possui o contexto dessa coluna, ela simplesmente é armazenada dentro dessa calculadora e só é utilizado em seus valores materializados quando inseridos em visuais. Já uma coluna calculada, que foi o que fizemos para solucionar esse problema, consegue calcular cada uma dessas linhas porque ela tem o contexto da tabela, ou seja, é materializada aqui dentro e temos os valores exibidos para cada uma dessas linhas; entretanto, é possível fazer com que uma medida tenha um comportamento de coluna criando o seu contexto dentro dela e obtendo esse mesmo resultado, mas sem materializar os valores. O nosso objetivo nessa aula é eliminar essa coluna obtendo o mesmo resultado de 130.000, mas dessa vez com uma medida. Esse valor da esquerda está errado, foi uma tentativa de criação de uma medida para multiplicar o preço unitário com desconto, mas como mencionei, uma medida não possui o contexto da tabela. Mas agora vamos criá-lo. Para isso na tabela pedidos detalhes, vou clicar com o botão direito e criar uma nova medida. Essa medida vou chamá-la de valor desconto 3, que é a nossa terceira medida sobre esse mesmo assunto. Quando utilizamos a função SUM, eu não sei se você percebeu que temos uma SUMX aqui no final que é bastante peculiar, ela vai retornar a soma de uma expressão avaliada para cada linha de uma tabela e se selecionarmos ela, os seus argumentos são: tabela e expressão. O primeiro argumento dela é a criação do contexto, então aqui vamos informar a tabela em que quero criar esse contexto que é pedidos detalhes. Então vou inserir pedidos detalhes; Ponto e vírgula para irmos para a expressão. Então dentro dessa tabela quero fazer uma multiplicação entre a coluna preço unitário, vezes a coluna desconto. Cliquei nas duas para auto preencher, vou fechar a função SUMX e dizemos que é uma função da categoria iterante porque ela vai verificar cada uma das linhas que tem dentro dessa tabela e vai realizar esses cálculos para cada linha que possuir aqui dentro, ou seja, vai

iterar por cada uma dessas linhas até realizar o cálculo de todas e posteriormente fará uma soma dos valores encontrados, mas tudo isso internamente nessa função. Pressionar ENTER depois dos parênteses, selecionar a medida nova e trazer para esse visual. Olha o valor que encontramos: 130.821, exatamente igual ao valor desconto 2 que é a nossa coluna calculada. Consigo ver pelo símbolo, já que aquela primeira medida que tentamos fazer utilizando apenas uma função normal de soma resultou em mais de 19 bilhões e estava incorreto, portanto não precisamos mais dessa coluna de valor desconto 2. Se você está se perguntando por que temos dois valores e aprendemos uma função extremamente complexa para resolver esse problema, o motivo é o seguinte: toda vez que materializamos valores dentro de tabelas como o valor de desconto 2, estamos usando mais processamento do Power BI e do computador; então a melhor prática é não utilizar esse tipo de coluna calculada quando podemos substituir o seu cálculo por uma medida. Em alguns momentos você vai precisar de uma coluna calculada, mas em outros, como esse, podemos criar uma medida. Portanto vou selecionar a medida valor desconto 2 e excluir essa coluna do modelo; excluo novamente e, quando voltar no visual, ele estará com um erro porque usa uma coluna que não existe mais. Vou clicar no X e temos o valor anterior. Esse valor é o errado e esse valor é o certo utilizando a função iterante SUMX. Conceitualmente falando, na medida valor desconto 3 fizemos uma alteração do contexto de filtro. O normal de medidas é fazer o cálculo de toda a coluna sem identificar cada uma das linhas, assim como aprendemos. Mas em funções iterantes, como a SUMX, a medida cria um cálculo diferente observando cada uma das linhas do contexto inserido. Existem muitas outras funções iterantes. Pode ver clicando nova medida apenas para mostrar algumas. Então nós temos a MAXX, que também pede uma tabela no início; e sempre quando uma função pede uma tabela, é iterante. Temos também a FILTER, que pede uma tabela no início. A AVERAGEX, que também pede uma tabela no início, e muitas outras.

