Evaluating the Impact of Contextual Factors on Shot Success in Basketball

Group 5
Authors: Riham Abdu, Max Fleischer, Gregory Lederer, Rachel Secan, Robert Trenkamp

1. Introduction to Data

This paper will explore how different levels of defensive pressure affect outcomes in basketball to identify the most effective defensive strategy for reducing shot success. In our analysis, we specifically focused on shot contention, defined as situations where the defender applies pressure to the shooter either through their proximity or by raising their hand to obstruct the shot. Advances in basketball analytics have highlighted the importance of understanding defensive techniques for a team's success. We were particularly inspired by *Lucy et al.'s "How to Get an Open Shot,[1]"* which analyzed the success of different defensive actions using spatiotemporal tracking data in the three seconds before a three-point shot. Our team sought to further their analysis by broadening the study of shot contention to more areas of the court, without utilizing spaciotemporal tracking data. We hypothesized that defender distance and whether they raised their hand to block the shot would be the most influential factors in shot outcome. These factors were intuitively hypothesized because they result in immediate pressure or deflection of the ball. Defender distance was found to be highly significant in three-point shots in the study discussed above. Below is a preview of the dataset. The descriptions of variables can be found beneath the table.

Shot Type	Defender Distance	Hand	Result	Location	Position
Pull Up	Tight	1	Make	Mid	Big
Floater	Moderate	0	Make	Paint	Guard
Floater	Moderate	0	Miss	Mid	Guard
Catch & Shoot	Tight	1	Make	Three	Guard
Catch & Shoot	Tight	1	Miss	Three	Big
Catch & Shoot	Loose	0	Make	Three	Big
Pull Up	Tight	1	Make	Mid	Big
Pull Up	Tight	1	Make	Paint	Big
Pull Up	Tight	1	Make	Paint	Big
Catch & Shoot	Moderate	1	Make	Three	Guard
Catch & Shoot	Tight	1	Miss	Three	Big
Catch & Shoot	Loose	0	Miss	Three	Guard
Pull Up	Moderate	1	Make	Three	Guard
Pull Up	Loose	0	Make	Three	Big
Catch & Shoot	Tight	1	Miss	Three	Big
Pull Up	Loose	0	Make	Paint	Guard
Catch & Shoot	Moderate	1	Miss	Three	Guard
Floater	Tight	1	Miss	Paint	Big
Pull Up	Tight	1	Miss	Paint	Big

Our dataset was constructed manually by collecting shot-level data from five NBA games in the 2025 season. Specifically, we watched the following five games: Knicks vs. Bulls (11/2/25), Clippers vs.

[1] Lucey, P., Bialkowski, A., Carr, P., Yue, Y., & Matthews, I. (2014). *How to get an open shot: Analyzing team movement in basketball using tracking data*. MIT Sloan Sports Analytics Conference.

Heat (11/3/25), Knicks vs. Bucks (11/1/25), Suns vs. Spurs (11/2/25), and Mavericks vs. Pistons (11/4/25). These games were accessed through a group member's work repository of uninterrupted games. Each member of the group watched an assigned game and documented at least the first 50 contested shots that they observed. Before we began formal data collection, we watched a short clip to review real examples of the observations we would be recording, such as defender distance (tight, moderate, loose) and shot type, so everyone understood and applied the same criteria.

All observations were collected from the first half of the game, where player intensity and fatigue levels are often lower for all players. The first observations were analyzed to assess the best defensive strategy at a more even player level. Each row in the data represents a single shot attempt, regardless of the outcome or the shot's location. Shots that resulted in a foul were not documented in our data unless the shot was successful. In total, our dataset has 253 observations, highlighting the following six metrics:

- *Shot Type*: A categorical variable identifying the offensive player's type of shot.
    - Pull-Up: A dribbling player stops and immediately jumps and shoots
    - Floater: A shot taken while driving to the basket, typically with a high arc or the shooter off balance to avoid defenders
    - Catch & Shoot: A shot taken as soon as it is received from a pass, without dribbling the ball
- *Defender Distance*: A categorical variable quantifying the defender's distance from the shooter at the time of the shot attempt.
    - Tight: Defender within two feet of the shooter
    - Moderate: Defender between two and four feet from the shooter
    - Loose: Defender four or more feet away from the shooter

