

Conception d'une Bases de Données

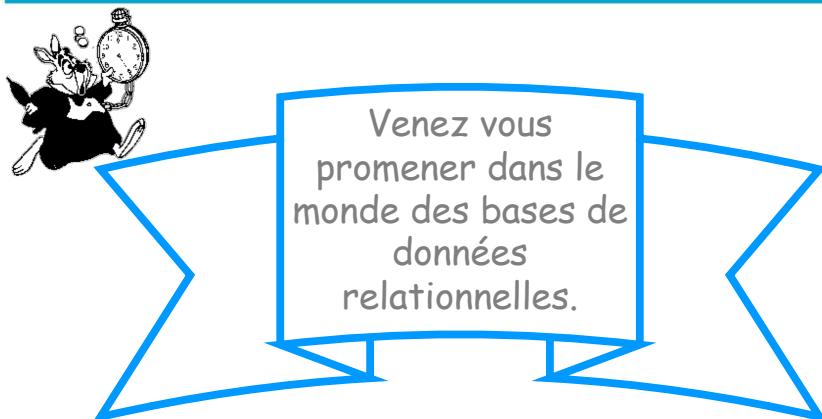
3IRC
2018/2019

F. PERRIN

membre de UNIVERSITÉ DE LYON



I have a dream today (2)



I have a dream today (1)



Ne serait-ce pas merveilleux s'il existait un monde dans lequel je saurais à qui j'ai prêté mon DVD préféré, à quelle heure part le train demain, ..., sans effectuer une fouille archéologique dans tous les post-it collés sur mon bureau ? Mais ce n'est probablement qu'un rêve...

Qu'est-ce qu'une Base de Données (1) ?



Chaque fois que vous faites une recherche sur Internet, recevez une amende pour excès de vitesse, etc., une **base de données** reçoit une demande d'informations, que l'on appelle une **requête**.



Qu'est-ce qu'une Base de Données (2) ?

- Une **base de données** est un récipient qui contient des **tables** et d'autres structures SQL liées à ces tables.
- Bases de données est une collection de données :
 - **Non redondante**.
 - Ayant souvent des **liens** entre elles.
 - **Cohérentes** (respectent des contraintes d'intégrité).
 - **Partagées** par plusieurs utilisateurs simultanément.

I have a dream today (1)



Ne serait-ce pas merveilleux s'il existait un moyen d'organiser les données et les liens qu'elles tissent entre elles de manière à les retrouver rapidement ?
Mais ce n'est probablement qu'un rêve...

Comment est gérée une BD ?

▪ SYSTÈME DE GESTION DE BASES DE DONNÉES

- Permet description, création de la base, l'utilisation et la mise à jour des données.
- Contrôle cohérence, non redondance, sécurité.



I have a dream today (2)



Découvrez les secrets des tables et de leurs lignes et colonnes.



