

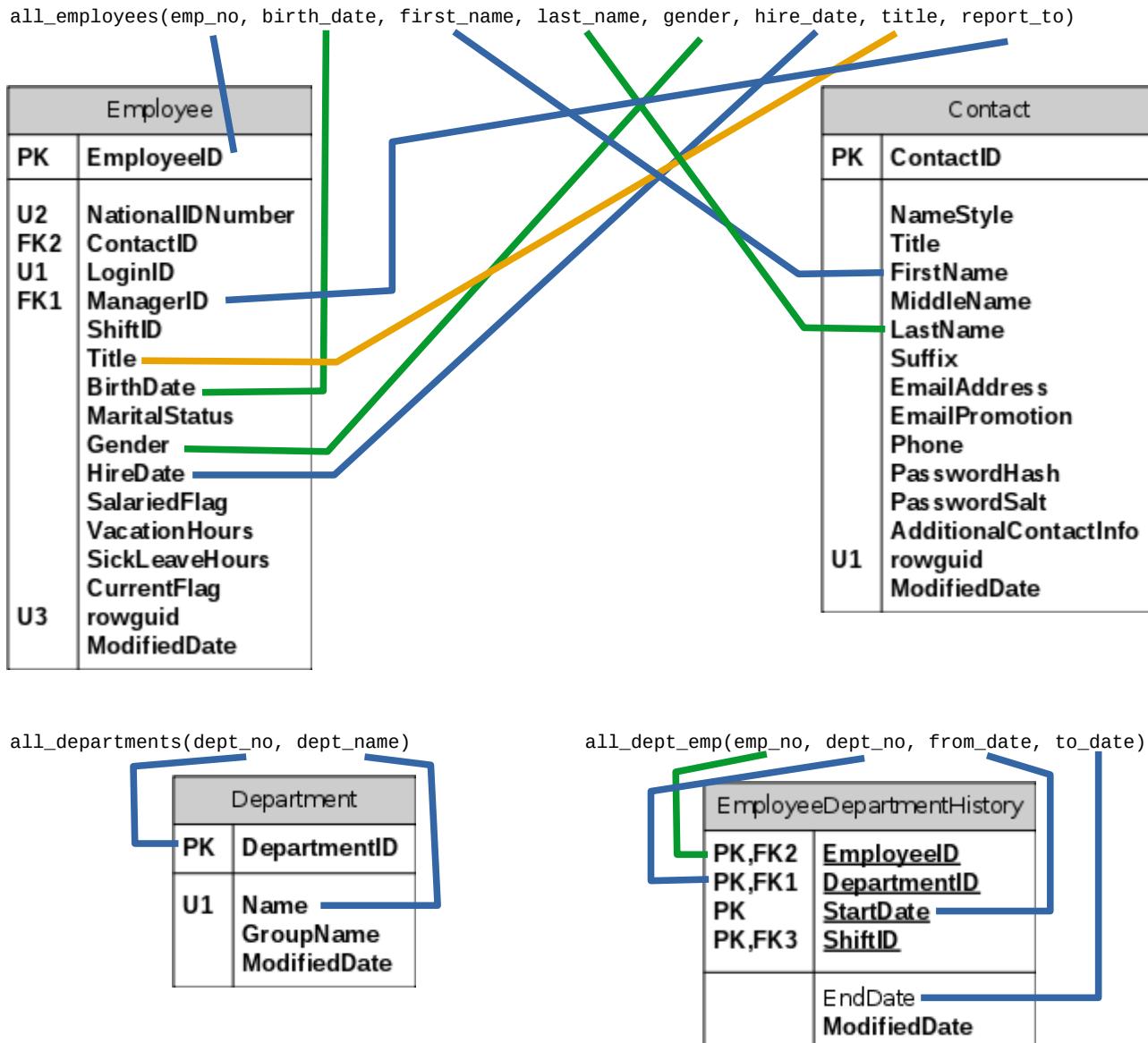
Grupo 16

78375 – João Pirralha
84725 – Henrique Almeida
84773 – Tomás Oliveira

Contents

1) Schema matching between the AdventureWorks database and the mediated schema.....	2
2) SQL views that define the schema mapping between the two data sources and the mediated schema.....	2
3) PDI transformation to detect approximate duplicates between the job titles in both databases.....	3
4) SQL instructions to create the data warehouse tables.....	7
5) Extract-Transform-Load (ETL) process that populates the data warehouse.....	8
dim_territory.....	8
dim_product.....	9
dim_time.....	12
fact_order.....	13
job.....	15
6) SQL Queries.....	16
7) SQL/OLAP query.....	16
8) PSW.....	17
9) Saiku Queries.....	18
a) Sales and quantity by country region code.....	18
b) Sales and quantity by country region code and product category.....	18
c) Sales and quantity by country region code and product category and year.....	19
10) MDX.....	19
11) PRD.....	20

1) Schema matching between the AdventureWorks database and the mediated schema



2) SQL views that define the schema mapping between the two data sources and the mediated schema

```
-- Auxiliar views
create or replace view employees.curr_dept_emp(emp_no, dept_no, from_date, to_date) as
  select emp_no, dept_no, from_date, to_date
  from employees.dept_emp
  where from_date <= current_date and to_date >= current_date;

create or replace view employees.curr_dept_manager(emp_no, dept_no) as
  select emp_no, dept_no
  from employees.dept_manager
```

```

where from_date <= current_date and to_date >= current_date;

create or replace view adventureworks.curr_dept_emp(EmployeeID, DepartmentID, ShiftID, StartDate,
EndDate, ModifiedDate) as
  select EmployeeID, DepartmentID, ShiftID, StartDate, EndDate, ModifiedDate
  from adventureworks.employeedepartmenthistory
  where StartDate <= current_date and (EndDate is NULL or EndDate >= current_date);

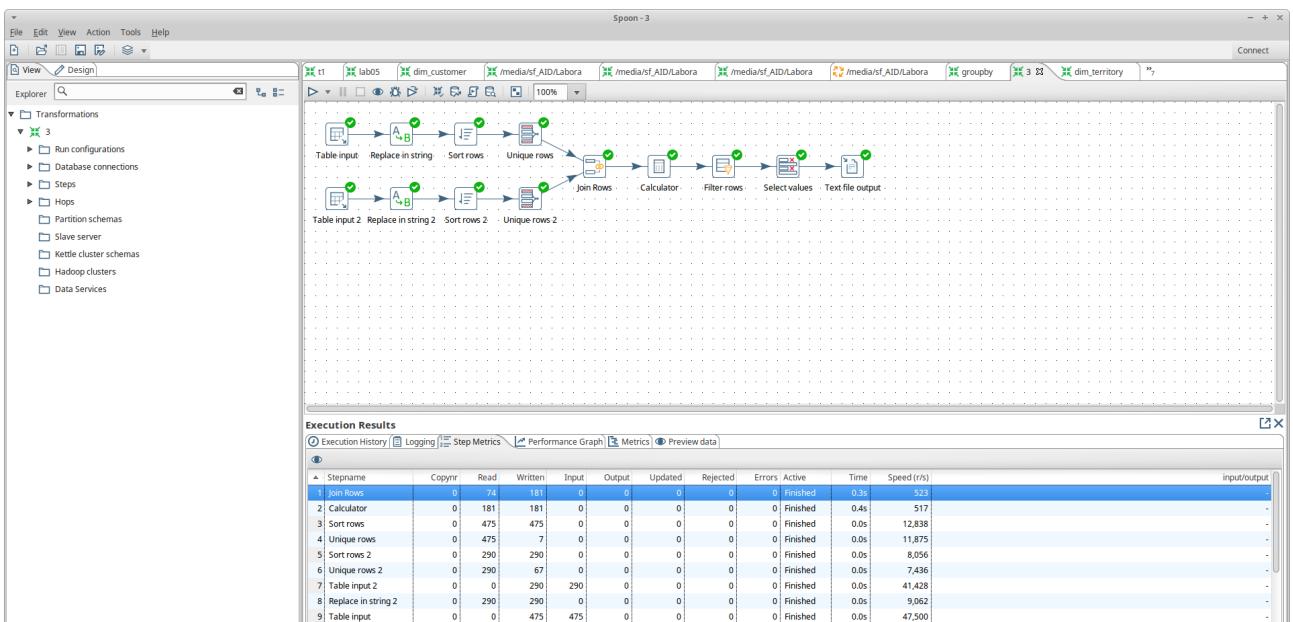
-- Main views
-- The maximum emp_no from adventureworks is 290 and the minimum from employees is 10721, so
there's no overlap
create or replace view all_employees(emp_no, birth_date, first_name, last_name, gender, hire_date,
title, report_to) as
  (select a.emp_no, a.birth_date, a.first_name, a.last_name, a.gender, a.hire_date, d.title,
c.emp_no as report_to
  from employees.employees as a,
       employees.curr_dept_emp as b,
       employees.curr_dept_manager as c,
       employees.titles as d
  where a.emp_no = b.emp_no and b.dept_no = c.dept_no
    and a.emp_no = d.emp_no)
union
  (select a.EmployeeID, a.BirthDate, b.FirstName, b.LastName, a.Gender, a.HireDate, a.Title,
a.ManagerID
  from adventureworks.employee as a,
       adventureworks.contact as b
  where a.ContactID = b.ContactID);

-- dept_no's from adventureworks are numeric and dept_no's from employees have a "d" prefix, so
there's no overlap
create or replace view all_departments(dept_no, dept_name) as
  (select dept_no, dept_name from employees.departments)
union
  (select DepartmentID, Name from adventureworks.department);

create or replace view all_dept_emp(emp_no, dept_no, from_date, to_date) as
  (select emp_no, dept_no, from_date, to_date from employees.curr_dept_emp)
union
  (select EmployeeID, DepartmentID, StartDate, EndDate from adventureworks.curr_dept_emp);

