Liste des variables

Name(Nominale) <- pas besoin de l’inclure dans les graphs de data viz

Quantitatives

Age(Discrete)

Overall(Discrete)

Potential(Discrete)

Value(Continue)

Wage(Continue)

Skill Moves(Discrete)

Height(Continue)

Weight(Continue)

Release Clause(Continue)

Qualitatives

Position(Nominale)

Nationality(Nominale)

Club(Nominal)

Preferred Foot(Nominale)

Contract Valid Until(Nominale)

Liste Graphs : <https://datavizcatalogue.com/>

Analyse Univariée

|  |  |  |  |
| --- | --- | --- | --- |
| Type de variable | Quantitative Discrète | Quantitative  Continue | Qualitative |
| Graphiques utilisables | Barplot / Histogram  Boxplot / Violinplot | Densityplot  Boxplot / Violinplot | Pie Chart  Donut Chart |

Nationalité : Choropleth Map

Analyse Bivariée

|  |  |  |
| --- | --- | --- |
| Type de variable | Quantitative | Qualitative |
| Quantitative | Scatter Plot | Bubble Chart |
| Qualitative | Bubble Chart | Heat Map |

Cas particuliers :

* Overall vs Potentiel : Ordered candlechart
* Position ou Nationalité impliqué : Code couleur particulier ( Bubble Graph ? )

Couples intéressants :

* Age vs Potentiel
* Height vs Overall
* Height vs Position
* Overall vs Wage
* Position vs Wage
* Nationality vs Wage
* Preferred Foot vs Skill Set
* Nationality vs Skill Set
* Value vs Contract Valid Until
* Age vs Wage

Prédiction

Variable à prédire : Wage ( Régression )

Prédicteurs :

* Age
* Overall
* Potential
* Value
* Height
* Weight
* Skill Moves
* Position
* Nationality
* Preferred Foot

Modèles sélectionnés :

* Régression Linéaire Multiple
* SVR
* GAM
* Arbre CART
* Random Forest

Graphes affichés :

Régression Linéaire : T-value des variables

GAM : Effets estimés

Arbre CART : Arbre

Résidus ( Line Graph pour chaque modèle ) +

Nuage de points observés / estimés ( Avec Noms )

RMSE / MAE / Pourcentage de Variance Expliquée / R2