

Licence d'informatique

Bases de Données 1

Partie 1 : Initiation aux Bases de Données

Florence Petit

25/01/2009



Horaires

- Cours :
 - Sylvain CHERRIER
- TD : 18h, soit 12 semaines x 1h30
début la même semaine que le cours (CM)
- TP : 24h, soit 12 semaines x 2h00
début la semaine suivant le cours (CM)



Contenu de l'enseignement "Bases De Données 1"

- Initiation aux Bases de Données (BD) relationnelles
 - Définition et vocabulaire des bases de données
 - Types de données
 - Dictionnaire de données
 - Utilisation d'un SGBD (ex : MySQL)
 - Langage d'interrogation : SQL
 - Méthodes de conception de BD relationnelles
 - Graphe des Dépendances fonctionnelles
 - Entité/association
- Approfondissement du langage HTML
 - XHTML
 - CSS
- Langages de script pour le web
 - script côté serveur : PHP
 - Interfaçage PHP & BD
 - script côté client : JavaScript





Base de données (BD) = DataBase (DB)

- Définition :
Ensemble d'informations, cohérent, structuré, mémorisé sur support informatique.
- Représentation d'un aspect du réel (fonctionnement d'une organisation, relations entre personnes, objets, évènements...).
- BD présentes dans tous les domaines professionnels (gestion de salariés, gestion de production, commerce, transports, santé...) et personnels (carnet d'adresses, agenda, recherche d'horaires...)
- BD : un des grands domaines de l'informatique
- BD accessible par un logiciel (SGBD).



Qualités d'une BD

- Données non redondantes
- Données cohérentes
- Consultable et modifiable aisément
- Fort degré d'indépendance entre données et matériel
- Fort degré d'indépendance entre données et programmes



SGBD : Système de Gestion de Bases de Données

DBMS : Data Base Management System

- Ensemble de services permettant de gérer des bases de données :
 - décrire et modifier la structure des données
 - manipuler les données : insérer, modifier, supprimer
 - consulter les données de façon simple (requête)
 - assurer sécurité, confidentialité (authentification), intégrité, cohérence
 - permettre des accès simultanés (multi-utilisateurs)
- Exemples :
 - MySQL (Open Source, pour tous systèmes)
 - PostgreSQL (Open Source, université de Berkeley)
 - Oracle
 - DB2 (IBM)
 - Access (Microsoft)
 - SQL Server (Microsoft)



MySQL



- SGBDR Open Source pour réseau
- Dernière parue : version 5
- Format MyISAM (data .MYD, index .MYI, structure .FRM)
- Systèmes Linux, Unix, Windows, Mac OS X...
- Interface d'administration la plus connue : phpMyAdmin
- Taille max. de table: 8 millions de téraoctets
- Souvent utilisé comme SGBD pour faire des sites web dynamiques
- Fait partie du LAMP (Linux, Apache, MySQL, PHP)



Environnements de travail

LAMP
Linux Apache MySQL PHP

Ou aussi
WAMP Server

EN 2008	UMLV	EasyPHP v. 2.0b1
		http://www.easyphp.org
OS	Linux	Windows
Serveur web	Apache	Apache v. 2.2
PHP	v. 5	v. 5.2
Serveur BD	sqletud.univ-mlv.fr	localhost
MySQL	v. 5	v. 5
Base de données	monlogin_db	plusieurs
PhpMyAdmin	2.9.1.1	2.9.1.1
Autre SGBD	PostgreSQL avec phpPgAdminv.	



Comment accéder à l'interface de MySQL à UMLV ?

- Accès à l'activation et gestion de la BD MySQL
- <http://www.univ-mlv.fr>
puis à gauche ESPACE ETUDIANTS puis Services informatiques

The screenshot shows the UMLV website interface. On the left, the 'ESPACE ETUDIANTS' menu is expanded, with 'Services informatiques : pages web et interventions' highlighted. In the center, the 'connexion' section contains a link to 'Activation et administration de sa base de données MySQL', which leads to 'Activer sa base de données MySQL' and 'Administrer sa base de données MySQL'. On the right, the 'services informatiques avec identification' section contains links for 'se connecter à sa webmail', 'Activer et administrer sa base de données MySQL', 'Activer et administrer sa base de données PostgreSQL', 'Hébergement des pages web', and 'Respectez la charte informatique!'. A banner for 'Le LMD Les temps changent !!!' is visible at the bottom right.

