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

1. Criação de uma instância na Azure para MySQL ok
2. Criar o Banco de dados com base disponível no github ok
3. Integração do Power BI com MySQL no Azure Não consegui. Não consegui fazer com que aparecesse a janela de conexão no Power BI. Baixei a ligação, mas não funcionou. Baixei a base de dados do MySQL e puxei no Power BI como arquivos em CSV
4. Verificar problemas na base a fim de realizar a transformação dos dados ok

Diretrizes para transformação dos dados

1. Verifique os cabeçalhos e tipos de dados ok. Alterei o nome das colunas Dnumber para padronizar. Estavam nos arquivos employee (Dno), Project (Dnum).
2. Modifique os valores monetários para o tipo double preciso. Ok. Transformei em texto as colunas Ssn, MgrSsn, Essn dos arquivos. Verifiquei que Mgr\_Ssn, Essn e Ssn são os mesmos códigos utilizados para relacionar os empregados.
3. Verifique a existência dos nulos e analise a remoção. Existe valor número em superior do empregado James Borg.
4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente. Tem valor de gerente no Super\_Ssn vazio do empregado James Borg, porque não tem gestor acima dele. Coloquei a matrícula dele mesmo para eliminar o valor vazio. Ele é o único colaborar que possuía valor vazio.
5. Verifique se há algum departamento sem gerente. Não existe nenhum departamento sem gerente.
6. Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas. ok
7. Verifique o número de horas dos projetos. ok
8. Separar colunas complexas. Separei a coluna de Address via Merge Columns, via delimitador “-“. Como um dos nomes havia “-“ entre ele, criou mais uma coluna. Ajustei os nomes e eliminei a coluna criada a mais.
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. ok
10. Neste processo elimine as colunas desnecessárias. ok
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. ok, feito no Power BI
12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores ok – coluna Superior
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. ok
14. Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir. Isto ocorre porque o departamento Research possui mais de uma 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

A screenshot of a computer

Description automatically generated

Mecanismo via groupby

1. Elimine as colunas desnecessárias, que não serão usadas no relatório, de cada tabela

Ok.