

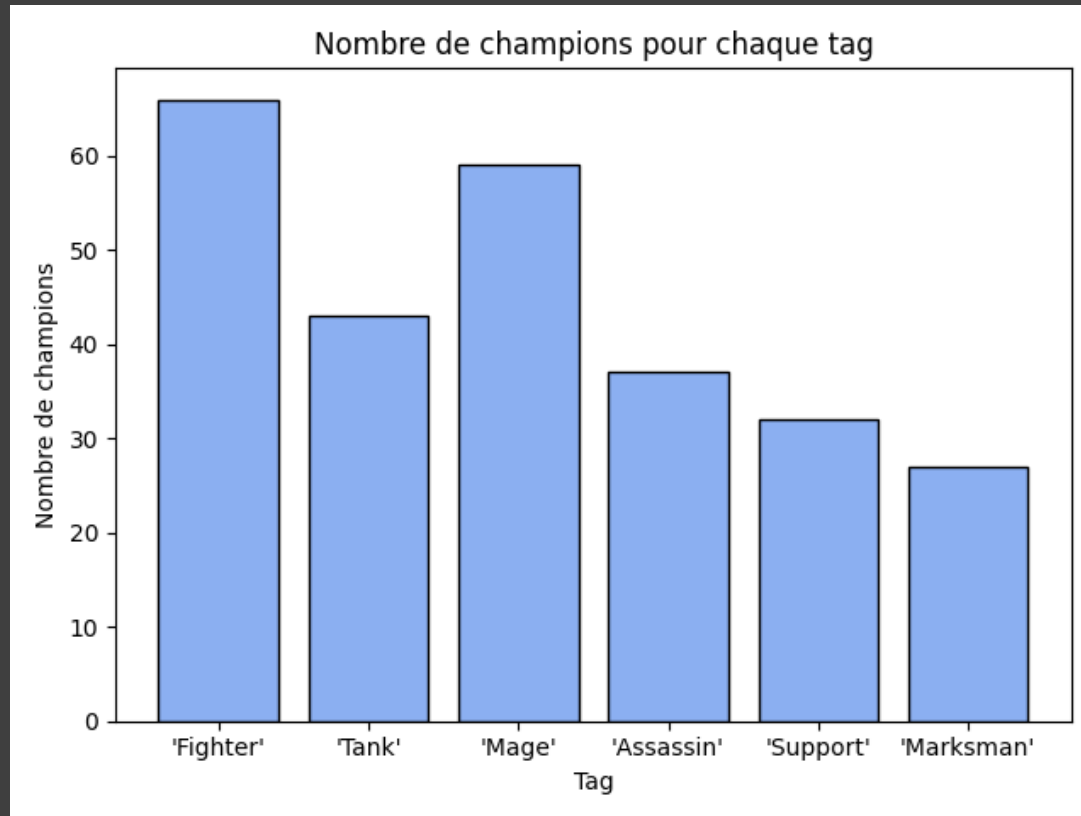
# Projet de Machine Learning

Rita Abou Samra, Léa Bucher, Tifenn Primet et Martin Rialhe-Badet

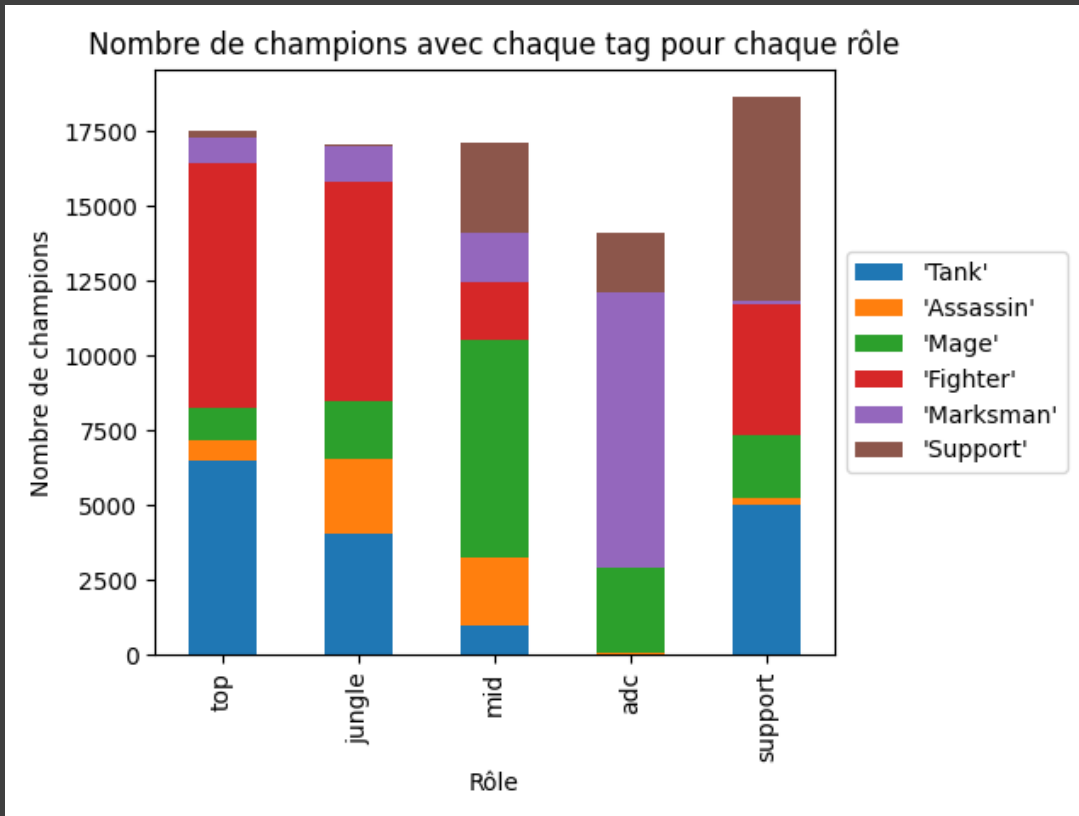
# Présentation du sujet

- ▶ Prédire le vainqueur d'une partie de League of Legends
- ▶ Composition de l'équipe :
  - ▶ 'top'
  - ▶ 'jungle'
  - ▶ 'mid'
  - ▶ 'adc'
  - ▶ 'support'
- ▶ Caractéristiques des personnages

