Primary Questions:

- What team variables are most predictive of the following season's success?
- Can we correctly predict the following season's top contenders?

Team Variables Used

- **Current season win %**
    - Season wins / total season games
- **Offensive Rating**
    - Measures how many points a player or team scores per 100 possessions
- **Defensive Rating**
    - The number of points a player or team allows per 100 possessions
- **Net Rating**
    - Shows the difference between points scored and points allowed per 100 possessions
- **Assist %**
    - The % of team field goals that were assisted
- **Assist-Turnover Ratio**
    - Dividing a team's total assists by their total turnovers
- **Offensive Rebound %**
    - The % of the time a team secures an offensive rebound
- **Defensive Rebound %**
    - The % of the time a team secures their defensive rebound
- **Effective Field Goal %**
    - Measures a player's or team's shooting efficiency by adjusting for the extra value of three-point shots
    - EFG % = (2pt_FGM + 1.5 * 3pt_FGM) / FGA
- **True Shooting %**
    - Measures scoring efficiency by incorporating field goals, three-point shots, and free throws
    - TS % = Total Points / 2 * (FGA + 0.44 * FTA)
    - Players with 60+ TS % are typically considered efficient
    - Accounts for free throws unlike EFG %
- **Pace**
    - How many possessions a team has in a game
- **Opponent Field Goal %**
    - The opponent's FG %

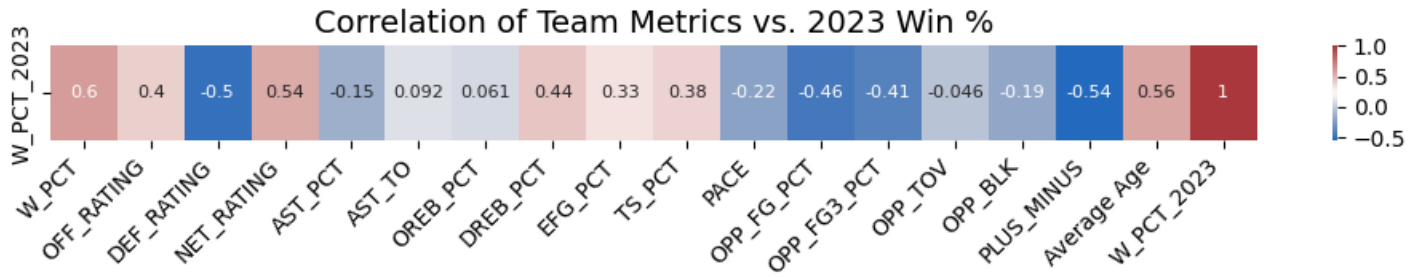
- **Opponent 3pt %**
    - The opponent's 3pt %
- **Opponent Turnovers**
    - How many turnovers they force onto the other team
- **Opponent Blocks**
    - How many blocks they allow opponents to gather
- **Plus Minus**
    - In our data, this is expressed as the opponent's point differential, meaning negative values are better
- **Average Age**
    - The average of the team roster's age

The Logic:

1. Good teams score well on advanced metrics
2. We can identify "good" teams by analyzing these variables
3. Good teams will likely stay good next season

The Approach:

1. Compile data for each team from the 2021-2022 season
2. Analyze the correlation between this data and the 2022-2023 season winrate