```

3) PDI transformation to detect approximate duplicates between the job titles in both databases



<p>Table input</p> <p>Step name: Table input</p> <p>Connection: employees</p> <p>SQL:</p> <pre>SELECT title as title_employees FROM titles</pre> <p>Line 1 Column 0</p> <p>Enable lazy conversion <input type="checkbox"/></p> <p>Replace variables in script? <input type="checkbox"/></p> <p>Insert data from step <input type="button" value=""/></p> <p>Execute for each row? <input type="checkbox"/></p> <p>Limit size: <input type="text" value="0"/></p> <p>OK Preview Cancel</p>	<p>Table input</p> <p>Step name: Table input 2</p> <p>Connection: adventureworks</p> <p>SQL:</p> <pre>SELECT Title as title_adventureworks FROM employee</pre> <p>Line 1 Column 0</p> <p>Enable lazy conversion <input type="checkbox"/></p> <p>Replace variables in script? <input type="checkbox"/></p> <p>Insert data from step <input type="button" value=""/></p> <p>Execute for each row? <input type="checkbox"/></p> <p>Limit size: <input type="text" value="0"/></p> <p>OK Preview Cancel</p>																																				
<p>Replace in string</p> <p>Step name: replace in string</p> <p>Fields string:</p> <table border="1"> <thead> <tr> <th>In stream field</th> <th>Out stream field</th> <th>use RegEx</th> <th>Search</th> <th>Replace with</th> <th>Set empty string?</th> <th>Replace with field</th> <th>Whole Word</th> <th>Case sensitive</th> </tr> </thead> <tbody> <tr> <td>title_employees</td> <td>title_1</td> <td><input checked="" type="checkbox"/></td> <td>[^A-Za-z0-9]</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>OK Get fields Cancel</p>	In stream field	Out stream field	use RegEx	Search	Replace with	Set empty string?	Replace with field	Whole Word	Case sensitive	title_employees	title_1	<input checked="" type="checkbox"/>	[^A-Za-z0-9]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Replace in string</p> <p>Step name: replace in string 2</p> <p>Fields string:</p> <table border="1"> <thead> <tr> <th>In stream field</th> <th>Out stream field</th> <th>use RegEx</th> <th>Search</th> <th>Replace with</th> <th>Set empty string?</th> <th>Replace with field</th> <th>Whole Word</th> <th>Case sensitive</th> </tr> </thead> <tbody> <tr> <td>title_adventureworks</td> <td>title_2</td> <td><input checked="" type="checkbox"/></td> <td>[^A-Za-z0-9]</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>OK Get fields Cancel</p>	In stream field	Out stream field	use RegEx	Search	Replace with	Set empty string?	Replace with field	Whole Word	Case sensitive	title_adventureworks	title_2	<input checked="" type="checkbox"/>	[^A-Za-z0-9]	<input type="checkbox"/>				
In stream field	Out stream field	use RegEx	Search	Replace with	Set empty string?	Replace with field	Whole Word	Case sensitive																													
title_employees	title_1	<input checked="" type="checkbox"/>	[^A-Za-z0-9]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																													
In stream field	Out stream field	use RegEx	Search	Replace with	Set empty string?	Replace with field	Whole Word	Case sensitive																													
title_adventureworks	title_2	<input checked="" type="checkbox"/>	[^A-Za-z0-9]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																													
<p>Sort rows</p> <p>Step name: Sort rows</p> <p>Sort directory: %{java.io.tmpdir}%</p> <p>TMP-file prefix: out</p> <p>Sort size (rows in memory): 1000000</p> <p>Free memory threshold (in %): 1000000</p> <p>Compress TMP Files? <input type="checkbox"/></p> <p>Only pass unique rows? (verifies keys only) <input type="checkbox"/></p> <p>Fields:</p> <table border="1"> <thead> <tr> <th>Fieldname</th> <th>Ascending</th> <th>Case sensitive compare?</th> <th>Sort based on current locale?</th> <th>Collator Strength</th> <th>Presorted?</th> </tr> </thead> <tbody> <tr> <td>title_employees</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>OK Cancel Get Fields</p>	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?	title_employees	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<p>Sort rows</p> <p>Step name: Sort rows 2</p> <p>Sort directory: %{java.io.tmpdir}%</p> <p>TMP-file prefix: out</p> <p>Sort size (rows in memory): 1000000</p> <p>Free memory threshold (in %): 1000000</p> <p>Compress TMP Files? <input type="checkbox"/></p> <p>Only pass unique rows? (verifies keys only) <input type="checkbox"/></p> <p>Fields:</p> <table border="1"> <thead> <tr> <th>Fieldname</th> <th>Ascending</th> <th>Case sensitive compare?</th> <th>Sort based on current locale?</th> <th>Collator Strength</th> <th>Presorted?</th> </tr> </thead> <tbody> <tr> <td>title_adventureworks</td> <td><input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>OK Cancel Get Fields</p>	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?	title_adventureworks	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>												
Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?																																
title_employees	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																
Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?																																
title_adventureworks	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																

Unique rows

Step name **Unique rows**

Settings

Add counter to output? Counter field

Redirect duplicate row Error description

Fields to compare on (no entries means: compare complete row)

Fieldname	Ignore case
1 title_employees	N

(?) Help **OK** **Cancel** **Get**

Unique rows

Step name **Unique rows 2**

Settings

Add counter to output? Counter field

Redirect duplicate row Error description

Fields to compare on (no entries means: compare complete row)