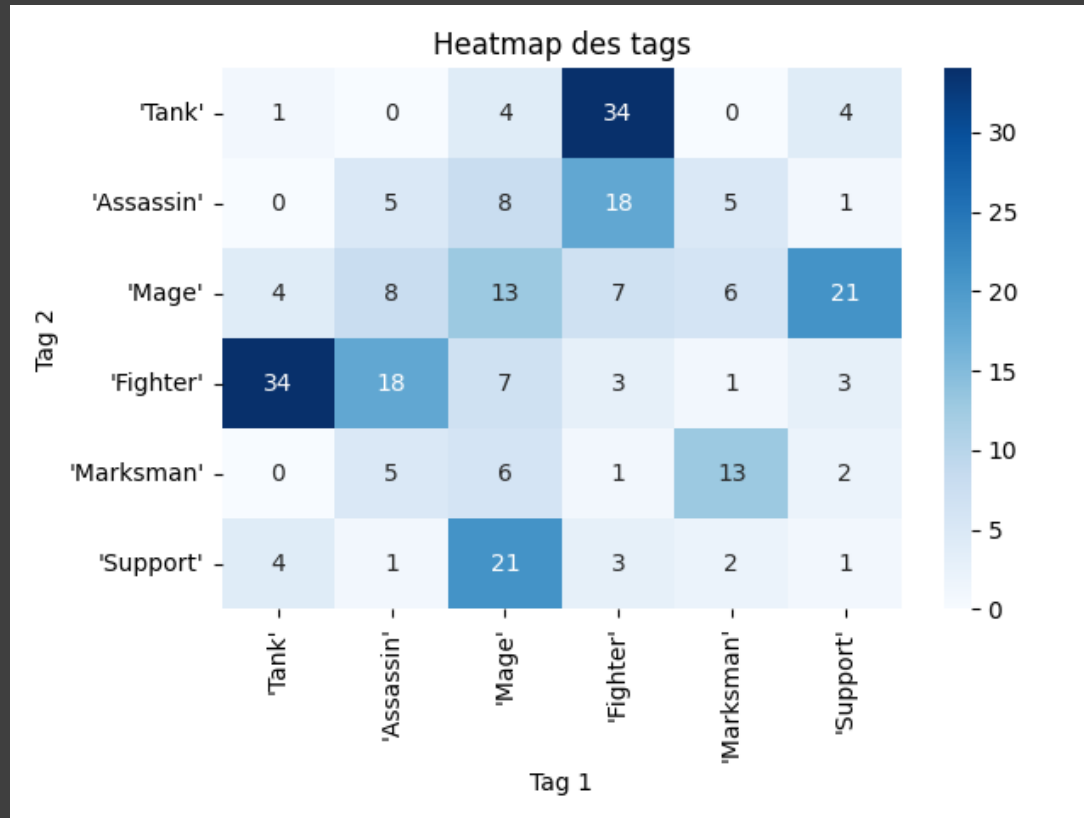
# Exploration des données



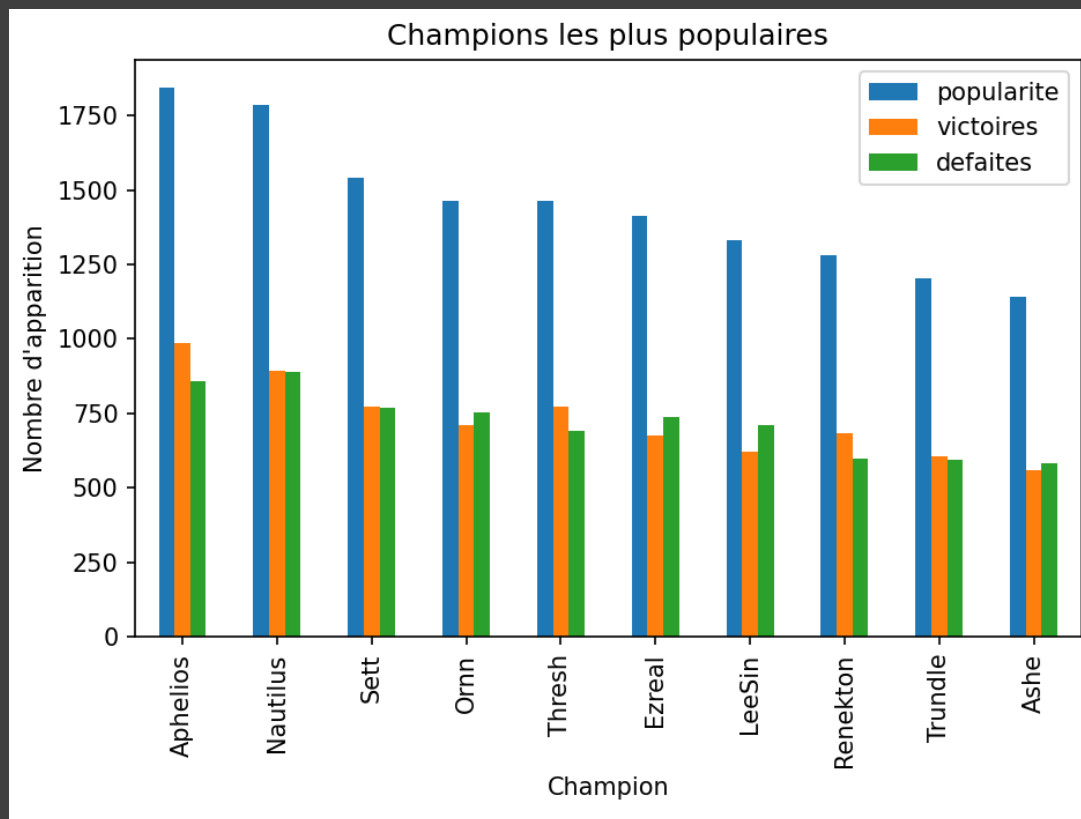
- ▶ Il existe 6 tags
- ▶ Un champion a 1 ou 2 tags



