

Exercício 1_07 - Importar dados de uma pasta¹

Um cenário bastante comum, nas organizações, consiste em importar dados com uma determinada periodicidade. Admita, por exemplo, que é o responsável pelo departamento de vendas da empresa “**Vende Bem, SA**” e que pretende fazer uma análise anual da evolução das suas vendas. Por norma, colocará numa pasta o mapa mensal de vendas, pretendendo que o ficheiro, anual, no “**Power BI**” seja automaticamente atualizado. O ficheiro inicial “**exercicio_1_07_jan.csv**”, relativo ao mês de janeiro, é trabalhado na pasta “**exercicio_1_07 (mapas_em_curso)**”. Quando fechado, este ficheiro, é transferido para a pasta “**exercicio_1_07 (mapas_fechados)**”, dando-se início ao tratamento de um novo ficheiro relativo às vendas realizadas no mês de fevereiro “**exercicio_1_07 fev.csv**”.

Vamos carregar no “**Power BI**” o ficheiro “**exercicio_1_07_jan.csv**”, depois de colocado na pasta “**exercicio_1_07 (mapas_fechados)**”, seguindo o caminho “**home → get data → folder**”, e proceder a algumas transformações nos dados importados: eliminar a primeira coluna; reorganizar colunas (colocar a data no início) e converter a margem para **valores decimais**.

Depois de importado e tratado o ficheiro, relativo ao mês de janeiro, vamos atualizar a pasta “**exercicio_1_07 (mapas_fechados)**” com cada um dos restantes ficheiros mensais, provenientes da pasta “**exercicio_1_07 (mapas_em_curso)**”, à medida que vão sendo fechados. Não esquecer de ir fazendo o refresh do “**Power BI**” por cada ficheiro transferido para a pasta “**exercicio_1_07 (mapas_fechados)**”.

Concluído o exercício, grave o ficheiro criado no “**Power BI**” com a designação “**exercicio_1_07_R.pbix**”.

¹ Nogueira, N. (2018). POWER BI – para Gestão e Finanças. FCA – Editora de Informática. Lisboa

Exercício 1_08 - Importar dados de diversas fontes²

Um outro cenário, também bastante comum, consiste em importar para o “**Power BI**” dados de diversas fontes. Considere a situação em que necessita de analisar dados relativos ao desempenho de um hotel e que os dados estão disponíveis em ficheiros diferentes:

- “exercicio_1_08 (regiões).txt”;
- “exercicio_1_08 (hotéis).accdb” e
- “exercicio_1_08 (reservas).xlsx”.

Após a importação destes ficheiros, grave o ficheiro criado no “**Power BI**” com a designação “exercicio_1_08_R.pbix”.

² Carvalho, A. (2019). Exercícios de Power BI – Importação, Edição e Visualização de dados. FCA – Editora de Informática. Lisboa

Exercício 1_09 – Editar dados e usar as funções “trim” e “clean”³

Continuando a analisar, situações com que os utilizadores de dados organizacionais se possam deparar com frequência, vamos considerar a situação em que é necessário remover espaços que possam existir em vários campos de dados, após a sua importação para o “**Power BI**” e que possam ter impacto na sua análise.

Vamos importar para o “**Power BI**” o ficheiro “**exercicio_1_09.csv**”. Para isso podemos seguir o seguinte caminho: “**home → get data → Text/CSV**” e seleccionar o ficheiro desejado. Como pretendemos fazer alterações aos dados importados, vamos trabalhar com o “**Power Query Editor**”, para isso, clicamos no botão “**transform data**”.

O “**Power Query Editor**”, oferece um ambiente muito rico e poderoso com algumas funções de limpeza muito úteis. Olhando para os dados na coluna “**name**” podemos verificar que os mesmos apresentam problemas relacionados com a existência de espaços extra: espaços precedendo e sucedendo algumas entradas e também entre palavras.

Para corrigir estas anomalias podemos recorrer à função “**trim**”, que remove todos os espaços em branco à esquerda e à direita do texto “**transform → format → trim**”, e à função “**clean**”, que devolve o texto com todos os caracteres não imprimíveis removidos “**transform → format → clean**”. Após estas transformações verificamos que ainda subsistem espaços em branco que podem ser resolvidos recorrendo ao comando “**home → replace values**”.

Vamos ainda proceder a uma alteração que obrigue a primeira letra da coluna “**gender**” a ser maiúscula “**transform → format → capitalize each word**”. Após concluir estas transformações, grave o ficheiro com a designação “**exercicio_1_09_R.pbix**”.

³ Gamble, G. (2018). *Power BI Step-by-Step Part 3: Power Query, Parameters, Templates & Custom Functions*

Exercício 1_10 – Limpeza e metadados⁴

Utilizando o ficheiro “**exercicio_1_09_R.pbix**”, criado no exercício anterior, entre novamente no “**Power Query Editor**”, seguindo o caminho “**home → transform data**”. No “**Power Query Editor**”, podemos ver a consulta final que criámos no exercício anterior no painel etapas aplicadas.