- ou <http://cri.univ-mlv.fr>
puis Présentation générale
puis Base de données



MySQL à UMLV : activer sa base MySQL

- L'activation n'est à faire qu'une SEULE fois.
- Elle crée un répertoire utilisateur et met à jour la table utilisateur
- Elle initialise le mot de passe de la BD avec celui de l'étudiant
- Ne pas confondre MySQL et PostGreSQL
- <http://etudiant.univ-mlv.fr/activmysql.php>

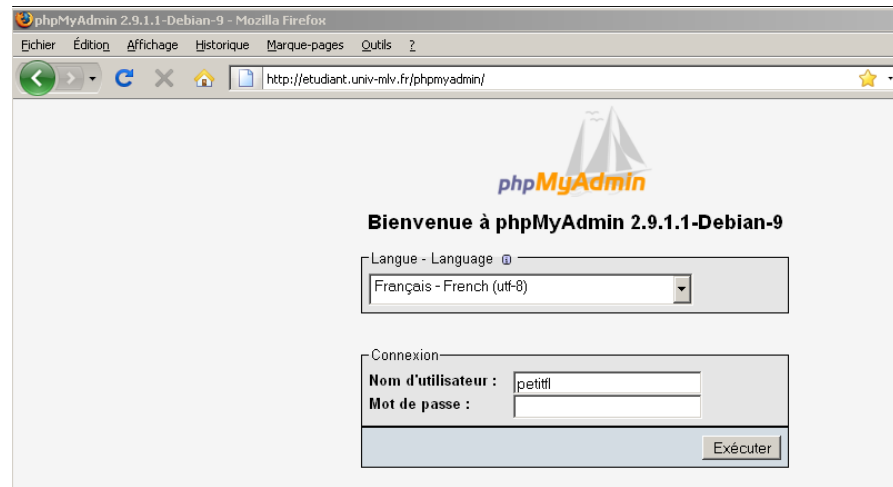
The screenshot shows a Mozilla Firefox browser window with the address bar displaying <http://etudiant.univ-mlv.fr/activmysql.php>. The page content is a form titled "Activer sa base de données MySQL:" with the following fields and controls:

- Login:
- Mot de passe:
- Encodage: LATIN1 UTF-8
- Buttons:



MySQL à UMLV : gérer sa base de données MySQL

- Ne pas confondre PHPMysqlAdmin et phpPgAdmin
- Entrée dans l'interface PHPMysqlAdmin avec login et mot de passe :
<http://etudiant.univ-mlv.fr/phpmyadmin/>





Utiliser MySQL à UMLV : accéder à sa base de données MySQL

The screenshot shows the phpMyAdmin interface in a web browser. The browser address bar displays `http://etudiant.univ-mlv.fr/phpmyadmin/`. The page title is **phpMyAdmin - 2.9.1.1-Debian-6**.

Left sidebar:

- phpMyAdmin logo
- Navigation icons: Home, Exit, SQL, Help, Logout
- Base de données: `petitfl_db (23)`
- Choisissez une base de données

Main content area:

sqlletud.univ-mlv.fr

- Version du serveur: 5.0.32-Debian_7etch1
 - Version du protocole: 10
- Serveur: sqlletud.univ-mlv.fr via TCP/IP
 - Utilisateur: petitfl@kapouer.univ-mlv.fr
- Jeu de caractères pour MySQL: **UTF-8 Unicode (utf8)**
- Interclassement pour la connexion MySQL: `utf8_unicode_ci`
- [Afficher l'état du serveur](#)
- [Afficher les variables du serveur](#)
- [Processus](#)
- [Jeu de caractères et interclassement](#)
- [Moteurs de stockage](#)
- [Bases de données](#)
- [Exporter](#)
- [Importer](#)
- [Quitter](#)

phpMyAdmin - 2.9.1.1-Debian-6

- Version du client MySQL: 5.0.32
- Extensions PHP utilisées: mysql
- Langue - Language: Français - French
- Thème / Style: Original
- Taille du texte: 100%
- [Documentation de phpMyAdmin](#)
- [Site officiel de phpMyAdmin](#)
 - [\[ChangeLog\]](#) [\[CVS\]](#) [\[Lists\]](#)

