

Projet - Analyse de données

Analyse des facteurs impactant le niveau d'un joueur de tennis



Table des matières

I – Les données de notre étude	3
II – Les résultats de notre sondage	4
III – Analyse du sondage	6
1 - Y a-t-il un lien entre le classement et le nombre d'année de pratique ?	7
2 – Y a-t-il un lien entre le classement et le nombre d'heures d'entraînement/semaine ?	12
3 – Y a-t-il un lien entre le classement et le nombre de matchs par an ?	16
4 - Ébauche d'ACM à l'aide XLSTAT	20
Bilan de notre projet.....	22



Dans le cadre de notre 2^{ème} année d'étude au sein d'EFREI Paris, nous avons réalisé un projet d'analyse de données dans le cadre du département de mathématiques.

Ce projet a pour but d'étudier les corrélations entre les différentes variables d'un sondage que nous avons réalisé nous-même. Cela nous permet d'effectuer, en situation réelle, une analyse de données complète et en donner une interprétation.

Ce rapport explicitera l'ensemble de notre travail, et comprendra donc la création du sondage, le choix du thème et des différentes variables de celui-ci, dans le but d'avoir le plus d'éléments intéressants à étudier sur un sujet qui l'est pour nous. Nous étudierons aussi le choix de la méthode d'analyse de ces données, pourquoi nous pensons que celle choisie est la plus adaptée pour ces données, ainsi que les résultats des différentes analyses que nous proposons.

I – Les données de notre étude

Nous sommes passionnés de sport et plus particulièrement de tennis. En effet, Benjamin est classé 15/5 et Pierrick est 3/6. Nous avons donc souhaité, dans le cadre de ce projet, récolter nos propres données et non de récupérer des données open source, directement disponibles sur internet.

Il fallait donc réfléchir à quelles variables pourrait potentiellement avoir des corrélations avec le niveau de tennis des individus, ou plus particulièrement avec leur classement.

Nous avons donc réalisé un sondage avec 14 facteurs (à l'aide de Google Forms) :

- Classement actuel,
- Meilleur classement,
- Sexe,
- Âge,
- Nombre d'années de pratique,
- Marque de la raquette,
- Type de cordage,
- Fréquence de casse du cordage,
- Tension du cordage de la raquette,
- Type de jeu,
- Nombre de raquettes,
- Nombre de matchs en 2019/2020 ou pendant la meilleur saison (en effet depuis 1 ans et demi le nombre de tournois a grandement été impacté par la pandémie),
- Nombre d'heures d'entraînement par semaine,
- Type d'entraînement (loisir, collectif, individuel).

Pour chaque variable, nous avons proposé différentes modalités. Par exemple, pour le classement nous avons listé l'ensemble des classements afin d'éviter des différences d'orthographe (30/5 et 30.5 par exemple) et des différences de valeurs pour pouvoir regrouper les informations de manière plus simple.

Nous avons par la suite mis en place une campagne de communication sur 4 jours afin de pouvoir analyser un grand nombre de variable pour obtenir une analyse significative. La difficulté a été de toucher tous les types de joueurs, et donc tous les types de classement afin que chaque niveau soit représenté de manière similaire.

Avec l'aide du club de l'Us Cagnes Tennis, de l'Union sportive de tennis de Villejuif, et du centre sportif Ultra ainsi que des contacts de Pierrick et Benjamin, nous avons pu obtenir 181 réponses sur les 4 jours d'ouverture du sondage. Le nombre de réponses étant suffisant pour avoir des données objectives, nous avons choisi de continuer dans cette voie.

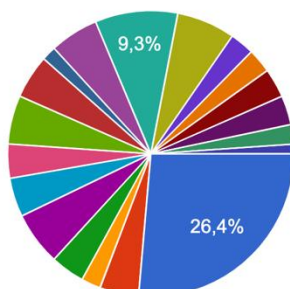
Il faut noter que le classement peut, dans certains cas, ne pas représenter le niveau de jeu d'un joueur. Il s'agit simplement d'un indicateur. En effet, une personne peut s'entraîner et garder un très bon niveau de jeu, sans pour autant participer à des tournois, et donc ne pas faire évoluer son classement. C'est pour cela que nous avons choisi d'exploiter le meilleur classement des joueurs et non leurs classements actuels. Nous verrons tout cela plus tard dans notre analyse.

II – Les résultats de notre sondage

Voici les résultats obtenus sur le sondage. Ces données sont extraites de Google Forms.

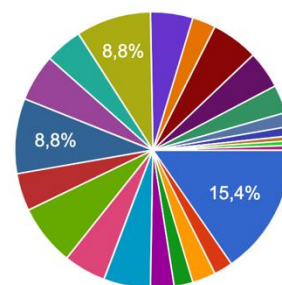
Mon classement actuel

182 réponses



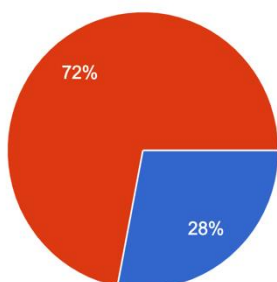
Mon meilleur classement

182 réponses



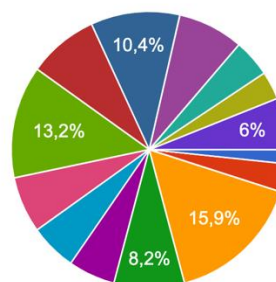
Sexe

182 réponses



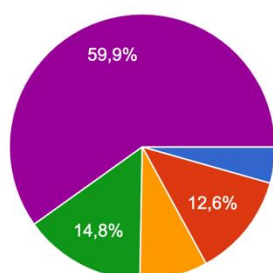
Age

182 réponses



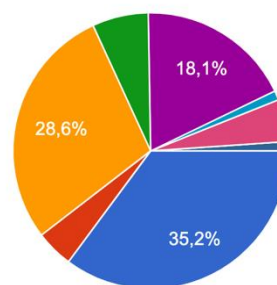
Fréquence de casse du cordage

182 réponses



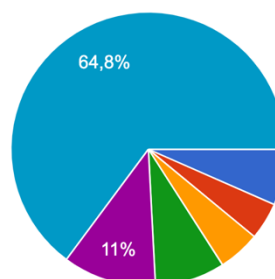
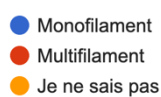
Marque de ma raquette

182 réponses



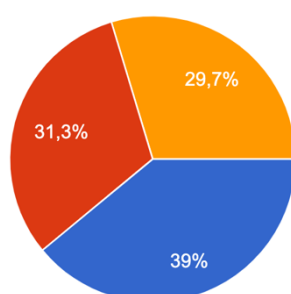
Nombre d'année(s) de pratique

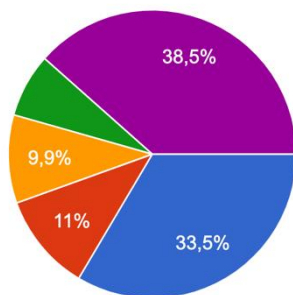
182 réponses



Type de cordage

182 réponses

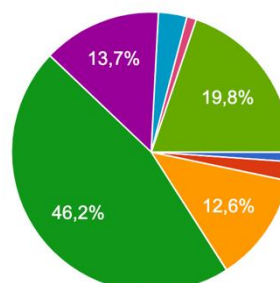




- Attaquant(e) de fond de cours
- Attaquant(e) volleyeur(se)
- Contreur(se)
- Défenseur(se)
- Polyvalent(e)

Tension de ma raquette

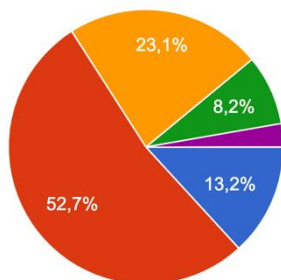
182 réponses



- Moins de 18 kg
- 18 - 20 kg
- 21 - 22 kg
- 23 - 24 kg
- 25 - 26 kg
- 27 - 28 kg
- Plus de 28 kg
- Je ne sais pas

