

## Consolidando o seu conhecimento

Chegou a hora de você pôr em prática o que foi visto na aula. Para isso, execute os passos listados abaixo.

---

- 1) Vá no seu MySQL e crie um banco de dados, chamado **TesteLookup**.
- 2) Abra um editor de execução de comandos SQL, e execute o script abaixo:

```
USE TesteLookup;
```

```
CREATE TABLE `dim_pessoa` (  
  `Id_Pessoa` INT(11) NOT NULL,  
  `Cod_Pessoa` VARCHAR(50) NULL DEFAULT NULL,  
  `Nome_Pessoa` VARCHAR(250) NULL DEFAULT NULL,  
  `Departamento` VARCHAR(250) NULL DEFAULT NULL,  
  PRIMARY KEY (`Id_Pessoa`)  
)  
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
;
```

```
CREATE TABLE `dim_pessoa_data` (  
  `Id_Pessoa` INT(11) NOT NULL,  
  `Cod_Pessoa` VARCHAR(50) NULL DEFAULT NULL,  
  `Desc_Pessoa` VARCHAR(250) NULL DEFAULT NULL,  
  `Departamento` VARCHAR(250) NULL DEFAULT NULL,  
  `Versao` INT(11) NULL DEFAULT NULL,  
  `Data_Inicial` DATE NULL DEFAULT NULL,  
  `Data_Final` DATE NULL DEFAULT NULL,  
  PRIMARY KEY (`Id_Pessoa`)  
)  
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
;
```

```
CREATE TABLE `dim_pessoa_incremento` (  
  `Id_Pessoa` INT(11) NOT NULL AUTO_INCREMENT,  
  `Cod_Pessoa` VARCHAR(50) NULL DEFAULT '0',  
  `Nome_Pessoa` VARCHAR(250) NULL DEFAULT '0',  
  `Departamento` VARCHAR(250) NULL DEFAULT '0',  
  PRIMARY KEY (`Id_Pessoa`)  
)  
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
AUTO_INCREMENT=6  
;
```

```
CREATE TABLE `fato` (  
  `Id_Pessoa` INT(11) NOT NULL,  
  `Data` DATE NOT NULL,  
  `Valor` DOUBLE NULL DEFAULT NULL,  
  PRIMARY KEY (`Id_Pessoa`, `Data`)  
)
```

```
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
;  
  
CREATE TABLE `fato_data` (  
  `Id_Pessoa` INT(11) NOT NULL,  
  `Data` DATE NOT NULL,  
  `Valor` FLOAT NULL DEFAULT NULL,  
  PRIMARY KEY (`Id_Pessoa`, `Data`)  
)  
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
;  
  
CREATE TABLE `fato_incremento` (  
  `Id_Pessoa` INT(11) NOT NULL,  
  `Data` DATE NOT NULL,  
  `Valor` DOUBLE NULL DEFAULT NULL,  
  PRIMARY KEY (`Id_Pessoa`, `Data`)  
)  
COLLATE='utf8_general_ci'  
ENGINE=InnoDB  
;
```

