

Statistical Basketball Player Position Analysis

Anna Wu

Mission San Jose High School
Fremont, CA 94539, USA
anna.dong.wu@gmail.com

Mason Chen

Stanford Online High School
San Jose, CA 95132, USA
mason.chen.training@gmail.com

Abstract

The purpose of this STEM project is to develop the Sports Analytics concept for playing Basketball by forming a special Basketball Team. Authors have defined the key functional metrics for each of the three player positions: Guard – Shooting and Assists; Forward – Shooting and Rebounds; and Center – Blocks and Rebounds. The objective is to maximize team performance through effective performance across all player positions. In order to maximize team performance, Guard will try to pass the ball to Forward. Center should focus on Rebounding and Blocking, not Steals. Each player's Statistics are recorded by two different individuals. Hypothesis tests are conducted (1) to verify whether players can fulfill their position within their team, and (2) to perform a position comparison between the two teams. Sample sizes are determined in order to conduct a 2-Proportions Test at 95% confidence. In the Analyze phase, the authors decided to add Steals (for all three players), Turn Overs (for Guard), and Offensive/Defensive Rebounds into the Project Scope. Statistical analysis software (JMP and Minitab) was used to study the multi-correlation patterns. The winning team demonstrates better team communication among players by playing according to this unique Basketball Game player position model.

Keywords

STEM, Basketball, Sports Analytics, Hypothesis Testing (2-Proportions Test)

1. Introduction and Literature Research

In basketball, there are many positions and unique roles. While other authors have used many advanced analytic techniques to assess player performance (Oliver, Hollinger), these methods have not provided tangible insights which determine overall team performance (Goldsberry). In fact, a review of performance by position using foundational statistical methods could not be readily found in the authors' review of published literature. A team that is properly assigned per position will have maximum efficiency because every player is in their best-matched position in terms of their ability to meet the positions' performance criteria (Trinic and Dizdar). But a team that is incorrectly assigned will be poorly coordinated with players not knowing how to play in their designated position. Therefore, in this project we want to find out how we can correctly assign a player's position. To solve this problem, data collected after each game is analyzed to determine whether or not each player is correctly assigned in each position. For example, if Center is scoring more than Forward, then these two players are not in their correct positions. In a classic 3-on-3 basketball game, there are 3 players, each in a different position. There is one Point Guard, one Center, and one Forward. Each one of these positions has their own role that they must play. The Point Guard's role is to be the leader and coordinator of the team. The Point Guard does a lot of dribbling and passing/assists. Next is the Forward. The Forward's responsibility consists mainly of shooting and scoring. The Forward also does a lot of dribbling but will receive the ball from the Point Guard and shoot. Lastly, the Center's

role is to block the opponent's attempts to score and to get rebounds. When the Forward misses, the Center's job is to recover the ball and pass it back to the team. The Center is also the team's defense, and should attempt to block as many shots and passes as possible. In our project, we have assigned players to each position and will analyze their statistics to see if they were properly assigned.

2. Define Project Scope and Baseline Analysis

In the following section, four main areas will be covered: (1) Project Challenges, (2) Project Hypotheses, (3) Data Collection Plan and Game Rules, and (4) Collect Baseline Data.

2.1 Project Challenges

The main challenge of conducting this STEM project is to facilitate the team communication. Each player needs to be familiar with each member's position and function. They need to split the work on both offense and defense. Guard needs to pass the ball to Forward. Forward needs to find all the opportunities to shoot. Center needs to get rebounds. Altogether, the three players must synchronize their ball and body movement in order to maximize teamwork and to improve efficiency on both offense and defense. Specially designed game rules will help facilitate such team communication during real games.

2.2 Project Hypotheses

There are two main project hypotheses: (1) within the same team, is the position assignment appropriate, and (2) between two teams: which team has the better player at each of the three positions?

The first hypothesis investigates whether or not the coach has assigned each player to the correct position in order to maximize his potential and contribution to the team. The second hypothesis determines, out of the three positions, which player is comparatively better than the one from the other team at the corresponding position.

2.3 Data Collection Plan and Game Rules

Based on the project objectives, the recorder will collect two main player statistics at each position. Guard will be judged by Assists and Goals; Forward will be judged by Goals and Rebounds; Center will be judged by Rebounds and Blocks. An assist is defined by when Guard passes the ball to Forward and Forward can score successfully. Center will not be credited if they score. On the other hand, Guard will not earn any credit if they rebound the ball or block any shots. These special game rules will reinforce each players' fulfillment of their position, by emphasizing better team collaboration during each ball movement. During their play on offense and defense, each player needs to understand their roles and position as well as that of their team members'. Also, during their play on defense, players need to remember the opponent teams' player positions and primary functions within them. For example, each team will try to minimize the chance that their opponent Guard can pass the ball to their Forward and score.

