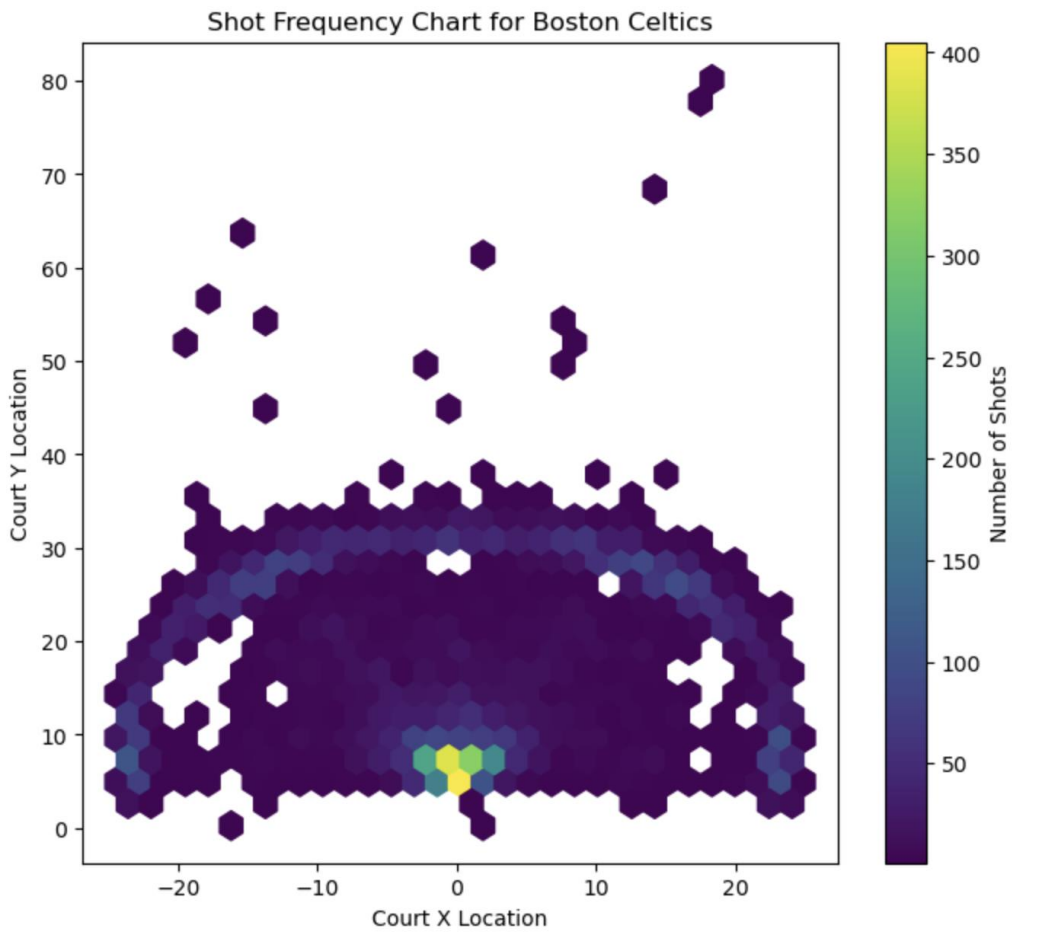
SPORTS STRATEGY OPTIMIZATION

KEVIN, LUKA, NATE, ERIC, SAHAT

CMSE 202 SECTION 4



WHAT WE'RE LOOKING FOR

- What key team performance metrics can an NBA team focus on to maximize their chance to win the most games?