Nombre de raquette(s)

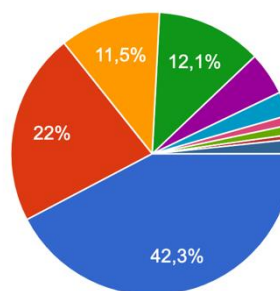
182 réponses



- 1
- 2
- 3
- 4
- Plus de 4

Nombre de match(s) en 2019/2020 ou à ma meilleure saison

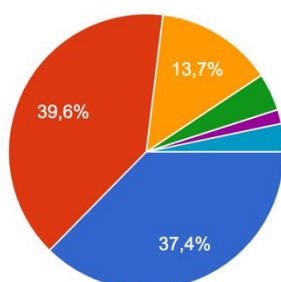
182 réponses



- Moins de 10
 - 10 - 20
 - 20 - 30
 - 30 - 40
 - 40 - 50
 - 50 - 60
 - 60 - 70
 - 70 - 80
- ▲ 1/2 ▼

Nombre d'heure(s) d'entraînement par semaine

182 réponses



- Moins de 2h
- 2h - 4h
- 4h - 6h
- 6h - 8h
- 8h - 10h
- Plus de 10h

III – Analyse du sondage

Le choix de notre méthode d'analyse

Une fois l'ensemble des données récoltées, nous avons dû choisir la méthode d'analyse de ces dernières : Analyse des Correspondances Principales (ACP) ou Analyse des factorielles des correspondances (AFC), deux méthodes que nous avons étudiées dans le cadre de notre module *Analyse de données*.

Comme les données que nous avons obtenues ne sont pas quantitatives (exemple : marque de la raquette ou encore des tranches d'heures), l'ACP n'aurait donc pas été une méthode correcte pour les analyser.

C'est donc pour avoir choisi de réaliser plusieurs AFC afin d'étudier le lien entre les différentes modalités de chaque facteur et les différents classements. Nous avons fait des tranches de 6 classements afin de regrouper les individus par niveau (Non-classé à 30/5, 30/4 à 30/2, 30/1 à 15/5, 15/4 à 15/2, 15/1 à 5/6 et 4/6 et plus – dans l'ordre croissant de niveau).

Il est cependant important de préciser que la meilleure des méthodes d'analyse de donnée pour analyser nos variables aurait été l'analyse en composante multiple (ACM). Avec cette méthode, nous aurions pu analyser l'ensemble des modalités de toutes nos variables.

En effet, l'analyse des correspondances multiples permet d'analyser les corrélations entre différentes variables de nature (qualitative ou quantitative) différentes. L'ACM s'applique donc de manière particulièrement efficace pour les sondages.

L'ACM est une généralisation de L'AFC. N'ayant pas étudié la méthode de l'ACM, nous avons décidé de continuer notre analyse avec 3 AFC en posant 3 problématiques :

- 1 – Y a-t-il un lien entre le classement et le nombre d'année de pratique ?**
- 2 – Y a-t-il un lien entre le classement et le nombre d'heures d'entraînement par semaine ?**
- 3 – Y a-t-il un lien entre le classement et le nombre de matchs par an ?**

L'AFC nous permettra d'étudier les correspondances entre les différentes modalités du classement et les autres modalités d'une autre variable. Nous avons dans une quatrième partie réalisé une ébauche d'ACM avec XLSTAT.

Nous analyserons ainsi seulement certaines variables de notre sondage.

1 - Y a-t-il un lien entre le classement et le nombre d'année de pratique ?

On effectue une AFC sur une population de joueurs de tennis, les deux variables étudiées sont le classement des joueurs et leur nombre d'années de pratique.

Avec les résultats de notre sondage, nous obtenons le tableau suivant :

	Plus de 11 ans	6 - 11 ans	0 - 6 ans	TOTAL
NC	5	7	15	27
40	2	1	1	4
30/5	0	2	3	5
30/4	0	3	1	4
30/3	2	2	1	5
30/2	6	1	3	10
30/1	9	0	0	9
30	8	5	0	13
15/5	8	0	0	8
15/4	10	5	1	16
15/3	6	2	2	10
15/2	7	1	0	8
15/1	12	3	1	16
15	9	0	0	9
5/6	4	1	0	5
4/6	9	1	0	10
3/6	8	0	0	8
2/6	5	1	0	6
1/6	3	0	0	3
0	2	0	0	2
Plus	3	0	0	3
TOTAL	118	35	28	181

On peut traduire ce tableau en le tableau de contingence suivant :

	Plus de 11 ans	6 - 11 ans	0 - 6 ans	TOTAL
NC - 30/5	7	10	19	36
30/4 - 30/2	8	6	5	19
30/1 - 15/5	25	5	0	30
15/4 - 15/2	23	8	3	34
15/1 - 5/6	25	4	1	30
4/6 et plus	30	2	0	32
TOTAL	118	35	28	181

En termes de fréquence, nous avons le tableau ci-dessous :

	Plus de 11 ans	6 - 11 ans	0 - 6 ans	TOTAL
NC - 30/5	0,04	0,06	0,10	0,20
30/4 - 30/2	0,04	0,03	0,03	0,10
30/1 - 15/5	0,14	0,03	0,00	0,17
15/4 - 15/2	0,13	0,04	0,02	0,19
15/1 - 5/6	0,14	0,02	0,01	0,17
4/6 et plus	0,17	0,01	0,00	0,18
TOTAL	0,65	0,19	0,15	1,00

Analysons la valeur en gras ci-dessus dans le tableau des fréquences. La fréquence de joueur classé 4/6 et plus dont leur nombre d'années de pratique est supérieure à 11 ans est de 17 %.

Ainsi, nous avons la problématique suivante : **le nombre d'années de pratique de tennis influe-t-il sur le classement du joueur ?**

Désormais, intéressons-nous aux profil-lignes (PFL) et profil-colonnes (PFC).

Réalisons les tableaux des profil-lignes et des profil-colonnes à partir des formules $\frac{k_{ij}}{k_i}$ et $\frac{k_{ij}}{k_j}$ respectivement pour les PFC et PFL.

Nous obtenons le tableau des PFL suivant :

Les PFL sont les fréquences des modalités en colonne par rapport au total de chaque ligne.

PFL	Plus de 11 ans	6 - 11 ans	0 - 6 ans	
NC - 30/5	0,19	0,28	0,53	1,00
30/4 - 30/2	0,42	0,32	0,26	1,00
30/1 - 15/5	0,83	0,17	0,00	1,00
15/4 - 15/2	0,68	0,24	0,09	1,00
15/1 - 5/6	0,83	0,13	0,03	1,00
4/6 et plus	0,94	0,06	0,00	1,00
PFL moyen	0,65	0,19	0,15	1,00

Analyse des PFL → On sélectionne les individus remarquables, qui se situent en dehors de l'intervalle de plus ou moins 10 % du PFL moyen.

Plus de 11 ans : [0,59 ; 0,72]

6 – 11 ans : [0,17 ; 0,21]

0 – 6 ans : [0,14 ; 0,17]

En rose, nous avons les individus sous-représentés et en vert, ceux qui sont sur-représentés.

Nous obtenons le tableau des PFC suivant :

Les PFC sont les fréquences de modalité en ligne par rapport au total de chaque colonne.

PFC	Plus de 11 ans	6 - 11 ans	0 - 6 ans	PFC moyen
NC - 30/5	0,06	0,29	0,68	0,20

30/4 - 30/2	0,07	0,17	0,18	0,10
30/1 - 15/5	0,21	0,14	0,00	0,17
15/4 - 15/2	0,19	0,23	0,11	0,19
15/1 - 5/6	0,21	0,11	0,04	0,17
4/6 et plus	0,25	0,06	0,00	0,18
	1,00	1,00	1,00	1,00

Analyse des PFC → On sélectionne les individus remarquables, qui se situent en dehors de l'intervalle de plus ou moins 10 % du PFC moyen.

