# Soccer Predictive Analytics

By Chizoba Obasi

PREMIER LEAGUE PREDICTOR

# Motivation

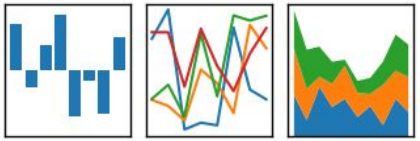

- **English Premier League (EPL)**
  - Most watched and most lucrative soccer league worldwide
  - Revenue of € 2.2 billion/yr in domestic and international TV rights
- **High level of uncertainty, unpredictability and variability in soccer**
- **Sports betting industry - a major global financial industry**

https://en.wikipedia.org/wiki/Premier_League

# Technologies

pandas

$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



python™

dmlc XGBoost

jupyter

scikit learn

NumPy

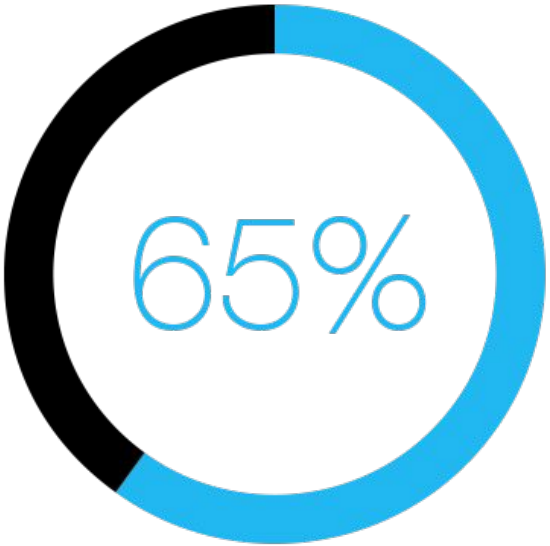
matplotlib

SciPy

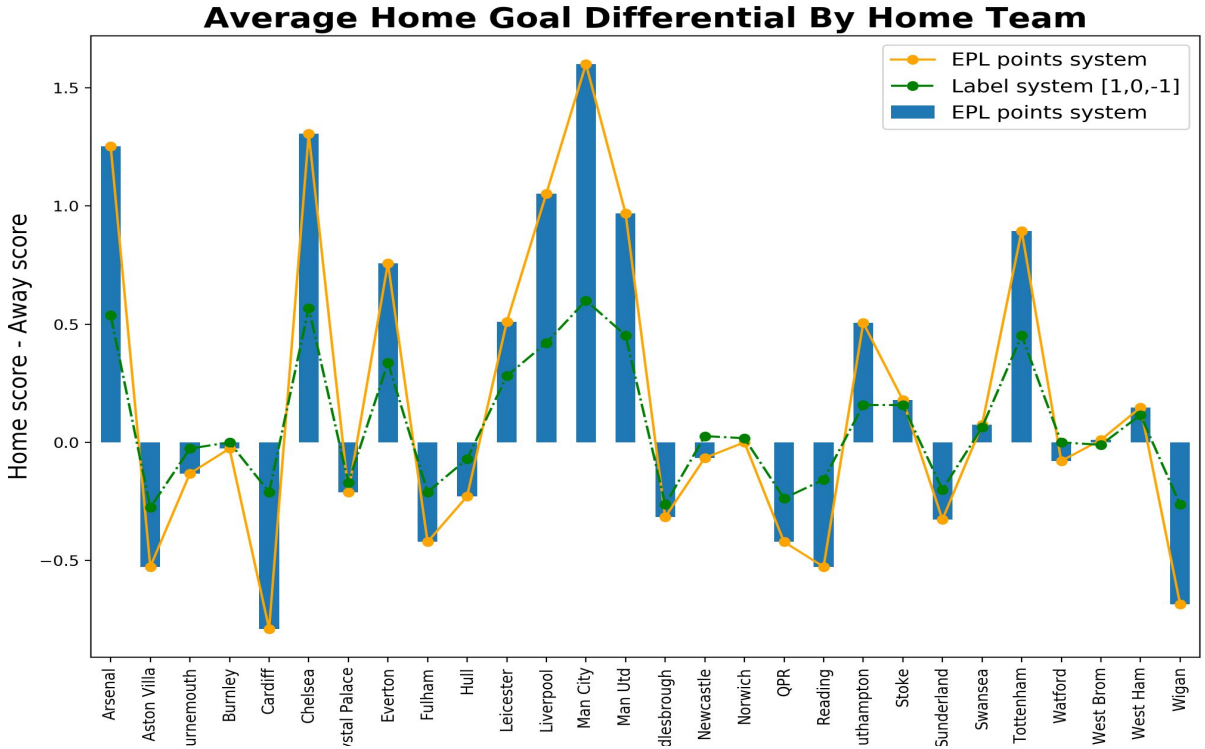
SM StatsModels Statistics in Python

# Win/Loss/Draw classification

- Random Forest Classifier (90+ features)



# Observations - Home Advantage Bias



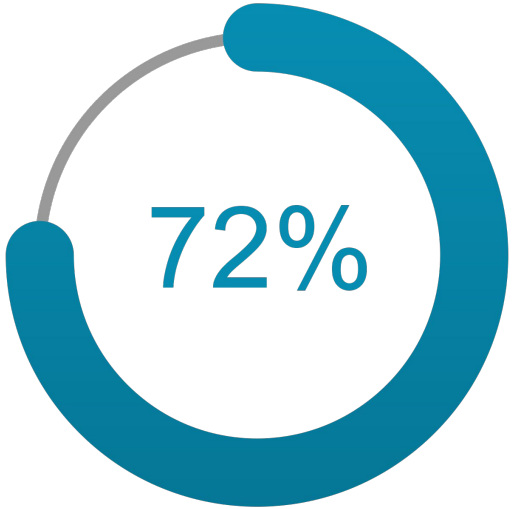

- **EPL point system**
  - Win: 3 points
  - Draw: 1 point
  - Loss: 0 points

- **Label Class system**
  - Win: 1
  - Draw: 0
  - Loss: -1

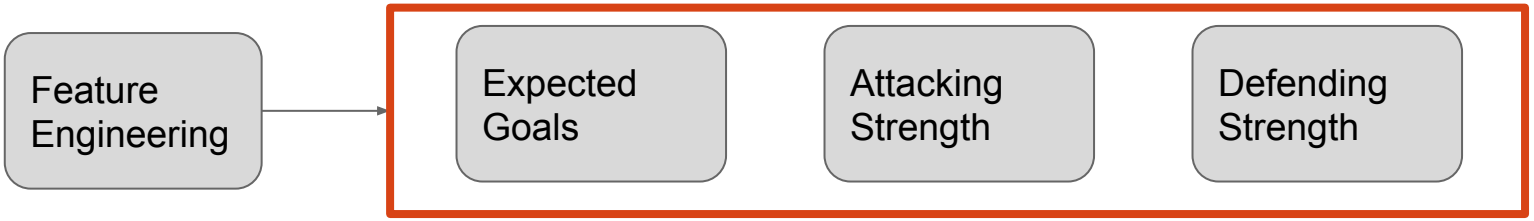
