

Bases de données

Le modèle relationnel

Bourrez les gens de données incombustibles, gorguez-les de « faits », qu'ils se sentent gavés, mais absolument « brillants » côté information. Ils auront l'impression de penser, ils auront le sentiment du mouvement tout en faisant du sur-place.

Ray Bradbury¹

Thème : Bases de données

Contenus	Capacités attendues	Commentaires
Modèle relationnel : relation, attribut, domaine, clef primaire, clef étrangère, schéma relationnel.	Identifier les concepts définissant le modèle relationnel.	Ces concepts permettent d'exprimer les contraintes d'intégrité (domaine, relation et référence).
Base de données relationnelle.	Savoir distinguer la structure d'une base de données de son contenu. Repérer des anomalies dans le schéma d'une base de données.	La structure est un ensemble de schémas relationnels qui respecte les contraintes du modèle relationnel. Les anomalies peuvent être des redondances de données ou des anomalies d'insertion, de suppression, de mise à jour. On privilégie la manipulation de données nombreuses et réalistes.

I. Un peu d'histoire

Le modèle relationnel est une manière de modéliser les relations existantes entre plusieurs informations, et de les ordonner entre elles. Cette modélisation qui repose sur des principes mathématiques mis en avant par E.F. Codd (laboratoire de recherche d'IBM) est souvent implémentée dans une base de données.

1970	Edgar Franck Codd pose les bases du modèle relationnel
1974	Création du langage SQL
1979	Création du premier système de gestion de base de données Oracle
1980	Le volume mondial de données stockées est estimé à 10^{18} octets
1990	Le volume mondial de données stockées est estimé à 10^{19} octets
1995	Première version du langage MySQL
2002	Le volume mondial de données stockées est estimé à 10^{20} octets
2010	Le volume mondial de données stockées est estimé à 10^{21} octets
2014	Le volume mondial de données stockées est estimé à 10^{22} octets

1. Fahrenheit 451 - <https://www.franceculture.fr/emissions/fictions-theatre-et-cie/fahrenheit-451-de-ray-bradbury>

II. Les limites des outils traditionnels

II. 1. L'exemple d'une médiathèque

Prenons l'exemple d'une médiathèque. Elle souhaite recenser les ouvrages qu'elle possède. Voici le tableau qu'elle a réalisé.

titre	nom auteur	prenom auteur	editeur	langue	annee pu- blication	reference
1984	Orwell	George	J'ai Lu	anglais	1949	120125
Dune	Herbert	Frank	Pocket	anglais	1965	02154
Fondation	Asimov	Isaac	Pocket	anglais	1951	95321
Le meilleur des mondes	Huxley	Aldous	J'ai Lu	anglais	1931	1254
Fahrenheit 451	Bradbury	Ray	J'ai Lu	anglais	1953	2365
Ubik	K. Dick	Philip	Denoël	anglais	1969	4523669
Chroniques martiennes	Bradbury	Ray	Denoël	anglais	1950	24568
La nuit des temps	Barjavel	René	Folio SF	français	1968	724867
Blade runner	K Dick	Philip	J'ai Lu	anglais	1968	24967
Les robots	Asimov	Isaac	Folio SF	anglais	1950	3588
La planète des singes	Boulle	Pierre	Folio SF	français	1963	89547
Ravage	Barjavel	René	Folio SF	français	1943	96541
Le maître du haut chateau	K.Dick	Philip	Denoël	anglais	1962	34587
Le monde des A	Van Vogt	Alfred Elton	Folio SF	anglais	1945	26974
La fin de l'éternité	Asimov	Isaac	Folio SF	anglais	1955	885236

Activité 1.

