

Yiyang WANG
Vincent MANUEL
M2 MIAGE IPM
22/02/2019



Projet Données, Intégration, Analyse

DOSSIER DE CONCEPTION



Table des matières

1. Contexte.....	2
1.1. Présentation synthétique du projet :	2
1.2. Expression des besoins	2
2. Architecture fonctionnelle du SIAD	4
2.1. Magasins de données :	4
2.3. Architecture globale du SAD	4
3. Magasins	5
3.1 MD_DRH	5
3.1.1 Données nécessaires :	5
3.1.2 Modélisation conceptuelle	5
3.1.3 Modélisation logique	7
3.2 MD_RGP	12
3.2.1 Données nécessaires :	12
3.2.2 Modélisation conceptuelle	12
3.2.3 Modélisation physique	14
3.3 MD_Marketing.....	16
3.3.1 Données nécessaires :	16
3.3.2 Modélisation conceptuelle	16
3.3.3 Modélisation logique	19
4. Entrepôt de données	22
4.1. Modélisation des données :	22
5. Conclusion.....	23

1. Contexte

1.1. Présentation synthétique du projet :

L'objectif de notre projet est de déployer un système d'aide à la décision pour analyser les différentes activités du pôle presse du groupe Lagardère afin de répondre aux attentes de ses différents décideurs. Concrètement, nous devons répondre aux besoins de trois décideurs ou groupe de décideurs :

- La Direction des Ressources Humaines
- Les responsables des groupes de presse
- Le service Marketing

Le groupe Lagardère dispose déjà d'un entrepôt de données, basé sur BD Oracle.

Le planning prévisionnel de notre projet vous a été remis en début de projet, nous vous avons joint le planning final en annexe.

1.2. Expression des besoins

Référence du besoin	Classe de décideur	Description	Restitutions souhaitées	Type de restitution
B1	Direction des Ressources Humaines	Analyses hebdomadaires, mensuelles et annuelles sur l'activité des journalistes dans les différentes publications	B1.1 : Nombre annuel d'articles écrits par type de journalistes (classés par ordre décroissant du nombre d'articles)	Etat Requête (via oracle)
			B1.2 : Montants totaux versés mensuellement pour payer les journalistes en fonction des publications (caractérisées par leur nom et le groupe de presse l'éditant)	Etat Requête (via oracle)
			B1.3 : Montants totaux payés annuellement pour chaque publication (nom et périodicité)	Etat Requête (via oracle)
B2	Responsables des groupes de presse	Analyser mensuellement et annuellement les nombres de numéros de publications distribués aux clients en fonction de caractéristiques sur les publications	B2.1 : Nombre de livraisons par type de publications pour chaque année	Rapport BO (via Web I)
			B2.2 : Evolution des ventes de publications en 2017 et 2018 par périodicité	Rapport BO (via Web I)
			B2.3 : Analyse par ville de centre distributeur du nombre de publications vendues par mois et année (par groupe de presse)	Rapport BO (via Web I)
B3	Service Marketing	Analyses sur les abonnements des clients à des publications en fonction du temps	B3.1 : Montant des bénéfices par ville de centre de distributeur	Graphique (via Tableau)
			B3.2 : Meilleurs clients par année et par ville	Graphique (via Tableau)
			B3.3 : Analyse des bénéfices et des montants par Groupe de presse et types de publications	Graphique (via Tableau)

Les responsables des groupes de presse souhaitent effectuer des analyses sur les nombres de numéros de publications distribués aux clients. Le nombre de livraisons par type de publications pour chaque année est intéressant car il donne une vue globale et une vue par groupe de presse ainsi que sur la situation annuelle de livraison, les responsables peuvent donc savoir le type de publication le plus populaire de manière globale. Le rapport de l'évolution des ventes de publications en 2017 et 2018 par périodicité permet de savoir quelle périodicité de publication est de plus en plus populaire pour les clients et donne une orientation prévisionnelle des ventes futures.

Le service souhaite faire des analyses sur les abonnements des clients à ses publications en fonction du temps. Donc on propose trois rapports préétablis : Le montant des bénéfices par ville de centre de distributeur permet de donner une vue géographique sur l'analyse des abonnements et il va donc aider le service à suivre des politiques sur différentes zones géographiques. Les graphiques des meilleurs clients par année et par ville concernent plutôt l'analyse des clients de différentes villes. Il lui offre un classement de montants payés par client et un classement de bénéfices par client. De plus, on va lui proposer des filtres, par exemple le filtre de temps, permettant lui de trouver les meilleurs clients. Ainsi le service Marketing peut repenser sa politique de fidélisation de client. Le rapport d'analyse des bénéfices et des montants par Groupe de presse et types de publications permet de voir la capacité de rentabilité de groupes de presses et le poids sur le marché des différents types de publication et il donne donc une vue prévisionnelle du marché.

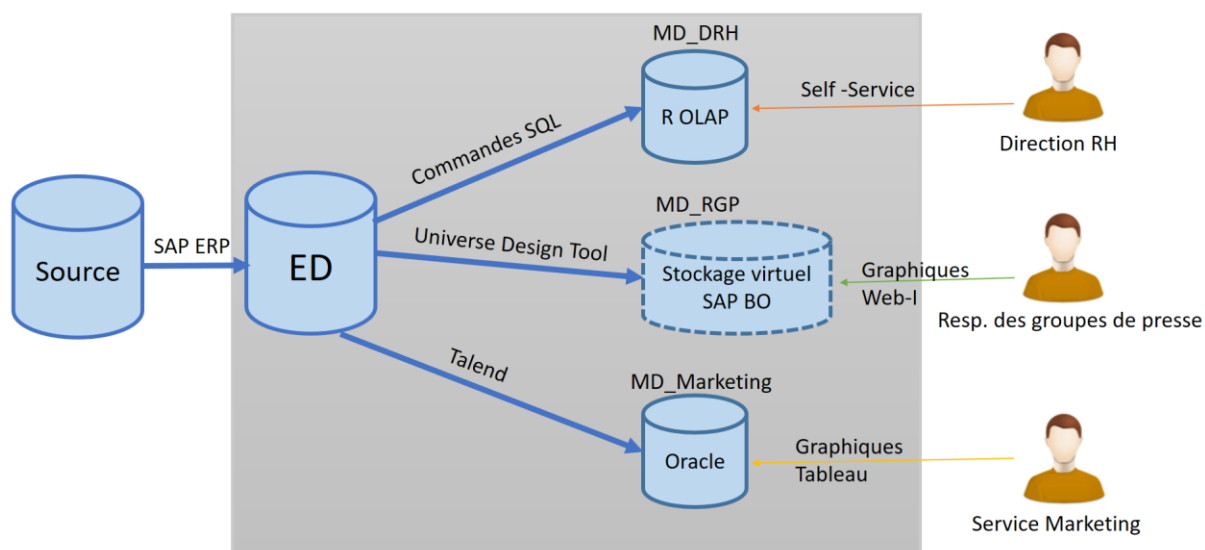
2. Architecture fonctionnelle du SIAD

2.1. Magasins de données :

Identifiant	Description	Utilisateurs	Outils de stockage	Outils de restitution
MD_DRH	Permet d'analyser en « self-service » les activités des journalistes et les montants versés aux journalistes dans les différentes publications en fonction du temps	Direction des Ressources Humaines	Oracle (R-OLAP)	Oracle 10
MD_RGP	Permet d'analyser les nombres de numéros de publications distribués aux clients en fonction de caractéristiques sur les publications et en fonction du temps	Responsables des groupes de presse	SAP BO (Stockage virtuel)	Web Intelligence Rich Client
MD_Marketing	Permet d'analyses sur les abonnements des clients à des publications en fonction du temps	Service Marketing	Talend	Tableau 2018.1

2.3. Architecture globale du SAD

Schéma global :



