

The study on the
increase in orders on the
products of the past 2
years

Company × Client

Directed by:

Mrs BOUFARES Nadeya

Created by:

Ms MIMOUNI Kamilia



Dataset:



augmentation des commandes

Data Code (0) Discussion (0) Settings

0 New Notebook Download (35 kB)

Product (6 files)

About this directory

This file does not have a description yet.

File Name	Size
README.txt	49 B
customer.csv	4.46 kB
orderitem.csv	45.11 kB
orders.csv	35.11 kB
product.csv	3.77 kB
supplier.csv	2.27 kB

Data Explorer

Version 1 (90.78 kB)

- Product
 - README.txt
 - customer.csv
 - orderitem.csv
 - orders.csv
 - product.csv
 - supplier.csv

Summary

- 6 files
- 29 columns

+ New Version

The link of dataset: <https://www.kaggle.com/datasets/mimounikamilia/product>

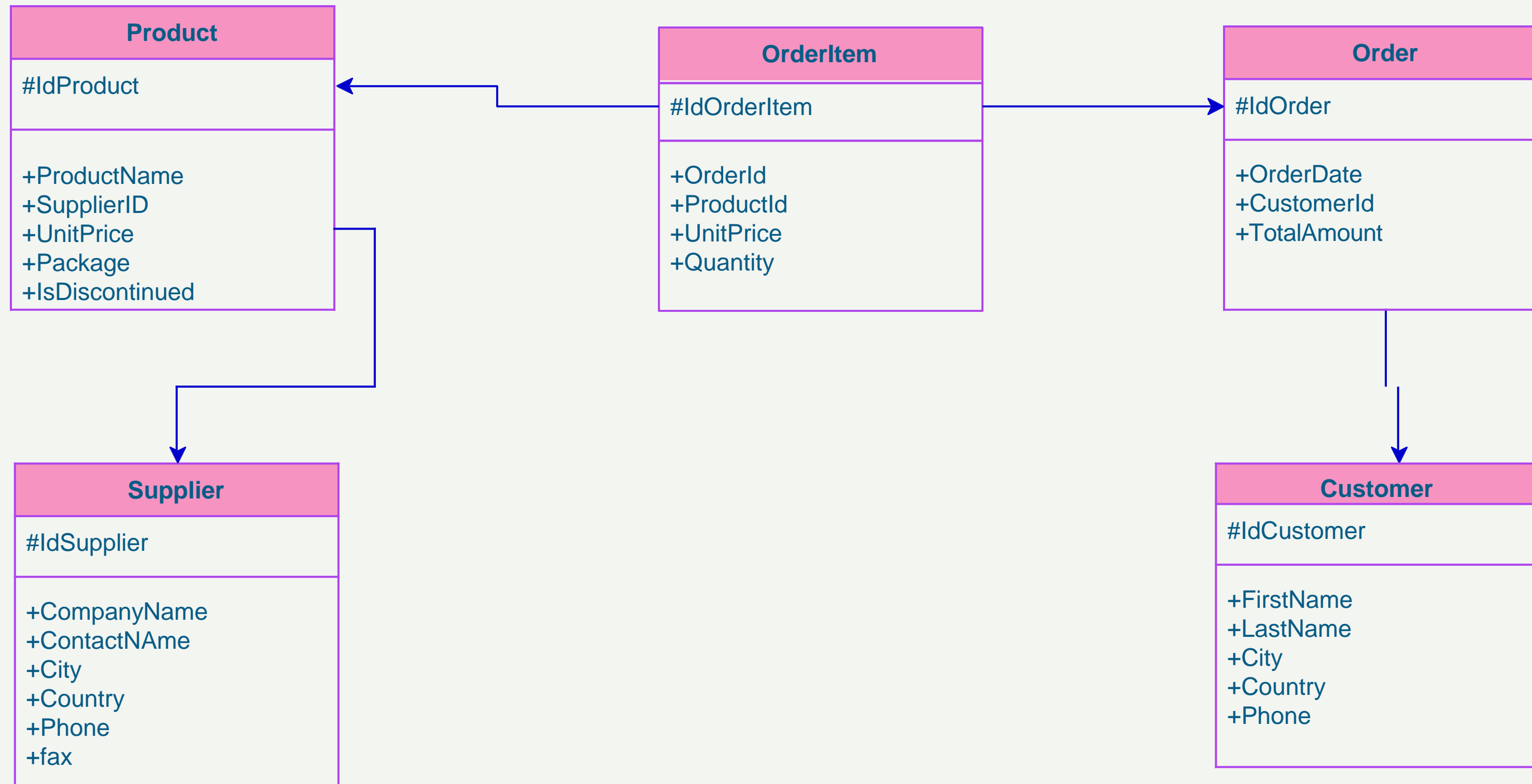
Problem statement

we are going to do a large study on the increase in orders for products from the past 2 years in order to take new strategic directions:

- Set appropriate days to increase advertisements and offer discounts to attract more visitors.
- Rank sellers and offer new features to the most active sellers.
- Maintain the customer and encourage him to buy several times.
- Give more privileges to loyal customers.
- Promote best-selling products for good customer orientation.



Class hierarchy



This diagram shows an **OrderItem** is related to classes representing **Product** and **Order**. The latest is related to **Customer** as well as the **Product** is related to **Supplier**.

The steps how to create a delimited file in Talend

Fichier - Étape 2 de 4

Ajouter un fichier de métadonnées au référentiel
Définissez le chemin d'accès au fichier et les paramètres de format



Paramètres du fichier

Serveur	Localhost 127.0.0.1	
Fichier	C:/Users/hp/Desktop/Product/customer.csv	Parcourir...
Format	WINDOWS	

Visualiseur de fichier

```
Id;FirstName;LastName;City;Country;Phone
1;Maria;Anders;Berlin;Germany;030-0074321
2;Ana;Trujillo;México D.F.;Mexico;(5) 555-4729
3;Antonio;Moreno;México D.F.;Mexico;(5) 555-3932
4;Thomas;Hardy;London;UK;(171); 555-7788
5;Christina;Berglund;Luleå;Sweden;0921-12 34 65
6;Hanna;Moos;Mannheim;Germany;0621-08460
7;Frédérique;Citeaux;Strasbourg;France;88.60.15.31
8;Martín;Sommer;Madrid;Spain;(91); 555 22 82
9;Laurence;Lebihan;Marseille;France;91.24.45.40
10;Elizabeth;Lincoln;Tsawassen;Canada;(604) 555-4729
11;Victoria;Ashworth;London;UK;(171); 555-1212
12;Patricio;Simpson;Buenos Aires;Argentina;(1); 135-5555
13;Francisco;Chang;México D.F.;Mexico;(5) 555-3392
```

Fichier - Étape 1 de 4

⚠ Il est déconseillé de laisser vide le champ Objectif.



