



Université
de Lille

UNIVERSITÉ DE LILLE

SAÉ S2.04
STATISTIQUES

RAPPORT

Analyse de données

Élèves :

Baptiste LAVOGIEZ

Enseignant :

Marie DELETOMBE

18 juin 2025

A decorative graphic in the bottom right corner of the page, consisting of a grey, folded paper-like shape.

Table des matières

1	Introduction	3
1.1	Contexte	3
1.2	Le déroulement de ce portfolio	3
2	SAÉ 1.02 - Réalisation d'un jeu en iJava	4
2.1	Présentation	4
2.2	Quelle était l'application ?	4
2.3	En images	4
2.4	Qu'ai-je appris ?	5
3	SAÉ 1.03 - Installation de machines virtuelles	7
3.1	Présentation	7
3.2	Quelle était l'application ?	7
3.3	Qu'ai-je appris ?	7
4	SAÉ 1.04 - Base de données Access	8
4.1	Présentation	8
4.2	Quelle était l'application ?	8
4.3	En images	8
4.4	Qu'ai-je appris ?	10
5	SAÉ 1.05-06 - Réalisation d'un site Web	11
5.1	Présentation	11
5.2	Quelle était l'application ?	11
5.3	En images	11
5.4	Qu'ai-je appris ?	11
6	SAÉ 2.01-02 - Logiciel de gestion d'échanges en Java	12
6.1	Présentation	12
6.2	Quelle était l'application ?	12
6.3	En images	13
6.4	Qu'ai-je appris ?	14
7	SAÉ 2.03 - Installation de services réseaux	15
7.1	Présentation	15
7.2	Quelle était l'application ?	15
7.3	Qu'ai-je appris ?	15
8	SAÉ 2.05 - Gestion d'un projet	16
8.1	Présentation	16
8.2	Quelle était l'application ?	16
8.3	En images	16
8.4	Qu'ai-je appris ?	18

9	SAÉ 2.06 - Réalisation d'une vidéo sur un point de science informatique	19
9.1	Présentation	19
9.2	Quelle était l'application ?	19
9.3	Qu'ai-je appris ?	19

1 Contexte

Cette partie décrit le modèle de bases de données utilisé. Elle ne constitue pas une réponse au sujet statistique mais un rappel du contexte choisi. AAA

1.1 Introduction

Ce **rapport** présentera la mise en forme d'une base de données en nous appuyant sur les Jeux Olympiques (environ 270 000 lignes, 15 colonnes) au format *CSV*. Nous allons d'abord importer les données, pour ensuite les *ventiler*, soit les ranger dans des tables plus adaptées afin de décomposer les données au mieux. Lors de la partie *SQL*, nous travaillerons sur un serveur *PostgreSQL*.

1.2 Données brutes

Nous obtenons le fichier suivant, pesant près de 40 Mo.


 `img/fichier_explorateur.png`

FIGURE 1 – Aperçu dans l'explorateur

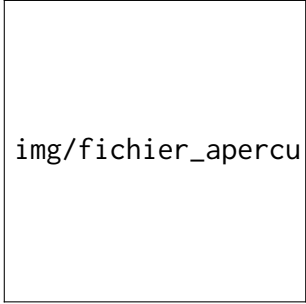
 `img/fichier_aperçu.png`

FIGURE 2 – Aperçu du fichier

1.3 Observations

Nous constatons alors que bon nombre de données sont répétées et c'est normal, car nous nous situons dans un fichier brut.

Nous pouvons distinguer plusieurs tables possibles :

- Des athlètes,
- Des jeux,
- Des sports,
- Des délégations...

Passons maintenant à la conception de ce modèle !

1.4 Le Modèle Logique de Données (MLD)

Un Modèle Logique de Données est une façon de représenter un schéma de tables sous une forme à la fois graphique et textuelle.

