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
2. Modifique os valores monetários para o tipo double preciso
3. Verifique a existência dos nulos e analise a remoção
4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente
5. Verifique se há algum departamento sem gerente
6. Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas
7. Verifique o número de horas dos projetos
8. Separar colunas complexas
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
10. Neste processo elimine as colunas desnecessárias.
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.
12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores
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.
14. Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir.   
    R: Mesclar combina duas tabelas com base em uma coluna comum, criando uma tabela onde os dados de ambas as tabelas são incorporados. No caso citado, seria incorporado informações para criar uma visualização coerente de departamento com a localização e com isso permite que mantenha os registros de departamento e adicionando informações correlacionadas da localização.

Atribuir criaria apenas uma nova tabela como cópia de departamento sem adicionar as novas informações de localização.

Interface gráfica do usuário, Tabela

Descrição gerada automaticamenteTabela

Descrição gerada automaticamente

1. Agrupe os dados a fim de saber quantos colaboradores existem por gerente
2. Elimine as colunas desnecessárias, que não serão usadas no relatório, de cada tabela