Nom	customer	
Objectif		
Description		
Créé par :	user@talend.com	
Verrouillé par		
Version	0.1	M m
Statut		
Chemin d'accès		Sélectionner

Fichier - Étape 4 de 4

Ajouter un schéma au référentiel
Définissez le schéma

Nom

metadata

Commentaire

Schéma

Cliquez pour mettre à jour la prévisualisation du schéma.

Détecter

Description du schéma

Colonne	Clé	Type	<input checked="" type="checkbox"/> N..	Modèle de date (Ctrl...	Longueur	Précision	Par défaut	Commentai...
Id	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
FirstName	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
LastName	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
City	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
Country	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
Phone	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
Column6	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		

Fichier - Étape 3 de 4

Ajouter un fichier de métadonnées au référentiel
Définissez les paramètres du Job de passage

Paramètres du fichier

Encodage

UTF-8

Séparateur de champs

Semicolon

Caractère correspondant

;"

Séparateur de lignes

Standard EO

Caractère correspondant

"\n"

Lignes à ignorer

Si des lignes doivent être ignorées, spécifiez les paramètres suivants

En-tête

☒

1

Pied de page

☐

☐ Ignorer les lignes vides

Paramètres du caractère d'échappement

☐ CSV☒ Délimité

Caractère d'échappement

Vider

Entourage du texte

Vider

☐ Scinder la ligne avant le champ

Limite de lignes

Si le nombre de lignes doit être limité, spécifier ce nombre

Limite

☐

Aperçu

Sortie

☒ Définir la ligne d'en-tête comme nom de colonnes

Actualiser l'aperçu

Id	FirstName	LastName	City	Country	Phone	Colonne 6	
1	Maria	Anders	Berlin	Germany	030-0074321		
2	Ana	Trujillo	México D.F.	Mexico	(5) 555-4729		
3	Antonio	Moreno	México D.F.	Mexico	(5) 555-3932		
4	Thomas	Hardy	London	UK	(171)	555-7788	

Exporter en tant que contexte

Revenir au contexte précédent

1. Create a Job and add a tMap component

The screenshot displays the Talend Studio interface for a job named "Job filtre 0.1".

Left Panel (Project Explorer):

- Jobs: filtre 0.1 (highlighted with a red box)
- Contextes
- Code
- Modèles SQL
- Métadonnées
- Documentation
- Corbeille

Bottom Left Panel (Outline):

- tFileInputDelimited_1 (Product)
- tFileOutputXML_1 (unitPrice<=20)
- tFileOutputXML_2 (20<initPrice<=50)
- tFileOutputXML_3 (50<unitPrice)
- tMap_1

Designer View:

The job flow is as follows:

- Product** (tFileInputDelimited_1) outputs 78 rows in 0,02s at 3900 rows/s.
- tMap_1** splits the data into three paths:
- inf20 (Main order:1)** outputs 41 rows in 0,02s at 1708,33 rows/s to **unitPrice<=20** (tFileOutputXML_1).
- inf50 (Main order:2)** outputs 30 rows in 0,02s at 1250 rows/s to **20<initPrice<=50** (tFileOutputXML_2).
- sup50 (Main order:3)** outputs 7 rows in 0,02s at 304,35 rows/s to **50<unitPrice** (tFileOutputXML_3).

Bottom Right Panel (Job filtre):

Exécution simple

- Exécution en mode Debug
- Paramètres avancés
- Cible d'exécution
- Exécution pour la mémoire

Execution Log:

```
Démarrage du Job filtre à 08:44 28/07/2022.  
[statistics] connecting to socket on port 3356  
[statistics] connected  
[statistics] disconnected  
Job filtre terminé à 08:44 28/07/2022. [Code de sortie =
```


2. Add a filter to select only products that have unit prices less than 20.

The screenshot displays a data transformation tool interface with three main panels: 'Product', 'Var', and 'inf20'. The 'Product' panel lists columns: Id, ProductName, SupplierId, UnitPrice, Package, and IsDiscontinued. The 'Var' panel is empty. The 'inf20' panel shows a filter expression 'Product.UnitPrice <= 20' and a table of columns: Product.Id, Product.ProductName, and Product.UnitPrice. Below this, the 'inf50' panel shows a filter expression 'Product.UnitPrice > 20 && Product.UnitPrice < 50' and the same column table. The 'sup50' panel shows a filter expression 'Product.UnitPrice > 50' and the same column table. A red arrow points from a file explorer window to the 'inf20' filter. The file explorer shows the path 'Ce PC > Windows (C:) > Programmes (x86) > TOS_DI-8.0.1 > studio > workspace' and lists files: .Java, .JETEmitters, .metadata, temp, TP1, inf20, inf50, and sup50. The files inf20, inf50, and sup50 are highlighted with a red box.

Column
Id
ProductName
SupplierId
UnitPrice
Package
IsDiscontinued

Expression	Column
Product.Id	Id
Product.ProductName	ProductName
Product.UnitPrice	UnitPrice

Expression	Column
Product.Id	Id
Product.ProductName	ProductName
Product.UnitPrice	UnitPrice

Expression	Column
Product.Id	Id
Product.ProductName	ProductName
Product.UnitPrice	UnitPrice

Nom	Modifié le	Type	Taille
.Java	25/07/2022 09:50	Dossier de fichiers	
.JETEmitters	25/07/2022 09:17	Dossier de fichiers	
.metadata	25/07/2022 09:39	Dossier de fichiers	
temp	25/07/2022 09:16	Dossier de fichiers	
TP1	25/07/2022 09:16	Dossier de fichiers	
inf20	28/07/2022 08:44	Microsoft Edge H...	5 Ko
inf50	28/07/2022 08:44	Microsoft Edge H...	4 Ko
sup50	28/07/2022 08:44	Microsoft Edge H...	1 Ko

3. Create a second filter with two conditions, products that have unit prices strictly greater than 20 and less than 50 then add a 3rd filter to select only products that have unit prices greater than 50.

4. Store the results in an XML file (add a tFileOutputXML component to the Job).

Products that have unit prices below 20.

