

Des médailles aux JO !

Compte-Rendu du projet d'Analyse Exploratoire

Duc Hau NGUYEN; Anaïs RABARY

7 Décembre 2018

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 modernes (Sport Reference 2018). Ce dernier recense des informations sur tous les athlètes ayant participés aux JO, les médailles qu'ils ont gagnées, 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 mises à 0. D'autre fois, on a préféré les ignorer.

Problématique

Au-delà du caractère spectaculaire et sportif, les Jeux Olympiques sont un sujet intéressant pour les scientifiques : Le dernier siècle avait prédit que l'Homme atteindrait bientôt ses limites physiques. Pourtant, on voit toujours aujourd'hui des sportifs 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/optiment la performance physique des athlètes?

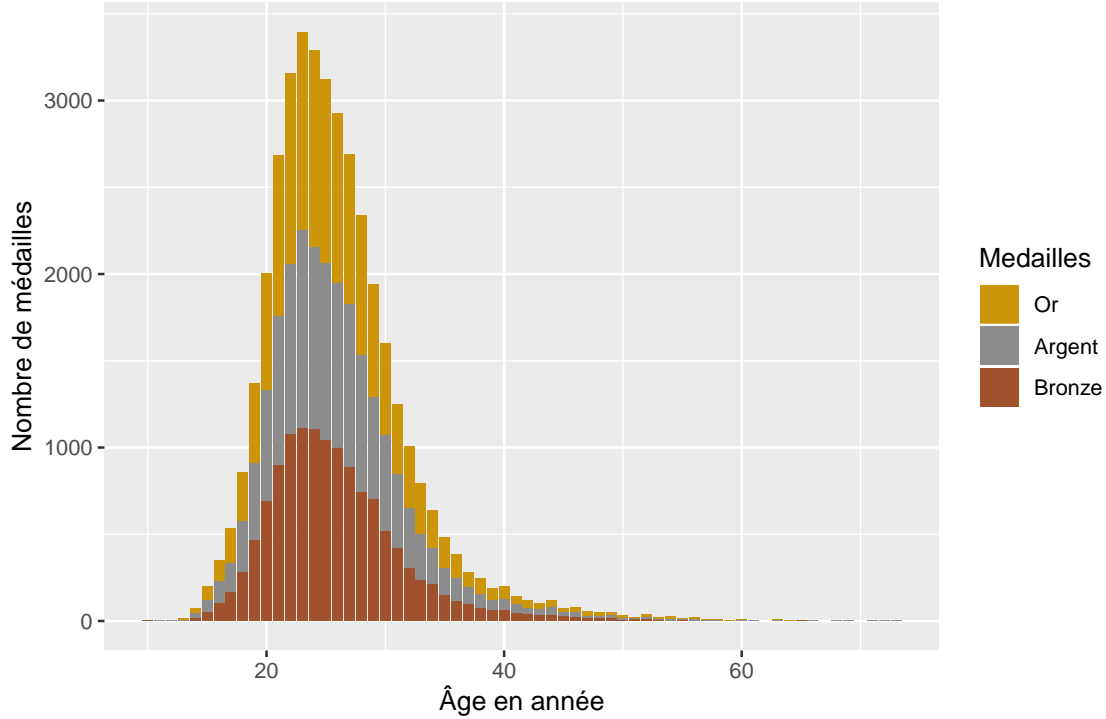
On peut aussi se demander quelle est la place de la femme, encore trop souvent controversée dans notre société actuelle.

1. 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, quel que soit l'année des Jeux. La performance est mesurée 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.

Fig.1 : Qualité des médailles obtenues selon l'âge des athlètes



D'après ce graphe, on peut considérer que la maturité sportive est atteinte par les athlètes entre 20 et 27 ans. La maturité sportive est définie ici comme étant la tranche d'âge pour laquelle le nombre de médailles remportées est le meilleur. Cependant, cette déduction est réalisée sans prendre en compte le nombre de participants. Cela n'est donc pas suffisant pour représenter la performance des groupes d'athlètes. 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 qu'il se peut que des athlètes plus âgés soient moins nombreux à participer mais aient un ratio de médailles plus élevé.

