

# Présentation Projet BigData

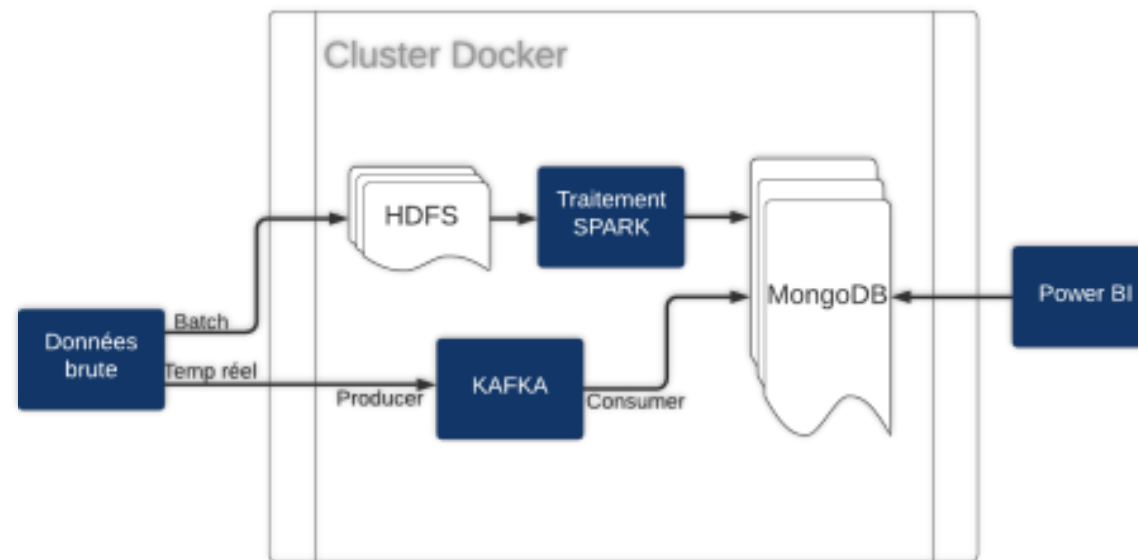
# Executive Summary

- Problématiques
- Etat de L'art
- I. Sources de données
- II. Technologies Utilisés
- III. Traitement des données
- IV. Visualisation des données avec Power BI
- Conclusion

# Problématiques

- Le club est-il un facteur important pour qu'un joueur puisse gagner un ballon d'or ?
- Le nombre de buts marqué par un joueur augmente t-il ses chances de gagner un ballon d'or ?

# Etat de L'art



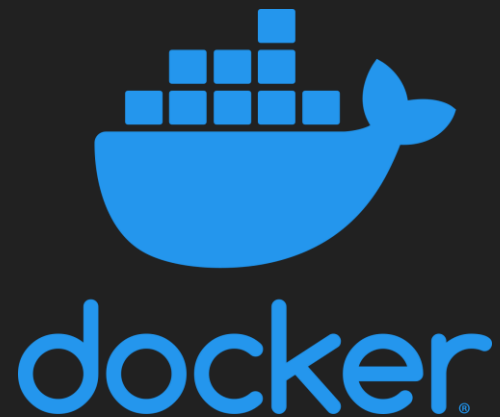
# I. Sources de données

- Site Web : Data World
- Champions\_league.csv et BallonDOr.csv
- <https://data.world/johayes13/ballon-dor-voting-history>
- <https://data.world/sportsvizsunday/sports-viz-sundays-2018/workspace/file?filename=champions+league.xlsx>

## II. Technologies Utilisés



PySpark

The PySpark logo features the word "Py" in orange and "Spark" in black, with an orange star icon above the "k".

docker

The Docker logo consists of a blue whale icon carrying a stack of blue containers on its back, with the word "docker" in blue lowercase letters below it.

Power BI

The Power BI logo shows three vertical bars of increasing height in shades of yellow and orange, with the text "Power BI" in black below them.

jupyter

The Jupyter logo features an orange circular icon with a white crescent shape inside, and the word "jupyter" in grey lowercase letters below it.

