

## How to Stop Stephen Curry?

Wei Gao, Oklahoma State University; Vrushali Walde, Oklahoma State University

### ABSTRACT

Stephen Curry is a professional basketball player for the Golden State Warriors of the National Basketball Association (NBA). With 30.2 points per game for the 2015 -2016 season, he is the top NBA 3-point shooter and seems to be an unstoppable player in NBA. He led the Golden State Warriors to an astonishing 73 win and 9 lose record for 2015-2016. So the focus of this research is how to defend him.

In the 2015-2016 season, among the 73 winning games, his 3- point percentage is 46.6%. However, among the nine losing games, the 3-point percentage dropped to 35.4%. What makes such a huge difference between winning and losing games performance? Some of the questions answered through this research are:

- Does Curry have any shooting pattern?
- Can anybody effectively defend Curry?
- Is help defense effective when guarding Curry at various ranges?
- Will taller and stronger players better defend Curry?

The NBA official website offers statistics, which included almost all the variables needed, such as players' height, weight, speed, total rebounds, stats of different shooting area etc. In addition, Curry's shooting statistics were recorded manually by watching games played during the study time frame. Shooting statistics included how Curry shot (catch-and- shot, lay-up, etc.), distance from the rim, and shot outcome. All the games that Golden State lost and won against the same opponent were included in the data. Logistic regression and decision tree were used for data modeling.

### INTRODUCTION

In 1979, when the 3-point line was introduced for the first time in basketball games, the players were not sure how to score maximum 3-poniter shots. Only 3% of all shots were 3- pointers in that season. The percentage of 3- pointers increased over the next three decades to 22% and was fairly constant for quite a while("The Golden State Warriors Have Revolutionized Basketball" by Ben Cohen). Then, the Golden State Warriors revolutionized the basketball game and much of the credit goes to Stephan Curry. The percentage of 3-pointers surged to 35%(NBA official statistics). This motived us to analysis the performance factor of Stephan Curry. Is it just the hard work of Curry which led his team to such a spectacular performance or is it a team-planned strategy. If it is a strategy, then we wanted to decode it and find a certain playing pattern of Curry. This inspired us to analyze all the associated factors which can influence 3-point shots. Our analysis could help defenders from other teams plan specific strategies to defend Stephan Curry and to perform on par with the Golden State Warriors.

### A. DATA PREPARATION

Data was prepared and analyzed mainly using SAS® Visual Analytics, which is a very powerful data visualization tool. The poster includes a bar chart of the frequency of Curry's shooting type (jump shot, layup, stepback, etc.), a decision tree analyzing how likely will Curry take a jump shot when facing different defender heights, and many other graphics generated by SAS® Visual Analytics. Readers will have a better understanding on how SAS® Visual Analytics works with sports data, and also how visual data exploration builds beautiful charts and models by dragging and clicking.

SAS® Visual Analytics was used for data preparation. The Golden State Warriors played 106 games in 2015-2016 season. They lost to 10 different teams such as the OKC Thunder and the Cleveland Cavaliers. The official site of NBA provides video clips of all the shots in the game. Every single shot of Stephen Curry in 44 different games against the team which they lost was recorded one by one, manually. The recorded data table contained variables related to the shooting pattern of Stephen Curry such as shooting type, makes or misses, shooting range, shot position from 3-point line, game results etc. Various shooting types recorded were jump shot, layup, step back, hook shot, fade away and push shoot. Shooting range described a 2 point shot or a 3 point shot. Shot position from 3- point line was defined as close, one step away or two steps away from 3 -point line and game results were win or lose. Other variables which could have impacted the shooting pattern of Curry were also recorded which were game location (home game in Oakland, CA or away game) and was the shot made with the help defense.( Help defense means when the match-up defense player loses his position, other players can move and help defend Curry. Along with above recorded variables, this table contained information about the player who defended Stephen Curry while shooting.

