Projet analyse de données Analyse factorielle Classification

Comment les facteurs mécaniques, physiologiques et environnementaux influent sur la performance en course à pied ?



Table des matières

I)	Introduction et contexte	3
II)	Construction et nettoyage de la base de données	3
1)	Collecte et exportation	3
2)	Nettoyage de la base de données	3
III)	Exploration de la base de données	4
1)	Définitions des indicateurs	4
2)	Statistique descriptive	5
IV)	Analyse en Composantes Principales (ACP)	6
1)	Justification de la méthode factorielle	6
2)	Conditions d'applications	6
3)	Variables actives et illustratives	7
4)	Choix du nombre d'axes et valeurs propres	7
5)	Interprétation des variables	8
	a) Sens des axes	8
	b) Facteurs mécaniques 1	1
	c) Facteurs physiologiques	3
	d) Facteurs environnementaux / Conditions entraînements	3
6)	Tableau récapitulatif des variables1	7
V)	Classification des activités1	7
1)	Classification Ascendante Hiérarchique1	7
2)	Caractérisation des classes1	9
	a) Classe 12	0
	b) Classe 22	1
	c) Classe 3	2
	d) Classe 42	2
3)	Les caractéristiques d'une course performante2	3
VI)	Synthèse de l'étude2	3

I) Introduction et contexte

Cette étude est centrée sur le domaine sportif, elle porte sur mes activités de course à pied de 2019. En effet, il est intéressant de se focaliser sur les performances de plusieurs activités car cela permet d'évaluer où le coureur doit progresser (technique de foulée, gain de puissance, ...). La course à pied est l'objet de nombreux travaux de recherche de plus en plus perfectionnées. Des techniques ont été mises au point pour mesurer la performance dans les dimensions biomécaniques et physiologiques. C'est pourquoi, la course à pied est un enjeu scientifique dans les deux aspects suivants : la coordination gestuelle et l'optimisation de la dépense énergétique occasionnée.

Plusieurs indicateurs peuvent être récoltés pendant une séance. Ces paramètres concernent la longueur de foulée, la cadence, l'oscillation verticale, le temps de contact au sol pour les paramètres mécaniques. La fréquence cardiaque et la dépense énergétique constituent les facteurs physiologiques. On peut également retrouver des facteurs externes au corps comme le type de chaussure, la température ou encore la vitesse du vent.

L'enjeu de l'étude est de déterminer comment les facteurs mécaniques, physiologiques et environnementaux influent sur la performance en course à pied ?

Cependant, avant d'interpréter les résultats d'analyse, il faut expliquer la notion de performance dans le domaine de la course à pied. Le running se compose de phases différentes d'intensité dans un but de performance et de minimisation de la perte d'énergie. La performance en course à pied est lorsque le coureur réussi à maximiser sa vitesse tout en minimisant ses efforts. De multiples facteurs biomécaniques, physiques et extérieurs rentrent donc en jeu. Nous allons ainsi analyser le lien entre les variables, voir quelles sont celles qui participent ou non à la performance et leurs degrés de contribution. La performance est une notion importante pour un coureur de compétition. L'objectif de l'étude est de voir dans quelle mesure les facteurs de course à pied pourraient évaluer l'économie de course et la vitesse.

Cette étude est réalisée à partir des données de mes entraînements et course donc l'interprétation des résultats sera à nuancer pour d'autres coureurs.

II) Construction et nettoyage de la base de données

1) Collecte et exportation

Grâce à une montre connecté portée pendant l'intégralité de l'entrainement ou de la course, des informations sont récoltées. Ces données sont ensuite transférées sur une plateforme nommée Garmin Connect. Pour récupérer les données, il suffit d'exporter au format CSV un tableau de toutes les activités avec les variables associées.

2) Nettoyage de la base de données

Une fois la base de données chargée, un nettoyage de données est nécessaire pour pouvoir l'analyser de manière pertinente.

Voici quelques corrections établis sur la base :

- Conversion des variables quant en type nombre
- Remplacement des séparateurs décimales en virgules
- Renommage de variables
- Conversion du fichier en .xlsx
- Ajout d'une colonne ID pour rendre unique 1 activité
- Suppression des colonnes vides

III) Exploration de la base de données

Pour répondre du mieux possible à la problématique, il faut d'abord commencer par décrire la base de données.

On a donc 98 observations qui correspondent au nombre d'activité pour un coureur pour l'année 2019.

Variable	Type
ID activité	Qualitatif ordinal
Mois	Qualitatif ordinal
Moment du jour	Qualitatif ordinal
Météo	Qualitatif cardinal
Type de chaussure	Qualitatif cardinal
Température extérieure	Quantitatif continue
Vitesse moyenne	Quantitatif continue
FC moyenne	Quantitatif continue
FC maximale	Quantitatif continue
TE aérobie	Quantitatif continue
Cadence moyenne	Quantitatif discret
Longueur foulée moyenne	Quantitatif continue
Calories	Quantitatif continue
Rapport vertical moyen	Quantitatif continue
Oscillation verticale moyenne	Quantitatif continue
Temps contact sol moyen	Quantitatif continue

Tableau des variables avec leurs types

1) Définitions des indicateurs

16 variables (5 qualitatives et 11 quantitatives).

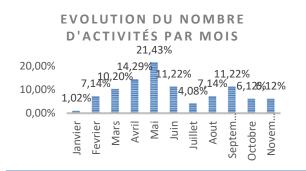
De quoi est-ce qu'on parle?

