DESAFIO DE PROJETO

Integrando Dados com MySQL Azure e Transformando com Power BI

Ementa - Etapas do Desafio

1. Descrevendo o desafio de projeto
2. Criando uma instância do MySQL na Azure
3. Explorando o Recurso - Instância do MySQL
4. Se conectando ao Banco de Dados com Cloud Shell – corte: manter início-9:06
5. Criando Regra no Firewall na Azure para Acesso ao banco de dados
6. Conectando ao MySQL na Azure utilizando Workbench
7. Integrando Power BI com MySQL na Azure

Descrição do desafio de projeto

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. **Todas as tabelas não mostravam os valores no Power Query, sendo que validei cada atributo importado com os tipos de dados reais através do Workbench**
2. Modifique os valores monetários para o tipo double preciso
   1. **Alterado todos os valroes, aparecendo os dados no Power Query**
3. Verifique a existência dos nulos e analise a remoção
   1. **Encontrado na tabela EMployee o valor Nulo em Super\_ssn.**
4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente
   1. **O único nulo era do próprio gerente com o código 888665555**
5. Verifique se há algum departamento sem gerente
   1. **A tabela departamento tem três descrições com total preenchimento**
6. Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas
   1. **Substitui o valor nulo pelo código 888665555**
7. Verifique o número de horas dos projetos
   1. **Numero de horas = 66 -> fiz uma medida HP = sum (Numero de horas do projeto = SUM('azure\_company project'[Pnumber])**
8. Separar colunas complexas
   1. **Na analise não encontrei colunas complexas para separa, o que não irei utilizar será eliminada nas próximas etapas**
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. **Criado o merge das duas tabelas**
10. Neste processo elimine as colunas desnecessárias.
    1. **Eliminado as colunas para aparecer apenas as associações.**
11. Realize a junção dos colaboradores e respectivos nomes dos gerente . 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. **Ajustei através da mesclagem entre a tabela employee e emprego gestores (criei como referência) para a mesclagem entre SSN e Super SSN**
12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores
    1. **Feito a mescla selecionando as duas colunas e selecionando Merge**
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. **Criado o merge como new e adicionado as duas tabelas relacionando pelo Dnumber**
14. Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir.

Interface gráfica do usuário, Tabela

Descrição gerada automaticamenteTabela

Descrição gerada automaticamente

**Podemos utilizar apenas o *mesclar* neste caso porque os dados desejados (como o nome do gerente) já existem em outra tabela e podem ser recuperados por meio de uma chave de relacionamento (como o Mgr\_ssn). O *atribuir* não é necessário, pois não estamos criando ou calculando um novo valor, apenas relacionando informações já existentes por meio de um join entre tabelas.**

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

**Agrupado através do recurso Group By selecionando a coluna gerentes, na tabela que foi feito o merge entre empregados e colaboradores**

1. Elimine as colunas desnecessárias, que não serão usadas no relatório, de cada tabela
   1. **Eliminado as colunas que acreditei ser desnecessárias neste momento.**