

## SUJETS DE COLLES 08

Le sujet 2 est plus difficile que les autres.

### SUJET 1.

**Exercice 1 : chaîne de Markov.** Dans un certain pays, il ne fait jamais beau 2 jours de suite. Si un jour il fait beau, le lendemain, il peut neiger ou pleuvoir avec autant de chances. Si un jour il pleut ou il neige, il y a une chance sur deux qu'il y ait changement de temps le lendemain et s'il y a changement de temps, il y a une chance sur deux pour que ce soit pour du beau temps.

Au premier jour il fait beau.

On note  $p_n$  la probabilité qu'il fasse beau au jour  $n$ ,  $q_n$  la probabilité qu'il pleuve et  $r_n$  la probabilité qu'il neige.

1. On considère le vecteur  $X_n = (p_n, q_n, r_n)$ . Montrer que

$$X_{n+1} = X_n \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

2. En notant  $A$  la matrice de transition de la chaîne de Markov (la matrice précédente), montrer que 1,  $-1/4$  et  $1/4$  sont valeurs propres de la matrice.
3. En déduire la proportion de jours de beau temps, de pluie et de neige sur le long terme.
4. Refaire la question précédente dans le cas où la chaîne n'a que deux états : beau temps et mauvais temps.

**Exercice 2 : SQL.** Inspiré d'ECRICOME 2023 sujet zéro

Sur le marché des véhicules d'occasion, on observe en général une baisse du prix de revente (ou décote) d'un véhicule lorsque le nombre de kilomètres parcourus augmente. Une bonne estimation de cette baisse de prix permet au vendeur de fixer avec précision le prix de revente du véhicule.

On dispose d'une base de données comportant deux tables `vehicule` et `annonce` décrite ci-dessous.

- La table `vehicule` recense des informations sur les modèles de véhicules en vente sur le marché. Elle est composée des attributs suivants :
    - `id_vehicule` (de type `INTEGER`) : un code permettant d'identifier de manière unique chaque référence de véhicule (marque et modèle).
    - `marque` (de type `TEXT`) : le nom du constructeur du véhicule.
    - `modele` (de type `TEXT`) : le modèle du véhicule, un constructeur proposant en général plusieurs modèles de véhicule à la vente.
    - `prix_neuf` (de type `INTEGER`) : prix de vente du véhicule neuf.
  - La table `annonce` regroupe des informations sur un grand nombre d'annonces de véhicules d'occasion. Chaque enregistrement correspond à une annonce et possède les attributs suivants :
    - `id_annonce` (de type `INTEGER`) : un code permettant d'identifier chaque annonce de manière unique.
    - `id_vehicule` (de type `INTEGER`) : l'identifiant du modèle du véhicule vendu, qui correspond à l'identifiant utilisé dans la table `vehicule`.
    - `annee` (de type `INTEGER`) : année de première mise en circulation du véhicule.
    - `km` (de type `INTEGER`) : nombre de kilomètres parcourus par le véhicule au moment de la revente.
    - `prix_occasion` (de type `INTEGER`) : prix de vente du véhicule d'occasion.
1. En justifiant brièvement, identifier une clé primaire dans chacune des tables `vehicule` et `annonce`, ainsi qu'une clé étrangère dans la table `annonce`.
  2. Écrire une requête SQL permettant d'extraire les noms de tous les modèles mis en vente par le constructeur Dubreuil Motors.
  3. Expliquer le fonctionnement de la requête SQL suivante et préciser l'effet de cette requête sur chacune des tables `vehicule` et `annonce`.

```
1 UPDATE annonce
2 SET prix_occasion = prix_neuf
3 FROM vehicule
4 WHERE vehicule.id_vehicule = annonce.id_vehicule
5 AND vehicule.prix_neuf < annonce.prix_occasion
```

4. À l'aide d'une jointure, écrire une requête SQL permettant d'obtenir, sur une même table, la liste de toutes les annonces de la table annonce avec les attributs suivants
  - o l'identifiant de l'annonce id\_annonce.
  - o le kilométrage km.
  - o le prix de vente du véhicule prix\_neuf.
  - o le prix de l'annonce d'occasion prix\_occasion.
5. Faire apparaître toutes les annonces des véhicules ayant moins de 100 000 kilomètres et vendus moins de 5 000 euros.
6. Faire apparaître la ou les annonces du ou des véhicules dont le prix a le plus baissé entre sa mise sur le marché et la date de l'annonce.
7. Trouver toutes les annonces de véhicules construits par Dubreui Motors et mis en circulation après 2020.

## SUJET 2.

**Exercice 1 : chaîne de Markov.** Un triathlon comporte 3 épreuves : la course  $C$ , la nage  $N$  et le vélo  $V$ . Un triathlète prépare son entraînement de la manière suivante.

