Paper ID # 33

"Optimizing NBA Lineups: A Net Rating and Win Probability Analysis"

Santiago Soto, Patrick Lucey, Morehouse College/Stats Perform

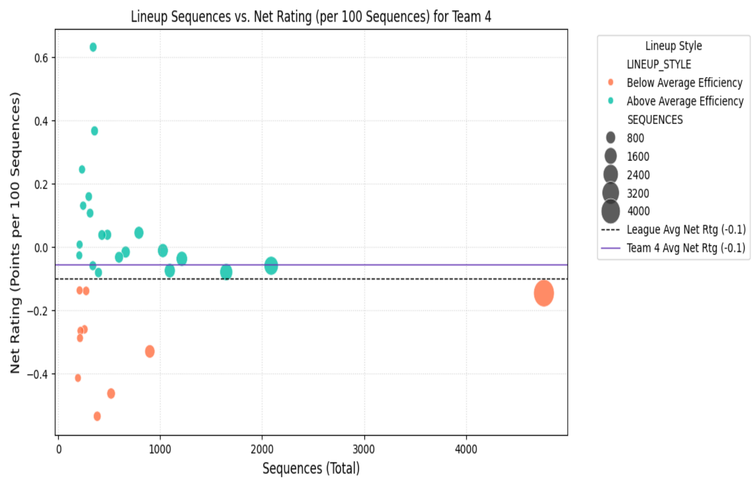
- **The Foundation: "What Makes a Lineup Optimal?"**

➢ Though it may seem very straightforward, not all high-scoring lineups are great. The best lineups are efficient on both offense and defense. Historical data from all 30 NBA teams over the past 20 years were used to examine what made certain lineups the most effective.

➢ Key Metric: Net Rating
- To understand the true value of a five-man unit, we must break down its efficiency into two core components:
- Offensive Rating (OFF_RTG): The total points scored by the lineup per 100 sequences.
- Defensive Rating (DEF_RTG): The total points allowed by the lineup per 100 sequences.
- The final Net Rating is simply the difference between the offensive and defensive ratings.
- A sequence is defined as an event within a basketball game, whether that be an out-of-bounds call, a made basket, an attempted basket, or a pass. This is because I used play-play data that was measured in terms of sequences.

➢ Interpreting Efficiency through Net Rating:

➢ A positive Net Rating signifies an Above Average Efficiency unit that successfully outscores its opponents. Conversely, a negative or zero value identifies a Below Average Efficiency unit based on the league-wide median.



➢ As shown in the analysis of "Team 4," a majority of their lineups fell below the league average efficiency, indicating that while they may score points, they often lack the defensive consistency required to be a top-tier team.

Visualizing Team and Lineup Efficiency: Consistency and Quality

➢ You can see the direct relationship between a team's efficiency and its success. Not all lineups within a team are created equal. Data suggests that elite teams may score fewer total points than average teams in some scenarios, but they are the only ones that consistently allow fewer points defensively.

➢ The image below shows a chart indicating some of the best lineups by the calculated net rating, and how often good vs average vs below average lineups play over the course of the game.

Top 5 Lineups from Other Teams (Benchmarking):

ONCOURT_TEAM_ID	LINEUP_ID	SEQUENCES	NET_RTG
29	Aldama-Bane-Clarke-Green-Konchar	22	2.94698
27	Avdija-Davis-Morris-Nunn-Porzingis	20	2.78522
15	Carter-Holiday-Mamukelashvili-Nwora-Portis	22	2.78417
15	Antetokounmpo-Beauchamp-Connaughton-Matthews-Wigginton	38	2.61605
14	Butler-Martin-Oladipo-Robinson-Strus	20	2.49532

--- 3. Strength of Opponent Analysis for Team 4 ---
Avg Net Rating vs. Stronger Opponents (>1.0 Factor): **0.03**
Avg Net Rating vs. Weaker Opponents (<=1.0 Factor): **-0.05**
Performance Delta (Strong - Weak): **0.09**

--- 4. Lineup Quality Distribution for Team 4 ---

Percentage of Total Sequences Played by Lineup Quality (Based on Team Percentiles):

Quality	Sequences %
Good	16.38%
Average	64.43%
Weak	19.19%

Average Key Metrics by Lineup Quality (Per Sequence):

QUALITY	Avg_Off_Rtg	Avg_Def_Rtg	Avg_Reb_Pct	Avg_Stl_Rate
Good	108.855	108.222	0.684	0.035
Average	111.816	111.848	0.693	0.04
Weak	112.273	112.921	0.704	0.032