1. Tables

Nous en déduisons, après réflexion, près de 7 tables différentes :

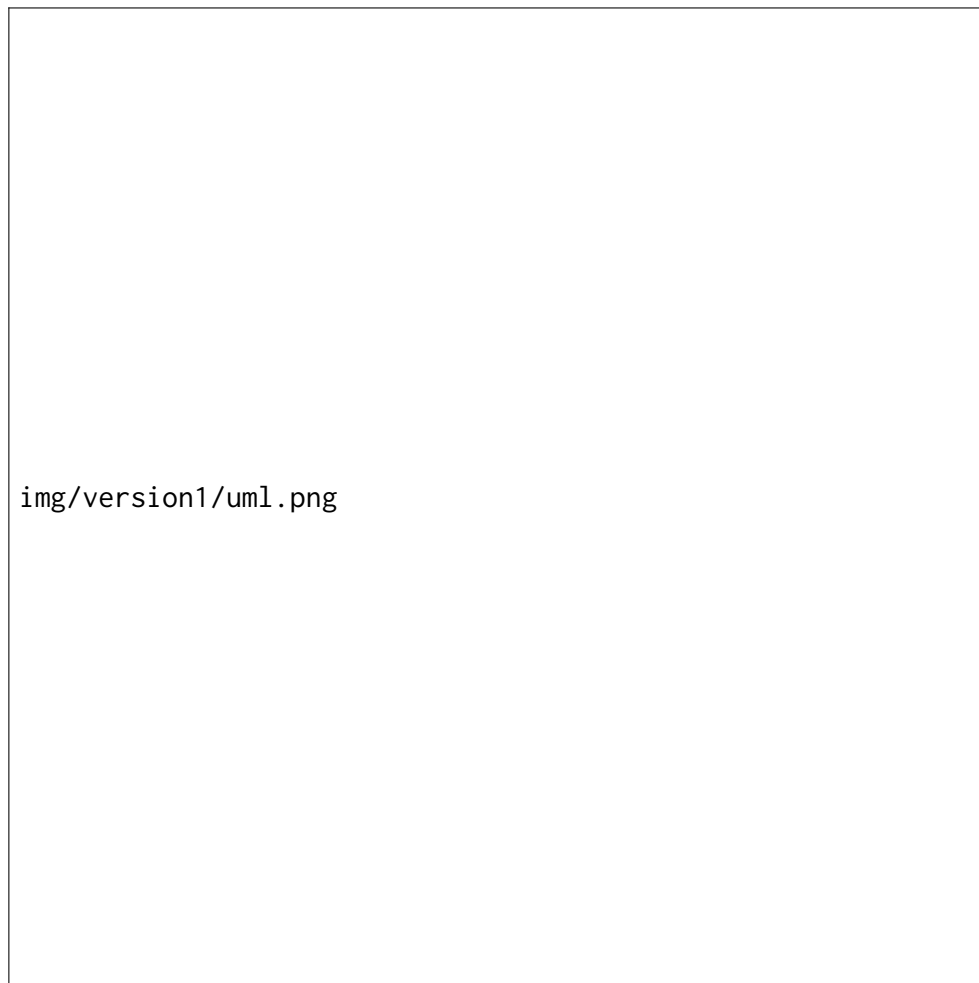


FIGURE 3 – Diagramme UML

La logique utilisée est :

- Une clé primaire est nommée par la première lettre du nom de la table, suivi de "no",
- Une clé primaire est soulignée,
- Une clé étrangère est précédée d'un dièse "#".

2. Raccourcis

On peut penser à Participation(#ano, #noc), ce qui serait possible, mais en réalité

cette table est inutile car un athlète qui fait partie d'une nation revient au même qu'un athlète qui participe pour une nation (dans Participation). De plus, un athlète peut concourir pour une nation une fois, puis pour une autre une autre fois.

3. Précisions

Quelques explications sur des nommages qui peuvent ne pas être explicites :

- Un **NOC** veut dire **National Olympic Committee**. Exemple : États-Unis → USA. Il est unique pour le cas des Jeux Olympiques.
- Une discipline est un sous-genre d'un sport. Exemple : Natation → 4x100m relais. Ce sera utile pour les épreuves.
- Pour le reste, les raisons seront comprises lors de l'exécution des requêtes.

2 Présentation du sujet

2.1 Énoncé

Ce sujet traitera des Jeux Olympiques, avec pour objectif d'analyser leurs données avec une démarche statistique.

2.2 Traitement des questions

Nous aurons donc à répondre à différentes questions statistiques, autant à l'aide de requêtes sur la base de données que d'outils statistiques (formules, graphiques...).

Enfin, afin d'aller plus loin, nous imaginerons d'autres études statistiques en **allant plus loin que le sujet initial**.

Toutes les réponses seront illustrées par des graphiques appropriés.