- Mois : le mois à laquelle l'activité a été réalisé en 2019.
- Moment du jour : la phase de la journée de l'activité (matin, après-midi ou soir à partir de 18 heures).
- Météo : pluie, nuageux ou soleil.
- Type de chaussure : 3 marques différentes pour observer des possibles variations de performance (Nike, New Balance et Salomon).

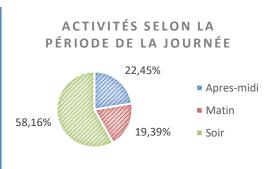
- Température extérieure : exprimé en degré Celsius.
- Vitesse moyenne : Exprimé en km/h.
- FC moyenne : la fréquence cardiaque en bpm (battement par minute).
- FC maximale : la fréquence cardiaque en bpm (battement par minute) à un instant t.
- TE aérobie : la Training Effect aérobie est un indicateur construit par Garmin. C'est la contribution de l'effort en endurance sur une échelle de 1 à 5.
- Cadence moyenne : le nombre de pas par minute (pied gauche et pied droit). Généralement, l'optimum d'un point de vue biomécanique est de 180 ppm (pas par minute).
- Longueur foulée moyenne : la distance moyenne en mètre entre deux appuis.
- Calories : quantité d'énergie pour produire un effort (kilocalorie).
- Rapport vertical moyen : rapport coût/efficacité (oscillation verticale/longueur foulée). Indicateur de l'efficacité de la technique de course.
- Oscillation verticale moyenne : la hauteur de rebond propre à la foulée de course à pied exprimé en millimètres.
- Temps contact sol moyen : le temps pendant lequel le pied reste en contact avec le sol exprimé en millisecondes.

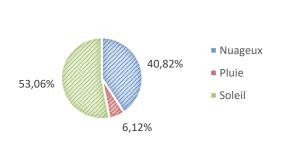
2) Statistique descriptive

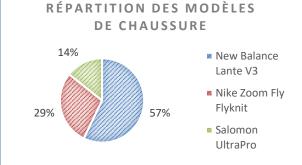
Variables qualitatives



RÉPARTITION DE LA MÉTÉO







Variables quantitatives

Variable	Moyenne	Ecart-type	Minimum	Maximum
Température (°C)	17	7	3	34
Vitesse vent (km/h)	16	10	0	50
Calories (kcal)	713	402	191	2283
Vitesse moyenne (km/h)	13,2	1,6	7,5	16,1
FC moyenne (bpm)	160	12	126	183
FC maximale (bpm)	184	11	147	201
TE aérobie (1 à 5)	3,7	1,0	2,0	5,0
Cadence moyenne (ppm)	168	3	156	174
Longueur foulée moyenne (m)	1,3	0,1	0,8	1,6
Rapport verticale moyen	72,2	8,6	58,3	113,9
Temps contact sol moyen (ms)	231	7	205	267

Tableau des indicateurs et statistiques associés

IV) Analyse en Composantes Principales (ACP)

1) Justification de la méthode factorielle

La méthode factorielle permet de synthétiser la base de données sur des composantes principales. L'objectif est d'évaluer des liens de corrélations entre les variables actives et la vitesse. La contribution des variables sur les axes permettra d'interpréter les composantes principales et donc de répondre à la problématique.

Pour cela, nous allons utiliser la méthode ACP (Analyse en Composantes Principales) sur les variables quantitatives. Enfin, les calculs ont été faits au moyen du logiciel de statistiques SPAD. Le logiciel SPAD permet d'obtenir les tableaux des valeurs propres (variances des axes), des coordonnées des variables actives sur les axes (contributions) et les valeurs des corrélations entre les variables actives. Le cercle de corrélations des variables ainsi que le graphique des individus sont générés par SPAD.

2) Conditions d'applications

Pour avoir une ACP et classification la plus fiable possible, il faut au minimum 50 observations et au moins 10 variables. Avec cette base de données, les conditions sont respectées. Nous pouvons donc poursuivre notre analyse sur le lien entre les facteurs clé de performance pour expliquer la vitesse et le gain d'énergie en course à pied.

3) Variables actives et illustratives

Nous décidons de mettre toutes les variables quantitatives en actives et illustratives les variables qualitatives. Les variables qualitatives expliqueront s'il y a un lien (si oui lequel) entre la performance en les facteurs extérieurs aux corps.

4) Choix du nombre d'axes et valeurs propres

Le pourcentage de variance expliquée nous indique que l'axe 1 retient 38% de l'inertie totale. L'axe 2 quant à lui retient tout de même 29% de la variance, ce qui n'est pas négligeable, et qui conduit à un taux d'inertie expliquée de 67%, ce qui est un très bon résultat. Il est donc intéressant d'étudier les variables et individus sur ces deux axes qui présentent un éventuel intérêt.

Cependant, d'après SPAD et le critère de Kaiser, il faudrait retenir les 3 premières dimensions car les valeurs propres de ces axes sont supérieures à 1 (voir tableau 1 cidessous).

