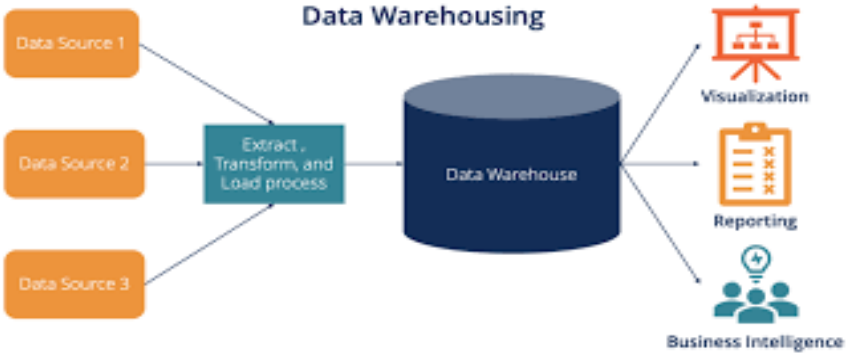
RAPPORT BUSINESS INTELEGENT PROJECT



Réaliser par :

- ❖ Dhaibi soukayna
- ❖ El mazani imane
- ❖ El alami chaimaa

Encadrante :

Soussi nassima

introduction:

Les systèmes de BI sont utilisés par les décideurs pour obtenir une connaissance approfondie de l'entreprise et de définir et soutenir leurs stratégies d'affaires, par exemple :

- d'acquérir un avantage concurrentiel,
- d'améliorer la performance de l'entreprise,
- de répondre plus rapidement aux changements,
- d'augmenter la rentabilité, et d'une façon générale la création de valeur ajoutée pour l'entreprise.

Pour notre cas nous allons réaliser un système décisionnel de la base de données AdventureWorks qui stocke les scénarios de traitement des transactions en ligne standard pour un fabricant de vélos fictif. Les scénarios incluent la fabrication, les ventes, les achats, la gestion des produits

Étude des besoins et de l'existant

Objectif :

Nous intéressants à l'analyse de ventes en fonction des clients, des produits ,des offre spécial ,des vendeurs …on va les détailler après .

Les questions analytiques auxquelles devra répondre notre magasin de données :

- quelles produits sont les plus vendus et qui génèrent plus de revenu.
- quelles catégories des produits sont les plus vendus et qui génèrent plus de revenu.
- quelles classes des produits sont les plus vendus et qui génèrent plus de revenu.

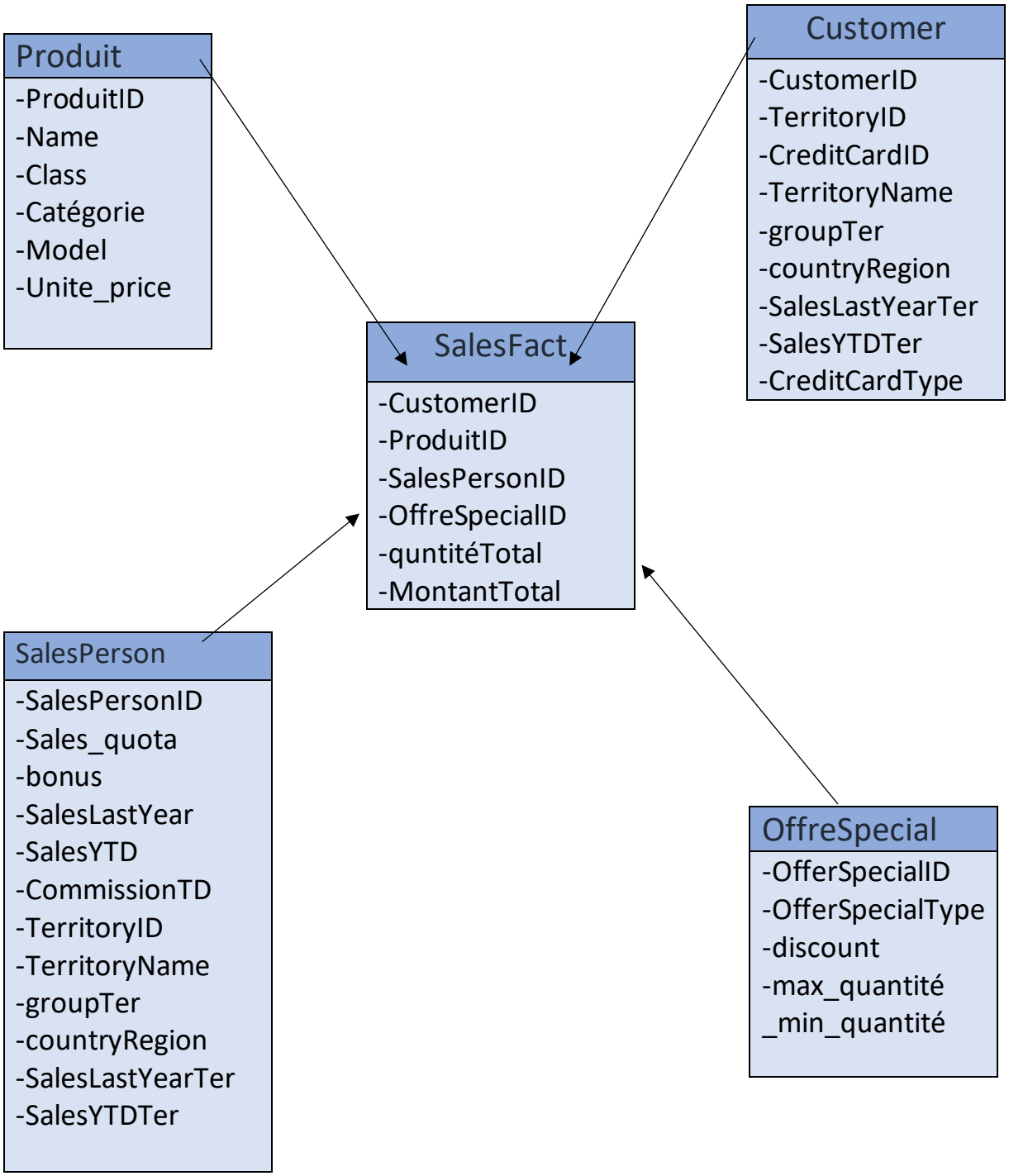
- quelles modèles des produits sont les plus vendus et qui génèrent plus de revenu.
- Quelles sont les clients fidèles
- Quelles sont les territoriens des clients fidèles
- Quelles sont les groupes des territoriens des clients fidèles
- Quelles sont les régions pays des territoriens des clients fidèles
- Visualiser les revenus des territoriens des clients selon les ventes de l'année précédentes
- Visualiser les revenus des territoriens des clients selon les ventes effectuer jusqu' à maintenant
- Quelles sont les types des crédites cartes des clients fidèles
- Quelles sont les offres spéciales les plus populaire
- Quelles sont les types des offres spéciales les plus populaires
- Quelles sont les catégories des offres spéciales les plus populaires
- Quelles sont les maximum de quantité des offres spéciales les plus populaires
- Quelles sont les minimum de quantité des offres spéciales les plus populaire
- Quelles sont les unités commerciales qui génèrent le plus de revenus?
- Quelles sont les territoriens des unités commerciales qui génèrent le plus de revenus?
- Quelles sont les bonus des unités commerciales qui génèrent le plus de revenus?
- Quelles sont les commissionTD des unités commerciales qui génèrent le plus de revenus?

- Quelles sont les sales quota des unités commerciales qui génèrent le plus de revenus?
- Quelles sont les ventes de la dernière année des unités commerciales qui génèrent le plus de revenus?
- Quelles est le total des ventes des unités commerciales qui génèrent le plus de revenus?

Modélisation et conception :

Modélisation dimensionnelle d'un magasin de données :

Model en étoile :



Model en flocon de neige :

Les hierarchies :

Table Produits :

ProduitID → Name → Class → Categorie → Model

ProduitID → unite_price

Table Customer:

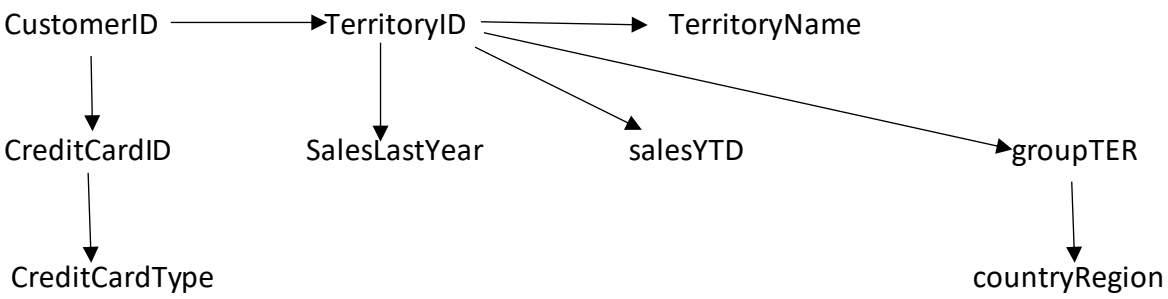


Table SalesPerson:

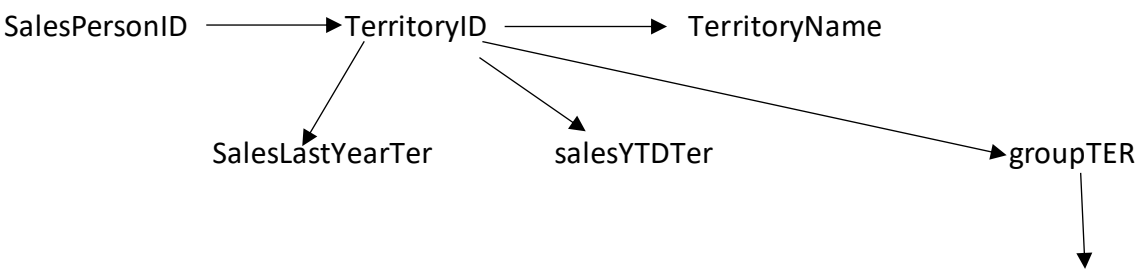
SalesPersonID → sales_quota

SalesPersonID → commissionTD

SalesPersonID → bonus

SalesPersonID → salesLastYear

SalesPersonID → salesYTD



countryRegion

Table OffreSpecial:

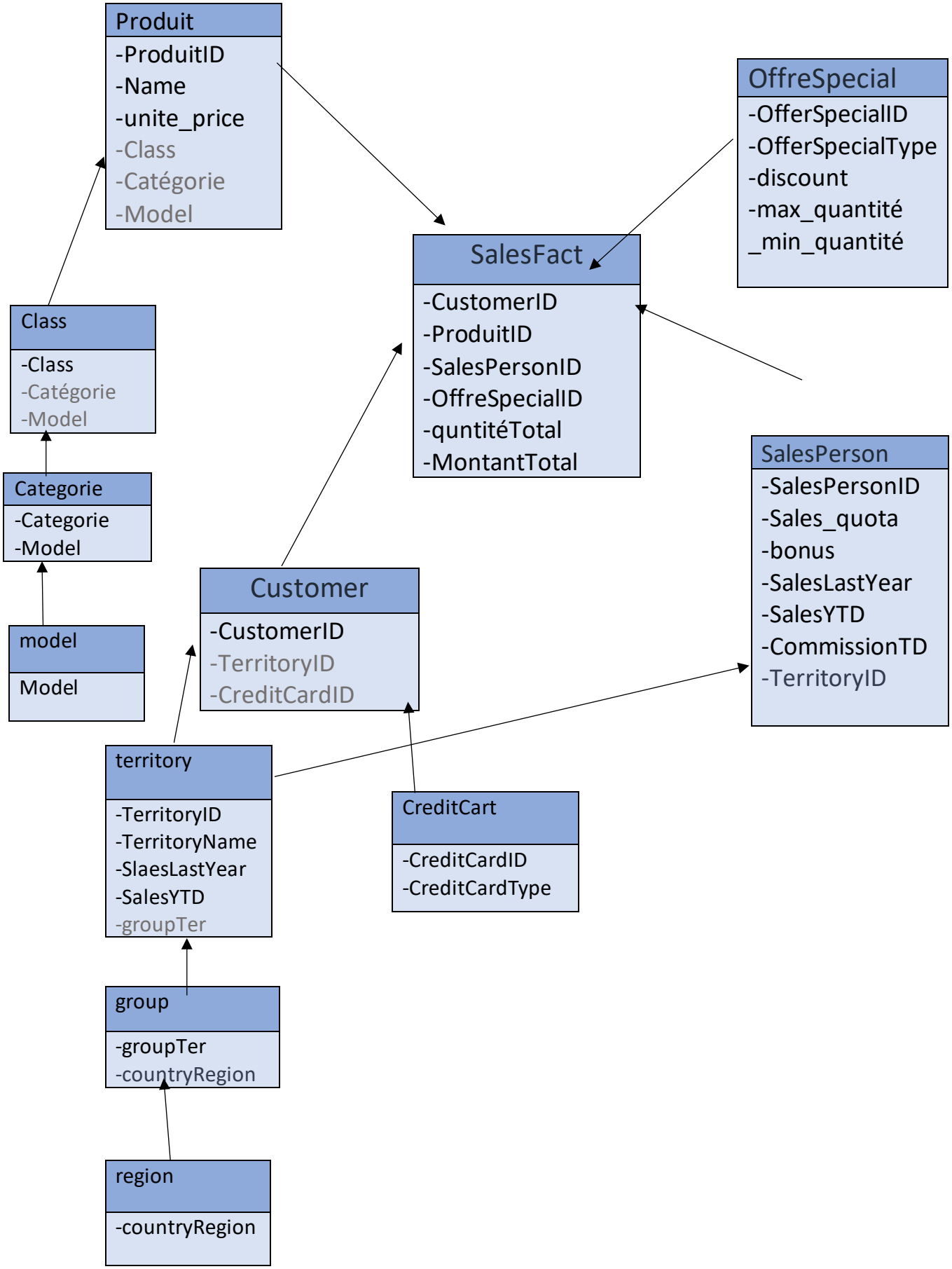
OfferSpecialID ⟶ OffreType

OfferSpecialID ⟶ discount

OfferSpecialID ⟶ max_cuantité

OfferSpecialID ⟶ min_quantité

Schéma en flocon de neige :


Produit
-ProduitID
-Name
-unite_price
-Class
-Catégorie
-Model
OffreSpecial
-OfferSpecialID
-OfferSpecialType
-discount
-max_quantité
_min_quantité
SalesFact
-CustomerID
-ProduitID
-SalesPersonID
-OffreSpecialID
-quntitéTotal
-MontantTotal
Class
-Class
-Catégorie
-Model
Categorie
-Categorie
-Model
model
Model
Customer
-CustomerID
-TerritoryID
-CreditCardID
SalesPerson
-SalesPersonID
-Sales_quota
-bonus
-SalesLastYear
-SalesYTD
-CommissionTD
-TerritoryID
territory
-TerritoryID
-TerritoryName
-SlaesLastYear
-SalesYTD
-groupTer
CreditCart
-CreditCardID
-CreditCardType
group
-groupTer
-countryRegion
region
-countryRegion

Implémentation du data Mart et Mise en place de l'ETL

Les trois phases du processus d'ETL pour l'alimentation d'un

dataMart sont les suivantes :

- ## Extraction des données

Sélection et extraction des données sources utiles et judicieuse.

Les données des produits :

```
SELECT PR.ProductID, PR.Name, PR.class
,PC.Name,PM.Name,SOD.UnitPrice
FROM
Production.Product PR , Production.ProductSubCategory PSC
,Production.ProductCategory PC , Production.ProductModel PM ,
Sales.SalesOrderDetail SOD

where PC.ProductCategoryID in ( SELECT PSC.ProductCategoryID
FROM Production.ProductSubCategory WHERE
PR.ProductSubCategoryID=PSC.ProductSubCategoryID)

AND (PR.ProductModelID = PM.ProductModelID and SOD.ProductID =
PR.ProductID)
```

Les données de Customer :

```
SELECT
  CS.CustomerID
, CS.TerritoryID,IND.CreditCardID
, ST.Name
,ST. CountryRegionCode
, ST."Group"
, ST.SalesYTD
, ST.SalesLastYear
, SCD.CardType
FROM Sales.Customer CS ,Sales.SalesOrderHeader IND
,Sales.SalesTerritory ST , Sales.CreditCard SCD
WHERE
IND.CreditCardID in(select CreditCardID from
Sales.SalesOrderHeader where  CS.CustomerID = IND.CustomerID)
AND( (ST.TerritoryID=CS.TerritoryID)
AND(IND.CreditCardID=SCD.CreditCardID))
```

Les données de OffreSpecial :

```
SELECT
  SpecialOfferID
, DiscountPct
, Type
, Category
, MinQty
, MaxQty
FROM Sales.SpecialOffer
```

Les données de SalesPerson :

```
SELECT
  SP.BusinessEntityID
, SP.TerritoryID
, SP.SalesQuota
, SP.Bonus
, SP.CommissionPct
, SP.SalesYTD
, SP.SalesLastYear
, ST.Name
,ST. CountryRegionCode
, ST."Group"
, ST.SalesYTD
, ST.SalesLastYear
FROM Sales.SalesPerson SP ,Sales.SalesTerritory ST where
SP.TerritoryID IS NOT NULL AND (ST.TerritoryID=SP.TerritoryID)
```

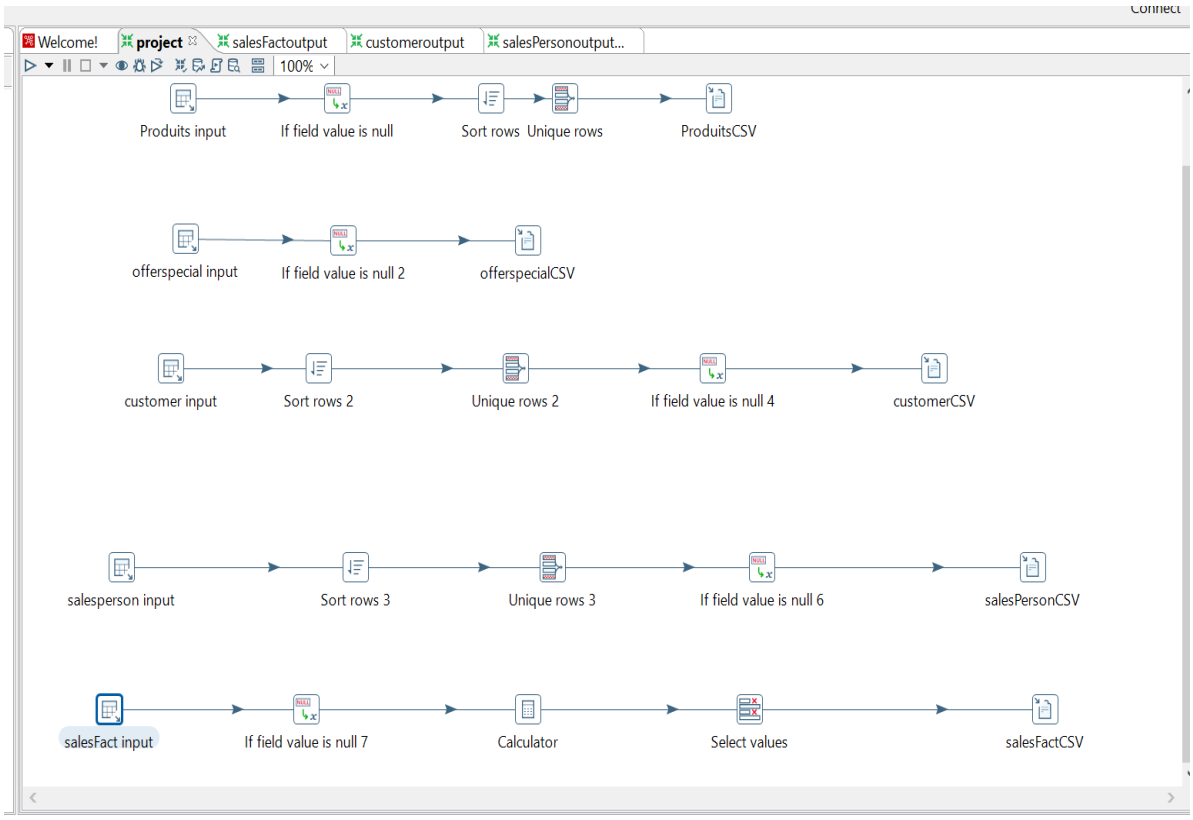
Les données de SalesFact :

```
SELECT
OrderQty
, ProductID
, SpecialOfferID
, UnitPriceDiscount*UnitPriceDiscount
,UnitPrice*OrderQty,SOH.CustomerID, SOH.SalesPersonID
FROM Sales.SalesOrderDetail SOD, Sales.SalesOrderHeader SOH
where (SOD.SalesOrderID = SOH.SalesOrderID and SOH.CustomerID
IS NOT NULL)
and( (ProductID IS NOT NULL and SpecialOfferID IS NOT NULL )
and SOH.SalesPersonID  IS NOT NULL)
```