3. Magasins

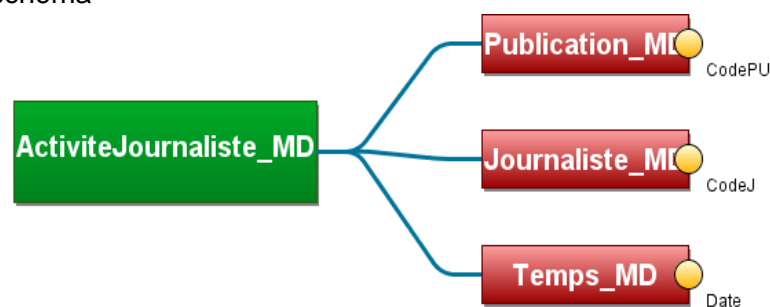
3.1 MD_DRH

3.1.1 Données nécessaires :

Besoin couverts	Données nécessaires	Où ?			Format Restitution
		Source (ED)	Calculée	Saisie	
B1.1 : Nombre annuel d'articles écrits par type de journalistes (classés par ordre décroissant du nombre d'articles)	Année		to_char(Numero.DateNO,'yyyy')		Analyse OLAP (oracle)
	Nombre d'Article		SUM(EcrireArticle.NBArt)		
	Type de Journaliste	ED(TypeJour.NomTYJ)			
B1.2 : Montants totaux versés mensuellement pour payer les journalistes en fonction des publications (caractérisées par leur nom et le groupe de presse l'éditant)	Montant versé aux journalistes		SUM(TypeJour.PrixCar*EcrireArticle.Nbcar+ CoutArticle.ForfaitArt * EcrireArticle.NbArt)		Analyse OLAP (oracle)
	Nom du groupe de presse	ED(Publication.GroupePresse)			
	Nom de la publication	ED(Publication.No mPU)			
	Mois		to_char(Numero.DateNO,'mm/yyyy')		
B1.3 : Montants totaux payés annuellement pour chaque publication (nom et périodicité)	Montant versé aux journalistes		SUM(TypeJour.PrixCar*EcrireArticle.Nbcar+ CoutArticle.ForfaitArt * EcrireArticle.NbArt)		Analyse OLAP (oracle)
	Nom du publication	ED(Publication.No mPU)			
	Périodicité de la publication Année	ED(Publication.périodicite)	to_char(Numero.DateNO,'yyyy')		

3.1.2 Modélisation conceptuelle

- Structure du schéma



- Dictionnaire de mesure

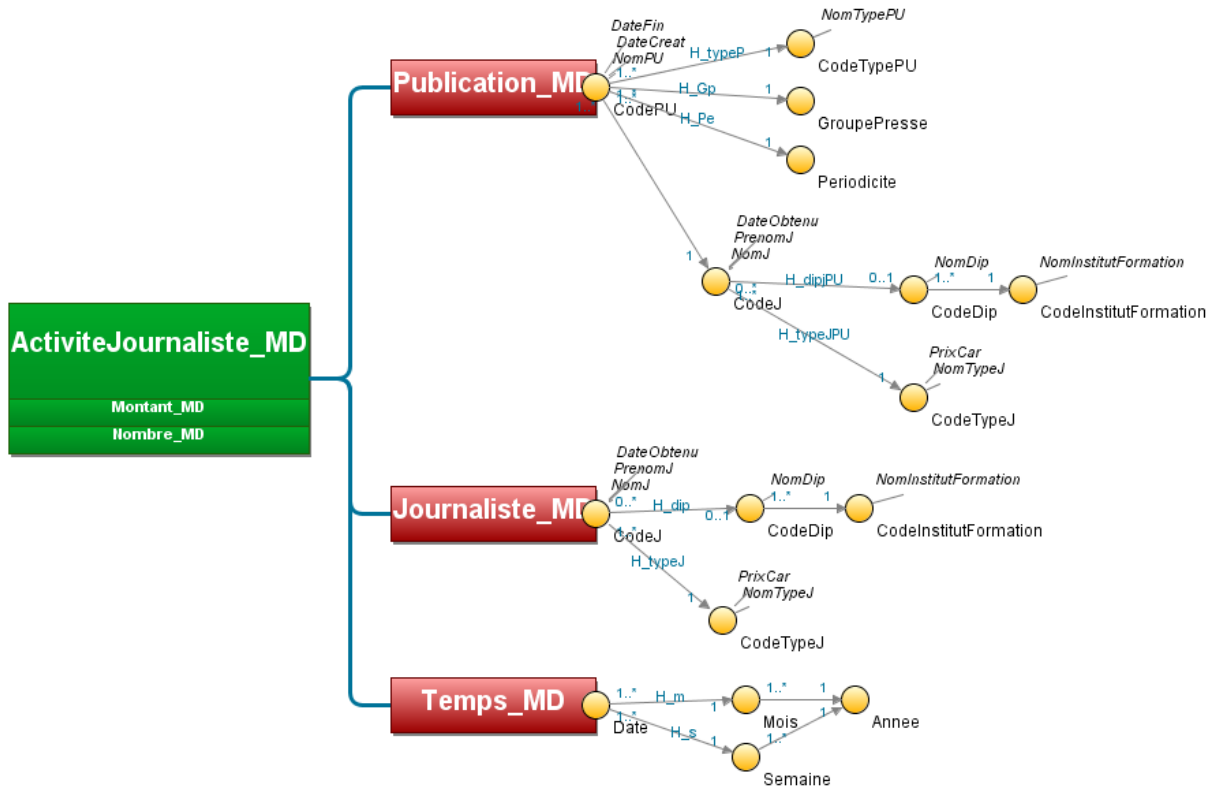
Code	Désignation	Type	Formule d'extraction
Montant_MD	La somme de montant versé à un journaliste donné pour une publication donnée dans un mois donné	Réel	=SUM(TypeJour.PrixCar*EcrireArticle.Nbcar + CoutArticle.ForfaitArt * EcrireArticle.NbArt)

Nombre_MD	La somme de nombre d'articles écrits par un journaliste donné pour une publication donné dans un mois donné	Entier	=SUM(EcrireArticle.NBArt)
------------------	---	--------	---------------------------

● Dictionnaire des attributs

Code	Désignation	Type	Formule d'extraction
Annee	Année d'un numéro de publication donnée	CC(4)	=to_char(Numero.DateNO,'yyyy')
CodeDip	Code unique d'un diplôme de journalisme donné	Entier	=Journaliste.CodeDI
CodeInstitutFormation	Code unique d'un institut de formation donné	Entier	=InstitutFormation.CodeIF
CodeJ	Code unique d'un journaliste donné	Entier	=Journaliste.CodeJ
CodePU	Code unique d'une publication donnée	Entier	=Publication.CodePU
CodeTypeJ	Code unique d'un type de journaliste donné	Entier	=Journaliste.CodeTY
CodeTypePU	Type d'une publication donnée	Entier	=Publication.CodeTY
Date	Date d'un numéro de publication donnée	Date	=Numero.DateNO
DateCreat	Date de création d'une publication donnée	Date	=Publication.DateCreatPU
DateFin	Date éventuelle de fin d'une publication donnée	Date	=Publication.DateFinPU
DateObtenu	Date d'obtention d'un diplôme de journalisme par un journaliste donné	Date	=Journaliste.Dateobt
GroupePresse	Groupe Presse d'une publication donnée	CC(100)	=Publication.GroupePresse
Mois	Mois d'une année d'un numéro de publication donnée	CC(7)	=to_char(Numero.DateNO,'mm/yyyy')
NomDip	Nom d'un diplôme de journalisme donné		=Diplome.NomDI
NomInstitutFormation	Nom d'un institut de formation de journalistes donné	CC(100)	=InstitutFormation.NomIF
NomJ	Nom d'un Journaliste donné	CC(50)	=Journaliste.NomJ
NomPU	Nom d'une publication donnée	CC(75)	=Publication.NomPU
NomTypeJ	Nom d'un type de journaliste donné	CC(75)	=TypeJour.NomTYJ
NomTypePU	Nom d'un type de publication donnée	CC(75)	=TypePU.NomTY
Periodicite	Périodicité d'une publication donnée	CC(30)	=Publication.Periodicite
PrenomJ	Prénom d'un Journaliste donné	CC(50)	=Journaliste.PrenomJ
PrixCar	Prix du caractère pour un type de journaliste donné (en €)	Réel	=TypeJour.PrixCar
Semaine	Numéro de semaine d'une année d'un numéro de publication donnée	CC(7)	=to_char(Numero.DateNO,'ww/yyyy')