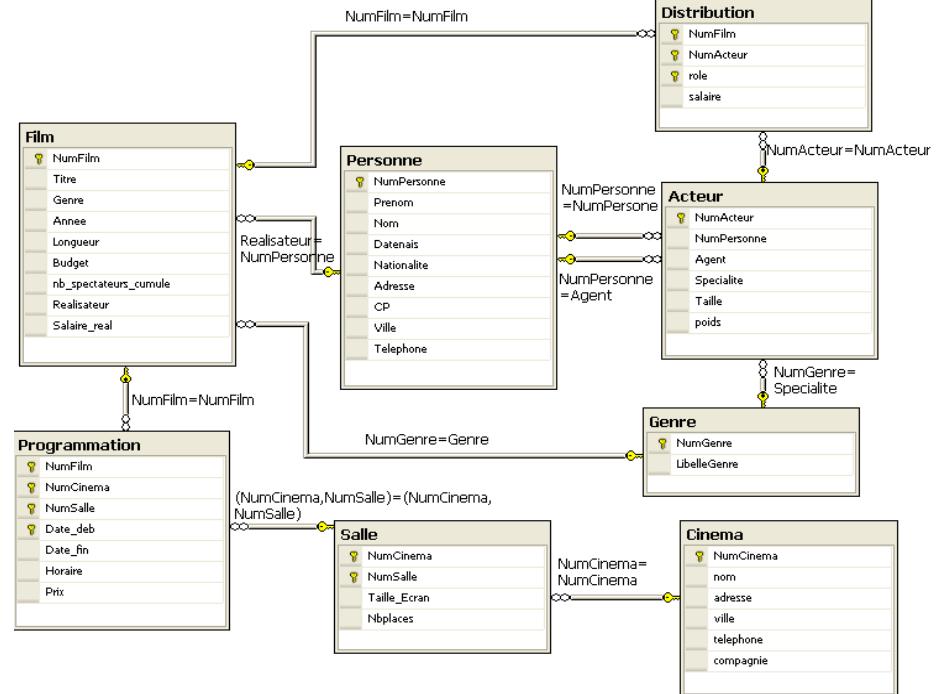
Exemple : BD Cinemas

- Soit une BD qui gère la programmation de film dans des salles de cinéma ainsi que les caractéristiques des films, acteurs et réalisateurs.
- Les utilisateurs peuvent faire des requêtes pour connaître les films programmés à 20h au 'Pathé', mais aussi la liste des films 'fantastique', réalisés par 'Spielberg', etc.
- Elle est composée des tables FILM, PERSONNE, GENRE, DISTRIBUTION, CINEMA , SALLE , PROGRAMMATION dont la description et les liens sont modélisés page suivante.



Schéma d'une BD

- Une description des données (les colonnes et les tables) de votre base de données, ainsi que de tout autre objet en dépendant et de la façon dont ils sont tous liés entre eux s'appelle un **schéma**.



I have a dream today



Qu'est-ce qu'une table ?

- Une **table** est une structure qui contient des données, organisées en **colonnes** et en **lignes** (PERSONNE, etc.).
- Une **colonne** est une donnée conservée par votre table (Nom, Nationalité, etc.).
- Une **ligne** est un ensemble unique de colonnes qui décrit les attributs d'un objet particulier.



	NumPersonne	Prenom	Nom	Datenaïs	Nationalité	Adresse	CP	Ville	Telephone
1	1	Steven	Spielberg	1947	Americain	5é Avenue	NULL	New York	555111

Qu'est-ce qu'une clé primaire (Primary Key) ?

- Une **clé primaire** est une colonne de votre table qui rend chaque enregistrement **unique** (NumPersonne, NumFilm+NumActeur+Role, etc.).
- La valeur d'une clé primaire permet **d'identifier**, de manière unique, les autres attributs d'une ligne (t-uplet, enregistrement) de la table.

	NumPersonne	Prenom	Nom	Datenaïs	Nationalité	Adresse	CP	Ville	Telephone
1	1	Steven	Spielberg	1947	Americain	5é Avenue	NULL	New York	555111

Relationnel, vous avez dit Relationnel ?

- Pour définir une table vous devez vous interroger :
 - Quelle est la **seule chose** que décrit votre table (des personnes, des films, des cinémas, etc.) ?
 - Comment allez-vous **concevoir** cette table pour qu'elle soit facile à **interroger** ?
- Une table est centrée sur ces **relations** :
 - Quelles sont les relations des données entre elles à l'intérieur d'une table ?
 - Quelles sont les relations que la table entretient avec les autres tables ?

Quelles sont les règles régissant une clé primaire ?

- Elle doit **avoir une valeur** qui doit être **unique**.
- Elle doit être **compacte** pour accélérer les temps d'accès : la plupart du temps un Id auto incrémenté est créé (surtout en remplacement d'une clé longue NumFilm+NumActeur+Role).
- Ses valeurs **ne doivent pas être modifiées** :
 - vous risqueriez de la définir à une valeur déjà utilisée. Rappelez-vous, elle doit rester unique !
 - Elle pourrait être référencée dans une autre table.



I have a dream today (1)



Ne serait-ce pas merveilleux
s'il existait un moyen
d'identifier les différents
types de relations entre les
tables ?
Mais ce n'est probablement
qu'un rêve...