Fieldname	Ignore case
1 title_adventureworks	N

(?) Help **OK** **Cancel** **Get**

Join rows

Step name **Join Rows**

Temp directory **%%java.io.tmpdir%%** **Browse...**

TMP-file prefix **out**

Max. cache size (in rows) **500**

Main step to read from **title_employees**

The condition:

title_employees < title_adventureworks

(?) Help **OK** **Cancel**

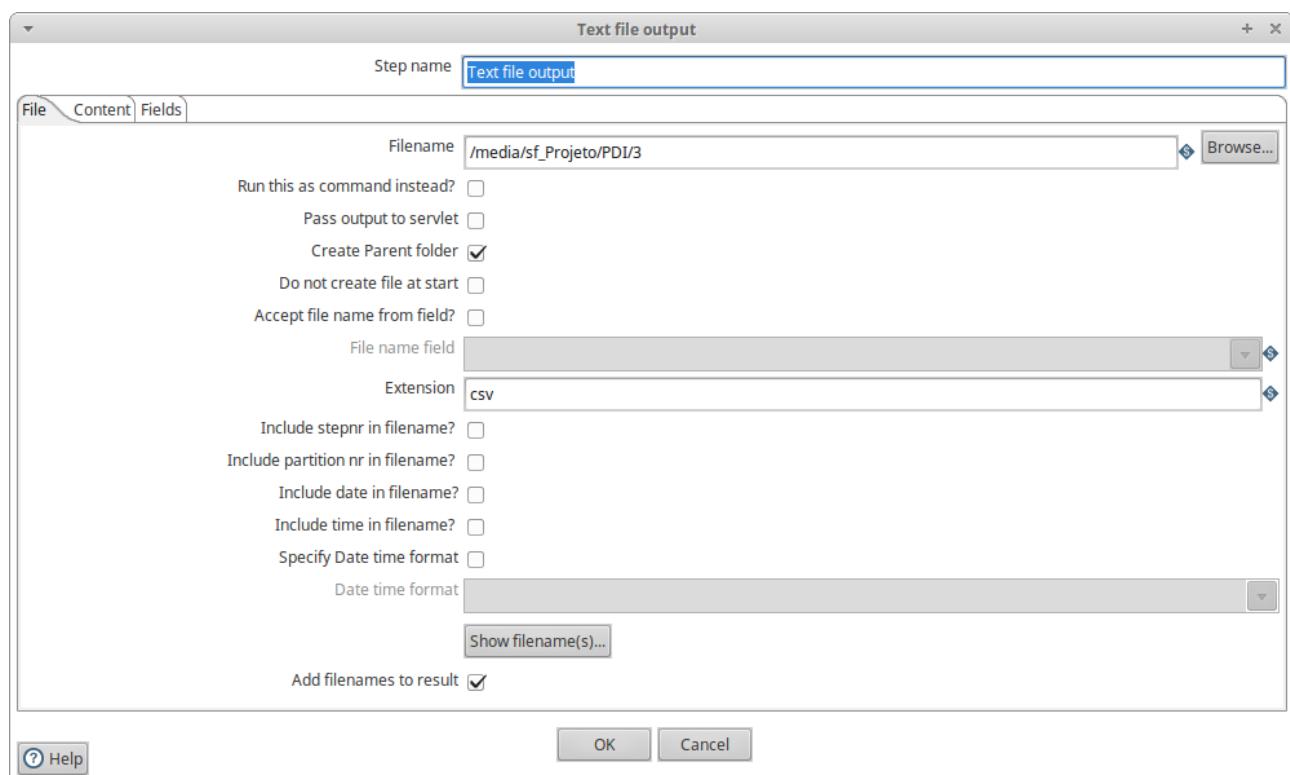
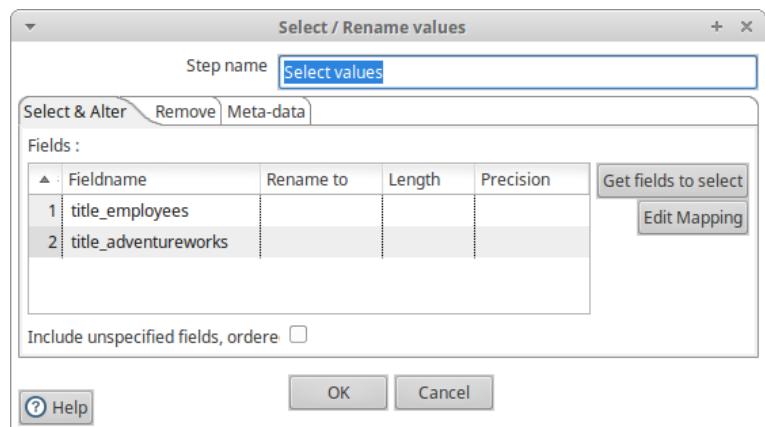
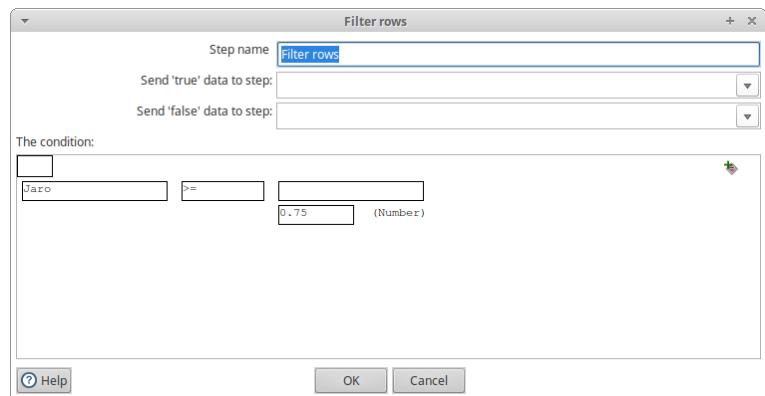
Calculator

Step name **Calculator**

Fields:

New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove
1 Levenshtein	Levenshtein Distance (source A and target B)	title_1	title_2		Number			N
2 DamerauLevenshtein	DamerauLevenshtein distance between String A and String B	title_1	title_2		Number			N
3 NeedlemanWunch	NeedlemanWunsch distance between String A and String B	title_1	title_2		Number			N
4 Jaro	Jaro similitude between String A and String B	title_1	title_2		Number			N
5 JaroWinkler	JaroWinkler similitude between String A and String B	title_1	title_2		Number			N

(?) Help **OK** **Cancel**



Examine preview data

Rows of step: Text file output (3 rows)

	title_employees	title_adventureworks
1	Assistant Engineer	Design Engineer
2	Engineer	Engineering Manager
3	Senior Engineer	Senior Tool Designer