WHAT WE'RE LOOKING AT

- Stats from NBA that influences winning % -
  - Shot types (2p, 3p, layups, etc.) – shot chart
  - Turnover %
  - True shooting %
  - Team adjusted offensive/defensive ratings
  - Height/Wingspan

OUR PROCESS

- Get data from official NBA API
- Setup Github and organize
- Create graphs and charts
- Run regressions and create a model that is a good indicator of what creates wins/losses
- Train/test model with data
- Make a correlation matrix
- Analyze results

OUR VARIABLES

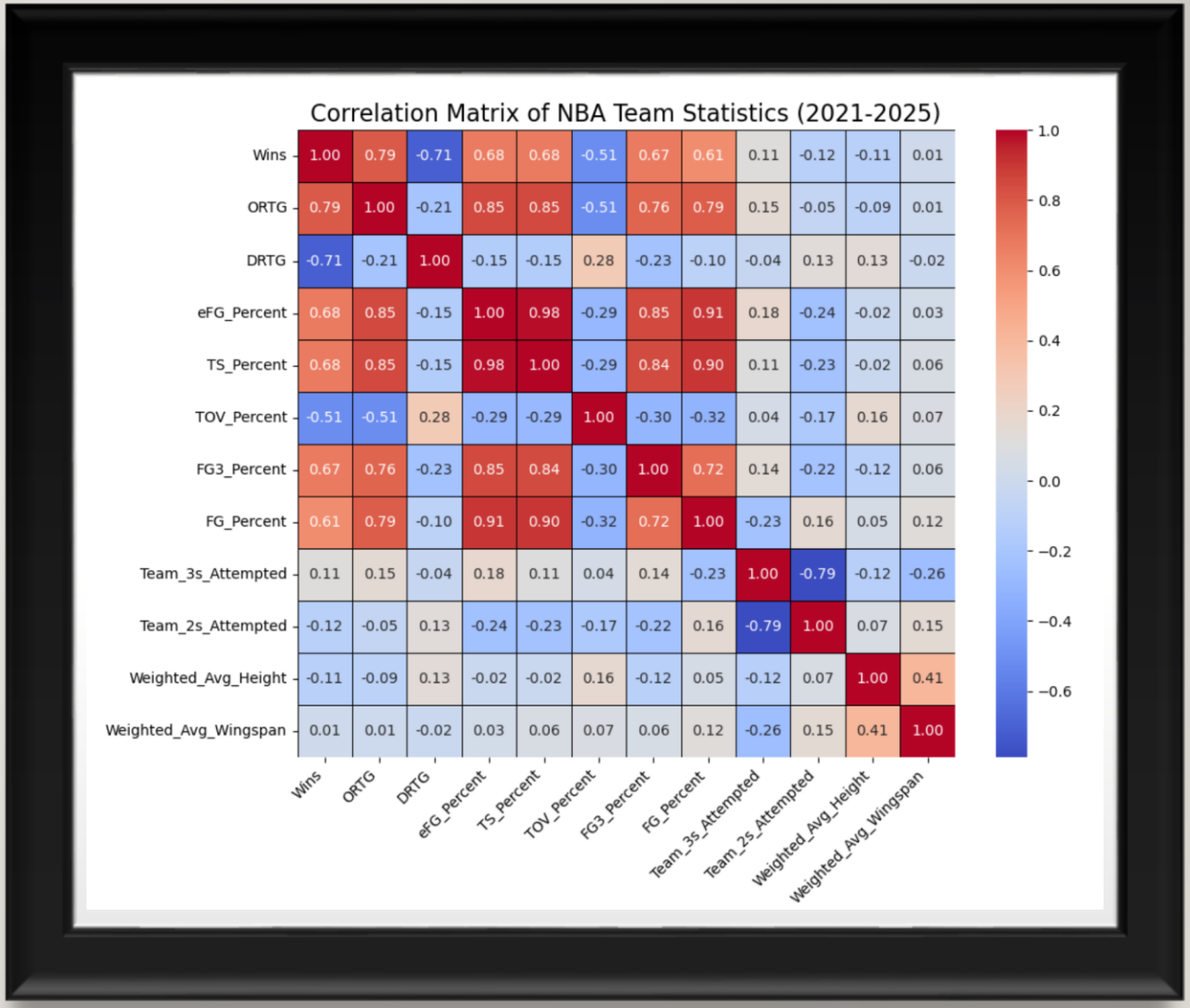
- Offensive Rating(ORTG)
  - Points per 100 possessions
- Defensive Rating(DRTG)
  - Points allowed per 100 possessions
  - The lower the rating, the better the defense is
- True Shooting Percentage