2.3 Forme des requêtes

Par convention, les requêtes et les tables sont écrites en anglais. La commande `copy`, ou `copy` en SQL nous permettra de générer des fichiers CSV à partir de la sortie des requêtes. Malheureusement, PostgreSQL oblige à tout écrire sur une seule ligne sans pouvoir indenter, alors qu'il serait plus lisible d'aérer les requêtes, surtout les plus denses.

SQL Code

```
1 \copy (SELECT AVG(Participations.age) AS age_moyen_global FROM
    Participations;) TO 'average_global_age.csv' CSV HEADER
```

Ici, on requête l'âge global et on le copie dans un fichier.

De cette façon, dans le rapport, les requêtes seront aérées, sans la commande, et les fichiers, eux, auront la forme ci-dessus.

Ces fichiers seront utiles pour mieux analyser les données, en utilisant notamment l'outil <https://www.rawgraphs.io/> qui, prenant en entrée un jeu de données, permet sa visualisation sous une multitude de formats. Tous ces fichiers sont disponibles dans la racine de ce document.

3 Questions du sujet

Cette partie traitera des questions relatives au sujet.

3.1 Partie 1

1. Déterminer les 20 athlètes affichant le plus de participations aux J.O. (tout sexe confondu)

20 athlètes affichant le plus de participations

```
1      SELECT Athletes.name, COUNT(Participations.ano) AS  
      nombre_participations  
2      FROM Athletes JOIN Participations USING(ano)  
3      GROUP BY Athletes.name  
4      ORDER BY nombre_participations DESC  
5      LIMIT 20;
```


Résultat:

img/version1/20ath.png

FIGURE 4 – 20 athlètes les plus actifs

3.2 Partie 2

Cette partie **statistique** traitera d'une édition précise, et unique vis-à-vis de la promotion, des Jeux Olympiques.

Notre édition sera les Jeux d'Été de 1984, à Los Angeles, aux États-Unis. Les JO d'été étant plus suivis et avec plus d'épreuves, il y aura donc plus de données à visualiser qu'en hiver, et ce afin d'avoir une étude statistique plus précise et plus permissive en terme d'options !

Précision

Afin de sélectionner les jeux de 1984 dans la table, chaque requête utilisera la vue suivante (cette partie ne s'appuie que sur la table *Participations*)

Vue exclusive à 1984

```

1      CREATE VIEW Participations1984 AS (
2          SELECT Participations.*
3          FROM Participations
4          JOIN Events USING(eno)
5          JOIN Games USING(gno)
6          WHERE Games.games = '1984 Summer'
7      );

```

1. Remplir le tableau statistique

Tableau statistique complet

```

1      SELECT Delegations.region, AVG(Participations.age) AS
2          average_age, COUNT(DISTINCT Participations.ano) AS
3          number_of_athletes, MIN(Participations.age) AS min_
4          age, MAX(Participations.age) AS max_age
5      FROM Delegations JOIN Participations USING(noc)
6      WHERE Delegations.region IS NOT NULL
7      GROUP BY Delegations.region
8      ORDER BY average_age DESC;

```

Résultat:

img/version2/tableau_statistique.png

FIGURE 5 – Tableau statistique généré

Pour l'anecdote, les Émirats Arabes Unis apparaissent en haut car l'âge de leurs participants (au nombre de 7) n'est pas renseigné. C'était leur première participation aux JO, et ils n'ont pas eu de médailles.

UAE en 1984

Note : Les âges non renseignés sont exclus lors du calcul de la moyenne, mais pas du nombre d'athlètes. La moyenne d'âge n'est donc pas entièrement représentative, mais elle est plus juste que si on avait pris en compte les âges non renseignés.

2. Comparer l'âge moyen des médaillés et l'âge moyen des participants (tout sexe confondu)

On cherche d'abord l'âge des médaillés, puis celui des participants. *Les participants incluent donc tout le monde, y compris les médaillés.*

Âge moyen des médaillés

```
1 SELECT AVG(Participations.age) AS age_moyen_medailles
2 FROM Participations1984
3 WHERE medal IS NOT NULL ;
```

Résultat: 24.24

Âge moyen des participants

```

1      SELECT AVG(Participations.age) AS age_moyen_participants
2      FROM Participations1984 ;

```

Résultat: 24.00

Comparaison: Les médaillés sont en moyenne 6 mois plus jeunes que les participants.

Comparaison: Les participants sont 2% plus âgés que les médaillés.

