

1) Na pasta **C:\treinamento\designer-tools\data-integration\simple-jndi**, confira no arquivo **jdbc.properties** se a configuração da conexão com o banco MySQL está correta:

```
dwsucos/type=javax.sql.DataSource  
dwsucos/driver=com.mysql.jdbc.Driver  
dwsucos/url=jdbc:mysql://localhost:3306/dwsucos  
dwsucos/user=root  
dwsucos/password=root
```

[COPIAR CÓDIGO](#)

Observação: Se você está configurando a conexão **no horário brasileiro de verão**, use:

```
dwsucos/type=javax.sql.DataSource  
dwsucos/driver=com.mysql.jdbc.Driver  
dwsucos/url=jdbc:mysql://localhost:3306/dwsucos?useTimezone=true&serverTimezone=America/Sao_Paulo  
dwsucos/user=root  
dwsucos/password=root
```

[COPIAR CÓDIGO](#)

2) Abra o **Spoon**, executando o **Spoon.bat**, no diretório **C:\treinamento\designer-tools\data-integration**.

3) Crie uma nova transformação, clicando com botão direito do mouse sobre a pasta **Transformations** e clicando na opção **New**.

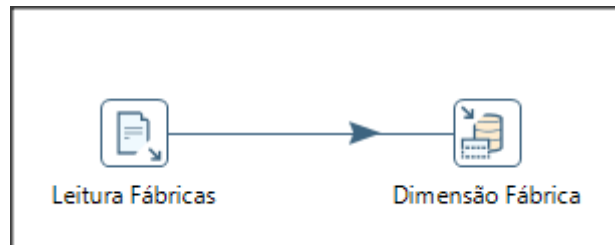
4) Na aba **View**, expanda a transformação, clique com o botão direito do mouse sobre **Database connections** e clique em **New**, para criar uma nova conexão, com as seguintes propriedades:

- **Connection name:** Datawarehouse Sucos
- **Connection type:** MySQL
- **Access:** JNDI
- **JNDI Name:** \${banco}

5) Verifique se antes o MySQL está no ar, teste a conexão e depois clique no botão **OK**.

6) Salve a transformação com o nome **DimensaoFabrica**.

7) Com os componentes **CSV file input**, que terá nome **Leitura Fábricas**, e **Insert / update**, que terá nome **Dimensão Fábrica**, construa a seguinte transformação:



8) Dê um duplo clique em **Leitura Fábricas** e preencha as suas propriedades:

CSV file input

Step name:

Filename:

Delimiter:

Enclosure:

NIO buffer size:

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional)

Running in parallel? ☐

New line possible in fields? ☐

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	CodigoFabrica	String	#	50	0	R\$,	.	both
2	DescritorFabrica	String		250		R\$,	.	both

9) Faça o mesmo com **Dimensão Fábrica**, preenchendo as seguintes propriedades:

Insert / update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates: ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	Cod_Fabrica	=	CodigoFabrica	

Update fields:

#	Table field	Stream field	Update
1	Cod_Fabrica	CodigoFabrica	N
2	Desc_Fabrica	DescritorFabrica	Y

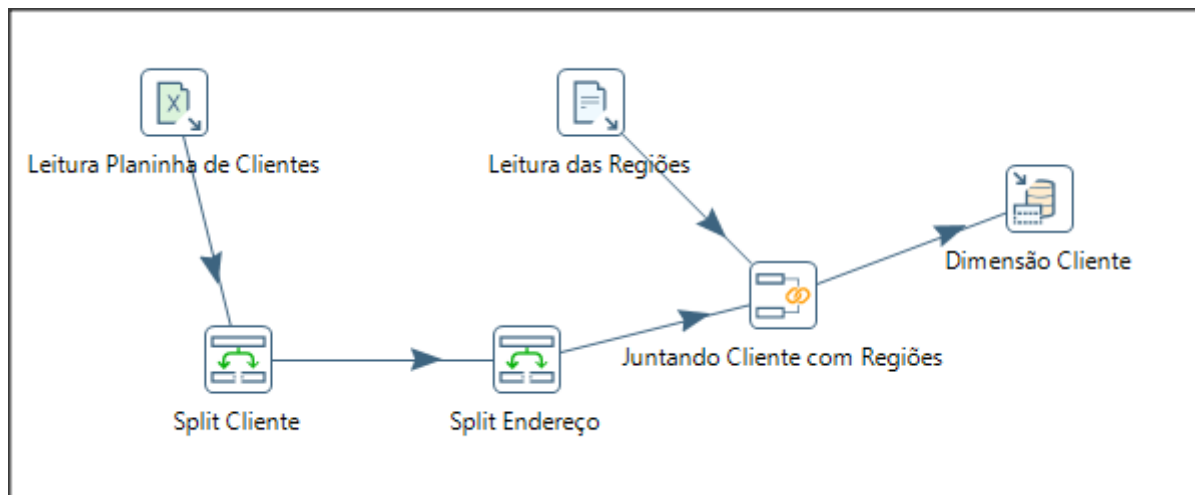
10) Salve a transformação e execute-a. Confira se os dados da **Dim_Fabrica** foram carregados no *Data Warehouse*.