I have a dream today (2)



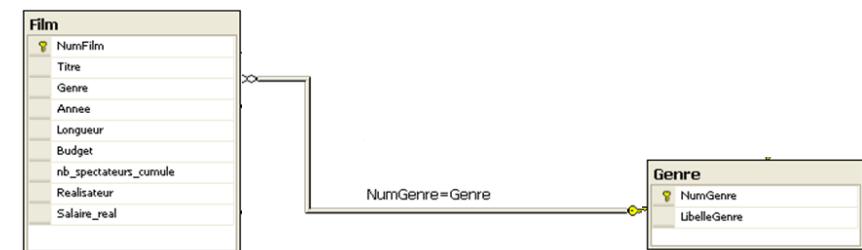
Venez à la
rencontre des
patterns de
conception des
données.

Qu'est-ce que le pattern de données Un-à-Un ?

- Une ligne table parente reliée à une ligne table enfant.
- Il est généralement plus sensé de laisser vos données un-à-un dans votre table principale, mais vous pouvez parfois trouver quelques avantages à extraire ces colonnes :
 - Ecrire des requêtes plus rapides :
 - Colonne très souvent accédée/modifiée <-> peu souvent.
 - Données de type BLOB.
 - Isoler certaines données pour restreindre leur accès.

Qu'est-ce que le pattern de données Un-à-Plusieurs ?

- Une ligne d'une table peut avoir plusieurs lignes correspondantes dans la seconde table, mais la seconde table ne peut avoir qu'une ligne correspondante dans la première.



Qu'est-ce qu'une clé étrangère (foreign key) ?

- La **CLÉ ÉTRANGÈRE** est une colonne de la table qui **référence la CLÉ PRIMAIRE** d'une autre table (Genre dans table Film référence NumGenre dans table Genre).
- Elle peut avoir un nom différent de la clé primaire à laquelle elle correspond.
- La table d'où vient la clé primaire est appelée la table parente.
- Une colonne clé étrangère **peut ne pas avoir** de valeur mais si elle a une valeur, celle-ci doit être **clé primaire** dans la table parente.



Qu'est-ce que le pattern de données Plusieurs-à-Plusieurs (2) ?

- Peut on appliquer le même principe de clé étrangère ajoutée à la table 'Femme' ?

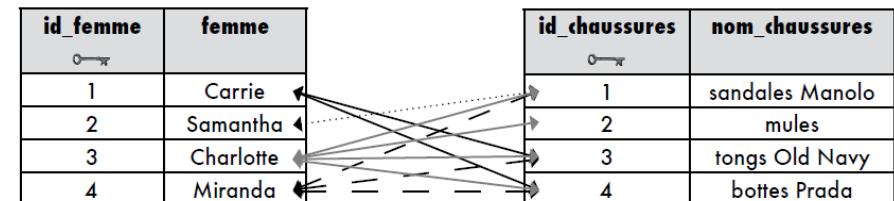
id_femme	femme	id_chaussures	nom_chaussures
1	Carrie	3	sandales Manolo
2	Samantha	1	mules
3	Charlotte	1	tongs Old Navy
4	Miranda	1	botte Prada
5	Carrie	4	
6	Charlotte	2	
7	Charlotte	3	
8	Charlotte	4	
9	Miranda	3	
10	Miranda	4	

Les deux tables sont maintenant reliées par la colonne id_chaussures.

- Pb : on a de la **redondance** sur la colonne 'femme'.

Qu'est-ce que le pattern de données Plusieurs-à-Plusieurs (1) ?

- Une ligne d'une table peut avoir **plusieurs** lignes correspondantes dans la seconde table, et dans la seconde table **une ligne** peut avoir **plusieurs** lignes correspondante dans la première.



Qu'est-ce que le pattern de données Plusieurs-à-Plusieurs (3) ?