Close

4) SQL instructions to create the data warehouse tables

```

DROP DATABASE IF EXISTS adventureworks_dw;
CREATE DATABASE adventureworks_dw;

USE adventureworks_dw;

-- salesterritory
CREATE TABLE dim_territory (
    TerritoryID int(11) NOT NULL,
    Name varchar(50) NOT NULL,
    CountryRegionCode varchar(3) NOT NULL,
    `Group` varchar(50) NOT NULL,
    PRIMARY KEY (TerritoryID)
);

-- product, productssubcategory & productcategory
-- Slowly changing dimension!
CREATE TABLE dim_product (
    ProductSK int(11),
    ProductID int(11),
    Name varchar(50),
    ProductSubcategoryName varchar(50),
    ProductCategoryName varchar(50),
    Version int,
    DateFrom DATETIME,
    DateTo DATETIME,
    PRIMARY KEY (ProductSK)
);

-- salesorderheader
CREATE TABLE dim_time (
    OrderDate timestamp NOT NULL,
    `Day` int,
    MonthID int,
    MonthName varchar(255),
    `Year` int,
    PRIMARY KEY (OrderDate)
);

-- salesorderdetail & salesorderheader
CREATE TABLE fact_order (
    SalesOrderID int(11) NOT NULL,
    SalesOrderDetailID int(11) NOT NULL,
    TerritoryID int(11) NOT NULL,
    ...
);

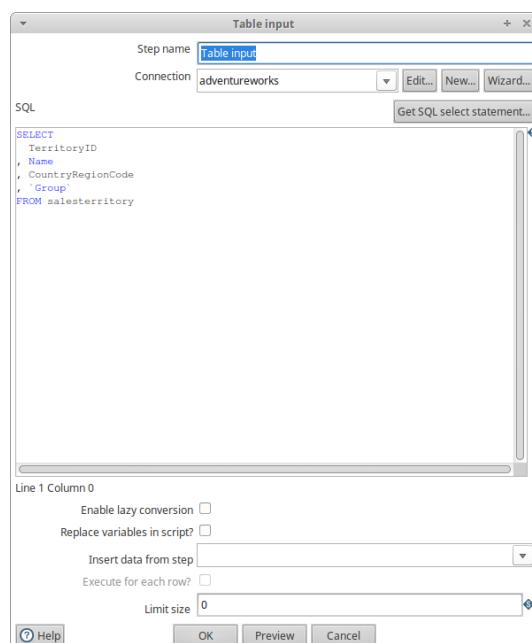
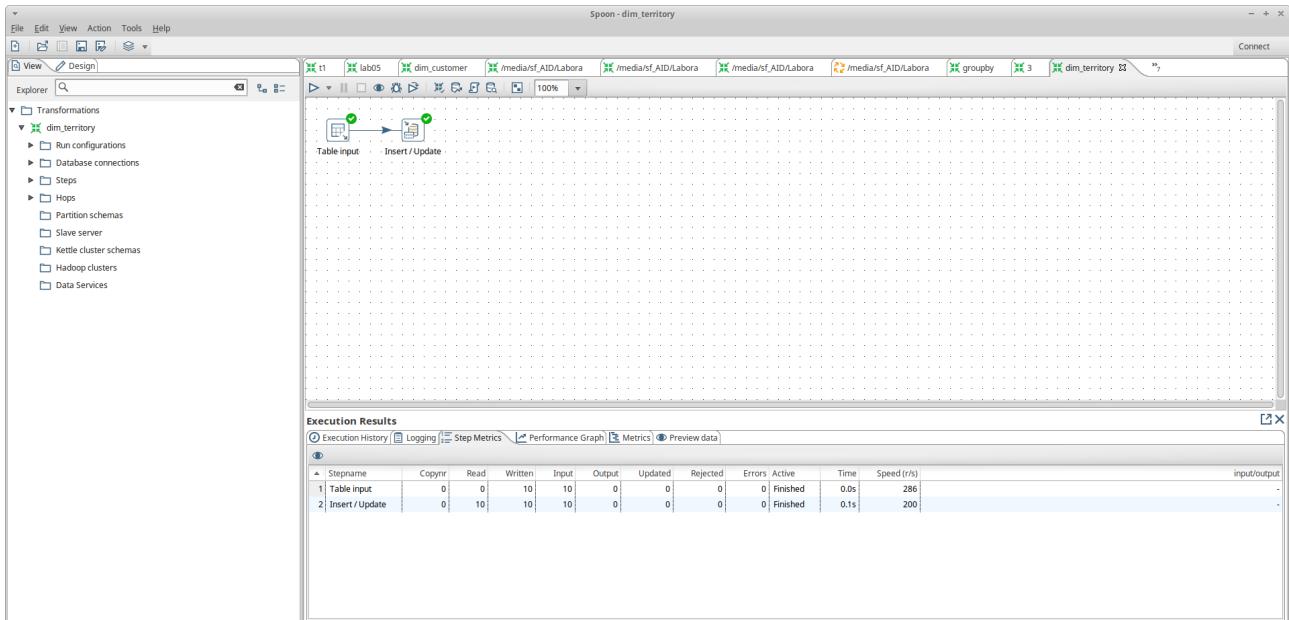
```

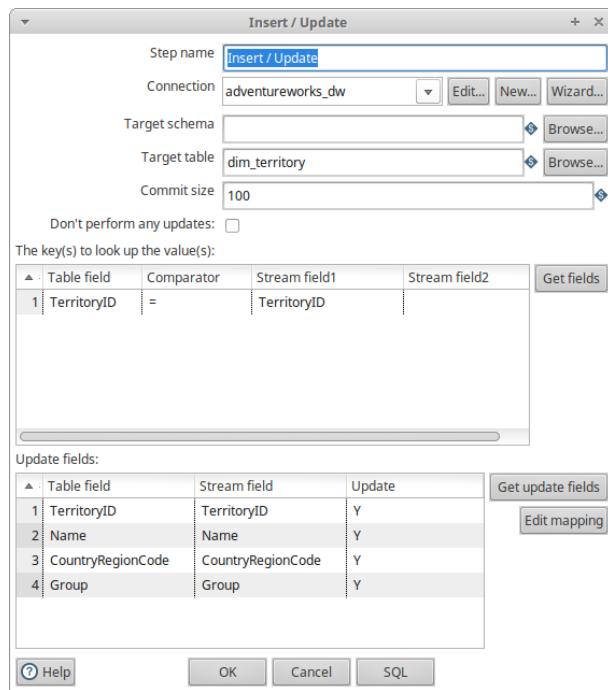
```

ProductSK int(11) NOT NULL,
OrderDate timestamp NOT NULL,
LineTotal double NOT NULL,
OrderQty smallint(6) NOT NULL,
PRIMARY KEY (SalesOrderDetailID, SalesOrderID),
FOREIGN KEY (TerritoryID) REFERENCES dim_territory (TerritoryID),
FOREIGN KEY (ProductSK) REFERENCES dim_product (ProductSK),
FOREIGN KEY (OrderDate) REFERENCES dim_time (OrderDate)
);

```

5) Extract-Transform-Load (ETL) process that populates the data warehouse dim_territory





dim_product

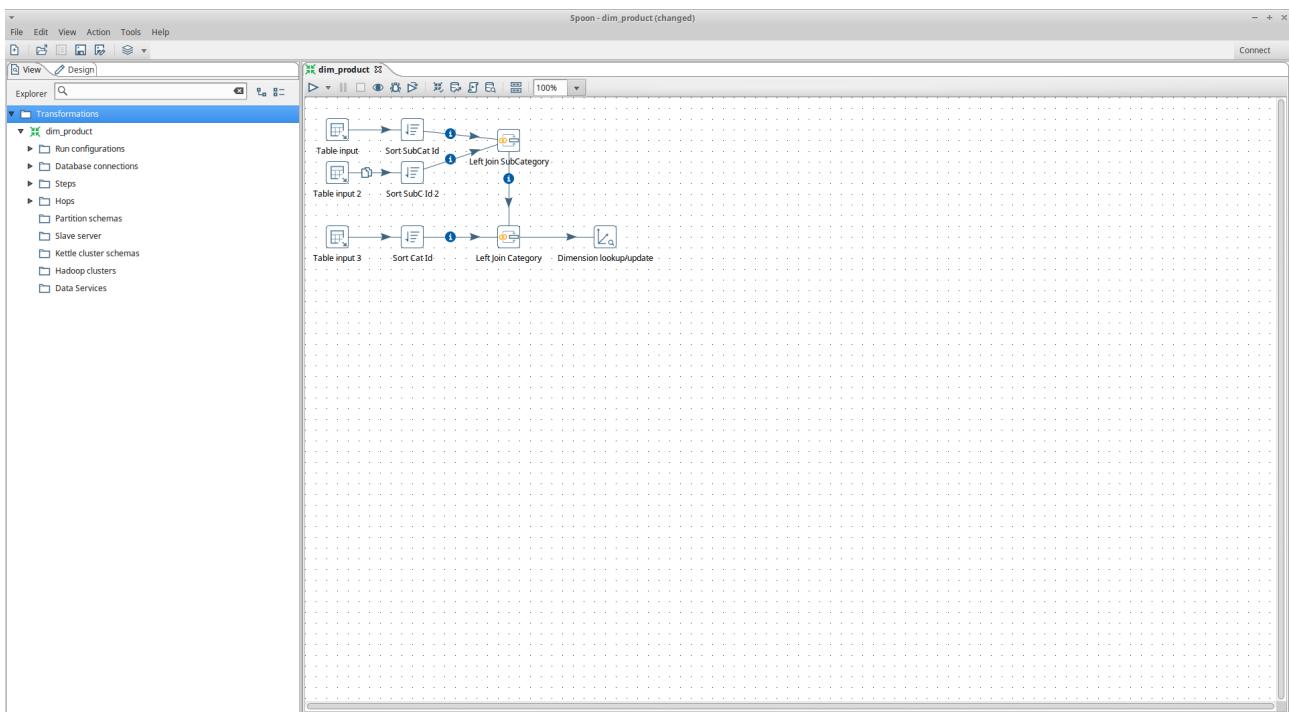


Table input

Step name: Table input

Connection: adventureworks

SQL:

```
SELECT
    ProductID,
    Name,
    ProductSubcategoryID
FROM product
```

Line 1 Column 0

Enable lazy conversion

Replace variables in script?

