

Classification des meilleurs joueurs mondiaux de Call Of Duty : Warzone

STA101 – STATISTIQUES DESCRIPTIVES
ALEXANDRE MOURET

Abstract

Le présent document se propose d'étudier les performances des meilleurs joueurs mondiaux de Call Of Duty : Warzone et d'en tenter une classification. Le travail, ici présenté, a consisté à récupérer 104 profils parmi les 250 joueurs les mieux classés en mode *Battle Royal* sur le site cod.tracker.gg puis à procéder à une analyse statistique en vue de trouver des profils de performances au sein du groupe de joueurs.

Pour cela, on a appliqué une Analyse en Composantes Principales sur des variables quantitatives centrées-réduites relatives aux performances des joueurs en mode *Battle Royal*. On procède ensuite à une classification tandem. On a d'abord cherché à déterminer un nombre optimal de classes en appliquant une Classification Ascendante Hiérarchique sur des centres de classe déterminés par la méthode des k-means à 30 classes. Une fois le nombre de classes déterminés, on obtient les 4 groupes finaux par une méthode des k-means à 4 classes.

L'étude a permis de déterminer que les caractéristiques qui séparent le plus les meilleurs joueurs du top mondial entre eux sont : le temps passé sur le jeu, la capacité de survie et les compétences dans les combats. Ainsi, on distingue 4 sous-groupes parmi notre échantillon :

- Un groupe d'individus qui jouent moins que la moyenne mais gagnent beaucoup grâce à leur capacité à survivre.
- Un groupe d'individus probablement trop agressifs qui gagnent souvent les affrontements mais peinent à survivre une partie entière.
- Un groupe d'individus qui compensent leur faible niveau de jeu par un temps de jeu plus important.
- Un groupe, plus nombreux que les autres, d'individus moyens qui ont des performances et des temps de jeu proches de la moyenne de l'échantillon.

Table of Contents

Introduction.....	3
Contexte	3
Objectifs de l'étude	3
Les données	3
Présentation des données	3
Pré-traitement des données	6
Analyse univariée.....	7
Analyse bivariée	9
Analyse exploratoire multivariée	12
Choix de la méthode et objectifs	12
Choix du nombre d'axes pertinents pour l'analyse.....	12
Analyse des variables	13
Analyse des individus.....	17
Conclusion.....	17
Classification (Tandem).....	18
Choix de la méthode.....	18
Réalisation de la classification mixte.....	18
Interprétation des classes	19
Conclusion.....	23
Conclusion.....	23
Annexes	24
Annexe 1 : Algorithme Python pour la récupération des données des joueurs.....	24
Annexes 2 à 21 : Histogrammes des variables actives	26
Annexe 22 : Résumé de la variance expliquée en fonction des valeurs propres	29
Annexe 23 : Table des corrélations entre la première dimension de l'ACP et les variables du jeu de données.....	30
Annexe 24 : Table des corrélations entre la deuxième dimension de l'ACP et les variables du jeu de données.....	31
Annexe 25 : Table des corrélations entre la troisième dimension de l'ACP et les variables du jeu de données.....	32
Annexe 26 : Caractérisation de la première classe par les variables quantitatives	33
Annexe 27 : Caractérisation de la deuxième classe par les variables quantitatives	34
Annexe 28 : Caractérisation de la troisième classe par les variables quantitatives.....	35
Annexe 29 : Caractérisation de la troisième classe par les variables quantitatives.....	35

Introduction

Contexte

Call of Duty : Warzone est un jeu vidéo sorti le 10 mars 2020 sur PC, PS4 et Xbox One et développé conjointement par les studios Infinity Ward et Raven Software pour le compte de l'éditeur Activision-Blizzard-King. Initialement intégré au jeu *Call Of Duty : Modern Warfare* (2019), il a ensuite été intégré à *Call Of Duty : Black Ops Cold War* (2020) et *Call Of Duty : Vanguard* (2021). Le jeu compte, depuis juin 2022, plus de 125 millions de joueurs¹.

Il s'agit d'un jeu de type *battle royal* où 150 joueurs s'affrontent en ligne dans une ville fictive jusqu'à ce qu'un seul d'entre eux survive. Le territoire disponible se réduit avec le temps et les joueurs sont donc forcés de converger vers un point commun. Pour s'affronter, les joueurs disposent d'armes à feu et de véhicules. Le jeu a été enrichi par d'autres modes de jeu mais on s'intéressera principalement au mode *battle royal* dans ce document.

Le site cod.tracker.gg² permet de suivre les performances des joueurs via diverses statistiques. Il propose notamment des [classements des meilleurs joueurs](#)³.

Objectifs de l'étude

Dans ce document, on s'intéressera aux profils des joueurs les mieux classés du mode *battle royal* de *Call Of Duty : Warzone*.

On essaiera en particulier de répondre aux questions suivantes :

- Existe-t-il des ressemblances entre des joueurs du mode *battle royal* ? Si oui, peut-on les caractériser ?
- Existe-t-il des liaisons linéaires entre les variables de notre jeu de données ?

Les données

Présentation des données

Dans le but de répondre à ces questions, on utilisera un jeu de données original constitué de 104 des 250 profils les mieux classés dans le mode principal du jeu : le mode *battle royal* (classement mondial par nombre de victoires en *battle royal* au 01 avril 2023).

Le jeu de données a été constitué par mes soins à l'aide d'un algorithme reposant sur une API Python non-officielle nommée [cod-api](#)⁴. L'algorithme Python utilisé est fourni en annexe 1.

Cette API ne permet d'accéder qu'aux profils utilisateurs placés en "public" par leurs utilisateurs. En conséquent, il n'était pas possible de récupérer les 250 premiers profils.

Les données récupérées proviennent de deux tables de données différentes. La première contient des données spécifiques au mode *battle royal* de *Call Of Duty : Warzone*, la seconde contient des données sur les performances générales du joueur (tous modes confondus).

Les 104 individus (lignes) du tableau de données sont des joueurs. Chaque joueur est identifié par son identifiant en ligne (valeur unique). Il est ensuite décrit par un ensemble de 32 variables (colonnes) dont la description figure dans le tableau 15.

¹ <https://twitter.com/charlieINTEL/status/1534595837325676554>

² <https://cod.tracker.gg/>

³ <https://cod.tracker.gg/warzone/leaderboards/battle-royale/atvi/Wins?page=1&mode=1>

⁴ <https://pypi.org/project/cod-api/>

Tableau 15 : Description des variables du jeu de données

Nom de la variable	Type	Description	Catégorie
G_totalXp	double	Nombre total de points d'expérience gagnés par le joueur (le joueur gagne des points d'expérience à chaque partie). Tous modes confondus.	Générale
G_accuracy	double	Ratio G_hits/G_shots. Mesure de la capacité du joueur à toucher son adversaire. Tous modes confondus.	Générale
G_gamesPlayed	double	Nombre de parties jouées tous modes confondus.	Générale
G_headshots	double	Nombre de tirs à la tête (les tirs à la tête sont plus difficiles à réaliser et font plus de dégâts). Tous modes confondus.	Générale
G_hits	double	Nombre de fois où le joueur a touché un adversaire avec un tir. Tous modes confondus.	Générale
G_misses	double	Nombre de fois où le joueur a tiré sans toucher un adversaire. Tous modes confondus.	Générale
G_kdRatio	double	Ratio G_kills/G_deaths. Tous modes confondus. Un joueur avec un ratio kills/deaths supérieur à 1 est tue plus d'adversaires qu'il ne se fait tuer.	Générale
G_suicides	double	Nombre total de fois où le joueur s'est suicidé. Tous modes confondus.	Générale
G_shots	double	Nombre total de tirs du joueur. Tous modes confondus.	Générale
G_scorePerMinute	double	Cette variable modélise le nombre de points marqués par minute en moyenne. Tous modes confondus.	Générale
G_hkRatio	double	Ratio G_headshots/G_kills. Tous modes confondus. Cet indicateur a été construit lors de la récupération de la base de données.	Générale
G_wgRatio	double	Ratio G_wins/G_gamesPlayed. Cet indicateur a été construit lors de la récupération de la base de données.	Générale
contracts	double	Nombre de contrats accomplis dans le mode <i>Battle Royal</i> . Les contrats sont des missions à accomplir en cours de partie pour gagner des avantages.	Battle Royal
kills	double	Nombre total d'adversaires abattus dans le mode Battle Royal.	Battle Royal
downs	double	Nombre de fois où le joueur a mis un adversaire à terre sans le tuer dans le mode Battle Royal.	Battle Royal
deaths	double	Nombre total de fois où le joueur est mort dans le mode Battle Royal.	Battle Royal
kdRatio	double	Ratio kills/deaths. Mode <i>Battle Royal</i> .	Battle Royal
revives	double	Quand le joueur meurt une première fois dans la partie en mode <i>Battle Royal</i> , il peut affronter un autre joueur éliminé en combat singulier. S'il gagne, il peut revenir dans la partie. Cette variable indique le nombre total de fois où le joueur est revenu dans la partie.	Battle Royal
score	double	Nombre total de points marqués en mode <i>Battle Royal</i> .	Battle Royal
scorePerMinute	double	Cette variable modélise le nombre de points marqués par minute en moyenne en mode <i>Battle Royal</i> .	Battle Royal

timePlayed	double	Temps passé en mode <i>Battle Royal</i> , en secondes.	Battle Royal
gamesPlayed	double	Nombre de parties jouées en mode <i>Battle Royal</i> .	Battle Royal
topFive	double	Nombre de fois où le joueur a terminé dans les 5 premiers en mode <i>Battle Royal</i> .	Battle Royal
topTen	double	Nombre de fois où le joueur a terminé dans les 10 premiers en mode <i>Battle Royal</i> .	Battle Royal
topTwentyFive	double	Nombre de fois où le joueur a terminé dans les 25 premiers en mode <i>Battle Royal</i> .	Battle Royal
wins	double	Nombre de victoires en mode <i>Battle Royal</i> .	Battle Royal
averageLifetime	double	averageLifeTime = timePlayed/deaths Mesure le temps moyen entre deux morts du joueur, en seconde, dans le mode <i>Battle Royal</i> . Cet indicateur a été construit au moment de la récupération de la base de données.	Battle Royal
wgRatio	double	Ratio G_wins/G_gamesPlayed. Cet indicateur a été construit lors de la récupération de la base de données.	Battle Royal
killsPerGame	double	Ratio kills/gamesPlayed Mesure le nombre moyen de d'adversaires tués par partie en mode <i>Battle Royal</i> . Cet indicateur a été construit lors de la récupération de la base de données.	Battle Royal
downsPerGame	double	Ratio downs/gamesPlayed Mesure le nombre moyen d'adversaires mis à terre par partie en mode <i>Battle Royal</i> . Cet indicateur a été construit lors de la récupération de la base de données.	Battle Royal
revivesPerGame	double	Ratio revives/gamesPlayed Mesure le nombre moyen de fois où le joueur est revenu à la vie par partie en mode <i>Battle Royal</i> . Cet indicateur a été construit lors de l'importation des données et ne figure pas dans le fichier csv.	Battle Royal
contractsPerGame	double	Ratio contracts/gamesPlayed Mesure le nombre moyen de contrats réalisés par le joueur par partie en mode <i>Battle Royal</i> . Cet indicateur a été construit lors de l'importation des données et ne figure pas dans le fichier csv.	Battle Royal

Notons que toutes les variables sont quantitatives mais qu'elles proviennent de 2 sources (tables) différentes. On pourra donc effectuer une Analyse en Composantes Principales (ACP) qui est bien adaptée à la description de jeu de données quantitatives.

Pré-traitement des données

Données manquantes

Des données sont manquantes dans les colonnes G_accuracy, G_misses, G_hits, G_shots, G_scorePerMinute et G_wgRatio.

Ces données concernent seulement 4 individus. Compte-tenu du faible nombre d'individus concernés (environ 3.8% des individus du jeu de données), on pourrait retirer ces individus de notre jeu de données mais, ici, on préférera imputer les données manquantes par ACP à l'aide du package R *missMDA*.