Crie tabelas de parâmetros

Nessa aula vamos utilizar alguns conceitos aprendidos em aulas anteriores para depois fazer uma tabela de parâmetro, criar uma variação de custo e depois descobrir o seu lucro a partir da variação inserida no parâmetro. Vamos começar primeiro calculando o custo. Vamos avaliar que na tabela de produtos tem uma coluna chamada custo padrão e na tabela pedidos detalhes temos outra coluna chamada quantidade, para fazer o cálculo do custo para cada uma dessas categorias precisamos multiplicar a quantidade que está impedido os detalhes com um custo que está em outra tabela, e também sabemos que é preciso ver cada uma dessas linhas porque é preciso ponderadamente fazer o cálculo entre as quantidades dos produtos comprados e também o seu custo respectivo. Utilizaremos uma função iterante chamado SUMX que consegue criar um contexto de tabela onde vamos referenciar ambas. Em modelagem crio uma nova medida que se chamará soma custo, que será igual a uma SUMX. O primeiro argumento da SUMX é uma tabela. Vou referenciar a pedidos detalhes, que é onde tenho uma dessas colunas para usar. Posteriormente, ponto e vírgula para determinarmos qual será expressão, que é a quantidade vezes o custo padrão. Entretanto, quando tento puxar a custo padrão escrevendo CUS, que é o início de custo, ela não aparece e, como mencionei em outra aula, é uma dica importante para todos os cálculos realizados no DAX, quando Intellisense, ou seja, essas seleções não sugerem as colunas que estamos buscando significa que tem algo errado e, nesse caso tem, pois temos uma coluna que está em outra tabela, ou seja, não temos o contexto dela inserido na função, na verdade precisamos buscar a sua relação. Para esse cálculo funcionar, utilizar outra função chamada related, já que estamos com outra tabela sendo relacionada aqui dentro, utilizar a função related, que vai exibir todas as colunas relacionadas com pedidos detalhes, que foram aquelas relações que criamos nesse botão daqui em uma aula anterior. Ao tentar escrever o custo padrão, ela é buscada com sucesso por meio da relação criada entre pedidos detalhes e produtos. Selecionar duas vezes em produtos custo padrão, fechar os parênteses da related, fechar os parênteses da SUMX e acabamos de criar um contexto de filtro. Observando cada uma das linhas dessa tabela “pedidos detalhes” e cada uma de suas quantidades foram multiplicadas pelos custos relacionados na tabela produtos, na coluna custo padrão. Com a função finalizada, vou selecionar o último parênteses, pressionar ENTER e teremos a criação da medida. Selecionar e colocar no visual e sabemos que o custo de acessórios é 635.000 e assim por diante. Vou formatar essa medida para ficar no mesmo formato anterior. Então vou selecioná-la e, em moeda geral, vou transformar para número decimal. Agora fica muito mais fácil de entender que essa empresa tem um custo de 100.000.000 para todas as vendas que ela já executou quando vendeu 109.000.000, ou seja, o lucro é mais ou menos de 9.000.000. Vamos calculá-lo. Esse cálculo é bem simples, basta diminuir ou subtrair as vendas do custo e para isso vou criar uma nova medida. Em modelagem, nova

medida, que se chamará lucro, que será igual às vendas, menos a soma custo. Podemos utilizar o operador de menos da matemática para realizar esse cálculo. Assim como mais para somar; a barra para dividir e o asterisco para multiplicar. Pressionar ENTER para finalizar essa medida, selecioná-la na tabela, colocar no visual e temos o lucro para cada uma dessas categorias. Gerar uma tabela de parâmetro para criar a variação que aplicaremos na soma custo para calcular o lucro de acordo com o cenário escolhido pelos usuários desse relatório. Para criar essa tabela de parâmetro, aqui em modelagem, vou clicar em novo parâmetro e nessa janela vou preenchê-lo vou deixar o nome de parâmetro o mesmo, mas pode alterar para aquele que mais se adequar à necessidade que está desenvolvendo. O tipo de dados será número decimal porque vou colocar percentuais. Então o mínimo será 0,001 e o máximo será 0,09. Vou incrementar de 0,001 e o padrão será. 0, deixar essa opção marcada, pressionar OK. Temos esse parâmetro e quando posiciono o mouse em cima e altero esses valores, temos de 0,01 até 0,09. Vou clicar em dados para verificar que esse parâmetro gerou uma nova tabela e é uma tabela virtual dentro do DAX, que antes não existia e foi criada por essa funcionalidade, vai de 0,01 até o valor que colocamos, que foi 0,09. Pode formatar essa coluna, mas vou deixá-la na ordem crescente, selecionando em modelagem colocar como percentual e temos de 0,1 até 8,9. Vamos supor que as variações de custo que serão inseridos, o máximo a 0,90. e o mínimo é 0,01. Voltar no relatório, quando selecionamos esse parâmetro nada acontece ainda neste visual, apesar de tentar fazer uma atualização, mas nenhum valor é alterado. Na próxima aula vamos fazer uma relação entre os valores do parâmetro com os valores da soma custo.

