

# Rapport domaine d'application

## Projet d'analyse approfondie de la relation âge-performance

Par Louise THOMAS et Jeanne FOND

# Sommaire

Sommaire	1
<b>Introduction</b>	<b>2</b>
Présentation des bases de données	2
Athlétisme	2
Echecs	2
<b>Structurer les données</b>	<b>3</b>
Calcul de l'âge	3
Calcul de la vitesse	3
Création de bases de données	3
Graphiques performances maximales par âge	4
<b>Ajustement par modèles</b>	<b>6</b>
Modèle linéaire	6
Modèle polynomial	6
Equation de Moore	7
<b>Critères d'ajustement des modèles</b>	<b>8</b>
Résidus	8
Coefficient de détermination ajusté	8
AIC	8
BIC	8
Conclusion	8
<b>Pic de performance</b>	<b>9</b>
Résultats	9
Conclusion sur le pic de performance	9
<b>Relation âge-performance aux échecs</b>	<b>10</b>
Calcul de l'âge	10
Création base de données	10
Graphique performances maximales par âge	10
Ajustement par modèles	11
Modèle linéaire	11
Modèle polynomial	11
Equation de Moore	12
Conclusion ajustement des modèles	12
Critères d'ajustement des modèles	13
Résultats	13
Conclusion sur les critères d'ajustement	13
Pic de performance	13
<b>Conclusion générale</b>	<b>14</b>
<b>Question bonus n°2</b>	<b>15</b>
<b>Annexe</b>	<b>16</b>

# Introduction

Dans le cadre de notre projet SAE Domaine d'application, nous avons réalisé une analyse approfondie, sur R, de la relation entre l'âge et la performance sportive. Dans un premier temps, nous avons examiné cette relation à travers diverses disciplines de l'athlétisme, puis nous l'avons étendue aux résultats dans le domaine des échecs. Pour mener à bien cette étude, nous avons eu accès à trois bases de données : les deux premières contenant les résultats masculins et féminins de huit épreuves d'athlétisme, tandis que la troisième portait sur les performances des joueurs d'échecs.

## Présentation des bases de données

### Athlétisme

Concernant les données liées à l'athlétisme, nous avons 2 bases de données, l'une pour les femmes et l'autre pour les hommes. Celle des hommes contient 132 569 observations contre 122 319 pour celle des femmes. Les deux bases sont structurées pareils, elles contiennent 12 variables avec des informations sur les athlètes et les compétitions réalisées. Pour notre étude, nous n'avons pas besoin de toutes ces informations ainsi nous ne garderons que 4 variables : Mark (le temps réalisé), DOB (la date de naissance), Date (la date de l'épreuve) et Dis (la discipline).

### Echecs

La base concernant les échecs contient 139 119 observations et 7 variables portant les informations de différents joueurs d'échecs et leurs résultats. Pour notre étude, nous n'utilisons pas toutes les informations, nous gardons 4 variables : Année (année du résultat du joueur), Mois (mois du résultat du joueur), Points (résultat du joueur) et Date.naissance (date de naissance).

# Structurer les données

## Calcul de l'âge

Nous allons calculer l'âge des athlètes lors de leur compétition. Pour cela, nous traitons les colonnes "DOB" et "Date". Premièrement, nous supprimons les individus ayant des cases vides dans l'une des deux variables car ces derniers ne sont pas utilisables pour notre problématique age-performance. Deuxièmement, nous modifions les dates étant incomplètes. Nous avons deux cas, les dates n'ayant pas de jour, pour lesquelles nous leur assignons le milieu du mois c'est-à-dire le 15, et celles ayant juste l'année à qui nous assignons le milieu de l'année soit le 02 juillet. Une fois que tous les individus des deux variables ont tous une date complète et au même format, nous les transformons au format date de R. Nous pouvons à présent procéder au calcul de l'âge. Pour cela nous soustrayons la date de naissance à la date de compétition et nous divisons le résultat par 365,2. Nous réalisons ces opérations sur les 2 dataframes.

## Calcul de la vitesse

Afin de calculer la vitesse nous devons traiter les colonnes "Mark" (temps réalisé) et "Dis" (Discipline).

Premièrement, nous nous occupons de la variable "Mark". Nous devons formater cette colonne afin d'obtenir chaque durée en seconde. Nous avons trois cas de figure dans notre variable : des observations en hh:mm:ss, en mm:ss ou en mm:ss.ms. Nous avons donc réalisé 3 fonctions, une pour chaque cas. Le principe de ces fonctions est de séparer les chiffres en fonction des ":" ou "." et de multiplier les heures par 3600, les minutes par 60 et de diviser les millisecondes par 100 afin d'obtenir des temps en seconde.

Deuxièmement, une fois notre nouvelle variable "resultat\_seconde" créée, nous nous occupons du traitement de la variable "Dis". Dans cette variable, nous avons la distance et le type d'épreuve séparé par un "/". Dans les distances, nous avons 3 cas de figure : une distance en mètre, une distance en kilomètre et le marathon et semi-marathon. Dans le cas d'une distance en mètre, nous n'avons qu'à retirer le mot clé "metres". Dans le cas d'une distance en kilomètre, il nous faut retirer le mot clé "kilometres" et multiplier la valeur par 1000. Dans le cas du semi-marathon et du marathon, nous assignons aux athlètes la distance en mètre correspondante. Nous obtenons donc une nouvelle variable "distance" contenant la distance en mètre.