Tableau 1 : Liste des données manquantes		
Nom d'utilisateur	row	col
yeezee.8865145	60	2
OHeyItsMike.4652848	83	2
Gyrodabest.8413713	93	2
yeezee.8865145.1	60	5
OHeyItsMike.4652848.1	83	5
Gyrodabest.8413713.1	93	5
yeezee.8865145.2	60	6
OHeyItsMike.4652848.2	83	6
Gyrodabest.8413713.2	93	6
yeezee.8865145.3	60	9
OHeyItsMike.4652848.3	83	9
Gyrodabest.8413713.3	93	9
yeezee.8865145.4	60	10
OHeyItsMike.4652848.4	83	10
Gyrodabest.8413713.4	93	10
TAZ.9459019	49	12
yeezee.8865145.5	60	12
OHeyItsMike.4652848.5	83	12
Gyrodabest.8413713.5	93	12

Choix des individus

Le nombre d'individus étant faible, on cherchera à garder l'ensemble des individus.

Construction de variables supplémentaires

A partir des données récupérées dans la base de données d'Activision, on a construit un certain nombre de variables additionnelles qui sont soit des ratios classiques dans les jeux de tirs à la première personne, soit ont pour but de mieux appréhender la performances moyennes des joueurs au cours d'une partie. Voici la liste de ces variables (décrites dans le tableau 15) :

- killsPerGame
- downsPerGame
- revivePerGame
- contractsPerGame
- averageLifetime
- wgRatio
- G_wgRatio
- G_hkRatio

Choix des variables actives/illustratives.

De façon générale, puisque notre objectif est de caractériser les performances des joueurs en mode *battle royal*, on placera toutes les variables liées aux performances générales en variables supplémentaires (toutes les variables préfixées d'un "G_"). Ces variables permettront d'enrichir notre interprétation des axes de l'ACP et des groupes dans notre classification.

Toutes les autres variables seront utilisées comme variables actives dans l'analyse.

Analyse univariée

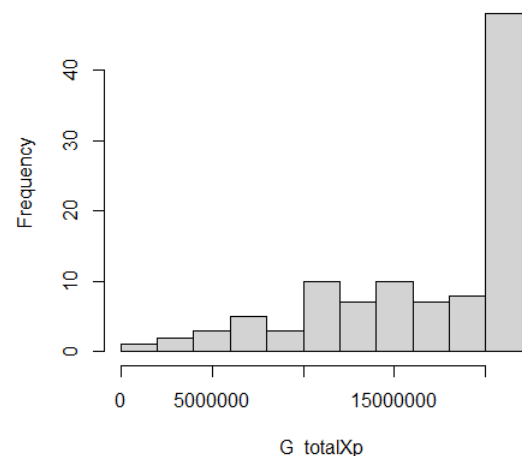
La variable G_totalXp

La variable G_totalXp (figure 1) possède une valeur bloquée à 21 978 000 points d'expérience. En conséquence, la distribution de cette variable est fortement biaisée et n'est pas représentative des profils des joueurs et de leurs performances. On la retirera donc de l'étude dans la suite du document.

Indicateurs statistiques

Les indicateurs statistiques univariés (tableau 16) montrent que les variables ont toutes des unités différentes et que les écart-types ont des ordres de grandeurs très différents. A partir de ce point du document, on choisira donc de travailler sur le tableau des données centrées-réduites afin que les variables ayant la plus forte variance n'aient pas un poids disproportionné dans l'analyse factorielle.

Figure 1: Histogramme de la variable G_totalXp




```

##
## Tableau 16 : Indicateurs statistiques
## =====
## Statistic      Pctl(25)      Median      Mean      Pctl(75)      St. Dev.
## -----
## G_accuracy      0.2          0.2          0.2          0.2          0.1
## G_gamesPlayed    9,957.8      13,015.5     13,125.8     16,561.8     4,658.5
## G_headshots     13,984.5     19,758.5     23,208.7     31,449.2     12,291.9
## G_hits          20,207       69,469       86,077.6     112,852      87,424.5
## G_misses        70,424       253,455      307,061.6    397,366      282,381.9
## G_kdRatio        1.7          2.3          2.5          3.0          1.0
## G_suicides       9,360        16,673       19,045.2     24,939.8     14,239.7
## G_shots          86,365       313,898      393,139.2    502,964      365,993.4
## G_scorePerMinute 4,126.3      8,794.1      150,358.2    27,793.1     1,175,318.0
## G_hkRatio        0.2          0.2          0.2          0.2          0.04
## G_wgRatio        0.02         0.05         0.1          0.1          0.1
## contracts        8,059.2      10,834       12,185.2     14,181.8     6,173.8
## kills           43,717.5     61,574.5     69,089.1     76,596       38,714.2
## downs           41,941       58,968.5     66,015.5     77,942.8     36,710.4
## deaths          16,238.2     21,264       22,353.6     26,689       9,440.1
## kdRatio          2.3          3.1          3.2          3.9          1.3
## revives          2,756.2      3,517        3,706.1      4,314.8      1,546.8
## score            43,298,335   55,747,526   63,899,872.0 81,785,798.0 28,934,401.0
## scorePerMinute   334.5        382.5        407.3        475.7        105.5
## timePlayed       7,263,736.0  8,974,964    9,484,262.0  10,985,657.0 3,718,589.0
## gamesPlayed      6,483.5      8,344.5      9,107.8      10,981       3,891.2
## topFive          1,823.2      2,233.5      2,549.4      3,044        1,010.2
## topTen           2,768.8      3,356        3,735.5      4,283.8      1,481.5
## topTwentyFive    4,811.5      5,879.5      6,137.8      6,962        2,435.7
## wins             753.8        946          1,177.1      1,480.8      597.0
## averageLifetime  987.7        1,066.6      1,074.5      1,154.6      154.8
## wgRatio          0.1          0.1          0.1          0.2          0.1
## killsPerGame     5.7          7.4          7.8          9.8          2.9
## downsPerGame     5.5          7.2          7.6          9.6          3.1
## rivivesPerGames  0.3          0.4          0.4          0.5          0.2
## contractsPerGames 1.1          1.3          1.4          1.6          0.6
## -----

```

Les diagrammes en boîtes (figures 2, 3, 4 et 5) quant à eux, nous permettent de constater qu'il y a un grand nombre de valeurs extrêmes pour presque toutes les variables de la série. Au totale, 42 individus sont valeur extrême pour au moins une variable. On peut notamment constater que certains individus admettent des valeurs très au-dessus des autres (Par exemple, TAZ#9459019 a un score par minute 10 écarts-types au-dessus de la moyenne et Jesus2b#1012410 a un nombre de contrats 5 écarts-types au-dessus de la moyenne). Cependant, compte-tenu du faible nombre d'individus et du fait que de telles performances extrêmes ne sont pas étonnantes pour des individus parmi les meilleurs mondiaux, on choisira de ne pas les retirer du jeu de données.

Figure 2: Diagrammes en boîte de 11 variables actives centrées-réduites

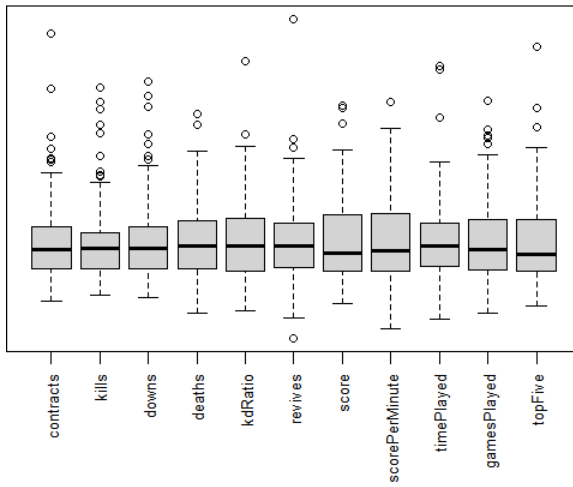


Figure 3: Diagrammes en boîte de 9 variables actives centrées-réduites

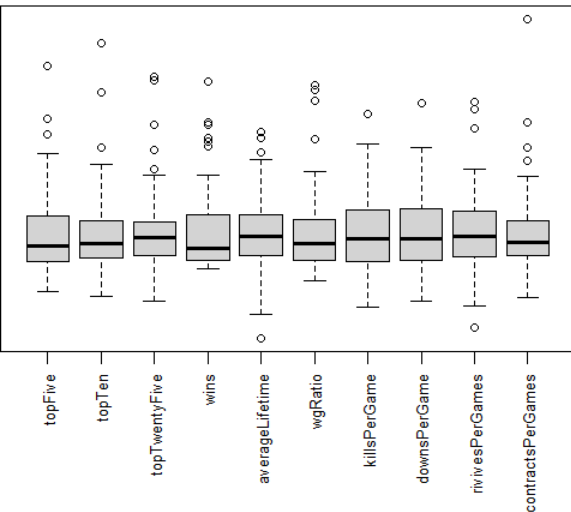


Figure 4: Diagrammes en boîte de 10 variables inactives centrées-réduites

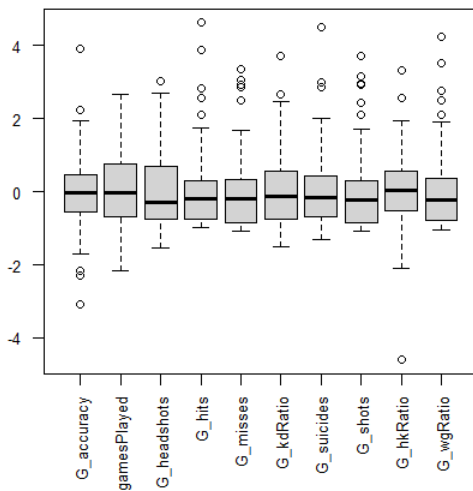
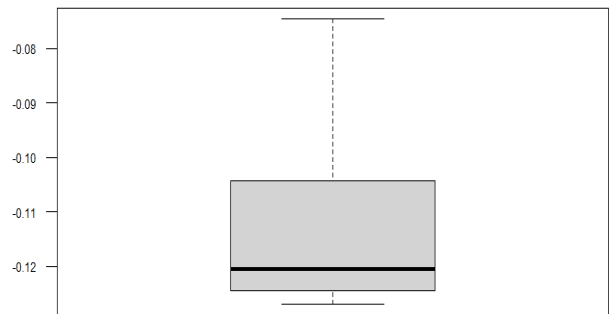


Figure 5: Diagrammes en boîte de la variable inactive centrée-réduite G_scorePerMinute sans les valeurs extrêmes



Analyse bivariée

Beaucoup de variables (actives ou illustratives) du jeu de données n'ont pas une distribution normale mais sont biaisées à droite (à asymétrie positive). On peut s'en rendre compte en visualisant les histogrammes des différentes variables actives (voir figure 6 et annexes 2 à 21).