➢ Nettoyage et Transformation

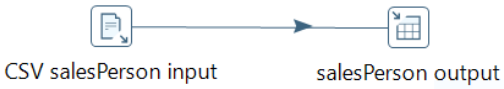
L'objectif majeur de cette phase est de résoudre le problème de consistance des données

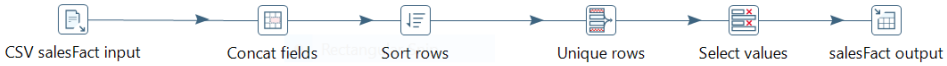
- Gérer les valeurs nulles des attributs
- Ordonner les lignes des tables
- Supprimer les ligne redondantes
- Calculer le montant total en fonction de quantité et de prix unitaire et de l'offre discount
  montantTotal = Prix_unit*quantité*(1-discount)



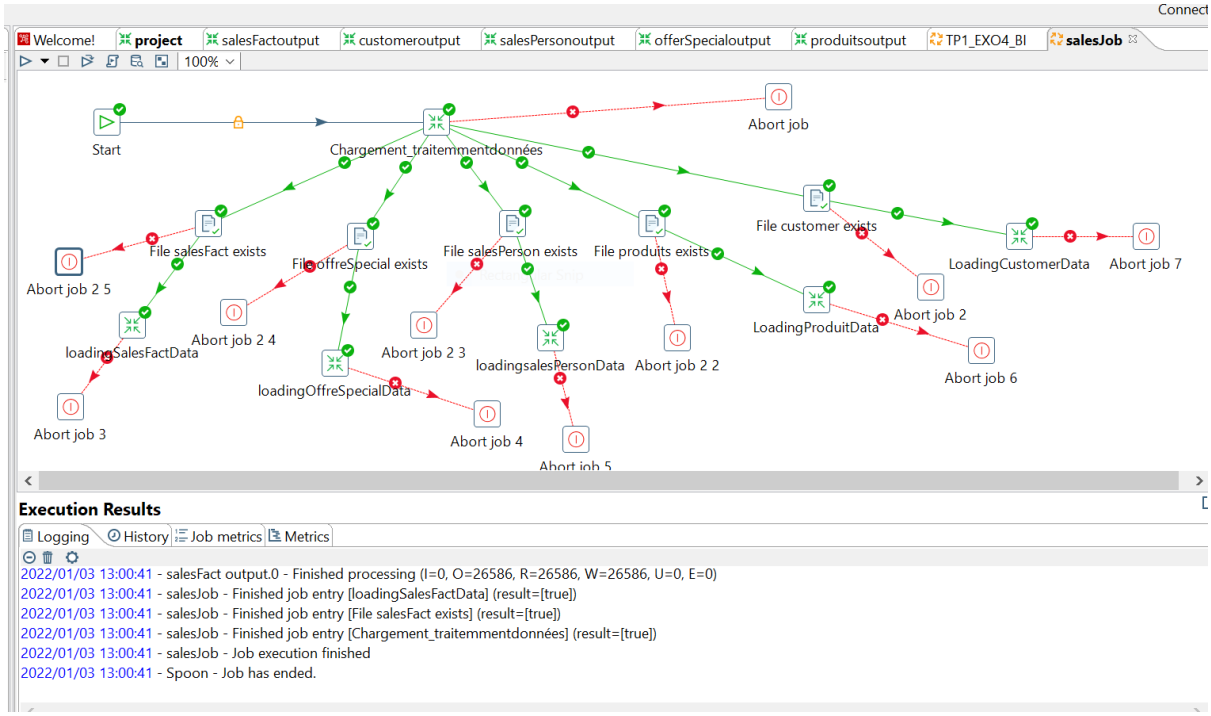
➢ Chargement

Objectif : charger les données nettoyées et préparées dans l'ED.





Job qui rassemble le processus ETL

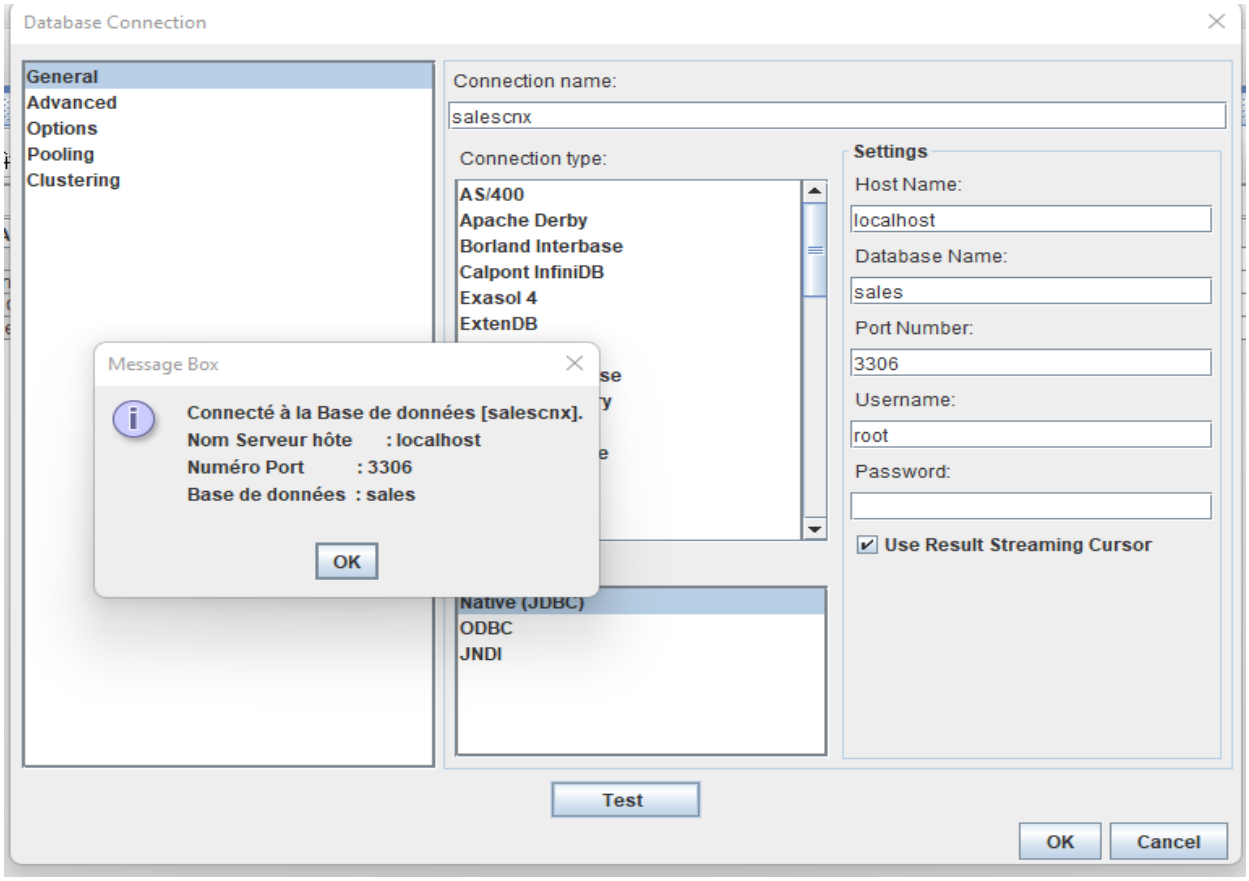


Exploitation du data mart :

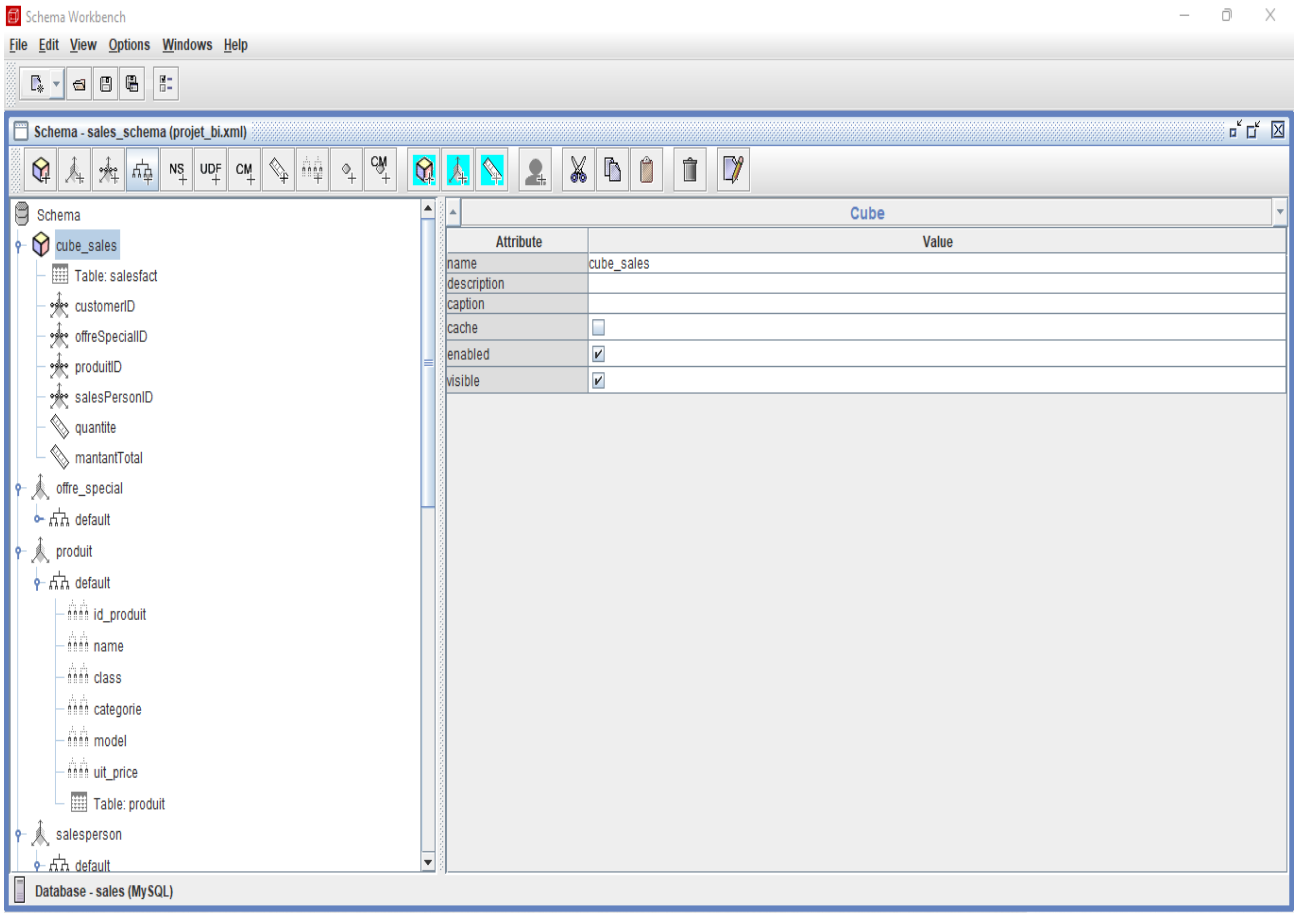
Phase 1 :

modélisation du schéma de cube avec pentaho schéma workbensh

Création d’une connexion avec le data mart

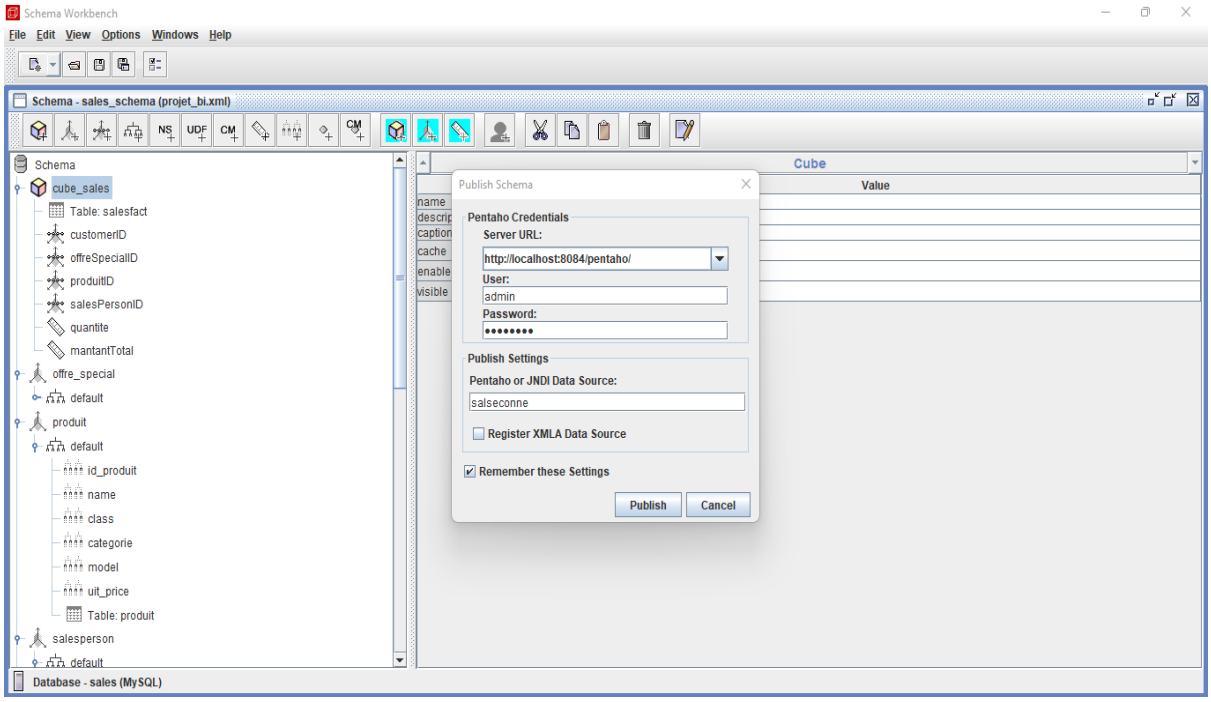


Schema xml du cube de notre data sales:

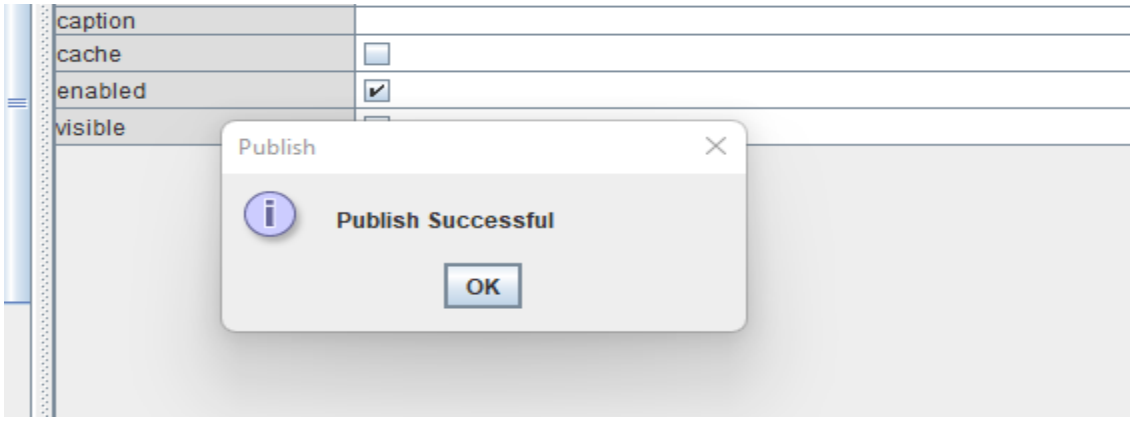


Phase 2 : publication du schema Mondrian du cube sur le serveur Pentaho

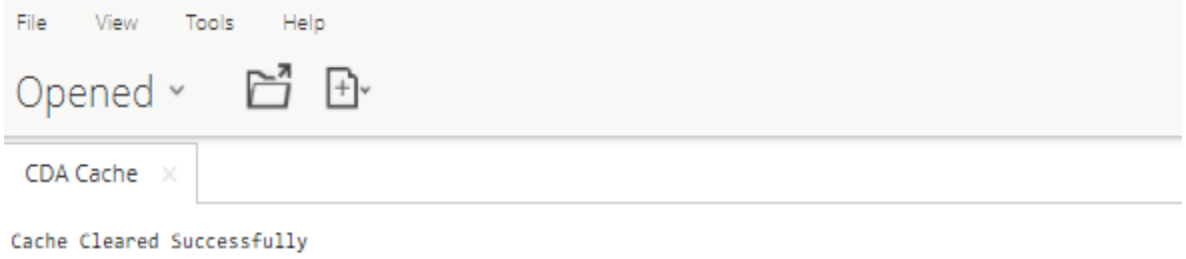
File ---->publish



Apres la reussite de publication de notre schema par PSW

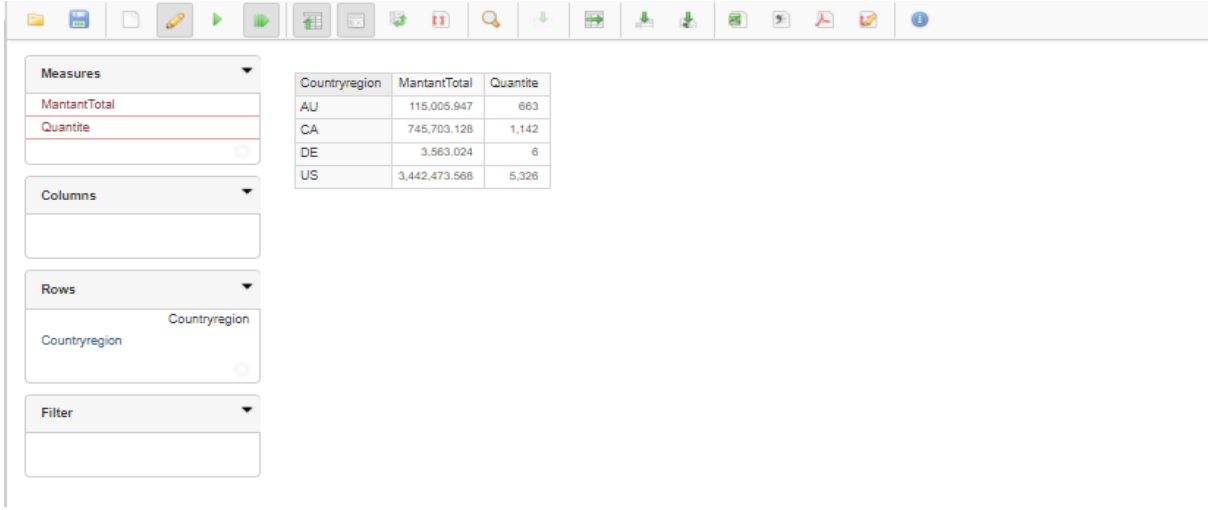


Puis vider le cache du schema

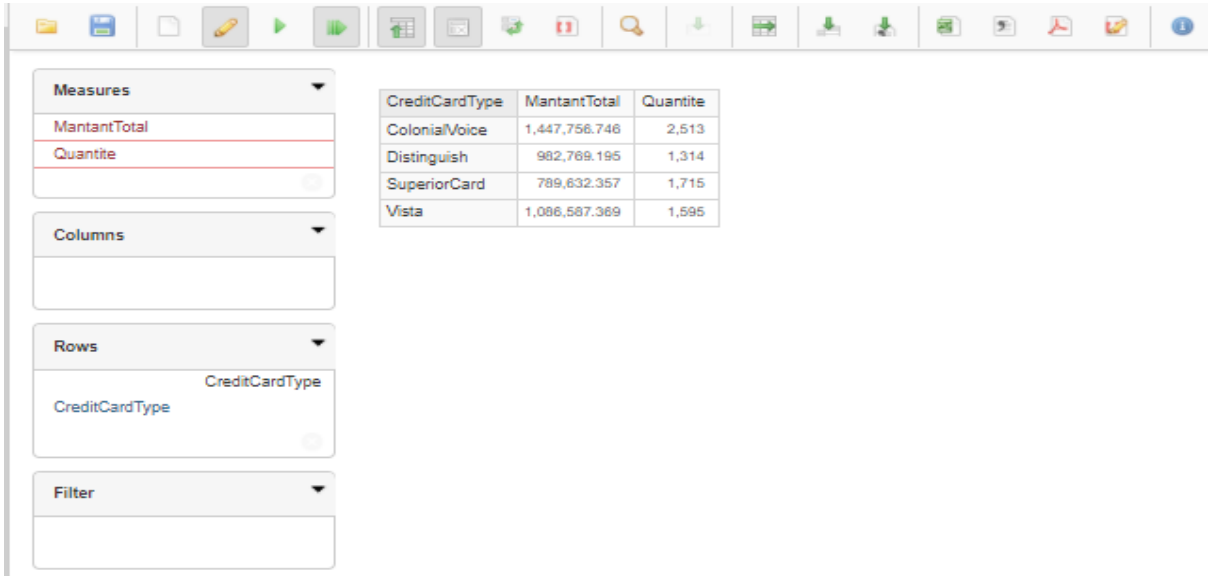


Restitution des vues métiers a l'aide du plugins saiko :

Le montant total et quantité totale par countryregion:

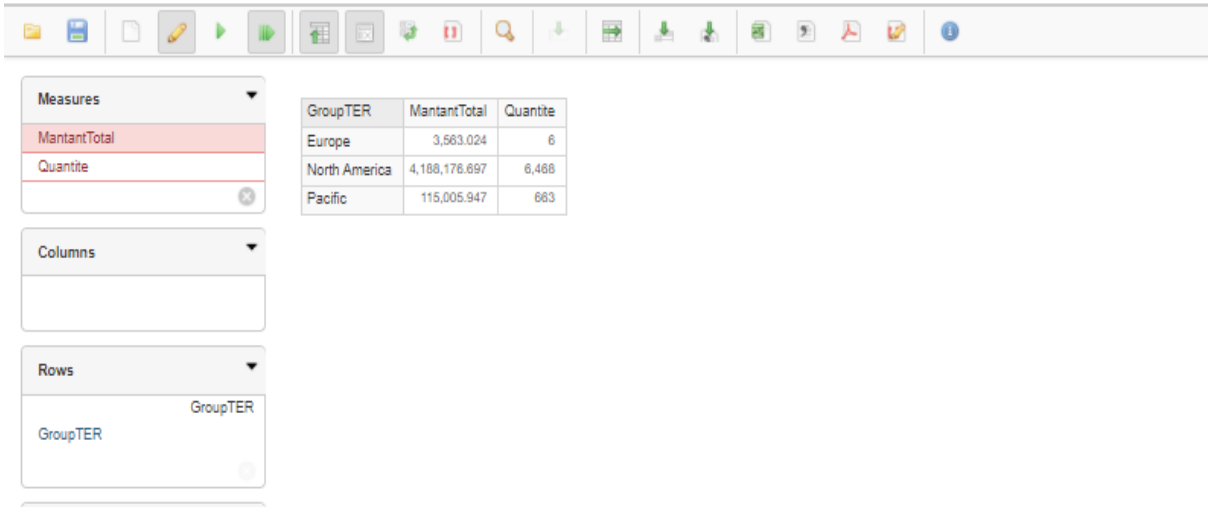


Le montant total et quantité totale par creditCardType :



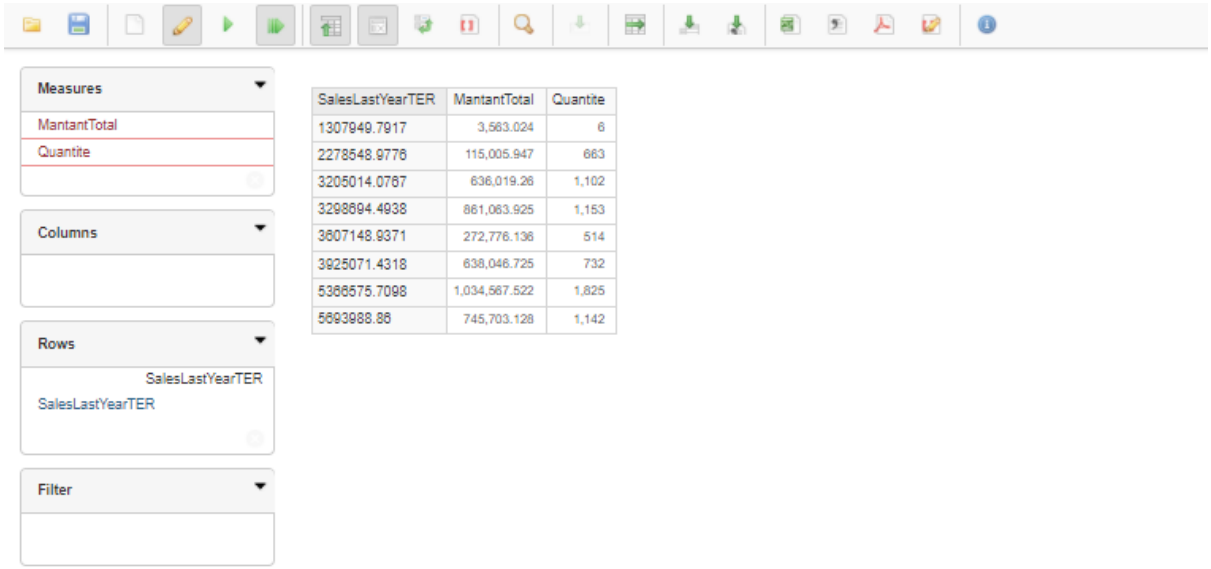
CreditCardType	MantantTotal	Quantite
ColonialVoice	1,447,756.746	2,513
Distinguish	982,769.195	1,314
SuperiorCard	789,632.357	1,715
Vista	1,086,587.369	1,595

Le montant total et quantité totale par groupTER :



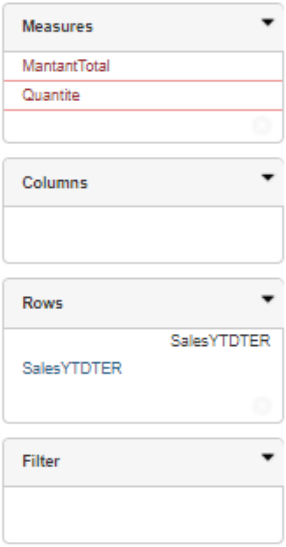
GroupTER	MantantTotal	Quantite
Europe	3,563.024	6
North America	4,188,176.697	6,468
Pacific	115,005.947	663

Le montant total et quantité totale par salesLastYearsTER :



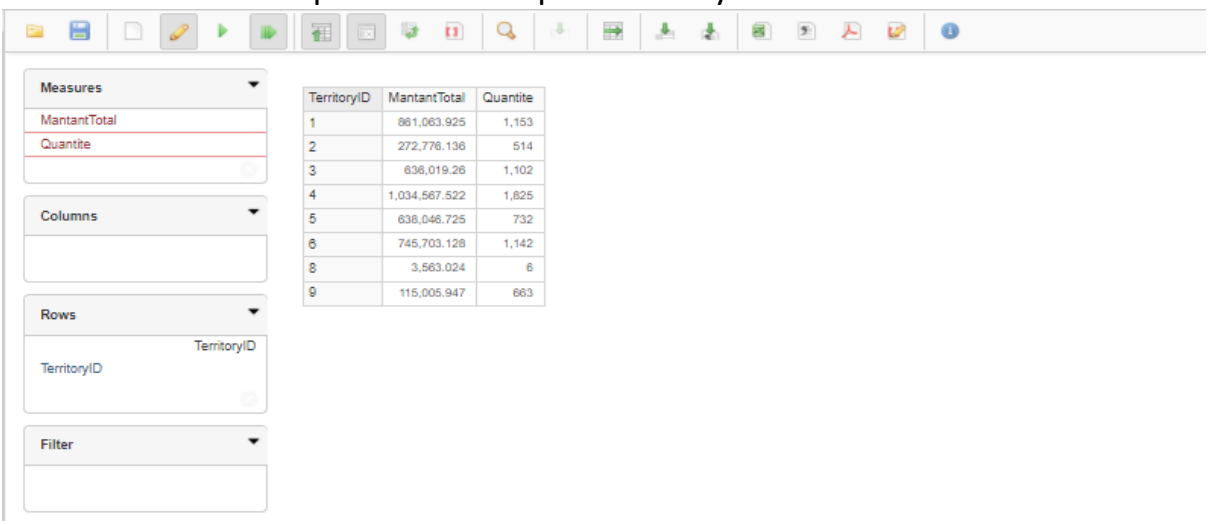
SalesLastYearTER	MantantTotal	Quantite
1307949.7917	3,563.024	6
2278548.9776	115,005.947	663
3205014.0767	636,019.26	1,102
3298694.4938	861,063.925	1,153
3607148.9371	272,776.136	514
3925071.4318	638,046.725	732
5366575.7098	1,034,567.522	1,825
5693988.86	745,703.128	1,142

Le montant total et quantité totale par SalesYTDTER:



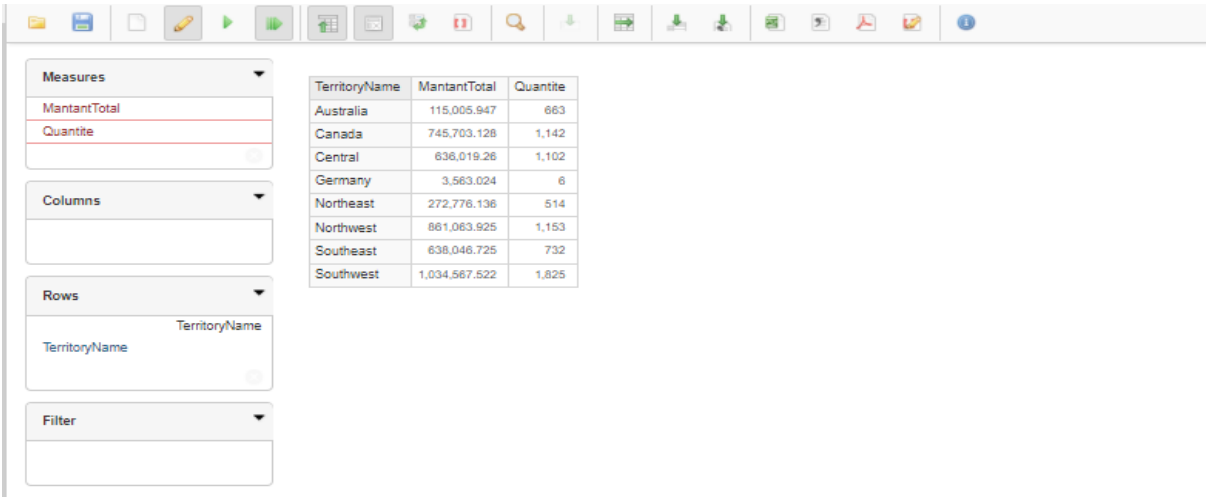
SalesYTDTER	MantantTotal	Quantite
2402176.8476	272,776.136	514
2538667.2515	638,046.725	732
3072175.118	636,019.26	1,102
3805202.3478	3,563.024	6
5977814.9154	115,005.947	663
6771829.1376	745,703.128	1,142
7887186.7882	861,063.925	1,153
1.05108538739E7	1,034,567.522	1,825

Le montant total et quantité totale par Territeryid



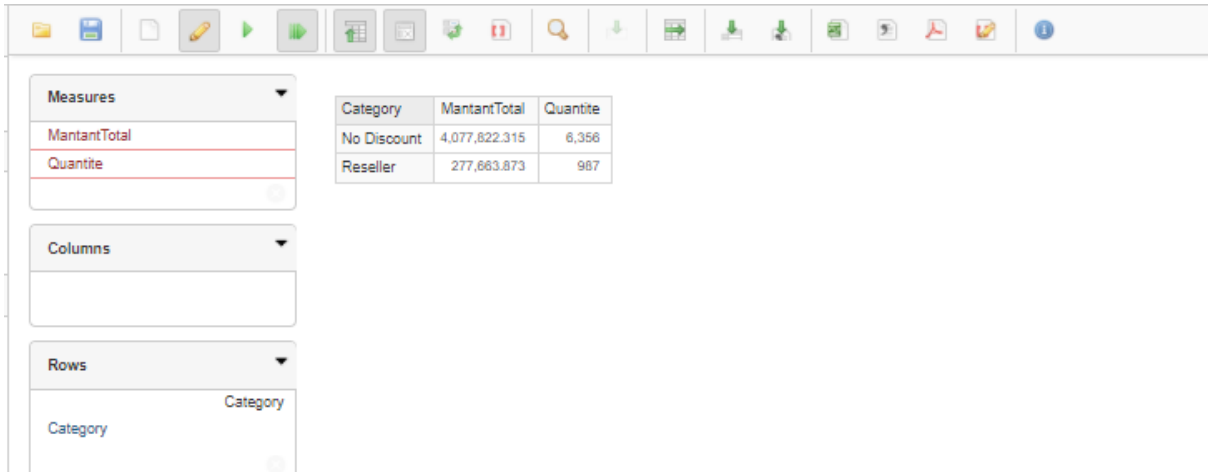
TerritoryID	MantantTotal	Quantite
1	861,063.925	1,153
2	272,776.136	514
3	636,019.26	1,102
4	1,034,567.522	1,825
5	638,046.725	732
6	745,703.128	1,142
8	3,563.024	6
9	115,005.947	663

Le montant total et quantité totale par territeryName :

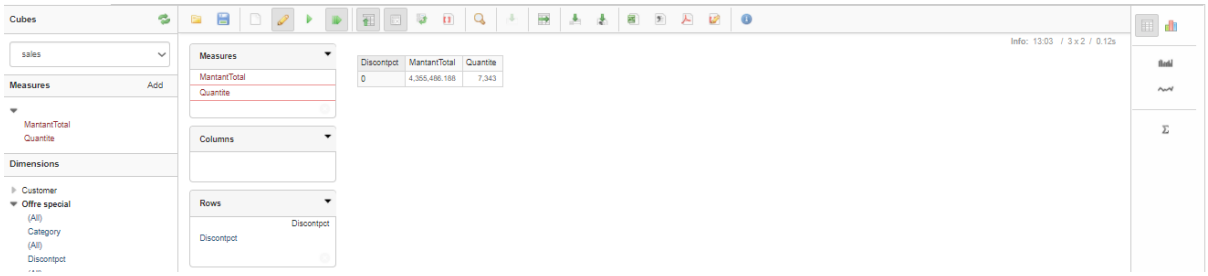


TerritoryName	MantantTotal	Quantite
Australia	115,005.947	663
Canada	745,703.128	1,142
Central	636,019.26	1,102
Germany	3,563.024	6
Northeast	272,776.136	514
Northwest	861,063.925	1,153
Southeast	638,046.725	732
Southwest	1,034,567.522	1,825

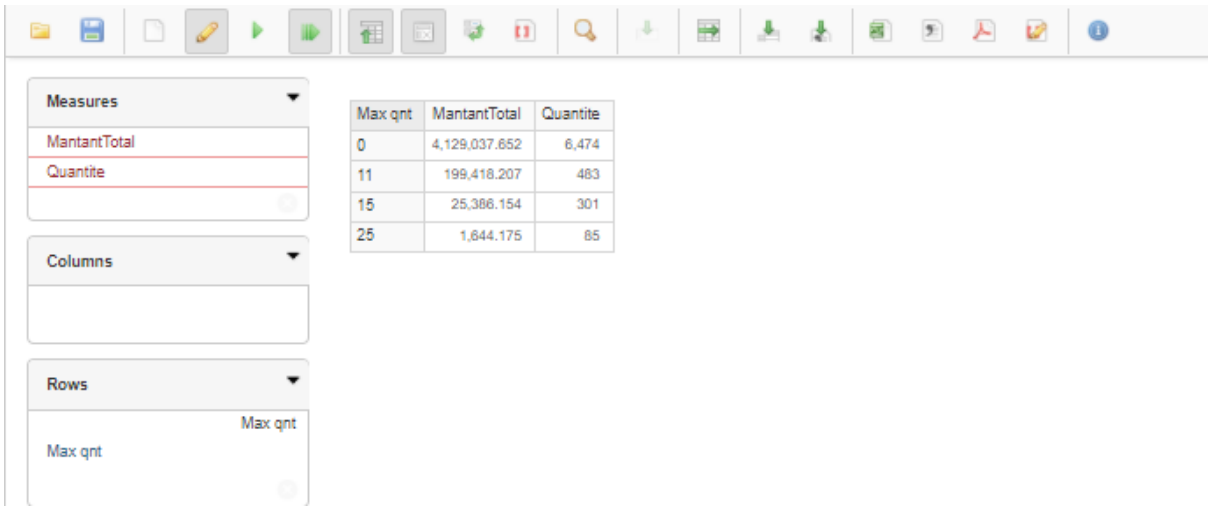
Le montant total et quantité totale par Category :