3. Comparer le poids moyen des médaillés et le poids moyen des participants (distinguer les hommes et les femmes)

Hommes

Poids moyen des médaillés hommes

```

1      SELECT AVG(Participations.weight)
2      FROM Participations1984
3      JOIN Athletes USING (ano)
4      WHERE medal IS NOT NULL
5      AND sex= 'M'

```

Résultat: 78.89

Poids moyen des hommes

```

1      SELECT AVG(Participations.weight)
2      FROM Participations1984
3      JOIN Athletes USING (ano)
4      AND sex= 'M'

```

Résultat: 74.41

Comparaison: On observe que les hommes médaillés sont plus lourds que les hommes participants, de 4480 grammes en moyenne.

Comparaison: Soit une différence de 6%.

Femmes

Poids moyen des médaillées femmes

```

1      SELECT AVG(Participations.weight)
2      FROM Participations1984
3      JOIN Athletes USING (ano)
4      WHERE medal IS NOT NULL
5      AND sex= 'F'
  
```

Résultat: 62.16

Poids moyen des femmes

```

1      SELECT AVG(Participations.weight)
2      FROM Participations1984
3      JOIN Athletes USING (ano)
4      AND sex= 'F'
  
```

Résultat: 58.02

Comparaison: On observe que les femmes médaillées sont plus lourdes que les femmes participantes, de 4140 grammes en moyenne.

Comparaison: Soit une différence de 7,13%.

Observation: On observe plus ou moins la même tendance chez les deux genres, avec une différence étant sensiblement la même ; les médaillés sont plus lourds que les participants.

3.3 Partie 3

Sur la période 1992 à 2016, faire la liste des 15 pays qui ont gagné le plus de médailles cumulées sur cette période (jeux d'été et hiver confondus)

Top 15 régions par médailles

```

1  SELECT region, COUNT(medal) AS medals_number
2  FROM Participations
3  JOIN Delegations USING(noc)
4  JOIN Events USING(eno)
5  JOIN Games USING(gno)
6  WHERE Games.year BETWEEN 1992 AND 2016
7  GROUP BY region
8  ORDER BY medals_number DESC
9  LIMIT 15;

```

Résultat:

img/version3/top15.png

FIGURE 6 – Classement

Choisir 5 pays parmi ces 15 et, pour chacun d'eux, représenter l'évolution

Nous choisirons :

- ☐ Australie
- ☐ Chine (RPC)
- ☐ Italie
- ☐ Espagne
- ☐ Royaume-Uni

En s'appuyant sur la vue suivante, les requêtes seront raccourcies.

On ne va pas inclure le résultat à chaque fois pour ne pas surcharger, car l'étude statistique sera présente et inclura ces résultats.

Vue exclusive à 1992-2016 et aux 5 pays

```

1      CREATE VIEW ParticipationsPart3 AS (
2      SELECT Participations.*, year, region
3      FROM Participations
4      JOIN Delegations USING(noc)
5      JOIN Events USING(eno)
6      JOIN Games USING(gno)
7      WHERE Games.year BETWEEN 1992 AND 2016
8      AND region IN ('Australia', 'China', 'Italy', 'Spain', '
9      UK')
10     );

```

Représentation des évolutions

Les résultats des requêtes sont sauvegardées dans des fichiers CSV correspondants. Ces fichiers sont utilisés en entrée d'un petit programme Python (matplotlib) dessinant un nuage de points correspondant aux données regroupées par pays et par jeu. Ce programme calcule également une droite de régression afin de représenter l'évolution globale des données selon leur corrélation. Dans certains cas et selon les paramètres, nous aurons une extrapolation ou une interpolation et il en sera jugé en commentaire. Cette droite ne vaut pas pour vérité absolue et parfois elle ne sera pas fiable.

Il est à noter que la largeur des cases est différente car entre un Jeu d'Hiver et un Jeu d'Été, il y a 1 an et demi d'écart alors que dans le cas contrainte, il y a 2 ans et demi d'écart.

1. Du nombre de participants

Nombre de participants

```

1      SELECT region, games, COUNT(DISTINCT ano) AS contestants
2      _number
3      FROM ParticipationsPart3
4      GROUP BY region, games
5      ORDER BY games ASC, contestants_number DESC;

```



FIGURE 7 – Graphique d'évolution - B1

Observation: Une observation intéressante est de constater que lorsque le pays est hôte, comme l'Australie en 2000 à Sydney ou la Chine en 2008 à Pékin, il concentre beaucoup plus de participants (et c'est assez logique car en marge de ces événements beaucoup d'investissements sportifs sont lancés). En dehors de ces événements, on distingue peu d'évolution.