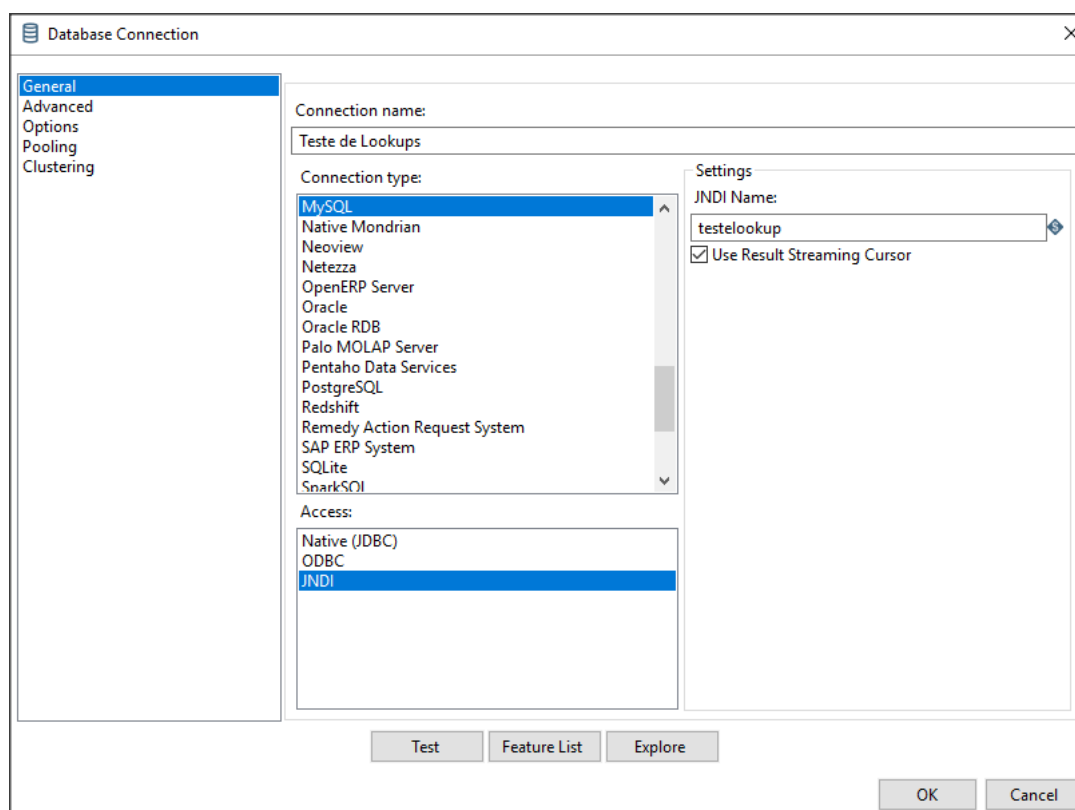
3) Em C:\treinamento\designer-tools\data-integration\simple-jndi, abra o arquivo **jdbc.properties** e inclua os comandos de conexão:

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

4) Copie os arquivos **Dimensao.xlsx** e **Fato.xlsx** para o diretório **arquivos**, da pasta **treinamento**.

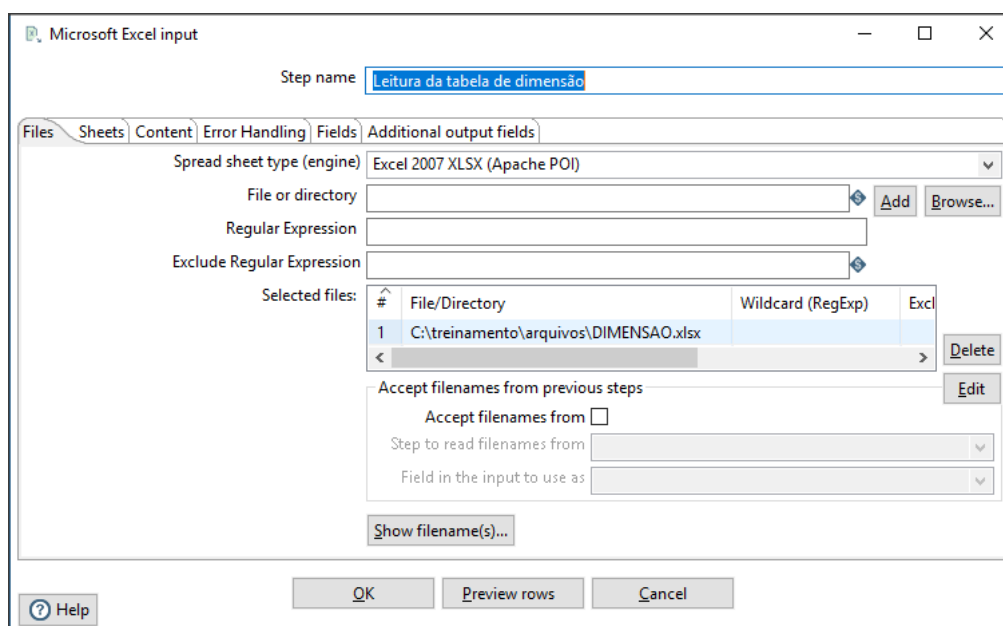
5) Abra o **Pentaho Data Integration** e crie uma transformação, chamada **TesteLookupDimensao**.