- Schéma complète



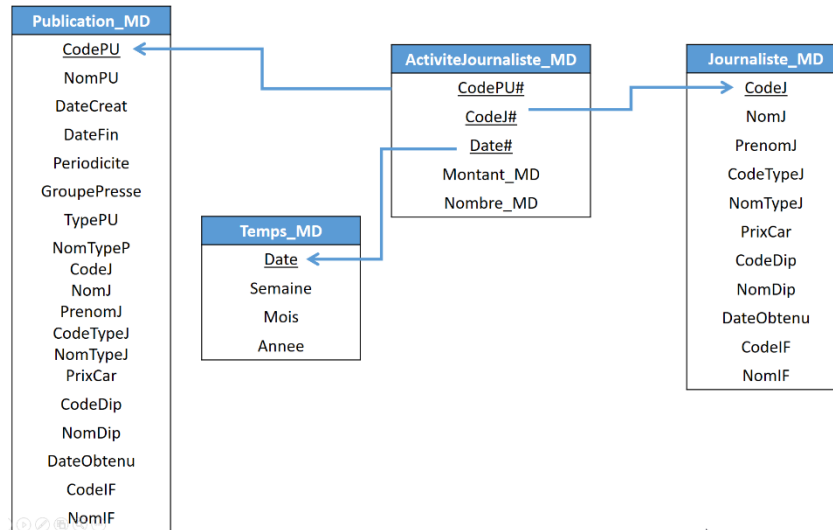
Ce schéma complet contient un fait (les activités des journalistes) et trois dimensions (Publication, Journaliste, Temps) permettant d'analyser les montants versés aux journalistes pour l'écriture d'articles et le nombre d'articles écrits par les journalistes. En raison de l'unicité de type, de groupe de presse et de périodicité pour une publication, ils sont considérés comme une hiérarchie dans la dimension de publication. La hiérarchie Journaliste est traduite par la relation CIF sur le schéma d'entrepôt de données. Dans la dimension Journaliste, les deux hiérarchies sont basées sur la relation CIF sur le schéma d'entrepôt de données et particulièrement la hiérarchie *H_dip* est une hiérarchie incomplète avec la cardinalité (0..*, 0..1). Le numéro de semaine dans un an peut être partagée par deux mois et on donc les distingue comme deux hiérarchies dans le domaine de temps.

3.1.3 Modélisation logique

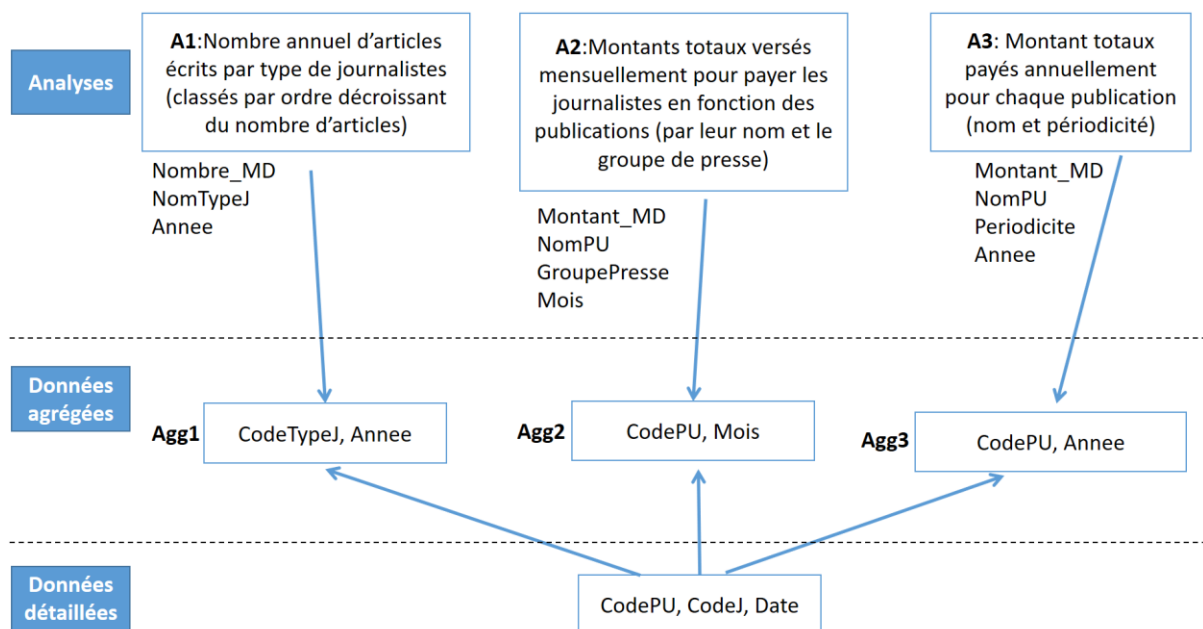
On choisit la modélisation logique R-OLAP dénormalisé en raison de l'adaptation. La modélisation logique R-OLAP dénormalisé possède peu de tables et puis les jointures limitent les analyses décisionnelles. Cela rend ce magasin des données plus propre, s'adaptant bien aux besoins d'analyse de la Direction des Ressources Humaines. Par ailleurs, une BD R-OLAP dénormalisé peut être implémenté sur Oracle et cela répond au besoin du directeur des ressources humaines.

- Statique :

Le schéma R-OLAP dénormalisé graphique pour données détaillées représente ci-dessous.



Le treillis partiel avec requêtes de construction des agrégats pour les données agrégées est représenté ci-dessous.



-- Création du Agg1

```
CREATE MATERIALIZED VIEW Agg1
BUILD IMMEDIATE REFRESH COMPLETE ON DEMAND
AS SELECT  t.Annee,j.NomTypeJ as Type_Journaliste, SUM(a.Nombre_MD) AS NbArticle
FROM Journaliste_MD j, Temps_MD t, ActiviteJournaliste_MD a
WHERE j.CodeJ = a.CodeJ AND t.Date_MD = a.DATEACTIVITE
GROUP BY j.NomTypeJ, t.Annee
ORDER BY SUM(a.Nombre_MD) DESC;
```

-- Création du Agg2

```

CREATE MATERIALIZED VIEW Agg2
BUILD IMMEDIATE REFRESH COMPLETE ON DEMAND
AS SELECT p.NomPU as Nom_Publication, p.GroupePresse as Groupe_Presse,t.Mois,
        SUM(a.Montant_MD) AS Montant_Mensuelle
FROM Publication_MD p,Temps_MD t,ActiviteJournaliste_MD a
WHERE p.CodePU = a.CodePU AND t.Date_MD = a.DATEACTIVITE
GROUP BY p.NomPU, p.GroupePresse,t.Mois;

-----
-- Création du Agg3
-----

CREATE MATERIALIZED VIEW Agg3
BUILD IMMEDIATE REFRESH COMPLETE ON DEMAND
AS SELECT p.NomPU as Nom_Publication, p.Periodicite,t.Annee, SUM(a.Montant_MD) AS Montant_Annuelle
FROM Publication_MD p,Temps_MD t,ActiviteJournaliste_MD a
WHERE p.CodePU = a.CodePU AND t.Date_MD = a.DATEACTIVITE
GROUP BY p.NomPU, p.Periodicite,t.Annee;

```

● Dynamique :

Cette partie représente les commandes SQL pour données détaillées.

```

-----
-- Créer la dimension temporelle
-----

CREATE MATERIALIZED VIEW TEMPS_MD
BUILD IMMEDIATE
REFRESH COMPLETE ON DEMAND
AS SELECT distinct DateNO AS DATE_MD,
        TO_CHAR(DateNO,'ww/yyyy') AS Semaine,
        TO_CHAR(DateNO,'mm/yyyy') AS Mois,
        TO_CHAR(DateNO,'YYYY') AS Annee
FROM PJDIA_YV.Numero;

ALTER TABLE TEMPS_MD ADD CONSTRAINT pk_TEMPS_MD PRIMARY KEY(DATE_MD);

CREATE DIMENSION TEMPS_DIM
    LEVEL N_Date IS (TEMPS_MD.DATE_MD)
    LEVEL N_Semaine IS (TEMPS_MD.SEMAINE)
    LEVEL N_Mois IS (TEMPS_MD.MOIS)
    LEVEL N_Annee IS (TEMPS_MD.ANNEE)
HIERARCHY H_s (N_Date
    CHILD OF N_Semaine
    CHILD OF N_Annee)
HIERARCHY H_m (N_Date
    CHILD OF N_Mois
    CHILD OF N_Annee) ;

-----
-- Créer la dimension de Journaliste
-----

CREATE MATERIALIZED VIEW Journaliste_MD
BUILD IMMEDIATE
REFRESH COMPLETE ON DEMAND
AS SELECT j.CodeJ,j.NomJ,j.PrenomJ,j.Dateobt AS DateObtenu, t.CodeTYJ AS CodeTypeJ,t.nomtyj as NomTypeJ,t.PrixCar,
        d.CodeDI AS CodeDip,d.nomdi as NomDip, i.codeif as CodeInstitutFormation,i.nomif as NomInstitutFormation
FROM PJDIA_YV.Journaliste j, PJDIA_YV.TypeJour t,PJDIA_YV.diplome d,PJDIA_YV.INSTITUTFORMATION i
WHERE j.codetyj = t.codetyj and d.codedi = j.codedi and d.codeif = i.codeif;

ALTER TABLE Journaliste_MD ADD CONSTRAINT pk_Journaliste_MD PRIMARY KEY(CodeJ);

CREATE DIMENSION Journaliste_DIM
    LEVEL N_Codej IS (Journaliste_MD.CodeJ)

