Une image contenant Emblème, symbole, cheval, logo

Description générée automatiquement

Application streamlit

stage été 2024

Nathan TALBOT

Clermont foot

2024

***Présentation document***

Ce document a pour but de présenter l’application Streamlit que j’ai développé lors de mon stage au Clermont Foot durant l’été 2024. Elle a pour but d’étudier des données et d’obtenir diverses informations sur des clubs de différents championnats, sur différentes saisons.

Ces données sont obtenues car à deux fournisseurs : Stats Bomb et Skill Corner.

Streamlit est une librairie Python qui permet de développer des applications/sites web, en proposant des outils permettant une interaction agréable et pratique pour l’utilisateur.

***INTRODUCTION***

Les pages sont divisées en 2 principaux thèmes :

1. L’étude des métriques.

Cette partie est divisée en 3 pages : les métriques différenciants des groupes d’équipe d’un même championnat, l’évolution des métriques pour un groupe d’équipe ou une équipe en particulier au cours d’une saison, ainsi que l’évolution des métriques au cours de plusieurs saisons, pour un groupe d’équipe.

Les données de Stats Bomb que j’ai utilisées proviennent de l’API « Team season stats », qui sont spécifiques aux équipes et agrégées sur une saison, tandis que les données de Skill Corner sont divisées en 4 catégories :

* Physiques : <https://skillcorner.crunch.help/en/glossaries/physical-data-glossary>
* Courses sans ballon : <https://skillcorner.crunch.help/en/glossaries>
* Passes : <https://skillcorner.crunch.help/en/glossaries/passing-metrics-1>
* Pressing : <https://skillcorner.crunch.help/en/glossaries/overcoming-pressure>

Ces données sont spécifiques aux joueurs et étudiées sur un match. Pour ramener ces données sur des équipe, j’ai donc, pour chaque métrique, regroupé et additionné les données des joueurs d’une même équipe sur un même match (e.g. : métrique basée sur la distance parcourue par un joueur durant un match : la métrique pour l’équipe est donc la somme de toutes les distances parcourues par les joueurs d’une équipe durant un match).

Nb : Pour les métriques de type « ratio » (pourcentage de réussite d’une action durant un match), j’ai calculé la moyenne pour tous les joueurs d’une même équipe sur un match, et non additionné les pourcentages de chaque joueur.

Pour chaque catégorie de métrique, il est possible d’étudier les métriques en faisant une moyenne sur une durée, une période avec ou sans possession du ballon, ou encore un nombre d’évènement. Par exemple, pour les métriques de passes, il est possible de faire une moyenne d’une métrique par 100 opportunités de passe, par 30 minutes de jeu avec la possession du ballon, ainsi que par match.

1. Les Heatmaps.
2. Bonus : Étude du nombre de passes sur les actions menant à un but.

***Métriques discriminantes***

Cette première page a pour but d’étudier les métriques différenciant le plus des groupes d’équipes d’un même championnat.

Les données étudiées sont les données de différentes métriques en moyenne par match.

Un groupe d’équipe correspond peut correspondre à un Top (ex : Top 5 de Ligue 2), au Middle, ou au Bottom.

Pour étudier la différence entre 2 groupes d’équipes pour une métrique, j’ai calculé la différence en % entre la moyenne des 2 groupes pour cette métrique.

L’application permet de :

* Sélectionner les saisons sur lesquelles étudier les métriques.
* Choisir la taille des différents groupes.
* Filtrer les métriques.
* Choisir le nombre final de métriques gardées.
* Afficher un tableau avec diverses informations sur les métriques gardées pour les différents groupes (Moyenne pour un groupe, différence entre 2 groupes, écart-type d’un groupe…)
* Sélectionner une métrique dans ce tableau et afficher les valeurs pour chaque équipe sur les saisons sélectionnées.