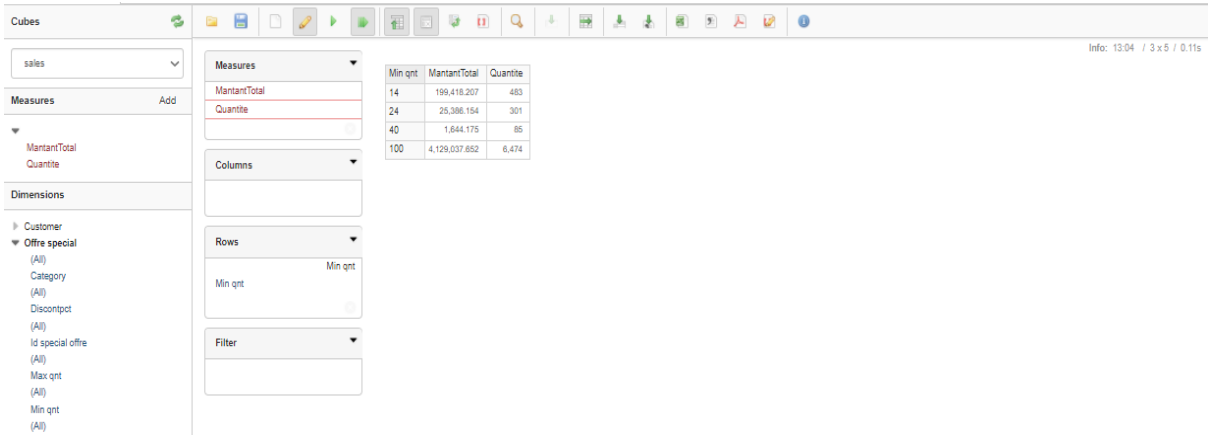
Le montant total et quantité totale par Discountpnt :



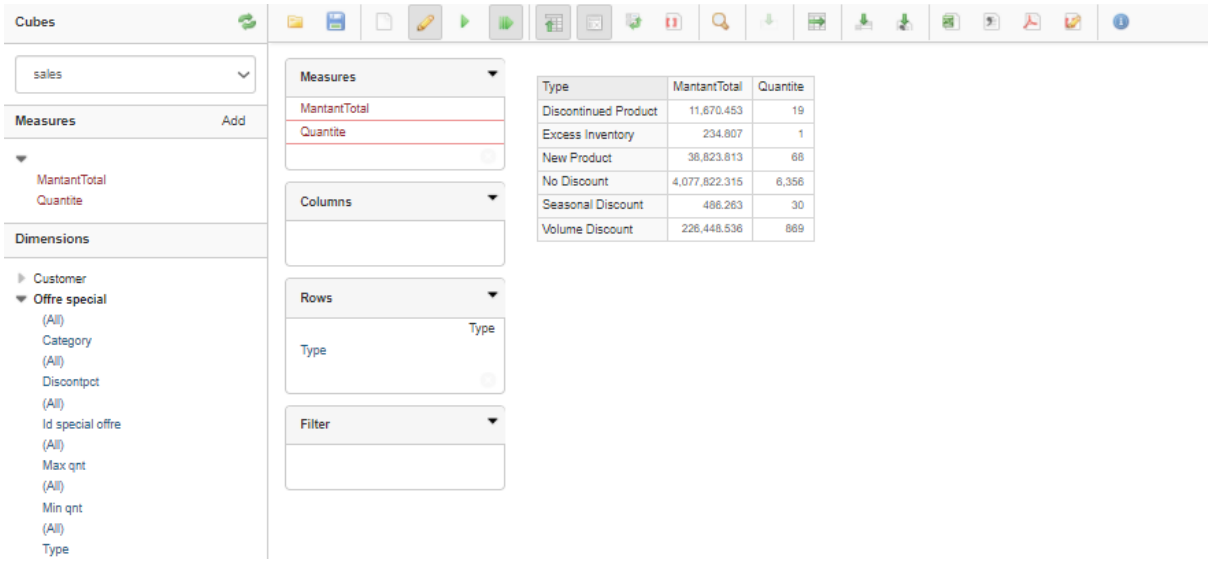
Le montant total et quantité totale par Max quantiter :



Le montant total et quantité totale par Min quantite :

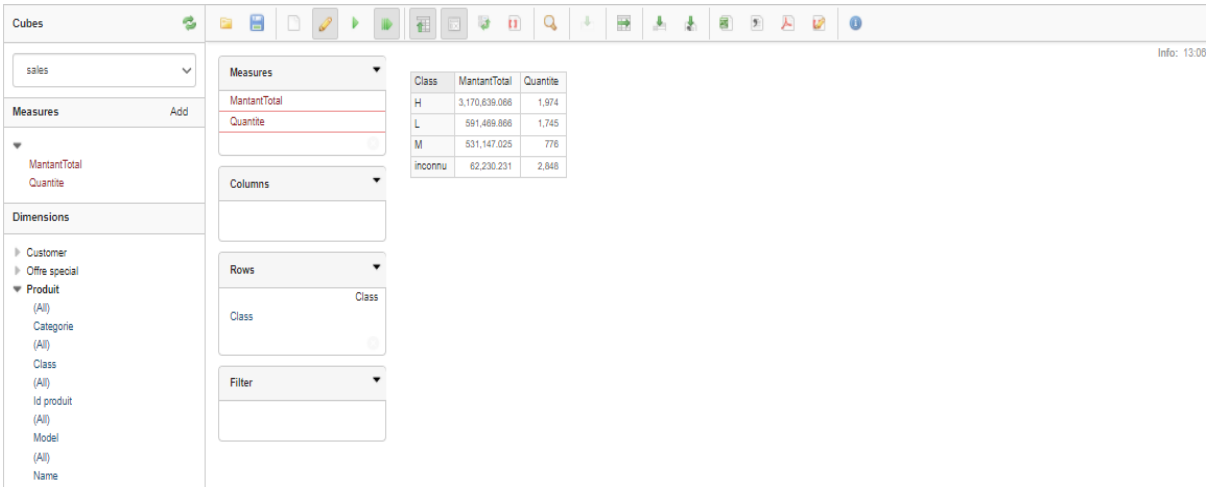


Le montant total et quantité totale par Type :

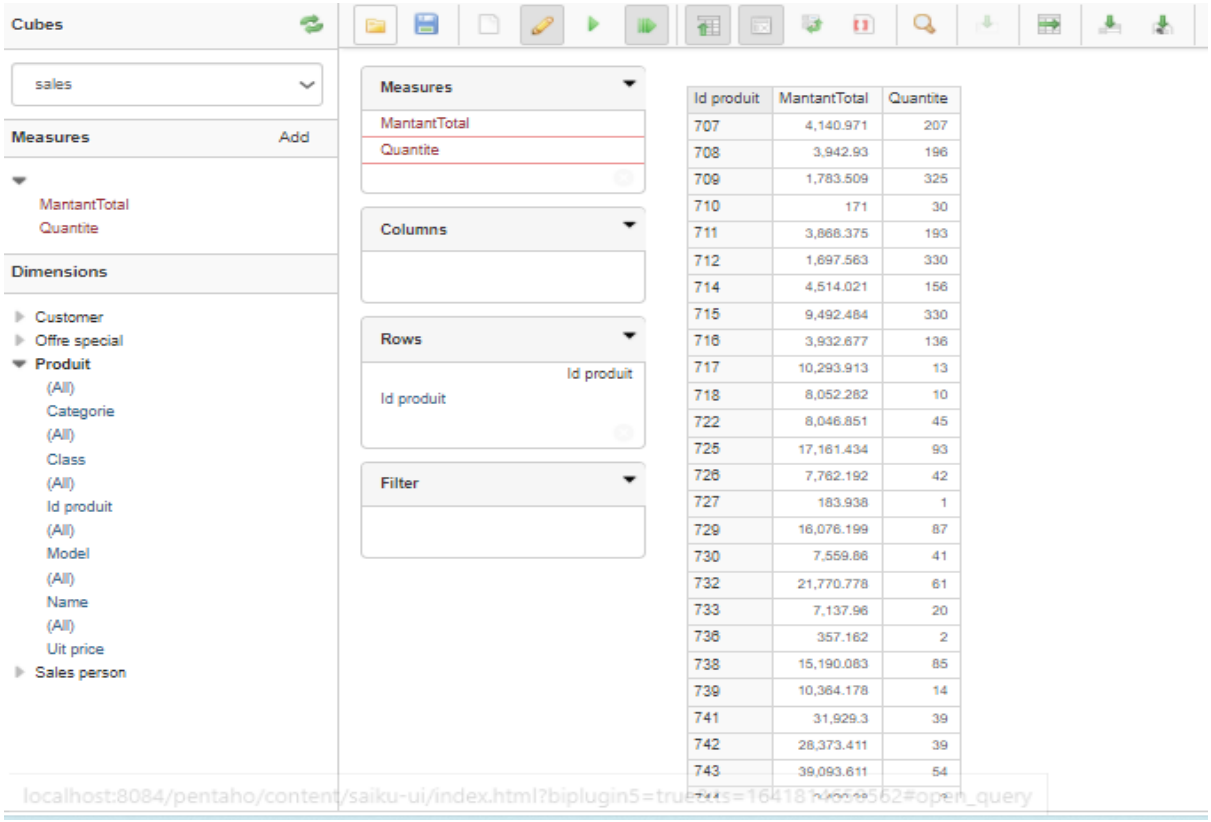


Dimension produit :

Le montant total et quantité totale par Class:

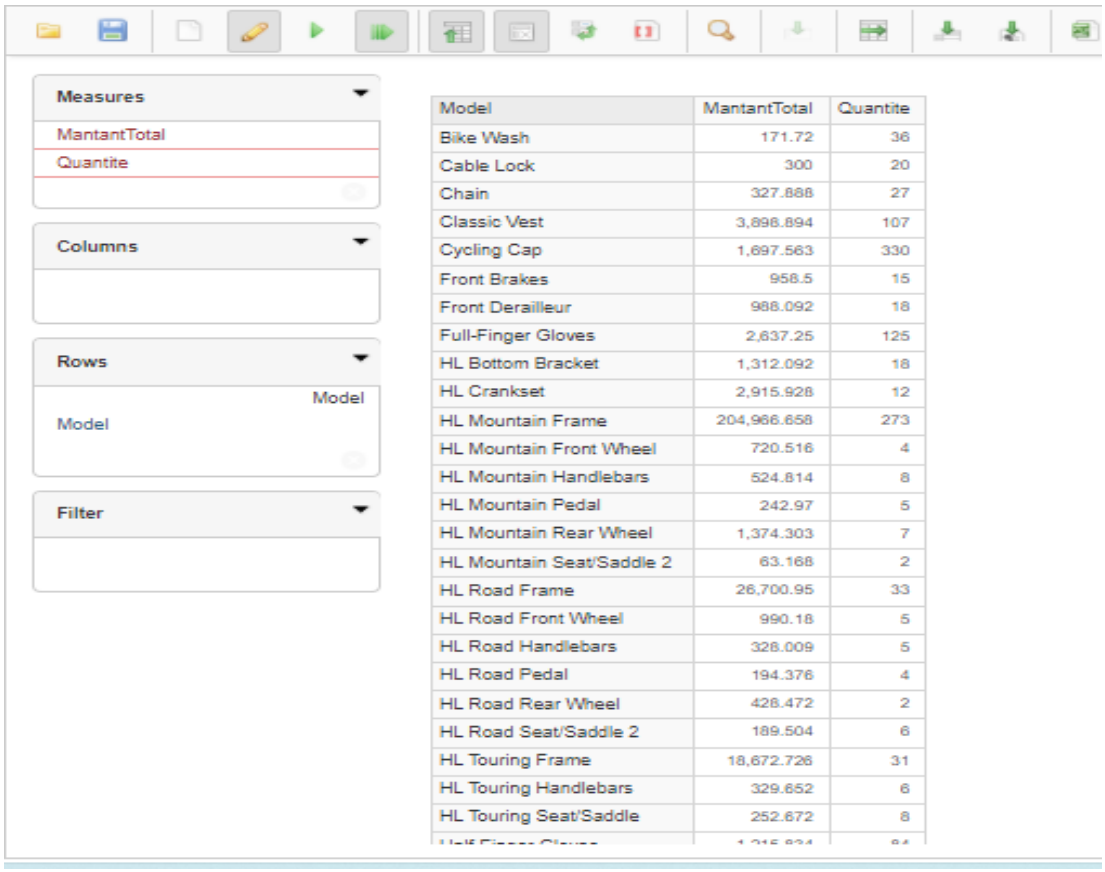


Le montant total par Id_produit:



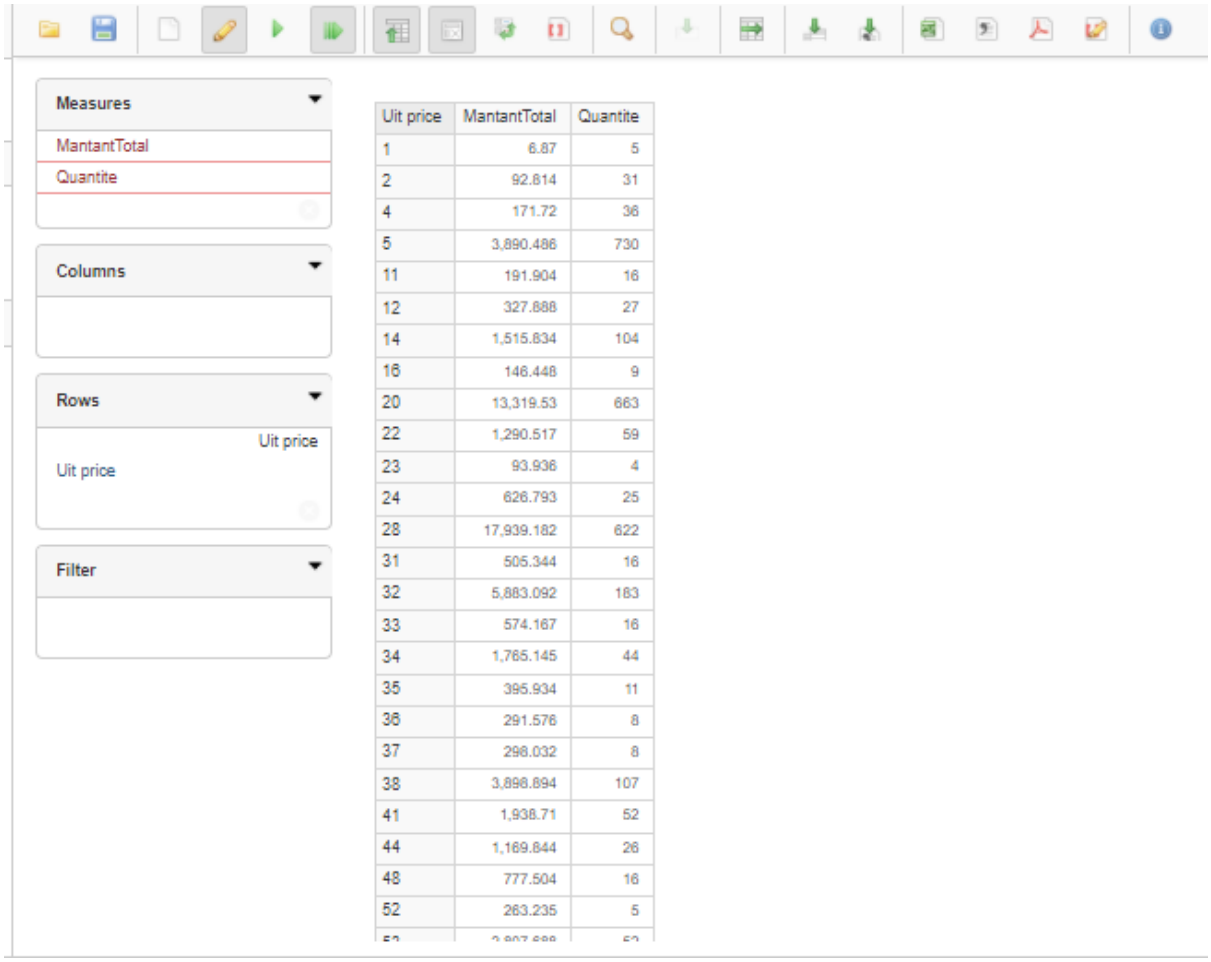
Id produit	MantantTotal	Quantite
707	4,140.971	207
708	3,942.93	196
709	1,783.509	325
710	171	30
711	3,868.375	193
712	1,697.563	330
714	4,514.021	156
715	9,492.484	330
716	3,932.677	136
717	10,293.913	13
718	8,052.282	10
722	8,046.851	45
725	17,161.434	93
726	7,762.192	42
727	183.938	1
729	16,076.199	87
730	7,559.86	41
732	21,770.778	61
733	7,137.96	20
736	357.162	2
738	15,190.083	85
739	10,364.178	14
741	31,929.3	39
742	28,373.411	39
743	39,093.611	54

Le montant total par Model :



Model	MantantTotal	Quantite
Bike Wash	171.72	36
Cable Lock	300	20
Chain	327.888	27
Classic Vest	3,898.894	107
Cycling Cap	1,697.563	330
Front Brakes	958.5	15
Front Derailleur	988.092	18
Full-Finger Gloves	2,637.25	125
HL Bottom Bracket	1,312.092	18
HL Crankset	2,915.928	12
HL Mountain Frame	204,986.658	273
HL Mountain Front Wheel	720.516	4
HL Mountain Handlebars	524.814	8
HL Mountain Pedal	242.97	5
HL Mountain Rear Wheel	1,374.303	7
HL Mountain Seat/Saddle 2	63.168	2
HL Road Frame	26,700.95	33
HL Road Front Wheel	990.18	5
HL Road Handlebars	328.009	5
HL Road Pedal	194.376	4
HL Road Rear Wheel	428.472	2
HL Road Seat/Saddle 2	189.504	6
HL Touring Frame	18,672.726	31
HL Touring Handlebars	329.652	6
HL Touring Seat/Saddle	252.672	8

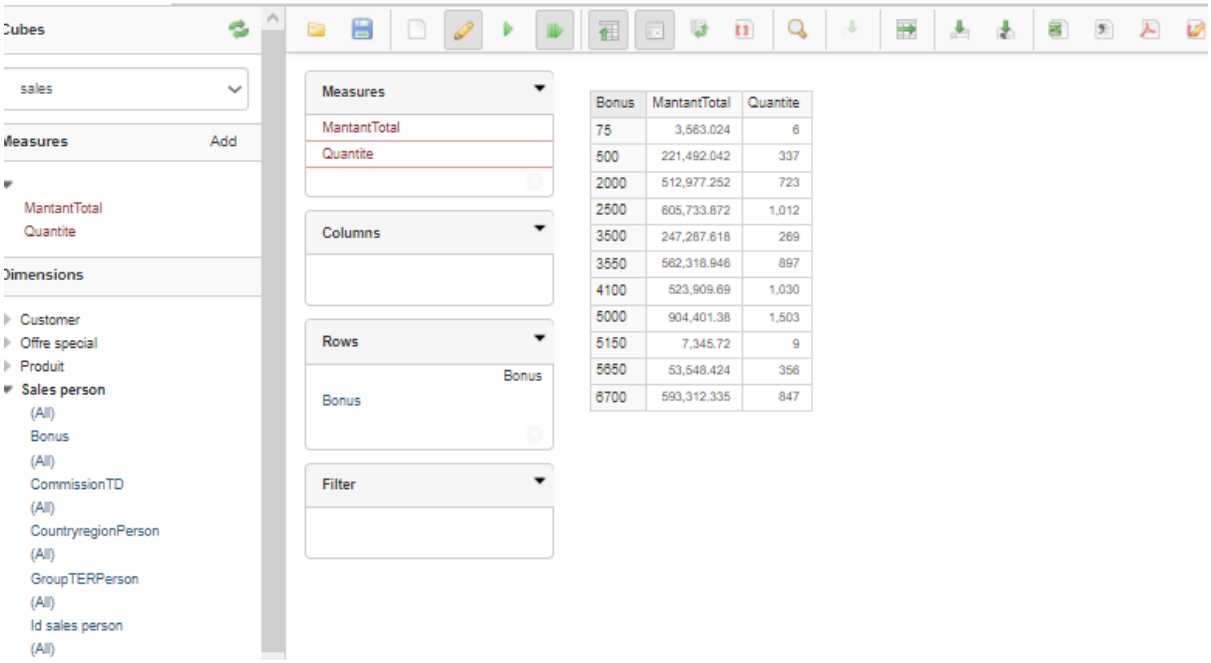
Le montant total et quantité totale par uitPrice :



