

Dax e ETL com Power Bi: Carregando Dados

Aula 01 – Entendo processo de ETL

[00:00] Meu nome é Victorino Vila e esse vai ser um curso já um pouco mais especializado na parte de Power BI, vimos um curso passado, um curso introdutório, que inclusive é um misto entre algumas coisas que foram vistas no curso de Power BI, especializado na carreira Business Intelligence e no curso passado que é o primeiro curso de uma carreira de Power BI que nós estamos produzindo aqui na Alura, agora sim esse curso vai ser um pouco mais especializado.

[00:35] E vamos tratar nesse curso, da primeira parte da construção de um relatório em Power BI que é a parte de carga dos dados, nem sempre as nossas fontes, elas estão preparadas para o nosso relatório, ao ler os dados para dentro do Power BI, podemos precisar fazer algumas transformações e além disso precisamos fazer alguns cálculos. O foco desse nosso treinamento vai ser tratar somente dessa parte inicial, que é a parte de leitura de dados e de transformações e cálculos.

[01:18] Vamos olhar um pouco os tópicos que são vistos nesses cursos? Então claro, vamos passar pela parte de instalação do Power BI, porque pode ser que você aluno, não tenha feito o curso introdutório, porque você já tem conhecimento do Power BI, ou já fez o curso do Power BI na carreira de Business Intelligence e sua máquina está vazia, você precisa instalar o Power BI.

[01:50] E vamos discutir um pouco sobre quando precisamos aplicar o ETL dentro do Power BI, quando já temos o ETL pronto vindo do DataWarehouse, quando queremos construir relatórios gerenciais e falar um pouco sobre o conceito, o que é o DAX e o que ele tem de diferente com as funções do Excel.

[02:20] Depois na parte 2 do nosso treinamento, vamos fazer uma conexão com diversos tipos de bancos de dados diferentes, na verdade, vou corrigir a expressão que eu falei agora, vamos fazer a conexão com diversas fontes de dados diferentes, então vamos ver como lemos de um arquivo texto, como lemos de Excel, como lemos de um folder, vamos ver por exemplo como lemos de um banco de dados relacional e por exemplo quando eu leio dados diretos da internet, da web.

[03:00] Depois na parte 3 do nosso curso, vamos estar falando aqui sobre diversos tópicos, entre eles na parte de algumas transformações que podemos

fazer, através de algumas funções previamente já definidas dentro do menu do Power BI, no momento da carga, na aula 4 vamos dedicar o nosso treinamento na parte da modelagem do nosso ambiente e vamos ter o primeiro contato com DAX.

[03:40] Quando eu falo de modelagem, eu estou falando sobre melhorar o meta dados que foi carregado no Power BI e também sobre o relacionamento das tabelas, e vamos ter o primeiro contato de como é que eu, as vezes ao criar colunas calculadas eu já uso o DAX. Bem as aulas 5 e 6 vão estar dedicadas ao DAX.

[04:08] Basicamente aula 5 nós vamos ter o primeiro contato com DAX vendo algumas funções soltas e a aula 6 vamos estar se dedicando a um tópico específico, que é a construção de um relatório que nós chamamos de relatório Pareto de 80/20, é um relatório muito usado geralmente nas empresas e vamos usar ele como problema prático para treinar um pouco e ver algumas outras funções do DAX.

[04:35] Então é isso, o objetivo desse treinamento é esse, é tratar a parte de carga e transformação dos dados, antes de começar a construir o dashboard, espero que vocês gostem, vamos seguindo em frente, um forte abraço a todos.

[00:00] Para você que está começando a carreira de Power BI direto nesse curso, ou seja, não fez o curso de introdução, eu vou mostrar agora um vídeo que eu gravei no curso introdutório, falando um pouco os conceitos da arquitetura do Power BI e vale a pena você que não viu esse vídeo assistir ele novamente para entender um pouco como é a arquitetura do produto, se você já fez o curso introdutório e já assistiu esse vídeo, se você quiser não precisa assistir o resto do vídeo atual.

[00:35] Então vamos lá, vamos assistir o vídeo sobre os conceitos do Power BI. Já que o nosso assunto é Power BI, vamos falar um pouco sobre o conceito da arquitetura do Power BI, porque ele não é apenas uma aplicação, ele é um conceito de arquitetura que engloba todo o ferramental de visualização de dados fornecidos pela Microsoft e ele já carrega consigo um conceito novo, de relatórios gerenciais que tem como objetivo a distribuição fácil de qualquer informação para as pessoas dentro da empresa.

[01:30] No começo tínhamos relatórios gerencias que precisavam da área de tecnologia da empresa, a área de TI, para preparar as informações e depois entregar esses relatórios para os usuários, é o que eu chamaria aqui de um BI técnico. Depois entramos para um segundo patamar, que seria ferramentas que analistas da informação exploram as mesmas e entregam para o usuário a

informação, ou o próprio usuário mesmo, com um pouco mais de treinamento, ele consegue fazer isso.

[02:10] O que nós chamamos hoje de self-service BI, hoje estamos em uma onda a cima de isso tudo, que é End-User BI, ou seja, a informação fácil para todos, que o próprio usuário, possa ele mesmo, extrair a informação, preparar, visualizar e depois compartilhar aquela informação, o Power BI tem a pretensão de ser uma arquitetura que forneça instrumentos para construirmos o que eu chamaria aqui de End-User BI.

[02:55] Algumas características do Power BI, primeiro, ele facilita o acesso ao seu dado, seja ele um dado que esteja na sua empresa, seja dado que esteja na nuvem, seja qualquer outro tipo de dados que você precise capturar para massagear eles e transformar em relatório. Então ele fornece um acesso rápido e fácil do seu dado, ele permite o que nós chamamos em uma visão 360° da organização, isso porque você consegue dentro do próprio relatório, juntar coisas diferentes da organização e visualizá-las de maneira completamente aberta.

[03:48] De forma fácil e intuitiva, e já falei um pouco no terceiro tópico que é a exploração dos dados, isso permite que você consiga explorar os dados de qualquer maneira, a colaboração é muito importante, você pode compartilhar seus relatórios com outros usuários e criar por exemplo, comentários onde você vai colocando suas conclusões e o usuário quando olha seu relatório ou o relatório que está compartilhado, ele consegue também visualizar conclusões que você deixou registrado nesses relatórios.

[04:30] E também uma característica importante é a visualização desses relatórios em qualquer lugar, através da web, computador, telefone/celular, dependendo da forma com que você libere essa informação, você pode visualizá-la de qualquer lugar.

[04:50] Alguns diferenciais, o Power BI já fornece uma série de análises pré-construídas que podem ser disponibilizadas como um serviço, ou seja, você em vez de começar um relatório do 0, você já pode pegar coisas que ou estejam publicadas na sua própria organização, ou então serviços externos, que já te forneçam análises previamente construídas para você começar, a partir dali, construir as suas próprias.

[05:30] Ele permite também, que o seu relatório possa, salvo algumas condições claro, fazer atualizações das informações em tempo real, ou seja, aquilo que eu estou olhando em meu relatório é o que realmente está acontecendo e não há necessidade de esperar por exemplo, um pré-

processamento, alguma coisa a ser executada durante o dia ou a noite, para eu visualizar a informação detalhada.

[06:00] Um diferencial do Power BI é a segurança, através do seu serviço de nuvem, as informações que estão armazenadas e visualizadas por você, são completamente seguras no que diz respeito a visualizações externas, através de hackers, ou algum tipo de outras pessoas que queiram invadir seu ambiente para visualizar aquilo que elas não podem ver.

[06:28] A exploração do dado é extremamente intuitiva, quando você publica um relatório, a forma com que ele é apresentando, faz com que você intuitivamente consiga explorar o dado de maneira fácil, sem a necessidade de entendimentos profundos técnicos, todas as informações que o Power BI produz, elas são integradas naturalmente com os produtos da família Microsoft, mais precisamente com o Office 360.

[07:00] E claro, tem como objetivo você de maneira fácil e rápida já construir e distribuir relatórios através do Power BI, temos mais ou menos a arquitetura de como é o ambiente do Power BI, então nós temos aqui no canto superior esquerdo, o que nós chamamos de ferramentas analíticas, ou é o seu Excel onde você tem planilhas com uma série de informações que você pode publicá-las dentro do ambiente do Power BI da nuvem.

[07:40] Ou então, uma aplicação que aqui eu estou chamando de Power BI design, onde eu construo os meus próprios relatórios localmente e depois publico eles na nuvem, esses dados visualizados pelo Excel ou pelo Power BI design, podem estar olhando seus dados que nós chamamos de On-Premises, ou seja, são dados que estão armazenados nos computadores de redes da sua empresa.

[08:15] Seja, banco de dados SQL server, ou outras fontes de dados, como banco de dados hora como MY SQL, arquivos textos, qualquer outro tipo de banco de dados relacional, SAP, e uma série de outras fontes de dados que o Power BI consegue ler, mas também eu posso estar lendo coisas, da nuvem Microsoft, coisas que estão particularmente salvas em ambientes de nuvem específicos da Microsoft, já tem conectividades naturais com o Power BI.

[08:55] E você publica isso no Power BI em nuvem, onde seu relatório fica publicado e ele pode fazer atualizações em real time, em tempo real, ou então, de tempos em tempos, lendo as informações on presence ou as informações da nuvem, e também você pode ter aplicações que nós chamamos de fora da nuvem Microsoft, são aplicações que já fornecem conectores com o Power BI, que você pode por exemplo puxar informações.

[09:30] Como por exemplo Salesforce, que é uma aplicação de gerenciamento de vendas, que já fornece dados automaticamente para o Power BI e uma série de outras, chamamos isso de SSAS application, ou seja, aplicações com um serviço que fornece dados ao Power BI.

[09:54] E depois que essas informações estão, todas dentro das suas análises do Power BI, você pode distribuí-las e visualizá-las através de browsers da web, ou então por exemplo através de aplicações remotas, no seu próprio computador, aplicações mobiles, tablet e assim por diante.

[10:20] Claro que para você acessar a nuvem e publicar os seus relatórios na nuvem, você precisa se inscrever, você precisa fazer uma inscrição, ter uma assinatura na nuvem Microsoft do Power BI e você pode utilizar o seu próprio e-mail escolar de estudo ou então o seu e-mail da sua organização para fazer isso, existe uma série de tipos de licenciamento desse Power BI na nuvem, inclusive há um gratuito onde você pode estar publicando e compartilhando relatórios.

[10:55] Claro que a opção gratuita te dá algumas informações, em termo de volume, de formas que você pode compartilhar isso, mas já é possível pela forma gratuita, fazer bastante coisa com o Power BI. O Power BI ele já fornece para você como fonte de dados uma série de pré-conteúdos já pronto.

[11:25] Então você abre lá o painel, ele vai te mostrar uma série de aplicações externas que já distribuem dados para o Power BI, então você pode por exemplo, através desses pacotes criar relatórios ou painéis já previamente configurados, preparados para aquele negócio, modelo de dados e até mesmo consultas embutidas que já estão semiprontas lá nessa aplicação parceira, você pode consumir e estar analisando.

[12:05] Você também constrói o que nós chamamos de seu dashboard, o dashboard ou painel de instrumentos, traduzindo para o português mais literal, ele vai mostrar todas as informações importantes do seu negócio em um único local, e o dashboard ele pode ter na mesma visualização consulta de diferentes fontes de dado ou ao contrário, eu posso ter diferentes relatórios dentro de um único dashboard, e você vai consolidar em seu dashboard informações que são pertinentes a um determinado assunto.

[12:47] Exemplo, ter um dashboard de vendas, de custo, ou então um dashboard da presidência, um dashboard corporativo. Os dashboards são compostos de relatórios, e quando você constrói o seu relatório você tem um painel de desenho que começa vazio e você vai adicionando componentes, como por exemplo são mostrados aqui em cima, uma série de componentes já

pronto que você vai visualizando e esses componentes vão te mostrando determinados filtros ou propriedades.

[13:30] Que você vai configurando na medida em que você vai arrastando esses componentes, e você tem aqui do lado campos, que são os campos das suas fontes de dado e você associa quais campos você quer utilizar dentro das propriedades do componente que você está adicionando no seu painel, e você também pode criar além de componentes de visualização, filtros que vão ser aplicados automaticamente a todos os componentes de visualização que você adicionou no seu painel.

[14:05] E você pode por exemplo como eu falei, ter menus de propriedade, cada componente vai ter propriedade diferente e você pode estar colocando propriedades básicas como por exemplo, cor, tamanho, fonte ou campos que estão na sua fonte de dados.

[14:32] Você também pode, na parte de filtros, criar filtros dinâmicos nas suas visualizações, esses filtros podem ser aplicados em campos textos, numéricos ou campos de data, que estão vindo da sua fonte de dados e você pode visualizar esses filtros, disponibilizar esses filtros para os usuários através de uma lista, através de campos mais avançados onde você pode fazer seleções múltiplas, onde você pode ordenar a saída da informação ou através de expressões lógicas que você pode mesclar and e or para jogar as informações no seu filtro.

[15:25] A visualização, pode ser feita através de centenas de componentes diferentes, tem alguns componentes de gráficos padrões, linha, barra, ou então alguns gráficos especiais, como donuts, funil, gauges que são aqueles instrumentos de valores, mapas onde conseguimos colorir mapas através do resultado das nossas análises, ou então outros componentes externos.

[15:58] Vários componentes externos são oferecidos para você adicionar dentro do Power BI, às vezes são componentes pagos, às vezes componentes gratuitos. A disponibilidade dos dados através do mobile é uma coisa importante, inclusive hoje em dia é o que os dispositivos moveis já estão se tornando praticamente os seus computadores pessoais, então vocês podem disponibilizar essas visualizações independente do sistema operacional do seu telefone, seja um Windows Phone, IOS ou Android.

[16:40] Você pode através da aplicação telefone celular, selecionar as aplicações mais importantes para você, para que você tenha atalhos para visualizar elas de maneira mais rápida, você pode fazer zoom para visualizar melhor a informação, às vezes a tela do telefone é muito pequena para você conseguir visualizar o resultado que está ali, você pode criar anotações

associadas aquela visualização que depois vão ser compartilhadas com outros usuários e configurar alertas para quando algum determinado tipo de meta é atingida.

[17:20] A colaboração é muito importante dentro do ambiente moderno de informações gerenciais, é importante que você compartilhe as suas análises para todos os usuários dentro da sua organização. A área de TI tem um papel importante nesse ambiente que ele tem um painel onde ele pode controlar os acessos dos usuários as informações, e até que ponto eles podem ou não compartilhar elas com outros usuários.

[17:53] E você pode criar o que nós chamamos de área de trabalho ou workplace, onde dentro dessa área você pode juntar Datasets que são conjunto de dados, relatórios e dashboard, que vão ser visualizados ao mesmo tempo por um grupo de usuários. A atualização do dado dentro do Power BI pode ser feita de diversas maneiras, você pode ter o dado atualizado em tempo real, ou seja, na medida que eu estou visualizando a informação ela está acontecendo, é o dado que está lá.

[18:32] Você pode visualizar isso através de um schedule onde você agenda as atualizações e você pode por exemplo, fazer atualizações na medida em que o usuário interage com o usuário, na medida em que o usuário abre uma árvore, seleciona um filtro, essas atualizações são feitas, ou seja, a consulta vai lá e busca aquela informação.

[18:58] Existe um componente do Power BI chamado Gateway, que ele é instalado normalmente nos servidores da sua empresa, que faz uma ponte entre os dados que estão dentro da sua rede e as aplicações Power BI que estão na nuvem e você consegue fazer a conexão entre o dado publicado na nuvem e o seu dado interno na empresa.

[19:30] Essas atualizações do Gateway, podem ser feitas pontualmente, podem ser feitas manualmente, ou seja, eu quero que atualize agora e a configuração desse Gateway é uma coisa muito simples, que normalmente nem necessitaria a interferência de alguém da área de tecnologia para fazê-lo. E claro, toda essa transferência de dados ela é garantida através das ferramentas de segurança do serviço da Azure que é a nuvem da Microsoft.

[20:00] O Power BI Desktop é a aplicação onde você vai construir os seus relatórios que depois posteriormente você vai aplicar na nuvem, e ele é uma aplicação livre, ou seja, instalar o Power BI na sua máquina não é necessário ter nenhuma licença e você pode inclusive produzir relatórios e consumi-los pessoalmente na sua máquina, sem nenhum custo, ele combina 3 serviços que

no decorrer não só desse treinamento, mas nos próximos treinamentos de Power BI vamos estar falando mais detalhadamente.

[20:40] Que é o Power BI query, Power BI pivô e Power BI view, ele resulta em relatórios interativos e visuais e através dele eu consigo publicá-los na nuvem. Quer dizer, construo meus relatórios no Power BI Desktop, faço toda a conectividade, depois que eles estão prontos eu faço upload para minha conta na nuvem para depois compartilhar com os outros usuários.

[21:10] Existe uma parte do Power BI que é mais para desenvolvedores, onde você pode embutir a arquitetura Power BI dentro das suas publicações customizadas, ou seja, tenho uma página web que eu estou desenvolvendo, ou uma aplicação que eu estou desenvolvendo em C, ou em DOT NET, não importa, que eu posso embutir nela as minhas consultas do Power BI.

[21:40] E além disso, eu posso criar Datasets que vão ser enviados para o Power BI através da minha aplicação. Então é isso, é isso que eu queria mostrar para vocês, falar um pouco assim, bem rápido sobre a arquitetura do Power BI e o que vamos encontrar de forma mais detalhada nesse e nos próximos cursos dessa nova carreira aqui na Alura, valeu, um grande abraço.

[00:00] Pode ser que você não tenha feito o curso de introdução ao Power BI, porque ou já conhece o Power BI, ou então você fez o curso de Power BI na carreira de Business Intelligence e pulou o curso introdutório, eu mesmo lá na apresentação do curso introdutório eu falo que se você já fez o curso do Power BI na carreira de Business Intelligence não precisaria fazer o curso introdutório da carreira de Power BI.

[00:32] Apesar de nesse curso introdutório ter alguns assuntos que não são comuns, mas ok, se você então não fez o curso introdutório e está usando uma máquina que você não tem Power BI instalado, não tem problema, você assiste esse vídeo que eu estou gravando agora, eu vou repetir o vídeo que eu mostrei no curso introdutório, mostrando passo a passo a instalação do Power BI Desktop.

[01:00] Se você já tem o Power BI Desktop instalado na sua máquina, se está tudo legal, não tem problema, você pode parar esse vídeo e ir para o próximo, então vamos assistir o vídeo de instalação do Power BI Desktop.

[01:20] Para começarmos, precisamos baixar na nossa máquina o Power BI Desktop, como eu falei no vídeo anterior o Power BI Desktop é aquela aplicação que vai ser instalada na nossa máquina, onde vamos construir os relatórios que depois serão publicados na nuvem.

[01:40] Então vou abrir aqui uma janela no meu browser e vou colocar aqui para buscar, Power BI Desktop download, tem aqui o site da Microsoft, “powerbi.microsoft.com”, vou entrar nele, e eu tenho aqui vários downloads para serem feitos, mas o que eu vou fazer é esse aqui, Microsoft Power BI Desktop, vou clicar no botão download, ele me pede para abrir a Microsoft Store, então eu tenho a aplicação Power BI Desktop e eu tenho o botão aqui instalar.

[02:35] Vou clicar em instalar, e ele está baixando o Power BI e instalando na minha máquina, então eu vou aguardar alguns minutos, vou interromper o vídeo e quando eu terminar o processo de download, se tiver mais alguma coisa que tenha que ser feita eu vou estar voltando e vamos ver passo a passo a instalação, pronto, terminou o download, fiz a instalação, você vê que como o Power BI Desktop é uma aplicação Windows, a forma de baixar e instalar é bem diferente de um software normal.

[03:20] Como se fosse um aplicativo de celular, você simplesmente clica e ele já instala a aplicação na sua máquina, eu vou clicar em iniciar, então o Power BI Desktop vai ser inicializado, e pronto, eu tenho aqui a minha análise inicial, eu vou fechar essa caixa de diálogo e nós temos pronto a nossa interface do Power BI, pronta para ser utilizada.

[00:00] Eu vou falar agora aqui com vocês uma questão de conceito e metodologia, conceito e metodologia de implementação de um sistema de informações gerenciais, ou um sistema de relatórios, que é o objetivo do Power BI, que é te mostrar informações sobre a empresa, quem assistiu já o curso de Business Intelligence, os conceitos de Business Intelligence, que é a área que trata desse assunto de fornecer informações gerenciais para os executivos da empresas.

[00:40] Sabe que o processo de Business Intelligence possui uma série de fases que estão representadas no slide que eu estou mostrando aqui em cima, então basicamente temos um modelo de negócio que é desenhado através de entrevistas com os usuários, entendimento do que a empresa faz e a partir desse modelo de negócios, através de algumas técnicas, construímos o que chamamos de um Data Warehouse. O Data Warehouse vem do inglês, armazém de dados.