- ▷ Identifiez les redondances dans ce tableau, et les éléments qui semblent uniques.
- ▷ Le livre de Ray Bradbury intitulé « Chroniques martiennes » est régulièrement indisponible car il est trop souvent emprunté. La médiathèque achète donc un deuxième exemplaire, paru chez le même éditeur. Il portera le numéro d'inventaire 236984. Complétez le tableau avec ce nouvel ouvrage.
- ▷ Existe-t-il un élément unique qui va référencer totalement l'ouvrage ?
- ▷ La médiathèque s'ouvre depuis peu à la bande dessinée. Elle souhaite enregistrer l'album de bande dessinée en langue française **L'incal noir**, scénarisé par **Alejandro Jodorowsky**, dessiné par **Moebius** et édité par **Les humanoïdes associés** en **1982**. Quel est (sont) le(s) problème(s) rencontré(s) ?

II. 2. Notion de base de données

Une donnée est une information représentée sous une forme conventionnelle, afin de pouvoir être traitée automatiquement. Une base de données (BDD) représente un ensemble ordonné de données dont l'organisation est régie par un modèle.

Les données sont généralement regroupées selon leur appartenance à un objet du monde réel.

- ▷ Le regroupement d'objets homogènes constitue une **entité**.
- ▷ Une entité est décrite par un ensemble d'**attributs**.
- ▷ Chacun de ces attributs prend une **valeur** pour chaque objet.

Exemple 1

- Entité : un être humain
- Attributs : nom, prénom, date de naissance, lieu de naissance
- Valeurs : Young, Neil, 12 novembre 1945, Toronto

Activité 2.

On souhaite représenter les données issues de la carte nationale d'identité. À partir de votre carte personnelle, donnez le nom de l'entité correspondante, les attributs qui lui sont associés et les valeurs qui vous sont propres.

III. Le modèle relationnel

III. 1. Les principes de base

Un des modèles de données le plus courant est le modèle relationnel. Les principes de base de ce modèle sont les suivants :

- ▷ séparer les données dans plusieurs tables

Activité 3.

1. Établissez, pour chaque entité, la liste des attributs.
2. Représentez chaque entité par une table contenant les attributs et la liste de leurs valeurs.
3. Pensez-vous qu'il y ait des redondances ? S'il y en a, modifiez les tables pour les supprimer.
4. En réalité, la table **Produits** n'est pas pertinente. Elle ne devrait contenir que les informations associées à un produit. Quels seraient ses attributs ?
5. Construisez une quatrième table intitulé **Details** reprenant les attributs manquants.
6. Comment éditera-t-on le bon de commande ?

III. 3. Définition

Déf.1

Une **relation** (on parle aussi de table) est composée d'un en-tête (le libellé des **attributs**) et d'un corps composé d'un ou plusieurs **t-uplets** (on parle aussi d'enregistrement).

id	titre	auteur	ann_publi
1	1984	Orwell	1949
2	Dune	Herbert	1965
3	Fondation	Azimov	1951
4	Le meilleur des mondes	Huxley	1931
5	Fahrenheit 451	Bradbury	1953
6	Ubik	K.Dick	1969
7	Chroniques martiennes	Bradbury	1950
8	La nuit des temps	Barjavel	1968
9	Blade Runner	K.Dick	1968
10	Les Robots	Azimov	1950
11	La Planète des singes	Boulle	1963
12	Ravage	Barjavel	1943
13	Le Maître du Haut Château	K.Dick	1962
14	La fin de l'éternité	Azimov	1955

en-tête

un t-uplet

attribut : titre

corps

FIGURE 2.3 – en-tête, attributs, et t-uplets de la relation **Livres**

Bonnes Pratiques

Même si la notation n'est pas normalisée, nous écrirons les tables en utilisant l'écriture CamelCase^a, c'est-à-dire sans espace (tous les mots sont attachés), et chaque mot commence par une majuscule.

On écrira **AnalyseEauxParis** plutôt que analyse eau Paris.

Les attributs sont l'écriture camelCase, c'est-à-dire sans espace (tous les mots sont attachés), et seul le premier mot commence par une minuscule.

On écrira **idAnalyse** plutôt que id_analyse.

a. https://fr.wikipedia.org/wiki/Camel_case

III. 4. Domaine de valeurs d'un attribut

Chaque valeur possède un type et un ensemble fini ou non des valeurs possibles. Cela constitue le **domaine de valeurs** d'un attribut.