Axe	Variance de l'axe (valeur propre)	% de variance expliquée	% de variance expliquée cumulé	Retenu selon le critère de Kaiser
1	3,816	38,2	38,2	X
2	2,920	29,2	67,4	X
3	1,017	10,2	77,5	X
4	0,961	9,6	87,1	
5	0,560	5,6	92,7	
6	0,381	3,8	96,5	
7	0,198	2,0	98,5	
8	0,111	1,1	99,6	
9	0,035	0,4	100,0	
10	0,001	0,0	100,0	
Total	10,000	100,0	100,0	

Tableau 1 – Données des variances des axes factoriels

Lors de l'ACP, on retiendra les deux premiers axes afin de faciliter la représentation des variables et l'analyse. En effet, le but est d'obtenir le maximum d'inertie avec le minimum de facteurs.

De plus, le critère du coude *(voir figure 1 ci-dessous)* nous conduit à retenir les deux premières car on observe une cassure (changement abrupt de la pente) au niveau de k=2 puis une décroissance régulière à partir du troisième axe. La chute est importante dès le troisième axe qui ne conserve plus que 10% de l'inertie totale.

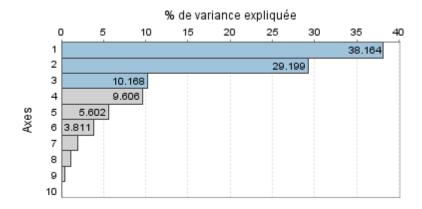


Figure 1 - Courbe de décroissance des valeurs propres

Avec le critère de Cattell, la différence secondes en millièmes entre les pallier 2 et 3 est de 1846 (très élevé). C'est la plus grande différence en comparaisons aux autres paliers.

Dans cette étude, on retiendra donc les deux premières composantes principales (premier plan factoriel) suite à la validation des critères de Kaiser et du coude. Ces deux dimensions nous permettrons d'apporter des éléments pour répondre à la problématique.

5) Interprétation des variables

a) Sens des axes

Les deux facteurs à expliquer (de référence) sont la vitesse moyenne et les calories dépensés. Les variables explicatives vont donc influencer soit positivement, soit négativement sur les variables exogènes. Elles se révèlent plus ou moins importantes sur le niveau de performance.

La corrélation de chaque point sur un axe exprime la qualité de représentation du point sur l'axe.

REPRESENTATION DE LA VITESSE

Tout d'abord, on remarque grâce aux coordonnées des variables que la vitesse moyenne (voir figure 2 ci-dessous) est fortement corrélée (0,95) avec l'axe 1 car il est proche du cercle de corrélation (extrême). Cette variable intervient dans la construction de l'axe à hauteur de 24%. Cependant, la vitesse est éloignée de l'axe 2. En effet, elle est très peu corrélée (0,20) avec la deuxième dimension. Sa contribution est de 1%, ce qui est négligeable sur l'axe 2.

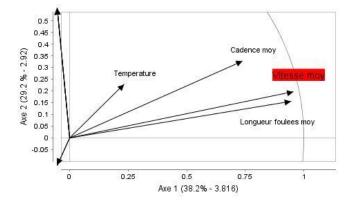


Figure 2 – Variable vitesse sur les axes 1 et 2

FONDAMENTAUX SUR LE FONCTIONNEMENT D'UN MUSCLE.

Pour fonctionner, les muscles ont besoins d'énergie. Pour fournir l'énergie nécessaire à la contraction musculaire, le coureur utilise deux types filières : la filière aérobie (avec apport d'oxygène) et la filière anaérobie (sans apport d'oxygène). Dans les courses de durée (d'endurance), l'énergie provient essentiellement de la filière aérobie. La molécule pour contracter un muscle est l'ATP. L'oxygène intervient dans la synthèse de cette molécule dans les courses de longue distance.

On peut mesurer l'aérobie, c'est-à-dire le volume d'oxygène que le corps recrute pour fournir un effort avec le VO2max. Plus le VO2max est élevé et plus le corps à d'endurance car la durée de contraction des muscles pour un même effort est augmentée.

REPRESENTATION DE LA DEPENSE ENERGETIQUE.

L'autre variable de mesure de la performance avec la vitesse est la Training Effect aérobie. Cette donnée mesure la capacité à résister à l'effort sur une échelle de 1 à 5 (indicateur calculé par Garmin). Cette donnée peut être comparée au VO2max donc on va s'en servir dans l'analyse pour évaluer le niveau de performance.

La corrélation de TE aérobie (voir figure 3 ci-dessous) est fortement corrélée (0,88) avec l'axe 2 donc la variable est très bien représentée sur l'axe. Elle n'intervient pas dans la construction de l'axe 1 car sa contribution n'est pas significative (-0,153).

De ces observations, nous pouvons dire que l'axe 1 correspond plutôt à l'appréciation de la vitesse alors que l'axe 2 correspond plutôt à l'interprétation de la dépense énergétique et notamment de la TE aérobie.

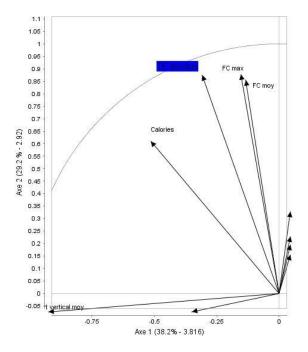


Figure 3 – Variable TE aérobie sur les axes 1 et 2

L'angle quasi-droit formé par la vitesse moyenne et la TE aérobie (voir figure 4 ci-dessous) indique que ces deux variables sont indépendantes. On peut augmenter sa vitesse sans forcément dépenser une grande quantité d'énergie, cela dépend d'autres facteurs de course comme la distance et le type de terrain (chemin de terre, asphalte, ...).

