

Compte-Rendu du projet d'Analyse Exploratoire

Des médailles aux JO !

Duc Hau NGUYEN; Anaïs RABARY

4 Décembre 2018

Contents

| | |
|---|----------|
| Introduction | 2 |
| Contexte | 2 |
| Problématique | 2 |
| 1. Premier axe: Les caractéristiques physiques des athlètes influencent leur performance | 2 |
| 1.1 L'âge, gage de réussite | 2 |
| 1.2 La taille compte-elle vraiment ? | 4 |
| 2. Quelle est la place des femmes dans les Jeux Olympiques ? | 5 |
| 2.1 La montée fulgurante des femmes | 5 |
| 2.2 Les femmes et l'âge de la sagesse sportive | 7 |
| Conclusion | 8 |
| References | 8 |

```
## # A tibble: 271,116 x 15
##       ID Name Sex   Age Height Weight Team NOC   Year_Games Season
##   <int> <chr> <fct> <int>  <int>  <dbl> <chr> <fct> <chr>    <chr>
## 1     1  1 A Di~ M     24    180    80 China CHN   1992    Summer
## 2     2  2 A La~ M     23    170    60 China CHN   2012    Summer
## 3     3  3 Gunn~ M     24     NA    NA Denm~ DEN   1920    Summer
## 4     4  4 Edga~ M     34     NA    NA Denm~ DEN   1900    Summer
## 5     5  5 Chri~ F     21    185    82 Neth~ NED   1988    Winter
## 6     6  5 Chri~ F     21    185    82 Neth~ NED   1988    Winter
## 7     7  5 Chri~ F     25    185    82 Neth~ NED   1992    Winter
## 8     8  5 Chri~ F     25    185    82 Neth~ NED   1992    Winter
## 9     9  5 Chri~ F     27    185    82 Neth~ NED   1994    Winter
## 10    10  5 Chri~ F     27    185    82 Neth~ NED   1994    Winter
## # ... with 271,106 more rows, and 5 more variables: Year <int>,
## #   City <chr>, Sport <fct>, Event <chr>, Medal <fct>
```

Introduction

Contexte

Les **Jeux Olympiques** (JO) sont une compétition que nous avons hérité de la Grèce Antique. On lit souvent que seuls les citoyens les plus riches pouvaient y participer. La trêve olympique permettait à la civilisation d'oublier, le temps de 12 travaux, les guerres et troubles de l'époque. Après quelques maigres tentatives, c'est seulement depuis 1896 que les Jeux Olympiques modernes voient le jour, à Athènes.

Dans le cadre du projet d'**Analyse Exploratoire** à l'**INSA Toulouse**, avec l'accompagnement de professeur **Gilles Tredan**, nous avons récupéré sur le site Kaggle un dataset sur ces JO moderne (Sport Reference 2018). Ce dernier recense des informations sur tous les athlètes ayant participé aux JO, les médailles qu'ils ont gagnés, leur équipe, l'année de participation, etc.

Sauf précision contraire, les données sont exploitées ici dans leur totalité, depuis la création des Jeux Olympiques modernes (1896). Suivant l'analyse, les données incomplètes ont été traitées différemment. Parfois elles ont été mise à 0. D'autre fois, on a préféré les ignorer.

Problématique

Au delà du caractère spectaculaire et sportif, les Jeux Olympics sont un sujet intéressant pour les scientifiques : Le dernier siècle a prédit que l'humain atteindra bientôt ses limites physiques, tandis que l'on voit toujours aujourd'hui des sportifs continuer d'établir de nouveaux records.

Avec plusieurs centaines d'années d'existence des JO, des facteurs optimisant les performances des athlètes ont été identifiés. L'historique des JO pourraient surement fournir une indication à la question: Quels sont les facteurs qui influent/optimisent la performance physique des athlètes?