6) Crie uma conexão chamada **Teste de Lookups**, com as seguintes propriedades:



7) À área de Canvas, adicione um *step*, do tipo **Microsoft Excel input**, com o nome **Leitura da tabela de dimensão**, com as seguintes propriedades:

- Aba **Files**:



- Aba **Sheets**:

The screenshot shows the "Microsoft Excel input" dialog box. At the top, there's a title bar with standard window controls. Below it, a "Step name" field contains the text "Leitura da tabela de dimensão". The main area has four tabs: "Files", "Sheets", "Content", and "Error Handling". The "Files" tab is selected, showing a table titled "List of sheets to read". This table has four columns: "#", "Sheet name", "Start row", and "Start column". The first row contains the values "1", "Planilha1", "0", and "0". Below the table is a button labeled "Get sheetname(s)...". At the bottom of the dialog are three buttons: "OK", "Preview rows", and "Cancel". A "Help" button with a question mark icon is located at the bottom left corner.

- **Aba *Content*.**

Microsoft Excel input

Step name: Leitura da tabela de dimensão

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

Header ☒

No empty rows ☒

Stop on empty row ☐

Limit: 0

Encoding: [dropdown]

Result filenames

Add filenames to result ☒

OK Preview rows Cancel

Help

- Aka *Fields*:

[illegible]

8) Adicione um *step* **Insert / update**, chamado **Atualiza Dimensao Incremento**. Ligue o *step* **Leitura da tabela de dimensão** a este.

9) Nas propriedades deste novo *step*, inclua:

Step name: **Atualiza Dimensao Incremento**

Connection: **Teste de Lookups**

Target schema:

Target table: **dim\_pessoa\_incremento**

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_Pessoa	=	CODIGO	

Update fields:

#	Table field	Stream field	Update
1	Cod_Pessoa	CODIGO	N
2	Nome_Pessoa	NOME	Y
3	Departamento	DEPTO	Y

10) Inclua um novo *step*, do tipo **Combination lookup/update**, chamado **Atualiza Dimensao**. Ligue o *step* **Leitura da tabela de dimensão** a este.

11) Nas propriedades deste novo *step*, inclua:

Step name: **Atualiza Dimensao**

Connection: **Teste de Lookups**

Target schema:

Target table: **dim\_pessoa**

Commit size: **100**

Cache size: **9999**

Pre-load the cache? ☐

Key fields (to look up row in table):

#	Dimension field	Field in stream
1	Cod_Pessoa	CODIGO
2	Nome_Pessoa	NOME
3	Departamento	DEPTO

Technical key field: **Id\_Pessoa**

Creation of technical key:

- ☒ Use table maximum + 1
- ☐ Use sequence
- ☐ Use auto increment field

Remove lookup fields? ☐

Use hashcode? ☐

Hashcode field in table:

Date of last update field (optional):

12) Inclua um novo *step*, do tipo **Dimension lookup/update**, chamado **Atualiza Dimensão Data**. Ligue o *step* **Leitura da tabela de dimensão** a este.