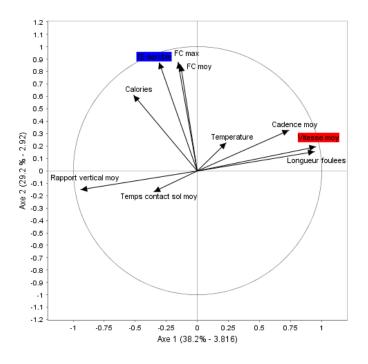


Figure 4 – Graphique des variables actives sur les axes 1 et 2

b) Facteurs mécaniques

LONGUEUR DE FOULEE ET CADENCE.

Intéressons-nous au facteur longueur de foulée, on peut voir sur le graphique des variables que la vitesse moyenne est fortement corrélée avec la variable longueur de foulée (coefficient de corrélation entre les deux variables = 0,99). De plus, graphiquement (voir figure 5 ci-dessous), la variable longueur de foulée est très proche du cercle de corrélation donc elle contribue fortement à la construction de l'axe 1 à hauteur de 23%. Ainsi, le facteur longueur de foulée influe fortement et positivement sur la vitesse.

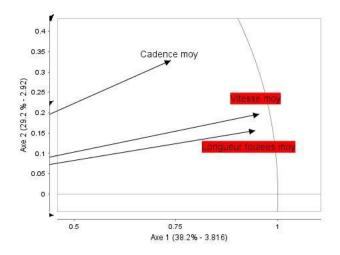


Figure 5 – Proximité entre longueur de la foulée avec la vitesse

L'autre facteur influant sur la vitesse est la cadence de course. Elle est aussi corrélée positivement avec la vitesse (0,73). La cadence produit une action sur la vitesse

Lorsque ces deux facteurs augmentent, alors la vitesse augmente aussi. En effet, la vitesse est le produit de la cadence et de la longueur de foulée.

$$Vitesse (m/min) = Cadence (ppm) \times Longueur de foulée (m)$$

Donc pour augmenter sa vitesse, il suffit d'accroître sa cadence de pas et sa distance entre deux appuis.

Plusieurs études menés par des scientifiques ont conduit a montré que la cadence optimale est de 180 ppm. C'est la cadence où l'efficacité du tendon d'Achille (voir figure 6 cicontre) est la plus forte. A partir de 180 ppm, le coureur est plus économique en énergie car on utilise au maximum l'efficacité mécanique liée à l'impact au sol. Au lieu d'utiliser uniquement notre énergie pour se propulser, on utilise l'impact au sol pour rebondir grâce à l'élasticité du pied et de la cheville.



Figure 6 – Tendon d'Achille (en blanc)

TEMPS DE CONTACT AU SOL.

Nous avons vu plus haut qu'il suffisait d'augmenter sa fréquence de pas par minute et sa longueur de foulée. Cependant, accroître ces paramètres ne se fait pas sans dépense d'énergie supplémentaire. D'autres facteurs mécaniques rentrent alors en jeu dans le but de diminuer ces dépenses énergétiques.

Le temps de contact au sol est opposé à la vitesse moyenne sur l'axe 1. Le temps au sol est corrélé négativement avec la vitesse (-0,23). Plus on va vite et plus le temps de contact au sol est faible. Logique car lorsque le pied est au sol, il n'y a aucun mouvement vers l'avant, la vitesse est donc nulle. Ainsi, pour augmenter sa vitesse, il faut diminuer son temps de contact au sol en augmentant sa cadence.

Cependant, la projection du vecteur temps de contact au sol est assez proche du centre (coordonnées de la variable est de -0,35 sur l'axe 1) donc ce n'est pas un facteur indispensable à la vitesse. Son lien avec la vitesse est faible ne garantit pas nécessairement une augmentation de la performance.

RAPPORT VERTICAL.

Un autre facteur biomécanique contribue davantage à la vitesse, c'est le rapport vertical moyen. Sur le graphique des variables, le rapport vertical est différent de la vitesse. La corrélation entre le rapport et la vitesse est de -0,96. La contribution sur l'axe 1 est élevé (23%). Donc la vitesse est liée négativement à cet indicateur, sa qualité de représentation est élevée sur l'axe, on peut l'interpréter comme un facteur qui influe sur les variables. Le rapport vertical montre l'efficacité de la propulsion vers l'avant. Plus le rapport est faible et plus le corps convertie les forces sur le plan horizontal, ce qui engendre une augmentation de la vitesse. De plus, le mouvement du corps est porté vers l'avant et il y a moins de perte d'énergie vers le haut.

Ce rapport est mesuré à partir de l'oscillation verticale et de la longueur de la foulée. On peut expliquer le rapport vertical par la vitesse à l'aide de cette formule :

$$Rapport\ vertical = \frac{Oscillation\ verticale\ (mm)}{Longueur\ de\ foul\'ee\ (m)}$$

Grâce à cette formule, on voit que le rapport vertical varie en fonction de l'oscillation verticale et de la vitesse. Comme vu précédemment, pour augmenter sa vitesse, il faut diminuer le rapport vertical. Donc, la minimisation de l'oscillation verticale est indispensable pour augmenter sa vitesse. Pour réduire l'oscillation verticale, l'attaque du pied au moment de la réception doit se faire sous le bassin. En effet, l'oscillation sera amoindrie car l'alignement du bassin, du genou et de la cheville permettra d'amortir l'impact du pied au sol. La technique de course est donc un élément à ne pas négliger pour garantir un meilleur rendement lors de la propulsion.

c) Facteurs physiologiques FREQUENCE CARDIAQUE.

