

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté d'Informatique



Rapport de Projet
Module : Business Intelligence

Analyse de la base de données NorthWind

Spécialité : Big Data Analytics

Master 2 / Group : 2

Réaliser par :

- ACHACHI Elhadj-Ali
- YAHIA mohamed amine

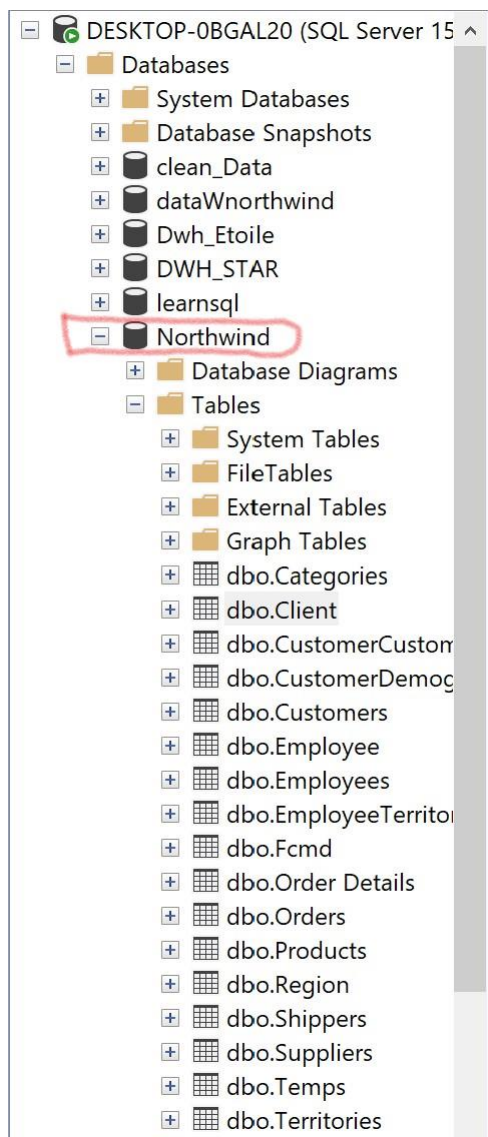
1. Introduction :

La Business Intelligence (BI) ou l'Informatique Décisionnelle permet à tous les acteurs d'une entreprise d'avoir une vision globale sur leurs activités pour les aider dans leurs prises de décisions. Ce projet a pour objectif la réalisation d'une solution Bi permettant d'analyser l'import et l'export des données de la société Northwind Traders.

Les données sont stockées dans deux bases de données relationnelles différentes, la base SQL Server et la base Access, dans ce projet on va utiliser : SQL Server pour l'alimentation de la base , Talend open source pour l'intégration de données.

2. Réalisation :

- Installions de la base de données Northwind sous SQL server :



- La vérification de l'état de SQLserver est en cour :

Gestionnaire de configuration SQL Server (Local)		
Services SQL Server	Nom	État
Configuration réseau SQL Server (32 bits)	SQL Server (SQLEXPRESS)	En cours d'exécution
Configuration de SQL Native Client 11.0	SQL Server Agent (SQLEXPRESS)	Arrêté
Configuration du réseau SQL Server	SQL Server Browser	Arrêté

- La requête : (nombre des commandes livrer , non_livrer par clientid , client nom ,employeeID,nom ,prenom, mois , annee)

```

select MONTH(a.OrderDate) as mois, YEAR(a.OrderDate) as annee, b.CustomerID, b.CompanyName,b.City,
b.Country, c.EmployeeID,c.LastName , c.FirstName, c.City, c.Country,
sum(case when a.ShippedDate IS NULL then 1 else 0 end) commande_non_livrer, count(a.ShippedDate) commande_livrer
from Orders a, Customers b, Employees c
where a.CustomerID = b.CustomerID and a.EmployeeID = c.EmployeeID
group by month(a.OrderDate) , YEAR(a.OrderDate), b.CustomerID, b.CompanyName, b.City, b.Country, c.LastName,
c.EmployeeID,c.FirstName,c.City,c.Country

