

Le TP sera ramassé et noté. L'objectif est de fournir le code nécessaire pour réaliser les diagrammes (et par conséquent, comment obtenir les données) et les diagrammes. Il y a deux parties distinctes.

Partie 1 : des diagrammes en pagaille

Le plus simple est de travailler sur un Jupyter Notebook.

L'ensemble des datasets se trouve ici :

https://figshare.com/collections/Soccer_match_event_dataset/4415000

Il y a 12 différents datasets mais tous ne vont pas nous servir.

Pour le dataset *coaches.json*, faites un diagramme donnant la répartition entre les entraîneurs encore en activité et ceux qui ne le sont plus. Fournissez aussi un diagramme donnant la répartition par pays. Demandez-vous s'il s'agit d'être exhaustif dans la visualisation ou si les pays les plus représentés sont les seuls donnés.

Pour le dataset *teams.json*, il faut ressortir l'ensemble des équipes pour la Premier League. A partir de là, recherchez l'ensemble des villes qui est concerné par une équipe de Premier League ainsi que les coordonnées GPS afin de pouvoir les placer par la suite sur une carte. Peut-être choisir plutôt les coordonnées du stade quand il y a plusieurs équipes dans la même ville. La réalisation de la carte se fera à l'aide de BaseMap. Cf. les exemples suivants : <https://python-graph-gallery.com/310-basic-map-with-markers/> et <https://peak5390.wordpress.com/2012/12/08/matplotlib-basemap-tutorial-plotting-points-on-a-simple-map/>

Pour le dataset *matches.json*, nous allons nous intéresser au Big 6 (Liverpool, Manchester City, Manchester United, Chelsea, Arsenal et Tottenham) et l'évolution du nombre de points par journée sur toute la saison. Il s'agit donc de créer un diagramme qui montre l'évolution pour ces 6 clubs. Le choix du diagramme semble évident mais il faut maintenant le mettre en œuvre.

Pour le dataset *events.json*, comme vous allez vous en apercevoir, ce dataset donne les événements durant un match. Il existe deux niveaux de finesse : les *event data* que nous avons ici et les *tracking data*, qui donne la position de chaque joueur sur le terrain durant tout le match. L'objectif est de mieux connaître les équipes, mais nous nous concentrons sur les positions du buteur lors d'un but (histoire de voir si l'histoire se répète). Choisissez une équipe, plutôt une prolifique en buts, et récupérez l'ensemble des buts sur tous les matchs. Il faudra ensuite les placer sur une représentation du terrain de foot. Nous disposons pour cela de *mplsoccer* (<https://mplsoccer.readthedocs.io/en/latest/index.html>) et des nombreux exemples.

Partie 2 : Analyse de l'équipe adverse

Supposons que nous en sommes à la dixième journée du championnat et l'équipe X (celle que vous aidez) rencontre l'équipe Y. Il est nécessaire de comprendre l'équipe adverse pour la battre. Pour cela, vous allez récupérer les *events* pour tous les matchs précédents de cette équipe Y et faire apparaître deux informations : le système de jeu (comment le ballon remonte du gardien vers les attaquants) et comment les buteurs marquent. Vous pourrez fournir autant de diagrammes que vous voulez (voire une petite vidéo réalisée avec `matplotlib.ArtistAnimation`).