- ▶ Répartition des tags par rôle n'est pas uniforme
- ▶ Première idée pour créer un arbre



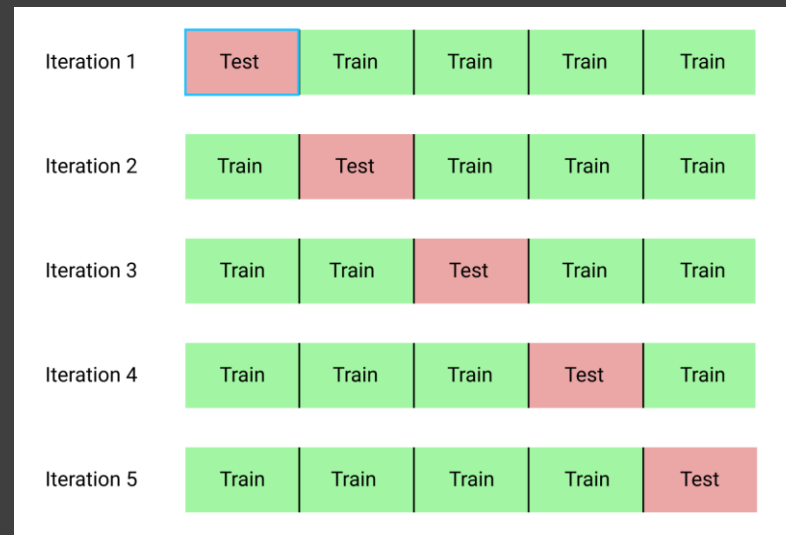
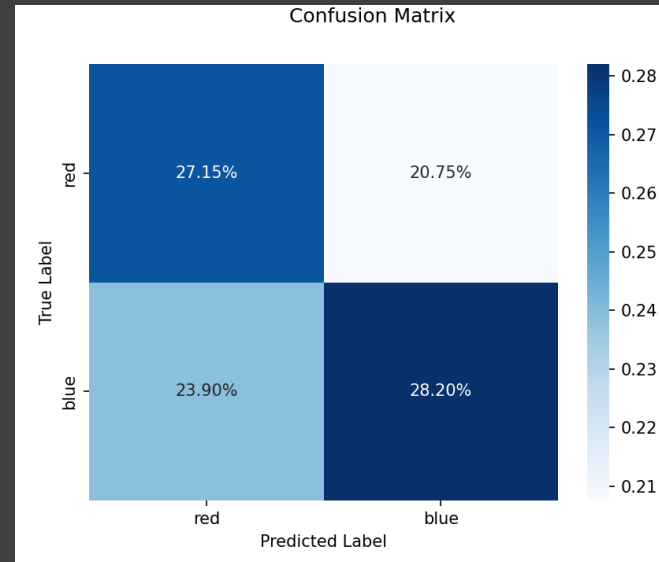
► Certains tags sont corrélés