```

119 %

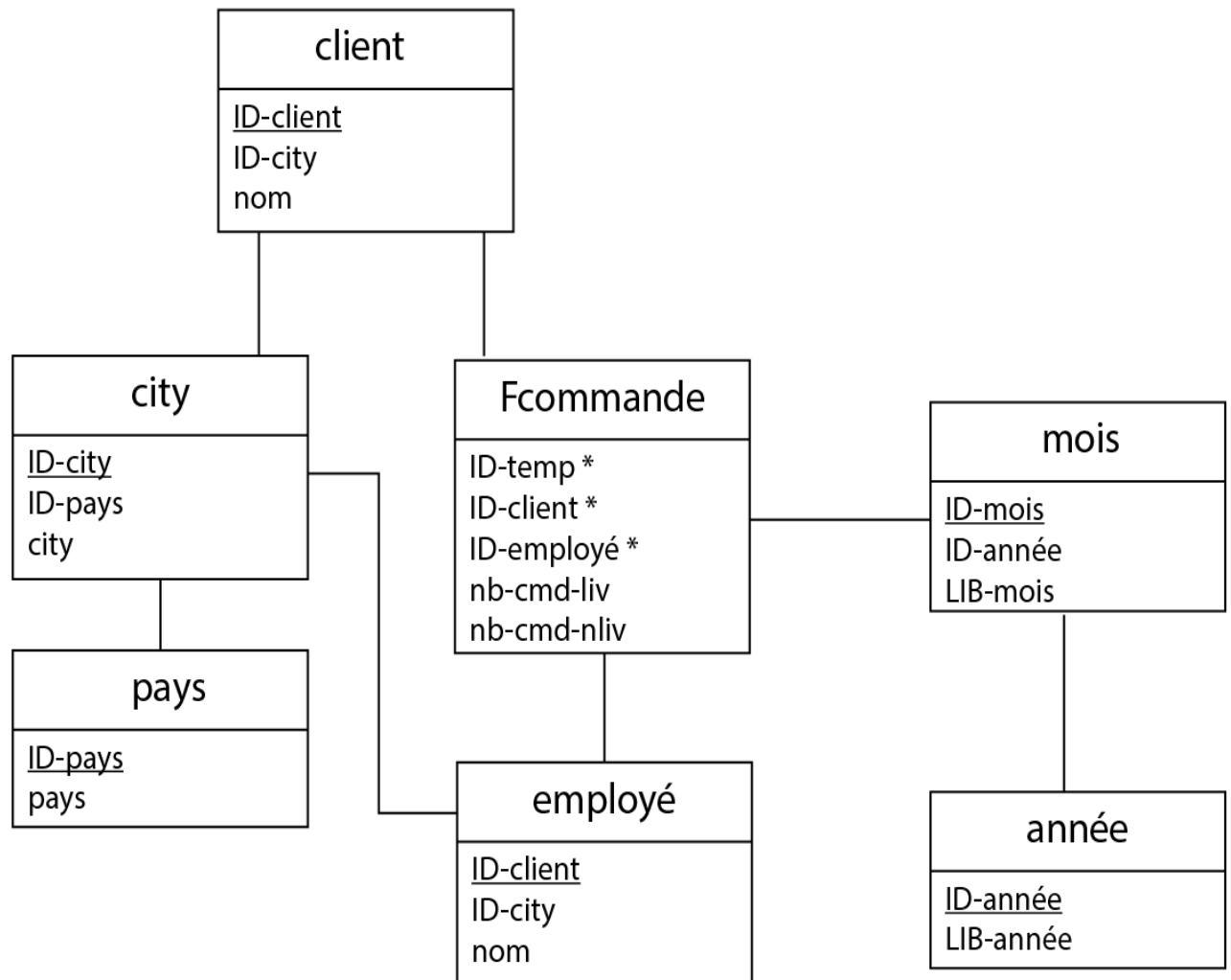
Results Messages

	mois	annee	CustomerID	CompanyName	City	Country	EmployeeID	LastName	FirstName	City	Country	commande_non_livrer	commande_livrer
786	12	1997	LILAS	LILA-Supermercado	Barquisimeto	Venezuela	2	Fuller	Andrew	Tacoma	USA	0	1
787	12	1997	MAGAA	Magazzini Alimentari Riuniti	Bergamo	Italy	4	Peacock	Margaret	Redmond	USA	0	1
788	12	1997	MAISD	Maison Dewey	Bruxelles	Belgium	4	Peacock	Margaret	Redmond	USA	0	1
789	12	1997	MORGK	Morgenstern Gesundkost	Leipzig	Germany	3	Leverling	Janet	Kirkland	USA	0	1
790	12	1997	OTTIK	Ottlies Käseladen	Köln	Germany	4	Peacock	Margaret	Redmond	USA	0	1
791	12	1997	QUEDE	Que Delicia	Rio de Jane...	Brazil	6	Suyama	Michael	London	UK	0	1
792	12	1997	QUEEN	Queen Cozinha	Sao Paulo	Brazil	8	Callahan	Laura	Seattle	USA	0	1
793	12	1997	QUICK	QUICK-Stop	Cunewalde	Germany	1	Davolio	Nancy	Seattle	USA	0	1
794	12	1997	QUICK	QUICK-Stop	Cunewalde	Germany	3	Leverling	Janet	Kirkland	USA	0	1
795	12	1997	RATTC	Rattlesnake Canyon Grocery	Albuquerque	USA	5	Buchanan	Steven	London	UK	0	1
796	12	1997	SEVES	Seven Seas Imports	London	UK	1	Davolio	Nancy	Seattle	USA	0	1
797	12	1997	SEVES	Seven Seas Imports	London	UK	6	Suyama	Michael	London	UK	0	1
798	12	1997	SIMOB	Simons bistro	Kobenhavn	Denmark	4	Peacock	Margaret	Redmond	USA	0	1
799	12	1997	SUPRD	Suprêmes délices	Charleroi	Belgium	4	Peacock	Margaret	Redmond	USA	0	1
800	12	1997	THEBI	The Big Cheese	Portland	USA	2	Fuller	Andrew	Tacoma	USA	0	1
801	12	1997	THECR	The Cracker Box	Butte	USA	7	King	Robert	London	UK	0	1
802	12	1997	VAFFE	Vaffeljernet	Århus	Denmark	3	Leverling	Janet	Kirkland	USA	0	1
803	12	1997	VICTE	Victualies en stock	Lyon	France	3	Leverling	Janet	Kirkland	USA	0	1
804	12	1997	WARTH	Wartian Herikku	Oulu	Finland	2	Fuller	Andrew	Tacoma	USA	0	1
805	12	1997	WELLI	Wellington Importadora	Resende	Brazil	4	Peacock	Margaret	Redmond	USA	0	1
806	12	1997	WOLZA	Wolski Zajazd	Warszawa	Poland	1	Davolio	Nancy	Seattle	USA	0	1

Query executed successfully.

DESKTOP-0BGAL20 (15.0 RTM) DESKTOP-0BGAL20\Surfac... Northwind 00:00:00 806 rows

- La conception de la datawarehouse en flocon (SnowFlake) :



- La conception et la creation de la Datawarehouse en Etoile (Star warehouse) sous SQLserver :
 - Les tables dimensions sont : - DClient - DEmployee - DTemps
 - Table Fait (Fait_command) avec tous les clés primaire d'autre tables + nb_livrer et nb_non_livrer

```
SQLQuery1.sql - DE...AL20\Surface (51))*  X
-- creer la data warehouse
-- utiliser datawarehouse

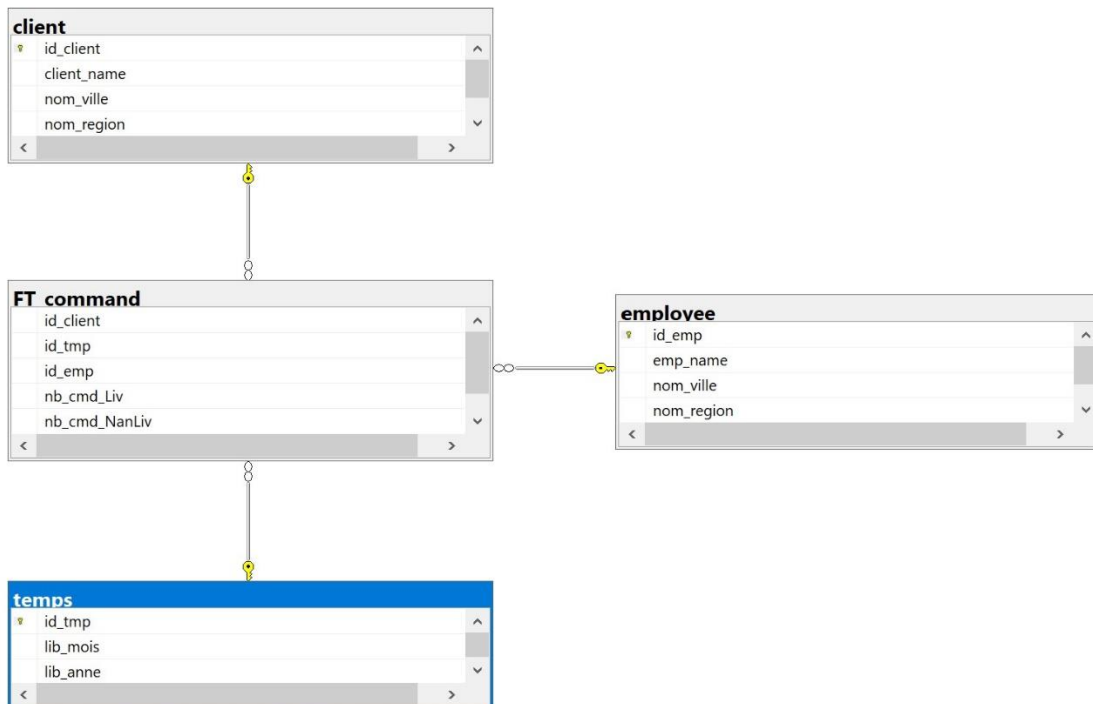
-- creation les dimensions et la table Fait

create table Client
(id_client varchar(10) NOT NULL, nom_cl varchar(10) ,
id_ville varchar(10) , nom_vil varchar(20), id_region varchar(10) ,
nom_region varchar(20))
GO

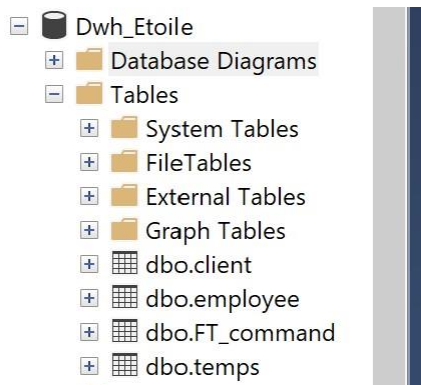
create table Employee
(id_emp varchar(10) NOT NULL, nom_emp varchar(10) ,
id_ville varchar(10) , nom_vil varchar(20), id_region varchar(10) ,
nom_region varchar(20))
GO

create table Temps
(id_tmp varchar(10) NOT NULL, id_mois varchar(10) ,
lib_mois varchar(10) , id_anne varchar(10), lib_anne varchar(20))
GO

create table Fcmd
(id_tmp varchar(10) NOT NULL,
id_emp varchar(10) NOT NULL,
id_client varchar(10) NOT NULL,
nb_cmd_liv varchar(20),
nb_cmd_nliv varchar(20)
)
GO
alter table Temps add constraint pk_tmp primary key (id_tmp)
GO
alter table Client add constraint pk_cl primary key (id_client)
GO
alter table Employee add constraint pk_emp primary key (id_emp)
GO
alter table Fcmd add
constraint fk_tmp foreign key (id_tmp) references Temps(id_tmp),
constraint fk_cl foreign key (id_client) references Client(id_client),
constraint fk_emp foreign key (id_emp) references Employee(id_emp)
```



