

Rapport Automatisé - Rmarkdown

V. Tolon & V. Payet

03 mai 2022

Contents

Introduction:	1
Présentation des données	2
Les données	2
Une première analyse	3
Histogrammes	3
Corrélations	3
Analyse en Composante Principale	4
Rappels	4
Les valeurs propres	4
Le partitionnement	4
Le cercle des corrélations	4
Le nuage des individus	5
Description des groupes	5
Conclusions	5
Références	5

Introduction:

Dans ce rapport nous analyserons des profils de joueurs offensifs du championnat de France sur la saison 2017-2018. Les analyses seront faites avec R (R Core Team (2019)).

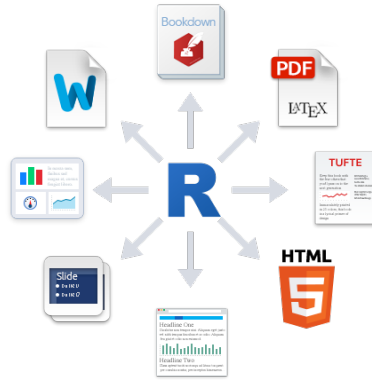


Figure 1: Logo Rmarkdown

Présentation des données

Nous transformons les données en fréquence d'événements par heure de jeu.

```
jouA <- read.csv2("joueurs.csv",header=TRUE,sep=";",row.names=1,encoding = "latin1")
jou1<-jouA[,c(1:2,4:8)]
jou1b<-jou1/(jouA$Minutes.jouees/60)
jou<-data.frame(jou1b,Minutes.jouees=jouA$Minutes.jouees)
jou<-na.omit(jou)
```

Les données

```
names(jou)
```

```
## [1] "Buts.marques"      "Passes.decisives" "Tirs"              "Tirs.cadres"
## [5] "Pourcent.buts"    "Fautes.subies"    "Fautes.commises"  "Minutes.jouees"
```

```
dim(jou)
```

```
## [1] 27  8
```

```
head(round(jou,3),3)
```

```
##      Buts.marques Passes.decisives  Tirs Tirs.cadres Pourcent.buts
## Neymar           0.635           0.434 3.040         1.403         0.601
## M'bappé          0.360           0.194 2.185         0.941         0.415
## Fekir            0.436           0.170 2.156         1.187         0.509
##      Fautes.subies Fautes.commises Minutes.jouees
## Neymar           3.441           0.802         1796
## M'bappé          1.079           0.249         2169
## Fekir            2.471           1.696         2477
```

Le jeu de données est complété pour 27 joueurs et contient 8 variables. La fréquence moyenne de buts marqués par heure de jeu et par attaquant est égale à 0.32 ± 0.17 (moyenne \pm écart type).

