

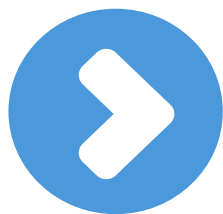


15th November 2023

“Predicting shot success in NBA”

By Anastasios Pliatsikas

DeptEconResearchSeminars
UNIVERSITY OF CRETE



Data Scientist
– EWORX S.A.



MSc in Business
Analytics – AUEB



BSc in Statistics
– AUEB



ANASTASIOS PLIATSIKAS

an.pliatsikas@gmail.com





Objectives



Create a large, complex, multi-purpose dataset of NBA shot logs



Identify the best classification method for predicting shot success



Determine the indicators mostly influencing shooting effectiveness



Investigate separately all field goals, two-pointers & three-pointers



AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

04 Methodology

05 Analysis Variants

06 Discussion and Conclusions



AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

04 Methodology

05 Analysis Variants

06 Discussion and Conclusions



NBA History



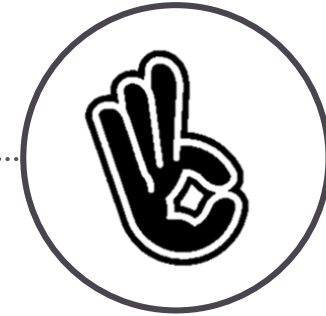
NBA founded
on 06-06-1946

1946



1891

Basketball invented by
James Naismith on 21-12-1891



1979

Three-point line
adopted by NBA



NBA History



2013

Stats Perform
provider

Second Spectrum
provider

2017



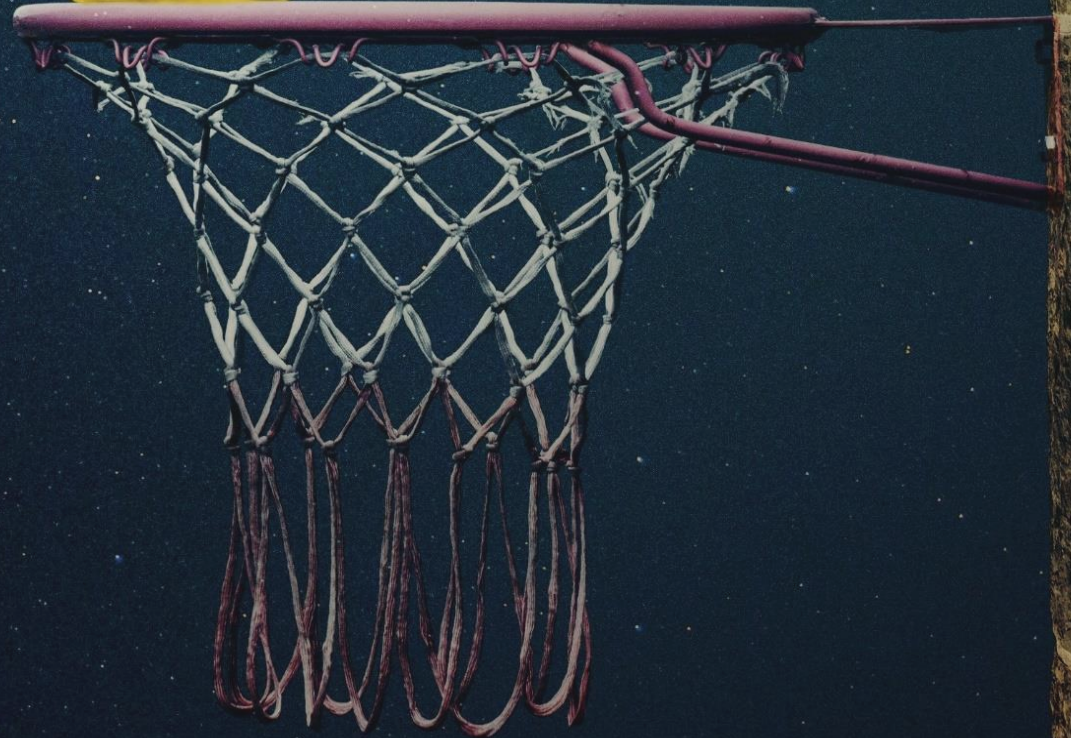
Today

3rd wealthiest sport league in
2023 with \$10B annual revenue

“

“NBA is the strongest professional sports league in the world.”

— Michael Jordan





AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

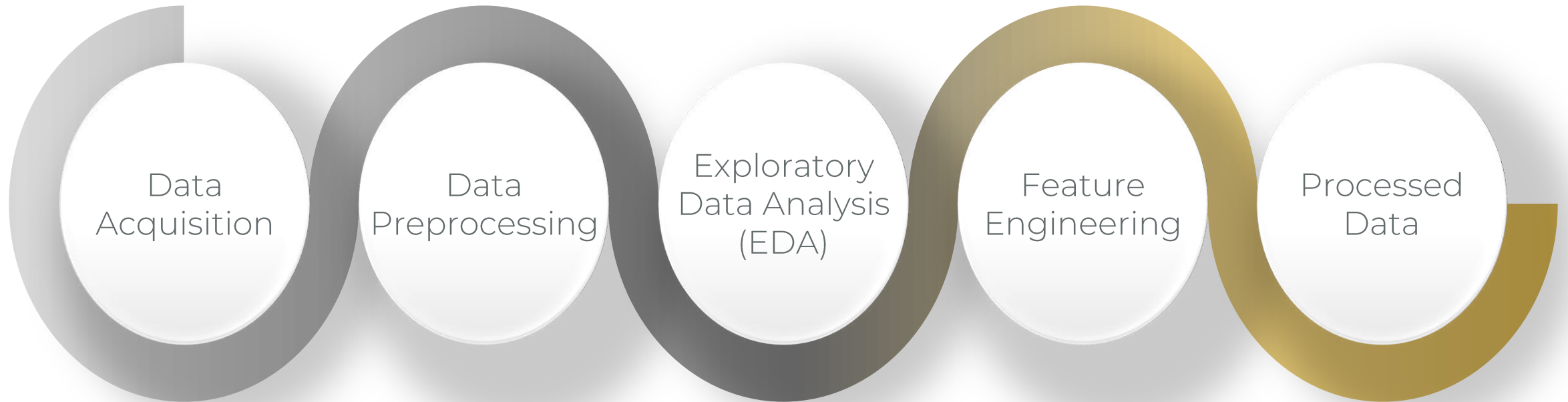
04 Methodology

05 Analysis Variants

06 Discussion and Conclusions



Data Preparation





Data Acquisition

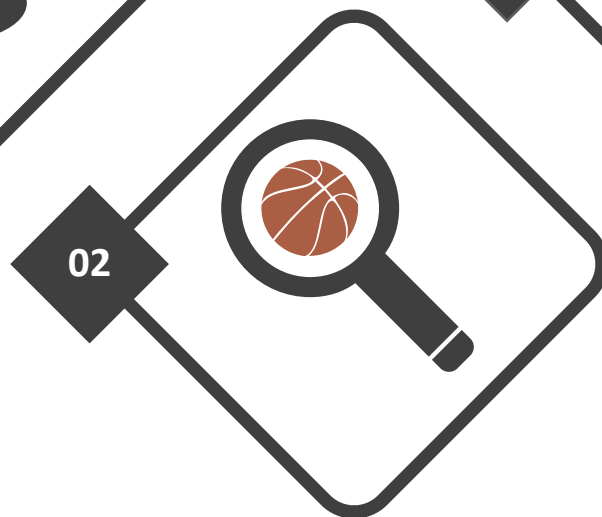
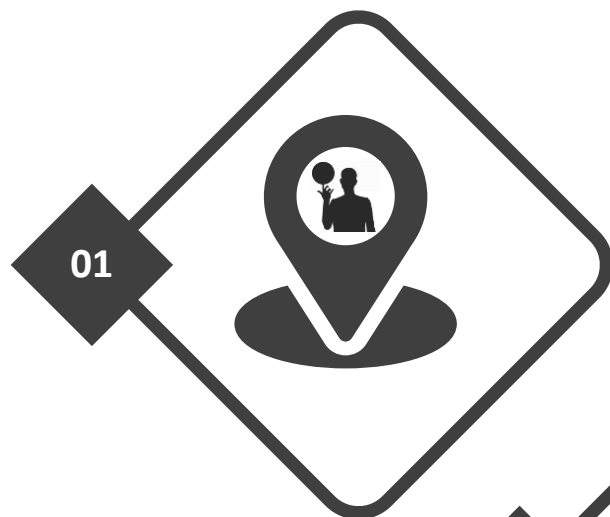