```

```

LEVEL N_Typej IS (Journaliste_MD.CodeTypeJ)
LEVEL N_codedip IS (Journaliste_MD.CodeDip)
LEVEL N_codeif IS (Journaliste_MD.CodeInstitutFormation)

```

```

HIERARCHY H_typej (N_Codej CHILD OF N_Typej)
HIERARCHY H_dip (N_Codej CHILD OF N_codedip CHILD OF N_codeif )

```

```

ATTRIBUTE N_Codej DETERMINES Journaliste_MD.NomJ
ATTRIBUTE N_Codej DETERMINES Journaliste_MD.PrenomJ
ATTRIBUTE N_Codej DETERMINES Journaliste_MD.DateObtenu
ATTRIBUTE N_Typej DETERMINES Journaliste_MD.NomTypeJ
ATTRIBUTE N_Typej DETERMINES Journaliste_MD.PrixCar
ATTRIBUTE N_codedip DETERMINES Journaliste_MD.NomDip
ATTRIBUTE N_codeif DETERMINES Journaliste_MD.NomInstitutFormation ;

```

 -- Créer la dimension de Publication

```

CREATE MATERIALIZED VIEW Publication_MD
BUILD IMMEDIATE
REFRESH COMPLETE ON DEMAND

```

```

AS SELECT p.CodePU,p.NomPU,p.DateCreatPU AS DateCreat,p.DateFinPU AS DateFin,
           p.Periodicite,p.GroupePresse,p.CodeTY AS CodeTypePU,t.NomTY AS NomTypePU,
           j.CodeJ,j.NomJ,j.PrenomJ,j.Dateobt AS DateObtenu, t.CodeTYJ AS CodeTypeJ,t.nomtyj as NomTypeJ,t.PrixCar,
           d.CodeDI AS CodeDip,d.nomdi as NomDip, i.codeif as CodeInstitutFormation,i.nomif as NomInstitutFormation
FROM PJDIA_YV.Publication p,PJDIA_YV.TypePU t,PJDIA_YV.Journaliste j,
       PJDIA_YV.TypeJour t,PJDIA_YV.diplome d,PJDIA_YV.INSTITUTFORMATION i
WHERE p.CodeTY = t.CodeTY and p.codej = j.codej
       and j.codetyj = t.codetyj and d.codedi = j.codedi and d.codeif = i.codeif;

```

```

ALTER TABLE Publication_MD ADD CONSTRAINT pk_Publication_MD PRIMARY KEY(CodePU);

```

```

CREATE DIMENSION Publication_DIM
LEVEL N_CodePU IS (PUBLICATION_MD.CodePU)
LEVEL N_TypePU IS (PUBLICATION_MD.CodeTypePU)
LEVEL N_Periodicite IS (PUBLICATION_MD.Periodicite)
LEVEL N_GroupePresse IS (Publication_MD.GroupePresse)
LEVEL N_j IS (Publication_MD.CodeJ)
LEVEL N_dipj IS (Publication_MD.CodeDip)
LEVEL N_ifj IS (Publication_MD.CodeInstitutFormation)
LEVEL N_typejr IS (Publication_MD.CodeTypeJ)

```

```

HIERARCHY H_typeP (N_CodePU CHILD OF N_TypePU)
HIERARCHY H_Gp(N_CodePU CHILD OF N_GroupePresse)
HIERARCHY H_Pe(N_CodePU CHILD OF N_Periodicite)
HIERARCHY H_dipjPU(N_CodePU CHILD OF N_j
                  CHILD OF N_dipj
                  CHILD OF N_ifj)
HIERARCHY H_typeJPU(N_CodePU CHILD OF N_j CHILD OF N_typejr)

```

```

ATTRIBUTE N_CodePU DETERMINES Publication_MD.NomPU
ATTRIBUTE N_CodePU DETERMINES Publication_MD.DateCreat
ATTRIBUTE N_CodePU DETERMINES Publication_MD.DateFin
ATTRIBUTE N_TypePU DETERMINES Publication_MD.NomTypePU
ATTRIBUTE N_j DETERMINES Publication_MD.NomJ
ATTRIBUTE N_j DETERMINES Publication_MD.PrenomJ
ATTRIBUTE N_j DETERMINES Publication_MD.DateObtenu
ATTRIBUTE N_dipj DETERMINES Publication_MD.NomDip
ATTRIBUTE N_ifj DETERMINES Publication_MD.NomInstitutFormation
ATTRIBUTE N_typejr DETERMINES Publication_MD.NomTypeJ
ATTRIBUTE N_typejr DETERMINES Publication_MD.PrixCar ;

```

-- Création du fait

```
CREATE MATERIALIZED VIEW ActiviteJournaliste_MD
BUILD IMMEDIATE
REFRESH COMPLETE ON DEMAND
AS SELECT p.CodePU, j.CodeJ, n.DateNO AS DateActivite,
          SUM(t.PrixCar*e.Nbcar+c.ForfaitArt * e.NbArt) AS Montant_MD,
          SUM(e.NbArt) AS Nombre_MD
FROM PJDIA_YV.Publication p, PJDIA_YV.Journaliste j, PJDIA_YV.Numero n,
     PJDIA_YV.EcrireArticle e, PJDIA_YV.CoutArticle c, PJDIA_YV.typejour t, PJDIA_YV.rubrique r
WHERE r.codepu = p.codepu
    and j.codetyj = j.codetyj
    and e.codej = j.codej
    and e.codeno = n.codeno
    and e.coderu = r.coderu
    and c.coderu = j.codej
GROUP BY p.CodePU, j.CodeJ, n.DateNO;

ALTER TABLE ActiviteJournaliste_MD ADD CONSTRAINT pk_ActiviteJournaliste_MD
PRIMARY KEY(CodePU, CodeJ, DateActivite);

ALTER TABLE ActiviteJournaliste_MD ADD CONSTRAINT fk_ActiviteJ_publications
FOREIGN KEY(CodePU) REFERENCES Publication_MD(CodePU);
ALTER TABLE ActiviteJournaliste_MD ADD CONSTRAINT fk_ActiviteJ_journaliste
FOREIGN KEY(CodeJ) REFERENCES Journaliste_MD(CodeJ);
ALTER TABLE ActiviteJournaliste_MD ADD CONSTRAINT fk_ActiviteJ_temps
FOREIGN KEY(DateActivite) REFERENCES TEMPS_MD(DATE_MD);
```

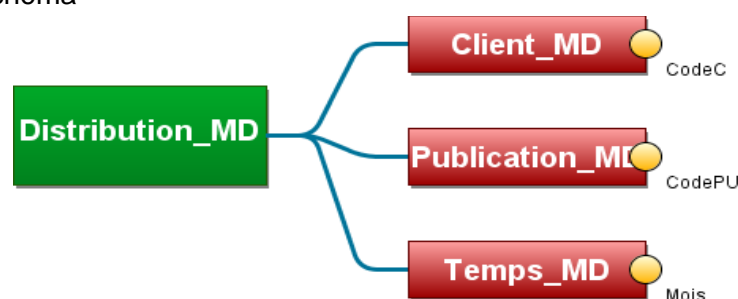
3.2 MD_RGP

3.2.1 Données nécessaires :

Besoin couverts	Données nécessaires	Où ?			Format Restitution
		Source (ED)	Calculée	Saisie	
B2.1 : : Nombre de livraisons par type de Publications pour chaque année	Année		to_char(Numero.DateNO,'yyyy')		Tableau, Graphique (Web-I)
	Mois		to_char(Numero.DateNO,'mm/yyyy')		
	Nom de type de publication	ED(TypePU.NomTY)			
	Nombre de numéro distribué		SUM(Count(Livrer.CodeNO))		
B2.2 : Evolution des ventes de publications en 2017 et 2018 par périodicité	Année		to_char(Numero.DateNO,'yyyy')		Tableau croisé, Graphique (Web-I)
	Mois		to_char(Numero.DateNO,'mm/yyyy')		
	Nombre de numéro distribué		SUM(Count(Livrer.CodeNO))		
	Nom de publication	ED(Publication.NomPU)			
	Périodicité de publication	ED(Publication.périodicité)			
B2.3 : Analyse par ville de centre distributeur du nombre de publications vendues par mois et année (par groupe de presse)	Année		to_char(Numero.DateNO,'yyyy')		Tableau croisé, Graphique (Web-I)
	Nombre de numéro distribué		SUM(Count(Livrer.CodeNO))		
	Groupe de presse	ED(Publication.GroupePresse)			
	Ville de centre de distributeur	ED(CentreDistributeur.VilleCD)			
	Nom de type de publication	ED(TypePU.NomTY)			