Insert data from step:

Execute for each row?

Limit size: 0

Table input

Step name: Table input 2

Connection: adventureworks

SQL:

```
SELECT
    ProductSubcategoryID AS ProductSubcategoryID2,
    ProductCategoryID AS ProductCategoryID2,
    Name AS ProductSubcategoryName
FROM productssubcategory
```

Line 1 Column 0

Enable lazy conversion

Replace variables in script?

Insert data from step:

Execute for each row?

Limit size: 0

Table input

Step name: Table input 3

Connection: adventureworks

SQL:

```
SELECT
    ProductCategoryID AS ProductCategoryID3,
    Name AS ProductCategoryName
FROM productcategory
```

Line 1 Column 0

Enable lazy conversion

Replace variables in script?

Insert data from step:

Execute for each row?

Limit size: 0

Sort rows

Step name: Sort SubCat Id

Sort directory: %java.io.tmpdir%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies keys only)

Fields:

Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
ProductSubcategoryID	Y	N	N	0	N

Sort rows

Step name: Sort SubC Id 2

Sort directory: %java.io.tmpdir%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies keys only)

Fields:

Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
ProductSubcategoryID2	Y	N	N	0	N

Sort rows

Step name: Sort SubCat Id

Sort directory: %java.io.tmpdir%

TMP-file prefix: out

Sort size (rows in memory): 1000000

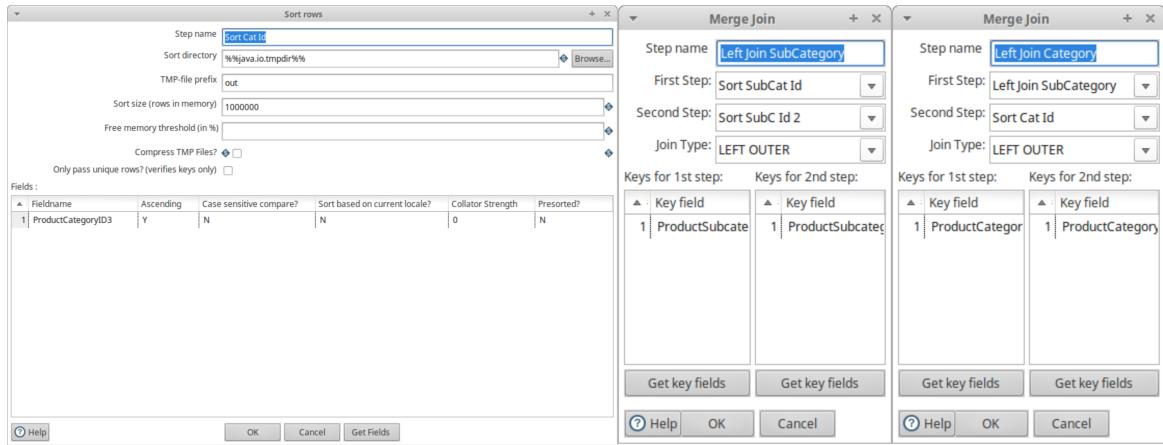
Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies keys only)

Fields:

Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
ProductSubcategoryID2	Y	N	N	0	N



Dimension Lookup / Update

Step name: Dimension lookup/update
Update the dimension?
Connection: adventureworks_dw
Target schema:
Target table: dim_product
Commit size: 100
Enable the cache?
Pre-load the cache?
Cache size in rows (0 = cache all): 5000

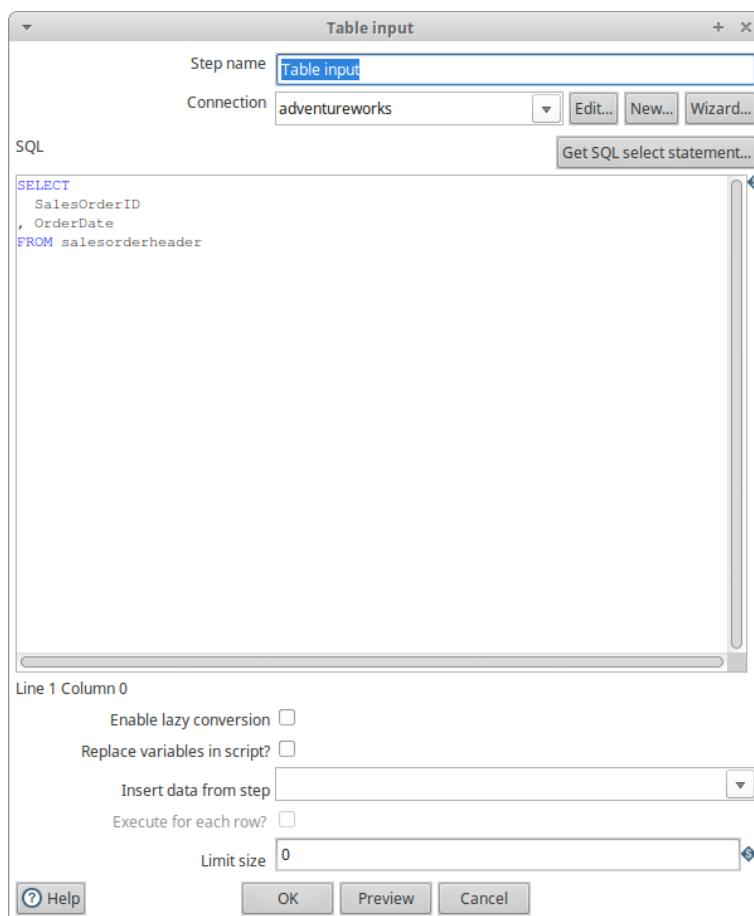
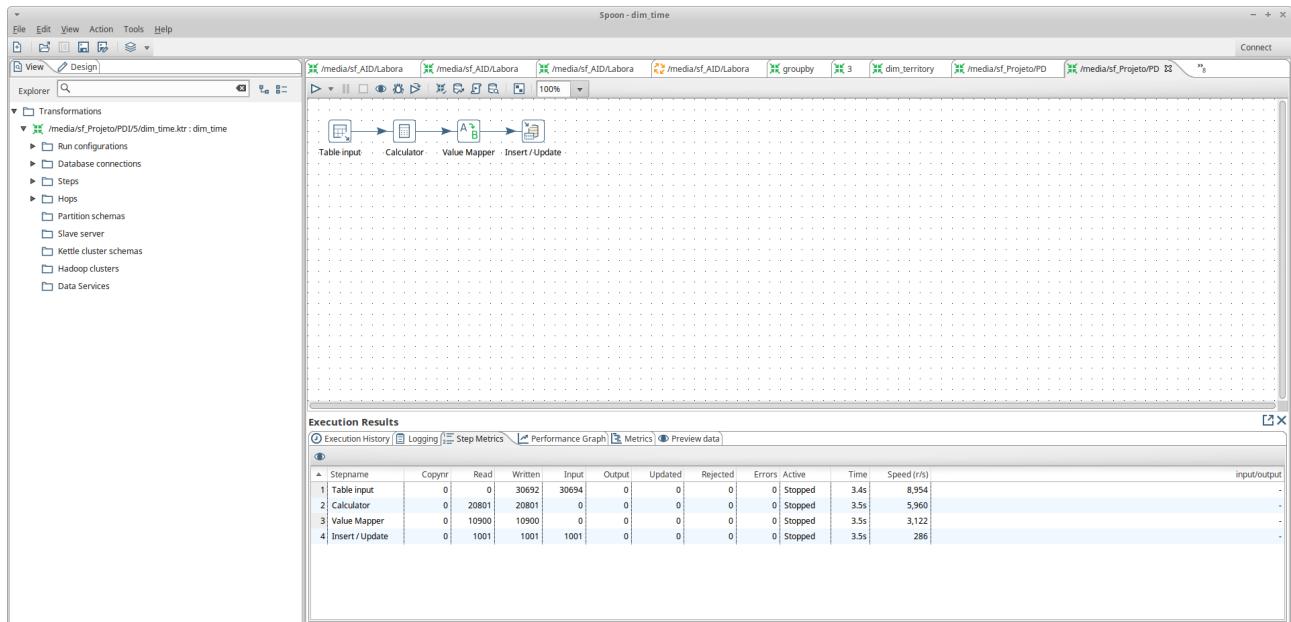
Keys **Fields**

