

Historical Analysis of NBA Shooting

Group 10

Calvin Crosby, Samuel Hughes, Tarun Kumar Reddy Etikala,
and Prathiksha Mohan Raje Urs

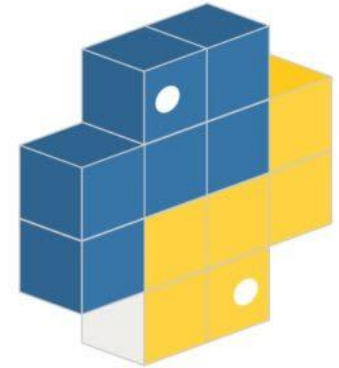


Project Goals

1. Visualize the recent historical evolution of basketball strategy
2. Explain analytics of NBA shooting
3. Create a machine learning model to predict shooting performance

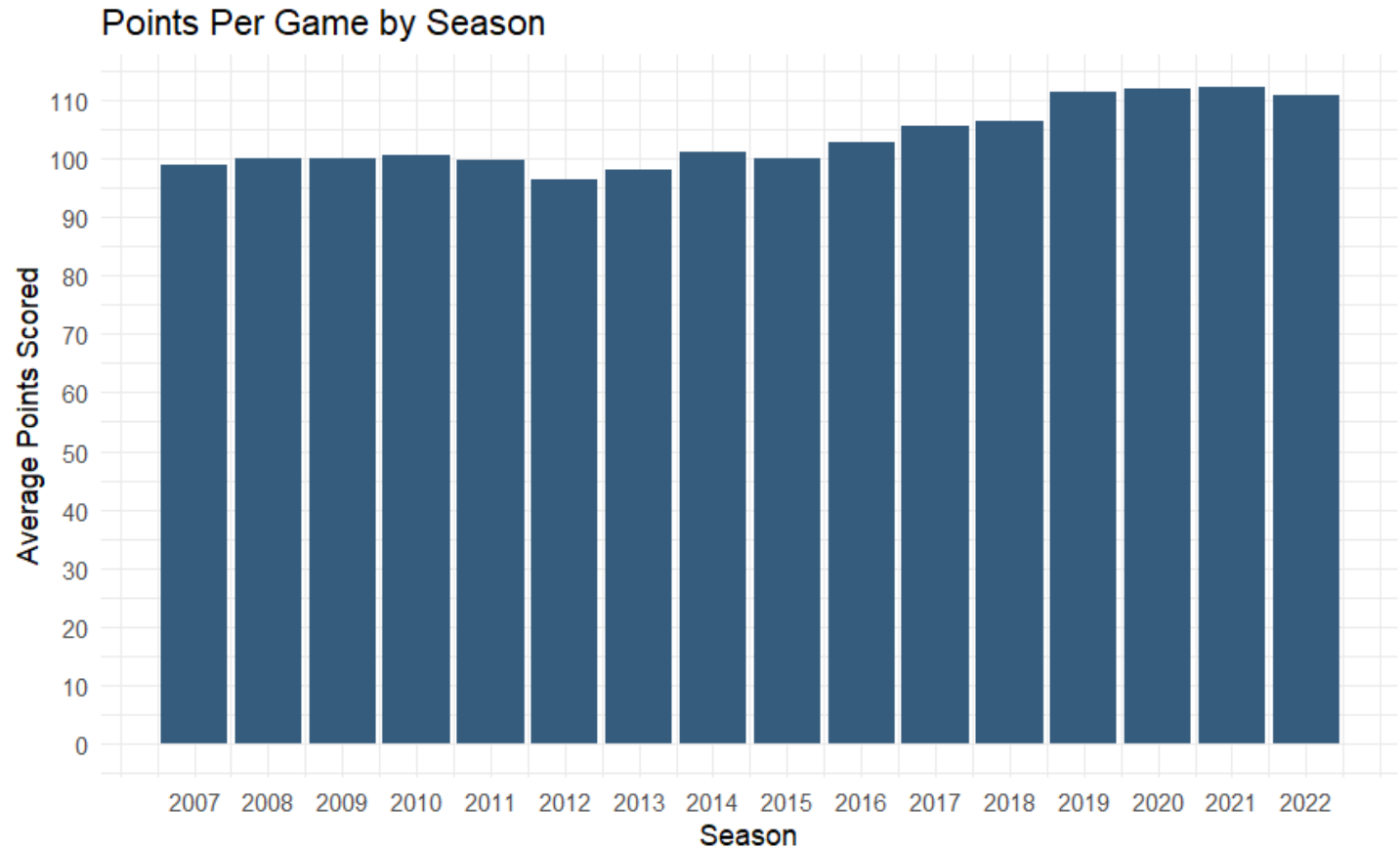
Data Preprocessing

- Dataset 1: NBA Shot Locations
 - NBA shots since 2007 containing shot technique, geocoordinates, and result of shot
 - Pulled from <https://www.nba.com/stats> using python nba_api package