Another table containing defenders' statistics was joined with the created table. Defenders statistics included height, weight, age, experience, average speed and salary, all of which was available from [www.nba.com](http://www.nba.com). The final data- set was obtained by merging these two tables, which resulted in 17 total variables for analysis. The target variable for analysis was shot outcome (make or miss).



Figure 1. Tables used in data preparation

One more table obtained from [www.nba.com](http://www.nba.com) was of games statistics. It contained performance parameters of Stephan Curry in all the games. This table consisted of field goals made (FGM), field goals attempted (FGA), three-point field goals made (3PM), three-point field goals attempted (3PA), etc. This table helped to capture performance variation of Curry from game to game.

Match Up	WL	MIN	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	STL	BLK	TOV	PF
<a href="#">APR 13, 2016 - GSW vs. MEM</a>	W	30	46	15	24	62.5	10	19	52.6	6	6	100	1	3	4	6	2	0	2	2
<a href="#">APR 10, 2016 - GSW @ SAS</a>	W	35	37	13	22	59.1	4	9	44.4	7	7	100	0	5	5	5	2	0	4	4
<a href="#">APR 09, 2016 - GSW @ MEM</a>	W	34	17	7	22	31.8	3	14	21.4	0	0	0	0	9	9	8	1	0	2	1
<a href="#">APR 07, 2016 - GSW vs. SAS</a>	W	36	27	11	19	57.9	3	7	42.9	2	2	100	0	5	5	9	2	0	3	2
<a href="#">APR 05, 2016 - GSW vs. MIN</a>	L	43	21	7	25	28	4	14	28.6	3	3	100	2	4	6	15	3	0	3	4
<a href="#">APR 03, 2016 - GSW vs. POR</a>	W	35	39	13	21	61.9	9	13	69.2	4	4	100	2	4	6	7	2	1	4	2
<a href="#">APR 01, 2016 - GSW vs. BOS</a>	L	37	29	9	19	47.4	8	14	57.1	3	3	100	0	5	5	6	1	0	9	2
<a href="#">MAR 30, 2016 - GSW @ UTA</a>	W	42	31	11	23	47.8	5	10	50	4	5	80	2	5	7	4	2	0	5	4
<a href="#">MAR 29, 2016 - GSW vs. WAS</a>	W	35	26	9	20	45	6	8	75	2	2	100	0	7	7	7	5	0	5	3
<a href="#">MAR 27, 2016 - GSW vs. PHI</a>	W	33	20	7	16	43.8	2	9	22.2	4	5	80	0	3	3	8	4	0	4	0

Table 1. Snapshot of game log table containing performance statistics of Stephan Curry

DATA STANDARDIZATION

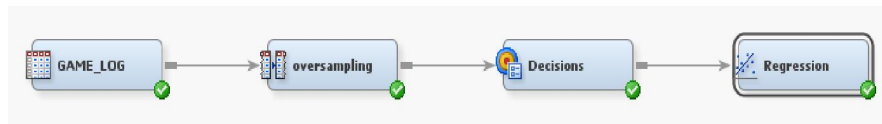
Defenders' data including age, height, weight and experience were not on the same scale. In order to make the data proportional with one another, normalization of all these variables was implemented before data mining. The following equation was used to implement unity based normalization.

score or not	Stdz_age	Stdz_DFG	Stdz_experience	Stdz_height	Stdz_speed	Stdz_weight
1	0.29	0.42	0.21	0.60	0.74	0.54
0	0.29	0.42	0.21	0.60	0.74	0.54
1	0.43	0.49	0.37	0.53	0.50	0.36
0	0.43	0.49	0.37	0.53	0.50	0.36
1	0.43	0.49	0.37	0.53	0.50	0.36
0	0.43	0.49	0.37	0.53	0.50	0.36
1	0.43	0.49	0.37	0.53	0.50	0.36
1	0.43	0.49	0.37	0.53	0.50	0.36
0	0.43	0.49	0.37	0.53	0.50	0.36
1	0.43	0.49	0.37	0.53	0.50	0.36