- Le premier jour, chaque sport a la même probabilité d'être choisi.
- La course étant l'épreuve la plus fatigante, le triathlète ne court jamais 2 jours de suite.
- Si le triathlète à couru un jour donné, il choisira  $N$  ou  $V$  le jour suivant avec autant de chances.
- S'il nage ou fait du vélo un jour donné, il y a une chance sur deux qu'il change d'activité le lendemain et dans ce cas, une chance sur 2 pour qu'il court.

On note, pour  $n \in \mathbb{N}^*$ ,  $X_n$  la variable aléatoire qui prend la valeur 1 (resp. 2, resp. 3) si l'activité  $C$ ,  $N$  ou  $V$  a été choisie au  $n$ ème jour d'entraînement. On note, pour  $n \in \mathbb{N}^*$ ,  $U_n$  le vecteur de  $\mathbb{R}^3$  défini par

$$U_n = (P(X_n = 1), P(X_n = 2), P(X_n = 3)).$$

1. a. Quelle est la loi de  $X_1$ ? Préciser son espérance et sa variance.  
b. Proposer une instruction Python permettant de simuler  $X_1$ .
2. Représenter, en justifiant, le diagramme de transition de la chaîne  $(X_n)$ .
3. Déterminer soigneusement, en citant les résultats utilisés, une matrice  $A$  telle que, pour tout  $n \in \mathbb{N}^*$ , on ait

$$U_{n+1} = U_n A.$$

4. Les commandes

```
1 import numpy.linalg as al
2
3 (16*al.matrix_power(A,2)-eye(3,3))*(A-eye(3,3))
```

renvoie

```
1 [[0.  0.  0.],
2  [0.  0.  0.],
3  [0.  0.  0.]]
```

En déduire les valeurs propres possibles de  $A$ .

5. Vérifier que les valeurs précédentes sont bien valeurs propres de  $A$ , puis exhiber, en rédigeant soigneusement, une matrice  $D$  diagonale (dont les éléments diagonaux sont rangés dans l'ordre croissant) et une matrice  $P$  inversible telles que

$$A = PDP^{-1}.$$

6. Inverser  $P$ .
7. Calculer explicitement  $A^n$  et en déduire la loi de  $X_n$ .

**Exercice 2 : SQL.** Un cuisinier note ses recettes dans une base de données. Il tient à jour trois tables de données, recettes, ingrédients et association, chacune d'entre elles possède les attributs suivants.

- Dans la table `recettes`, il décrit les types de recettes qu'il sait faire et renseigne leurs caractéristiques :
  - Le nom de la recette (de type `TEXT`) qui la décrit de manière unique.
  - Le temps (de type `INTEGER`) qui donne le temps nécessaire à la réalisation de la recette, en minutes.
  - La complexité (donnée par une note sur 20 de type `INTEGER`) qui indique la difficulté technique de la recette.
  - La satisfaction estimée de ses clients lorsqu'ils goûtent la recette, donnée par un note sur 20 de type `INTEGER`.
- La table `ingrédients` regroupe les matières premières que le cuisinier utilise et les caractéristiques suivantes :
  - Le nom (de type `TEXT`) qui le décrit de manière unique.
  - Le prix de l'ingrédient (de type `INTEGER`).
  - La saison (de type `TEXT`) qui indique à quel moment de l'année on peut se procurer l'ingrédient.

- L'attribut `vegetarien` de type `BOOLEAN` qui permet d'indiquer si l'ingrédient convient aux végétariens : il prend la valeur 0 si l'ingrédient ne convient pas et 1 s'il convient.

- La table `association` indique de quels ingrédients le cuisinier se sert pour réaliser une recette.

1. Rappeler ce qu'est une table d'association et justifier qu'il est indispensable de s'en servir ici.
2. Quelles sont les clés primaires et secondaires de ces trois tables.
3. Expliquer le fonctionnement de la requête SQL suivante

```
1 DELETE FROM ingredients
2 WHERE
3 saison <> 'printemps'
4 OR
5 saison <> 'ete'
6 OR
7 saison <> 'automne'
8 OR
9 saison <> 'hiver'
```

Que se passe-t-il sur les trois tables lorsqu'on appelle cette requête.

