

Desafio: Limpeza e Transformação de Dados com Power BI

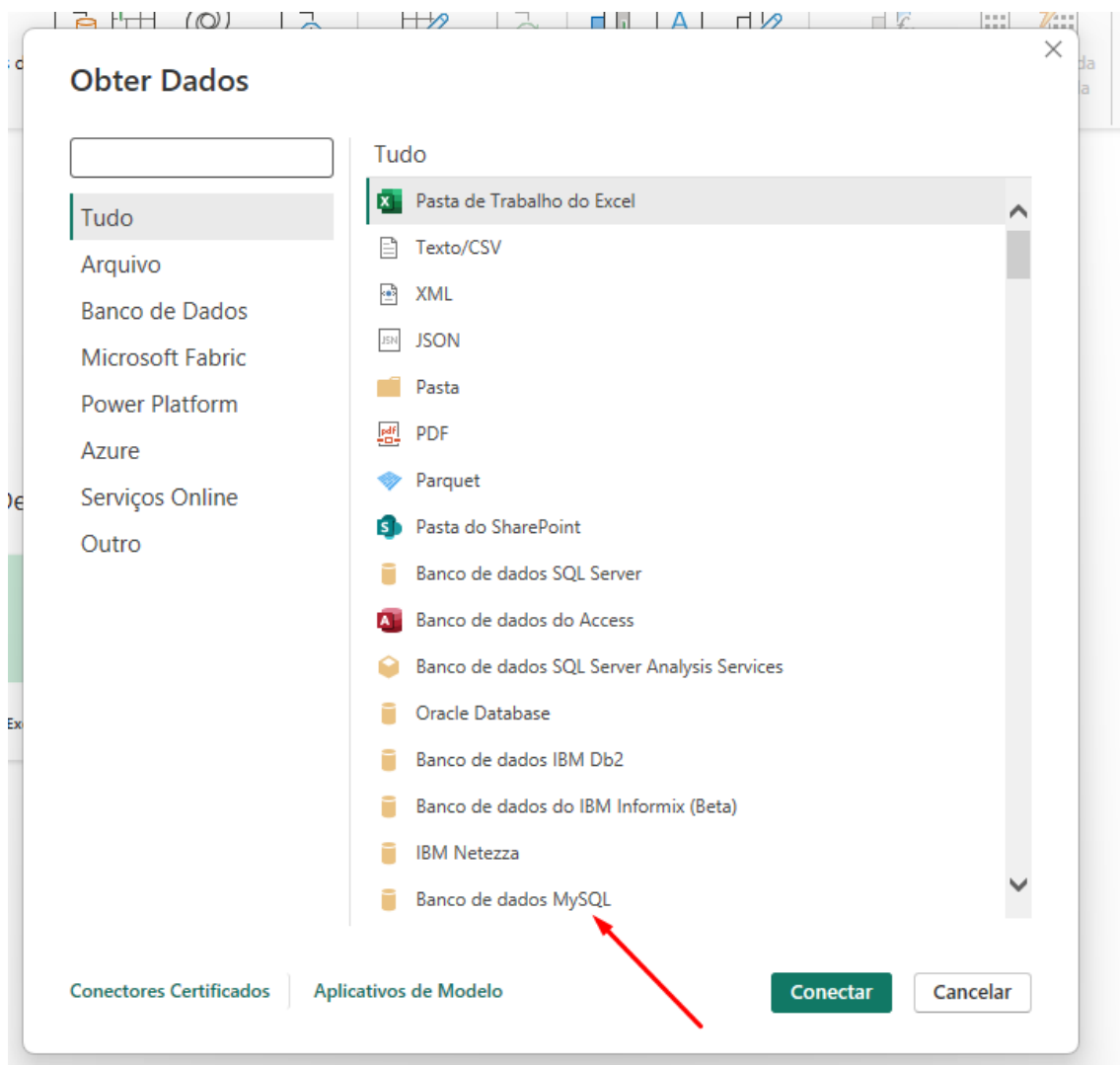
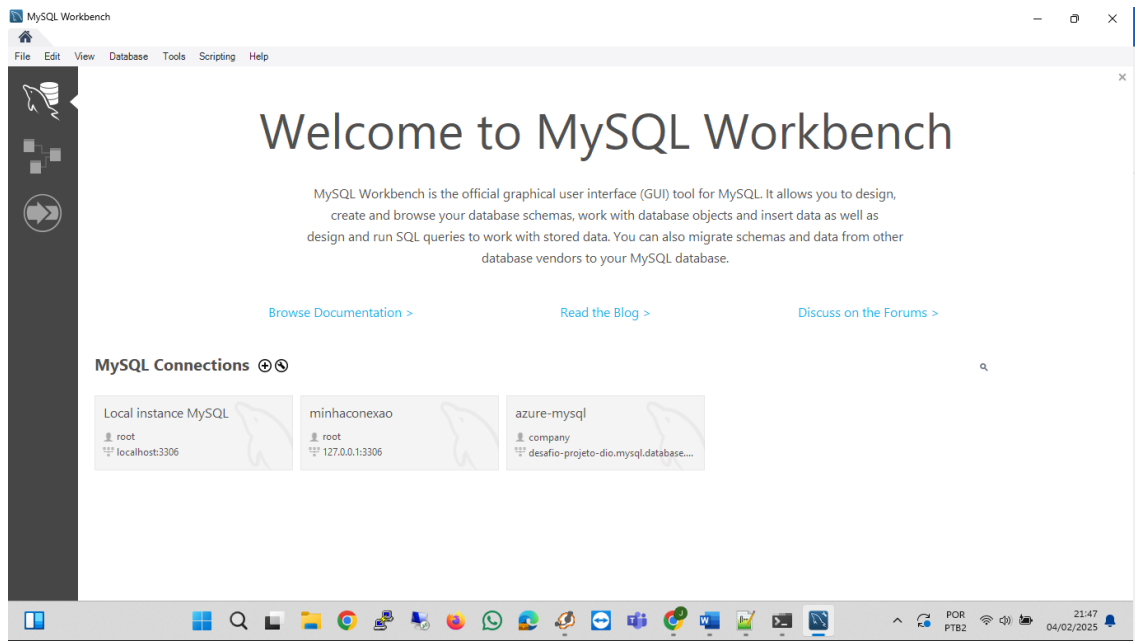
The image shows two overlapping screenshots. The top screenshot is the Microsoft Azure portal, and the bottom screenshot is the SQL Server Enterprise Manager 'Manage Server Connections' dialog box.