NBASavant.com

Player Tracking

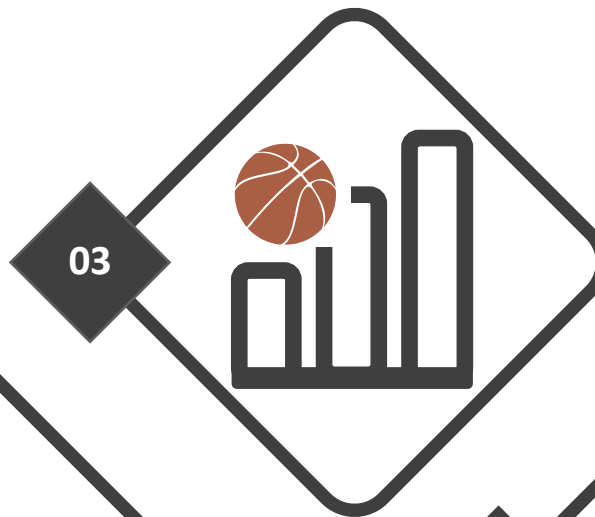
NBA Stats

Team Rankings



Play-by-Play

Package hoopR



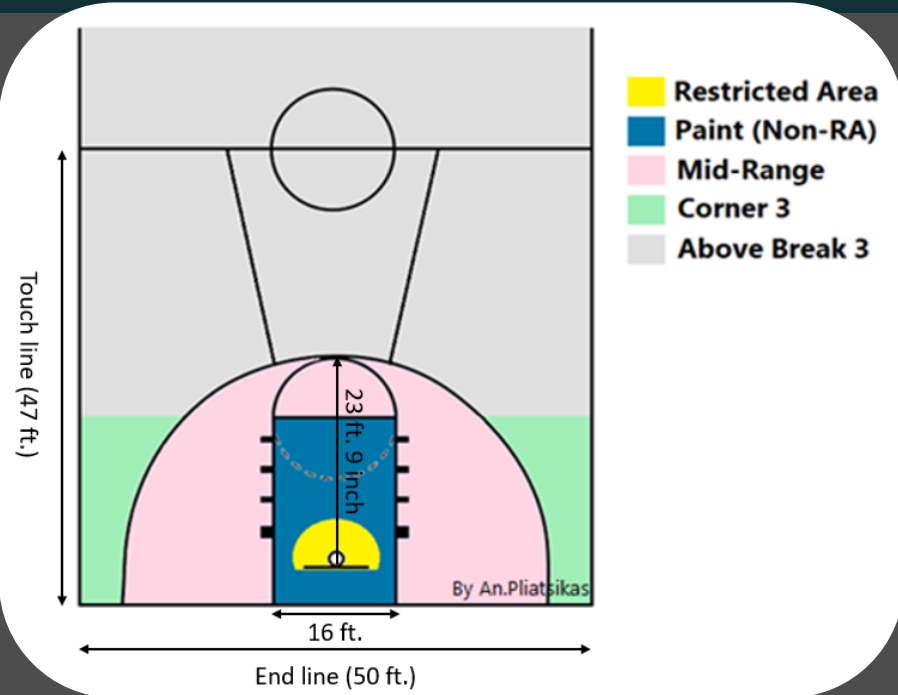
Player Bios



Basketball-Reference.com
Bball-index.com
NBA.com
NBASTuffer.com
Package nbastatR

Data Processing





Shot
Segmentation



Distance & Angle



Shooting Zones



Shooting Ranges



Shooting Side



Shooting Type

Data Processing

Score Margin

Defensive Pressure



Shot Segmentation



Defensive Pressure



Defensive Intensity

$$\text{Pressure} = \begin{cases} \text{No, if closest defender distance} = 6 + \text{ ft. (1.8 + m)} \\ \text{Low, if closest defender distance} = 4 - 6 \text{ ft. (1.2 - 1.8 m)} \\ \text{Moderate, if closest defender distance} = 2 - 4 \text{ ft. (0.6 - 1.2 m)} \\ \text{High, if closest defender distance} = 0 - 2 \text{ ft. (0 - 0.6 m)} \end{cases}$$

$$\text{Intensity} = \begin{cases} \text{Open, if closest defender distance} \geq 5 \text{ ft. (1.5 m)} \\ \text{Guarded, if } 3 < \text{closest defender distance} < 5 \text{ ft. (1.5 m)} \\ \text{Contested, if closest defender distance} \leq 3 \text{ ft. (0.9 m)} \end{cases}$$

Data Processing

Score Margin



$$\text{Period status} = \begin{cases} \text{Close, if } |\text{quarter margin}| \leq 8 \\ \text{Balanced, if } 9 \leq |\text{quarter margin}| \leq 17 \\ \text{Unbalanced, if } |\text{quarter margin}| \geq 18 \end{cases}$$

$$\text{Shot tension (quarterly)} = \begin{cases} \text{High Advantage, if } \text{quarter margin} \geq 21 \\ \text{Moderate Advantage, if } 11 \leq \text{quarter margin} < 21 \\ \text{Low Advantage, if } 3 \leq \text{quarter margin} < 11 \\ \text{Balanced, if } -5 \leq \text{quarter margin} \leq 2 \\ \text{Low Disadvantage, if } -5 \leq \text{quarter margin} < -13 \\ \text{Moderate Disadvantage, if } -13 \leq \text{quarter margin} < -24 \\ \text{High Disadvantage, if } \text{quarter margin} \geq -24 \end{cases}$$

Defensive Pressure



Shot Segmentation



Physical Attributes



Data Processing

Score Margin



Time-related information



Shot Clock

Possession
Type

$$= \begin{cases} \text{Late, if } 0 \leq \text{shot clock} < 5 \\ \text{Structured, if } 5 \leq \text{shot clock} < 10 \\ \text{Early, if } 10 \leq \text{shot clock} < 15 \\ \text{Quick, if } 15 \leq \text{shot clock} < 20 \\ \text{Hasty, if } \text{shot clock} \geq 20 \end{cases}$$