C'est une indication forte du fait que la plupart des variables du jeu de données ne suivent pas une loi normale. On peut confirmer notre hypothèse à l'aide d'un test de normalité de Shapiro-Wilk (tableau 2). Dans ce test, on conclut à la normalité si la p-value est non-significative (ici, on choisit un seuil de 5%). On fournit ci-dessous la table des résultats des tests pour lesquels la p-value est non-significative :

**Figure 6: Un exemple de distribution biaisée à droite
Histogramme de la variable deaths
skewness = 0.8**

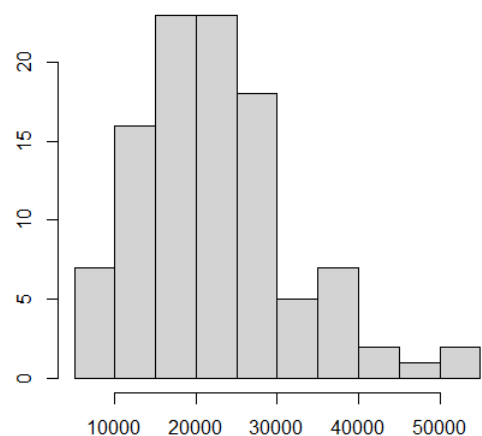


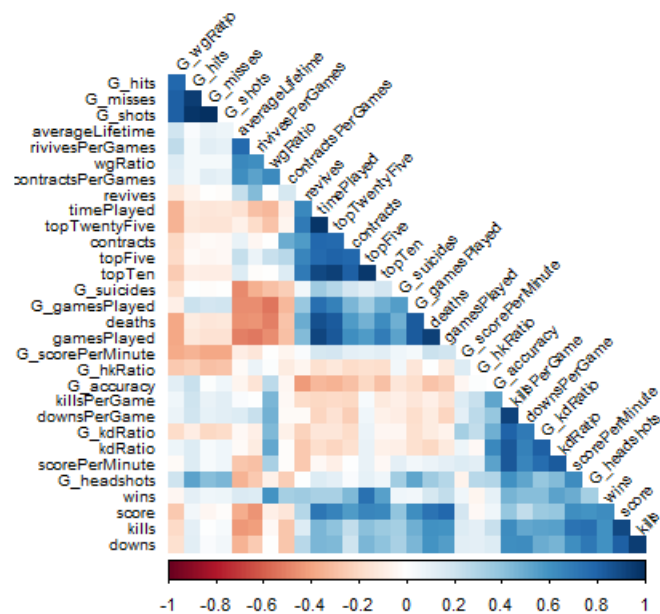
Tableau 2 : Tests de normalité de Shapiro-Wilk

Variable	statistique	p.value
G_gamesPlayed	0.99	0.6045
averageLifetime	0.977	0.0717
killsPerGame	0.977	0.0651

Seules 3 variables peuvent être considérées comme suivant une loi normale : G_gamesPlayed, averageLifeTime et killsPerGames.

Les coefficients de corrélation de Pearson peuvent malgré tout être calculés (figure 7) mais il ne sera pas possible de procéder à un test de corrélation de Pearson qui requiert une hypothèse de normalité des données (ici non satisfaite pour la quasi-totalité des variables). Un coefficient proche de 1 signifie qu'il y a une liaison linéaire positive forte entre les deux variables, tandis qu'un coefficient proche de -1 signifie une corrélation linéaire négative forte (et 0 une absence de corrélation linéaire).

Figure 7: Matrice des corrélations de Pearson



On peut constater qu'il n'y a aucune liaison linéaire négative forte entre les variables. En revanche, il existe de nombreuses liaisons linéaires positives fortes. En particulier:

- Entre les variables G_hits, G_misses et G_shots. (plus on tire, plus on rate et on touche)
- Entre les variables timePlayed, gamePlayed, contracts, topFive, topTen, topTwentyFive, deaths, G_gameplayed, score, revives. Un grand nombre d'indicateurs bruts sont donc linéairement liés au temps de jeu.
- Entre les variables averageLifetime, contractsPerGame, rivivesPerGame et wgRatio (la capacité à rester en vie devrait donc avoir un lien fort avec la capacité à gagner).

- Entre les variables killsPerGame, downsPerGame, kdRatio, G_kdRatio. On considérera dans la suite que ces variables représentent le skill⁵ du joueur.

Ces liaisons linéaires sont visibles sur les représentations des nuages de points (figure 8, 9 et 10).

Figure 8: Exemple de liaison linéaire entre les variables averageLifeTime et rivesPerGame

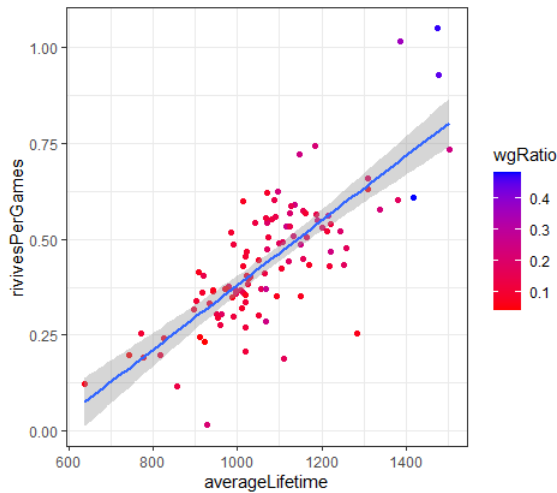


Figure 9: Exemples de liaisons linéaires entre les variables killsPerGame et downsPerGame

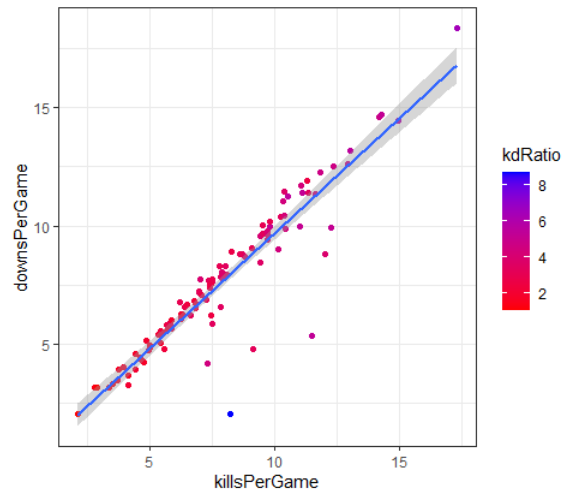
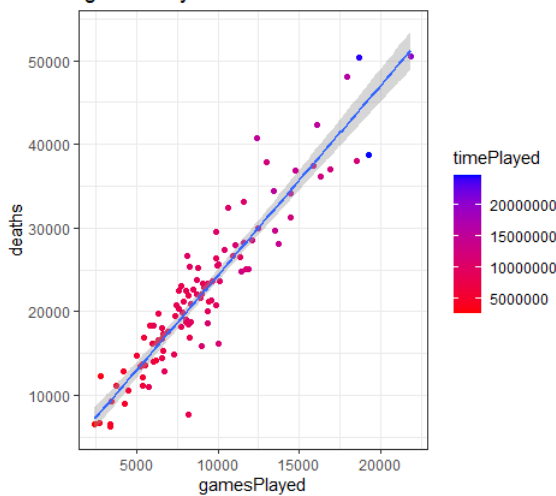


Figure 10: Exemple de liaison linéaire entre les variables deaths et gamesPlayed



Puisque les données ne sont pas normalement distribuées pour la plupart des variables, on peut aussi représenter les coefficients de corrélation de Spearman (figure 11). Ce coefficient permet de détecter des associations monotones (mais non nécessairement linéaires) entre deux variables. Un coefficient proche de 1 indique une association positive, un coefficient proche de -1 indique une association négative.

On constate que la variable G_scorePerMinute est négativement corrélée avec les variables G_wgRatio, G_hits, G_misses et G_shots. Cette corrélation n'est pas linéaire comme l'atteste le graphe du nuage de point des variables G_scorePerMinute et G_shots.

⁵ Dans un jeu de tir à la première personne, le skill représente la capacité du joueur à gagner ses combats contre un autre adversaire.

Figure 12: Exemple de liaison non linéaire entre les variables G_shots et G_scorePerMinute

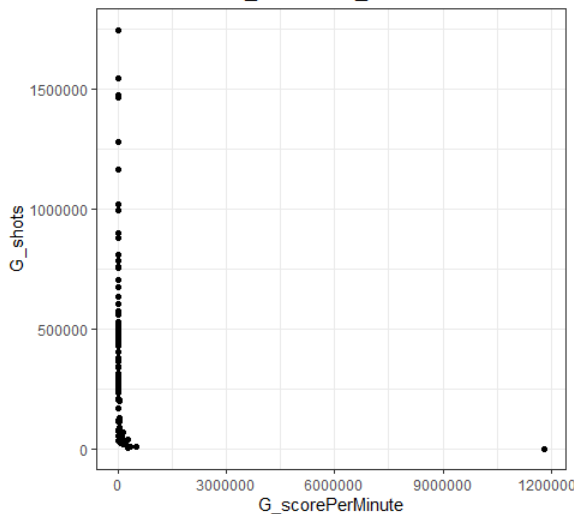
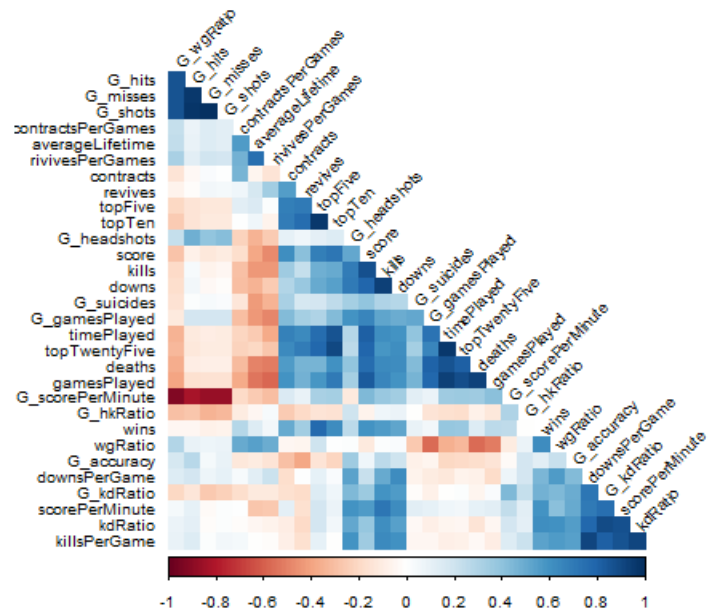


Figure 11: Matrice des corrélations de Spearman



Analyse exploratoire multivariée

Choix de la méthode et objectifs

Le jeu de données est constitué des données de deux sources (statistiques générales du joueur et statistiques du joueur en mode *battle royal*). Dans ce document, on souhaite étudier les performances des joueurs en mode *battle royal*. On ne conservera donc comme variables actives que les variables liées à ce mode.

Toutes les variables étant quantitative, on choisira alors de réaliser une Analyse en Composantes Principales (ACP) dont l'objectif est de condenser l'information contenu dans un jeu de données. L'ACP propose à la fois une visualisation des liaisons linéaires entre les variables et de la distance entre les individus.

Compte-tenu du fait que les variables n'ont pas les mêmes unités, l'ACP sera réalisée sur le tableau de données centré-réduit.

Par ailleurs, pour des raisons déjà évoquées, les variables préfixées d'un "G_" ne seront incluses à l'analyse que comme variables quantitatives supplémentaires.

Enfin, puisque certaines données sont manquantes et peu nombreuses, elles seront imputées par ACP via le package R *missMDA*.

Choix du nombre d'axes pertinents pour l'analyse

L'ACP est faite à partir des 20 variables actives. Le diagramme en barres, représentant le pourcentage d'inertie par composante de l'ACP, montre que les deux premières composantes représentent environ 66% de l'inertie (voir tableau complet en annexe 22).

Figure 13: Diagramme en barre de l'éboullis des valeurs propres

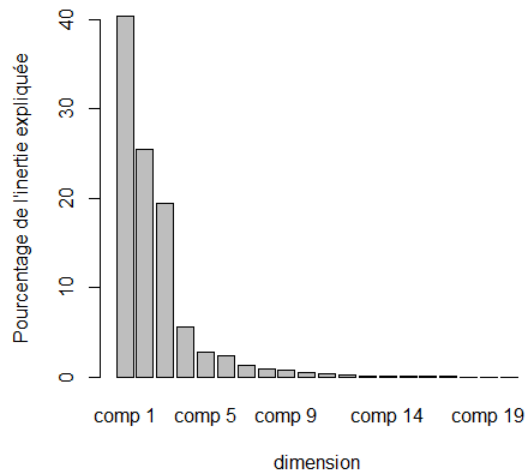
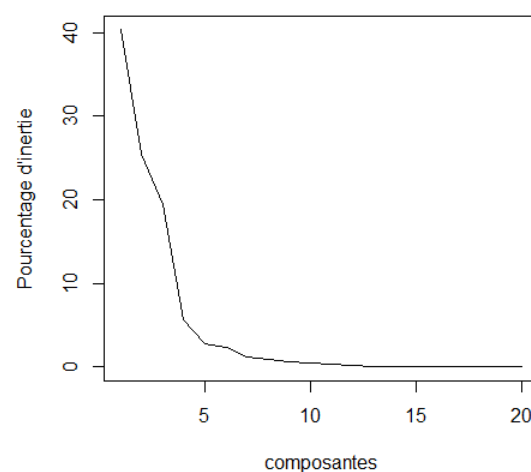


Figure 14: Graphe de l'éboullis des valeurs propres



Deux critères sont couramment utilisés pour le choix du nombre de dimensions à retenir dans l'analyse :

- La règle de Kaiser : On ne considère que les axes correspondant à une valeur propre supérieure à 1. Cela conduirait à retenir les 4 premiers axes.
- La règle du coude : On recherche une cassure dans le graphe des éboulis. Une cassure nette peut être observée entre les composantes 3 et 4.