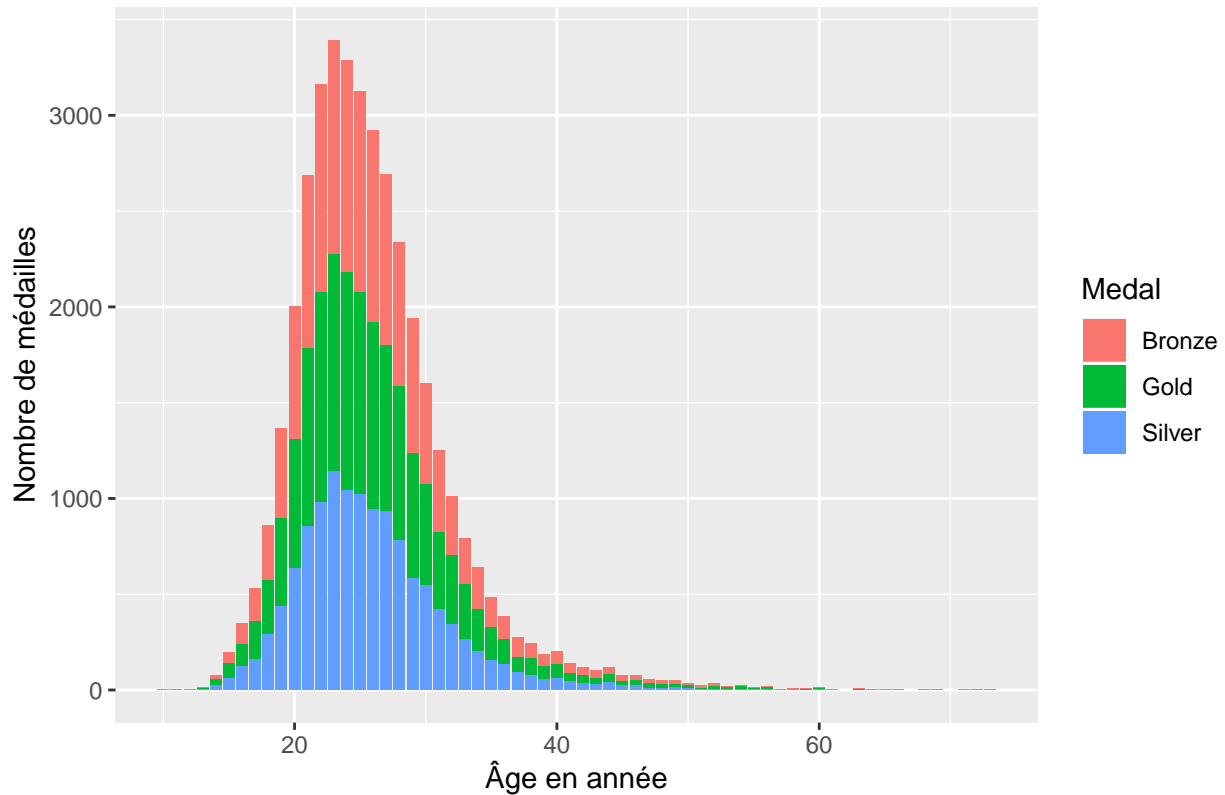
On peut aussi se demander quelle est la place de la femme, encore trop souvent controversé dans notre société actuelle.

1. Premier axe: Les caractéristiques physiques des athlètes influencent leur performance

1.1 L'âge, gage de réussite

L'approche naïve permet ici de présenter le nombre de médailles que chaque tranche d'âge a pu obtenir, quelque soit l'année des Jeux. La performance est mesurée ici en fonction du nombre de médailles gagnées ainsi qu'en fonction de la "couleur" de la médaille. C'est en effet un bon indicateur pour discriminer les performances entre les athlètes sur le podium et ceux n'ayant pas remporté de médaille.

