

# TP 4 Transformation et Visualisation

*MI - IA / GL*

Ilyas Bambrik

# Table des matières



|  |          |
|--|----------|
| <b>I - Exercice : Transformation des données</b> | <b>3</b> |
| <b>II - Exercice : Visualisation des données</b> | <b>4</b> |

# Exercice : Transformation des données



Lisez le Dataset "*players.csv*" présenté dans les cours et TPs précédants.

## Question 1

Transformez la colonne *current\_club\_id* en variable catégorique et affichez la description statistique approprié.

## Question 2

Affichez la description statistique la colonne *height\_in\_cm* et trouvez l'anomalie dans celle-ci.

## Question 3

- Filtrez les lignes de votre Dataframe pour retrouvez seulement les joueurs avec un contrat qui expire en 2024 ou plus tard. Placez le résultat dans un nouveau Dataframe.
- Pour le résultat précédant, ajoutez une nouvelle colonne '*year\_remaining*' qui prend le nombre d'années restants dans le contrat du joueur.

## Question 4

- Créez une colonne *age* contenant l'age du joueur (nombre d'années depuis la date de naissance).
- Créez une nouvelle colonne '*age\_bracket*' correspondant à l'une des tranches d'age suivantes : (12,17] , (17,20], (20,24], (24,28] ,(28,35] et (35,60].
- Quelle est la tranche d'age la plus fréquente dans le Dataset.
- Quelle est la valeur marchande (*market\_value\_in\_eur*) maximal pour chaque tranche d'age.
- Créez une nouvelle colonne '*market\_value\_bracket*' qui divise d'une manière approximativement uniforme les joueurs sur six catégories selon la colonne *market\_value\_in\_eur*. Affichez le résultat trié selon les intervalles.

# Exercice : Visualisation des données



## Question 1

Tracez *market\_value\_in\_eur* en fonction de *year\_remaining*. Que remarquez vous (la relation entre le deux variables tracées)?

## Question 2

- Tracez *market\_value\_in\_eur* et *highest\_market\_value\_in\_eur* en fonction de la colonne *age* dans la même figure.
- Utilisez *FacetGrid* pour tracez *market\_value\_in\_eur* en fonction de la colonne *age* pour chaque tranche d'age définie dans *age\_bracket*.