This XML file does not appear to have any style information associated with it. The document tree is shown

```
▼<root>
  ▼<unit>
    <Id>1</Id>
    <ProductName>Chai</ProductName>
    <UnitPrice>18.0</UnitPrice>
  </unit>
  ▼<unit>
    <Id>2</Id>
    <ProductName>Chang</ProductName>
    <UnitPrice>19.0</UnitPrice>
  </unit>
  ▼<unit>
    <Id>3</Id>
    <ProductName>Aniseed Syrup</ProductName>
    <UnitPrice>10.0</UnitPrice>
  </unit>
  ▼<unit>
    <Id>13</Id>
    <ProductName>Konbu</ProductName>
    <UnitPrice>6.0</UnitPrice>
  </unit>
  ▼<unit>
    <Id>15</Id>
    <ProductName>Genen Shouyu</ProductName>
    <UnitPrice>15.5</UnitPrice>
  </unit>
  ▼<unit>
    <Id>16</Id>
    <ProductName>Pavlova</ProductName>
    <UnitPrice>17.45</UnitPrice>
  </unit>
  ▼<unit>
    <Id>19</Id>
```

Products with unit prices strictly above 20 and below 50.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

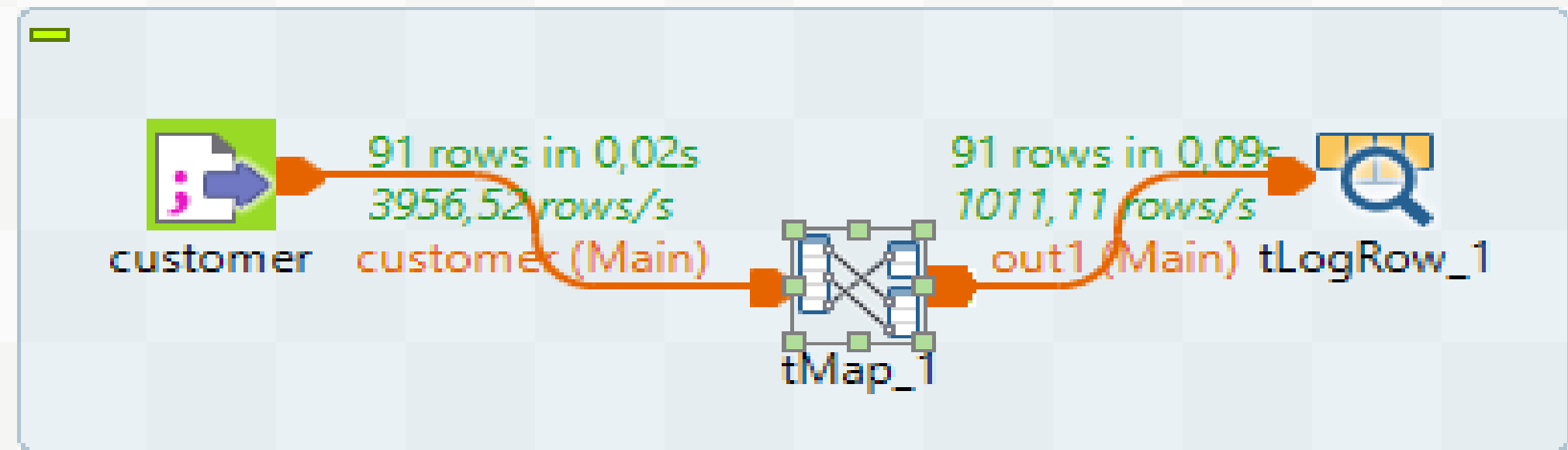
```
▼<root>
  ▼<unit1>
    <Id>4</Id>
    <ProductName>Chef Anton's Cajun Seasoning</ProductName>
    <UnitPrice>22.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>5</Id>
    <ProductName>Chef Anton's Gumbo Mix</ProductName>
    <UnitPrice>21.35</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>6</Id>
    <ProductName>Grandma's Boysenberry Spread</ProductName>
    <UnitPrice>25.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>7</Id>
    <ProductName>Uncle Bob's Organic Dried Pears</ProductName>
    <UnitPrice>30.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>8</Id>
    <ProductName>Northwoods Cranberry Sauce</ProductName>
    <UnitPrice>40.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>10</Id>
    <ProductName>Ikura</ProductName>
    <UnitPrice>31.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>11</Id>
    <ProductName>Queso Cabrales</ProductName>
    <UnitPrice>21.0</UnitPrice>
  </unit1>
  ▼<unit1>
    <Id>12</Id>
    <ProductName>Queso Manchego La Pastora</ProductName>
    <UnitPrice>38.0</UnitPrice>
  </unit1>
```


Products that have unit prices greater than 50.