# Feature Selection


- **Random Forest Classifier (~30 features)**
  - **Poor accuracy of "tie" games' predictions**

| 72.40% | Predicted | | |
|---|---|---|---|
| TRUE | Loss | Draw | Win |
| Loss | 80.70 | 14.04 | 5.26 |
| Draw | 21.28 | 36.17 | 42.55 |
| Win | 5.23 | 8.14 | 86.63 |



# Model Improvement




- **XGBoost Classifier (~20 features)**
  - Data engineering
  - More feature selection

| 99.80% | Predicted | | |
|---|---|---|---|
| TRUE | Loss | Draw | Win |
| Loss | 99.72 | 0.00 | 0.28 |
| Draw | 0.00 | 99.58 | 0.42 |
| Win | 0.00 | 0.00 | 100.00 |

# Conclusion & Recommendations

- ★ **Feature selection** is very important for modeling improvement
- ★ More **informed feature engineering** based on target variable definition
- ★ **Individual player statistics** and impact on team tactics, and results
  - “Star” player effect
  - Player injuries, tiredness/rest
- ★ **Coaching tactics’** impacts on results
- ★ Accuracy evaluation based on **high win/loss chance** (>= +/- 2 goal differentials)


# Questions ?

linkedin.com/in/chizoba-obasi

co14@utexas.edu

github.com/chizkidd/dsi-CapstoneProj