

- 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  
;
```

[COPIAR CÓDIGO](#)

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

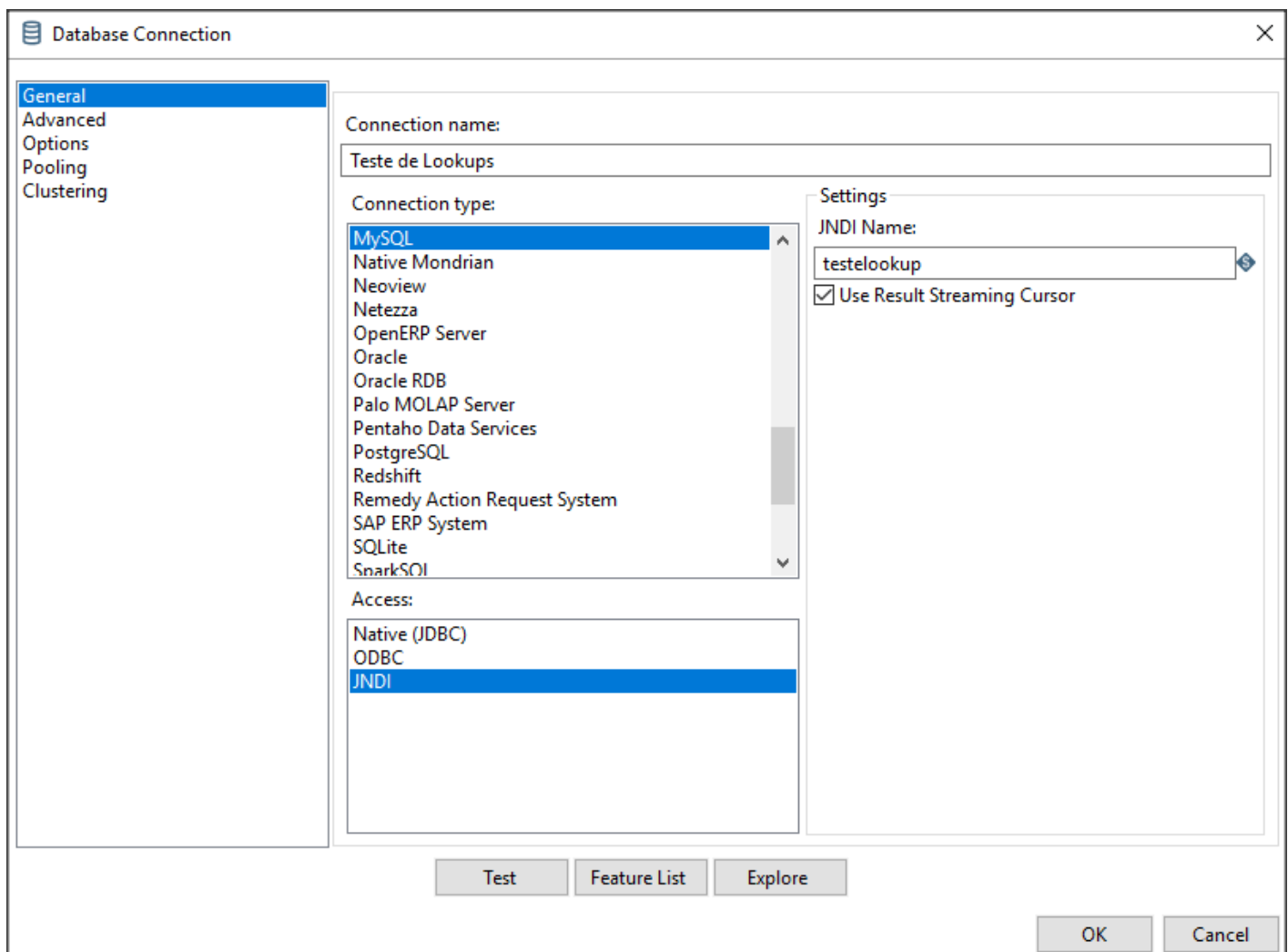
```
testelookup/type=javax.sql.DataSource  
testelookup/driver=com.mysql.jdbc.Driver  
testelookup/url=jdbc:mysql://localhost:3306/testelookup?useTimezone=1  
testelookup/user=root  
testelookup/password=root
```

[COPIAR CÓDIGO](#)

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**:

Microsoft Excel input

Step name: **Leitura da tabela de dimensão**

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)	Excl
1	C:\treinamento\arquivos\DIMENSAO.xlsx		

Accept filenames from previous steps: ☐

Step to read filenames from:

Field in the input to use as:

Show filename(s)...