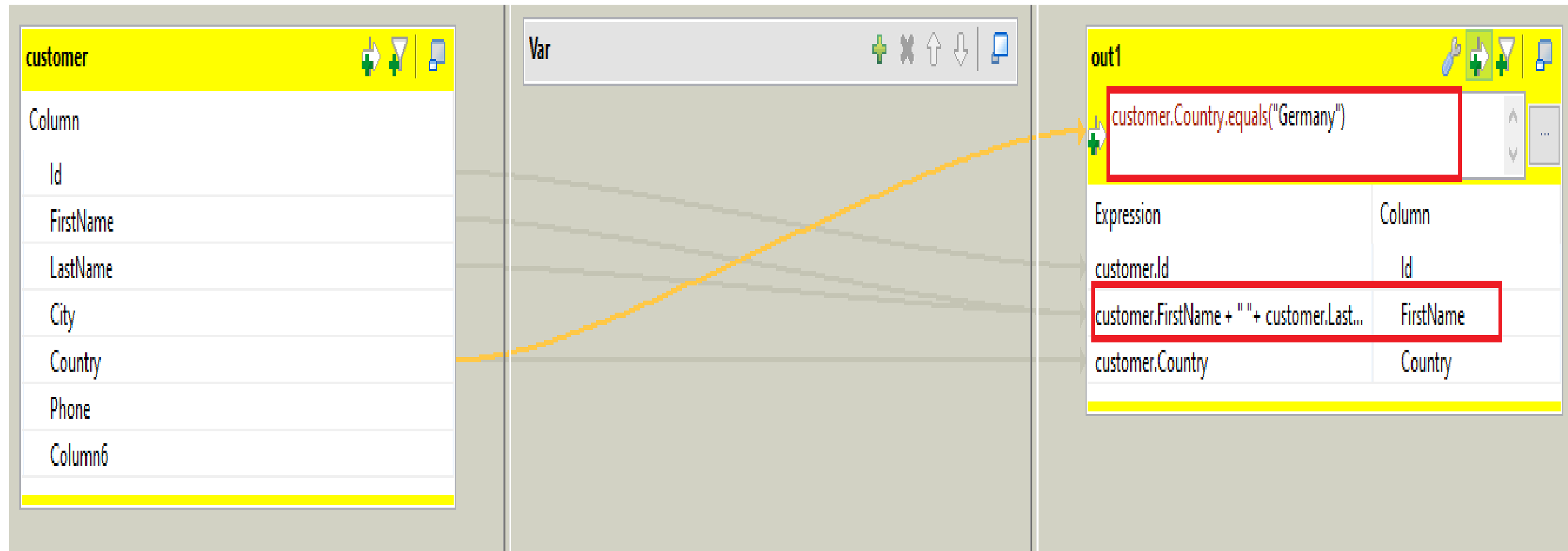
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0"?>
<root>
  <unit2>
    <Id>9</Id>
    <ProductName>Mishi Kobe Niku</ProductName>
    <UnitPrice>97.0</UnitPrice>
  </unit2>
  <unit2>
    <Id>18</Id>
    <ProductName>Carnarvon Tigers</ProductName>
    <UnitPrice>62.5</UnitPrice>
  </unit2>
  <unit2>
    <Id>20</Id>
    <ProductName>Sir Rodney's Marmalade</ProductName>
    <UnitPrice>81.0</UnitPrice>
  </unit2>
  <unit2>
    <Id>29</Id>
    <ProductName>Thüringer Rostbratwurst</ProductName>
    <UnitPrice>123.79</UnitPrice>
  </unit2>
  <unit2>
    <Id>38</Id>
    <ProductName>Côte de Blaye</ProductName>
    <UnitPrice>263.5</UnitPrice>
  </unit2>
  <unit2>
    <Id>51</Id>
    <ProductName>Manjimup Dried Apples</ProductName>
    <UnitPrice>53.0</UnitPrice>
  </unit2>
  <unit2>
    <Id>59</Id>
    <ProductName>Raclette Courdavault</ProductName>
    <UnitPrice>55.0</UnitPrice>
  </unit2>
</root>
```


1. Create a Job, add Customers metadata as input source and add a tMap_1 component



2. Configure the tMap_1 component to filter customers residing in Germany



**3. we only find
customers who
reside in germany**

Démarrage du Job collection2 à 16:27 28/07/2022.

[statistics] connecting to socket on port 3368

[statistics] connected

tLogRow_1		
Id	FirstName	Country
1	Maria Anders	Germany
6	Hanna Moos	Germany
17	Sven Ottlieb	Germany
25	Peter Franken	Germany
39	Philip Cramer	Germany
44	Renate Messner	Germany
52	Alexander Feuer	Germany
56	Henriette Pfalzheim	Germany
63	Horst Kloss	Germany
79	Karin Josephs	Germany
86	Rita Müller	Germany