$$\text{TS\%} = \frac{\text{Points}}{2 \times (\text{FG Attempts} + 0.44 \times \text{FT Attempts})} \times 100$$

- Turnover%
  - The % of a teams possesions that end in a turnover

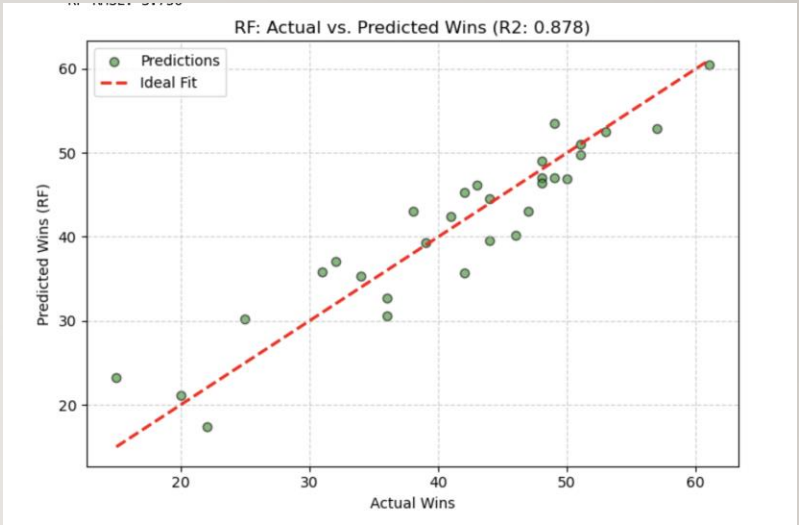
CORRELATION MATRIX WITH TEAM DATA:

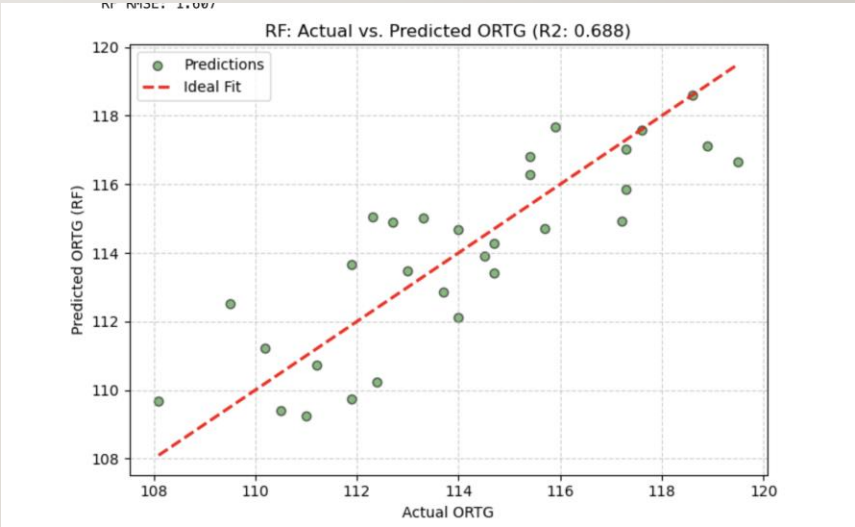
- Notable Pairwise Correlations
  - Wins/ORTG 0.79
  - Wins/DRTG -0.71
  - Wins/Height -0.11