- Création du la DataWarehouse :



- La connexion de la base de données Northwind avec Talend :



Nouvelle connexion à une base de données dans le référentiel - Etape2/2

Définir les paramètres de connexion



Type de BdD Microsoft SQL Server

Version de la base de données Open source JTDS

Chaîne de caractères de connexion jdbc:jtds:sqlserver://127.0.0.1:1433/Northwind;

Identifiant

Mot de passe

Serveur

Port

Base de données

Schéma

Paramètres supplémentaires

127.0.0.1

1433

Northwind

Vérifier la connexion



"SQLserverr" Connexion établie.

OK

Tester la connexion

Exporter en tant que contexte

Revenir au contexte précédent

[Comment installer un pilote](#)

< Back

Next >

Finish

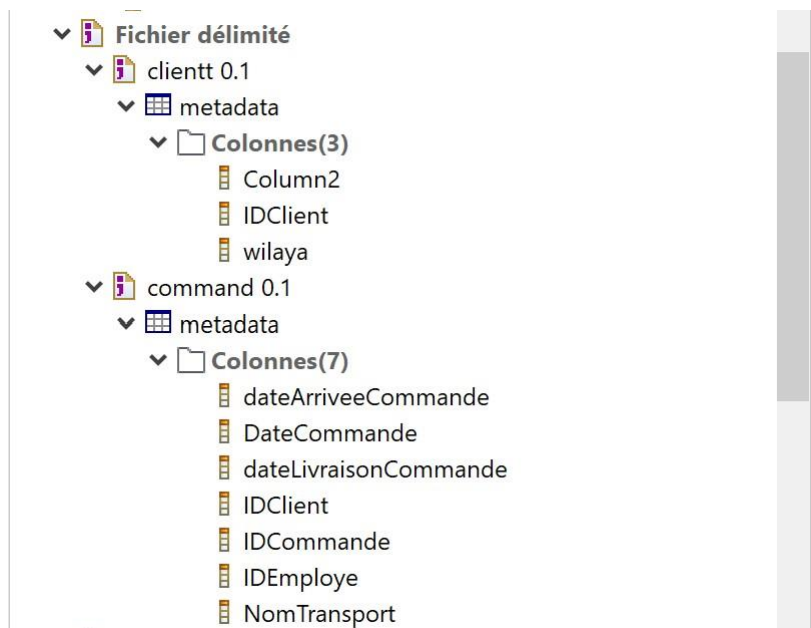
Cancel

- Récupérer les tables qui nous allons besoin pour l'ETL :

Nom	Type	Nombre de colonnes	Statut de création
▼ <input checked="" type="checkbox"/> dbo	SCHEMA		
<input type="checkbox"/> Categories	TABLE		
<input type="checkbox"/> Client	TABLE		
<input type="checkbox"/> CustomerCustomerDemo	TABLE		
<input type="checkbox"/> CustomerDemographics	TABLE		
<input checked="" type="checkbox"/> Customers	TABLE		
<input type="checkbox"/> Employee	TABLE		
<input checked="" type="checkbox"/> Employees	TABLE		
<input type="checkbox"/> EmployeeTerritories	TABLE		
<input type="checkbox"/> Fcmd	TABLE		
<input type="checkbox"/> Order Details	TABLE		
<input checked="" type="checkbox"/> Orders	TABLE		
<input type="checkbox"/> Products	TABLE		
<input type="checkbox"/> Region	TABLE		
<input type="checkbox"/> Shippers	TABLE		
<input type="checkbox"/> Suppliers	TABLE		
<input type="checkbox"/> sysdiagrams	TABLE		



- Connexion avec les fichiers plats :