- Ou bien, peut on appliquer le principe de clé étrangère ajoutée à la table 'Chaussure' ?

id_chaussures	nom_chaussures	id_femme	femme
1	sandales Manolo	3	Carrie
2	sandales Manolo	2	Samantha
3	tongs Old Navy	1	Charlotte
4	botte Prada	1	Miranda
5	mules	3	
6	tongs Old Navy	3	
7	botte Prada	3	
8	sandales Manolo	4	
9	tongs Old Navy	4	
10	botte Prada	4	

- Pb : on a de la **redondance** sur la colonne 'nom_chaussure'.

Qu'est-ce que le pattern de données Plusieurs-à-Plusieurs (3) ?

- Solution pour éviter la redondance dans une relation n-à-m : les tables sont reliées par une **table de jonction**.
- La clé primaire de la table de jonction est la **concaténation des clés primaires** des autres tables, souvent remplacée par un Id auto-incrémenté.

id_femme	id_chaussures
1	3
1	4
2	1
3	1
3	2
3	3
3	4
4	1
4	3
4	4

I have a dream today (1)

I have a dream today (1)

Ne serait-ce pas merveilleux s'il existait un moyen d'être certaine que les données dans ma BD sont non redondantes et que les MAJ les laisseront dans un état cohérent ? Mais ce n'est probablement qu'un rêve..



I have a dream today (2)



Normalisation des relations (1)

- L'objectif de la normalisation est d'éviter les **redondances de données**, et donc :
 - Limiter l'**espace disque** nécessaire.
 - Limiter les **incohérences** de données qui pourraient les rendre inutilisables.
 - Eviter les processus de **mise à jour** (réécritures pour retrouver cohérence).
- L'inconvénient étant :
 - Des **temps d'accès** potentiellement plus longs si les requêtes sont trop complexes (risque de lectures plus lentes, en général compensé par utilisation d'index).

Normalisation des relations (2)

- Vos relations seront normalisées si elles sont au moins en **3^e Forme Normale de Boyce Codd** (majorité des BD) voire en 4^e et en 5^e FN.
- Les FN s'appuient sur 2 concepts :
 - Dépendance fonctionnelle** : B dépend de A si pour une valeur de A correspond 1 unique valeur de B.
 - Clé primaire** d'une relation : ensemble minimum d'attribut qui détermine tous les autres.



1^{ère} Forme Normale – 1FN (2)

Enfin en 1FN

id_jouet	jouet
0	
5	raquette
6	frisbee
9	cerf-volant
12	yoyo

id_jouet	couleur
5	blanc
5	jaune
5	bleu
6	vert
6	jaune
9	rouge
9	bleu
9	vert
12	blanc
12	jaune

1^{ère} Forme Normale – 1FN (1)

1[°] FN : tout attribut est **atomique** : pas de notion de liste, ni de colonnes dupliquées.

Pas en 1FN

id_jouet	jouet	couleurs
5	raquette	blanc, jaune, bleu
6	frisbee	vert, jaune
9	cerf-volant	rouge, bleu, vert
12	yoyo	blanc, jaune



Toujours pas en 1FN

id_jouet	jouet	couleur1	couleur2	couleur3
5	raquette	blanc	jaune	bleu
6	frisbee	vert	jaune	
9	cerf-volant	rouge	bleu	vert
12	yoyo	blanc	jaune	

2^{ème} Forme Normale – 2FN (1)

2[°] FN : être en 1FN + tout attribut dépend d'une clé par 1 **DF élémentaire** : un attribut non clé ne dépend pas que d'une partie de la clé.

id_jouet	id_magasin	couleur	inventaire	adresse_magasin
0	+			
5	1	blanc	34	23 rue des érables
5	3	jaune	12	100 boulevard du nord
6	1	bleu	5	23 rue des érables
6	2	vert	10	1902 allée de l'ambre
6	4	jaune	24	17 route moulin de pont
9	1	rouge	50	23 rue des érables
9	2	bleu	2	1902 allée de l'ambre
12	2	vert	18	1902 allée de l'ambre
12	4	blanc	28	17 route moulin de pont
9	4	jaune	11	17 route moulin de pont