[statistics] disconnected

Job collection2 terminé à 16:27 28/07/2022. [Code de sortie

Creating a database connection

Nom	<input type="text" value="DataWarehouse"/>		
Objectif	<input type="text"/>		
Description	<div><div></div><div></div></div>		
Créé par :	<input type="text" value="user@talend.com"/>		
Verrouillé par	<input type="text"/>		
Version	<input type="text" value="0.1"/>	<input type="button" value="M"/>	<input type="button" value="m"/>
Statut	<div><div></div><div></div></div>		
Chemin d'accès	<input type="text"/>	<input type="button" value="Sélectionner"/>	

Nouvelle connexion à la base de données dans le référentiel - Étape 2/2

⚠ Le serveur de la base de données doit être spécifié



Type de Bdd MySQL

Version de la base de données

MySQL 8

Chaîne de caractères de connexion

jdbc:mysql://:3306/order?noDatetimeStringSync=true&enabledTLSProtocols=TLSv1.2,TLSv1.1,TLSv1

Connexion

root

Mot de passe

Serveur

Port

3306

Base de données

order

Paramètres supplémentaires

noDatetimeStringSync=true&enabledTLSProtocols=TLSv1.2,TLSv1.1,TLSv1

Tester la connexion

Exporter en tant que contexte

Revenir au contexte précédent

Nouvelle connexion à la base de données dans le référentiel - Étape 2/2

Définir les paramètres de connexion



Type de BddMySQL

Version de la base de donnéesMySQL 8

Chaîne de caractères de connexionjdbc:mysql:///3306/order?noDatetimeStringSync=true&enabledTLSProtocols=TLSv1.2,TLSv1.1,TLSv1.0

Connexionroot

Mot de passe

Serveur

Port


Base de données

Paramètres supplémentaires

OK

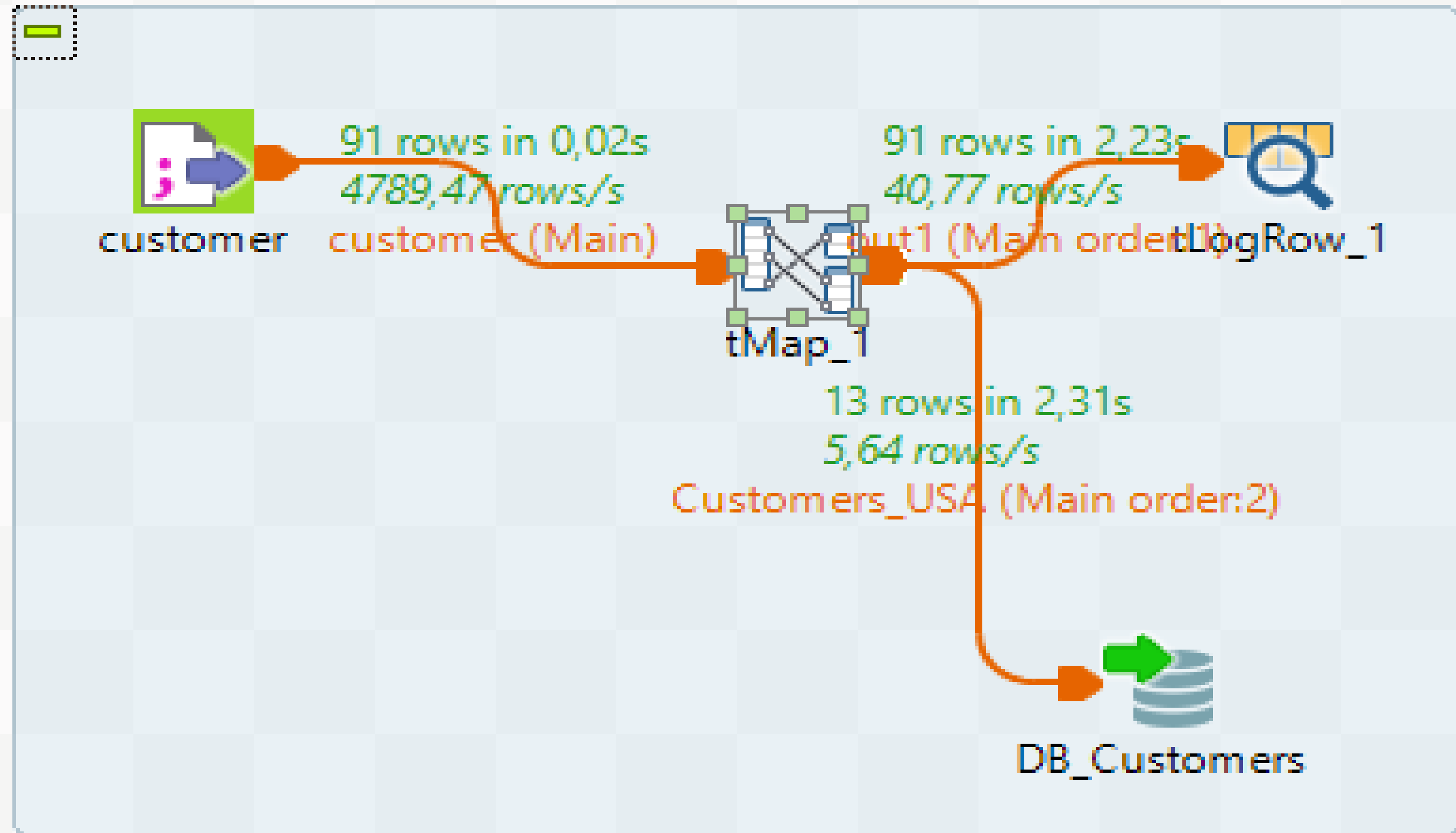
Tester la connexion