3.2.2 Modélisation conceptuelle

- Structure du schéma



- Dictionnaire de mesure

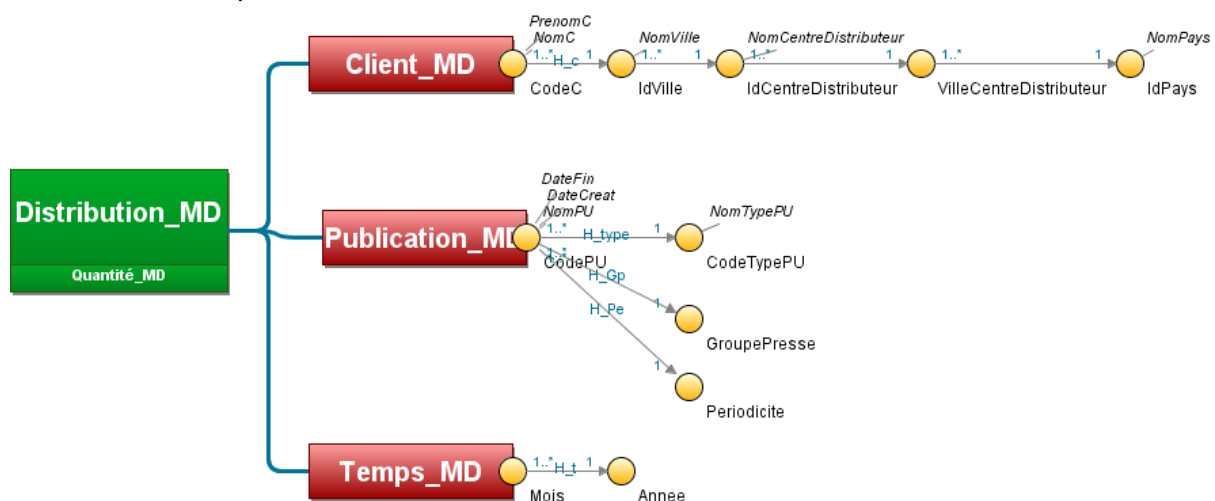
Code	Désignation	Type	Formule d'extraction
Quantité_MD	Le nombre de numéros d'une publication distribuées à un client dans un mois donné.	Entier	=Count(Livrer.CodeNO)

- Dictionnaire des attributs

Code	Désignation	Type	Formule d'extraction
------	-------------	------	----------------------




Annee	Année d'un numéro de publication donnée	CC(4)	=to_char(Numero.DateNO,'yyyy')
CodeC	Code unique d'un client donné	Entier	=Client.CodeC
CodePU	Code unique d'une publication donnée	Entier	=Publication.CodePU
CodeTypePU	Type d'une publication donnée	Entier	=Publication.CodeTY
DateCreat	Date de création d'une publication donnée	Date	=Publication.DateCreatPU
DateFin	Date éventuelle de fin d'une publication donnée	Date	=Publication.DateFinPU
GroupePresse	Groupe Presse d'une publication donnée	CC(100)	=Publication.GroupePresse
IdCentreDistributeur	Code unique d'un centre distributeur de courriers	Entier	=CentreDistributeur.CodeCD
IdPays	Code unique d'un pays fourni par l'INSEE	Entier	=Pays.CodeP
IdVille	Code unique d'une ville d'un client fourni par l'INSEE	Entier	=Ville.CodeV
Mois	Mois d'une année d'un numéro de publication donnée	CC(7)	=to_char(Numero.DateNO,'mm/yyyy')
NomC	Nom d'un client donné	CC(50)	=Client.NomC
NomCentreDistributeur	Nom d'un centre distributeur donné	CC(50)	=CentreDistributeur.NomCD
NomPays	Nom d'un Pays donné	CC(75)	=Pays.NomP
NomPU	Nom d'une publication donnée	CC(75)	=Publication.NomPU
NomTypePU	Nom d'un type de publication donnée	CC(75)	=TypePU.NomTY
NomVille	Nom d'une ville d'un client donné	CC(50)	=Ville.NomV
Periodicite	Périodicité d'une publication donnée	CC(30)	=Publication.Periodicite
PrenomC	Prénom d'un client donné	CC(50)	=Client.PrenomC
VilleCentreDistributeur	Ville d'un centre distributeur donné	CC(50)	=CentreDistributeur.VilleCD

● Schéma complète

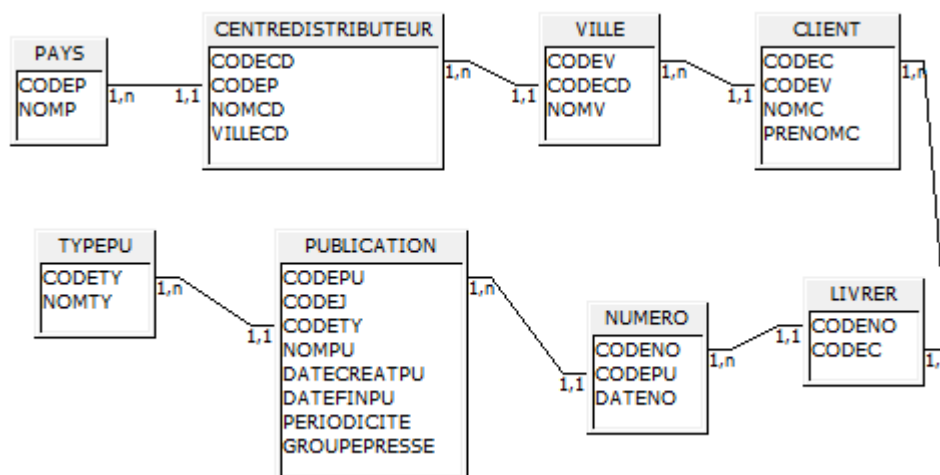


Ce magasin de donnée comprend un fait (Distribution des publications) et trois dimensions (Client, Publication et Temps). Selon le schéma E/A d'entrepôt de donnée, la dimension Client contient une seule hiérarchie car étant basé sur la relation CIF. La ville de résidence d'un client dépend d'un centre de distribution du courrier mais ces deux villes ne sont pas nécessairement dans la même ville et on donc considère que la ville de centre de distributeur peut servir plusieurs villes de résidence de clients. Pour bien voir cette différence, on met la ville de centre de distributeur comme paramètre plutôt que l'attribut faible. Les responsables des groupes de presse souhaitent analyser la livraison en fonction des caractéristiques sur les publications et des caractéristiques proposées qui sont le type, la périodicité, le groupe de presse ainsi que le nom et les dates de début et de fin de la publication puisque l'on ne met plus la hiérarchie Journaliste ici. De même, ils ont besoin d'analyser mensuellement et annuellement il faut donc ajouter mois et année dans la dimension temps.

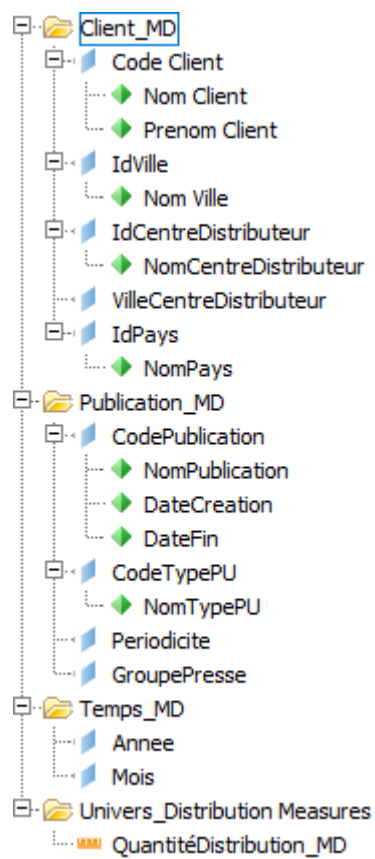
3.2.3 Modélisation physique

Conceptuel	SAP-BO
Schéma étoile	Univers
Fait	Class d'indicateur
Mesure	Indicateur 
Dimension	Class d'objet
Hiérarchie	Hiérarchie
Paramètre	Dimension 
Attribut faible	Information 

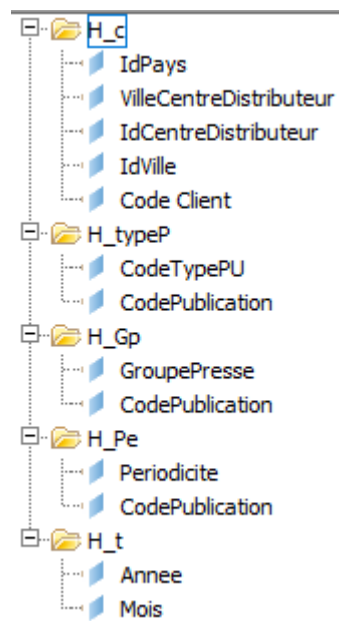
Le schéma BO déduit est ci-dessous.