Microsoft Azure Portal:

- Page title: Página inicial - Microsoft Azure
- URL: portal.azure.com/#home
- Search bar: Pesquisar recursos, serviços e documentos (G+)
- Services section: Criar um recurso, Bancos de dados SQL, Centro de Início Rápido, Azure AI services, Serviços do Kubernetes, Máquinas virtuais, Serviços de Aplicativos, Contas de armazenamento, Azure Cosmos DB, Mais serviços.
- Recursos section: Recentemente, Favorito. A table with columns Nome, Tipo, and Última visualização. A message states: 'Nenhum recurso foi visualizado recentemente' with a button 'Exibir todos os recursos'.
- Navegar section: Layout, Referências, Correspondências, Revisão, Exportar, Ajuda, Atualizar.

SQL Server Enterprise Manager - Manage Server Connections:

- Left pane: MySQL Connections, Local instance MySQL, minhaconexao, azure-mysql.
- Connection Name: azure-mysql
- Connection Method: Standard (TCP/IP) (Method to use to connect to the RDBMS)
- Parameters tab selected.
- Fields: Hostname (desafio-projeto-dio.mysql.da), Port (3306), Username (company), Password (Store in Vault ... / Clear), Default Schema.
- Buttons: New, Delete, Duplicate, Move Up, Move Down, Test Connection, Close.



egador

desafio-projeto-dio.mysql.database.azure.com....

Departamento

dependent

dept_locations

employee

project

works_on

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-Houstor
Alicia	J	Zelaya	999887777	19/01/1968	3321-Castle-Spring

Carregar

Transformar Dados

Cancelar

Foram ajustados os cabeçalhos os tipos de dados

Table.RenameColumns("#Tipo Alterado",{"Fname", "First Name"}, {"Lname", "Last Name"}, {"Bdate", "Date"}, {"Dno",

First Name

Minit

Last Name

Ssn

Date

Address

1	John	B	Smith	123456789	09/01/1965	731-Fondren-Houston-TX
2	Franklin	T	Wong	333445555	08/12/1955	638-Voss-Houston-TX
3	Joyce	A	English	453453453	31/07/1972	5631-Rice-Houston-TX
4	Ramesh	K	Narayan	666884444	15/09/1962	975-Fire-Oak-Humble-TX
5	James	E	Borg	888665555	10/11/1937	450-Stone-Houston-TX
6	Jennifer	S	Wallace	987654321	20/06/1941	291-Berry-Bellaire-TX
7	Ahmad	V	Jabbar	987987987	29/03/1969	980-Dallas-Houston-TX
8	Alicia	J	Zelaya	999887777	19/01/1968	3321-Castle-Spring-TX