13) Nas propriedades deste novo *step*, inclua:

Step name:

Update the dimension? ☒

Connection:

Target schema:

Target table:

Commit size:

Enable the cache? ☒

Pre-load the cache? ☐

Cache size in rows (0 = cache all):

**Keys** **Fields**

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	Cod_Pessoa	CODIGO

Technical key field:  New name:

Creation of technical key:

☒ Use table maximum + 1

☐ Use sequence

☐ Use auto increment field

Version field:

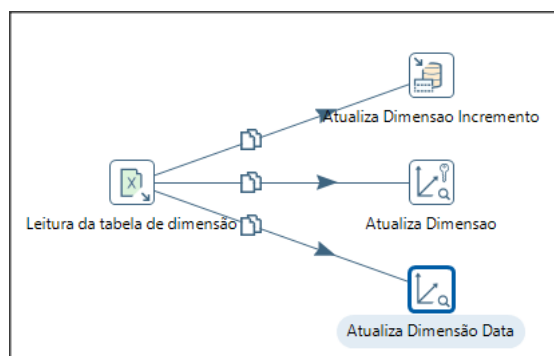
Stream Datefield:

Date range start field:  Min. year:

Use an alternative start date? ☐

Table date range end:  Max. year:

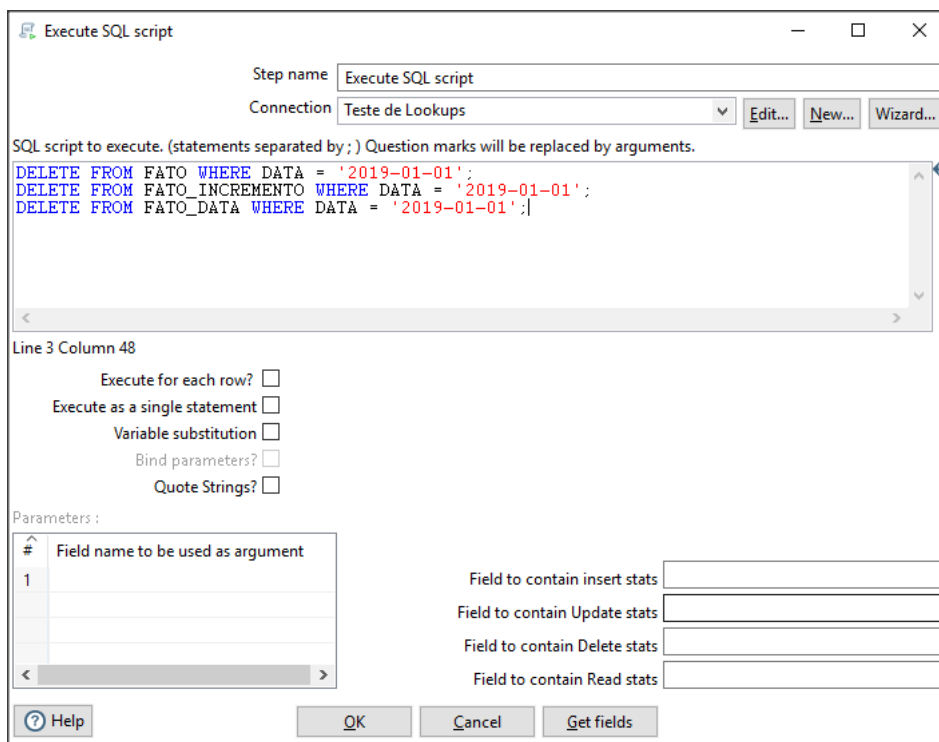
14) Você terá o seguinte esquema:



15) Salve e execute o processo. Verifique a atualização das dimensões no banco MySQL.

16) Crie a transformação **TesteLookupFato**.