Pour vérifier ce cas de figure, on calcule la "Performance" des participants. Cet indice calcule, par tranche d'âge, la somme des médailles gagnées sur le nombre total d'athlètes. Formellement :

$$Performance_{Age} = \frac{\#Medals_{Age}}{\#Participants_{Age}} = \frac{\sum (Gold+Silver+Bronze)_{Age}}{\sum (all)_{Age}}$$

Cette formule n'est pas tout le temps représentative d'une tranche d'âge. Par exemple, on retrouve une performance de 100% pour les athlètes âgés de 10 ans. Or, dans notre dataset, nous avons un seul participant âgé de 10 ans, **Dimitrios Loundras** en 1896 et il a remporté une médaille. Mais cela ne veut pas dire que si nous avions N athlètes âgés de 10 ans, ils remporteraient N médailles et donc que le groupe aurait une performance de 100%.

On décide donc de filtrer ce cas, considéré comme un bruit. Dans le diagramme de droite ainsi recentré, on voit que l'âge idéal pour gagner une médaille est d'avoir dépassé 25 ans mais sans entamer la trentaine.

Fig.2 : Probabilité de gagner une médaille suivant l'âge du participant

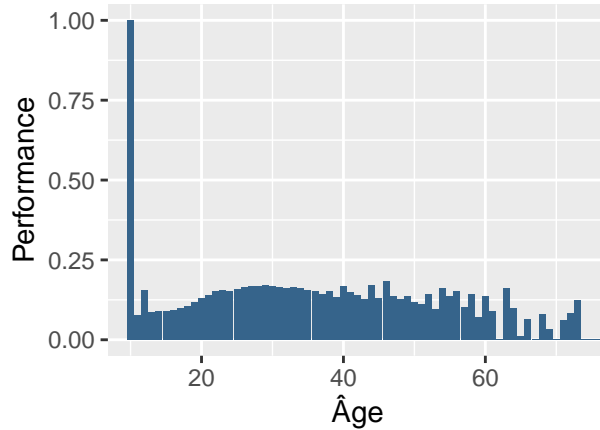
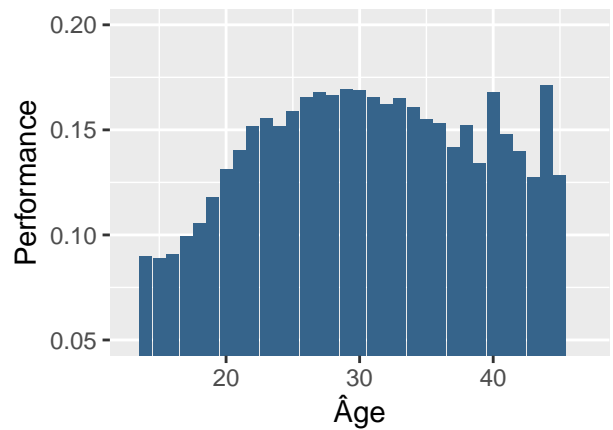