Game Clock



Touch Time



Data Processing

Score M

Defensive
Pressure

Shot
Segmentation



Time-related
information



Hot Hand



Physical
Attributes

Data Processing



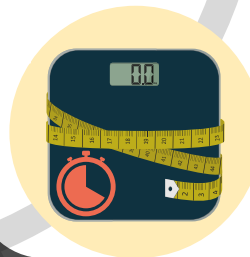
Height



Weight



Speed



Physical
Attributes

Score M

Defensive
Pressure

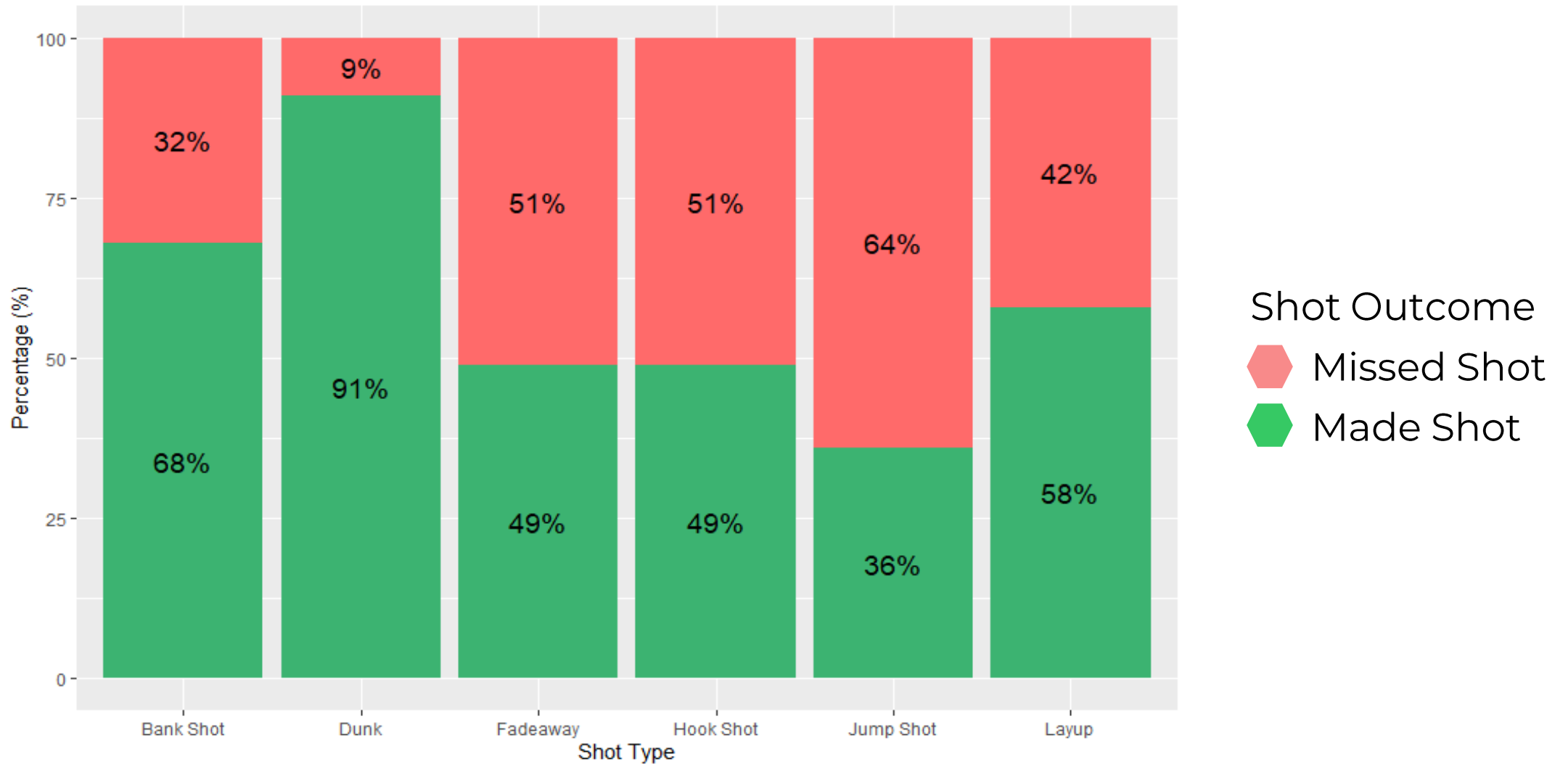
Shot
Segmentation

Time related
information

Hot Hand

EDA

Shots made per shot type

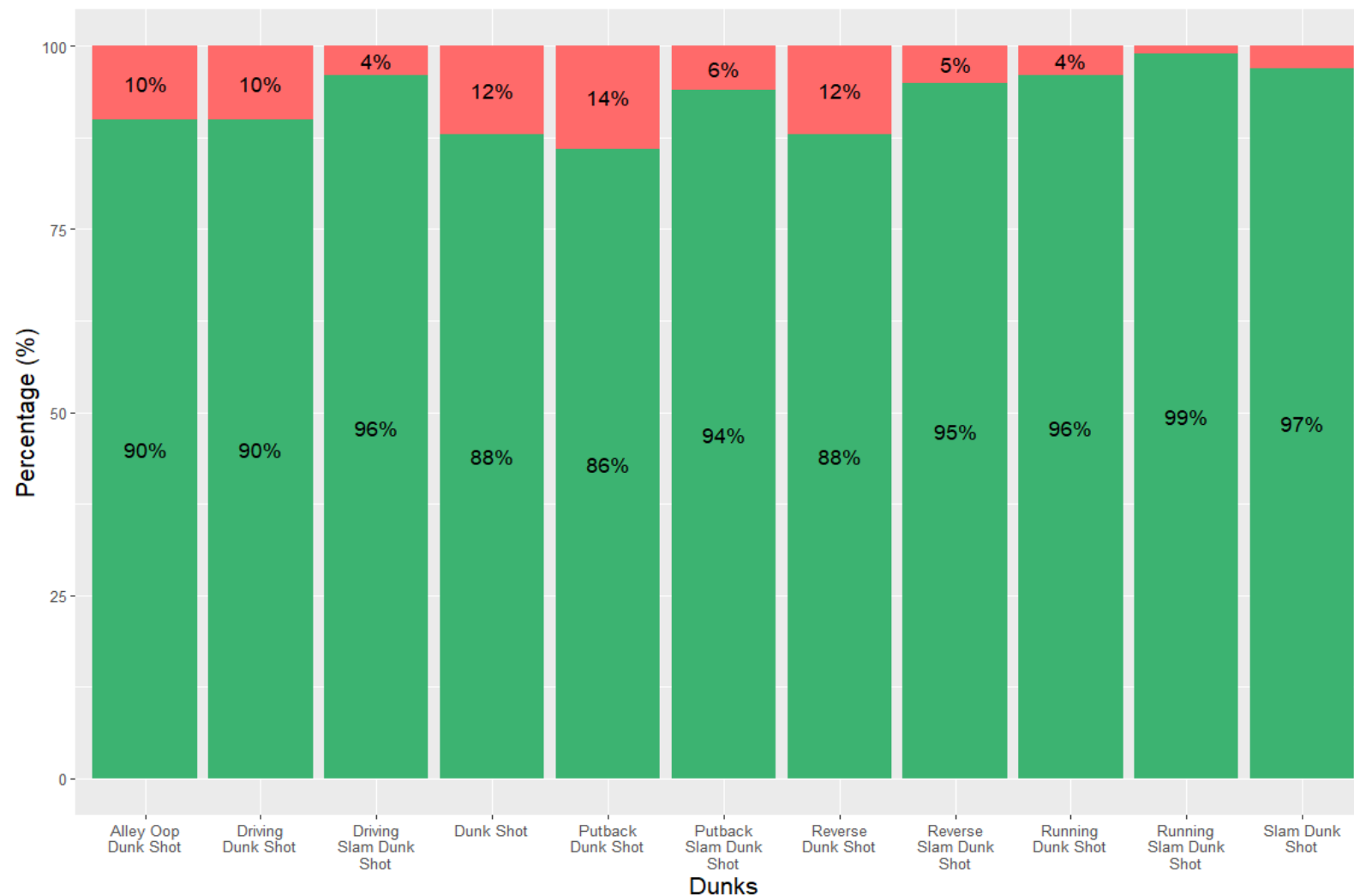




EDA (Cont'd)