- ▶ Personnages les plus populaires
- ▶ Taux de victoires
- ▶ Taux de défaites

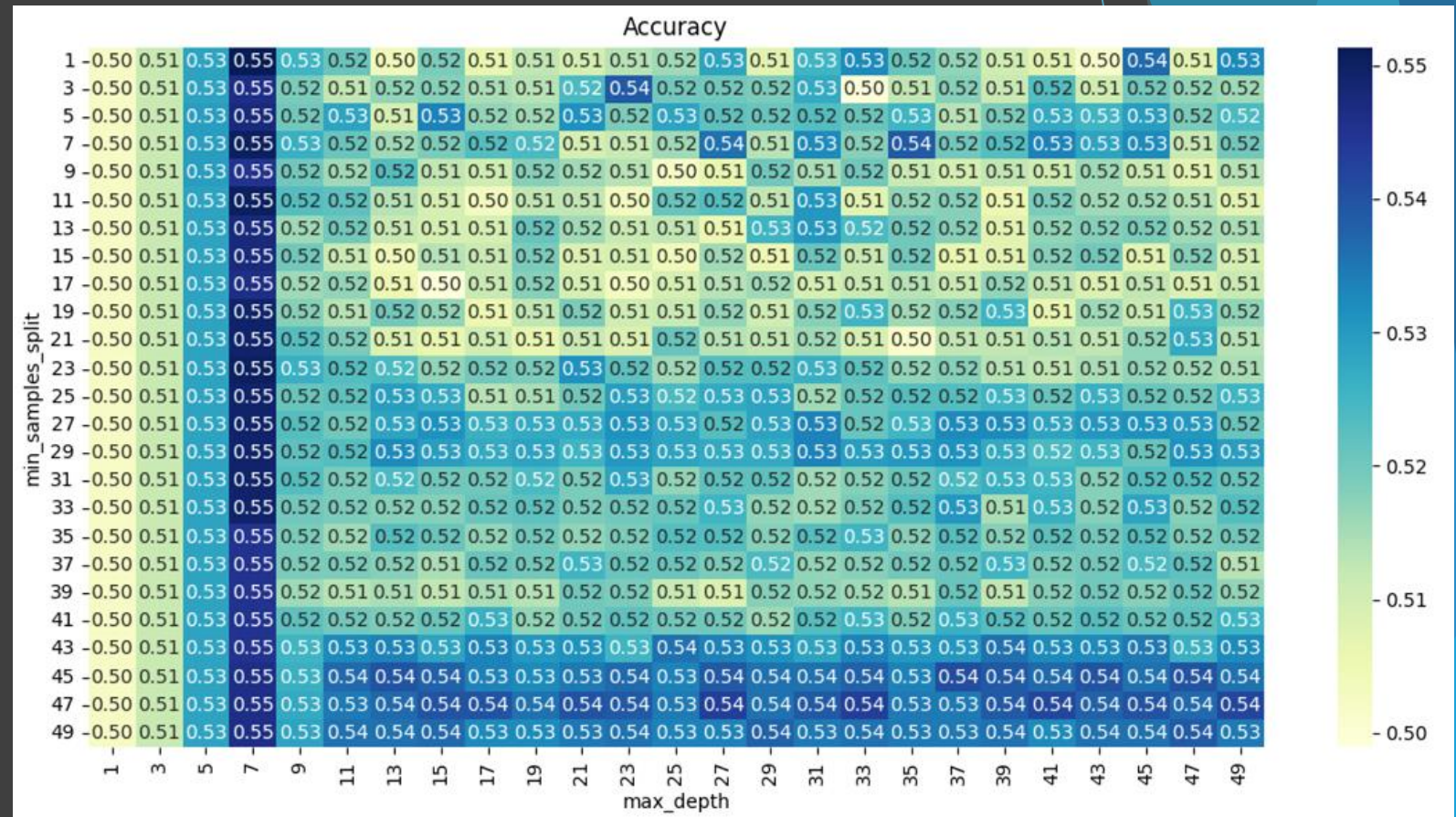
# Outils de validation

- ▶ Score de précision
- ▶ Matrice de confusion
- ▶ Validation croisée



# Trouver les paramètres optimaux

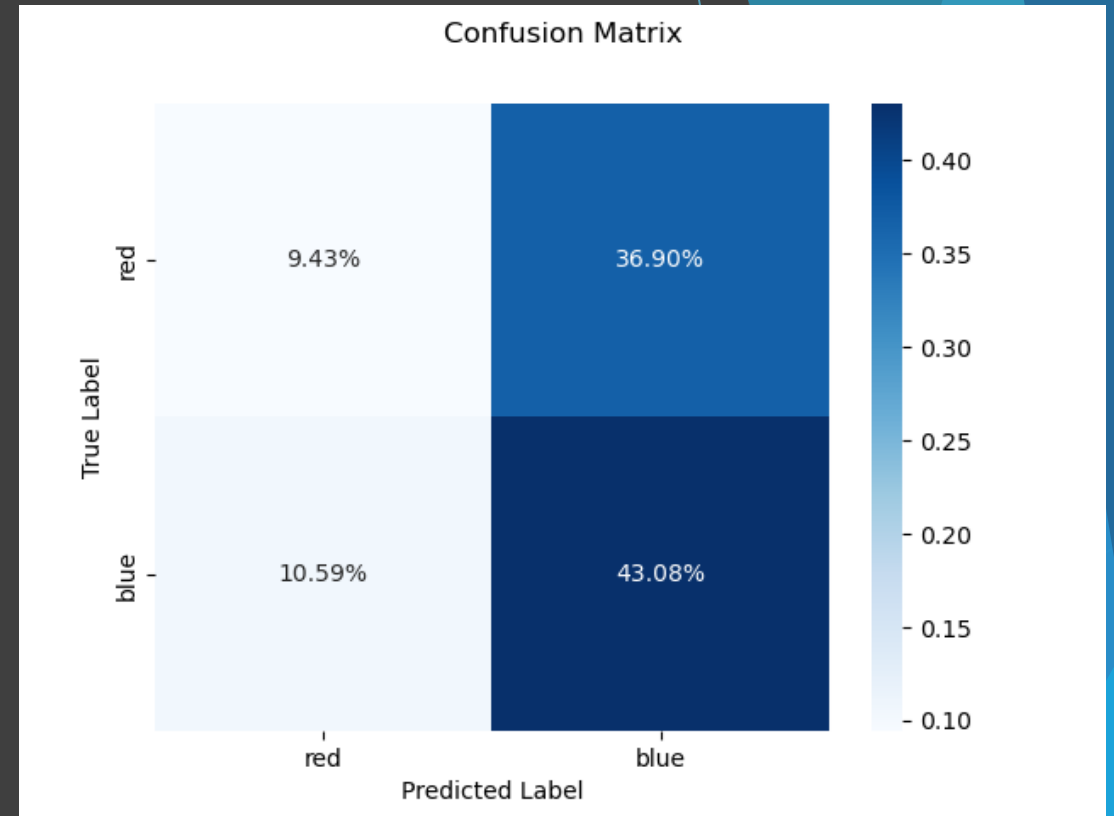
- Heatmap avec le nombre d'échantillon minimal par feuille et la profondeur maximale de l'arbre