11) Crie uma nova transformação, chamada **DimensaoCliente**.

12) Ela terá os seguintes componentes:

- **Microsoft Excel input**, que terá nome **Leitura Planilha de Clientes**
- **Split fields**, que terá nome **Split Cliente**
- **Split fields**, que terá nome **Split Endereço**
- **CSV file input**, que terá nome **Leitura das Regiões**
- **Join rows**, que terá nome **Juntando Cliente com Regiões**
- **Insert / update**, que terá nome **Dimensão Cliente**

A transformação ficará com a seguir:



13) Em **Leitura Planilha de Clientes**, na aba **Files**, faça:

Microsoft Excel input

Step name: **Leitura Planilha de Clientes**

Files | Sheets | Content | Error Handling | Fields | Additional output fields

Spread sheet type (engine): Excel 2007 XLSX (Apache POI)

File or directory: Add Browse...

Regular Expression:

Exclude Regular Expression:

Selected files:

#	File/Directory	Wildcard (RegExp)
1	\${diretorio}CADASTRO DE CLIENTES.xlsx	

Delete Edit

Accept filenames from previous steps

Accept filenames from ☐

Step to read filenames:

Field in the input to use as:

Show filename(s)...

Help OK Preview rows Cancel

14) Em **Leitura Planilha de Clientes**, na aba **Sheets**:

[illegible]

15) Em **Leitura Planilha de Clientes**, na aba **Content**, faça:

Microsoft Excel input

Step name **Leitura Planilha de Clientes**

Files Sheets **Content** Error Handling Fields Additional output fields

Header ☒

No empty rows ☒

Stop on empty row ☐

Limit 0

Encoding

Result filenames

Add filenames to result ☒

Help OK Preview rows Cancel

16) Em **Leitura Planilha de Clientes**, na aba **Fields**, faça:

Microsoft Excel input

Step name:

Files | Sheets | Content | Error Handling | Fields | Additional output fields

#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency
1	ID_CLIENTE	String	50	-1	both	N	#	
2	CLIENTE	String	250	-1	both	N		
3	ENDERECO	String	250	-1	both	N		
4	VENDEDOR	String	50	-1	both	N	#	
5	NOME VENDEDOR	String	250	-1	both	N		
6	SEGMENTO	String	50	-1	both	N		

< >

OK Preview rows Cancel

Help

17) Em **Split Cliente**, faça:

Split fields

Step name:

Field to split:

Delimiter:

Enclosure:

Fields

#	New field	ID	Remove ID?	Type	Length	P	F	I	C	N	D	Trim type
1	NOME CLIENTE		N	String	250							both
2	ESTADO CLIENTE		N	String	50							both

Help OK Cancel

18) Em **Split Endereço**, faça:

Split fields

Step name: Split Endereço

Field to split: ENDERECO

Delimiter: -

Enclosure:

Fields

#	New field	ID	Remove ID?	Type	Length	F	F	G..	D	C	N	D.	Trim type
1	RUA ENDERECO		N	String	250								both
2	NOME ESTADO ENDERECO		N	String	250								both
3	ESTADO ENDERECO		N	String	50								both

Help OK Cancel

19) Em **Leitura das Regiões**, faça:

CSV file input

Step name: Leitura das Regiões

Filename: \${diretorio}REGIOES DOS ESTADOS.csv

Delimiter: ;

Enclosure:

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

File encoding: UTF-8

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	CODIGO DO ESTADO	String		50		R\$,	.	both
2	NOME DO ESTADO	String		250		R\$,	.	both
3	REGIAO	String		50		R\$,	.	both

Help OK Get Fields Preview Cancel

20) Em **Juntando Clientes com Regiões**, faça:

Join rows (cartesian product)

Step name:

Temp directory:

TMP-file prefix:

Max. cache size (in rows):

Main step to read from:

The condition:

=

21) Por fim, em **Dimensão Cliente**, faça:

Insert / update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates: ☐

The key(s) to look up the value(s):