Como criar medidas dinâmicas

Na aula anterior criamos um cenário com "soma vendas", "soma custo" e o lucro, e mais um parâmetro que ao ser alterado precisa modificar essa medida daqui, de "soma custo" para, conseqüentemente, alterar o lucro. Para realizar essa influência, vamos utilizar uma medida que é criada automaticamente quando selecionamos um novo parâmetro. Procurar pela tabela de parâmetro e dentro temos uma medida chamada "Parâmetro Valor" que foi criada automaticamente ao usar essa funcionalidade. Dentro dessa medida, quando selecionei na barra de fórmulas, temos uma função chamada "selected value", da "parâmetro - parâmetro". Ou seja, da tabela que se chama "parâmetro", da coluna que também tem o mesmo nome. Se selecionar essa medida e trazer para um visual em branco colocando como cartão, ao invés de uma coluna, temos um valor de 0,04, que é exatamente a seleção da tabela. Formato essa medida para o mesmo formato de dados, que é o percentual, para verificarmos que é exatamente o mesmo valor. Se alterar para 6,70, também vai alterar, e assim por diante. Portanto, acho que você consegue imaginar que podemos utilizar essa medida de parâmetro valor, para automaticamente coletar o valor selecionado dentro desse segmento de dados e multiplicar pela soma custo. Então selecionar a soma custo, procurar por ela dentro de "pedidos detalhes". Se estiver difícil de procurar por uma medida, pode também utilizar o campo de pesquisa, digitar "soma custo". Encontrei! Então, vou seleciona-la, aqui dentro podemos fazer uma multiplicação por essa medida que se chama "parâmetro valor", que utiliza uma função chamada "selected value", que coleta o valor selecionado em segmentadores de dados, que utilizam a coluna "informado" dentro dela. Dentro dessa "soma custo", depois desse parêntese, "adicionar essa multiplicação", entretanto quero manter com essa medida. Então crio uma nova, referencio a soma custo e posteriormente multiplico por esse parâmetro. Em "nova medida" crio "soma custo variação", depois do igual chamo pela soma custo, recém mostrada, foi a medida criada na aula anterior. Somar pela multiplicação de soma custo vezes o parâmetro valor. Quando utilizamos esse novo parâmetro, o Power BI, automaticamente, já insere o nome da tabela antes da medida "parâmetro valor", mas não é necessário, já que medidas não possui um contexto fixo, elas podem estar em qualquer tabela. Esse parâmetro é inútil de ser informado. Mas como é uma funcionalidade padrão do Power BI, aí ele já insere. Fechar o parêntese da "parâmetro valor" e o que estamos fazendo aqui, selecionando o valor de soma custo, depois fazendo uma adição pelo percentual de 6,50, por exemplo, dessa seleção atual, sobre esse valor de custo também. Então, depois do parêntese, pressionar "ENTER" para criar essa nova medida, selecionar esse visual e inserir a "soma custo variação". E a variação de 6,50%, nesse custo, gerou esse valor e agora podemos selecionar outros tipos de variações, por exemplo, 1,70% para resultar em 102 mil. Formatar essa "soma custo variação" para o mesmo formato desses dados porque aqui nós temos um "R\$" e depois quatro decimais. Dessa forma vai ficar muito mais fácil de entender. Selecionei a soma custo variação e aqui em "moeda geral" vou selecionar "decimal", e vou deixar 2 casas decimais, assim como as outras. Também faz mais sentido trazer essa coluna para depois da soma de custo. Então, vou selecionar o visual, em "valores" buscar por "soma custo variação" e colocar acima de lucro. Pronto! temos

a redistribuição dessas medidas dentro do visual. Vamos supor que um diretor está fazendo simulações de novos investimentos em pesquisa na empresa, vamos aumentar o custo em 5,80%, e ele está avaliando, neste cenário, se compensa ou não aumentar o custo em 5,80% e terá apenas um determinado montante de lucro. Entretanto, é preciso observar que essa "lucro" ainda está apontando para "soma de custo" e não para o custo variação. Precisamos alterar essa medida de lucro ou criar uma nova com o lucro variação. Escolher criar nova em "modelagem", "nova medida", que se chamará "lucro variação"; será igual a "soma vendas" menos "soma custo variação". Então "soma vendas" menos essa medida, pressionar "ENTER". Depois procurar por essa medida na pesquisa com nome "lucro". "Lucro variação", trazendo para dentro do visual e temos a variação. Algumas categorias estão negativas e isso não pode acontecer. O diretor pode ajustar esse investimento em pesquisa na empresa, para garantir que todas as categorias estejam dando lucro e que também o valor final seja aceitável. Você acabou de aprender nessa aula a criar uma medida dinâmica utilizando a "selected value", que foi criada automaticamente pela "novo parâmetro." Essa medida também pode ser criada manualmente, basta clicar em "nova medida" e posteriormente dizer "parâmetro selecionado", determinando que o nome da medida será esse e utilizando a função "selected value", que é essa daqui. Sua descrição fala o seguinte: ela vai retornar o valor, quando há apenas um valor na coluna especificada, caso contrário retornará um valor alternativo. Então, como dentro dessa coluna estamos com um valor selecionado, que é 4,20%, retorna esse valor. Utilize esse recurso e técnica sempre que você precisar fazer variações e cenários.