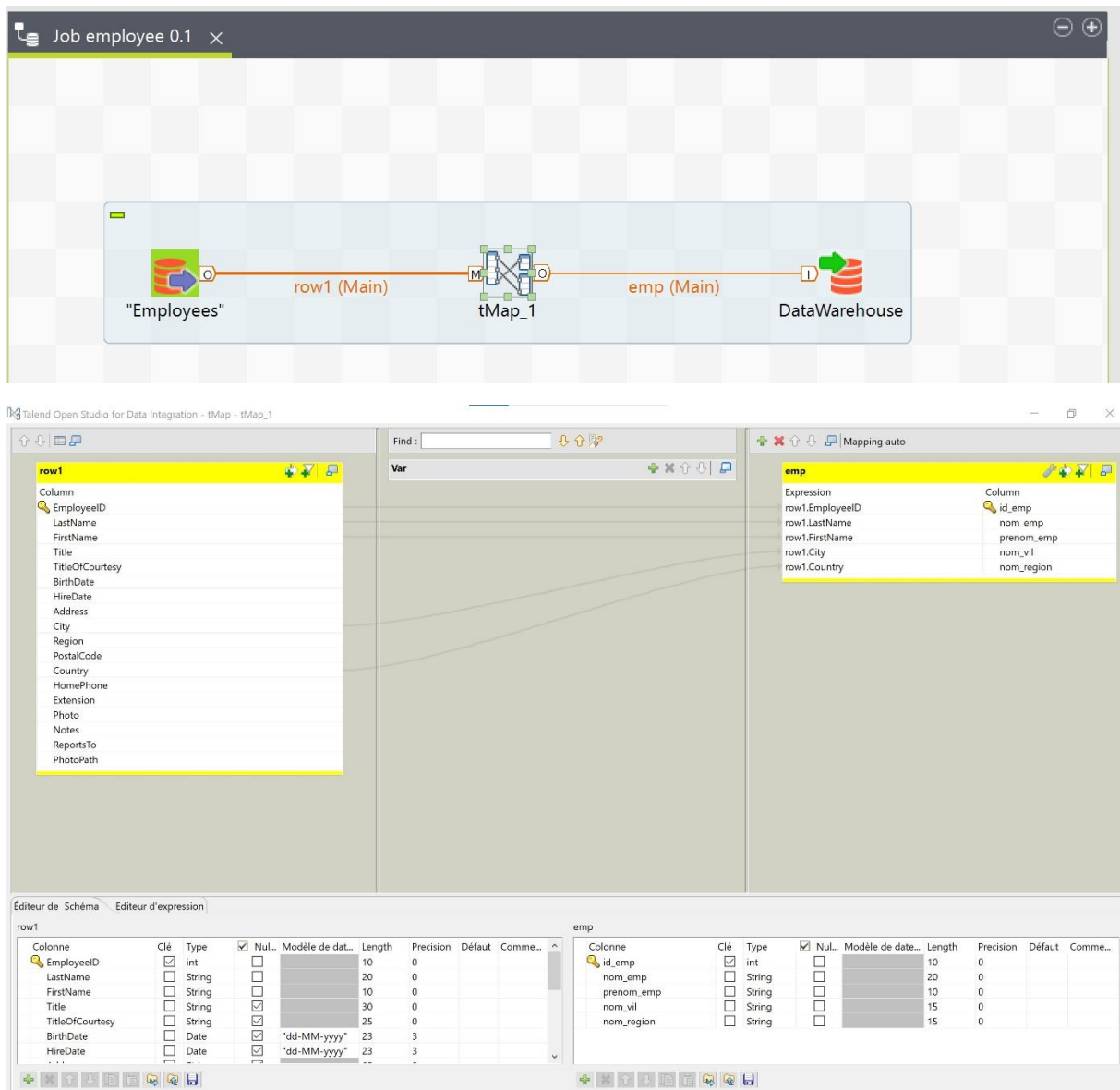


- ETL process ou l'alimentation de notre datawarehouse :

Note : pour l'alimentation avec Talend en utilisant :

- **tMap** : pour le join entre le input et le output (entre les table)
- **tUniqRow** : pour supprimer la redondance
- **tUnite** : pour merger les table par exemple le fichier plat client et la bd customer
- **tAggregateRow** : pour le count des command livrer et non livrer

1) Alimenter la dimension Employee :



- Résulta :

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP (1000) [id_emp]
      ,[nom_emp]
      ,[nom_vil]
      ,[nom_region]
      ,[prenom_emp]
FROM [StarWarehouse].[dbo].[Employee]

```

Results Messages

id_emp	nom_emp	nom_vil	nom_region	prenom_emp
1	Davolio	Seattle	USA	Nancy
2	Fuller	Tacoma	USA	Andrew
3	Leverling	Kirkland	USA	Janet
4	Peacock	Redmond	USA	Margaret
5	Buchanan	London	UK	Steven
6	Suyama	London	UK	Michael
7	King	London	UK	Robert
8	Callahan	Seattle	USA	Laura
9	Dodsworth	London	UK	Anne

2) Alimenter la dimension client :

Schéma detUnite_1

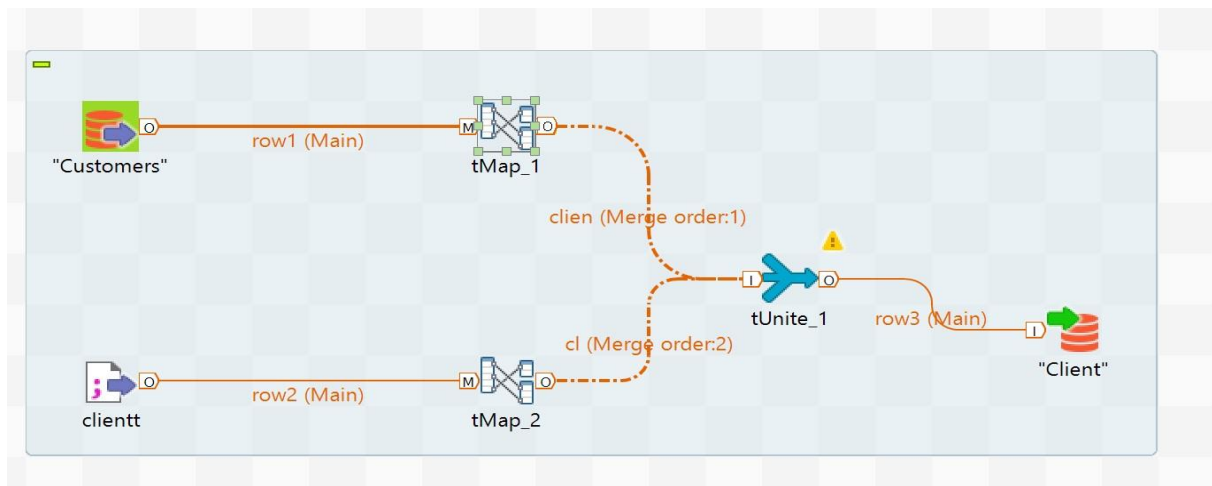
tMap_2 tMap_1

Colonne	Clé	Type	<input checked="" type="checkbox"/> Nul...	Modèle ...	Len...	Prec...	Dé...	Co...
id_client	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		5	0		
nom_cl	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>					
nom_vil	<input type="checkbox"/>	String	<input type="checkbox"/>		15	0		
nom_region	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>					

tUnite_1

Colonne	Clé	Type	<input checked="" type="checkbox"/> Nul...	Modèle ...	Len...	Prec...	Dé...	Co...
id_client	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		10	0		
nom_cl	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		10	0		
nom_vil	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		20	0		
nom_region	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		20	0		

OK Cancel



- Résulta :

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP (1000) [id_client]
      ,[nom_cl]
      ,[nom_vil]
      ,[nom_region]
FROM [StarWarehouse].[dbo].[Client]
  