Table.RenameColumns("#Tipo Alterado",{"Dname", "Department name"}, {"Dnumber", "Department number"})

Department name

Department number

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

	1 ² ₃ Department number	A ^B _C Department location
1	1	Houston
2	4	Stafford
3	5	Bellaire
4	5	Houston
5	5	Sugarland

	1 ² ₃ Essn	A ^B _C Dependent_name	A ^B _C Sex	Date	A ^B _C Relationship
1	123456789	Alice	F	30/12/1988	Daughter
2	123456789	Elizabeth	F	05/05/1967	Spouse
3	123456789	Michael	M	04/01/1988	Son
4	333445555	Alice	F	05/04/1986	Daughter
5	333445555	Joy	F	03/05/1958	Spouse
6	333445555	Theodore	M	25/10/1983	Son
7	987654321	Abner	M	28/02/1942	Spouse

= Table.RenameColumns("#Tipo Alterado",{{"Pname", "Projct name"}, {"Pnumber", "Projct number"}, {"Plocation", "Projct location"}, {"Pdepartment", "Projct department"}})

	A ^B _C Projct name	1 ² ₃ Projct number	A ^B _C Projct location	1 ² ₃ Department num
1	ProductX	1	Bellaire	5
2	ProductY	2	Sugarland	5
3	ProductZ	3	Houston	5
4	Computerization	10	Stafford	4
5	Reorganization	20	Houston	1
6	Newbenefits	30	Stafford	4

= Table.RenameColumns("#Tipo Alterado",{{"Pno", "Projct no"}})

	1 ² ₃ Essn	1 ² ₃ Projct no	1 ² ₃ Hours
1	123456789	1	325
2	123456789	2	75
3	333445555	2	100
4	333445555	3	100
5	333445555	10	100
6	333445555	20	100
7	453453453	1	200
8	453453453	2	200
9	666884444	3	400
10	888665555	20	0
11	987654321	20	150
12	987654321	30	200
13	987987987	10	350
14	987987987	30	50
15	999887777	10	100
16	999887777	30	300

Foram modificados os valores monetários para o tipo double preciso

= Table.TransformColumnTypes("#Cabeçalhos Promovidos",{{"Fname", type text}, {"Minit", type text}, {"Lname", type text},						
Bdate	Address	Sex	Salary	Super_ssn	Dno	
09/01/1965	731-Fondren-Houston-TX	M	300.000,00	333445555		5
08/12/1955	638-Voss-Houston-TX	M	400.000,00	888665555		5
31/07/1972	5631-Rice-Houston-TX	F	250.000,00	333445555		5
15/09/1962	975-Fire-Oak-Humble-TX	M	380.000,00	333445555		5
10/11/1937	450-Stone-Houston-TX	M	550.000,00	123456789		1
20/06/1941	291-Berry-Bellaire-TX	F	430.000,00	888665555		4
29/03/1969	980-Dallas-Houston-TX	M	250.000,00	987654321		4
19/01/1968	3321-Castle-Spring-TX	F	250.000,00	987654321		4

Sim, há alguns colaboradores que estão sem gerentes

= Table.ExpandTableColumn(Fonte, "dependent", {"Essn", "Dependent_name", "Sex", "Date", "Relationship"}, {"dependent.Essn",						
ment number	dependent.Essn	dependent.Dependent_name	dependent.Sex	dependent.Date	dependent.Relationship	
1	5	123456789 Alice	F	30/12/1988	Daughter	
2	5	123456789 Elizabeth	F	05/05/1967	Spouse	
3	5	123456789 Michael	M	04/01/1988	Son	
4	5	333445555 Alice	F	05/04/1986	Daughter	
5	5	333445555 Joy	F	03/05/1958	Spouse	
6	5	333445555 Theodore	M	25/10/1983	Son	
7	5	null	null	null	null	null
8	5	null	null	null	null	null
9	1	null	null	null	null	null
10	4	987654321 Abner	M	28/02/1942	Spouse	
11	4	null	null	null	null	null
12	4	null	null	null	null	null