Fig.3 : Probabilité de gagner une médaille

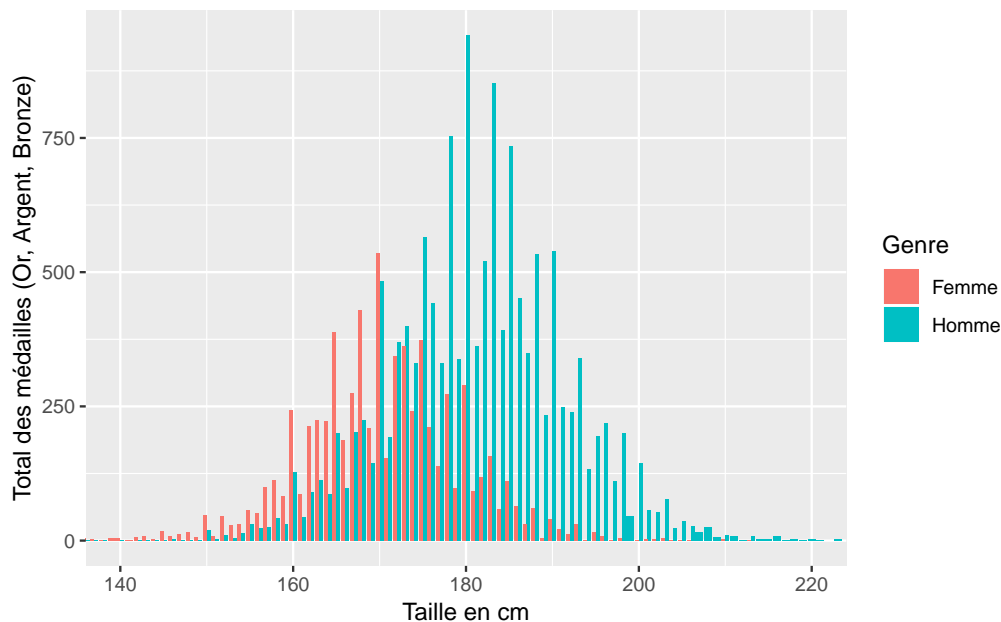
Plus de 500 observations



1.2 La taille compte-elle vraiment ?

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

Fig.4 : Nombre de médailles gagnées par les athlètes suivant leur taille et leur genre 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 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 de chaque sexe, que les femmes sont généralement plus petites que les hommes. C'est encore un fait que l'on peut constater dans notre société.

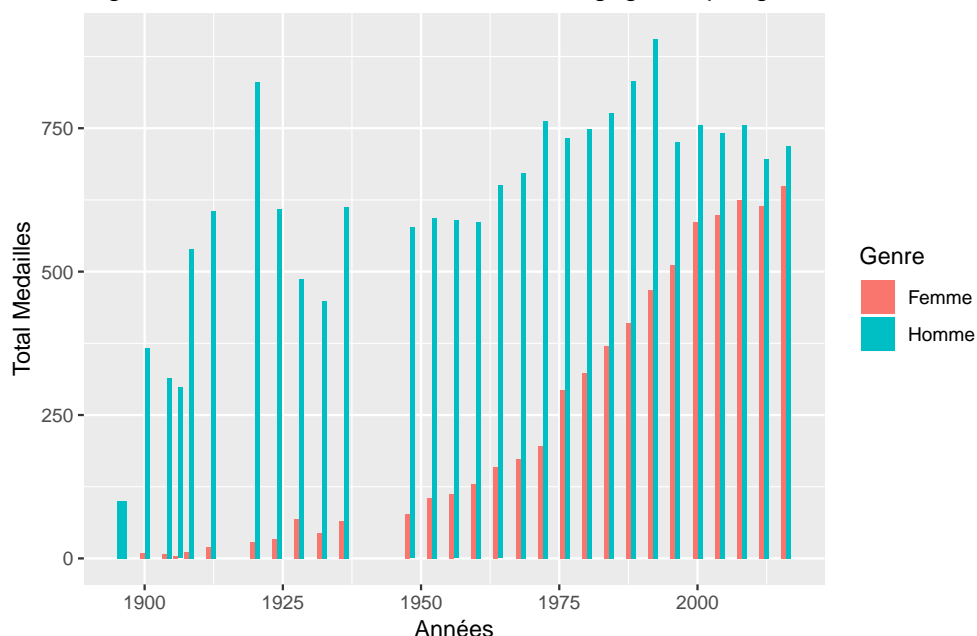
2. Quelle est la place des femmes dans les Jeux Olympiques ?

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 concourraient.

2.1 La montée fulgurante des femmes

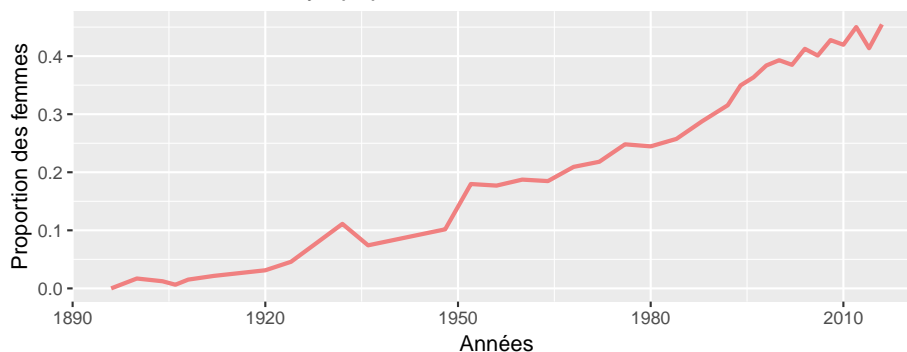
Dans notre société actuelle, on parle de la place de la femme dans tous les domaines. On aborde surtout le problème de la différence de rémunération. Dans les JO, la rémunération est calculée suivant les médailles gagnées. Dans le graphe suivant, on peut observer l'évolution du nombre de médailles remportées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

Fig.5 : Evolution du nombre de médailles gagnées par genre



On retrouve l'absence de femmes aux JO avant les années 1920. Mais ensuite, on remarque surtout la progression réalisée en termes de nombre de médailles gagnées par les femmes. Cela correspond à une évolution positive de la place de la femme dans 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 participantes 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.