Calcule o ano anterior com a SAMEPERIODLASTYEAR

A partir dessa aula começaremos a estudar as funções de inteligência de tempo. Vamos começar com uma que se chama "sameperiodlastyear", que é o mesmo em português de "mesmo período do ano anterior". Imagine que temos uma coluna com valores de data de 01/01/2019 até 06/01/2019. A sameperiodlastyear retornará ao valor do mesmo período do ano anterior, isso quer dizer que se colocarmos 01/01/2019 dentro do eixo de um gráfico, ou então dentro de uma tabela, a sameperiodlastyear retornará aos valores no ano anterior desse mesmo período. Ou seja, do dia 01/01/2018 para o dia 01/01/2019 e assim por diante. O mais incrível das funções de tempo é que se alterarmos esse contexto, por exemplo, se tivermos apenas os meses completos de 1º de janeiro até dia 31 de janeiro, a sameperiodlastyear se adapta e traz os valores de janeiro de 2018 também completo. Vamos ver na prática esse conceito lá no Power BI. Nesse arquivo nós temos uma tabela com o ano, o nome do mês e a soma vendas, que é uma medida que se soma o total de vendas realizadas em cada um desses anos e meses. Para calcular qual foi o valor realizado no ano anterior e comparar se minha performance atual está melhor ou pior. Para mostrarmos o valor do ano anterior nessa mesma linha, precisamos alterar o seu contexto, porque temos um filtro de 2019 e janeiro, e preciso mostrar janeiro de 2018. Temos que usar a função calculate. Em "Modelagem", "Nova Medida", ou então clicando aqui com o botão direito, em cima de "Calendário", "Nova Medida". Vamos criar uma medida que se chama "Vendas Ano Passado". Depois do igual, vou colocar uma calculate para alterar esse contexto de filtro e chamar uma expressão. Para calcular a soma de vendas, vou utilizar dentro de "Expressão", a medida "soma vendas", ponto e virgula, vamos filtrar a soma vendas com uma função chamadas "sameperiodlastyear", que retorna um conjunto de datas da seleção atual do ano anterior. Se a seleção atual é janeiro de 2019, vai retornar um conjunto de datas igual mas do ano anterior. Selecionar essa função e o único argumento que ela me pede é uma coluna de data, para isso informar a coluna de data que tenho dentro da minha tabela calendário. Clicar 2 vezes nela, depois fechar o parênteses da sameperiodlastyear, outro parênteses para fechar o da calculate, e pressionar "ENTER" para criar essa medida. Com ela criada, selecionar e trazer para o visual. A informação que temos é que em janeiro de 2018 o valor vendido foi 3 milhões. Se viermos para janeiro de 2020, o valor vendido em 2019, no mesmo período, ou seja, janeiro, foi 1.000.309. Vamos verificar se foi em janeiro de 2019, 1.000.309 e está correto! A função sameperiodlastyear criou um filtro dentro da função calculate, trazendo só o período do contexto atual, que nesse caso era janeiro/2020, mas considerou o ano passado. Quando calculate calculou essa medida, considerou esse filtro realizado, alterando, portanto, os valores exibidos aqui no visual. Se retirar o contexto atual de mês desse visual, selecionar essa tabela, retirar o nome do mês, teremos apenas 2019, onde foram vendidos 30 milhões arredondados. Esses 30 milhões também são exibidos aqui em baixo, que é o período anterior de 2020. Exibe 37 milhões porque foram os valores vendidos em 2018. Então retirando a seleção do filtro, voltando com 2019, em "vendas do ano passado de 2018" não aparece nada porque antes de 2018 não tinham vendas nesse modelo, por isso está em branco. Com a função

sameperiodlastyear rapidamente conseguimos fazer uma seleção do mesmo período do ano anterior, juntamente com a função calculate, conseguimos alterar o contexto de filtro para exibir os valores nesse contexto e possibilitar rápida comparação entre 2019 e 2018, e 2020 contra 2019. Nas próximas aulas vamos continuar explorando as funções de inteligência de tempo.

