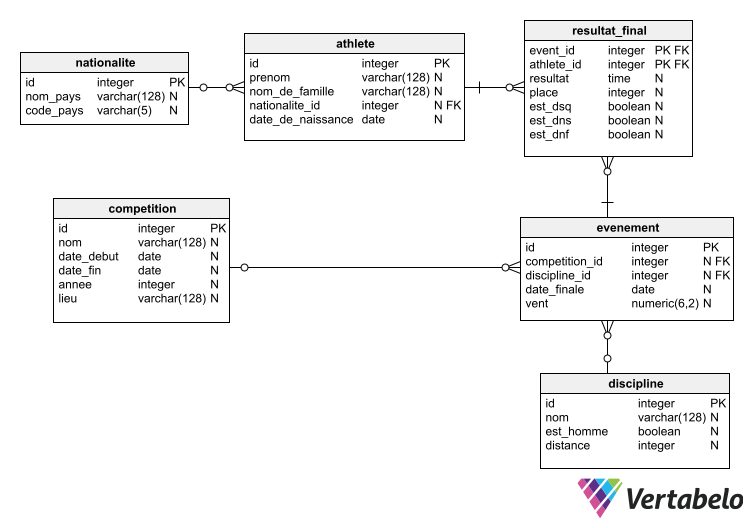
# Requêtes SQL : Sélection, agrégation et regroupement

|  |  |
| --- | --- |
| Objectifs | Révision sur les sélections simples à savoir :   * La sélection de lignes * La sélection de colonnes * Ordonner les lignes * Renommer les colonnes * Limiter le nombre de lignes * Supprimer les doublons |
| Méthodologie | Travail individuel |
| Durée estimée | 45 min (avec le corrigé) |
| Type de base de données : | MySQL |

Dans cet exercice, nous allons travailler avec des données concernant les finales de différentes disciplines de course sur piste lors de quelques championnats d'athlétisme : Jeux Olympiques de Rio de Janeiro en 2016, Championnats du monde d'athlétisme de l'IAAF de Londres en 2017, et Championnats du monde d'athlétisme de l'IAAF de Doha en 2019.

Jetons un coup d'œil aux données que nous allons traiter.



La base de données comporte six tables :

**competition**, **evenement** , **discipline**, **resultat**\_**final**, **athlete** et **nationalite**.

Passons en revue chacune d'entre elles.

La table **competition** comporte les colonnes suivantes :

**id** - un identifiant.

**nom** - le nom de la compétition.

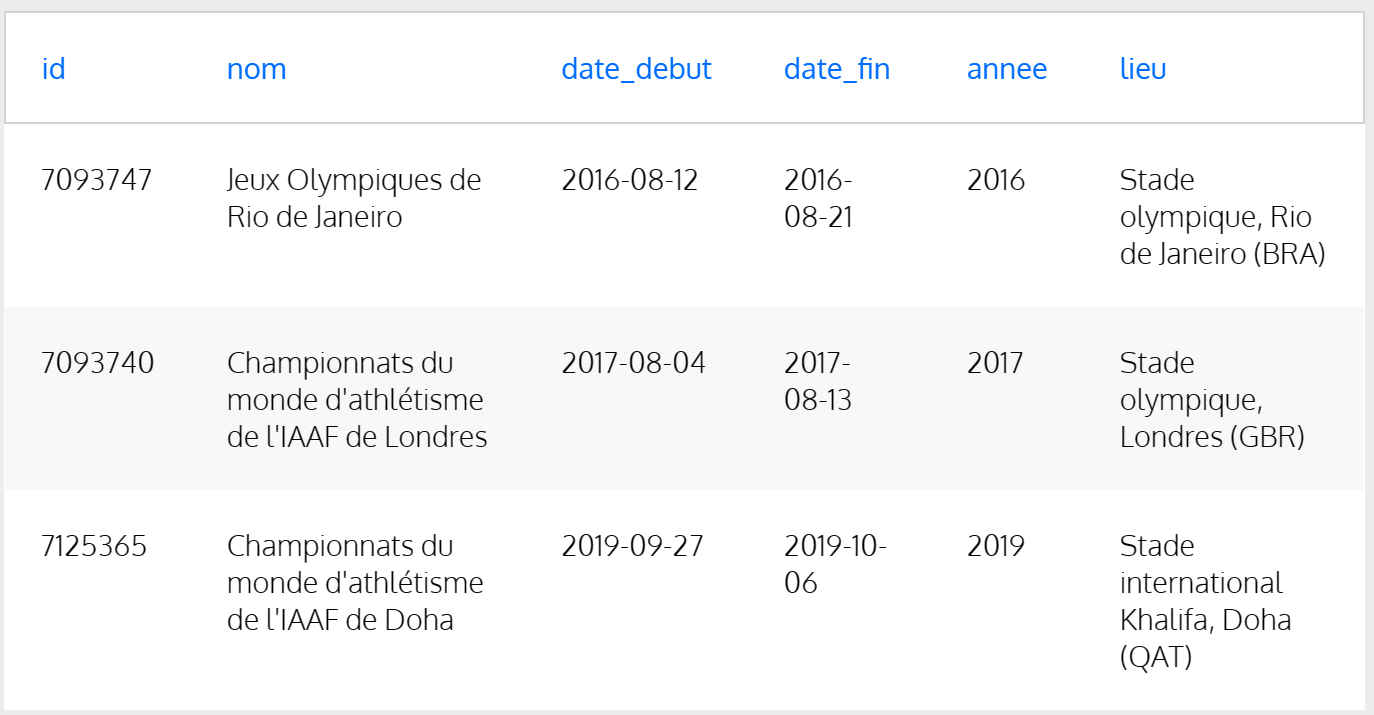
**date**\_debut - le premier jour de la compétition.