Key fields (to look up row in dimension):

Dimension field	Field in stream
ProductID	ProductID

Technical key field: ProductSK
Creation of technical key:
 Use table maximum + 1
 Use sequence
 Use auto increment field
Version field: Version
Stream Datefield:
Date range start field: DateFrom 1900
Use an alternative start date? <Select Option> 2199
Table date range end: DateTo 2199

dim_time



Calculator

Step name: **Calculator**

Fields:

New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping symbol	Currency
1 Day	Day of month of date A	OrderDate			Integer			N				
2 MonthID	Month of date A	OrderDate			Integer			N				
3 Year	Year of date A	OrderDate			Integer			N				

OK Cancel

Value Mapper

Step name: **Value Mapper**

Fieldname to use: **MonthID**

Target field name (empty=over): **MonthName**

Default upon non-matching:

Field values:

Source value	Target value
1	Jan
2	Oct
3	Nov
4	Dec
5	Feb
6	Mar
7	Apr
8	May
9	Jun
10	Jul
11	Aug
12	Sep

OK Cancel

Insert / Update

Step name: **Insert / Update**

Connection: **adventureworks_dw**

Target schema:

Target table: **dim_time**

Commit size: **100**

Don't perform any updates:

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

Table field	Comparator	Stream field1	Stream field2
OrderDate	=	OrderDate	

Update fields:

Table field	Stream field	Update
OrderDate	OrderDate	Y
Day	Day	Y
MonthID	MonthID	Y
MonthName	MonthName	Y
Year	Year	Y

Get update fields Edit mapping

OK Cancel SQL

fact_order

Spoon - fact_order

File Edit View Action Tools Help

Explorer

- Transformations
 - mediaf_Projeto/PDU/5/fact_order.ktr : fact_order
- Run configurations
- Database connections
- Steps
- Hops
- Partition schemas
- Slave server
- Kettle cluster schemas
- Hadoop clusters
- Data Services

Table input → Database lookup → Insert / Update

Execution Results

Stepname	Copyin	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1 Table input	0	0	11852	11854	0	0	0	0	Stopped	26.7s	445
2 Database lookup	0	1959	1958	1959	0	0	0	0	Stopped	26.6s	74
3 Insert / Update	0	1001	1001	1001	0	0	0	0	Stopped	26.7s	38

Table input

Step name: Table input
Connection: adventureworks
Get SQL select statement...

```
SELECT
    a.SalesOrderID
    , SalesOrderDetailID
    , TerritoryID
    , ProductID
    , OrderDate
    , LineTotal
    , OrderQty
FROM salesorderheader AS a
JOIN salesorderdetail AS b
ON a.SalesOrderID = b.SalesOrderID;
```

Line 1 Column 0

Enable lazy conversion
Replace variables in script?
Insert data from step
Execute for each row?
Limit size: 0

OK Preview Cancel

Database Value Lookup

Step name: Database lookup
Connection: adventureworks_dw
Lookup schema:
Lookup table: dim_product
Enable cache?
Cache size in rows (0=cache everything): 0
Load all data from table

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

Table field	Comparator	Field1	Field2
1 ProductID	=	ProductID	
2 DateFrom	<=	OrderDate	
3 DateTo	>	OrderDate	

Values to return from the lookup table:

Field	New name	Default	Type
1 ProductSK			Integer

Do not pass the row if the lookup fails
Fail on multiple results?
Order by

OK Cancel Get Fields Get lookup fields

Insert / Update

Step name: Insert / Update
Connection: adventureworks_dw
Target schema:
Target table: fact_order
Commit size: 100
Don't perform any updates:

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

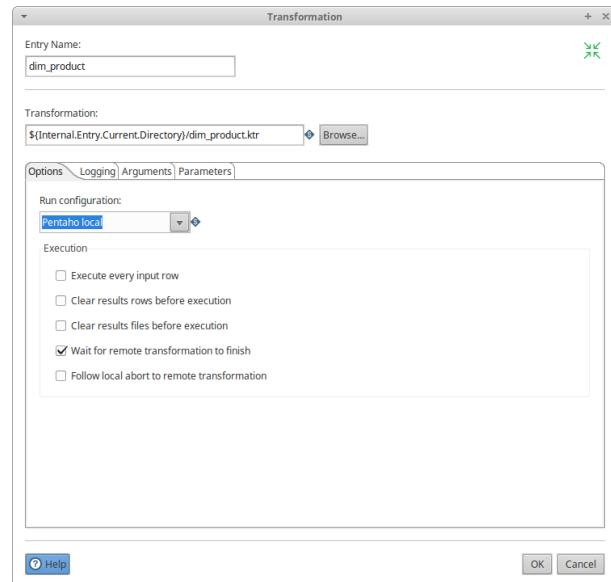
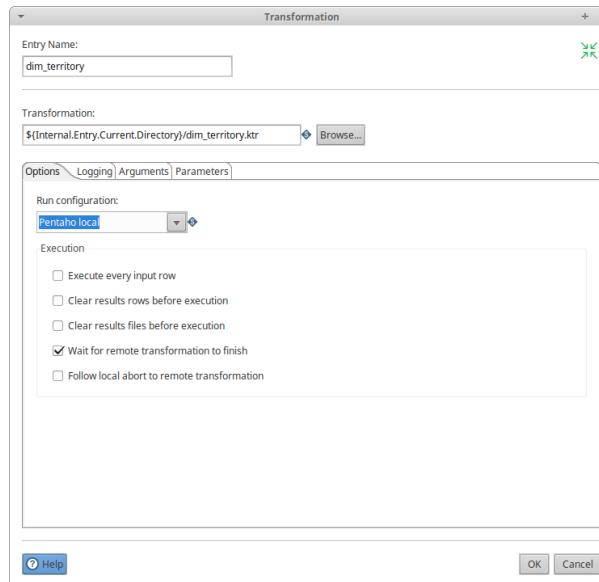
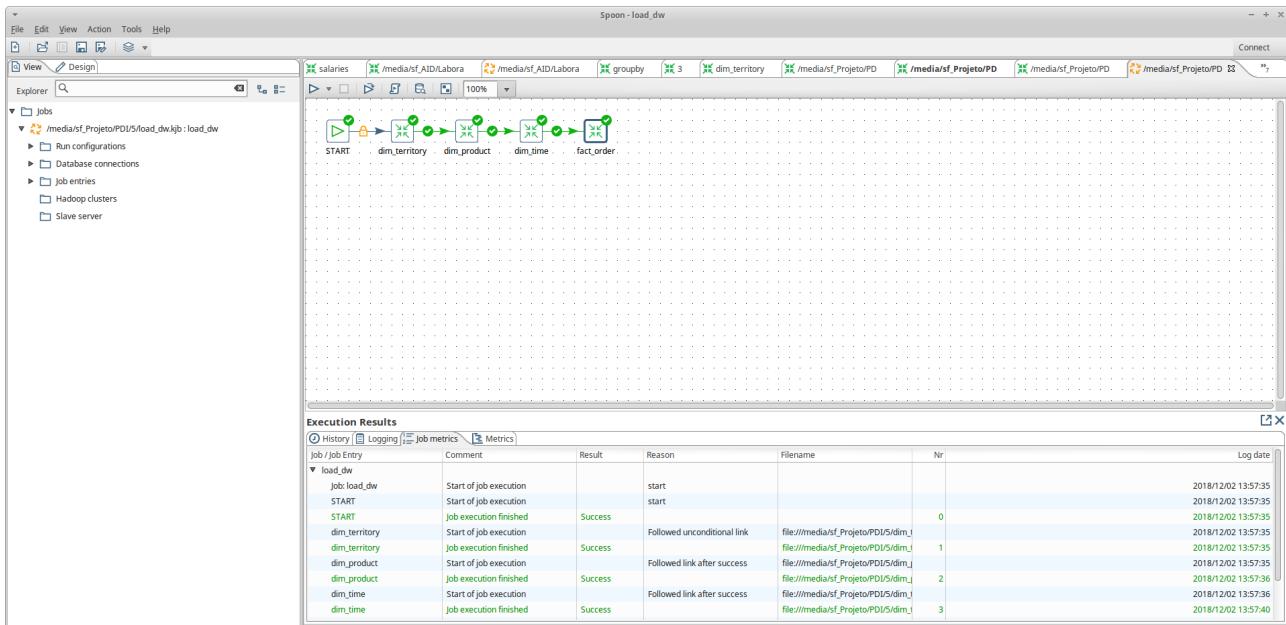
Table field	Comparator	Stream field1	Stream field2	Get fields
1 SalesOrderID	=	SalesOrderID		
2 SalesOrderDetailID	=	SalesOrderDetailID		