Dunks made per shot type



Shot Outcome

- Missed Dunk
- Made Dunk



AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

04 Methodology

05 Analysis Variants

06 Discussion and Conclusions

2014-15 Season Insights



2014-15 Season Insights

Field Goal
Types



Team
Strategies



Shooting
Performance



Two-way
ers



Defensive
Protectors



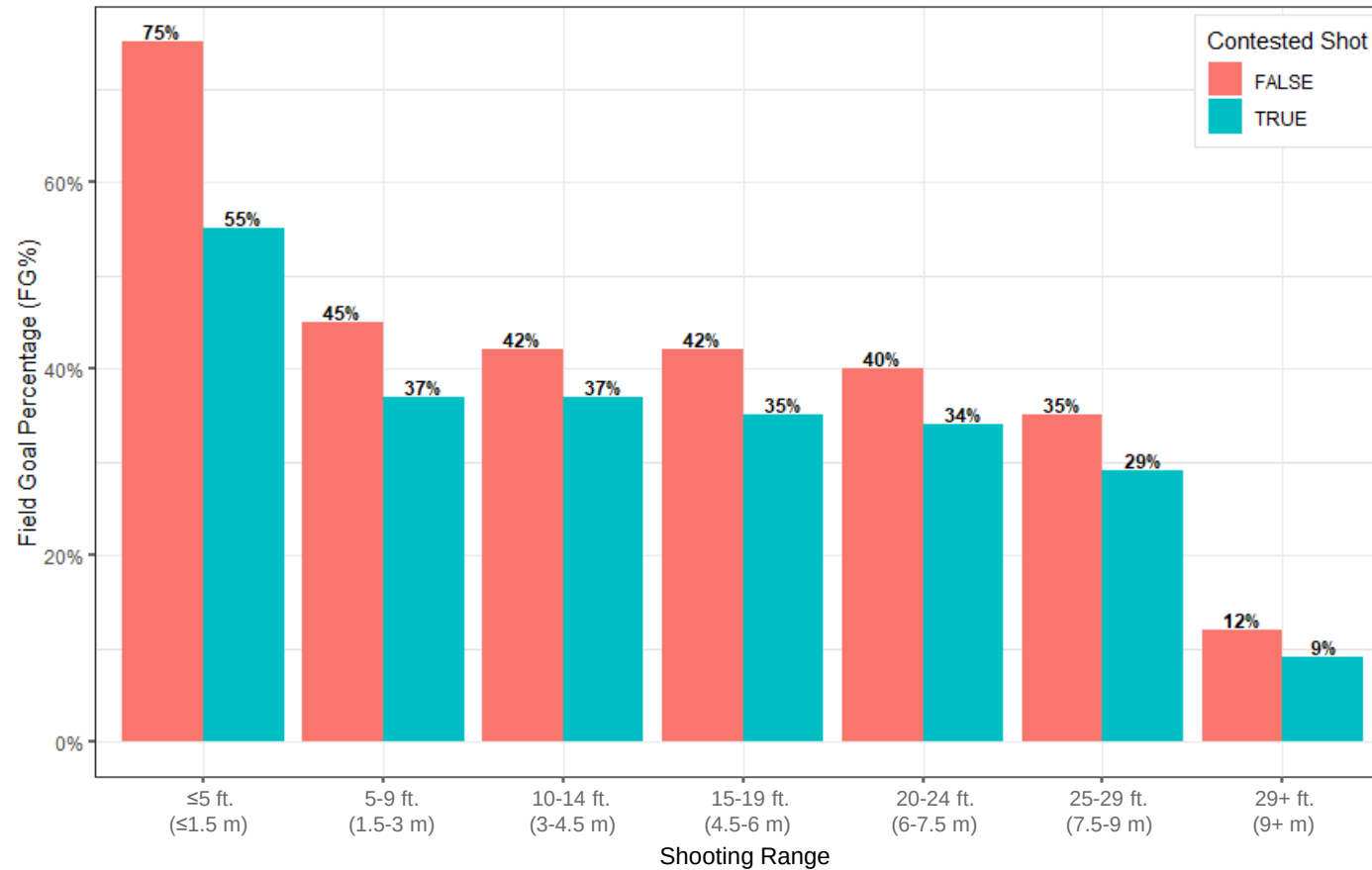


FG% per Distance



FG% per range of shooting distance

2014-15 NBA Season



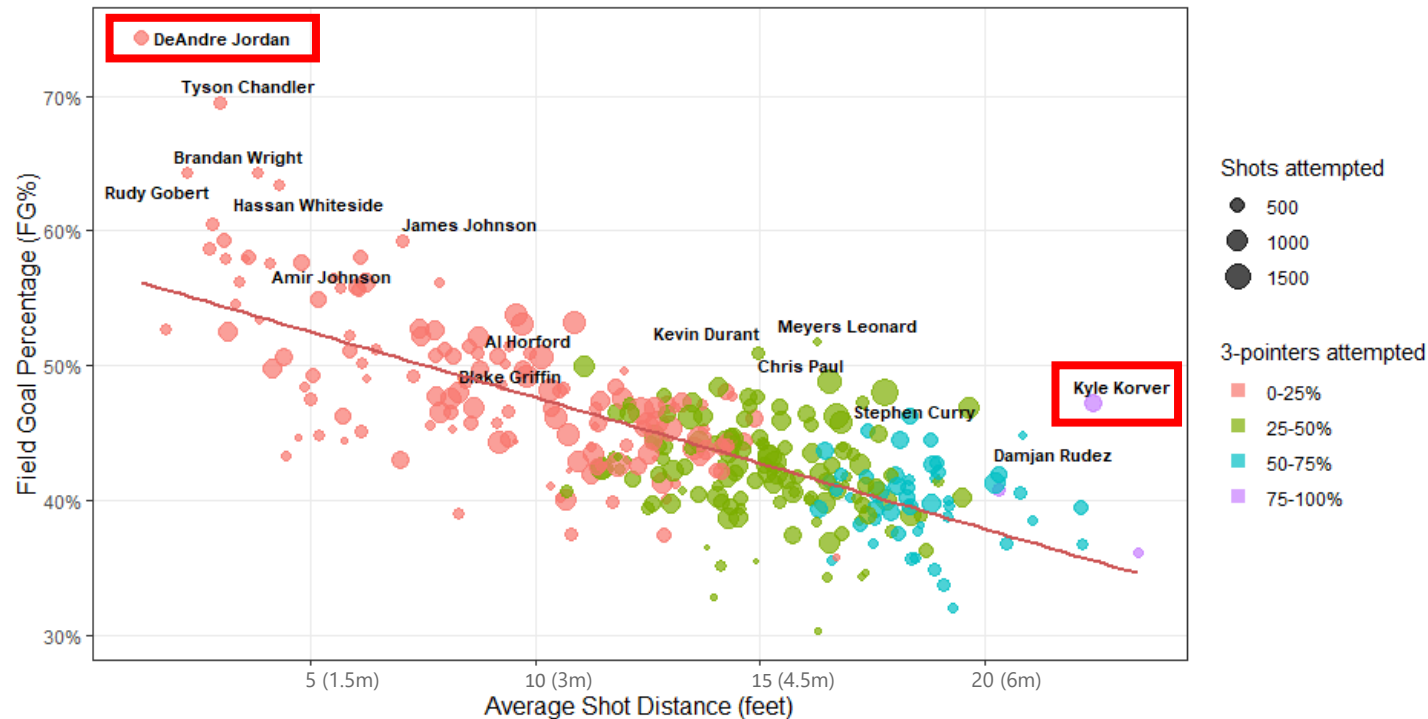


FG% vs. Distance



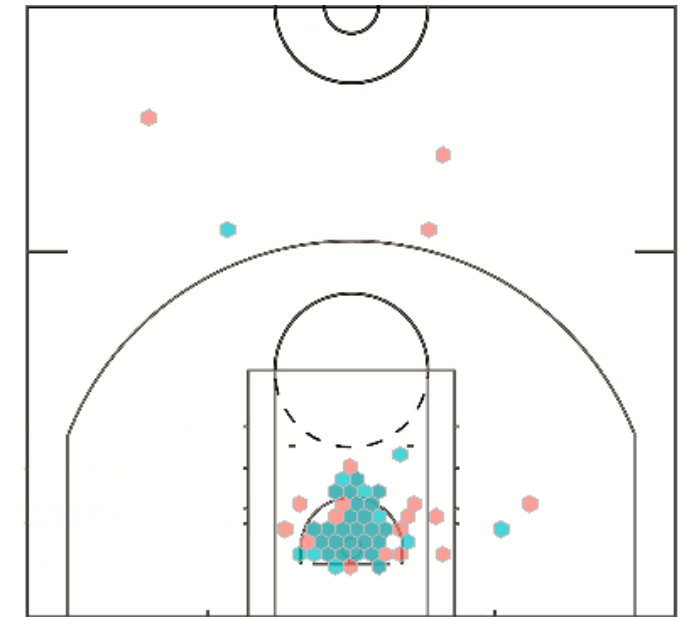
Evaluation of players' efficiency per shooting distance

2014-15 NBA Season (Minimum: 200 shots attempted)



