

1.1

Utiliser la fonction pandas loc sur le dataframe players.

Utiliser sample pour afficher des exemples.

1.2

Définir une fonction renvoyant le type de poste, puis appliquer la méthode `apply` à la colonne "Position" du dataframe.

1.2b

Utiliser les fonctions Pandas groupby et idxmax.

2.1

Il faut filtrer en utilisant le tableau `skills_ratings` puis utiliser la fonction `fill_na`

2.2

Initialiser le modèle TSNE du package manifold de sklearn/
Utiliser fit et Transform sur ce modèle

2.3

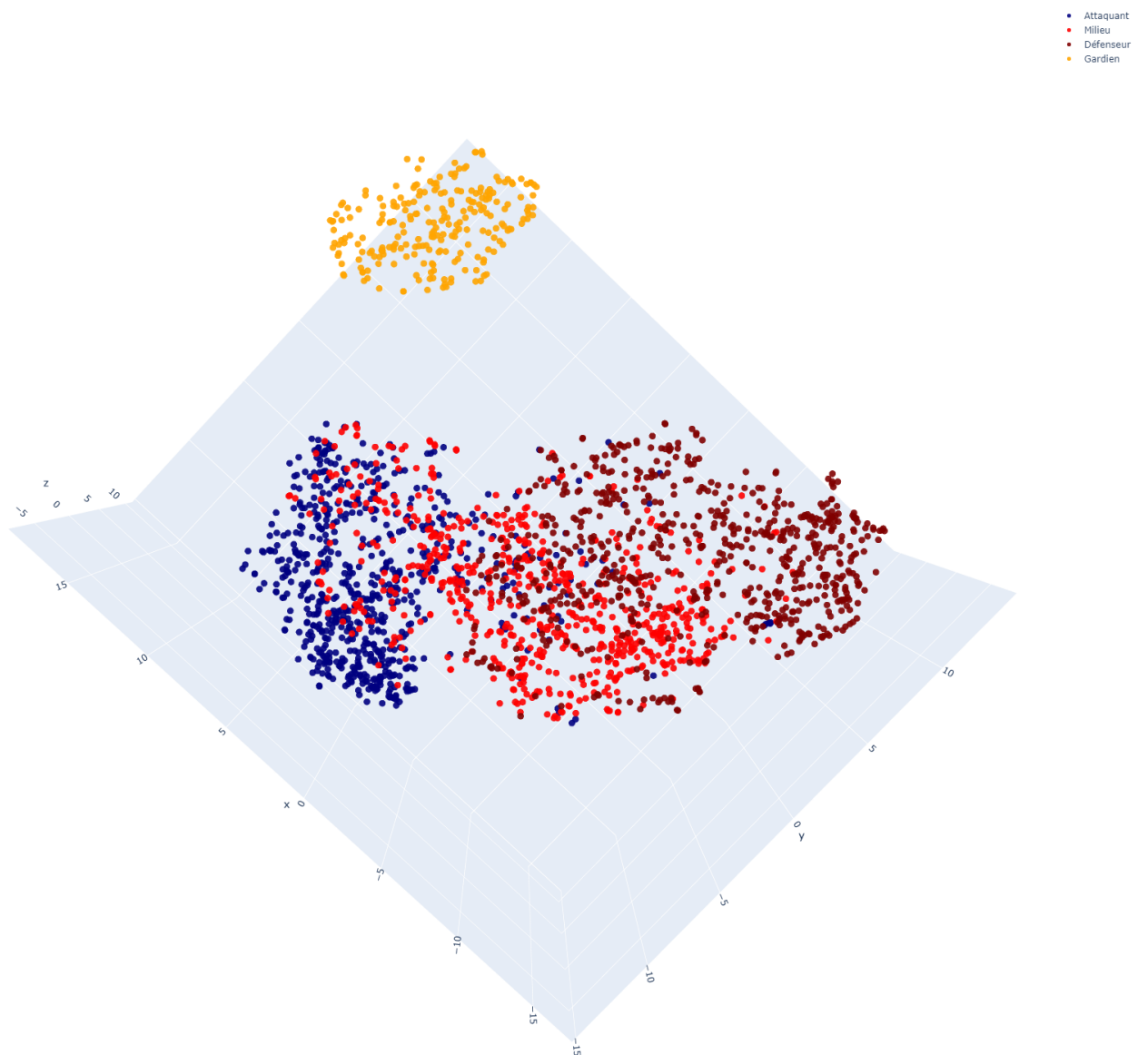
On peut faire un test '==' sur la colonne "Type_poste".