Qualité des médailles obtenues selon l'âge des athlètes

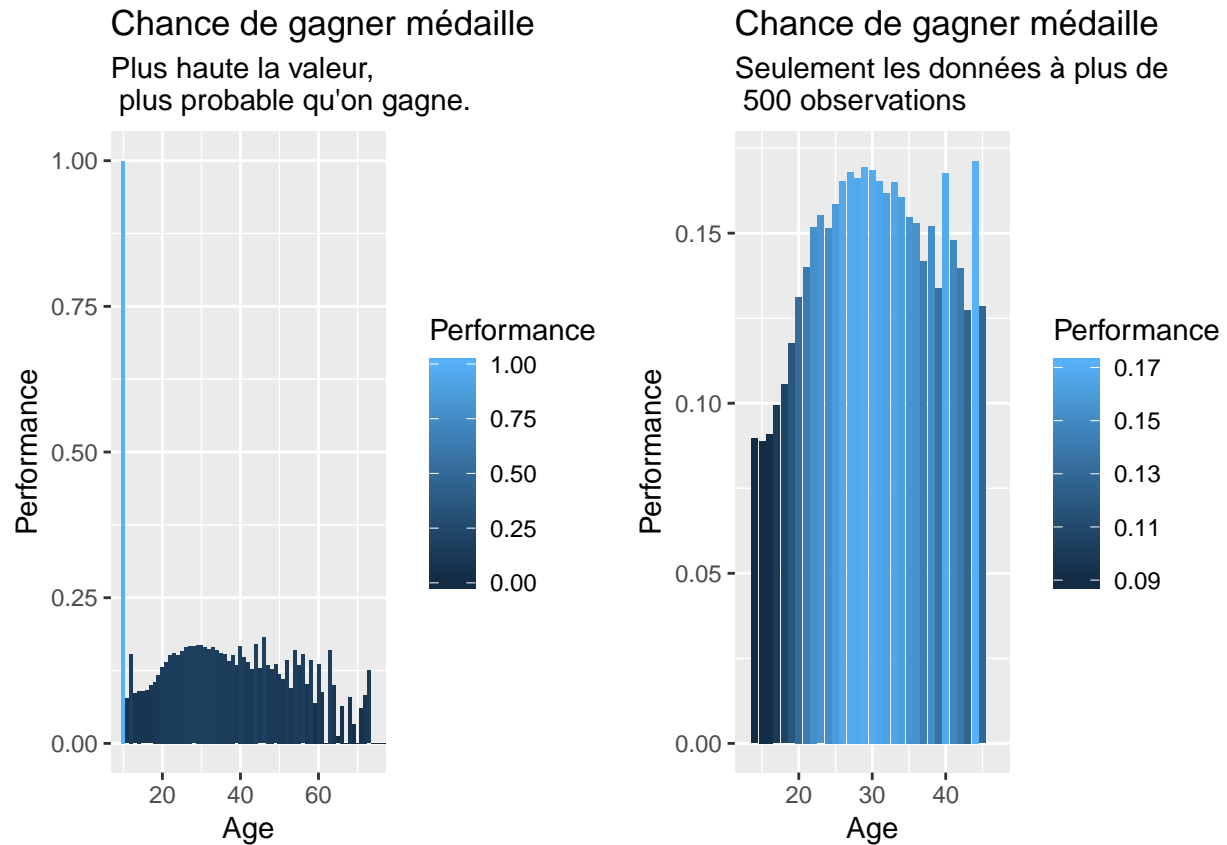


D'après le résultat, on constate que le pic de maturité des sportifs est atteint entre 20 et 27 ans. Cependant, cette déduction est faite sur l'hypothèse pour laquelle le nombre de participants n'entre pas en jeu. En effet, il peut arriver que l'on ait beaucoup de participants âgés de 20 à 27 ans, mais que seulement une faible proportion d'entre eux aient gagné, alors que peut-être ceux qui sont plus âgés sont moins nombreux à participer mais ont un ratio de médailles plus élevé.