NC – 30/5 : [0,18 ; 0,22]

30/4 – 30/2 : [0,09 ; 0,11]

30/1 – 15/5 : [0,15 ; 0,19]

15/4 – 15/2 : [0,17 ; 0,21]

15/1 – 5/6 : [0,15 ; 0,19]

4/6 et plus : [0,16 ; 0,20]

En rose, nous avons les individus sous-représentés et en vert, ceux qui sont surreprésentés.

En analysant les deux tableaux de PFC et PFL, on en tire les conclusions suivantes :

Il y a une surreprésentation des individus ayant peu d'expérience avec un classement faible.

Il y a une surreprésentation des individus ayant beaucoup d'expérience avec un classement fort.

Il y a une sous-représentation des individus ayant peu d'expérience avec un classement fort.

Il y a une sous-représentation des individus ayant beaucoup d'expérience avec un classement faible.

On peut calculer les distances du Khi-2 selon le PFC et le PFL. On obtient les tableaux suivants que l'on peut interpréter :

Distance entre les PFC j et k

$$d_{\chi^2}^2 = \sum_{i=1}^n \frac{1}{f_{i\bullet}} \left(\frac{f_{ij}}{f_{\bullet j}} - \frac{f_{ik}}{f_{\bullet k}} \right)^2 \quad j, k = 1, \dots, p$$

Distance Khi-2 colonnes	Plus de 11 ans	6 - 11 ans	0 - 6 ans
Plus de 11 ans	0	0,820	1,706
6 - 11 ans	0,820	0	1,017
0 - 6 ans	1,706	1,017	0

On observe déjà que les personnes pratiquant le tennis depuis plus de 11 ans sont plus proches des personnes le pratiquant déjà depuis au moins 6 ans ce qui semble cohérent.

Distance entre les PFL i et l

$$d_{\chi^2}^2 = \sum_{j=1}^p \frac{1}{f_{\bullet j}} \left(\frac{f_{ij}}{f_{i\bullet}} - \frac{f_{lj}}{f_{l\bullet}} \right)^2 \quad i, l = 1, \dots, n$$

Distance Khi-2 lignes	NC - 30/5	30/4 - 30/2	30/1 - 15/5	15/4 - 15/2	15/1 - 5/6	4/6 et plus
NC - 30/5	0	0,734	1,578	1,271	1,521	1,699
30/4 - 30/2	0,734	0	0,907	0,576	0,880	1,090
30/1 - 15/5	1,578	0,907	0	0,335	0,114	0,270
15/4 - 15/2	1,271	0,576	0,335	0	0,333	0,556
15/1 - 5/6	1,521	0,880	0,114	0,333	0	0,223
4/6 et plus	1,699	1,090	0,270	0,556	0,223	0

Ici aussi on observe des résultats logiques, les classements proches entre eux ont des distances plus proches.

Les tableaux des PFL et PFC sont de mêmes dimensions. Ainsi, de manière arbitraire, nous réaliserons l'AFC sur le tableau des PFL.

On calcule $T = X^t * X$ avec $X = \frac{f_{ij}}{\sqrt{f_i f_j}}$ avec $1 \leq i \leq n$ et $1 \leq j \leq p$

				0,11	0,31	0,60
				0,16	0,22	0,25
	0,75	0,30	0,15	0,42	0,17	0
On obtient donc T =	0,30	0,23	0,25	0,27	0,21	0,12
	0,15	0,25	0,43	0,42	0,11	0,06
				0,50	0,05	0

avec : $X =$

Le nombre de valeurs propres obtenu est égal au nombre de lignes ou de colonnes de notre matrice T soit 3 valeurs propres dont 1 qui est la valeur propre triviale égale à 1.

On a $\lambda_1 = 1$; $\lambda_2 = 0,393$; $\lambda_3 = 0,017$

On peut donc calculer l'inertie totale $R = \sum \lambda_i - 1$

$$R = \lambda_1 + \lambda_2 + \lambda_3 - 1 = 1 + 0,393 + 0,017 - 1 = 0,41$$

On peut donc calculer le Khi-2 afin de tester l'indépendance des modalités.

Soit l'hypothèse H_0 qui dit « il y a une indépendance entre le nombre d'années de pratique et le classement du joueur. »

Notre effectif total est $N = 181$.

$$\chi_{obs}^2 = N * R = 181 * 0,41 = 74,21$$

On teste l'indépendance de ces deux variables au seuil de risque de 5 %.

Calculons le degré de liberté : $DDL = (nbLignes - 1) \times (nbColonnes - 1)$

$$\text{On a : } (6 - 1) \times (3 - 1) = 10$$

D'après la table, le $\chi_{0,95}^2 = 18,3 < \chi_{obs}^2 = 74,21$

Ainsi, on rejette l'hypothèse H_0 : on peut affirmer **qu'il y a un lien entre le classement d'un joueur et son nombre d'années d'expérience**. La probabilité de se tromper ici est inférieure à 5 %. L'écart observé est significatif.

On calcule la qge à l'inertie totale (qualité de représentation de l'information par les nouveaux axes) :

$$\text{Contribution de l'axe 1 à l'inertie totale : } \frac{\lambda_2 * 100}{R} = \frac{0,393 * 100}{0,41} = 95,85 \%$$

$$\text{Contribution de l'axe 2 à l'inertie totale : } \frac{\lambda_3}{R} * 100 = \frac{0,017}{0,41} * 100 = 4,15 \%$$

On néglige par la suite l'axe 2 car la perte d'information est inférieure à 5%.

Avec les formules suivantes :

Les coordonnées des points lignes i sur l'axe α sont données par : $coord_{\alpha}(i) = \sum_{j=1}^p \frac{f_i}{f_i \sqrt{f_j}} * u_{\alpha j}$

Avec p : le nombre de colonne

$u_{\alpha j}$: le vecteur propre associé

Les coordonnées des points colonnes j se calculent de la même manière en inversant i par j .

La qualité de représentation de l'individu i sur l'axe α est $qlt_{\alpha}(i) = \frac{coord_{\alpha}^2(i)}{\sum_{\alpha=1}^{p-1} coord_{\alpha}^2(i)}$ dans le cas d'un point ligne. Pour un point colonne, on fait le même calcul en remplaçant i par j .

La contribution de l'individu i sur l'axe α est $cr_{\alpha}(i) = \frac{f_i * coord_{\alpha}^2(i)}{\lambda_{\alpha}}$

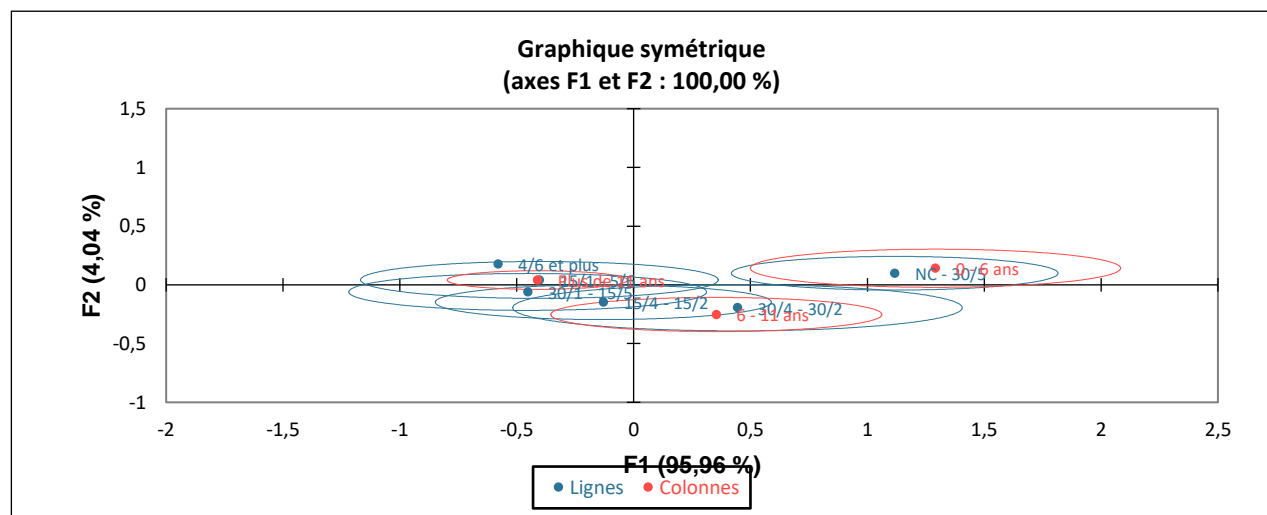
On obtient le tableau des composantes principales des lignes :