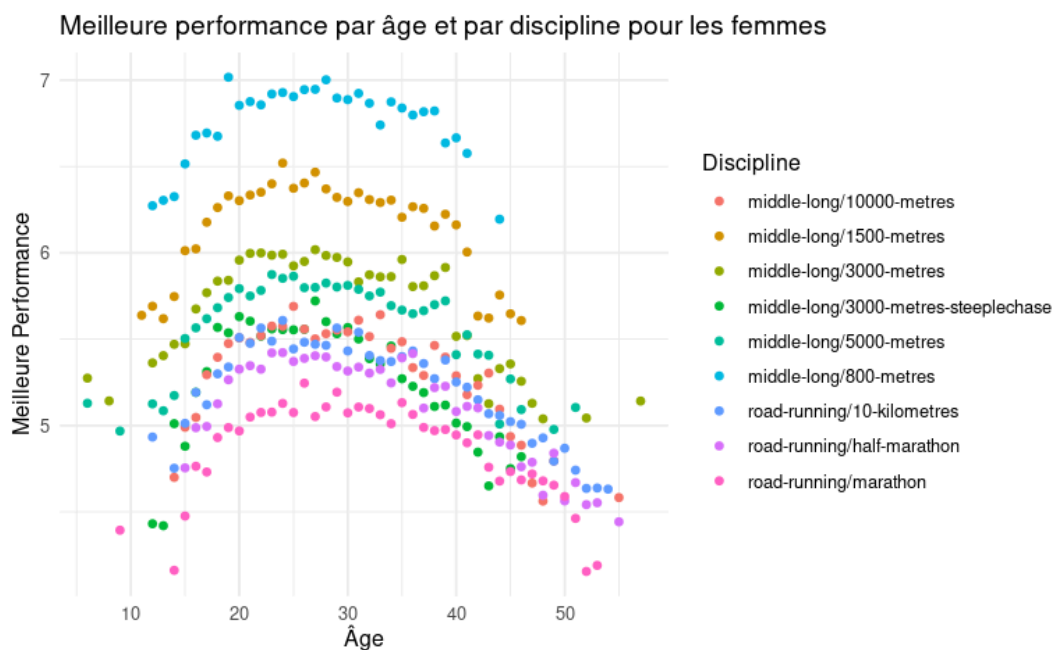
Nous pouvons maintenant, pour finir, calculer la vitesse réalisée. Pour cela, nous divisons la variable "distance" par la variable "resultat\_seconde" pour obtenir la variable "vitesse".

## Création de bases de données

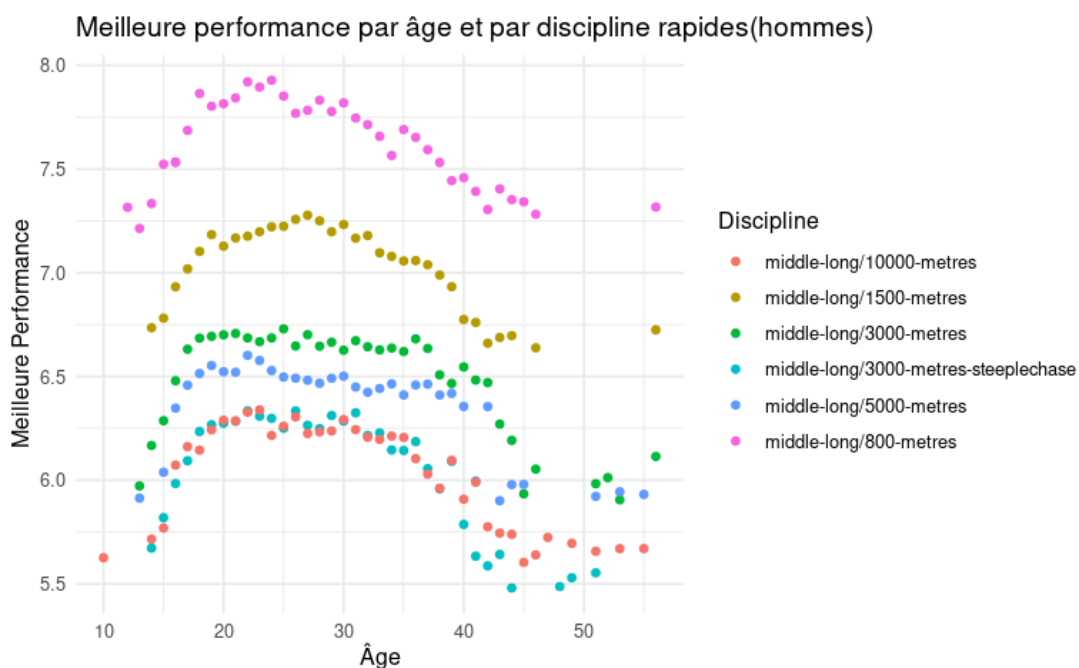
Pour chaque discipline, nous créons une base contenant la meilleure vitesse par âge (entier). Nous obtenons donc 8 bases par genre soit 16 au total. Nous avons décidé de regrouper les résultats des disciplines 10km et 10000m. En effet, nous avons au préalable calculé les moyennes de ces deux disciplines, étant très similaires, nous n'avons fait qu'un dataframe commun. Nous avons par ailleurs séparé les disciplines 3000 m haies et 3000 m.

## Graphiques performances maximales par âge

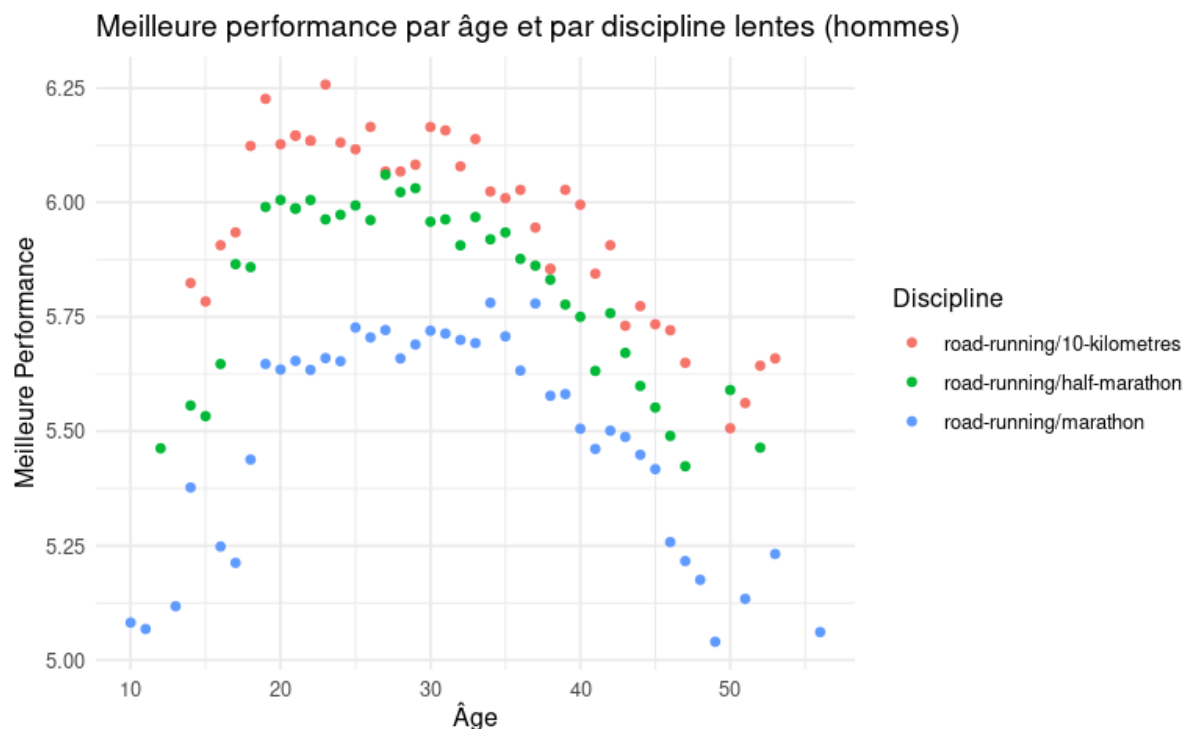
Ci-dessous se trouve le graphique de la meilleure performance par âge pour chaque épreuve féminine. On observe comme nous l'avons dit précédemment que les résultats du 10km (points rouges) et du 10000m (points bleus foncés) se mélangent. Toutes les disciplines se distinguent avec évidemment la plus courte distance (800m) en haut du graphique et la plus longue (marathon) en bas.



