

5PBI + 5BDA Rattrapage

➤ Module	5BDA
⌵ Type	Projet

Analyse des performances des joueurs de tennis.

Introduction

Depuis 1968, l'ère Open tennis a vu des performances spectaculaires de nombreux athlètes de la petite balle verte au travers du monde. L'AFP lance un challenge et fait appel aux professionnels de la data afin d'analyser les performances des joueurs de tennis sur base des données collectées depuis 1968 avec pour principal objectif de ressortir le classement des 100 meilleurs joueurs de tennis de tous les temps.

Subject

Afin de réaliser au mieux cette tâche, vous avez les données de ces 15 dernières années, que vous analyserez, gardant à l'esprit l'objectif principal.

Dataset: tennis dataset

Collecte des données (5BDA - 4pts)

A partir du dataset fourni, écrivez des scripts python de cleaning, de preprocessing, et de sauvegarde.

N.B: Vous pouvez tout à fait créer un fichier unique à partir des data.

Analyse des performances (5BDA - 8pts)

Créer des métriques spécifiques permettant d'analyser les performance des joueurs à partir des données .

N.B: Faites bien attention à ce que ces métrique vous permettent de répondre à toutes les questions suivantes.

Lister les meilleurs de tous les temps (5BDA - 8pts)

A partir des métriques précédentes:

- lister les 100 meilleurs joueurs de tous les temps,
- lister les 10 meilleurs joueurs pour les nationalités (Serbie, France, Canada)
- Lister les 10 meilleurs joueurs par tranche d'âge ([18-24], [25-30], [31-35],]35, +[)

Data Viz (5PBI - 20pts)

A partir des datasets choisies (15 années), créer moi un rapport multipage ayant:

N.B: le dataset utilisé pourra être préprocessé à l'aide de script python avant leur import sur POWERBI.

N.B: Pour le rendu, vous devez exporter votre travail final au format pdf afin de l'importer sur exams.

- Une page qui liste les meilleur joueur par tranche d'âge : [18-24], [25-30], [31-35],]35, +[⇒ 3 pts
- Une page représentant les nationalités les plus représentées dans le tennis (top 25), ⇒ 3 pts
- Une page représentant confrontations les plus longues (top 10), ⇒ 3 pts
- Une page reprenant les statistiques des joueurs par années (NB de compet, NB de victoire, NB d'echec, Total de points, Meilleur rang, ...), ⇒ 3 pts
- Une page reprenant les statistiques joueurs par origine (NB de compet, NB de victoire, NB d'echec, Total de points, Meilleur rang, ...) ⇒ 4pts

- Une page représentant les evolutions de points par année par nationalité (graphique). ⇒ 4pts

Rules

This project must be your creation and done by your own. Plagiarism or copy / past are forbidden and will result of a 0.

This is project can be done in group of 2 students maximum.

Each group containing more than 2 members will be discarded and not evaluated.

Delivery

You must deliver a your code source files or notebooks in a git (github, gitlab) containing:

- Technical Documentation (.pdf).
- The technical document must be uploaded in exams.

⚠ File extension should be strictly respected. If you deliver file in a different format than expected, the file will be entirely discarded.