On choisit ici de retenir 3 composantes comme suggéré par la règle du coude. A elles 3, ces 3 dimensions cumulent 85% de l'inertie (tableau 3).

Tableau 3: Résumé de la variance expliquée en fonction des 5 plus grandes valeurs propres

composante	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	8.069	40.343	40.343
comp 2	5.091	25.457	65.800
comp 3	3.894	19.471	85.271
comp 4	1.117	5.584	90.854
comp 5	0.544	2.722	93.577

Analyse des variables

Corrélations entre les variables et les axes factoriels

On procède à présent à l'analyse des corrélations entre les axes et les variables de notre jeu de données.

Le tableau suivant montre les corrélations entre la première dimension et les 10 variables de notre jeu de données les plus corrélées au premier axe (voir annexe 23 pour le tableau complet).

Tableau 4 : Table des corrélations entre la première dimension de l'ACP et les 10 variables les plus corrélées à l'axe 1 du jeu de données

variable	correlation	p.value
timePlayed	0.93	0
topTwentyFive	0.91	0
gamesPlayed	0.91	0
score	0.90	0
deaths	0.87	0
topTen	0.86	0
topFive	0.81	0
contracts	0.74	0
G_gamesPlayed	0.73	0
kills	0.73	0

On peut constater que :

- Le premier axe est fortement corrélé positivement au temps de jeu en mode *battle royal* (corrélation > 0.7) ainsi qu'à de nombreuses variables qui sont elles-mêmes linéairement corrélées au temps de jeu (gamesPlayed, score, contracts). Cette interprétation est confortée par le fait que la variable supplémentaire G_gamesPlayed est aussi fortement corrélée positivement au 1er axe.
- Il est aussi fortement corrélé positivement au nombre total de morts, d'adversaires tués ou mis à terre et au nombre de top 25, top 10 et top5 du joueur (corrélations > 0.7)
- Il est plus faiblement corrélé positivement (corrélations > 0.6) au nombre de fois où le joueur est revenu dans la partie.

La première dimension représente donc avant tout l'engagement du joueur dans le mode *battle royal* et plus largement son engagement sur le jeu.

On peut faire de même pour les 2e et 3e dimensions (tableaux des corrélations disponibles en annexes 24 et 25) et on peut ainsi conclure que :

- Le second axe est fortement corrélé positivement (corrélations > 0.7) aux performances du joueur au cours d'une partie (killsPerGame, scorePerMinute, kdRatio, downsPerGame). Cette dimension peut s'interpréter comme mesurant le niveau de jeu du joueur. Cette interprétation est confortée par la forte corrélation positive de l'axe avec G_kdRatio et la faible corrélation positive à G_accuracy (la précision générale) qui sont des variables supplémentaires.
- Le 3e axe est fortement corrélé positivement (corrélations > 0.7) au temps moyen en vie (averageLifetime), au nombre de résurrections par partie (rivesPerGames), au nombre de contrats effectués en moyenne par partie (contractsPerGame) et au ratio de victoires (wgRatio) en *battle royal*. Cet axe peut s'interpréter comme représentant les performances de survie d'un joueur en mode *battle royal*.

Contributions des variables à la construction des axes

On peut à présent analyser la contribution des variables à la construction des 3 premiers axes factoriels.

Si chaque variable contribuait autant à la construction d'un axe alors sa contribution vaudrait $\frac{1}{19} \approx 0.05$ pour chaque axe. On considérera comme forte une contribution supérieure à $1.5 \times 0.05 = 0.075$

Ci-dessous, on donne les tableaux, triés par ordre décroissant de contribution, des contributions des différentes variables à la construction de chacune des 3 dimensions d'analyses :

Tableau 5 : Contributions des variables les plus contributrices à la construction de la dimension 1

variable	Dim.1
timePlayed	0.11
topTwentyFive	0.10
gamesPlayed	0.10
score	0.10
deaths	0.09
topTen	0.09
topFive	0.08
contracts	0.07

Tableau 6 : Contributions des variables les plus contributrices à la construction de la dimension 2

variable	Dim.2
killsPerGame	0.18
kdRatio	0.16
scorePerMinute	0.15
downsPerGame	0.15
wgRatio	0.07

Tableau 7 : Contributions des variables les plus contributrices à la construction de la dimension 3

variable	Dim.3
averageLifetime	0.19
revivesPerGames	0.17
contractsPerGames	0.16
wgRatio	0.13
topFive	0.07

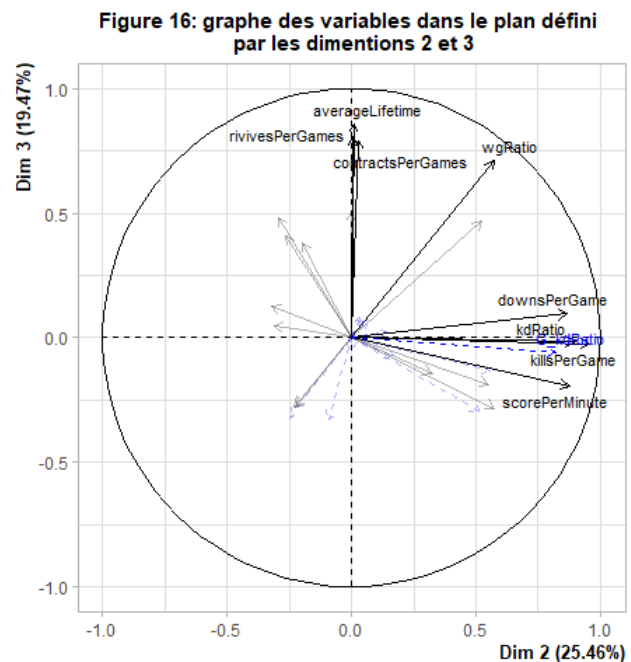
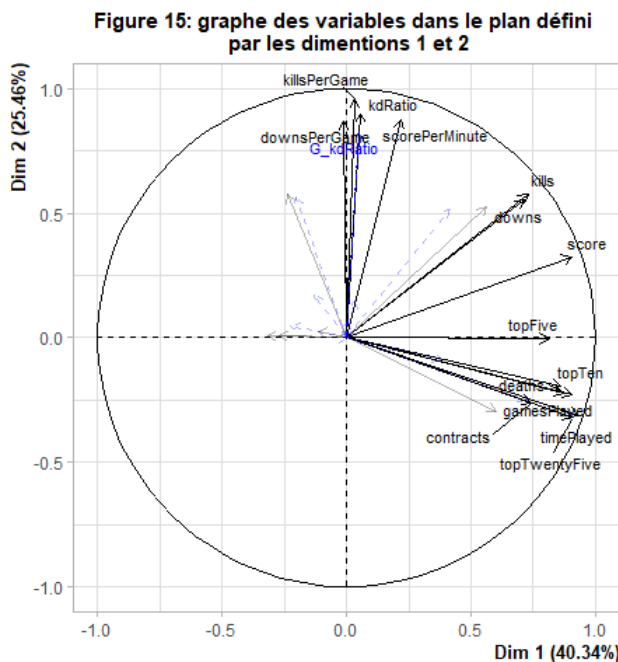
Les tableaux précédents permettent donc de voir que :

- Pour le premier axe, le plus important contributeur est le temps de jeu (timePlayed) ainsi que divers indicateurs qui augmentent naturellement avec le temps de jeu (topTwentyFive, gamesPlayed, score, deaths, topTen, topFive).
- Pour le second axe, on retrouve, comme plus forts contributeurs, les indicateurs de niveau de jeu (killsPerGame, kdRatio, scorePerMinute, downsPerGame).
- Enfin, pour la construction du 3e axe, les variables qui ont contribué le plus sont celles liées aux performances en mode battle royal (wgRatio, contractsPerGames) et à la survie (averageLifetime, revivesPerGames).

Cela conforte notre interprétation initiale des 3 premiers axes.

Analyse des variables projetées

On fournit ci-dessous les projections des variables dans le premier (dimensions 1 et 2) et le second plan (dimensions 2 et 3) de l'ACP. On ne représente que les variables bien représentées ($\cos^2 > 0.6$).



On peut noter que :

- La variable supplémentaire G_totalXp n'est bien représentée sur aucun des 3 premiers axes tandis que la variable supplémentaire G_kdRatio est liée au 2e axe.
- De nombreuses liaisons linéaires fortes sont visibles sur les différents plans d'analyse :
 - Plan défini par les dimensions 1 et 2 :
 - gamesPlayed, deaths, timePlayed, contracts, topTwentyFive, topTen (Jouer beaucoup permet d'avoir un grand nombre de top25 et 10. Surtout les top25).
 - kills et downs. (Un joueur qui tue beaucoup met aussi beaucoup d'adversaires à terre)
 - killsPerGame, downsPerGame, kdRatio. (Ces variables définissent des joueurs efficaces dans les duels)
 - Plan défini par les dimensions 2 et 3 :

- `rivivesPerGames`, `averageLifetime` (Rester en vie et ressusciter en mode *Battle Royal* sont corrélés, ce qui est trivialement vrai.) ainsi que `contractsPerGames`.

Il est aussi à noter que les variables `kdRatio`, `killsPerGame` et les variables `rivivesPerGame`, `averageLifetime`, `contractsPerGames` sont linéairement indépendantes.

Analyse des individus

On peut commencer par représenter les individus dans les différents plans de l'ACP. On a mis en évidence les 10 premiers contributeurs à chaque plan :

Figure 17: Graphe des individus dans le plan défini par les dimensions 1 et 2

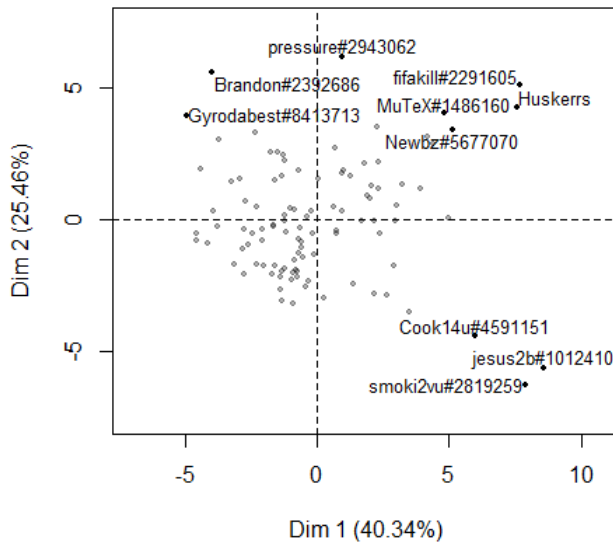
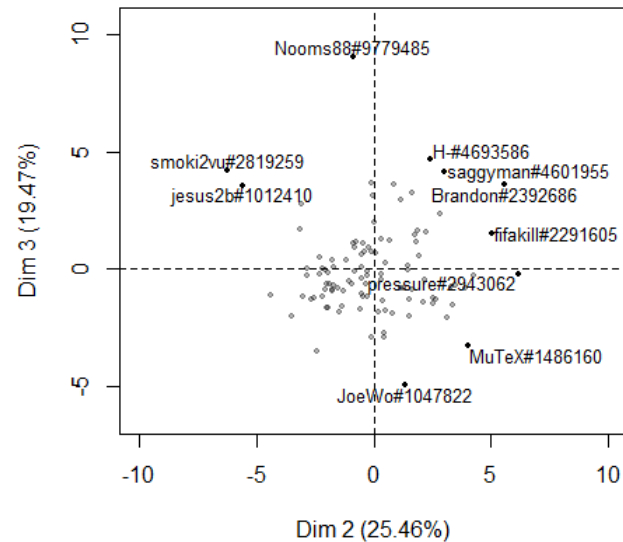


Figure 18: Graphe des individus dans le plan défini par les dimensions 2 et 3



Sur l'axe 1, les individus qui ont le plus contribué à la construction de l'axe sont `Jesus2b#1012410`, `smoki2vu#2819259`, `fifakill#2291605` et `Huskerrs`. Ce sont des joueurs qui jouent beaucoup en mode *battle royal* ce qui confirme notre hypothèse que l'axe 1 représente l'investissement des joueurs en temps de jeu. Un joueur ayant une coordonnée positive et éloignée de 0 sur l'axe 1 est un joueur qui joue beaucoup.