Uit price	MantantTotal	Quantite
1	6.87	5
2	92.814	31
4	171.72	36
5	3,890.486	730
11	191.904	16
12	327.888	27
14	1,515.834	104
16	146.448	9
20	13,319.53	663
22	1,290.517	59
23	93.936	4
24	626.793	25
28	17,939.182	622
31	505.344	16
32	5,883.092	183
33	574.167	16
34	1,765.145	44
35	395.934	11
36	291.576	8
37	298.032	8
38	3,898.894	107
41	1,938.71	52
44	1,169.844	26
48	777.504	16
52	263.235	5

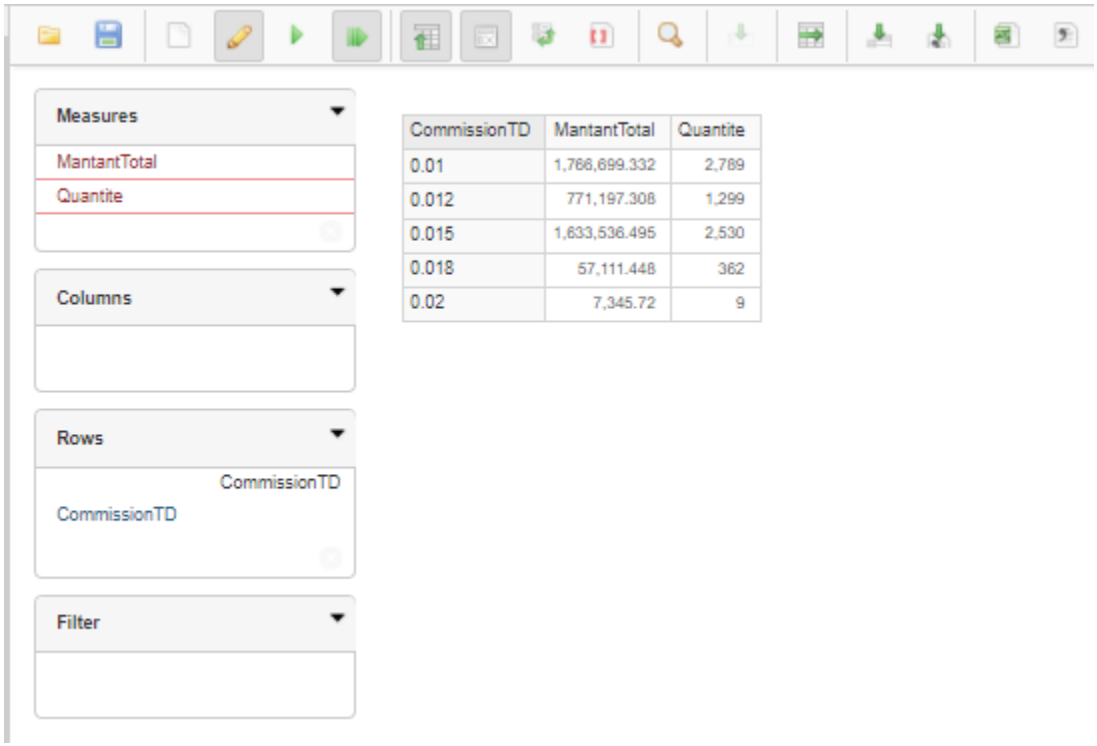
Dimension salesPerson :

Le montant total et quantité totale par Bonus :



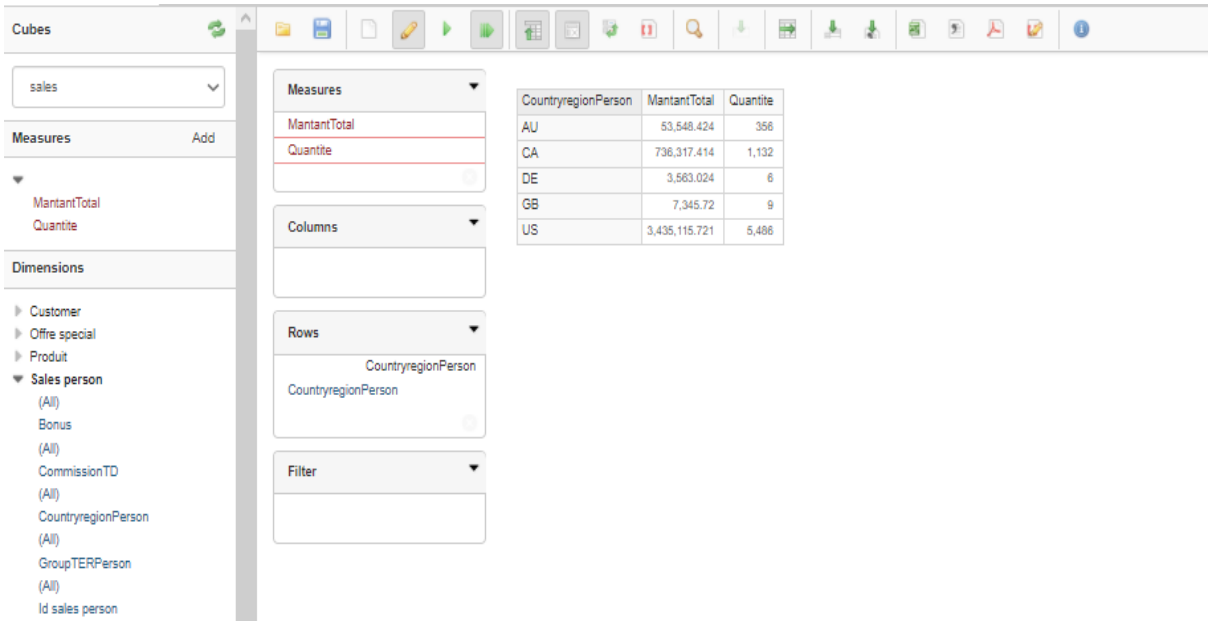
Bonus	MantantTotal	Quantite
75	3,563.024	6
500	221,492.042	337
2000	512,977.252	723
2500	605,733.872	1,012
3500	247,287.618	269
3550	562,318.946	897
4100	523,909.69	1,030
5000	904,401.38	1,503
5150	7,345.72	9
5650	53,548.424	356
6700	593,312.335	847

Le montant total et quantité totale par commissionTD :



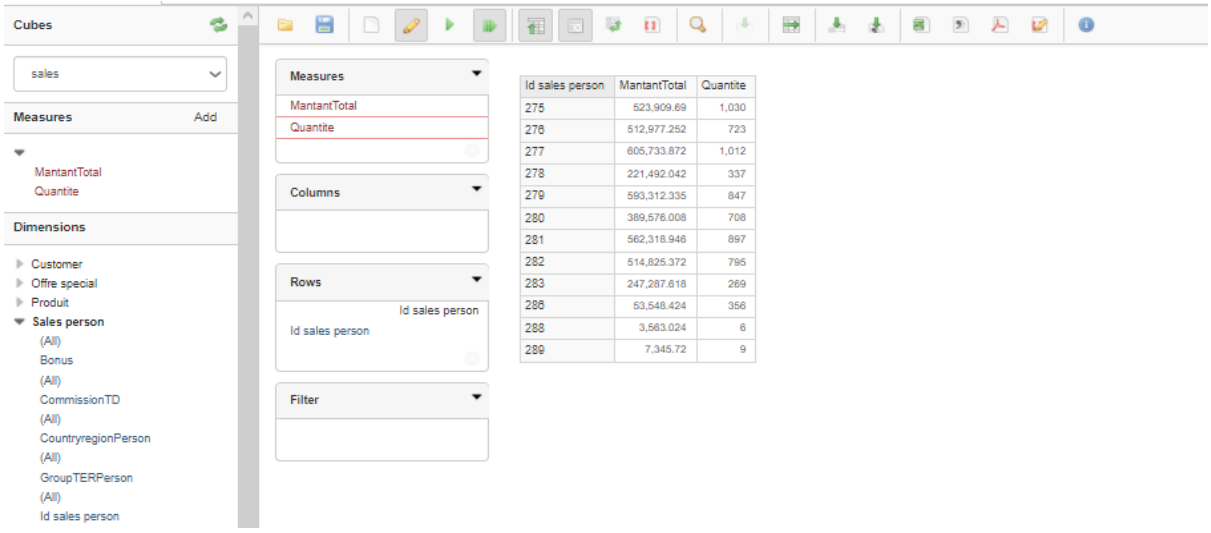
CommissionTD	MantantTotal	Quantite
0.01	1,766,699.332	2,789
0.012	771,197.308	1,299
0.015	1,633,536.495	2,530
0.018	57,111.448	362
0.02	7,345.72	9

Le montant total et quantité totale par contryRegionPerson :



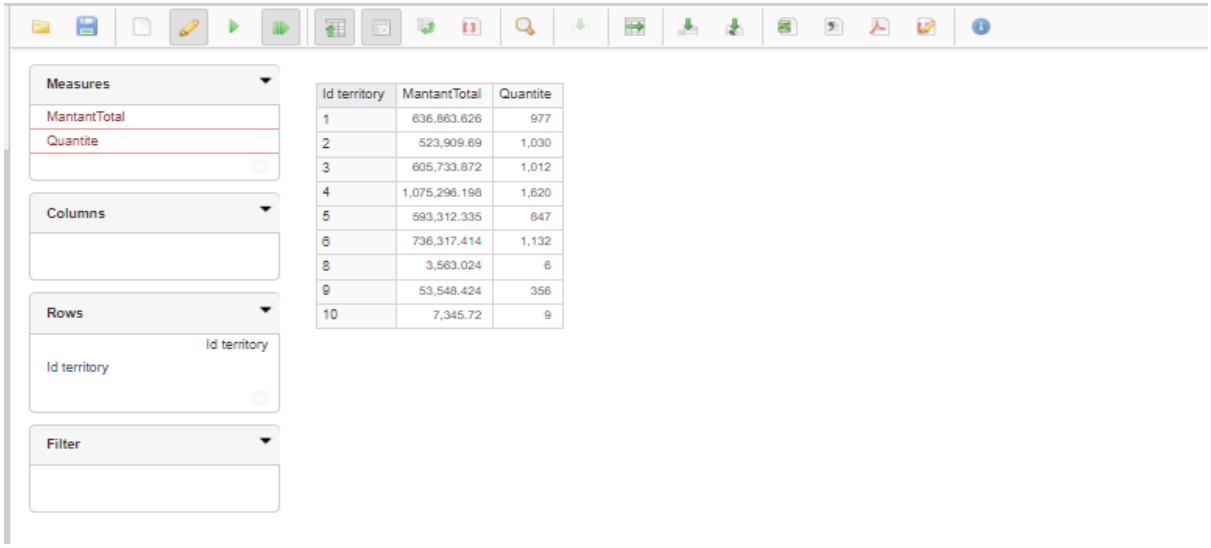
CountryregionPerson	MantantTotal	Quantite
AU	53,548.424	356
CA	736,317.414	1,132
DE	3,563.024	6
GB	7,345.72	9
US	3,435,115.721	5,486

Le montant total et quantité totale par idSalesPerson :



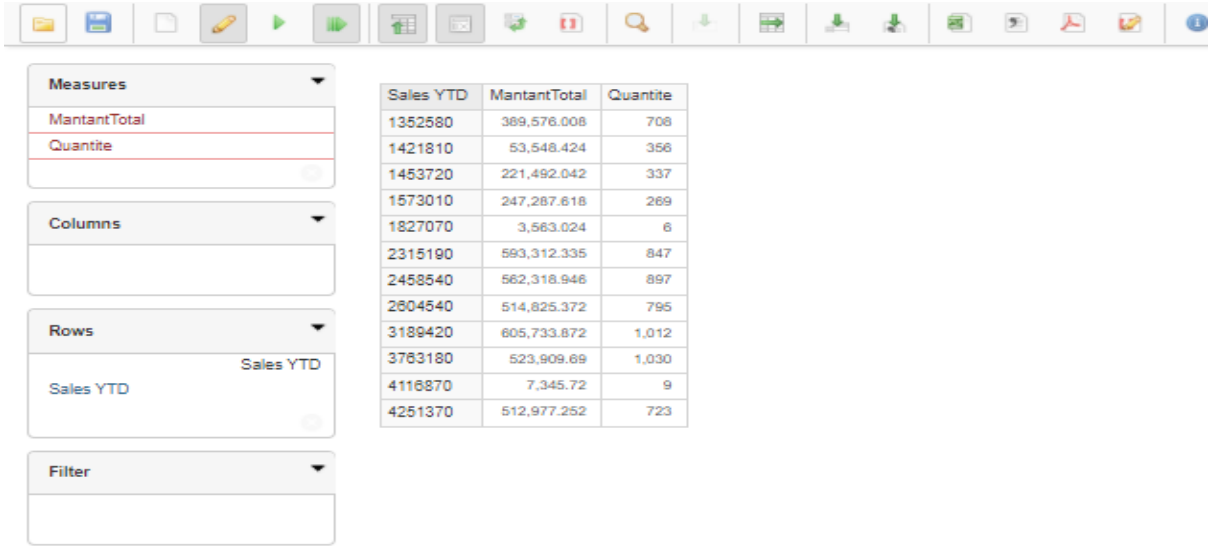
Id sales person	MantantTotal	Quantite
275	523,909.69	1,030
276	512,977.252	723
277	605,733.872	1,012
278	221,492.042	337
279	593,312.335	847
280	389,576.008	708
281	562,318.946	897
282	514,825.372	795
283	247,287.618	269
286	53,548.424	356
288	3,563.024	6
289	7,345.72	9

Le montant total et quantité totale par idTerritory :



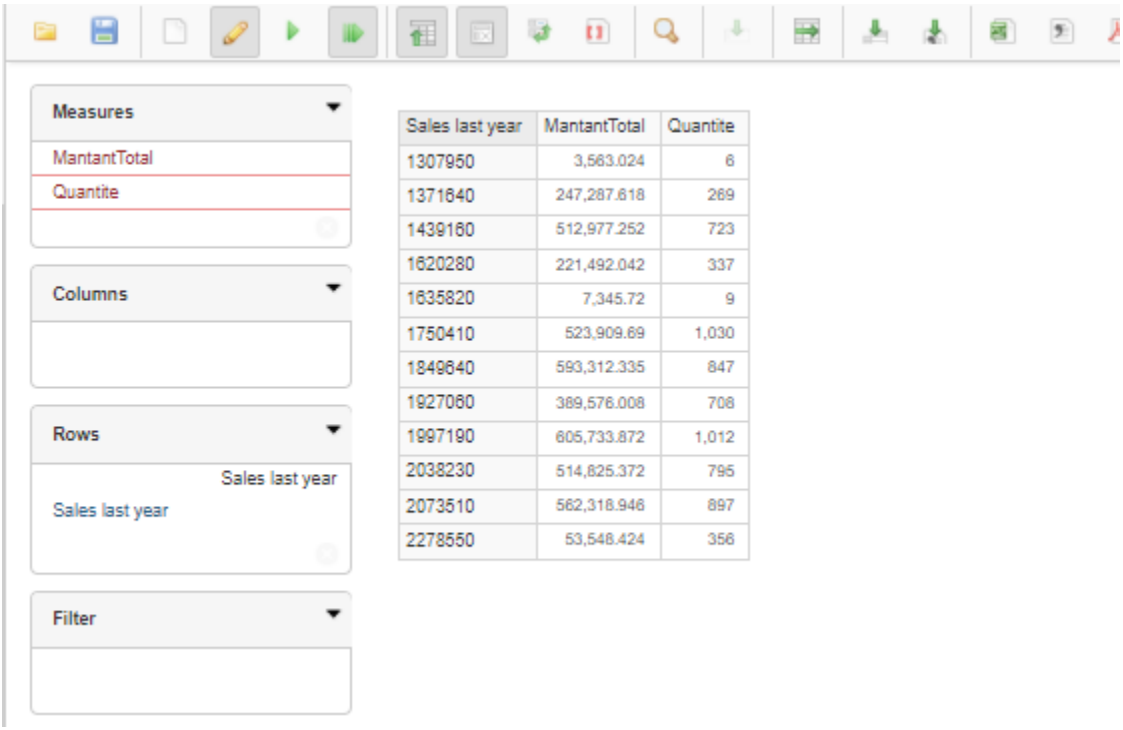
Id territory	MantantTotal	Quantite
1	636,863.626	977
2	523,909.69	1,030
3	605,733.872	1,012
4	1,075,296.198	1,620
5	593,312.335	847
6	736,317.414	1,132
8	3,563.024	6
9	53,548.424	356
10	7,345.72	9

Le montant total et quantité totale par SalesYTD :



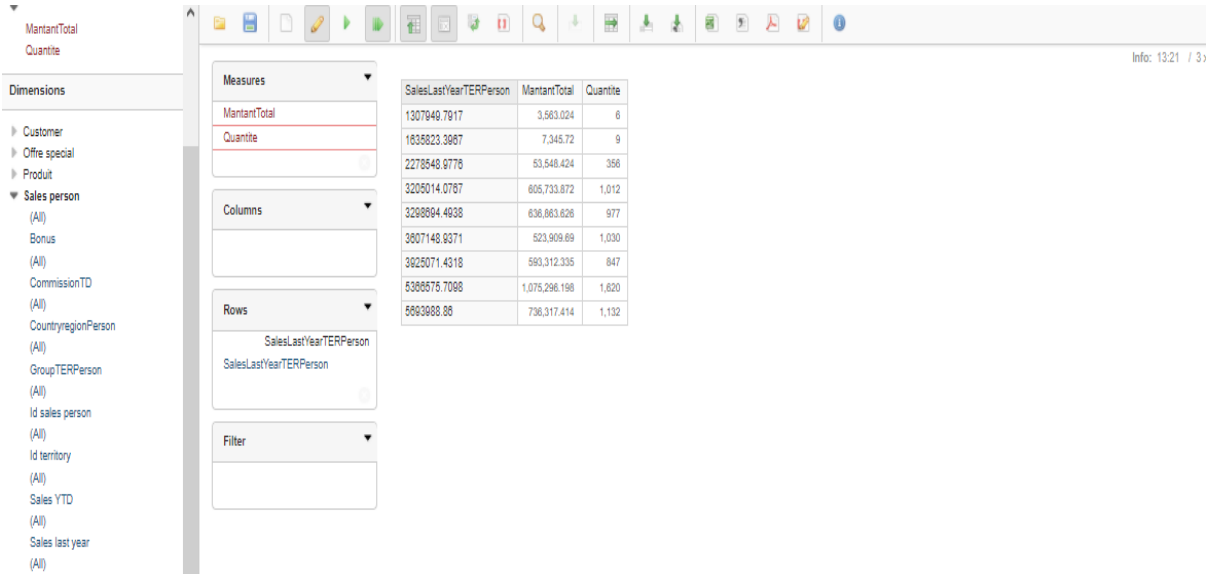
Sales YTD	MantantTotal	Quantite
1352580	389,576.008	708
1421810	53,548.424	356
1453720	221,492.042	337
1573010	247,287.618	269
1827070	3,563.024	6
2315190	593,312.335	847
2458540	562,318.946	897
2604540	514,825.372	795
3189420	605,733.872	1,012
3763180	523,909.69	1,030
4116870	7,345.72	9
4251370	512,977.252	723