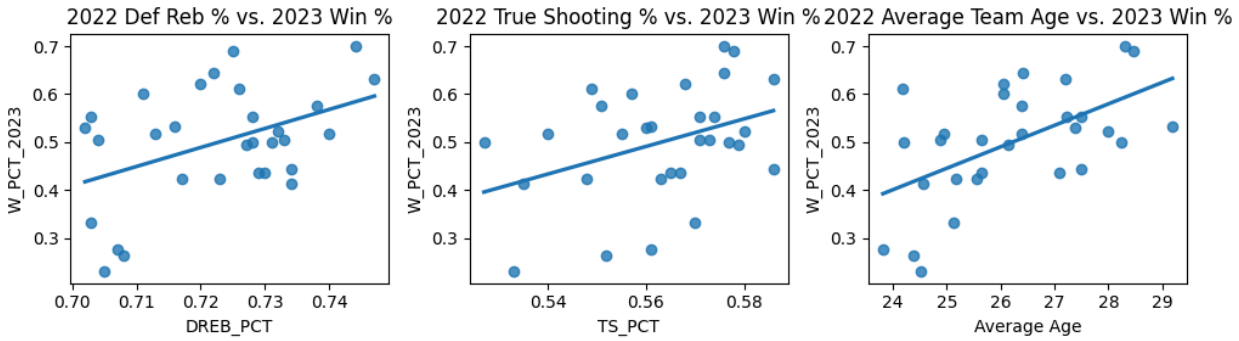
Top Positive Correlations: Current season win percentage, net rating, offensive rating, average age, defensive rebound percentage, and true shooting percentage.

Top Negative Correlations: Defensive rating, plus_minus, and opponent field goal percentage.

Which Teams Are Most Influential?

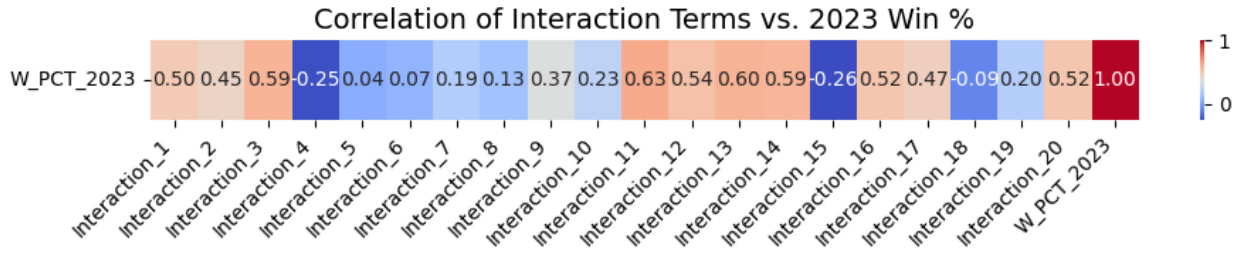
Recall our logic is premised on "good teams staying good" for next season. This means the meaningful correlations in our data are driven by teams that performed well

in 2022 and again in 2023. The teams that stayed in the overall top 8 from 2022 to 2023 are the Bucks, Celtics, 76ers, Cavaliers, Nets, and Heat. These teams were all offensively sound However, the Bucks (Giannis and Jrue), Celtics (Marcus Smart and Robert Williams), and Heat (Jimmy and Bam) were better regarded as defensive specialists. This dynamic is likely responsible for defensive rating having a greater weight than offensive rating. Similarly, none of these teams were considered especially young or inexperienced. They all were experienced and had ample veteran presence such as the 76ers with Paul Milsap. Lastly, these teams played at a moderately slower pace. Defensive stops helped create transition opportunities, but they hesitated to intentionally push the gas. Playing steady allows your defense to dictate your offense.



Interactions

Combining existing variables can create new ones with higher correlation values than any individual variable. We created and tested 20.



Our focus is on *three* specific interactions.

Interaction 11: Formula is (Current Season Win %) * (Average Age).
This interaction achieves a very high correlation value of 0.63. Win % as a standalone variable is inherently correlated with other important metrics such as shooting % and defensive rating. So here, a high win % is also bringing those other values. By adding

age, the interaction seems to suggest that older teams have a better ability to close out games and deliver wins.