Calcule o mês anterior com a DATEADD

Continuando a analisar as funções de inteligência de tempo no Power BI, veremos sobre a "dateadd", que permite personalizar períodos e quantidade deles em seus argumentos. Imagine a tabela onde temos de Janeiro de 2019 a Junho de 2019. Com "dateadd" podemos, por exemplo, fazer uma análise do mês anterior, comparando a sua performance. Do lado de Janeiro/2019 consigo colocar Dezembro/2018 usando a "dateadd". O mesmo com Fevereiro/2019, buscando Janeiro/2019; portanto, ela desloca um período, de acordo com os argumentos inseridos nela. Conseguimos comparar esses dois valores alterando o seu contexto de filtro utilizando uma "Calculate" e trazendo o valor de Janeiro para a mesma linha de Fevereiro. Vamos ver na prática isso tudo acontecendo. No Power BI, com o mesmo projeto finalizado na aula anterior sobre a "sameperiodlastyear", vou adicionar de volta o nome do mês aqui. Para isso, em "Calendário" selecionar a "Nome do mês" e colocar abaixo de "Ano". Agora eu quero comparar, além das vendas do ano passado, se a performance deste mês, Abril de 2019, foi melhor ou pior que o mês anterior, que é Março de 2019. Para isso, a função "sameperiodlastyear" não funciona porque só busca o período do ano passado; então criar uma nova medida em "Modelagem", "Nova medida" ou clicando com o botão direito em cima de "Calendário", "Nova medida". O nome dessa medida será "Vendas mês passado", que será igual a "Calculate", porque precisamos alterar o contexto de filtro e a expressão será a "Soma vendas", porque queremos mostrar a "Soma vendas", mas do mês anterior. Vou colocar um ponto e vírgula para irmos ao argumento de filtrar e, posteriormente, vou chamar a função "dateadd". A sua descrição diz: ela vai mover um conjunto determinado de datas, de acordo com o intervalo especificado. Vou selecioná-la, ela tem 3 argumentos: primeiro, uma coluna de data; depois, o número de intervalos e, posteriormente, o intervalo. A minha coluna de datas será a data do calendário, que é onde tenho datas completas sem pular finais de semana, feriados e sei que é uma tabela íntegra. Para funções de tempo, sempre prefira usar a tabela "Calendário", justamente por esse motivo de ter datas contínuas sem pular eventuais calendários em que uma empresa pode não trabalhar. Colocar ponto e vírgula para irmos ao segundo argumento que é o número de intervalos. Nesse caso, vou preencher menos um, porque quero menos um mês para trás; mas informo o intervalo apenas no próximo argumento; coloco um ponto e vírgula e esse intervalo será o mês. Posso escolher o dia anterior para fazer essa comparação, o trimestre anterior ou o ano anterior. Se utilizasse o ano anterior, seria idêntico resultado da "sameperiodlastyear". Mas utilizo o novo, o month. Vou fechar o parêntese da "dateadd", fechar o da "Calculate", pressionar ENTER para criar a medida e procura-la para trazer à tabela. Ao fazer isso, vamos avaliar o retorno, em 2019, em Janeiro, temos 1.000.309; se verificarmos em Fevereiro, o valor que aparece aqui na "Vendas" do mês passado foi exatamente o valor de Janeiro. Então consigo comparar 2.000.451 com 1.000.309; e vendi muito mais em Fevereiro do que em Janeiro. Nessa aula você acabou de aprender a função "datedd" e utilizamos uma "Calculate" para alterar o contexto de filtro. Como mencionei na aula sobre "Calculate", utilizaríamos ela em várias funções durante esse curso; por isso ela é uma das mais importantes do DAX. Na "dateadd" podemos colocar vários intervalos, como vimos, então lembre-se disso na hora de fazer suas análises. Vou deixar o mês representado pela palavra "Month", em inglês, que significa mês e na próxima aula continuaremos a explorar as funções.