```

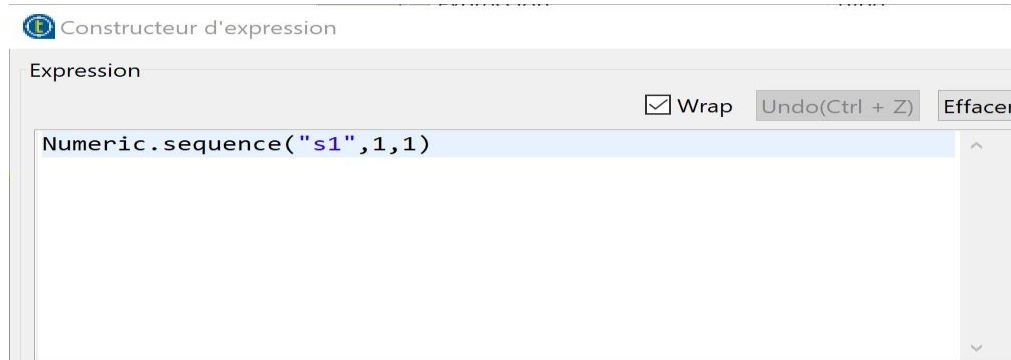
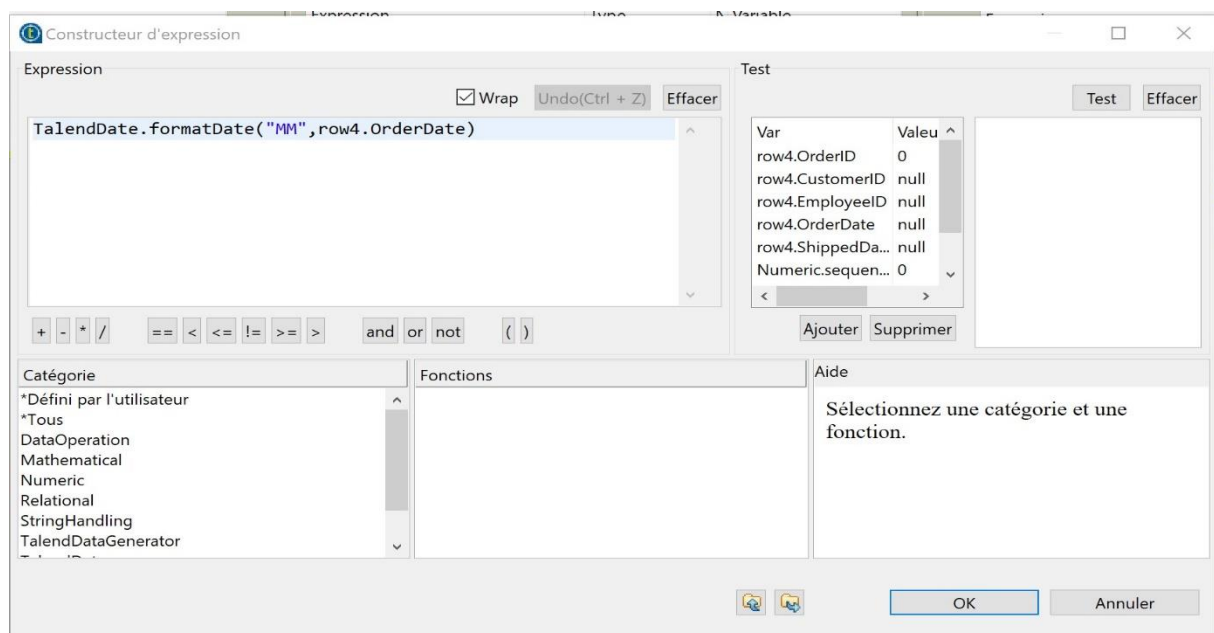
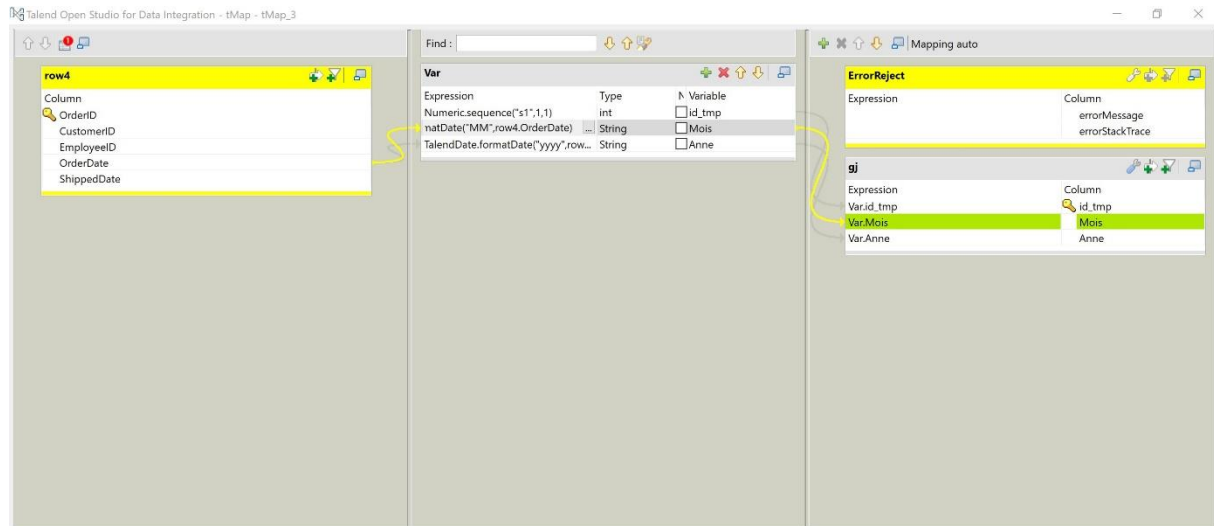
119 %

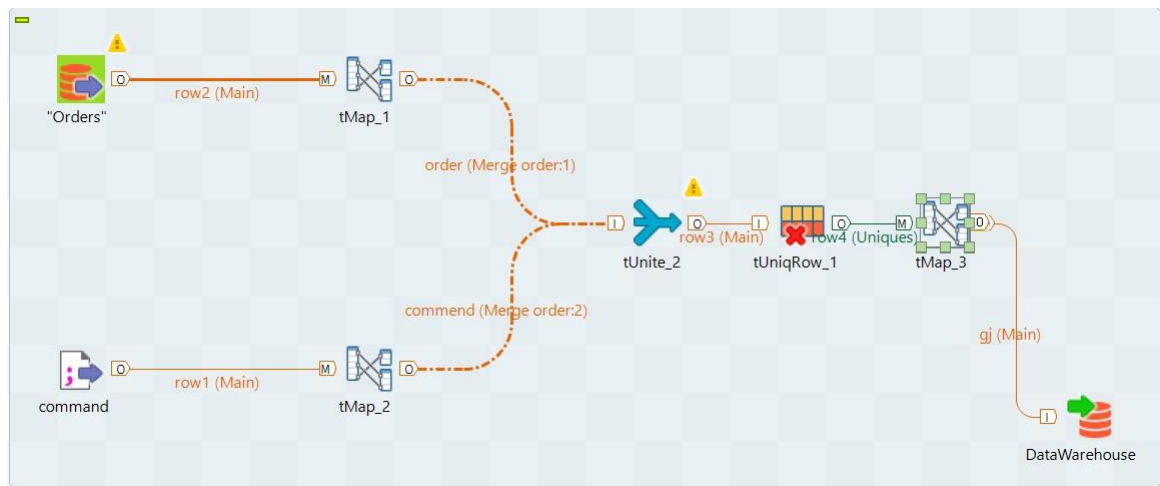
Results Messages

	id_client	nom_cl	nom_vil	nom_region
74	SAVEA	Jose Pavarotti	Boise	USA
75	SEVES	Hari Kumar	London	UK
76	SIMOB	Jytte Petersen	Kobenhavn	Denmark
77	SPECD	Dominique Perrier	Paris	France
78	SPLIR	Art Braunschweig...	Lander	USA
79	SUPRD	Pascale Cartrain	Charleroi	Belgium
80	THEBI	Liz Nixon	Portland	USA
81	THECR	Liu Wong	Butte	USA
82	TOMSP	Karin Josephs	Münster	Germany
83	TORTU	Miguel Angel Pa...	México D.F.	Mexico
84	TRADH	Anabela Doming...	Sao Paulo	Brazil
85	TRAIH	Helvetius Nagy	Kirkland	USA
86	VAFFE	Palle Ibsen	Århus	Denmark
87	VICTE	Mary Saveley	Lyon	France
88	VINET	Paul Henriot	Reims	France
89	WANDK	Rita Müller	Stuttgart	Germany
90	WARTH	Pirkko Koskitalo	Oulu	Finland
91	WELLI	Paula Parente	Resende	Brazil
92	WHITC	Karl Jablonski	Seattle	USA
93	WILMK	Matti Karttunen	Helsinki	Finland
94	WOLZA	Zbyszek Piestrze...	Warszawa	Poland

3) Alimentation de la dimension Temps :

- On va prendre les dates depuis OrdersDate et en filtrer just le besoin : mois et année -> utilisant Talend.FormatDate (MM) et (yyyy).
Pour la clé primaire on va créer un séquence numeric