La fréquence cardiaque moyenne ainsi que la fréquence cardiaque maximale sont tous deux fortement corrélé (coefficient de corrélation = 0,74). Sur le graphique des variables, la fréquence cardiaque moyenne et maximale est très proche de la TE aérobie. En effet, 81% de la variabilité de la TE aérobie dépend de la fréquence cardiaque moyenne.

Pour augmenter la quantité d'oxygène apporté vers les muscles, il faut augmenter la FC. La FC ne peut augmenter indéfiniment. Une fois la FC maximale atteinte la quantité d'oxygène apportée aux muscles n'augmente plus. A chaque contraction, la quantité de sang (donc d'O2) amenée vers les muscles est plus grande. La FC est représentative du niveau de sollicitation du système aérobie.

CALORIES.

Pour se contracter, un muscle a donc besoin d'oxygène mais aussi de l'énergie fourni par les aliments dont les glucides, les lipides et les protéines. Grâce aux corrélations des variables, pour avoir une force de contraction des muscles, il faut une grande quantité d'énergie. La variable calorie a un coefficient de corrélation de 0,73 avec la TE aérobie. Les calories ont donc une influence positive sur la performance. Plus un individu réalise une activité physique, plus il dépense de l'énergie, donc son besoin en calories est conséquent. Pour avoir un rendement meilleur, le coureur doit améliorer sa foulée pour plus économique et donc gagner en vitesse.

d) Facteurs environnementaux / Conditions entraînements **TEMPERATURE EXTERIEURE.**

Sur l'axe 1, la température influe peu sur la vitesse (voir figure 7 ci-dessous) avec un pourcentage de contribution de 1%. Le vecteur de la température est proche du centre du cercle de corrélation (indépendance des autres variables actives). Le coefficient de corrélation avec la vitesse est de seulement 0,20. Donc on ne peut pas dire que la température a une influence significative sur la performance mais dans une certaine mesure, elle peut influer positivement sur la vitesse. Avec la chaleur, les muscles s'échauffent plus rapidement. Ils sont plus souples et le processus de contraction est amélioré.

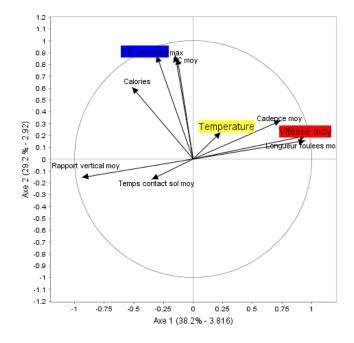


Figure 7 – La température est proche du centre

METEO.

Pour analyser la météo en ACP, nous avons choisi de la mettre en variable illustrative car elle est de type qualitatif. Aux vues des coordonnées sur les axes, les modalités soleil, nuageux et pluie montrent qu'elles ne participent pas à la construction des axes (voir figure 8 cidessous).

Libellé	Effectif	Poids	Distance à l'origine	Axe 1	Axe 2	Axe 3	Axe 4	Axe 5
Nuageux	40	40,000	0,805	0,037	-0,725	0,011	-0,327	0,043
Pluie	6	6,000	1,586	-0,283	-1,059	-0,863	-0,630	-0,377
Soleil	52	52,000	0,762	0,004	0,680	0,091	0,324	0,011

Figure 8 – Coordonnées des points moyens des modalités de la météo

La pluie a une légère influence opposée à la vitesse car sa coordonnée est négative sur l'axe 1 (-0,283).

MOMENT DE LA JOURNEE.

Le tableau ci-dessus *(voir figure 9 ci-dessous)* représente les coordonnées des points moyens des modalités des variables nominales.

Coordonnées des points moyens des modalités des variables nominales (illustratives)

Libellé	Effectif	Poids	Distance à l'origine	Axe 1	Axe 2	Axe 3	Axe 4	Axe 5
Mois								
Aout	7	7,000	2,543	1,995	1,316	0,625	0,319	0,430
Avril	14	14,000	0,782	-0,066	-0,620	-0,243	-0,114	-0,167
Fevrier	7	7,000	2,068	-1,563	0,333	-0,485	-1,061	0,485
Janvier	1	1,000	3,672	-2,471	-1,454	-1,678	0,075	1,504
Juillet	4	4,000	3,692	-3,112	0,469	1,081	1,486	0,528
Juin	11	11,000	1,937	1,169	0,957	0,934	0,301	-0,700
Mai	21	21,000	0,704	0,134	0,108	-0,457	0,264	-0,426
Mars	10	10,000	1,196	0,041	0,688	-0,524	-0,624	0,434
Novembre	6	6,000	2,121	0,135	-1,884	-0,684	-0,667	0,050
Octobre	6	6,000	2,022	-0,153	-1,638	0,823	-0,520	0,555
Septembre	11	11,000	0,922	-0,287	-0,168	0,318	0,480	0,089
Moment journee								
Apres-midi	22	22,000	1,012	0,048	-0,667	0,487	0,553	-0,028
Matin	19	19,000	0,887	-0,157	0,738	-0,002	-0,450	-0,014
Soir	57	57,000	0,222	0,034	0,011	-0,187	-0,063	0,015

Figure 9 – Coordonnées des points moyens des modalités mois et moment journée

On peut voir que les mois n'influent pas le coureur car les coordonnées des points moyens dans les axes 1 et 2 sont toujours proches de 0.