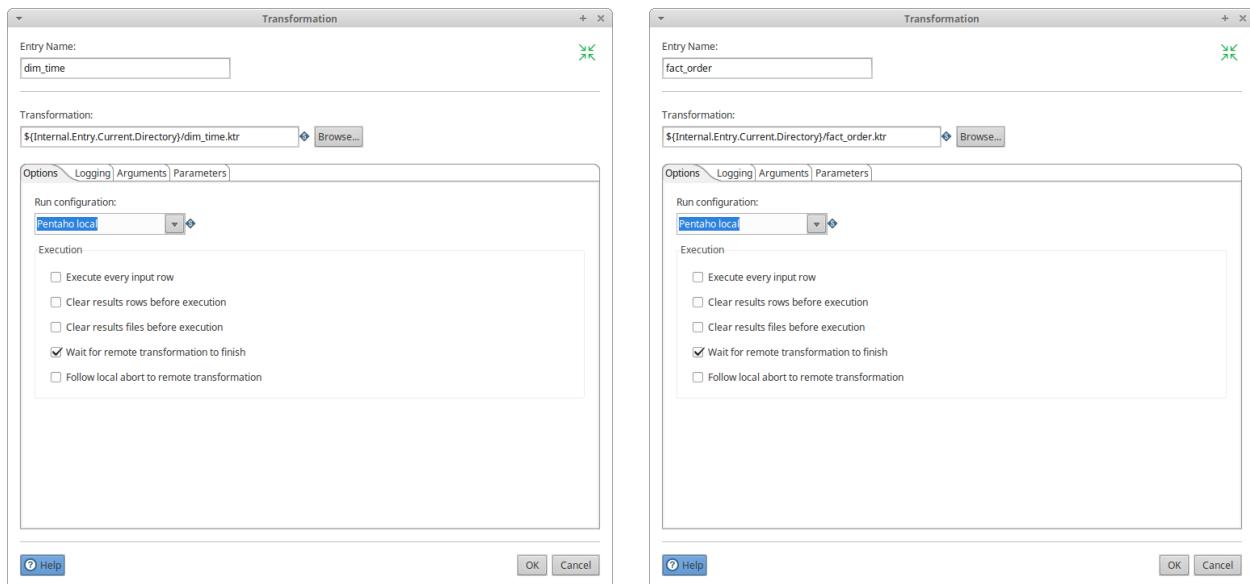
Update fields:

Table field	Stream field	Update	Get update fields	Edit mapping
1 SalesOrderID	SalesOrderID	Y		
2 SalesOrderDetailID	SalesOrderDetailID	Y		
3 TerritoryID	TerritoryID	Y		
4 ProductSK	ProductSK	Y		
5 OrderDate	OrderDate	Y		
6 LineTotal	LineTotal	Y		
7 OrderQty	OrderQty	Y		

OK Cancel SQL

job





6) SQL Queries

```
Use adventureworks_dw;

-- a)

Select sum(OrderQty) as TotalQ
From fact_order;

-- b)

Select sum(OrderQty) as TotalQ, CountryRegionCode
From fact_order natural join dim_territory
Group by CountryRegionCode;

-- c)

Select sum(OrderQty) as TotalQ, CountryRegionCode, ProductCategoryName
From fact_order natural join (Select TerritoryID, CountryRegionCode From
dim_territory) as dim_Region natural join dim_product
Group by CountryRegionCode, ProductCategoryName
```

7) SQL/OLAP query

```
Use adventureworks_dw;

Select sum(OrderQty) as TotalQ, CountryRegionCode, ProductCategoryName
From fact_order natural join (Select TerritoryID, CountryRegionCode From
dim_territory) as dim_Region natural join dim_product
Group by CountryRegionCode, ProductCategoryName with Rollup;
```

8) PSW

```
<Schema name="adventureworks_dw">
  <Cube name="Orders" visible="true" cache="true" enabled="true">
    <Table name="fact_order">
    </Table>

    <Dimension type="StandardDimension" visible="true" foreignKey="ProductSK" name="Product">
      <Hierarchy name="Product Hierarchy" visible="true" hasAll="true" allMemberName="All Products" primaryKey="ProductSK">
        <Table name="dim_product">
        </Table>
        <Level name="Product Category" visible="true" column="ProductCategoryName" type="String" uniqueMembers="false" levelType="Regular">
          </Level>
        <Level name="Product Subcategory" visible="true" column="ProductSubcategoryName" type="String" uniqueMembers="false" levelType="Regular">
          </Level>
        <Level name="Product Name" visible="true" column="Name" type="String" uniqueMembers="false" levelType="Regular">
          </Level>
        </Hierarchy>
      </Dimension>

      <Dimension type="TimeDimension" visible="true" foreignKey="OrderDate" name="Time">
        <Hierarchy name="Time Hierarchy" visible="true" hasAll="true" allMemberName="All Years" primaryKey="OrderDate">
          <Table name="dim_time">
          </Table>
          <Level name="Year" visible="true" column="Year" type="Integer" uniqueMembers="false" levelType="TimeYears">
            </Level>
          <Level name="Month" visible="true" column="MonthName" ordinalColumn="MonthID" type="String" uniqueMembers="false" levelType="TimeMonths">
            </Level>
          <Level name="Day" visible="true" column="Day" type="Integer" uniqueMembers="false" levelType="TimeDays">
            </Level>
          </Hierarchy>
        </Dimension>

        <Dimension type="StandardDimension" visible="true" foreignKey="TerritoryID" name="Territory">
          <Hierarchy name="Territory Hierarchy" visible="true" hasAll="true" allMemberName="All Territories" primaryKey="TerritoryID">
            <Table name="dim_territory">
            </Table>
            <Level name="Group" visible="true" column="Group" type="String" uniqueMembers="false" levelType="Regular">
              </Level>
            <Level name="Region" visible="true" column="CountryRegionCode" type="String" uniqueMembers="false" levelType="Regular">
              </Level>
            <Level name="Name" visible="true" column="Name" type="String" uniqueMembers="false" levelType="Regular">
              </Level>
            </Hierarchy>
          </Dimension>

          <Measure name="Sales" column="LineTotal" datatype="Numeric" formatString="$ #,###.00" aggregator="sum" visible="true">
          </Measure>
          <Measure name="Quantity" column="OrderQty" datatype="Integer" formatString="#,###" aggregator="sum" visible="true">
          </Measure>
        </Cube>
      </Schema>
```