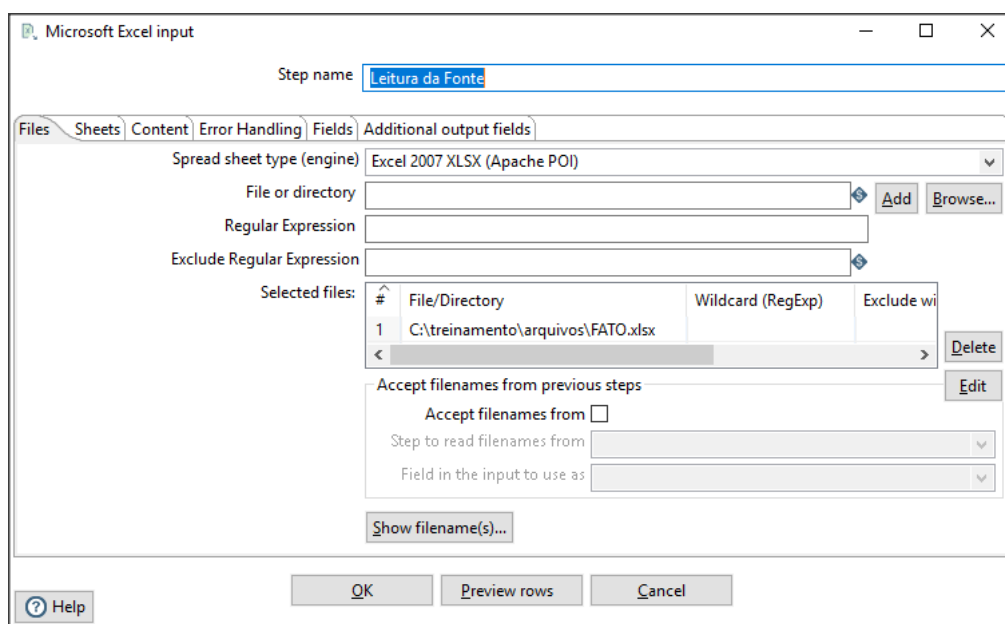
17) Acrescente o *step* do tipo **Execute SQL script**, mantendo o seu nome e incluindo as seguintes propriedades:



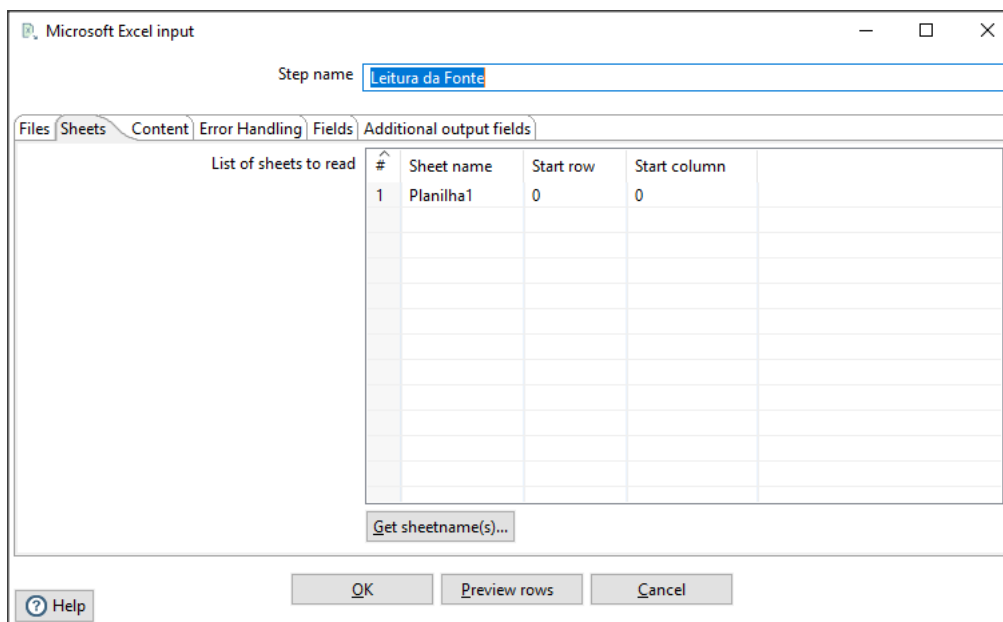
18) Inclua o *step* **Microsoft Excel input**, com o nome **Leitura da Fonte**, e ligue o **Execute SQL Script** a este.