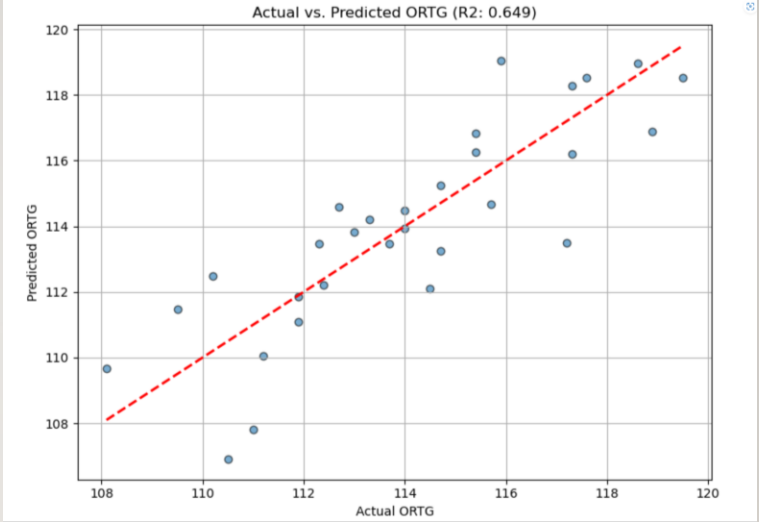
RANDOM FOREST PREDICTING OFFENSIVE RATING AND WINS

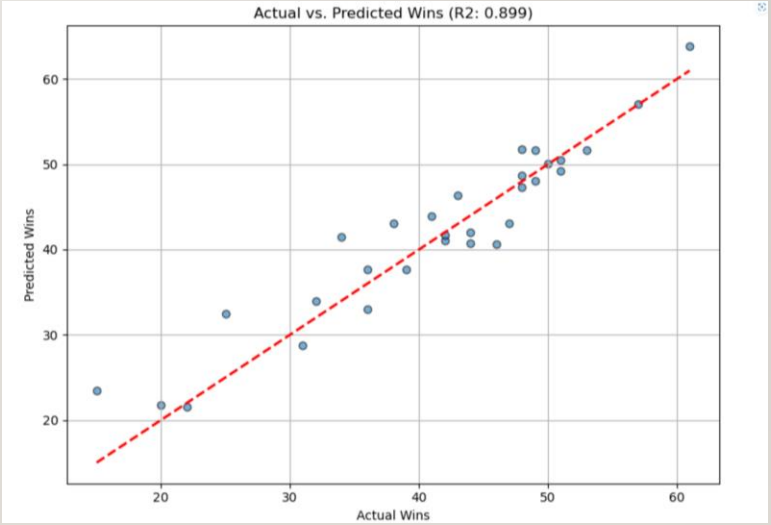
- Why random forest?
  - Random forest excels at handling non-linear relationships