Acumule valores com funções de TOTALYTD, QTD e MTD

A Microsoft atualiza o Power BI com uma frequência bastante alta, todos os meses e sempre surgem novidades. Dentre elas, existe uma função de inteligência de tempo que foi aprimorada chamada "Total Year To Date". Com ela, não precisamos mais utilizar a função "Calculate", apesar de que dentro dela internamente isso já acontece utilizando uma função que altera o contexto de filtro. A "Total Year To Date" é capaz de acumular valores então, por exemplo, se estamos em Janeiro, ela vai mostrar apenas Janeiro; em Fevereiro, vai mostrar a somatória de Janeiro e Fevereiro e assim por diante, ou seja, ela vai acumular esses valores de acordo com o contexto de filtro atual. A sigla "YTD", que tem no final do nome da função "Total Year To Date" significa no inglês um indicador chamado "Year To Date", que é o ano até a data. Para compreender melhor esse conceito, vamos verificar diretamente no Power BI. Utilizando o mesmo arquivo, com o projeto que desenvolvemos nas últimas aulas, eu vou criar uma nova medida utilizando a "Total Year

To Date" para fazer um acumulado de todos esses valores, mas apenas até o final do ano. Para isso, eu vou clicar com o botão direito em cima da "Calendário" e depois "Nova medida". Nessa nova medida, eu vou chamar de "Acumulado Year To Date", que é o mesmo que "Year To Date" - vou escrever por extenso para ficar mais claro. Depois do igual, eu vou chamar a função "Total Year To Date" e temos muitas outras, como "QTD", que significa "Quarter To Date", ou seja, em português, "Trimestre até a data" e a MTD ou, então, "Month To date", que é o mês até a data. Eu vou escolher a "Total Year To Date". Ela me pede alguns argumentos. O primeiro é expressão e, nesse caso, eu quero mostrar a somatória de vendas; portanto, vou informar a medida "Soma de vendas". Então vou chamar a "Soma Vendas", vou pressionar o ponto e vírgula para informar o argumento de Dates, que é a minha coluna de data e vou inserir na "Calendário data", que possui datas lineares. A total "Total Year To Date" também possui outros argumentos como "Filtrar" e "YearAndDates". A "Filtrar", como o próprio nome já diz, pode realizar um filtro dentro dessa expressão e "YearAndDates", essa sua empresa utiliza um outro calendário fiscal, por exemplo, se ela terminar o ano em Junho ou, então, em Julho e não em Dezembro, como o calendário civil normal. Eu não vou preencher esses argumentos, vou simplesmente fechar a função "Total Year To Date" com parênteses aqui no final, pressionar ENTER e teremos a medida criada. Vou colocar no visual e agora vamos analisar o resultado. Eu vou usar a barra de rolagem até aqui em cima para verificar que em 2018, Janeiro, vendemos três milhões, arredondando, e é esse valor que aparece um acumulado "YearToDate". Já em Fevereiro, vendemos quatro milhões, arredondando, e o valor que aparece no acumulado "Total Year To Date" é sete milhões, ou seja, três + quatro. Ele acumulou de Janeiro até Fevereiro e isso vai acontecer em todos esses meses até Dezembro. Então aqui temos dois milhões e o que a "Total Year To Date" mostra é 37 milhões, que é a somatória de 2018/Dezembro até 2018/Janeiro. O mesmo acontece nos outros anos. Então quando um novo ano se reinicia, estou agora em Janeiro de 2019, o mesmo valor de Janeiro de 2019 é exibido: um milhão, trezentos e nove mil, que é o que temos aqui em "Vendas". Em Fevereiro, ele acumula Janeiro com Fevereiro, então aqui temos três milhões e quando chegar em Dezembro, ele novamente exibe o valor da somatória de Janeiro até Dezembro, e reinicia no próximo ano; isso que é um cálculo "Year To Date". Se alterássemos a função utilizada, por exemplo, eu vou selecionar novamente a medida e ao invés de utilizar a "Total Year To Date", agora eu vou usar a "Total QTD", ou seja, "Quarter To Date". Alterei a função, vou pressionar ENTER para aplicar a alteração e agora teremos um outro cenário. Aqui em Janeiro, Fevereiro e Março, temos a somatória desses três meses; só que em Abril de 2018, reinicia a contagem. Temos 3.000.820, que é o mesmo valor aqui desse mês. Então ele vai somar de três em três meses, ou seja, a cada trimestre. Mas eu não quero "QTD", "Quarter To Date", vou alterar de volta para "Year To Date" porque esse é o nome da medida que colocamos. Vou selecionar depois dos parênteses, pressionar ENTER para aplicar a alteração e temos novamente o acumulado do ano rapidamente realizado. Então nessa série de aulas, você acabou de aprender a utilizar várias funções de inteligência de tempo. Eu acredito que você também achou que elas são tão incríveis quanto eu falei no início. Existem muitas outras funções de inteligência de tempo que podem ser consultadas na documentação da linguagem DAX.

