

Devoir maison - Analyse de données

Benmaddi Yanis

Jasny Hamza

L3 MIASHS

Vous souvenez-vous des vainqueurs de la coupe du monde de football 2014 au Brésil ? L'Allemagne bien évidemment avec des performances hors normes (7-1 contre le Brésil en demi-finale...).

Mais saviez vous que cette sélection avait recours à l'analyse des données biométriques et les déplacements des joueurs ainsi que l'historique des matchs joués par l'Allemagne et ses adversaires d'après le journal L'Express.

La Mannschaft a ainsi démontré que l'étude ainsi que l'analyse de données pouvaient avoir un impact vraiment positif sur les performances réalisées.

De cela, nous avons décidé de réaliser notre étude sur la saison de Premier League 2018-2019 pour expliquer, à travers quelques paramètres tels que les buts marqués à l'extérieur, les fautes ou encore les corners tirés, le succès de Manchester City talonné de très près (1 point) par Liverpool lui-même à 25 points du troisième : Chelsea.

On se demande ainsi comment expliquer la domination de ces 2 équipes sur le championnat.

Pour répondre à cette question, nous avons scindé les variables en 2 catégories dans la base de données étudiée ; les variables actives et supplémentaires que nous avons préalablement renommés pour plus de lisibilité

Variables actives : But(s) Domicile Fin Temps Réglementaire, But(s) Extérieur Fin Temps Réglementaire, But(s) Domicile Mi-Temps, But(s) Extérieur Mi-Temps, Tir(s) tenté(s) Domicile, Tir(s) tenté(s) Extérieur, Tir(s) tenté(s) cadré(s) Domicile, Tir(s) tenté(s) cadré(s) Extérieur, Faute(s) Domicile, Faute(s) Extérieur, Corner(s) Domicile, Corner(s) Extérieur, Carton(s) Jaune Domicile, Carton(s) Jaune Extérieur, Carton(s) Rouge Domicile, Carton(s) Rouge Extérieur

Variables supplémentaires : Division, Date, Domicile, Extérieur, Résultat final, Résultat mi-temps

Dans ce cadre nous avons plus d'informations qualitatives à étudier donc nous avons décidé de partir sur une ACP (Analyse en Composantes Principales).

Comme précisé dans notre code, nous avons un total de 16 axes à étudier (car 16 variables actives). Nous n'allons évidemment pas les étudier tous mais uniquement les plus pertinents. Grâce à la règle de Kaiser qui exprime l'utilité des axes à étudier en fonction des valeurs propres de chaque axe ; chaque axe doit avoir une valeur propre $\lambda_k > 1$.

On retient donc seulement les axes 1 à 6. Après étude de chaque axe on constate que l'axe 4, 5 et 6 ont à peu près les mêmes variables contributives et par conséquent expriment à peu près la même chose. On va donc s'intéresser aux axes 1 à 4 inclus.

Résultats et interprétation :

- *Axe 1 : Les frappes tentées*

Ce premier axe représente 23% de l'inertie totale. Il va de soi que le nombre de frappes tentées est un indicateur de l'efficacité offensive d'une équipe. Qu'elles soient cadrées ou non, cela est un bon indicateur d'une attaque organisée qui sait se procurer des occasions et donc possiblement des buts. Sur cet axe, les variables les plus contributives sont les [tirs tentés à domicile](#) et les [tirs tentés à l'extérieur](#). On constate que les valeurs les plus extrêmes dans le sens des axes cités sont bien sûr Manchester City et Liverpool avec leur système offensif très efficace tandis qu'à l'opposé nous avons les clubs avec un faible rendement offensif sur la saison comme Cardiff ou Huddersfield

- *Axe 2 : L'efficacité offensive à l'extérieur*

Ce deuxième axe représente 11% de l'inertie totale. Un autre facteur déterminant de la réussite des équipes en Premier League est son efficacité lorsqu'elle joue en déplacement. Un exercice difficile car jouer dans le stade adverse et son public n'est pas toujours aisé. Cet axe quant à lui a pour variables les plus contributives les [buts à l'extérieur à la mi-temps](#) et les [buts à l'extérieur à la fin du temps réglementaire](#). On constate encore une fois les extrêmes Liverpool et Man City qui marquent beaucoup à l'extérieur mais à l'opposé nous avons Burnley qui au contraire a eu beaucoup du mal lors de cette saison à trouver le chemin des filets chez l'adversaire.