LINEAR REGRESSION PREDICTING OFFENSIVE RATING AND WINS





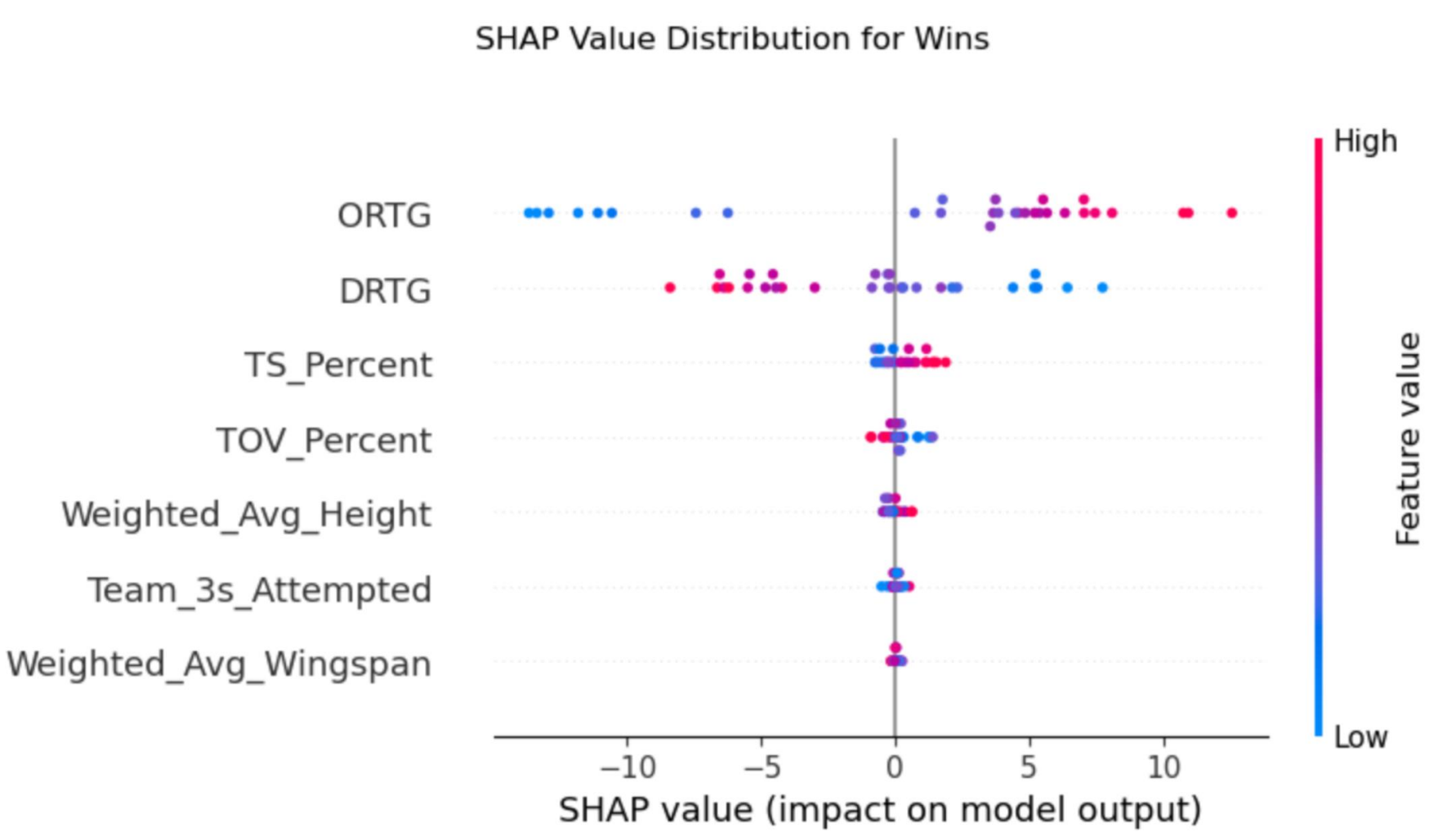
DIFFICULTIES AND COMPLICATIONS

- Importing the Python shap module and utilizing its methods for SHAP analysis was more complicated than the other parts, as this went beyond the material taught in class.
- Finding data beyond the NBA.com API
- Basketball changes over time, so it's hard to get a good sample size