8. Organizando os visuais e compartilhando o projeto

Organização e publicação do relatório

No início desse curso vimos que existe um processo no Power BI. Primeiro obtemos os dados no Power BI Desktop utilizando o Power Query; posteriormente, modelamos esses dados utilizando também o Power Query para tratá-los; os relacionamentos e depois das funções DAX para realizar cálculos. Fazemos visualizações com gráficos e tabelas e, no fim, temos que publicar esse projeto para consumo das informações; e isso fazemos no Power BI Serviço, ou seja, no site do Power BI. Vamos finalizar agora o projeto que desenvolvemos durante esse curso, organizando um pouco os visuais e compartilhando no Power BI Serviço. Esse é o arquivo final que você pode encontrar na pasta de exercícios desse curso, e temos algumas páginas aqui embaixo e essa primeira daqui que será onde colocaremos o relatório. Então aqui eu vou clicar duas vezes em "Página um" e renomear para "Relatório" e remover esse gráfico porque esses dois representam os mesmos dados; são simplesmente vendas versus meta por ano. Então vou remover esse e reorganizá-lo para cá, na esquerda, e redimensionar. Agora, vou verificar se essas outras páginas precisam aparecer. Na "Calculate" apenas estávamos aprendendo um conceito novo, então eu posso clicar com o botão direito em cima dela e ocultar essa página. Eu ainda vou mantê-la mas para o consumo dessas informações lá no Power BI Serviço, ela não aparecerá. Vou fazer o mesmo com "ALL e ALLSELECTED", cliquei com o botão direito e vou ocultar; as opções interessantes, também vou fazer o mesmo, ocultar. Em

"Parâmetros", eu vou querer trazer essa visualização lá para o "Relatório". Então eu vou selecionar esse visual, pressionar CTRL+X no teclado para recortar, selecionar o "Relatório", colar aqui e posicioná-lo mais abaixo. Vou voltar na página "Parâmetros", avaliar se preciso trazer esse daqui, e não preciso porque é apenas uma seleção que era apenas esse visual de parâmetro que faz a alteração desse campo daqui selecionado. Então vou pressionar CTRL+X no "Parâmetro", voltar para o "Relatório", colar aqui e permitir que os usuários que acessaram este relatório utilizem esse parâmetro para modificar o lucro e a sua variação. E essa página de "Parâmetros", eu vou excluir porque não tem mais nada nela; então eu cliquei no "X" que apareceu acima dela, aqui, e depois vou clicar no botão "Excluir". Em "Funções de inteligência de tempo", eu quero mostrá-las como uma segunda página do meu relatório; então vou selecionar e posicioná-la antes da "Calculate" e depois da "Relatório". Dessa forma, temos uma análise temporal dessas informações de vendas do ano passado, no mês passado e no acumulado Year To Date. Aqui em "Relatório" vou organizar melhor esses visuais, colocar um pouco para cima, reduzir um pouco. Aqui em baixo vou diminuir esse tamanho e colocá-lo um pouco mais para baixo. Eu vou tentar colocar mais alguns visuais aqui que façam sentido para este relatório, por exemplo, fazer uma análise dessas vendas por país; então eu posso selecionar um visual de mapa, posteriormente, procurar pelas regiões, e buscar o país e colocar em "Localização". Agora eu quero trazer a soma vendas para a legenda. Na verdade eu escolhi o visual errado. Esse mapa daqui é o coroplético, eu quero o de bolha que é esse daqui. Posteriormente, vou trazer a "Soma Vendas" para "Tamanho" e teremos o tamanho da representatividade dessas vendas em cada um desses países; posso aumentar um pouco esse gráfico, reposicioná-lo. Vou colocar esse gráfico um pouco mais para baixo e colocar um título aqui em cima. Para isso, aqui na página inicial vou clicar em "Caixa de texto", posicioná-la aqui do lado, aumentar um pouco seu tamanho, a sua fonte para 20 e escrever "Relatório gerencial". Posso alterar essa fonte, aqui no "Segoe UI" é a fonte atual, quero colocar Segoe (Bold). Posso aumentar um pouco também, se eu achar que ficou pequeno. E temos um relatório que compara as vendas versus a meta, que é uma informação super importante; qual é a representatividade dessas vendas pelos países e posso utilizar a rodinha do mouse para dar zoom ou retirá-lo e também temos uma variação do custo com a soma de vendas por categorias, por custo e o seu lucro e avaliação, conforme a seleção do usuário. Agora que finalizamos a construção desse relatório, apesar de simples, porque focamos mais no Power Query e no DAX, que são técnicas fundamentais para realmente dominar o Power BI, eu vou clicar em "Arquivo", "Salvar como" e vou dar o nome de "Adventure publicação finalizado". Vou salvar e agora vou me certificar que estou logada na minha conta - e eu estou -, se você não estiver, vai estar escrito "Entrar" e aí você tem que entrar com uma conta, seja ela gratuita ou PRO e aqui na página inicial, vou clicar em "Publicar". Vou escolher um Workspace, que será os "Relatórios executivos"; posteriormente, selecioná-lo e fazemos a publicação diretamente no Power BI Serviço, que é app.powerbi.com. Na próxima aula vamos verificar meios de consumo dessas informações diretamente no Power BI Serviço.