(ESPN Staff)
Tight: Defender is within two feet of the shooter

("Basketball Shot")
Moderate: Defender is between two and four feet from the shooter

(O'Donnell)
Loose: Defender is four or more feet away from the shooter

- *Hand*: A binary variable indicating whether the defender raised their hand during the shot attempt
    - 0: No hand was raised
    - 1: Hand was raised
- *Result*: A binary variable indicating whether the shot was successful
    - 0: The shot was unsuccessful

    - 1: The shot was successful
- *Location*: A categorical variable noting where the shot was attempted on the court
    - Paint: The rectangular shaded area between the baseline and the free-throw line
    - Mid: The area between the paint and the three-point line
    - Three: Any area beyond the three-point line
- *Position*: A categorical variable identifying the position of the shooter
    - Big: A taller player who normally plays near the basket, such as a center or power forward
    - Guard: A smaller player who primarily plays on the perimeter, such as a guard or a small forward

2. Summary of the Data

Table 1

Variable	Category	# Count	# Percent (%)
Shot Type	Catch & Shoot	120	47.4
Shot Type	Floater	60	23.7
Shot Type	Pull Up	73	28.9
Defender Distance	Loose	43	17
Defender Distance	Moderate	79	31.2
Defender Distance	Tight	131	51.8
Hand	Down	61	24.1
Hand	Up	192	75.9
Result	Miss	133	52.6
Result	Make	120	47.4
Location	Mid	48	19
Location	Paint	62	24.5
Location	Three	143	56.5
Position	Big	80	31.6
Position	Guard	173	68.4

Table 1 gives a snapshot of the categorical variables in our dataset. Catch & Shoot attempts showed up the most, and defenders were usually recorded as being at a Tight distance from the shooter. In most of the contests, defender(s) had a hand up. Additionally, missed shots occurred slightly more often than shots made, and most shots came from beyond the three-point line. We also concluded that guards took more shots overall than bigs.

Figure 1

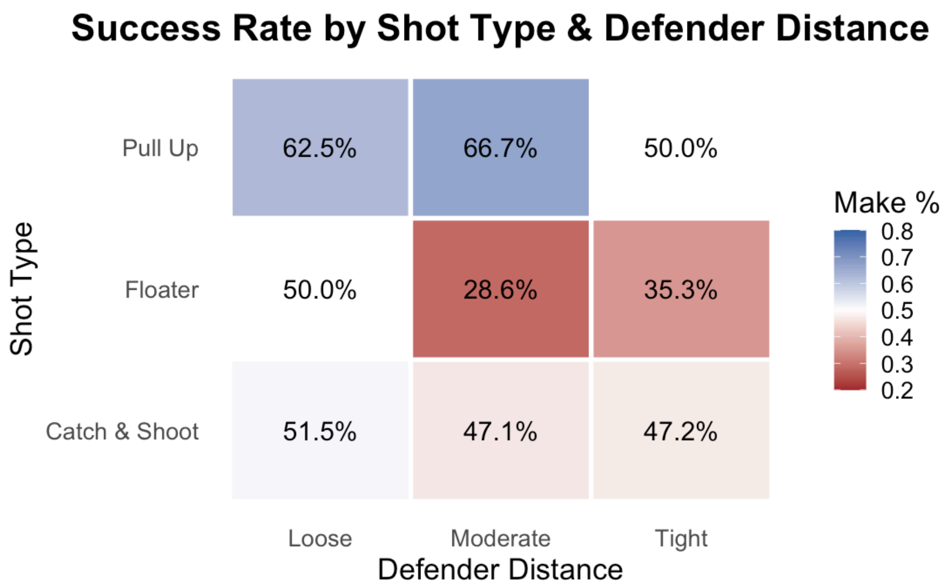


Figure 1 depicts shooting success across two key categorical variables: shot type and defender distance. The heatmap shows the make percentage for each combination.
Pull-up shots performed the best overall. Players made about 62.5% when the defender distance was loose, and 66.7% when it was moderate, before dropping to 50% under tight defense. Floaters were the most affected by defensive pressure. Their success rate dropped from 50% when defender distance was loose, to 28.6% at moderate distance, and then rose slightly to 35.3% when defender distance was tight. Catch & shoot shots were the most consistent, staying between 47–51% regardless of defender distance. Overall, the figure shows that each shot type responds differently to defender distance: pull-ups still perform well with moderate pressure, floaters decline the most, and catch & shoot attempts show little fluctuation.