 - Removed data entry errors (e.g. two-points shots from beyond three-point line and vice versa)
 - Collapsed original shot technique variable into more intuitive categories
- Dataset 2: Points Per Game
 - Contains average points scored in a game by season
 - Pulled from <https://www.basketball-reference.com/> using rvest package



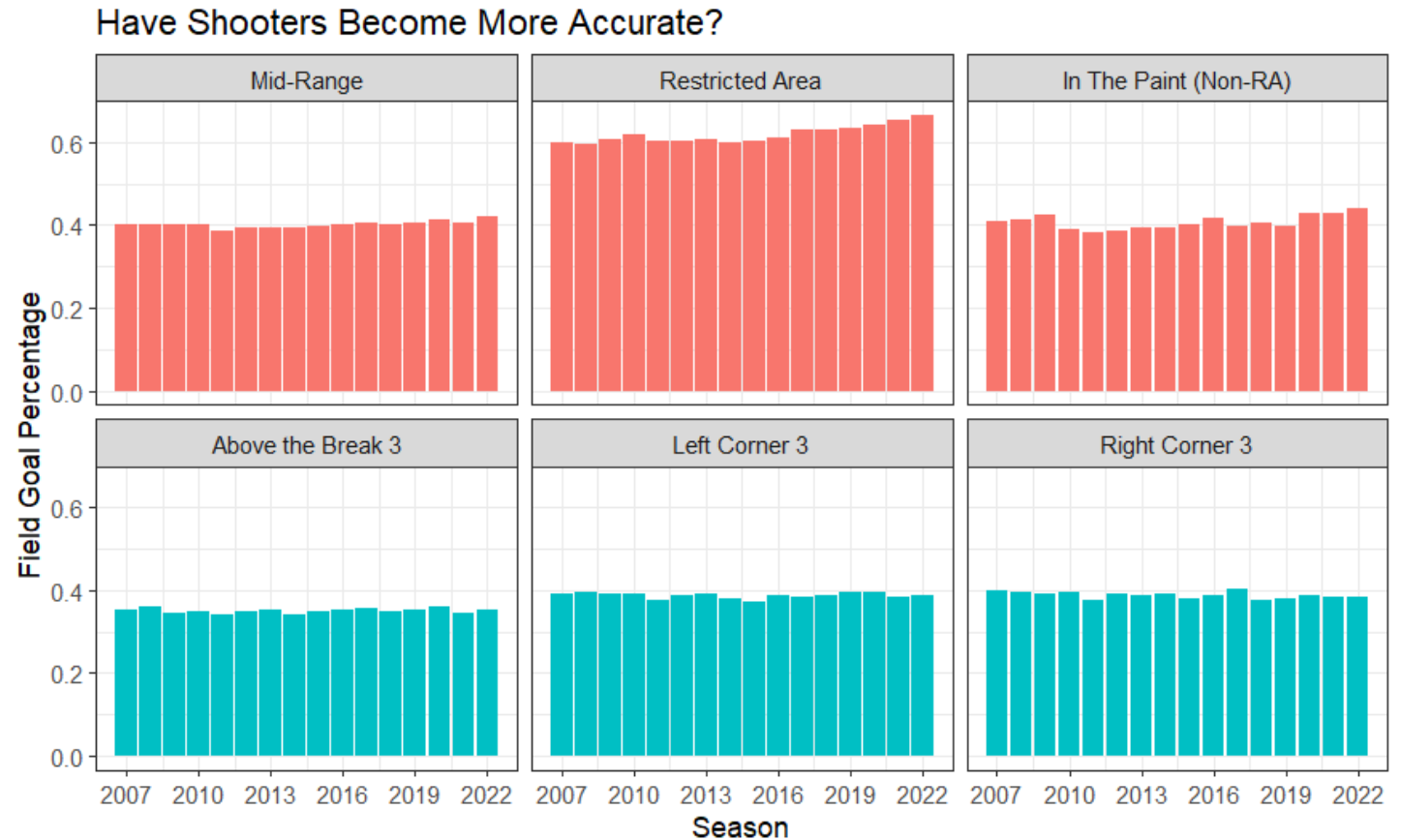
Evolution of NBA Scoring

- Teams score about 10 more points per game now than 15 years ago
- Intuitively, we would think this is because NBA players learned to shoot the ball better...