Formas de consumo de dados no Power BI

Na aula anterior fizemos a publicação desse projeto, e agora verificaremos formas de compartilhá-lo com outras pessoas. Para abrir o Power BI Serviço vou clicar aqui em "Abrir", "Adventure", "Publicação", "Finalizado". O site app.powerbi.com está sendo aberto, com o relatório exibido aqui nesta página. É bastante importante ressaltar que esse relatório foi publicado para dentro da minha conta; você consegue ver que eu estou logada porque aparece a minha foto aqui e isso aconteceu porque lá no Power BI Desktop - vou voltar para ele agora -, eu já estava com a conta logada aqui também, então ao publicar, ele foi diretamente para a minha conta no Power BI Serviços. O relatório é completamente interativo, por exemplo, nós temos aqui essa análise de custo. Se eu aumenta-lá, ela vai alterar também esses valores aqui embaixo; se eu diminuir, também, da mesma forma e agora, eu quero compartilhá-lo com outras pessoas. Para isso, eu preciso ter uma conta PRO, então aqui nesse botão chamado "compartilhar" eu posso inserir os e-mails das pessoas que podem ter acesso a esse relatório, escrever também uma mensagem para elas, caso eu queira, e permitir ou não que esses destinatários compartilhem esse relatório, utilizem o conjunto de dados que foi tratado por mim nesse relatório e se eles receberão ou não uma notificação por e-mail dizendo que eu compartilhei com eles. Não vou publicar com ninguém, vou simplesmente pressionar "Cancelar". Caso você não tenha uma conta PRO no Power BI, você também pode publicar esse relatório utilizando o "Exportar para PowerPoint", mas esses dados ficarão estáticos, ou seja, você não poderá utilizar filtros ou, então, PDF que acontecerá o mesmo processo e até mesmo imprimir. Existe uma outra forma de publicar que são nesses três pontinhos. Aqui podemos escolher "Incorporar" e, posteriormente, publicar na

web. É bastante importante que você saiba que, ao clicar em "Publicar na web", essas informações ficaram públicas, ou seja, não terá a camada de segurança de solicitar login e senha. Então utilize essa opção com muita atenção e cautela, porque qualquer pessoa com esse link poderá ver essas informações. Se você possuir a conta PRO, pode compartilhar no SharePoint online, que é um serviço da Microsoft para gerenciamento de documentos e integração e organizações ou então publicar em um site ou portal. Eu vou escolher essa opção e agora eu tenho uma URL com direto acesso a esse relatório. Se eu seleciono essa URL e copio, depois clico aqui em uma nova aba e colo essa URL e pressiono ENTER, teremos o acesso diretamente ao relatório e não ao Power BI Serviços. Verifique que teremos um visual bastante diferente, focado apenas nesse relatório. Se eu volto nessa outra aba, que é onde temos o Power BI Serviço, há um menu aqui na esquerda, nós temos algumas ferramentas aqui em cima; já essa publicação de incorporar no site ou portal temos diretamente o relatório pronto para ser utilizado, com as páginas aqui abaixo. Portanto, acabamos de ver várias formas de publicar esse relatório e compartilhar com outras pessoas. Algumas que não são tão recomendadas porque deixam o relatório estático como PowerPoint, PDF e imprimir; outras que você deve utilizar com cautela como "Publicar na web" e algumas que são bastante interessantes como o "Site ou portal" e "SharePoint", mas você precisa antes ter a conta PRO do Power BI.

CONCLUSÃO

Próximos passos para explorar seus conhecimentos e continuar aprendendo

Parabéns, você chegou ao fim desse curso. Falamos de muitos conceitos complexos sobre o Power Query e a linguagem DAX, por isso, é fundamental que você continue praticando para dominar o assunto. O Power BI é possui atualizações mensais, então é recomendado que você fique por dentro de todas elas, diretamente no site oficial, principalmente, na seção do blog. Como mencionei durante esse curso, leia as documentações das linguagens DAX e M. Meu desejo era te ensinar todas elas, mas teríamos um custo enorme e nem todas serão relevantes para todos os cenários que você utilizará; e eu te convido a se conectar comigo no LinkedIn. Basta procurar pelo meu nome, Karine Lago. Nos vemos nos próximos cursos!