	Axe 1			Axe 2		
	coord	qlt	cr	coord	qlt	cr
NC - 30/5	1,117	0,992	0,632	0,099	0,008	0,118
30/4 - 30/2	0,444	0,839	0,053	-0,195	0,161	0,241
30/1 - 15/5	-0,453	0,983	0,087	-0,060	0,017	0,036
15/4 - 15/2	-0,129	0,434	0,008	-0,148	0,566	0,248
15/1 - 5/6	-0,403	0,989	0,069	0,042	0,011	0,018
4/6 et plus	-0,580	0,914	0,152	0,178	0,086	0,338

Et le tableau des composantes principales des colonnes :

	Axe 1			Axe 2		
	coord	qlt	cr	coord	qlt	cr
Plus de 11 ans	-0,411	0,990	0,281	0,041	0,010	0,067
6 - 11 ans	0,354	0,663	0,062	-0,252	0,337	0,745
0 - 6 ans	1,291	0,988	0,657	0,142	0,012	0,188

On sélectionne en bleu Les individus ayant une qlt importante et des coordonnées extrêmes.



En analysant les résultats obtenus, comme évoqué lors de l'étude du choix des axes, l'axe 2 est négligeable car il représente une part très faible de l'information.

On remarque que les individus souvent moins bien classés ont beaucoup moins d'années d'expérience tandis que les joueurs les mieux classés ont beaucoup plus d'années de pratique.

2 – Y a-t-il un lien entre le classement et le nombre d'heures d'entraînement/semaine ?

Pour cette analyse nous utiliserons les mêmes méthodes que précédemment. Tout d'abord, voici le tableau de donnée obtenu en croisant les données du classement et celles du nombre d'heures d'entraînement par semaine :

	Moins de 2h	2h - 4h	4h - 6h	6h-8h	Plus de 8h	Total
NC - 30/5	15	17	3	0	1	36
30/4 - 30/2	9	6	1	1	2	19
30/1 - 15/5	12	12	5	1	0	30
15/4 - 15/2	16	9	7	2	0	34
15/1 - 5/6	3	14	8	3	2	30
4/6 et plus	13	13	1	1	4	32
TOTAL	68	71	25	8	9	181

En termes de fréquence, nous obtenons alors le tableau suivant :

	Moins de 2h	2h - 4h	4h - 6h	6h-8h	Plus de 8h	Total
NC - 30/5	0,08	0,09	0,02	0,00	0,01	0,20
30/4 - 30/2	0,05	0,03	0,01	0,01	0,01	0,10
30/1 - 15/5	0,07	0,07	0,03	0,01	0,00	0,17
15/4 - 15/2	0,09	0,05	0,04	0,01	0,00	0,19
15/1 - 5/6	0,02	0,08	0,04	0,02	0,01	0,17
4/6 et plus	0,07	0,07	0,01	0,01	0,02	0,18
TOTAL	0,38	0,39	0,14	0,04	0,05	1,00

Ainsi, nous pouvons nous poser la question suivante : **le classement d'un joueur influe-t-il sur son nombre d'heures d'entraînement par semaine ?**

Pour cela, nous allons nous intéresser aux **profils-colonnes (PFC)** et aux **profils lignes (PFL)**. Nous obtiendrons dans un premier temps le tableau de PFC suivant :

Profil-colonne

PFC	Moins de 2h	2h - 4h	4h - 6h	6h-8h	Plus de 8h	PFC Moyen
NC - 30/5	0,22	0,24	0,12	0,00	0,11	0,20
30/4 - 30/2	0,13	0,08	0,04	0,13	0,22	0,10
30/1 - 15/5	0,18	0,17	0,20	0,13	0,00	0,17
15/4 - 15/2	0,24	0,13	0,28	0,25	0,00	0,19
15/1 - 5/6	0,04	0,20	0,32	0,38	0,22	0,17
4/6 et plus	0,19	0,18	0,04	0,13	0,44	0,18
PFC Total	1,00	1,00	1,00	1,00	1,00	1,00

Puis nous obtiendrons le tableau des PFL ci-dessous :

PFL	Moins de 2h	2h - 4h	4h - 6h	6h-8h	Plus de 8h	PFL Total
NC - 30/5	0,42	0,47	0,08	0,00	0,03	1,00
30/4 - 30/2	0,47	0,32	0,05	0,05	0,11	1,00
30/1 - 15/5	0,40	0,40	0,17	0,03	0,00	1,00
15/4 - 15/2	0,47	0,26	0,21	0,06	0,00	1,00
15/1 - 5/6	0,10	0,47	0,27	0,10	0,07	1,00
4/6 et plus	0,41	0,41	0,03	0,03	0,13	1,00
PFL Moyen	0,38	0,39	0,14	0,04	0,05	1,00

Analyse des profils lignes et colonnes :

Nous aurons en **vert** les individus surreprésentés (+ de 10 % du PFC ou PFL Moyen) et en **violet** ceux qui sont sous-représentés (- 10 % du PFC ou PFL Moyen)

Nous ne remarquons pas de tendances particulières ici, on peut donc se demander si le nombre d'heures d'entraînement par semaine n'aurait en fait aucune influence sur le classement.

Pour cela, nous allons calculer le Khi-2. Pour ce faire, commençons par calculer nos matrices T et X de la même manière que précédemment.

On obtient la matrice X suivante :

$$X = \begin{pmatrix} 0,30 & 0,34 & 0,10 & 0,00 & 0,06 \\ 0,25 & 0,16 & 0,05 & 0,08 & 0,15 \\ 0,27 & 0,26 & 0,18 & 0,06 & 0,00 \\ 0,33 & 0,18 & 0,24 & 0,12 & 0,00 \\ 0,07 & 0,30 & 0,29 & 0,19 & 0,12 \\ 0,28 & 0,27 & 0,04 & 0,06 & 0,24 \end{pmatrix}$$

Puis nous obtenons la matrice T :

$$T = \begin{pmatrix} 0,42 & 0,37 & 0,20 & 0,11 & 0,13 \\ 0,37 & 0,41 & 0,23 & 0,13 & 0,14 \\ 0,20 & 0,23 & 0,19 & 0,10 & 0,06 \\ 0,11 & 0,13 & 0,10 & 0,07 & 0,05 \\ 0,13 & 0,14 & 0,06 & 0,05 & 0,10 \end{pmatrix}$$

On en ressort les valeurs propres suivantes :

$$\lambda_1 = 1 \quad \lambda_2 = 0,095 \quad \lambda_3 = 0,061 \quad \lambda_4 = 0,0230 \quad \lambda_5 = 0$$

$$\text{On calcule ensuite l'inertie totale } R = \sum_2^5 \lambda_i = 0,179$$

On va réaliser le test du Khi-2 pour étudier la corrélation entre le classement et le nombre d'heures d'entraînement par semaine. On va donc poser l'hypothèse H0 : « Les lignes et les colonnes du tableau sont indépendantes »

On calcule ensuite le $\chi_{obs}^2 = 181 * R = 32,34$ et on teste ensuite l'indépendance de ces deux variables au seuil de risque de 5 %

Pour cela, nous avons besoin de notre DDL = $(n-1)(p-1) = 5 * 4 = 20$. Nous trouvons ainsi $\chi^2_{0,95} = 31,41$ en regardant la table.