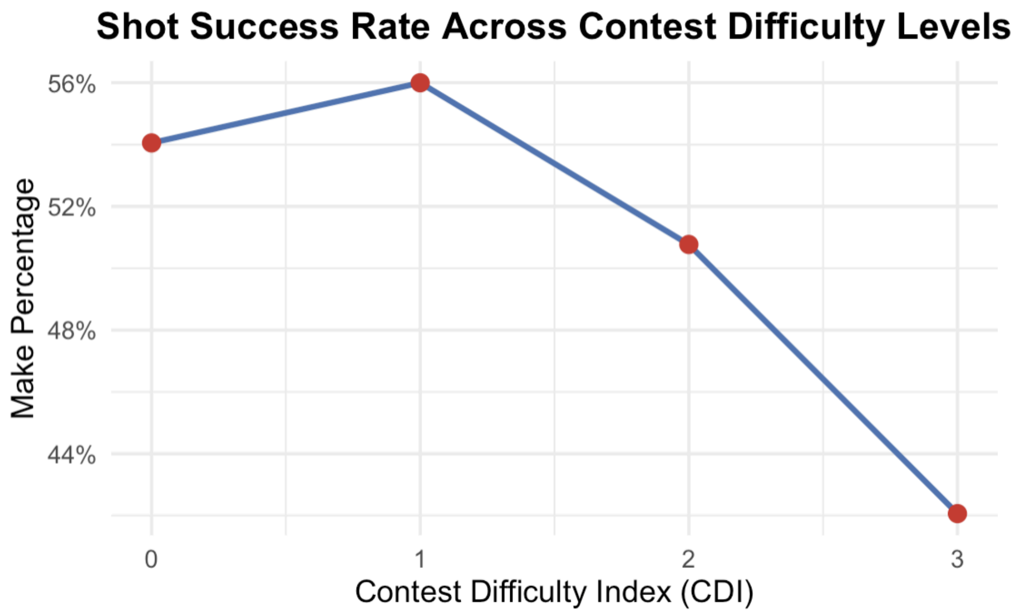
Given these distinct responses to defender pressure, we created a single metric that captures both distance and hand contesting. The Contest Difficulty Index (CDI) is defined below as:

$$\text{CDI} = \text{Defender Distance} + \text{Hand}$$

Defender distance is defined as *Loose = 0, Moderate = 1, Tight = 2*. Hand position is defined as *Hand Up = 1* and *Hand Down = 0.* This metric ranges from 0 (easiest contest) to 3 (hardest contest).

At CDI = 0, players make 54% of shots.
At CDI = 1, success peaks at 56%.
At CDI = 2, success falls to about 51%.
At CDI = 3, success drops to 43%.

Figure 2



The pattern in Figure 2 suggests that combining tight defender distance and a defender's hand-up reduces shot success. While the effect size is modest, the decline across CDI levels shows that defensive pressure plays a role in shot success.

Our broader goal was to understand how different shot contexts and defensive pressures affect the value of offensive decisions. Instead of only asking whether a shot goes in, we examined which combinations of shot type, location, and defensive contest produce the highest value. To do this, we focused on expected points per shot (EPPS) rather than only on make percentage. For each shot, we defined *points_scored* as 2 or 3 (depending on whether the shot qualified as a three-pointer) if the attempt was made and 0 otherwise. Expected points per shot (EPPS) in any group of attempts is then the average of *points scored* in the given group. The following is the formula for EPPS:

$$Expected\ Points\ per\ Shot\ =\ \frac{Total\ Points\ Scored\ in\ the\ Group}{Total\ Shots}$$

Essentially, we are calculating the expected value of a shot type. We are answering the question: "On average, how many points are earned every time we take this type of shot?" The initial interest was in how the value changed based on where the shot was taken. The table below reports the expected value of each shot based on whether it was in the paint, a mid-range (between the paint and the three-point line), or a three-point shot:

Table 2

Location	Attempts	Makes	Field Goal Percentage	Expected Points per Shot
Paint	62	22	0.452	0.903
Mid	48	28	0.458	0.917
Three	143	70	0.49	1.47