Temos sempre a oportunidade de substituir nomes codificados por nome explícitos. Podemos, por exemplo, renomear a própria consulta. Para tal clique com o botão direito do rato na consulta e escolha a opção “**renomear**”. Considerando que o conjunto de dados, utilizado neste exercício, corresponde a uma lista de assinantes, podemos renomear a consulta para “**assinantes**”, premindo de seguida “**ENTER**”. O mesmo resultado pode ser obtido, com recurso ao painel de propriedades em que, bem no topo, temos um campo “**nome**” que pode ser alterado.

Podemos igualmente, renomear colunas, clicando simplesmente duas vezes no nome da coluna existente e digitar um novo nome. Por vezes, basta inserir espaços entre as palavras para as tornar mais legíveis; por exemplo, alterando “**username**” para “**user name**”. Na coluna identificada como “**DOB**”, temos uma abreviatura padrão, que podemos substituir por “**Date of Birthday**”.

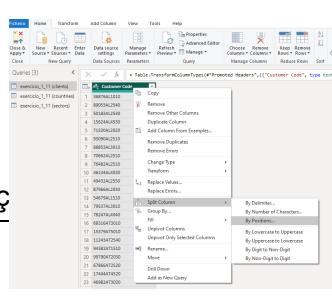
Há também a possibilidade de renomear etapas da consulta, clicando com o botão direito do rato, sobre a designação da etapa que pretendemos renomear, e seleccionar “**renomear**” no menu de contexto. Podemos renomear a etapa “**Trimmed text**” para “**Transform/format/trim**”, para dar indicação a outros utilizadores da localização do comando que foi usado para criar a etapa. De forma similar, podemos alterar “**Cleaned text**” para “**Transform/format/clean**”.

Finalmente, podem ser associadas descrições a cada uma das etapas, clicando com o botão direito do rato na etapa pretendida e seleccionar “**propriedades**” no menu de contexto. Assim serão inseridos metadados beneficiando qualquer utilizador que esteja a trabalhar no projeto. Por exemplo, podemos associar à primeira etapa a descrição “**automaticamente aplicada pelo editor de consultas**”.

Após concluir estas transformações, grave o ficheiro com a designação “**exercicio_1_10_R.pbix**”.



⁴ Gamble, G. (2018). *Power BI Step-by-Step Part 3: Power Query, Parameters, Templates & Custom Functions*



Split Column by Positions

Specify the positions at which to split the text column.

Positions

0, 5, 7

Advanced options

Exercício 1_11 – Divisão de colunas⁵

Neste exercício, vamos organizar dados, recorrendo à sua distribuição por várias colunas. Para isso vamos utilizar os seguintes ficheiros: “exercício_1_11 (clients).txt”, “exercício_1_11 (countries).txt”, e “exercício_1_11 (sectors).txt”. Seguindo o caminho **home → get data → text/CSV** vamos importar os dados armazenados no ficheiro “exercício_1_11 (clients).txt” fazendo a sua edição no “Power Query Editor”. Vamos promover o cabeçalho “Customer Code” com recurso ao comando “**home ou transform → use first row as headers**”.

Examinando o código do cliente, verificamos que o mesmo é composto por três partes: primeiro, temos um **ID do cliente** que consiste nos primeiros **cinco caracteres**; de seguida temos um **código de país de duas letras**; e, finalmente, temos um **código de setor de negócio de quatro caracteres**. Assim, vamos criar três colunas, uma para o ID do cliente, outra para o código do país e outra para o setor de negócio. Podemos fazer isto clicando com o **botão direito do rato no atributo e escolher “split column”, no menu de contexto, ou em “home ou transform → split column”**.

Começamos então por dividir a coluna num determinado número de caracteres, o mais à esquerda possível, o número de caracteres pretendido é cinco. Este comando cria uma etapa chamada “**split column by position**”, criando também uma segunda etapa chamada “**changed type**”. Como os dados são um ID de cliente, não precisamos de proceder a qualquer alteração, **podemos eliminar esta etapa clicando no X** à esquerda da sua designação.

Precisamos fazer uma nova divisão, tornando os dois caracteres da nova coluna “customer code.2” o código do país. Vamos, então, **clicar com o botão direito no título da coluna** e escolher a opção “**split column**” e de seguida “**by number of characters**”. Novamente o “Power Query Editor” tenta ser útil, reconhecendo dados numéricos e sugerindo a sua conversão para número inteiro. Como não pretendemos fazer isto descartamos esta etapa. Finalmente, renomear as colunas: “**código do cliente**”, “**código do país**” e “**código do setor**”.

Agora, podemos importar os dois ficheiros de texto restantes. Seguimos o caminho “**home → new source → text/CSV**” e importamos o ficheiro “exercício_1_11 (countries).txt”. Porque as colunas deste ficheiro são texto, a promoção dos cabeçalhos não é feita automaticamente, devendo-se seguir o caminho “**home ou transform → use first row as headers**”. Finalmente, vamos importar o ficheiro “exercício_1_11 (sectors).txt”. Desta vez, a promoção dos cabeçalhos é gerada automaticamente. Após concluir estas transformações, grave o ficheiro com a designação “exercício_1_11_R.pbix”.

⁵ Gamble, G. (2018). *Power BI Step-by-Step Part 3: Power Query, Parameters, Templates & Custom Functions*