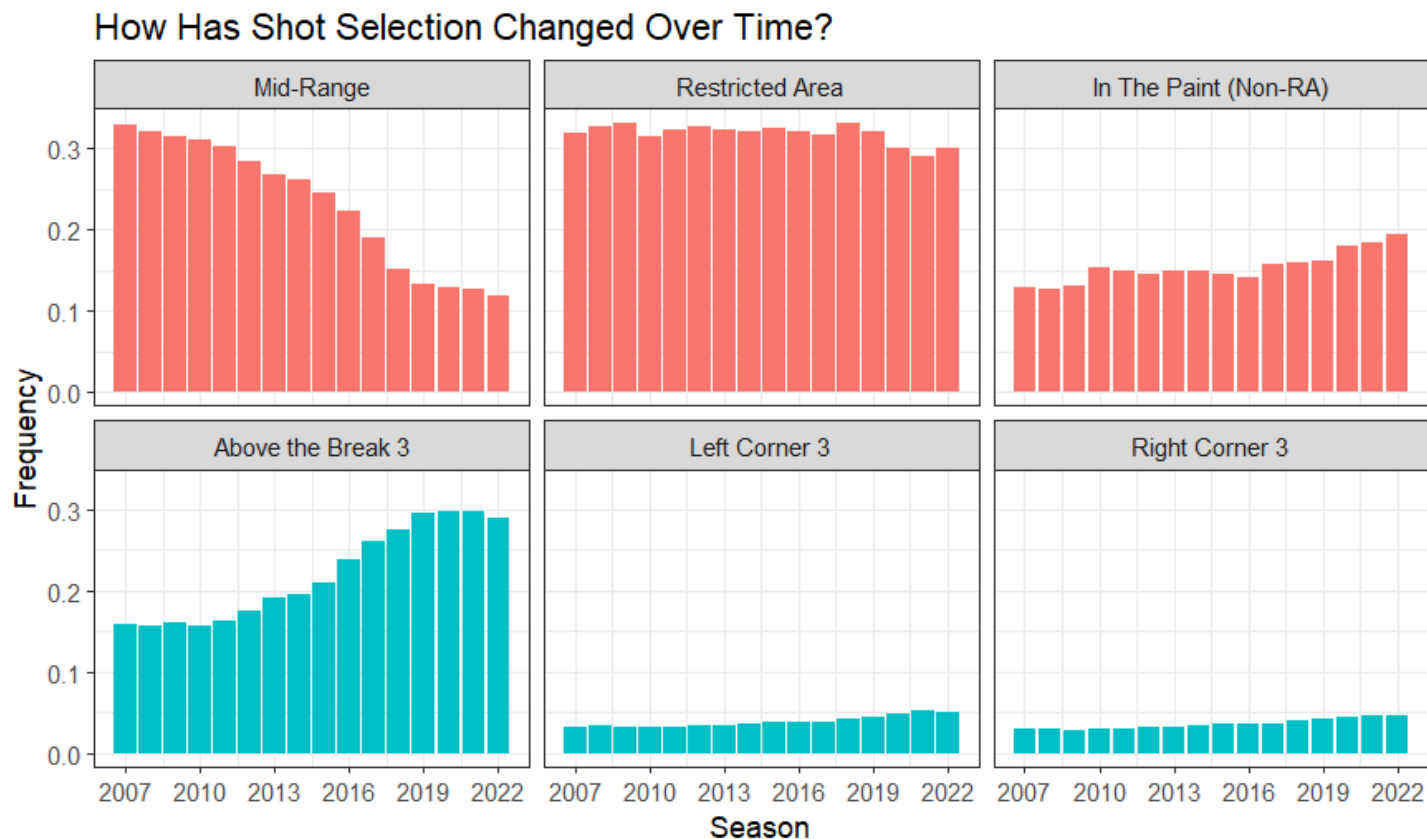
Shooting Accuracy

- Shooting percentages have remained relatively stagnant
- If shooters are not more accurate, how are teams scoring more points than ever before?