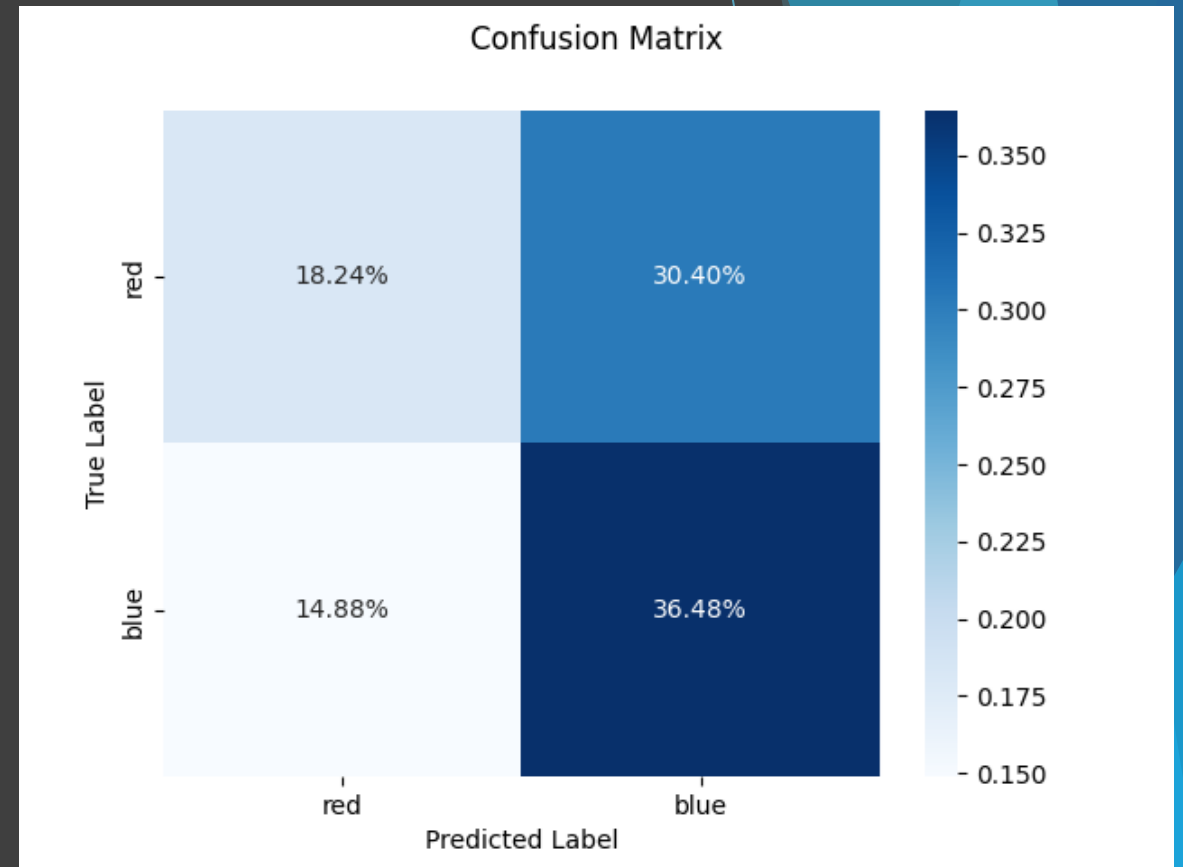


# Premiers arbres et résultats

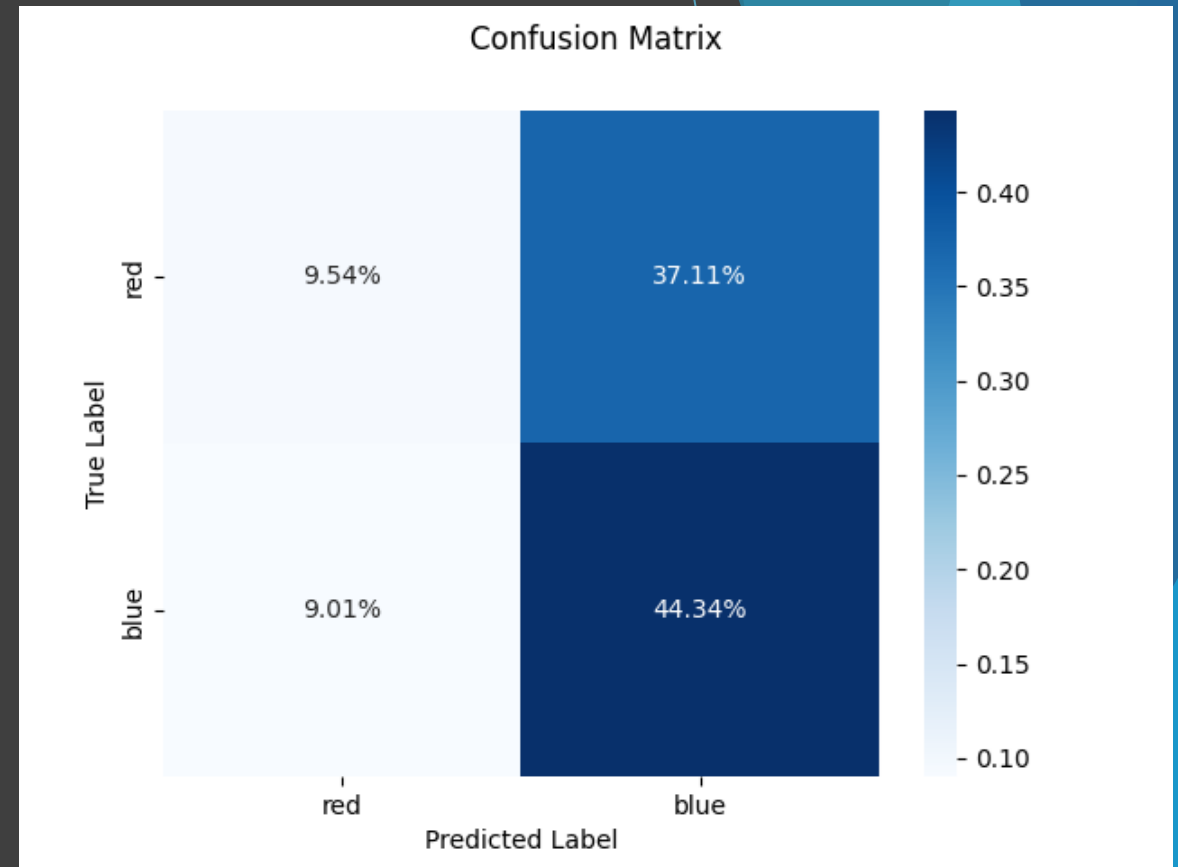
- ▶ On regarde le ou les tags les plus représentés par rôle
- ▶ La moyenne des scores est de 48.0922 %
- ▶ L'écart type des scores est de 0.0325
- ▶ La précision est de 52.5157 %
- ▶ Prédit trop de victoire pour l'équipe bleue