2.4

méthode Scatter de l'objet plotly.graph, en mode "markers"

Résultat :

t-SNE - Fifa 2019



2.5

méthode Scatter de l'objet plotly.graph, en mode "text"

3.1

Créer un dictionnaire avec les valeurs du joueurs parfait

Utiliser la fonction `append` (avec `ignore_index` à `true`) pour ajouter ce joueu au `dataFrame`

utiliser `fillna` pour compléter les valeurs manquantes à 50/

3.2

Utiliser KMeans de Scikit Learn, avec un nombre de cluster de 5

3.3

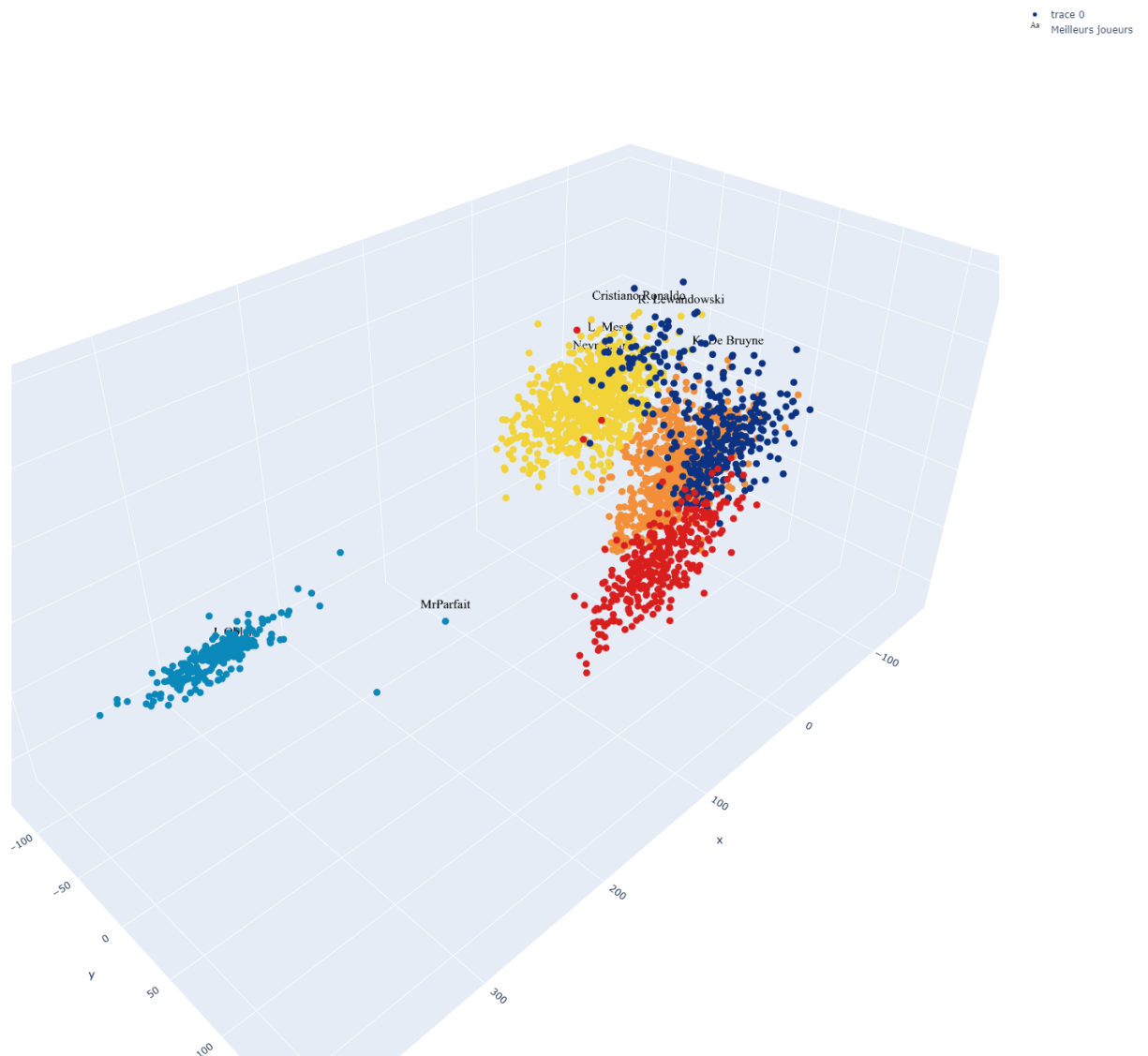
Utiliser décomposition de de Scikit Learn, avec un nombre de composants de 2

3.4

Pour colorer les points, récupérer le cluster de chaque joueur grâce à la propriété "labels_" de l'objet kmean.

Résultat :

ACP + Clustering



3.5

A vous de définir votre approche :-)

Vous pouvez par exemple utiliser la colonne “club” pour comparer les équipes en présence.