2^{ème} Forme Normale – 2FN (2)

infos_jouets				
id_jouet	jouet	couleur	prix	poids
1	raquette	blanc	1.95	0.3
2	raquette	jaune	2.20	0.4
3	frisbee	bleu	1.95	0.3
4	frisbee	vert	3.50	0.5
5	frisbee	jaune	1.50	0.2
6	cerf-volant	rouge	5.75	1.2
7	cerf-volant	bleu	5.75	1.2
8	yoyo	vert	3.15	0.8
9	yoyo	blanc	4.25	0.4
10	cerf-volant	jaune	1.50	0.2

inventaire_magasins		
id_jouet	id_magasin	inventaire
5	1	34
5	3	12
6	1	5
6	2	10
6	4	24
9	1	50
9	2	2
12	2	18
12	4	28
9	4	11

infos_magasins

id_magasin	adresse	telephone	directeur
1	23 rue des érables	0112233445	Jean
2	1902 allée de l'ambre	0198877665	Isabelle
3	100 boulevard du nord	0223345678	Liane
4	17 route moulin de pont	0455443322	Grégoire

3^{ème} Forme Normale de Boyce-Codd – BCNF

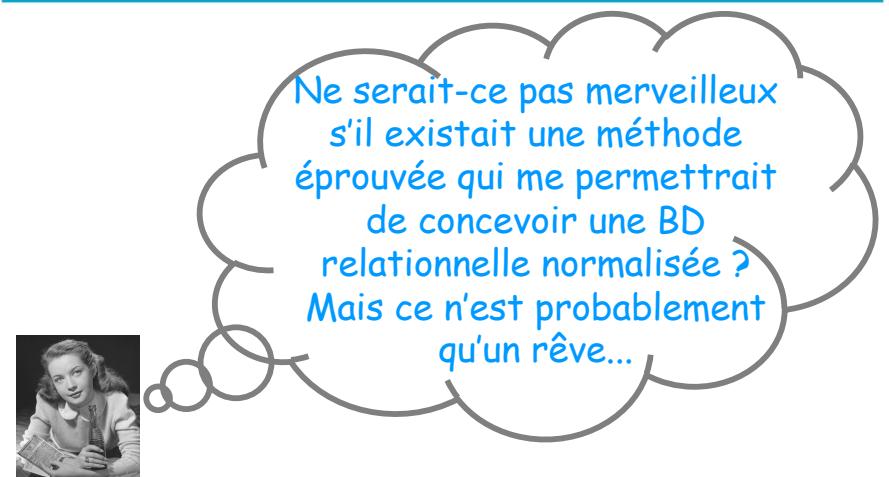
- BCNF : être en 3FN + un attribut non clé ne détermine pas 1 **partie** de la clé.
- Pas en BCFN :
 - Vins (Cru, Pays, Région)
 - Cette relation est bien en 3FN car aucun attribut non clé ne dépend d'une partie de la clé ou d'un attribut non clé et pourtant elle présente des redondances.
- En BCFN :
 - Vins_Cru (Cru, Région), Région_Pays (Région, Pays)

Cru	Pays	Région
Chenas	France	Beaujolais
Juliénas	France	Beaujolais
Morgon	France	Beaujolais
Brouilly	France	Beaujolais
Chablis	Etats-Unis	Californie

