

DIGITAL INNOVATION ONE



Reinaldo Zimmer Wendt

RELATÓRIO DE RESOLUÇÃO DO LAB

Transformação de dados com Power BI

Uruguaiana/RS
2024

1. Adaptação e criação da base de dados *MySQL* utilizando a ferramenta *HeidiSQL*:

1.1:

The screenshot shows the HeidiSQL interface with the following SQL code in the editor:

```
1 CREATE TABLE employee(  
2   Fname varchar(15) not null,  
3   Minit char,  
4   Lname varchar(15) not null,  
5   Ssn char(9) not null,  
6   Bdate date,  
7   Address varchar(30),  
8   Sex char,  
9   Salary decimal(10,2),  
10  Super_ssn char(9),  
11  Dno int not null,  
12  constraint chk_salary_employee check (Salary > 2000.0),  
13  constraint pk_employee primary key (Ssn)  
14 );  
15  
16  
17 alter table employee  
18   add constraint fk_employee  
19   foreign key (Super_ssn) references employee (Ssn)  
20   on delete set null  
21   on update cascade;  
22 alter table employee modify Dno int not null default 1;  
23  
24  
25 create table department(  
26   Dname varchar(15) not null,  
27   Dnumber int not null,  
28   Mgr_ssn char(9) not null,  
29   Mgr_start_date date,  
30   Dept_create_date date,  
31   constraint chk_date_dept check (Dept_create_date < Mgr_start_date),  
32   constraint pk_dept primary key (Dnumber),  
33   constraint unique_name_dept unique (Dname),  
34   foreign key (Mgr_ssn) references employee (Ssn)  
35 );  
36  
37 alter table department  
38   add constraint fk_dept foreign key (Mgr_ssn) references employee (Ssn)  
39   on update cascade;  
40  
41  
42  
43 create table dept_locations(  
44   Dnumber int not null,
```

The bottom status bar shows: r9: c14 (2.3 KiB) | Conectado: 00:00 h | MySQL 8.2.0 | Ativo durante: 00:09 h | Tempo de servidor: 0 | Ocupado.

1.2:

The screenshot shows the HeidiSQL interface with the table structure view for the 'department' table. The table structure is as follows:

Nome	Registros	Tamanho	Criado em	Atualizado em	Motor	Comentário	Tipo
department	3	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela
dependent	7	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela
dept_locations	5	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela
employee	8	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela
project	6	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela
works_on	16	1.0 KiB	2024-03-11 00:42:32	2024-03-11 00:42:32	MyISAM		Tabela

The bottom status bar shows: Conectado: 00:04 h | MySQL 8.2.0 | Ativo durante: 00:11 h | Tempo de servidor: 0 | Ocupado.

2. Realização da conexão com o banco de dados no *Power BI*:

2.1:

Banco de dados MySQL

Servidor

localhost:3306

Banco de Dados

company

Opções avançadas

OK

Cancelar

2.2:

Navegador

Opções de Exibição

localhost:3306: company [6]

company.departament

company.dependent

company.dept_locations

company.employee

company.project

company.works_on

company.employee

Fname	Minit	Lname	Ssn	Bdate	Address
John	B	Smith	123456789	09/01/1965	731-Fondren-Hous
Franklin	T	Wong	333445555	08/12/1955	638-Voss-Houston-
Joyce	A	English	453453453	31/07/1972	5631-Rice-Houston
Ramesh	K	Narayan	666884444	15/09/1962	975-Fire-Oak-Huml
James	E	Borg	888665555	10/11/1937	450-Stone-Houstor
Jennifer	S	Wallace	987654321	20/06/1941	291-Berry-Bellaire-
Ahmad	V	Jabbar	987987987	29/03/1969	980-Dallas-Housto
Alicia	J	Zelaya	999887777	19/01/1968	3321-Castle-Spring

<

>

Selecionar Tabelas Relacionadas

Carregar

Transformar Dados

Cancelar

3. Transformação dos dados

3.1- Verifique os cabeçalhos e tipos de dados.

Os cabeçalhos estão postos corretamente, não há nenhuma linha a mais em nenhuma tabela, para que fosse necessário alterar algum cabeçalho.

Tipos modificados:

- Departament.Dnumber: texto → inteiro
- Project.Pnumber: texto → inteiro;
- Dept_locations.Dnumber: texto → inteiro;
- Employee.Dno: texto → inteiro.