2.4 Collect Baseline Data

Two 3-player teams have been formed and trained on the game rules. The baseline data has been collected as shown in Table 1. The end of the game was determined by the first team to score 10 Goals.

Table 1. Baseline Data

	G-Goal	G-Assist	F-Goal	F-Rebound	C-Rebound	C-Block
Team A	1	2	7	7	11	1
Team B	4	2	6	2	9	3

3. Analysis and Results

After collecting the baseline data, authors conducted two hypothesis tests to verify the team and the players' efficiency.

3.1 1st Hypothesis: Player Position Assignment

As shown in Table 2, the total row has combined the player statistics from both teams. The 2-Proportion hypothesis test has been used to check whether (1) Forwards score better than Guards and (2) Centers can rebound more than Forwards, with combined statistics taken across both teams.

Per Table 2 statistics, based on the results of a 2-Proportion Test, the calculated P-value comparing Guard and Forward is 0.003. The P-value of Forward and Center is 0.002. Since both P-values are smaller than 0.05, we reject the null hypothesis at 95% confidence. As expected, Forward can score better than Guard and Center can Rebound better than Forward.

Table 2. Team combined statistics.

	G-Goal	G-Assist	F-Goal	F-Rebound	C-Rebound	C-Block
Team A	1	2	7	7	11	1
Team B	4	2	6	2	9	3
Total	5	4	13	9	20	4

3.2 2nd Hypothesis: which player is better?

The objective was to compare the number of Goals, Rebounds, and Blocks of players of the same position across the two teams. A 2-proportion test was used to compare the total number of Goals, Rebounds, Assists and Blocks by position player position (Guard, Forward, Center) across Team A and Team B. For each test, the resultant p-value is higher than 0.05. Therefore, we fail to reject the null hypothesis that the proportion difference is zero in the population, with 95% confidence. So, we can say that, based on the p-values, each player position produces equivalent output (Goals, Assists, Rebounds and Blocks) among teams in the population.

Table 3. 2-Proportion Hypothesis Tests

	G-Goal	G-Assist	F-Goal	F-Rebound	C-Rebound	C-Block
Team A	1	2	7	7	11	1
Team B	4	2	6	2	9	3
Team A	6%	50%	39%	24%	38%	25%
Team B	22%	50%	33%	7%	31%	75%
2-Proportions P-Value	0.136	1.000	0.728	0.062	0.783	0.102

4. Additional Analysis and Root Cause Investigation

Even though the 2-proportion tests (Figure 1) showed that there is no statistical difference between players of the same position across the two teams, looking at the data in the sample, there is a clear practical difference (see groups highlighted in green). (22% of goals in Team B is different than 6% in Team A). The current data-collection method cannot assess the position assignment. As such, all player statistics categories need to be collected on each player, regardless of which position he is playing. Also, Turnover and Steal criteria should be added.

4.1 Additional Data Collection Plan

All six player statistics will be collected for players in every position. In assessing player position statistics, the green highlighted cells in Figure 2 will be calculated for each position.

Table 4. Player Position Assessment

	Guard	Forward	Center
Assist			
Goal			
Rebound			
Block			
Steal			
Turn Over			

Table 5. Player Position Data

Team A	Guard	Center	Forward	Team A
Point	1	2	5	8
Rebound	5	19	9	33
Block	0	0	0	0
Assist	2	2	2	6
Steal	5	5	3	13

Turnover	7	3	1	11
Team B	Guard	Center	Forward	Team B
Point	2	5	3	10
Rebound	1	7	8	17
Block	2	0	0	2
Assist	0	0	0	0
Steal	5	6	6	17
Turnover	1	5	2	8

4.2. Team Comparison

Statistically, Team A has better performance for Rebounds and Assists. The P-values for both Rebounds and Assists are less than 0.05, which means that there is an observed difference in the population. Statistically, Team A and Team B are equivalent when it comes to Points, Blocks, Steals, and Turnovers (where the P-values for these group comparisons are greater than 0.05).

Table 6. P-Value Comparison of Rebound and Assist Ability

	Team A	Team B	P-Value
Point	8	10	0.502
Rebound	33	17	0.001
Block	0	2	0.333
Assist	0	3	0.002
Steal	13	17	0.297
Turnover	11	8	0.324

4.3 Hypothesis #1 Analysis

Position assignment for Points and Rebounds is appropriate. Center has more rebounds than Forward (26 vs. 17 rebounds), and Forward scores more than Guard (8 vs. 3 points), and these differences are statistically significant ($p < 0.05$). Note: P-Values for Steals and Turnovers (Forward vs. Guard) are greater than 0.05, indicating no significant difference for these.