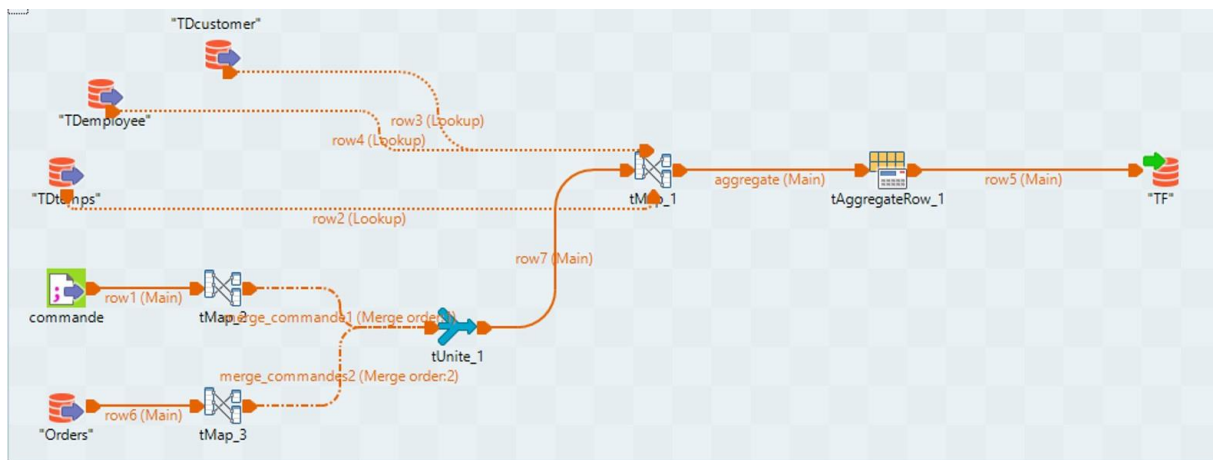
- Résultat :

```
select * from Temps
```

%		
Results	Messages	
IDtemps	Mois	Annee
1	07	1996
2	08	1996
3	09	1996
4	10	1996
5	11	1996
6	12	1996
7	12	1997
8	01	1997
9	02	1997
10	03	1997
11	04	1997
12	05	1997
13	06	1997
14	07	1997
15	08	1997
16	09	1997
17	10	1997
18	11	1997
19	12	1998
20	01	1998
21	02	1998
22	03	1998
23	04	1998
24	05	1998
25	01	2022
26	02	2022
27	05	2022

4) Alimentation de la Table_Fait :

Note : pour tUnit_ ici en l'utilise pour récupérer les dates depuis le fichier plat command



row1 Column IDCommande IDClient IDEmploye DateCommande dateArriveeCommande dateLivraisonCommande NomTransport	Find : Var	merge_commande1 Expression row1.IDCommande row1.IDClient row1.IDEmploye row1.DateCommande row1.dateLivraisonCommande Column OrderID CustomerID EmployeeID OrderDate ShippedDate
--	---------------	--

row6 Column OrderID CustomerID EmployeeID OrderDate RequiredDate ShippedDate ShipVia Freight ShipName ShipAddress ShipCity ShipRegion ShipPostalCode ShipCountry	Find : Var	merge_commandes2 Expression row6.OrderID row6.CustomerID row6.EmployeeID row6.OrderDate row6.ShippedDate Column OrderID CustomerID EmployeeID OrderDate ShippedDate
--	---------------	--

Job(Job_Faits 0.1) Contexts(Job_Faits) Composant Exécuter (Job Job_Faits)			
tAggregateRow_1			
Paramètres simples Paramètres avancés Paramètres dynamiques Vue Documentation	Schéma Group by Colonne de sortie employee_id customer_id order_date	Référentiel Bases de données (MSSQL):dwh1 - TF Modifier le schéma Sync colonnes Position de la colonne d'entrée employeeid idcustomer codeTemps	
Opérations	Colonne de sortie nb_cmd_livree nb_cmd_nonLivree	Fonction somme somme	Position de la colonne d'entrée nb_livrees nb_non_livrees <input type="checkbox"/> Ignorer les valeurs nulles <input type="checkbox"/>

- Résultat :

```
select * from Fait_cmd;
```

119 %

Results Messages

	IDclient	IDemp	IDtemps	nbr_commande_nonlivre	nbr_commande_livre
774	WELLI	1	21	0	1
775	WELLI	3	1	0	1
776	WELLI	3	8	0	1
777	WELLI	3	15	0	1
778	WELLI	4	19	0	1
779	WELLI	4	22	0	1
780	WELLI	7	14	0	1
781	WELLI	7	20	0	1
782	WELLI	9	21	0	1
783	WHITC	1	10	0	1
784	WHITC	2	23	0	1
785	WHITC	3	17	0	2
786	WHITC	3	21	0	1
787	WHITC	4	5	0	1
788	WHITC	4	11	0	1
789	WHITC	4	18	0	1
790	WHITC	4	20	0	1
791	WHITC	5	1	0	1
792	WHITC	7	10	0	1
793	WHITC	7	24	0	1
794	WHITC	8	14	0	1
795	WHITC	8	17	0	1
796	WILMK	1	21	0	1
797	WILMK	2	14	0	1
798	WILMK	2	16	0	1
799	WILMK	2	23	0	1
800	WILMK	3	21	0	1
801	WILMK	4	21	0	1
802	WILMK	7	17	0	1
803	WOLZA	1	6	0	1
804	WOLZA	1	7	0	1
805	WOLZA	4	21	0	1
806	WOLZA	4	23	0	1
807	WOLZA	5	21	0	1
808	WOLZA	6	14	0	1
809	WOLZA	8	23	0	1

3. Réalisation d'un tableau de bord qui analyser le besoin :

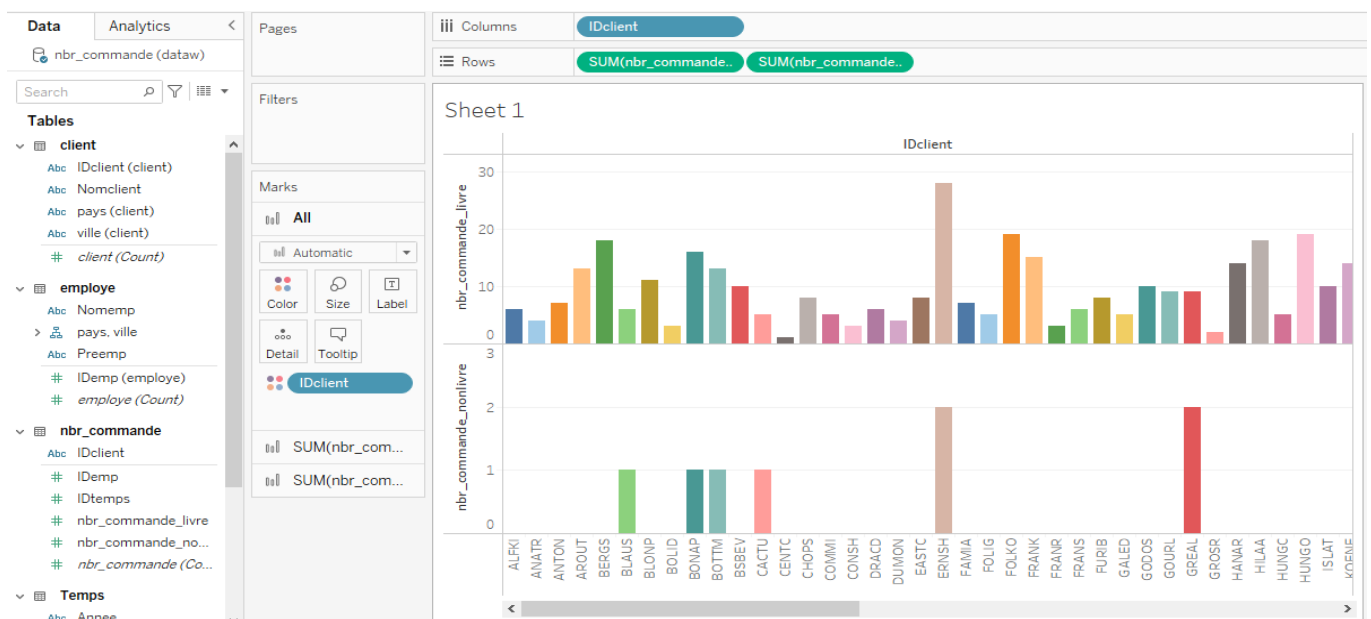
Utilisant l'outil Tableau :