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

Tipo modificado:

- Employee.Salary: texto → decimal;

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

Não houve necessidade de remoção de nenhum registro, não há registros duplicados ou nulos (que não possuam justificativa, como será especificado no item a seguir).

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

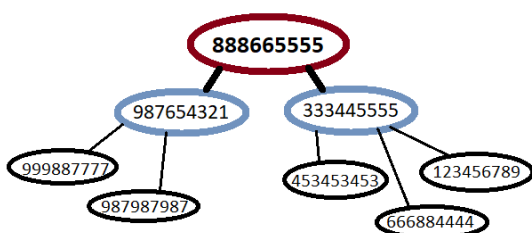
Sim, no entanto, faz sentido hierárquico que o mesmo seja nulo.

3.4.1:

= Table.ReorderColumns("#Linhas Classificadas",{"Fname", "Minit", "Lname", "Bdate", "Address", "Sex", "Salary", "Ssn", "Super_ssn", "Dno"})										
	A ^B Fname	A ^B Minit	A ^B Lname	A ^B Bdate	A ^B Address	A ^B Sex	1.2 Salary	A ^B Ssn	A ^B Super_ssn	
1	James	E	Borg	10/11/1937	450-Stone-Houston-TX	M	55000	888665555		null
2	Joyce	A	English	31/07/1972	5631-Rice-Houston-TX	F	25000	453453453	333445555	
3	Ramesh	K	Narayan	15/09/1962	975-Fire-Oak-Humble-TX	M	38000	666884444	333445555	
4	John	B	Smith	09/01/1965	731-Fondren-Houston-TX	M	30000	123456789	333445555	
5	Jennifer	S	Wallace	20/06/1941	291-Berry-Bellaire-TX	F	43000	987654321	888665555	
6	Franklin	T	Wong	08/12/1955	638-Voss-Houston-TX	M	40000	333445555	888665555	
7	Alicia	J	Zelaya	19/01/1968	3321-Castle-Spring-TX	F	25000	999887777	987654321	
8	Ahmad	V	Jabbar	29/03/1969	980-Dallas-Houston-TX	M	25000	987987987	987654321	

Pois tal registro atua como um “super-gerente” como pode ser visualizado graficamente na imagem a seguir.

3.4.2:



3.5- Verifique se há algum departamento sem gerente.

Não, como pode ser visto na imagem a seguir:

3.5.1:

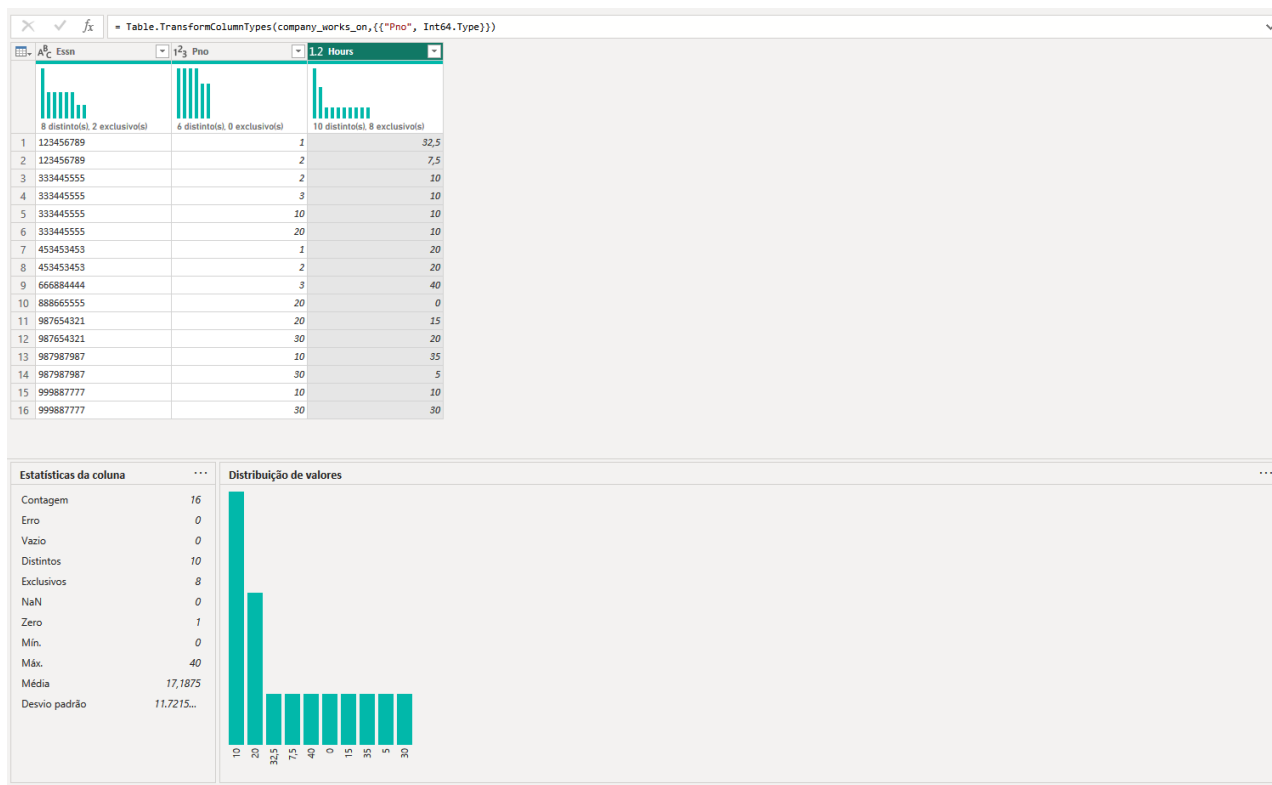
= Table.TransformColumnTypes(company_departament,{{"Dnumber", Int64.Type}})					
	A ^B _C Dname	1 ² ₃ Dnumber	A ^B _C Mgr_ssn	Mgr_start_date	Dept_create_date
1	Headquarters	1	888665555	19/06/1981	19/06/1980
2	Administration	4	987654321	01/01/1995	01/01/1994
3	Research	5	333445555	22/05/1988	22/05/1986