Table 7: 2-Proportions Test

	Guard	Center	Forward	P-Value
Point	3	7	8	0.171
Rebound	7	26	17	0.000
Assist	2	0	0	-0.333
Steal	10	11	9	0.784
Turnover	8	8	3	0.074

4.4 Hypothesis #2 Analysis

It is hard to tell which team player is better in all three positions (p value > 0.05). The reason it is hard to tell is because for each player position (across all the statistics: Points, Rebounds, Blocks, Assists, Steals, Turnovers on Team A vs Team B) there are not more than one group comparisons where statistical significance is achieved and not more than 3 group comparisons where practical significance is achieved.

Table 8: Guard Comparison

Guard	Team A	Team B	P-Value
Point	1	2	0.544
Rebound	5	2	0.236
Block	0	2	0.333
Assist	2	0	0.083
Steal	5	5	1.000
Turnover	7	1	0.010

Table 9: Forward Comparison

Guard	Team A	Team B	P-Value
Point	5	3	0.418
Rebound	9	8	0.790
Block	0	0	1.000
Assist	2	0	0.083
Steal	3	6	0.273
Turnover	1	2	0.546

Table 10: Center Comparison

Guard	Team A	Team B	P-Value
Point	2	5	0.196
Rebound	19	7	0.004
Block	0	0	1.000
Assist	2	0	0.083
Steal	5	6	0.738
Turnover	3	5	0.422

4.5 MVP Analysis

The Center on Team A scored the highest overall points (“Sum by Category” = Points, Rebounds, Blocks, Assists, Steals). Also, he had the highest rebounds among all players across teams by a significant margin (19 total), and fewer turnovers than his counterpart on Team B (3 vs 5, Table 10). Therefore, Center on Team A was named MVP!

Table 11: MVP Data

Overall sum	Team A	Position →			Sum by Category
		Guard	Center	Forward	
	Point	1	2	5	8
	Rebound	5	19	9	33
	Block	0	0	0	0
	Assist	2	2	2	6
	Steal	5	5	3	13
	Turnover	7	3	1	11
	MVP	20	31	20	71

Overall sum	Team B	Position →			Sum by Category
		Guard	Center	Forward	
	Point	2	5	3	10
	Rebound	2	7	8	17
	Block	2	0	0	2
	Assist	0	0	0	0
	Steal	5	6	6	17
	Turnover	1	5	2	8
	MVP	12	23	19	54

5. Conclusions and Future Work

The Authors successfully designed a special Basketball Game position model which can facilitate team communication. The baseline analysis sample size was too small and was missing critical sample statistics for Steals and Turnovers. The improved analysis can facilitate team communication on both Offense and Defense. The use of statistical analysis to make inferences on position and team performance in the population at large can help each team study their performance pattern and to study each player's' strengths/ weaknesses relative to their position & across teams.

Future efforts will be focused on studying a larger sample size in improve decision accuracy. For example, a future study could compare three or more teams. This way, there will be a higher chance of inferring the correct conclusion of a true performance difference (or lack thereof) between players of the same position & across positions across teams.

Acknowledgements

Authors would like to acknowledge Dr. Charles Chen for his mentorship and project guidance.

References

- Dean Oliver, *Basketball on Paper: Rules and Tools for Performance Analysis*, Potomac Books Inc., 2004.
- De Rose Jr., Dante, *Statistical Analysis of Basketball Performance Indicators According to Home/Away Games and Winning and Losing Teams*, *Journal of Human Movement Studies*, v.47, pp.327-336, 2004.
- John Hollinger, *Pro Basketball Forecast: 2005-2006*, Potomac Books Inc., 2005.
- Goldsberry, K., *CourtVision: New Visual and Spatial Analytics for the NBA*, MIT Sloan Sports Analytics Conference 2012, Boston, United States, 2012.
- S. Trninic and D. Dizdar: *Performance Evaluation Criteria*, *Coll. Antropol.* 24 (2000) 1: 217–234
- SAS Institute Inc. 2017. *JMP® 13 Basic Analysis*. Cary, NC: SAS Institute Inc. Ch. 7 (p. 223, 233); *Two Sample Test for Proportions*.
- © 2013 Minitab Inc. *Minitab Help Directory* Minitab v.16, Stat > Basic Statistics > 2 Proportions, *Test of two binomial proportions*.

Biographies

Anna Wu is a junior at Mission San Jose High School and is currently taking AP Statistics. Anna is familiar with IBM SPSS/Modeler software and is an active user of SAS JMP software for applications in Biology/ Chemistry.

Mason Chen is a certified Six Sigma Black Belt enrolled in Stanford's OHS Program. He has also certified IBM SPSS Statistics and /Modeler Data Analyst and Data Miner, and familiar with SPSS/Modeler, JMP, and Minitab Statistics Software.