

Transformações de dados no Power BI

- 1- Verifique os cabeçalhos e tipos de dados.

Durante o projeto, os cabeçalhos e tipos de dados foram ajustados conforme necessário.

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

Após carregar os dados, as colunas de valores monetários foram convertidas para o tipo double de precisão.

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

Verifiquei a presença de valores nulos e implementei procedimentos para sua remoção.

- 4- Os employees com nulos em Super_ssn podem ser os gerentes. Verifique se há algum colaborador sem gerente.

Identifiquei um colaborador (James E Borg) com valor de gerente nulo. Contudo, o mesmo é um gerente.

- 5- Verifique se há algum departamento sem gerente e

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

Todos os departamentos estão associados a um gerente, conforme visualização no relatório.

- 7- Verifique o número de horas dos projetos

Visualização disponível no relatório.

- 8- Separar colunas complexas

Na tabela employee o endereço foi subdividido em: Nº, Rua, Cidade, UF

Ssn	Bdate	Sex	Salary	Super_ssn	Dno	Nº	Rua	Cidade	UF
123456789	1965-01-09	M	30000	333445555	5	731	Fondren	Houston	TX
333445555	1955-12-08	M	40000	888665555	5	638	Voss	Houston	TX
453453453	1972-07-31	F	25000	333445555	5	5631	Rice	Houston	TX
666884444	1962-09-15	M	38000	333445555	5	975	Fire Oak	Humble	TX
888665555	1937-11-10	M	55000		1	450	Stone	Houston	TX
987654321	1941-06-20	F	43000	888665555	4	291	Berry	Bellaire	TX
987987987	1969-03-29	M	25000	987654321	4	980	Dallas	Houston	TX
999887777	1968-01-19	F	25000	987654321	4	3321	Castle	Spring	TX

- 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 e
- 10- Neste processo elimine as colunas desnecessárias.

Foi desenvolvida uma consulta utilizando SQL que integra as tabelas 'employee' e 'department' no banco de dados. Essa consulta foi posteriormente importada no Power BI para análise.

```
-- agrupamento da coluna employee com departament
select concat(Fname, " ", Minit, " ", Lname) as Name, Dname as Departament from employee as e
inner join departament as d on e.Dno = d.Dnumber;
```

Name	Departament
John B Smith	Research
Franklin T Wong	Research
Joyce A English	Research
Ramesh K Narayan	Research
James E Borg	Headquarters
Jennifer S Wallace	Administration
Ahmad V Jabbar	Administration
Alicia J Zelaya	Administration

- 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.

Para unir os colaboradores e seus respectivos gerentes, uma consulta foi elaborada no Power BI, com o suporte de uma query SQL. Isso possibilitou a combinação das informações dos colaboradores e seus gerentes, incluindo seus nomes.

```
-- agrupamento da coluna employee com manager
select concat(e.Fname, " ", e.Minit, " ", e.Lname) as Employee, concat(m.Fname, " ", m.Minit, " ", m.Lname) as Manager
from employee as e
left join employee as m on e.Super_ssn = m.Ssn;
```

Employee	Manager
John B Smith	Franklin T Wong
Franklin T Wong	James E Borg
Joyce A English	Franklin T Wong
Ramesh K Narayan	Franklin T Wong
James E Borg	
Jennifer S Wallace	James E Borg
Ahmad V Jabbar	Jennifer S Wallace
Alicia J Zelaya	Jennifer S Wallace

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

Na tabela employee a coluna Nome Colaborador é a mescla das colunas Nome e Sobrenome

Ssn	Bdate	Sex	Salary	Super_ssn	Dno	Nº	Rua	Cidade	UF	Nome Colaborador
123456789	1965-01-09	M	30000	333445555	5	731	Fondren	Houston	TX	John B Smith
333445555	1955-12-08	M	40000	888665555	5	638	Voss	Houston	TX	Franklin T Wong
453453453	1972-07-31	F	25000	333445555	5	5631	Rice	Houston	TX	Joyce A English
666884444	1962-09-15	M	38000	333445555	5	975	Fire Oak	Humble	TX	Ramesh K Narayan
888665555	1937-11-10	M	55000		1	450	Stone	Houston	TX	James E Borg
987654321	1941-06-20	F	43000	888665555	4	291	Berry	Bellaire	TX	Jennifer S Wallace
987987987	1969-03-29	M	25000	987654321	4	980	Dallas	Houston	TX	Ahmad V Jabbar
999887777	1968-01-19	F	25000	987654321	4	3321	Castle	Spring	TX	Alicia J Zelaya

- 13- Mescla 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.

As tabelas 'department' e 'dept_locations' foram combinadas para a posterior fusão das colunas 'Dname' e 'Dlocation'. Além disso, qualquer tabela desnecessária foi excluída do processo.

Dname_location	Dnumber	Mgr_ssn	Mgr_start_date	Dept_create_date
Houston Headquarters	1	888665555	1981-06-19	1980-06-19
Stafford Administration	4	987654321	1995-01-01	1994-01-01
Bellaire Research	5	333445555	1988-05-22	1986-05-22
Houston Research	5	333445555	1988-05-22	1986-05-22
Sugarland Research	5	333445555	1988-05-22	1986-05-22

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

O método 'atribuir' não pode ser empregado na operação anterior, uma vez que preenche os espaços em branco com valores nulos, o que não é desejado na junção. Foquei exclusivamente na fusão da coluna 'Dnumber', que é comum a ambas as tabelas.

A _C Dname	1 ₂ Dnumber	1 ₂ Mgr_ssn	Mgr_start_date	Dept_create_date	A _C Dlocation
1 Headquarters	1	888665555	19/06/1981	19/06/1980	null
2 Administration	4	987654321	01/01/1995	01/01/1994	null
3 Research	5	333445555	22/05/1988	22/05/1986	null
4 null	1	null	null	null	Houston
5 null	4	null	null	null	Stafford
6 null	5	null	null	null	Bellaire
7 null	5	null	null	null	Houston
8 null	5	null	null	null	Sugarland

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

Foi realizada uma consulta em SQL no Power BI para agrupar os dados e determinar quantos colaboradores estão sob a supervisão de cada gerente. Visualização disponível no relatório.

```
select concat(m.Fname, " ", m.Minit, " ", m.Lname) as Gerente,  
count(*) as Total_de_colaboradores  
from employee as e  
inner join employee as m on e.Super_ssn = m.Ssn group by Gerente;
```

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

Após o carregamento dos dados, procedi com a remoção das colunas de metadados de todas as tabelas. Além disso, durante a análise, identifiquei e eliminei quaisquer tabelas ou colunas que não eram necessárias para o projeto.