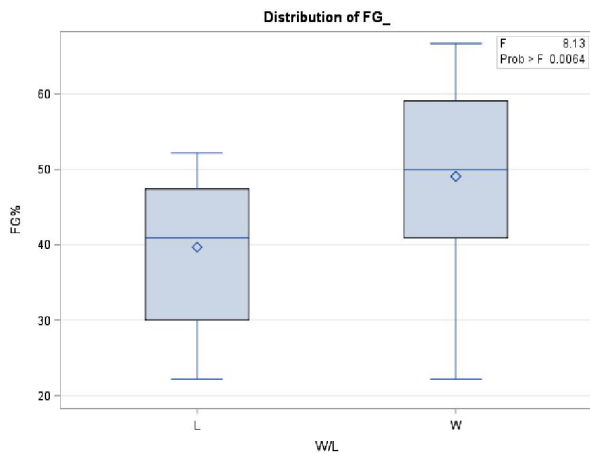
Table 2. Snapshot of standardized table

## B. DATA ANALYSIS

### PERFORMANCE EVALUATION MODEL



Parameter	W_L	DF	Estimate	Chi-Square	p-value
Intercept	W	1	-2.0016	1.76	0.1846
FG_	W	1	0.0658	3.85	0.0497
FTA	W	1	-0.00109	0	0.9896



Display 1. Performance evaluation model

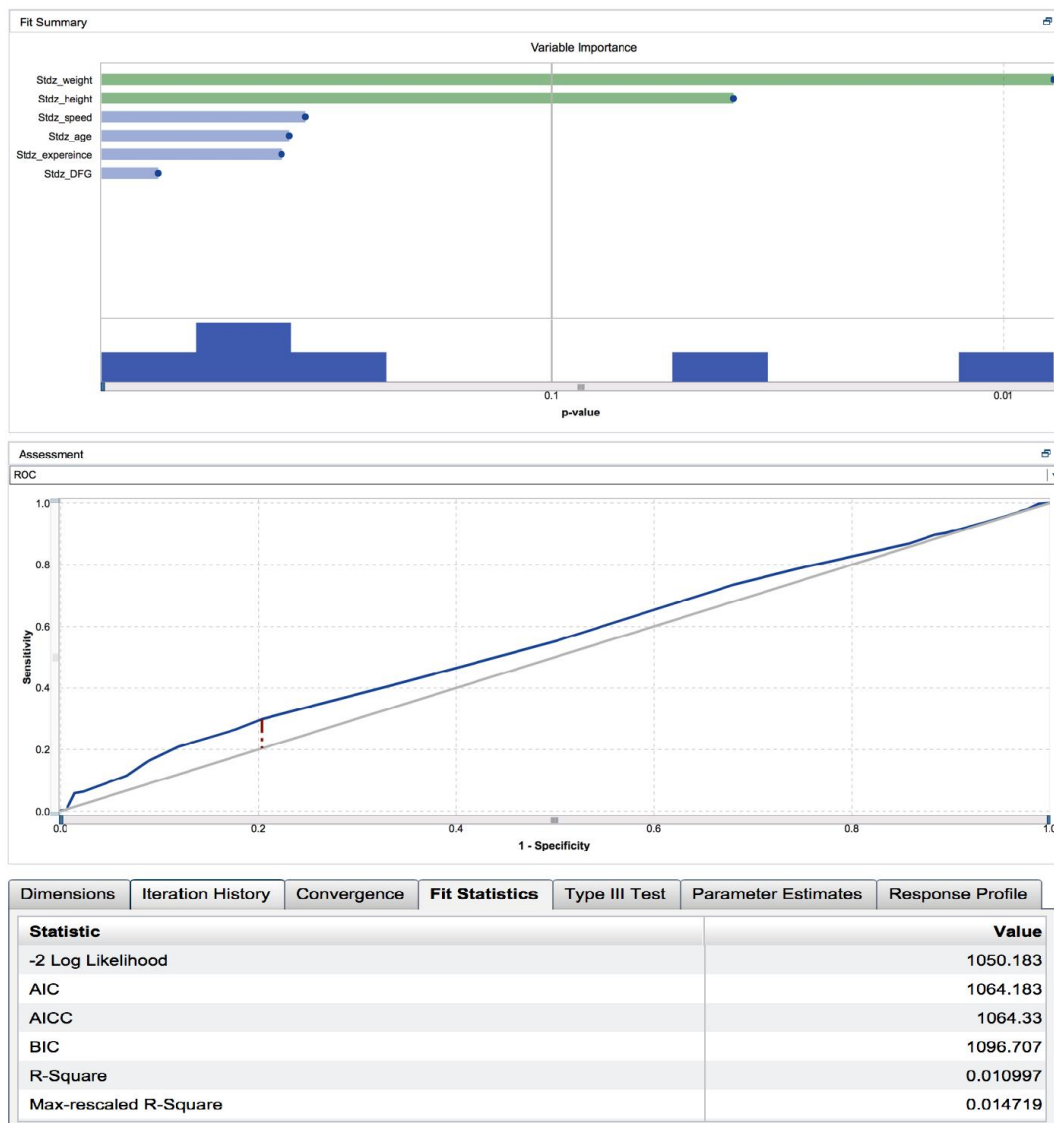
The fundamental analysis was to evaluate Curry's effectiveness in the game. Does Curry's performance make a difference in game result? Table "Game\_log" included variables about all the performance parameters of Curry. Since the Warriors only lost 15 games in the 2015-2016 season, rare event sampling was used to balance the proportion of lost