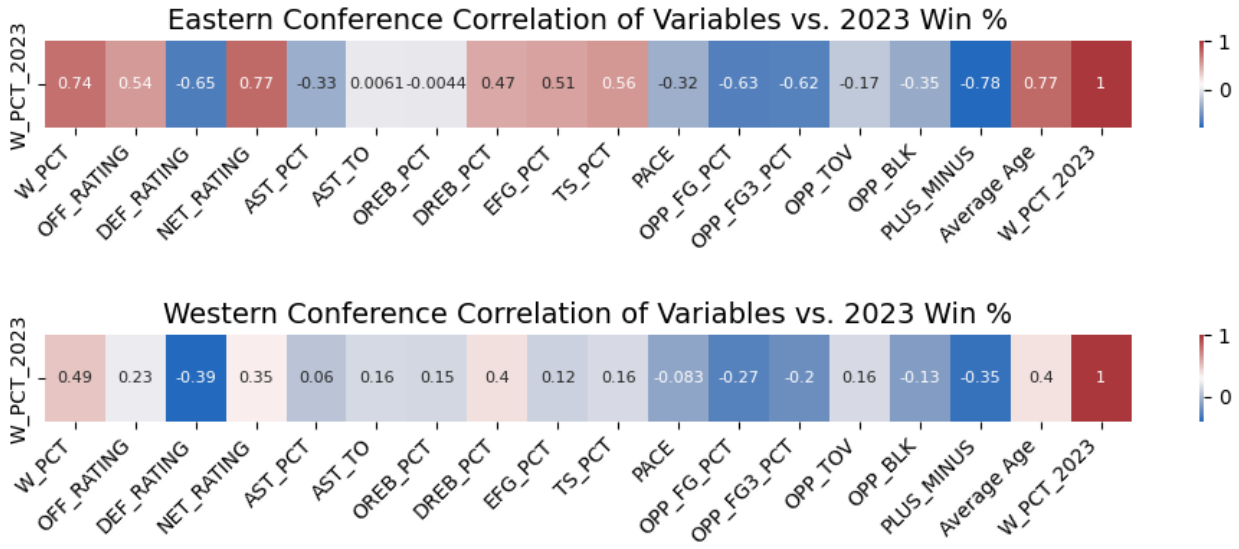
Interaction 13: Formula is (Defensive Rebound %) * (True Shooting %) / (Def Rating). Teams that score efficiently, secure possession ending rebounds, and play solid defense perform well next season. This logic is intuitive and self-explanatory. This interaction covers basketball's 3 core tenets: offense, defense, and rebounding.

Interaction 19: Formula is (Offensive Rebound %) / (Opp FG %).
A single offensive rebound can swing momentum and completely change a game, but their statistical importance is not captured. This interaction's correlation was only 0.20. This interaction figured that teams with the hustle to grab offensive rebounds also played lockdown defense.

Segmenting Data

Eastern v. Western Conference

Splitting our data by conferences reveals stark differences.