On remarque que $\chi^2_{0,95} < \chi^2_{obs}$, ainsi nous pouvons dire que l'hypothèse H_0 est rejetée et qu'il **existe un lien entre le classement d'un joueur et son nombre d'heures d'entraînement par semaine**. La probabilité de se tromper ici est inférieure à 5 %.

On calcule ensuite les contributions de chaque axe et on remarque que les axes 1 et 2 sont suffisants car ils représentent à eux deux 87% de l'information.

	F1	F2	F3	F4
Inertie (%)	53,170	33,934	12,808	0,087
% cumulé	53,170	87,105	99,913	100,000

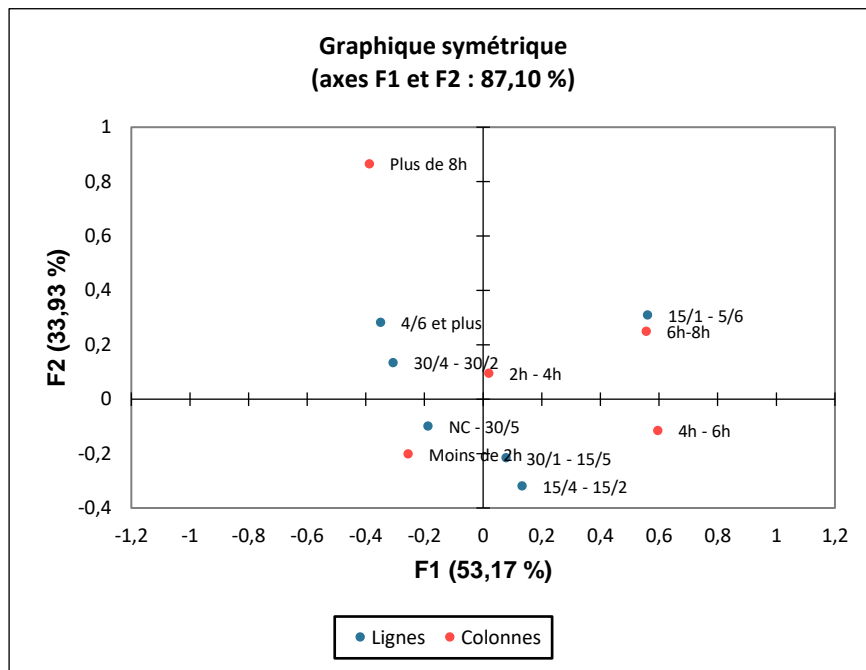
On obtient le tableau des composantes principales des lignes :

	Axe1			Axe 2		
	coord.	qlt	contribution	coord.	qlt	contribution
NC - 30/5	-0,187	0,363	0,073	-0,099	0,101	0,032
30/4 - 30/2	-0,306	0,597	0,104	0,134	0,115	0,031
30/1 - 15/5	0,078	0,103	0,011	-0,216	0,778	0,128
15/4 - 15/2	0,133	0,116	0,035	-0,319	0,663	0,315
15/1 - 5/6	0,562	0,765	0,55	0,31	0,233	0,262
4/6 et plus	-0,349	0,6	0,227	0,282	0,391	0,232

On obtient le tableau des composantes principales des colonnes :

	Axe1			Axe 2		
	coord.	qlt	contribution	coord.	qlt	contribution
- 2h	-0,255	0,578	0,257	-0,201	0,063	0,251
2h - 4h	0,02	0,011	0,002	0,096	0,75	0,059
4h - 6h	0,597	0,955	0,518	-0,116	0,008	0,031
6h-8h	0,557	0,606	0,144	0,25	0,269	0,046
Plus de 8h	-0,387	0,158	0,078	0,865	0,051	0,614

On sélectionne en vert les individus ayant une grande qlt et des coordonnées extrêmes.



On remarque grâce à ce graphique que les classements les plus forts ont en général une corrélation avec des heures d'entraînement plutôt grande.

Les joueurs classés plus de 4/6 s'entraînent soit entre 2h et 4h soit plus de 8h.

Les joueurs classés entre 15/1 et 5/6 ont s'entraînent en majorité entre 7 et 8h par semaine.

⇒ Les joueurs un peu moins bien classé (15/2 à NC) s'entraînent en général moins de 2h.

Cependant, on observe de manière surprenante que les 30/2-30/4 s'entraînent presque autant que les 4/6 et plus.

Ainsi, notre analyse nous a montré que la quantité d'heure d'entraînement par semaine n'est peut-être pas un des facteurs les plus importants concernant l'entraînement sportif. On peut donc en déduire que l'efficacité de l'entraînement impacte plus que la quantité. (Une analyse que nous n'avons pas décidé de retenir dans ce rapport, n'ayant pas eu assez de réponse pour faire ressortir une dépendance entre ces variables.)

3 – Y a-t-il un lien entre le classement et le nombre de matchs par an ?

Pour cette analyse nous utiliserons les mêmes méthodes que précédemment. Tout d'abord, voici le tableau de donnée obtenu en croisant les données du classement et celles du nombre de match par ans :

	Moins de 10	10--20	20--30	30--40	40--60	Plus de 60	Total
NC - 30/5	27	5	0	1	0	3	36
30/4 - 30/2	13	1	2	1	2	0	19
30/1 - 15/5	15	7	3	3	1	1	30
15/4 - 15/2	13	7	6	5	2	1	34
15/1 - 5/6	5	8	7	6	2	2	30
4/6 et plus	3	12	3	6	7	1	32
Total	76	40	21	22	14	8	181

On obtient ainsi le tableau des fréquences suivants :

	Moins de 10	10--20	20--30	30--40	40--60	Plus de 60	Total
NC - 30/5	0,15	0,03	0,00	0,01	0,00	0,02	0,20
30/4 - 30/2	0,07	0,01	0,01	0,01	0,01	0,00	0,10
30/1 - 15/5	0,08	0,04	0,02	0,02	0,01	0,01	0,17
15/4 - 15/2	0,07	0,04	0,03	0,03	0,01	0,01	0,19
15/1 - 5/6	0,03	0,04	0,04	0,03	0,01	0,01	0,17
4/6 et plus	0,02	0,07	0,02	0,03	0,04	0,01	0,18
Total	0,42	0,22	0,12	0,12	0,08	0,04	1,00

On s'intéressera comme pour les analyses précédentes au profil-ligne et au profil-colonne :

Profil-ligne :

PFL	Moins de 10	10--20	20--30	30--40	40--60	Plus de 60	
NC - 30/5	0,75	0,14	0,00	0,03	0,00	0,08	1
30/4 - 30/2	0,68	0,05	0,11	0,05	0,11	0,00	1
30/1 - 15/5	0,50	0,23	0,10	0,10	0,03	0,03	1
15/4 - 15/2	0,38	0,21	0,18	0,15	0,06	0,03	1
15/1 - 5/6	0,17	0,27	0,23	0,20	0,07	0,07	1
4/6 et plus	0,09	0,38	0,09	0,19	0,22	0,03	1
PFL-moyen	0,43	0,21	0,12	0,12	0,08	0,04	1

Analyse des PFL → On sélectionne les individus remarquables, qui se situent en dehors de l'intervalle de plus ou moins 10 % du PFL moyen.

- Moins de 10 matchs : [0,39 ; 0,47]
- 10-20 matchs : [0,19 ; 0,23]
- 20-30 matchs : [0,11 ; 0,13]
- 30-40 matchs : [0,11 ; 0,13]
- 40-60 matchs : [0,07 ; 0,09]
- Plus de 60 matchs : [0,036 ; 0,044]

En rose, nous avons les individus sous-représentés et en vert, ceux qui sont surreprésentés.