Shot Chart of DeAndre Jordan

2014-15 NBA Season



Missed Shot Made Shot



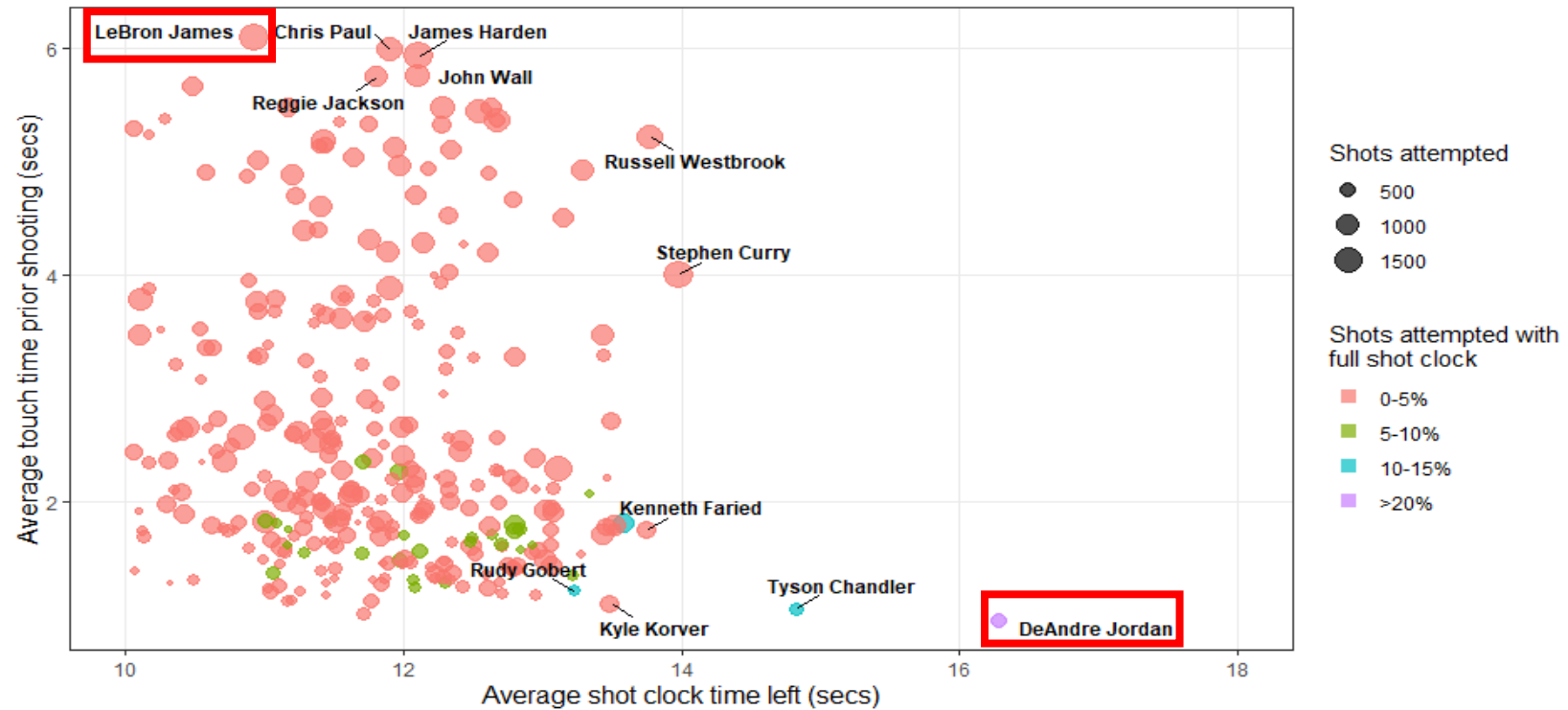


Touch Time vs. Shot Clock



Players' shot attempts based on touch time and shot clock

2014-15 NBA Season (Minimum: 200 shots attempted)