Exemple 2

- L'attribut **idAuteur** de la relation **Auteurs** est un entier positif. Son type est **entier**. Le domaine de valeur correspond à **tous les entiers de 0 à $+\infty$** .
- L'attribut **titre** de la relation **Livres** est une chaîne de caractères. Son type est **chaîne de caractères**. Le domaine de valeur correspond à **toutes les chaînes de caractères**.

Activité 4.

Déterminez le domaine de valeurs des attributs de la relation **Produits**.

reference	libelle	prix
1900464K	Calcium chlorure 1 mol/L	77
31535.292	Sodium chlorure 1 mol/L (1N)	105
30024.290	Acide chlorhydrique 1 mol/L (1N)	41
30917.320	Iode 0,05 mol/L (0,1 N)	936

III. 5. Notion de clé primaire

À l'intérieur d'une relation, deux t-uplets identiques ne sont pas autorisés. Il faut pouvoir identifier de façon unique un t-uplet.

Déf.2

Une **clé primaire** permet d'identifier un t-uplet de **manière unique**.

Il faut déterminer, parmi les attributs, lequel permet d'identifier de manière unique un t-uplet. Cet attribut sera considéré comme la **clé primaire** de la relation.

Dans la figure 2.3 :

- ▷ L'attribut **auteur** ne peut pas jouer le rôle de clé primaire (deux ouvrages pouvant avoir le même auteur).
- ▷ De même pour les attributs **titre** et **annPubli**.
- ▷ Il reste donc l'attribut **idLivres** (pour identifiant), qui a été ajouté ici pour jouer le rôle de clé primaire.

Remarque

Ici, nous avons créé artificiellement une clé primaire, car aucun des autres attributs ne pouvait convenir (ce n'est pas toujours le cas).

III. 6. Clé étrangère

Dans la relation **Auteurs**, chaque auteur est identifié par l'attribut **idAuteur** (clé primaire de la relation).

Dans la relation **Livres**, on a rajouté un attribut **idAuteur** qui est la clé primaire de la relation **Auteurs**.

Déf.3

Une **clé étrangère** référence une clé primaire d'une autre table.

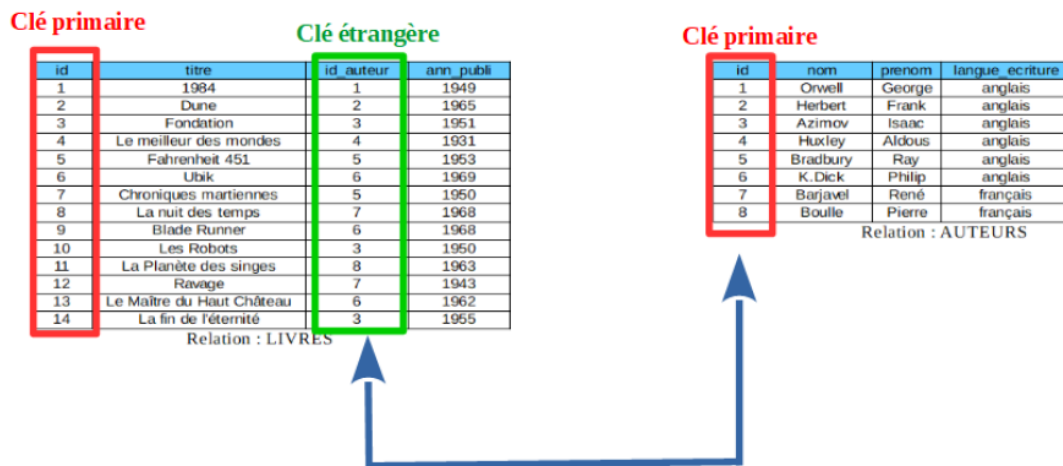


FIGURE 2.4 – La clé primaire de la relation **Auteurs** est aussi une clé étrangère de la relation **Livres**

L'attribut **idAuteur** est ce que l'on nomme une **clé étrangère** de la relation **Livres**, elle permet de faire le lien entre les deux relations.