OK Preview rows Cancel

Help

- Aba **Sheets**:

[illegible]

- **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 ☒

Help OK Preview rows Cancel

- **Aba *Fields*:**

Insert / update

Step name: Atualiza Dimensao Incremento

Connection: Teste de Lookups Edit... New... Wizard...

Target schema: Browse...

Target table: dim_pessoa_incremento Browse...

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		

Get fields

Update fields:

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

Get update fields

Edit mapping

Help OK Cancel SQL

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:

Combination lookup/update

Step name:

Connection:

Target schema:

Target table:

Commit size: Cache size:

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:

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:

Dimension lookup/update

Step name: **Atualiza Dimensão Data**

Update the dimension? ☒

Connection: **Teste de Lookups** [Edit...] [New...] [Wizard...]

Target schema: [Browse...]

Target table: **dim_pessoa_data** [Browse...]

Commit size: **100**

Enable the cache? ☒

Pre-load the cache? ☐

Cache size in rows (0 = cache all): **5000**

Keys Fields

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	Cod_Pessoa	CODIGO

Technical key field: **Id_Pessoa** [New name:]

Creation of technical key:

☒ Use table maximum + 1

☐ Use sequence []

☐ Use auto increment field

Version field: **Versao**

Stream Datefield: []

Date range start field: **Data_Inicial** [Min. year: **1900**]

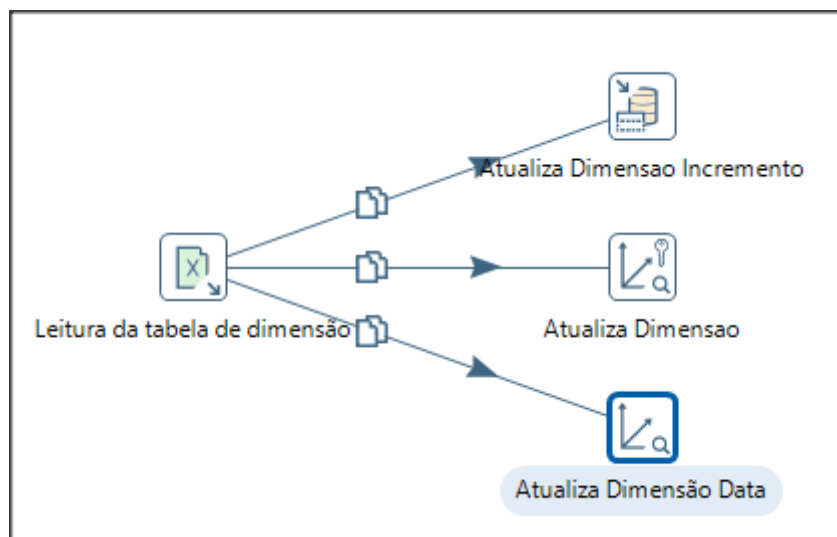
Use an alternative start date? ☐ [<Select Option> []]

Table date range end: **Data_Final** [Max. year: **2199**]

[OK] [Cancel] [Get Fields] [SQL]

[?] Help

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:

Execute SQL script

Step name: Execute SQL script

Connection: Teste de Lookups [Edit...] [New...] [Wizard...]