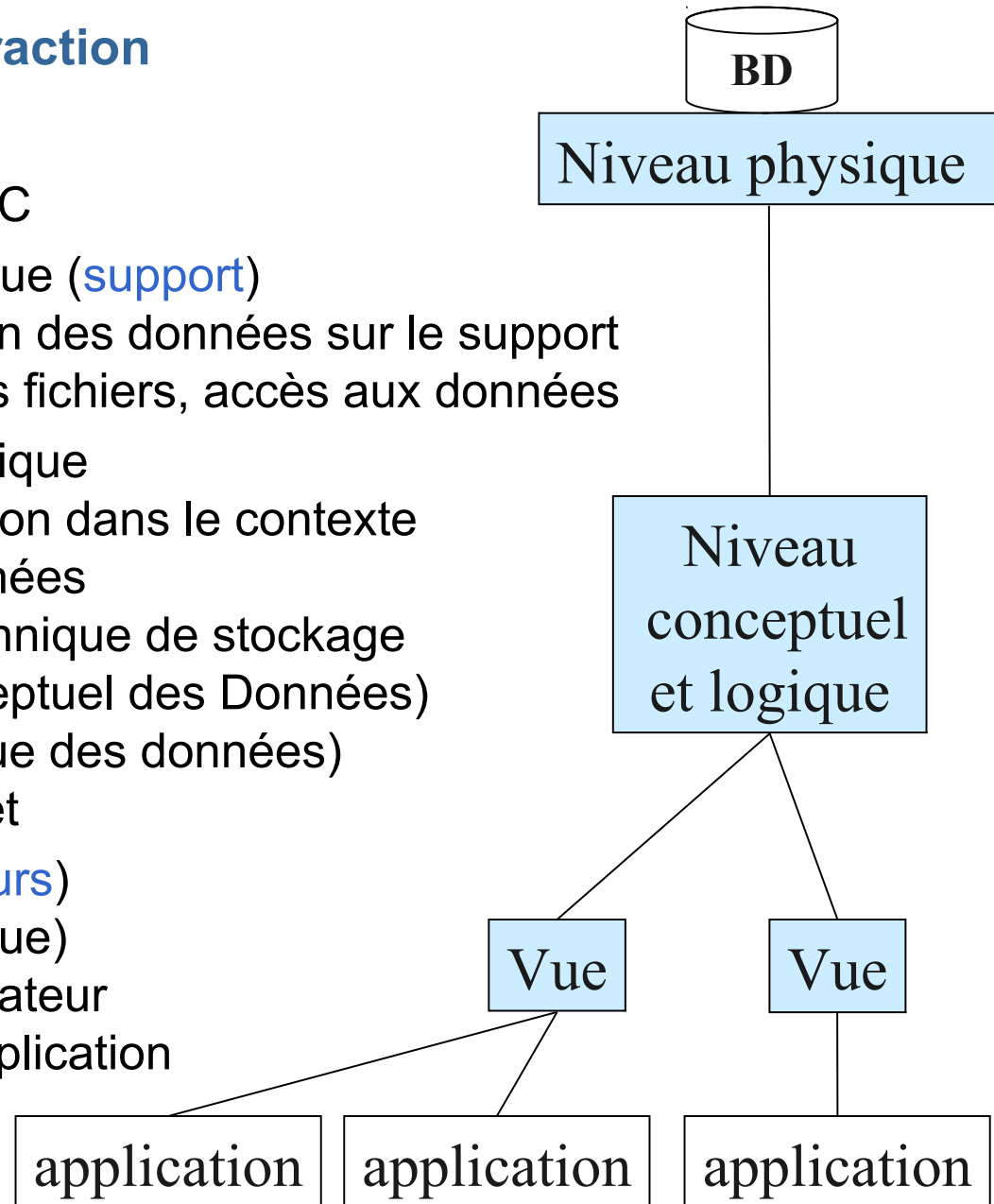
Footer:

[Ouvrir une nouvelle fenêtre phpMyAdmin](#)



SGBD : 3 niveaux d'abstraction

- Architecture ANSI/SPARC
- Niveau interne ou physique (**support**)
 - > structure d'implantation des données sur le support
 - > stockage/structure des fichiers, accès aux données
- Niveau conceptuel et logique
 - > structure de l'information dans le contexte
 - > modélisation des données
 - > indépendant de la technique de stockage
 - ex : MCD (Modèle Conceptuel des Données)
 - MLD (Modèle Logique des données)
 - Modèle Orienté Objet
- Niveau externe (**utilisateurs**)
 - > Schéma externe (ou vue) de chaque type d'utilisateur
 - > accessible par une application





Historique des modèles de BD

BD basées sur les modèles d'accès

- Modèle hiérarchique (1965)
- Modèle réseau (Codasyl 1971)

BD 2e génération

- **Modèle relationnel** (Codd 1970)