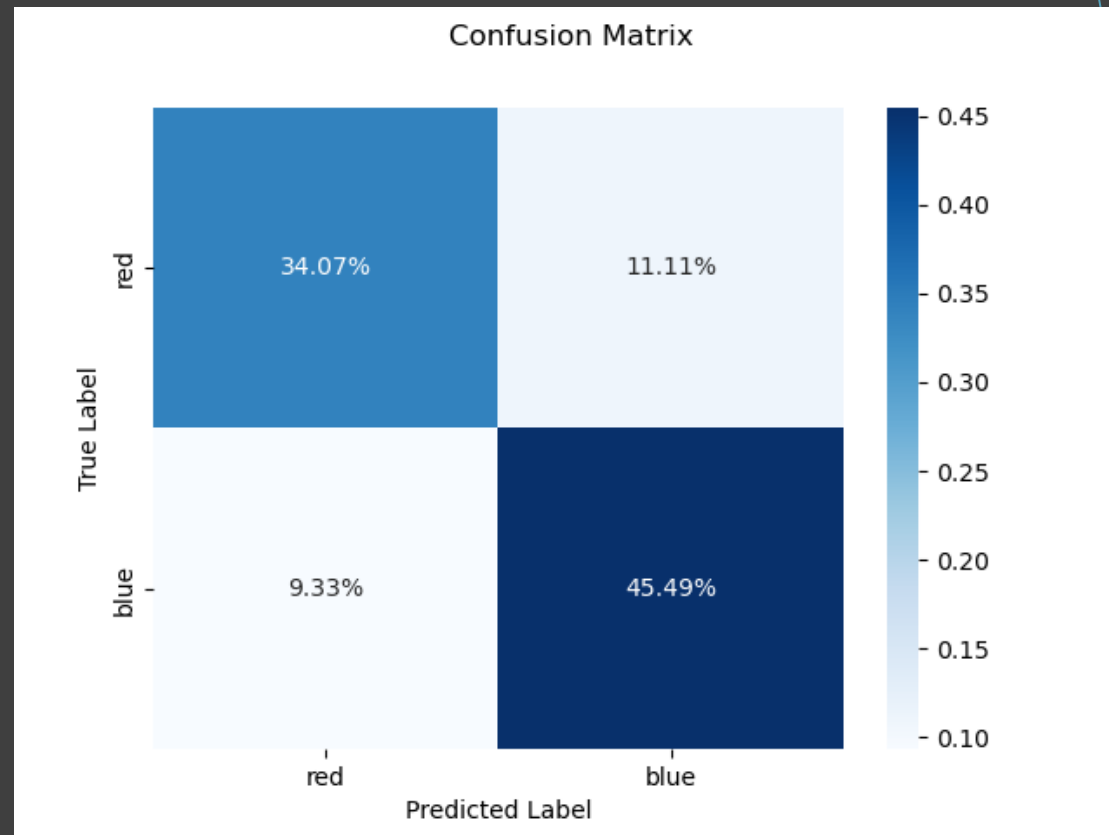
- ▶ On fait la différence des stats entre les deux équipes
- ▶ La moyenne des scores est de 51.09%
- ▶ L'écart type des scores est de 0.0172
- ▶ La précision est de 54.72%
- ▶ Prédit trop de victoire pour l'équipe bleue

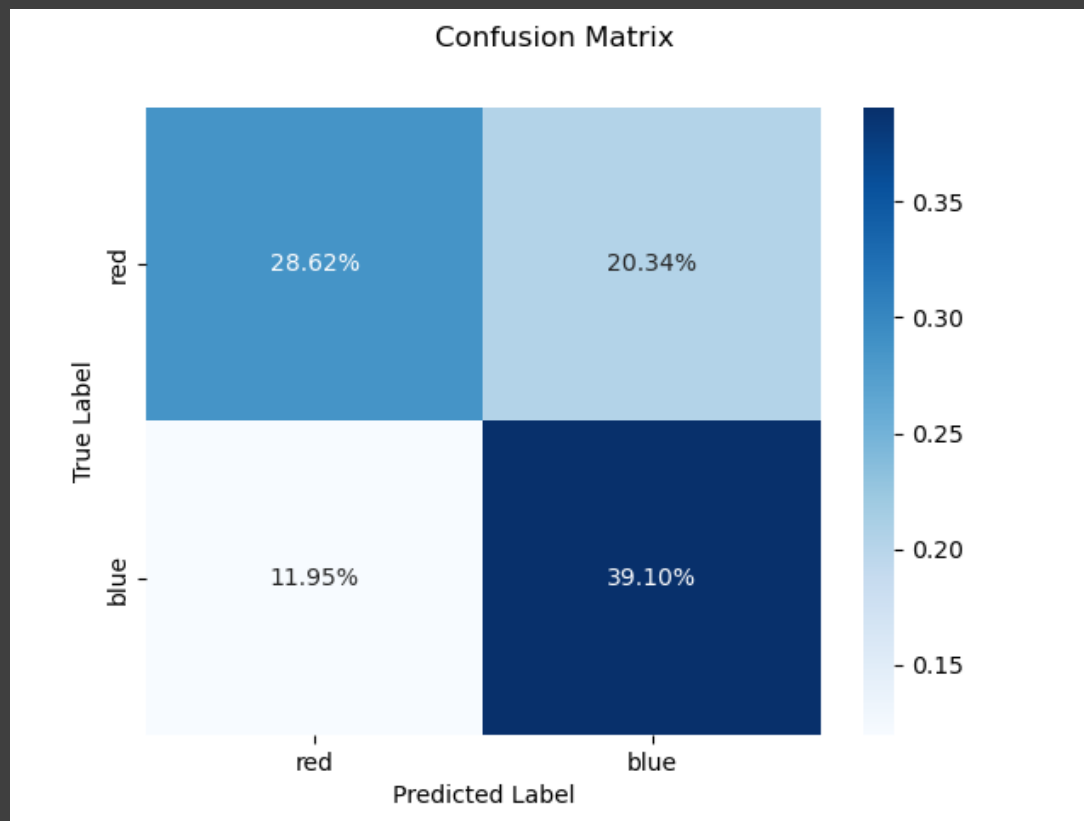


- ▶ On fait le ratio des stats entre les deux équipes
- ▶ La moyenne des scores est de 51.25%
- ▶ L'écart type des scores est de 0.0170
- ▶ La précision est de 53.88%
- ▶ Prédit trop de victoires pour l'équipe bleue

