



Kobe Bryant's NBA career analysis

MIDTERM PROJECT REPORT

VIPUL MUNOT, SIDDHARTH JAYASANKAR, ANIRUDH K MURALIDHAR

Contents

Introduction	2
Motivation	2
Relevant Work	2
Interesting story of Analytics in Basketball:	3
Visualization Techniques	4
Hypothesis	4
Research questions	4
Process.....	4
Data Collection	4
Data Cleaning	4
Analysis	4
Visualization.....	5
Inference.....	5
RESULTS AND INSIGHTS	6
Shot Accuracy Across Seasons	6
Short Accuracy Based on Shot Types.....	6
Shot Accuracy Based on Shot Zone Area	7
Shot Accuracy Based on Shot Range	7
Shot Accuracy Against different opponents.....	8
Conclusion and Future Work.....	9
REFERENCES	9

Introduction

Nowadays big data is being used in almost all the domains. Sports informatics and analytics is also venturing into big data to perform in-depth analysis in order to come up with better strategies and team composition. Through this project we would like to would wanted to explore how useful effective visualization of data regarding player's performance could help develop better strategies and also help the player become better at the sport.

Motivation

Basketball is the second favorite sport in USA. It is third with regards to the money being spent. Basketball is the most played sport in America. The NBA has 30 teams and there are 82 games per season. Team USA (men) have won 15 gold medals in 19 Olympics. With so many games being played combined with the popularity and interest which this sport has, the amount of data at hand is huge. Added to these the team management want data driven strategies and solution to enhance their performance.

This project would help,

1. Analyze the player's performance
2. Develop a player by improving his weak zones.
3. Come up with better game strategies for different opponents and different players in the opponent team.
4. In drafting new players into the team.

Baseball is a sport where advanced analytics is already being used to great extent (The movie Money Ball is a great example of how advanced analytics can be used to develop better strategies and team composition which would lead to the success of the team.). Through this project we would like to contribute to such analytics where a whole team can be built in the game of basketball just by looking at the statistics.

Relevant Work

- Baseball is the game which is primarily data driven and other sports such as NFL, soccer and basketball are catching up. Sabermetrics is the application of statistical analysis of baseball that measure in-game activity.
- The NBA recently hosted its first hackathon on Sep 24th, 2016. This is an indication that NBA is turning towards the data side of that game.

Interesting story of Analytics in Basketball:

Daryl Morey is the General Manager of the NBA team Houston Rockets. He is predominantly a statistician. Through detailed analysis of how the game is played, Morey figured out that the average points scored from the 3-point range and close range shots were more than the number of points scored from midrange shots.