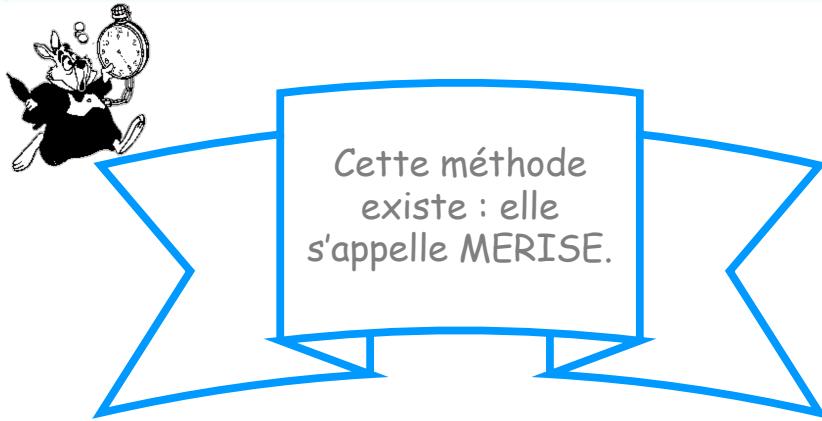
3^{ème} Forme Normale – 3FN

- 3° FN : être en 2FN + tout attribut dépend d'une clé par 1 DF élémentaire **directe** : pas de transitivité.
- Pas en 3° FN :
 - Voiture (NV, marque, type, puissance, couleur)
- En 3° FN :
 - Voiture (NV, type, couleur)
 - Modèle (type, marque, puissance)

I have a dream today (1)



I have a dream today (2)



Exemple : Top Chef

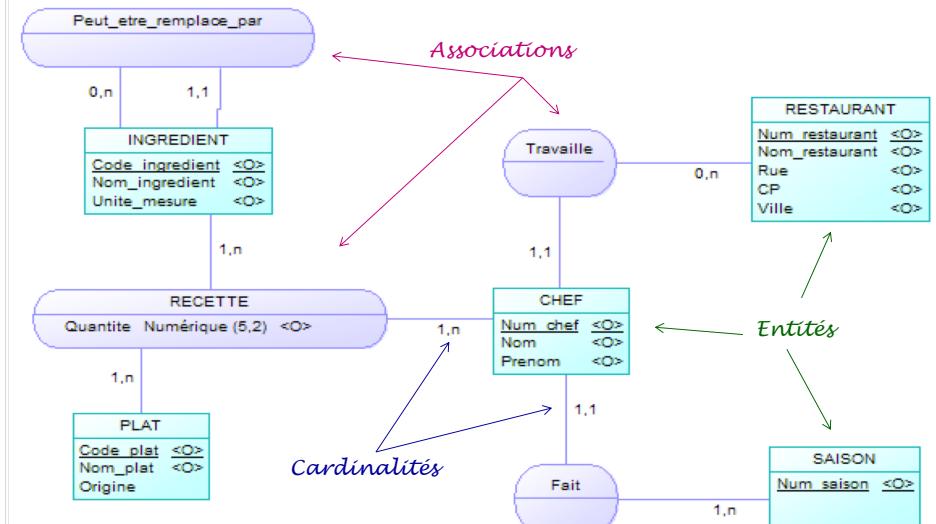
- Suite à la demande de nombreux spectateurs de l'émission Top Chef d'avoir accès aux recettes réalisées par les candidats, vous devez concevoir la BD nécessaire.
- Pour un plat imposé (paella valencienne, bortsch ukrainien, etc.), les chefs vont réaliser des recettes différentes pouvant utiliser des ingrédients différents en quantités différentes (150g, 3 cuillérées, etc.).
- Certains ingrédients étant parfois difficiles à trouver pour le cuisinier amateur qui refera la recette chez lui, il est possible de proposer un ingrédient de remplacement.
- Un chef appartient à un restaurant (dont on connaît le nom et l'adresse) et participe à une saison numérotée de 1 à 7 (car il y a eu 7 saisons !).



1^{ère} étape : établissez le MCD

- **Le Modèle Conceptuel des Données :**
 - Est un schéma qui représente la **structure** du Système d'Information,
 - C'est-à-dire les dépendances (relations) entre les différentes données du SI.
- Il s'intéresse aux **données** du SI sans préfigurer des **traitements** qu'elles vont subir.
- Il peut être traduit ensuite en Modèle Logique de Données Relationnel, mais pas exclusivement.