Pour vérifier notre hypothèse, on calcule la "Performance" des participants. L'indice de performance détermine la probabilité qu'un participant figurant dans une catégorie de gagner une médaille ou pas. Formellement:

$$Performance = \frac{\#Medal}{\#Participant} = \frac{\sum(Gold+Silver+Bronze)}{\sum(all)}$$

Cette formule n'est pas adaptée lorsque l'on a un seul participant. Car s'il gagne une médaille, la probabilité sera de 100%. On peut voir dans la figure à gauche un pique à l'âge de 10: le seul participant à cet âge **Dimitrios Loundras** en 1896 qui a gagné 1 médaille.

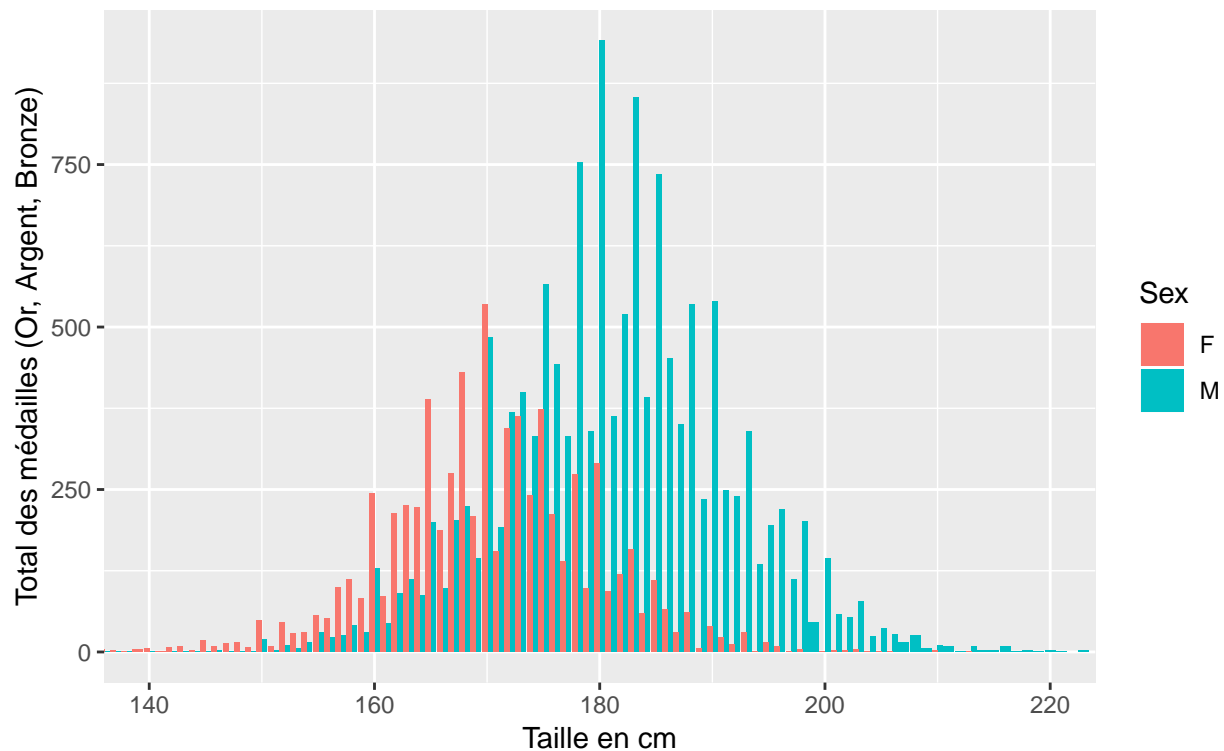


Dans le diagramme à droit, en filtrant les catégories dont le nombre d'observation n'est pas "suffisant", on voit que l'âge idéale pour gagner de médaille est à la fin de 20 devant 30 ans.

1.2 La taille compte-elle vraiment ?

A travers l'étude suivante, on veut essayer de comprendre si la taille des athlètes est corrélées avec leur performance. On présente ici les médailles gagnées par les athlètes féminines et masculins en fonction de leur taille.