SQL script to execute. (statements separated by ;) Question marks will be replaced by arguments.

```
DELETE FROM FATO WHERE DATA = '2019-01-01';  
DELETE FROM FATO_INCREMENTO WHERE DATA = '2019-01-01';  
DELETE FROM FATO_DATA WHERE DATA = '2019-01-01';
```

Line 3 Column 48

☐ Execute for each row?

☐ Execute as a single statement

☐ Variable substitution

☐ Bind parameters?

☐ Quote Strings?

Parameters :

#	Field name to be used as argument
1	

Field to contain insert stats: []

Field to contain Update stats: []

Field to contain Delete stats: []

Field to contain Read stats: []

[?] Help [OK] [Cancel] [Get fields]

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**:

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	Wildcard (RegExp)	Exclude wi
1	C:\treinamento\arquivos\FATO.xlsx		

Accept filenames from previous steps ☐

Accept filenames from ☐

Step to read filenames from:

Field in the input to use as:

- **Aba *Sheets*:**

[illegible]

- **Aba Content:**

Microsoft Excel input

Step name: Leitura da Fonte

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

- **Aba *Fields*:**

[illegible]

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:

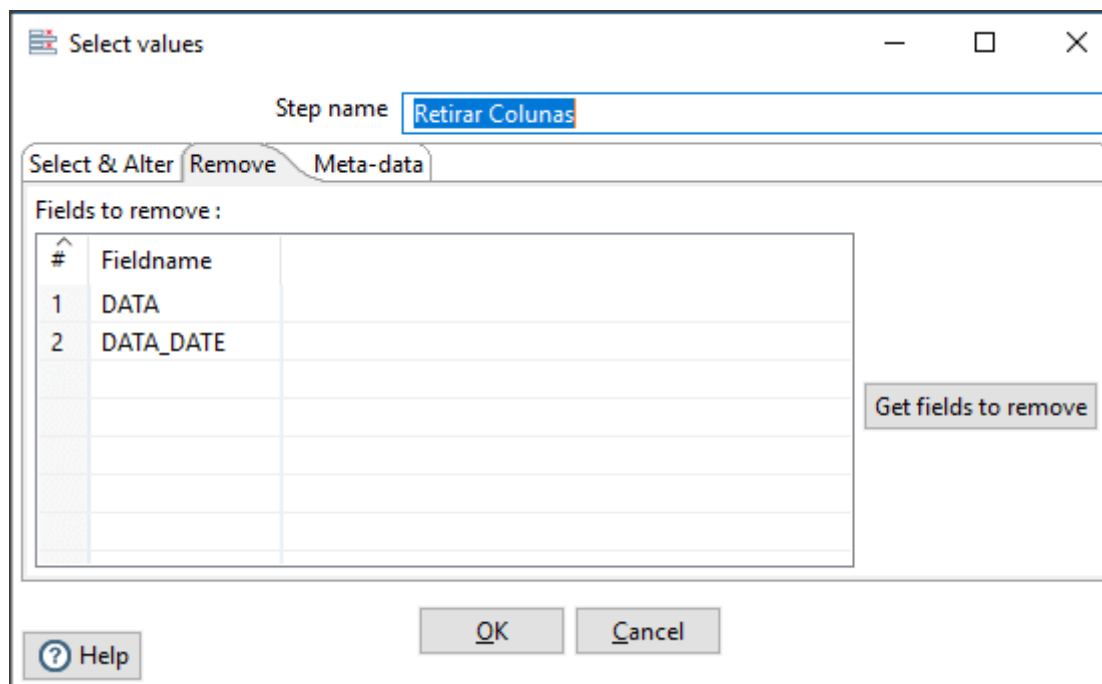
☒ Throw an error on non existing files

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:



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:

Connection:

Lookup schema:

Lookup table:

Enable cache? ☐

Cache size in rows (0=cache all rows):

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

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:

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	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	

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 ▼ Edit... New... Wizard...