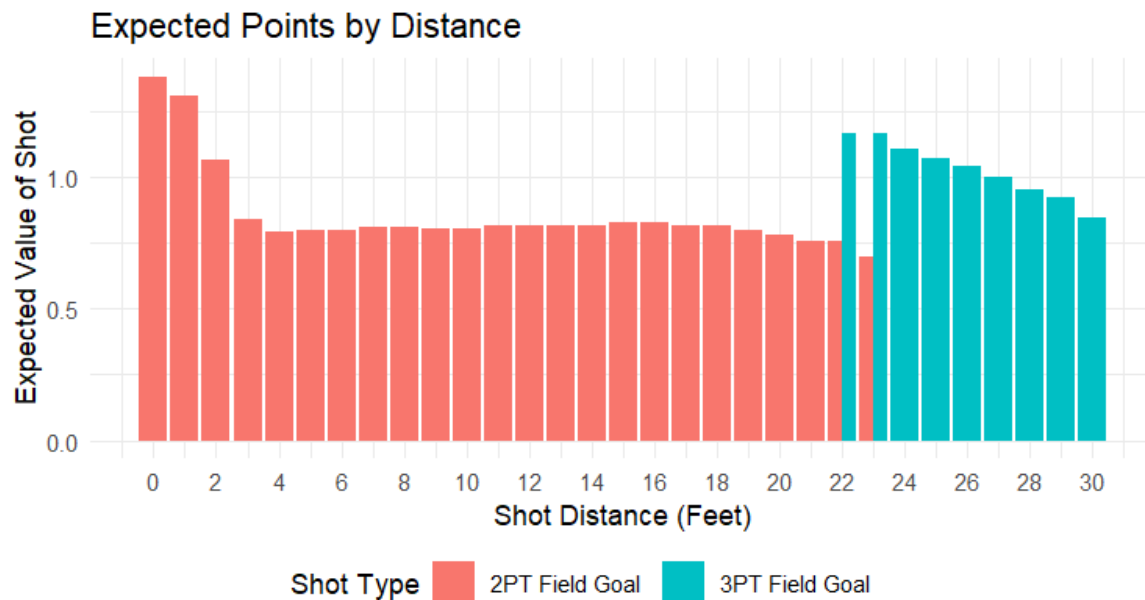
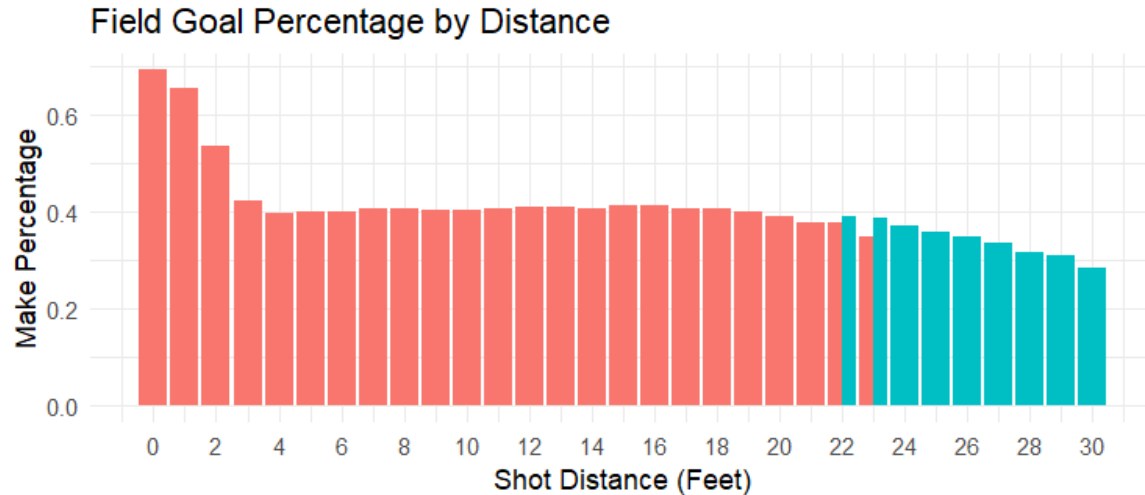
Shot Selection