[01:13] E significa uma série de tabelas e um banco de dados que estão arrumadas para armazenar informações ao longo do tempo, parece que é um detalhe não muito importante, mas normalmente os bancos de dados que estão associados aos sistemas que controlam as transações da empresa, o sistema que controla a nota fiscal, o sistema que faz a folha de pagamento da empresa,

o sistema que controla os funcionários, isso tudo nós chamamos de sistemas transacionais.

[01:50] Que eles organizam as transações dentro das empresas. Os bancos de dados gerenciais eles tem um esquema de desenho diferente, que são feitos para armazenar informações para tomada de decisão e normalmente extraímos os dados que estão nas fontes de dados dessas empresas, dos sistemas transacionais, fazemos uma série de transformações, manipulações dos dados que nós chamamos de ETL esse processo de manipulação de dados.

[02:28] Que vem do inglês Extract Transform and Load, e eventualmente podemos ter um banco auxiliar, que chamamos de ODS, que é um banco de dados auxiliar para colocar informações dentro do banco de dados gerenciais e depois extraímos desse banco de dados gerenciais, uma série de base de dados que são representados por esses cubos, que são bases de dados que nós chamamos de OLAPs (Online Analytical Process).

[03:05] Por que eu estou colocando essa teoria toda? Isso tudo é visto com bastante detalhe no curso de Business Intelligence, mas eu estou falando isso, porque quando você fornece relatórios para os executivos das empresas, você vai estar lendo as informações desses cubos, onde a informação já está arrumada, já está mastigada, ela já está processada, por esse processo todo aqui, que aconteceu e você tem aqui, representado com 3 usuários de formas diferentes.

[03:40] Você pode ter, um usuário que quer ver o dashboard bonito, o gráfico de pizza, o gráfico de barra, aquele relóginho, coisas bem pontuais para ele olhar e tomar a decisão, mas em compensação ele não pode mudar aquela análise, análises mais ou menos flexíveis ou aquele cara que vai explorar o dado, explorar informação, até achar algum tipo de conclusão baseado em algum estudo dos números que estão sendo fornecidos por essas bases de dados.

[04:17] O Power BI ele entra aqui, ou seja, ele vai fornecer relatórios, dashboards, planilhas, pegando como fonte de dados esses cubos, quando isso acontece, podemos dizer que o Power BI vai fazer uma leitura direta, ou seja, eu já tenho o dado nesses 3 banquinhos por exemplo aqui arrumados, já estão prontos para serem usados nos relatórios para exibirem para os usuários.

[04:53] Porque esse processo todo aqui de trás ele já pegou o dado, já mastigou, já tirou o lixo, já normalizou tudo e colocou o dado todo legal aqui, só que o que vemos na prática é que você às vezes acaba usando o Power BI também assim, ou seja, eu quero fazer um relatório mas direto do Data

Warehouse, ou do banco de dados ODS ou até mesmo, eu quero fazer um relatório pegando os dados transacionais.

[05:32] Quando você vai fazer esse tipo de coisa nem sempre o dado está arrumado, ele só vai estar arrumado quando vir esses cubos aqui, mas se eu estou pegando o dado em pontos intermediários, fazendo dashboards, fazendo relatórios nesses pontos intermediários, eu preciso arrumar o dado antes de colocar no dashboard.

[05:57] E como esse processo não está ocorrendo quando eu utilizo essas fontes, precisamos dentro do Power BI fazer essa arrumação, ou seja, o processo de ETL que é feito dentro da teoria do Business Intelligence por fora e disponibilizando no banco de dados a informação já arrumada, quando você quer plugar o Power BI direto nessas fontes, você tem que se preocupar em fazer a limpeza pelo Power BI.

[06:30] E vem a pergunta, isso é certo ou errado? É relativo, fica a critério do cliente, decidir se é certo ou errado, o que acontece muito e é uma questão de prática, é que projetos de Data Warehouse, ou seja, fazer esse processo de ler as informações, transformar, criar o Data Warehouse, fazer as entrevistas, criar essas bases, ou seja, esse processo que vai até onde estão os cubos, ou seja, arrumar os dados, é um processo que as vezes custa dinheiro para a empresa, que nada é de graça, é claro.

[07:10] Você precisa ter consultores, você precisa ter ferramental, e é um dinheiro investido pela empresa que não tem uma aparência, o resultado não é palpável, o que é palpável é o dashboard, o que um usuário final que normalmente é o que está assinando o cheque para pagar a consultoria que vai fazer esse trabalho todo na empresa, ele quer ver é o gráfico dele, ele quer ver é o dashboard dele bonito para ele consultar.

[07:45] E quando você gasta um monte de dinheiro, ou um monte de tempo para preparar informação, para depois dizer para ele que só daqui a 6 meses, só daqui a 1 você vai ter os dashboards, muitos usuários vão dizer não, eu não vou assinar um cheque, para só daqui a 1 ano ter os meus dashboards, eu quero ter os meus dashboards agora, eu quero ter os meus relatórios de decisão amanhã.

[08:10] E muitas vezes você acaba plugando o Power BI direto no banco transacional para fornecer as informações para o seu usuário. Felizmente ou infelizmente as empresas fazem muito isso aqui de baixo, eu particularmente defendo a ideia de cima, de você preparar toda uma base de dados arrumada para depois começar a fornecer as informações gerenciais para os seus usuários.

[08:45] Mas vivemos no mundo real e isso aqui de baixo acontece, então esse mecanismo de fazermos o próprio ETL dentro do Power BI ele é importante e é importante que saibamos fazer isso, é o que vai ser alvo desse treinamento. Vamos falar muito, não sobre como construir o dashboard, mas sim, como fazer esse tipo de transformação.

[09:15] Primeiro eu queria falar um pouco sobre essa teoria, sobre as diferentes abordagens quando estamos implementando um projeto de Business Intelligence, ou implementando uma coleção de dashboards em Power BI, de saber por exemplo, se você vai realmente preparar o dado e o seu Power BI só vai ler a informação direto, ou se você vai ter que fazer ETL específico dentro do próprio Power BI.

[00:00] No vídeo anterior eu parei nesse slide aqui e eu gostaria de falar mais algumas observações sobre essa situação em que você vai ter que construir o ETL dentro dos relatórios do Power BI. Existe um cuidado que tem que se tomar quando você faz esse tipo de abordagem, vamos supor por exemplo que eu tenho uma base de vendas e eu preciso carregar esses dados para o Power BI, e eu preciso fazer transformações, manipulações de dados, formulas, criações de novas colunas, limpezas de dados.

[00:40] Eu tenho coisas que eu preciso expurgar, deletar da base, ou seja, processos que em um ambiente normal estariam sendo feitas lá pelo Data Warehouse, pelo processo de ETL, mas como eu estou lendo direto da fonte de dados, eu preciso fazer isso dentro do Power BI.

[00:58] E essas transformações estão representadas aqui com esse $AX + B$, coloquei uma forma matemática para representar uma transformação nos dados, quer dizer que o dado que está aqui, ele acaba saindo diferente aqui, problema é que se eu faço um relatório fazendo essas transformações e depois outro usuário, outra pessoa, outro analista vai usar o Power BI para fazer outro relatório, mas usando o mesmo indicador.

[01:25] Por exemplo, o usuário 1 fez o relatório do Power BI mostrando vendas da empresa, mas precisou fazer uma série de transformações, $AX + B$, transformar o dado para ter as vendas finais dentro do dashboard 1, e um outro analista de negócios da empresa, vai fazer um dashboard, também de vendas, lendo da mesma fonte, ele precisa tomar cuidado e também colocar aqui nesse segundo relatório, as mesmas transformações que o outro usuário fez no relatório original.

[02:05] Porque se não esse segundo relatório vai dizer que vendas foi 1 milhão e 200, enquanto o relatório de cima pode dizer que foi 1 milhão, porque as transformações que foram feitas entre a origem e o destino, foram

feitas de maneiras diferentes, porque como essas transformações elas ficam associadas ao conjunto de dados que estão associados ao Dashboard, ou ao relatório, é dentro da construção do relatório que elas são feitas.

[02:40] Então se eu não tomar cuidado e não repetir as mesmas transformações, eu posso gerar problemas, ou seja, dois dashboards que teoricamente deveriam mostrar as mesmas coisas, acabam mostrando números diferentes, então isso é um cuidado que deve ser tomado.

[02:55] Mas independente de que eu vou tomar esse cuidado e eu quero fazer essas transformações no Power BI como eu as faço? Eu tenho dois passos específicos para fazer isso, primeiro eu carrego os dados para dentro do ambiente Power BI, então o Power BI ele fornece uma série de conectores a vários tipos de bases de dados externos, eu trago esses dados para dentro do ambiente de Power BI e dentro desse ambiente de Power BI eu faço a modelagem do banco de dados.

[03:35] Eu junto as tabelas, eu relaciono informações que são linkadas, são conectadas, mesmo que venham de fontes diferente e depois eu faço funções e transformações de dados que eu posso fazer direto, ou seja, através de alguns comandos dentro do próprio menu do Power BI, na parte de modelagem, ou então utilizo a linguagem DAX que vem do Data Analysis eXpression.

[04:10] Então por exemplo, quando eu estou lá manipulando o meu dado eu tenho por exemplo uma série de funcionalidades, aqui tem uma figura mostrando algumas delas, então dividir coluna, formatar coluna, tem algumas funções que eu não consigo ver, mas funções de estatísticas, científicas, trigonometria, inverter linhas e assim por diante. E que eu posso fazer essas transformações.

[04:40] Ou então eu posso fazer fórmulas usando funções DAX, Data Analysis eXpression, agora o DAX ele é o que? Ele é complicado? É difícil de aprender? Se você já conhece Excel e já trabalha com formas do Excel vai ser fácil você aprender o DAX, não é a mesma coisa, tem muitas diferenças, mas o princípio do DAX é muito parecido com formas do Excel, eu vou construir operações em cima das colunas e linhas da tabela.

[05:27] Da mesma maneira que no Excel eu construo transformações, fórmulas em cima de linhas e colunas da planilha, só que no caso do Power BI eu tenho a tabela já carregada dentro do ambiente Power BI e aplico funções DAX ali naquela informação tabular que está a minha disposição. Então assim as formas do DAX são bem parecidas com a do Excel, o processo de criação delas é praticamente o mesmo, você digita o sinal de igual, seguindo o nome da função e coloca ou de uma expressão mais complexa.

[06:15] E você depois coloca dentro dos parênteses os argumentos necessários, ou seja, bem semelhante como se faz no Excel, assim como no Excel o DAX ele fornece diversas funções que você pode usar para trabalhar com por exemplo caracteres textos, trabalhar com cálculo usando por exemplo datas, usando horas, criar valores condicionais, se acontecer isso faz uma coisa, se não, faz a outra.

[06:47] Ou seja, ele fornece uma gama enorme de funções que você pode trabalhar, porém um detalhe, as fórmulas do DAX são diferentes do Excel em alguns aspectos, por exemplo, a função DAX sempre faz referência a uma coluna completa de uma tabela, por exemplo, no Excel você sempre faz referência a sua fórmula a uma célula, você pega o valor de uma célula e soma com o valor de outra célula, depois se você quiser repetir isso dentro de todas as linhas você copia a forma de uma célula e traz para as outras.

[07:35] No DAX não, eu não tenho uma formula célula a célula, eu tenho uma formula que está associada a coluna e tudo que eu disser vai se refletir naquela coluna, você pode por exemplo dentro de uma coluna fazer algumas fórmulas que vão se referenciar a apenas alguns valores delas, assim você adiciona o que nós chamamos de filtros nas fórmulas, ou seja, dentro dos parâmetros da fórmula do DAX você aplica filtros.

[08:07] Vamos dar um exemplo hipotético, digamos que eu queira criar uma coluna que seja uma fórmula matemática, mas que só vai ser aplicada se um campo condicional por exemplo for valor A, então eu aplico nessa coluna a fórmula DAX que teoricamente vai ser aplicada para todo mundo e depois eu aplico um filtro sobre o resultado dela dizendo, olha essa fórmula DAX você só vai aplicar para linhas cuja o valor da outra coluna que eu estou usando como referência seja valor igual a A.

[08:47] Então você tem maneiras de personalizar os cálculos linhas a linhas, ou por exemplo nós podemos fazer cálculos que são somatórios das linhas, cálculos que por exemplo uma linha se referêcia a linha anterior e assim por diante. Outra coisa interessante é que as funções DAX elas fornecem formas de inteligência do tempo, o que significa isso? Eu posso criar cálculos usando intervalos de datas, do tipo, não somente quantos dias eu tenho entre duas datas, quantas horas eu tenho entre dois horários.

[09:28] Mas também cálculos do tipo, qual é o valor das vendas no ano 2019 a variação em relação ao ano de 2018, ou seja, variações ao longo do tempo que eu posso estar aplicando com algumas funções de DAX. Uma outra coisa importante que podemos estar falando sobre a diferença entre o DAX e as funções de Excel é que muitas das funções às vezes tem o mesmo nome.

[10:05] O mesmo nome da função que eu uso no DAX eu uso no Excel, mas normalmente os parâmetros que eu vou encontrar internamente vão estar diferente, justamente porque as funções DAX estão sempre se referenciando a toda linha ou a toda a tabela. E também eu posso ter em algumas funções cujo nomes são iguais, mas os tipos de dados que eu coloco como parâmetro, ou os tipos de dados que eu tenho como retorno podem serem um pouco diferentes.

[10:40] Outra coisa, o DAX ele possui funções de matriz que são semelhantes as funções de pesquisa do Excel, aquelas funções por exemplo PROCV funções que te pedem para achar determinada condição dentro de uma coluna que eu tenho no Excel, no DAX também eu tenho funções parecidas desse tipo, mas que realmente são mais poderosas do que eu tenho no Excel, porque elas se aplicam não somente a um intervalo de células que eu tenho que especificar nessas funções, mas isso vai ser aplicado em toda coluna.

[11:20] De toda extensão da tabela, cujo o limite vai ser o número de linhas que a tabela está servindo como fonte, para o Power BI possui. Então é isso que eu queria falar um pouco sobre o DAX e o Excel, então vamos agora para os nossos vídeos práticos desse treinamento, quando vamos falar bastante de carga de dados, de transformações que eu posso fazer através dos menus, e o uso de algumas funções DAX.

[11:52] Infelizmente, pelo tamanho do curso não vai ser possível percorrer todo o espectro de funções DAX existente, só vamos estar falando claro, das mais importantes.

Aula 02 – Leitura de Dados

[00:00] Vamos continuar estudando? Nós vamos, nessa aula, mostrar a vocês a carga de dados do Power BI e nós vamos usar como fonte, diversos lugares diferentes. Claro que eu não teria tempo hábil de mostrar a vocês um exemplo de carga de dados de todas as fontes disponíveis no Power BI. Teríamos que fazer horas e horas de aulas.

[00:25] Nós vamos estar mostrando as fontes de dados, talvez, mais usadas. Vamos começar por ler dados de um arquivo do formato CSV. Para quem não sabe, o arquivo CSV é um arquivo texto onde tenho informações separadas por ponto e vírgula “;” onde cada informação separada por “;” representa uma coluna.

[00:53] Aqui no link desse vídeo você tem a possibilidade de baixar um arquivo na sua máquina e eu vou abrir esse arquivo para vocês. No caso da minha máquina, ele está aqui: “...\\Download\\Tabela_Clientes.csv”. Note que,

no meu Windows, ele coloca o símbolo do Excel porque, por padrão, um arquivo no formato CSV eu consigo abri-lo pelo Excel.

[01:27] Vamos olhar o conteúdo desse arquivo, mas não usando o Excel, vamos usar um editor de texto. Então faça o seguinte: botão direito do mouse sobre o arquivo, “Abrir com” e eu vou escolher “Bloco de Notas”. Note: isso aqui é um arquivo CSV. Nós temos, na primeira linha, um cabeçalho onde tenho o nome das colunas e, depois, os dados separados por “;”.

[02:12] Eu tenho esse formato CSV. Vamos lê-lo para o Power BI. Vou fechar aqui e abrir o Power BI. No caso, vou usar o Power BI Desktop. Tenho aquela página inicial de apresentação, vou fechar ela e eu tenho a minha interface do Power BI.

[02:45] Nós vamos vir aqui e clicar em Página Inicial e eu tenho a opção Obter Dados. Ao clicar na setinha que está bem do lado de Obter Dados eu tenho alguns tipos de dados, digamos assim, mais usados e eu vou escolher essa opção: “Texto/CSV”. Ele vai me perguntar a localização do arquivo, então já vou escolher o arquivo “Tabela_Clientes.csv” e clicar em Abrir.

[03:32] Ele me apresenta o conteúdo do arquivo já em formato de tabela. Note que ele colocou aqui como título da tabela aquela primeira linha, cada coluna separada por vírgula, e temos o conteúdo. Já podemos notar um problema na carga desse arquivo. Note que tenho aqui, nos nomes, um símbolo estranho. Aqui, por exemplo, era para ser “César Mattos” e tem esse “Ã©”, aqui era para ser “Heitor Beltrão”. Enfim, por que isso?

[04:18] Se vocês não sabem, internamente todos os caracteres dentro do sistema operacional utilizam uma tabela que chamamos de tabelas ASCII. Essas tabelas ASCII convertem, para o número hexadecimal interno do computador, cada símbolo que você vai usar.

[04:39] Quando os primeiros sistemas operacionais foram construídos, as primeiras tabelas ASCII só comportavam os caracteres da língua inglesa. Na língua inglesa os caracteres de acentuação (Ex.: á, Â, ç, Ç) não são comportados. Então, com o tempo, foi-se criando outras tabelas ASCII que comportar caracteres diferentes da língua inglesa.

[05:15] Quando falo de outras tabelas ASCII, não estou falando somente da língua latina. Por exemplo, se você estivesse usando o Power BI na Grécia, teríamos o alfabeto grego que possui símbolos diferentes do alfabeto que utilizamos no português. Ou o alfabeto russo, ou até mesmo, se fosse uma

língua oriental com aqueles ideogramas diferenciados do chinês, do japonês, aqueles símbolos que achamos que é mentira que eles entendem.

[05:50] O que acontece é que o Power BI pegou para mim um padrão. O padrão que ele usou está aqui em cima: “1252: Europeu Ocidental (Windows)”. Não sei bem o que significa essa tabela ASCII, mas ela não se adaptou bem para os símbolos que estão vindo em português. Preciso mudar isso.

[06:13] Felizmente, já sei qual é a resposta. Tenho aqui várias tabelas ASCII: alemão, árabe, chinês, coreano, e eu tenho essa aqui: “65001: Unicode (UTF-8)”. Essa tabela UTF-8 comporta os caracteres de acentuação da língua portuguesa. Então vou escolher ela. Note que, ao fazer isso, tenho os acentos de maneira correta. Vou carregar os dados.

[07:12] Os dados foram carregados. Note que no meu canto direito ele já escreveu o nome da tabela. Esse nome é pego de uma característica específica do arquivo-fonte. No caso do CSV, ele usa o próprio nome do arquivo CSV como nome da tabela.

[07:37] Se vocês observarem, em alguns campos ele colocou este símbolo que representa o tipo do campo. Quando não tenho nada, o campo é de caracteres textos. Quando tenho o símbolo de somatória significa que aquele caractere é um número. Se notarmos, ele disse que é número tanto o CPF quanto o CEP, e sabemos que isso pode causar problemas. Por que isso pode causar problemas?

[08:20] Deixa-me abrir um editor de texto, o Bloco de Notas. Por que é o seguinte: digamos que eu tenho o CPF “12345678901”. Se eu converto esse conjunto de caracteres para números, vou ter um número idêntico “1234567801”. Até aí tudo bem, não vi problema. No caso, estou vendo CPF como número.

[09:06] Agora, se eu tiver meu CPF assim: “00456789003”, quando eu converter esse CPF para número, vou ter isso: “456789003”. Ou seja, esses dois primeiros zeros (00) que estão na frente do número do CPF vão sumir porque nenhum número pode começar com “00” na frente. Então há um problema: o Power BI, ao converter o CPF para número, ele pode suprimir informações que são importantes.

[09:54] A mesma coisa vale para o CEP. Eu posso ter CEPs que começam com “00”. Então eu preciso mudar isso, preciso dizer para o Power BI: “Olha, Power BI, essa coluna é um caractere e não um número.”

[10:10] Há um conceito que precisamos entender. Quando tenho essa frase: “MINHA IDADE É 40 ANOS ...”, meus caracteres são esses mesmos, mas esse número “40” é considerado um caractere. O número “4” não é um número, é um caractere “4”. Porque quando ele é número, eu posso fazer operações matemáticas com ele. Quando é um caractere, eu não posso fazer operações matemáticas com ele. Então tenho que converter, para texto, a coluna CEP e CPF, mas isso vamos fazer no próximo vídeo.