**date**\_fin - le dernier jour de la compétition.

**annee** - l'année au cours de laquelle une compétition donnée a eu lieu.

**lieu** - le lieu où s'est déroulée une compétition donnée.

Voici quelques exemples de lignes :



La table **discipline** comporte les colonnes suivantes :

**id** - un identifiant.

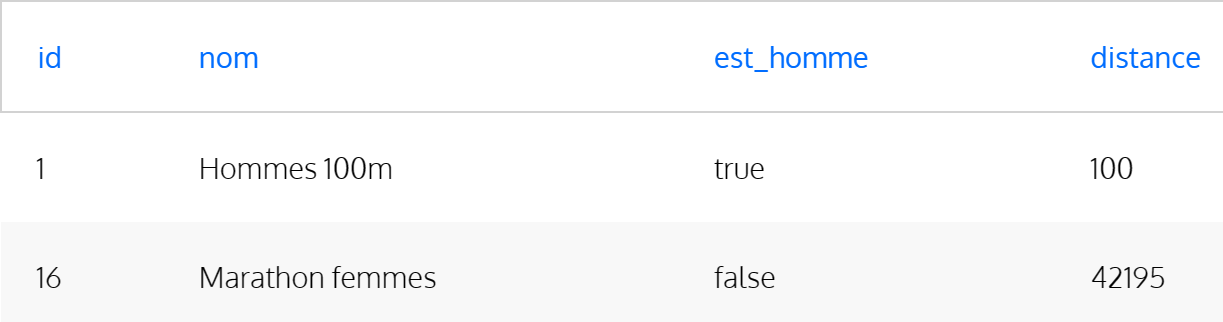
**nom** - le nom de la discipline donnée.

**est\_homme** - TRUE s'il s'agit d'une discipline masculine, FALSE s'il s'agit d'une discipline féminine.

**distance** - la distance de la discipline donnée, exprimée en mètres.

Nous stockons des informations pour toutes les disciplines de la course à pied - des courses de courte distance comme le 100 mètres aux courses de longue distance comme le marathon.