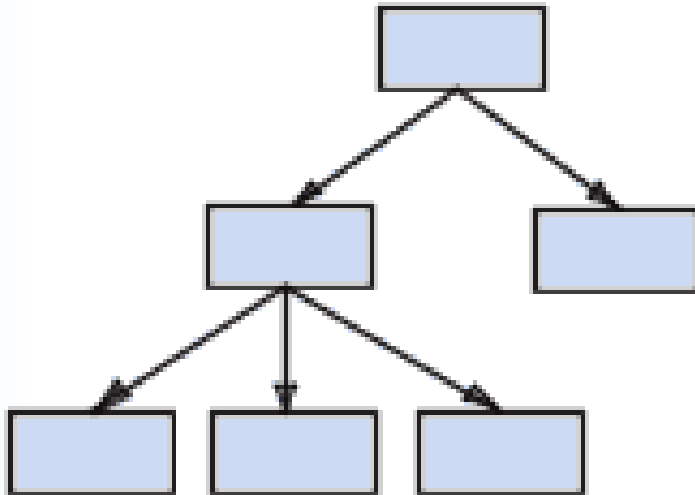
BD 3e génération

- Modèle déductif
- Modèle objet (OMG 1991)
- Multidimensionnelles
- Semi-structurée (XML...)



Modèle hiérarchique

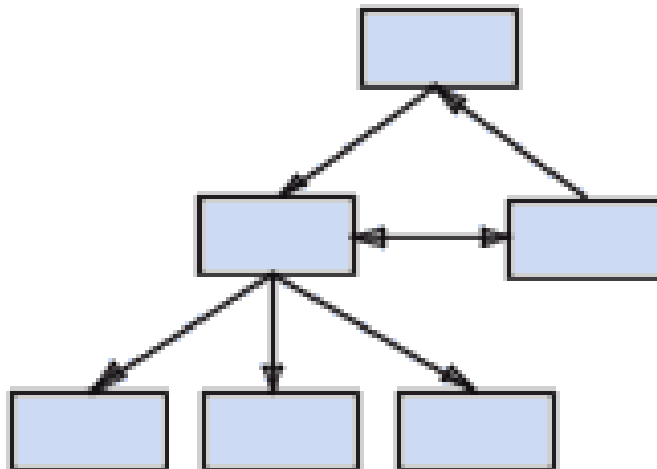
- Premier modèle de SGBD
- Données classées hiérarchiquement
- Arborescence descendante
- Pointeurs entre les différents enregistrements





Modèle réseau

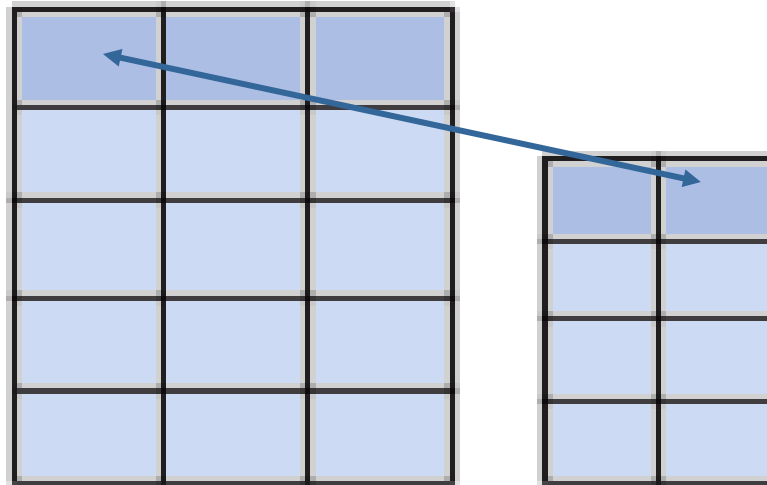
- Proche du modèle hiérarchique
- Pointeurs entre les différents enregistrements
- Structure pas forcément descendante, référence cyclique possible