[00:00] Vamos, então, converter as colunas CEP e CPF, que foram carregadas como números, para caracteres. Eu vou clicar em Página Inicial e escolher esta opção: Editar Consultas. O Editar Consultas vai pegar toda a transformação que ele fez ao carregar os dados de CSV para o formato atual e eu posso embutir novas transformações.

[00:42] Então eu vou ter a minha Tabela_Clientes com as colunas e seus conteúdos e note que, onde tenho CPF, tenho do lado um símbolo “123”. Esse símbolo me diz que ele converteu isso para um número inteiro. Aqui do lado, por exemplo, tenho esse símbolo “ABC” que diz que essa coluna é um texto.

[01:15] Se formos olhar aqui veremos que, na coluna CEP, ele também especificou como um número inteiro. Note que na coluna LIMITE_DE_CREDITO ele colocou “1.2”, que significa que é um número decimal. Tanto é que aqui eu tenho “85000,9”, que seria 85 mil e 90 centavos.

[01:51] Vamos fazer então a conversão. Ela é muito simples. Eu clico sobre o número que está sobre o formato da coluna e tenho os tipos. Então vou converter ele para: “Texto”, para converter a coluna, que é um número, em texto.

[02:19] Nessa caixa de diálogo, vou escolher que tipo de transformação eu vou fazer: se vou aplicá-la sobre a coluna atual que já existe, ou se vou adicionar uma nova etapa. Vou escolher a opção “Substituir atual” e, ao fazer

isso, notem que a linha 5, que tinha um CPF que começava com “00”, agora está aparecendo. Então eu fiz minha transformação de forma correta.

[02:52] Eu vou repetir esse processo, também, para a coluna CEP. Vou clicar sobre “123”, vou escolher “texto” e “Substituir atual”. Então eu já converti as colunas CPF e CEP, que originalmente o Power BI sugeriu que fossem números, em textos.

[03:22] Vamos entender um pouquinho mais essa tela. Temos, aqui em cima, um menu onde podemos fazer uma série de transformações nos dados. Nós não vamos ver isso nesse momento. Mais à frente, em outros vídeos, estaremos olhando sobre essas funções. Do lado direito, nós temos o nome da nossa tabela. Lembra que eu falei para vocês que ele já sugere um nome que está associado, no caso do CSV, ao nome do arquivo texto? Então eu vou manter esse “Tabela_Clientes”.

[04:00] Aqui em baixo eu tenho as etapas aplicadas na transformação. Note que tenho três linhas escritas: Fonte, Cabeçalhos Promovidos e Tipo Alterado. Cada uma dessas linhas representa fases da transformação desde a leitura do arquivo até o produto final, que é o que estou vendo agora aqui.

[04:29] Se eu clicar sobre Fonte, notem, vou ver o conteúdo do arquivo CSV original. Note que, por exemplo, na primeira linha, eu tenho os nomes dos títulos e note também que as colunas ainda não têm nomes. Essa é a forma com que ele lê o arquivo. Já na segunda opção: Cabeçalhos Promovidos, é onde ele especifica que essa primeira linha tem que subir e vir aqui para cima. Se eu clicar no segundo, note que agora tenho o arquivo com as colunas com seus nomes.

[05:15] Observe que CPF ainda está como texto, está correto, mas todo mundo está como texto. Por exemplo: DATA_DE_NASCIMENTO ainda está como texto, VOLUME_DE_COMPRA também. Quando eu seleciono Tipo Alterado é que o Power BI, ou seja: entre a fase 2 e a fase 3, converte as colunas para o tipo que ele acha ser o correto.

[05:54] A diferença agora é que nessa fase 3, por exemplo, VOLUME_DE_COMPRA (que ele converteu para um número inteiro), ou LIMITE_DE_CREDITO (que ele converteu para um número decimal), isso o Power BI fez sozinho, mas eu interfeirei nesse terceiro passo para dizer ao Power BI que: CPF, que originalmente tinha sido carregado como número, tem que ser carregado como texto. A mesma coisa para o CEP.

[06:28] Então, nesse momento, eu tenho essa carga. A princípio estou satisfeito com ela, então venho aqui em Página Inicial e seleciono “Fechar e Aplicar”. Ele vai fazer a carga de novo e note, na direita, que aqueles símbolos que tínhamos aqui de somatório, tanto em CEP quanto em CPF, sumiram porque agora esses dois caracteres são textos e, não mais números.

[07:08] Vamos fazer o seguinte: salvem esse arquivo na máquina de vocês vindo em Arquivo, Salvar Como, e coloquem um nome qualquer. Porque é interessante, a cada final de vídeo, irmos salvando os arquivos, pois no vídeo seguinte, caso você tenha algum problema, algum erro, você consegue recuperar o status do vídeo anterior correto para tentar repetir os passos novamente. Então eu vou salvar ele como: “Carga_Dados_002.pbix”.

Utilizem o nome que vocês queiram.

[00:00] Vamos fazer agora uma leitura de dados de uma planilha de Excel. Para entendermos como o Power BI entende uma planilha de Excel, vamos fazer o seguinte: baixe o arquivo que está relacionado com o link deste vídeo, salve na máquina, e vocês vão ter duas planilhas: “Tabela_Vendedores.xlsx” e

“Tabela_Vendedores_Formatada.xlsx”. Ignorem a segunda planilha por enquanto.

[00:44] Abram, com seu Excel, a opção “Tabela_Vendedores.xlsx”. Então vou dar um duplo-clique e eu tenho uma planilha onde na primeira linha os títulos de cada coluna e os dados de cada coluna: MATRICULA (número), NOME. A DATA_ADMISSAO está em um formato estranho, isso aqui não é bem um número que eu conheça, mas, na verdade, é um campo Data. Nós temos aqui as informações.

[01:29] Vamos mexer um pouquinho nessa tabela. Primeiro, vou pegar esse campo data, selecionar essa área e vou formatar ele como uma data. Vou pegar essa área onde tenho a planilha e vou “Formatar como Tabela”.

[01:57] Por que estou fazendo isso tudo? Não é necessário, claro, fazer essas transformações dentro do Excel para o Power BI ler uma planilha, mas eu quero simular algumas situações nessa planilha para mostrar a vocês como o Power BI enxerga uma planilha de Excel. Então vamos continuar. Eu vou formatar essa área como tabela, então isso aqui virou uma tabela interna do Excel.

[02:30] Vou vir com o botão direito do mouse sobre Planilha1, vou “Mover ou Copiar” e criar uma outra aba clicando em “Criar uma cópia”. Vou clicar com o botão direito do mouse sobre essa nova pasta e vou copiar ela para Planilha2. Então eu tenho o mesmo conteúdo tanto na aba Planilha1, quanto na aba Planilha2.

[03:14] Vou criar uma nova aba que vou chamar de Planilha3. Nessa nova aba, vou fazer o seguinte: vou copiar dados da Planilha1 ou da Planilha2, tanto faz, dar “CTRL + C” e na Planilha3, vou escolher “Colar Valores”. Note que a data de admissão voltou a vir com esse número não formatado, não tem problema.

[03:59] Vou fazer outra coisa: vou selecionar essa área de novo, mas ao invés de formatar como tabela, como fiz nas outras planilhas, eu vou vir aqui nessa área e dar um nome para esse conjunto de dados. Vou chamar de “Conjunto_Dados” e dar um Enter. Para quem não sabe, quando faço isso, Conjunto_Dados é essa área aqui.

[04:30] O que nós temos nessa planilha? Temos 3 abas, 2 tabelas, e temos na aba 3 dados sem ser tabelas, mas associados a um conjunto de dados. Note que, no Excel, ele escreve “Tabela13” que é, justamente, a tabela da aba

Planilha2. O que é a “Tabela1”? É a tabela da Planilha1 original, que foi aquela que nós formatamos.

[05:17] Eu fiz isso tudo para dizer para vocês o seguinte: quanto o Power BI for ler essa planilha, ele vai capturar 3 tipos de coisas: abas, tabelas e a terceira coisa que ele vai capturar é conjunto de dados. No Power BI eu vou dizer: “Eu quero carregar o que: dados da tabela, da aba ou de um conjunto de dados?”. Inclusive posso carregar vários ao mesmo tempo.

[06:00] Esta planilha de Excel eu vou salvar com outro nome: “Tabela_Vendedores_Formatada.xlsx”. Essa segunda planilha que vocês baixaram, simplesmente, é a planilha que eu fiz a transformação.

[06:25] Caso você não tenha o Excel, ou não quis seguir os passos que eu fiz, não tem problema, vocês vão carregar no Power BI a tabela “Tabela_Vendedores_Formatada.xlsx” que vocês baixaram. Se vocês fizeram tudo que eu fiz agora no Excel, salvem ou aqui, ou com outro nome, não importa. Eu vou salvar em cima e vai ser essa planilha “Tabela_Vendedores_Formatada.xlsx” que eu vou carregar.

[06:59] Então eu fiz tudo isso no Excel só para vermos como é que o Power BI enxerga a planilha. Vamos voltar para nosso Power BI Desktop. Estou com o Power BI que salvei no final do vídeo passado com a Tabela_Clientes carregada. Agora vou vir em Página Inicial, Obter Dados, vou escolher Excel e vou selecionar a minha “Tabela_Vendedores_Formatada.xlsx”.

[07:47] Ao fazer isso, note: ele me mostrou 6 coisas que achou na planilha. Pelo símbolo que está na frente, eu consigo descobrir o que é. Esse símbolo representa tabelas, então ele achou, dentro daquela planilha, 2 tabelas. Ele achou, também dentro da planilha, três abas. Se vocês notarem, o ícone tem, em baixo, o símbolo de uma aba, e o Power BI achou um conjunto de dados.

[08:32] Eu posso escolher o que eu quiser para carregar. Se eu selecionar Tabela1 consigo ver, do lado direito, o que ele vai carregar. Note, inclusive, uma coisa: quando seleciono Tabela1, veja que DATA_ADMISSAO está no formato de Data. Lembra que, quando transformei a planilha original em tabela, eu formatei, no Excel, a data? Então ela já veio formatada aqui.

[09:04] Quando escolho o conjunto de dados, veja que a DATA_ADMISSAO não vem formatada. Vamos usar o pior caso: vamos carregar do conjunto de dados que não tem nenhum formato. Tudo aquilo que fiz antes foi só para mostrar para vocês as três formas em que os dados são visualizados pelo

Power BI no Excel, mas vamos carregar do conjunto de dados, então vou selecionar esta opção e vou carregar.

[09:50] Aqui do lado eu, agora, tenho meu Conjunto_Dados. Veja: ele pegou o nome automaticamente daquela opção que apareceu. Quando criamos o conjunto de dados no Excel, nós demos esse nome. Então foi ele que ficou para a tabela. Aqui embaixo temos a Tabela_Clientes.

[10:12] Vamos alterar a transformação da informação Conjunto_Dados. Vou clicar em Página Inicial, Editar Consultas, e note que agora aqui do lado eu tenho as duas tabelas: Tabela_Cliente e Conjunto_Dados. Vou clicar sobre o conjunto de dados e a primeira coisa: vou mudar esse nome. Conjunto_Dados não é apropriado. Então, em nome, vou colocar “Tabela_Vendedores”.

[10:57] Outra coisa: DATA_ADMISSAO não é para ser número, e sim data. Então eu cliço sobre “123”, digo que é uma Data e “Substituir Atual”. Note: ele fez a mesma coisa que eu faria no Excel quando formatei ele para cá. A matrícula, apesar de estar com “235”, “236”, “237” e “238”, como é um código, eu vou transformar ele também para texto.

[11:49] Então eu tenho minha formatação correta, cliço em “Fechar e Aplicar” e já tenho minhas duas tabelas importantes: Tabela_Clientes e Tabela_Vendedores. Antes de terminar vou vir em Arquivo e Salvar Como “Carga_Dados_003.pbix”.

[00:00] Agora nós vamos fazer uma carga de Folder. O que significa isso? Vamos entender o problema primeiro. Baixem, novamente, arquivos que estão relacionados abaixo do link do vídeo. Vocês vão baixar esse arquivo: Carga_Folder.zip.

[00:27] Vou descompactá-lo aqui mesmo e vou ter um diretório chamado “Diretório” e um arquivo chamado “Tabela_Produtos3.csv”. Vamos esquecer esse CSV primeiro e vamos entrar nesse diretório. Eu tenho 2 arquivos: Tabela_Produtos1.csv e Tabela_Produtos2.csv. Vou abrir Tabela_Produtos1.csv com meu editor de texto para ver o conteúdo dele e vou abrir esse segundo, também com editor de texto.

[01:19] Se colocarmos eles um perto do outro, note que eles possuem os mesmos formatos: possuem os mesmos cabeçalhos, as mesmas colunas, mas o conteúdo é diferente. Aqui eu tenho uma lista de produtos e aqui eu tenho outra lista de produtos. Então esse é o caso, por exemplo, em que tenho vários arquivos CSVs onde cada arquivo tem um pouco do cadastro que eu quero carregar. É importante que os dois arquivos tenham o mesmo formato.

[01:59] Eles estão gravados dentro de um único diretório. Isso é uma coisa muito normal, que pode acontecer, de eu ter um diretório onde arquivos vão sendo colocados dentro. Cada arquivo, por exemplo, está vindo de uma fonte diferente: um arquivo que vem do “Sistema A”, outro que vem do “Sistema B”. Esses arquivos vão se acumulando nesse diretório e eu posso, depois, fazer uma carga completa deles através da carga por Folder.

[02:35] Vou voltar para o Power BI, naquele arquivo que salvamos no vídeo anterior. Vou vir em Pagina Inicial, Obter Dados e vou clicar na opção “Mais...” aqui de baixo. Ele vai me mostrar essa caixa de diálogo com todas as fontes que o Power BI pode ler. Tenho aqui diversas fontes. Quando falei para vocês que não íamos olhar todas essas fontes nesse vídeo, acho que agora conseguem entender o porquê.

[03:20] Vamos usar essa opção: Pasta. Eu, às vezes, costumo chamar de Folder por que é o nome em inglês, mas é pasta, diretório. Vou clicar em Pasta, Conectar e vou dizer qual é a pasta que vai ser a fonte. Vou clicar em Procurar e vou selecionar aquela pasta “Diretório”. Vocês vão escolher o diretório onde salvaram o arquivo que descompactaram através do link que foi feito download.

[04:20] Então clico em “Diretório”, vou clicar em “OK”, tenho o caminho completo e clico novamente em “OK”. Ele vai me mostrar quantos arquivos ele achou dentro do Folder. Ele achou 2 arquivos: Tabela_Produtos1.csv e Tabela_Produtos2.csv, e tenho os caminhos.

[04:47] Tenho a seguinte opção: “Combinar e Transformar Dados” e “Combinar e Carregar”. Simplesmente, Combinar vai combinar os dois arquivos. Se eu escolher “Combinar e Transformar Dados”, ele vai abrir a área de transformação, caso eu queira fazer alguma modificação, e “Combinar e Carregar” vai carregar direto os arquivos. Então vou escolher “Combinar e Transformar Dados”, vai me mostrar o formato final.

[05:25] Relembrando: ele tem, em “Arquivo de Amostra”, “Primeiro arquivo” ou eu posso olhar arquivo por arquivo existente no diretório. Vou escolher, em “Origem do Arquivo”, “65001: Unicode (UTF-8)” no “Primeiro arquivo”. Vamos olhar a Tabela_Produtos1.csv, como ela vem. Vou dizer que ela é UTF-8. A Tabela_Produtos2.csv, também, e vou dar OK.

[06:16] Ele vai buscar, do diretório, os dois arquivos. Vai juntar um com o outro e disponibilizar para mim. Caso não aconteça nada, eu tenho aqui: “Existem alterações pendentes em suas consultas que ainda não foram aplicadas”. Vou clicar em “Aplicar alterações” e ele, agora, vai fazer a carga.

[06:38] Note que, ali no canto, ele criou uma tabela chamada Diretório com várias colunas. Se formos em Pagina Inicial, Editar Consultas, eu tenho os meus dados onde tenho minhas transformações. Porém, diferentemente das outras, ele coloca o nome da tabela origem. Vou, então, mudar para Tabela_Produtos. Vou “Fechar e Aplicar” e pronto, já tenho as minhas 3 tabelas carregadas no meu arquivo. Relembrando, vamos salvar nosso Projeto: “Arquivo”, “Salvar Como”, e eu vou salvar o projeto como “Carga_Dados_004.pbix”.

[00:00] Vamos fazer um exemplo mostrando como que o Power BI tem um dinamismo na leitura dos dados quando ele vai buscar arquivos dentro de um Folder (diretório). Então, vocês se recordam que, nesse projeto aqui, nós temos a nossa “Tabela_Produtos” que está vindo de um diretório. E nesse diretório, vamos abrir aqui, eu tenho 2 arquivos: “Tabela_Produtos1.csv” e “Tabela_Produtos2.csv”.

[00:36] Pois bem, nós vamos fazer o seguinte. Para podermos ver isso acontecendo, eu vou clicar nesse componente aqui: “Gráfico de barras empilhadas”. Vou colocar o gráfico de barras aqui. Não é nosso objetivo, nesse treinamento, ver com detalhes a construção de relatórios, mas, para vermos esse dinamismo, nós vamos criar um relatório aqui de maneira rápida.

[01:08] Arrastei aqui o gráfico de barras empilhadas e, em “Tabela_Produtos”, eu vou colocar SABOR no Eixo e também no Valor. Vou tirar essa coluna Filtros para ficar um pouco maior. Eu coloquei SABOR no Eixo e SABOR como contagem (Valor). Se eu colocar o mouse aqui, eu vejo que o SABOR Manga tem 6 produtos, o SABOR Laranja tem 5, o SABOR Melancia tem 4 produtos e assim por diante.

[01:55] Ok, o que eu vou fazer agora? Eu vou vir aqui no meu arquivo onde eu tenho o Diretório. E eu tenho outra tabela chamada “Tabela_Produtos3.csv” que tem o mesmo formato da 1 e da 2. O que eu vou fazer é: vou pegar esse cara, vou copiar e colar aqui dentro. Agora, dentro do diretório da tabela, eu tenho três arquivos e não só dois.

[02:31] Se eu voltar para o Power BI, eu posso vir aqui em Página Inicial e Atualizar. Quanto eu clicar em atualizar, o Power BI vai fazer uma nova carga. E quando ele for fazer carga de todos os arquivos que estão dentro do Diretório, ele vai encontrar 3 arquivos e vai carregar, agora, os 3 e não apenas os 2 que haviam anteriormente.

[03:02] Então, se eu vier aqui em Página Inicial e Atualizar. Note que ele está atualizando todas as 3 tabelas: Tabela_Clientes, Tabela_Vendedores e Tabela_Produtos. Não sei se vocês notaram, houve uma mudança no gráfico.

Você vê que, antes, Manga estava lá no topo e agora é Laranja. Laranja eu tenho 8, Manga eu tenho 6, depois vem Melancia 5 e assim por diante.

[03:30] Ou seja, quanto fazemos carga de Folder, há um dinamismo. E, conforme vão entrando novos arquivos, nós vamos atualizando o conteúdo na medida em que for dando refresh. Era esse ponto específico que eu queria mostrar nesse vídeo para vocês.

[00:00] Quando trabalhamos com arquivos externos, que é o caso de CSV, que nós já carregamos, arquivos Excel, Folder ou outros tipos de arquivos que nós nem vamos ver neste curso como, por exemplo, arquivos XML e arquivos JSON. São coisas que estão localizadas em um diretório específico seja da minha máquina, seja de um diretório de rede de uma empresa onde estou buscando essas informações.

[00:30] Isso está registrado aqui na Carga, mas pode acontecer de eu, por algum motivo, mudar a posição desses arquivos. Como eu faço para resolver esse problema, caso isso aconteça com meu projeto de Power BI?

[00:46] Vamos fazer um exemplo na prática. Então voltei ao status atual do meu projeto. Tenho aquele meu gráfico de barras que eu criei. Mas ele não vai ser importante para este nosso teste. Eu simplesmente vou vir aqui em Página Inicial, vou clicar em Editar Consultas, e eu tenho aqui no canto as três transformações que nós temos até agora: a Tabela_Clientes, Tabela_Vendedores e Tabela_Produtos.

[01:20] Se eu clico na tabela de vendedores, note que eu tenho todos os passos que nós fizemos: a leitura, a navegação, promover os cabeçalhos e Tipo Alterado normalmente é onde eu coloco os tipos das colunas.