Voici quelques exemples de lignes :



La table **evenement** stocke les informations relatives à un événement, par exemple "100m femmes", et comporte les colonnes suivantes :

**id** - un identifiant.

**competition\_id** - correspondant à l'identifiant de la compétition.

**discipline\_id** - correspondant à l'identifiant de la discipline.

**date\_finale** - la date à laquelle s'est déroulée la finale d'une épreuve donnée.

**vent** - les points de vent pendant la finale.

La table **athlete** comporte les colonnes suivantes :

**id** - un identifiant.

**prenom** - le prénom de l'athlète.

**nom\_de\_famille** - le nom de famille de l'athlète

**nationalite\_id** - la nationalité de l'athlète.

**date\_de\_naissance** - la date de naissance de l'athlète.

La table **nationalite** stocke les informations relatives aux pays et comporte les colonnes suivantes :

**id** - un identifiant.

**nom\_pays** - le nom du pays.

**code\_pays** - l'abréviation à trois lettres du pays.

La table **resultat\_final** contient des informations sur tous les participants à un événement particulier :

**event\_id** - l'identifiant d'un événement correspondant.

**athlete\_id** - l'identifiant d'un athlète.

**resultat** - le temps réalisé par un athlète donné. Peut être manquant.

**place** - la place d'un athlète donné. Peut être manquant.

**est\_dsq** - TRUE en cas de disqualification.

**est\_dnf** - TRUE si un athlète donné n'a pas terminé la course (did not finish).

**est\_dns** - TRUE si un athlète donné n'a pas pris le départ de la course (did not start).

Les trois dernières colonnes peuvent prêter à confusion, mais cette notation est largement utilisée. Par exemple, voici à quoi ressembleraient les résultats du 5000 mètres masculin en 2019, par exemple, à la télévision :



Muktar Edris s'est imposé avec le résultat de 12 minutes et 58,85 secondes, son meilleur résultat de la saison.

L'un des frères Ingebrigtsen - Filip - n'a pas terminé la course. A un moment donné, il est sorti de la piste. Le statut de son résultat est DNF - n'a pas terminé.

Les autres statuts possibles sont :

DNS - n'a pas pris le départ

DSQ - disqualifié.

Les indicateurs booléens est\_dnf, est\_dns, et est\_dsq correspondent à ces statuts possibles.

Voici ce que cela donne dans notre base de données, dans la table **resultat**\_final:



Pour obtenir les noms, dates de naissance et nationalités des athlètes, vous devrez utiliser d'autres tables. Comme vous pouvez le constater, la dernière ligne correspond à une course non terminée, de sorte que les valeurs de temps et de place sont manquantes.

Requête n°1 :

Où se sont déroulées toutes les compétitions ? Indiquez tous les lieux.

|  |
| --- |
| SELECT lieu  FROM competition; |

Requête n°2 :

Affichez les dates finales de tous les événements et les points de vent.

|  |
| --- |
| SELECT date\_finale, vent  FROM evenement; |

Requête n°3 :

Sélectionnez tous les noms et prénoms, ainsi que les dates de naissance de tous les athlètes.

|  |
| --- |
| SELECT nom\_de\_famille, prenom, date\_de\_naissance  FROM athlete; |

Requête n°4 :

Sélectionnez tous les noms de compétitions ainsi que leurs dates de début et de fin.

|  |
| --- |
| SELECT nom, date\_debut, date\_fin  FROM competition; |

Requête n°5 :

Sélectionnez toutes les colonnes pour les lignes dont le résultat est en première position.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE place = 1; |

Requête n°6 :

Affichez toutes les données relatives aux compétitions qui se sont déroulés avant l'année 2018.

|  |
| --- |
| SELECT \*  FROM competition  WHERE annee < 2018; |

Requête n°7 :

Affichez tous les résultats (temps) pour lesquels la place était égale ou supérieure à 100.

|  |
| --- |
| SELECT resultat  FROM resultat\_final  WHERE place >= 100; |

Requête n°8 :

Affichez toutes les dates finales où le vent était supérieur à 0,5.

|  |
| --- |
| SELECT date\_finale  FROM evenement  WHERE vent > 0.5; |

Requête n°9 :

Affichez tous les noms de disciplines dont la distance est supérieure à 3000 mètres.

|  |
| --- |
| SELECT nom  FROM discipline  WHERE distance > 3000; |

Requête n°10 :

Affichez les noms et prénoms des athlètes dont le prénom commence par K et se termine par a.

|  |
| --- |
| SELECT nom\_de\_famille, prenom  FROM athlete  WHERE prenom LIKE 'K%a'; |

Requête n°11 :