Remarque

Il peut y avoir plusieurs clés étrangères dans une relation.

Activité 5.

1. Reprenez les relations de l'activité 3, en identifiant une clé primaire pour chacune des tables, ou en introduisant une si elle n'existe pas.
2. Justifiez que dans la relation **Details**, les attributs **reference** et **numero commande** ne peuvent pas constituer individuellement une clé primaire.
3. Justifiez que dans la relation **Details**, le couple (**reference**, **numero commande**) constitue une clé primaire.
4. Identifiez les clés étrangères.

III. 7. Représentation du modèle relationnel

Le schéma d'une relation définit cette relation. Il est composé :

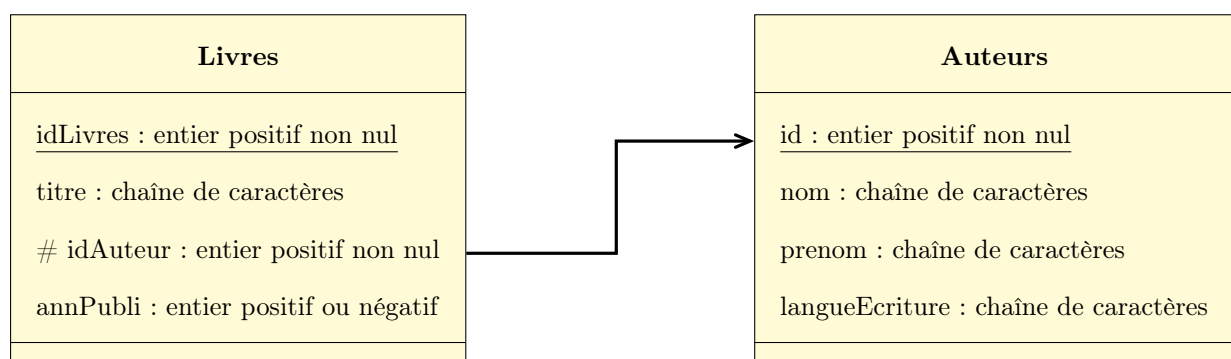
- ▷ du nom de la relation (on la place en tête, et en gras),
- ▷ de la liste de ses attributs avec les domaines respectifs dans lesquels ils prennent leurs valeurs (cette liste est placée entre parenthèses),
- ▷ de la clé primaire (elle est soulignée),
- ▷ des clés étrangères (on place un # entre la clé étrangère et la référence),
- ▷ des autres valeurs.

L'ensemble des relations peut être représenté soit par un schéma, soit par une notation textuelle.

Exemple 3 (notation textuelle pour la figure 2.4)

Livres (idLivres, titre, idAuteur # Auteurs(idAuteur), annPubli)
Auteurs (idAuteur, nom, prenom, langueEcriture)

Exemple 4 (schéma relationnel pour la figure 2.4)



La flèche représente le lien entre la clé étrangère et la clé primaire.

Activité 6.

Les relations de l'activité 3 sont représentées ci-dessous :

Produits

reference	libelle	prix
1900464K	Calcium chlorure 1 mol/L	77
31535.292	Sodium chlorure 1 mol/L (1N)	105
30024.290	Acide chlorhydrique 1 mol/L (1N)	41
30917.320	Iode 0,05 mol/L (0,1 N)	936

Client

numero client	nom	adresse	localite
BD2014	Le Coin	24 rue Saint Jean	Caen

Details

qte	reference	numero commande
10	190464K	14010
15	31535.292	14010
3	30024.290	14010
8	30917.320	14010

Commande

numero commande	# numero client	date
14010	BD2014	24/09/2014

Les clés primaires et étrangères ont été définies à l'activité 5.

1. Écrivez la notation textuelle du modèle relationnel.
2. Réalisez le schéma relationnel.