Vérifier la connexion




 "DataWarehouse" connexion établie.













OK

Create a Job, add Customers metadata as input source and add a tMap_1 component and TDBOutput

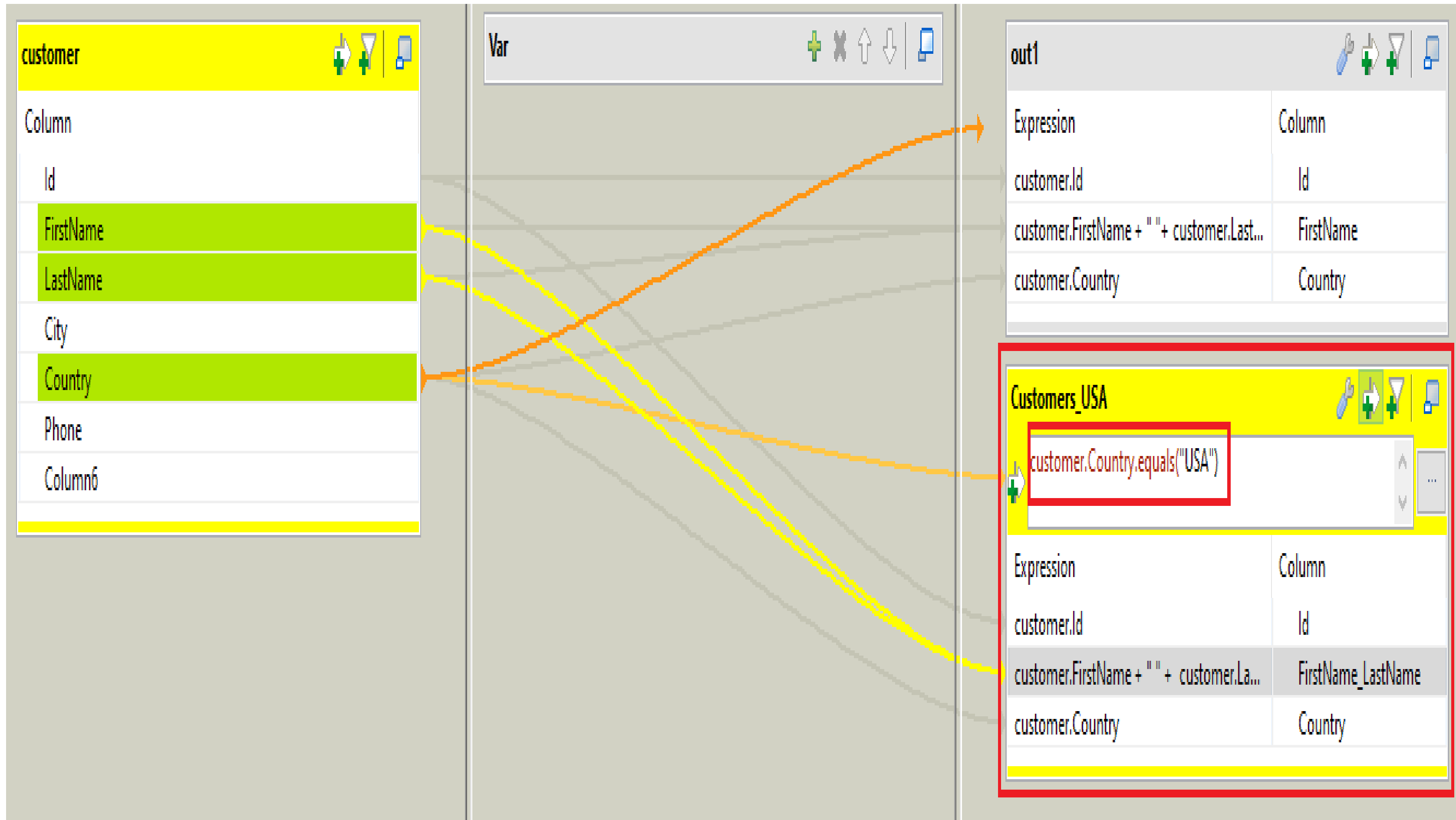


Configure TDBOutput as follow

 **DB_Customers(tDBOutput_1)(MySQL)**  

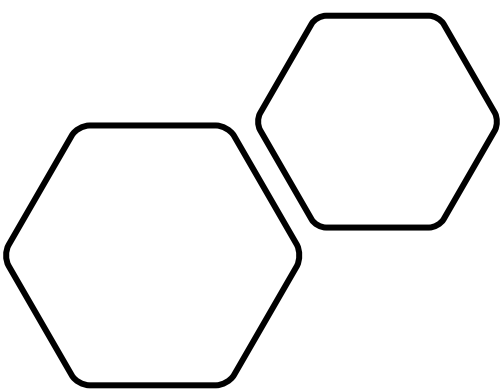
Paramètres simples	Database MySQL  Appliquer
Paramètres avancés	Type de propriété Référentiel  Bases de données (MYSQL):DataWarehouse 
Paramètres dynamiques	Version de la base de données MySQL 8  
Vue	<input type="checkbox"/> Utiliser une connexion existante
Documentation	Hôte ""  * Port "3306" 
	Base de données "order"
	Utilisateur "root"  * Mot de passe ***** 
	Table "Customer_USA"
	Action sur la table Supprimer la table si elle existe et la créer  Action sur les données Insert
	Schéma Built-in  Modifier le schéma  Sync colonnes
	Source de données This option only applies when deploying and running in the Talend Runtime <input type="checkbox"/> Spécifier un alias de source de données
	<input type="checkbox"/> Arrêter en cas d'erreur

Tmap

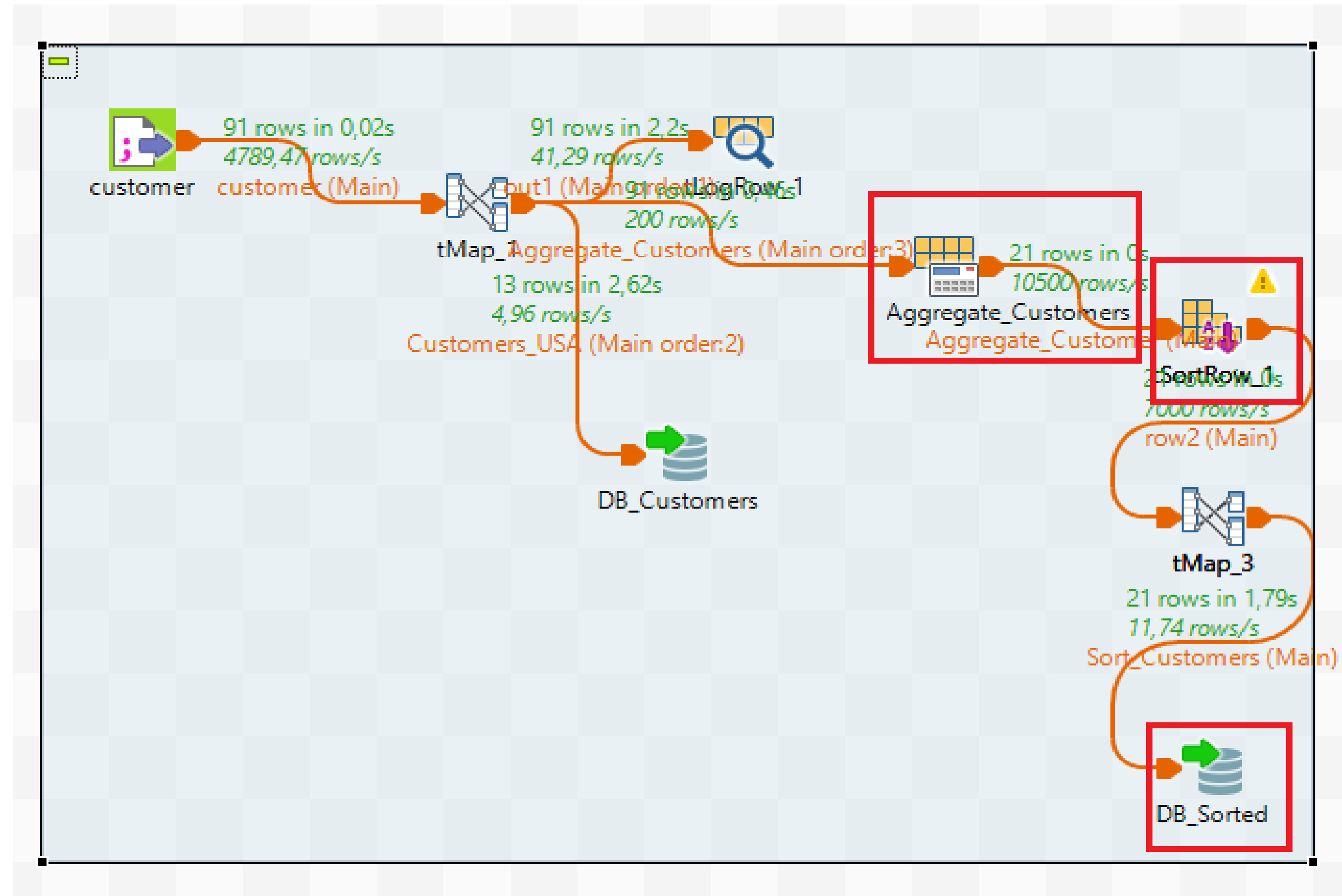


MySQL Output

Id	FirstName_LastName	Country
32	Howard Snyder	USA
36	Yoshi Latimer	USA
43	John Steel	USA
45	Jaime Yorres	USA
48	Fran Wilson	USA
55	Rene Phillips	USA
65	Paula Wilson	USA
71	Jose Pavarotti	USA
75	Art Braunschweiger	USA
77	Liz Nixon	USA
78	Liu Wong	USA
82	Helvetius Nagy	USA
89	Karl Jablonski	USA

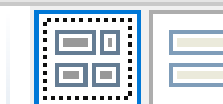


- Create a Job, add Customers metadata as input source and add a tMap_1 component, tAggregation, Tsort and TDBOutput