- *Axe 3 : L'efficacité offensive à domicile*

Ce troisième axe représente 10% de l'inertie totale. Passons à présent à l'efficacité offensive à domicile, bien évidemment l'efficacité devant les cages devant ses spectateurs est gage de la force de l'équipe. d'après ce dit axe, on constate que Fulham a eu un peu de mal dans cet exercice ce qui explique leur place par rapport à cet axe qui a pour variables les plus contributives [buts à domicile à la mi-temps](#) et les [buts à domicile à la fin du temps réglementaire](#)

- Axe 4 : Le sang-froid des joueurs

Ce quatrième axe représente 9% de l'inertie totale. Et oui, même le sang-froid des joueurs et par conséquent des équipes impactera les résultats de cette dernière. Un groupe qui a du self-contrôle et sait gérer la pression d'un match aura l'ascendant sur ses adversaires et cela se reflétera dans ses résultats. C'est ce que représente l'axe 4 avec pour variables les plus contributives, les fautes et les cartons jaunes. En effet ces actions provoquent l'arrêt du jeu et sont (en général) une forme de relâchement ou d'erreur qui peut avoir de lourdes conséquences sur l'issue de la rencontre. Ici nous avons notamment Manchester United qui dans cet aspect est assez présent et contributeur pour cet axe avec plus de 70 cartons jaunes sur la saison.

Ainsi, pour conclure, autant d'indicateurs permettent de justifier la réussite de certaines équipes mais il y a bien évidemment beaucoup d'autres variables qui entrent en considération en réalité.

Axes à afficher

Histogramme des valeurs propres

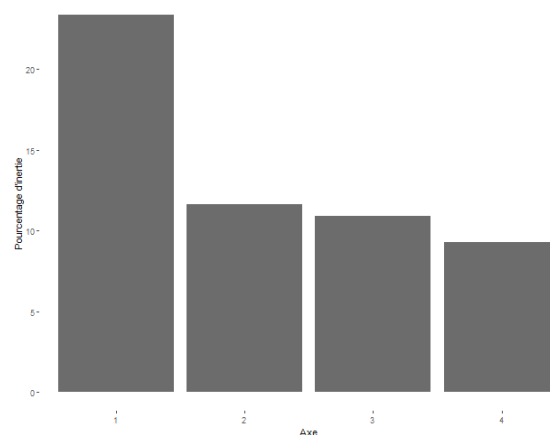


Tableau des valeurs propres

Axe	%	Cum. %
1	23.3	23.3
2	11.6	35.0
3	10.9	45.9
4	9.3	55.2

Axe X
Axis 1 (23.35%)

Axe Y
Axis 4 (9.26%)

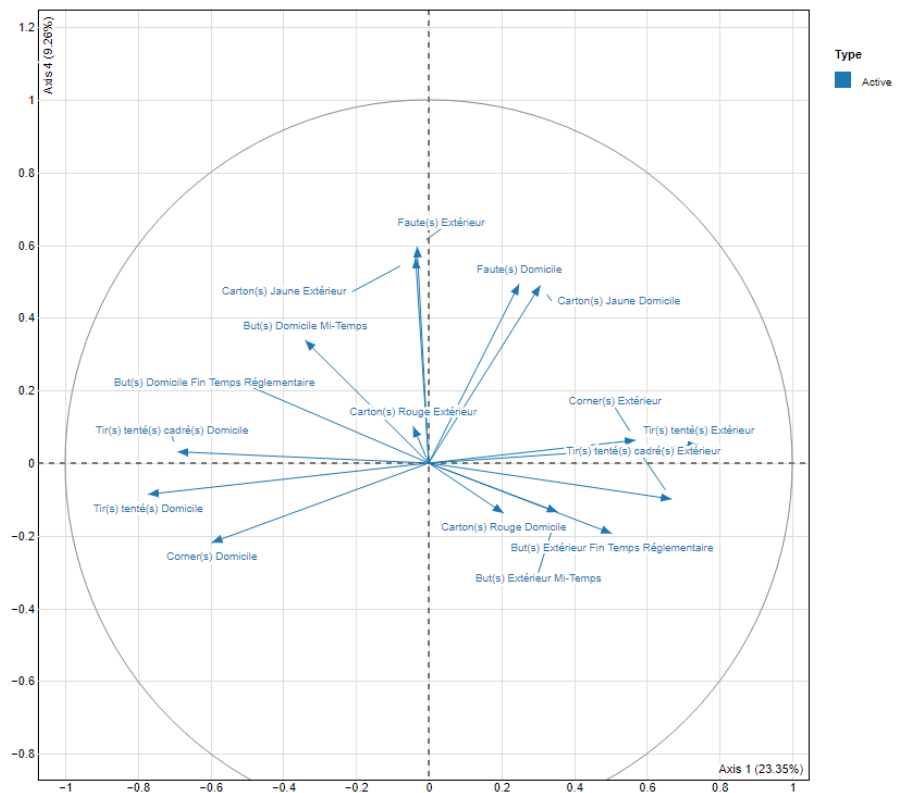
Taille des étiquettes
4 10 20

Contribution min pour afficher l'étiquette
0

Couleur :
Type de variable

☐ Variables supplémentaires

☒ Animations



Axe X
Axis 1 (23.35%)

Axe Y
Axis 2 (11.64%)

Taille des points
8 64 128

Opacité des points :
Constante

Opacité des points
0 0.5 1

☐ Afficher les étiquettes

Couleur :
None

☒ Ellipses

☒ Animations

