

# BD51

## Business Intelligence & Data Warehouse

**Examen Final du 17/1/2011**

Rédacteur : Christian FISCHER

Département Informatique

Automne 2010

Conditions de réalisation :

Documents autorisés : Supports de Cours et TP de BD51

Durée : 2h

Barème : Le sujet comprend 2 dossiers indépendants.

Dossier 1 8 pts	1. Requêtes SQL OLAP Rédiger les ordres SQL	5pts
	2. Requêtes MDX Rédiger les ordres MDX	3 pts
Dossier 2 12 pts	Modélisation du Data Warehouse	
	1. Modèle dimensionnel	5 pts
	2. Partitionnement de table	2 pts
	3. Partitionnement de table	1 pt
	4. Vue matérialisée	2 pts
	5. Architecture du système d'information décisionnel	2 pts

## **Dossier 1 : Interrogation du système d'information de la distribution**

La société Distrisys est une société européenne, spécialisée dans la distribution de grands et petits appareils d'électroménager. La société est distributeur officiel et exclusif d'appareils qu'elle fait fabriquer, importe puis vend sous ses propres marques : Lagon, Cuccina et Pierre-Michel.

Distrisys ne vend pas ses produits en direct aux consommateurs. Elle dispose d'un réseau de revendeurs et distributeurs agréés.

Son réseau est de plusieurs types : réseau de distributeurs généraliste, réseau de magasins spécialisés en enfin réseau de sites marchands.

Distrisys est issue historiquement de plusieurs sociétés acquises par croissance externe. La société dispose de 5 sites géographiques :

3 en France : le siège à Paris, Aix-en-provence et Bordeaux

2 à l'étranger : Barcelone et Munich

Chaque site dispose de son propre progiciel de gestion intégré (PGI ou ERP), qui a été toutefois homogénéisé pour en faciliter la gestion et la maintenance.

La direction est depuis très longtemps confrontée à des problèmes de qualité des données et souhaite se doter d'outils permettant d'appréhender leurs données de manière globale.

Une première étude a été menée pour définir les contours du projet décisionnel. Le périmètre a été défini. Il s'agit dans l'ordre des priorités :

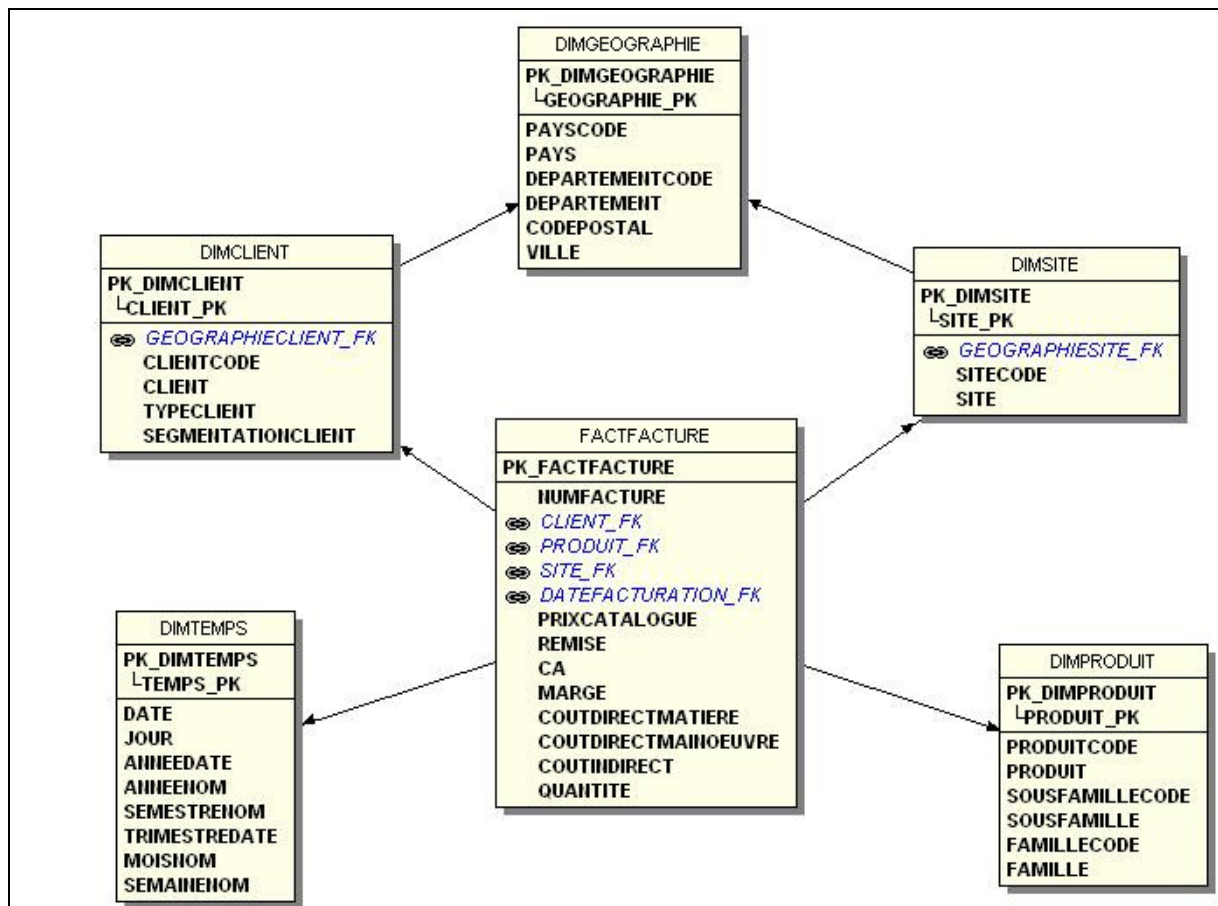
1. Facturation (facture, commande et budget)
2. Stocks
3. Achats (commandes fournisseurs, commandes en attentes et facturation fournisseurs)
4. Service après-vente
5. Marketing