Dans les 2 graphiques ci-dessous, nous avons séparé les épreuves courtes (middle-long) et les épreuves longues (road-running) de la base de données des hommes. On observe étonnamment que sur les distances courtes, les performances sur le 10000m et le 3000m haies sont similaires. Pour les autres disciplines, les performances sont croissantes en fonction de la distance\*



Quant aux âges on observe pour les distances courtes, les meilleures performances se situent entre 20 et 25 ans alors que pour les longues distances, le pic se situe entre 25 et 30 ans. Par ailleurs, l'étendue des âges est différente entre les épreuves courtes et longues. Par exemple, peu de résultats sont présents dans les longues distances pour les âges inférieurs à 20 ans alors que pour les épreuves rapides, on observe beaucoup plus de résultats. En revanche, pour les épreuves les plus longues, il y a davantage de données disponibles après 45 ans que pour les épreuves les plus courtes.



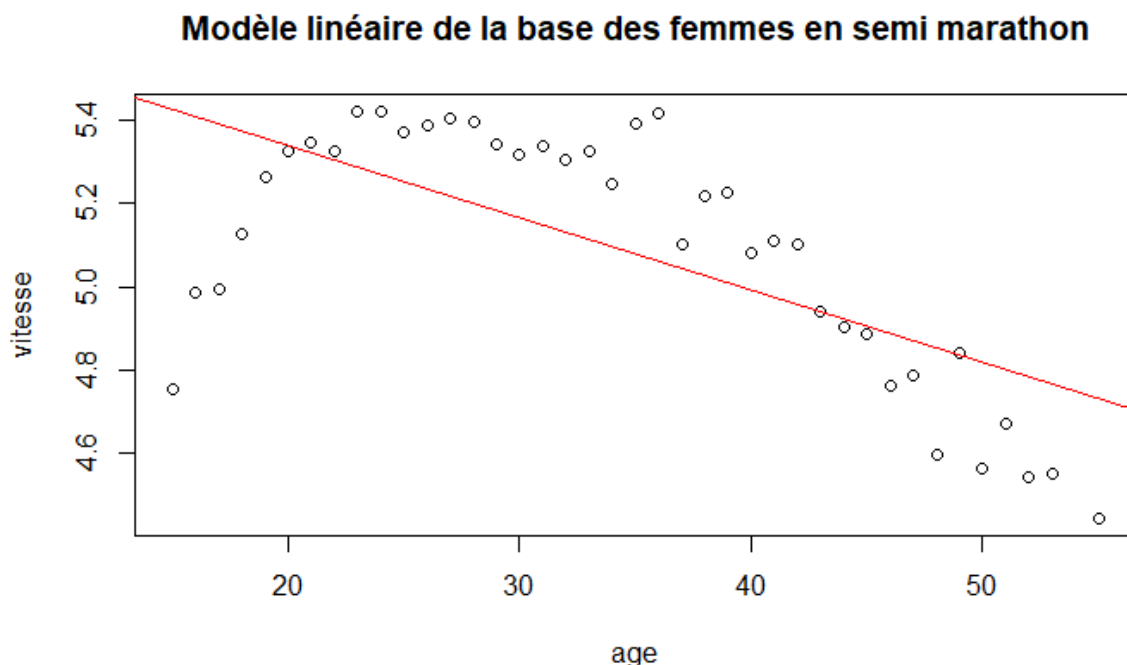
# Ajustement par modèles

Pour chaque épreuve de chaque genre, nous allons ajuster les données par 3 modèles : le modèle linéaire, le modèle polynomial et l'équation de moore. Nous réalisons donc 24 modèles

## Modèle linéaire

Pour chacun des 16 dataframes, nous ajustons un modèle linéaire grâce à la fonction "lm" de R puis nous traçons tous les graphiques correspondants.

Voici ci-dessous le graphique du modèle linéaire approchant le mieux les données parmi les 16 dataframes (8 épreuves pour chacun des 2 genres). On remarque qu'il reste tout de même très médiocre avec des différences très élevées entre les données réelles et celles ajustées par le modèle.