Sur l'axe 2, les plus gros contributeurs sont `smoki2vu#2819259`, `pressure#2943062`, `Jesus2b#1012410` et `Brandon#2392686`. `smoki2vu#2819259` et `Jesus2b#1012410` sont des joueurs ayant des `kdRatio`, `scorePerMinute`, `killsPerGame` et `downsPerGame` très inférieurs à la moyenne de l'échantillon, tandis que `pressure#2943062` et `Brandon#2392686` sont très au-dessus de la moyenne. On peut donc en conclure qu'une coordonnée positive sur l'axe 2 indique un skill élevé du joueur.

Sur l'axe 3, les plus gros contributeurs sont `Nooms88#9779485`, `H-#4693586`, `saggyman#4601955`, `JoeWo#1047822`, et `smoki2vu#2819259`. Ces joueurs se démarquent par leurs valeurs pour les variables `averageLifetime`, `rivivesPerGame` et, dans une moindre mesure, `contractsPerGame`. Une coordonnée élevée sur cet axe signifie que le joueur reste en vie plus longtemps que la moyenne.

Conclusion

A ce stade de l'étude, nous pouvons conclure que 3 éléments semblent fortement résumer le jeu de données et caractériser les performances des joueurs : le temps de jeu, le skill, la capacité à rester en vie.

Certains indicateurs, plutôt conçus, en théorie, pour évaluer les performances des joueurs sont en fait surtout liés au temps de jeu (nombre de morts, nombre d'adversaires tués, de résurrections, score, ...) et ne remplissent donc pas leur rôle supposé.

Classification (Tandem).

Choix de la méthode

On a effectué l'ACP en utilisant une métrique M_{1/s^2} , c'est à dire sur le tableau de données centré-réduit. Par soucis de consistance, nous conserverons cette métrique ici.

Même si le jeu de données comporte peu d'observations, on choisit quand même de procéder par une approche mixte qui est théoriquement plus robuste. On procédera donc de la façon suivante :

- Une classification par k-means pour un grand nombre de clusters sur le résultat de l'ACP (permet un choix simple du nombre de classes).
- Une CAH en partant de la partition précédemment obtenue (permet un choix simple du nombre de classes)
- Une consolidation des classes obtenues par k-means (permet d'optimiser la partition du point de vue du critère de l'inertie)

Réalisation de la classification mixte.

On commence donc par appliquer une classification par la méthode des k-means avec un nombre élevé de centres (ici 30 centres pour 104 observations). La méthode des k-means est utilisée directement à partir des coordonnées des individus obtenues dans sur les axes de l'ACP.

On peut ensuite appliquer une classification ascendante hiérarchique (CAH) par la méthode de Ward sur la matrice des distances euclidiennes entre les centres de gravités des classes obtenus par la méthode des k-means, en tenant compte des effectifs des classes.

Le dendrogramme suggère un découpage en 4 classes, de même que le diagramme des gains d'inertie qui suggère aussi 4 classes (On peut observer un coude dans le passage de 4 à 3 classes) ou 5 classes. On retiendra donc 4 classes dans ce qui suit.

Figure 18: Dendrogramme

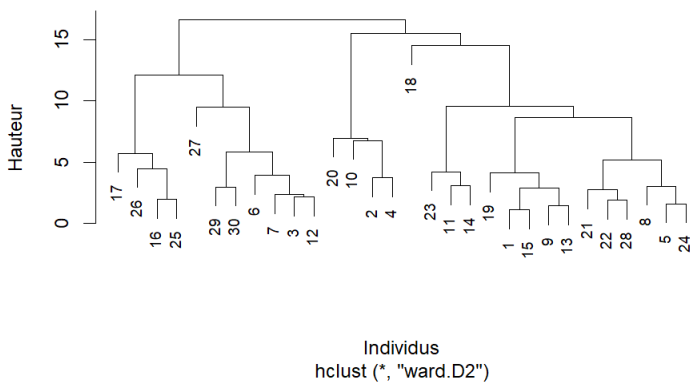
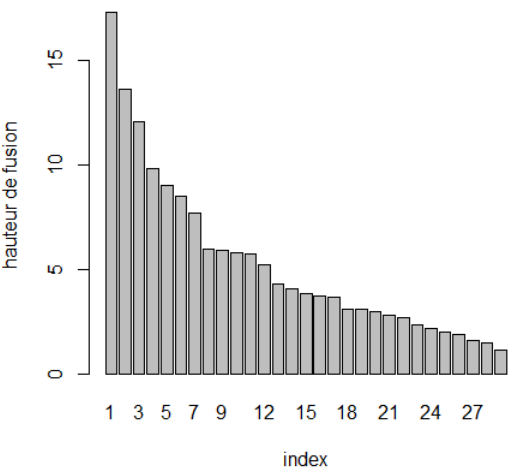


Figure 19: Diagramme des gains d'inertie



On réalise enfin une classification par la méthode des k-means avec 4 classes pour consolider le résultat précédent.

Tableau 8 : Nombre d'individus par classe :

Classe	Nombre d'individus dans la classe
1	64

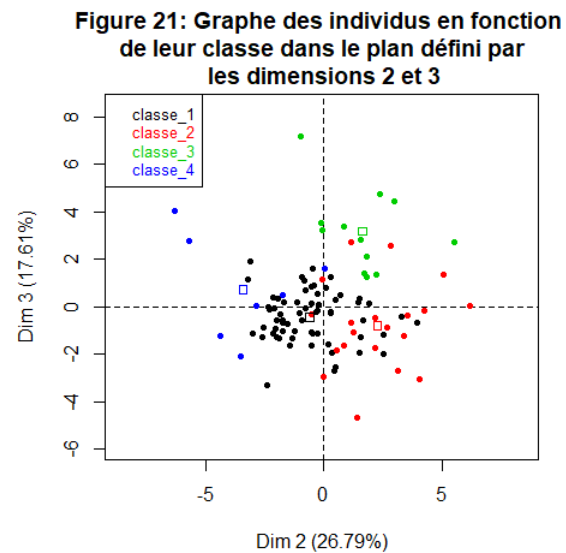
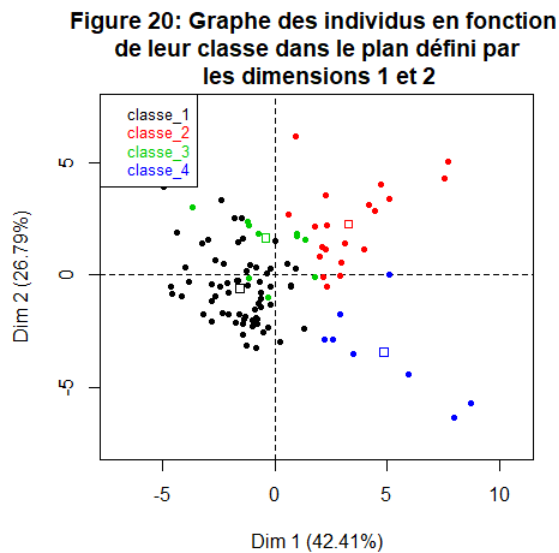
Tableau 8 : Nombre d'individus par classe :

Classe	Nombre d'individus dans la classe
2	20
3	12
4	8

Interprétation des classes

Description à partir des plans de l'ACP

On peut à présent représenter les individus en fonction de leur classe dans les plans de l'ACP.



On peut voir ainsi que :

- Les individus de la classe 2 sont des joueurs qui jouent beaucoup et ont un skill plus élevé que la moyenne.
- Les individus de la classe 1 se distinguent par le fait qu'ils jouent moins, en moyenne, que le reste de l'échantillon et semblent avoir un skill plus faible que la moyenne. Ils sont les plus nombreux.
- Les individus du groupe 3 sont des joueurs qui ont un skill plus élevé que la moyenne, jouent peu mais survivent longtemps en mode *battle royal*.
- Les individus de la classe 4 sont des joueurs qui jouent beaucoup mais ont un skill en-dessous de la moyenne. Ce sont les individus les plus rares.

Les classes 2 et 4 semblent s'opposer fortement par leurs capacités à gagner des duels.

Description à partir des variables

On peut aussi décrire les classes par les variables en analysant les résultats de la fonction *catdes* de FactoMineR (résultats disponibles en annexes 26 à 29). On peut ainsi remarquer que :

- La 1ère classe se caractérise en fait par des performances faibles dans tous les domaines et, en particulier, ce sont des joueurs qui ont du mal à gagner leurs duels (les moyennes des variables *downsPerGame*, *killsPerGame* et *kdRatio* sont significativement plus faibles que dans l'échantillon). Ce sont aussi des joueurs qui gagnent et jouent moins que la moyenne de l'échantillon.

- La 3e classe se caractérise par leur ratio de victoires élevé et leur capacité à rester en vie (et à effectuer des contrats). Ce sont aussi des joueurs qui jouent significativement moins que la moyenne de l'échantillon. On peut remarquer qu'ils ont aussi une bonne capacité à gagner leurs duels.
- La 2e classe se caractérise par un nombre élevé d'adversaires tués ou mis à terre, un score élevé, un grand nombre de victoires et un nombre de parties important. Cela au prix d'un temps de survie moyen plus faible que la moyenne de l'échantillon.
- Enfin, la 4e classe est caractérisée par des temps de jeu très élevés et un grand nombre de tops mais aussi une grande difficulté à gagner les duels.

Tableau 9 : Tests des rapports de corrélations		
Variable	Eta2	P-value
kills	0.68	0
score	0.66	0
timePlayed	0.64	0
downs	0.63	0
topTwentyFive	0.63	0
gamesPlayed	0.59	0
deaths	0.58	0
wins	0.54	0
topTen	0.53	0
wgRatio	0.52	0

Description à partir des individus

Parangons

On peut s'intéresser aux parangons de chaque classe, c'est-à-dire à l'individu le plus proche (au sens de la distance euclidienne), du centre de chaque classe.

Tableau 10: Individus les plus proches du centre de la classe 1	
Individus	distance.centre\$`1`
AA is Cheating#1235247	1.65
mrhaannn#7209929	1.67
s1mple#9853606	1.91
trav3rse#2241205	2.03
twitch lil2pac ツ #9802649	2.16
jayhawk#5140064	2.24

Tableau 11: Individus les plus proches du centre de la classe 2	
Individus	distance.centre\$`2`

Tableau 11: Individus les plus proches du centre de la classe 2

Individus	distance.centre\$`2`
iZundGod#7322165	1.68
Rias#9013253	2.25
tony wh1te#4150630	2.59
twitch_elatoo#2608520	2.84
mussee1#6257754	3.25
Newbz#5677070	3.28

Tableau 12: Individus les plus proches du centre de la classe 3

Individus	distance.centre\$`3`
Blue#7558256	2.78
H-#4693586	2.88
malvingill#9693869	2.96
nYakir_#4066522	3.02
Downtown Dieter#3252059	3.04
minenmoritz#2824891	3.47

Tableau 13: Individus les plus proches du centre de la classe 4

Individus	distance.centre\$`4`
Bombastic#7870958	3.33
SwagaShotz#8604004	3.78
itz essdoubleu#9690132	4.06
ghost#5040040	4.38
Cook14u#4591151	5.12
Bir_Hayat#6532438	5.16

Pour la classe 1, le parangon est AA is Cheating#1235247 qui est un individu qui joue peu mais ressuscite souvent.