Les données de la facturation seront analysables par les utilisateurs selon 4 axes :

- L'axe produit
- L'axe client
- L'axe site (pour connaître l'origine de la vente)
- L'axe temps

## Dossier 1 Première partie : Ecriture de requêtes SQL - OLAP

Modèle logique de données relationnel de la base de données DistrisysDW pour le suivi de la facturation



La table FactFacture ne comporte pas de clé de primaire.

## Contenu de la table Géographies

	Geographie_PK	PaysCode	Pays	DepartementCode	Departement	CodePostal	Ville
1	0	INC	Inconnu	INC	Inconnu	Inconnu	Inconnu
2	1	FR	France	13	Bouches-du-Rhône	13001	Marseille
3	2	FR	France	69	Rhône	69001	Lyon
4	3	FR	France	75	Paris	75001	Paris
5	4	FR	France	33	Gironde	33001	Bordeaux
6	5	FR	France	59	Nord	59001	Lille
7	6	FR	France	31	Haute Garonne	31001	Toulouse
8	7	FR	France	67	Bas-Rhin	67000	Strasbourg
9	8	FR	France	13	Bouches-du-Rhône	13100	Aix-en-Provence
10	9	FR	France	25	Franche-Comté	25000	Besançon
11	10	DE	Allemagne	091	Haute-Bavière	58352	Munich
12	11	ES	Espagne	08	Catalogne	08006	Barcelone

## Contenu de la table des clients

	Client_PK	GeographieClient_FK	ClientCode	Client	TypeClient	SegmentationClient
1	0	0	INC	Inconnu	Inconnu	Inconnu
2	1	1	C1	LaBoutiqueOnLine.com	Site Marchand	Bon Client
3	2	3	C2	Maison Discount	Discounteur	Bon Client
4	3	8	C3	Cuisine du sud	Spécialiste	Tiède
5	4	4	C4	Discount plus	Discounteur	Tiède
6	5	2	C5	EquiperSaMaison.com	Site Marchand	Très Bon Client
7	6	3	C6	Hypermarché Youpi	Grande surface	Très Bon Client
8	7	10	C7	EineKüche	Spécialiste	Bon Client
9	8	11	C8	Mercado Del Sol	Grande surface	Bon Client
10	9	1	C9	ElectroYoupa	Spécialiste	Bon Client
11	10	5	C10	Toutmoinscher.com	Site Marchand	Tiède

## Contenu de la table des sites

	Site_PK	GeographieSite_FK	SiteCode	Site
1	1	3	D001	Siège social
2	2	8	D002	Agence Sud
3	3	4	D003	Agence Ouest
4	4	11	D004	Agence Europe Sud
5	5	10	D005	Agence Europe Est