Observation: La régression est probablement fortement influencée par le fait des pays hôtes concentrant plus de participants à cette occasion. Désigner une évolution globale pour ces pays n'est donc pas totalement vrai à en voir ces écarts.

2. Du nombre de médaillés

Nombre de médaillés

```

1      SELECT region, games, COALESCE(COUNT(NULLIF(medal IS
2      NULL, true)), 0) AS medals_number
3      FROM ParticipationsPart3
4      GROUP BY region, games
5      ORDER BY games ASC, medals_number DESC;
```




FIGURE 8 – Graphique d'évolution - B2

Observation: Dire qu'il y a plus de médailles est une affirmation dépendant en réalité des pays que l'on traite et de l'évolution de leurs niveaux sportifs. L'Italie a relativement stagné alors que la Chine a beaucoup augmenté (en 2008 à Pékin notamment) ainsi que l'Australie (en 2000 à Sydney).

3. Du nombre de femmes participantes

Nombre de femmes participantes par pays et année

```

1      SELECT region, games, COUNT(DISTINCT ano) AS women_
        contestants_number
2      FROM ParticipationsPart3
3      JOIN Athletes USING(ano)
4      WHERE sex = 'F'
5      GROUP BY region, games
6      ORDER BY region ASC, games ASC, women_contestants_number
        DESC;
```

Résultat:

img/version3/nbfemmes.png

FIGURE 9 – Nombre de femmes participantes

img/version3/b3.png

FIGURE 10 – Graphique d'évolution - B3

Observation: Il y a aujourd'hui de plus en plus de femmes participantes ! Cependant, les jeux d'hiver ont moins participants, donc moins de participantes. Il faut veiller à bien analyser les points pour cette affirmation.

4. De la proportion de femmes participantes (en comparaison de la proportion d'hommes)

Proportion de femmes participantes par rapport aux hommes par pays et année

```

1      SELECT region, games, COUNT(CASE WHEN sex = 'F' THEN 1
2          END) * 1.0 NULLIF(COUNT(CASE WHEN sex = 'M' THEN 1
3          END), 0) AS women_over_men
4      FROM ParticipationsPart3
5      JOIN Athletes USING(ano)
6      GROUP BY region, games
7      ORDER BY region ASC, games ASC, women_over_men DESC;
```



FIGURE 11 – Graphique d'évolution - B4

Observation: Les points sont assez rapprochés ; nous pouvons ici compter sur une régression de qualité nous informant que le rapport femmes/hommes parmi les médaillés a pu augmenter de 0.71 (35% de femmes et 65% d'hommes) à 1.01 en moyenne (50% des deux côtés). La situation est, en 2016, dans les pays et selon l'évolution, égalitaire.

Observation: Les prochaines questions traiteront du nombre de personnes médaillées et non de médailles : j'ai cru avoir fait une erreur donc j'ai beaucoup regardé car les comptes ne suivaient pas mais en réalité, il s'agit souvent d'épreuves collectives (rugby, relais, football...). Par conséquent, il y a donc très souvent plus de personnes médaillées que de médailles

5. De la proportion de médaillées parmi les femmes

Proportion de médaillées parmi les femmes

```

1      WITH WomenParticipation AS (
2          SELECT region, games, COUNT(DISTINCT ano) AS total_women
3          FROM ParticipationsPart3
4          JOIN Athletes USING(ano)
5          WHERE sex = 'F'
6          GROUP BY region, games
7      ),
8
9      MedalsWomen AS (
10         SELECT region, games, COUNT(DISTINCT ano) AS medal_women
11         FROM ParticipationsPart3
12         JOIN Athletes USING(ano)
13         WHERE sex = 'F' AND medal IS NOT NULL
14         GROUP BY region, games
15     )
16
17     SELECT wp.region, wp.games, COALESCE(mw.medal_women *
18         100.0 NULLIF(wp.total_women, 0) 100.0, 0) AS
19         medal_women_proportion_over_women
20     FROM WomenParticipation wp
21     LEFT JOIN MedalsWomen mw USING(region, games)
22     ORDER BY region ASC, wp.games ASC, medal_women_
23         proportion_over_women DESC;

```



FIGURE 12 – Graphique d'évolution - B5

Observation: Les résultats sont très dispersés et on ne peut pas vraiment constater une vraie dynamique. La régression nous informe néanmoins que la proportion a quand même augmenté de 0.05 sur la période donnée.

