NBA Shot Log Analysis

By Jesse Zamora

TABLE OF CONTENTS

Introduction 2

Product/Service/Methodology 3

Key Findings 4

Key Findings #1 4

Key Findings #2 5

Key Findings #3 5

Visual Data 6

Conclusion 7

Key Takeaways 7

Appendix…………………………………………………………………………………………………………………………………………………8

Additional Questions……………………………………………………………………………………………………………………………….9

Jesse Zamora



# Introduction

The business problem this project is attempting to solve would be to assist NBA teams in drilling down on if a certain player should even be attempting a 3-point shot or if utilizing them elsewhere would be more efficient. I am getting my data set from Kaggle. The data set is a csv file titled shot\_logs.

Table

Description automatically generated with medium confidence

# Product/Service/Methodology

Initially I wanted to focus on the EDA portion of the data and really drill down on the important information. I did this by creating a new data frame and dropped a lot of columns that I did not feel were necessary. Having the bulk cut out of the project assisted in being able to drill down on the focus of the project and that was verifying the accuracy of 3-point shots versus 2-point shots.

The data shown here is clearing out the columns that were not necessary and then viewing the data which led to the realization that there was a negative amount in touch time which should not be possible. I then moved on to negate the negative value and then implanted the mean of the column for touch\_time to avoid having null values. I believe that implanting the data with the mean is the best way to go around removing bad data as long as it is clear that the bad data was dealt with in that particular fashion.

# Key Findings

## Shot percentage by player

Graphical user interface, table

Description automatically generated with medium confidence

This was really a shocking discovery. The player with the highest percentage of shots made or “hits” was Deandre Jordan at over 71%.

## Key Findings #2

Chart, scatter chart

Description automatically generated

This data shows the percentage of “hits” or shots made in relation to how many seconds are left on the shot clock.

## Key Findings #3

Graphical user interface, text, application

Description automatically generated

This data shows that the 3-pointer is made 13% less than a 2-pointer.

Table

Description automatically generated

# Conclusion

I think that overall, the project was done in a way that leaves the door open to additional examination. There were many interesting columns in the csv file that were not used, and I would like to return to this project eventually and work on it some more. Some of the things I would add would be the 3 and 2 point shot percentages in relation to a defender within three and five feet ranges to see if that increases or decreases the “hit” percentage.

As far as ethical concerns of this project I believe this will not have a negative impact and most of the information was known but the margins were unknown. For example, I don’t think many people would guess that the percentage of 3-point shots versus 2-point shots is only 13%. This data does not include shots like dunks or layups that may have an unproportionally high percentage of shots made as the player is inches from the hoop.

## Key Takeaways

* Deandre Jordan is the NBA’s most reliable scorer at a staggering 71.25%.
* 20-22 seconds on the shot clock is the most reliable scoring percentage.

**Reference:**

Becker, D. (2016, August 18). *NBA shot logs*. Kaggle. Retrieved April 3, 2022, from https://www.kaggle.com/datasets/dansbecker/nba-shot-logs

**Appendix**

**Analysis of Defenders**

The additional data has not been analyzed as of this point; however, this is definitely a very interesting aspect of the project. I believe that this information will shed some light on the information that was seen in this project in ways that were not seen before.

10 Questions

1. In relation to the best shooters – who are also the NBA’s highest scorers?
2. Is there a positive or negative correlation between shots taken and shots made?
3. Who is the NBA’s best defender?
4. Why were specific columns cut from the project?
5. Is Deandre Jordan the highest scorer in the NBA?
6. If he is not, how come?
7. Should the team that Deandre Jordan plays for focus on how to center the offense around him?
8. Is Deandre Jordan a statistical anomaly?
9. Why is 20-22 seconds on the shot clock the highest percentage of shots made?
10. Will you accept a position with our organization?