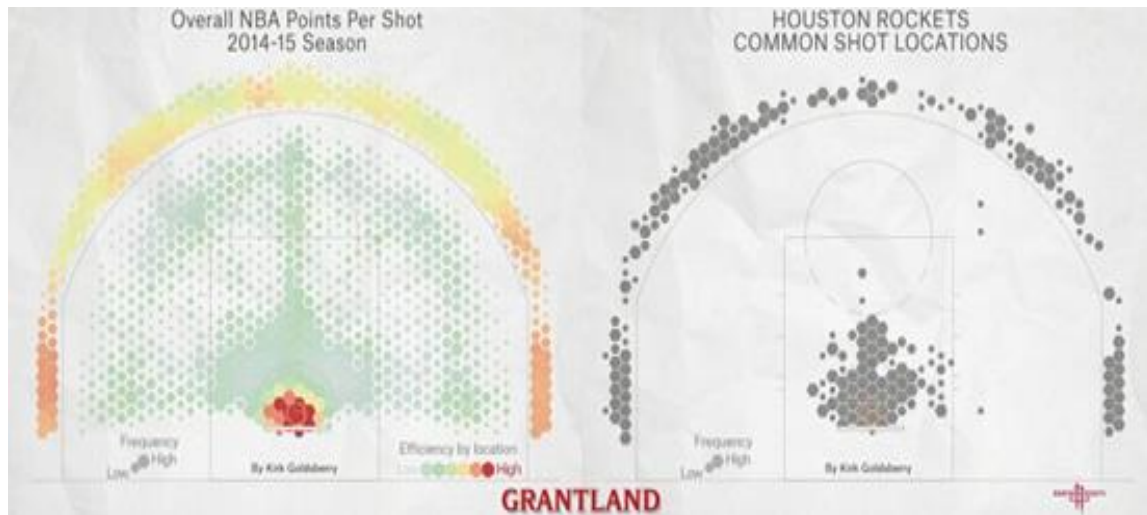


Figure 1 (a) Visualization showing efficiency of different range shots. (b) Shot map of Houston Rockets

Based on his analysis, he trained his team to shoot 3-pointers and close range shots. He ordered his team to not shoot from midrange and just concentrate on converting 3-pointers and close range shots.

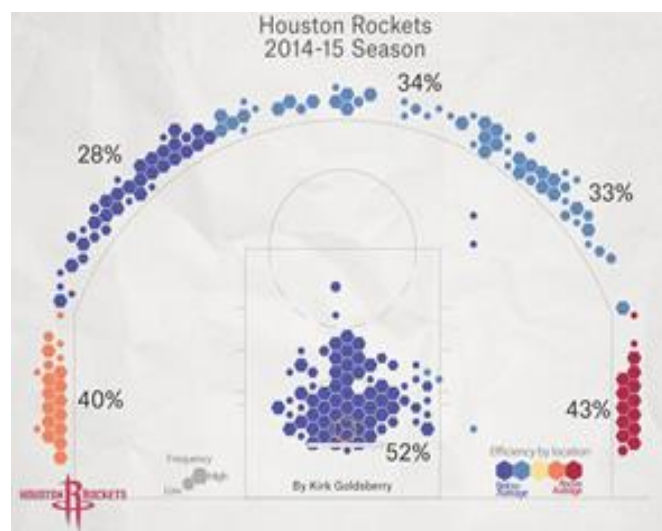


Figure 2. Efficiency of Daryl Morey tactics

As a result of this tactics, Houston Rockets reached the western conference for the first time since 1997.

This story proves that effective statistical analysis will help basketball teams perform better.

Visualization Techniques

We have used Python packages such as Matplotlib, Seaborn & Plotly and D3 module in Javascript to build rich visualizations to enable effective statistical analysis.

Hypothesis

If statistics can be used in other domains and in other sports such as baseball, why not make use of it in basketball given that huge amounts of data are available.

Research questions

- Analysis of Kobe Bryant's strong and weak zones.
- Performance analysis across seasons.
- Performance against various opponents.
- Develop a predictive model, in order to find if Kobe can convert a shot or not.

The broad-view of this analysis is going to be how well he has performed for the team and how he did in specific conditions and how this can be extended to other players in the team as well.

Process

Data Collection

The data has been downloaded from kaggle.com. (Kaggle had extracted the data NBA stats website)

	season	combined_shot_type	game_id	lat	loc_x	loc_y	lon	minutes_remaining	period	playoffs	seconds_remaining	shot_distance	shot_made_flag	shot_type	shot_zone_area	shot_zone_basic	shot_zone_range
1	1996/97	Jump Shot	29600027	33.9283	-140	116	-118.4098	0	1	0	42	18	0	2PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
2	1996/97	Jump Shot	29600031	33.9473	-131	97	-118.4008	10	2	0	8	16	0	2PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
3	1996/97	Jump Shot	29600044	33.8633	-142	181	-118.4118	8	2	0	37	23	1	3PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
4	1996/97	Jump Shot	29600044	34.0443	0	0	-118.2698	6	2	0	34	0	0	3PT Field Goal	Center(C)	Restricted Area	Less Than 8 ft.
5	1996/97	Jump Shot	29600044	33.9063	-10	138	-118.2798	5	2	0	27	13	1	2PT Field Goal	Center(C)	In The Paint (Non-RA)	8-16 ft.
6	1996/97	Jump Shot	29600057	33.8213	-64	223	-118.3338	2	2	0	16	23	1	3PT Field Goal	Center(C)	Mid-Range	16-24 ft.
7	1996/97	Jump Shot	29600057	33.8673	-79	177	-118.3488	1	3	0	53	19	0	2PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
8	1996/97	Jump Shot	29600057	33.8373	-103	207	-118.3728	1	3	0	14	23	1	3PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
9	1996/97	Layup	29600057	34.0443	0	0	-118.2698	0	3	0	2	0	0	2PT Field Goal	Center(C)	Restricted Area	Less Than 8 ft.
10	1996/97	Jump Shot	29600057	33.8693	-155	175	-118.4248	9	4	0	9	23	0	3PT Field Goal	Left Side Center(LC)	Mid-Range	16-24 ft.
11	1996/97	Layup	29600057	34.0443	0	0	-118.2698	8	4	0	36	0	0	2PT Field Goal	Center(C)	Restricted Area	Less Than 8 ft.