Nombre de médailles gagnées par les athlètes suivant leur taille et leur sex Depuis le début des JO



Comme on pouvait s'y attendre, le nombre de médailles suivant la taille a une allure de distribution gaussienne. La taille des hommes où le pic du nombre de médailles gagnées est le plus important est à 1m80 pour 941 médailles. Chez les femmes, le pic est à 1m70, pour 535 médailles. Ces tailles correspondent bien à des tailles normales pour des hommes et des femmes. D'autre part, on visualise, par l'allure des courbes par sexe, que les femmes sont généralement plus petites que les hommes

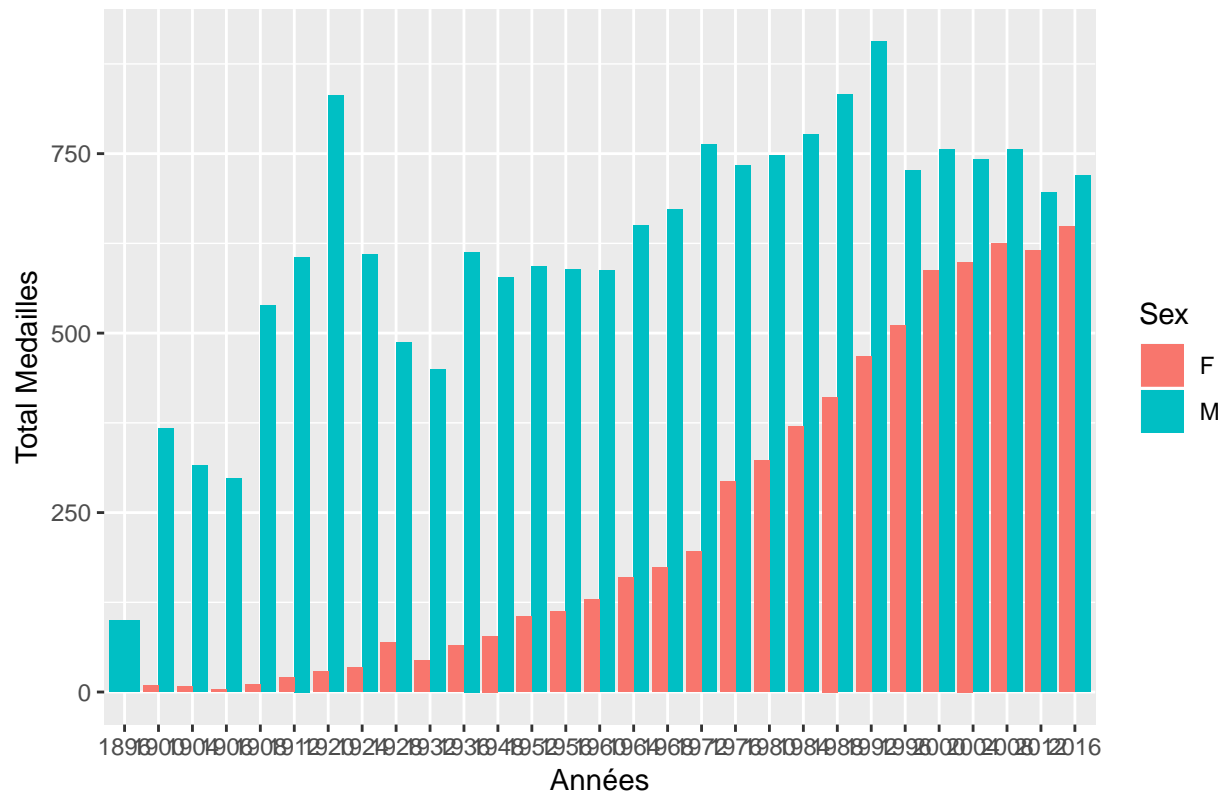
2. Quelle est la place des femmes dans les Jeux Olympiques ?