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

3.7- Verifique o número de horas dos projetos.

O número de horas e sua distribuição de acordo com projeto podem ser vistos na Power Query da imagem abaixo.

3.7.1:



3.8- Separar colunas complexas.

A única coluna complexa que fez sentido ao meu ver separar, foi Address da tabela Employee. Pois a mesma consta com o endereço completo, separado por hifens, que podem ser separados em Address.number (número de endereço), Address.street (rua), Address.city (cidade) e Address.state (estado). Como pode ser visualizado nas imagens abaixo.

3.8.1:

×

Dividir Coluna por Delimitador

Especifique o delimitador usado para dividir a coluna de texto.

Selecione ou insira o delimitador

--Personalizado--

-

Dividir em

☐ Delimitador da extremidade esquerda

☐ Delimitador da extremidade direita

☒ Cada ocorrência do delimitador

▶ Opções avançadas

Caractere de Aspas

"

☐ Dividir usando caracteres especiais

Inserir caractere especial

OK

Cancelar

3.8.2:

×

✓

f_x

= Table.RenameColumns("#Tipo Alterado1",{"Address.1", "Address.number"}, {"Address.2", "Address.street"}, {"Address.3", "Address.city"}, {"Address.4", "Address.state"})

	A ₀ Lname	Bdate	A ₂ Address.number	A ₀ Address.street	A ₀ Address.city	A ₀ Address.state	A ₀ Sex	1.2 Salary	A ₀ Ssn
1	Borg	10/11/1937	450	Stone	Houston	TX	M	55000	88865555
2	English	31/07/1972	5631	Rice	Houston	TX	F	25000	453453453
3	Narayan	15/09/1962	975	Fire Oak	Humble	TX	M	38000	666884444
4	Smith	09/01/1965	731	Fondren	Houston	TX	M	30000	123456789
5	Wallace	20/06/1941	291	Berry	Bellaire	TX	F	43000	987654321
6	Wong	08/12/1955	638	Voss	Houston	TX	M	40000	333445555
7	Zelaya	19/01/1968	3321	Castle	Spring	TX	F	25000	999887777
8	Jabbar	29/03/1969	980	Dallas	Houston	TX	M	25000	987987987

3.9- 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.

3.9.1:

Mesclar

Selecione a tabela e as colunas correspondentes para criar uma tabela mesclada.

company employee

ess.number	Address.street	Address.city	Address.state	Sex	Salary	Ssn	Super_ssn	Dno
450	Stone	Houston	TX	M	55000	888665555	null	1
5631	Rice	Houston	TX	F	25000	453453453	333445555	5
975	Fire Oak	Humble	TX	M	38000	666884444	333445555	5
731	Fondren	Houston	TX	M	30000	123456789	333445555	5

company departament

Dname	Dnumber	Mgr_ssn	Mgr_start_date	Dept_create_date
Headquarters	1	888665555	19/06/1981	19/06/1980
Administration	4	987654321	01/01/1995	01/01/1994
Research	5	333445555	22/05/1988	22/05/1986

Tipo de Junção

Externa esquerda (todas a partir da primeira, correspo...