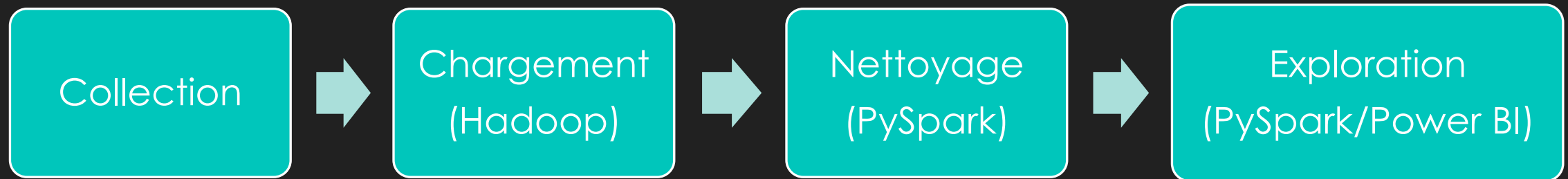
phpMyAdmin

The phpMyAdmin logo includes a stylized sailboat icon on the water, with the text "php" in blue and "MyAdmin" in orange below it.

GitHub

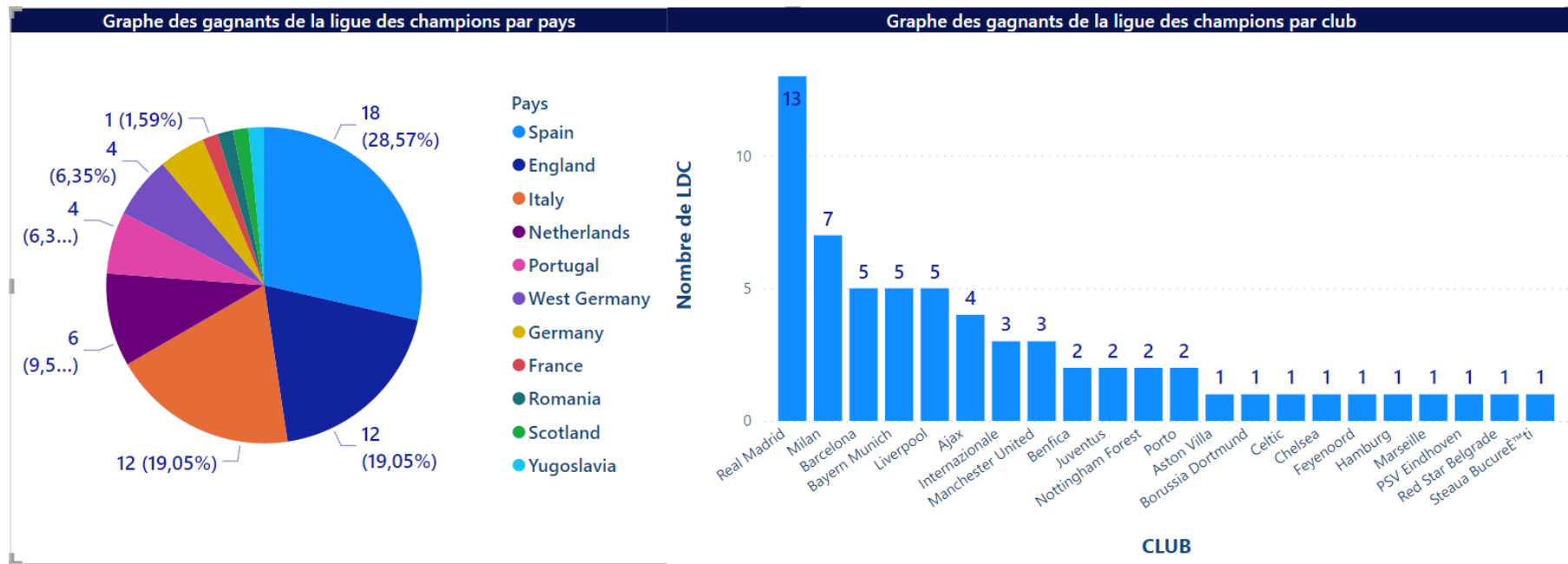
The GitHub logo features a black octocat icon inside a black circle, with the word "GitHub" in black bold letters below it.