games and won games. After, the oversampling decision node was added to adjust the posterior probabilities. A regression model with win/lose as a target revealed that FG% was significant in the model. In addition, FG% was higher in the won games compared to lost games.

**DEFENDER CHARACTERISTICS EFFECTS**

In this model, defender's standardized weight, height, average speed, age, experience and DFG% (defended field goal percentage, measures how effective a player defense) were used as input variables. Make/Miss was used as the target variable.

By the logistic regression analysis in SAS VA. Variable stdz\_weight, stdz\_height, are significantly affecting Make/Miss. However, by the fit statistics in the output, this model was not fitting most of the data (R-squared=0.01). Also, the small area under ROC curve indicated the model's poor performance.

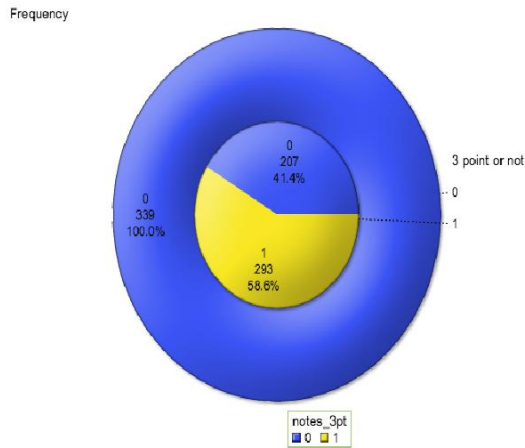
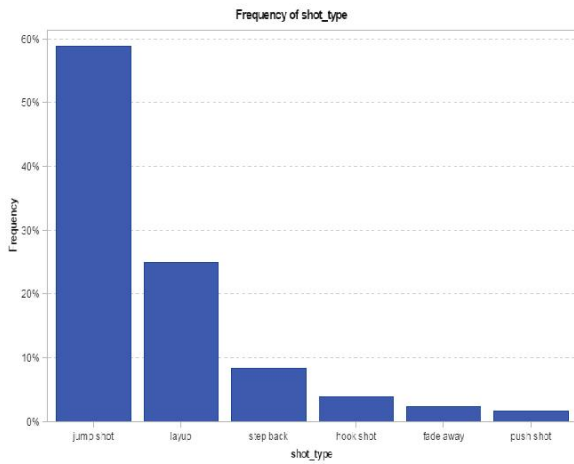


**Display 2. Regression model from SAS Visual Analytics**

Our analysis of this sample data did not find that Stephen Curry could be adequately defended in order to statistically change his scoring outcome. Therefore, no defender or defender characteristics variables had significant effect on Make/Miss. Nevertheless, does Curry have any shooting patten? Will he take different shooting type when facing different defenders?