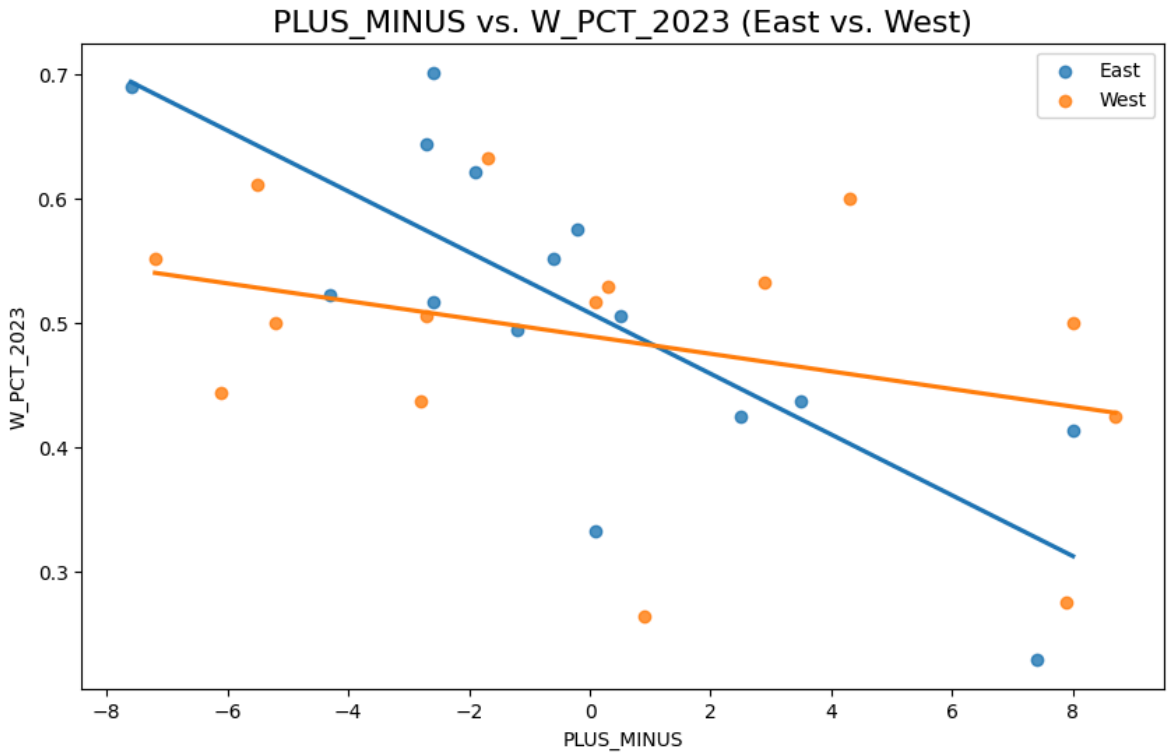


--- ABSOLUTE DIFFERENCE (Sorted) ---

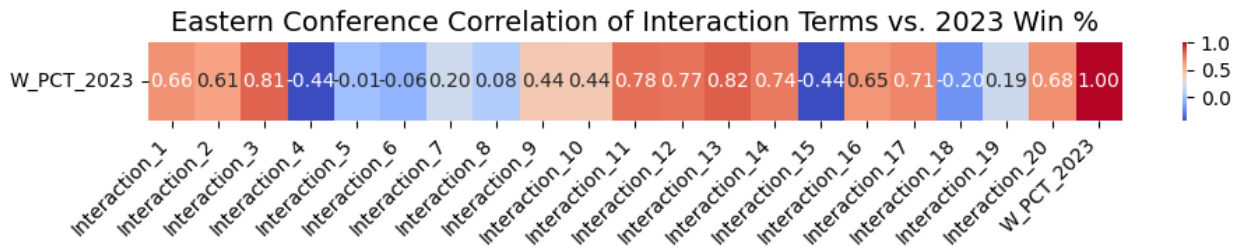
PLUS_MINUS	**0.426895**
OPP_FG3_PCT	**0.424483**
NET_RATING	**0.416905**
TS_PCT	**0.399401**
AST_PCT	**0.390847**
EFG_PCT	**0.390216**

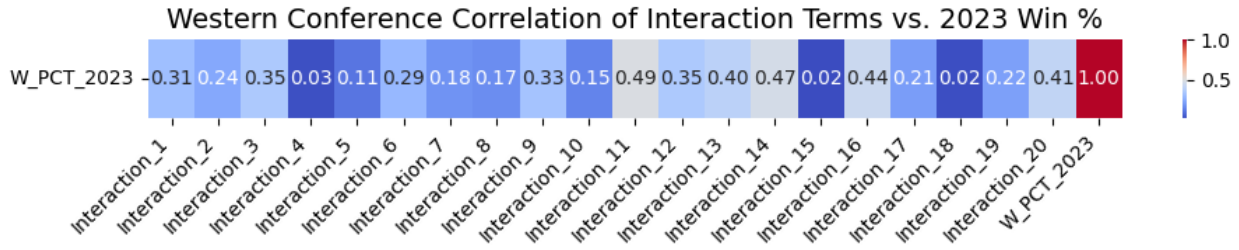
Average Age	**0.374548**
OPP_FG_PCT	**0.355385**
OPP_TOV	**0.330018**
OFF_RATING	**0.313876**
DEF_RATING	**0.255145**
W_PCT	**0.247701**
PACE	**0.238621**
OPP_BLK	**0.221350**
OREB_PCT	**0.150023**
AST_TO	**0.149120**
DREB_PCT	**0.072181**
W_PCT_2023	**0.000000**