Affichez tous les noms de pays commençant par 'U'.

|  |
| --- |
| SELECT nom\_pays  FROM nationalite  WHERE nom\_pays LIKE 'U%'; |

Requête n°12 :

Affichez tous les noms de pays composés d'au moins trois mots.

|  |
| --- |
| SELECT nom\_pays  FROM nationalite  WHERE nom\_pays LIKE '% % %'; |

Requête n°13 :

Quel est le pays dont l'abréviation est 'ART'? Affichez son nom.

|  |
| --- |
| SELECT nom\_pays  FROM nationalite  WHERE code\_pays = 'ART'; |

Requête n°14 :

Affichez toutes les lignes de la table discipline relatives aux marathons (contenant 'Marathon' dans le nom de la discipline).

|  |
| --- |
| SELECT \*  FROM discipline  WHERE nom LIKE 'Marathon%' |

Requête n°15 :

Affichez toutes les lignes de resultat\_final où le coureur a terminé sans place.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE place IS NULL; |

Requête n°16 :

Affichez toutes les lignes de la table resultat\_final où une disqualification s'est produite.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE est\_dsq IS TRUE; |

Requête n°17 :

Affichez toutes les lignes de la table resultat\_final où le coureur n'a pas pris le départ.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE est\_dns IS TRUE; |

Requête n°18 :

Affichez tous les noms de disciplines féminines.

|  |
| --- |
| SELECT nom  FROM discipline  WHERE nom LIKE '%femmes%'; |

Requête n°19 :

Affichez toutes les lignes de la table resultat\_final où un coureur a été disqualifié mais a commencé et terminé la course.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE est\_dsq IS TRUE  AND est\_dnf IS FALSE  AND est\_dns IS FALSE; |

Requête n°20 :

Affichez toutes les lignes de resultat\_final pour tous les médaillés d'or dont le temps est supérieur à une heure.

|  |
| --- |
| SELECT \*  FROM resultat\_final  WHERE place = 1  AND resultat > '01:00:00.00';  -- Alternative avec le mot clé INTERVAL  SELECT \*  FROM resultat\_final  WHERE place = 1  AND resultat > INTERVAL '1 hour'; |

Requête n°21 :

Affichez tous les noms des disciplines masculines si la distance est inférieure à 500 mètres.

|  |
| --- |
| SELECT nom  FROM discipline  WHERE nom LIKE '%hommes'  AND distance < 500; |

Requête n°22 :

Affichez tous les noms et prénoms des athlètes s'ils contiennent un 'z' (un ou plusieurs, en minuscules ou en majuscules) dans leur nom et leur prénom.

|  |
| --- |
| SELECT nom\_de\_famille, prenom  FROM athlete  WHERE nom\_de\_famille LIKE '%z%'  AND prenom LIKE '%z%'; |

Requête n°23 :

Indiquez le prénom et le nom de tous les athlètes s'ils contiennent un 'x' (un ou plusieurs, en minuscules ou en majuscules) dans l'un ou l'autre de leurs noms.

|  |
| --- |
| SELECT prenom, nom\_de\_famille  FROM athlete  WHERE prenom LIKE '%x%'  OR nom\_de\_famille LIKE '%x%'; |

Requête n°24 :

Affichez les noms et abréviations des pays. Trier les lignes par nom de pays, dans l'ordre croissant.

|  |
| --- |
| SELECT nom\_pays, code\_pays  FROM nationalite  ORDER BY nom\_pays;  -- equivalent  SELECT nom\_pays, code\_pays  FROM nationalite  ORDER BY nom\_pays ASC; |

Requête n°25 :

Affichez tous les noms des disciplines féminines possibles, triés par la distance, dans l'ordre décroissant.

|  |
| --- |
| SELECT nom  FROM discipline  WHERE nom LIKE '%femmes%'  ORDER BY distance DESC;  -- equivalent  SELECT nom  FROM discipline  WHERE est\_homme IS FALSE  ORDER BY distance DESC; |

Requête n°26 :

Affichez tous les temps et places des résultats. Trier les temps par ordre croissant. Si plusieurs résultats sont identiques, affichez d'abord celui qui a la meilleure place.