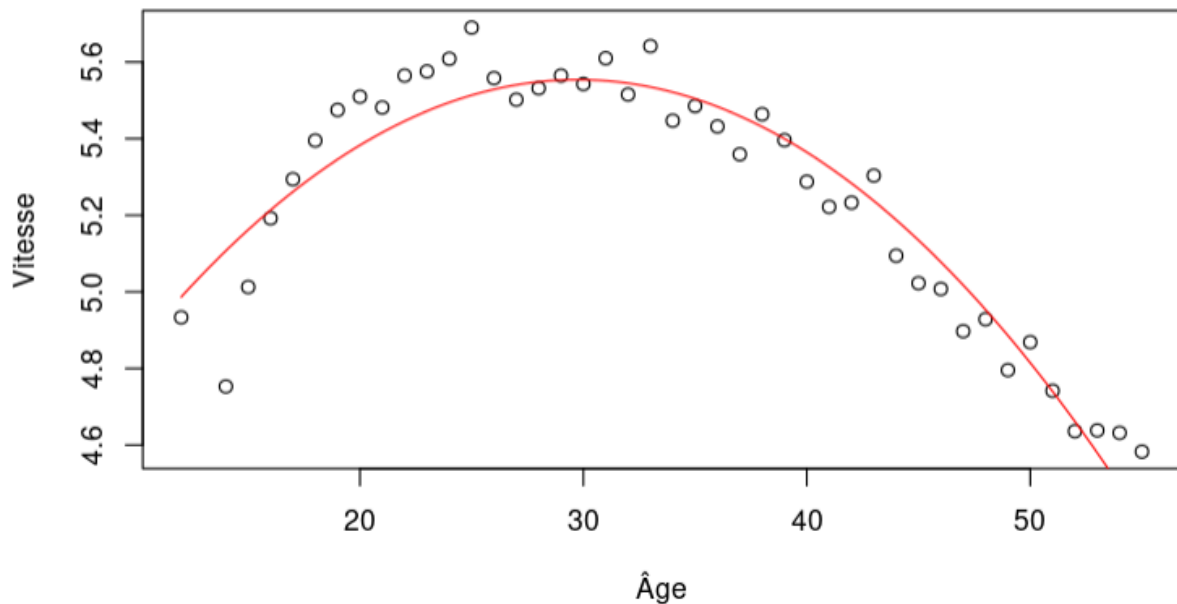


## Modèle polynomial

Pour chacun des 16 dataframes, nous ajustons un modèle polynomial grâce à la fonction "lm" de R puis nous traçons tous les graphiques correspondants.

Ci-dessous se trouve le graphique du modèle polynomial étant le plus adapté aux données. On observa qu'à la différence du modèle linéaire, la courbe suit les tendances principales des données. Les résidus sont, ici, plus faibles qu'avec le modèle linéaire.

### Modèle Polynomial 10km femmes

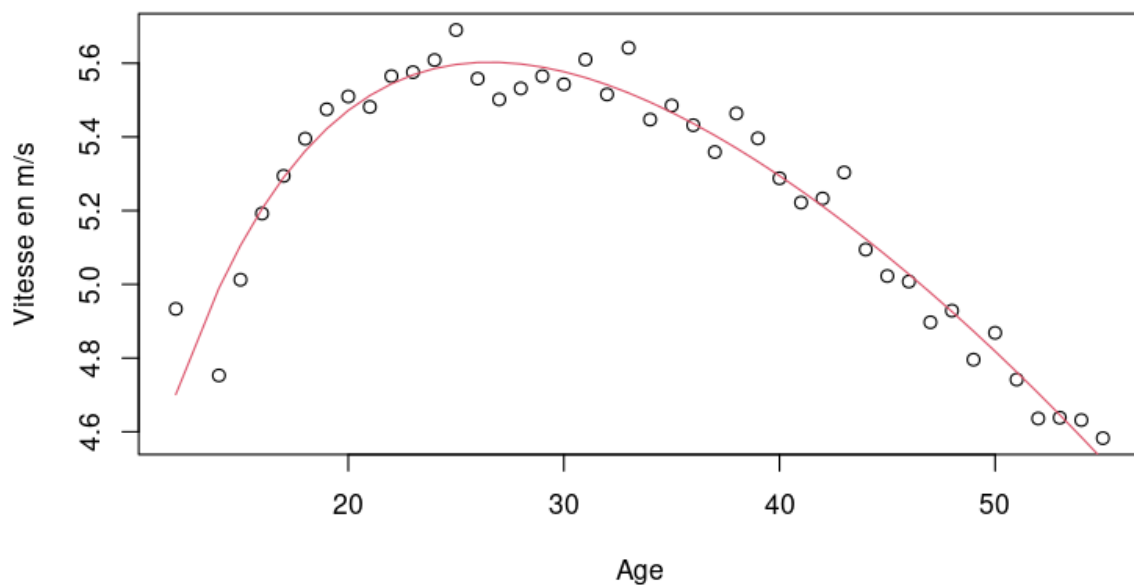


### Equation de Moore

Enfin nous utilisons l'équation de Moore, pour la résoudre nous faisons une régression non linéaire, effectuée à l'aide de la fonction `MMC()`. Nous appliquons à chacun de nos dataframes cette équation.

Le graphique présenté ci-dessous correspond au modèle de Moore approchant le mieux les données. On peut voir que ce modèle, représenté par cette courbe rouge, approche, encore mieux, les données que le modèle polynomial. Les points sont quasiment sur la courbe ou s'en approchent, les résidus sont donc faibles.

### Modèle de moore femmes 10km





# Critères d'ajustement des modèles

Afin de déterminer quel modèle est le plus proche des données, nous calculons, après avoir observé la normalité des résidus, les critères suivants sur tous nos modèles:  $R^2$ , AIC et BIC

## Résidus

Nous affichons pour chaque modèle les graphiques permettant l'étude de la normalité des résidus. Étudier la normalité des résidus d'un modèle est important pour vérifier si les erreurs entre les valeurs observées et celles prédites suivent une répartition aléatoire autour de zéro, comme le ferait une distribution normale. Si c'est le cas, cela indique que le modèle traduit efficacement les tendances des données et ce sans biais. Si les résidus ne sont pas normalement distribués, cela peut remettre en question la validité du modèle et des conclusions tirées. Ainsi, examiner la normalité des résidus aide à choisir les meilleurs modèles pour analyser les relations entre les variables.

## Coefficient de détermination ajusté

On calcule ensuite le  $R^2$  qui est le coefficient de détermination ajusté. Il mesure à quel point un modèle s'adapte bien aux données sans prendre en compte la complexité du modèle. Il est exclusivement compris entre 0 et 1. Un score proche de 1 indique un bon ajustement, tandis qu'un score proche de 0 suggère que le modèle pourrait être amélioré. Cependant, bien qu'il soit utile pour évaluer l'ajustement, il ne considère pas la balance entre la précision et la complexité du modèle.

## AIC

L'AICc (corrected Akaike information criterion) est un outil d'évaluation des modèles statistiques qui, contrairement au  $R^2$ , prend en compte à la fois la précision et la complexité du modèle. Ainsi, un score bas d'AIC indique un bon compromis entre la précision et la simplicité du modèle.