- NewSheet → pour visualiser un besoin
- NewDashboard → pour mettre tous les visualisation de chaque besoin

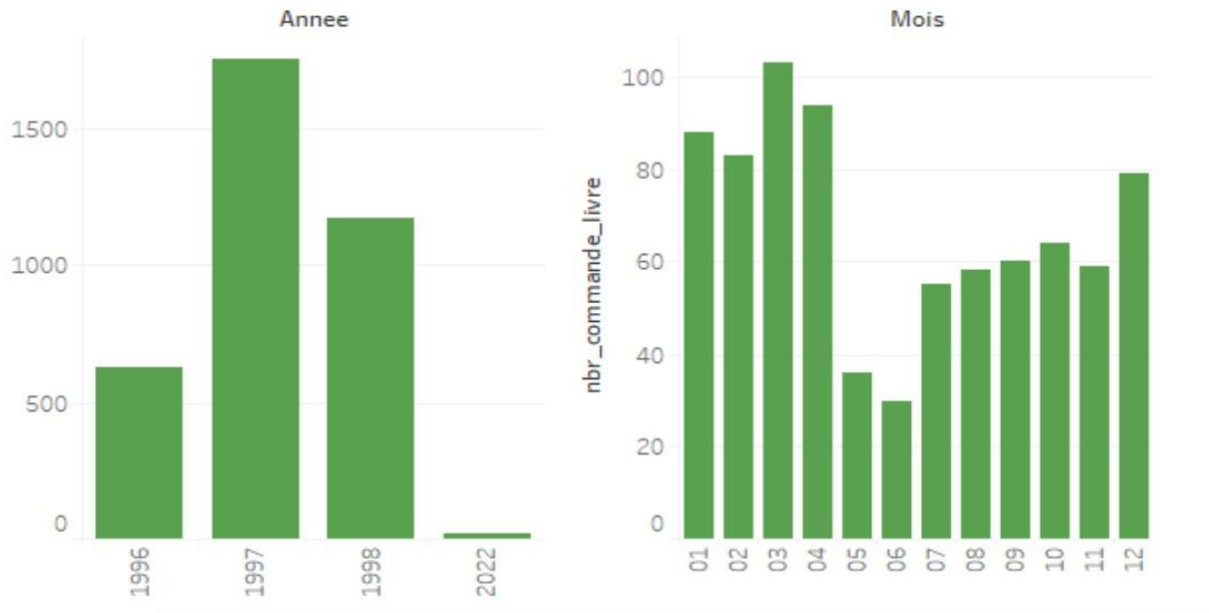
- Connexion de la Datawarehouse alimenter (résulta) avec Tableau :

The screenshot shows the Tableau Desktop interface with the 'nbr_commande (dataw)' connection. The 'client' table is selected, and its fields are being mapped to the 'nbr_commande' table. The 'client' table has fields: IDclient (client), Nomclient, ville (client), pays (client). The 'nbr_commande' table has fields: IDclient (client), client, IDclie...

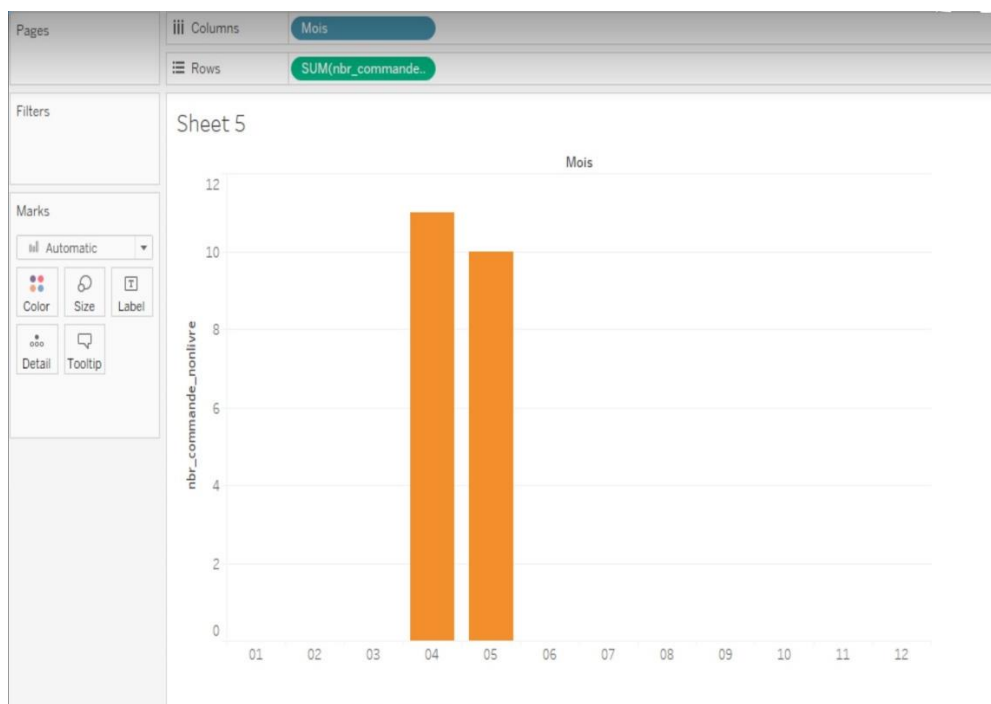
- Nombre commande livrer et non_livrer par client :