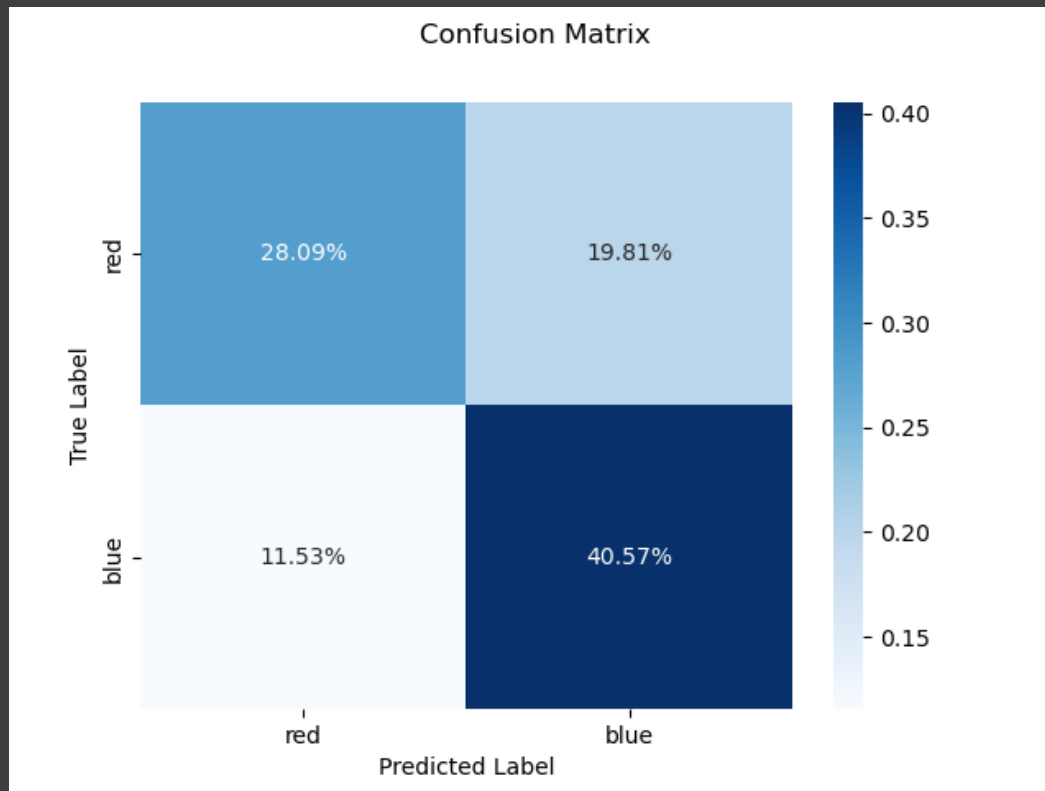


- ▶ On regarde certaines stats par champions en s'aidant des connaissances d'un joueur de LOL
- ▶ La moyenne des scores est de 51.0901 %
- ▶ L'écart type des scores est de 0.0293 %
- ▶ La précision est de 53.2495 %





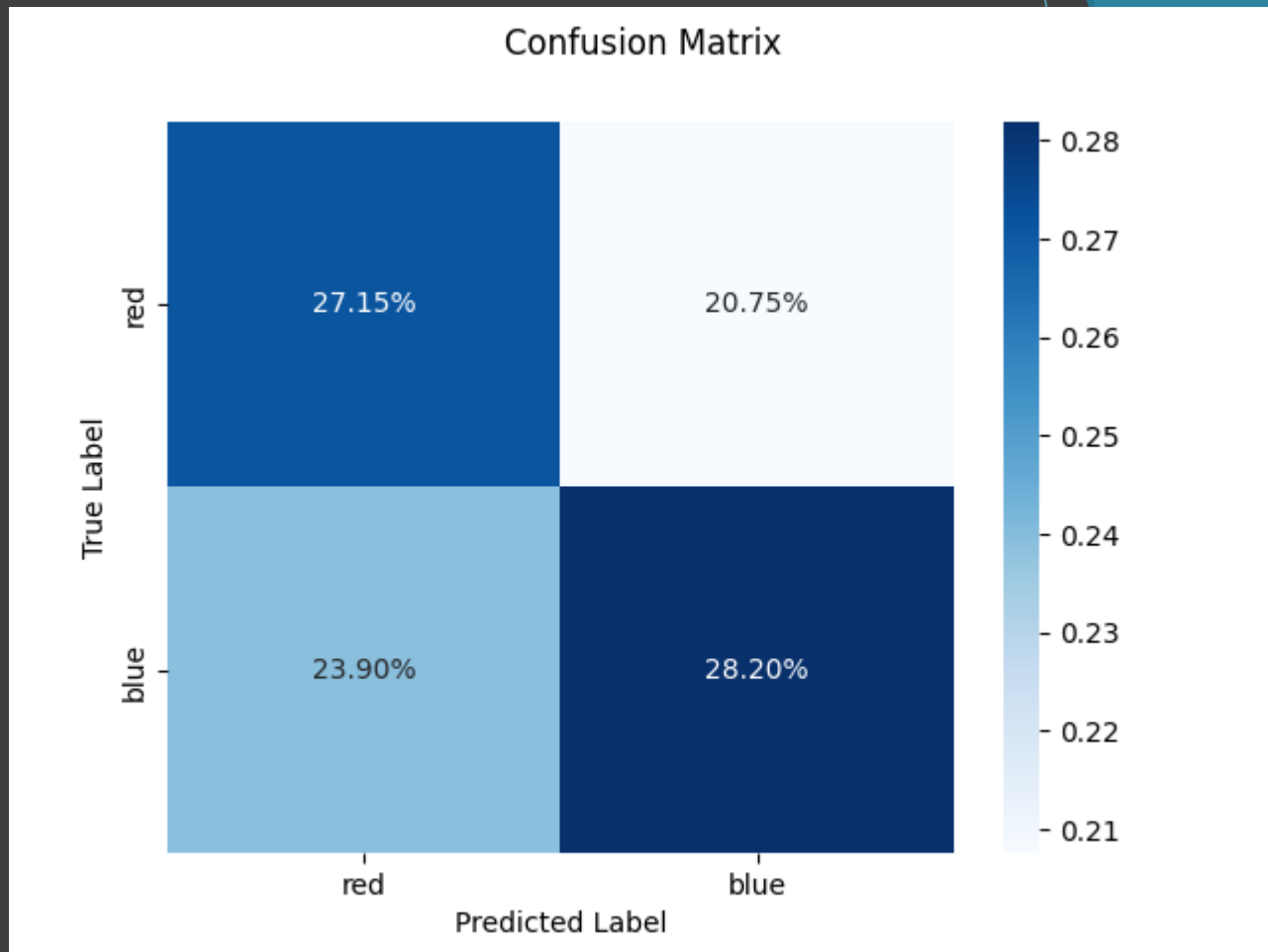
- ▶ On regarde le ratio nb victoire/ nb de parties jouées par équipe
- ▶ Deux manières de considérer cette information
- ▶ La moyenne des scores est de 0.5849
- ▶ L'écart type des scores est de 0.0277
- ▶ Précision = 0.6771



