Descrição do desafio módulo 3 – Processamento de Dados Simplificado com Power BI

1. Criação de uma instância na Azure para MySQL
2. Criar o Banco de dados com base disponível no github
3. Integração do Power BI com MySQL no Azure
4. Verificar problemas na base a fim de realizar a transformação dos dados

Diretrizes para transformação dos dados

1. Verifique os cabeçalhos e tipos de dados
   1. Campo Essn da tabela works\_on modificado para tipo numérico.
2. Modifique os valores monetários para o tipo double preciso
   1. Modificado para valor monetário. O Power BI não possui tipo double preciso.
3. Verifique a existência dos nulos e analise a remoção
   1. Não houve necessidade de exclusão de nulos
4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente
   1. James Borg é o único colaborador que não possui gerente
5. Verifique se há algum departamento sem gerente
   1. Não há departamentos sem gerentes
6. Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas
   1. Não há departamentos sem gerentes.
      1. 333445555 é gerente do Research.
      2. 987654321 é gerente do Administration.
      3. 888665555 é gerente do Headquarters
7. Verifique o número de horas dos projetos
   1. Há uma coluna de horas no tipo decimal.
8. Separar colunas complexas
   1. Coluna Adress da tabela Employee separada em: Number, Street, City e State
9. Mesclar consultas employee e departament para criar uma tabela employee com o nome dos departamentos associados aos colaboradores. A mescla terá como base a tabela employee. Fique atento, essa informação influencia no tipo de junção
   1. Mescla realizada no Power Query adicionando a consulta “azure\_company employee X Departament
10. Neste processo elimine as colunas desnecessárias.
    1. Não há informações sobre o que é necessário ou não. Não há delimitação de escopo. Para atender à solicitação da questão 9, apenas importei o nome do departamento, reduzindo a quantidade de etapas aplicadas no Power BI.
11. Realize a junção dos colaboradores e respectivos nomes dos gerentes . Isso pode ser feito com consulta SQL ou pela mescla de tabelas com Power BI. Caso utilize SQL, especifique no README a query utilizada no processo.
    1. Consultas mescladas no Power Query gerando a “azure\_company Employee X Manager Name”. O nome do gerente foi adicionado ao lado da coluna Super\_ssn
12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores
    1. Realizado na consulta “azure\_company Employee X Manager Name”, tanto com o nome do funcionário, quanto com o nome do gerente.
13. Mescle os nomes de departamentos e localização. Isso fará que cada combinação departamento-local seja único. Isso irá auxiliar na criação do modelo estrela em um módulo futuro.
    1. Realizado na consulta “azure\_company departament”. Existe um erro nesta etapa, pois a tabela consta 3 localidades para o mesmo setor. Foi
14. Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir.
    1. A mescla dos setores funciona como um Inner Join, utilizando o campo Dnumber como chaves. É a única maneira de responder a questão supracitada (13).
15. Agrupe os dados a fim de saber quantos colaboradores existem por gerente
    1. Realizado na consulta “azure\_company Employee X Manager Name count”
16. Elimine as colunas desnecessárias, que não serão usadas no relatório, de cada tabela
    1. Esta é uma etapa com escopo indefinido. Não foi passado o tipo de análise que será executado. Eliminar colunas cria etapas aplicadas no Power Query, caso seja necessário utilizar novamente estes dados, e houver uma ação posterior, algumas métricas e relacionamentos efetuados após a exclusão podem ficar comprometidas. Para eliminar as colunas, é necessária a definição do **ESCOPO DE PROJETO**.