4. Une journée particulièrement intense, le cuisinier n'a pas le temps de mettre à sa carte les plats dont soit la complexité dépasse 10, soit le temps de préparation dépasse 30 minutes. Quelle requête faut-il faire pour afficher les plats que le cuisinier met à la carte.
5. Un soir de service, tous les clients du restaurant sont végétariens. Quelle requête permet de faire afficher les recettes qui conviennent ?
6. Proposer une méthode pour faire afficher le prix de la recette 'couscous', c'est-à-dire la somme des prix des ingrédients dont le cuisinier se sert pour faire un couscous.
7. Le cuisinier veut faire évoluer sa carte, il décide de supprimer tous les plats dont la note donnée par les clients est inférieure à 10 et dont le prix est supérieur à 20 euros. Comment faire afficher ces plats ?
8. Après quelques années de labeur et grâce à ses succès bien mérités, le cuisinier a maintenant économisé suffisamment d'argent pour s'agrandir. Il décide d'acheter un food truck, en plus de son restaurant original. Les deux cuisines ne sont pas du tout équipées de la même façon et le cuisinier ne peut donc pas faire toutes ses recettes dans chacun des deux établissements. Indiquer comment modifier la base de données pour tenir compte de ces nouvelles contraintes.

## SUJET 3

**Exercice 1 : chaîne de Markov.** Un mobile se déplace aléatoirement dans l'ensemble des sommets d'un triangle. On note ces sommets 1,2 et 3. Les déplacements obéissent à la règle suivante : si à l'instant  $n$ , le mobile est sur l'un quelconque des sommets, soit il y reste au moment  $n + 1$  avec probabilité  $2/3$ , soit il bouge sur l'un quelconque des sommets adjacents, et ceci avec la même probabilité.

On note  $X_n$  la variable aléatoire égale au numéro du sommet sur lequel se trouve le mobile au moment  $n$ . On suppose qu'à l'instant 0, il se trouve sur le sommet numéro 1 (ou encore  $X_0 = 1$ ).

1. Représenter le diagramme de transition de la chaîne de Markov ( $X_n$ ).
2. En déduire la matrice de transition  $A$ .
3. Quelles instructions Python permettent de simuler et d'afficher la trajectoire des  $n$  premiers termes de la chaîne, où  $n$  est rentré par l'utilisateur ?
4. On pose  $B = 6A$ . Calculer  $B^2 - 9B + 18I_3$ .
5. En déduire la diagonalisation de  $B$ , puis les puissances de  $B$ .
6. Déterminer la loi de  $X_n$  et commenter sur la position du mobile à long terme.

**Exercice 2 : SQL.** Un cinéphile tient à jour une base de données sur le thème du cinéma. Cette base de données contient deux tables, films et acteurs dont les attributs sont les suivants.

- La table films possède les attributs
    - id\_films (de type INTEGER) qui donne à chaque enregistrement un numéro unique.
    - nom (de type TEXT) indique le nom du film.
    - annee (de type INTEGER) donne l'année de production du films.
    - nationalite (de type TEXT) donne le pays dans lequel le film a été produit.
    - realisateur (de type TEXT) indique la personne qui a fait le film.
    - acteur\_principal (de type TEXT) donne le nom de l'acteur qui a le premier rôle dans le film.
    - acteur\_secondaire (de type TEXT) donne le nom de l'acteur qui a le second rôle.
  - La table acteurs possède les attributs
    - id\_acteur (de type INTEGER) est un numéro unique attribué à chaque acteur.
    - nom (de type TEXT) donne le nom de l'acteur ou de l'actrice.
    - nationalite (de type TEXT) donne le pays de naissance de l'acteur ou de l'actrice (on ne tient donc pas compte des doubles nationalités).
    - annee\_naissance (de type INTEGER) donne l'année de naissance de l'acteur ou de l'actrice.
1. Quelles sont les clés primaires et les clés secondaires de ces tables de données ? Pourquoi le nom du film ne peut-il pas être utilisé pour servir de clé primaire (*Indication* : "La vie est belle"). Trouver deux acteurs qui ont le même nom. Qu'en déduit-on ?
  2. Quel type de film apparaît si on demande

```
1 SELECT * FROM films
2 WHERE acteur_secondaire = NULL
```

3. On forme la requête suivante

```
1 SELECT films.nom
2 FROM
3 films JOIN acteur
4 ON films.acteur_principal = acteurs.nom
5 WHERE
6 films.annee < acteur.annee_naissance
```

Qu'est-ce qui est retourné par votre interpréteur ?

4. Comment faire afficher tous les films qui ont été produits après votre naissance ?
5. Proposer une requête qui sert à afficher tous les films français dont l'acteur principal est japonais.

6. Compter le nombre de films ayant été réalisés par Christopher Nolan. Pourquoi ne peut-on pas procéder de la même manière pour compter les films de Claude Lellouch ?
7. Comment faire pour trouver le réalisateur qui fait tourner le plus d'acteurs espagnols.
8. On souhaite faire évoluer la base de données de manière à ce que tous les acteurs qui ont joué dans chaque film soient accessibles, et pas seulement les acteurs principaux. Comment modifications faut-il apporter ?