Stats like plus_minus and true shooting % appear to matter much more in the east than the west. Meanwhile, stats including defensive rebound percentage are equally important in either conference. Why? Recall our list of influential teams (Bucks, Celtics, 76ers, Cavaliers, Nets, and Heat). These are all teams in the eastern conference. Therefore, shooting efficiently (true shooting %) and dominating opponents (plus_minus) is much more important in a stronger conference. Simple actions like a defensive rebound are less likely to push the needle. Evidently, western teams within this dataset were either weak or changed too much season-season to be impactful.



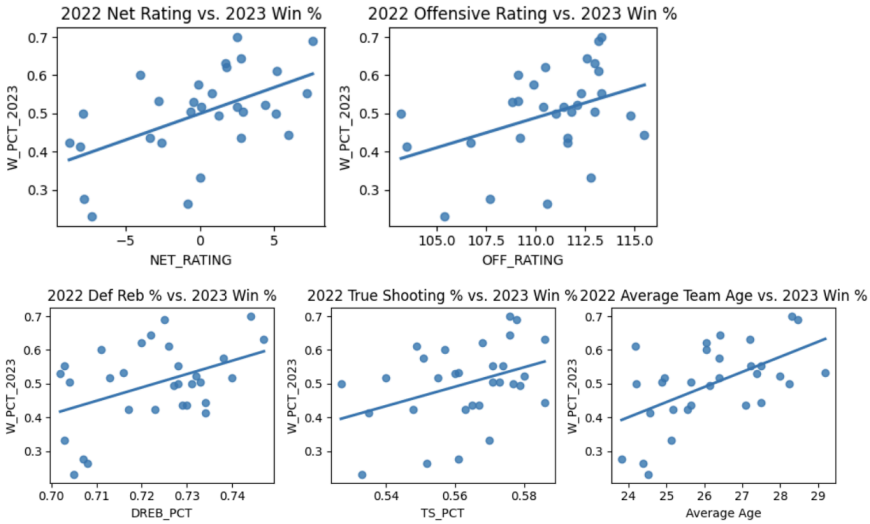
Interactions Across Conferences





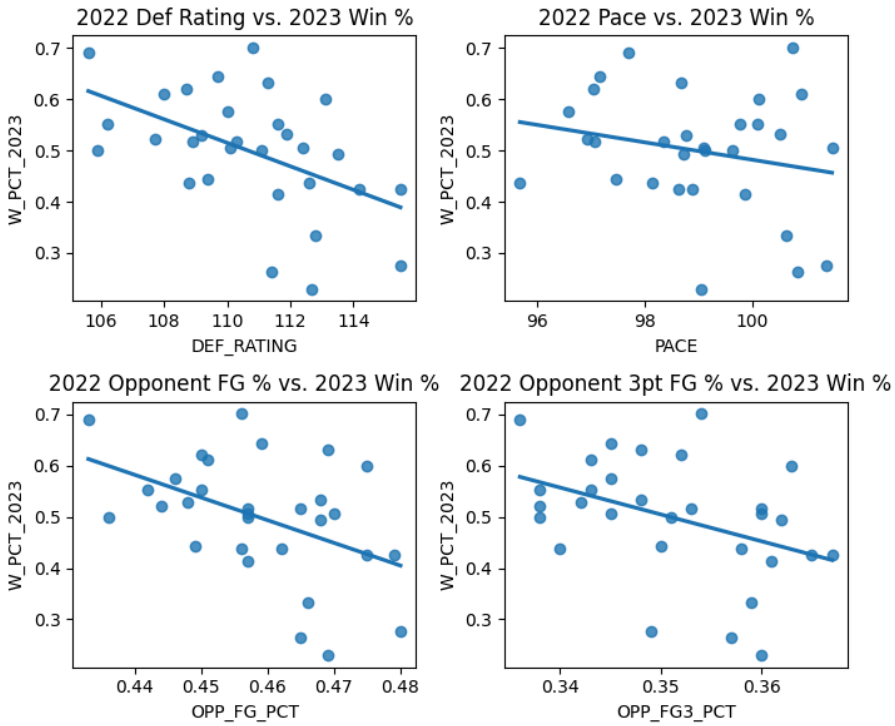
Again, the correlations are much stronger in the east for reasons previously mentioned. Amongst the interactions we previously discussed, interaction 13 performs well in both conferences. Remember that this interaction captures offense, defense, and rebounding; I expect it to be robust no matter the era or context.