Profil-colonne :

PFC	Moins de 10	10--20	20--30	30--40	40--60	Plus de 60	PFC-moyen
NC - 30/5	0,36	0,13	0,00	0,05	0,00	0,38	0,15
30/4 -30/2	0,17	0,03	0,10	0,05	0,14	0,00	0,08
30/1 -15/5	0,20	0,18	0,14	0,14	0,07	0,13	0,14
15/4 -15/2	0,17	0,18	0,29	0,23	0,14	0,13	0,19
15/1 - 5/6	0,07	0,20	0,33	0,27	0,14	0,25	0,21
4/6 et plus	0,04	0,30	0,14	0,27	0,50	0,13	0,23
	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Analyse des PFC → On sélectionne les individus remarquables, qui se situent en dehors de l'intervalle de plus ou moins 10 % du PFC moyen.

- NC – 30/5 : [0,135 ; 0,165]
- 30/4 – 30/2 : [0,07 ; 0,09]
- 30/1 – 15/5 : [0,13 ; 0,15]
- 15/4 – 15/2 : [0,17 ; 0,21]
- 15/1 – 5/6 : [0,19 ; 0,23]
- 4/6 et plus : [0,21 ; 0,25]

En rose, nous avons les individus sous-représentés et en vert, ceux qui sont surreprésentés.

En analysant les deux tableaux de PFC et PFL, on en tire les conclusions suivantes :

- Il y a une surreprésentation des individus jouant peu de matchs avec un classement faible.
- Il y a une surreprésentation des individus jouant entre 10 et 60 matchs par an avec un classement fort.
- Il y a une sous-représentation des individus jouant peu de matchs avec un classement fort.
- Il y a une sous-représentation des individus avec un classement moyen jouant beaucoup de matchs.

Comme pour les analyses précédentes, on calcul T et X :

T=	0.55	0.23	0.16	0.16	0.10	0.14	X=	0.51	0.13	0	0.04	0	0.18
	0.23	0.25	0.17	0.19	0.16	0.10		0.35	0.04	0.1	0.05	0.13	0
	0.16	0.17	0.16	0.15	0.10	0.06		0.31	0.2	0.12	0.12	0.02	0.06
	0.16	0.19	0.15	0.15	0.12	0.06		0.25	0.19	0.22	0.18	0.09	0.06
	0.11	0.16	0.10	0.12	0.14	0.04		0.1	0.23	0.28	0.23	0.1	0.13
	0.14	0.10	0.06	0.06	0.04	0.06		0.06	0.33	0.11	0.22	0.33	0.06

Le nombre de valeurs propres obtenu est égal au nombre de lignes ou de colonnes de notre matrice T soit 6 valeurs propres dont 1 qui est la valeur propre triviale égale à 1.

On a $\lambda_1 = 1$; $\lambda_2 = 0.266$; $\lambda_3 = 0.052$; $\lambda_4 = 0.028$; $\lambda_5 = 0.005$; $\lambda_6 = 0$;

On peut donc calculer l'inertie totale $R = \sum \lambda_i - 1$
 $R = 1 + 0.266 + 0.052 + 0.028 + 0.005 + 0 - 1 = 0.351$

On va ici aussi réaliser un test du Khi-2 pour étudier la corrélation entre le classement et le nombre de matchs par an. On pose l'hypothèse H_0 : "Le classement du joueur est indépendant du nombre de matchs joués par ce joueur par an."

Ensuite, on calcule le Khi-2 de la même manière que précédemment, on trouve :

$$\chi_{obs}^2 = 63.403$$

De même on a un DDL = $5 \times 5 = 25$.

D'après la table on a :

$$\chi_{0,95}^2 = 37,7 < \chi_{obs}^2 = 63,403$$

Ainsi l'hypothèse H_0 est rejetée, **il y a une corrélation entre le nombre de matchs joués par an et le classement du joueur**. La probabilité de se tromper ici est inférieure à 5 %. L'écart observé est significatif.

Taux de contribution des axes à l'inertie totale :

Taux de contributions axe 1 : $\frac{0.266 \times 100}{0.351} = 75.8\%$.

Taux de contribution axe 2 : $\frac{0.052 \times 100}{0.351} = 14.8\%$.

Contribution axe1 + axe2 = $14.8 + 75.8 = 90.6\%$.

On peut désormais calculer les différents tableaux des composantes principales.

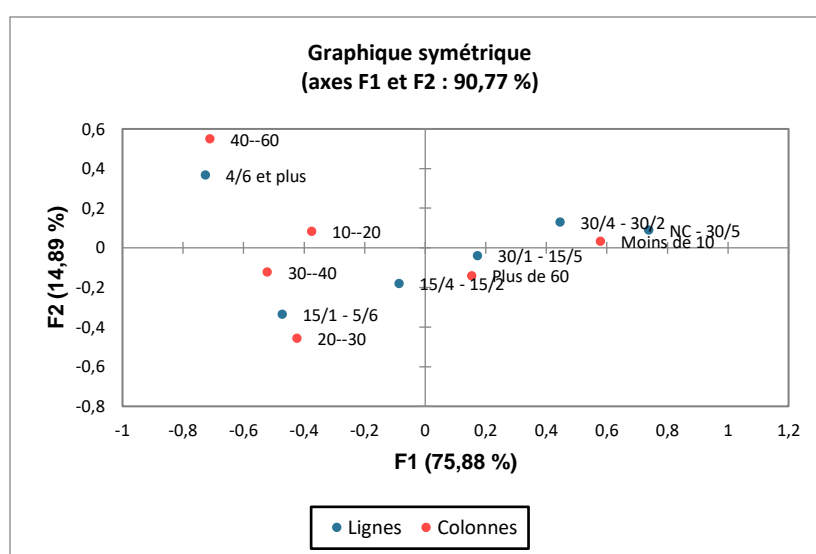
On obtient le tableau des composantes principales des lignes :

	Axe 1			Axe 2		
	Coord	qlt	cr	coord	qlt	cr
NC - 30/5	0,739	0,925	0,409	0,089	0,013	0,030
30/4 - 30/2	0,446	0,512	0,079	0,129	0,043	0,034
30/1 - 15/5	0,173	0,605	0,019	-0,040	0,032	0,005
15/4 - 15/2	-0,085	0,144	0,005	-0,181	0,645	0,118
15/1 - 5/6	-0,472	0,647	0,139	-0,335	0,326	0,357
4/6 et plus	-0,726	0,794	0,350	0,367	0,203	0,457

On obtient le tableau des composantes principales des colonnes :

	Axe 1			Axe 2		
	Coord	qlt	cr	coord	qlt	cr
Moins de 10	0,579	0,990	0,530	0,032	0,003	0,008
10--20	-0,375	0,768	0,117	0,082	0,037	0,029
20--30	-0,423	0,416	0,078	-0,457	0,485	0,464
30--40	-0,521	0,946	0,124	-0,122	0,052	0,035
40--60	-0,710	0,579	0,147	0,549	0,346	0,447
Plus de 60	0,154	0,068	0,004	-0,141	0,057	0,017

On sélectionne en bleu les individus ayant une qlt importante et des coordonnées extrêmes. On peut négliger l'axe 2 car les valeurs n'y sont pas bien représentées. L'axe 1 portant 75% de l'information.



Conclusion

On observe un groupement entre les joueurs 4/6 et plus autrement dit haut-classés font 40 à 60 matchs par an. On observe également que les joueurs les moins bien classés entre NC et 15/5 jouent moins de 10 matchs par an.

Cette observation est logique car pour augmenter son classement il faut gagner des matchs en tournoi pour accumuler un certain nombre de point (qui dépend du classement de l'adversaire du match gagné). Ainsi en réalisant plus de match, il y a plus de chance de gagner contre des joueurs étant un peu mieux classés que soi.

Cependant, étonnement, on observe que certains joueurs moins bien classés réalisent beaucoup de match (plus de 60).