## Contenu de la table des produits

	Produit_PK	ProduitCode	Produit	SousFamilleCode	SousFamille	FamilleCode	Famille	Valide
1	0	INC	Inconnu	INC	Inconnu	INC	Inconnu	1
2	1	LL1100	LAGON LL1100	GMLL	Lave-Linge	GM	Gros Menager	1
3	2	LL1200	LAGON LL1200	GMLL	Lave-Linge	GM	Gros Menager	1
4	3	LV1620	LAGON LV 1620	GMLV	Lave-Vaisselle	GM	Gros Menager	1
5	4	SL1000	LAGON SL 1000	GMSL	Seche-Linge	GM	Gros Menager	1
6	5	F120	Pierre Michel F120	GMF	Four	GM	Gros Menager	1
7	6	R080	Pierre Michel R 080	GMR	Refrigerateur	GM	Gros Menager	1
8	7	GP700	Cuccina GP 700	PMGP	Grille-Pain	PM	Petit Menager	1
9	8	C470	Cuccina C 470	PMC	Cafetière	PM	Petit Menager	0
10	9	RC3000p	Cuccina RC 3000+	RC	Robot Cuisine	PM	Petit Menager	1
11	10	C260	Cuccina C 260	PMC	Cafetière	PM	Petit Menager	0
12	11	C270	Cuccina C 270	PMC	Cafetière	PM	Petit Menager	0
13	12	C470	Cuccina C 470	CC	Cafetière	C	Cafetière	1
14	13	C260	Cuccina C 260	CC	Cafetière	C	Cafetière	1
15	14	C270	Cuccina C 270	CC	Cafetière	C	Cafetière	1

## Extrait de la table Temps

	Temps_PK	Date	Jour	AnneeCode	AnneeDate	AnneeNom	SemestreCode	SemestreDate
1	0	1900-01-01 00:00:00	inconnu	19000101	1900-01-01 00:00:00	inconnu	19000101	1900-01-01 00:00:00
2	20060101	2006-01-01 00:00:00	dimanche, janvier 01 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
3	20060102	2006-01-02 00:00:00	lundi, janvier 02 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
4	20060103	2006-01-03 00:00:00	mardi, janvier 03 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
5	20060104	2006-01-04 00:00:00	mercredi, janvier 04 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
6	20060105	2006-01-05 00:00:00	jeudi, janvier 05 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
7	20060106	2006-01-06 00:00:00	vendredi, janvier 06 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
8	20060107	2006-01-07 00:00:00	samedi, janvier 07 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
9	20060108	2006-01-08 00:00:00	dimanche, janvier 08 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
10	20060109	2006-01-09 00:00:00	lundi, janvier 09 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
11	20060110	2006-01-10 00:00:00	mardi, janvier 10 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00
12	20060111	2006-01-11 00:00:00	mercredi, janvier 11 2006	20060101	2006-01-01 00:00:00	Calendrier 2006	20060101	2006-01-01 00:00:00

## Extrait de la table des factures

	NumFacture	DateFacturation_FK	Site_FK	Produit_FK	Client_FK	PrixCatalogue	Remise	CA	Marge	CoutDirectMatiere	CoutDirectMainOeuvre	CoutIndirect	Quantite
1	F1	20100117	3	7	1	1500.00	0.00	1500.00	230.00	870.00	380.00	20.00	20.00
2	F2	20100115	3	8	5	1620.00	84.00	450.00	42.00	240.00	162.00	6.00	18.00
3	F2	20100121	3	7	7	1980.00	60.00	1920.00	370.00	1000.00	510.00	40.00	20.00
4	F3	20100117	3	10	9	504.00	24.00	480.00	112.00	248.00	112.00	8.00	8.00
5	F4	20100121	3	8	6	558.00	12.00	546.00	36.00	312.00	186.00	12.00	6.00
6	F5	20100117	3	10	3	728.00	24.00	704.00	160.00	384.00	144.00	16.00	8.00
7	F6	20100119	3	9	10	368.00	8.00	360.00	40.00	216.00	96.00	8.00	4.00
8	F7	20100121	3	3	3	19475.00	0.00	19475.00	7975.00	8950.00	2325.00	225.00	25.00
9	F8	20100126	3	3	9	37875.00	0.00	37875.00	6475.00	19325.00	11325.00	750.00	75.00
10	F9	20100122	3	6	8	560.00	7.00	553.00	217.00	189.00	140.00	7.00	7.00
11	F10	20100129	3	5	5	41919.00	1413.00	40506.00	15543.00	14915.00	9263.00	785.00	471.00

## **Travail à faire**

Q1. Rédiger les requêtes SQL ci-dessous.