Offensive and Defensive Variables are Indicative of Future Success



As depicted in the graph above, Offensive Rating (OffRtg) and True Shooting % (TS%) both show clear positive correlations with next season win percentage (r ~ 0.4 – 0.55).

At the same time, the scatterplots for 2022 vs. 2023 OffRtg vs Win % and TS% vs Win % have steep upward slopes supporting the consensus that teams that create efficiency in shot selection tend to sustain or increase their winning levels. We believe this is due to the fact that the way today's game is played is massively different from the past. Modern NBA offenses are built around spacing, pace, and isolation scoring, leading to a disproportionate impact from star players. The data may be derived from the league's evolution after 2013/14, with breakout seasons from 3pt shooters like Stephen Curry, Klay Thompson, and Damian Lillard (albeit to a lesser extent). There are also arguments to be made about a strong correlation between a team's win percentage and its 3-point shooting, as teams that shoot a higher percentage of 3-pointers in today's game tend to win more often, especially when outshooting their opponent. It should be noted that the simplicity of shooting more 3-pointers does not automatically lead to more wins. While this is a well-known theory, the argument requires further research and our data is insufficient to support either side.



A lower defensive rating in basketball represents a lower number of points an opponent scores per every 100 possessions

The data reveals that defensive performance serves as the stronger indicator of future team success. Defensive Rating shows a clear negative correlation with next-season win percentage (r ~ -0.5), meaning that teams allowing fewer points per 100 possessions tend to maintain higher win rates in the following year. Similarly, both Opponent Field Goal Percentage and Opponent Three-Point Percentage display strong negative relationships with win percentage, highlighting that limiting the quality of shots opponents take (particularly from beyond the arc)is more predictive of long-term success than factors such as pace.

For pace, findings are mixed. Faster pace inflates counting stats and can correlate weakly with next-year W% if a team plays sloppy and out of control rather than taking advantage of fast break transitions. For instance, Sacramento used pace as a means to better shots (early drag screens for Fox/Sabonis DHOs), not as an end. We are curious to see if controlling for Age and Pace raises the correlations of EFG_PCT and AST_TO, suggesting pace is more of a mediator rather than a driver.

From a basketball standpoint, the data underscores a shift in what effective defense means. In today's era, contesting perimeter shots may be more critical than accumulating blocks or steals. Because most offensive possessions now begin with multiple shooters spread across the floor, preventing open looks from three point range has a far greater impact on opponent efficiency than traditionally packed in interior defense. Historically elite defenses like the 2022 Celtics ranked among the top five in opponent three-point defense and transition coverage (which then went on to win the championship) illustrate this pattern.

https://www.espn.com/nba/stats/_/season/2022/seasontype/2/view/team?

