

## **Diretrizes para transformação dos dados**

### 1. Verifique os cabeçalhos e tipos de dados.

Colunas e tipos de dados incorretos:

- Tabela works\_on: Hours - texto;
- Tabela employee: Salary - inteiro;

### 2. Modifique os valores monetários para o tipo double preciso.

Modificação realizada em:

- Tabela employee:
  - Coluna: Salary - inteiro para decimal fixo;

### 3. Verifique a existência dos nulos e analise a remoção.

Verificação e análise realizada em:

- Tabela works\_on:
  - Coluna: Hours - texto; (erro .0), substituição do . por , ; alteração de texto para decimal.

Tabela employee: James - Super\_ssn (null) - verificação do Dno (1) department Dname = Headquarters;

### 4. Os employees com nulos em Super\_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente.

Tabela employee: James - Super\_ssn (null) - verificação do Dno (1) department Dname = Headquarters;

### 5. Verifique se há algum departamento sem gerente.

Dno (1) department Dname = Headquarters;

### 6. Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas.

(Substituição de null por 888665555);

### 7. Verifique o número de horas dos projetos.

Correção e substituição do valor ,0 para 15 hs (projeto com mesmo Pno que outro) Essn pertencente ao gerente James(Headquarters);

### 8. Separar colunas complexas.

Separação da coluna employee.Address em 3; (substituição de logradouro composto, com separação por hífen, para remoção do delimitador, para realizar a separação da coluna corretamente);

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

Mesclar Consultas como novas -> tabela employee como base, junção realizada com base na coluna employee.Dno e departament.Dnumber; Nova tabela salva como employee\_departament;

10. Neste processo elimine as colunas desnecessárias.

Colunas eliminadas: departament.Dnumber; departament.Mgr\_start\_date; departament.Dept\_create\_date;

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.

Realização feita no Power BI; Mesclar Consultas como novas -> tabela employee\_department como base, junção realizada com a tabela employee, tendo como indicadores as colunas employee\_department.Super\_ssn e employee.Ssn, gerando uma nova tabela: employee\_department\_manager;

12. Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores

Colunas de Nome e sobrenome mescladas, (colaborador), para colaboradores e para os gerentes também (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.

Mesclar Consultas como novas -> tabela departamento como base, junção realizada com a tabela dept\_locations, tendo como indicadores departamento.Dnumber e dept\_locations.Dnumber;

14. Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir.

Quando utilizamos o atribuir não ocorre a relação entre os dados, apenas a adição dos dados das duas tabelas, gerando uma terceira, com campos/colunas diferentes, gerando muitos valores nulos e sem relações dos dados.

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

Reutilização da tabela employee\_manager para determinar a quantidade de funcionários por gerente.

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

Colunas desnecessárias eliminadas e realização de relatório para teste e exibição das tarefas realizadas.