- Configure TAggregateRow, Tsort and TDBOutput as follow





Aggregate_Customers(tAggregateRow_1)



Paramètres simples

Paramètres avancés

Paramètres dynamiques

Vue

Documentation

Schéma

Built-in



Modifier le schéma



Sync colonnes

Group by

Colonne de sortie

Position de la colonne d'entrée

Country

Country



Opérations

Colonne de sortie

Fonction

Position de la colonne ...



Igr

FirstName

liste

FirstName

LastName

liste

LastName



tSortRow_1

Paramètres simples

Paramètres avancés

Paramètres dynamiques

Vue

Documentation

Built-in

Modifier le schéma

Sync colonnes

Colonne du schéma

tri num ou alpha ?

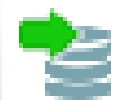
Ordre asc ou desc ?

FirstName

alpha

asc





DB_Sorted(tDBOutput_2)(MySQL)



Paramètres simples

Paramètres avancés

Paramètres dynamiques

Vue

Documentation

Database

MySQL



Appliquer

Type de propriété

Référentiel



Bases de données (MYSQL):DataWarehouse



Version de la base de données

MySQL 8



☐ Utiliser une connexion existante

Hôte

""



Port

"3306"

Base de données

"order"

Utilisateur

"root"



Mot de passe

Table

"Sort_Customers"

Action sur la table

Supprimer la table si elle existe et la créer



Action sur les données

Insert

Schéma

Built-in



Modifier le schéma



Sync colonnes

Source de données

This option only applies when deploying and running in the Talend Runtime

☐ Spécifier un alias de source de données

☐ Arrêter en cas d'erreur

MySQL Output

Récentes

Préférées

Nouvelle base de données

information_schema

mysql

order

Nouvelle table

customer_germany

customer_usa