Tous les paramètres et les attributs faibles



La hiérarchie :



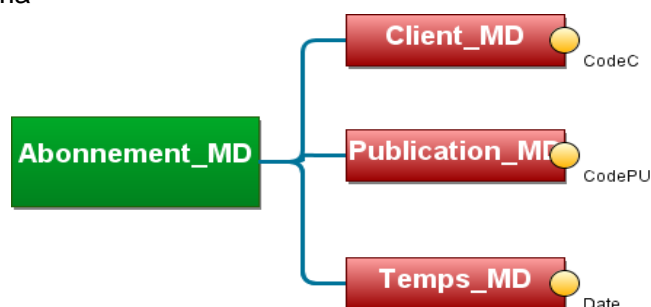
3.3 MD_Marketing

3.3.1 Données nécessaires :

Besoin couverts	Données nécessaires	Où ?			Format Restitution
		Source (ED)	Calculée	Saisie	
B3.1 : Montant des bénéfices par ville de centre de distributeur	Année		to_char(Numero.DateNO,'yy yy')		Graphique (Tableau)
	Montant des bénéfices		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		
	Ville de centre de distributeur	ED(CentreDistributeur.VilleCD)			
	Bénéfice réalisé		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		
B3.2 : Meilleurs clients par année et par ville	Nom de client	ED(Client.NomC)			Graphique (Tableau)
	Ville de centre de distributeur	ED(CentreDistributeur.VilleCD)			
	Année		to_char(Numero.DateNO,'yy yy')		
	Nombre de numéro abonné		SUM(Abonner.Nbnum)		
	Montant payé		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		
	Bénéfice réalisé		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		
B3.3 : Analyse des bénéfices et des montants par Groupe de presse et types de publications	Type de publication	ED(TypePU.NomTY)			Graphique (Tableau)
	Groupe de presse	ED(Publication.GroupePresse)			
	Périodicité	ED(Publication.Periodicite)			
	Montant payé		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		
	Bénéfice réalisé		SUM(Abonner.Nbnum*(Payer.PrixNumPublic - Abonner.PrixNumAb))		

3.3.2 Modélisation conceptuelle

- Structure du schéma



- Dictionnaire de mesure

Code	Désignation	Type	Formule d'extraction
------	-------------	------	----------------------

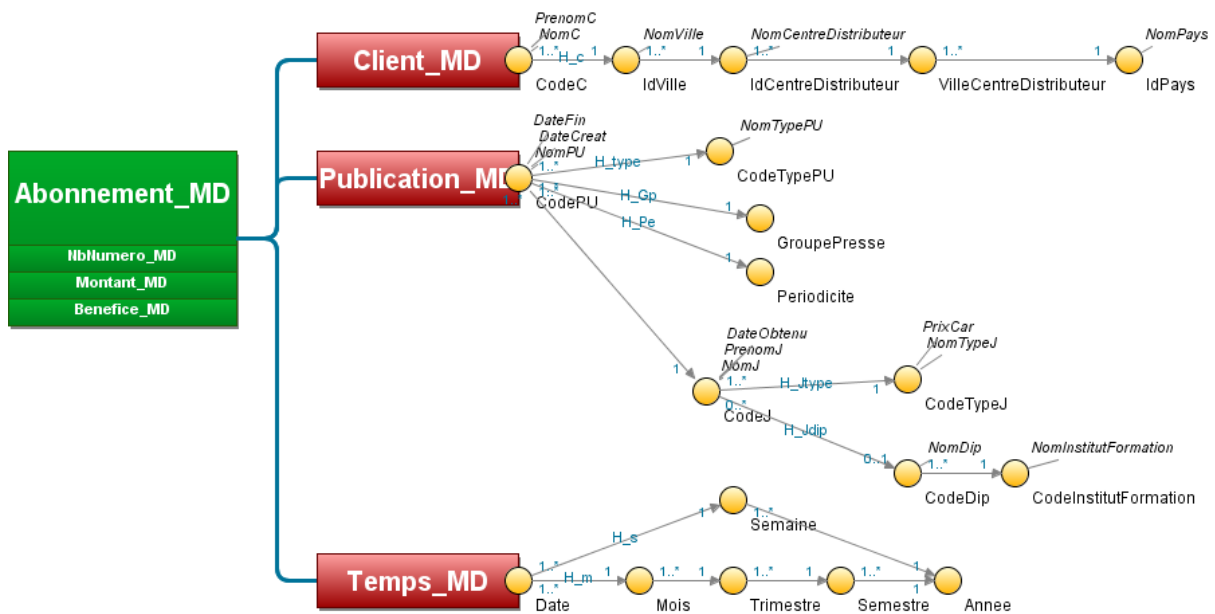
NbNumero_MD	Le nombre de numéros de publications abonnés par un client donné dans une date donnée.	Entier	=Abonner.Nbnum
Montant_MD	Le montant versé d'un publication abonnés par un client dans une date donnée.	Réel	=Abonner.PrixNumAb * Abonner.Nbnum
Benefice_MD	Le bénéfice réalisé d'un publication abonnés par un client dans une date donnée.	Réel	=Abonner.Nbnum*(Payer.PrixNumPublic -Abonner.PrixNumAb)

● Dictionnaire des attributs

Code	Désignation	Type	Formule d'extraction
Annee	Année de l'abonnement pour un client donné	CC(4)	=to_char(Abonner.DateDeb,'yyyy')
CodeC	Code unique d'un client donné	Entier	=Client.CodeC
CodeDip	Code unique d'un diplôme de journalisme donné	Entier	=Journaliste.CodeDI
CodeInstitutFormation	Code unique d'un institut de formation donné	Entier	=InstitutFormation.CodeIF
CodeJ	Code unique d'un journaliste donné	Entier	=Journaliste.CodeJ
CodePU	Code unique d'une publication donnée	Entier	=Publication.CodePU
CodeTypeJ	Code unique d'un type de journaliste donné	Entier	=Journaliste.CodeTY
CodeTypePU	Type d'une publication donnée	Entier	=Publication.CodeTY
Date	Date d'un numéro de publication donnée	Date	=Numero.DateNO
Date	Date de début de l'abonnement pour un client donné	Date	=Abonner.DateDeb
DateCreat	Date de création d'une publication donnée	Date	=Publication.DateCreatPU
DateFin	Date éventuelle de fin d'une publication donnée	Date	=Publication.DateFinPU
DateObtenu	Date d'obtention d'un diplôme de journalisme par un journaliste donné	Date	=Journaliste.Dateobt
GroupePresse	Groupe Presse d'une publication donnée	CC(100)	=Publication.GroupePresse
IdCentreDistributeur	Code unique d'un centre distributeur de courriers	Entier	=CentreDistributeur.CodeCD
IdPays	Code unique d'un pays fourni par l'INSEE	Entier	=Pays.CodeP
IdVille	Code unique d'une ville d'un client fourni par l'INSEE	Entier	=Ville.CodeV
Mois	Mois d'une année d'un numéro de publication donnée	CC(7)	=to_char(Numero.DateNO,'mm/yyyy')
Mois	Mois d'une année d'un numéro de publication donnée	CC(7)	=to_char(Abonner.DateDeb,'mm/yyyy')