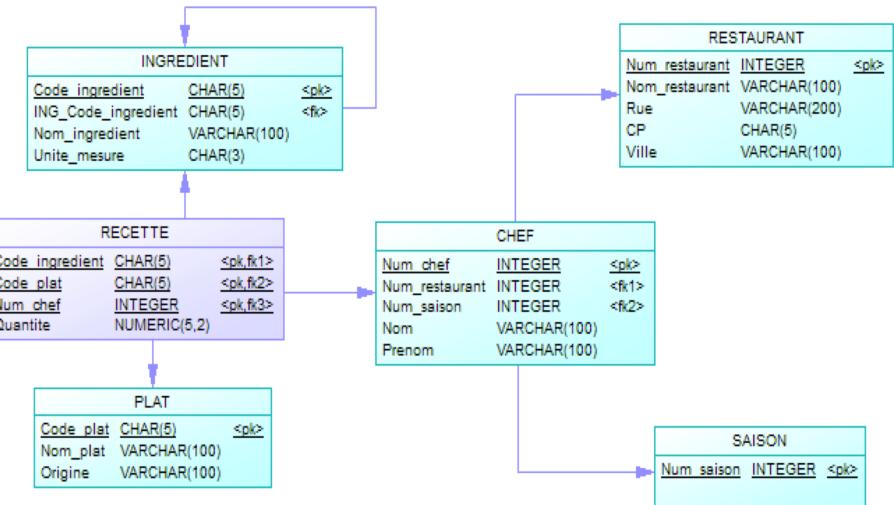
MCD Top Chef



2^{ème} étape : établissez le MLD Relationnel

- **Le Modèle Logique des Données Relationnel :**
 - Est une traduction (dérivation) du MCD dans un formalisme adapté à une implémentation ultérieure, au niveau physique, sous forme de base de données relationnelle.
- La transcription d'un MCD en modèle relationnel s'effectue selon quelques règles simples qui consistent à **transformer les entités et associations en relations** en fonction des **cardinalités** des associations.

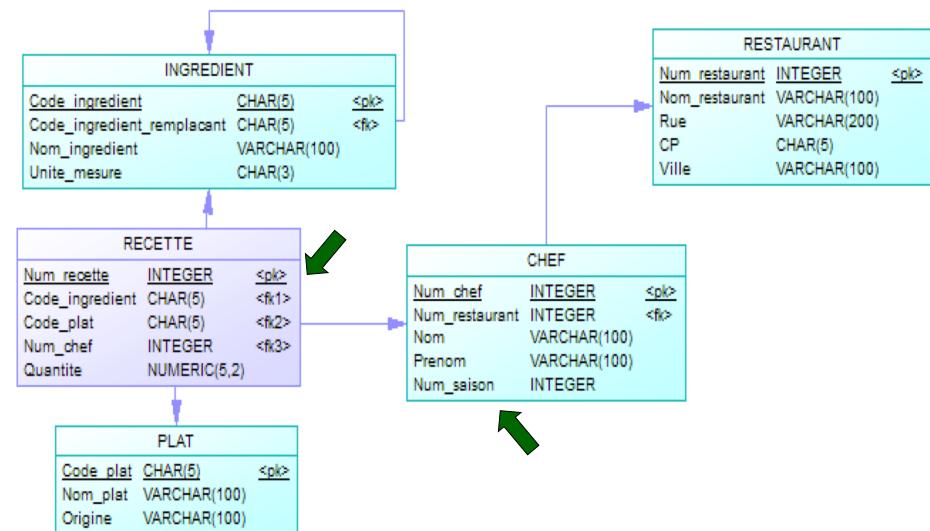
MLD Relationnel Top Chef



3^{ème} étape : établissez le MPD optimisé

- **Le Modèle Physique des Données** = optimisation du MLD (relations deviennent des tables) qui s'intéresse :
 - Aux **traitements** subis par les données en vue d'accélérer leur temps d'accès,
 - Et au mode de **stockage** des tables.
- On peut dans un souci d'optimisation :
 - Créer des **index** pour accélérer les recherches.
 - Remplacer de longues clés primaires (composées) par un **Id auto-incrémenté**.
 - Supprimer tables de paramètres et ajouter **contraintes** de vérification de valeur à la création des tables.
 - etc.

MPD optimisé Top Chef



Kit de survie



© CPE Lyon - Françoise PERRIN - 2018-2019

44

Comment élaborer le MCD ?