☐ Usar a correspondência difusa para executar a mesclagem

Opções de correspondência difusa

☒ A seleção corresponde a 8 de 8 linhas da primeira tabela.

OK

Cancelar

3.9.2:

Consultas [6]															= Table.NestedJoin("#Colunas Renomeadas", ("Dno"), "#company departament", ("Dnumber"), "company departament", JoinKind.LeftOuter)														
company departa...	1	50	Stone	Houston	TX	M	55000	888665555		888665555		null		1	Table														
company dept_loc...	2	131	Rice	Houston	TX	F	25000	453453453		333445555				5	Table														
company dependent	3	175	Fire Oak	Humble	TX	M	38000	666884444		333445555				5	Table														
company employee	4	131	Fondren	Houston	TX	M	30000	123456789		333445555				5	Table														
company project	5	191	Berry	Bellaire	TX	F	43000	987654321		888665555				4	Table														
company works_on	6	138	Voss	Houston	TX	M	40000	333445555		888665555				5	Table														
	7	121	Castle	Spring	TX	F	25000	999887777		987654321				4	Table														
	8	180	Dallas	Houston	TX	M	25000	987987987		987654321				4	Table														

3.10- Neste processo elimine as colunas desnecessárias.

3.10.1:

Table.NestedJoin("#Colunas Removidas", ("Dno"), "#company departament", ("Dnumber"), "company departament", JoinKind.LeftOuter)

Preparar Colunas para Expandir

Expandir ☐ Agregar ☐

☒ (Selecionar Todas as Colunas)

☐ Dname

☐ Dnumber

☐ Mgr_ssn

☐ Mgr_start_date

☐ Dept_create_date

☒ Use o nome da coluna original como prefixo

OK Cancelar

3.10.2:

Table.ExpandTableColumn("#Consultas Mescladas1", "company departament", ("Dname"), {"company departament.Dname"})

Address.street	Address.city	Address.state	Sex	Salary	Ssn	Super_ssn	Dno	company departament.Dname
one	Houston	TX	M	55000	888665555			1 Headquarters
ce	Houston	TX	F	25000	453453453	333445555		5 Research
re Oak	Humble	TX	M	38000	666884444	333445555		5 Research
ndren	Houston	TX	M	30000	123456789	333445555		5 Research
rry	Bellaire	TX	F	43000	987654321	888665555		4 Administration
ss	Houston	TX	M	40000	333445555	888665555		5 Research
stle	Spring	TX	F	25000	999887777	987654321		4 Administration
allas	Houston	TX	M	25000	987987987	987654321		4 Administration

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

3.11.1:

Table.RenameColumns("#company departament Expandido1", {"company departament.Dname", "departament.Dname"})

Mesclar Colunas

Escolha como mesclar as colunas selecionadas.

Separador

Espaço

Novo nome da coluna (opcional)

Mesclado

OK Cancelar

3.11.2:

Table.RenameColumns("#Colunas Mescladas", {"Mesclado", "Name"})

Name	Bdate	Address.number	Address.street	Address.city	Address.state	Sex	Salary	Ssn	company.departament.Dname
James E Borg	10/11/1937	450	Stone	Houston	TX	M	55000	888665555	
Joyce A English	31/07/1972	5631	Rice	Houston	TX	F	25000	453453453	3334
Ramesh K Narayan	15/09/1962	975	Fire Oak	Humble	TX	M	38000	666884444	3334
John B Smith	09/01/1965	731	Fondren	Houston	TX	M	30000	123456789	3334
Jennifer S Wallace	20/06/1941	291	Berry	Bellaire	TX	F	43000	987654321	8886
Franklin T Wong	08/12/1955	638	Voss	Houston	TX	M	40000	333445555	8886
Alicia J Zelaya	19/01/1968	3321	Castle	Spring	TX	F	25000	999887777	9876
Ahmad V Jabbar	29/03/1969	980	Dallas	Houston	TX	M	25000	987987987	9876

3.12- Realize a junção dos colaboradores e respectivos nomes dos gerentes.