Il y a tout de même quelques facteurs qui peuvent expliquer la différence des courses de chaque mois comme : la fatigue musculaire, le manque de temps, le repos ou la préparation de course.

Mois.

Par exemple : on remarque que les courses de mai sont les plus élevées. Il s'agit d'un mois avec des entrainements intensifs pour préparer une course pour le mois de juin.

Concernant les moments de la journée, on peut voir que les coordonnées des points moyens sont aussi proches de 0. Cela signifie que les moments de la journée n'influent pas non plus la performance de l'athlète.

On remarque tout de même que les courses du soir sont plus fréquentes (57 activités) que les courses de l'après-midi (22 activités) et que celles du matin (19 activités). Cela s'explique car le fait que le coureur ne s'entraine que pendant son temps libre et qu'il n'y a pas de différence de performance.

TYPE DE CHAUSSURE.



Tableau comparatif des chaussures sur la longueur de la foulée moyenne

Dans le tableau ci-dessus, on compare les longueurs de foulée moyenne de chaque chaussure. On voit nettement que la Nike Zoom Fly Flyknit est beaucoup plus efficace pour la performance du coureur. Cela se remarque aussi sur le graphique des individus avec les variables illustratives (voir figure 10 ci-dessous).

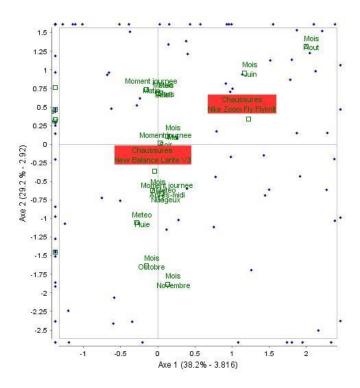


Figure 10 - Variables illustratives (chaussures)

La Nike se situe à droite de l'axe 1 (axe de la vitesse) avec une coordonnée de 1,22 sur l'axe 1 tandis que la New Balance a une coordonnée de -0,04 sur l'axe 1. Les deux modalités sont tout de même proches du centre donc elles ne contribuent pas beaucoup à la construction des axes mais la Nike a plus d'influence sur la vitesse.

Cela s'explique par les caractéristiques de la chaussure. En effet, il a été démontré que la Nike permet une économie de course de 4%. Cette innovation réside dans la semelle. En effet, celle-ci comporte des caractéristiques incomparables. Elle est ultra-légère et très réactive lors de la propulsion sur le sol. De plus, cette chaussure est dotée de la technologie Nike Zoom X (voir figure 11 ci-dessous) Figure 11 – Décomposition de la semelle intégrant une plaque de carbone dans la semelle



intermédiaire. Un réel avantage car la plaque offre un excellent retour d'énergie, supérieur à tous les autres modèles. Elle renforce aussi la propulsion à chaque foulée. Un drop de 5 millimètres (différence entre la semelle avant et arrière) permet une propulsion vers l'avant. Il faut savoir que le record du monde du Marathon a été fait à l'aide de ces chaussures.

6) Tableau récapitulatif des variables

Variables	Axe factoriel 1	Axe factoriel 2
Sens positif	Vitesse moyenne	TE aérobie
	Longueur de foulée	FC moyenne
	Cadence moyenne	FC maximale
	Température	Calories
		Vitesse moyenne
		Longueur de foulée
		Cadence moyenne
		Température
Sens négatif	Rapport vertical moyen	Rapport vertical moyen
	Temps de contact au sol	Temps de contact au sol
	moyen	moyen
	TE aérobie	
	FC moyenne	
	FC maximale	
	Calories	

Tableau récapitulatif des variables actives

V) Classification des activités

1) Classification Ascendante Hiérarchique

Dans cette partie, nous allons réaliser une classification avec la méthode de classification ascendante hiérarchique (CAH) dans le but de rassembler les activités similaires. Pour cela, nous utilisons l'algorithme de SPAD « Classification CAH » pour repérer les classes et observer les caractérisations. Ainsi, on pourra voir quelles sont les courses qui ont été les plus performantes.

Commençons par déterminer le nombre de classe où l'inertie intra-classe est minimisé et l'inertie inter-classe maximiser avec le critère de Ward. Voici le dendrogramme (voir figure 12 ci-dessous) qui va servir à choisir le nombre de classe à retenir.

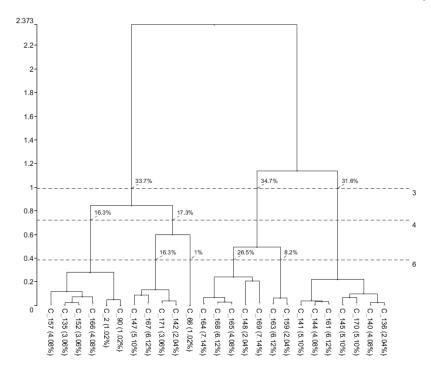


Figure 12 – CAH dendrogramme

Ce dendrogramme nous permet de trouver le nombre de classe qui va détenir la plus grande inertie possible tout en gardant une inertie intra faible et une inertie inter haute. De plus, il faut minimiser le nombre de classe pour faciliter l'interprétation.

