

MASTER 1

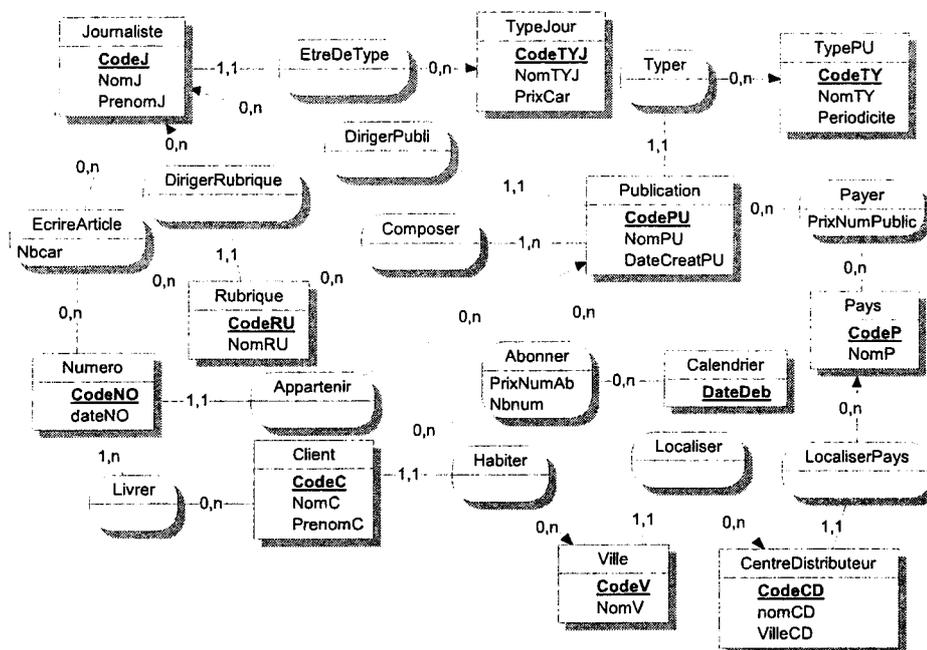
BASES DE DONNÉES
(durée 1h30)

Vendredi 11 janvier 2013 ~ 14h00 -15h30

R. TOURNIER

Les supports de cours et TD ainsi que les appareils électroniques sont interdits..

Une maison d'édition a élaboré un entrepôt de données pour étudier ses publications. Cet entrepôt permet de suivre les productions (publications) des journalistes ainsi que les abonnés au journal. En outre, l'entrepôt permet de connaître de quel centre les abonnés dépendent. Le schéma conceptuel dans le formalisme entité association de cet entrepôt est le suivant :



Chaque journaliste appartient à un type (grand reporter, directeur...) qui indique le prix payé au caractère rédigé. Le salaire d'un journaliste dépend du nombre de caractères rédigés dans leurs articles. Les journalistes écrivent des articles dans rubriques (politique, économie, international...), ces articles font un certain nombre de caractères et ils paraissent dans des numéros d'une publication. Une publication (Le Monde, Elle, Maison&Travaux...) est composée de rubriques et a été créée à une date donnée. Chaque rubrique et chaque publication sont dirigées par un journaliste. En outre, chaque publication a une catégorie (informations, mode...) avec une certaine périodicité (hebdomadaire, bi-mensuelle...).

Les publications ont un prix public qui dépend du pays dans lequel la publication est vendue. Les abonnés sont des clients qui s'abonnent pour une année à une date (DateDeb) et qui payent un prix réduit par numéro (PrixNumAb) pour un abonnement qui représente un certain nombre de numéros (Nbnum). Ce dernier dépend de la périodicité de la publication. Les clients, en fonction de leur ville dépendent d'un centre de distribution (non nécessairement dans la même ville) et ce centre est localisé dans un pays.

Le schéma relationnel de l'entrepôt de données est le suivant :

- ◆ **Journaliste**(CodeJ, NomJ, PrenomJ, CodeTYJ#)
- ◆ **TypeJour**(CodeTYJ, NomTYJ, PrixCar)
- ◆ **Numero**(CodeNo, DateNO, CodePU#)
- ◆ **Rubrique**(CodeRU, NomRU, CodeJ#)
- ◆ **Publication**(CodePU, NomPU, DateCreatPU, CodeJ#, CodeTYJ#)
- ◆ **TypePU**(CodeTY, NomTY, Periodicite)
- ◆ **Client**(CodeC, NomC, PrenomC, CodeV#)
- ◆ **Ville**(CodeV, NomV, CodeCD#)
- ◆ **CentreDistributeur**(CodeCD, NomCD, VilleCD, CodeP#)
- ◆ **Pays**(CodeP, NomP)

- ◆ **EcrireArticle**(CodeJ#, CodeNO#, CodeRU#, NBCar)
- ◆ **Livrer**(CodeNo#, CodeC#)
- ◆ **Composer**(CodeRU#, CodePU#)
- ◆ **Abonner**(CodeC#, CodePU#, DateDeb, PrixNumAb, Nbnum)
- ◆ **Payer**(CodePU#, CodeP#, PrixNumPublic)

Dans ce schéma, les attributs clés primaires sont soulignés et les attributs clés étrangères sont suffixés d'un #.