**SHOOTING PATTERN ANALYSIS**

By the distribution of the shot type (jump shot, layups, step back, etc.), Curry tends to take jump shot more often (nearly 60% of the time) compared to other type of shots. The least attempted shots are the push shot and fade away shot. (Figure 2)



**Figure 2. Distribution of Shot types**

**Figure 3. Frequency of 3-pointers**

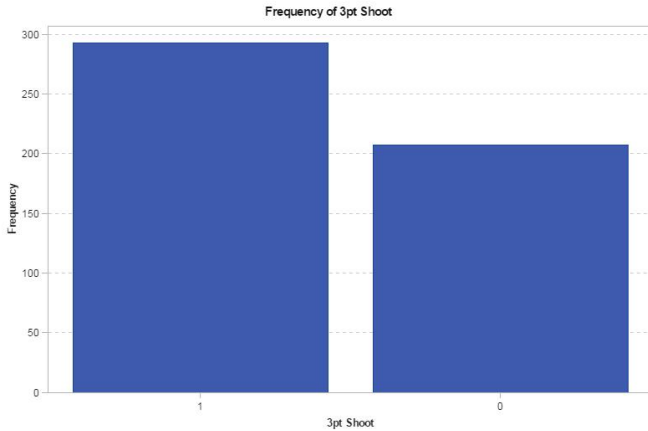
Curry’s shooting pattern was very constant. Figure 3 shows that nearly 60% of his jump shots were 3 pointer shots. And among those attempted 3 point shots, 58.6% was a perfect basket. Figure 4 shows that the ratio of 3 -point shots to a 2- point shot was 2:3, in both home region and away from home region. Therefore, the frequency that Curry attempted a 3- point shoot was very constant.

**Figure 4. 3-point frequency by home or away**

The distinction in Curry’s performance in win and lose game can be easily understood by comparing important statistics in the game. Table 1. shows that his performance was hampered in games that were lost because his FG% and 3-point percentage was significantly less compared to the games Golden State won. This indicates that Curry’s performance could be a main determinate in the Golden State Warrior’s game results.

STATISTICS	LOSE	WIN
AVERAGE POINTS	20.13	29.87
FIELD GOAL %	33.33	50.44
3 POINTS %	33.33	45.25

**Table 3. Stats of win/lose performance indicators**



**Figure 5. Frequency of 3-point**

Curry is very likely to make 3-point shots continuously. Curry made 500 3-point shots in the 44 games analyzed. Approximately, 60% of them were immediately followed by another 3p-point shot in next attacking round.

distance from 3point line ▲		close	one step away	two steps away
notes_3pt ▲	score or not ▲	Frequency	Frequency	Frequency
0	0	38	27	12
	1	46	14	7
1	0	89	36	12
	1	83	32	11

**Table 4. Cross table of back to back shot**

By the crosstab above, Curry’s back to back - point shots have a very high FG% when shooting close to the 3-point line (54.7%). This was much higher than the non-continuous shot (48.2%). When he shot from a very long distance (one step or two steps away from 3-point line), back to back shots had a significant FG % (41%) than discrete shots (34%). In consequence, when Curry made a 3-point shot, defenders may have paid more attention to his back to back shoot as the distance varied from 3-point line to 2 steps away.

Make or Miss	With help defense or not		Total
	0	1	
0	135 (35.45%)	45 (11.9%)	179 (47.35%)
1	103 (27.25%)	96 (25.4%)	199 (52.65%)
Total	237 (62.7%)	141 (37.3%)	378 (100%)

**Table 5. Cross table of help defense by make or miss**

Help defense was not always possible while other defenders are far away from Curry. However, 378 shots where help defense was possible were recorded. Of those, 237 shots were made under help defense and 141 were not. FG%

under help defense was 43.46%, without help defense 68.09%. With a likelihood ratio Chi-square of 21.87 (p-value<0.0001), help defense was significantly affecting Curry's shooting outcome.