Utilisation des jointures ANSI et des alias de tables et des alias de colonnes selon le résultat attendu

Pour chaque requête le résultat doit être trié selon le résultat.

### **Requête 1 :**

Calculer le chiffre d'affaires total et quantité totale des ventes pour chaque segment de client

Résultat attendu

	Segmentation	Chiffre d'affaires	Quantité
1	Bon Client	24822623.48	98664.00
2	Tiède	12595672.88	55328.00
3	Très Bon Client	10556659.52	38442.00

### **Requête 2 :**

Calculer le chiffre d'affaires total et quantité totale des ventes pour chaque segment de client

Utilisation d'un case pour trier le résultat

Résultat attendu

	Segmentation	Chiffre d'affaires	Quantité
1	Très Bon Client	10556659.52	38442.00
2	Bon Client	24822623.48	98664.00
3	Tiède	12595672.88	55328.00

### **Requête 3 :**

Calculer le chiffre d'affaires total et quantité totale des ventes pour chaque pays et segment de client

Résultat attendu

	Pays Client	Segmentation	Chiffre d'affaires	Quantité
1	Allemagne	Bon Client	3618063	13822
2	Espagne	Bon Client	5398768	20400
3	France	Bon Client	15805791	64442
4	France	Tiède	12595672	55328
5	France	Très Bon Client	10556659	38442

**Requête 4 :**

Calculer le chiffre d'affaires total et quantité totale des ventes pour chaque pays et segment de client

Utilisation d'une clause Grouping Sets + Décodification des valeurs nulles

**Résultat attendu**

	Pays Client	Segmentation	Chiffre d'affaires	Quantité
1	France	Très Bon Client	10556659	38442
2	France	Tiède	12595672	55328
3	France	Bon Client	15805791	64442
4	France	_Total Segment	38958123	158212
5	Espagne	Bon Client	5398768	20400
6	Espagne	_Total Segment	5398768	20400
7	Allemagne	Bon Client	3618063	13822
8	Allemagne	_Total Segment	3618063	13822
9	_Total Pays	_Total Segment	47974955	192434

**Requête 5 :**

Calculer le chiffre d'affaires total et quantité totale des ventes pour chaque pays.

Ajouter une fonction de classement sur le chiffre d'affaires et la quantité.

Tri sur le rang du CA.

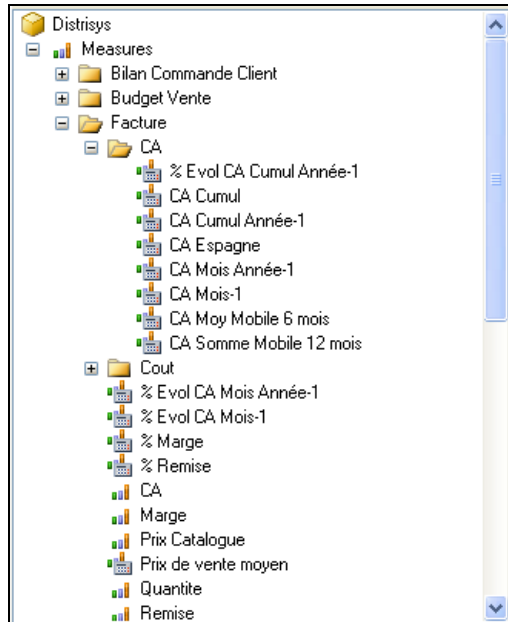
**Résultat attendu**

	Pays Client	Chiffre d'affaires	Rang CA	Quantité	Rang Qté
1	France	38958123	1	158212	1
2	Espagne	5398768	2	20400	2
3	Allemagne	3618063	3	13822	3

## Dossier 1 Seconde partie : Ecriture de requêtes MDX

Vous disposez d'un cube OLAP DISTRISYS

Présentation des mesures du cube Distrisys avec les indicateurs du groupe de mesures Facture à utiliser pour l'écriture des requêtes.