Modèle relationnel

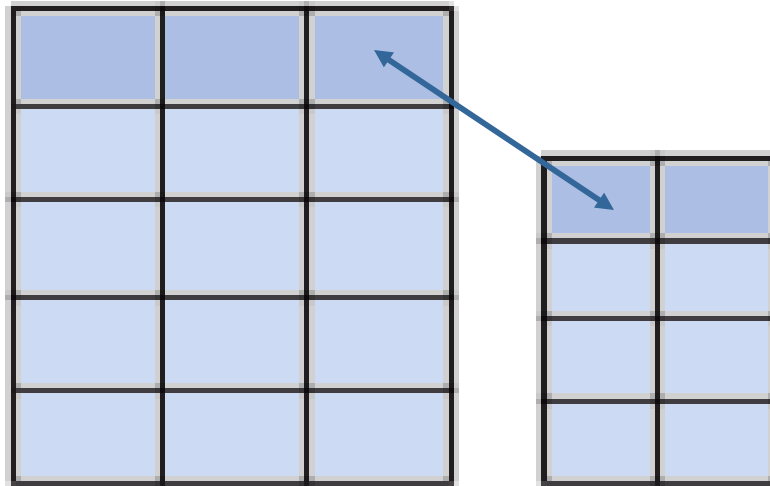
- Utilisé par les SGBDR (Système de Gestion de BD Relationnelles)
- Représentation par tableaux à 2 dimensions : lignes et colonnes
- Tableau = Table = Relation
- Données manipulées selon la théorie mathématique des relations.
- Une base de données comporte une ou plusieurs tables.
- Lien entre tables par concept de clé : une clé étrangère se réfère à une clé primaire





Modèle déductif

- Proche du modèle relationnel
- Données enregistrées dans des tableaux (ou tables) à deux dimensions (lignes et colonnes)
- Manipulation des données par calcul de prédicats





Modèle objet

- SGBDO : Système de gestion de BD objet
- Données stockées sous forme d'objets
- Structures appelées classes présentant des données membres.
- Les champs sont des instances de ces classes

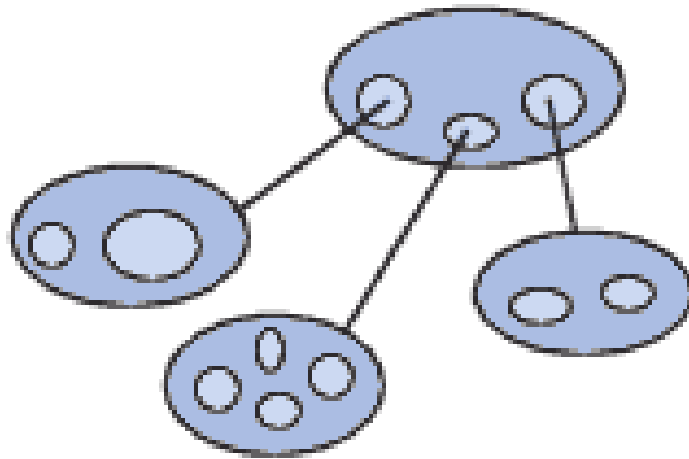




Table (ou relation)

Ensemble d'enregistrements structurés par champs (=colonnes).
 -> L'ordre des colonnes et des lignes de la table n'a pas importance.

Exemple : Table ETUDIANT

Colonne = Champ = Attribut = Propriété = Caractéristique

Ligne

=

Objet

=

N-uplet

=

Tuple

=

Enregistrement

=

Occurrence

nom	prenom	date_n
Dupont	René	
Durand	Paul	1990-12-31
Martin	Paul	1990-12-30
...

Donnée = valeur de l'attribut
 prise dans un domaine de valeurs

Domaine = ensemble de valeurs
 possibles d'un champ



Algèbre relationnelle : projection et restriction

Projection : sélection de colonnes

nom	prenom	date_n
Dupont	René	
Durand	Paul	1990-12-31
Martin	Paul	1990-12-30

Restriction (ou Sélection) : sélection de lignes

-> application pratique de l'algèbre relationnelle : le [SQL](#)



Schéma d'une relation

- Schéma en **extension**
 - état de la table, exprimé sous forme tabulaire (tableau)
 - comporte un exemple de données à un instant t

num	nom	prenom	date_n
00111	Dupont	René	
00112	Durand	Paul	1990-12-31
00113	Martin	Paul	1990-12-30

- Schéma en **intention** (ou compréhension)
 - exemple : Etudiant (num, nom, prenom, date_n)
 - exprimé de manière générale sous la forme :
nomTable (nomChamp1, nomChamp2, nomChamp3,...)



Schéma relationnel d'une base de données

- Ensemble des schémas des relations de la base de données
- Exemple très simplifié de gestion de garage

CLIENTS (numC, nomC, prenomC, adresseC)

SALARIE (numS, nomS, prenomS, adresseS, dateNaissS)

VOITURE (numV, marqueV, modeleV, numC)

REPARATION (numR, dateArriveeR, numV)

REPARER (numS, numR, nbHeuresTravail)