## BIC

Le critère d'information bayésien (BIC) est un autre outil d'évaluation des modèles statistiques, similaire à l'AIC mais avec une pénalisation plus forte de la complexité du modèle. En sélectionnant le modèle avec le score BIC le plus bas, nous optons pour celui qui explique au mieux les données avec la plus grande simplicité.

## Conclusion

Dans ce tableau en annexe se trouve toutes les valeurs des indicateurs pour tous les modèles (3 par épreuve de chaque genre soit 24 modèles). On remarque que quasiment pour toutes les épreuves, le modèle qui approche le mieux les données, c'est-à-dire qui a un  $R^2$  proche de 1 et un BIC et AIC faible, est le modèle de Moore.

# Pic de performance

Grâce aux graphiques vus précédemment dans l'ajustement par modèles on a pu constater que la relation âge-performance ressemble à un "U" inversé, ce qui signifie qu'on admet un âge où la performance va être maximale.

Pour retrouver l'estimation de cet âge lié à la meilleure performance par épreuve il nous faut prendre le peak de la fonction "MMC" que l'on a utilisé lors de l'ajustement des modèles avec l'équation de Moore. Celle-ci va nous donner l'estimation de la meilleure vitesse de l'épreuve selon son modèle, à partir de cette donnée on peut retrouver une estimation de l'âge facilement.

## Résultats

Le tableau suivant permet de rendre compte de l'estimation de l'âge à laquelle on a la meilleure performance réalisée selon l'épreuve.

Epreuve	Vitesse	Âge
800m - hommes	7.810 m/s	20
1500m - hommes	7.196 m/s	23
3000m - hommes	6.722 m/s	25
3000m haies - hommes	6.313 m/s	29
5000m - hommes	6.540 m/s	24
10km - hommes	6.275 m/s	21
Semi marathon - hommes	6.010 m/s	20
Marathon - hommes	5.712 m/s	31
800m - femmes	6.962 m/s	27
1500m - femmes	6.436 m/s	27
3000m - femmes	5.939 m/s	30
3000m haies - femmes	5.591 m/s	28
5000m - femmes	5.777 m/s	22
10km - femmes	5.602 m/s	24
Semi marathon - femmes	5.397 m/s	28
Marathon - femmes	5.134 m/s	35

## Conclusion sur le pic de performance

Grâce au tableau, on peut constater que l'âge estimé de la meilleure performance réalisée est très variable d'une épreuve à l'autre. Cependant on constate quand même que pour la moitié des épreuves, tout sexe confondu, le pic de performance se situe entre 20 et 25 ans. Il n'y a que pour les épreuves de marathon hommes et femmes où l'âge de la meilleure performance se situe après 30 ans avec respectivement 31 et 35 ans. Ainsi pour la majorité des épreuves le pic de performance a lieu entre 20 et 30 ans.

# Relation âge-performance aux échecs

## Calcul de l'âge

Nous allons calculer l'âge des joueurs d'échecs lors de leurs différents résultats. Pour cela, nous traitons les colonnes "Année", "Mois" correspondant aux dates de compétition et "Date.naissance" pour la date de naissance du joueur. Premièrement, nous supprimons les individus ayant des cases vides dans l'une des deux variables car ces derniers ne sont pas utilisables pour notre problématique âge-performance. Deuxièmement, nous modifions les dates étant incomplètes pour la naissance. Nous avons deux cas, les dates n'ayant pas de jour, pour lesquelles nous leur assignons le milieu du mois c'est-à-dire le 15, et celles ayant juste l'année à qui nous assignons le milieu de l'année soit le 02 juillet. Ensuite, on s'intéresse aux dates des résultats, on commence par garder uniquement les 3 premiers caractères du mois, puis on regroupe les variables "Année" et "Mois" et enfin on rajoute le jour, on décide ici de mettre le 15 du mois.

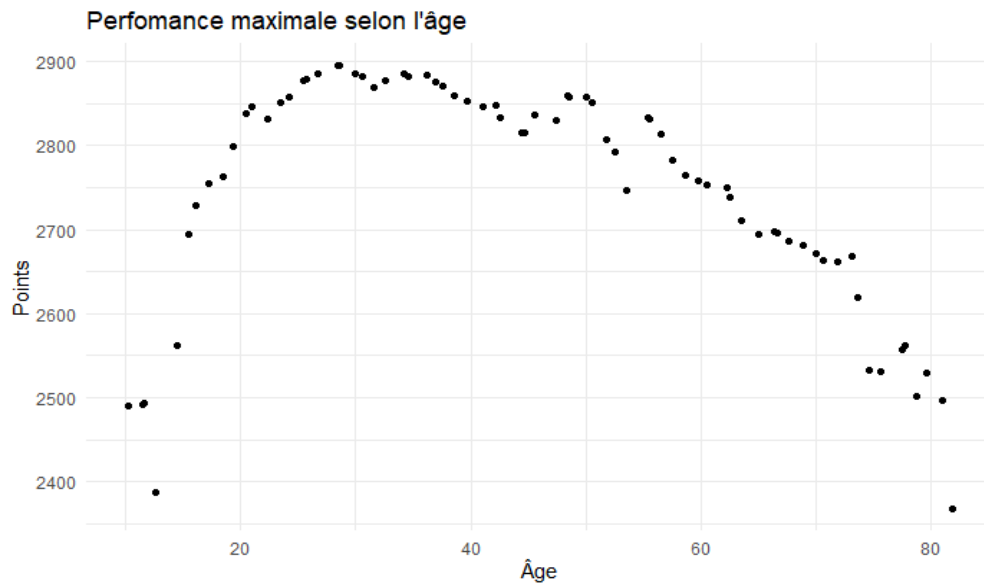
Une fois que tous les individus des deux variables ont tous une date complète et au même format, nous les transformons au format date de R. Nous pouvons à présent procéder au calcul de l'âge. Pour cela nous soustrayons la date de naissance à la date des résultats et nous divisons le résultat par 365,2.