The table suggests that three-point shots are significantly more valuable than other shots. This conclusion largely aligns with the analytical trend in basketball over the last decade - shooting three-pointers is more beneficial than two-pointers. Additionally, we evaluated which shot types are the most successful through examining the expected points by category of shot type:

Table 3

Location	Shot Type	Attempts	Makes	Field Goal Percentage	Expected Points Per Shot
Paint	Catch & Shoot	4	3	0.75	1.5
Paint	Floater	41	18	0.439	0.878
Paint	Pull Up	17	7	0.412	0.824
Mid	Pull Up	27	16	0.593	1.19
Mid	Catch & Shoot	7	7	0.571	1.14
Mid	Floater	2	2	0.143	0.286
Three	Pull Up	18	18	0.621	1.86
Three	Catch & Shoot	109	109	0.268	1.4
Three	Floater	5	1	0.2	0.6

At each shot location, an interesting result is uncovered. Floaters are less valuable from mid-range and behind the three-point line, while catch & shoot and pull-up shots are generally more valuable.

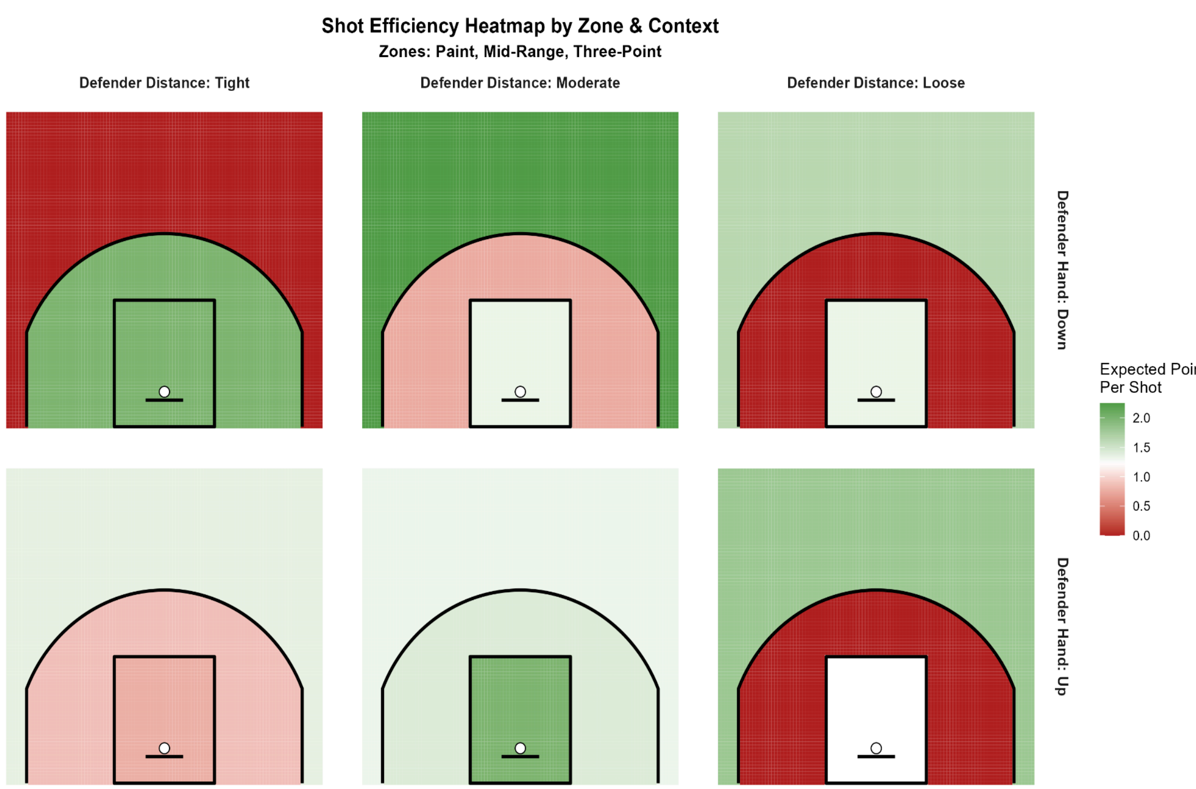
Finally, we needed to evaluate the impact of defense on expected value. One of the most fundamental aspects of defense is having a hand up while guarding a player with the ball. Our data supports this fundamental, as shooters were always better when the defender had their hand down at each defender distance:

Table 4

Defender Distance	Hand	Attempts	Makes	Field Goal Percentage	Expected Points per Shot
Tight	Down	5	4	0.8	1.6
Tight	Up	126	53	0.421	1
Moderate	Down	19	11	0.579	1.47
Moderate	Up	60	29	0.483	1.35
Loose	Down	37	20	0.541	1.57
Loose	Up	6	3	0.5	1.5

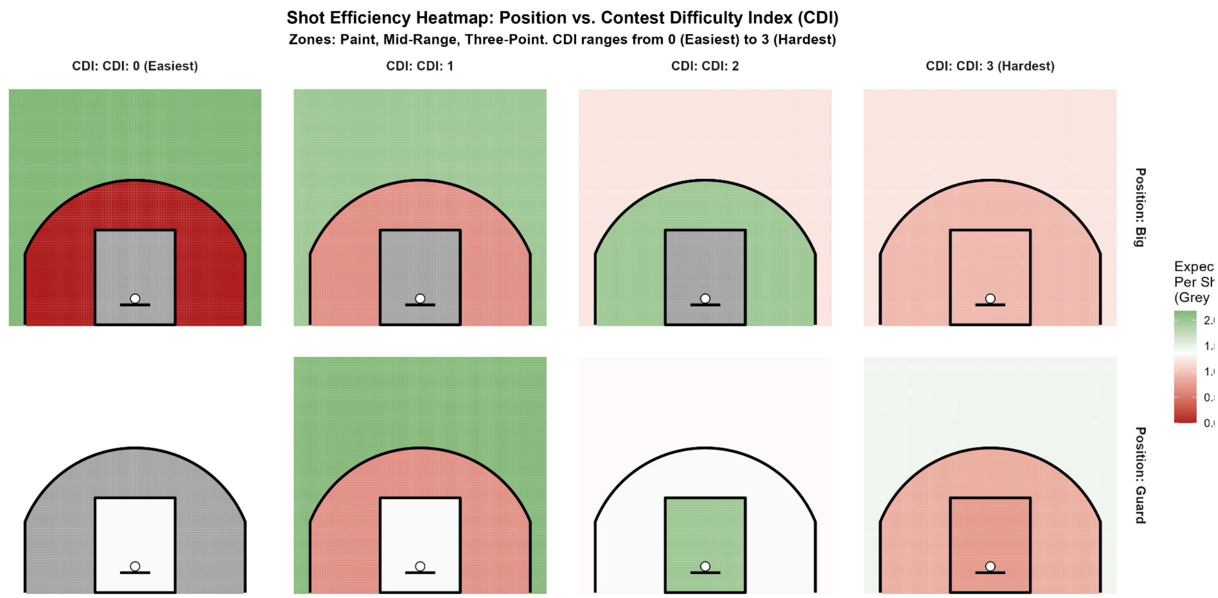
The quality of the defense is also dependent on where the shot is being taken on the floor. Three-point shots are always above average in value when the defender is in loose or moderate coverage. When the defender is playing the best defense (tight distance with a hand up), it is hard for a team to shoot a valuable shot inside the three-point arc:

Figure 3



Using our contest difficulty index, we can examine how each type of player (Big, guard) does against each quality of defense:

Figure 4



As shown in the graph, a player labeled as "big" never has a valuable shot when the defenders are in "perfect" coverage (CDI = 3). It is also worth noting that three-point shots are extremely valuable when the contest difficulty is a 0 or a 1. Again, this figure supports the trend in basketball - players should be looking for open three-point shots.

3. Insights from the Data

Insight 1 - Basic Context Alone Does Not Predict Shot Success

A core question in our project was which defensive actions reduce shot success. To test whether broad contextual factors meaningfully influence outcomes, we fit a logistic regression model predicting made versus missed shots using shot type, defender distance, defender hand, shot location, and position. As shown in Table 5, none of these variables reached standard significance levels (all p-values > 0.05). The strongest directional effect, Hand = 1 (defender's hand up), had a marginal p-value of 0.099, indicating that although hand-up contests trend toward reducing shot success, this effect is not consistently measurable with our coarse categories. The model's test accuracy was 48.7%, nearly identical to always guessing the majority class, reinforcing that these broad descriptors alone do not reliably predict shots made.