Clé primaire

- Identificateur d'un enregistrement de la table
-> permet de différencier 2 enregistrements d'une table
- Naturel ou artificiel (code, numéro incrémenté automatiquement)
- Choix parfois entre plusieurs clés (clés candidates)
- Clé simple (un seul attribut)
ou clé composée/multiple/concaténée (plusieurs attributs)
- doit être stable dans le temps
- souvent représenté en soulignant le ou les attributs
- contrainte d'unicité (ou contrainte d'entité) quand un attribut doit avoir une valeur unique dans la table.

num	NIR	Login	nom	prenom	date_n
00001	1630706054781	rdupont	Dupont	René	
00002	1801277013111	pdurand	Durand	Paul	1990-12-31
00003	1801275012141	pmarti01	Martin	Paul	1990-12-30
00004	1801275012142	pmarti02	Martin	Pierre	1990-12-30



Aspect juridique : NIR (N°Sécurité sociale)

- **Numéro d'Inscription au Répertoire national d'identification**
- Site de la CNIL : www.cnil.fr
- " *L'enregistrement du numéro de sécurité sociale dans les fichiers de paie et de gestion du personnel n'est autorisé que pour :
- l'établissement des bulletins de paie
et des différentes déclarations sociales obligatoires
- la tenue des comptes d'épargne salariale "*
- " *Le numéro de sécurité sociale d'un employé ne peut donc pas être utilisé comme numéro de matricule unique pour l'identifier dans tous les fichiers de gestion des ressources humaines de son entreprise ou de son administration "*.



Phases de conception et réalisation d'une base de données

- Conception (structure de la base de données)
 - Établir la liste des champs et leur type (dictionnaire de données)
 - Définir les tables et les liens entre tables (clés)
- Implémentation
 - Créer la base de données à l'aide du SGBD
 - Créer les tables (structure de colonnes, vides au début) à partir du dictionnaire de données
- Test (jeu d'essai)
 - Insérer les enregistrements (ensemble de données)
 - Modifier les enregistrements, en supprimer
 - Interroger la base de données (sélectionner, afficher, imprimer...)



Dictionnaire des données (liste des champs)

abréviation ou mnémonique	Libellé	Type de données	Remarques : longueur, format, contraintes, valeur par défaut, exemple, calcul...
numEtu	numéro de l'étudiant	nombre entier	clé primaire incrémentation auto
nomEtu	nom de l'étudiant	chaîne variable	35 max
prenomEtu	1er prénom de l'étudiant	chaîne variable	30 max
date_n	date de naissance	date	AAAA-MM-JJ
Date_ins	date d'inscription	date	Date d'enregistrement
note_math	moyenne annuelle de math	nombre décimal	12.5 (de 0 à 20)

Eviter les mélanges comme ici -> choisir un système de nommage et un seul pour un projet :
soit nommage avec _
soit nommage de type Dromadaire



Conseils pour nommer bases, tables et champs

- Ex : nom_etud, nom_etudiant, nomE, nomEtud, nomEtudiant
- Utiliser les lettres de A à Z et les chiffres de 0 à 9 et _
- On peut utiliser la notation "Dromadaire" avec lettres maj/min
- Eviter les espaces, accents, cédilles, ponctuation... (migration)
- Préférer le trait de soulignement _ (underscore) au trait d'union - (pris pour un signe de soustraction)
- On peut préfixer/suffixer les noms des éléments d'une BD
 - base de données : BD_ (ex: BD_GEST_ZOO)
 - table : T_ (ex : T_ANIMAL)
- Eviter les mots-clés du SQL (sinon à utiliser entre guillemets, ce qui est fastidieux)

ACTION
DATE
ISOLATION
LOCAL
MATCH

MAX, MIN
MINUTE
MODULE
NATIONAL

USAGE
USER
SCHEMA
SECTION

SESSION
SITE
TYPE
ZONE



Les types de données

- Chaque champ est caractérisé par un type de données.
 - prénom : chaîne de n caractères maximum
 - prix : nombre décimal
 - ordre d'arrivée à une course : nombre entier
 - date de naissance : date
 - titre de livre : chaîne de n caractères maximum
- Le type de données doit permettre de stocker toutes les données du champ, quel que soit les enregistrements à venir.
 - prénom : Luc ; Anne ; Jean-Christophe ; ...
 - prix : 0 ; 1 ; 1.50 ; 99.99 ; ...
 - titre : PHP pour les nuls ; 1984 ; C++ ; ...