Le montant total et quantité totale par SalesLAstYear:



Sales last year	MantantTotal	Quantite
1307950	3,563.024	6
1371640	247,287.618	269
1439160	512,977.252	723
1620280	221,492.042	337
1635820	7,345.72	9
1750410	523,909.69	1,030
1849640	593,312.335	847
1927060	389,576.008	708
1997190	605,733.872	1,012
2038230	514,825.372	795
2073510	562,318.946	897
2278550	53,548.424	356

Le montant total et quantité totale par SalesLastTERPerson :



SalesLastYearTERPerson	MantantTotal	Quantite
1307949.7917	3,563.024	6
1635823.3967	7,345.72	9
2278548.9776	53,548.424	356
3205014.0767	605,733.872	1,012
3298694.4938	636,663.626	977
3607148.9371	523,909.69	1,030
3925071.4318	593,312.335	847
5366575.7098	1,075,296.198	1,620
5693988.88	736,317.414	1,132

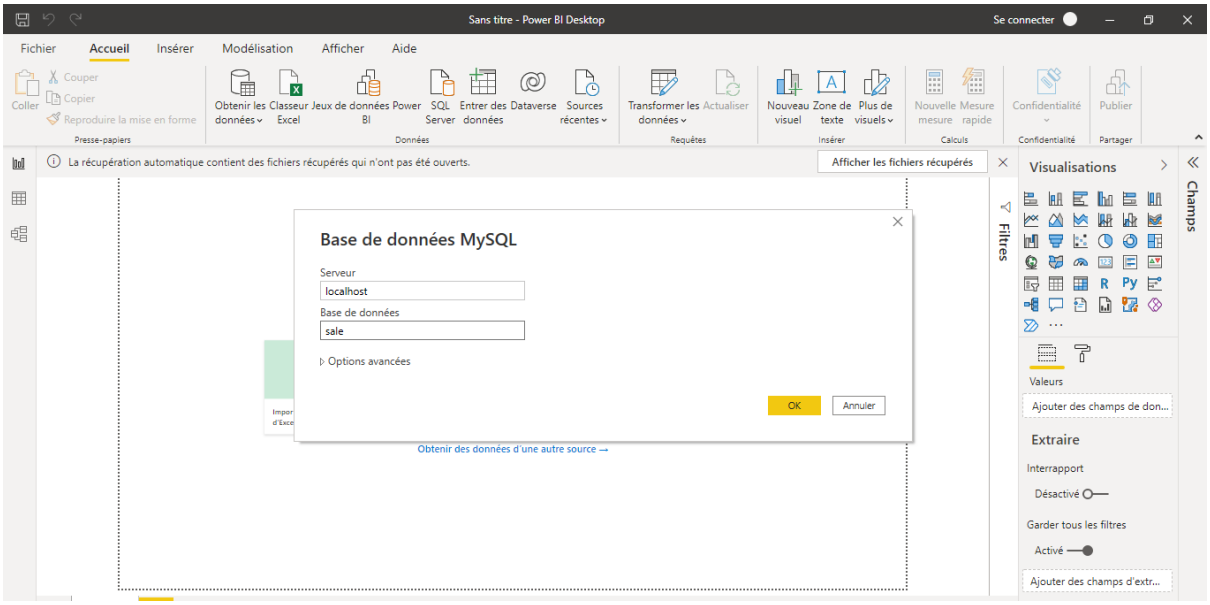
visualisation des donnees a l'aide de power BI

Pourquoi power BI ?

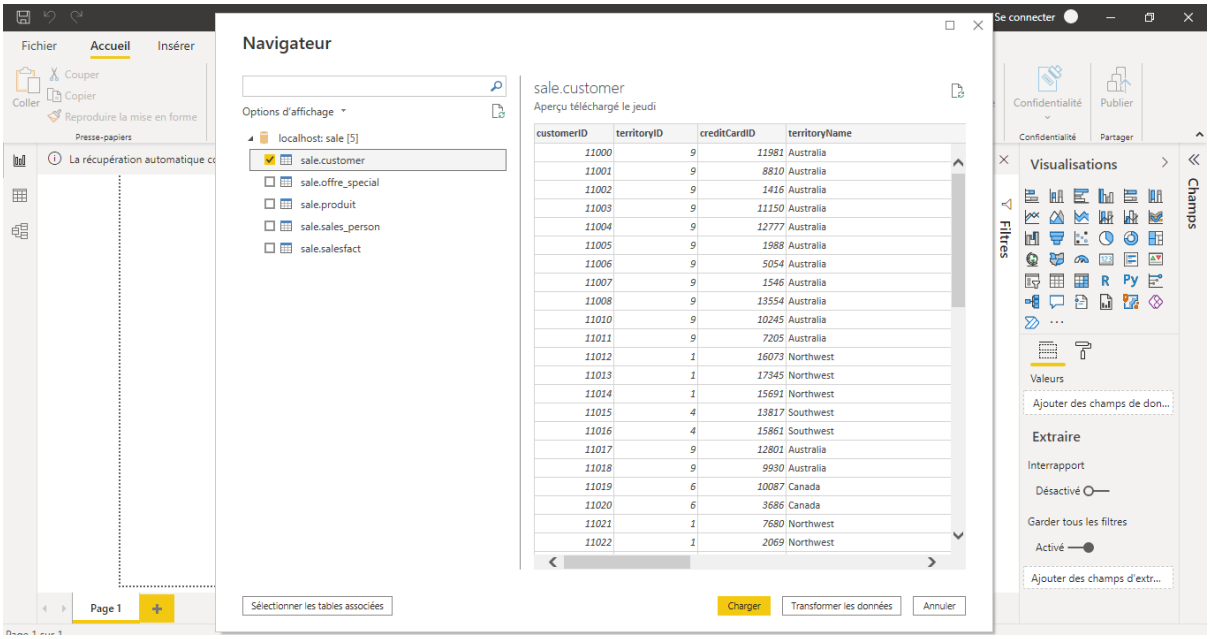
Power BI est un outil de visualisation de données et d'aide à la décision qui convertit les données de différentes sources de données en tableaux de bord interactifs et en rapports d'aide à la décision. La suite Power BI fournit plusieurs logiciels, connecteurs et services : Power BI Desktop, le service Power BI basé sur Saas et les applications Power BI mobiles disponibles pour différentes plateformes. Cet ensemble de services est utilisé par les utilisateurs professionnels pour consommer des données et créer des rapports d'aide à la décision.

Dans notre cas on a utiliser power BI pour faire des visualisations des ventes (quantité ,montant total) en fonction des dimensions .

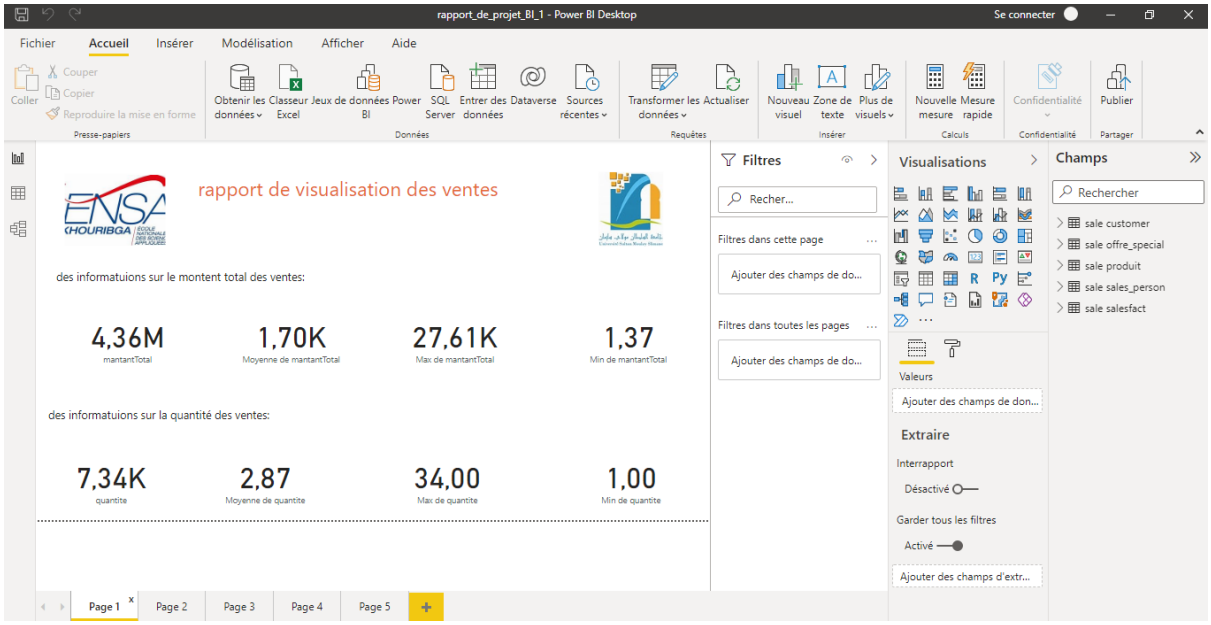
1. Obtenir les donnees apartir de la base de donnees mysql :



Selectionner et charger les données .

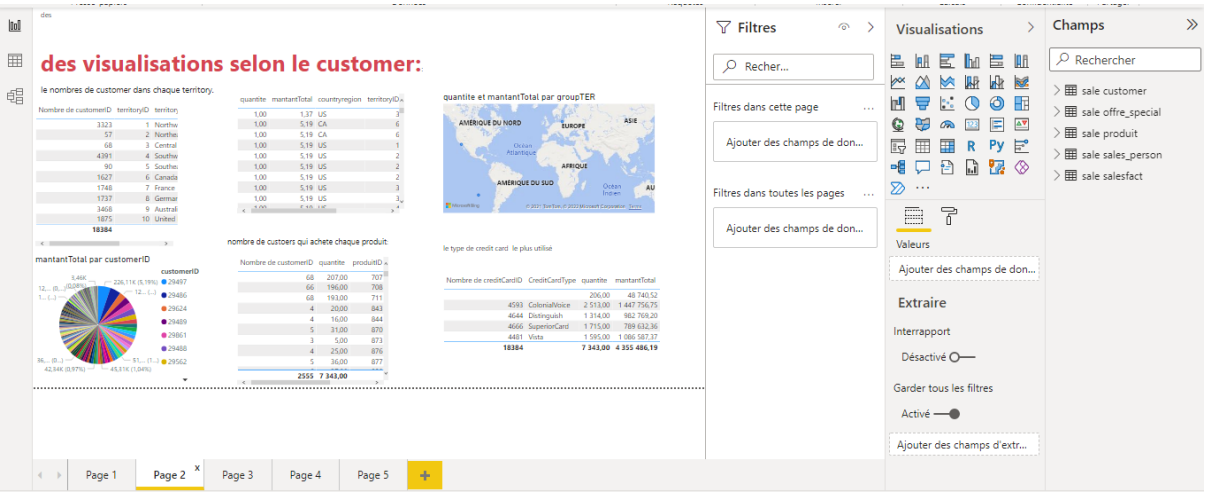


2. Extractions des informations global sur la quantité de vente et le montant total de vente(somme,moyenne,max ,min) :



3. Faire des visualisation selon chacune des dimensions :

   a. Selon le client (customer) :



- On determine le nombre de client par territory :

Retour au rapport

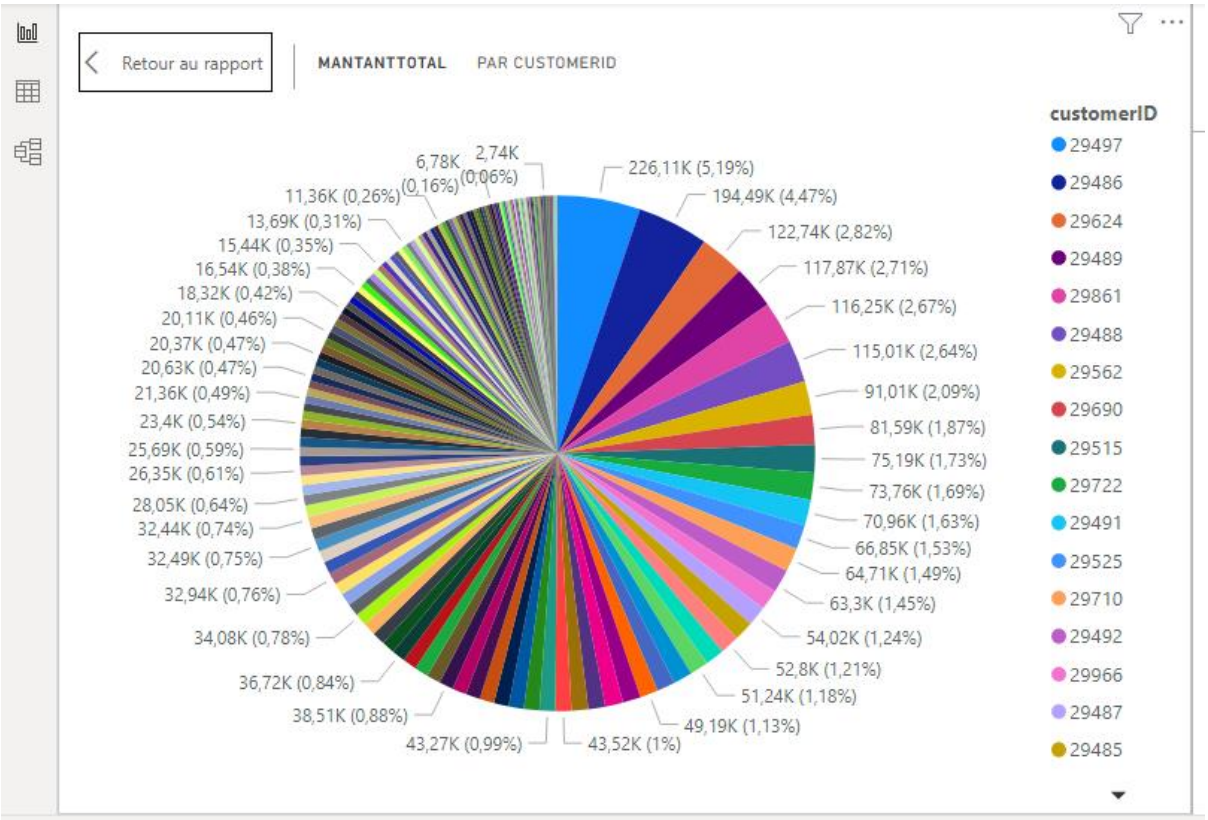
Nombre de customerID	territoryID	territoryName	countryregion
3323	1	Northwest	US
57	2	Northeast	US
68	3	Central	US
4391	4	Southwest	US
90	5	Southeast	US
1627	6	Canada	CA
1748	7	France	FR
1737	8	Germany	DE
3468	9	Australia	AU
1875	10	United Kingdom	GB
18384			

- La quantité total et le montant total de vente pour chaque client :

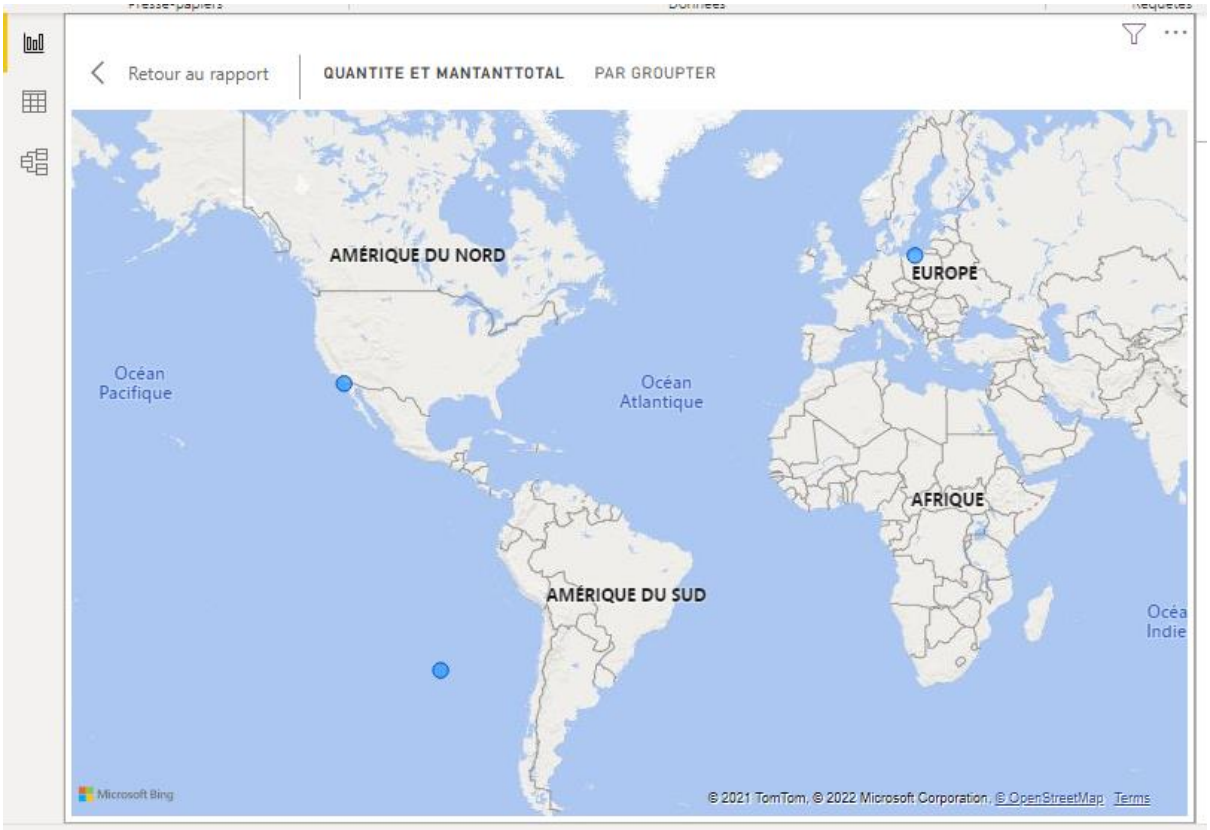
Presse-papiers | Données | Requêtes

Retour au rapport

quantite	mantantTotal	countryregion	territoryID	territoryName	customerID
206,00	48 740,52				29484
212,00	53 081,00	US	4	Southwest	29485
459,00	194 494,00	US	3	Central	29486
75,00	54 023,31	US	2	Northeast	29487
663,00	115 005,95	AU	9	Australia	29488
316,00	117 870,38	US	4	Southwest	29489
13,00	6 378,12	US	1	Northwest	29490
71,00	70 960,11	US	5	Southeast	29491
162,00	63 300,92	US	3	Central	29492
6,00	5 670,23	US	5	Southeast	29493
14,00	16 198,29	CA	6	Canada	29494
6,00	3 563,02	DE	8	Germany	29495
99,00	50 288,22	US	4	Southwest	29496
477,00	226 112,44	US	1	Northwest	29497
1,00	874,79	US	2	Northeast	29498
30,00	16 374,90	US	5	Southeast	29504
53,00	52 803,45	US	5	Southeast	29507
19,00	6 783,77	US	2	Northeast	29509
83,00	21 720,07	CA	6	Canada	29510
40,00	39 450,79	US	3	Central	29511
46,00	75 191,96	CA	6	Canada	29515
7 343,00	**4 355 486,19**				