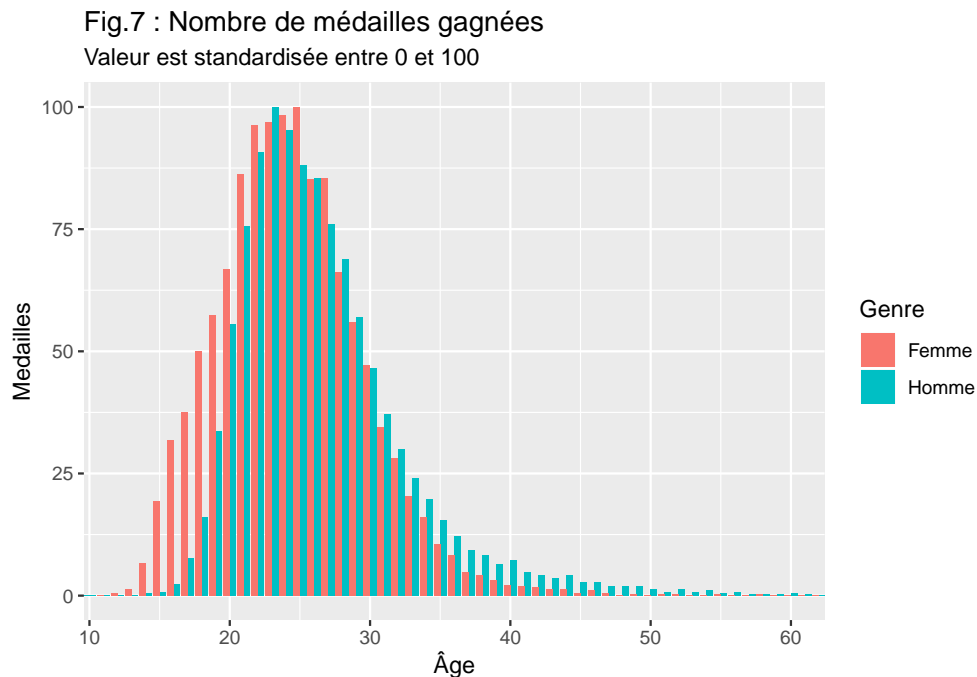
Fig.6 : Evolution dans les temps de la proportion des femmes parmi les athlètes des Jeux Olympiques



2.2 Les femmes et l'âge de la maturité sportive

Revenons sur l'étude réalisée dans la partie 1.1, mais cette fois-ci portée sous un angle féminin. On considère le nombre total de médailles gagnées, selon l'âge des participants, mais en distinguant les résultats des athlètes féminines et masculins. On voit que les femmes, entre 12 et 25 ans, gagnent plus de médailles que les hommes, sauf à 23 ans).

Reprenant le terme de maturité sportive, on peut en déduire que les femmes sont plus matures que les hommes pour cette tranche d'âge. Mais alors, y aurait-il un lien entre cette maturité sportive et la maturité scolaire ? En effet, on dit souvent que les filles à l'école sont plus matures scolairement, que les garçons. Surement car elles sont plus exigeantes dans le travail. Peut-être est-ce un constat équivalent dans leur exigence sportive. Mais cette analyse ne permet pas de conclure à ce sujet.



Conclusion

Les résultats des jeux olympiques modernes sont une ressource intéressante pour déterminer plusieurs caractéristiques. On constate qu'être jeune et donc en bonne forme physique, est un atout pour remporter une médaille. La taille n'est pas un facteur clé à première vue. Mais en analysant les tailles des athlètes suivant les sports, on trouverait peut-être un schéma particulier.

Ce qui est surtout frappant à travers ce dataset, c'est l'évolution de la place des femmes dans le monde sportif des Jeux Olympiques. Elle reflète l'évolution de la place de la femme dans notre société mondiale, même si les efforts mis en place par les JO sont parfois en avance sur la situation dans certaines régions du monde. On pourrait se demander aussi si la richesse d'un pays influe sa capacité de remporter des médailles aux jeux. Ou si les guerres et conflits internationaux se reflètent dans les résultats.

Références

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>.