IV. Les contraintes d'intégrité

Il est important d'assurer la cohérence et donc l'intégrité des données présentes dans une base de données. Cela consiste à s'assurer que les données stockées sont cohérentes entre elles, c'est à dire qu'elles respectent toutes les règles exigées par le concepteur de la base de données. C'est une assertion vérifiée par les données de la base, à tout moment.

Déf.4

Une contrainte d'intégrité est une règle appliquée à un attribut ou une relation et qui doit toujours être vérifiée.

Les contraintes d'intégrité sont vérifiées par le système de gestion des bases de données (SGBD). Si l'une des règles n'est pas respectée, le SGBD signalera cette erreur et n'autorisera pas l'écriture de cette nouvelle donnée.

IV. 1. La contrainte de domaine

Déf.5

Chaque attribut doit prendre une valeur dans son domaine de valeurs.

Exemple 5

- La note obtenue dans une matière doit être comprise entre 0 et 20 ;
- la quantité commandée est obligatoire ET doit être strictement supérieure à 0.

IV. 2. La contrainte de relation

Déf. 6

Chaque relation dans le modèle relationnel est identifiée par une clé primaire qui doit être **unique** et **non nulle**. Donc, chaque t-uplet est également identifié par une clé primaire.

IV. 3. La contrainte de référence

Déf. 7

Une clé étrangère dans une relation doit être une clé primaire dans une autre. De plus, le domaine de valeurs de ces deux clés doit être identique. Enfin, la valeur d'une clé étrangère doit exister dans la clé primaire qui y fait référence.

Exemple 6

Dans la figure ci-dessous, la **valeur** de la clé étrangère **idAuteur** ne peut pas être supérieure à 8.

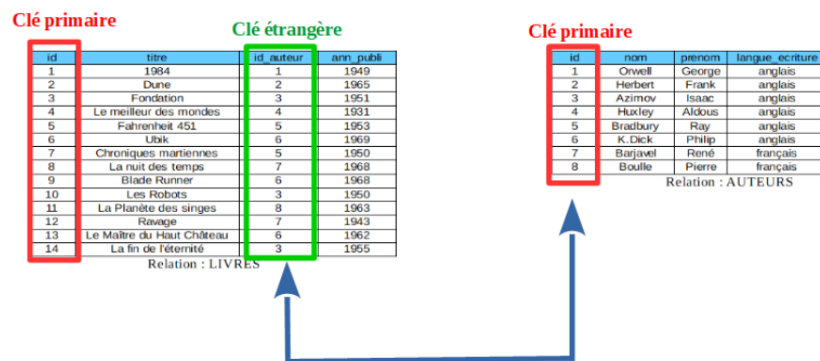


FIGURE 2.5 – contrainte de référence entre la relation Livres et la relation Auteurs

V. Références

- ▷ http://www.ecole.ensicaen.fr/~alebret/_data/teaching/bdd/snir/bdd.pdf
- ▷ https://isn-icn-ljm.pagesperso-orange.fr/T-NSI/co/Bases_de_donnees.html
- ▷ https://www.lecluse.fr/nsi/NSI_T/bdd/modrel/#quelques-lments-de-vocabulaire

VI. Exercices

Exercice 1.

Dans l'exemple de la Médiathèque II. 1., on souhaite créer 3 relations :

- ▷ **Auteur**, contenant toutes les informations sur un auteur
- ▷ **Livre**, contenant toutes les informations sur un livre
- ▷ **Ouvrage**, contenant toutes les informations sur un ouvrage disponible à l'emprunt.

1. Pour chacune des relations, indiquez les attributs avec leur type et leur domaine de valeurs.
2. Pour chacune des relations, indiquez la clé primaire et éventuellement la ou les clé(s) étrangère(s).
3. Représentez le schéma relationnel correspondant, ainsi que la notation textuelle qui lui est associée.

Exercice 2.