NomC	Nom d'un client donné	CC(50)	=Client.NomC
NomCentreDistributeur	Nom d'un centre distributeur donné	CC(50)	=CentreDistributeur.NomCD
NomDip	Nom d'un diplôme de journalisme donné		=Diplome.NomDI
NomInstitutFormation	Nom d'un institut de formation de journalistes donné	CC(100)	=InstitutFormation.NomIF
NomJ	Nom d'un Journaliste donné	CC(50)	=Journaliste.NomJ
NomPays	Nom d'un Pays donné	CC(75)	=Pays.NomP
NomPU	Nom d'une publication donnée	CC(75)	=Publication.NomPU
NomTypeJ	Nom d'un type de journaliste donné	CC(75)	=TypeJour.NomTYJ
NomTypePU	Nom d'un type de publication donnée	CC(75)	=TypePU.NomTY
NomVille	Nom d'une ville d'un client donné	CC(50)	=Ville.NomV
Periodicite	Périodicité d'une publication donnée	CC(30)	=Publication.Periodicite
PrenomC	Prénom d'un client donné	CC(50)	=Client.PrenomC
PrenomJ	Prénom d'un Journaliste donné	CC(50)	=Journaliste.PrenomJ
PrixCar	Prix du caractère pour un type de journaliste donné (en €)	Réel	=TypeJour.PrixCar
Semaine	Numéro de semaine d'une année d'un numéro de publication donnée	CC(7)	=to_char(Numero.DateNO,'ww/yyyy')
Semaine	Numéro de semaine d'une année de l'abonnement pour un client donné	CC(7)	=to_char(Abonner.DateDeb,'ww/yyyy')
Semestre	Semestre d'une année de l'abonnement pour un client donné	CC(6)	=to_char((to_char(Abonner.DateDeb,'Q')+1)/2,varchar) to_char('ww/yyyy')
Trimestre	Trimestre d'une année de l'abonnement pour un client donné	CC(6)	=to_char(Abonner.DateDeb,'Q/yyyy')
VilleCentreDistributeur	Ville d'un centre distributeur donné	CC(50)	=CentreDistributeur.VilleCD

- Schéma complète

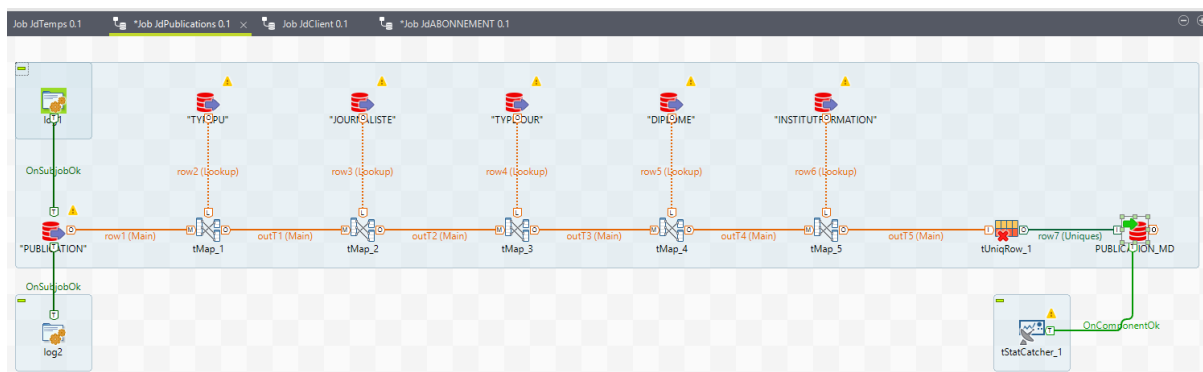


Ce magasin de données permet de calculer le nombre numéros, le montant total versé et le bénéfice réalisé en matière d'abonnement. Pour les publications et les clients, le service Marketing demande d'extraire toutes les informations possibles de l'entrepôt de données puisque l'on convertit toutes les relations CIF sur le schéma E/A d'entrepôt de données et toutes les règles métier en hiérarchie. Le trimestre peut être calculé par le mois et celui-ci permet de calculer le semestre, on les met donc dans une même hiérarchie. Le numéro de semaine dans un an peut exister dans deux mois, on distingue en deux hiérarchies différentes.

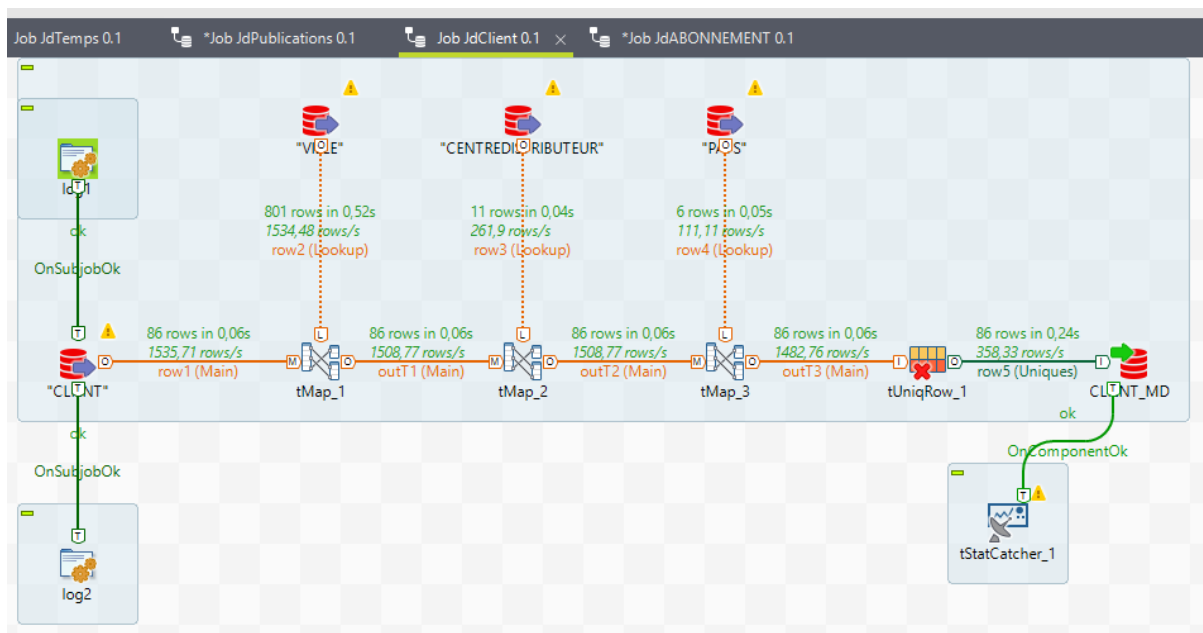
3.3.3 Modélisation logique

L'outil d'intégration de données type ETL Talend est utilisé pour déployer ce magasin de donnée. Les maquettes de réalisation sur Talend sont ci-dessous.

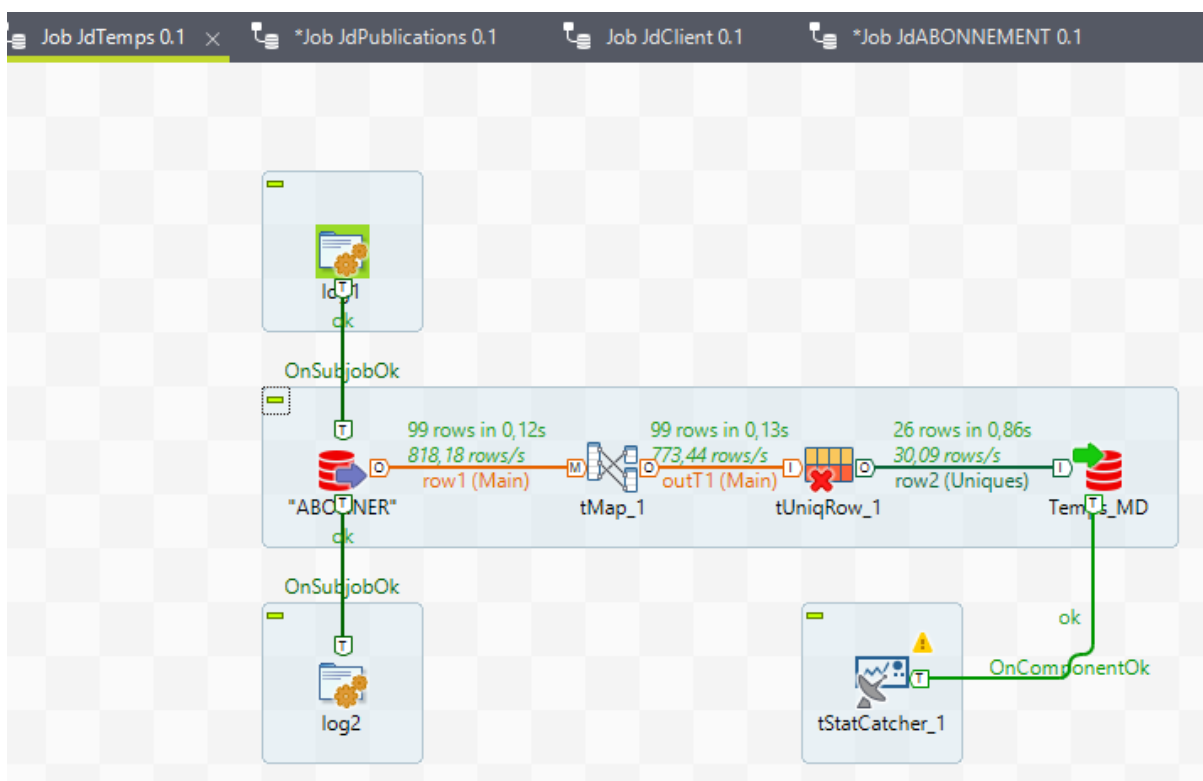
La dimension de publication :