**Incremental Response Model:**

Incremental modelling, also known as uplift modelling, true lift modelling, or net modelling is a predictive modelling technique that directly models the incremental impact of a treatment (such as a direct marketing action) on an individual's behavior.

**Discouraged from buying if received offer**



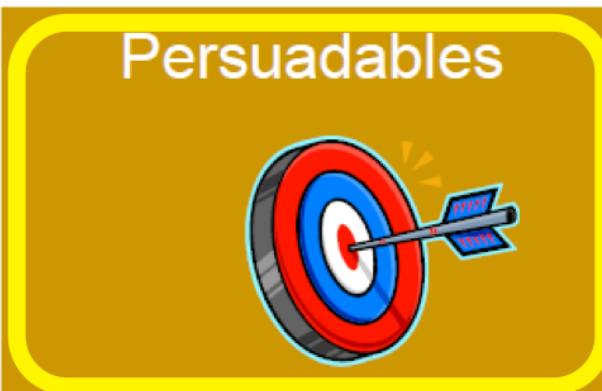
**Will not buy whether received offer or not**



**Sure Things**



**Persuadables**



**Likely to buy whether received offer or not**

**Likely to buy if received offer**

The Persuadables : customers who only respond to the marketing action because they were targeted

The Sure Things : customers who would have responded whether they were targeted or not

The Lost Causes : customers who will not respond irrespective of whether or not they are targeted

The Do Not Disturbs or Sleeping Dogs : customers who are less likely to respond because they were targeted

Incremental response model can not only measure the effectiveness of a marketing action but also to build a predictive model that predicts the incremental response the the market action.

**Apply incremental response model on sport analytics**

For this project:

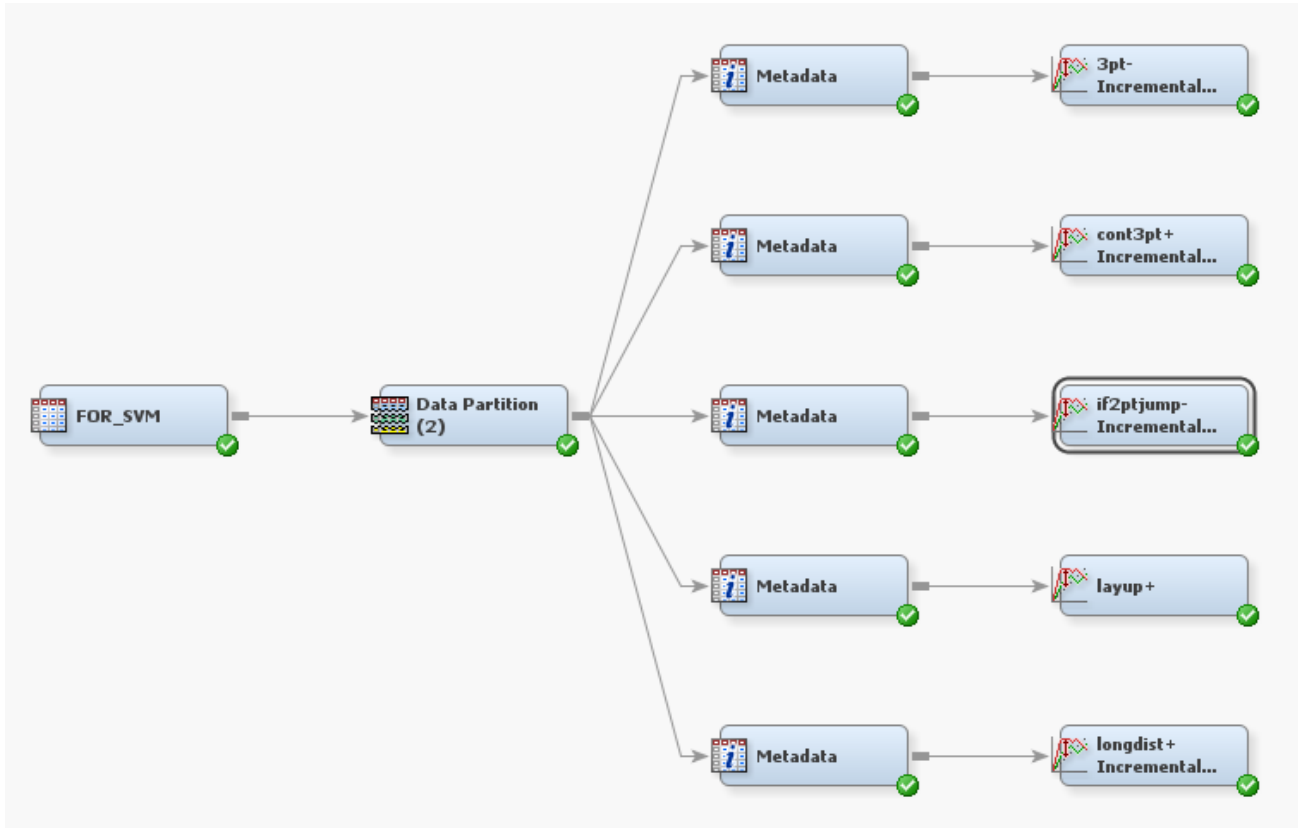
<p><b>The Do Not Disturbs or Sleeping Dogs :</b>  <b>Curry will make the shot only if under defense. (Negative defense effect).</b></p>	<p><b>The Lost Causes :</b>  <b>Curry will miss the shot no matter whether he is under defense or not.</b></p>
<p><b>The Sure Things :</b>  <b>Curry will make the shot no matter whether he is under defense or not.</b></p>	<p><b>The Persuadable :</b>  <b>Curry will miss the shot only if under defense.</b></p>