Un laboratoire souhaite gérer les médicaments qu'il conçoit :

- ▷ Un médicament est décrit par un nom, qui permet de l'identifier. En effet il n'existe pas deux médicaments avec le même nom.
- ▷ Un médicament comporte une description courte en français, ainsi qu'une description longue en latin.
- ▷ On gère aussi le conditionnement du médicament, c'est à dire le nombre de pilules par boîte (qui est un nombre entier).
- ▷ À chaque médicament on associe une liste de contre-indications, généralement plusieurs, parfois aucune.
- ▷ Une contre-indication comporte un code unique qui l'identifie, ainsi qu'une description.
- ▷ Une contre-indication est toujours associée à un et un seul médicament.

Voici deux exemples de données :

1. Le Chourix a pour description courte « Médicament contre la chute des choux » et pour description longue « Vivamus fermentum semper porta. Nunc diam velit, adipiscing ut tristique vitae, sagittis vel odio. Maecenas convallis ullamcorper ultricies. Curabitur ornare. ». Il est conditionné en boîte de 13.

Ses contre-indications sont :

- ▷ CI1 : Ne jamais prendre après minuit.
- ▷ CI2 : Ne jamais mettre en contact avec de l'eau.

2. Le Tropas a pour description courte « Médicament contre les dysfonctionnements intellectuels » et pour description longue « Suspendisse lectus leo, consectetur in tempor sit amet, placerat quis neque. Etiam luctus porttitor lorem, sed suscipit est rutrum non. ». Il est conditionné en boîte de 42.

Ses contre-indications sont :

- ▷ CI3 : Garder à l'abri de la lumière du soleil.

1. Donnez la représentation sous forme de tables des relations **Médicament** et **ContreIndication**.
2. Écrivez le schéma relationnel permettant de représenter une base de données pour ce laboratoire.

Exercice 3.

Proposez un domaine approprié pour les attributs représentant :

- ▷ Une adresse e-mail
- ▷ Une nationalité
- ▷ Un âge
- ▷ Une date
- ▷ Un mot de passe
- ▷ Un classement à un concours

Exercice 4.

1. La paire d'attributs (nom, prénom) constitue-t-elle toujours une clé primaire pour une relation de schéma **Eleve** (nom, prénom) ?
2. Comment peut-on étendre ce schéma relationnel pour mieux coller à la réalité ?
3. Comment assure-t-on que cette clé sera bien primaire ?

Exercice 5.

Le camping « Le Bougalou » souhaiterait informatiser la location de ses emplacements. Il a recensé dans 3 tables les informations nécessaires à la gestion des locations.

N°	Nom	Prénom	Adresse
1247	BERT	Rene	22, av. de Picardie – 69000 LYON
1248	SALI	Mehdi	14, rue des Armées – 13000 MARSEILLE
1249	GIRARD	Veronique	22, bd Pasteur – 75000 PARIS

FIGURE 2.6 – relation Client

Code	Date Début Loc	Date Fin Loc	N°	N° Emplacement
06JUI258	10/07/06	24/07/06	1249	147MH
06JUI259	10/07/06	17/07/06	146T	1247
06JUI260	11/07/06	31/07/06	1248	152T

FIGURE 2.7 – relation Contrat

N°	Surface	Px Location par Nuitée
147MH	45	70
148MH	32	50
151T	19	Entre 15 & 18
152MH	47	65

FIGURE 2.8 – relation Emplacement

1. Distinguez dans chaque table, les clés primaires et étrangères.
2. Relevez dans les tables les erreurs de saisie.
3. Représentez le schéma relationnel de cette base de données.

Exercice 6.

Voici ci-dessous 3 relations d'une base de données de l'Université Maurice Thorez.