A travers notre jeu de données, on peut noter que les Femmes ne sont autorisées à participer aux JO qu'à partir de la 7ème édition, en 1920. Jusqu'alors, seuls des hommes concouraient.

2.1 La montée fulgurante des femmes

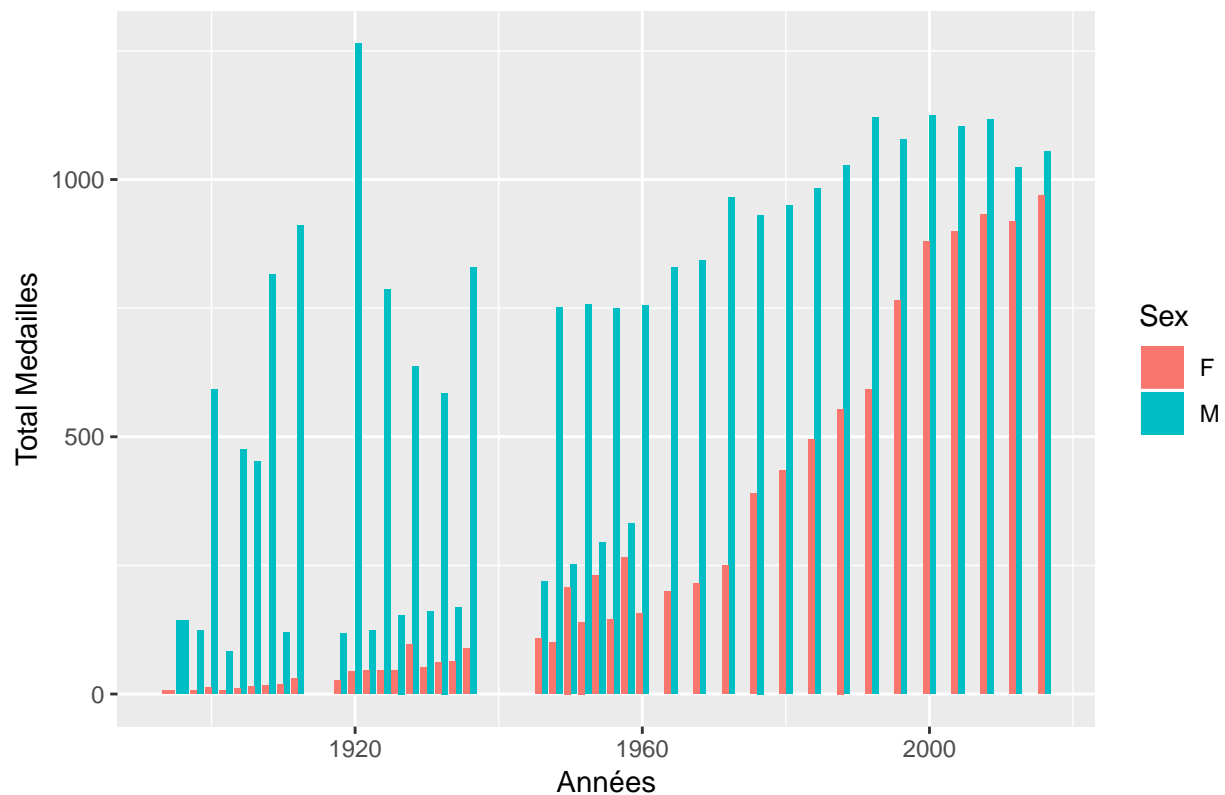
Dans notre société actuelle, on aborde la place de la femme dans tous les domaines. On aborde surtout le problème de la différence de rémunération. Dans le graphe suivant, on peut observer l'évolution du nombre de médailles remportés respectivement par les hommes et les femmes, au cours des sessions olympiques.