Foram mesclados consultas Employee e Departament para criar uma nova tabela

= Table.ExpandTableColumn(Fonte, "Departamento", {"Department name", "Department number", "Mgr_ssn", "Mgr_start_date",						
First Name	Minit	Last Name	Ssn	Date	Address	
John	B	Smith	123456789	09/01/1965	731-Fondren-Houston-TX	
Franklin	T	Wong	333445555	08/12/1955	638-Voss-Houston-TX	
Joyce	A	English	453453453	31/07/1972	5631-Rice-Houston-TX	
Ramesh	K	Narayan	666884444	15/09/1962	975-Fire-Oak-Humble-TX	
James	E	Borg	888665555	10/11/1937	450-Stone-Houston-TX	
Jennifer	S	Wallace	987654321	20/06/1941	291-Berry-Bellaire-TX	
Ahmad	V	Jabbar	987987987	29/03/1969	980-Dallas-Houston-TX	
Alicia	J	Zelaya	999887777	19/01/1968	3321-Castle-Spring-TX	

= Table.ExpandTableColumn(Fonte, "Departamento", {"Department name", "Department number", "Mgr_ssn", "Mgr_start_date",						
Sex	Salary	Super_ssn	Department number	Departamento.Department na...	Departament	
M	300.000,00	333445555	5	Research		
M	400.000,00	888665555	5	Research		
F	250.000,00	333445555	5	Research		
M	380.000,00	333445555	5	Research		
M	550.000,00	123456789	1	Headquarters		
F	430.000,00	888665555	4	Administration		
M	250.000,00	987654321	4	Administration		
F	250.000,00	987654321	4	Administration		

.ExpandTableColumn(Fonte, "Departamento", {"Department name", "Department number", "Mgr_ssn", "Mgr_start_date",				
Departamento.Department numb...	Departamento.Mgr_ssn	Departamento.Mgr_start_date	Departamento.Dept_create_date	
5	333445555	22/05/1988	22/05/1986	
5	333445555	22/05/1988	22/05/1986	
5	333445555	22/05/1988	22/05/1986	
5	333445555	22/05/1988	22/05/1986	
1	888665555	19/06/1981	19/06/1980	
4	987654321	01/01/1995	01/01/1994	
4	987654321	01/01/1995	01/01/1994	
4	987654321	01/01/1995	01/01/1994	

Foi realizada a junção dos colaboradores e os respectivos nomes dos gerentes.

	A ^B _C First Name	A ^B _C Last Lname	A ^B _C dependent.Dependent_name
1	John	Smith	Alice
2	John	Smith	Elizabeth
3	John	Smith	Michael
4	Franklin	Wong	Alice
5	Franklin	Wong	Joy
6	Franklin	Wong	Theodore
7	Jennifer	Wallace	Abner

Foram muclados os nomes de Department e Localização

= Table.ExpandTableColumn("#dept_locations.1 Expandido", "dept_locations", {"Department number", "Departmer						
	A ^B _C Department name	1 ² ₃ Department number	1 ² ₃ 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	
4	Research	5	333445555	22/05/1988	22/05/1986	
5	Research	5	333445555	22/05/1988	22/05/1986	
6	Research	5	333445555	22/05/1988	22/05/1986	
7	Research	5	333445555	22/05/1988	22/05/1986	
8	Research	5	333445555	22/05/1988	22/05/1986	
9	Research	5	333445555	22/05/1988	22/05/1986	
10	Research	5	333445555	22/05/1988	22/05/1986	
11	Research	5	333445555	22/05/1988	22/05/1986	

= Table.ExpandTableColumn("#dept_locations.1 Expandido", "dept_locations", {"Department number", "Department location"},					
1 ² ₃ dept_locations.Department number	A ^B _C dept_locations.Department location	1 ² ₃ dept_locations.1.Department number	A ^B _C dept_locations.1.Department location		
1	Houston	1	Houston		
4	Stafford	4	Stafford		
5	Bellaire	5	Bellaire		
5	Houston	5	Bellaire		
5	Sugarland	5	Bellaire		
5	Bellaire	5	Houston		
5	Houston	5	Houston		
5	Sugarland	5	Houston		
5	Bellaire	5	Sugarland		
5	Houston	5	Sugarland		
5	Sugarland	5	Sugarland		

Foram agrupados os dados a fim de saber quantos colaboradores existem por gerente.

	A ^B _C Dependent_name	1 ² ₃ Colaboradores
1	Alice	2
2	Elizabeth	1
3	Michael	1
4	Joy	1
5	Theodore	1
6	Abner	1