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

1. Criação de uma instância na Azure para MySQL

Foi utilizado o ElephantSQL para hospedar o banco de dados. Abaixo estão os dados da instância criada:

|  |  |
| --- | --- |
| Server | silly.db.elephantsql.com (silly-01) |
| Region | amazon-web-services::sa-east-1 |
| Created at | 2023-09-19 23:15 UTC+00:00 |
| URL | postgres://iwewrtff:6jl1TChG7y-z5zJ56BmncCPHBq50ei4w@silly.db.elephantsql.com/iwewrtff |

1. Criar o Banco de dados com base disponível no github

Algumas modificações foram realizadas no código para ser utilizado no ElephantSQL, pois a linguagem difere em alguns quesitos. O código modificado encontra-se no repositório do Github.

1. Integração do Power BI com MySQL no Azure.

Foi realizada a integração mas o ElephantSQL só permite 5 conexões de tabelas no plano gratuito, então, como temos 6 tabelas, a tabela dept\_locations foi inserida manualmente no Power BI.

1. Verificar problemas na base a fim de realizar a transformação dos dados.

Transformações realizadas no Power BI.

Diretrizes para transformação dos dados

1. Verifique os cabeçalhos e tipos de dados OK
2. Modifique os valores monetários para o tipo double preciso OK
3. Verifique a existência dos nulos e analise a remoção OK(Nenhum valor nulo removido)
4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente OK(Havia 1 e foi preenchido manualmente)
5. Verifique se há algum departamento sem gerente OK(Não havia)
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 OK(Separado a coluna “Adress” em “Adress” e “AdressNumber”)
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(Mesclado e mantido na mesma tabela utilizando as colunas “employee.super\_ssn” e “department.mgr\_ssn”. Após mesclar foram remoovidas as duplicidades)
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(Realizado através de mescla de tabelas no Power BI)
12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores OK
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. OK(Neste caso é preciso primeiro realizar a mescla de tabelas para trazer a coluna referente à localização do departamento e depois realizar a mescla de colunas, tornando os valores únicos e associados na tabela. Isso porém causa uma duplicidade de valores na tabela Department.)

Interface gráfica do usuário, Tabela

Descrição gerada automaticamenteTabela

Descrição gerada automaticamente

1. Agrupe os dados a fim de saber quantos colaboradores existem por gerente. OK

|  |  |
| --- | --- |
| **Gerente** | **Contagem** |
| Franklin Wong | 3 |
| James Borg | 2 |
| Jennifer Wallace | 3 |

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