On remarque également que les joueurs 15/1 - 5/6 se comportent de manière différente, ils réalisent entre 10 et 40 matchs par an.

Comme évoqué au début de ce rapport, dans certains cas, le classement peut ne pas représenter le niveau de jeu d'un joueur. Il s'agit simplement d'un indicateur. En effet, une personne peut s'entraîner et garder un très bon niveau de jeu, sans pour autant participer à des tournois, et donc ne pas faire évoluer son classement.

4 - Ébauche d'ACM à l'aide XLSTAT (hors programme)

Le nombre de modalité et le nombre d'individus est très grand, l'ACM est alors la méthode la plus appropriée dans ce cas. Cependant, cette méthode d'analyse est hors programme. Nous allons cependant en faire ici une rapide ébauche à l'aide de XLSTAT.

Tout d'abord on code les données en remplaçant les modalités des variables par des chiffres. Ensuite, on associe chaque modalité à une variable ne prenant que des valeurs binaires en fonction que la modalité soit vérifiée pour chaque individu. En regroupant cela on obtient le tableau disjonctif complet suivant :

Individus	Classement				Fréquence de jeu				etc ...
	NC	30/5	30/4		-15	- de 2h	2h - 4h	+ de 10h	
1	1	0	0		0	1	0	0	
2	0	0	0		1	0	0	0	
3	0	0	1	...	0	0	0	1	
4	0	1	0		0	0	0	1	
5	0	1	0		1	0	0	0	
6	0	0	1		0	1	0	0	
				...					
181	1	0	0		1	0	0	0	

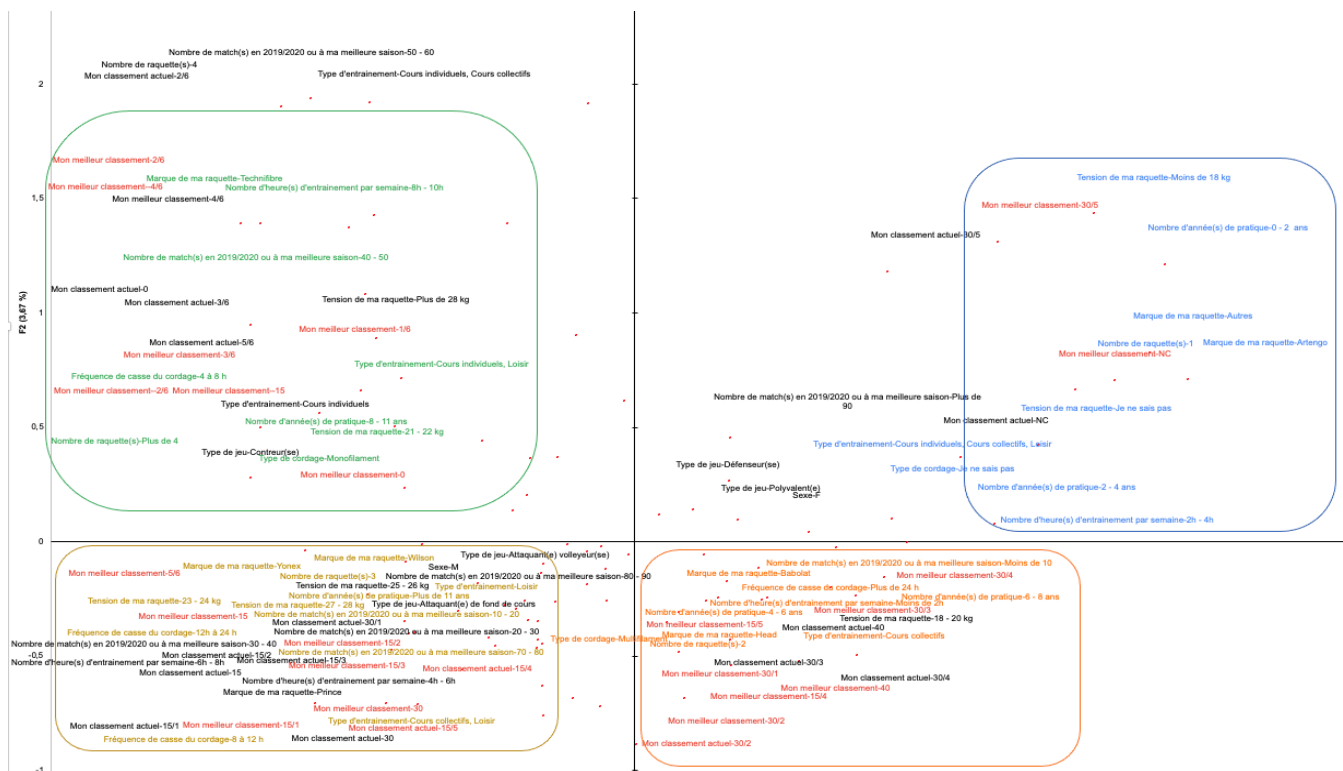
Par la suite, comme lors d'une étude AFC, on aurait pu analyser les distances entre les profils lignes et les profils colonne. Or, on ne s'intéresse pas ici à l'indépendance des variables.

Puis, on calcule les fréquences marginales de chaque modalité.

On diagonalise ensuite le tableau disjonctif complet afin d'obtenir les valeurs propres. Grâce à XLSTAT, on obtient 95 valeurs propres. En déterminant le port de l'information par chaque modalité et la contribution de chaque modalité à l'aide de la formule : $\frac{\lambda_i}{95 \text{ (nombre de valeur propre)}} * 100$

En en déduit que 5 axes sont nécessaires pour perdre au maximum 25% de l'information.

Grâce à XLSTAT on obtient le tableau des composantes principales sur les 5 axes et en plaçant chacune des modalités, on obtient grâce à XLSTAT le graphique suivant :



Interprétation de l'analyse (ACM) :

On peut distinguer 4 groupes d'individus :

- Les individus classés entre NC et 30/5 :

Ces individus sont débutants et ont le classement le plus faible. On remarque que leur nombre d'années de pratique ne dépasse pas les 4 ans et ils ne s'entraînent pas plus de 4h par semaine. Leur type d'entraînement est très varié. Ils ne savent pas en général la marque de leur raquette, la tension de leur cordage ou encore le type de cordage. Ils ont une fréquence de casse du cordage supérieure à 24h. Ils ont donc besoin théoriquement d'une seule raquette. On observe donc bien que chaque individu possède une seule raquette.

- Les individus classés entre 30/4 et 15 /5 :

Ces individus pratiquent le tennis depuis plus de 4 ans mais moins de 8 ans. Ils jouent le plus souvent avec des raquettes de marque Babolat ou Head. Ils ont 2 raquettes et la tension de leur cordage est entre 18 et 20 kg. Ils font moins de 10 matchs en tournoi par an. Ils s'entraînent en général en cours collectifs pendant 2h maximum par semaine. On notera tout de même des différences pour les individus étant 30 ou 15/5, et qui ont des comportements plus proches des individus ci-dessous, notamment en ce qui concerne les types d'entraînements.

- Les individus classés entre 15/4 et 5/6 :

Ce groupe d'individu s'entraîne entre 4 et 8h par semaine (en loisir ou en cours collectifs) et pratique en moyenne le tennis depuis 11 ans. Les raquettes les plus utilisées par ce groupe sont des raquettes de marque Wilson et Yonex, et sont utilisées par des joueurs très offensifs, ce qui n'est pas étonnant puisque ces deux marques possèdent des caractéristiques favorisant ces types de jeu (même si cela dépend des modèles de raquettes et que là nous n'analysons que la marque). Nous remarquons que les tensions varient beaucoup dans ce groupe, mais que plus un joueur est offensif, plus celui-ci aura une tension élevée. Les attaquants de ce groupe auront donc une tension située entre 25 et 28kg. Cela est logique puisque qu'une tension élevée signifie plus de contrôle et moins de puissance, et qu'en général, les joueurs offensifs ont besoin d'avoir plus de contrôle dans leurs frappes de balles pour ne pas faire des fautes tout le temps. Enfin, nous remarquons que ceux-ci possèdent en général 3 raquettes, et qu'ils cassent leur cordage en moins de 24h de pratique. Il y a donc un lien entre la fréquence de cassage de cordage et la nécessité pour un joueur de posséder plusieurs raquettes.