Our findings show that defensive stability is more likely to translate from one season to the next. While offenses fluctuate more due to shooting variance, new rosters (take chemistry on the floor for example which is definitely more important on the offensive end than defense which is mostly effort based), and new coaching schemes, defenses tend to carry over and sustain success over time.

Predicting Next Season's Top Contenders

Our variables are from 2021-22 and we correlated it to 2022-23's winrate. In order to truly test how good interaction 13 is, we will apply it to data from the 2017-2018

season. Our hope is to pick the top 8 in each conference for the subsequent season (2018-2019).

2018-2019 Eastern Predictions

1. 76ers
2. Raptors
3. Celtics
4. Heat
5. Hornets
6. Cavs
7. Wizards
8. Pistons

2018-2019 Eastern Results

1. Bucks
2. Raptors
3. 76ers
4. Celtics
5. Pacers
6. Nets
7. Magic
8. Pistons

East Summary

- Included 4/8 teams correctly.
- Incorrect: Heat, Hornets, Cavs, Wizards
- Missing: Bucks, Pacers, Nets, Magic

2018-2019 Western Predictions

1. Rockets
2. Jazz
3. Warriors

4. Trail Blazers
5. Pelicans
6. Spurs
7. Nuggets
8. Mavericks

2018-2019 Western Results

1. Warriors
2. Nuggets
3. Trail Blazers
4. Rockets
5. Jazz
6. Thunder
7. Spurs
8. Clipper

West Summary

- Included 6/8 teams correctly.
- Incorrect: Pelicans, Mavericks
- Missing: Thunder, Clippers

Overall

Our predictions were more accurate in the east than the west. Considering interaction 13 had higher correlation values in the east than west with its training data, this is unsurprising.

Limitations

Roster Changes

Our approach is sensible, but limitations exist. A team may perform horribly one season, replace players during the preseason, then perform significantly better. For example, our Western Conference prediction completely misses the Clippers. The Clippers the previous season were a bad team. They became better because they acquired Shai and Zubac after the regular season. There was no way for our model to account for this change and adjust accordingly.

Injuries

A team may be missing their star player for a season, making their team stats appear worse than normal. Then next season, that player may return and flip the team's trajectory entirely. This dynamic did not play a role in our prediction, but there are many teams where this occurs. Looking ahead to the current season, models may expect the Pacers to top the east, but that is extremely unlikely without Haliburton available.

Coaching and Culture

Coaching is very commonly overlooked by basketball fans. Success follows some, and failure follows others. Some coaches are simply better at unlocking a team's full potential. Similarly, some organizations (team franchises) are notorious for having a "winning" culture. Meanwhile, others have never come close to a championship. Teams are more than a starting five and bench. There's an entire bureaucracy of front office workers, coaches, trainers, and even janitors that contribute to a team's culture. For example, our Eastern Conference prediction misses the Bucks. They were an okay team the previous season, but coach Mike "Bud" joined the following season and they ended the season as first seed. Eventually, he led them to a championship, highlighting how a coach plays a key role in unlocking their team's potential. Our model could not account for this with its existing parameters.

Conclusion

The best predictors for future success are kept simple. Teams that score efficiently (TS %), stop their opponents from scoring (OPP FG %), and secure the ball (DEF RB %) perform well. However, our data was susceptible to biases from its training set and outsized influence amongst the league's top teams. One solution is likely increasing the timeframe from where data is pulled. Therefore, the methodology is not only exposed to a single season's trends and schemes. Where our approach fell short is its inability to capture certain factors. More sophisticated models that successfully consider intangibles such as culture and organizational competence likely perform much better. However, even with these limitations, our predictions were reasonably in-line. Between both conferences, we predicted 10/16 of the next season's top contenders correctly. This result affirms our assumption that "good teams will likely stay good next season."

All analysis and conclusions based off of 2021-22 and 2022-23 data, may not be transferable to other past or future seasons