Type de données : ensemble de valeurs possibles du champ

■ Chaîne (caractères)

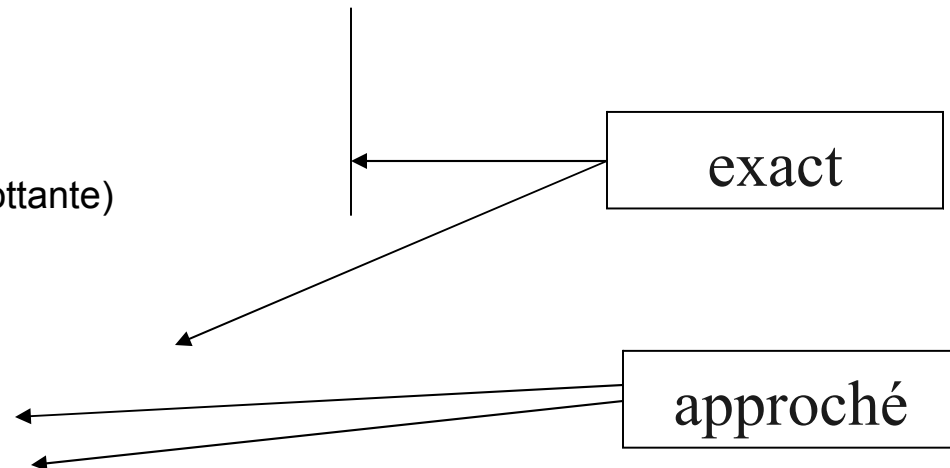
- CHAR (=CHARACTER)
- VARCHAR (=CHARACTER VARYING)
- Liste de valeurs : ENUM, SET
- TEXT : TINYTEXT, TEXT, MEDIUMTEXT, LONGTEXT
- BLOB : TINYBLOB, BLOB, MEDIUMBLOB, LONGBLOB

■ Temporel

- DATE, YEAR
- TIME, TIMESTAMP

■ Numérique

- **Entier** signé (+,-) ou non signé (+)
 - TINYINT, BIT, BOLLEAN, BOOL
 - SMALLINT
 - MEDIUMINT
 - INT (=INTEGER)
 - BIGINT
- **Réel** (Nombre à virgule flottante)
 - DEC (=DECIMAL)
 - FLOAT
 - DOUBLE (=REAL)





Données MySQL 4 : chaînes, énumération, ensemble

Chaînes fixes et variables				
Désignation	octets	Longueur	nombre de caractères	optionnel
CHAR(n)	n	fixe de n	$1 \leq n \leq 255$	BINARY
VARCHAR(n)	Lg+1	n max Lg variable	$1 \leq n \leq 255$	BINARY

Énumération et ensemble de valeurs				
Désignation	octets	Liste de valeurs	choix de valeurs	nombre max. de valeurs
ENUM	1 ou 2	'val1','val2'	une	65 635
SET	1,2,3,4 ou 8	'val1','val2'	plusieurs	64

- CHAR(n) : traitement plus rapide
- VARCHAR(n) : optimisation de la taille de la base de données
- Option BINARY : comparaisons et tris sensibles à la casse (maj, min)



Données MySQL 4 : Chaînes BLOB et TEXT

Blob et Text				
Désignation	octets	valeurs	longueur maximale de la chaîne	soit
TINYBLOB TINYTEXT	1	2^8	255	$(2^8) - 1$
BLOB TEXT	2	2^{16}	65 535	$(2^{16}) - 1$
MEDIUMBLOB MEDIUMTEXT	3	2^{24}	16 777 215	$(2^{24}) - 1$
LOB LONGTEXT	4	2^{32}	4 294 967 295	$(2^{32}) - 1$

- Stockage de volume important de données
- TEXT : objet texte (insensible à la casse)
- BLOB : Binary Large Object = objet binaire (image,....)
ressemble à un TEXT mais BINARY (sensible à la casse)



Données MySQL : date, temps

Désignation	octet	format	de	à
DATE	3	YYYY-MM-DD	0001-01-01	9999-12-31
YEAR(4)	1	YYYY	1901	2155
YEAR(2)	1	YY	70 (1970)	69 (2069)
TIME	3	HHH:MM:SS	-838:59:59	838:59:59
DATETIME	8	YYYY-MM-DD HH:MM:SS	0001-01-01 00:00:00	9999-12-31 23:59:59
TIMESTAMP	4	YYYY-MM-DD HH:MM:SS	1970-01-01 00:00:00	2037-12-31 23:59:59
TIMESTAMP(14)		YYYYMMDDHHMMSS	19700101000000	20371231235959
TIMESTAMP(12)	4	YYYYMMDDHHMM	1970-01-01 00:00	2037-12-31 23:59
TIMESTAMP(10)	4	YYYYMMDDHH	1970-01-01 00	2037-12-31 23
TIMESTAMP(8)	4	YYYYMMDD	1970-01-01	2037-12-31
TIMESTAMP(6)	4	YYMMDD	70-01-01	37-12-31
TIMESTAMP(4)	4	YYMM	70-01	37-12
TIMESTAMP(2)	4	YY	70	37

v. ≥ 4.1

Les données temporelles affichent des zéros 0 si la valeur n'est pas valide (hors intervalle, mois > 12, jour > 31)