Figure 3 Sample dataset showing various features

Data Cleaning

The data obtained was cleaned by removing incomplete records and irrelevant features.

Analysis

The data has been analyzed to build a predictive model using machine learning algorithms to predict if a Kobe converted a particular shot or not

Visualization

Various visualizations are created in order to support the analysis and also to make easy and effective inferences from the data.

Inference

The final inferences are derived from the analysis and the graphs.

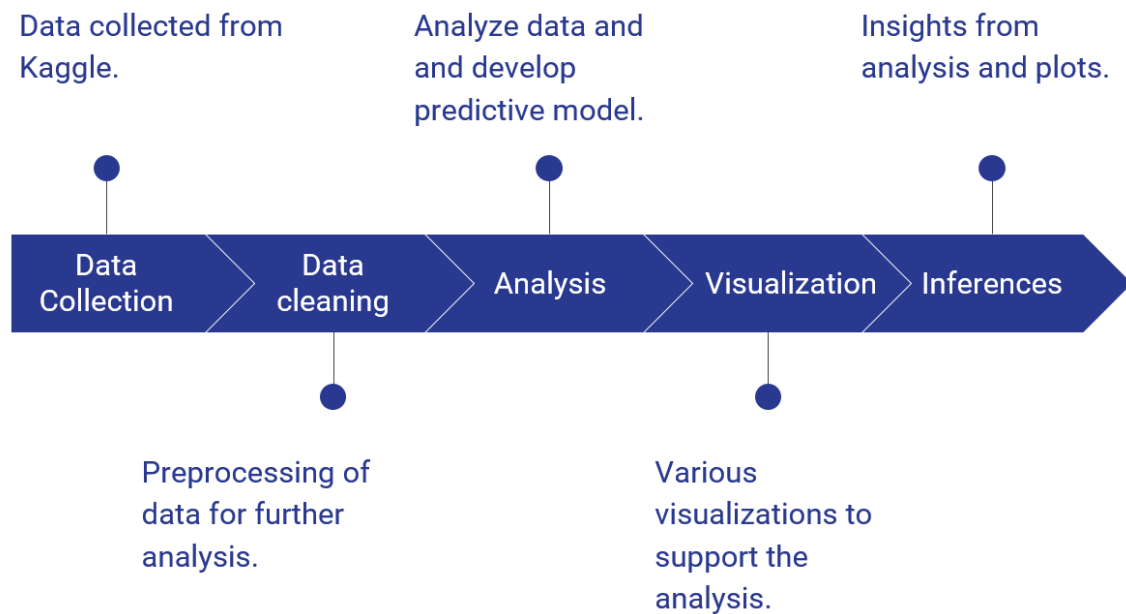


Figure 4 Image showing the overall process of the analysis