D'après dendrogramme et les variances intra, inter-classes, SPAD choisi de conserver 6 classes d'un point de vue statistique. En effet, les indicateurs de qualité montrent une variance intra-classes de 2,239 et une variance inter-classes de 5,514 après consolidation. Par rapport aux autres classes, les inerties inter et intra sont meilleures que 3 ou 4. Donc la qualité de la classe 6 est bonne puisque qu'elle retient 70% de l'inertie totale (voir figure 13 ci-dessous). La différence d'inertie entre 4 classes et 6 classes (14%), alors que la différence d'inertie expliqué entre 3 classes et 4 classes est de 10%.

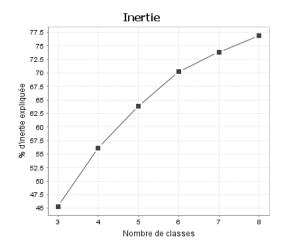


Figure 13 – Graphique de l'inertie des classes

En revanche, d'un point de vue métier, il est préférable de conserver 4 classes car les interprétations des groupes seront plus faciles. Sur le dendrogramme, on observe un saut de variance (voir figure 14 ci-dessous) entre la partition 3 et 4.

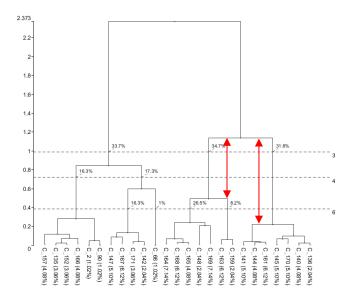


Figure 14 – CAH dendrogramme (saut important entre les classes 3 et 4)

De plus, les données sont réparties de façon homogène dans chaque classe de la partition 4.

- Classe 1 retient 16,3% des données
- Classe 2 retient 17,3% des données
- Classe 3 retient 34,7% des données
- Classe 4 retient 31,7% des données

Il est donc judiciable de se focaliser sur 4 classes.

2) Caractérisation des classes

Les activités réalisées au cours de l'année 2019 n'ont pas vraiment de grandes différences entre-elles. C'est pourquoi, sur le graphique des individus avec la partition en 4 classes (voir figure 15 ci-dessous), les classes se chevauchent légèrement.

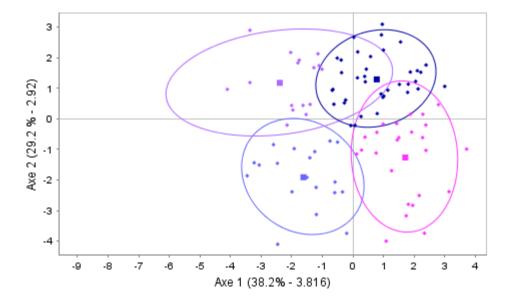


Figure 15 – Graphique des individus 4 classes (bleu clair : classe 1, violet : classe 2, rose : classe 3 et bleu foncé : classe 4)

Maintenant, nous allons interpréter les 4 classes de la classification en apportant une caractérisation pour chacune des classes.

La répartition des observations (activités) est homogène dans chaque classe. La classe 1 comporte 19% des données, la classe 2 détient 17%, la classe 3 a 26% des données et la classe 4 comprend 38% des activités.

a) Classe 1

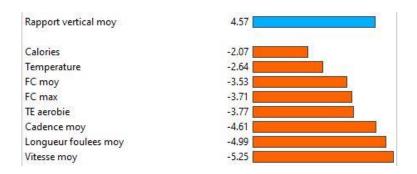


Figure 16 – Description de la classe 1 (active)

Dans cette classe, sont regroupés les activités avec une faible vitesse et dépense énergétique. La valeur-test de la moyenne est de -5,25 tandis que la valeur du rapport vertical moyen est de 4,57. La grandeur du rapport vertical montre qu'il a une influence négative sur la vitesse car plus le rapport vertical est haut et plus la vitesse est faible. De plus, les facteurs associés à la vitesse (longueur de foulée et cadence) diminuent quand ce rapport augmente. Les variables énergétiques diminuent aussi donc avoir un rapport vertical élevé ne s'apparente pas à des courses de vitesse et de performance. En termes de chiffres, la moyenne du rapport vertical moyen dans cette classe est de 80 mm, la vitesse est de 11

km/h et la fréquence cardiaque est de 151 bpm. Ces activités sont donc des entraînements à faible allure.

b) Classe 2



Figure 17 – Description de la classe 2 (active)

La classe 2 concerne les activités moyenne en termes de performance. Comme vu précédemment, le rapport vertical joue négativement sur la vitesse, la longueur de foulée et la cadence. La vitesse moyenne dans cette classe est de 11,5 km/h. Cette fois le rapport vertical est corrélé positivement avec les calories, l'aérobie et la fréquence cardiaque. Ces activités sont donc très physiques. La dépense énergétique est forte (FC moyenne de 168 et calories moyennes de 1141) alors que la vitesse reste faible. On peut donc dire que ses activités ressemblent à des courses de montagne sur des longues distances et à fort dénivelé. Afin d'améliorer la performance dans les courses de trail, le facteur que nous pouvons améliorer à court terme est de diminuer le rapport vertical. Pour cela, il faut améliorer sa technique de course et l'impact au sol en utilisant une meilleure posture et en renforçant ses chaînes musculaires.



Figure 18 – Description de la classe 2 (illustrative)

Le type de chaussure adapté à ce type de course est la Salomon UltraPro spécifique au trail. La valeur-test de la Salomon est de 5,4 donc elle participe à la performance en trail.