[01:41] Se eu clicar sobre Fonte, note que eu tenho aqui, uma fórmula que se aplica à tabela toda. Ainda não vamos falar sobre o que é esta sintaxe desses comandos, mas se clicarmos sobre este botão aqui, conseguimos expandir esse comando e ver ele com um detalhe um pouco maior.

[02:12] Então note: eu tenho aqui um comando do tipo "Excel.Workbook" que está buscando um arquivo neste diretório grande com uma série de parâmetros que não vêm ao caso. Agora notem que a localização do arquivo está expressa nesse comando. Eu vou fazer o seguinte: vou fechar esse Projeto e, sem fechá-lo, vou vir aqui no diretório. Esse diretório depende de onde você salvou esses arquivos, mas a fonte daquele dado é esse arquivo "Tabela_Vendedores_Formatada.xlsx".

[03:12] E eu vou fazer o seguinte: vou movê-lo, vou recortar ele e copiar aqui, por exemplo, para esse meu diretório "C:\Tempo". Voltando ao Power BI, se eu clicar em Página Inicial e Atualizar, ele vai me mostrar esta mensagem que não foi possível localizar o arquivo porque ele mudou de posição.

[03:48] Como é que eu resolvo, então, esse problema? É simples: eu vou de novo em Página Inicial e Editar Consultas. Eu tenho a minha consulta de vendedores aqui (Tabela_Vendedores) e tenho, aqui, o meu comando. Mas tenho que, primeiro, clicar na etapa Fonte, abrir, e eu tenho o caminho. Mas o caminho, agora, é o "C:\Tempo". Foi onde gravei o arquivo.

[04:22] Então vou copiar esse caminho: seleciono e dou um "Ctrl + C" e vou alterar o caminho no comando, menos o nome do arquivo, e vou colar. Agora estou dizendo que o arquivo está localizado em um outro diretório. Vou "Fechar e Aplicar". Note que ele, agora, está carregando a tabela de vendedores que estava pendente e agora deu tudo certo. Se eu fizer de novo Atualizar não vou ter o erro durante a carga da tabela de vendedores.

[05:04] Então muito cuidado. Quando trabalhamos com projetos onde estou olhando arquivos externos em um determinado diretório, é importante que eles permaneçam naqueles diretórios. Se, por algum motivo, eles forem movidos daquele diretório origem, temos que voltar no Power BI e modificar o caminho da localização onde os arquivos se encontram.

[00:00] Vamos mostrar como fazemos a conexão do Power BI com o banco de dados. Aqui, no caso, vamos usar o banco de dados SQL Server da Microsoft. Se você já tem o SQL Server instalado na sua máquina, você não precisa nem seguir os passos desse vídeo. No caso, eu vou ter que instalar o SQL Server e, também, um outro software chamado SQL Server Management Studio que faz a interface com o banco de dados.

[00:33] Vou abrir o browser e procurar por "SQL SERVER 2017". Temos este link "Downloads do SQL Server | Microsoft". Tenho aqui algumas opções e vou procurar "SQL Server Developer", que é uma versão de desenvolvimento em que posso usar desde que não coloque esse SQL Server em produção. Vou clicar em Download agora mesmo. O download vai ser inicializado e tenho o programa a ser executado.

[01:29] Ele vai ter três níveis de instalação: Básico, Personalizado ou Baixar Mídia. Vou usar o Personalizado. Ele vai me dizer o diretório onde vai ser instalado. Vou clicar em Instalar. Pronto, vou ter esta caixa de diálogo. Talvez, para você, a coisa não vá ser tão rápida como foi meu caso. Isso por que eu já tinha, na máquina, o instalador. O que vai acontecer no seu caso é

que ele vai fazer o download do instalador e, depois que terminar, vai abrir essa caixa de diálogo.

[02:13] Vamos continuar. Vou clicar em Instalação e selecionar "Nova Instalação Autônoma do SQL Server". Ele vai me especificar qual é a versão que vou usar: Developer. Vou clicar em Avançar e aceitar os termos de licença. Ele vai checar por updates antes de fazer a instalação. Pode ser que demore um pouquinho esse processo. Vou parar o vídeo um instante e volto quando terminar.

[02:53] Voltei. Ele me mostra aqui se a instalação tem os requisitos necessários. Vou clicar em Avançar. Aqui vou escolher o que vou instalar. No meu caso, só vou instalar "Serviços de Mecanismo de Banco de Dados". Deixa eu ver se tem mais alguma coisa interessante que eu precise instalar. Essa aqui: "Conectividade das Ferramentas de Cliente". Vou clicar em Avançar.

[03:40] Ele vai me mostrar o nome da Instância. Instância seria "ambiente". Eu poderia ter, numa mesma máquina, diversos ambientes diferentes de banco de dados. Vou clicar em Avançar. Aqui avanço novamente. Nesse ponto eu digo a forma de conexão do banco de dados. Vou escolher a "Modo Misto (autenticação do SQL Server e do Windows)". Existe um usuário chamado "sa", que é o administrador do SQL Server.

[04:16] Vou colocar qual é a senha desse usuário. Vou repetir, em baixo, a senha. Aqui vou dizer qual é o Usuário que vai ser responsável por administrar o ambiente. Vou clicar nesse botão "Adicionar Usuário Atual". Ponto, vou clicar em Avançar. As senhas digitadas são incompatíveis. Na verdade, essa senha precisa ter caracteres diferentes. Por exemplo: arroba (@), maiúsculos, misturar nome com números. Então vou colocar uma outra senha que, talvez essa, o SQL Server permita. Agora sim.

[05:21] Ele mostra tudo que vai ser instalado e eu clico em Instalar. A instalação é inicializada. Então vou parar um pouco o vídeo aqui e volto quando terminar. Pronto, fez a instalação do SQL Server. Vou fechar a caixa de diálogo e essa aqui também.

[05:42] Agora vou fazer a instalação do SQL Server Management Studio. Vou procurar ele aqui no browser: "sql server management studio". Vou clicar no link "Baixar o SQL Server Management Studio (SSMS)". Tenho esse link "Baixar o SQL Server Management Studio 18.2". Clico no link e ele vai fazer o download. Vou esperar um pouco e, daqui a pouco, volto quando o download terminar.

[06:24] Pronto, meu download terminou, executei o programa que foi baixado e tenho essa caixa de diálogo aqui. Vou clicar no botão Instalar e o SQL Server Management Studio, sua instalação, será inicializada agora. Bem, vou parar de novo o vídeo e volto quando terminar a instalação.

[06:48] Pronto, a instalação do Management Studio terminou. Vou clicar em Fechar. Vou procurar pelo "SQL SERVER". Aqui: "Microsoft SQL Server Management Studio 18". Vou clicar no link e eu entro, aqui, no Login. Vou selecionar "LOCALHOST", que é o nome do meu computador local. Aqui vou colocar o Usuário e a Senha que cadastrei durante a instalação. Clico em Conectar e já estou acessando o SQL Server local da minha máquina.

[00:00] Independente se você está usando o SQL Server ou instalou o SQL Server conforme o vídeo da aula anterior, vamos recuperar o backup da base de dados. O backup da base de dados é uma cópia dela, externa, que vamos incluir no nosso SQL Server. Para fazer isso, clique no link associado a esse vídeo, faça download do arquivo e salve na sua máquina. Eu, por exemplo, tenho o arquivo salvo aqui: "...\\Download". O nome dele é "SUCOS_VENDAS.BAK". Ele é o backup e vamos recuperar ele no SQL Server.

[00:44] Voltando para o SQL Server Management Studio, vou clicar com o botão direito do mouse sobre "Banco de Dados" e vou escolher "Restaurar Banco de Dados". Aqui vou clicar em Dispositivo e vou clicar nesses "..." que estão aqui nesse botão. Ao clicar aqui, eu clico em Adicionar e seleciono o arquivo no local onde salvei na minha máquina.

[01:18] Então estou com ele aqui: "SUCOS_VENDAS.BAK". Vou clicar nele, clico em OK, seleciono ele e clico em OK. Então eu tenho, agora, as propriedades do arquivo que será recuperado. Vou em Opções e clico em "Substituir o banco de dados existente (WITH REPLACE)" e clico em OK. O banco foi recuperado. Se eu abrir ele, consigo ver as tabelas da base de dados dentro do meu SQL Server. O que vamos copiar para o Power BI são estas duas tabelas: "dbo.NOTAS FISCAIS" e "dbo.ITENS NOTAS FISCAIS".

[02:12] Vamos agora para o Power BI. Eu o fechei por causa das instalações, então vou abri-lo de novo. Vou abrir o último arquivo que salvei para continuar a fazer as cargas dos dados. Vou procurar onde salvei ele. No caso ele está aqui: "...\\Projeto Inicial\\Carga_Dados_004.pbix". Pronto, carreguei o arquivo.

[03:00] Está com aquele gráfico que usamos para testar o dinamismo daqueles arquivos que fomos colocando dentro do diretório Folder. No canto direito tenho minhas três tabelas que foram carregadas anteriormente. Vamos

carregar, agora, as duas tabelas do bando de dados. Vou vir, então, em "Página Inicial", "Obter Dados" e vou selecionar a opção "SQL Server". Aqui vou colocar as propriedades de conexão: servidor é "localhost" e o banco de dados é "SUCOS_VENDAS".

[03:50] Eu tenho duas opções: ou eu uso o modo Importar, ou o modo Direct Query. O modo Direct Query faz uma conexão online com o banco. Ou seja, na medida que o banco for sendo modificado o dado é modificado no Power BI. Enquanto, no Importar, eu carrego tudo para dentro do relatório. Por enquanto vamos usar a opção Importar, então vou manter selecionado Importar.

[04:23] Vou clicar em OK e ele vai me mostrar algumas propriedades a mais da conexão. Então, por exemplo, vou manter "Usar minhas credenciais atuais". O servidor é "localhost" e vou clicar em Conectar. Ele vai dizer que não foi possível usar suporte à criptografia. Ignoramos isso.

[04:58] Ele vai me mostrar, aqui, justamente as cinco tabelas que tenho na base de dados e mais algumas informações importantes relacionadas com a base. Vou selecionar, então, "ITENS NOTAS FISCAIS" e "NOTAS FISCAIS". Vou selecionar esses dois arquivos. Na verdade, duas tabelas. Clico em Carregar e, agora, os dados estão sendo puxados do SQL Server para meu relatório do Power BI.

[05:39] Podemos ir em Página Inicial e Editar Consultas. Temos as nossas cinco consultas e podemos vir aqui e mudar. Vou chamar ela de "Tabela_Itens_Notas" e essa outra de "Tabela_Notas".

[06:13] Quando trazemos do banco de dados ele já traz os tipos automaticamente. Ou seja: na leitura do banco de dados, quando lemos direto uma tabela, os metadados (as características como os nomes das colunas, os tipos e os relacionamentos) já são herdados para o Power BI facilitando a nossa modelagem da base de dados.

[06:51] Pronto. Eu vou, então, "Fechar e Aplicar". Tenho, agora, as minhas cinco tabelas carregadas e vou salvar o arquivo como sendo "Carga_Dados_005.pbix".

[00:00] Vamos mostrar para vocês como consigo, no Power BI, carregar coisas que estão em páginas WEB. Para isso, vamos fazer o seguinte: no nosso Power BI, vamos clicar em Arquivo, Novo. Ou seja, vou abrir outro arquivo Power BI, que não vai ter relação com as nossas tabelas que estamos carregando desde o início dos vídeos dessa aula.

[00:31] Antes de continuar, vou abrir meu browser e digitar uma página: “<http://www.opendatani.gov.uk>”. Isso é um site inglês que possui uma série de estatísticas. Vou clicar em Datasets e ele vai me mostrar uma série de arquivos que posso acessar e vou escolher aleatoriamente um deles: “Northern Ireland Goods Vehicle Operator’s Licence Records (registros de licenças de veículos da Irlanda do Norte) ”.

[01:48] Então eu tenho aqui um arquivo do formato CSV. Ele me mostra o conteúdo do arquivo (que está carregando) e, em cima, eu tenho um link. Vou clicar sobre ele com o botão direito do mouse e vou “Copiar endereço do link”. Se eu abrir uma outra sessão de browser e colar o endereço do link ele, automaticamente, vai fazer o download de um arquivo “.csv”. Eu fiz o download do arquivo e posso carregar no Power BI, mas será que posso fazer essa carga direta?

[02:50] Vamos fazer o seguinte: vou manter essa página aberta e vou voltar para o Power BI que abri novo. Vou clicar em Página Inicial, Obter Dados e vou escolher essa opção: Web. Vou clicar em Web e ele vai me pedir um link. Vou voltar para a página e vou, novamente, clicar com o botão direito do mouse, “Copiar endereço do link”.

[03:53] Voltando no Power BI, eu colo o link em “URL”. Então eu tenho aqui a minha página WEB que vou acessar. Vou clicar em OK e ele vai me mostrar o link da página que vai acessar. Eventualmente pode ser um arquivo ou um acesso que exija, ou uma chave de criptografia, ou um login e senha. Nesse caso, é anônimo, ou seja, é aberto. Vou clicar em Conectar.

[04:39] O que ele está fazendo? Ele está indo no site, pegando aquele CSV e baixando direto para dentro do Power BI. Então não preciso baixar o arquivo para um diretório na minha máquina para, depois, fazer a leitura. Então esse link faz parte da conexão, ou seja, quando eu atualizar os dados, ele vai de novo na WEB buscar essa informação.

[05:16] Ele me trouxe uma série de informações associadas às carteiras de motoristas de pessoas ou de empresas na Irlanda do Norte. Eu posso clicar Carregar ou “Transformar Dados” onde posso mudar o tipo de cada dado. Se eu clicar em Carregar, novamente estou indo na WEB e estou trazendo o conteúdo daquele arquivo para meu relatório do Power BI. Vou fazer o seguinte: vou parar o vídeo e volto quando terminar de carregar.

[06:02] Pronto, fiz a carga e eu tenho, no meu canto direito, as informações referentes às colunas. Então eu posso usar essa informação no meu relatório. Essa é uma forma, mas existe outra forma mais esperta que o Power Bi consegue capturar informações de uma página WEB.

[06:26] Vou voltar para o browser, abrir uma outra aba e vou procurar: “LISTAS DE UNIDADES FEDERATIVAS DO BRASIL POR POPULAÇÃO”. Ele me trouxe algumas coisas e eu quero pegar esse link: “Lista de unidades federativas do Brasil por população”, que é um link que tenho na Wikipedia.org.

[07:03] Se eu abrir esse link, vou ver a página da Wikipedia e eu tenho algumas tabelas. Esse “Lista” é uma tabela. Possui posição, as unidades federativas do Brasil, a população, percentual da população e países em que essa população pode ser comparada. Mas não tenho só essa tabela, tenho essa outra que é a população concentrada por região. Enfim, é uma página WEB.

[07:45] Se formos olhar o código fonte dessa página WEB vamos ver que, no meio da HTML que forma essa página, eu tenho componentes que chamamos de Tabelas. Vou clicar com o botão direito do mouse sobre a página WEB e vou no link “Exibir código fonte da página”. Vou ter um montão de coisas que, para quem não entende HTML, é grego, mas se eu observar com calma, vou achar uma tag chamada “

□
”.

[08:33] Table, na HTML, significa a criação de uma tabela. A tabela tem um corpo e eu tenho a tag “

□

” para criar uma linha e a tag “

” para criar uma coluna. Então aqui (“

□

”) eu abro uma linha, nessa célula (“

”) escrevo Posição, Unidade federativa, população e depois, aqui “

”, eu fecho a linha. Depois abro outra linha (“

□

”) e vou ter o número 1 e assim por diante.

[09:17] Então eu tenho, internamente, o desenho da tabela. Esse desenho em HTML, quando expresso no browser, vai desenhar essa tabela. Vou pegar este link:

“[http://pt.wikipedia.org/wiki/Lista de unidades federativas do Brasil por população](http://pt.wikipedia.org/wiki/Lista_de_unidades_federativas_do_Brasil_por_populacao)”. Vou voltar para o Power BI e fazer uma nova conexão também do tipo WEB. Em URL vou colocar o endereço da página da Wikipedia. Note que esse endereço não tem um arquivo “.csv” no final. Estou dando o endereço da página WEB.

[10:04] Vou clicar em OK, ele vai me mostrar parâmetros parecidos: se eu, por acaso, tenho que passar credenciais para aquela página para buscar informação.

Vou clicar como modo Anônimo e Conectar. Ele agora não está trazendo o arquivo, mas está analisando o conteúdo da página web e vai me mostrar essas coisas que representam Tabelas que estão embutidas dentro do código fonte.

[10:47] Se eu for clicando nelas vou vendo o que significam. Essa, por exemplo, é uma tabela que não me diz respeito. É alguma coisa interna da página WEB. Essa “Lista” é legal, ela me dá justamente a tabela com as populações. Ou seja, aqui estou vendo, em forma de dados, o que estou vendo nessa tabela (WEB). Ele está indo capturar o conteúdo da tag HTML.

[11:25] Tenho algumas outras tabelas aqui que não me dizem respeito, mas as importantes são essas: “Lista” e “Por região”. Quando eu vejo a tabela, se eu clicar na aba “Exibição da Web”, eu consigo ver a página WEB e eu tenho, destacado, a área da página que representa a tabela. Eu não consegui fazer por um motivo de “Modo de Compatibilidade de Internet Explorer”. Não tem problema.

[12:59] Vou marcar a tabela “Lista” e a tabela “Por região”, coloco Carregar e trago, para meu relatório, mais duas novas tabelas. Tenho aqui a minha Lista, que é a lista por estado, e tenho minha lista por região. Então eu quis mostrar a vocês que o Power BI não necessariamente precisa somente carregar dados de tabelas ou de bancos de dados, mas também consigo carregar de páginas WEB.

Aula 03 – Funções de Transformações

[00:00] Então eu tenho no momento o meu arquivo Power BI com todas as cargas efetuadas, eu agora posso passar a construir os meus relatórios, infelizmente ainda não vamos poder construir os relatórios, nem sempre o dado que foi carregado está do jeito que eu preciso para o meu relatório, precisamos agora observar os dados carregados e ver se há necessidade de fazer transformações.

[00:35] Essas transformações ou limpezas de dados, vão ser o foco desse e dos próximos vídeos dessas aulas. Para podermos começar do mesmo ponto todo mundo igual, vamos fazer o seguinte, ao invés de usarmos esse arquivo aqui, que é o arquivo final da aula passada, eu vou pedir para vocês criarem junto comigo um novo arquivo Power BI onde vamos fazer uma carga todo mundo junto dos mesmos dados e fazer as transformações daqui para frente.

[01:08] Vamos supor que esses dados que nós vamos carregar foram frutos das diversas diferentes formas de carga que vimos na aula passada. Então vamos fazer o seguinte, vamos fechar aqui o Power BI

e associado aqui ao link desse vídeo façam o download dos arquivos que estão aqui, se eles vierem compactados descompacte-os e nós vamos ter esses dados aqui.

[01:40] Serão 5 arquivos textos, que nós vamos usar esses arquivos textos para a carga de um novo Power BI e a partir dele fazer algumas transformações, vamos copiar esses arquivos para um outro diretório na máquina de vocês, mas vamos criar esse diretório todo mundo junto com o mesmo nome, e por que isso? Porque depois se vocês quiserem recuperar parte do projeto que eu vou deixar publicado aqui no curso, vocês vão conseguir pegar esses arquivos e carregá-los.

[02:15] Porque a referência a esses arquivos textos serão iguais tanto no meu computador, quanto no computador de vocês, então façam o seguinte, eu vou aqui no driver C e eu vou criar aqui um diretório chamado “CursoPowerBI”, e eu vou copiar todos os txts que eu baixei para esse diretório, então vou voltar para o meu diretório origem, vou selecionar os arquivos, botão direito do mouse, copiar, vou para o meu diretório CursoPowerBI, onde ele está? Está aqui.

[03:10] E vou colar aqui, eu agora vou abrir o Power BI, e esse arquivo mesmo sem carga nenhuma, sem nada, eu vou botar arquivo, salvar como e eu vou salvá-lo nesse diretório que eu criei. Então vou chamar esse arquivo de “CursoPowerBi_001”, agora com esse arquivo salvo vamos carregar os dados. Página inicial, obter dados, arquivos textos ou csv, vamos selecionar arquivo por arquivo, peguei o primeiro arquivo.

[04:20] Sempre reparem se está usando o tf8 e eu vou dar carregar, vou repetir isso para o segundo, página inicial, obter fonte de dados, arquivo texto, vou agora carregar esses itens note, carregar, mais um, obter dados externos, arquivo texto, tabela notas, carregar, obter dados, arquivo texto, tabela de produtos e carregar.

[05:20] E finalmente, obter dados, arquivo texto, tabela de vendedores, carregar. Então vou ter aqui do lado direito as minhas 5 tabelas carregadas e eu vou simplesmente salvar esse arquivo e já tenho meu arquivo preparado para continuarmos aqui essa aula.