19) Nas propriedades deste novo *step*, inclua:

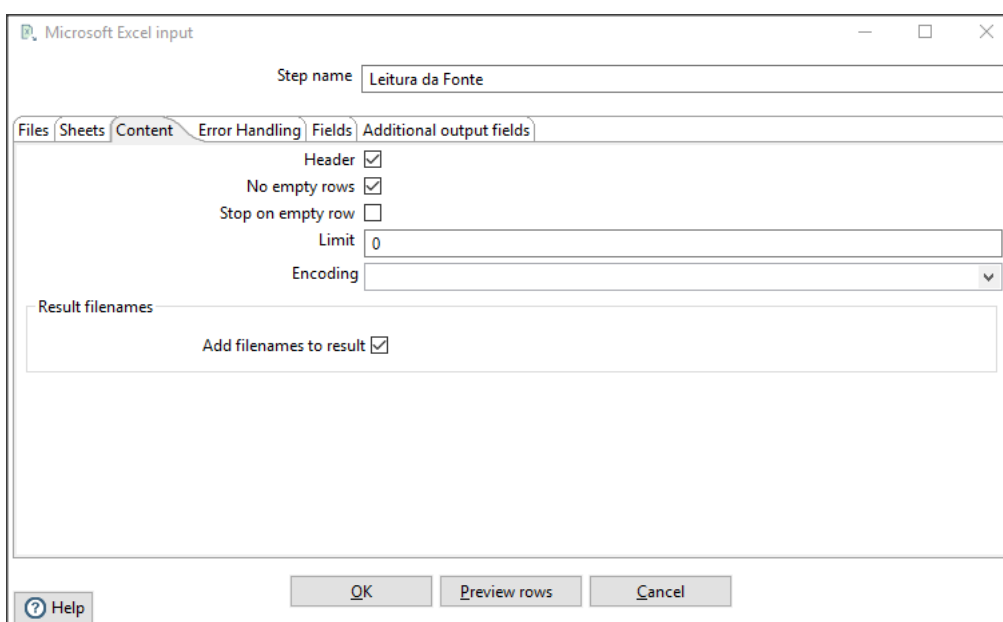
- Aba **Fontes**:



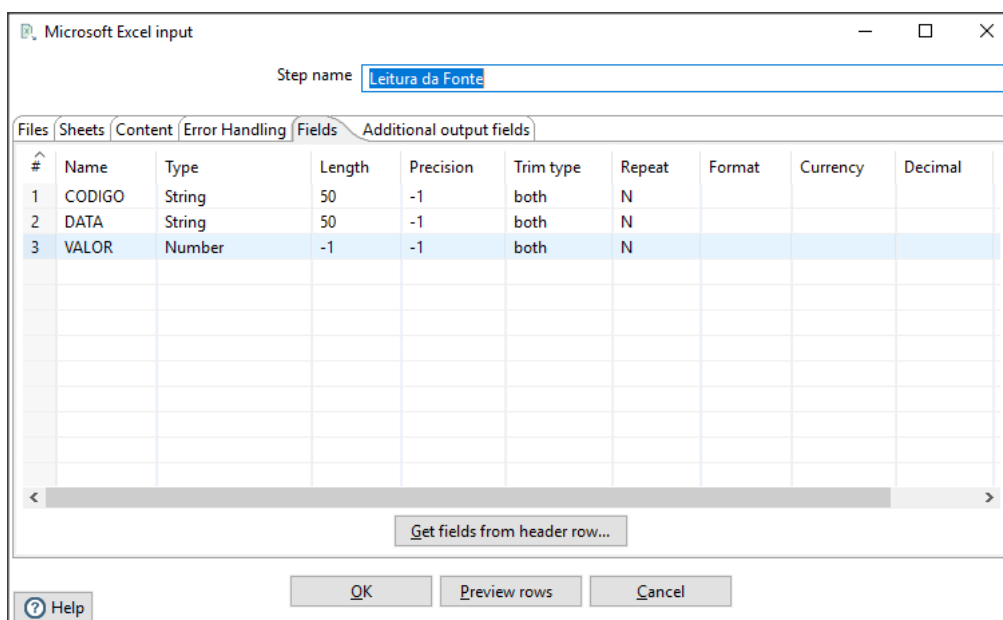
- Aba **Sheets**:



- Aba **Content**:



- Aba **Fields**:





20) Inclua um novo *step*, do tipo **Filter rows**, com o nome **Filtrar a Data**. Ligue o *step* **Leitura da Fonte** a este.