Observation: Cependant, la régression n'est pas vraiment fiable car il y a trop d'écart entre les valeurs. Une bonne régression doit habituellement concentrer beaucoup de points sur sa ligne et les points sont trop éloignés pour le permettre. Dans cette situation, analyser l'évolution a peu de sens.

6. De la proportion de femmes parmi les médaillés

Proportion de femmes parmi les médaillées

```

1      WITH AllCombinations AS (
2      SELECT DISTINCT region, games
3      FROM ParticipationsPart3
4      ),
5
6      MedalsPeople AS (
7      SELECT region, games, COUNT(DISTINCT ano) AS count_medal
8      FROM ParticipationsPart3
9      WHERE medal IS NOT NULL
10     GROUP BY region, games
11     ),
12
13     MedalsWomen AS (
14     SELECT region, games, COUNT(DISTINCT ano) AS count_medal
15         _women
16     FROM ParticipationsPart3
17     JOIN Athletes USING(ano)
18     WHERE sex = 'F' AND medal IS NOT NULL
19     GROUP BY region, games
20     )
21
22     SELECT ac.region, ac.games, COALESCE(mw.count_medal_
23         women * 100.0 NULLIF(mp.count_medal, 0) 100.0, 0)
24         AS medal_women_proportion_over_global
25     FROM AllCombinations ac
26     LEFT JOIN MedalsPeople mp USING(region, games)
27     LEFT JOIN MedalsWomen mw USING(region, games)
28     ORDER BY ac.region ASC, ac.games ASC, medal_women_
29         proportion_over_global DESC;
    
```



FIGURE 13 – Graphique d'évolution - B6

Observation: Le rendu ressemble un peu au précédent, néanmoins on peut distinguer plus clairement les écarts entre les pays, avec moins de droites se coupant. La progression par année tend vers une égalité parfaite (0.50), avec pour ces pays et dans cette période, une fin démontrant que la proportion de femmes parmi les médaillés est égale à celle des hommes.

Observation: Les jeux d'hiver sont en réalité un peu le cadeau empoisonné de ce jeu de données. En effet, il y a 5 à 10 fois moins de participants, donc les résultats sont très sensibles au hasard (exemple : 8 participants, 0 médailles, donc 0 alors que même une médaille aurait fait 0.125). Ainsi, on se retrouve à compter les proportions d'hiver et d'été au même niveau alors qu'en réalité elles concernent des nombres de participants bien différents. Les proportions à étudier devraient être les jeux d'été car ayant plus de données. Cependant, si l'objet d'études est l'évolution uniquement, combiner les deux saisons n'est pas un problème.

Observation: Sur ces exemples, on peut globalement constater :

- Une hausse du niveau des femmes
- Une proportion devenant égale à celle des hommes
- Une meilleure représentation parmi les participants

Observation: En conclusion, ces évolutions peuvent avoir du sens selon la qualité et la corrélation des données entre elles. Sans ces paramètres, ces évolutions peuvent parfois être utiles, mais en y prêtant une grande attention tout de même. Dans tous les cas, on préfère une régression de qualité. Il faut aussi rappeler que ces évolutions ne valent que pour les pays sélectionnés et pour la période. La Chine a toujours eu plus d'inclusion féminine aux Jeux, donc l'ensemble est un peu tiré vers le haut de ce fait.

4 Pour aller plus loin

Ici, nous allons plus loin que le sujet !

Nous nous baserons ici sur tous les pays.

1. L'impact du changement de poids des athlètes

J'aurais aimé faire une requête concernant les athlètes qui ont changé de poids et l'impact de ce changement sur leur poids, or la base n'admet qu'un seul poids unique pour chaque athlète. Il en va de même pour la taille (qui ne peut logiquement pas changer après 20 ans).

2. L'âge moyen des médaillés d'un sport

On pense souvent qu'il faut être jeune dans le sport. Ici nous analyserons quelles disciplines et quels sports demandent une préparation dès la jeunesse.