- Mid-range shots are much less common
- NBA teams shoot more three pointers than ever before



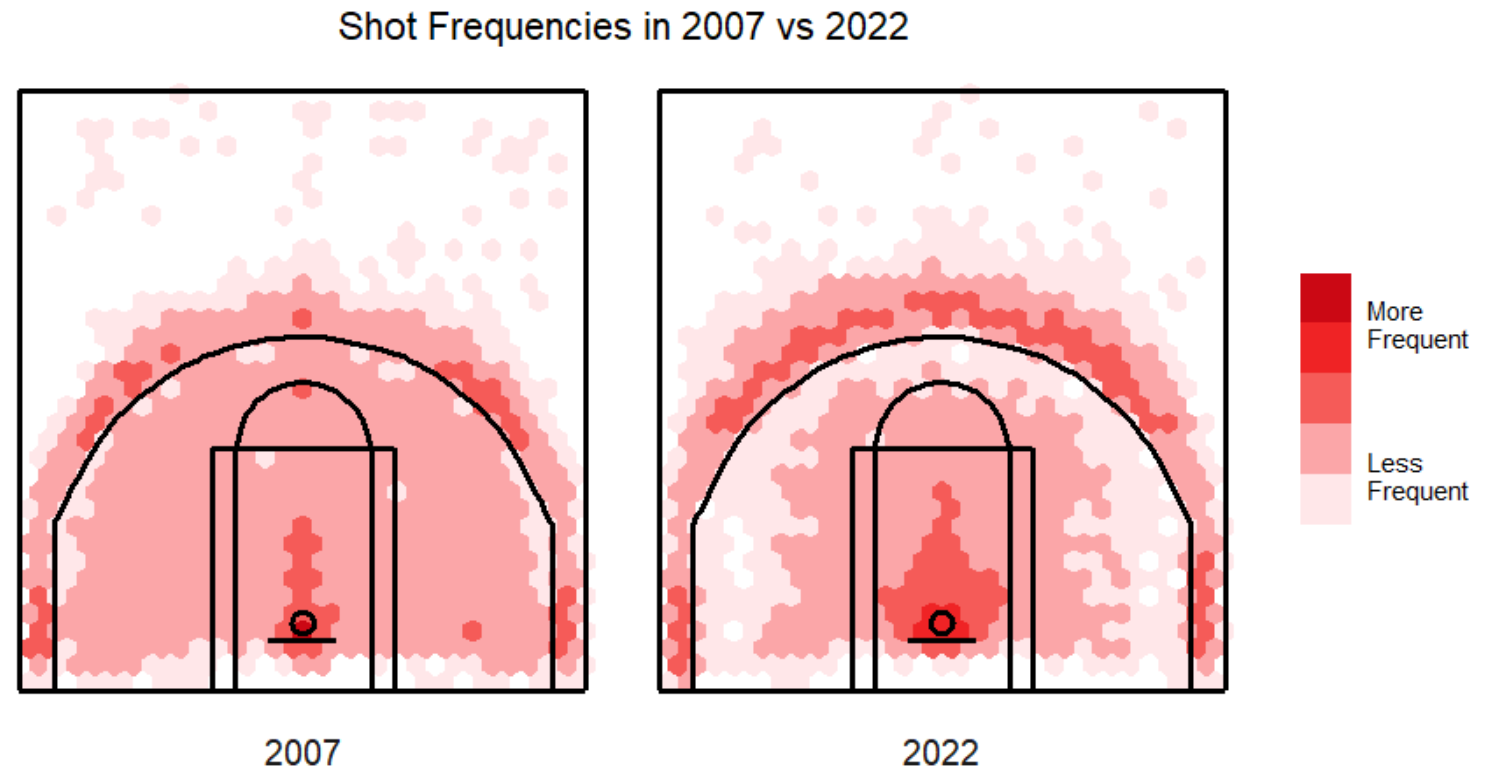
Shooting Analytics

- A medium ranged two-point shot is barely made at a higher percentage than a three-point shot, yet is worth 1 less point
- By decreasing mid-range shots and increasing three-point shots, teams have increased points scored per game