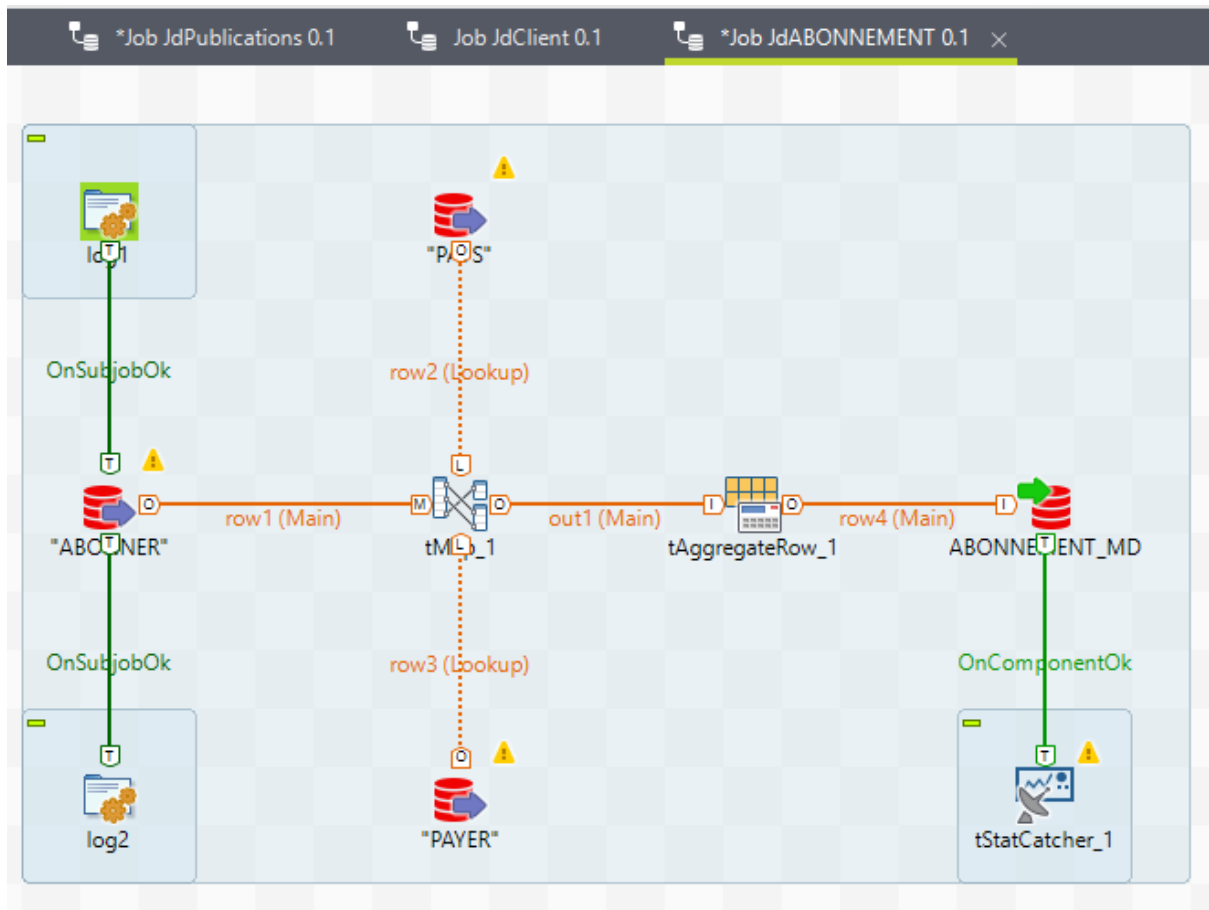
La dimension de client :



La dimension de temps :



Le fait d'abonnement :

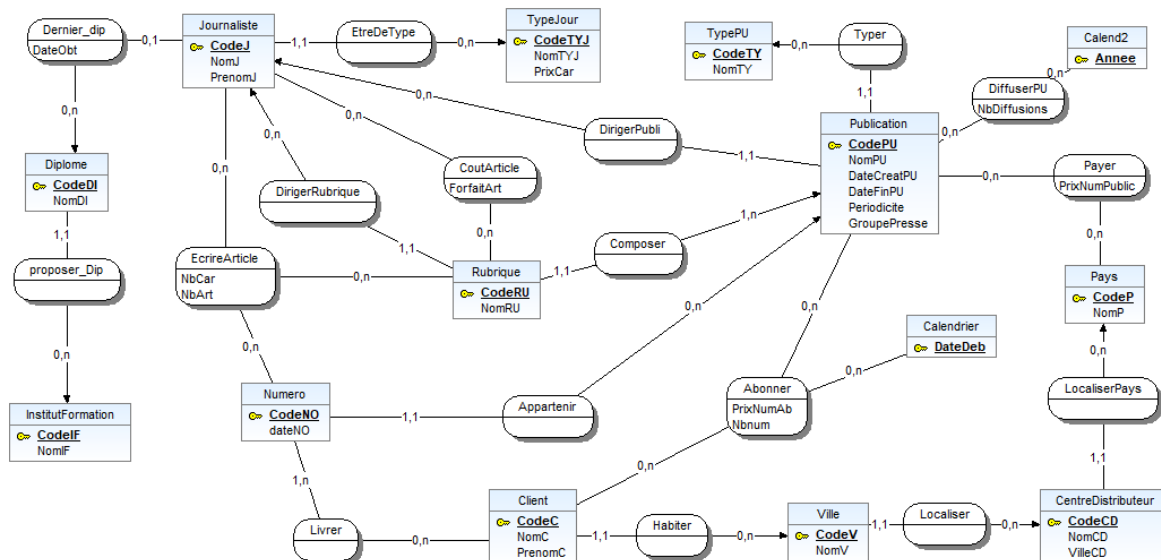


4. Entrepôt de données

4.1. Modélisation des données :

Conception détaillée

- Schéma conceptuel des données avec dictionnaire des données associé



- Schéma logique relationnel déduit
- Journaliste(CodeJ, NomJ, PrenomJ, CodeTYJ#, CodeDI#, Dateobt)
- Diplome(CodeDI, NomDI, CodeIF#)
- InstitutFormation(CodeIF, NomIF)
- TypeJour(CodeTYJ, NomTYJ, PrixCar)
- Numero(CodeNo, DateNO, CodePU#)
- Rubrique(CodeRU, NomRU, CodeJ#, CodePU#)
- Publication(CodePU, NomPU, DateCreatPU, GroupePresse, DateFinPU, Periodicite, CodeJ#, CodeTY#)
- TypePU(CodeTY, NomTY)
- DiffuserPU(CodePU#, Annee, NbDiffusions)
- Client(CodeC, NomC, PrenomC, CodeV#)
- Ville(CodeV, NomV, CodeCD#)
- CentreDistributeur(CodeCD, NomCD, VilleCD, CodeP#)
- Pays(CodeP, NomP)
- EcrireArticle(CodeJ#, CodeNO#, CodeRU#, NBCar, NbArt)
- Livrer(CodeNo#, CodeC#)
- Abonner(CodeC#, CodePU#, DateDeb, PrixNumAb, Nbrnum)
- Payer(CodePU#, CodeP#, PrixNumPublic)
- CoutArticle(CodeRU#, CodeJ#, ForfaitArt)

5. Conclusion

Besoin	Restitution	Magasin
B1	B1.1 Etat Requête - oracle B1.2 Etat Requête - oracle B1.3 Etat Requête - oracle	MD_DRH
B2	B2.1 Rapport BO - I B2.2 Rapport BO - I B2.3 Rapport BO - I	MD_RGP
B3	B3.1 Graphique - Tableau B3.2 Graphique - Tableau B3.3 Graphique - Tableau	MD_Marketing