Pour la classe 2, le parangon est iZundGod#7322165 qui est manifestement très bon dans les duels, joue beaucoup mais a du mal à survivre.

Pour la classe 3, le parangon est Blue#7558256 qui est lui aussi très bon dans les duels mais ne joue pas beaucoup et survit plus que la moyenne. C'est un excellent joueur du jeu et il gagne souvent.

Enfin, pour la classe 4, le parangon est Bombastic#7870958 qui joue beaucoup au jeu mais est plutôt mauvais dans les duels, ne gagne pas souvent et a du mal à survivre en partie.

Tableau 14 : Description des parangons

Variables	AA is Cheating#1235247	iZundGod#7322165	Blue#7558256	Bombastic#7870958
contracts	-0.5917372	0.3060933	1.081308	0.7248001
kills	-0.6390693	1.158462	0.2039805	0.1905487
downs	-0.5978547	1.318688	0.3395097	0.3520947
deaths	-0.6496363	0.4617903	-0.412669	0.7708966
kdRatio	-0.3683951	0.801672	0.7295477	-0.4855935
revives	-0.2379423	0.1370167	1.422868	1.700208
score	-0.8905327	1.060965	0.4706524	0.6298904
scorePerMinute	-0.5315846	1.070544	0.8446819	-0.4830206
timePlayed	-0.7985125	0.3832473	-0.03102652	1.16842
gamesPlayed	-0.8040679	0.4731811	-0.239971	1.137762
topFive	-0.7160998	0.6340923	1.132	0.9350149
topTen	-0.7158167	0.4694596	0.7597038	1.080322
topTwentyFive	-0.6925315	0.2882847	0.2316281	1.211213
wins	-0.5830503	0.8758074	1.416807	0.1572656
averageLifetime	0.09797402	-0.504256	0.4631672	-0.3405538
wgRatio	-0.09533853	0.1022049	1.198839	-0.6274116
killsPerGame	-0.1428908	0.8775271	0.5402796	-0.7464088
downsPerGame	-0.07090992	0.9352278	0.6576901	-0.5733161
revivesPerGames	0.6757292	-0.5046227	1.64364	0.1447667
contractsPerGames	0.02594159	-0.2191019	1.550775	-0.3137024
G_shots	0.3335519	0.1516825	-0.3783326	-0.3669991
G_accuracy	0.5101616	0.8845008	-0.3047923	-0.4664709
G_gamesPlayed	-0.05275692	0.0859137	-1.040624	0.4613547
G_headshots	-0.03633868	0.7303422	-0.6710632	-0.3505282
G_hits	0.4440219	0.3656685	-0.4096863	-0.4267867
G_misses	0.2948468	0.08338493	-0.3635169	-0.3435333
G_kdRatio	-0.6040096	0.681757	0.9283567	-0.5286152
G_suicides	0.1839777	1.247974	-0.2380815	-0.3281115
G_scorePerMinute	-0.1257374	-0.120466	-0.1120831	-0.1197011
G_hkRatio	-0.1099991	0.04288782	-0.8325903	-0.5431663
G_wgRatio	0.3337111	-0.1143952	0.0313672	-0.3795979
classe	1	2	3	4

Conclusion

On peut donc proposer une classification des meilleurs joueurs mondiaux de Call Of Duty : Warzone en 4 groupes :

- Un groupe d'individus significativement meilleurs que la moyenne au jeu. Ils n'ont pas besoin de jouer autant que les autres pour figurer dans le top mondial. Ils ont une bonne capacité à gagner leurs duels mais surtout, ils savent survivre en cours de partie (et font beaucoup de contrats). Leur capacité à survivre est plus importante que leur capacité à gagner des duels.
- Un groupe d'individus qui ont un bon skill mais ne parviennent pas à rester en vie. On peut supposer que ces joueurs sont trop agressifs et qui compensent leur faible ratio de victoires par un temps de jeu plus important.
- Un groupe d'individus qui compensent leur faible niveau de jeu (en particulier leur difficulté à gagner les combats) par un énorme temps de jeu.
- Un groupe qui joue moins que la moyenne et n'est pas particulièrement doué mais qui représente l'essentiel des joueurs du top mondial.

Conclusion

Au terme de ce document, l'on aura réussi à montrer que de nombreuses données brutes de performances des joueurs (victoires, défaites, morts, adversaires tués, score, ...) sont linéairement corrélées au temps de jeu et ne sont donc pas de très bons indicateurs du niveau de jeu d'un joueur.

On aura aussi montré que les 3 plus gros facteurs de variabilité au sein des joueurs du top mondial sont leur temps de jeu, les capacités à gagner les combats et leur capacité à survivre sur le champ de bataille.

Enfin, on aura le document fourni une classification des joueurs en 4 groupes :

- Les joueurs excellents qui gagnent beaucoup car ils sont très bons au jeu.
- Les joueurs qui ont un bon niveau de jeu mais sont probablement trop agressifs et finissent par mourir. Pour compenser, ils doivent jouer beaucoup.
- La majorité des joueurs qui ne sont pas particulièrement bons et ne jouent pas énormément mais survivent un peu mieux en moyenne que le second groupe.
- Enfin, des joueurs, plus rares, qui n'ont pas un très bon niveau de jeu et doivent donc jouer beaucoup pour compenser.

Il semble donc se dégager du travail effectué une stratégie qui permet d'améliorer son ratio victoires/défaites : éviter les affrontements aussi longtemps que possible, remplir des contrats, savoir bien jouer les affrontements qui permettent de ressusciter et s'entraîner pour être efficace dans les duels même s'il n'est pas nécessaire d'être incroyable non plus sur ce point. A l'inverse, la stratégie qui semble la moins efficace est celle qui consiste à privilégier les affrontements (même si on est très bon) et à être agressif.

Annexes

Annexe 1 : Algorithme Python pour la récupération des données des joueurs

Annexe 1 : Algorithme Python pour la récupération des données des joueurs

```
from cod_api import API, platforms
import pandas as pd

#For Spyder compatibility
import nest_asyncio
nest_asyncio.apply()

api = API()

# Login in with sso token
api.login('your SSO token')

def getProfiles(aListOfPlayersNames) :
    """
    Parameters
    -----
    aListOfPlayersNames : List[string]
        a list of strings with players usernames.
        They must be Activision platform's usernames.

    Returns
    -----
        a list of public profiles pulled from Activision database using the unofficial API cod-
    api.
    """
    n = 0
    profiles = []
    for i in aListOfPlayersNames :
        profile = api.Warzone.fullData(platforms.Activision, i)
        if profile['status'] == 'success' :
            profiles.append(profile['data'])
            n = n + 1
    return profiles

def makeDataFrame(profiles):
    """
    Takes a list of profiles from the Activision Database and return a structured dataframe.

    Parameters
    -----
    profiles : List[dict]
        a list of dict containing data from users pulled from the Activision database.

    Returns
    -----
        a dataframe of users datas.
    """
    players = []
    for i in range(0, len(profiles)) :
        #General datas about the player
        username = profiles[i]['username']
        G_totalXp = profiles[i]['totalXp']
```

```
#Type : dict : contains datas from all modes
lifetime = profiles[i]['lifetime']['all']['properties']
```

```
G_accuracy = lifetime['accuracy']
G_gamesPlayed = lifetime['gamesPlayed']
G_headshots = lifetime['headshots']
G_hits = lifetime['hits']
G_misses = lifetime['misses']
G_kdRatio = lifetime['kdRatio']
G_suicides = lifetime['suicides']
G_shots = lifetime['totalShots']
G_scorePerMinute = lifetime['scorePerMinute']
G_hkRatio = lifetime['headshots'] / lifetime['kills']
G_wgRatio = lifetime['wins'] / lifetime['gamesPlayed']
```

```
#Type : dict : contains datas from Battle Royal mode
br = profiles[i]['lifetime']['mode']['br_all']['properties']
```

```
contracts = br['contracts']
kills = br['kills']
downs = br['downs']
deaths = br['deaths']
kdRatio = br['kdRatio']
revives = br['revives']
score = br['score']
scorePerMinute = br['scorePerMinute']
# time in second
timePlayed = br['timePlayed']
gamesPlayed = br['gamesPlayed']
topFive = br['topFive']
topTen = br['topTen']
topTwentyFive = br['topTwentyFive']
wins = br['wins']
averageLifetime = timePlayed / gamesPlayed
wgRatio = wins / gamesPlayed
killsPerGame = kills / gamesPlayed
downsPerGame = downs / gamesPlayed
```

```
player = [username, G_totalXp, G_accuracy, G_gamesPlayed, G_headshots, G_hits,
          G_misses, G_kdRatio, G_suicides, G_shots, G_scorePerMinute, G_hkRatio,
          G_wgRatio, contracts, kills, downs, deaths, kdRatio, revives, score,
          scorePerMinute, timePlayed, gamesPlayed, topFive, topTen,
          topTwentyFive, wins, averageLifetime, wgRatio, killsPerGame,
          downsPerGame]
```

```
players.append(player)
```

```
names = ['username', 'G_totalXp', 'G_accuracy', 'G_gamesPlayed', 'G_headshots',
        'G_hits', 'G_misses', 'G_kdRatio', 'G_suicides', 'G_shots',
        'G_scorePerMinute', 'G_hkRatio', 'G_wgRatio', 'contracts', 'kills', 'downs',
        'deaths', 'kdRatio', 'revives', 'score', 'scorePerMinute', 'timePlayed',
        'gamesPlayed', 'topFive', 'topTen', 'topTwentyFive', 'wins',
        'averageLifetime', 'wgRatio', 'killsPerGame', 'downsPerGame']
```

```
return pd.DataFrame(players, columns = names)
```

```
def createCSV(df) :
```

```
    """
```

```
    Create a CSV from a dataframe.
```

Parameters

df : DataFrame

A Pandas dataframe.

Returns

None.

"""

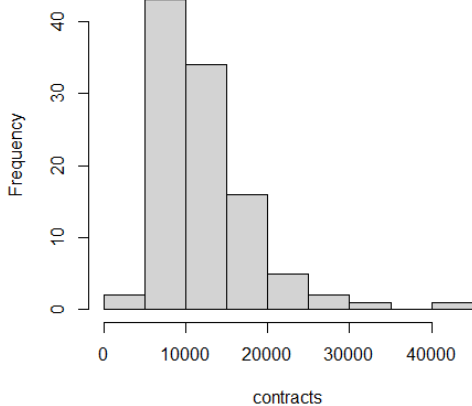
```
df.to_csv('codPlayers_01avril2023.csv',header = True, sep = ';', index = False)
```

```
activisionIds = ['a','list','of','Activision','username', 'such','as','Username#0001']
```

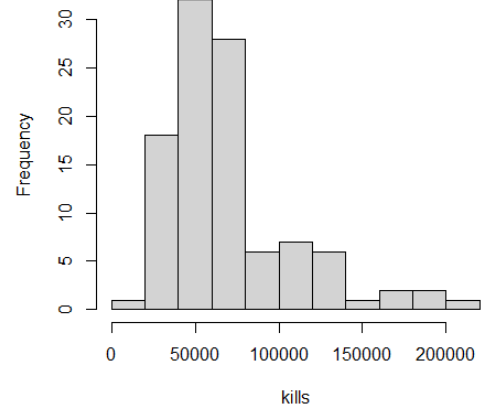
```
profiles = getProfiles(activisionIds)
print("Vous avez récupéré ",len(profiles)," profiles.")
df = makeDataFrame(profiles)
createCSV(df)
```

Annexes 2 à 21 : Histogrammes des variables actives

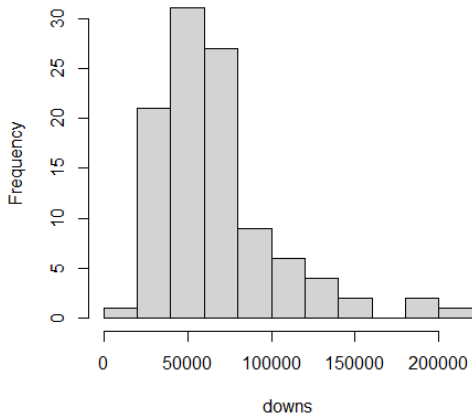
**Annexe 2 : Histogramme de la variable
contracts
skewness = 1.96**