[00:00] Vamos então começar agora a analisar o conteúdo nas nossas tabelas que foram carregadas, para podermos ver se são necessárias

transformações a serem feitas. Nós temos aqui no canto as nossas tabelas, e vamos analisar algumas delas e verificar se realmente preciso que sejam feitas transformações, vamos começar pela tabela de clientes.

[00:35] Então vou vir aqui em página inicial, editar consultar e nós temos aqui as nossas consultas e a primeira tabela selecionada é a tabela de cliente. Notem o seguinte, vamos olhar aqui a coluna endereço, se batermos o olho já vemos que está uma zona, está uma bagunça o conteúdo dessa coluna.

[01:00] Temos várias coisas aqui a ser tratado, temos aqui esse espaço em branco que começa antes da palavra, a palavra rua deveria estar colada, não deveria ter nenhum espaço em branco antes do número do endereço, tem lugar que eu começo com rua, tem lugar que é “r.” tem lugar que é “R.”. Então deveria ter uma unificação, nós temos um espaço em branco no meio aqui e temos por exemplo, alguns nomes de coluna que hora começa com minúsculo, hora começa com maiúsculo.

[01:38] Não respeitando um certo padrão no nome do endereço, então vamos tentar dar uma arrumada nisso. Vamos começar indo aqui, eu vou clicar em transformar, e eu tenho aqui uma opção chamada formato. Formato já tem algumas funções para formatar dados de uma coluna, então vou escolher a minha coluna endereço, e se eu vier aqui em formato eu tenho algumas funções.

[02:20] Primeira função que eu vou usar é a cortar, a função cortar tira os espaços em branco, tanto no início quanto no final do endereço, então se eu vier aqui, formato, cortar, prestem atenção no conteúdo do endereço, vejam que aqueles espaços em branco que estavam no começo da palavra já foram retirados. Então pronto, já fiz uma transformação.

[02:57] Aqui do lado, se vocês repararem quando começamos tínhamos essas 3 aplicações sobre a tabela, a primeira era a leitura do dado da fonte, a segunda era colocar a primeira linha como cabeçalho e a terceira era mudar os tipos conforme o entendimento do conteúdo do arquivo. E ele colocou agora uma quarta transformação sobre o arquivo, que foi esse que ela chamou de texto aparado, que foi aquela que tirou os espaços em branco na frente do nome.

[03:40] Eu tenho uns espaços aqui no meio, esses espaços no meio podem ser espaços mesmo, mas eu estou suspeitando que sejam outra coisa, olha o que eu vou fazer aqui, eu vou vir aqui naquele diretório que criamos, o C, “CursoPowerBI” eu tenho aqui o meu arquivo tabela cliente, só que eu vou dar botão direito do mouse com ele e vou abrir ele com esse editor de texto aqui, notepad++.

[04:25] Esse notepad++ é um editor de texto, um notepad melhorado e inclusive pode ser baixado da internet, é gratuito e eu gosto muito de usá-lo, vocês não precisam ter o notepad++ baixado na máquina de vocês, mas só observem, eu vou abrir o arquivo tabela cliente com esse editor, eu tenho aqui o conteúdo do arquivo, eu tenho uma tecla no notepad++ que ele me mostra os caracteres que estão escondidos no texto.

[05:00] Quando você lê um texto no Word, no Excel ou no editor de texto, às vezes tem coisas ali escondidas que tem a ver com a formatação do texto, que você não vê, com esse botão aqui você consegue enxergar isso, se eu clicar nele, notem que não estou vendo nenhum daqueles símbolos amarelos, se eu clico nele de novo estou vendo.

[05:25] Ele está me dizendo o seguinte, que a seta representa a tecla TAB que dá um espaçamento, os 3 pontinhos são os espaços, justamente esses 3 pontinhos que é um caractere escondido, mas na verdade o espaço está representado por vários pontinhos, não só 3, falei 3 pontinhos mas na verdade são 6, são os espaços que já foram eliminados.

[05:57] Se notarmos aqui, o que eu tenho escrito entre a palavra rua e a palavra iriquitia não é um espaço em branco, porque se não teria pontos ali, é uma seta, significa que tem uma tabulação e a tecla TAB aqui do teclado que serve para dar espaçamento, aqui também tem um espaçamento, um TAB, significa então o seguinte, que voltando aqui para o Power BI, esses caracteres escondidos aqui, na verdade são TABs.

[06:43] São partes daqueles caracteres que não enxergamos, como eu os elimino? Aqui em formato tem uma função que faz isso, é a função limpar, a função limpar ele tira o que existe no meio da palavra, daqueles caracteres que servem de formatação, então se eu clicar aqui

sobre formato limpar, note que o nome agora já ficou legal, já não tem mais os espaços em branco.

[07:23] Próxima transformação é a seguinte, eu tenho palavras que começam com maiúscula e outras com letras minúsculas, vamos colocar um padrão nisso, eu tenho outra função aqui em formato, chamada colocar cada palavra em maiúscula, na verdade a tradução do Power BI para português não foi muito feliz, o que ele quer dizer com essa função é colocar cada letra inicial da palavra em maiúscula, colocar todas as palavras em maiúscula na verdade é essa função aqui de cima.

[08:00] Então vamos rodar, formato, colocar cada palavra em maiúscula, agora já temos um padrão onde tudo começa com letra maiúscula, mas ainda preciso padronizar a representação da rua, em alguns lugares rua está escrito com a palavra RUA, e em outros lugares R. e da mesma maneira eu tenho aqui Av. Av sem ponto, e eu quero que escreva a palavra Avenida completa, vamos fazer isso.

[08:40] Só que agora não é na função formato, eu vou vir aqui onde está o nome da coluna endereço e botão direito do mouse eu tenho aqui substituir valores, e eu vou substituir, por exemplo, R. eu substituo por Rua, pronto, agora está tudo começando com Rua, eu tenho Av. que eu quero substituir por avenida, então substituir valores, vamos botar Av. por Avenida.

[09:25] E está faltando o Av sem ponto, vamos transformar ele também para avenida, substituir valores, Av sem ponto por Avenida. Tivemos um problema, tinha um Av aqui, que ele acabou botando Avenidaenida, porque na verdade ele pegou o Av sem ponto e substituiu por avenida, não devíamos ter feito isso, para eu desfazer essa transformação eu sei que ela foi escrita aqui por último, então eu posso clicar aqui e excluir.

[10:25] Note que ele voltou ao status que estava antes de eu aplicar aquela nova substituição, eu vou fazer o seguinte, eu vou substituir, mas assim, Av e um espaço, por Avenida e um espaço, e agora eu só vou pegar o Av que está isolado, que é o da linha 13, agora ficou legal. Então eu tenho já o meu endereço formatado, e foi aplicado sobre a tabela essas seis transformações, onde eu fui fazendo a limpeza do dado até ter a coluna endereço do jeito que eu preciso ter.

[11:20] Depois disso feito, vamos vir aqui em arquivo, fechar e aplicar, e agora para guardamos esse status final desse nosso vídeo, vou vir em arquivo, salvar como e ir no diretório C, "CursoPowerBI" eu vou salvar o 002.

[00:00] Vamos então agora fazer a transformação na tabela de produtos, vamos olhar o conteúdo dela, eu vou clicar aqui em tabela de produtos, clico aqui em tabela, na verdade agora sim eu tenho que clicar em tabela de produtos, e eu tenho aqui o nome do produto, o que eu tenho que fazer é o seguinte.

[00:25] No nome do produto eu tenho as seguintes informações, eu tenho a família do produto, então no caso aqui a família é sabor da montanha, aqui eu tenho o tamanho do produto que no caso é 700ml e aqui no final eu tenho o sabor do produto, que no caso é um produto sabor uva, é interessante ter uma coluna só para família, só com o tamanho e só com o sabor.

[01:04] Então eu preciso fazer o que chamamos de uma divisão da coluna, vou pegar essa coluna e dividir em 3, só que esse nome aqui tem que ser preservado, o nome do produto tem que continuar na tabela, eu simplesmente tenho que dividir. Então vamos trabalhar nessa transformação.

[01:20] Eu vou ver aqui em página inicial, editar, consultas, eu vou escolher a minha tabela produtos, antes de fazer isso, note uma coisa, eu tenho aqui o nome da origem que era aquele nome do arquivo original, vocês se recordam? Lemos um produto de um diretório onde eu tinha vários arquivos e o Power BI foi juntando-os, e ele cria uma coluna chamada nome de origem que é o nome original do arquivo de onde veio aquela linha.

[02:10] Esse nome origem não me interessa, na hora de eu construir o meu relatório, então eu vou fazer o seguinte, eu vou clicar sobre a coluna, botão direito do mouse e eu vou escolher a opção remover, assim eu vou remover a coluna. Essa coluna não existe mais, vamos então voltar ao nosso problema de eu dividir a coluna, eu tenho aqui em transformar, uma funcionalidade que se chama dividir coluna, só que essa funcionalidade ela tem uma característica, ela destrói a coluna original.

[02:48] Quando eu aplico ela sobre uma coluna, a coluna original é perdida, mas eu quero preservar a coluna original, porque essa coluna original é o nome do produto, então o que eu vou fazer é o seguinte, eu vou clicar aqui sobre nome do produto, botão direito do mouse, duplicar coluna, o duplicar coluna vai criar uma coluna ao lado, com o mesmo conteúdo da coluna original. Então é o que eu vou fazer agora, duplicar coluna.

[03:25] Notem que agora eu tenho uma coluna igual a coluna original, agora sobre essa nova coluna vamos dividir as informações em família, tamanho e sabor, eu vou então escolher o nome produto “-copiar” que é o nome da coluna e eu vou clicar aqui em dividir coluna, ao abrir eu tenho vários tipos, por delimitador, por número de caracteres ou por posições, por maiúsculo, minúscula, dígitos ou não dígitos, significa o seguinte, se é por número de caracteres por exemplo eu vou dividir as colunas.

[04:09] Pegar os 3 primeiro caracteres é uma coluna, do quarto ao sétimo é outra, do oitavo ao decimo é outra, isso é dividir por caracteres, por posição é muito parecido, só que por caracteres eu vou pegar 3 de cada vez e enquanto que em posição eu vou pegar a posição relativa dentro do texto, maiúscula ou minúscula ele vai identificar, achou a primeira maiúscula divide, e assim por diante.

[04:38] O que nós vamos usar é por delimitador, porque temos um delimitador aqui, é um traço, ele identifica a separação das minhas colunas. Então vamos lá, eu vou clicar aqui em dividir coluna por delimitador, e ele vai me dizer qual é o delimitador que eu vou usar, eu posso usar um delimitador padrão, que normalmente encontramos nos delimitadores de texto, ou um personalizado e ele mesmo já me sugere o traço como sendo personalizado.

[05:17] E eu posso ou pegar a coluna mais à esquerda, ou mais à direita, ou cada ocorrência, no caso eu vou pegar cada ocorrência, porque eu quero ter 3 colunas, uma para cada divisão que eu tenha. Então eu vou escolher essa aqui, cada ocorrência de delimitador e vou clicar em ok, olha o que ele fez, ele dividiu as colunas em 3, na primeira coluna eu tenho a família, na segunda eu tenho tamanho, na terceira coluna eu tenho o sabor.

[06:00] Uma coisa interessante é você ter nome de colunas coerentes, então eu vou poder vir aqui, se eu der um duplo clique sobre o nome da coluna eu posso mudar o nome para família, a segunda coluna eu vou chamar de tamanho, e a terceira coluna eu vou chamar de sabor. Então pronto, eu já tenho as minhas 3 colunas, família, tamanho, sabor, baseadas no delimitador desse campo original aqui.

[06:45] Então eu consegui fazer a transformação que eu queria, e óbvio todas as transformações que eu fiz estão registradas aqui, foram os passos que foram efetuados, clico em arquivo, fechar e aplicar, ele vai fazer a leitura, aplica, agora eu tenho as minhas colunas separadas, agora eu vou vir em arquivo, salvar como, eu vou salvar aqui agora como o arquivo 003. Pronto, está salvo.

[00:00] Vamos fazer mais uma transformação, de novo na tabela de clientes, vou abrir aqui página inicial, consultar e editar consultas e vamos olhar a tabela de clientes. Temos aqui 3 colunas, um é o título, outra é o nome e a outra é o sobrenome. Vamos juntar essas 3 colunas em uma só, mas antes de fazer isso vamos padronizar o título, você vê que eu tenho título começando com minúsculo título começando com maiúsculo, enfim, vamos manter um padrão único.

[00:40] Primeiro para vermos todas as combinações basta eu clicar aqui e eu vou ver que eu tenho essas combinações aqui. Eu quero passar tudo para ou Sr ou Sra mas com S maiúsculo, isso vai ser o padrão para Senhor ou Senhora que é o título da pessoa que faz parte do cadastro de cliente do dado que eu estou manipulando.

[01:05] Então vamos começar a fazer aquelas substituições, substituir valores, vamos pegar “sra” minúsculo, com “Sra” com a primeira letra maiúscula. Agora eu vou substituir “Sra.” que tem ponto no final por “Sra” normal, padronizei o Senhora.

[01:47] Agora o Senhor, vamos ver que combinação que eu tenho, eu só vou tirar o ponto desse aqui para ficar igual a esse aqui de cima. Então eu vou substituir valores, “Sr.” vai virar Sr normal, agora sim, já tenho a minha padronização. Vamos agora criar uma coluna nova que vai ser o nome completo, que vai ser o título, o nome e sobrenome. Eu tenho algumas maneiras diferentes de fazer isso, mas eu vou clicar aqui em adicionar coluna e eu vou usar uma funcionalidade que eu acho muito interessante.

[02:37] Que o Power BI aprende com o exemplo que você dá, eu vou escolher essa opção aqui, coluna de exemplos, e olha o que eu vou fazer, primeiro vou com o “Ctrl” apertado, eu vou selecionar a coluna título, nome e sobrenome, e eu vou vir aqui e selecionar coluna de exemplos, seleção, ele me mostra as 3 colunas aqui e eu vou meio que ensinar a ele o que eu quero fazer, eu vou escrever isso aqui, “Sra Érica Carvalho”, ou seja, eu escrevi aqui do lado o que eu tenho aqui, o que eu tenho no segundo e no terceiro.

[03:40] Ao fazer isso eu vou ensinar ao Power BI que eu quero que repita dessa linha para baixo, a mesma lei de formação que eu fiz na primeira linha, como se ele aprendesse comigo pelo o que eu fiz na primeira linha. Então se eu digitei os 3 olha o que ele faz aqui, ele começa a me sugerir juntar também, “Sr César Teixeira”, você vê que a união que está aqui é aplicada aqui.

[04:25] Então se eu estou de acordo com essa mudança eu cliço em “ok”, e ele me cria uma outra coluna que ele me chamou de mesclado, mas claro eu já sei fazer isso, dou um duplo clique aqui e vou chamar de nome completo. Então eu já tenho aqui o nome completo que é uma coluna que está juntando o título com o nome e o sobrenome.

[05:03] Então certo, vamos no início, fizemos essa modificação, vamos fechar e aplicar, e vamos dar arquivo, salvar como, “CursoPowerBI004”. Pronto, terminamos esse vídeo.

[00:00] Vamos agora manipular com datas, e eu vou mostrar o problema que eu tenho que resolver, vou vir aqui em página inicial, editar consultar e eu vou escolher a tabela de clientes, note que eu tenho aqui um campo chamado idade, esse campo idade ele me diz a idade do cliente, só que o valor desse campo eu não vou confiar nele, porque esse campo está vindo do arquivo texto, do arquivo original e aquele arquivo original ele está estático.

[00:40] A medida que o tempo vai passando, as pessoas vão ficando mais velhas, mas o valor que está lá no arquivo texto não muda, então eu queria construir um campo idade que fosse dinâmico, baseado na data do cliente, ou seja, quando eu fizesse uma carga a idade fosse recalculada, para eu fazer isso primeiro eu preciso converter essa coluna data em tipo data.

[01:12] Porque se nós olharmos, ela está com aquela letrinha ABC ali na frente significa que aquele campo ele é do tipo texto, e eu preciso converter para data. Então a primeira coisa que eu vou fazer é clicar sobre o texto e eu vou escolher a opção data, para converter ele para data. Ele até muda um pouco a forma de escrita, ele vai para a direita e eu consigo ver que ele agora é um campo data.

[01:50] Essa coluna idade já não vai servir mais para mim, então eu vou clicar sobre ela, botão direito do mouse, remover, agora eu vou calcular a idade, eu vou clicar sobre esse campo, data de nascimento e eu vou vir aqui, adicionar coluna, eu tenho a opção aqui data, e eu tenho uma opção idade, ou seja, ele já tem uma função que calcula para mim a idade baseada na data.

[02:28] Então eu vou clicar aqui em idade e eu já tenho a idade calculada, só que o que eu tenho aqui é no formato dias, horas, minutos e segundos. Então esse número que eu estou vendo aqui é a idade em dias, eu preciso converter essa idade para anos. Eu tenho aqui, transformar e note eu tenho aqui uma função que faz transformação sobre o campo do tipo duração, eu sei que é esse campo porque esse símbolo do relógio eu estou vendo-o, então há uma correspondência.

[03:10] Eu clicando sobre a coluna idade, clicando em duração eu quero ver isso em total anos. Pronto, tenho a idade. Só que ninguém tem 22,0198 anos, só tem ou 21 ou 22. Então eu vou clicar aqui sobre o tipo, que ele colocou automaticamente o tipo decimal, eu vou clicar aqui e vou converter ele para número inteiro, agora sim eu tenho a idade.

[03:53] Eu tenho a idade em anos dinamicamente calculada, baseada no conteúdo da idade que eu tenho aqui, que está vindo do arquivo e baseado na data do computador. Arquivo, fechar e aplicar, arquivo, salvar como, esse vai ser o “CursoPowerBI005”, pronto.

[00:00] Vamos aplicar o conhecimento adquirido até agora para resolver um outro problema de formatação no nosso modelo. Deixa-me mostrar o problema para vocês, eu vou vir aqui então, página inicial, editar consultas e eu vou selecionar a minha tabela notas. Na tabela notas eu tenho um campo que representa a data, só que ele está

expresso em ano/mês/dia e 00000 que é a hora, minuto e segundo que no caso não está sendo levado em conta.

[00:50] Veio assim do arquivo texto, tem um texto desse jeito, vamos aplicar aquilo que já sabemos, quero converter isso para data, clico aqui no ABC e vou escolher data. Vamos substituir a atual e note que ele ficou com essa mensagem “Error”, o que significa isso? Ele não conseguiu entender aquilo como data, talvez o tipo esteja errado, seria data hora? Não, continua com Error.

[01:40] Então vou voltar de novo para texto. Então não conseguimos converter isso em uma data, como eu faço para converter isso em uma data? Vamos aplicar o que já conhecemos. Eu sei que se eu tiver a data como dia/mês/ano, se eu tiver com a data nesse formato, eu consigo converter ele para data, foi isso que fizemos no vídeo anterior, o primeiro passo.

[02:30] Então voltando aqui para tabela de notas, como que eu consigo pegar o dia/mês/ano? Pelo Split de colunas, usando o delimitador, traço como delimitador. Eu vou selecionar a coluna data, vou vir em transformar, dividir coluna, por delimitador, o delimitador vai ser o traço e cada ocorrência do delimitador vai virar uma coluna.

[03:10] Feito. Eu tenho aqui a coluna do ano, do mês, só que a coluna do dia veio com esse lixo aqui, eu vou até aqui primeiro, vamos botar aqui, ano, mês, só que esse ainda não é o dia, mas eu também posso dar um outro Split, usando como delimitador o espaço.

[04:05] Então eu vou vir aqui, de novo dividir coluna, por delimitador, não é dois pontos, porque ele viu que tem muito dois pontos ali, então ele sugeriu os dois pontos, não, é personalizado, então eu vou colocar aqui um espaço, cliquei aqui na tecla espaço, só que eu quero só fazer um Split da extremidade da esquerda, porque eu só quero pegar os primeiros números que estão a esquerda. Eu clico aqui, delimitador da extremidade esquerda, clico em ok.

[04:45] Então está aqui, ficou aqui a hora, e aqui ficou a data. Só que esse cara eu vou chamar de dia. Essa coluna eu vou apagar, não preciso dela, então agora eu tenho a coluna de ano, de mês e de dia, vamos juntar para dia/mês/ano? Já fizemos uma coisa parecida em

vídeos passados quando juntamos o título do nome e do sobrenome do cliente.

[05:20] Usamos a coluna exemplo, então vamos pegar o ano, com o “Ctrl” apertado o mês, depois o dia e eu vou vir aqui em coluna de exemplos, da seleção, tenho os 3 campos marcados, então eu vou dizer para ele: você vai fazer “03/01/2015”, e ele vai aprender e vai me dar aqui a solução, eu tenho aqui a coluna mesclado, que eu vou chamar de data da venda.