La Nike a un rôle contraire à la Salomon, elle joue positivement sur la vitesse mais ce n'est pas une chaussure conçue pour le trail car le maintien du pied est moins important. Le pourcentage de la modalité dans la classe 2 de la Nike est de 0%.

c) Classe 3

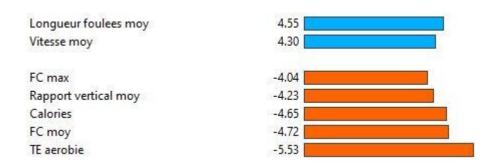


Figure 19 – Description de la classe 4 (active)

La classe 3 est lorsque la vitesse et la longueur de foulée sont élevées. Ces facteurs entraînent un faible rapport vertical. Ce sont des courses de courtes distances avec un dénivelé faible. Ce type d'entraînement est appelé le fractionnée qui alterne entre des allures rapides et des phases de repos. En effet, la FC moyenne reste assez faible (150 bpm) de même que la TE aérobie (2,7). Les courses à faible distance ont une fréquence cardiaque maximal de 176 bpm car c'est un effort qui se court avec une forte vitesse donc il faut apporter un maximum d'oxygène dans les muscles. Avec une grande longueur de foulée, les paramètres physiologiques diminuent tous. Donc il faut avoir une bonne puissance dans les muscles pour générer plus de vitesse et augmenter sa longueur de foulée. Il ne faut pas oublier la technique de la foulée car sans cette efficacité, on perd trop d'énergie.

d) Classe 4

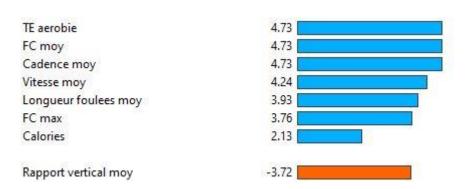


Figure 20 – Description de la classe 3 (active)

Les activités avec de la vitesse ont une grande longueur de foulée, une forte cadence et des facteurs énergétiques important ont un impact négatif sur le rapport vertical (valeur-test négative de -3,72). Plus le rapport est faible donc le corps propulsé vers l'avant et plus la vitesse est élevée. La technique de foulée joue sur la performance. Cette classe serait la meilleure au niveau du rendement énergétique car la fréquence cardiaque est moyennement élevée (167 bpm). La vitesse optimale est donc de 14 km/h. Il ne faut pas trop augmenter cette vitesse car le corps sera en surrégime. La technique de course est donc très

Comment les facteurs mécaniques, physiologiques et environnementaux influent sur la performance en course à pied ?

importante pour minimiser la perte d'énergie au moment de l'impact et utiliser l'élasticité des tendons.

Les chaussures, la météo, le moment de la journée et le mois n'influent pas sur des courses de performance. On peut être performant en courant soit le matin ou l'après-midi ou le soir.



Figure 21 – Description de la classe 4 (illustrative)

Le type de chaussure n'est pas représentatif dans cette classe sauf pour la Salomon qui ne s'apprête pas à ce type de course rapide.

3) Les caractéristiques d'une course performante

Quelles serait donc la classe qui favorisent la performance ? D'après la classification, la classe 4 présente les meilleures caractéristiques.

Variable caractéristique	Moyenne dans la classe		
TE aerobie	4,270		
FC moy	167,676		
Cadence moy	169,892		
Vitesse moy	14,143		
Longueur foulees moy	1,386		
FC max	189,892		
Calories	824,757		
Rapport vertical moy	67,997		

Figure 21 – Description de la classe 4 (illustrative)

VI) Synthèse de l'étude

Comme nous l'avons pu démontrer statistiquement, il y a trois grands types de facteurs, les facteurs mécaniques, facteurs physiologiques et les facteurs environnementaux. Il y a des facteurs qui sont plus indispensables que d'autres et il faut les améliorer pour une meilleure performance.

Pour les facteurs mécaniques, on a pu démontrer qu'il fallait une meilleure longueur de foulée pour avoir une bonne vitesse tout en dépensant moins d'énergie et donc moins de

calories, la cadence de course est aussi à améliorer pour une meilleure économie de course (utilisation de l'élasticité des tendons). Autre facteur à améliorer est la diminution du temps de contact au sol. Une minimaison de l'oscillation verticale est aussi indispensable pour la vitesse est l'amélioration du rapport vertical. Il est donc très important de travailler sa technique de course.

Concernant les facteurs physiologiques, il est conseillé au coureur d'améliorer sa fréquence cardiaque pour augmenter la quantité d'oxygène dans les muscles ainsi que son aérobie. Un bon apport en colories est important pour toutes activités physiques, il faut donc soigner son alimentation.

Pour en finir avec les facteurs environnementaux, ils n'ont pas beaucoup d'impact sur la performance d'un coureur, peu importe le mois ou le moment de la journée ou la météo lors de l'entrainement, la performance n'a pas de lien fort.

On a prouvé que le type de chaussure peut avoir un impact sur la performance grâce à ses caractéristiques comme la Nike Zoom Fly Vaporfly ou les chaussures de trail.

Il existe d'autres facteurs pouvant influencer la vitesse comme la capacité musculaire (explosivité, force, élasticité, souplesse), la fatigue due aux entraînements, le facteur mental (confiance en soi, motivation, gestion du stress) et le facteur social (hygiène de vie, ...).