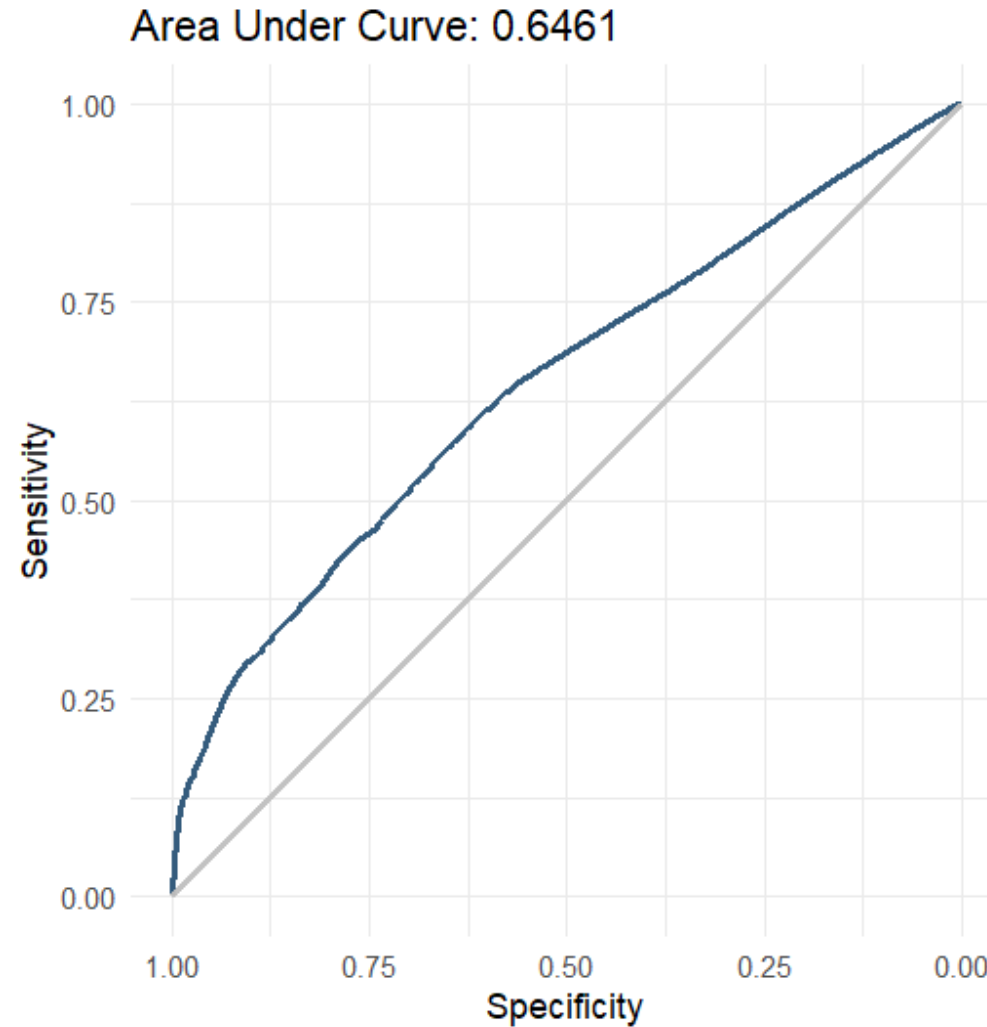
Shot Chart Comparison

- In 2007, shot selection was much more evenly distributed
- In 2022, many shots were either three-pointers or close to the basket



Model

- Predicts chance of scoring a given shot
- L1 Logistic Regression
 - Built-in feature selection
 - Handles multicollinearity
 - Easily interpretable variables
- 62% accuracy and 0.6461 AUC score on out-of-sample data