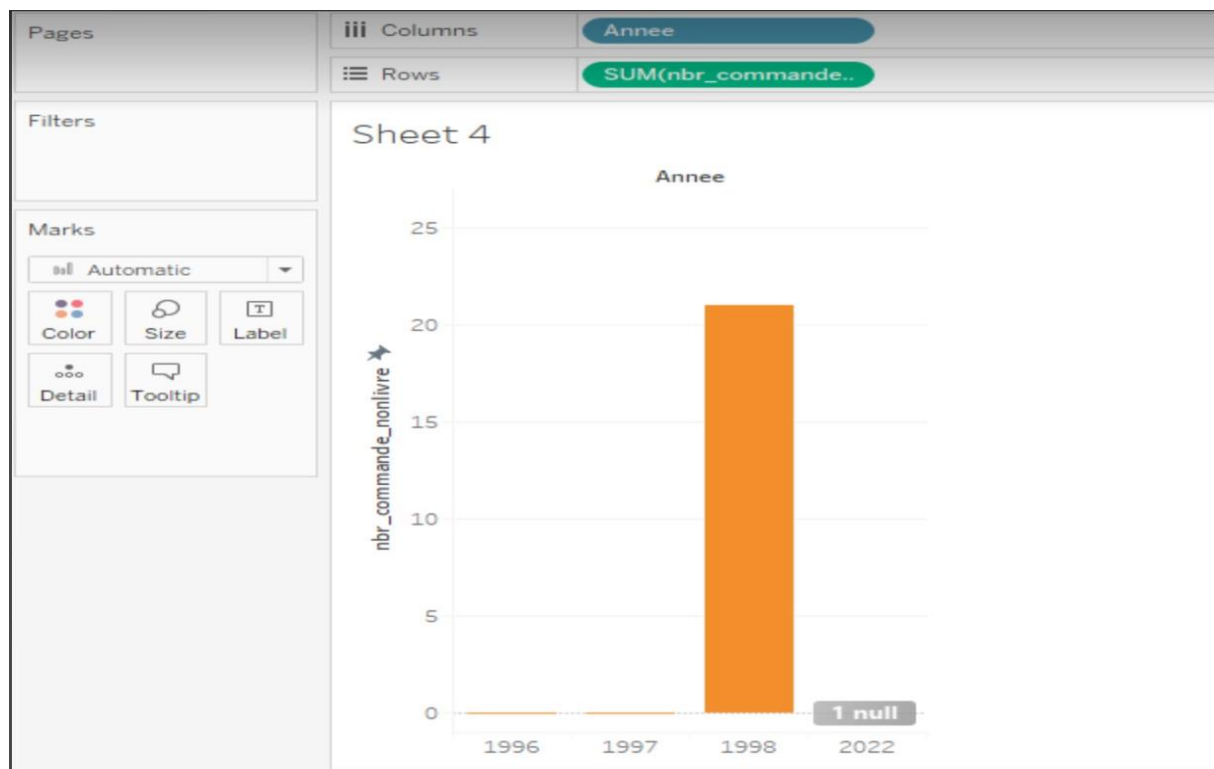
- Nombre commande livrer par Mois et Année :



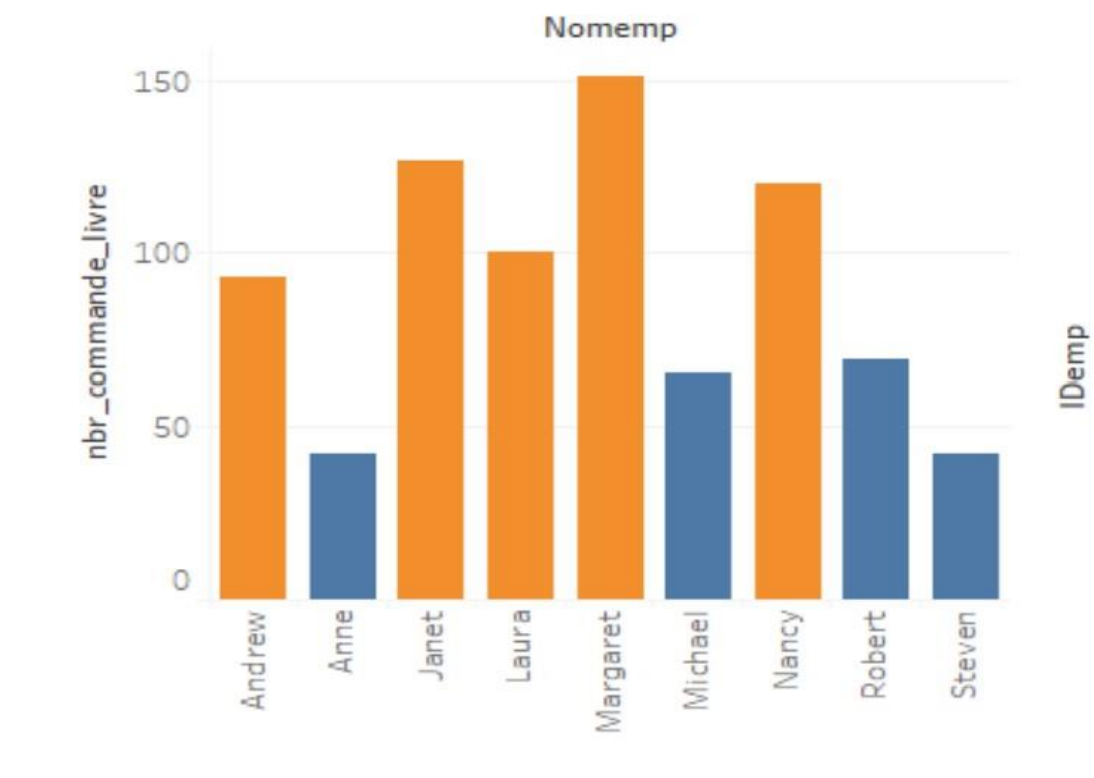
- Nombre commande Non_livrer par Mois et Année :
 - Par Mois :



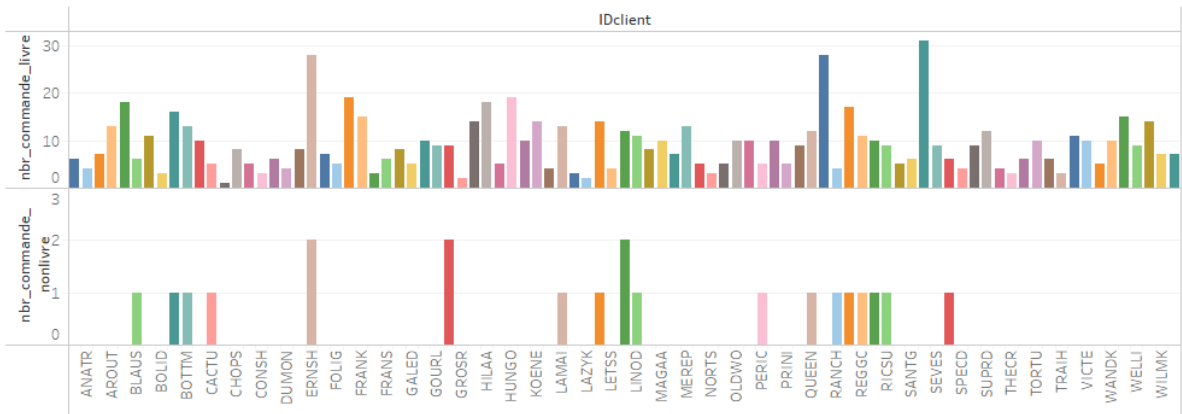
- Par Année :



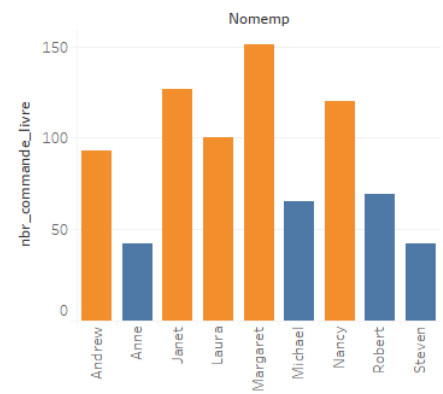
• Nombre commande livrer et non_livrer par employée :



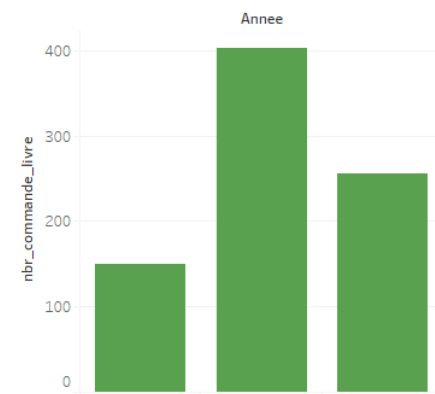
- Dashboard (résultat final) :



Sheet 8



Sheet 2



Sheet 3