[06:20] Só que esse campo ainda está como um string, eu clico nele e digo que ele é um campo data. Pronto, depois de vários passos, olha só quantos passos fizemos, depois de vários passos conseguimos transformar aquele campo original em um campo de data, e aplicamos aqui vários conhecimentos que tivemos durante algumas aulas desse curso.

[07:00] Então vamos agora vir em arquivo, fechar e aplicar, note que agora ele está demorando para ler o tabela notas que é uma tabela grande, porque agora que ele está aplicando a transformação sobre cada linha daquela tabela que está no txt, deve ter 80 a 90 mil linhas, vamos agora em arquivo, salvar como, vou salvar aqui como sendo o 006, pronto.

Aula 04 – Modelagem e primeiro contato com DAX

[00:00] Então o que eu tenho aqui é o meu projeto de Power BI com as transformações referentes ao último vídeo da aula passada, eu tenho carregada 5 tabelas, a tabela de clientes, a tabela de itens notas, a tabela de notas, a tabela de produtos e a tabela de vendedores. Vamos mexer um pouco no nosso modelo de dados, primeira coisa, eu vou mudar o nome dessas referências, não usar a tabela de clientes, mas só o nome clientes, itens, notas.

[00:45] Porque depois nós vamos ver que esses nomes são importantes quando construirmos os nosso relatórios e quanto mais fácil tiver as informações que estão carregadas no meu arquivo do Power BI para o usuário que vai construir o relatório melhor. Essas informações nós

chamamos de metadados do relatório, o metadados é a informação sobre o dado.

[01:17] Vamos começar clicando em página inicial, editar consultas e vamos começar pela tabela cliente, eu vou chamar de clientes, vai ser esse o nome dessa tabela. Outra coisa que eu vou modificar, é fazer uma revisão nos tipos de cada coluna, então eu vou começar por exemplo por CPF. CPF foi carregado automaticamente pelo Power BI como um campo numérico, mas na verdade eu quero que esse campo seja um campo texto.

[01:54] Porque CPF ele pode ser texto, porque às vezes eu posso ter um CPF que comece com um 0 na frente, isso foi algo que eu até já discuti em vídeos anteriores, para mudar isso já sabemos, eu clico sobre o símbolo 123 que representa número, e eu vou selecionar ele como texto. Vamos fazer outra revisão, CEP é a mesma coisa, CEP foi carregado na tabela e o Power BI o colocou como número, vamos clicar aqui sobre o símbolo 123 e eu vou colocar ele como texto.

[02:40] Uma outra coisa é o nome dos campos, vamos melhorar um pouco esse nome para que fique nomes mais fáceis de trabalhar. Então por exemplo, esse endereço 1 eu vou colocar endereço, sobrenome, vamos colocar aqui “Sobrenome”, vou tirar esses maiúsculos para ficar algo mais padronizado, “Nome”, “Título”, CPF eu vou manter, endereço 2 é um campo que está vazio, então vou até apagar ele, “Bairro”, “Cidade”, “Estado”, CEP eu vou manter maiúsculo.

[03:48] Data de nascimento eu vou escrever assim “Data de Nascimento”, “Limite de Crédito em RS”, dando um nome mais amigável para os campos, “Volume de Compra em Litros”, esse campo aqui 01 significa se o cliente já fez ou não a primeira compra, então vou colocar aqui, “Fez a Primeira Compra?”, “Nome Completo do Cliente”, idade está legal, vou manter idade.

[05:10] Vamos pra a tabela de itens de notas, eu vou chamar essa tabela de itens de notas fiscais, número da nota é número mesmo, é um número sequencial não vou mudar, código do produto apesar de ele estar representado como um número, mas ele na verdade também é um texto, então vou mudar ele para texto. Quantidade está legal, preço está legal, então vou colocar aqui assim, “Número da Nota

Fiscal”, “Código do Produto”, a quantidade é “A Quantidade em Litros”.

[06:20] Isso tudo claro, se eu estou fazendo isso para uma empresa eu já fiz entrevista com os usuários e eu sei como vai ser o nome dessas colunas, aqui vai ser por exemplo, “Preço Unitário de Venda”, vamos para a tabela de notas? Eu vou chamar a tabela de notas fiscais. O CPF ele também está como um número, eu vou mudar para texto. Matrícula do vendedor é o número interno do vendedor dentro da empresa, é o vendedor que fez a venda, ele também é um texto.

[07:05] Então vou mudar aqui de número para texto, ano vou manter, mês, dia, está legal, vamos para os textos, aqui eu vou chamar de “Matricula do Vendedor”, “Ano”, “Mês”, “Dia”, aqui é o “Número da Nota Fiscal”, “Imposto em %”, então podemos até dizer se é imposto em percentual, eu tenho essa formatação aqui, percentual, então isso vai ser percentual.

[08:20] Data da venda, vamos para a tabela de vendedores? Eu vou chamar de “Vendedores”, “Matrícula do Vendedor”, “Nome do Vendedor”, “Percentual de Comissão”, o próprio nome diz que é percentual, vamos colocar %, “Data da Admissão” e aqui esse campo é um campo para dizer se o vendedor está de férias ou não naquele momento, então eu vou botar aqui “De Férias?”, aqui é o “Bairro do Escritório do Vendedor”.

[09:37] Tabela produtos, vamos chamar a tabela de “Produtos”, aqui vamos chamar de “Código do Produto”, aqui “Nome do Produto”, código do produto vamos converter ele para texto, “Embalagem”, “Preço Unitário de Lista”, é o preço padrão do produto, “Família”, “Tamanho” e “Sabor”, uma outra coisa que é interessante no metadados, não vamos fazer isso porque se não vai dar muito trabalho, eu vou ficar muito tempo fazendo.

[10:45] Vamos fazer o seguinte, aqui foram feitas várias transformações, podemos mudar o nome dessa transformação para dizer o que ela está fazendo, porque depois na manutenção do metadados fica mais simples, então por exemplo, dividir coluna por delimitador, isso foi que pegamos a coluna nome do produto e dividimos.

[11:18] Então eu poderia vir aqui, botão direito do mouse, renomear e escrever “Dividir coluna do nome por delimitador”. Então eu tenho aqui no meu metadados, eu sei que o eu fiz aqui, coluna duplicada por exemplo, quem é que eu dupliquei, esse aqui por exemplo foi quando eu dupliquei a coluna nome do produto, então eu posso dizer aqui duplicar coluna nome do produto e ter aqui no meu metadados uma explicação mais coerente.

[12:15] Fizemos várias transformações e eu não vou estar aqui mudando isso nesse momento, eu vou manter assim só porque o mais importante realmente são o nome das tabelas, os nomes das colunas e os tipos das colunas que são mais importantes.

[12:33] Vamos fechar e aplicar para podermos estar implementando essa alteração, ele vai aplicar isso de novo, vai ler da fonte e vai estar aplicando essas mudanças, que demora mais sair na tabela de notas fiscais, você vê que itens ele já foi, clientes também, vendedores também e vamos agora salvar lá naquele nosso diretório de trabalho eu vou salvar como o arquivo 007.

[00:00] Uma outra coisa importante que nós temos que observar são os relacionamentos entre as tabelas, nós temos esses ícones aqui que pouco entramos, que mostra o relacionamento entre as tabelas que foram carregadas, quando carregamos as tabelas elas por teoria são isoladas, só que as informações às vezes, que nós vamos usar para montar o nosso relatório, estão em tabelas diferentes e às vezes eu tenho que juntar essas tabelas.

[00:30] Então é por um relacionamento entre elas que fazemos essa união, conseguimos buscar os dados de tabelas diferentes no mesmo gráfico, na mesma tabela, no mesmo relatório. Eu vou clicar aqui sobre esse ícone e eu vejo aqui o relacionamento entre as tabelas, como ele construiu esse relacionamento? Ele fez esse relacionamento no momento em que eu fiz a carga.

[01:05] Ele observou campos com nomes iguais e ele acabou efetuando esses relacionamentos porque ele observou que os campos tinham o mesmo nome. Por exemplo, se eu colocar o mouse sobre a ligação, note que ele está dizendo que tem uma ligação entre CPF de clientes com CPF em notas fiscais, está aqui a ligação, logo quando eu quiser montar um relatório que envolva a idade do cliente e o dia da compra,

eu vou conseguir, porque vai ser por essa ligação que eu vou conseguir juntas essas duas informações em um único relatório.

[01:53] A mesma coisa aqui, se eu colocar o mouse aqui, vamos descer um pouco aqui para baixo, ele faz a ligação entre a nota fiscal e o item da nota, através do número da nota fiscal, então ele observou que esses campos tinham os nomes iguais e colocou essa ligação. Agora temos aqui no canto uma tabela isolada, por que essa tabela está isolada?

[02:25] Na verdade, a ligação que ela deveria fazer é através de código do produto, tem código do produto aqui na tabela de itens notas fiscais, e código do produto aqui na tabela de produtos, agora porque nesse caso ele não fez a ligação? Os nomes são iguais, atenção, os nomes são iguais porque eu no vídeo passado eu mudei esses nomes, agora os nomes originalmente não eram iguais, quer ver?

[03:05] Olha só, se eu clicar aqui em tabela e for em página inicial, editar consultas, se eu pegar a tabela de produtos, o nome aqui está, Código do Produto, mas se eu clicar sobre essa transformação, ou seja, uma antes de eu ter alterado o nome das colunas, nessa aqui que eu tenho que ligar, ao clicar nessa aqui, note que o nome da coluna original era código do produto com um underscore, entre código e do, e outro underscore entre do e produto.

[04:00] Esse era o original que veio da tabela, agora se eu for na tabela de itens de notas fiscais e clicar aqui em cima, note que na tabela de itens o nome do campo é código do produto com um espaço entre as palavras, então esses nomes originalmente foram diferentes, o Power BI não fez essa conexão e ao renomear esses nomes ele não faz a conexão, porque a conexão ele faz no momento da carga, quando ele faz a primeira carga dos dados para dentro do projeto Power BI.

[04:45] Então já que entendemos agora porque a ligação não foi feita de forma automática, podemos fazer essa ligação de maneira manual. Cliquei em fechar e aplicar, vou clicar de novo aqui nas relações, e eu vou pegar aqui e vou clicar em código do produto, da tabela de produtos e mantendo o botão da esquerda do mouse apertado eu vou arrastar para “código do produto” que está na tabela itens de notas fiscais.

[05:23] E ele me mostra essa caixa de diálogos, que é o tipo de relação que eu vou fazer, vamos dar um “ok” e ele cria a ligação entre essas duas tabelas. Então agora o meu modelo está todo ligado, qualquer informação que eu possa ter, por exemplo se eu pegar uma informação aqui de embalagem e quiser juntar ela com nome completo do cliente, eu vou ter essa ligação, porque essa ligação vai vir por aqui.

[06:05] Vai vir daqui para cá e daqui para cá, e eu consigo chegar qualquer campo que eu escolha eu consigo chegar no outro, para montar o relatório quando eu quiser juntar dois ou mais campos que estão em tabelas diferentes, porque agora o modelo está todo conectado através do relacionamento entre as tabelas.

[00:00] Antes de continuar esse vídeo, vamos fazer uma coisa que eu esqueci de orientar vocês no vídeo anterior, vamos pegar esse projeto, arquivo, salvar como, e vamos salvar ele como arquivo 008.

[00:20] Vamos falar agora sobre inteligência do tempo, ou time intelligence, quando vamos fazer um relatório que envolva informações ao longo do tempo, é muito legal já termos tabelas que mostrem baseado em uma data, qual é o mês, o dia da semana, o trimestre, o ano e assim por diante. Essas informações baseadas na data original, nós chamamos de time intelligence ou inteligência do tempo.

[00:55] Isso porque quando vamos construir os relatórios no Power BI, existem algumas funcionalidades que usam essas informações de mês, trimestre, dia da semana, para nós dar alguns resultados, então vamos construir agora a nossa tabela de inteligência de tempo, que vai ser uma tabela que vai ter como colunas, todas essas informações que nós temos, baseado em uma data real.

[01:30] Não vamos carregar essa tabela direto da nossa fonte, vamos aproveitar a tabela que já carregamos para construir a inteligência do tempo, vamos vir então aqui e clicar aqui em dados, e depois eu vou vir em página inicial, editar consultas, nós temos a nossa tabela notas fiscais que tem essa coluna que é a data da venda e ela inclusive já tem alguma coisinha de inteligência de tempo aqui.

[02:10] Que é o ano, o mês e o dia, mas vamos aproveitar essa própria tabela para construir o nosso time intelligence. Primeiro eu vou duplicar a tabela, porque a tabela notas fiscais da maneira que ela está, ela tem que continuar existindo, então eu vou dar aqui botão direito do mouse sobre notas fiscais e eu vou escolher a opção duplicar, ao fazer isso eu vou duplicar a minha tabela.

[02:48] Eu tenho aqui uma outra tabela chamada “Notas Fiscais 2”, mas é claro que eu vou chamar ela de “Inteligência do Tempo”. Então esse vai ser o nome dela, vamos começar agora a mexer nisso, ela não tem CPF, não preciso de CPF nela, ela não precisa de matrícula do vendedor, vou manter aqui o ano, vou manter aqui o mês, vou manter o dia.

[03:25] Número da nota fiscal some, estou clicando na coluna apertando a tecla delete, imposto some e ficou só essas 4 colunas, ano, mês, dia e data da venda. Notem que as datas se repetem aqui e eu não preciso dessa repetição, se repetem porque no dia eu posso ter mais de uma nota fiscal, posso ter mais de uma venda, e eu usei a tabela de notas fiscais para construir a nossa tabela de inteligência do tempo.

[04:10] Só precisamos ter as linhas que não se repetem, como fazemos isso? Eu posso vir aqui em página inicial e eu escolho a opção remover linhas e ao selecionar aqui, remover linhas duplicadas, ao fazer isso a minha tabela ficou agora mais enxuta, e só mostra dias distintos, nenhum dia mais se repete, porque toda a informação duplicada foi retirada da base quando eu escolhi a opção remover duplicatas.

[05:02] Só que agora nós vamos criar uma série de indicadores nessa própria tabela baseado na data da venda e temos aqui em transformar com a data da venda selecionada, nós temos a opção data onde podemos, por exemplo, vamos começar no dia, eu tenho dia da semana, dia do ano, início do dia, eu vou escolher nome do dia. Então ele já escreveu aqui sábado, terça, quinta, só que eu acabei criando sobre a própria coluna.

[05:53] Eu fiz uma transformação errada, eu substitui a coluna que eu tinha e foi essa transformação aqui, para eu tirar ela já aprendemos a fazer isso, eu venho aqui em botão direito do mouse e excluir, e eu volto aqui a data da venda, eu vou clicar nela de novo, só que em vez de transformar eu tenho que escolher o adicionar coluna, porque o

transformar ele vai aplicar a função sobre a coluna que está selecionada.

[06:35] Enquanto que o adicionar coluna, vai aplicar transformação sobre uma nova coluna, então foi até bom eu ter cometido esse erro para falar para vocês essa diferença, entre o transformar e o adicionar coluna, transformar eu aplico a função de transformação sobre a própria coluna, enquanto que o adicionar colunas eu crio uma coluna nova eu quero criar colunas novas, então cliquei sobre adicionar coluna, tenho o mesmo ícone data, eu vou clicar nele, dia, dia da semana.

[07:10] Agora sim, só que eu escolhi a função dia da semana. Dia da semana ele me dá um valor numérico que começa com o 0 se eu não me engano é o domingo e vai do número de 1 a 6, eu não tenho muita certeza disso, mas vou fazer o seguinte para descobrirmos se o 0 é o domingo ou a segunda. Eu vou clicar de novo em data da venda, só que agora eu vou selecionar o nome do dia, tenho outra coluna aqui.

[07:57] Agora podemos descobrir, é isso mesmo, domingo é zero, segunda feira é 1, terça feira 2, então é isso mesmo o 0 é o domingo. Eu quero mais, data da venda, semana, semana do ano, às vezes é legal vermos o dado por semana dentro do ano, então vai ser um número de 1 a 44 a 45, dependendo do ano eu posso ter mais ou menos semanas, aqui vai me dizer o número da semana no ano.

[08:38] Então, eu tenho semana 7, semana 10, aqui é a semana 20, por quê? Porque o dia 10 de maio é a vigésima semana do ano 2015. Semana no mês, e vai de 1 a 4 ou de 1 a 5 dependendo do mês, então eu posso dar esses números, mas vamos passar para trimestre, trimestre é legal, trimestre do ano, ele mostra aqui, primeiro trimestre, segundo trimestre, mas pode ser que essa forma de representar com 1, 2, 3 e 4 não seja legal.

[09:35] Eu vou substituir o número por um nome, botão direita do mouse sobre trimestre, substituir valores. Então o 1 eu vou chamar de primeiro trimestre, note que ele não habilitou pra mim o botão de “ok” e me colocou esse símbolo aqui de atenção, por quê? Se eu passar o mouse em cima ele vai me dizer “insira um valor number”, por quê? A coluna ela é um número, então eu não posso inserir um texto.

[10:20] Como eu resolvo isso? Eu vou clicar em cancelar, clico sobre o número 123 e eu vou passar agora a coluna para texto, agora como ela é texto, eu posso vir aqui, botão direito do mouse, substituir valores e eu vou pegar o valor 1 e chamar de “Primeiro Trimestre”, de novo, o valor 2 “Segundo Trimestre”, mais uma vez, substituir valores, o valor 3 é o “Terceiro Trimestre” e novamente, substituir valores, o valor 4 é o “Quatro Trimestre”.

[11:25] De novo sobre data da venda eu vou agora pegar o mês, e eu vou pegar por exemplo o nome do mês, janeiro, fevereiro, março, então eu tenho aqui o nome do mês. Enfim, eu poderia estar acrescentando todas as opções existentes aqui, nessa tabela, vamos parar por aqui que já temos a nossa tabela e vamos fechar e aplicar essas modificações.

[12:03] Ele está agora aplicando a inteligência do tempo sobre essas novas colunas, vamos esperar terminar porque ainda falta uma coisinha, vamos aguardar a atualização do Power BI com essas novas transformações, que foram a criação dessa nova tabela e as inserções dessas novas colunas. Vamos voltar para modelagem de dados, então se eu clicar aqui em modelagem de dados, eu tenho a minha tabela inteligência de tempo perdida aqui no mundo.

[12:43] Eu preciso ligá-la com alguém, e ela vai estar ligada com quem? Com a tabela que deu origem a ela, que foi a tabela de notas fiscais. Então vou arrumar um pouco aqui o esquema das tabelas, vou colocar aqui e eu vou pegar a data da venda da tabela inteligência de tempo e vou arrastar sobre a data da venda em notas fiscais.

[13:18] Vamos olhar aqui? Eu passando o mouse em cima, ele mostra aqui a relação que foi construído, então a tabela inteligência de tempo já está ligada também ao nosso modelo, então eu consigo ver qualquer informação daqui, por qualquer coluna que seja uma propriedade do dia, que está contida na nossa tabela de inteligência de tempo. Vamos em arquivo, vou salvar, vai ser o nosso projeto 009.

[00:00] Até agora quando criamos novas colunas, fazer cálculos entre colunas, sempre utilizamos indo no nosso editar colunas, usamos uma série de botões com funções já embutidas em cima que faziam uma série de trabalhos para nós, seja na parte de transformar, seja na parte

de adicionar colunas. Não vimos todas as opções porque tem muitas opções para serem vistas.

[00:40] Mas também podemos estar criando coluna calculada através de linha de comando e essas linhas de comando nós utilizamos uma linguagem chamada DAX que vem de Data Analise Expression, o DAX seria uma coisa muito parecida com aquelas funções que você trabalha no Excel, só que elas não trabalham sobre um conjunto de células como por exemplo são as funções de Excel da planilha, eles agem sempre sobre uma coluna ou uma tabela.

[01:20] Vamos ver agora alguma coisa bem simples de DAX, e vamos começar criando algumas colunas calculadas. Eu vou fechar aqui, vou voltar para a nossa página principal e se eu clicar aqui no botão de dados, eu vejo os dados das minhas tabelas e eu vou escolher a tabela itens de notas fiscais. Eu tenho duas colunas é a coluna quantidade em litros e preço unitário de venda, o faturamento, ou seja, o valor em dinheiro que a empresa vai receber vai ser a multiplicação entre a quantidade de litros e preço unitário de venda.

[02:10] Vamos criar uma nova coluna calculada aplicando essa fórmula. Eu tenho aqui no menu modelagem a opção nova coluna, eu vou clicar aqui em modelagem, nova coluna, ele me monta a nova coluna e aqui em cima, como se fosse um Excel ele me abriu uma área para eu colocar a fórmula, aqui no começo onde eu tenho coluna eu vou escrever o nome dessa coluna, que é o nome do resultado final.