NOTE		
NumE	NumM	Note
1	1	10
1	2	12
2	3	5
2	4	10
3	1	14
3	2	15
3	3	
4	1	13
4	3	18
5	2	4
6	6	
6	2	15
6	4	12

MATIERE		
NuM	Intitulé	Contenu
1	ULIN 305	BD et java
2	XUT 402	Algo
3	ULIN 205	Statistique
4	TUT 108	Web

ETUDIANT		
NumE	PrenomE	NomE
1	Servane	Coppalle
2	Stephane	Felici
2	Florence	Mendes
3	Esther	Lefebvre
4	Jordan	Houari
5	Christophe	Guerville

FIGURE 2.9 – tables **Note**, **Matiers**, **Etudiant** de l'Université Maurice Thorez

1. Combien existe t-il d'enregistrements dans la table **étudiant** ?
2. Combien d'attributs possède la table Etudiant ?
3. Quelles sont les clés primaires de ces 3 tables ?
4. Trouve-t-on des clés étrangères dans ces 3 tables ?
5. Quelles sont les informations complémentaires de la table **Matiere** ?
6. Trouve-t-on des clés multi-composants dans ces 3 tables ?
7. Un étudiant peut-il avoir plusieurs notes pour une matière ?
8. Quel attribut serait susceptible d'être un identifiant secondaire dans la table **Matiere** ?
9. Quelles colonnes pourraient être facultatives ?
10. Si la colonne Note est obligatoire dans la table **Note**, y a-t-il un problème ?
11. Vérifier les contraintes d'unicité (pour les clés primaires) et les contraintes référentielles (pour les clés étrangères). Constatez-vous des erreurs ?

Exercice 7.

Le Lycée Claude François est le seul lycée de France à héberger toutes les CPGE, avec toutes les filières possibles, à raison de 2 classes pour chaque filière et pour chaque niveau. La liste des filières (1^{re} et 2^e année est consultable ici : https://www.sainte-anne-brest.fr/cpge_filières_scientifiques_439_310-jpg/). Pour la prochaine rentrée, elle ouvrira en plus la filière MPI (Mathématiques Physique Informatique), mais uniquement en première année, en attendant l'ouverture de la deuxième année pour la rentrée suivante.

Au sein de cet établissement, on peut distinguer à minima les entités suivantes : les élèves, les enseignants, les classes. Un stagiaire informaticien est chargé de créer une relation pour chacune de ces entités. Il propose le schéma suivant :

1. Pour la relation **Eleves**, les attributs peuvent se limiter dans un premier temps à nom, prenom, date_naiss.
 - a. Pour chacun de ces attributs, identifiez le domaine de valeurs.
 - b. Représentez cette relation sous la forme d'une table contenant quelques enregistrements respectant le domaine de valeurs de chaque attribut.
2. Pour la relation **Classes**, les attributs peuvent se limiter à filiere, numero et salle (chaque classe se voit affecter une salle pour l'année scolaire).
 - a. Pour chacun de ces attributs, identifiez le domaine de valeurs.
 - b. Représentez cette relation sous la forme d'une table contenant quelques enregistrements respectant le domaine de valeurs de chaque attribut.
3. Pour la relation **Enseignants**, les attributs peuvent se limiter à nom, prenom, filiere et salle (chaque enseignant se voit affecter une classe et une salle pour l'année scolaire).
 - a. Représentez cette relation sous la forme d'une table contenant quelques enregistrements respectant le domaine de valeurs de chaque attribut.
4. Existe-t-il une clé primaire dans la relation **Eleves** ? Sinon, modifiez la relation pour en ajouter une.
5. Existe-t-il une clé primaire dans la relation **Enseignants** ? Sinon, modifiez la relation pour en ajouter une.
6. Montrez qu'il existe une clé primaire dans la relation **Classes**. Quelle est-elle ?
7. La contrainte de référence est-elle respectée entre les tables **Classes** et **Enseignants** ? Sinon, modifiez l'une des deux relations pour qu'elle soit respectée. Énoncez la référence entre la clé étrangère et la clé primaire correspondante.
8. La contrainte de référence est-elle respectée entre les tables **Classes** et **Eleves** ? Sinon, modifiez l'une des deux relations pour qu'elle soit respectée. Énoncez la référence entre la clé étrangère et la clé primaire correspondante.
9. Écrivez le schéma relationnel permettant de représenter une base de données.