## Création base de données

Nous créons une base contenant le meilleur nombre de points par âge. Avant de pouvoir se faire nous avons dû convertir la variable "Points" de caractère à numérique, cependant la présence d'espaces non visibles nous a obligé à extraire les données de la variable pour en créer une nouvelle. A partir de là nous pouvons garder les meilleurs résultats en fonction de l'âge, en faisant cela nous avons obtenu des doublons avec des joueurs ayant le même nombre de points et le même âge. Ainsi nous avons supprimé les doublons pour n'avoir qu'un âge par meilleur résultat, on obtient donc notre base associant l'âge à la meilleure performance avec 61 individus.

## Graphique performances maximales par âge

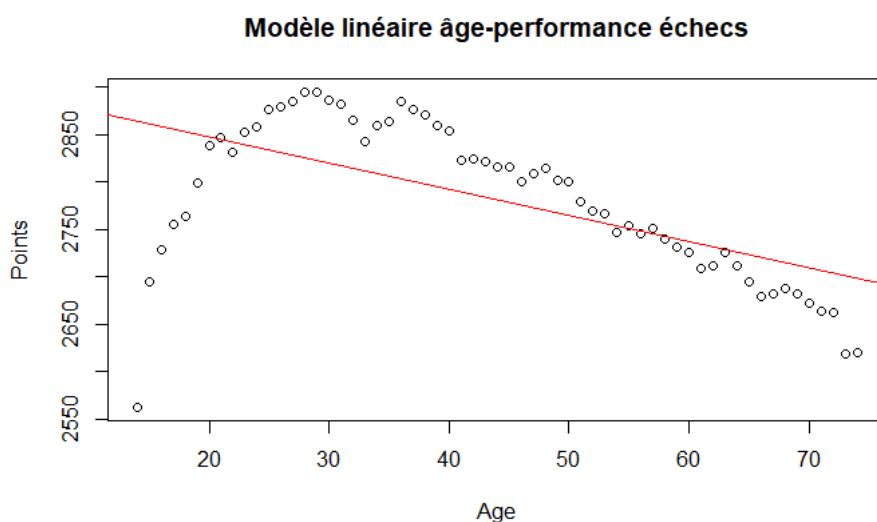
Ce graphique ci-dessous permet de mettre en avant le nombre maximum de points marqués par les joueurs d'échecs en fonction de l'âge, on a ici uniquement le meilleur score pour chaque âge. On remarque que jusqu'à environ 20 ans le nombre de points augmente significativement, tandis qu'entre 20 et 40 ans on observe une stabilisation avec notamment un pic de performance dans ces âges là, par la suite le nombre de points ne cesse de diminuer.



## Ajustement par modèles

Nous allons ajuster nos données par 3 modèles : le modèle linéaire, le modèle polynomial et l'équation de moore. Nous utilisons les mêmes fonctions expliquées précédemment pour la base d'athlétisme à notre base d'échecs.

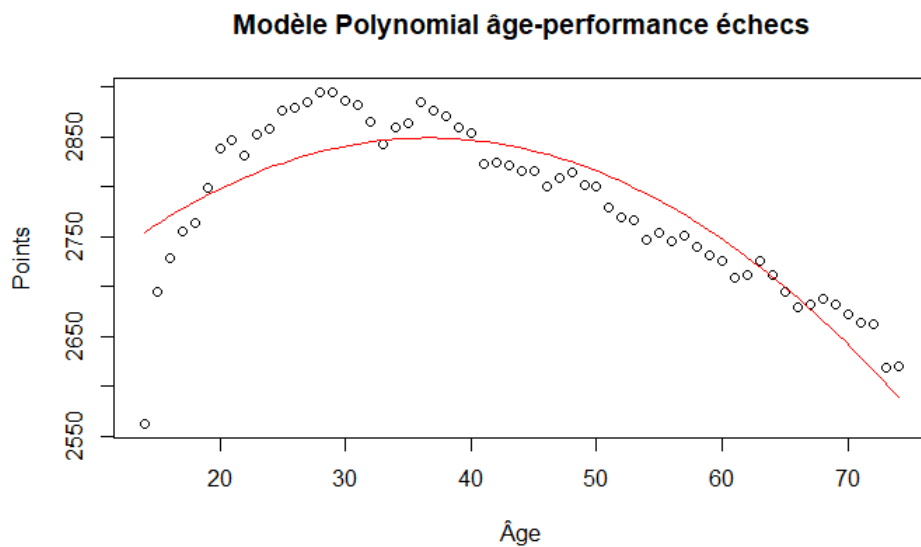
### Modèle linéaire



Le tracé du modèle linéaire sur notre graphique montre bien que ce modèle n'est pas ajusté à nos données, effectivement il ne suit pas les différents points tracés.

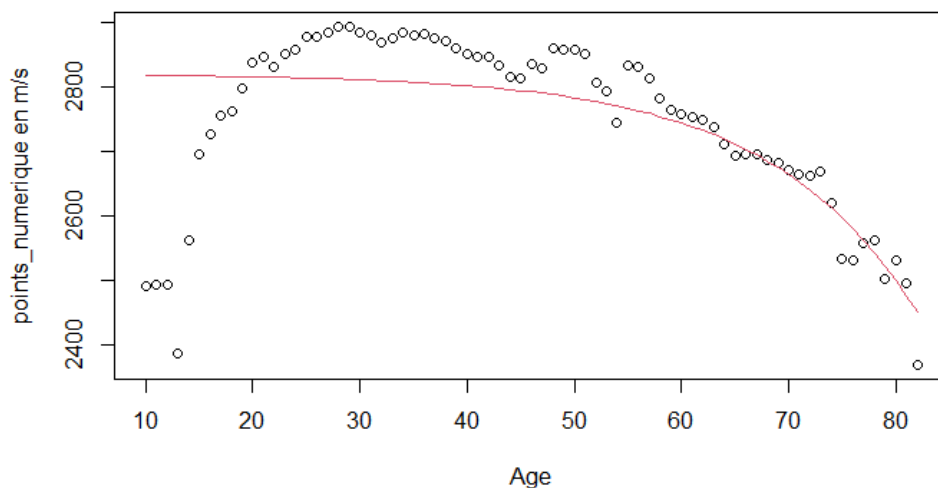
### Modèle polynomial

Dans le cadre du modèle polynomial, on remarque que celui-ci est plus ajusté à nos données par rapport au modèle linéaire, cependant il ne suit pas exactement la tendance de nos données.