9) Saiku Queries

a) Sales and quantity by country region code

The screenshot shows the Saiku interface with the following configuration:

- Cubes:** Orders
- Measures:** Sales, Quantity
- Dimensions:** Region

The resulting data table is:

Region	Sales	Quantity
DE	\$ 4,915,407.50	13,143
FR	\$ 7,251,555.65	19,906
GB	\$ 7,470,721.04	20,099
CA	\$ 16,355,770.45	49,381
US	\$ 62,997,590.71	154,092
AU	\$ 10,655,235.96	18,293

b) Sales and quantity by country region code and product category

The screenshot shows the Saiku interface with the following configuration:

- Cubes:** Orders
- Measures:** Sales, Quantity
- Dimensions:** Region, Product Category

The resulting data table is:

Region	Product Category	Accessories	Bikes	Clothing	Components			
DE	\$ 97,330.65	4,810	\$ 4,382,176.51	3,510	\$ 98,113.82	3,524	\$ 337,786.52	1,299
FR	\$ 111,438.51	5,390	\$ 6,114,241.36	5,738	\$ 155,127.44	5,153	\$ 870,748.34	3,625
GB	\$ 119,223.07	5,906	\$ 6,688,589.87	5,940	\$ 151,068.31	5,192	\$ 711,859.79	3,061
CA	\$ 221,505.20	10,763	\$ 13,457,682.98	14,414	\$ 432,112.25	14,808	\$ 2,244,470.02	9,396
US	\$ 559,357.30	27,196	\$ 53,832,611.25	54,638	\$ 1,170,944.85	41,614	\$ 743,097.31	30,644
AU	\$ 182,638.16	7,867	\$ 10,175,870.74	6,028	\$ 113,175.75	3,379	\$ 203,651.31	1,019

c) Sales and quantity by country region code and product category and year

Product Category	Region	Year	Sales	Quantity	Bk	Clothing	Components
Accessories	DE	2001	-	-	\$ 237,784.99	76	-
		2002	-	-	\$ 521,230.85	233	-
		2003	\$ 45,969.26	2,286	\$ 1,856,525.81	1,688	\$ 54,958.03
		2004	\$ 51,361.39	2,524	\$ 1,766,634.86	1,513	\$ 43,195.89
Bk	FR	2001	-	-	-	2,012	\$ 217,126.77
		2002	-	-	-	-	832
		2003	-	-	-	-	467
		2004	-	-	-	-	-
Clothing	FR	2001	-	-	\$ 180,571.69	59	-
		2002	\$ 5,096.23	296	\$ 1,169,180.21	1,071	\$ 27,843.63
		2003	-	-	-	-	900
		2004	-	-	-	-	\$ 169,945.12
Components	FR	2001	-	-	-	-	646
		2002	\$ 52,625.45	2,485	\$ 2,782,408.06	2,791	\$ 79,083.79
		2003	-	-	-	-	2,673
		2004	\$ 53,116.83	2,609	\$ 1,982,080.51	1,817	\$ 48,200.03
GB	Region	2001	-	-	\$ 291,590.52	96	-
		2002	\$ 3,880.07	218	\$ 1,238,275.75	1,016	\$ 24,160.21
		2003	-	-	-	-	800
		2004	\$ 57,558.75	2,798	\$ 2,949,475.84	2,859	\$ 76,227.03
CA	Year	2001	\$ 57,842.25	2,890	\$ 2,209,248.83	1,969	\$ 50,681.07
		2002	\$ 22,495.58	255	\$ 1,517,551.08	1,370	\$ 7,913.33
		2003	-	-	-	-	468
		2004	\$ 51,475.58	1,267	\$ 4,559,886.37	4,789	\$ 115,643.91
US	Region	2001	\$ 101,473.50	4,864	\$ 4,889,111.18	5,057	\$ 198,013.46
		2002	\$ 92,115.11	4,071	\$ 2,426,354.80	2,930	\$ 54,941.55
		2003	\$ 41,253.41	748	\$ 1,425,150.80	1,144	\$ 20,000.00
		2004	\$ 16,643,500.69	3,426	\$ 18,643,500.69	16,940	\$ 217,839.41
AU	Year	2001	\$ 259,397.51	12,269	\$ 18,817,233.02	21,495	\$ 548,366.46
		2002	\$ 224,198.09	10,753	\$ 11,046,110.74	11,088	\$ 278,175.97
		2003	-	-	-	-	9,286
		2004	-	-	-	-	5,323

10) MDX

```

SELECT NON EMPTY Product.[Product Category].Members ON COLUMNS,
CROSSJOIN(Territory.[Region].Members, {Time.Year.[2003], Time.Year.[2004]}) ON ROWS
FROM Orders
WHERE Measures.Quantity
    
```

-- OR

```

SELECT NON EMPTY Product.[Product Category].Members ON COLUMNS,
Territory.[Region].Members ON ROWS
FROM Orders
WHERE {(Time.Year.[2003], Measures.Quantity), (Time.Year.[2004],
Measures.Quantity)}
    
```

Comentário: Seguindo o conselho do docente, apresentamos duas soluções devido à ambiguidade da expressão: “for the years 2003 and 2004”.

11) PRD

Pentaho Report Designer - sales-by-category - /home/aid/Desktop/AID/Proj/AID/PRPT/sales-by-category.prt

Structure

- Master Report
 - Page Header
 - Report Header
 - Ab label: Sales By Category
 - Group Header
 - Details Body
 - Ab
 - D
 - Group Header
 - Details Header
 - D
 - number-field: [Measures].[Sales]
 - text-field: [Product.Product Hierarchy]
 - Group Footer
 - Report Footer
 - Page Footer
 - Watermark

Style

Name	Value	Expr
common	number-field	[Measures].[Sales]
label	label	
value	value	
name	name	
format	format	#,##0.00;(#,##0.00)
full	full	
style-class	style-class	
id	id	
query-meta...	query-meta...	
bar-format	bar-format	
style-format	style-format	
enable-style-bold	enable-style-bold	
enable-style-italics	enable-style-italics	
enable-style-underline	enable-style-underline	
enable-style-italic	enable-style-italic	
enable-style-color	enable-style-color	
enable-style-size	enable-style-size	
enable-style-value	enable-style-value	
enable-style-align	enable-style-align	
wizard	wizard	
aggregation-group	aggregation-group	

Pentaho Report Designer - sales-by-category - /home/aid/Desktop/AID/Proj/AID/PRPT/sales-by-category.prt

Structure

- Master Report
 - Page Header
 - Report Header
 - Ab label: Sales By Category
 - Group Header
 - Details Body
 - Ab
 - D
 - Group Header
 - Details Header
 - D
 - number-field: [Measures].[Sales]
 - text-field: [Product.Product Hierarchy]
 - Group Footer
 - Report Footer
 - Page Footer
 - Watermark

Style

Name	Value	Expr
common	number-field	[Measures].[Sales]
label	label	
value	value	
name	name	
format	format	#,##0.00;(#,##0.00)
full	full	
style-class	style-class	
id	id	
query-meta...	query-meta...	
bar-format	bar-format	
style-format	style-format	
enable-style-bold	enable-style-bold	
enable-style-italics	enable-style-italics	
enable-style-underline	enable-style-underline	
enable-style-italic	enable-style-italic	
enable-style-color	enable-style-color	
enable-style-size	enable-style-size	
enable-style-value	enable-style-value	
enable-style-align	enable-style-align	
wizard	wizard	
aggregation-group	aggregation-group	