Lookup schema

Browse...

Lookup table

dim_pessoa Browse...

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

? 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:

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	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

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:

Dimension lookup/update

Step name: Acha ID (Data)

Update the dimension? ☐

Connection: Teste de Lookups Edit... New... Wizard...

Target schema: Browse...

Target table: dim_pessoa_data Browse...

Commit size: 100

Enable the cache? ☒

Pre-load the cache? ☐

Cache size in rows (0 = cache all): 5000

Keys **Fields**

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	Cod_Pessoa	CODIGO

Technical key field: Id_Pessoa New name

Creation of technical key

☒ Use table maximum + 1

☐ Use sequence

☐ Use auto increment field

Version field: Versao

Stream Datefield: DATA_FINAL

Date range start field: Data_Inicial Min. year 1900

Use an alternative start date? ☐ <Select Option>

Table date range end: Data_Final Max. year 2199

OK Cancel Get Fields SQL

Help

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:

Insert / update

Step name: Atualiza Fato Data

Connection: Teste de Lookups Edit... New... Wizard...

Target schema: Browse...

Target table: fato_data Browse...

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	

Get fields

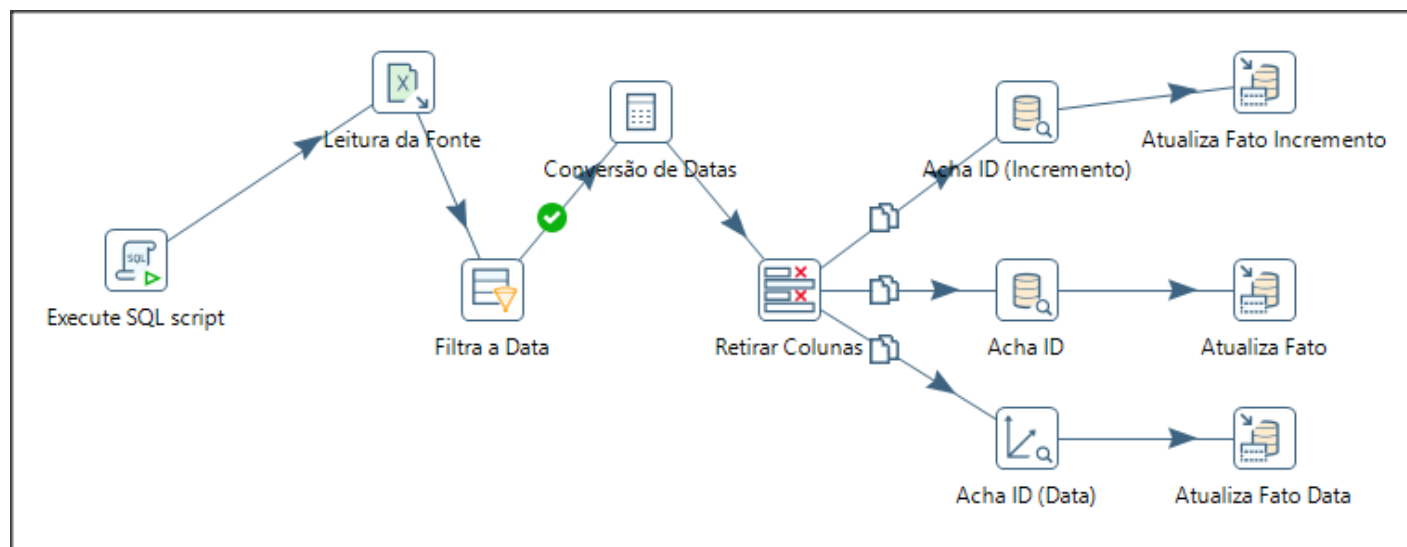
Update fields:

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

Get update fields Edit mapping

Help OK Cancel SQL

38) Você terá o seguinte esquema:



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