RESULTS AND INSIGHTS

Shot Accuracy Across Seasons

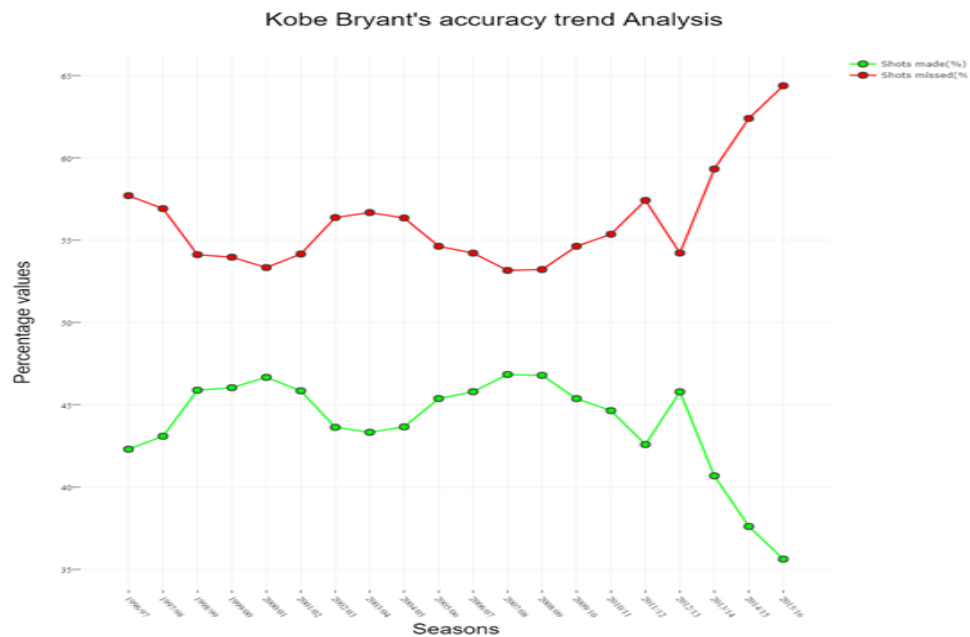


Figure 5 Trend showing Kobe's shot accuracy across seasons

Insights from the graph are as follows

- Kobe had a good start to his career.
- His peak was between the years 2007-2010.
- His performance declined from the year 2013.
- His overall shot accuracy is 44.6%

Short Accuracy Based on Shot Types

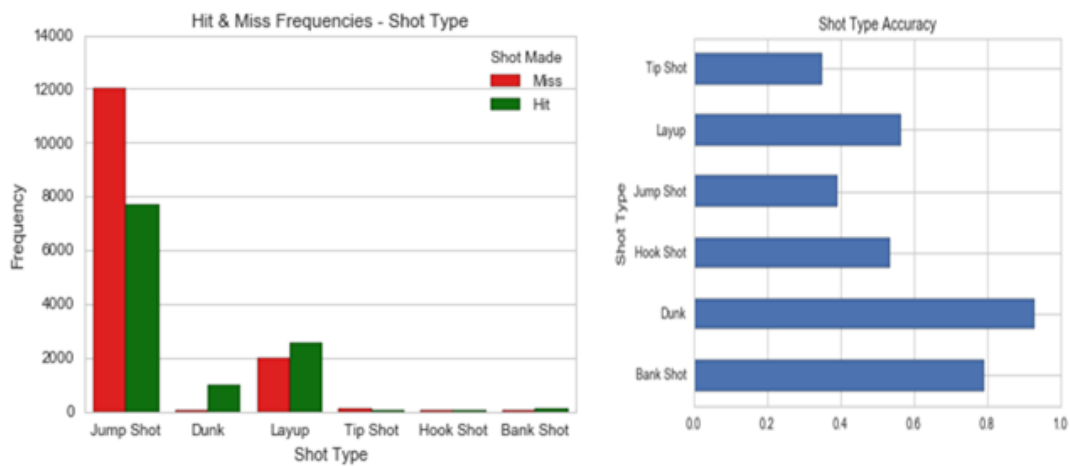


Figure 6 Plots showing Kobe's accuracy with various shots

Insights made from these graphs

- Kobe's default shot type is jump shot.
- He is most likely to succeed when he tries dunk shots.