Travail à réaliser

1. Modélisation d'un Magasin de Données (MD) multidimensionnel

Cet entrepôt de données sert de support à l'élaboration d'un Magasin de Données Multidimensionnelles pour le directeur des ventes. Ce directeur souhaite connaître le nombre et le montant des abonnements par année et par mois pour chaque publication et pour chaque client. Afin de lui offrir un maximum d'opportunités, les dimensions non temporelles de ce magasin doivent intégrer tous les niveaux d'agrégation qu'il est possible d'extraire de l'entrepôt et toutes les informations associées. Les montants correspondent aux prix au numéro multiplié par le nombre de numéros.

- 1.1. Donner le contenu du dictionnaire des mesures (les colonnes de ce dictionnaire sont Code, Description, Type et Formule d'extraction). Noter que la description est très importante ;
- 1.2. Donner une représentation graphique complète du schéma conceptuel de ce magasin de données (respecter le formalisme multidimensionnel étudié en cours).

2. Interrogation SQL de l'entrepôt de données

Écrire les requêtes SQL suivantes qui sont exécutées sur les tables de l'entrepôt (les tables de l'entrepôt sont décrites par le schéma relationnel).

- 2.1. Donner, le nombre d'articles écrits pour chaque journaliste et type de journaliste ;
- 2.2. Donner le nom et prénom des journalistes ne publiant que dans publications de périodicité « quotidienne » ;
- 2.3. Pour chaque publication ayant plus de 1000 abonnés, donner le nombre total d'abonnements, ainsi que le type de publication.

3. Classeur décisionnel

Le service de paye des journalistes utilise un classeur de données Excel. Répondre à la question 3.1 ci-après.

Code Journaliste	Nom Journaliste	Prénom Journaliste	Type Journaliste	Prix au caractère	Numéro	Nombre de caractères	Date Parution	Nom Rubrique
J1	Jean	Sairien	Reporter	0 50 €	LM4562	1450	01/02/2012	Infos
J1	Jean	Sairien	Reporter	0 50 €	EL8756	2340	01/03/2012	Reportages
J1	Marc	Dekafé	Grand Reporter	0 75 €	LM4562	2875	01/02/2012	Reportages
J2	Ella	Froid	Stagiaire	0 05 €	EL8755	255	01/02/2012	Politique
J2	Ella	Froid	Stagiaire	0 05 €	MT745	540	01/04/2012	Economie
J2	Ella	Froid	Stagiaire	0 05 €	MT745	185	01/04/2012	Jeux
J3	Charles	De Takikardie	Retraité	0 45 €	LM4562	835	01/02/2012	Infos
J3	Charles	De Takikardie	Retraité	0 45 €	MT745	460	01/04/2012	Jeux
J4	Al	Quollic	Indépendant	0 25 €	MT745	1860	01/04/2012	Reportages
J5	Ramon	Heur	Journaliste AFP	0 75 €	EL8755	980	01/02/2012	Infos
J6	Alain	Provist	Reporter	0 50 €	LM4562	1680	01/02/2012	International
J6	Alain	Provist	Reporter	0 50 €	MT745	2380	01/04/2012	Reportages

Ce classeur est composé d'une feuille appelée **Data** contenant un tableau universel (cf. image précédente) avec le code, prénom, nom et type d'un journaliste, ainsi que son salaire au caractère, le numéro dans lequel il a publié, la date de parution du numéro et la rubrique dans laquelle se trouve l'article.

Paye d'un journaliste			
Journaliste			
Code	J1	Nom	Sairien
		Prénom	Jean
Paye			
Nombre d'articles	3		
Montant à verser	4 051.25 €		

Le tableau universel défini précédemment sert de support à l'élaboration de la feuille **Paye** (cf. image précédente). Dans cette feuille, seule la cellule B4 est une cellule de saisie. La feuille permet d'afficher le journaliste dont le code est saisi en B4 et la paye qui lui est due.