- Listez les propriétés et leur domaine : [dictionnaire des données](#).
- Identifiez les [entités](#) et leur [clé](#).
- Identifiez les liens entre entités : [associations](#) et contraintes d'intégrité fonctionnelles ([CIF](#))
- Rattachez les propriétés aux entités en trouvant les dépendances fonctionnelles élémentaires directes par rapport aux clés : [3° BCNF](#). Une propriété n'apparaît [que 1](#) seule fois dans le MCD.
- Les propriétés qui dépendent de plusieurs clés sont portées par une [association](#) entre les entités concernées.
- Déterminez les [cardinalités](#) des associations.

© CPE Lyon - Françoise PERRIN - 2018-2019

46

Quelle démarche appliquer pour définir une BD



- Établissez le Modèle Conceptuel des Données (MCD) – Modèle Entité / Associations.
- Transformez le MCD en Modèle Logique de Données Relationnel (MLD).
- Optimisez le MLD en fonction des traitements pour constituer le modèle Physique de Données (MPD).
- Définissez les tables (identifier les contraintes d'intégrité), vues, index, procédures cataloguées et déclencheurs, rôles – utilisateurs et priviléges, etc.
- Définissez l'infrastructure de stockage et les procédures d'administration.

© CPE Lyon - Françoise PERRIN - 2018-2019

45

Quels sont les critères de normalisation ?



- 1° FN : tout attribut est [atomique](#) : pas de notion de liste, ni de colonnes dupliquées.
- 2° FN : être en 1FN + tout attribut dépend d'une clé par 1 [DF élémentaire](#) : un attribut non clé ne dépend pas que d'1 partie de la clé.
- 3° FN : être en 2FN + tout attribut dépend d'une clé par 1 DF élémentaire [directe](#) : pas de transitivité.
- BCNF : être en 3FN + un attribut non clé ne détermine pas 1 [partie](#) de la clé.

© CPE Lyon - Françoise PERRIN - 2018-2019

47

Passage MCD à MLD Relationnel



- Toute **entité** devient une **relation** (future table).
- Les **propriétés** deviennent les **attributs** de la relation.
- **L'identifiant** entité devient **clé primaire** de la relation.
- Les associations dont la cardinalité est **0-1(ou 1-1)** d'une part et **0-N(1-N)** d'autre part disparaissent et la clé du «côté» **0-N** devient **clé étrangère** de la relation «côté» **1-1** :
 - Client (N° client, raison sociale...)
 - Commande (N° commande, date commande, N° client)
- Autres types associations deviennent des relations dont les clés sont la **concaténation** des clés des entités, et les attributs les propriétés de l'association :
 - Notation (N° élève, Id matière, note)

48

© CPE Lyon - Françoise PERRIN - 2018-2019

Passage MLD à MPD optimisé



- **Clés primaires auto-incrémentées courte** pour éviter d'utiliser de gros index lors des jointures.
 - Ajouter une clause **UNIQUE** à la clé primaire naturelle.
 - Penser à clause **NOT NULL** sur les attributs qui la constituent.
- **Conservation ou non des tables de références** (valeurs peuvent alimenter une liste déroulante dans formulaire) :
 - Contrainte **CHECK** à définir sur la colonne si la liste des valeurs n'est pas susceptible d'être modifiée + select distinct pour alimenter liste déroulante.
 - Ou bien table de paramètres dont les valeurs sont accessibles rapidement et qui peuvent être modifiées, complétées, supprimées.
- **Renommage** éventuel des tables (noms commun et non verbes).

49

© CPE Lyon - Françoise PERRIN - 2018-2019



CONTACT

Domaine Scientifique de la Doua
43, bd du 11 Novembre 1918 - Bâtiment Hubert Curien
B.P. 2077 - 69616 Villeurbanne cedex - France

Tél. : (33) 04 72 43 17 00
Fax : (33) 04 72 43 16 84

www.cpe.fr

francoise.perrin@cpe.fr

membre de UNIVERSITÉ DE LYON