- Quantité et le montant total par groupTER :



- Le produit le plus acheté par les clients avec sa quantité et ses informations :

Retour au rapport

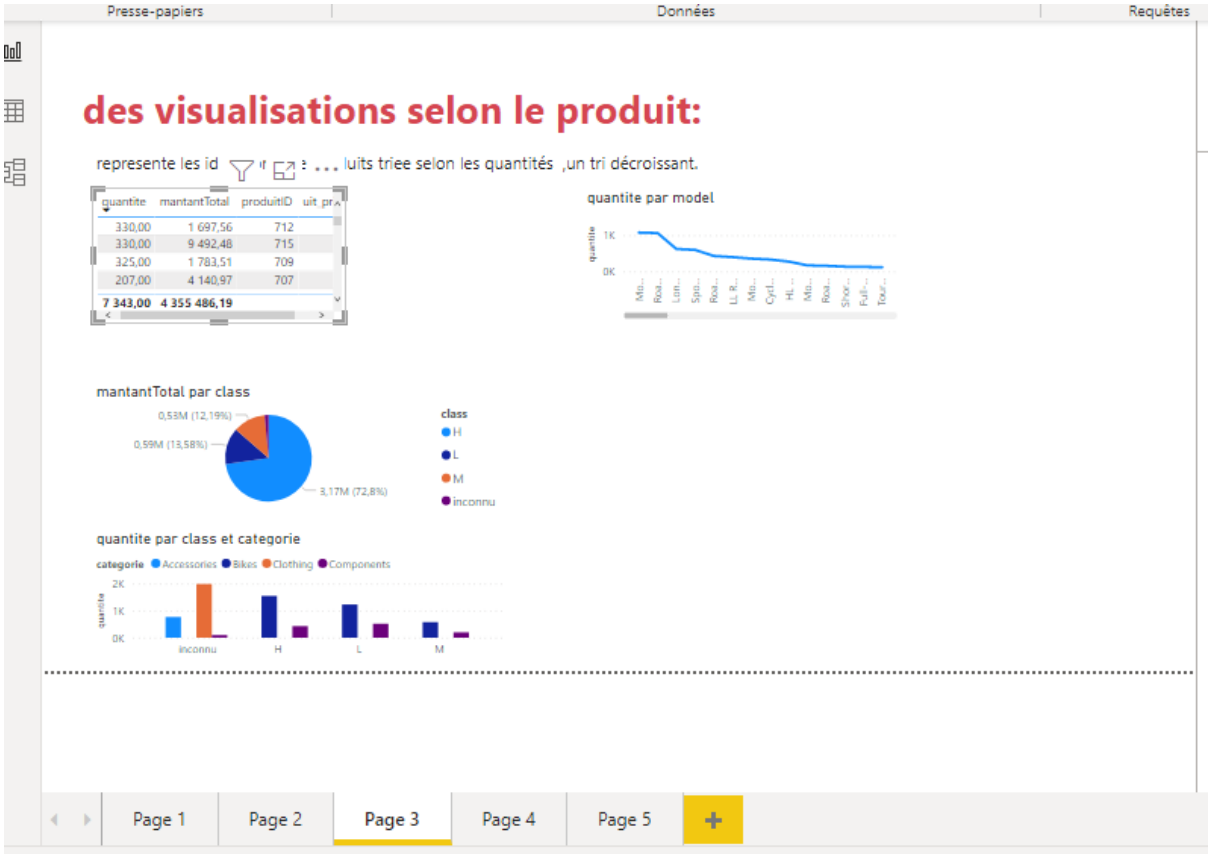
Nombre de customerID	quantite	produitID	categorie	class
68	207,00	707	Accessories	inconnu
66	196,00	708	Accessories	inconnu
68	193,00	711	Accessories	inconnu
4	20,00	843	Accessories	inconnu
4	16,00	844	Accessories	inconnu
5	31,00	870	Accessories	inconnu
3	5,00	873	Accessories	inconnu
4	25,00	876	Accessories	inconnu
5	36,00	877	Accessories	inconnu
6	37,00	880	Accessories	inconnu
20	40,00	749	Bikes	H
3	9,00	750	Bikes	H
5	11,00	751	Bikes	H
4	8,00	752	Bikes	H
47	88,00	753	Bikes	H
49	123,00	754	Bikes	M
37	71,00	755	Bikes	M
37	58,00	756	Bikes	M
5	11,00	757	Bikes	M
63	165,00	758	Bikes	M
5	5,00	759	Bikes	L
2555	**7 343,00**			

- Le type de credit card le plus utilisé par le client avec la quantité et le montant total acheté par cette carte :

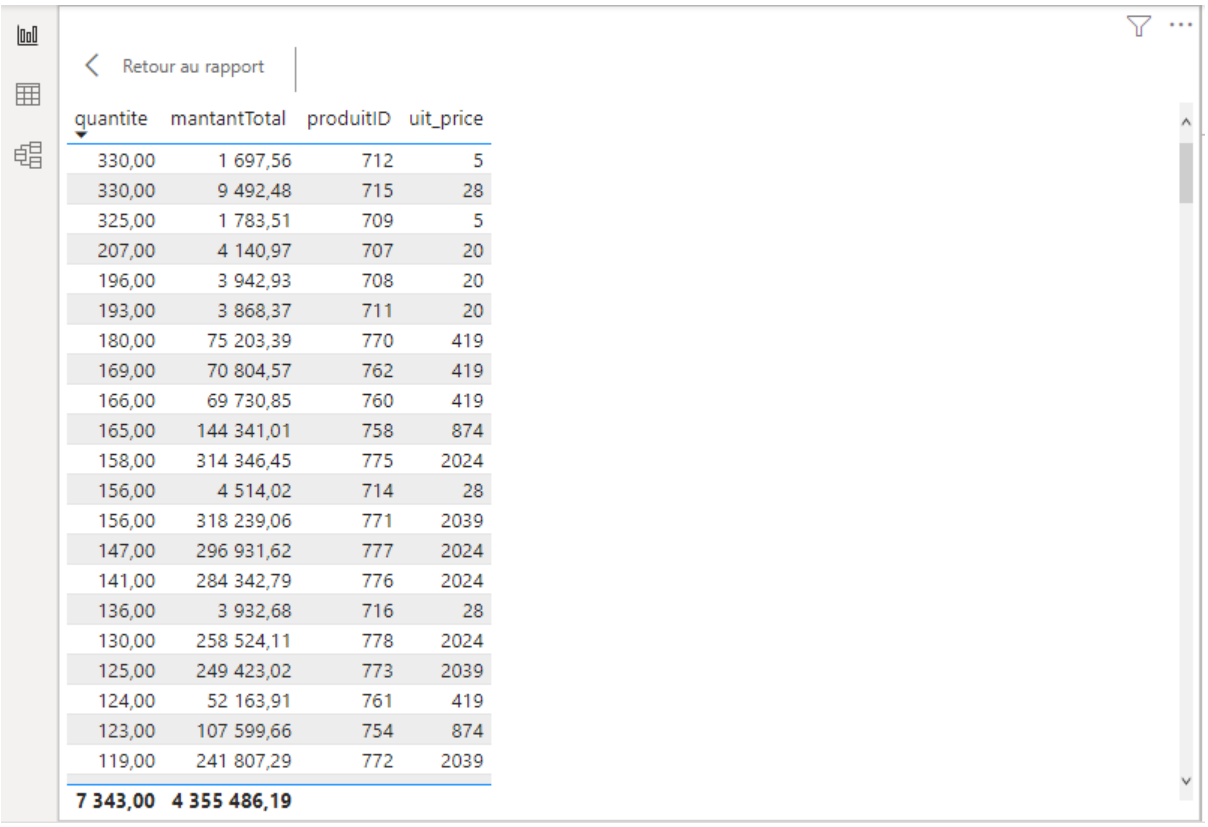
Retour au rapport

Nombre de creditCardID	CreditCardType	quantite	mantantTotal
		206,00	48 740,52
4593	ColonialVoice	2 513,00	1 447 756,75
4644	Distinguish	1 314,00	982 769,20
4666	SuperiorCard	1 715,00	789 632,36
4481	Vista	1 595,00	1 086 587,37
18384		**7 343,00**	**4 355 486,19**

b) selon le produit:



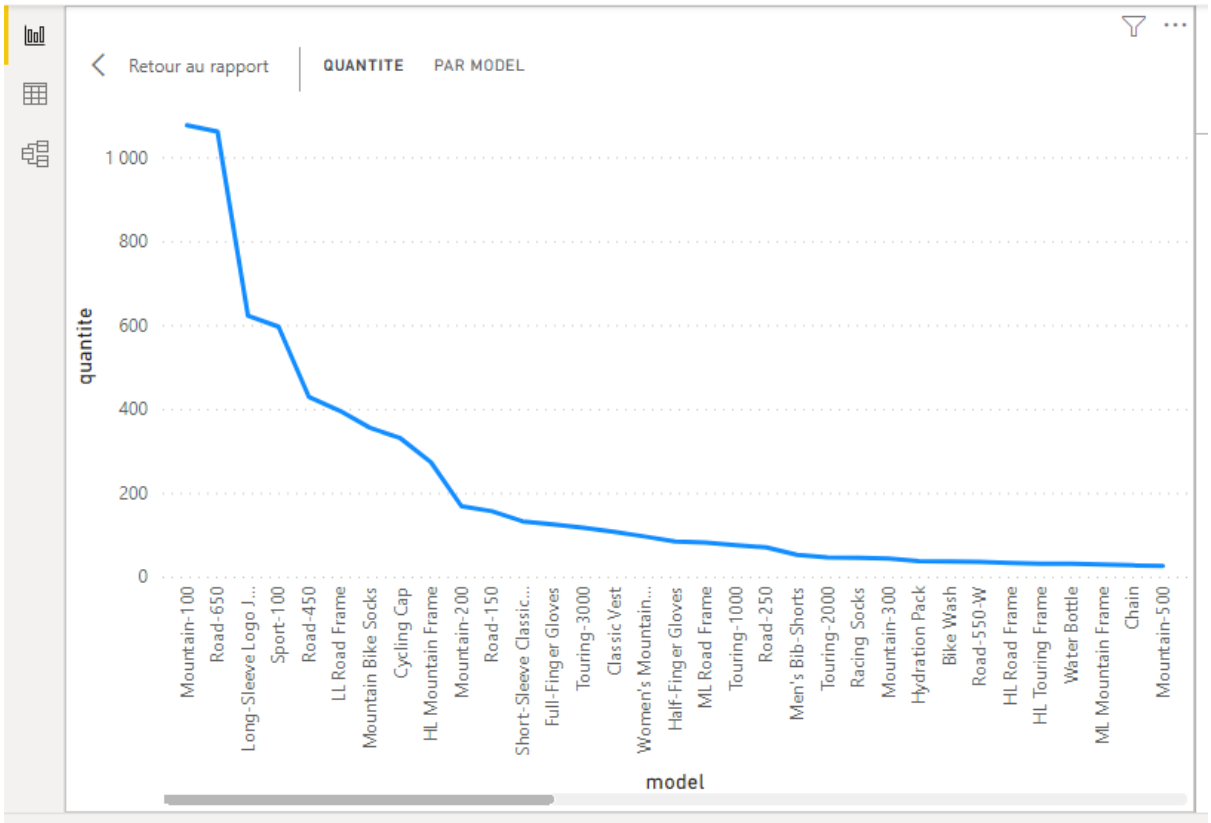
- quantité et le montant de produit avec un tri décroissant :



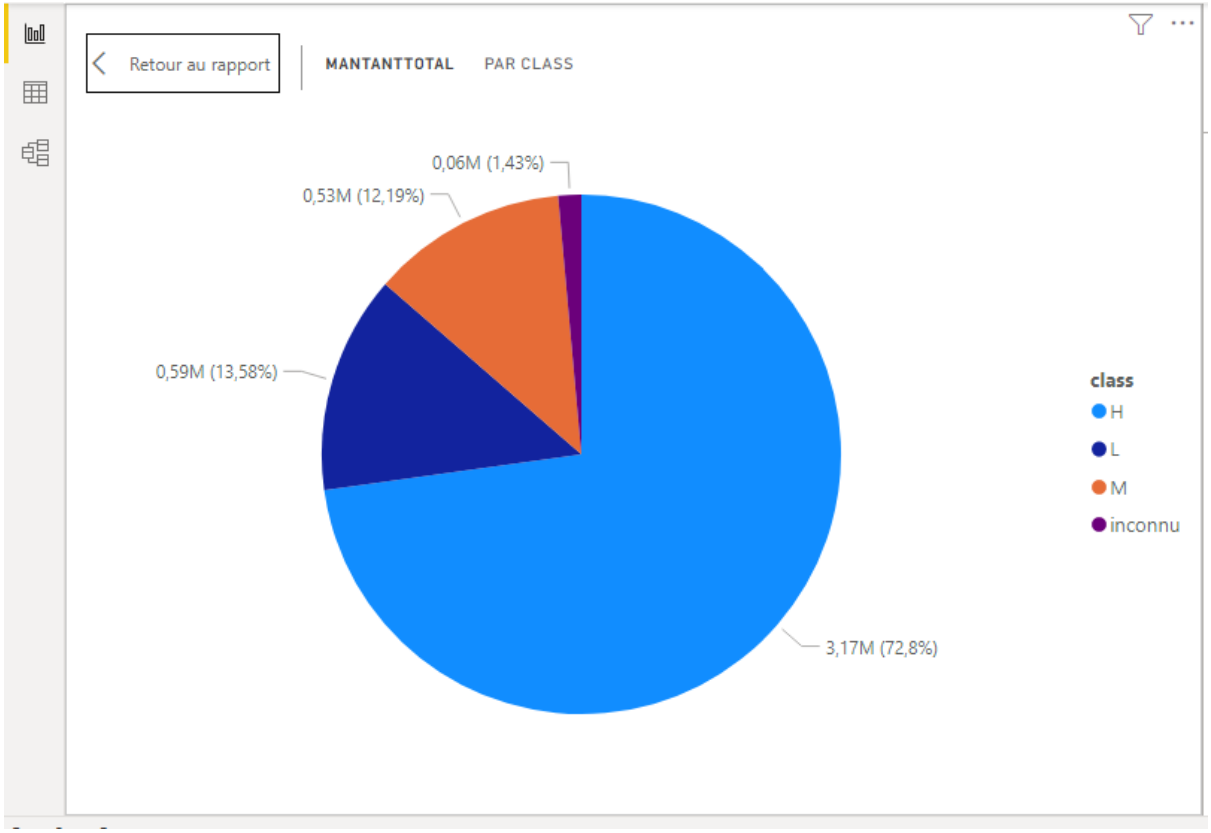
Retour au rapport

quantite	mantantTotal	produitID	uit_price
330,00	1 697,56	712	5
330,00	9 492,48	715	28
325,00	1 783,51	709	5
207,00	4 140,97	707	20
196,00	3 942,93	708	20
193,00	3 868,37	711	20
180,00	75 203,39	770	419
169,00	70 804,57	762	419
166,00	69 730,85	760	419
165,00	144 341,01	758	874
158,00	314 346,45	775	2024
156,00	4 514,02	714	28
156,00	318 239,06	771	2039
147,00	296 931,62	777	2024
141,00	284 342,79	776	2024
136,00	3 932,68	716	28
130,00	258 524,11	778	2024
125,00	249 423,02	773	2039
124,00	52 163,91	761	419
123,00	107 599,66	754	874
119,00	241 807,29	772	2039
7 343,00	**4 355 486,19**		

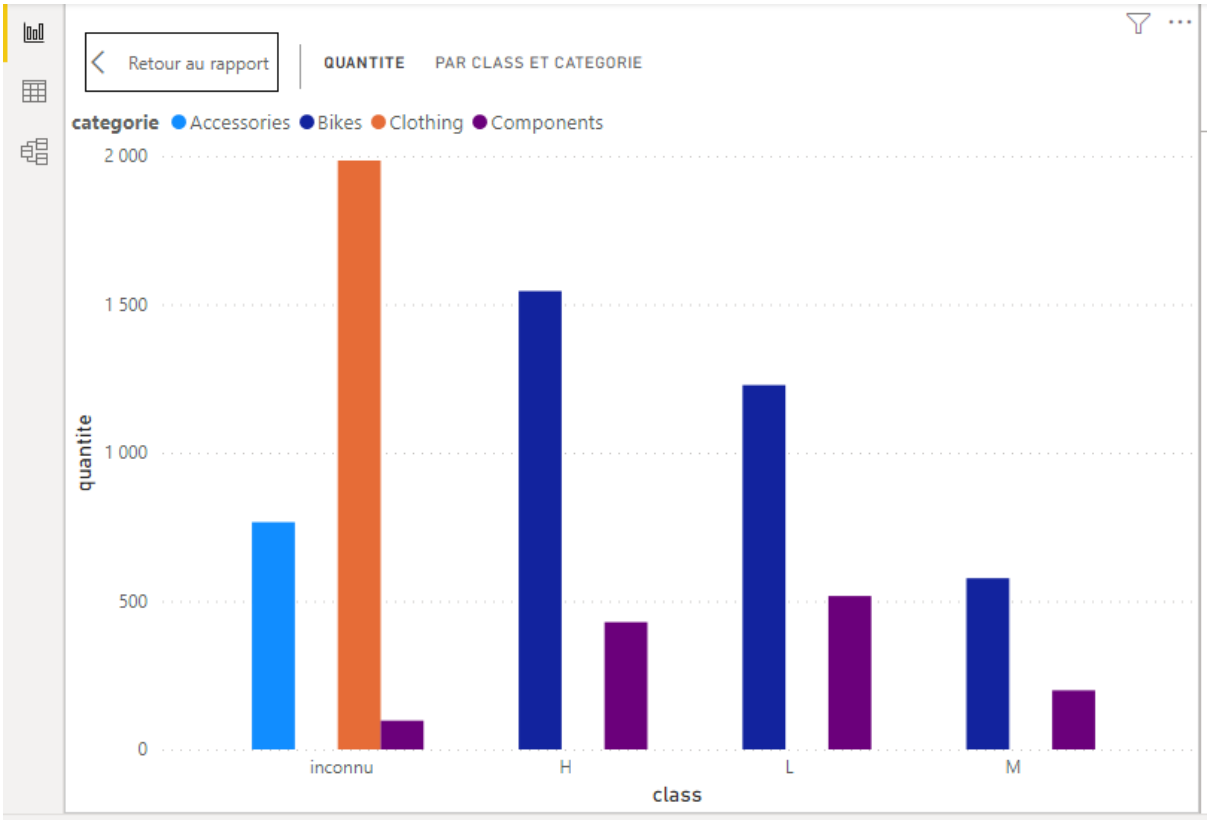
- quantité par model :