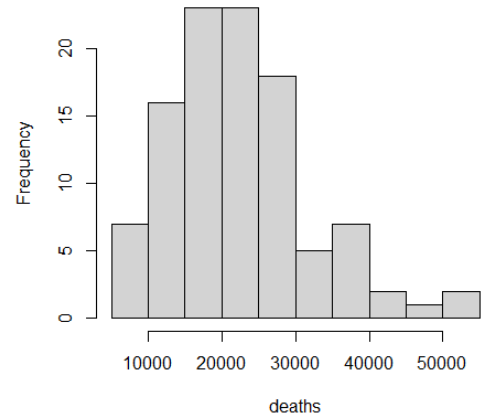
**Annexe 3 : Histogramme de la variable
kills
skewness = 1.57**



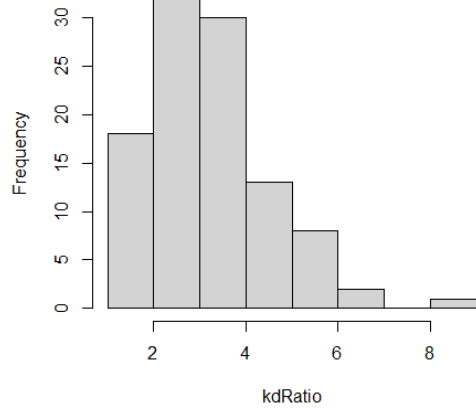
**Annexe 4 : Histogramme de la variable
downs
skewness = 1.61**



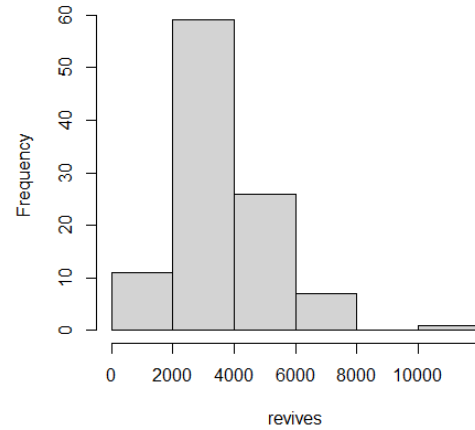
**Annexe 5 : Histogramme de la variable
deaths
skewness = 0.8**



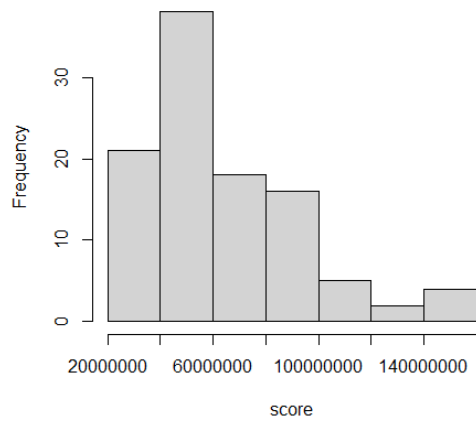
**Annexe 6 : Histogramme de la variable
kdRatio
skewness = 1.05**



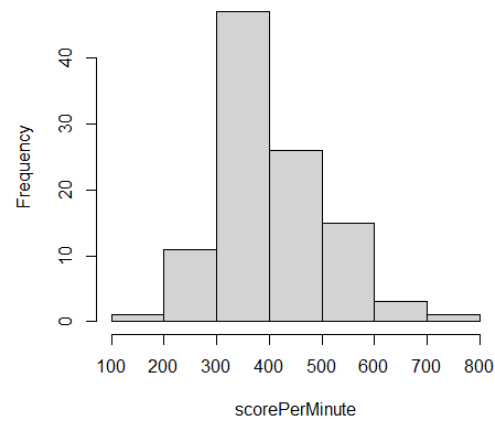
**Annexe 7 : Histogramme de la variable
revives
skewness = 1.54**



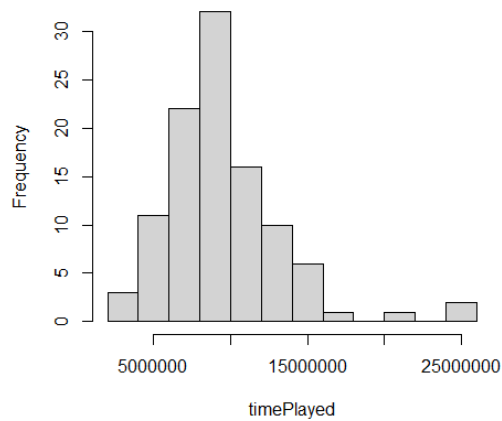
**Annexe 8 : Histogramme de la variable
score
skewness = 1.11**



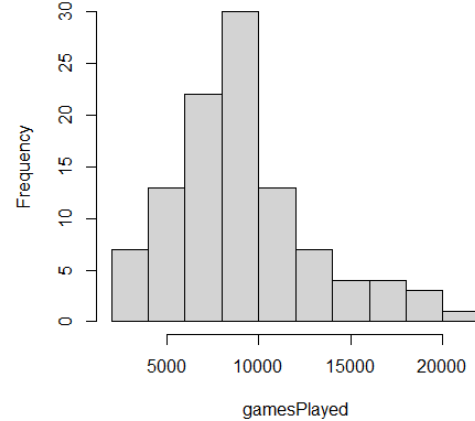
**Annexe 9 : Histogramme de la variable
scorePerMinute
skewness = 0.74**



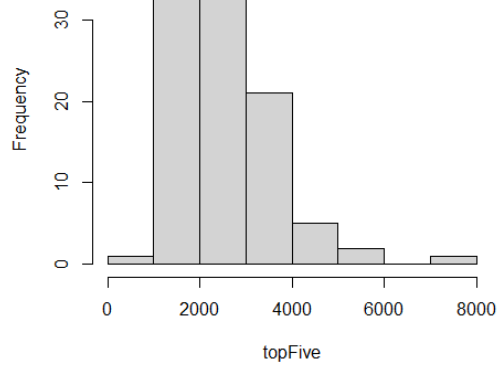
**Annexe 10 : Histogramme de la variable
timePlayed
skewness = 1.52**



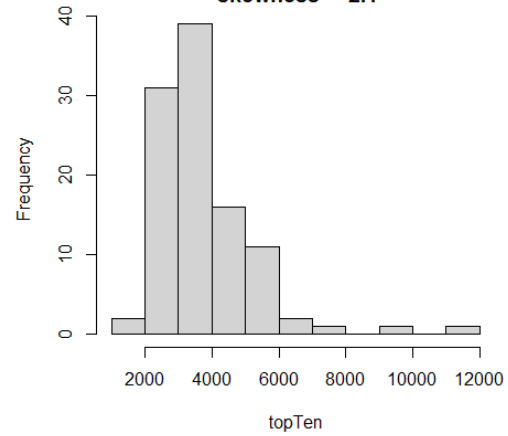
**Annexe 11 : Histogramme de la variable
gamesPlayed
skewness = 0.92**



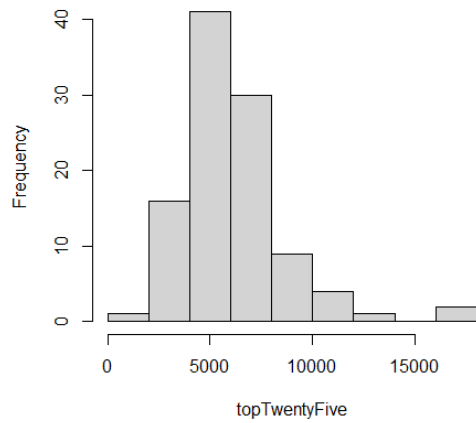
**Annexe 12 : Histogramme de la variable
topFive
skewness = 1.7**



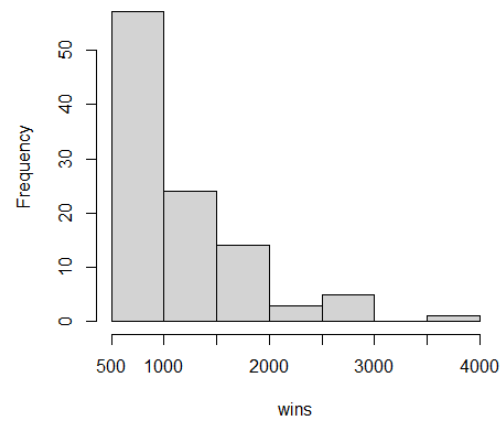
**Annexe 13 : Histogramme de la variable
topTen
skewness = 2.1**



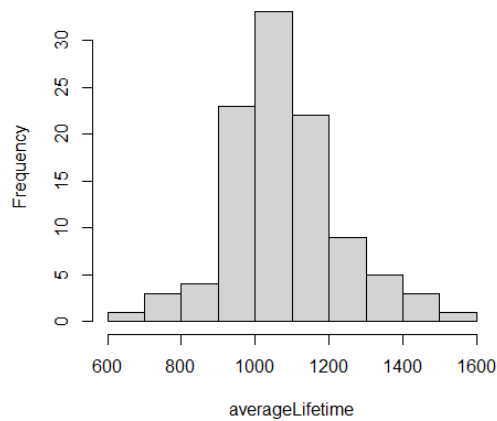
**Annexe 14 : Histogramme de la variable
topTwentyFive
skewness = 1.72**



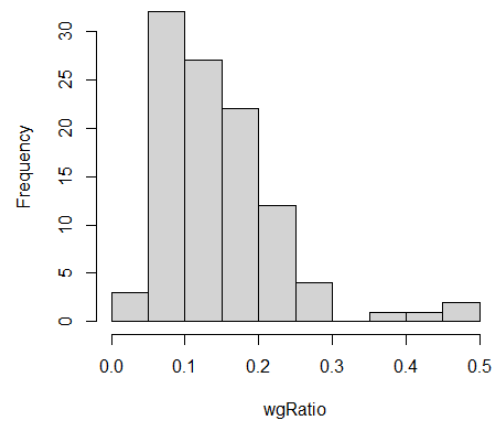
**Annexe 15 : Histogramme de la variable
wins
skewness = 1.71**



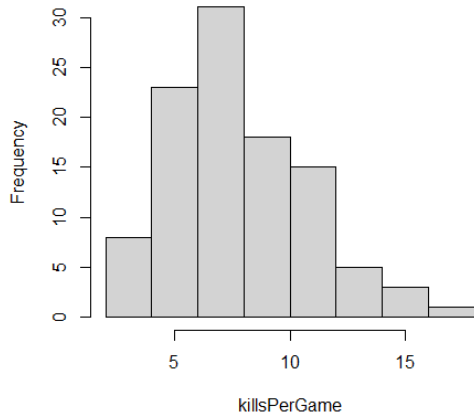
**Annexe 16 : Histogramme de la variable
averageLifetime
skewness = 0.34**



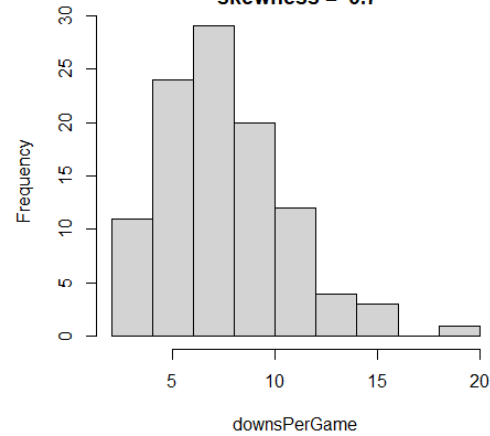
**Annexe 17 : Histogramme de la variable
wgRatio
skewness = 1.84**