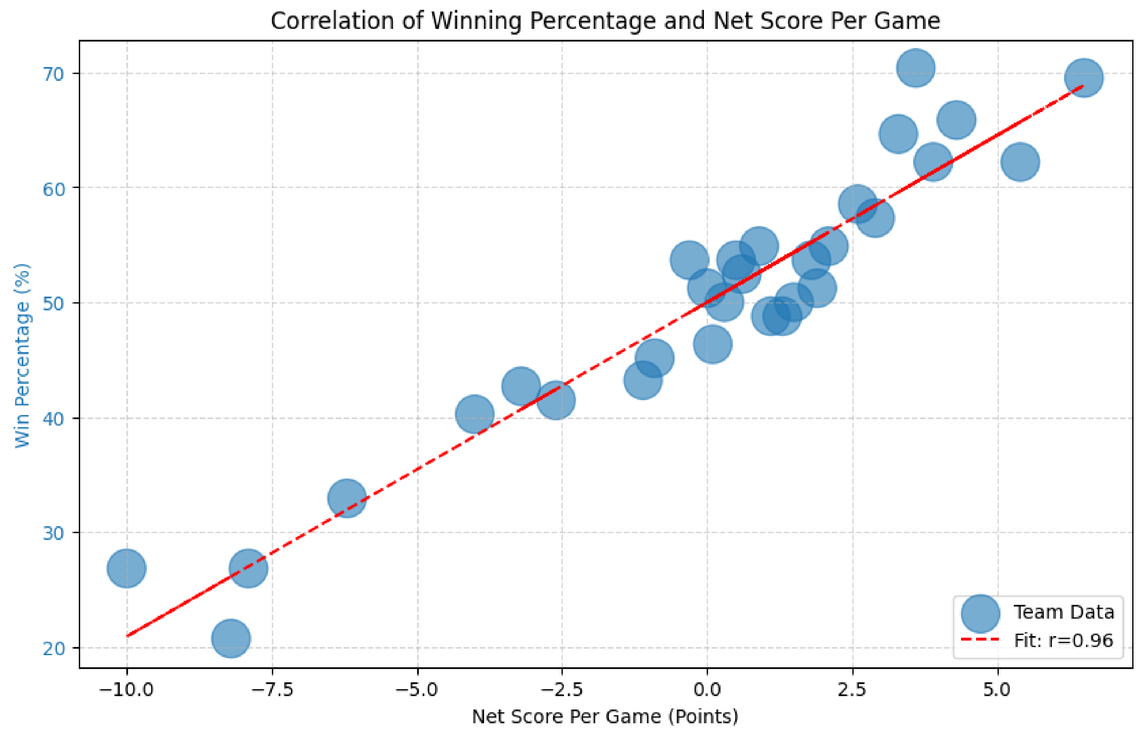
From Analysis to Prediction: The Win Probability Model

➢ We can use this efficiency data to build a simulation that predicts game outcomes in real-time. The model processes detailed play-by-play data to group statistics into unique 5-man lineups.

➢ Team strength is determined by taking a weighted average of these lineup metrics, ensuring that the players who play the most have the greatest impact on the prediction.

➢ To keep the simulation realistic, scaling factors are applied so that projected totals for points and rebounds align with specific league targets, such as an average of 114 points per game.

➢ The Flowchart functions something like this:



Summary/Conclusion

➢ This research demonstrates that Net Rating is a powerful, single metric for evaluating lineup performance.

➢ We found that defensive consistency is the primary differentiator between good and average teams. not just offensive output.

➢ The interactive Win Probability model translates historical efficiency into real-time game predictions.

➢ The ultimate goal of this research is to identify player combinations that maximize the spread between offense and defense. By focusing on Net Rating rather than raw scoring, teams can better understand which lineups provide the most stability. The simulation model demonstrates that by optimizing these styles, we can generate win probabilities and project game statistics with high confidence.

Next Steps

➢ Given that this is an ongoing project, the next steps would be to integrate Agentic AI Coding.

➢ We also want to be able to highlight the complexity of the problem.

➢ Ensure that the definition of an optimal lineup isn't difficult for the user to understand, all while highlighting how complex optimization is.

Libraries Used

➢ Openpyxl, Matplotlib, Seaborn, Ipywidgets, pandas, numpy, streamlit