[02:50] E depois aqui eu vou colocar a fórmula usando a linguagem DAX. A minha coluna eu vou chamar ela de faturamento em reais, e o que vai ser? Vai ser a coluna quantidade em litros x a coluna preço unitário de venda, se eu começo a escrever quantidade em litros ele já me apresenta uma opção que é a forma como representamos no DAX uma coluna, representamos no DAX o nome da tabela e entre colchetes o nome da coluna.

[03:44] Se eu começar a escrever o nome da tabela, itens, note que ele já começa a me mostrar aqui não somente a coluna como algumas funções em DAX que tem as letras “ite” contidas, então ele vai me dando sugestões na medida em que eu vou escrevendo, no caso eu não vou usar função nenhuma, vou pegar uma coluna e multiplicar pela

outra, então itens, eu vou pegar quantidade em litros e no DAX a multiplicação é o “*” vezes o preço unitário de venda.

[04:38] Então, a fórmula ficou assim, faturamento em reais vai ser o item nota fiscal, quantidade em litros, vezes itens notas fiscais, preço unitário de venda, se eu der o “Enter” ele já me escreve a nova coluna com a fórmula aplicada. Vamos fazer uma outra coluna calculada agora com a inteligência de tempo, no caso é o seguinte, eu tenho a coluna trimestre e a coluna no do mês, eu vejo junho, julho, terceiro trimestre, mas seria interessante eu ter uma coluna que mostrasse o nome do trimestre e o ano e o nome do mês e o ano também.

[05:37] Vamos fazer isso, vou vir aqui em modelagem, nova coluna, vou começar pelo trimestre, vou chamar essa nova coluna de trimestre ano. Vai ser inteligência do tempo trimestre, e eu vou concatenar, ou seja, eu vou juntar dois textos e o símbolo para concatenar é o & comercial, então eu vou concatenar inteligência do tempo trimestre com um espaço em branco e concatenar com inteligência do tempo ano e a fórmula, esse símbolo aqui o & comercial é a concatenação, estou juntando dois textos.

[06:35] Estou juntando o conteúdo dessa coluna com espaço em branco com o valor do ano, se eu der “Enter” tenho o conteúdo da primeira coluna, aqui tem um espaço em branco bem pequeno e depois vem o ano, vamos repetir isso agora para o mês. Então vou de novo, modelagem, nova coluna, eu vou chamar essa coluna de mês, ano, vai ser o nome do mês concatenado com espaço, concatenado com o ano.

[07:30] Vou dar um “Enter” e tenho também o meu mês junto com o nome do ano carregados. Então pronto, eu calculei duas colunas calculadas sem a necessidade de ir no editar consultas e construir uma função de transformação dentro dos menus, eu fui através de linhas de comando dentro da área de fórmulas, usando fórmulas na sintaxe da linguagem DAX para calcular três novas colunas.

[08:10] Antes de terminar, eu sempre esqueço, tenho que lembrar, vamos salvar o projeto, arquivo, salvar como, esse vai ser o arquivo 10, pronto, agora sim terminamos o vídeo.

[00:00] Vamos continuar? Vamos agora aplicar uma função chamada switch que faz parte da sintaxe do DAX e é uma forma um pouco mais complexa do que propriamente multiplicar duas colunas ou concatenar colunas. Vamos mostrar o problema e vamos ver como resolvemos isso através dessa função switch. Vamos olhar a tabela de clientes, clicar aqui em tabela, temos a tabela de clientes, eu quero criar uma coluna que mostre para mim o gênero do cliente.

[00:47] Se ele é masculino ou se ele é feminino, e eu vou usar isso, usando a coluna título, se eu tiver Sr vai ser masculino, se eu tiver Sra vai ser feminino. Eu vou vir em modelagem, nova coluna, ele me abriu a nova coluna e aqui em cima eu vou colocar o nome da coluna, “gênero” aqui eu vou colocar a função switch que ela vai servir como um if, se a coluna for Sr vai ser masculino, se não ela vai ser feminina.

[01:35] Só que o switch ele é mais próximo de um “case” do que de um “if”, ou seja, caso seja a opção 1 faça isso, caso a opção 2 faça aquilo, ou seja, o switch pode ter uma gama de opções e valores. Então pode ser que essa fórmula seja grande e eu posso expandir essa área de trabalho, basta eu clicar nessa seta do lado, ao fazer isso eu tenho uma área maior para trabalhar, aqui se eu der um “Ctrl” com “Enter” apertado eu avanço de linha para baixo ou eu subo de linha para cima.

[02:30] Então vamos usar a opção switch, na verdade eu vou escrevendo as primeiras letras ele vai me sugerindo a função, está aqui a função switch e ele me mostra aqui a sintaxe dela. Eu começo com uma expressão, na verdade a expressão é o nome da coluna, aqui eu vou escrever cliente, na verdade a coluna é clientes títulos, ponto e vírgula, primeiro valor, coloco aqui Sr que é o primeiro valor de teste, ponto e vírgula se for Sr eu vou escrever masculino, ponto e vírgula se for Sra eu vou escrever feminino.

[03:35] E sempre a função switch termina com um Else que é um valor caso eu não ache nem a primeira opção, nem a segunda, vou colocar aqui, “desconhecido” e eu fecho a fórmula do switch, vou dar um “Enter” sem o “Ctrl” apertado, significa que eu vou confirmar a edição dessa função, ele já escreveu aqui em baixo, já tenho a minha coluna de gênero, masculino e feminino, vamos fazer um switch um pouco mais complexo.

[04:28] Eu tenho aqui na tabela de clientes os bairros das cidades, e eu vou classificar esses bairros de zona A ou zona B, vamos supor que eu queira separar esses bairros em regiões, zona A ou zona B, se eu vier em bairro e clicar nessa seta eu vou conseguir ver todos os bairros que eu tenho, então tenho esses bairros todos. Eu vou até fazer o seguinte, eu vou deixar essa coluna aqui, botão direito do mouse, copiar, copiei o conteúdo da coluna.

[05:35] Vou manter na memória do clipboard da minha máquina, vou vir em modelagem e vou nova coluna, vou abrir com “Ctrl” eu vou descer aqui e vou dar um “Ctrl + V”, está lá os valores que eu copiei lá da coluna, então vamos usar aqui zona, é o nome coluna, vou usar o switch, abriu parênteses, eu vou pegar bairro, ponto e vírgula e eu vou colocar se o bairro for jardins eu vou chamar de zona A “Jardins”; “Zona A”; “Água Santa”; “Zona B”; “Tijuca” “Zona B”; “Inhauma”; “Zona B”; “Barra da Tijuca”; “Zona A”.

[07:20] “Humaitá”; “Zona A”, aqui está vendo uma coisa? Ele me marcou um erro aqui porque eu esqueci o ponto e vírgula aqui em cima, “Lapa”; “Zona B”; “Santo Amaro”; “Zona A”; “Brás”; “Zona B”; vamos olhando a fórmula e vamos vendo o que estamos esquecendo, tem uma marcação aqui é porque tem que ter um ponto e vírgula, “Cidade Nova”; “Zona B”; “Pieade”; “Zona B”; e no final é o Else, caso eu não ache nenhuma das opções, “A Classificar” vamos fazer assim e eu fecho parênteses.

[09:05] Não tem nenhuma linha vermelha então está me dizendo que a função switch está correta, então basta eu dar “Enter” e ela já vai para tabela. Tenho a minha coluna Zona A ou Zona B dependendo da classificação. Então vimos dois exemplos da função switch, uma bem simples e outra mais complexa, mas ela funciona como um “case” para podermos classificar cada linha da nossa coluna. Vamos salvar o nosso arquivo, esse vai ser o nosso projeto 11, agora sim, vamos encerrar.

[00:00] Vimos na aula passada um pouco de DAX, vamos agora nessa aula ver mais algumas funções, pelo menos as mais usadas, se você for na internet e abrir um site da Microsoft vocês vão ver um help sobre as funções DAX e tem muita coisa aqui, por exemplo, função de data, tem várias funções, Calendar, CalendarAuto, Date e assim por

diante, time-intelligence functions, data entre duas datas, tem aquelas funções de acumular do ano até a data, enfim, várias funções.

[01:00] Funções de filtro, funções de informação, funções lógicas, if, not, switch que nós vimos, funções matemáticas, várias funções. Enfim, você tem uma gama de funções, vamos ver só algumas e você pode ir depois pesquisando em fóruns da internet, documentação da Microsoft caso vocês queiram usar uma função específica dentro do seu projeto de Power BI.

[01:37] Vamos voltar para o nosso projeto do jeito que estava salvo no último vídeo da aula passada e vamos criar mais uma coluna calculada que no caso vai ser o faturamento sem impostos. Olha o nosso problema, eu tenho a tabela de itens nota fiscal, que eu tenho um faturamento, vou até olhar aqui na tabela que é melhor, o faturamento está na tabela de itens, ele não vai aparecer aqui claro porque é uma coluna calculada, mas está aqui.

[02:22] Só que eu quero tirar fora o imposto, ou seja, eu tenho um imposto que está na tabela de notas fiscais e eu quero pegar o faturamento e tirar o imposto para saber o faturamento líquido, é uma fórmula, só que eu estou usando uma fórmula em uma coluna de uma tabela usando o valor que está em outra tabela e como fazemos isso?

[02:55] Eu vou primeiro criar uma nova coluna que vai ser o imposto, mas vai ser o imposto aplicado a cada número de nota fiscal, porque há uma ligação pelo número da nota fiscal entre a tabela de notas e de itens, está aqui, vou destacar a ligação, é isso que vai juntar o faturamento que está aqui com o imposto que está aqui. Então eu vou vir aqui em modelagem, nova coluna, e eu vou chamar aqui de imposto % por item.

[03:45] Agora eu quero me referenciar a uma coluna da outra tabela, se eu escrever notas, eu só consigo me referenciar a campos da própria tabela, mas eu quero pegar o campo impostos que está na nova tabela, como eu faço isso? Eu uso a função related, ao usar o related eu dentro do parâmetro da related eu posso me referenciar a qualquer campo do meu modelo. Então eu vou pegar aqui na tabela notas o imposto.

[04:33] Se eu der Enter eu tenho a coluna aqui que é o imposto para cada nota, então essa linha aqui pega o número dessa nota fiscal, vai na tabela de notas e captura o imposto e coloca aqui. E agora eu posso ter a outra coluna que é o faturamento sem imposto. Modelagem, nova coluna, vou botar aqui: $1 \text{ Faturamento em R\$ sem Imposto} = \text{'Itens de Notas Fiscais' [Faturamento em R\$]} * (1 - \text{'Itens de Notas Fiscais' [Imposto \% por Item]})$.

[05:33] Lembrando que imposto é um valor menor que 1, se eu quero ter o faturamento líquido eu tiro de 1 sobre o imposto, então se o imposto é 0.10, 1 menos 0.10 vai dar 0.9, então só vou pegar 90% do faturamento, que vai me dar o faturamento sem o imposto. Então tenho aqui o faturamento sem o imposto, então com imposto o faturamento aqui foi 1.78, sem imposto foi 1.56.

[06:12] Então vimos como funciona a função related para buscarmos um cara que está em outra tabela, vamos salvar esse projeto antes de terminar o vídeo, vou colocar aqui 12, pronto.

Aula 05 – Trabalhando com medidas

[00:00] Vamos, agora, mostrar a vocês uma entidade do Power BI chamada Medida. Parece que não, mas ainda não vimos medida no Power BI. Quando vamos em Dados e olhamos, por exemplo, a tabela Itens de Notas Fiscais, esses valores que temos aqui não são medidas, elas são colunas numéricas. São valores que estão associados à cada linha da tabela.

[00:32] Posso, como fizemos nos vídeos anteriores, criar colunas calculadas, mas elas não são medidas. Medidas são fórmulas que não estão associadas, diretamente, as linhas das tabelas e que podem ser usadas em nossos relatórios.

[00:53] Vamos ver a diferença entre uma coluna e uma medida, como ela é definida, fazendo o seguinte exemplo prático: tenho a coluna Quantidade em Litros, vou criar uma medida chamada Quantidade em Litros Acumuladas que será sempre a soma da Quantidade em Litros, e vamos comparar as duas e ver que, no final, vão ter sempre os mesmos valores, mas vamos ver que a medida que vou estar criando não vai estar associada à nenhuma linha da tabela.

[01:32] Apesar de a medida não estar associada à nenhuma linha da tabela, ela tem que estar amarrada a uma tabela, mas temos uma boa prática que é a seguinte: vamos criar uma tabela vazia, que não vai ter dados, e que vai estar associadas a ela somente as medidas. Então vamos fazer isso.

[01:55] Como eu crio uma tabela vazia dentro do Power BI? Eu venho em Pagina Inicial e Inserir Dados. Vou chamar essa tabela de *“Medida”*. Por que o *“”*? Eu coloquei o underscore porque depois, alfabeticamente, essa tabela, que vou usar somente para associar as medidas, vai estar sempre no topo da lista de tabelas.

[02:29] Aqui em cima, vou chamar essa coluna de *“Leia me”*, que vai ser uma coluna só com texto informativo. Vou escrever nesse texto: *“Tabela usada apenas para associar medidas”*. Então tenho a tabela aqui, só com essa linha, e vou clicar em Carregar. Pronto, eu tenho minha tabela chamada *“Medida”* e agora vamos criar uma nova medida. Para fazer isso eu preciso ir em Dados, selecionar a tabela, ir em Modelagem e notem vou criar uma Nova Medida.

[03:31] Até agora, quando criamos aquelas colunas calculadas, nós criamos nova coluna. A coluna é um cálculo que vai estar associado, na tabela, linha a linha. A medida não. Nova Medida, tenho minha área de comandos DAX e vou criar a medida Quantidade em Litros Acumulado. Como a medida não está associada à nenhuma linha da tabela, eu tenho que usar, sobre ela, alguma fórmula de agregação. O que é uma fórmula de agregação? É, por exemplo: soma, média, valor máximo e valor mínimo. Isso tudo são fórmulas de agregação.

[04:30] Como são quantidades em litros acumulada, a palavra acumulada lembra o quê? Lembra soma. Então vou usar *“SUM()”* e note que, no parâmetro da agregação da medida, eu consigo ver todas as colunas independente da tabela. Vou escolher, claro, minha coluna Quantidade em Litros e vou fechar o parêntese para poder terminar a fórmula DAX desta medida. Eu confirmo isso, posso dar um *“Enter”* para confirmar e eu tenho agora associado à minha tabela *_Medida* a nova medida criada.

[05:23] Se eu venho aqui em Dados, note que não consigo ver nenhuma coluna com isso calculado por que a medida não está associada a nenhuma linha da tabela. Vou fazer um exemplo de uma visão. Vou selecionar esse ícone de tabela, vou abrir ela um pouco

mais e vou arrastar sobre os campos dessa tabela, por exemplo, família de produtos. Vou vir aqui na tabela Produtos e vou escolher Família. Arrasto a família de produtos para Valores. Eu tenho minhas famílias de produtos.

[06:11] Eu posso vir aqui em Formato, Valores e aumentar um pouco a fonte. Vou colocar 20. Eu tenho as famílias de produtos. Se eu vier aqui em Itens de Notas Fiscais, note que a Quantidade em Litros já tem uma definição prévia de somatório. Ou seja, intrinsecamente, essa coluna já tem como regra: somar quando junto as linhas.

[06:53] Então se eu vir em Campos, arrastar aqui em baixo, notem que ficou uma linha amarelinha mostrando a posição onde esta coluna vai ficar. Se a linha está embaixo da área Família é por que ela vai ficar à direita da coluna Família. Então eu tenho minha quantidade de litros.

[07:32] Agora vou arrastar a medida Quantidade em Litros Acumulado aqui para baixo. Notem que os valores foram os mesmos. Se eu venho na Quantidade em Litros e seleciono não Soma, mas Média, notem que eu tenho agora uma média de quantidade em litros por família, mas a quantidade em litros acumulada não mudou. Porque não mudou? Porque ela é uma medida que já está intrínseca que soma as linhas do relatório.

[08:24] É claro que média da quantidade não é um bom exemplo para usar. Vou voltar “Soma” e consigo ver a quantidade em litros via coluna e a quantidade em litros acumulada, que está pegando da mesma coluna, mas usando agrupando em soma.

[08:48] Nesse caso, não teve nenhuma vantagem em usar medida. Por que eu tive a necessidade de criar uma medida somando se eu já tenho isso na coluna Quantidade? É claro que fazer isso foi um trabalho que fiz somente a nível didático para mostrar a vocês como funciona o conceito de medida, mas usamos Medidas para cálculos um pouco mais complexos, que vamos usar nas tabelas.

No min [4:59], o instrutor inverte os número digitados. Ao digitar o valor **12.574.896,99**, ele inverte e digita **12.574.869,99**.

[00:00] Vamos agora ver uma função DAX chamada SUMX(). Qual é a diferença do SUMX() para o SUM()? Lembrando que quando viemos na tabela _Medidas e criamos a “Quantidade em Litros Acumulada”, nós usamos o SUM(). O SUMX() é quando vou somar uma expressão. Vamos dar um exemplo prático para ver o que significa isso.

[00:30] Voltando para minha tabela de itens das notas fiscais, vamos supor que eu queira criar uma medida que me calcule o imposto em Reais. O imposto em reais que vai ser o faturamento em R\$ menos faturamento em R\$ sem imposto. A diferença é o imposto que paguei. Eu poderia criar uma nova medida. Cliquei na tabela _Medida, vou selecionar “Nova Medida” e vou escrever “Imposto em R\$”. Então ela é a soma da diferença.

[01:27] Então vou colocar aqui: “SUM(‘Itens de Notas Fiscais’[Faturamento em R\$]) -”. Note que, quando coloquei o “-” ele nem deixou mais eu colocar uma outra coluna. Por quê? Porque a função SUM() só aceita um parâmetro. Eu souo uma coluna. Eu não consigo somar uma expressão numérica, uma fórmula. Na verdade, eu quero pegar a diferença de cada linha e depois somar. Para fazer isso, eu não uso o SUM(), eu uso SUMX(). O “X” vem de expression. Eu vou apagar aqui e vou usar o SUMX().

[02:35] Note que agora eu vou especificar uma tabela e dentro dela uma expressão. Que é a coluna faturamento com imposto menos o faturamento sem imposto. A tabela é a “Itens de Notas Fiscais” e a expressão é: “Faturamento em R\$” – “Faturamento em R\$ sem Imposto”. Ou seja, estou aplicando essa fórmula para cada linha e, depois, eu somo. Isso é o que vou fazer no caso dessa expressão.

[03:30] Então clico de novo em Relatório, tenho essa minha “Imposto e R\$”. Para podermos fazer a conta corretamente, vamos fazer o seguinte: vou clicar na tabela, note que ficou destacado. Vou tirar essas duas colunas clicando nesse X aqui do lado. Então vou colocar aqui, da tabela “Itens de Notas Fiscais”, primeiro o “Faturamento em R\$”, depois o “Faturamento em R\$ sem Imposto” e, agora, o “Imposto em R\$”.

[04:36] Vamos pegar uma calculadora e note que se eu pegar essa primeira linha, deu praticamente a mesma coisa. Deu um erro de uns 10, 15, 20 reais. Isso tem a ver com precisão numérica. Pois estou

fazendo a conta linha a linha e, na medida que ele vai agrupando, vai perdendo um pouco a precisão numérica, mas o resultado é esse. Então consigo ter a soma da diferença entre as duas colunas.

[05:54] Uma outra maneira de checar essa conta é fazer o seguinte: voltar na tabela Itens de Notas Fiscais, calcular uma coluna que faça essa conta e comparamos. Vou vir em Modelagem, Nova Coluna, vai ser: “Imposto em R\$ Tabela = SUMX(‘Itens de Notas Fiscais’; ‘Itens de Notas Fiscais’[Faturamento em R\$] - ‘Itens de Notas Fiscais’[Faturamento em R\$ sem Imposto]”. Aplico e, voltando, se eu pegar essa coluna nova, agora é uma coluna, não é uma medida, os números bateram.

[07:14] Agora o seguinte: a coluna, ou seja, essa da direita, fica guardada como um dado dentro do disco, enquanto que esta primeira, que é a medida, é calculada em memória no momento em que faço o relatório. Quando falamos de um relatório pequeno como esse, que tem poucos dados, não faz diferença, mas se estivermos trabalhando com volumes de dados muito grandes, você criar uma coluna calculada e deixar guardada fisicamente dentro do Power BI, vai dar problema de performance principalmente na atualização do relatório.