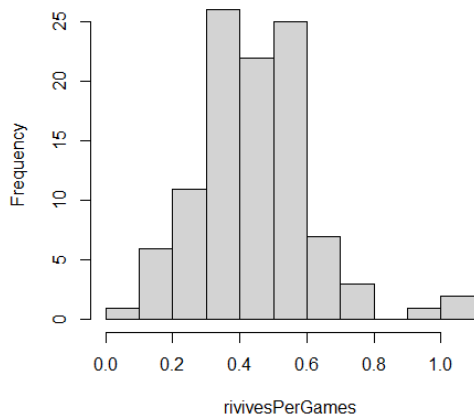
**Annexe 18 : Histogramme de la variable
killsPerGame
skewness = 0.57**



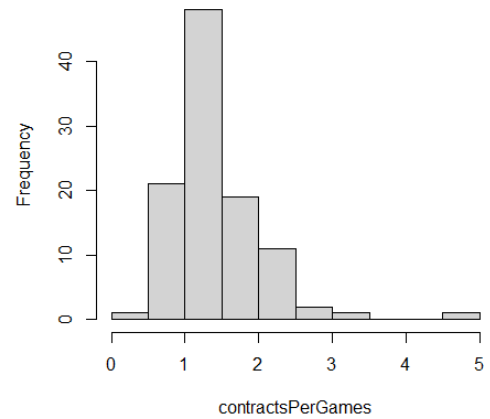
**Annexe 19 : Histogramme de la variable
downsPerGame
skewness = 0.7**



**Annexe 20 : Histogramme de la variable
rivesPerGames
skewness = 0.75**



**Annexe 21 : Histogramme de la variable
contractsPerGames
skewness = 2.39**



Annexe 22 : Résumé de la variance expliquée en fonction des valeurs propres

Annexe 22 : Résumé de la variance expliquée en fonction des valeurs propres

composante	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	8.069	40.343	40.343
comp 2	5.091	25.457	65.800
comp 3	3.894	19.471	85.271
comp 4	1.117	5.584	90.854
comp 5	0.544	2.722	93.577
comp 6	0.460	2.301	95.877
comp 7	0.249	1.243	97.121
comp 8	0.172	0.862	97.983
comp 9	0.139	0.697	98.680
comp 10	0.103	0.514	99.193
comp 11	0.058	0.288	99.481

Annexe 22 : Résumé de la variance expliquée en fonction des valeurs propres

composante	eigenvalue	percentage of variance	cumulative percentage of variance
comp 12	0.039	0.196	99.676
comp 13	0.024	0.119	99.796
comp 14	0.016	0.079	99.874
comp 15	0.010	0.048	99.923
comp 16	0.008	0.040	99.963
comp 17	0.004	0.019	99.982
comp 18	0.002	0.009	99.991
comp 19	0.001	0.006	99.997
comp 20	0.001	0.003	100.000

Annexe 23 : Table des corrélations entre la première dimension de l'ACP et les variables du jeu de données

Annexe 23 : Table des corrélations entre la première dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
timePlayed	0.93	0.00
topTwentyFive	0.91	0.00
gamesPlayed	0.91	0.00
score	0.90	0.00
deaths	0.87	0.00
topTen	0.86	0.00
topFive	0.81	0.00
contracts	0.74	0.00
G_gamesPlayed	0.73	0.00
kills	0.73	0.00
downs	0.72	0.00
revives	0.60	0.00
wins	0.56	0.00
G_headshots	0.41	0.00
G_suicides	0.38	0.00
scorePerMinute	0.22	0.03
kdRatio	0.06	0.56
G_kdRatio	0.05	0.60

Annexe 23 : Table des corrélations entre la première dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
G_hits	0.04	0.72
killsPerGame	0.03	0.75
G_scorePerMinute	0.01	0.94
G_shots	0.00	0.99
G_misses	-0.01	0.90
downsPerGame	-0.01	0.90
contractsPerGames	-0.11	0.25
G_hkRatio	-0.13	0.19
G_accuracy	-0.20	0.04
G_wgRatio	-0.22	0.02
wgRatio	-0.23	0.02
averageLifetime	-0.27	0.01
rivivesPerGames	-0.32	0.00

Annexe 24 : Table des corrélations entre la deuxième dimension de l'ACP et les variables du jeu de données

Annexe 24 : Table des corrélations entre la seconde dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
killsPerGame	0.96	0.00
kdRatio	0.90	0.00
scorePerMinute	0.88	0.00
downsPerGame	0.87	0.00
G_kdRatio	0.82	0.00
wgRatio	0.58	0.00
kills	0.58	0.00
G_accuracy	0.56	0.00
downs	0.55	0.00
wins	0.52	0.00
G_headshots	0.52	0.00
score	0.32	0.00
G_hkRatio	0.17	0.08
G_hits	0.15	0.13

Annexe 24 : Table des corrélations entre la seconde dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
G_shots	0.07	0.51
G_wgRatio	0.05	0.64
G_misses	0.04	0.69
contractsPerGames	0.03	0.78
averageLifetime	0.01	0.93
revivesPerGames	0.01	0.95
G_scorePerMinute	0.00	1.00
topFive	0.00	0.98
G_suicides	-0.09	0.37
topTen	-0.20	0.05
deaths	-0.22	0.02
gamesPlayed	-0.23	0.02
G_gamesPlayed	-0.25	0.01
contracts	-0.26	0.01
revives	-0.29	0.00
timePlayed	-0.31	0.00
topTwentyFive	-0.32	0.00

Annexe 25 : Table des corrélations entre la troisième dimension de l'ACP et les variables du jeu de données

Annexe 25 : Table des corrélations entre la troisième dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
averageLifetime	0.86	0.00
revivesPerGames	0.81	0.00
contractsPerGames	0.79	0.00
wgRatio	0.71	0.00
topFive	0.50	0.00
revives	0.48	0.00
wins	0.47	0.00
contracts	0.41	0.00
topTen	0.38	0.00
topTwentyFive	0.13	0.20

Annexe 25 : Table des corrélations entre la troisième dimension de l'ACP et les variables du jeu de données

variable	correlation	p.value
downsPerGame	0.10	0.32
G_misses	0.09	0.38
G_wgRatio	0.08	0.40
G_shots	0.07	0.46
timePlayed	0.05	0.64
G_hits	0.02	0.82
kdRatio	-0.01	0.88
killsPerGame	-0.02	0.81
G_scorePerMinute	-0.05	0.58
G_kdRatio	-0.06	0.56
G_hkRatio	-0.08	0.44
G_accuracy	-0.13	0.18
score	-0.14	0.15
downs	-0.19	0.06
scorePerMinute	-0.19	0.05
gamesPlayed	-0.28	0.00
deaths	-0.28	0.00
kills	-0.28	0.00
G_headshots	-0.29	0.00
G_gamesPlayed	-0.32	0.00
G_suicides	-0.33	0.00

Annexe 26 : Caractérisation de la première classe par les variables quantitatives

Annexe 26 : Caractérisation de la première classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
G_kdRatio	-2.68	-0.21	0	0.83	1	0.01
wgRatio	-3.10	-0.24	0	0.59	1	0.00
kdRatio	-3.27	-0.25	0	0.87	1	0.00
downsPerGame	-3.29	-0.26	0	0.81	1	0.00
killsPerGame	-3.39	-0.26	0	0.78	1	0.00
G_headshots	-3.67	-0.28	0	0.78	1	0.00

Annexe 26 : Caractérisation de la première classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
G_gamesPlayed	-3.73	-0.29	0	0.81	1	0.00
scorePerMinute	-4.03	-0.31	0	0.76	1	0.00
revives	-5.00	-0.39	0	0.65	1	0.00
deaths	-5.04	-0.39	0	0.61	1	0.00
contracts	-5.16	-0.40	0	0.55	1	0.00
gamesPlayed	-5.20	-0.40	0	0.62	1	0.00
topTwentyFive	-5.74	-0.45	0	0.52	1	0.00
timePlayed	-5.79	-0.45	0	0.53	1	0.00
kills	-6.12	-0.47	0	0.42	1	0.00
topTen	-6.22	-0.48	0	0.48	1	0.00
downs	-6.25	-0.48	0	0.43	1	0.00
topFive	-6.75	-0.52	0	0.47	1	0.00
wins	-6.75	-0.52	0	0.42	1	0.00
score	-7.06	-0.55	0	0.49	1	0.00

Annexe 27 : Caractérisation de la deuxième classe par les variables quantitatives

Annexe 27 : Caractérisation de la seconde classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
kills	8.22	1.65	0	0.90	1	0.00
downs	7.87	1.58	0	0.97	1	0.00
score	7.23	1.45	0	0.83	1	0.00
G_headshots	5.92	1.19	0	0.91	1	0.00
scorePerMinute	5.72	1.15	0	0.84	1	0.00
wins	5.16	1.04	0	1.20	1	0.00
killsPerGame	4.68	0.94	0	0.82	1	0.00
gamesPlayed	4.64	0.93	0	0.67	1	0.00
deaths	4.52	0.91	0	0.63	1	0.00
kdRatio	4.29	0.86	0	0.71	1	0.00
G_kdRatio	3.87	0.78	0	0.75	1	0.00
downsPerGame	3.86	0.78	0	0.94	1	0.00
timePlayed	3.43	0.69	0	0.48	1	0.00

Annexe 27 : Caractérisation de la seconde classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
G_gamesPlayed	3.41	0.68	0	0.67	1	0.00
topFive	3.35	0.67	0	0.93	1	0.00
topTwentyFive	2.99	0.60	0	0.51	1	0.00
topTen	2.78	0.56	0	0.71	1	0.01
G_suicides	2.28	0.46	0	1.13	1	0.02
contractsPerGames	-2.10	-0.42	0	0.64	1	0.04
rivivesPerGames	-2.95	-0.59	0	0.78	1	0.00
averageLifetime	-3.58	-0.72	0	0.76	1	0.00

Annexe 28 : Caractérisation de la troisième classe par les variables quantitatives

Annexe 28 : Caractérisation de la troisième classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
wgRatio	7.04	1.91	0	1.20	1	0.00
rivivesPerGames	5.82	1.58	0	1.09	1	0.00
averageLifetime	5.50	1.49	0	0.88	1	0.00
contractsPerGames	4.76	1.29	0	1.60	1	0.00
wins	4.26	1.16	0	0.73	1	0.00
downsPerGame	2.72	0.74	0	0.81	1	0.01
topFive	2.53	0.69	0	0.75	1	0.01
kdRatio	2.34	0.63	0	0.85	1	0.02
killsPerGame	2.13	0.58	0	0.89	1	0.03
revives	2.00	0.54	0	0.91	1	0.05
G_gamesPlayed	-2.36	-0.64	0	0.86	1	0.02
gamesPlayed	-2.37	-0.64	0	0.53	1	0.02
deaths	-2.51	-0.68	0	0.66	1	0.01

Annexe 29 : Caractérisation de la troisième classe par les variables quantitatives

Annexe 29 : Caractérisation de la quatrième classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
G_kdRatio	-2.68	-0.21	0	0.83	1	0.01
wgRatio	-3.10	-0.24	0	0.59	1	0.00
kdRatio	-3.27	-0.25	0	0.87	1	0.00

Annexe 29 : Caractérisation de la quatrième classe par les variables quantitatives

Variable	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
downsPerGame	-3.29	-0.26	0	0.81	1	0.00
killsPerGame	-3.39	-0.26	0	0.78	1	0.00
G_headshots	-3.67	-0.28	0	0.78	1	0.00
G_gamesPlayed	-3.73	-0.29	0	0.81	1	0.00
scorePerMinute	-4.03	-0.31	0	0.76	1	0.00
revives	-5.00	-0.39	0	0.65	1	0.00
deaths	-5.04	-0.39	0	0.61	1	0.00
contracts	-5.16	-0.40	0	0.55	1	0.00
gamesPlayed	-5.20	-0.40	0	0.62	1	0.00
topTwentyFive	-5.74	-0.45	0	0.52	1	0.00
timePlayed	-5.79	-0.45	0	0.53	1	0.00
kills	-6.12	-0.47	0	0.42	1	0.00
topTen	-6.22	-0.48	0	0.48	1	0.00
downs	-6.25	-0.48	0	0.43	1	0.00
topFive	-6.75	-0.52	0	0.47	1	0.00
wins	-6.75	-0.52	0	0.42	1	0.00
score	-7.06	-0.55	0	0.49	1	0.00