Table 5

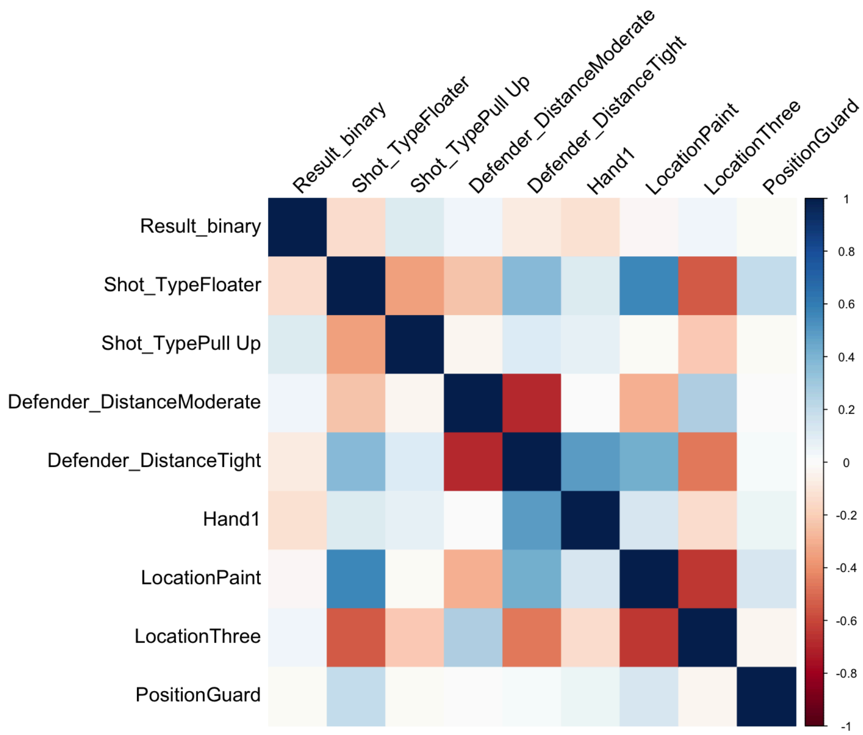
Result_binary ~ Shot_Type + Defender_Distance + Hand + Location + Position

	Estimate	P-Value
Intercept	0.529	0.418
Shot_TypeFloater	-0.644	0.267
Shot_TypePull Up	0.512	0.229
Defender_DistanceModerate	-0.046	0.936
Defender_DistanceTight	0.416	0.547
Hand1	-0.862	0.099
LocationPaint	0.117	0.816
LocationThree	-0.225	0.644
PositionGuard	-0.104	0.769

Shot_TypeCatch & Shoot, Defender_DistanceLight, LocationMid, & PositionBig are considered the "baseline" as all coefficients are compared to the baseline.

The correlation heatmap in Figure 5 supports this same conclusion. All correlations between *Result_binary* and our predictors fall between –0.14 and +0.11, showing only very weak linear relationships. For example, floaters have a slightly negative correlation with makes, while pull-ups show a slight positive correlation, but neither effect is large enough to matter. Similarly, defender distance categories (Loose, Moderate, Tight) exhibit correlations near zero, indicating that the simple labels we used do not capture the true complexity of a defensive contest.

Figure 5



Together, these results do not mean "nothing matters." Instead, they reveal a more important insight aligned with our original purpose:

Broad shot-context variables are too coarse to capture the defensive actions that actually impact shot success.

In other words, basic categorical labels: "tight," "moderate," "hand up," "pull-up," etc., do not explain outcomes. To find meaningful effects, shot context must be measured with greater detail and precision (e.g., exact contest angle, timing, shot difficulty). This directly motivated the development of our CDI and CDI2 metrics in Insight #2, where higher-resolution defensive intensity finally produced significant, actionable results.

Insight 2 - CDI Impact on Shot Outcome

To test whether CDI actually captures shot difficulty, we analyzed how well it predicts shot outcomes. First, we made a summary of the make percentages across CDI levels, and found that success rates were similar for CDI ratings of 0-2, but much lower when CDI had a rating of 3. This suggests that only the shots that are the hardest contests consistently disrupt a shooter.

We then fit a logistic regression of *Result_binary ~ CDI*, which produced a negative coefficient of about -0.19 and an odds ratio of about 0.83, which means each one-point increase in CDI is associated with roughly a 19% decrease in the odds of making a shot. The p-value for CDI was 0.110 >0.05, indicating the effect is not significant but suggestive.