Now we have two columns of variable. One is defenders' characteristics such as height, weight, age, experience, etc. Another is shooting information includes if 3pt or not, if jump shot or not, etc.



Incremental response node is used to run the model. The node requires a treatment. The variables on the left column is set as treatment. Shot outcome (make of miss) is set as target. Run the node, we get output tables and plots below.

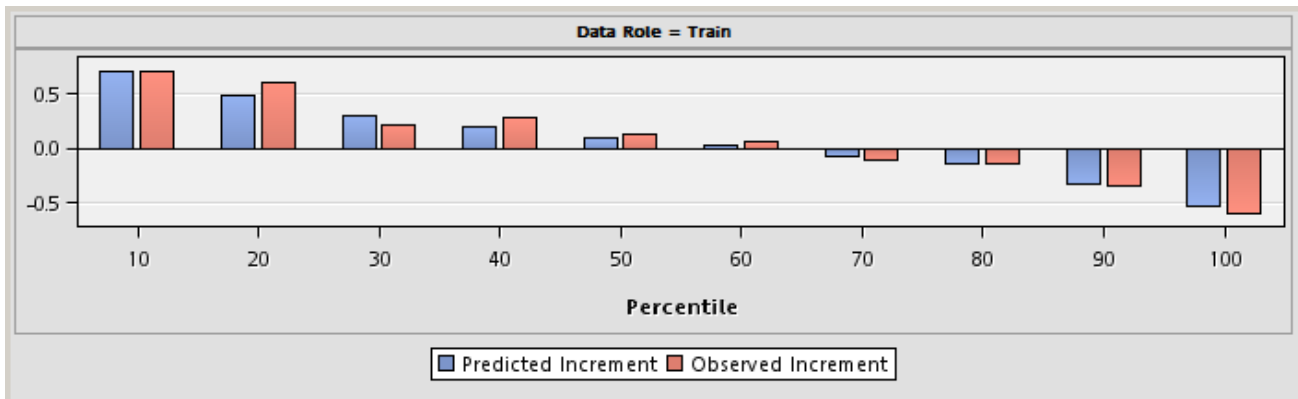




Demo 1: Incremental Response Node

**Example of increment response modeling output interpretation**

In this case, the treatment variable is if\_layups. Target variable is make\_or\_miss, the shot outcome.