Utilisation de la dimension Client	Utilisation de la dimension Temps
<p>Pays France = [FR]</p>	<p>Calendrier 2010 = [20100101]            Calendrier 2009 = [20090101]</p>



## **Travail à faire**

Q2. Rédiger les **requêtes MDX** ci-dessous.

### **Requête 1 :**

Calculer le chiffre d'affaires total, la quantité totale, la marge totale des ventes pour chaque pays de client

#### **Résultat attendu**

	CA	Quantite	Marge
Allemagne	3 618 063,38 €	13 822	1 104 101,26 €
Espagne	5 398 768,76 €	20 400	1 702 549,70 €
France	38 958 123,74 €	158 212	11 136 204,06 €
Inconnu	(NULL)	(NULL)	(NULL)

### **Requête 2 :**

Calculer le chiffre d'affaires total, la quantité totale, la marge totale des ventes pour les clients avec le pays = France et pour les années 2010 et 2009

#### **Résultat attendu**

	CA	Quantite	Marge
Calendrier 2010	18 831 423,38 €	76 955	5 319 533,83 €
Calendrier 2009	19 479 061,87 €	79 106	5 568 102,03 €

### **Requête 3 :**

Calculer le chiffre d'affaires total par mois pour l'année 2010. Ajoutez les mesures CA du Mois Précédent et le calcul du % d'évolution par rapport au mois précédent. Uniquement les clients français sont affichés.

#### **Résultat attendu**

	CA	CA Mois-1	% Evol CA Mois-1
janvier 2010	856 340,78 €	1 863 393,23 €	-54,04%
février 2010	1 099 151,18 €	856 340,78 €	28,35%
mars 2010	1 709 868,24 €	1 099 151,18 €	55,56%
avril 2010	1 970 710,95 €	1 709 868,24 €	15,26%
mai 2010	1 390 039,90 €	1 970 710,95 €	-29,47%
juin 2010	1 773 542,45 €	1 390 039,90 €	27,59%
juillet 2010	2 360 230,66 €	1 773 542,45 €	33,08%
août 2010	1 220 709,86 €	2 360 230,66 €	-48,28%
septembre 2010	1 315 169,61 €	1 220 709,86 €	7,74%
octobre 2010	1 591 123,77 €	1 315 169,61 €	20,98%
novembre 2010	1 681 142,75 €	1 591 123,77 €	5,66%
décembre 2010	1 863 393,23 €	1 681 142,75 €	10,84%