Evolution du nombre de médailles gagnées par sexe



TODO: Anais check out dans le code

Evolution du nombre de médailles gagnées par sexe



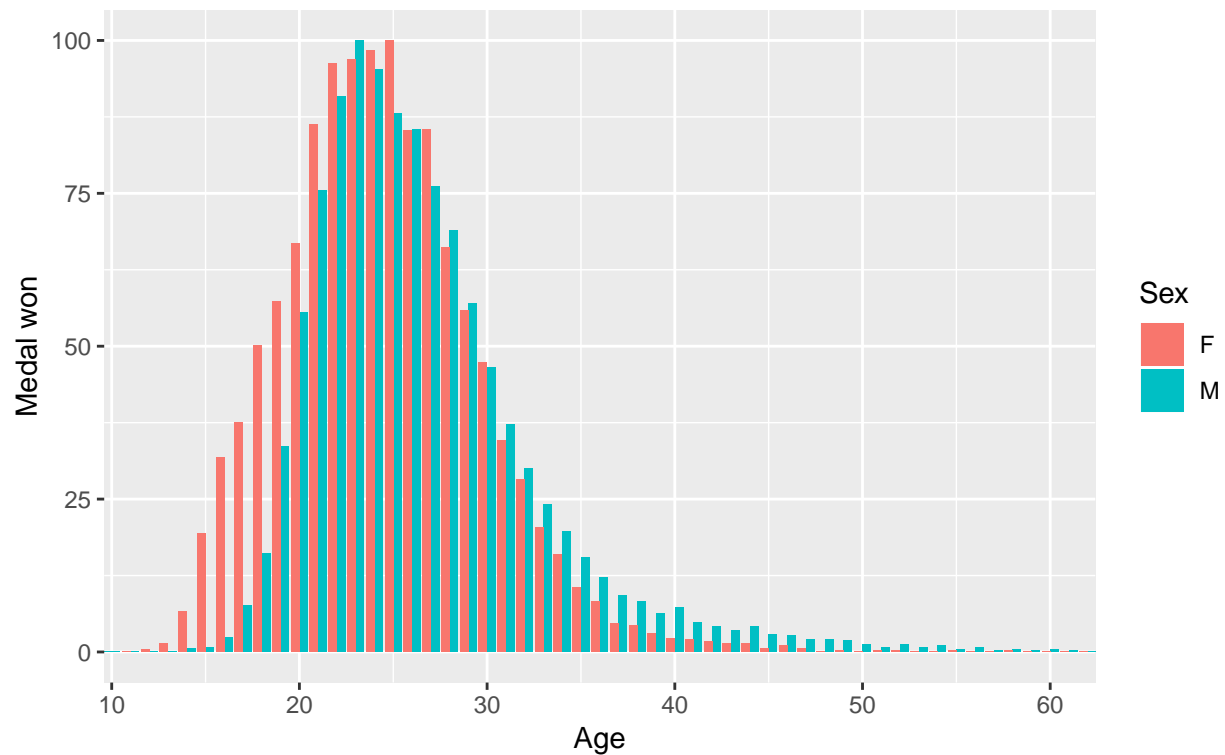
On retrouve l'absence de femmes aux JO avant les années 1920. Mais ensuite, on remarque surtout la progression réalisée en terme de nombre de médailles gagnées par les femmes. Cela correspond à une évolution positive de la place de la femme notre société. Une femme pouvant participer et s'affirmer dans une compétition sportive est un signe de liberté et d'égalité. On notera aussi l'augmentation croissante de la proportion de femmes participants parmi les athlètes. De 3% en 1920, les femmes représentent 18% des athlètes en 1960, 39% en 2000. En 2016, elles représentaient 45% de la population d'athlètes.

2.2 Les femmes et l'âge de la sagesse sportive

Projetons cette axes sur une autre dimension: Sex. On voit que, l'âge efficace est différent entre homme et femme: la condition physique des femmes maturise plus tôt que les hommes:

Nombre de médaille gagnant

Pour la raison de visibilité, cette valeur est standardisé entre 0 et 100



Conclusion

Axes d'ouverture : Richesse des pays

References

Sport Reference. 2018. "120 Years of Olympic History: Athletes and Results." Kaggle. <https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>.