## Equation de Moore

En effectuant l'équation de Moore on obtient la courbe suivante, on remarque que celle-ci s'ajuste bien aux données à partir de 60 ans, avant cet âge là on remarque que l'équation ne suit pas vraiment nos données. Les résidus restent assez importants, néanmoins l'équation de Moore reste le meilleur modèle d'ajustement par rapport au modèle linéaire et polynomial malgré quelques défauts ici.



## Conclusion ajustement des modèles

Pour ajuster nos modèles, on choisira ici de prendre l'équation de Moore comme meilleur modèle en termes d'estimation par rapport aux données réelles.

## Critères d'ajustement des modèles

Comme vu précédemment dans nos données sur l'athlétisme, nous cherchons à déterminer quel modèle est le plus proche des données, nous calculons les critères suivants :  $R^2$ , AIC et BIC.

## Résultats

Critères d'ajustement	Modèle linéaire	Modèle polynomial	Equation de Moore
$R^2$	0.1074692	0.8132819	0.4398413
AIC	921.3961	808.1549	43978.11
BIC	928.2675	928.2675	691.3065

## Conclusion sur les critères d'ajustement

Dans le tableau ci-dessus se trouve toutes les valeurs des indicateurs de critère d'ajustement pour notre base "Échecs". On remarque que le modèle qui approche le mieux les données, c'est-à-dire qui a un  $R^2$  proche de 1 et un BIC et AIC le plus faible par rapport aux autres modèles, est le modèle polynomial. Effectivement, sur les 3 critères c'est celui qui a les "meilleurs" en ayant 2 critères sur 3.

## Pic de performance

On a pu voir sur les graphiques de l'ajustement par modèles que la relation âge-performance ressemble à un "U" inversé, ce qui signifie qu'on admet un âge où la performance va être maximale. En utilisant la même méthode explicitée précédemment pour l'autre base de donnée, on calcule donc l'estimation de l'âge auquel la meilleure performance en termes de points à été réalisée.

On obtient le résultat suivant :

Epreuve	Points	Âge
Echecs	2819	44

Ainsi l'âge estimé, auquel la performance des joueurs d'échecs est maximale, se situe aux alentours de 44 ans.

## Conclusion générale

Le travail d'ajustement par modèles linéaires, polynomiaux et par l'équation de Moore a permis de mieux comprendre comment les performances varient avec l'âge dans ces disciplines. Les modèles de Moore se sont révélés être les plus adaptés pour représenter la relation âge-performance pour les différentes épreuves d'athlétisme. Cependant, dans la représentation de la relation entre l'âge et la performance pour les échecs, c'est le modèle polynomial qui s'est révélé être le plus adapté aux données.

De plus, les analyses révèlent que la performance atteint son pic à des âges variables selon la discipline et le sexe. Le pic de performance est plus précoce (entre 20 et 30 ans) dans l'athlétisme, alors que le pic est plus tardif (aux alentours de 44 ans) dans les échecs. Ces résultats suggèrent une différence notable dans la manière dont l'âge affecte la performance entre des activités physiques et intellectuelles.

Enfin, les interprétations de ces données montrent que la performance est intrinsèquement liée à l'âge. Cependant, l'âge n'est pas l'unique facteur déterminant la performance, cette dernière varie aussi selon le type d'activité et le genre.

## Question bonus n°2

Nous avons demandé à ChatGPT de nous répondre à la question 3 du devoir c'est-à-dire à celle pour laquelle il faut calculer le  $R^2$ , l'AIC et le BIC.

Le premier problème auquel nous nous sommes heurtés est le fait que notre sujet soit précis. Ainsi, il ne connaît pas nos données et ne connaît surtout pas le modèle de Moore. Pour autant, les réponses apportées pour calculer le  $R^2$ , l'AIC et le BIC d'un modèle linéaire ainsi que d'un modèle polynomial sont justes et ressemblent fortement à ce que nous avons réalisé.

Deuxième problème, il ne réalise pas les calculs ainsi, nous devons copier le code, l'exécuter sur notre R et lui donner ensuite les résultats à interpréter.

Une fois ces étapes réalisées, ses interprétations sont cohérentes et correspondent à nos interprétations précédentes. Il apporte même plusieurs pistes de réflexion afin qu'on puisse sélectionner les critères qui nous importent le plus ou qui correspondent à notre cas (cas de petit échantillon, préférence en un modèle simple ou plutôt d'un modèle le plus performant possible...).

Cependant, nous avons ensuite demandé à ChatGPT de nous rédiger des paragraphes sur l'AIC, sur le  $R^2$  et sur le BIC. Nous avons atteint ici ses limites. Il faut faire attention en demandant à ChatGPT de rédiger des textes correspondant à des termes techniques car il peut avoir des connaissances limitées.

Voici des extraits des paragraphes :

- $R^2$  : "il mesure à quel point un modèle s'adapte bien aux données, en tenant compte de sa complexité"
- AIC : "contrairement au coefficient de détermination ( $R^2$ ), il prend en compte à la fois la précision et la complexité du modèle"

Ainsi, on remarque, ici, qu'il est nécessaire de faire attention aux résultats car nous pourrions rédiger des informations contradictoires et donc fausses.

Pour conclure, ChatGPT est très efficace sur cette question de calcul d'indicateur et d'interprétation mais seulement dans le cas des modèles linéaires et polynomiaux. Cependant il faut toujours rester vigilant en utilisant cette IA pour des problèmes techniques et précis car ses connaissances sont parfois erronées.