# III. Traitement des données



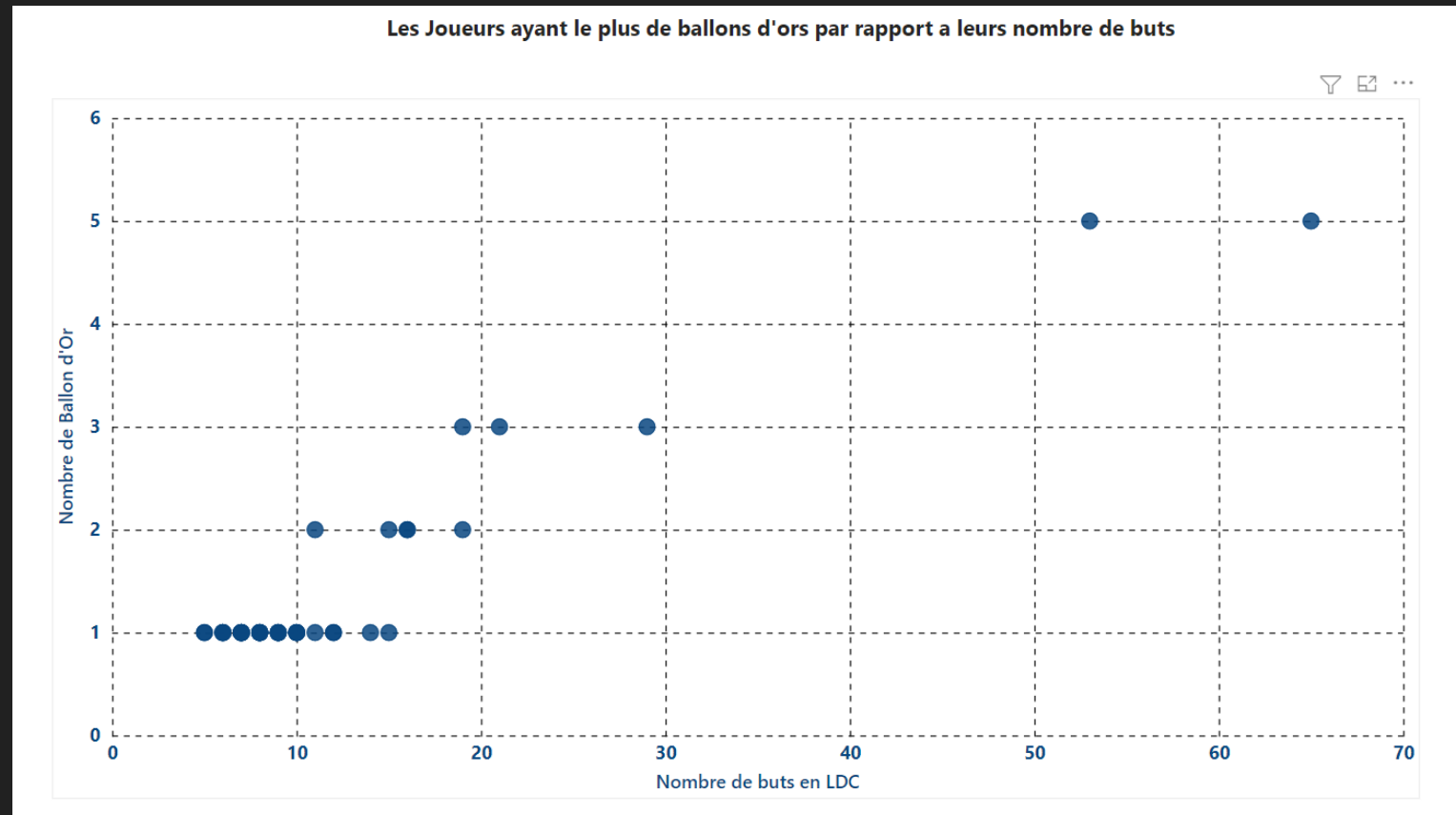
# IV. Visualisation des données avec Power BI

Les pays et les clubs où nous avons le plus de chance de gagner une ligue des champions





# IV. Visualisation des données avec Power BI



# Conclusion

- Difficultés
- Axes d'amélioration