Shot Accuracy Based on Shot Zone Area

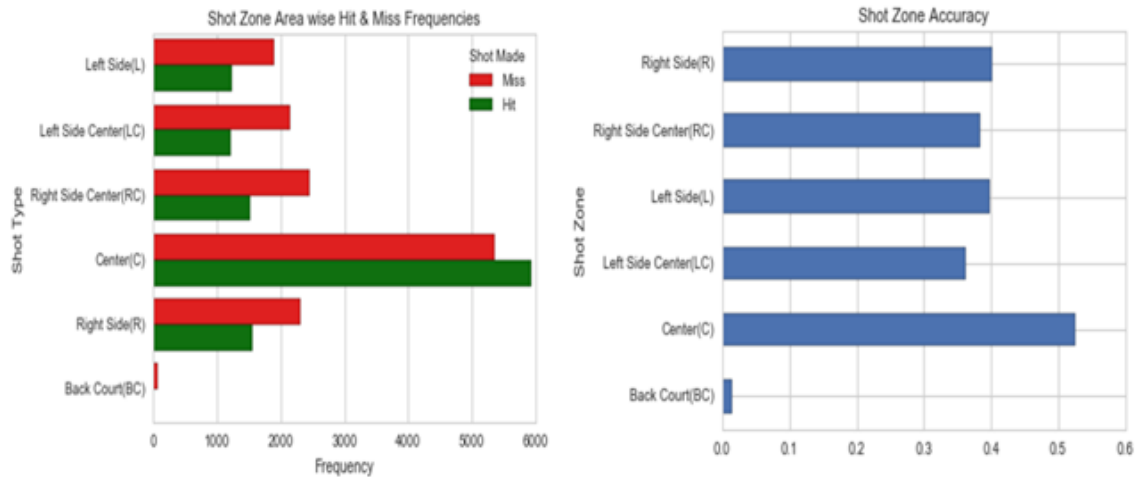


Figure 7 Plots showing Kobe's performance over various shot zones

Insights from these graphs are as follows

- Kobe has made more attempts from the center and has succeeded in that most of the times.
- He favors the left hand side compared to his right

Shot Accuracy Based on Shot Range

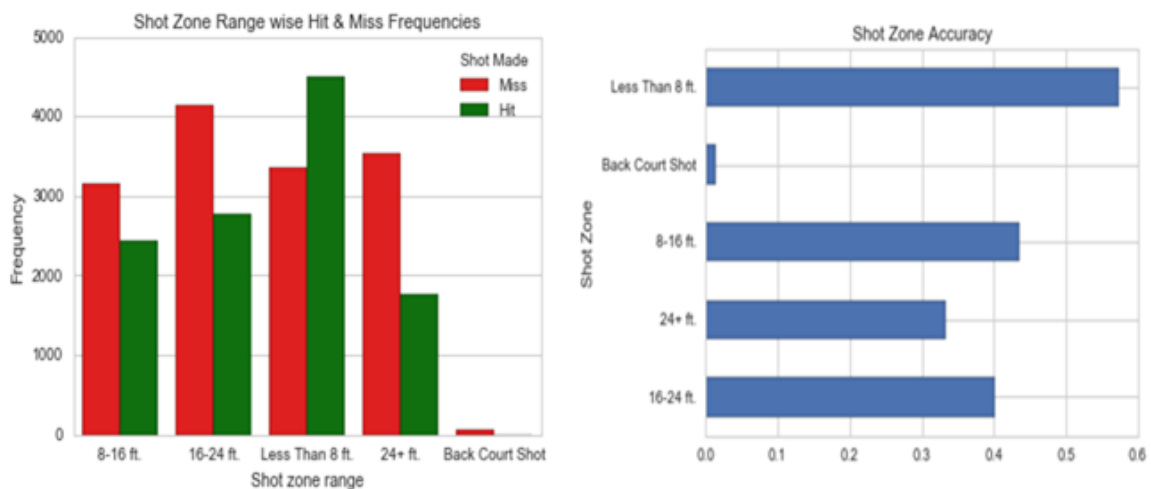


Figure 8 Plots showing Kobe's performance over various shot ranges

Insights from these graphs are as follows

- Kobe has the best go when the shot range is less than 8ft.
- Also, we could find the range is inversely proportional to his accuracy.