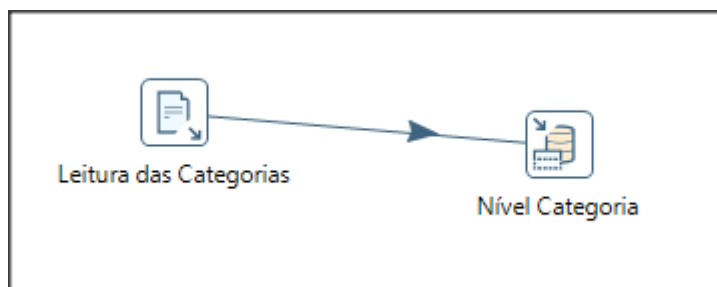
#	Table field	Comparator	Stream field1	Stream field2
1	Cod_Cliente	=	ID_CLIENTE	

Update fields:

#	Table field	Stream field	Update
1	Cod_Cliente	ID_CLIENTE	N
2	Desc_Cliente	NOME CLIENTE	Y
3	Cod_Cidade	NOME ESTADO ENDERECO	Y
4	Desc_Cidade	NOME ESTADO ENDERECO	Y
5	Cod_Estado	CODIGO DO ESTADO	Y
6	Desc_Estado	NOME DO ESTADO	Y
7	Cod_Regiao	REGIAO	Y
8	Desc_Regiao	REGIAO	Y
9	Cod_Segmento	SEGMENTO	Y
10	Desc_Segmento	SEGMENTO	Y

22) Salve e execute a transformação. Verifique no MySQL se a tabela **Dim_Cliente** foi carregada corretamente.

23) Na carga da dimensão Produto, faça-a em etapas. Primeiro, crie a transformação **DimensaoProduto_Categoria**. Insira na área de Canvas os componentes **CSV file input**, que terá nome **Leitura das Categorias**, e **Insert / update**, que terá nome **Nível Categoria**, construindo a seguinte transformação:



24) Em **Leitura das Categorias**, configure:

CSV file input

Step name: Leitura das Categorias

Filename: \${diretorio}MARCAS E CATEGORIAS.csv Browse...

Delimiter: ; Insert TAB

Enclosure:

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

File encoding: UTF-8

#	Name	Type	For...	Length	Precision	Currency	Decimal	Group	Trim type
1	MARCA	String		250		R\$,	.	both
2	CATEGORIA	String		250		R\$,	.	both

Help OK Get Fields Preview Cancel

25) Em **Nível Categoria**, faça:

Insert / update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates: ☐

The key(s) to look up the value(s):

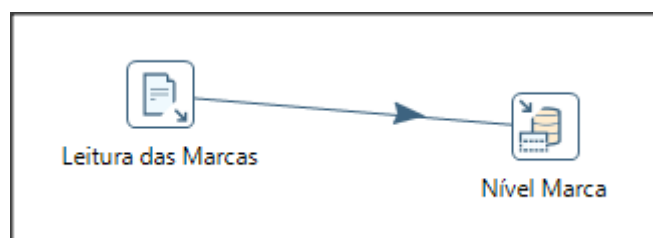
#	Table field	Comparator	Stream field1	Stream field2
1	Cod_Categoria	=	CATEGORIA	

Update fields:

#	Table field	Stream field	Update
1	Cod_Categoria	CATEGORIA	N
2	Desc_Categoria	CATEGORIA	Y

26) Salve a transformação e execute-a. Verifique se a tabela **Dim_Categoria** foi populada corretamente no MySQL.

27) Crie a transformação **DimensaoProduto_Marca**. Insira na área de Canvas os componentes **CSV file input**, que terá nome **Leitura das Marcas**, e **Insert / update**, que terá nome **Nível Marca**, construindo a seguinte transformação:



28) Em **Leitura das Marcas**, configure:

CSV file input

Step name:

Filename:

Delimiter:

Enclosure:

NIO buffer size:

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

File encoding:

#	Name	Type	Format	Length	Precis...	Currency	Decimal	Group	Trim type
1	MARCA	String		250		R\$,	.	both
2	CATEGORIA	String		250		R\$,	.	both

29) Em **Nível Marca**, faça:

Insert / update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates: ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	Cod_Marca	=	MARCA	