Par "discipline"

Age moyen des médaillés par discipline

```

1      SELECT event, coalesce(ROUND(AVG(age),2),0) as average_
      age
2      FROM Participations
3      JOIN Events USING(eno)
4      JOIN Disciplines USING(dno)
5      WHERE medal IS NOT NULL
6      GROUP BY event
7      ORDER BY average_age desc;
```

Résultat:

img/version4/disciplinerange.png

FIGURE 14 – Age moyen des médaillés par discipline

Par "sport"

Age moyen des médaillés par sport

```
1      SELECT sport, coalesce(ROUND(AVG(age),2),0) as average_
      age
2      FROM Participations
3      JOIN Events USING(eno)
4      JOIN Disciplines USING(dno)
5      WHERE medal IS NOT NULL
6      GROUP BY sport
7      ORDER BY average_age desc;
```

Résultat:

img/version4/sportrange.png

FIGURE 15 – Age moyen des médaillés par sport

Observation: Les sports avec les médaillés les plus âgés en moyenne sont des sports étant plus relatifs à une longue expérience.

3. Le physique fait-il la différence ?

Avons-nous réellement plus de chances d'être médaillé si nous avons le corps parfait pour notre sport ?

Le physique des médaillés

Physique des médaillés par sport

```

1 CREATE VIEW BodyMedalStats AS (
2   SELECT sport, COALESCE(AVG(height),0) AS avg_height,
3           COALESCE(AVG(weight),0) AS avg_weight
4   FROM Participations
5   JOIN Events USING(eno)
6   JOIN Disciplines USING(dno)
7   WHERE medal IS NOT NULL
8   GROUP BY sport
9 );

```

Résultat:

img/version4/bodymedal.png

FIGURE 16 – Physique des médaillés par sport

Le physique des non médaillés

Physique des non-médailleurs par sport

```

1 CREATE VIEW BodyNoMedalStats AS (
2   SELECT sport, COALESCE(AVG(height),0) AS avg_height,
3           COALESCE(AVG(weight),0) AS avg_weight
4   FROM Participations
5   JOIN Events USING(eno)
6   JOIN Disciplines USING(dno)
7   WHERE medal IS NULL
8   GROUP BY sport
9 );

```

Résultat:

img/version4/bodynomedal.png

FIGURE 17 – Physique des non-médailleurs par sport

Les plus grands favorisés

```

1      SELECT sport, b1.avg_height - b2.avg_height AS height_
      diff
2      FROM BodyMedalStats b1
3      JOIN BodyNoMedalStats b2 USING(sport)
4      WHERE b1.avg_height - b2.avg_height > 0
5      ORDER BY height_diff DESC;

```

Résultat:

img/version4/heightplus.png

FIGURE 18

Les plus petits favorisés

```

1      SELECT sport, b1.avg_height - b2.avg_height AS height_
      diff
2      FROM BodyMedalStats b1
3      JOIN BodyNoMedalStats b2 USING(sport)
4      WHERE b1.avg_height - b2.avg_height < 0
5      ORDER BY height_diff DESC;

```

Résultat:

img/version4/heightneg.png

FIGURE 19

Les plus lourds favorisés

```

1      SELECT sport, b1.avg_weight - b2.avg_weight AS weight_
      diff
2      FROM BodyMedalStats b1
3      JOIN BodyNoMedalStats b2 USING(sport)
4      WHERE b1.avg_weight - b2.avg_weight > 0
5      ORDER BY weight_diff DESC;
```

Résultat:

img/version4/weightplus.png

FIGURE 20

Les plus maigres favorisés

```

1      SELECT sport, b1.avg_weight - b2.avg_weight AS weight_
      diff
2      FROM BodyMedalStats b1
3      JOIN BodyNoMedalStats b2 USING(sport)
4      WHERE b1.avg_weight - b2.avg_weight < 0
5      ORDER BY weight_diff DESC;

```


Résultat:



img/version4/weightneg.png

FIGURE 21

Les résultats sont-ils cohérents ?

Observation: On observe sans surprise que le basketball ou la nage ont des médaillés en moyenne plus grands. L'inverse vaut également pour la gymnastique. Les autres différences peuvent également être sociologiques étant dû au mode de vie des médaillés, notamment pour le poids, avec des médaillés de tennis de tables en moyenne 1.73kg moins lourds que les non médaillés. Néanmoins aux Jeux quelqu'un de non médaillé reste tout de même un grand sportif qui a su se hisser aux plus grandes compétitions de son sport, donc il est déjà plus probable qu'il satisfasse les critères physiques de sa discipline que n'importe qui (Basketball : les non-médaillés font 1m90 en moyenne).

Merci d'avoir consulté cette présentation