Shot Accuracy Against different opponents

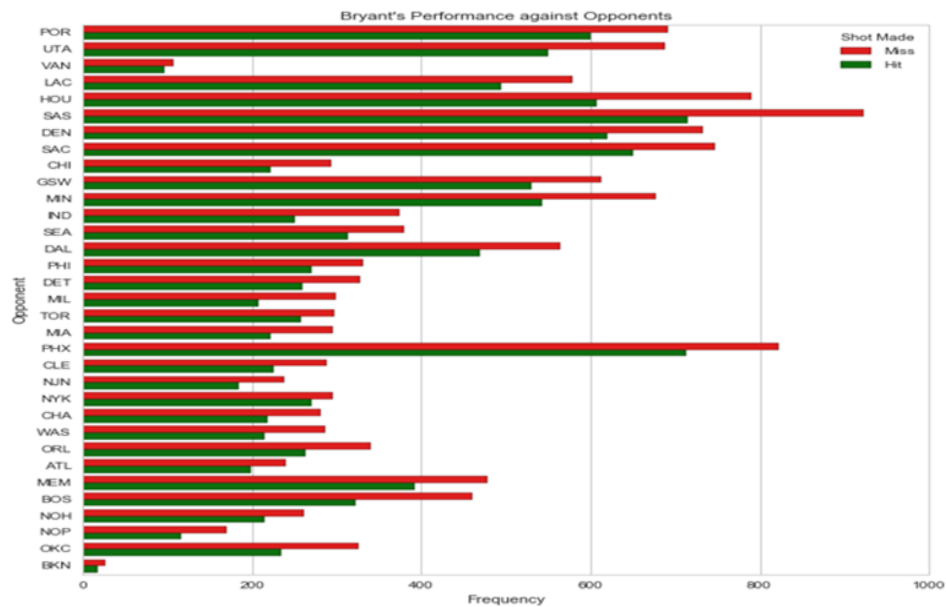


Figure 9 Performance against different opponents

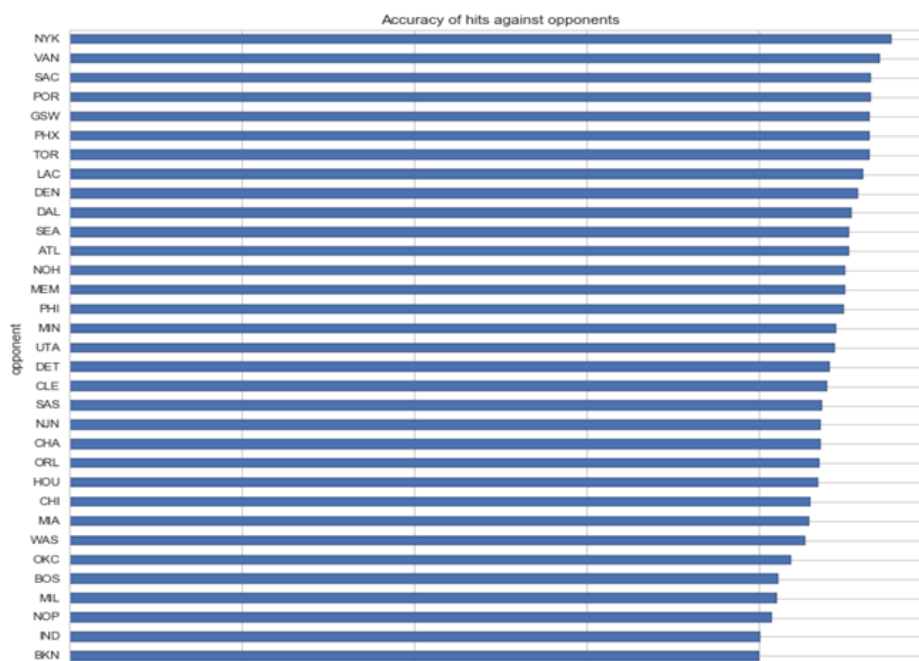


Figure 10 Plots showing Kobe's performance against various teams

Insights from these graphs

- We could see that he has done well against most of the teams, but he has the best accuracy against New York Knicks.
- Also, he has the worst accuracy against the team Brooklyn Nets.