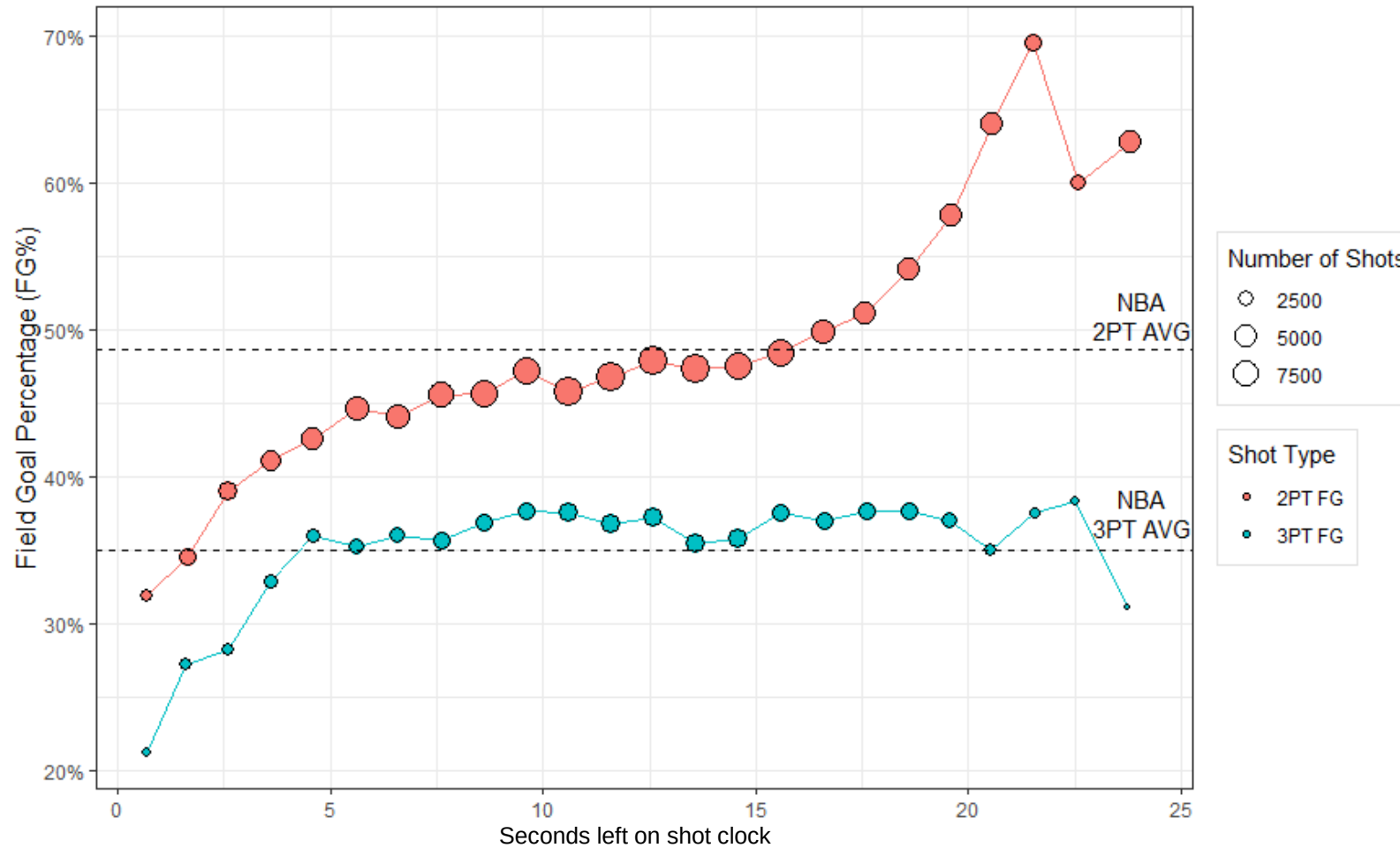


FG% vs. Shot Clock



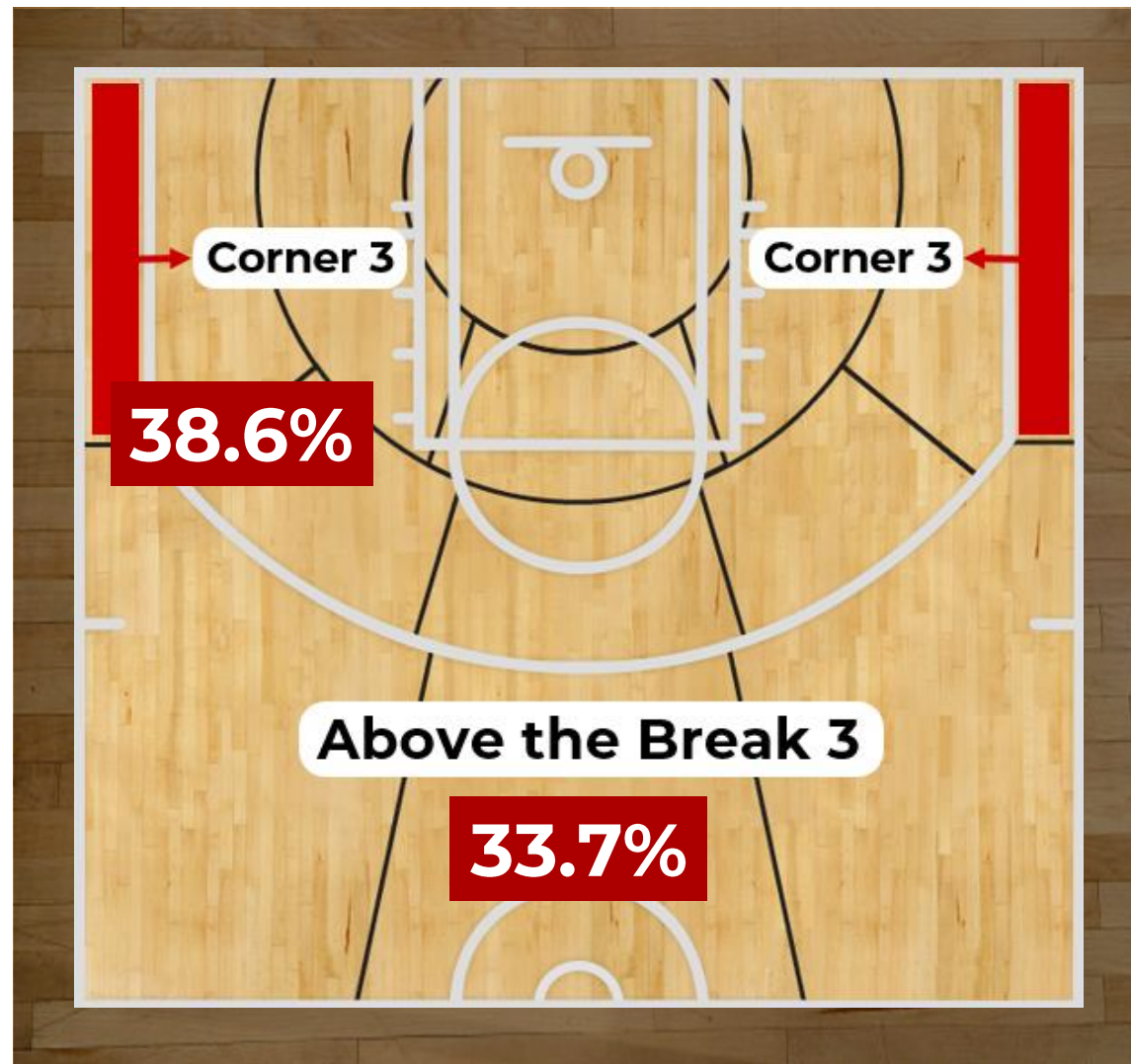
FG% by time left on shot clock

2014-15 NBA Season





FG% per 3-pt FG



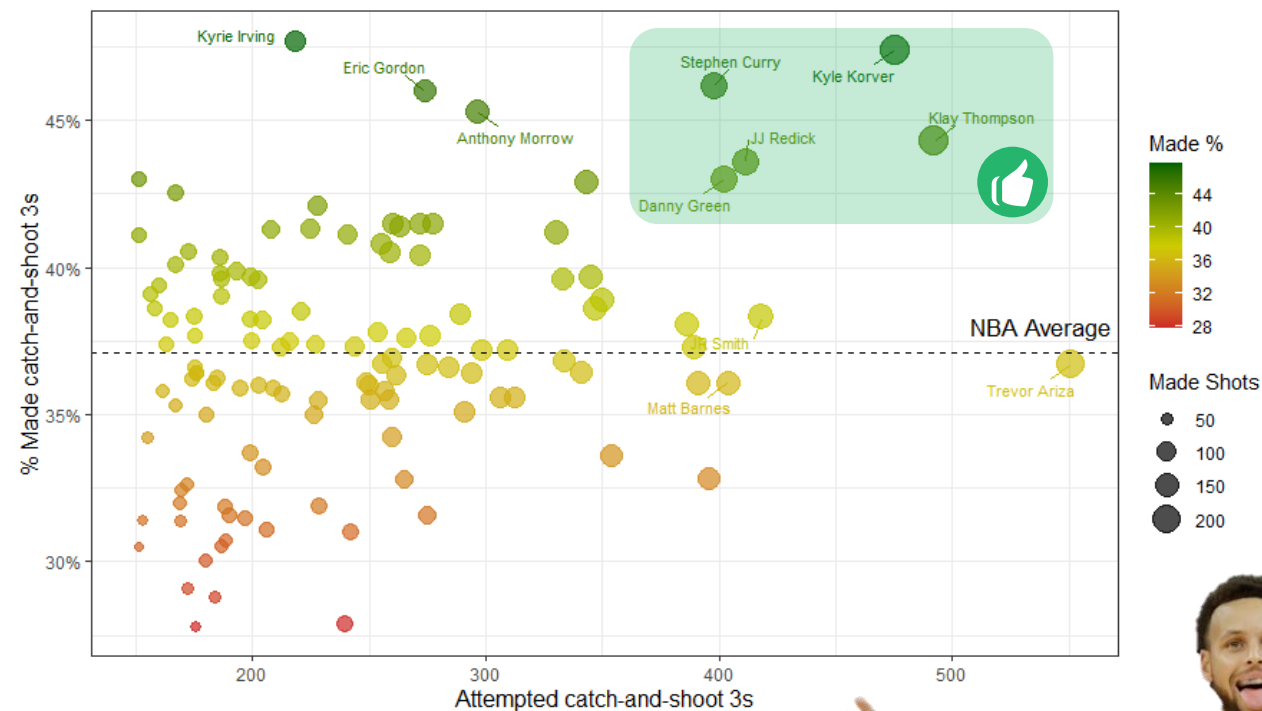


3-pt FG Shooting Dynamic



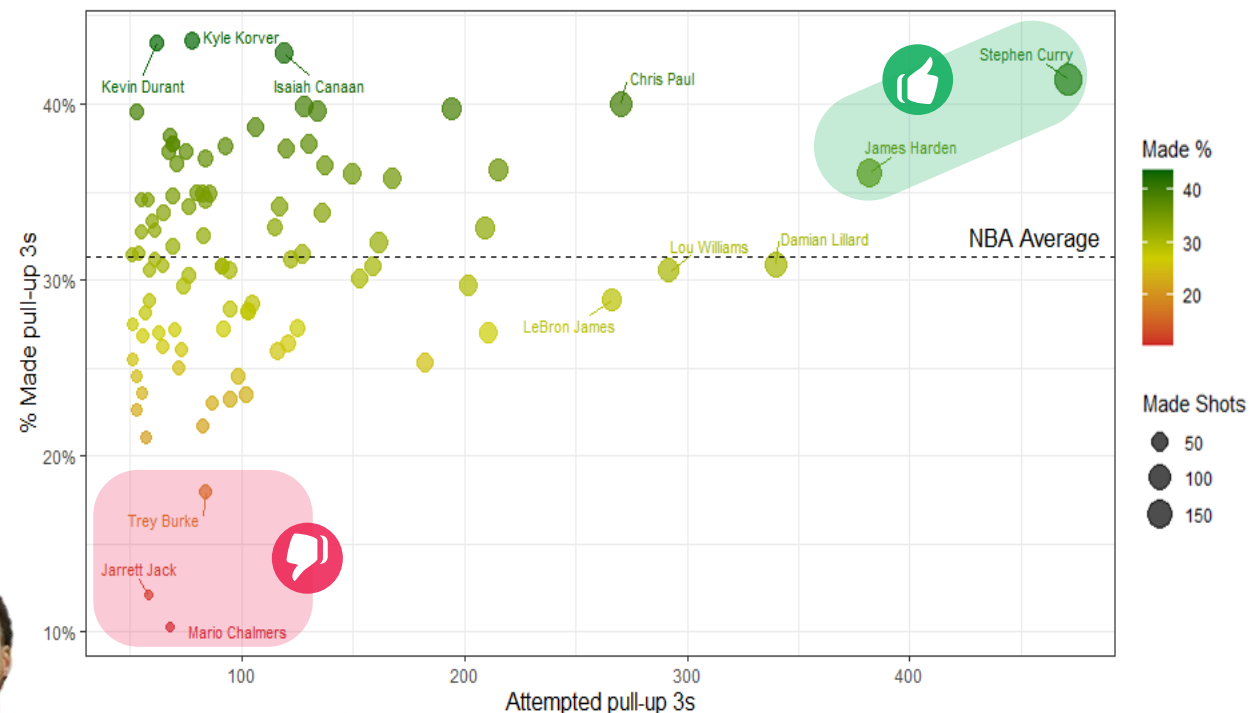
Catch-And-Shoot 3-point FG

2014-15 NBA Season (Minimum: 150 attempts)



Pull-up 3-point FG

2014-15 NBA Season (Minimum: 150 attempts)