[07:57] Lembra quando atualizamos o relatório? Se eu clico em Pagina Inicial e Atualizar, estou fazendo cargas e cálculos. Todas aquelas colunas calculadas que criamos estão sendo calculadas aqui, já as medidas não. Elas só serão calculadas quando puxarmos ela para um relatório.

[08:28] Então: criar colunas calculadas não é nenhum problema. O problema é que, dependendo do volume, não é uma boa prática. É melhor você criá-la e deixá-la para que seja exibida como medida. Então você precisa identificar em que momento vai usar o SUM() e em que momento vai usar o SUMX(). Antes de terminar, vamos salvar esse relatório para depois continuar nos próximos vídeos. Arquivo, Salvar como: “CursoPowerBI_014.pbix”.

[00:00] Uma coisa que fazemos muito quando usamos medidas é, por exemplo, calcular a participação percentual de um indicador sobre o total. Vamos fazer o seguinte, eu tenho nosso indicador “Imposto em R\$” e eu quero saber quantos por cento ele tem em relação ao total. Para fazer isso, primeiro eu preciso ter uma coluna que com o total

relacionado linha a linha. Eu já tenho minha expressão “Imposto em R\$”, que foi uma conta, e vamos usar o SUMX() para estar calculando uma outra medida que vai me dar o total agrupado deste indicador.

[01:00] Vamos clicar em Dados, _Medida e vou Criar uma Nova Medida. Vou chamar ela de “Total do imposto em R\$ = SUMX(‘Itens de Notas Fiscais’; ‘Itens de Notas Fiscais’[Faturamento em R\$] - ‘Itens de Notas Fiscais’[Faturamento em R\$ sem Imposto])”. Essa foi a expressão que fizemos no vídeo anterior, mas vou fazer o seguinte: na frente da tabela vou escrever: “ALL()”. Ao escrever ALL() eu estou agrupando a tabela para depois fazer a conta. Vamos ver o resultado? Eu vou tirar essas colunas e só deixar o Faturamento em R\$, que é nossa medida.

[02:40] Vou pegar, agora, o total do imposto. Note que, esse valor total se repete linha a linha. Agora eu posso aplicar uma outra medida que é: dividir “Imposto em R\$” por “Total do Imposto em R\$”. Vamos criar uma outra medida. Vou vir em Dados, Modelagem, Nova Medida, “Percentual de Imposto em R\$” e vou usar uma outra função DAX: DIVIDE(). Ela já trabalha a divisão por zero. Se eu encontrar um denominador zerado ela trata e não vai dar erro. Vou usar: “DIVIDE([Imposto em R\$];[Total do imposto em R\$])”.

[04:40] Vou voltar a minha tabela. Se eu arrastar o percentual aqui, para olhar o resultado, tenho a participação. Claro que essa aqui nós tiramos, pois ela não precisa aparecer. Ela só está sendo usada para calcular a participação do Imposto. Isso vale para qualquer Dimensão. Estou olhando aqui Família. Se eu jogar, por exemplo, Bairro, o percentual continua valendo, está sendo calculado corretamente.

[05:29] Ou seja: a medida, usando as expressões (SUMX, DIVIDE e SUM), está calculando, em memória, valores que eu não conseguiria se elas estivessem como colunas na minha tabela. Vamos salvar esse relatório: vou clicar em arquivo, salvar como “CursoPowerBI_015.pbix”.

Aula 04 – Análise de Pareto com DAX

[00:00] Uma análise que eu costumo a fazer quase sempre em todos os cursos que eu ministro aqui da Alura quando claro, o tipo do curso permite, é fazermos um relatório da análise de pareto, ou conhecido como análise de 80/20, o que diz a análise de 80/20? A análise 80/20, diz que normalmente 80% dos valores de um indicador são gerados por apenas 20% do número de geradores daquele indicador.

[00:40] Ficou complicado? Vou dar exemplos práticos, estatisticamente dizem que 80% da minha venda só está em 20% dos meus clientes, ou 80% dos meus custos normalmente estão em 20% dos meus departamentos, claro que a relação 80/20 ela não é mágica, ela é um parâmetro, se você vai fazer a sua análise de estatística e descobre que 80% das suas vendas só estão em 5% dos seus clientes, ou seja, em vez de você ter um 80/20, você tem um 80/05.

[01:25] Você tem um problema, por quê? Grande parte das suas vendas estão concentradas em poucos clientes, só 5% dos seus clientes representam 80% das suas vendas, se um daqueles clientes que fazem parte dos 5% deixam de comprar com você, você vai perder grande partes das suas vendas. Se eu tenho um 80/05, no caso dos custos, ou seja, se eu descobro de 80% dos meus custos só estão em 5% dos meus departamentos, isso é bom.

[02:08] Porque basta eu economizar dinheiro naqueles 5%, ou seja, eu vejo quais são os departamentos que me geram mais custos, faço um trabalho específico neles e eu consigo melhorar o todo. Agora, se eu tiver um 80/50, por exemplo, 80% das minhas vendas estão em 50% dos meus clientes, isso é bom, porque eu tenho os meus clientes bem distribuídos, se um desses clientes deixar de comprar comigo eu não perco muito, porque as minhas vendas estão bem distribuídas dentro da minha carteira de clientes.

[02:47] No caso dos custos isso não é muito bom, porque se meus custos estão bem distribuídos, para eu fazer uma economia de custos eu tenho que fazer uma ação em vários departamentos ao mesmo tempo para que o todo melhore, então assim, o 80/20 se aplica em vários tipos de análise, aqui por exemplo, eu estou vendo um exemplo

de um relatório no Power BI de 80/20, mostrando as vendas em reais sem impostos dos produtos.

[03:25] Então na primeira coluna eu tenho as vendas do maior produto para o menor, ou seja, esse produto aqui, festival de sabores 2 litros açaí, teve como venda 1.724.170,84. O segundo produto festival de sabores 1,5 litros açaí, teve como venda 1,258.542,25. Então eu tenho aqui um ranking do maior para o menor, aqui do lado eu tenho um percentual de participação, esse 1 milhão 720 mil representou 11% de todas as vendas, agora se eu pegar 1 milhão 258 mil e pegar essas duas vendas acumuladas, ou seja, primeiro lugar e o segundo.

[04:29] Elas duas vão representar 19,71% das vendas, e por exemplo, se eu venho aqui e pego essas 5 vendas acumuladas, todas essas 5 representam 38% das minhas vendas, então se eu vou aqui e eu olho a outra coluna, significa o seguinte, quantos produtos eu tenho no total? Eu não tenho o contador aqui, mas eu tenho de cabeça, eu tenho aqui 33 produtos, ou seja, eu tenho 33 linhas diferentes aqui, ou seja, um produto representa 3,03% de todos os produtos, seria 1 sobre 33.

[05:15] Dois produtos representam 6.6, 3 produtos representam 9.09, note que esse crescimento é constante porque o número vai crescendo de 1 em 1. Ou seja, quando eu olho essa linha aqui, eu digo que 11% das minhas vendas estão em 3% dos meus produtos, quando eu olho essa linha aqui, 43.90 que é as vendas acumuladas desses 6 primeiros produtos eu digo que 43% das minhas vendas estão em 18.18% dos meus produtos.

[06:00] Por quê? Quantos produtos eu tenho aqui? 6, porque 6 produtos dentro de um total de 33 representam 18.18%, logo, sabendo que significa isso, se eu arrastar aqui para baixo e achar o ponto 80, o ponto 80 está aqui, ou seja, 80% das minhas vendas estão em 51.52% dos meus produtos, ou seja, eu tenho um 80/50, isso é bom, eu tenho as minhas vendas bem distribuídas por produtos.

[06:47] Entendendo como conseguimos entender o 80/20 nesse tipo de relatório o nosso desafio agora nos próximos vídeos é construir esse relatório usando medidas e usando algumas funções DAX no nosso Power BI, então vamos nessa, vamos continuar no próximo vídeo.

[00:00] Então vamos fazer o seguinte, eu vou fechar esse relatório que está aqui, porque ele já está com o relatório montado e vou abrir o arquivo que eu salvei na penúltima aula, no penúltimo vídeo, que no último vídeo eu falei um pouco sobre a análise de pareto, mas na penúltima já estávamos criando algumas medidas, eu vou abrir aquele relatório que foi salvo.

[00:35] Só que claro, não vou aproveitar a tabela que está aparecendo aqui, porque a tabela está me mostrando a distribuição por bairros, eu vou clicar aqui sobre a tabela e aperto o botão “Delete”. Vamos aqui construir uma nova tabela, então eu vou clicar aqui, vou abrir ela um pouco mais, arrastar um pouco para baixo e vou arrastar para ela nome de produtos.

[01:17] Eu vou aqui em produtos, eu tenho aqui nome dos produtos e arrasto aqui para o lado, ficou um pouco pequeno a letra, mas não tem problema, clicamos na parte de formatação, vamos aqui em grade e aumentamos o tamanho da fonte para 14 por exemplo. Vou abrir um pouco mais aqui e eu vou trabalhar com aquele indicador que é imposto em reais, é uma medida, já criamos ela aqui, eu até errei a grafia, escrevi impostos e reais, mas na verdade é imposto em reais.

[00:02] Selecionando a tabela, eu arrasto ele para cá, já tenho o valor, se quisermos consertar o label que está errado, basta eu clicar em cima do imposto e reais, duplo clique, vou escrever aqui. Agora já tenho aqui o nome certo. Para construirmos análise de pareto, temos alguns passos, eu vou começar criando o posicionamento do ranking, posicionamento do ranking é o seguinte, eu tenho os valores, alguns são valores maiores, outros menores.

[02:50] Eu quero saber dentro dessa lista quem é o primeiro lugar, segundo lugar, terceiro lugar ou quarto lugar e assim por diante. Claro, isso vai ser uma nova medida, então eu vou clicar aqui em medida, onde eu tenho aqui 3 pontinhos e seleciono nova medida, aqui na minha medida eu vou escrever “Posição no Ranking” e eu vou usar uma função em DAX chamado “RANKX” onde eu passo aqui uma tabela.

[03:38] Por exemplo, eu quero ranquear o nome do produto, então eu vou colocar a tabela produtos, ponto e vírgula, qual é o indicador que eu estou ranqueando? É o imposto em reais, vou clicar nesse V para

confirmar a criação da medida, ela está aqui e eu vou arrastar para cá, note que tem uma coisa meio errada, está tudo 1. Na verdade é como se eu tivesse ranqueando cada produto que não está participando do ranking do total.

[04:30] Que na verdade, quando eu usei aqui a tabela como um todo eu estou ranqueando cada valor de produto por cada produto diferenciado. Então ele vai estar sempre um a um, imagina eu quero saber qual é a posição desse cara comparado só com ele, o conjunto de comparação é ele próprio, então vai estar só 1. Agora se eu vier aqui e escrever "ALL" e quando eu uso a função All eu tenho que colocar a tabela e um campo, ou seja, produtos e eu vou pegar aqui o nome do produto.

[05:29] Ao fazer isso eu tenho que fechar aqui um parênteses, vamos ver aqui onde que eu estou errando, rank All, esse parênteses fechou, tem um parênteses a mais aqui atrás, agora sim, a expressão ficou assim, RANKX(ALL(Produtos[Nome do Produto]);([Imposto e R\$])). O que eu estou vendo aqui é a posição, ou seja, em termos de impostos e reais esse primeiro produto aqui ele está no ranking 16 ele é o 16º mais vendido, esse aqui de baixo é o oitavo mais vendido.

[06:25] Esse daqui é o primeiro mais vendido, então essa função RANKX quando colocamos dentro do All o nome da coluna que eu estou ranqueando e o indicador, eu vejo aqui a posição do rank deles, claro que se eu quiser ver isso ordenado, basta eu clicar sobre a coluna imposto em reais e eu vejo a minha visão coordenada, do maior para o menor.

[07:00] E claro, ao fazer do maior para o menor o ranking ele fica crescente, ele começa no 1, 2, 3 e 4 e assim por diante, porque eu estou pegando do maior para o menor. Então esse foi primeiro passo que fizemos para construir a análise de pareto ou de 80/20, para não perdemos o que fizemos eu vou vir aqui em arquivo, salvar como e eu vou salvar esse como o arquivo 16, pronto, continuamos no próximo vídeo.

[00:00] Eu agora quero construir uma coluna aqui, que seja uma coluna que tenha o imposto em reais acumulativo, cada linha tenha o

seu valor acumulado, ou seja, na primeira linha eu vou ter o valor da primeira linha, na segunda linha a soma do primeiro e do segundo, na terceira linha a soma do primeiro, do segundo e do terceiro e assim por diante. Eu vou aplicar aqui uma nova medida e nessa nova medida eu vou chamar de imposto acumulado, eu vou usar a função DAX calculate.

[00:45] A função DAX calculate aplica uma operação matemática sobre um filtro, sobre um pedaço de um conjunto de linhas, no caso a minha expressão é o imposto em reais, e o filtro eu vou selecionar um conjunto de linhas, primeiro eu vou fazer isso aqui, eu vou selecionar o “ALL(Produtos[Nome do Produto])”, o que eu estou fazendo aqui? O meu filtro são todos os produtos e eu vou calcular esse valor para todos os produtos.

[01:45] Eu tenho aqui o imposto acumulado, vou clicar aqui na tabela, eu tenho o valor dele, só que o valor dele, de proposito é o valor total, por que? O imposto em reais ele já é um SUMX, ou seja, ele já é uma soma, então e estou aplicando no imposto acumulado a soma, ele já está somando, está somando todos os produtos, então na verdade o valor, cada linha que a função está sendo aplicada, na verdade está sendo aplicada a soma de todo mundo.

[02:30] Por isso que o valor é igual, não é isso que eu preciso, eu preciso filtrar esse conjunto de dados, eu vou usar então a função TOPN que funciona da seguinte maneira, eu passo uma posição, por exemplo, vou botar o número 3, o 3 significa o seguinte, eu vou pegar os 3 maiores, dentro de que conjunto? Dentro de todos os produtos e pegando os 3 maiores de que expressão? Do imposto. Vou aplicar.

[03:15] Olha só, deu um valor constante mas deu um valor menor, quem é esse 4.055.755,13? Ele é o seguinte, vou pegar a calculadora do Windows, eu vou somar, estou usando o TOPN(3) então vou somar esses 3 maiores, esses 3 maiores é o 1.724.170,84, eu vou somar 1.258.542,25, tenho esse subtotal e vou somar ele com 1.073.042,04, olha o valor, 4.055.755,13, então se eu colocar TOPN(3) eu tenho a soma dos 3 maiores, se eu colocar TOPN(5) eu vou ter a soma dos 5 maiores.

[04:38] Mas eu queria que na primeira linha eu tivesse a soma do primeiro, na segunda linha a soma dos dois maiores, na terceira linha a soma dos 3 maiores e assim por diante, onde que eu tenho a posição? Onde que esse número 5 está variando? Está variando aqui, então ao invés de usar o top 5 eu vou usar o valor da coluna posição do rank, ou seja, aqui em vez de eu fixar os 3 primeiros, é a posição do rank e a posição do rank está aqui.

[05:20] Então quando essa posição for aplicada nessa linha, posição do rank vai ser 1, quando aplicar ela sobre a segunda linha, posição do rank vai ser 2, quando aplicar ela na terceira linha, posição do rank vai ser 3, então se eu confirmar eu tenho aqui o valor acumulado. Aqui é a soma dos 3 maiores, é a soma dos 4 maiores, ou seja, tenho o acúmulo.

[05:57] Então consegui agora construir mais uma coluna necessária para podermos estar concluindo o relatório de 80/20, vamos antes de terminar o vídeo salvar como, vou salvar esse cara como 17, cliquei em salvar, pronto, tenho o meu vídeo salvo.

[00:00] Eu agora preciso ter o percentual de cada linha onde eu tenho imposto acumulado em relação ao total, vamos fazer a coisa por

partes, eu vou criar aqui uma medida que eu vou chamar ela de percentual do imposto acumulado, se eu pegar isso e fizer isso aqui, “CALCULATE(Imposto em R\$); ALL((Produtos [Nome de Produto]))”, lembra calculate é pegar uma expressão e aplicar um filtro, se eu pegar o imposto acumulado e jogar aqui na tabela, eu tenho o valor total.

[01:10] Mas eu preciso ter o quê? Esse dividido por esse, esse dividido por esse. Porque vai me dar o imposto acumulado, podemos até fazer o seguinte, podemos primeiro ter uma coluna com o percentual de cada linha e depois o percentual acumulado, vamos pegar aqui de novo, se eu por acaso pegar aqui, “DIVIDE” eu vou pegar o imposto em reais e dividir por esse cara aqui, eu tenho um percentual da linha, ou seja, esse cara aqui é esse valor dividido por esse final.

[02:13] Esse segundo, esse 0.8 vai ser 1.250 dividido por esse total, então eu tenho um percentual linha a linha, se eu quiser que essa coluna ela fique com formato de percentual eu clico nela e vou aqui em modelagem e clico nesse percentual aqui. Então tenho os percentuais linha a linha, só que eu preciso do acumulado, então o acumulado é o seguinte, em vez de eu usar essa coluna aqui eu vou usar essa coluna aqui. Então em vez de imposto em reais eu vou usar a coluna imposto acumulado.

[03:15] Vamos aplicar, claro que ficou em uma posição errada porque eu tenho que ordenar isso aqui por esse valor, agora sim, o primeiro representa 11%, o dois primeiros representam 19%, os três primeiros 26% e assim por diante, um crescente até a última linha que tem que representar 100%. Então a função final Imposto Acumulado = DIVIDE ([Imposto Acumulado];CALCULATE([Imposto em R\$]; ALL(Produtos[Nome do Produto]))). Vamos salvar como 18 e pronto, aos poucos o nosso relatório 80/20 está sendo construído.

[00:00] Agora eu preciso do percentual acumulado do número de produtos, vamos olhar aqui em baixo esse número 31 que está aqui, 31 é o maior rank dos produtos, esse 31 se repete porque quando eu aplico o ranking eu desprezo os 3 últimos produtos que estão com valores zerados, logo o que acontece? Quantos % o primeiro produto representa em termo de todos os produtos? Vai ser o 1 dividido pelo 31.

[00:50] Os dois primeiros produtos? Vai ser o 2 dividido pelo 31, então eu preciso na verdade ter uma coluna aqui onde eu tenho o número 31 repetido, toda linha, 31, 31 e depois eu faço a conta, um dividido pelo outro, como eu consigo fazer essa conta? Basta eu por exemplo aplicar uma função que eu tenho no DAX chamada MAXX a MAXX me dá o máximo de uma expressão, a expressão que eu tenho é a expressão do rank, então eu vou fazer o seguinte, eu vou vir aqui em medida, posição do ranking foi essa expressão que eu usei.

[01:40] Eu vou pegar essa expressão que está aqui e vou copiar ela, vou dar aqui um “Ctrl + C”, vou deixar na memória do meu clip board e eu vou vir aqui no leme e vou criar uma nova medida onde eu vou chamar de total produtos, vai ser o seguinte, vai ser o MAXX de uma tabela ou de um vetor de valores, que no caso é o “ALL(Produtos [Nome dos produtos])” e eu coloco uma expressão, qual a expressão que eu vou colocar? Agora eu dou o “Ctrl + V”.

[01:45] Eu coloco aquela expressão que foi usada, que é no caso a posição do ranking, fecho o parênteses, vou confirmar, vou pegar aqui total produtos e arrastar aqui em baixo, então tenho agora aqui o 31 repetido, com isso eu tenho como calcular o percentual dos produtos, nova medida, vou colocar aqui, percentual produtos, vai ser um “CALCULATE”, porque calculate vai ser calculado de linha a linha, eu vou pegar a posição do ranking e dividir pelo total produtos.

[03:53] Vou confirmar, vou arrastar esse último indicador, tenho ele aqui, vou selecionar essa coluna para poder ver aqui em modelagem e clicar em percentual, volto a clicar na primeira coluna para ficar ordenado do maior para o menor e tenho aqui o meu percentual acumulativo, claro que o relatório 80/20 não precisa ter todas as colunas exibidas aqui, então vou vir aqui e eu vou tirar essas. Vou manter só o valor original ranqueado e os percentuais.

[04:45] Então tiro essas. Na verdade o imposto em reais fica, eu até acabei tirando uma que não deveria ter tirado, pronto, tem que estar ordenado por essa, então pronto, está aqui o meu relatório de 80/20. O 80/20 está acontecendo aqui, 80% das minhas vendas estão em 54.84% dos meus produtos, então tenho aqui o meu relatório pronto, esse total aqui, até que esses valores estão legais, mas esse valor aqui não diz muita coisa.

[05:35] Então podemos por exemplo ao selecionar aqui a tabela, vir aqui na parte de formato total e aqui eu desativo a linha de total, e eu fico aqui com meu relatório 80/20 já montado, vamos salvar, arquivo, vou salvar esse como o 19. Pronto.