The original CDI used only defensive variables to test the shot contest: defender distance and hand. We discussed whether incorporating shot types could sharpen our metric. The type of shot you take is not always a choice by the offense, but an outcome of the interaction between offense and defense. When the defense contains the drive and stays attached to the perimeter, the offensive player with the ball is more likely to take a floater or a rushed shot than a balanced pull-up. Due to this, we thought it was appropriate to include the shot type in the equation:

$$CDI_2 = Defender\_Distance + Hand + Shot\_Type^2$$

where *Loose = 0, Moderate = 1, Tight = 2 and Hand Down = 0, Hand Up = 1 and Pull Up = 0, Catch & Shoot = 1, Floater = 2.*
This metric now ranges from 0 (easiest contest) to 7 (hardest contest).

Figure 6

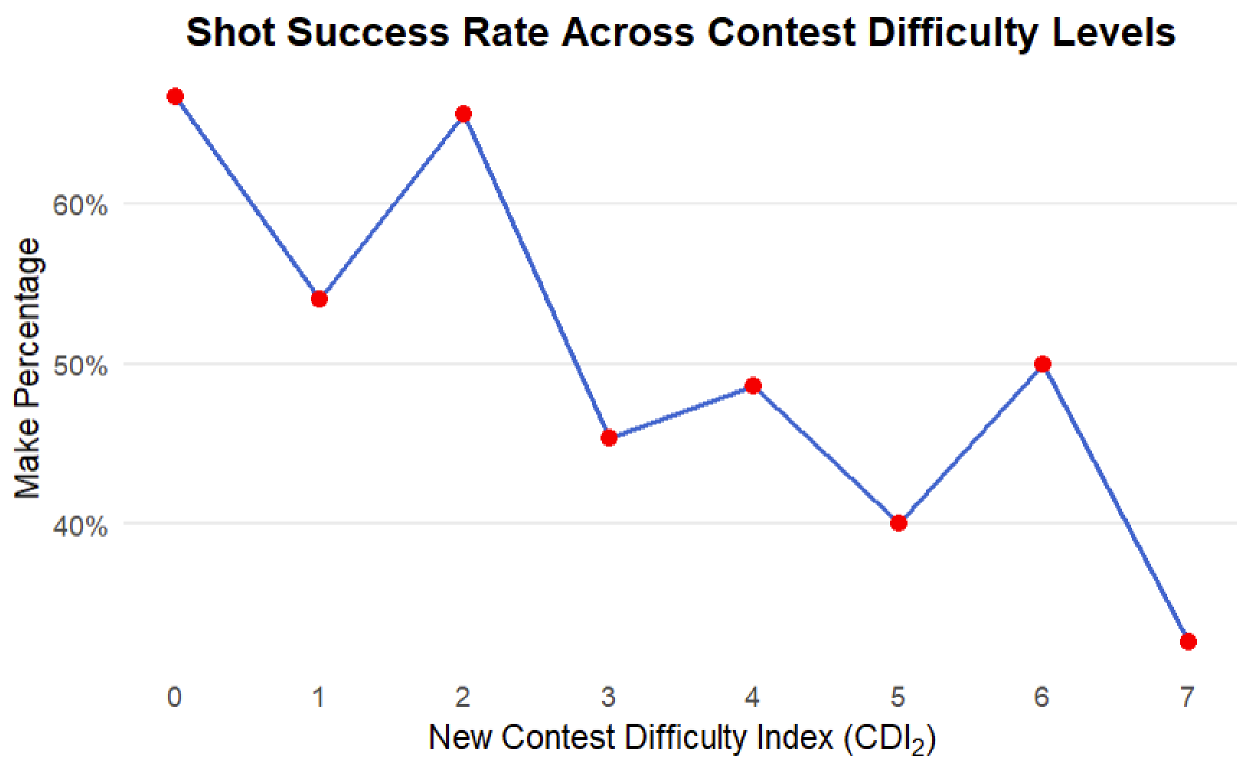


Figure 6 shows a similar, but stronger, pattern than Figure 2. As contests become tougher, the shooting success decreases. The line in Figure 6 is more jagged because the new CDI now spans from 0 to 7 and combines more information, but the overall pattern is that success rates are highest at low CDI and lowest at a high CDI.