Conclusion and Future Work

The initial exploratory analysis of Kobe Bryant's data gave us some insights about his performance based on several factors. Though these give us an overall idea about his performance, much deeper analysis can give us some worthy insights.

Based on the results obtained, the future works will be as follows

- Develop a predictive model to predict if Kobe will hit or miss a shot.
- Develop much deeper analysis against various teams and across various seasons based on several factors.
- Create much more enhanced visualizations.

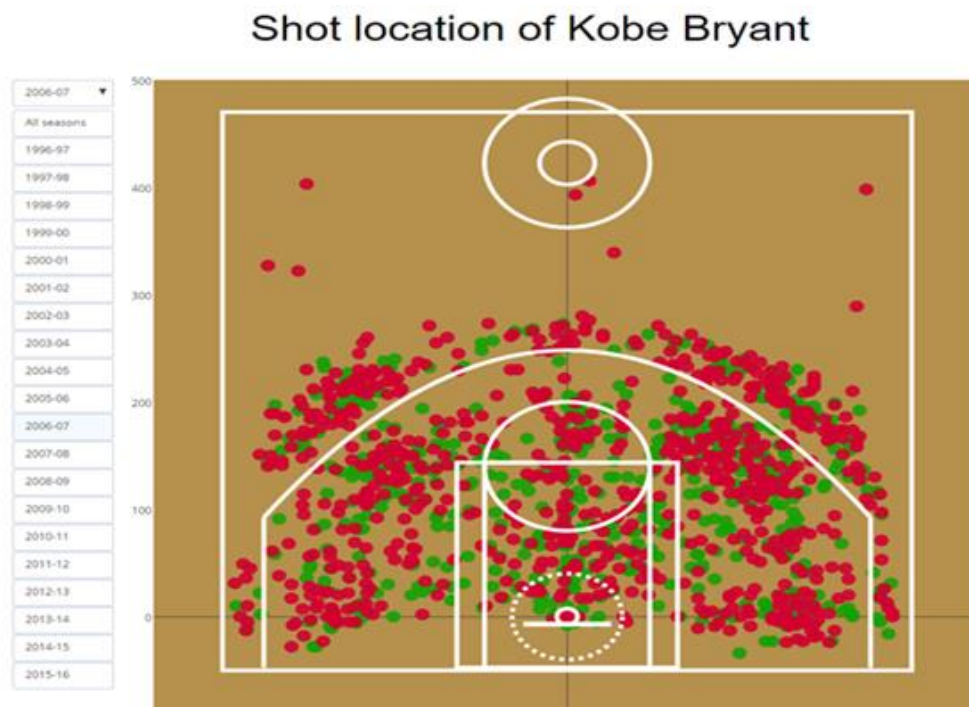


Figure 11 Shot location across seasons in a much better chart. [Link: https://goo.gl/MBU115](https://goo.gl/MBU115)

REFERENCES

1. Kaggle site: <https://www.kaggle.com/c/kobe-bryant-shot-selection>
2. NBA stats website: <http://stats.nba.com/>
3. Kobe Bryant: https://en.wikipedia.org/wiki/Kobe_Bryant
4. Python plotly: <https://plot.ly/>
5. Related work: <http://www.houstonchronicle.com/news/article/The-cult-of-analytics-with-Rockets-general-4342412.php>
6. Python Code: <https://github.com/vipmunot/Data-Science-Course/tree/master/Data%20Visualization/Project>