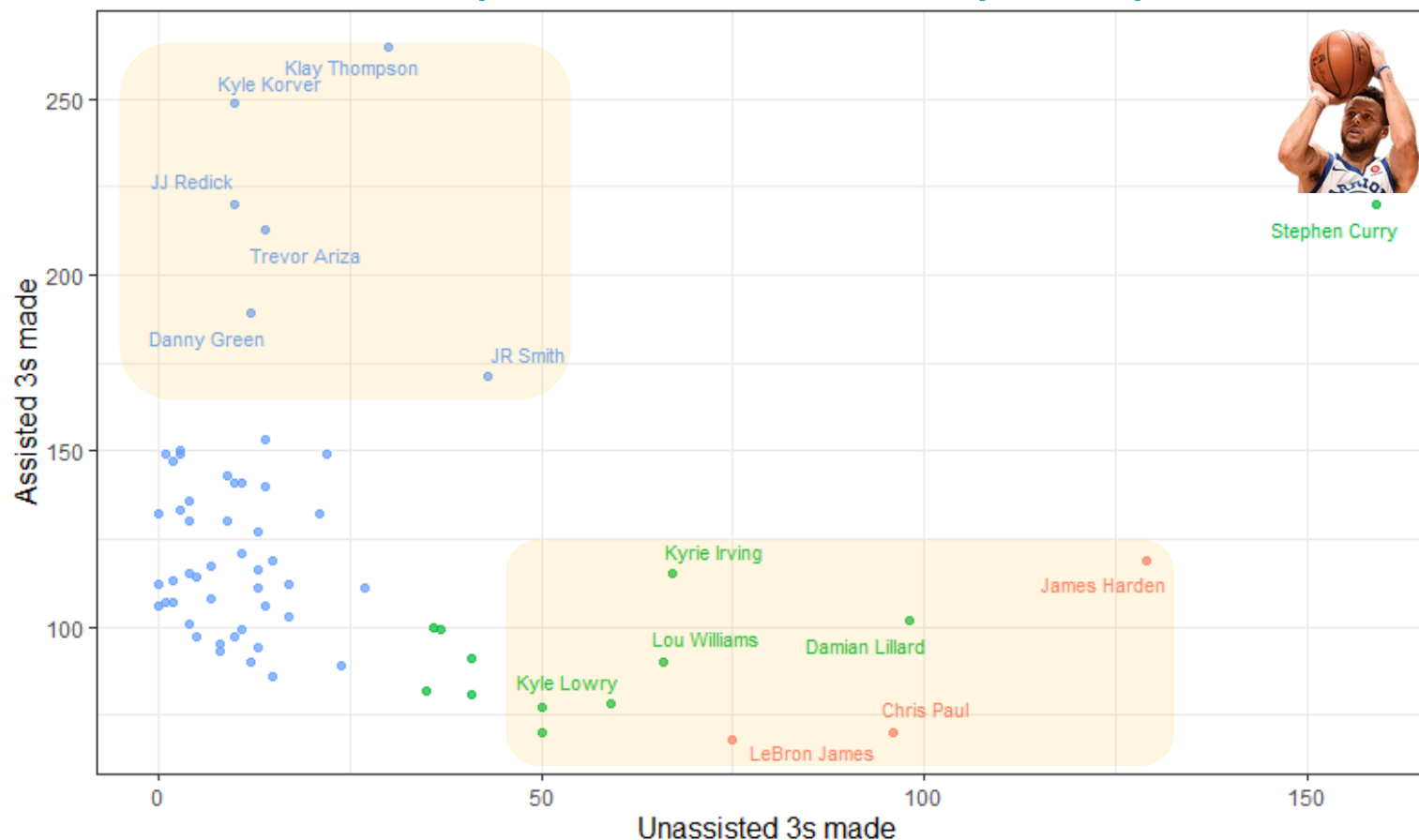


Assisted 3-pt FG



Three-point shooters

2014-15 NBA Season (Minimum: 100 made three-point FG)





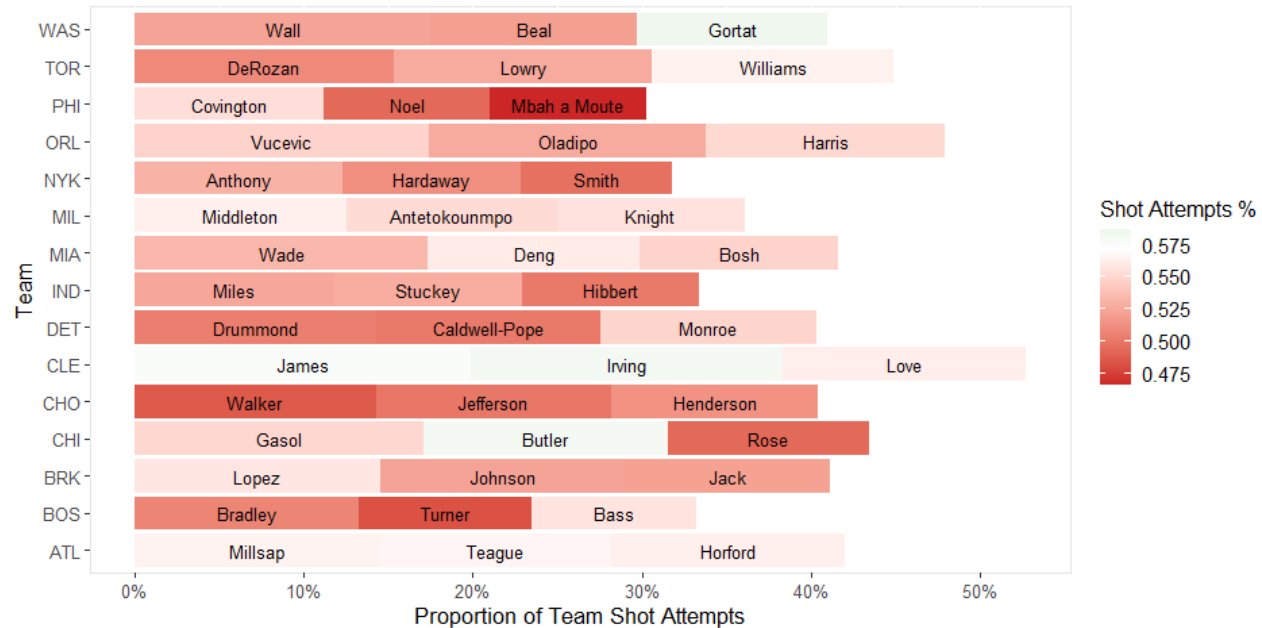
Big Threes



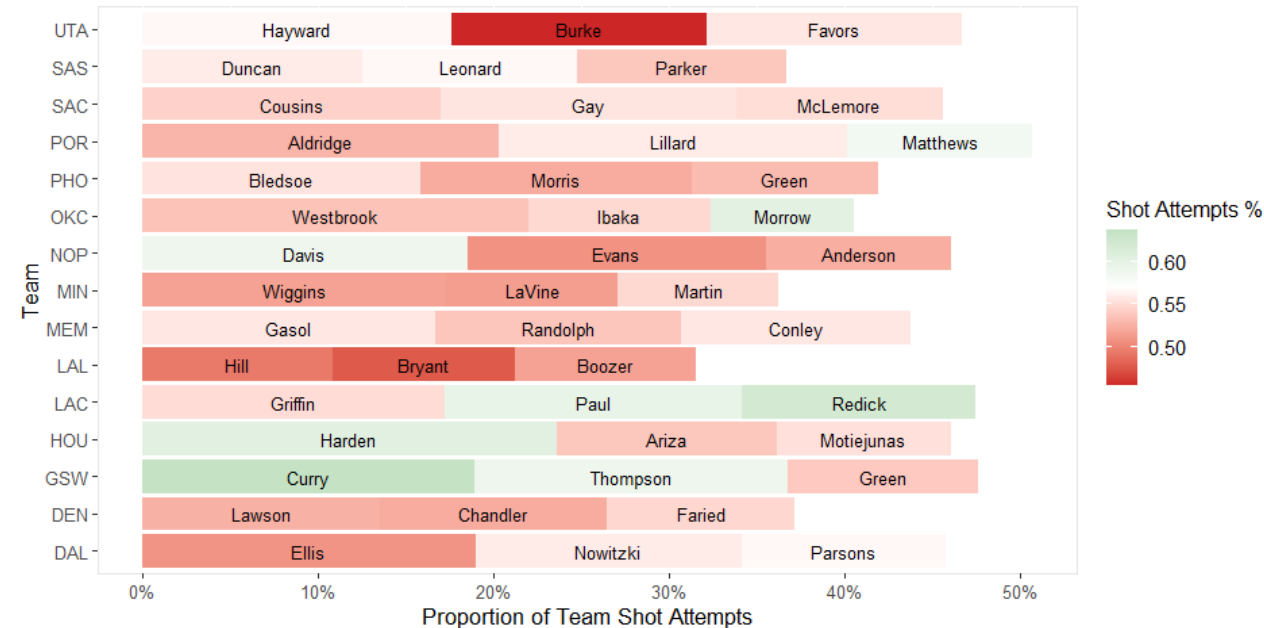
Top 3 players with most shot attempts per team