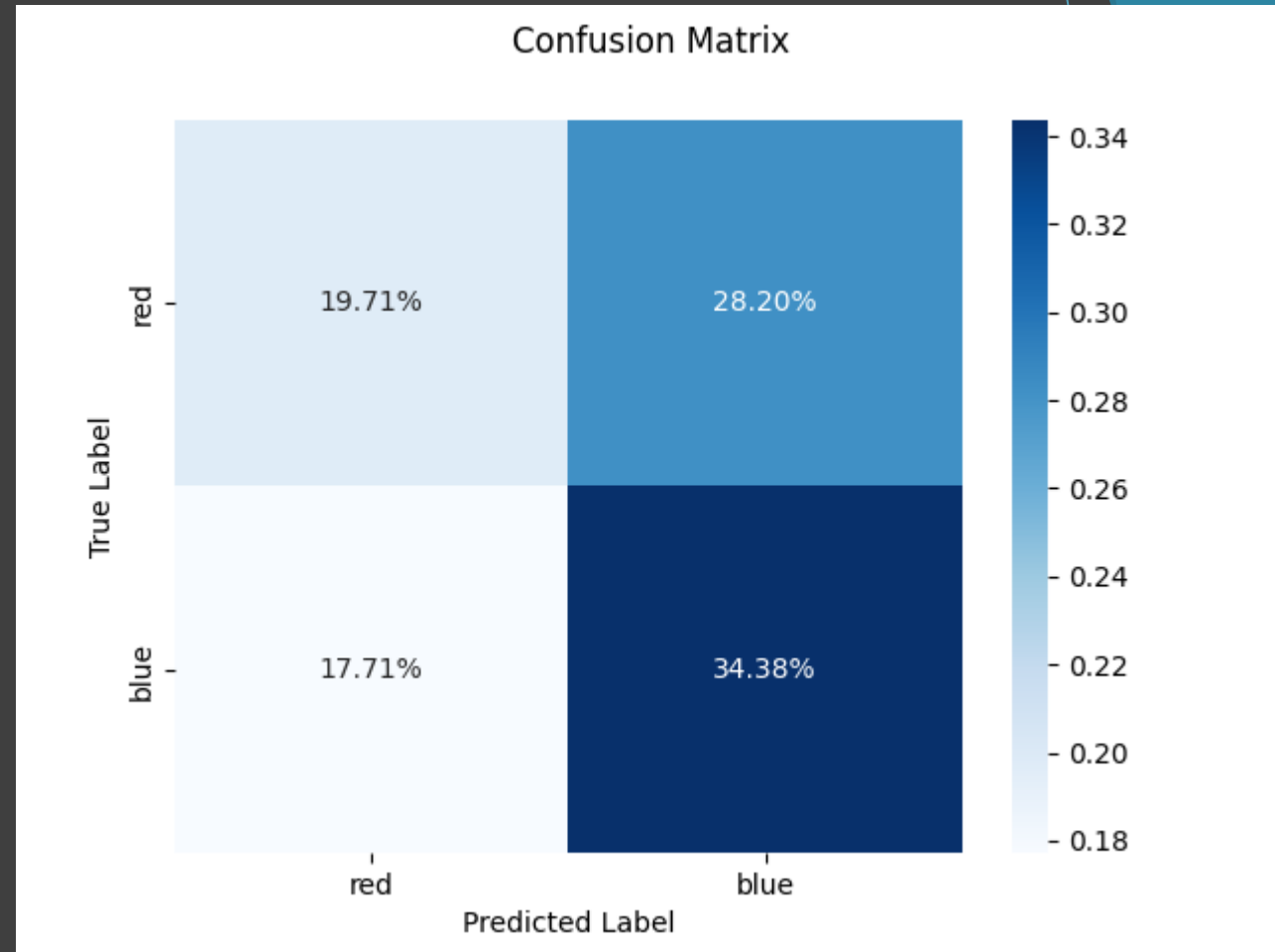
- ▶ Différence des ratios
- ▶ La moyenne des scores est de 0.5838
- ▶ L'écart type des scores est de 0.0261
- ▶ Précision = 0.6866

# Changement du critère de pureté

- ▶ Matrice de confusion avec le critère de gini :
- ▶ La précision avec gini est de 55.3459 %
- ▶ La moyenne des scores du critère gini est de 51.4675 %
- ▶ L'écart type des scores du critère gini est de 0.0264



- ▶ Matrice de confusion avec le critère d'entropie :
- ▶ La précision avec entropie est de 54.0881 %
- ▶ La moyenne des scores du critère entropie est de 49.4543 %
- ▶ L'écart type des scores du critère entropie est de 0.0237





# Conclusion

- ▶ Plusieurs critères pour prédire les résultats ont été utilisé
  - ▶ La méthode sur les champions ou les tags permet de prédire avec des nouveaux personnages
- ▶ Regarder la différence du ratio de victoire/total entre les deux équipes est la méthode la plus efficace, c'est celle que l'on a utilisé pour la prédiction avec les nouvelles données
  - ▶ 513 victoires pour l'équipe bleue
  - ▶ 329 pour l'équipe rouge
- ▶ Si on regarde les statistiques des joueurs, pas de très grande différence entre en utilisant quelques unes ou toutes

Merci de votre attention !

Avez-vous des questions ?