- Formats de saisie acceptés : chaîne avec délimiteurs divers, chaîne ou nombre sans délimiteurs sur un nombre de position valide (6, 8, 10, 12, 14) :
'YYYYMMDDHHMMSS', 'YYYY-MM-DD HH:MM:SS', 'YYYY-MM-DD-HH-MM-SS', 'YYYY/MM/DD HH-MM-SS',
'YY/MM/DD/HH/MM/SS', 'YY-M-D', YYYYMMJJ, YYMMJJ,
- TIMESTAMP prend la date/heure courante (pour transaction).



Données MySQL : entiers

Désignation	octets	valeurs	précision	de	à	soit
TINYINT	1	2 ⁸	signed	-128	127	(2 ⁷)-1
			unsigned	0	255	(2 ⁸)-1
SMALLINT	2	2 ¹⁶	signed	-32 768	32 767	(2 ¹⁵)-1
			unsigned	0	65 535	(2 ¹⁶)-1
MEDIUMINT	3	2 ²⁴	signed	-8 388 608	8 388 607	(2 ²³)-1
			unsigned	0	16 777 215	(2 ²⁴)-1
INT ou INTEGER	4	2 ³²	signed	-2 147 483 648	2 147 483 647	(2 ³¹)-1
			unsigned	0	4 294 967 295	(2 ³²)-1
BIGINT	8	2 ⁶⁴	signed	-9 223 372 036 854 775 808	9 223 372 036 854 775 807	(2 ⁶³)-1
			unsigned	0	18 446 744 073 709 551 615	(2 ⁶⁴)-1

■ **UNSIGNED** : option à préciser. Par défaut, les nombres sont signés (signed)

- **ZEROFILL** : ajoute les zéros à gauche pour unsigned (positif)
- Pas de booléen : utiliser un CHAR(1) ou TINYINT(1) ou ENUM



Données MySQL : réel (virgule flottante)



Désignation	octets	signe	de	à
FLOAT	4	signed unsigned	-3.402823466E+38 0	+3.402823466E+38 +3.402823466E+38
REAL ou DOUBLE	8	signed unsigned	-1.7976931348623157E+308 0	1.797693134862357E+308 1.797693134862357E+308
DECIMAL(M,D) ou DEC(M,D)	M + 1 ou 2	signed unsigned	-1.7976931348623157E+308 0	-1.797693134862357E+308 1.797693134862357E+308

- **ZEROFILL** : option pour mettre des zéro non significatifs à gauche.
- **DECIMAL** : enregistré comme une chaîne ; adapté aux données monétaires.
- **M** : nombre total de chiffres (avant et après la virgule)
- **D** : nombre de décimales (chiffres après la virgule)



Types de données : exemples et problèmes dans MySQL 4

Type de données	Valeur entrée		Valeur stockée
DATE	2004-12-00	=>	2004-01-00 (date valide, jour inconnu)
DATE	2004-13-31	=>	0000-00-00 (date invalide, mois 0 à 12)
DATE	2004-12-32	=>	0000-00-00 (date invalide, jour 0 à 31)
DATE	2004-02-31	=>	2004-02-31 (date invalide mais acceptée)
YEAR(4)	1020	=>	0000 (hors limite 1901-2155)
TIME	-120:62:01	=>	00:00:00 (0 < min/sec < 59)
TIME	900:00:00	=>	838:59:59 (le plus grand)
TIMESTAMP(8)	19501231	=>	20051231 (année courante, car < 1970-01-01)
INT(5)	12.567	=>	12 (troncature)
TINYINT(1)	1234.567	=>	255 (le plus grand)
DOUBLE(3,1)	12.59	=>	12.6 (arrondi)
FLOAT(3,1)	12.59	=>	12.6 (arrondi)
DECIMAL(3,1)	12.59	=>	12.5 (troncature)
DECIMAL(3,1)	12,59	=>	12.0 (virgule / point) idem FLOAT/DOUBLE
DECIMAL(3,1) sig	4312.59	=>	999.9 (le plus grand)
DECIMAL(3,1) sig	-4312.59	=>	-99.9 (le plus grand)