- Les individus classés plus de 4/6 :

Il s'agit ici du meilleur groupe, avec les meilleurs classements, on s'attend donc aux meilleures caractéristiques. Tout d'abord, ce groupe pratique le tennis depuis 8 à 11 ans, ce qui est logique. Ensuite, leur fréquence d'entraînement est de 8 à 10h (sous forme de cours individuels et loisir). Nous pouvons ici nous dire qu'une grande fréquence de cours individuels influe sur le niveau de jeu d'un individu. Les tensions de raquettes ici se trouvent entre 21 et 22kg, avec un cordage de type monofilament qui offre des sensations très appréciées des joueurs de haut niveau. Ce sont ces tensions très basses qui expliquent la fréquence de cassage de cordage se trouvant entre 4 et 6h. Ces fréquences rapides expliquent à leur tour le fait que les individus de ce groupe possèdent plus de 4 raquettes en général. Fait intéressant, nous remarquons que la marque Technifibre est la plus présente dans ce groupe. Il faut savoir que depuis quelques années, Technifibre mise beaucoup sur les jeunes dans le domaine du tennis, et offre souvent des contrats aux joueurs prometteurs (ce fut le cas de Pierrick plus jeune). La marque Technifibre a été aussi en partenariat pendant longtemps avec le club de l'US Cagnes Tennis dans lequel nous avons réalisé ce sondage. Il n'est donc pas étonnant de retrouver cette marque comme raquette principale du meilleur groupe d'individus.

Bilan de notre projet

Pour conclure, ce projet a été enrichissant pour l'ensemble du groupe. Cela nous a permis de voir et appliquer des méthodes d'analyse de données sur des situations réelles et comprendre la puissance des outils mathématiques pour étudier tout type d'information dans tous les domaines inimaginables. La liberté de choix du sujet nous a encore plus motivé pour réaliser ce projet de la plus sérieuse des manières. Ce TAI nous a aussi donné la possibilité de découvrir d'autres outils comme XLSTAT accouplé à Excel pour réaliser l'étude.

Synthèses individuelles

Raphaël Cadillat

Ce projet a permis de mieux s'approprier le cours étudié en classe, ce qui m'a aidé à comprendre davantage certaines parties du cours qui me laissaient perplexe. La liberté du choix du sujet a rendu ce TAI plus plaisant à réaliser car on a pu choisir un sujet assez intrigant et plus parlant pour nous. Transposer les mathématiques à la réalité m'a toujours fasciné car, facteur de compréhension et d'importance à nos yeux. Autre point positif, la découverte de l'outil XLSTAT qui permet de faciliter beaucoup de calculs et tracés de graphes bien qu'il ne faille pas trop en dépendre. Seul point négatif, le temps pour réaliser ce TAI étant assez restreint, il était compliqué d'obtenir un nombre important de réponses aux sondages, rendant les chiffres d'autant plus significatifs.

Pierrick Delrieu

Ce projet fut pour ma part très enrichissant. Je suis passionné de tennis et de sport depuis l'âge de 4 ans. Cette passion et le choix du sujet nous a poussé à choisir des données en lien avec ce domaine. Ne trouvant aucune données intéressantes exploitables sur internet, nous avons eu l'ambition de réaliser notre propre sondage. J'ai ainsi pu transmettre le sondage à l'ensemble de ma communauté tennistique. Surpris par le nombre de réponse, j'ai été enthousiaste de continuer dans cette voie.

L'étude des facteurs impactant le niveau de tennis des joueurs m'a poussé à comprendre dans les moindres détails les méthodes d'analyses. Le choix de la méthode d'analyse fut d'une grande difficulté et a été source d'une longue réflexion. De plus la découverte de différents outils professionnels comme XLSTAT nous a tout de suite permis d'avoir des éléments concrets d'analyse. Ce projet et cette matière nous a permis d'appréhender le monde de l'analyse de données en grande quantité à travers la découverte d'une nouvelle méthode d'analyse : ACM. Cette étude fut extrêmement motivante.

Faustin Dewas

Même si les exercices étudiés en TD étudient des statistiques réelles, le fait de regrouper nos statistiques nous-même lors de ce projet pour ensuite les étudier et en tirer des conclusions fut très intéressant. De ce fait, j'ai pu me confronter de manière tout de même plus concrète aux statistiques, chose souvent compliquée dans le cadre des cours. Les réponses au sondage, même si elles ne sont pas nombreuses par rapport à d'autres exercices étudiés ce semestre, furent tout de même beaucoup plus nombreuses que ce à quoi je m'attendais, ce qui fut une agréable surprise. De plus, le faire sur un sport aussi spécial que le tennis, ou le matériel peut varier énormément d'une personne à une autre, a été un choix extrêmement judicieux, en plus d'être instructif. L'utilisation d'Excel et de XLSTAT afin de trier et étudier les données recueillies par le sondage fut aussi une partie intéressante de ce projet, n'ayant jamais utilisé ces applications d'une manière aussi poussée. Finalement, ce projet fut pour moi une expérience aussi intéressante qu'enrichissante.

Benjamin Lesieux

J'ai toujours été très intéressé par les statistiques, et à quel point il est parfois possible de prédire un événement ou un comportement à l'aide des mathématiques. En plus de cela, je suis un féru de tennis, sport que je pratique depuis 10 ans. En cherchant notre thème pour ce TAI, Pierrick a tout de suite émis l'idée de créer notre propre sondage, d'aller nous-même chercher les variables que nous voulions analyser. Et le fait

d'avoir concrétisé ce choix sur le thème du tennis a fait de ce TAI un projet fortement intéressant et motivant. Ayant une connaissance plus ou moins profonde du sport, nous savions que nous serions capables d'analyser ces données et d'en ressortir des comportements en fonction du niveau des joueurs, de leur fréquence d'entraînement, etc.

Outre le thème, et le choix d'avoir créé notre propre sondage dans l'optique d'établir notre propre tableau de données, nous avons pu revoir notre cours et ainsi nous préparer au DE qui approche, mais aussi d'utiliser de nouveaux outils comme XLSTAT, très puissant, et nous permettant d'établir rapidement nos différentes AFC. Il n'y a donc que du positif qui ressort de ce TAI de mon côté.

Harold Molter

J'ai trouvé ce projet très intéressant. Utiliser les mathématiques, souvent synonyme d'activités abstraites pour les utiliser dans des cas particuliers réels et toujours très enrichissant pour comprendre certains contextes et événements. Bien que je ne sois pas un fan inconditionnel de tennis, j'ai tout de même apprécié le sujet. J'ai adoré utiliser la portée des mathématiques pour étudier des phénomènes sportifs. Le plus dur durant ce projet a été de choisir les bonnes analyses à faire. En effet, nous avons reçu de nombreux résultats à la suite de notre sondage. Il fallait donc réorganiser les données afin d'en tirer des informations intéressantes que l'on pouvait étudier. L'outil XLSTAT est aussi très efficace et simple d'utilisation ce qui nous a permis de gagner du temps et tout de suite aller dans le vif du sujet ce qui m'a énormément plu.

Ce TAI a été en somme assez difficile. Cependant, ce n'est pas un mauvais point car à la suite de celui-ci j'ai pu mieux comprendre l'AFC notamment et j'ai eu un avant-goût d'une nouvelle méthode qui est l'ACM que l'on découvrira sûrement dans le cadre de d'autres modules de mathématiques.

Merci à tous ceux qui ont répondu à notre sondage et au club de l'US Cagnes Tennis et au centre sportif Ultra d'avoir partagé notre sondage. Sans eux ce projet n'aurait pas été possible.