3.1. Donner la formule en E4 et sachant qu'un journaliste est payé au caractère donner la formule de la cellule B9.

*Remarque : une solution pour B9 nécessite la création de nouvelles colonnes calculées dans la feuille **Data**. Si une telle solution est envisagée, utiliser, dans la feuille **Data**, la colonne J, puis K, etc. En outre, donner, pour chaque nouvelle colonne créée ainsi, la formule qui se trouve en ligne 4 (c'est-à-dire les formules contenues dans les cellules J4, puis K4, etc.).*

Annexe 1 : extrait des fonctions Excel utiles pour l'examen

Fonctions de recherche :

Syntaxe	<p>RECHERCHEV(valeur_cherchée ; tableau ; no_index_col ; valeur_Proche)</p> <ul style="list-style-type: none"> ◆ valeur_cherchée : est la valeur à trouver dans la première colonne de la matrice ◆ tableau : est la table de données dans laquelle est exécutée la recherche de la valeur. ◆ no_index_col : est le numéro de la colonne du tableau dont la valeur correspondante doit être renvoyée. Si l'argument no_index_col est égal à 1, la fonction renvoie la valeur dans la première colonne de l'argument tableau et ainsi de suite. ◆ Valeur_proche : <ul style="list-style-type: none"> ◆ VRAI ou OMIS : si on souhaite une recherche sur une valeur proche (autrement dit, sur une valeur égale ou immédiatement inférieure à la valeur recherchée). Cette option <u>nécessite un classement croissant</u> des données de la première colonne. ◆ FAUX : si on souhaite que la fonction de recherche trouve exactement la valeur recherchée. Cette option <u>ne nécessite pas un classement croissant</u> des données de la première colonne.
Résultat	Cette fonction cherche une valeur donnée dans la colonne située à l'extrême gauche d'une matrice et renvoie une valeur dans la même ligne d'une colonne que vous spécifiez dans la matrice.

Syntaxe	<p>DECALER(adress ; nblig ; nbcoll)</p> <ul style="list-style-type: none"> ◆ adress adresse de référence ◆ nblig nombre de lignes à décaler (entier) ◆ nbcoll nombre de colonnes à décaler (entier)
Résultat	Cette fonction renvoie le contenu d'une cellule décalée de nblig lignes et nbcoll colonnes par rapport à l'adresse de référence donnée (adress)

EQUIV(valeur_cherchée ; tab ; type_recherche) renvoie la position relative de **valeur_cherchée** (valeur ou référence à une cellule) dans le tableau (plage de cellules adjacentes) intitulé **tab**. **Type_recherche** influence le type de la recherche (indique comment Excel doit effectuer des comparaisons entre la valeur recherchée et les valeurs du tableau). **Type_recherche** peut avoir 3 valeurs :

- ◆ **1** : la fonction **EQUIV** trouve la valeur la plus élevée qui est inférieure ou égale à celle de l'argument **valeur_cherchée**. Les valeurs de **tab** doivent être placées en ordre croissant.
- ◆ **0** : la fonction **EQUIV** trouve la première valeur exactement équivalente à celle de l'argument **valeur_cherchée**. Les valeurs de l'argument **tab** peuvent être placées dans un ordre quelconque.
- ◆ **-1** : la fonction **EQUIV** trouve la plus petite valeur qui est supérieure ou égale à celle de l'argument **valeur_cherchée**. Les valeurs de **tab** doivent être placées en ordre décroissant.

Fonctions conditionnelles

Syntaxe	Résultat
SOMME.SI(plageC;critère;plageS)	Additionne les cellules de la plage "PlageS" si elles répondent au critère de sélection ("Critère") appliqué aux cellules de "PlageC"
NB.SI(plage;critère)	Compte le nombre de cellules à l'intérieur d'une plage qui répondent à un critère donné.

Remarque : un critère peut-être une constante, une adresse de cellule ou une expression calculée. Le critère doit être une chaîne de caractères (exprimé entre guillemets), par exemple : ">10" ou "="&A1 (les références des cellules doivent être associées avec l'opérateur de concaténation "&").

Annexe 2 : rappel de la syntaxe d'une requête

La syntaxe d'une requête de sélection est la suivante :

```
SELECT R1.Att_Regroupement, Fonction(R2.Att_Agreg) AS [En_Tete],...  
FROM Relation1 R1, Relation2 R2, ...  
WHERE Condition_Jointure (...) AND Condition_Selection (...)  
GROUP BY R1.Att_regroupement,...  
HAVING Conditions_regroupement AND ...  
ORDER BY R1.Att type_d_ordre ... ;
```