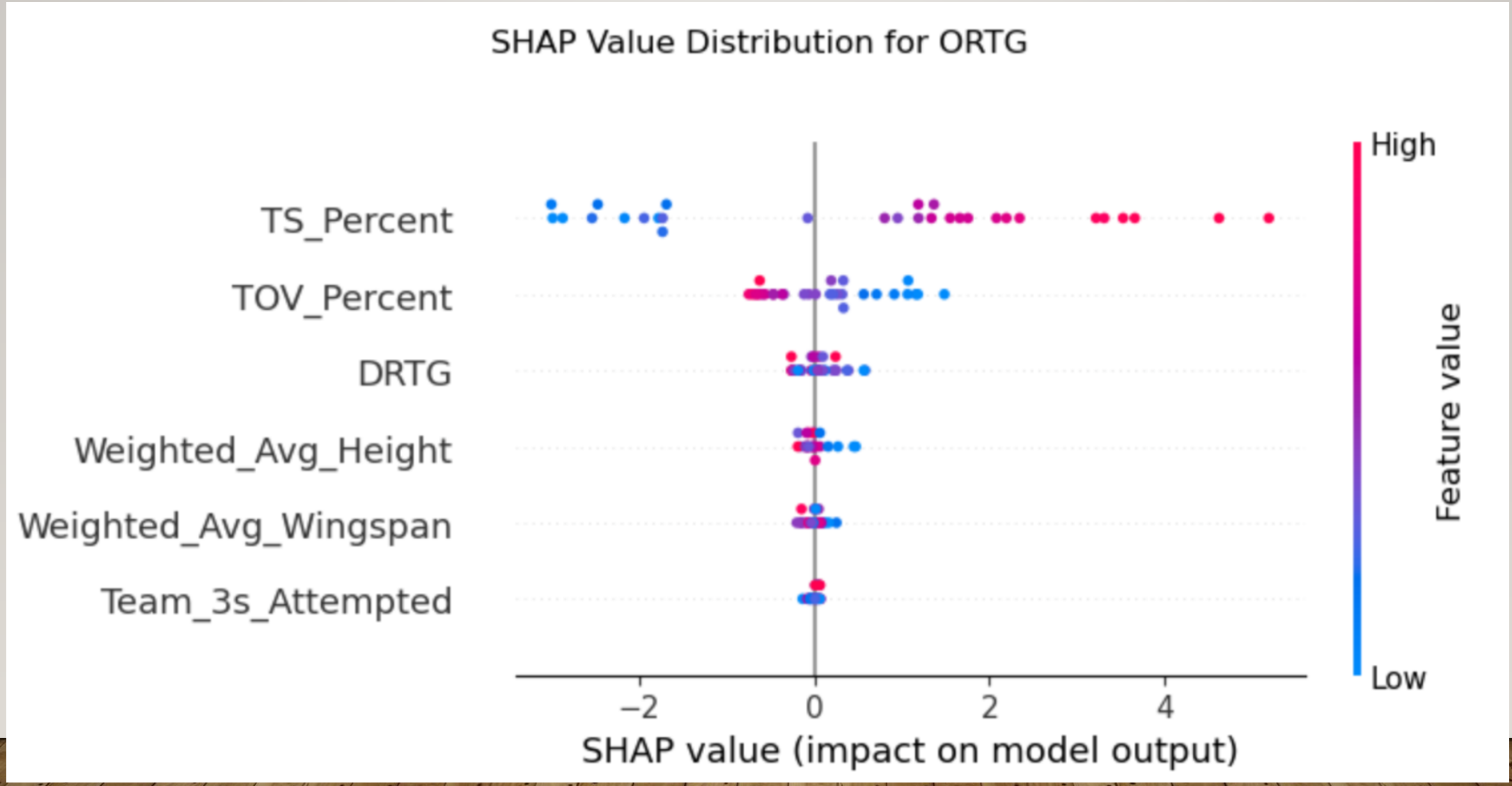
FACTORS PREDICTING WINS

The wider the spread in SHAP (SHapley Additive exPlanation) values, the more important the factor is when predicting wins:

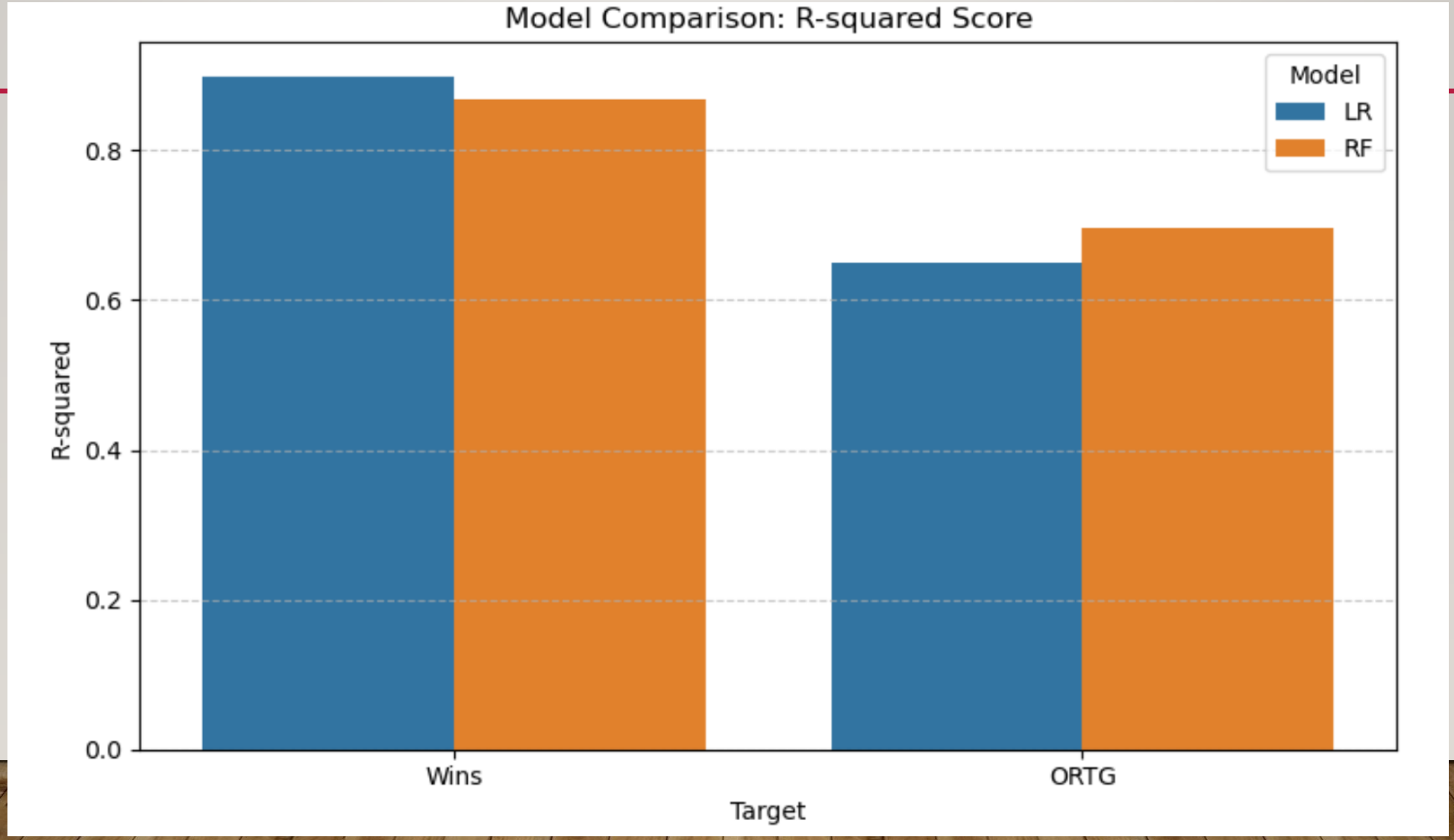
- Offensive rating is the strongest predictor for wins:



FACTORS PREDICTING OFFENSIVE RATING



RESULTS: LINEAR REGRESSION (LR) AND RANDOM FOREST (RF) MODEL COMPARISON



CONCLUSION

- Offensive efficiency should be prioritized over defense.
- This can be achieved by increasing shooting efficiency (high TS%) and minimizing turnovers (low TOV%).
- Other interesting finding: greater average height has a slightly negative correlation to wins and offensive rating for a team.