3.12.1:

Mesclar

Selecione a tabela e as colunas correspondentes para criar uma tabela mesclada.

company employee

treet	Address.city	Address.state	Sex	Salary	Ssn	Super_ssn	Dno	departament.Dname
	Houston	TX	M	55000	888665555	null	1	Headquarters
	Houston	TX	F	25000	453453453	333445555	5	Research
	Humble	TX	M	38000	666884444	333445555	5	Research
	Houston	TX	M	30000	123456789	333445555	5	Research

company employee (Atual)

treet	Address.city	Address.state	Sex	Salary	Ssn	Super_ssn	Dno	departament.Dname
	Houston	TX	M	55000	888665555	null	1	Headquarters
	Houston	TX	F	25000	453453453	333445555	5	Research
	Humble	TX	M	38000	666884444	333445555	5	Research
	Houston	TX	M	30000	123456789	333445555	5	Research

Tipo de Junção

Externa esquerda (todas a partir da primeira, corresp...

☐ Usar a correspondência difusa para executar a mesclagem

Opções de correspondência difusa

☒ A seleção corresponde a 7 de 8 linhas da primeira tabela.

OK Cancelar

3.12.2:

Colunas Renomeadas3

Pesquisar Colunas para Expandir

☒ Expandir ☐ Agregar

☒ (Selecionar Todas as Colunas)

☒ Name

☐ Bdate

☐ Address.number

☐ Address.street

☐ Address.city

☐ Address.state

☐ Sex

☐ Salary

☐ Ssn

☐ Super_ssn

☐ Dno

☐ departament.Dname

☒ Use o nome da coluna original como prefixo

OK Cancelar

3.12.3:

	Address.city	Address.state	Sex	Salary	Ssn	Super_ssn	Super_name	Dno	departament.Dname
1	Houston	TX	M	55000	888665555	null	null	1	Headquarters
2	Humble	TX	M	38000	666884444	333445555	Franklin T Wong	5	Research
3	Houston	TX	M	30000	123456789	333445555	Franklin T Wong	5	Research
4	Houston	TX	F	25000	453453453	333445555	Franklin T Wong	5	Research
5	Houston	TX	M	40000	333445555	888665555	James E Borg	5	Research
6	Bellaire	TX	F	43000	987654321	888665555	James E Borg	4	Administration
7	Spring	TX	F	25000	999887777	987654321	Jennifer S Wallace	4	Administration
8	Houston	TX	M	25000	987987987	987654321	Jennifer S Wallace	4	Administration

3.13- 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.

3.13.1:

Mesclar

Selecione a tabela e as colunas correspondentes para criar uma tabela mesclada.

company department

Dname	Dnumber	Mgr_ssn	Mgr_start_date	Dept_create_date
Headquarters	1	888665555	19/06/1981	19/06/1980
Administration	4	987654321	01/01/1995	01/01/1994
Research	5	333445555	22/05/1988	22/05/1986

company dept_locations

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Houston
5	Sugarland

Tipo de Junção

Externa direita (todas a partir da segunda, correspon...

☐ Usar a correspondência difusa para executar a mesclagem

Opções de correspondência difusa

☒ A seleção corresponde a 5 de 5 linhas da segunda tabela.

OK

Cancelar

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

Quando se mescla os nomes de departamentos e localizações, cria-se uma relação entre as duas tabelas com base em uma coluna comum, como um identificador único para cada departamento e localização. Isso garante que cada combinação departamento-local seja única, pois estará associada a um único registro na tabela resultante da mesclagem. Além de que, mesclar as tabelas de departamentos e localizações permite manter os relacionamentos entre elas. Enquanto a operação de atribuir valores geralmente é utilizada para adicionar uma coluna com valores estáticos a uma tabela existente. No entanto, se tentássemos apenas atribuir valores para os nomes de departamentos e localizações, perderíamos informações importantes, como os relacionamentos entre os departamentos e suas respectivas localizações.

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

3.15.1:

Agrupar por

Especifique a coluna a ser agrupada e a saída desejada.

☒ Básico ☐ Avançadas

Super_name

Nome da nova coluna: Contagem

Operação: Contar Linhas

Coluna: -

OK Cancelar

3.15.2:

	Super_name	Contagem
1	Franklin T Wong	3
2	James E Borg	2
3	Jennifer S Wallace	2
4	null	1

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

- Departament:
 - Dept_create_date;
 - Mgr_start_date;
 - Mgr_ssn;
- Dependent
 - Sex;
 - Bdate;
 - Relationship;
- Employee:
 - Bdate.