Une première analyse

Histogrammes

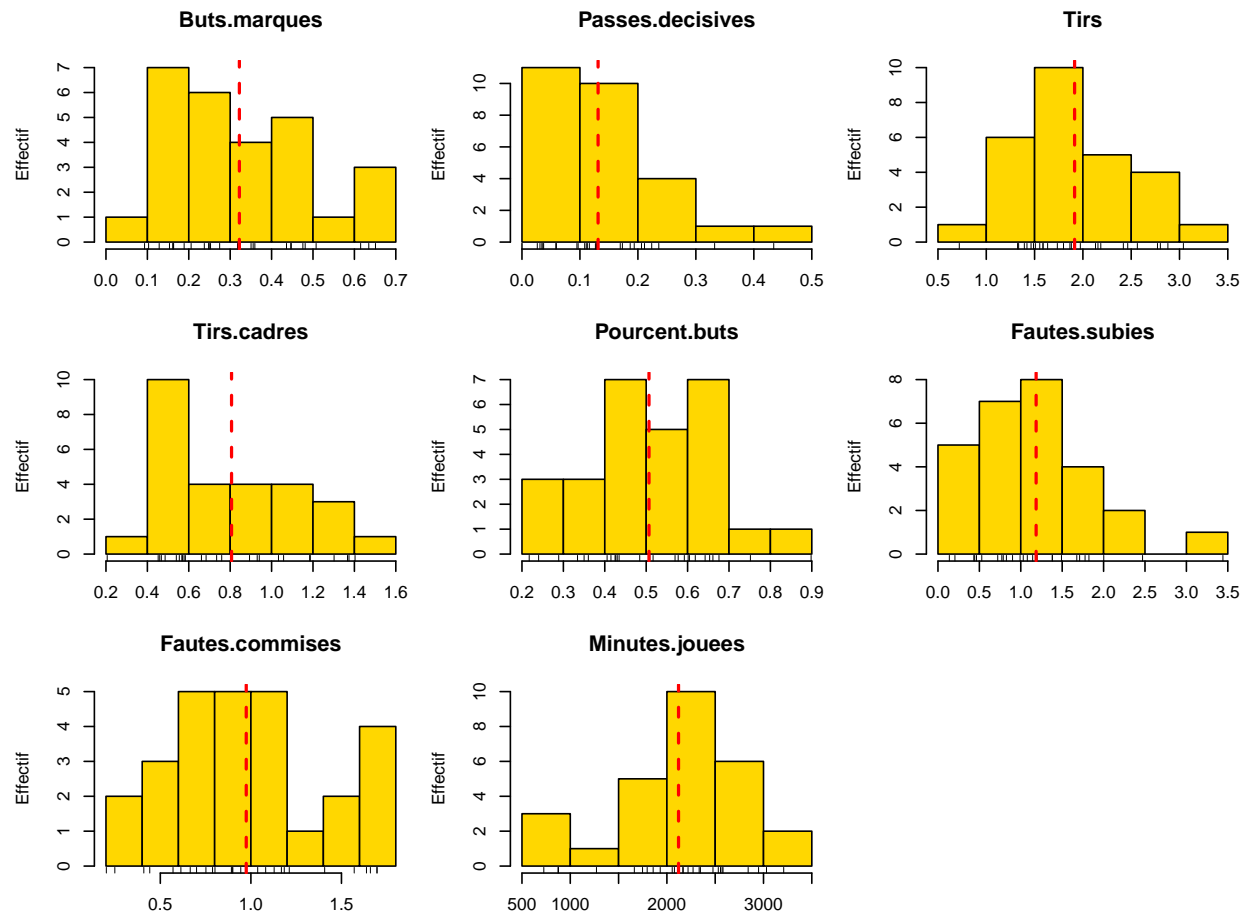


Figure 2: Histogrammes

Interprétation: Nous observons que

Corrélations

On peut les représenter comme ceci :

	Buts.marques	Passes.decisives	Tirs	Tirs.cadres	Pourcent.buts	Fautes.subies	Fautes.commises	Minutes.jouees
Buts.marques	1.000	0.228	0.786	0.888	0.636	0.015	-0.123	-0.097
Passes.decisives	0.228	1.000	0.327	0.239	-0.131	0.349	-0.376	0.096
Tirs	0.786	0.327	1.000	0.921	0.452	0.071	-0.136	-0.030
Tirs.cadres	0.888	0.239	0.921	1.000	0.487	0.053	-0.125	-0.083
Pourcent.buts	0.636	-0.131	0.452	0.487	1.000	0.061	0.173	-0.126
Fautes.subies	0.015	0.349	0.071	0.053	0.061	1.000	0.111	0.253
Fautes.commises	-0.123	-0.376	-	-0.125	0.173	0.111	1.000	-0.188
Minutes.jouees			0.136					

	Buts.marqués	Passes.decisives	Tirs	Tirs.cadres	Pourcent.but	Fautes.subies	Fautes.commises	Minutes.jouees
Minutes.jouees	-0.097	0.096	- 0.030	-0.083	-0.126	0.253	-0.188	1.000

Interprétation:

Analyse en Composante Principale

Rappels

Une ACP permet d'analyser simultanément les liens entre de multiples variables quantitatives et d'analyser les positions des individus vis à vis de l'ensemble de ces variables. Elle est basée sur la recherche d'axes principaux indépendants, chacun plus ou moins liés aux variables d'entrées. Pour rappel l'inertie totale se calcule par :

$$I = \sum_{i=1}^N \frac{1}{N} d_{(e_i;g)}^2$$

Avec $d_{e_i,g}^2 = \sum_{j=1}^p x_{ij}^2$ = Distance euclidienne au centre de gravité du nuage de point avec des données centrées et normées.

Les valeurs propres

Elles permettent de déterminer la proportion d'inertie expliquée par chacun des axes :

```
library(FactoMineR)

PCAs <- PCA(jou,graph=F)

#Choix des axes
par(mfrow=c(1,1),mar=c(4,4,3,3))
barplot(PCAs$eig[,2],ylab="Inertie %",names.arg = paste("Axe",1: nrow(PCAs$eig)),las=2,cex.axis=0.7,cex
```

Interprétation:

Le partitionnement

```
#Partitionnement
hcjou<-HCPC(PCAs,graph=F,nb.clust=4,consol = F)
plot(hcjou,choice="tree")
```

Interprétation: Nous observons que

Le cercle des corrélations

```
plot(PCAs,choix="var")
```

Interprétation: Nous observons que l'axe 1 est expliqué par

Le nuage des individus

```
plot(hcjou,choice="map",draw.tree = F)
```

Interprétation: Nous observons que

Description des groupes

```
lapply(hcjou$desc.var$quanti,function(x) return(round(data.frame(x)[-6],2)))
```

Interprétation:

Les groupes sont décrits dans la table suivante :

Table 2: Description des groupes

Numero	Nom	Description
1	A	Les mauvais joueurs provocateurs
2	B	Les mauvais buteurs
3	C	Les bons buteurs sans histoire
4	D	Les bons buteurs sous pression

Conclusions

Les joueurs de foot sont vraiment des gens très très forts (Wisloeff et al., 1998). Néanmoins nous savons que les performances ne sont pas les mêmes suivant leurs profils (*attaquants* : Little, Williams, 2003 ; et *défenseurs* : Mohr et al., 2003). Ceci pourrait fortement impacter leurs tactiques sur le terrain même si Wisloeff et al. (1998) suggère qu'au fond tout cela n'est qu'un jeu sans importance.

Références

Liens

<https://www.footballdatabase.eu/fr/>

<https://rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>
yihui/rmarkdown/

<https://bookdown.org/>

Bibliographie

LITTLE, Thomas et WILLIAMS, Alun, 2003. *Specificity of acceleration, maximum speed and agility in professional soccer players*. London, UK : Routledge.

MOHR, Magni, KRUSTRUP, Peter et BANGSBO, Jens, 2003. Match performance of high-standard soccer players with special reference to development of fatigue. In : *Journal of sports sciences*. 2003. Vol. 21, n° 7, p. 519-528.

R CORE TEAM, 2019. *R: A Language and Environment for Statistical Computing* [en ligne]. Vienna, Austria : R Foundation for Statistical Computing. Disponible à l'adresse : <https://www.R-project.org/>.

WISLOEFF, ULRIK, HELGERUD, JAN et HOFF, JAN, 1998. Strength and endurance of elite soccer players. In : *Medicine and science in sports and exercise*. 1998. Vol. 30, n° 3, p. 462-467.