Demo 2: Incremental Response Diagnostics

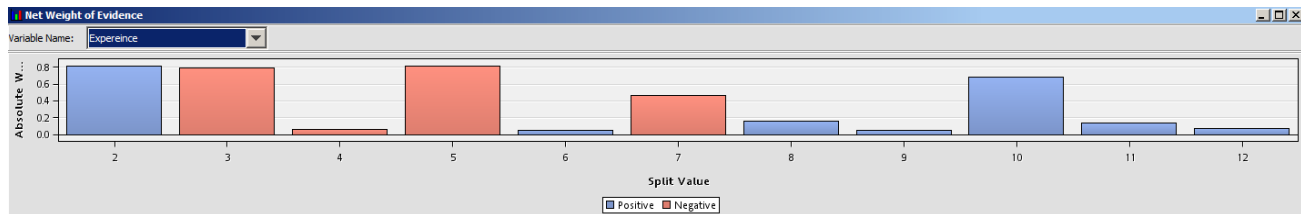
The first and second decile of defenders can effectively decrease Curry's FG%.

Variable Name	Adjusted Net Information Value	Net Information Value	Rank Percentile	Selection	Label
DFG_	15.67906	1573.591	11.11111	Yes	DFG%
with_help_defense...	-10.3857	205.6746	22.22222	Yes	
Defender_name	-275.068	1691.994	33.33333	Yes	
Expeirnce	-1021.37	189.5318	44.44444	Yes	
Height_m_	-1218.37	308.3141	55.55556	No	Height (m)
salary_in_million_	-1823.47	637.0697	66.66667	No	salary(in_million)
Age	-1878.86	323.9569	77.77778	No	
Average_speed_de...	-2072.98	437.7836	88.88889	No	Average_speed_d...
Weight_lb_	-3027.43	306.4445	100	No	Weight(lb)

Demo 3: Select Variable, DFG, help defense, defender name and experience are selected.

In Selected Variables by Adjusted NIV, notice that variable DFG, help defense, defender name and experience are selected.

Then in net weight evidence plot, we can refer that players with 3-5 years' experience are better defend Curry's layup shots.



Demo 4: Net weight of evidence, Experience

By similar research, we found some important factors that will affect Curry's FG%.

2 Point Jump Shot	Average Speed, height, age, experience
3 Point Shot	Some specific players, average speed, salary
Layups	Some specific players, DFG, salary, average speed
Back to back 3 point shot	Help defense, experience, average speed, age
Extra Long Shot	Help defense, experience, DFG, age

### CONCLUSION

- Help defense was very effective on all kinds of shooting at various ranges, except for wide open shots. In that case, defenders should not jump on Curry while he is taking a wide open shot to avoid an additional foul.
- Curry is very likely to shoot 3-point shots continuously. However, for short distance jump shots or layups, no obvious continuity was observed. So after he scores a 3 point shot, defenders should pay more attention for a 3-point shot next time he shoots.
- Defender's Salary is not a key factor to make the defense different. A head coach does not want waste his superstar player's energy on defend Curry.
- Players such as J.R Smith and Manu Ginobili are doing well on defending Curry's threes.
- According to this analysis, no one could absolutely stop Curry.

### **Future Work:**

Due to the difficulty of collecting shooting data, the time variable is not recorded in the data. I do believe the time is a factor that will affect Curry's shot type and field goal percentage.

### **CONTACT**

Your comments and questions are valued and encouraged. Contact the authors at:

Name: Vrushali Walde  
Enterprise: Oklahoma State University  
Address: 127 N Duck Street, Apt 23  
City, State ZIP: Stillwater, OK, 74075  
Phone: (918)850-8818  
E-mail: [vrushali.walde@okstate.edu](mailto:vrushali.walde@okstate.edu)

Name: Wei Gao  
Enterprise: Oklahoma State University  
Address: 213 W Elm, Apt 323  
City, State ZIP: Stillwater, OK, 74074  
Phone: (405)742-7723  
E-mail: [wei.gao@okstate.edu](mailto:wei.gao@okstate.edu)