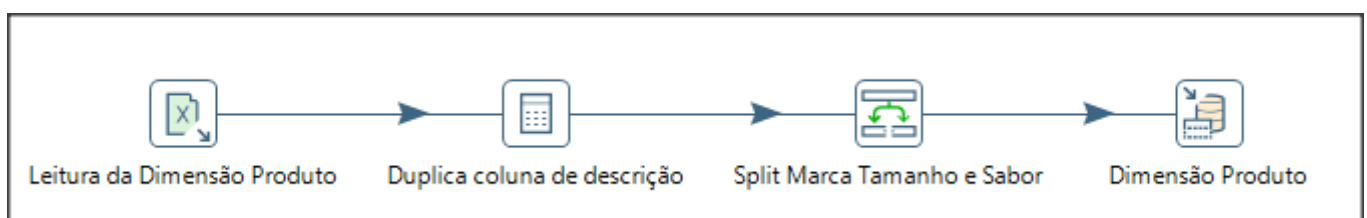
Update fields:

#	Table field	Stream field	Update
1	Cod_Marca	MARCA	N
2	Desc_Marca	MARCA	Y
3	Cod_Categoria	CATEGORIA	Y

30) Salva a transformação e execute-a. Verifique se a tabela **Dim_Marca** no MySQL foi populada com sucesso.

31) Cria a transformação **DimensaoProduto_Produto**. Arraste para a área de Canvas, os seguintes componentes:

- **Microsoft Excel input**, que terá nome **Leitura da Dimensão Produto**
- **Calculator**, que terá nome **Duplica coluna de descrição**
- **Split fields**, que terá nome **Split Marca Tamanho e Sabor**
- **Insert / Update**, que terá nome **Dimensão Produto**



32) Em **Leitura da Dimensão Produto**, na aba **Files**, faça:

Microsoft Excel input

Step name:

Files | **Sheets** | Content | Error Handling | Fields | Additional output fields

Spread sheet type (engine):

File or directory:

Regular Expression:

Exclude Regular Expression:

Selected files:

#	File/Directory	Wi...	Ex...	Required	Include
1	\${diretorio}PRODUTOS.xlsx			N	N

Accept filenames from previous steps

Accept filenames from ☐

Step to read filenames:

Field in the input to use as:

33) Em **Leitura da Dimensão Produto**, na aba **Sheets**, faça:

The screenshot shows a window titled "Microsoft Excel input". At the top, there's a "Step name" field containing the text "Leitura da Dimensão Produto". Below this are five tabs: "Files", "Sheets", "Content", "Error Handling", and "Fields". The "Files" tab is selected. Underneath the tabs, it says "List of sheets to read". To the right of this label is a table with four columns: "#", "Sheet name", "Start row", and "Start column". The first row of the table contains the values "1", "Sheet1", "0", and "0". There are several empty rows below. At the bottom left of the table area is a button labeled "Get sheetname(s)...". At the very bottom of the dialog are three buttons: "Help" (with a question mark icon), "OK", and "Cancel".

Microsoft Excel input

Step name Leitura da Dimensão Produto

Files Sheets Content Error Handling Fields Additional output fields

List of sheets to read

#	Sheet name	Start row	Start column
1	Sheet1	0	0

Get sheetname(s)...

? Help OK Preview rows Cancel

34) Em **Leitura da Dimensão Produto**, na aba **Content**, faça:

Microsoft Excel input

Step name: Leitura da Dimensão Produto

Files | Sheets | Content | Error Handling | Fields | Additional output fields

Header ☒

No empty rows ☒

Stop on empty row ☐

Limit:

Encoding:

Result filenames

Add filenames to result ☒

Help OK Preview rows Cancel

35) Em **Leitura da Dimensão Produto**, na aba **Fields**, faça:

[illegible]

36) Em **Duplica coluna de descrição**, faça as seguintes configurações:

[illegible]

37) Em **Split Marca Tamanho e Sabor**, faça:

Split fields

Step name: Split Marca Tamanho e Sabor

Field to split: DESCRITOR

Delimiter: -

Enclosure:

Fields

#	New field	ID	Remove ID?	Type	Len...	P.	F..	G	E	C	N	D	Trim type
1	MARCA		N	String	250								both
2	TAMANHO		N	String	250								both
3	SABOR		N	String	250								both

Help OK Cancel

38) Por fim, em **Dimensão Produto**, faça:

Insert / update

Step name: Dimensão Produto

Connection: Datawarehouse Sucos

Target schema:

Target table: dim_produto

Commit size: 100

Don't perform any updates: ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	Cod_Produto	=	SKU	

Update fields:

#	Table field	Stream field	Update
1	Cod_Produto	SKU	N
2	Desc_Produto	DESCRITOR PRODUTO	Y
3	Cod_Marca	MARCA	Y
4	Atr_Tamanho	TAMANHO	Y
5	Atr_Sabor	SABOR	Y

Help OK Cancel SQL

39) Salve a transformação e execute-a. Verifique se a tabela **Dim_Produto** foi populada com sucesso.