## Dossier 2 : Modèle dimensionnel de la fréquentation des musées

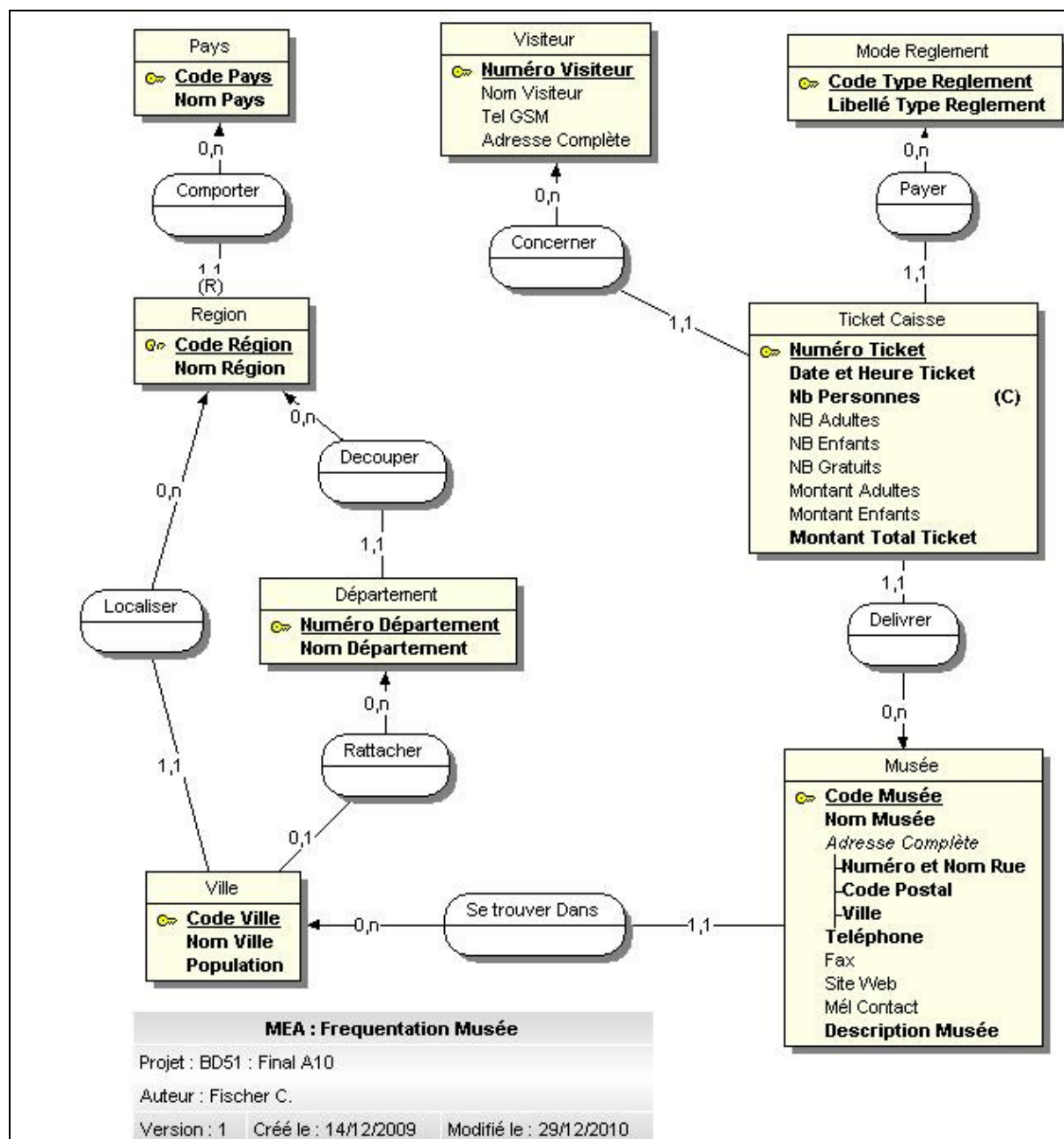
### Contexte général

Le ministère de la culture a décidé de la mise en place du système d'information décisionnel et de pilotage de la fréquentation des musées.

L'équipe chargée de l'entrepôt de données a récemment examiné la conception du système d'analyse actuel.

Mr Fischer C. a validé le modèle entité-association décrivant le système d'information relatif à la fréquentation des musées.

### Modèle entité-association validé



### **Travail à faire**

**Q1.** A partir du modèle entité-association présenté, vous devez fournir un modèle dimensionnel en étoile.

Dans le modèle fourni vous indiquerez quelles sont les tables de dimensions et les tables de faits.

**Q2.** La table « Ticket de Caisse » contient les données depuis le 1er janvier 2008. Cette table doit être partitionnée par année civile jusqu'à fin 2013.

Rédiger l'ordre SQL de création de la table partitionnée en utilisant la syntaxe d'Oracle 10G.

**Q3.** Quelle modification du modèle dimensionnel est nécessaire pour faire un partitionnement de la table « Ticket Caisse » par région et par année civile ?

**Q4.** Création d'une vue matérialisée

Rédiger l'ordre SQL de création de la vue matérialisée `agg_tk_an_mois_region_musee`

Structure de la vue matérialisée :

	Column_name	Type	Computed	Length	Prec	Scale
1	ANNEE	int	no	4	10	0
2	MOIS	int	no	4	10	0
3	CODE_MUSEE	char	no	6		
4	NOM_MUSEE	varchar	no	50		
5	CODE_VILLE	char	no	32		
6	VILLE	varchar	no	50		
7	NOM_REGION	varchar	no	50		
8	NB_Adultes	numeric	no	17	38	0
9	Nb_Enfants	numeric	no	17	38	0
10	Nb_Personnes	numeric	no	17	38	0
11	Montant_Adulte	numeric	no	17	38	2
12	Montant_Enfants	numeric	no	17	38	2
13	Montant_Total	numeric	no	17	38	2

**Q5.** Les directeurs de chaque musée doivent établir des statistiques sur la fréquentation par musée. Un directeur n'ayant accès qu'aux données du musée dont il est responsable.

Proposez une architecture technique et logicielle permettant, à ces personnes n'ayant qu'une connaissance très limitée du langage SQL, de réaliser les tableaux de bord nécessaires à leur métier.