# Annexe

Indicateurs :

<p>Modèle linéaire hommes_800m :  <math>R^2</math> ajusté : 0.111  AIC : -9.047726  BIC : -4.214972</p> <p>Modèle polynomial hommes_800m :  <math>R^2</math> ajusté : 0.493  AIC : -28.89866  BIC : -22.45499</p> <p>Modèle de Moore hommes_800m :  <math>R^2</math> ajusté : 0.739  AIC : -169.6366  BIC : -164.443</p>	<p>Modèle linéaire femmes_800m :  <math>R^2</math> ajusté : -0.011  AIC : -0.8787193  BIC : 3.423242</p> <p>Modèle polynomial femmes_800m :  <math>R^2</math> ajusté : 0.866  AIC : -62.63591  BIC : -56.89997</p> <p>Modèle de Moore femmes_800m :  <math>R^2</math> ajusté : 0.913  AIC : -144.9175  BIC : -140.72</p>
<p>Modèle linéaire hommes_1500m :  <math>R^2</math> ajusté : 0.23  AIC : -14.8969  BIC : -10.40737</p> <p>Modèle polynomial hommes_1500m :  <math>R^2</math> ajusté : 0.519  AIC : -29.51078  BIC : -23.52474</p> <p>Modèle de Moore hommes_1500m :  <math>R^2</math> ajusté : 0.683  AIC : -147.1412  BIC : -142.5838</p>	<p>Modèle linéaire femmes_1500m :  <math>R^2</math> ajusté : -0.016  AIC : 18.14326  BIC : 22.89382</p> <p>Modèle polynomial femmes_1500m :  <math>R^2</math> ajusté : 0.889  AIC : -60.77375  BIC : -54.43967</p> <p>Modèle de Moore femmes_1500m :  <math>R^2</math> ajusté : 0.954  AIC : -155.9131  BIC : -150.8694</p>
<p>Modèle linéaire hommes_3000m :  <math>R^2</math> ajusté : 0.243  AIC : 1.403376  BIC : 6.316134</p> <p>Modèle polynomial hommes_3000m :  <math>R^2</math> ajusté : 0.695  AIC : -32.27108  BIC : -25.72074</p> <p>Modèle de Moore hommes_3000m :  <math>R^2</math> ajusté : 0.791  AIC : -164.1369</p>	<p>Modèle linéaire femmes_3000m :  <math>R^2</math> ajusté : 0.086  AIC : 25.78756  BIC : 30.92828</p> <p>Modèle polynomial femmes_3000m :  <math>R^2</math> ajusté : 0.726  AIC : -22.70951  BIC : -15.85522</p> <p>Modèle de Moore femmes_3000m :  <math>R^2</math> ajusté : 0.072  AIC : -88.79931</p>

BIC : -158.7987	BIC : -83.05613
Modèle linéaire hommes_3000m_haies : R <sup>2</sup> ajusté : 0.346 AIC : 3.089127 BIC : 7.668208  Modèle polynomial hommes_3000m_haies : R <sup>2</sup> ajusté : 0.769 AIC : -31.41337 BIC : -25.30793  Modèle de Moore hommes_3000m_haies : R <sup>2</sup> ajusté : 0.817 AIC : -146.5318 BIC : -141.8057	Modèle linéaire femmes_3000m_haies : R <sup>2</sup> ajusté : 0.011 AIC : 30.83598 BIC : 35.50203  Modèle polynomial femmes_3000m_haies : R <sup>2</sup> ajusté : 0.773 AIC : -19.75777 BIC : -13.53638  Modèle de Moore femmes_3000m_haies : R <sup>2</sup> ajusté : 0.765 AIC : -139.5649 BIC : -134.6768
Modèle linéaire hommes_5000m : R <sup>2</sup> ajusté : 0.266 AIC : -8.859485 BIC : -4.193441  Modèle polynomial hommes_5000m : R <sup>2</sup> ajusté : 0.612 AIC : -30.22027 BIC : -23.99887  Modèle de Moore hommes_5000m : R <sup>2</sup> ajusté : 0.713 AIC : -148.0792 BIC : -143.1911	Modèle linéaire femmes_5000m : R <sup>2</sup> ajusté : -0.018 AIC : 20.18871 BIC : 25.1794  Modèle polynomial femmes_5000m : R <sup>2</sup> ajusté : 0.817 AIC : -45.74809 BIC : -39.09384  Modèle de Moore femmes_5000m : R <sup>2</sup> ajusté : 0.873 AIC : -108.6295 BIC : -103.1517
Modèle linéaire hommes_10km : R <sup>2</sup> ajusté : 0.335 AIC : -9.073129 BIC : -3.932413  Modèle polynomial hommes_10km : R <sup>2</sup> ajusté : 0.768 AIC : -51.25962 BIC : -44.40533  Modèle de Moore hommes_10km : R <sup>2</sup> ajusté : 0.914 AIC : -190.7845 BIC : -185.0413	Modèle linéaire femmes_10km : R <sup>2</sup> ajusté : 0.305 AIC : 15.84557 BIC : 21.12917  Modèle polynomial femmes_10km : R <sup>2</sup> ajusté : 0.902 AIC : -67.44284 BIC : -60.39804  Modèle de Moore femmes_10km : R <sup>2</sup> ajusté : 0.957 AIC : -214.38 BIC : -208.3878
Modèle linéaire hommes_mara :	Modèle linéaire femmes_mara :

<p>R<sup>2</sup> ajusté : 0.002 AIC : 2.410124 BIC : 7.55084</p> <p>Modèle polynomial hommes_mara : R<sup>2</sup> ajusté : 0.819 AIC : -66.71408 BIC : -59.8598</p> <p>Modèle de Moore hommes_mara : R<sup>2</sup> ajusté : 0.928 AIC : -169.1955 BIC : -163.4523</p>	<p>R<sup>2</sup> ajusté : 0.045 AIC : 15.86616 BIC : 21.00688</p> <p>Modèle polynomial femmes_mara : R<sup>2</sup> ajusté : 0.843 AIC : -57.13073 BIC : -50.27644</p> <p>Modèle de Moore femmes_mara : R<sup>2</sup> ajusté : 0.98 AIC : -150.252 BIC : -144.5088</p>
<p>Modèle linéaire hommes_semi_mara : R<sup>2</sup> ajusté : 0.126 AIC : -16.90327 BIC : -11.99051</p> <p>Modèle polynomial hommes_semi_mara : R<sup>2</sup> ajusté : 0.76 AIC : -65.14776 BIC : -58.59742</p> <p>Modèle de Moore hommes_semi_mara : R<sup>2</sup> ajusté : 0.079 AIC : -124.374 BIC : -119.0357</p>	<p>Modèle linéaire femmes_semi_mara : R<sup>2</sup> ajusté : 0.444 AIC : -2.197491 BIC : 2.869148</p> <p>Modèle polynomial femmes_semi_mara : R<sup>2</sup> ajusté : 0.885 AIC : -64.14691 BIC : -57.3914</p> <p>Modèle de Moore femmes_semi_mara : R<sup>2</sup> ajusté : 0.914 AIC : -202.3806 BIC : -196.768</p>