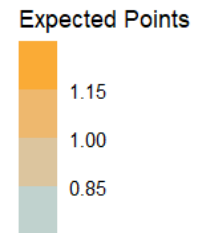
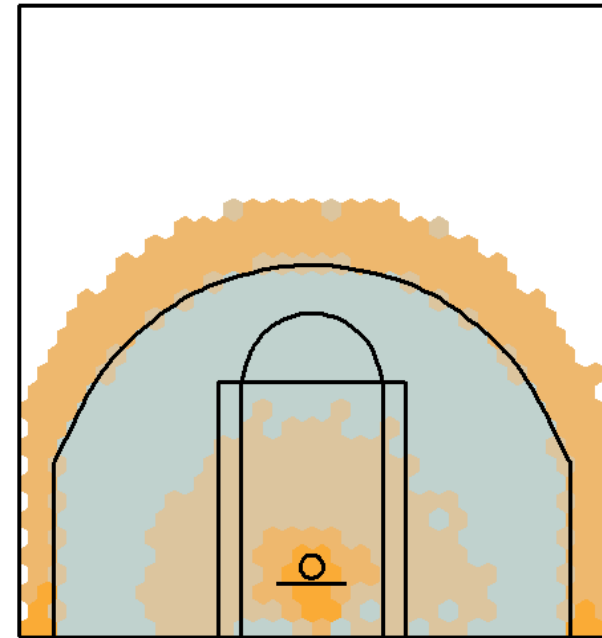
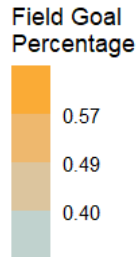
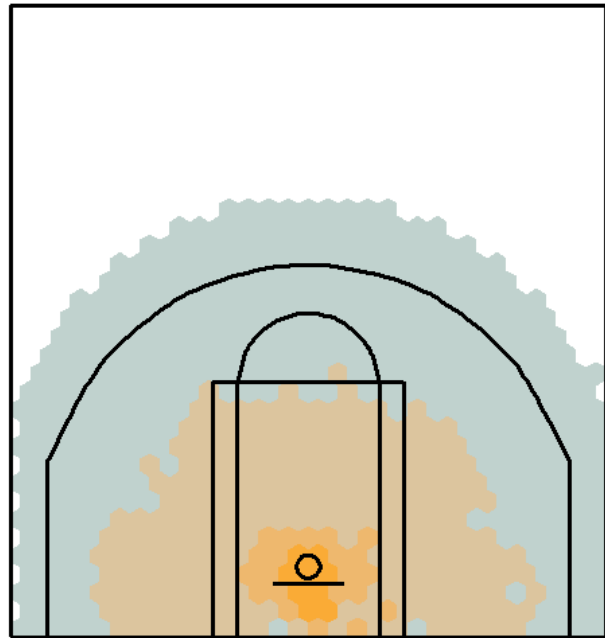
Feature Importance

- Dunks, layups, and driving shots lead to great scoring opportunities
- Horizontal distance and total distance of shot are negligible in model
- With perfect dataset, investigate closest defender and player velocity

| VARIABLE | DESCRIPTION | VALUE |
|--------------------------|---|-------|
| Dunk | Player pushes ball through basket | 0.52 |
| Layup | Finesse shot close to basket | 0.19 |
| Driving Shot | A shot while approaching basket | 0.15 |
| Cutting Shot | Player evades defender before receiving and shooting basketball | 0.14 |
| Running Shot | A shot on the run | 0.12 |
| Putback Shot | Player rebounds missed shot and immediately attempts new shot | 0.10 |
| Hook Shot | One-handed lob shot | 0.06 |
| Tip Shot | Player taps ball toward basket | 0.05 |
| Turnaround Shot | Shot where player is initially facing away from basket | 0.05 |
| Stepback Shot | Shot where ballhandler steps backwards before shooting | 0.05 |
| Reverse Shot | Shot where player traverses underneath basket | 0.05 |
| Alley-Oop Shot | Player catches the ball mid-air and shoots before landing | 0.05 |
| Fadeaway Shot | Shot where player is jumping away from basket | 0.04 |
| Floater Shot | Shot that is lobbed over a defender | 0.03 |
| Shot Horizontal Distance | Horizontal distance of shot | 0.00 |
| Shot Distance | Euclidean distance of shot | -0.01 |
| Shot Vertical Location | Vertical location of shot | -0.08 |

Shot Value According to Model

Using model output, we can evaluate the value of shot attempts





Summary

- The shift in NBA teams' strategy to take more three-point shots compared to medium range two-point shots has increased points per game
- Crucially, the league has realized that it is advantageous to exchange deeper two-pointers (with a low expected point value) for three-pointers (with a much higher expected value)
 - While these shots may have comparable odds of being made, the added point earned with a three-point shot more than compensates for the additional distance required for this shot



References

1. https://www.basketball-reference.com/leagues/NBA_stats_per_game.html
2. <https://www.nba.com/stats>
3. <https://pypi.org/project/nba-api/>
4. <https://cran.r-project.org/web/packages/rvest/index.html>

Code

https://github.com/shughes1000/nba_shot_visualization