In Figure 1, we find that pull-up shots are the most reliable option across defender distances, roughly 62%-67% under loose and moderate pressure, and remain around 50% under tight defense, whereas catch-and-shoot shots are around 47%-52%, and floater shots drop to 28-35% under tight coverage. We leveled the *Shot_type* according to this: Pull Up = 0, representing the easiest difficulty; Catch & Shoot = 1, representing moderate difficulty; and Floater = 2, representing the most difficult.

We squared the *Shot_Type* variable to reflect the nonlinear increase in shot difficulty, which aligns with intuition and our data. By squaring the variable so that pull-up shots remain at zero and catch-and-shoot and floaters receive higher values, with floaters receiving the largest penalty after squaring, the index becomes statistically significant as a predictor of makes. $P = 0.009 < 0.01$.

This improvement supports the idea that difficulty should be modeled with consideration for the player and context, rather than a linear sum of defensive variables. Our original CDI, which only added defender distance and hand position, resulted in the correct direction of the effect, but it failed to reach significance, which means the raw content data alone was not enough to determine makes from misses. Through incorporating shot type and emphasizing the non-linear impact through squaring the term, CDI captured the effect, as those shots occur when the defense has already mostly won the possession. The statistically significant relationship between CDI and shot outcome indicates that this revised index better measures true contest quality.

- **Going Forward**

Going forward, the biggest limitation of our project was that many of the variables we collected were too broad and also subjective, since they depended on the observer. Due to this shortcoming, some real patterns were likely missed. For example, defender distance was only labeled as loose, moderate, or tight. The "hand up" variable did not describe how the defender contested the shot - whether it was a late hand, a soft contest, or a strong challenge. Our shot location categories (mid-range, paint, three) were also general and did not capture specific spots on the floor. We also did not record the identity of the shooter, which is significant because different players have varying skill levels. Without knowing who took the shot, it's difficult to determine whether the outcome was driven by the player or by the shot's context. The methodology did not control for differences in shooter skill and was biased by only watching the first halves. Additionally, there was an imbalance in shot type and location. Some observations were also dependent, as multiple shots were taken by the same players.

If we had more time, we would add several important game-context variables that we did not track. These include time of game, score difference, shot clock, number of defenders involved, and whether the contest came from in front, the side, or behind. We would also track whether the shot came off a pass or off the dribble, and how many dribbles the player took before shooting. All of these factors play a big role in shot difficulty, and they would help to understand shot success more than our basic categories.

On the modeling side, we could have tried additional approaches like adding interaction terms, testing different types of models, or using better evaluation methods like cross-validation instead of one train/test split. We would also collect a larger and more varied sample of shots. Since defenders choose how tightly to guard a shooter based on unobserved factors, like shooter ability, variables like defender distance and hand up are probably correlated with omitted variables that affect the shot result. This limits our ability to make strong causal claims. Due to these endogenous factors, to estimate a true causal effect, we could use an Instrumental-Variable (IV) approach, which could isolate a causal effect of contest on shot success. Overall, collecting more detailed variables and improving our modeling process would give us clearer and more useful insights for coaches, players, managers, and other stakeholders.

References

Lucey, P., Bialkowski, A., Carr, P., Yue, Y., & Matthews, I. (2014, April 30). *"How to Get an Open Shot": Analyzing team movement in basketball using tracking data.* Disney Research. Retrieved from https://la.disneyresearch.com/publication/how-to-get-an-open-shot-analyzing-team-movement-in-basketball-using-tracking-data/

O'Donnell, R. (2020, October 10). *Danny Green missed a wide open three that could have won the Lakers the 2020 NBA Finals*. SBNATION. Retrieved December 2, 2025, from https://www.sbnation.com/nba/2020/10/10/21510378/danny-green-missed-three-video-lakers-heat-nba-finals-game-5

Thorpe, D. (2017, March 22). *Shot blocking 101 and the art of the contest*. ESPN. Retrieved December 3, 2025, from https://www.espn.com/blog/truehoop/post/_/id/74870/shot-blocking-101-and-the-art-of-the-contest

20,908 Basketball Shot Stock Photos - Free & Royalty-Free Stock Photos from Dreamstime. (n.d.). Dreamstime.com. Retrieved December 3, 2025, from https://www.dreamstime.com/photos-images/basketball-shot.html