21) Nas propriedades deste novo *step*, inclua:

Filter rows

Step name:

Send 'true' data to step:

Send 'false' data to step:

The condition:

=  (String)

22) Inclua um novo *step*, do tipo **Calculator**, com o nome **Conversão de Datas**. Ligue o *step* **Filtrar a Data** a este.

23) Nas propriedades deste novo *step*, inclua:

Calculator

Step name:

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	L...	P.	Remove	Conversion mask	Decimal symbol	Group
1	DATA_DATE	Create a copy of field A	DATA			Date			N	yyyyMMdd		
2	DATA_FINAL	Create a copy of field A	DATA_DATE			Date			N	dd/MM/yyyy		

24) Inclua um novo *step*, do tipo **Select values**, com o nome **Retirar Colunas**. Ligue o *step* **Conversão de Datas** a este.

25) Nas propriedades deste novo *step*, inclua:

Select values

Step name:

Select & Alter Remove Meta-data

Fields to remove:

#	Fieldname
1	DATA
2	DATA_DATE

26) Inclua um novo *step*, do tipo **Database lookup**, com o nome **Acha ID (Incremento)**. Ligue *step* **Retirar Colunas** a este.

27) Nas propriedades deste novo *step*, inclua:

Database lookup

Step name: Acha ID (Incremento)

Connection: Teste de Lookups

Lookup schema:

Lookup table: dim\_pessoa\_incremento

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

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

#	Table field	Comparator	Field1	Field2
1	Cod_Pessoa	=	CODIGO	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	Id_Pessoa		0	None

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

Buttons: Help, OK, Cancel, Get Fields, Get lookup fields

28) Inclua um novo *step*, do tipo **Insert / update**, com o nome **Atualiza Fato Incremento**. Ligue o *step* **Acha ID (Incremento)** a este.

29) Nas propriedades deste novo *step*, inclua:

Insert / update

Step name: Atualiza Fato Incremento

Connection: Teste de Lookups

Target schema:

Target table: fato\_incremento

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	Id_Pessoa	=	Id_Pessoa	
2	Data	=	DATA_FINAL	

Update fields:

#	Table field	Stream field	Update
1	Id_Pessoa	Id_Pessoa	N
2	Data	DATA_FINAL	N
3	Valor	VALOR	Y

Buttons: Help, OK, Cancel, SQL

30) Inclua um novo *step*, do tipo **Database lookup**, com o nome **Acha ID**. Ligue o *step* **Retirar Colunas** a este.

31) Nas propriedades deste novo *step*, inclua:

Database lookup

Step name: Acha ID

Connection: Teste de Lookups

Lookup schema:

Lookup table: dim\_pessoa

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

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

#	Table field	Comparator	Field1	Field2
1	Cod_Pessoa	=	CODIGO	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	Id_Pessoa		0	None

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by: ID\_PESSOA DESC

Buttons: Help, OK, Cancel, Get Fields, Get lookup fields

32) Inclua um novo *step*, do tipo **Insert / update**, com o nome **Atualiza Fato**. Ligue o *step* **Acha ID** a este.

33) Nas propriedades deste novo *step*, inclua:

34) Inclua um novo *step*, do tipo **Dimension lookup/update**, com o nome **Acha ID (Data)**. Ligue o *step* **Retirar Colunas** a este.

35) Nas propriedades deste novo *step*, inclua:

36) Inclua um novo *step*, do tipo **Insert / update**, com o nome **Atualiza Fato Data**. Ligue o *step* **Acha ID (Data)** a este.

37) Nas propriedades deste novo *step*, inclua:

38) Você terá o seguinte esquema:

39) Salve e teste o processo. Observe os dados nas tabelas de fato.