

# Examen de *Conception et traitement des données*

Claire Medrala (MESR), Fabien Coelho (Mines Paris – PSL)

Lundi 13 janvier 2025 – 9h00

## Conditions de l'examen

L'examen dure 1 heure 15 minutes. **Aucun document ni appareil électronique** n'est autorisé. **Aucune communication** entre élèves n'est permise. Un dictionnaire peut être consulté par les élèves non francophones. Le barème est indicatif, la notation sur 20. Des points de bonus sont susceptibles d'être attribués au loisir des correcteurs. L'exercice 3 dépend de l'exercice 2.

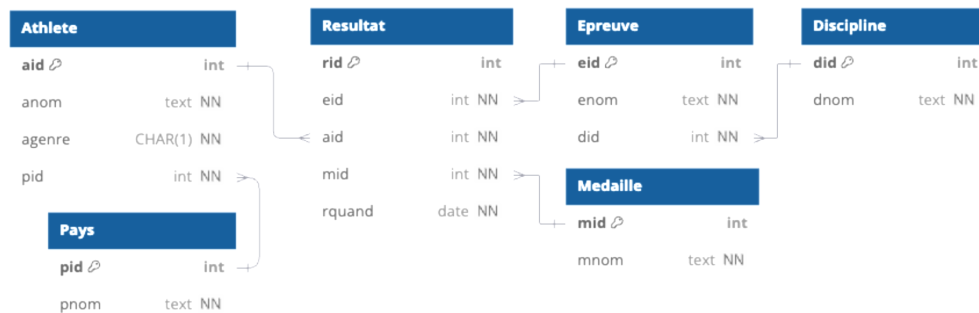
## 1 Citius, Altius, Fortius – Communiter

/7

L'année 2024 a été marquée par les Jeux Olympiques d'été à Paris.

Intéressons-nous d'abord au tableau des médailles. Le schéma de données suivant permet de représenter les médailles attribuées aux athlètes lors de chaque épreuve. Par exemple, le *2 août 2024, Léon Marchand*, représentant la *France*, a gagné la médaille d'*or* du *200 m 4 nages homme* en *natation*.

Un-e athlète représente un seul pays. Il ou elle peut gagner des médailles dans plusieurs épreuves. Une médaille est en *or*, *argent* ou *bronze*. Une discipline comprend plusieurs épreuves, par exemple l'*athlétisme* comprend des épreuves comme le *200 m homme*, le *200 m femme*, le *100 m haies femme*, le *marathon homme*, le *relais marche marathon mixte*...



Dans ce schéma, les attributs en gras sont les clés primaires et les arcs désignent les clés étrangères. Sont considérés comme unique : un-e athlète (*anom*), un pays (*pnom*), une couleur de médaille (*mnom*), un résultat (*eid*, *aid*), une épreuve dans une discipline (*enom*, *did*).

Proposez **une** requête (en cas d'ambiguïté choisir une interprétation possible) ou à défaut sa structure pour y répondre, pour les questions suivantes :

1. Quel-le-s athlètes *français·e·s* ont gagné une médaille d'*or* dans une épreuve de *natation* ?
2. Pour la requête précédente, quels (au moins trois) index (hors clés primaires ou uniques déjà signalées) sont potentiellement utiles pour en améliorer les performances ?
3. Comment afficher par ordre alphabétique les noms des disciplines qui ont au moins *12* épreuves, avec leur nombre d'épreuves ?

4. Comment afficher pour toutes les disciplines et tous les pays, le nombre de couleurs des médailles obtenues (de 0 à 3), par ordre des disciplines et des pays ?
5. Comment afficher pour tou·te·s les athlètes médaillé·e·s leur nombre de médailles par catégorie (or, argent, bronze), en ordre décroissant des couleurs puis par ordre alphabétique ?
6. En identifiant les épreuves mixtes comme celles ayant des médaillé·e·s des deux genres, et les épreuves individuelles comme celles ayant distribué 3 ou 4 médailles, quelles sont les épreuves individuelles mixtes (nom de la discipline et de l'épreuve) ?

Réponse : *Sports équestres, Concours complet individuel.*

## 2 L'important est de participer ! /6

Un modèle entité-association (E/A) se compose d'**entités** (rectangles) reliées entre elles par des **associations** (losanges). Chaque lien a une **cardinalité** (? 1 \* +). Entités et associations ont des **attributs** (ovales) qui représentent des données.

1. Reconstituer le modèle E/A du schéma relationnel précédent, **sans** les attributs existants.
2. Ajouter les notions suivantes avec des entités, associations ou attributs supplémentaires :
  - le type de l'épreuve : *individuelle* ou *collective*, avec un nombre de participant ;
  - le genre de l'épreuve : *masculin*, *féminin* ou *mixte* ;
  - les disciplines sont regroupées en sports, par exemple le sport *canoë-kayak* a pour disciplines *course en ligne* et *slalom* ;
  - le site de chaque épreuve (*Arena Bercy, Colline d'Élancourt, Marina de Marseille, Teahupo'o à Tahiti...*), les sites pouvant accueillir plusieurs épreuves en même temps ;
  - les compétitions de différents types (*tour préliminaire, séries, qualifications, repêchage, première et seconde demi-finales, petite finale, finale...*) qui composent les épreuves, se déroulent certains jours à certains horaires (début, durée) sur le site de l'épreuve ;
  - l'inscription des athlètes aux épreuves et leur participation aux compétitions associées.

Discuter éventuellement les motivations des choix de modélisation et les limites du modèle.

## 3 Vive les JO ! /3

À partir de ce nouveau modèle E/A, donner les commandes SQL pour construire le schéma relationnel correspondant en modifiant les tables existantes et en ajoutant les nouvelles tables, en portant attention aux types des données et aux contraintes. Vous prendrez soin de bien préciser les champs utiles et les contraintes pertinentes sur vos relations. Vous commenterez les contraintes que vous ne pourriez exprimer directement dans le modèle.

Pour gagner du temps de rédaction, vous utiliserez les abréviations suivantes : **CT** pour **CREATE TABLE**, **AT** pour **ALTER TABLE**, **I** pour **INT**, **S** pour **SERIAL**, **B** pour **BOOL**, **F** pour **FLOAT**, **T** pour **TEXT**, **D** pour **DATE**, **Ts** pour **TIMESTAMP**, **In** pour **INTERVAL**, **PK** pour **PRIMARY KEY**, **U** pour **UNIQUE**, **NN** pour **NOT NULL**, **R** pour **REFERENCES**, **C** pour **CHECK...**

## 4 À vos marques, prêts, partez ! /4

Choisissez un thème parmi les deux tirés aléatoirement en début d'examen dans la liste *Postgres, Relationnel, DCL (droits), Optimisation, Transactions, MVCC, Formes normales* et expliquez en 100 mots ce que vous en avez retenu.

Citez les noms de trois scientifiques ayant obtenu le prix Turing pour leurs travaux de recherche sur les bases de données relationnelles.