sort_customers

orderitem

Nouvelle table

customer

date_dim

orderitem_dim

product_dim

supplier_dim

performance_schema

phpmyadmin

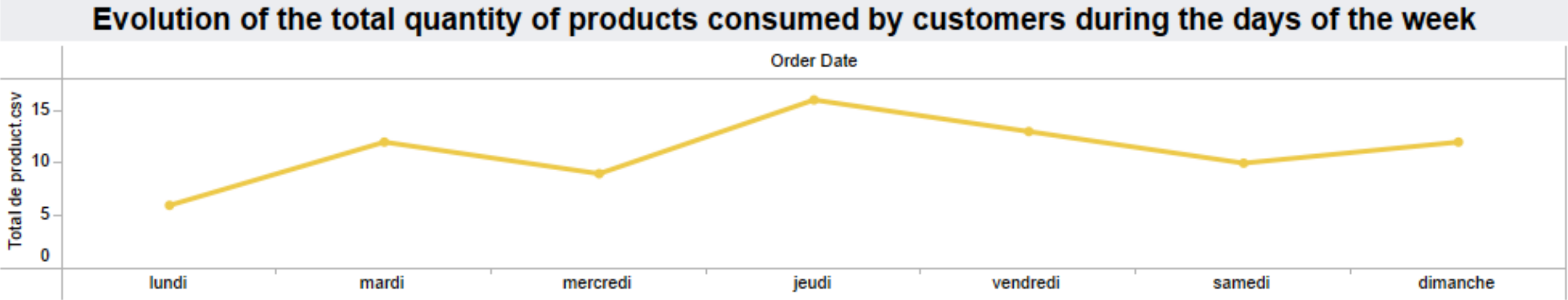
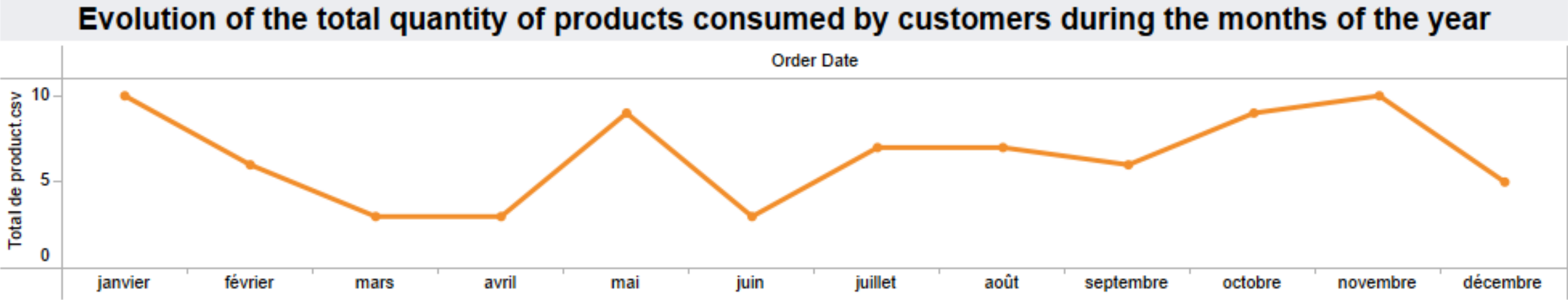
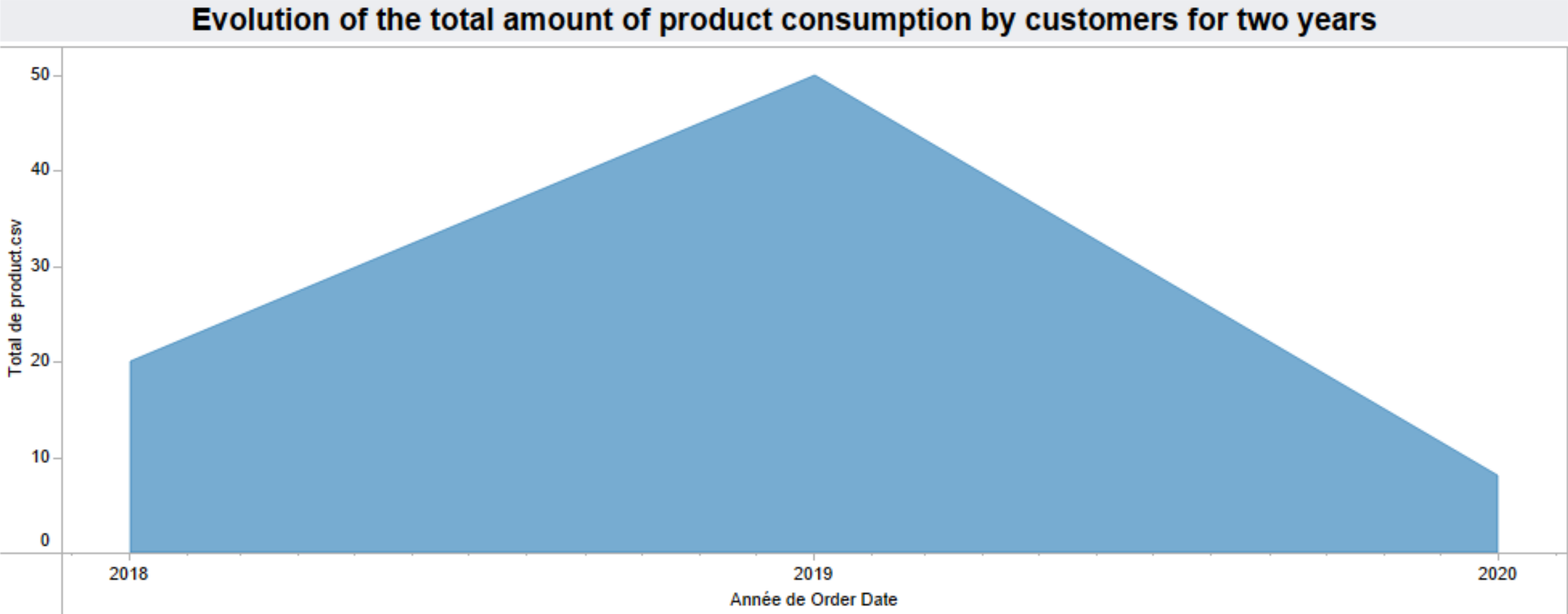
test

FirstName
ANA,ANTONIO,FRANCISCO,GUILLERMO,MIGUEL
CATHERINE,PASCALE
CHRISTINA,MARIA
ELIZABETH,YOSHI,JEAN
FRÉDÉRIQUE,LAURENCE,JANINE,MARTINE,CARINE,DANIEL,A...
HOWARD,YOSHI,JOHN,JAIME,FRAN,RENE,PAULA,JOSE,ART,L...
JONAS
JYTTE,PALLE
LINO,ISABEL
MANUEL,CARLOS,CARLOS,FELIPE
MARIA,HANNA,SVEN,PETER,PHILIP,RENATE,ALEXANDER,HEN...
MARTÍN,DIEGO,EDUARDO,JOSÉ,ALEJANDRA
PAOLO,GIOVANNI,MAURIZIO
PATRICIA
PATRICIO,YVONNE,SERGIO
PEDRO,ARIA,ANDRÉ,MARIO,BERNARDO,LÚCIA,JANETE,ANABE...
PIRKKO,MATTI
ROLAND,GEORG

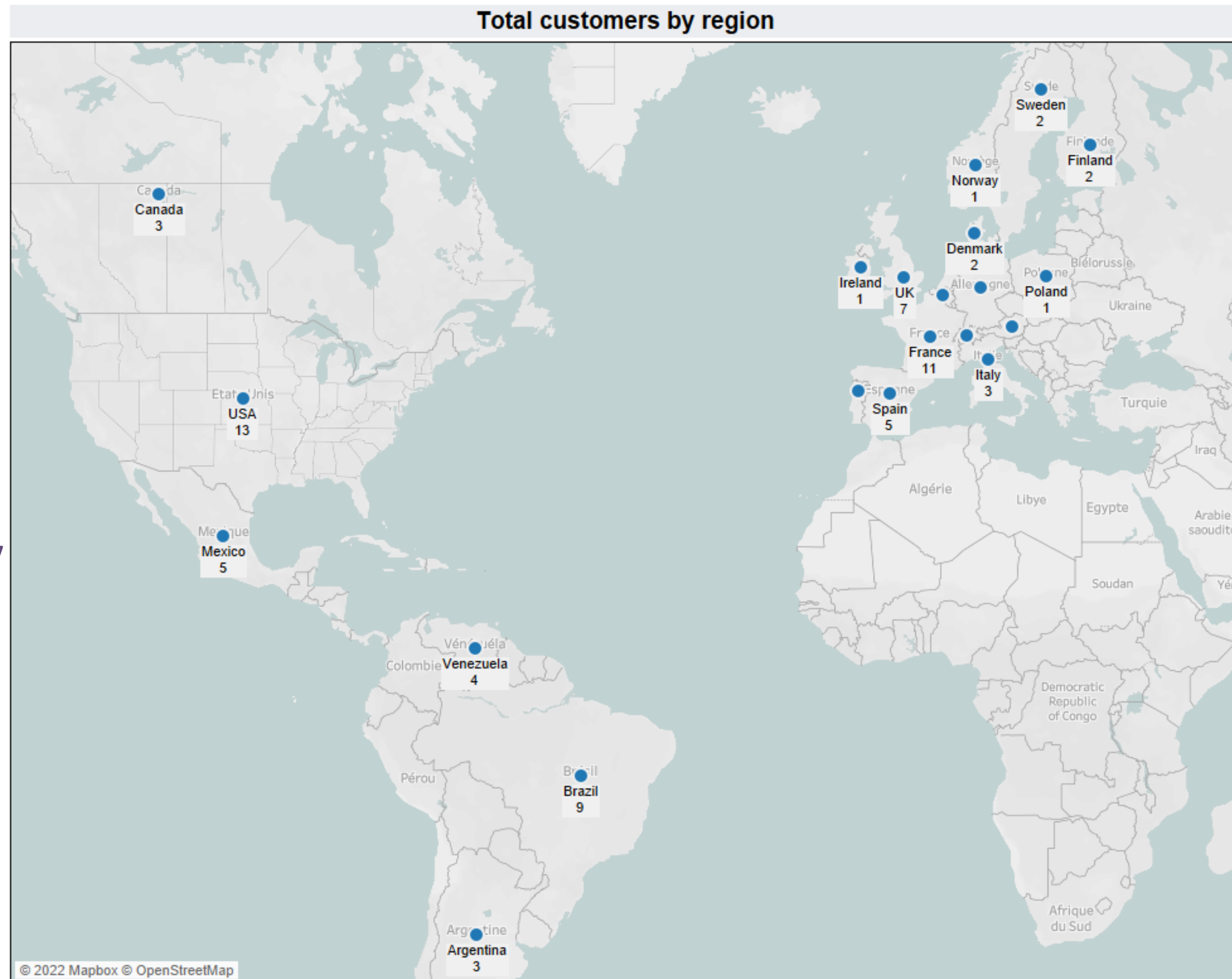
Visualisation



- We note that in 2019 is the peak of sales of products
- In May and November the evolution of sales is exponential, so it is necessary to make promotions and special offers in June, March and April to maintain the customer and encourage him to buy several times and of course give more privileges to loyal customers .
- During the weekend, the consumption rate is decreasing, so we see that on break days it is necessary to make special weekend offers



❑ The number of customers in poland, ireland and norway is very low, so special offers must be made in these countries to encourage them to buy products from other countries.



- ❑ we have classified sellers by region thus we will offer new features to the most active sellers who are located especially in France and USA.