|  |
| --- |
| SELECT resultat, place  FROM resultat\_final  ORDER BY resultat ASC, place ASC; |

Requête n°27 :

Pour tous les athlètes, indiquez leur nom et leur prénom. Triez par le nom de famille, et si quelques athlètes le partagent, triez par leur prénom (les deux dans l'ordre croissant).

|  |
| --- |
| SELECT nom\_de\_famille, prenom  FROM athlete  ORDER BY nom\_de\_famille, prenom; |

Requête n°28 :

Affichez le nom et le prénom de tous les athlètes dont le nom de famille commence par 'X'. Triez le résultat par le prénom dans l'ordre décroissant.

|  |
| --- |
| SELECT nom\_de\_famille, prenom  FROM athlete  WHERE nom\_de\_famille LIKE 'X%'  ORDER BY prenom DESC; |

Requête n°29 :

Pour tous les résultats finaux, affichez les résultats de temps qui sont d'au moins 3 heures. Triez les lignes par ordre décroissant des résultats.

|  |
| --- |
| SELECT resultat  FROM resultat\_final  WHERE resultat > '03:00:00'  ORDER BY resultat DESC;  -- equivalent  SELECT resultat  FROM resultat\_final  WHERE resultat > INTERVAL '3 hour'  ORDER BY resultat DESC; |

Requête n°30 :

Affichez tous les noms de pays commençant par 'T' ou 'Z' triés par ordre croissant..

|  |
| --- |
| SELECT nom\_pays  FROM nationalite  WHERE nom\_pays LIKE 'T%'  OR nom\_pays LIKE 'Z%'  ORDER BY nom\_pays; |

Requête n°31 :

Affichez toutes les disciplines masculines triées par distance dans l'ordre décroissant. Affichez deux colonnes : nom et distance.

|  |
| --- |
| SELECT nom, distance  FROM discipline  WHERE est\_homme IS TRUE  ORDER BY distance DESC; |

Requête n°32 :

Sélectionnez toutes les lignes pour tous les événements où les points de vent étaient positifs. Trier par ordre décroissant des points de vent.

|  |
| --- |
| SELECT \*  FROM evenement  WHERE vent > 0  ORDER BY vent DESC; |

Requête n°33 :

Comptons le nombre de nationalités que nous avons dans notre base de données. Affichez une colonne : compte\_nationalite.

|  |
| --- |
| SELECT COUNT(\*) AS compte\_nationalite  FROM nationalite; |

Requête n°34 :

Afficher le nombre d’événements que nous stockons et renommer la colonne : compte\_finales.

|  |
| --- |
| SELECT COUNT(\*) AS compte\_finales  FROM evenement; |

Requête n°35 :

Voyons combien de courses se sont terminées avec et sans résultat. Affichez deux colonnes : **courses**\_**reussies** et **courses\_non\_reussies**.

|  |
| --- |
| SELECT  COUNT(resultat) AS courses\_reussies,  COUNT(\*) - COUNT(resultat) AS courses\_non\_reussies  FROM resultat\_final; |

Requête n°36 :

Cherchons le prénom le plus populaire. Affichez deux colonnes : prenom et nombre\_de\_noms. Triez les résultats par ordre décroissant des prénoms..

|  |
| --- |
| SELECT prenom, COUNT(\*) AS nombre\_de\_noms  FROM athlete  GROUP BY prenom  ORDER BY prenom DESC; |

Requête n°37 :

Trouvons le prénom le plus populaire.

Affichez deux colonnes : prenom et nombre\_de\_noms, mais uniquement si le nombre est au moins égal à 5.

Triez le résultat par ordre décroissant.

|  |
| --- |
| SELECT  prenom,  COUNT(\*) AS nombre\_de\_noms  FROM athlete  GROUP BY prenom  HAVING COUNT(\*) >= 5  ORDER BY nombre\_de\_noms DESC; |

Requête n°38 :

Comptons combien il y a de pays qui contient ‘Républiques’ dans leur nom. N'affichez qu'une seule colonne : compte\_republiques.

|  |
| --- |
| SELECT COUNT(\*) AS compte\_republiques  FROM nationalite  WHERE nom\_pays LIKE '%République%'; |