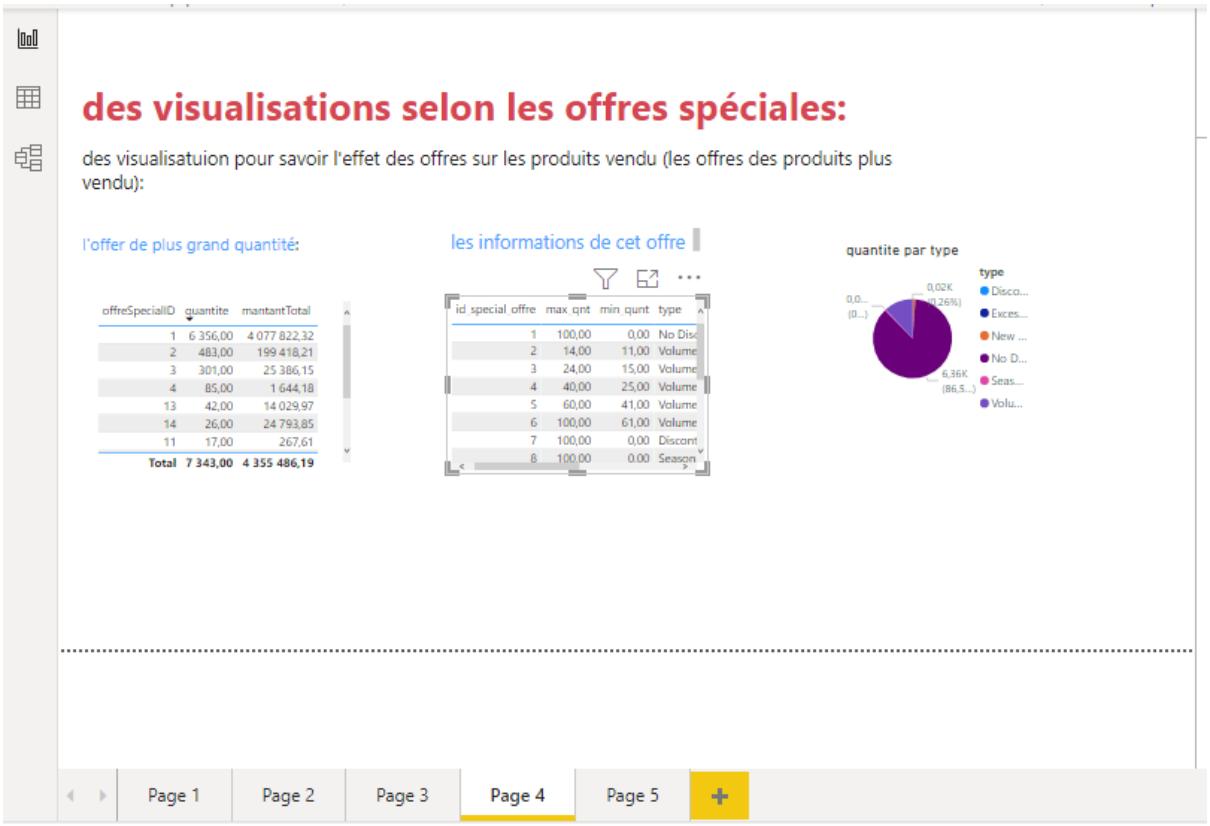
- montant total par class :



- quantité par classe et par catégorie :



c)selon les offres spéciales :



- l’offre de plus grande quantité vendue :

Retour au rapport

offreSpecialID	quantite	mantantTotal
1	6 356,00	4 077 822,32
2	483,00	199 418,21
3	301,00	25 386,15
4	85,00	1 644,18
13	42,00	14 029,97
14	26,00	24 793,85
11	17,00	267,61
7	13,00	10 993,11
8	13,00	218,66
16	6,00	677,35
9	1,00	234,81
Total	**7 343,00**	**4 355 486,19**

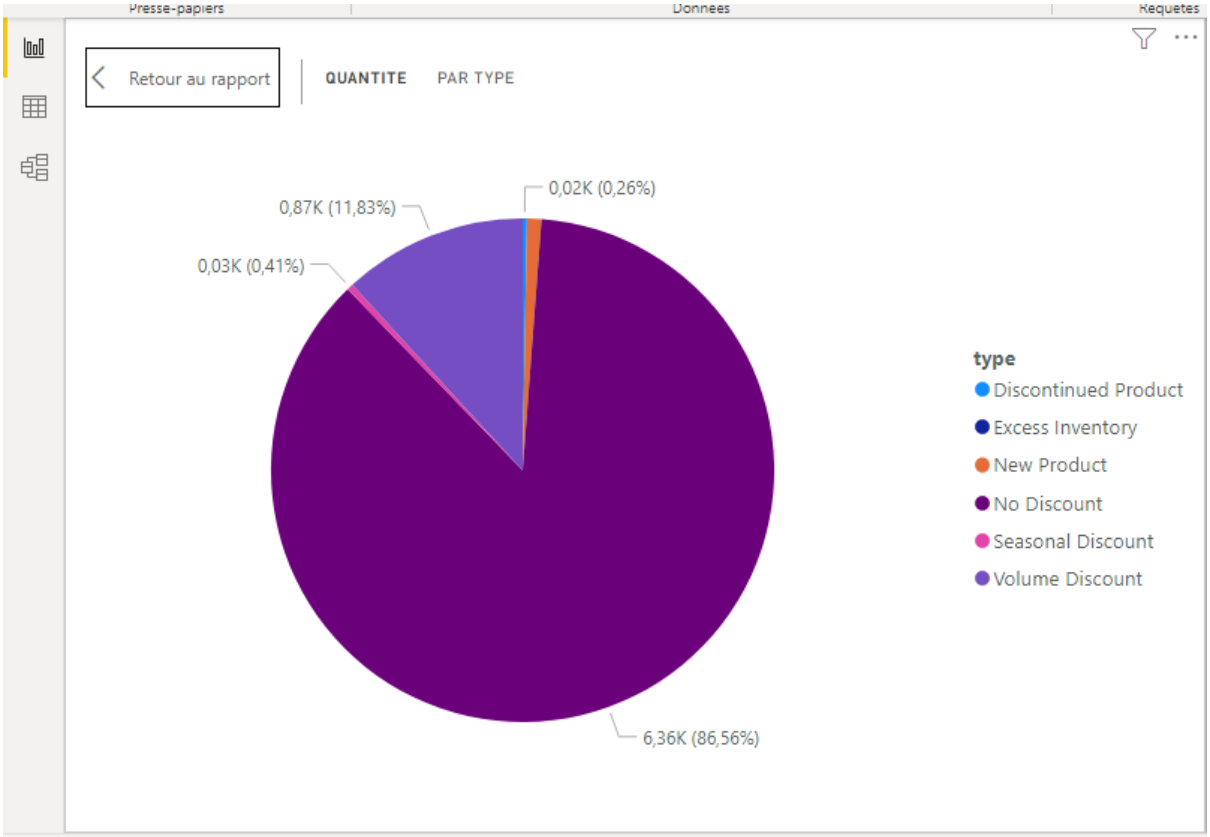
- les informations de cet offres :

Presse-papiers Données

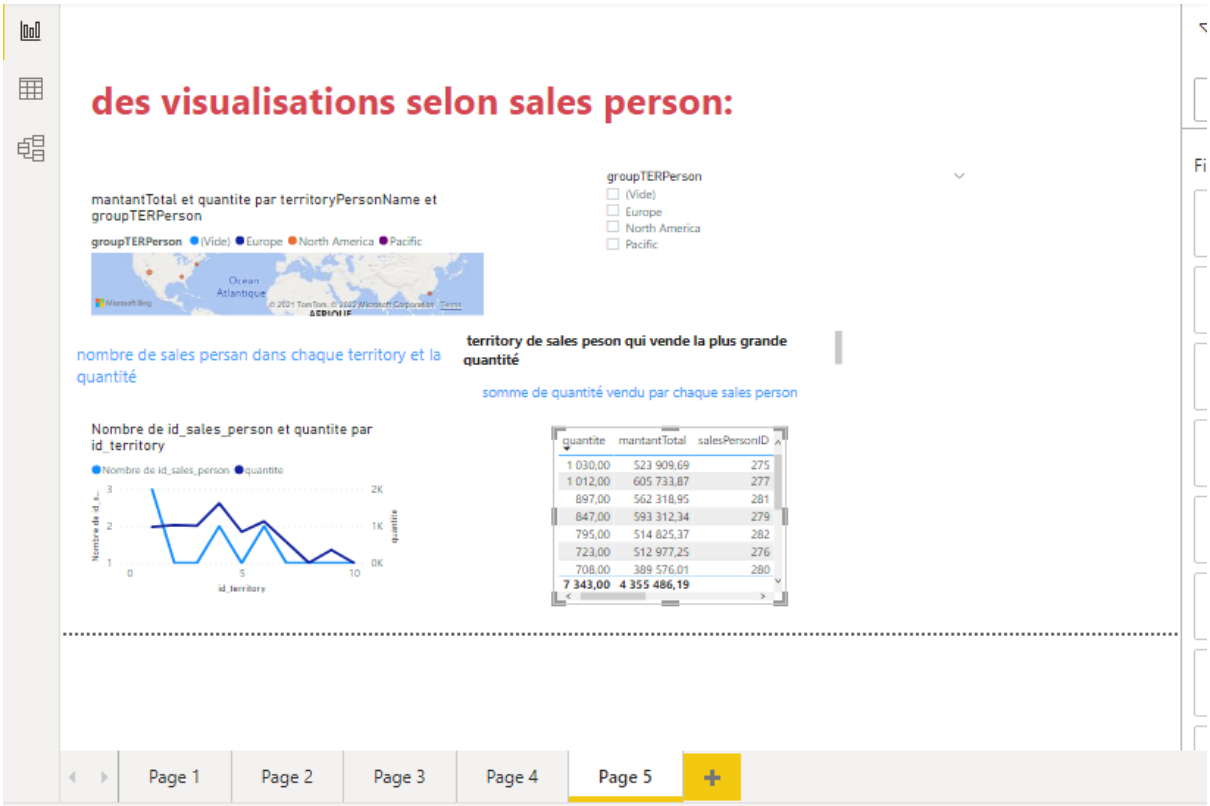
Retour au rapport

id_special_offre	max_qnt	min_qunt	type	discontpct	category
1	100,00	0,00	No Discount	0	No Discount
2	14,00	11,00	Volume Discount	0	Reseller
3	24,00	15,00	Volume Discount	0	Reseller
4	40,00	25,00	Volume Discount	0	Reseller
5	60,00	41,00	Volume Discount	0	Reseller
6	100,00	61,00	Volume Discount	0	Reseller
7	100,00	0,00	Discontinued Product	0	Reseller
8	100,00	0,00	Seasonal Discount	0	Reseller
9	100,00	0,00	Excess Inventory	0	Reseller
10	100,00	0,00	Excess Inventory	0	Customer
11	100,00	0,00	Seasonal Discount	0	Reseller
12	100,00	0,00	Excess Inventory	0	Reseller
13	100,00	0,00	New Product	0	Reseller
14	100,00	0,00	New Product	0	Reseller
15	100,00	0,00	Seasonal Discount	0	Customer
16	100,00	0,00	Discontinued Product	0	Reseller

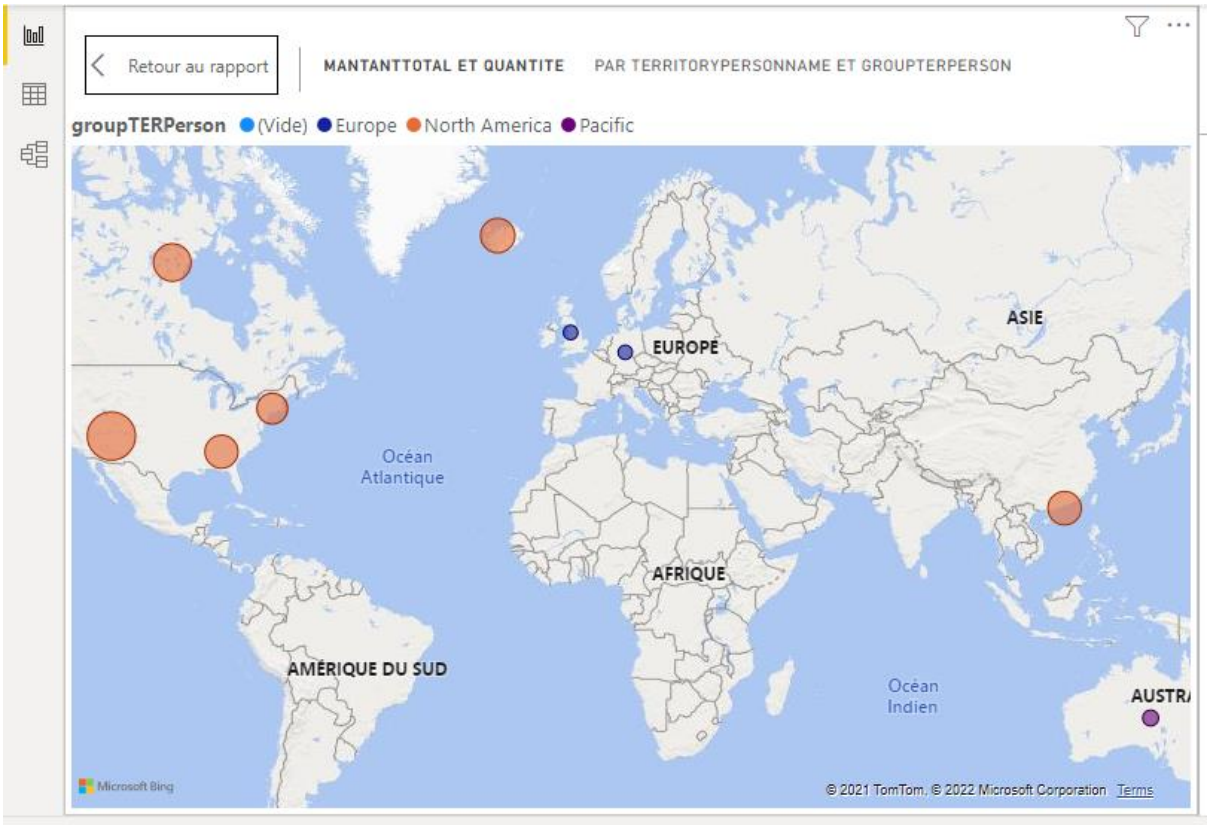
- quantité par type d’offre :



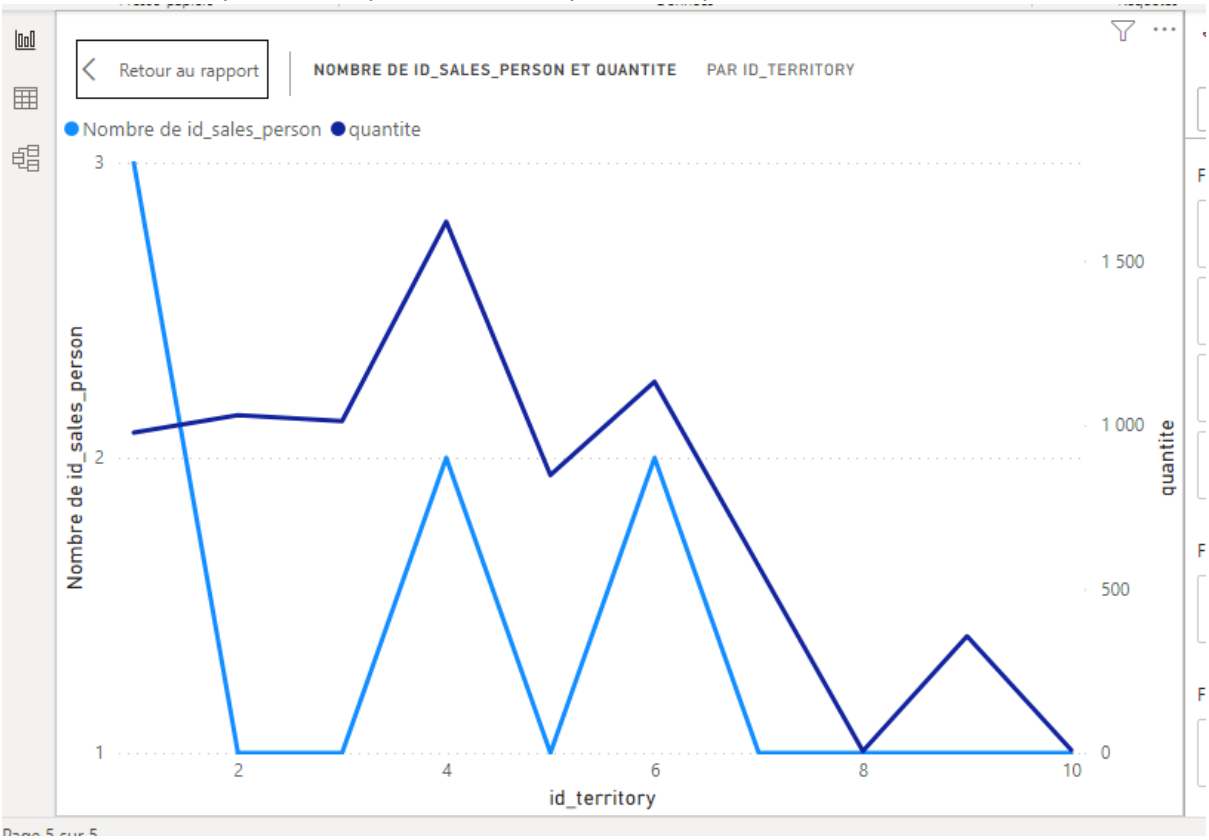
d) selon sales person :



- montant et quantité par territory person name et groupterperson :



- le nombre de sale peson et la quantité dan chaque territory :



- la somme de quantité vendu et le montant total par chaque sale person :

Retour au rapport

quantite	mantantTotal	salesPersonID	bonus	groupTERPerson	territoryPersonName	sales_last_year
1 030,00	523 909,69	275	4 100,00	North America	Northeast	1 750 410,00
1 012,00	605 733,87	277	2 500,00	North America	Central	1 997 190,00
897,00	562 318,95	281	3 550,00	North America	Southwest	2 073 510,00
847,00	593 312,34	279	6 700,00	North America	Southeast	1 849 640,00
795,00	514 825,37	282	5 000,00	North America	Canada	2 038 230,00
723,00	512 977,25	276	2 000,00	North America	Southwest	1 439 160,00
708,00	389 576,01	280	5 000,00	North America	Northwest	1 927 060,00
356,00	53 548,42	286	5 650,00	Pacific	Australia	2 278 550,00
337,00	221 492,04	278	500,00	North America	Canada	1 620 280,00
307,00	61 457,52	285				
269,00	247 287,62	283	3 500,00	North America	Northwest	1 371 640,00
47,00	58 138,36	274				
9,00	7 345,72	289	5 150,00	Europe	United Kingdom	1 635 820,00
6,00	3 563,02	288	75,00	Europe	Germany	1 307 950,00
7 343,00	**4 355 486,19**					

Conclusion :

Apres ces visualisations et les informations présenté par ce rapport on peut deduire que le developpement des ventes liées fortement au bon choix de lieu de vente et le nombre de sales person dans les lieux ou existes un grand nombre de clients , et ou les quantités vendus et les montants total plus élevés ,mais tous ces facteurs ne sont pas toujours homogènes , il faut parfois privilégier certains facteurs qui donne plus de rentabilité .