Eastern Conference



Western Conference



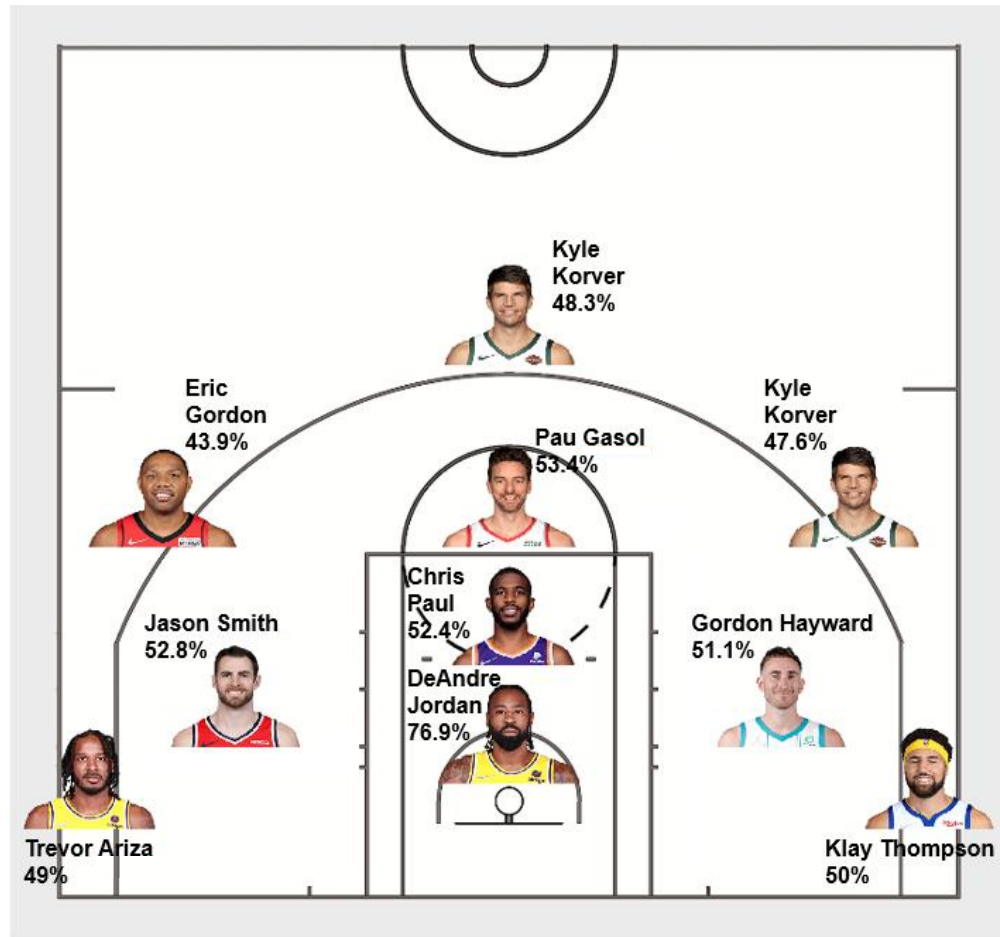


Best & Worst Shooters Per Zone



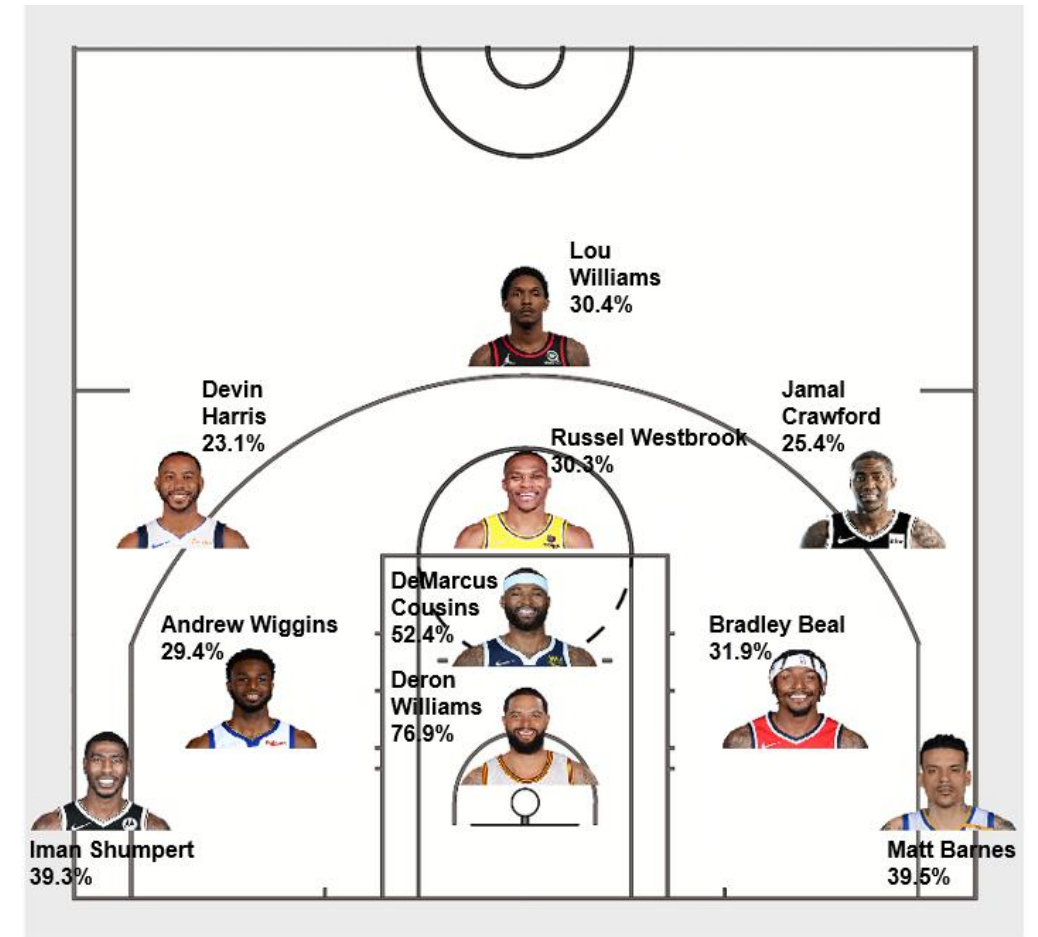
Most Effective Shooters per zone

2014-15 NBA Season (Minimum: 100 shots per zone & 150 shots in the paint)



Least Effective Shooters per zone

2014-15 NBA Season (Minimum: 100 shots per zone & 150 shots in the paint)





Evaluating Defense



Defensive Metrics

Calculated for the shooter's nearest defender



$$\rightarrow \text{Opponent field goal percentage} = \frac{\text{Shots made against defender}}{\text{Shots attempted against defender}}$$

$$\rightarrow \text{Opponent points per possession} = \frac{\text{Points per shot against defender}}{\text{Shots attempted against defender}}$$

$$\rightarrow \text{Average defenders' distance} = \frac{\text{Defender distance}}{\text{Shots attempted against defender}}$$

$$\rightarrow \text{Number of shots contested} = \text{Total shots attempted} \leq 3.5 \text{ ft. (1m) away from defender}$$

$$\rightarrow \text{Average shooting distance} = \frac{\text{Shot distance}}{\text{Shots attempted against defender}}$$

$$\rightarrow \text{Total number of shots closely guarded} = \text{Shots attempted against defender}$$

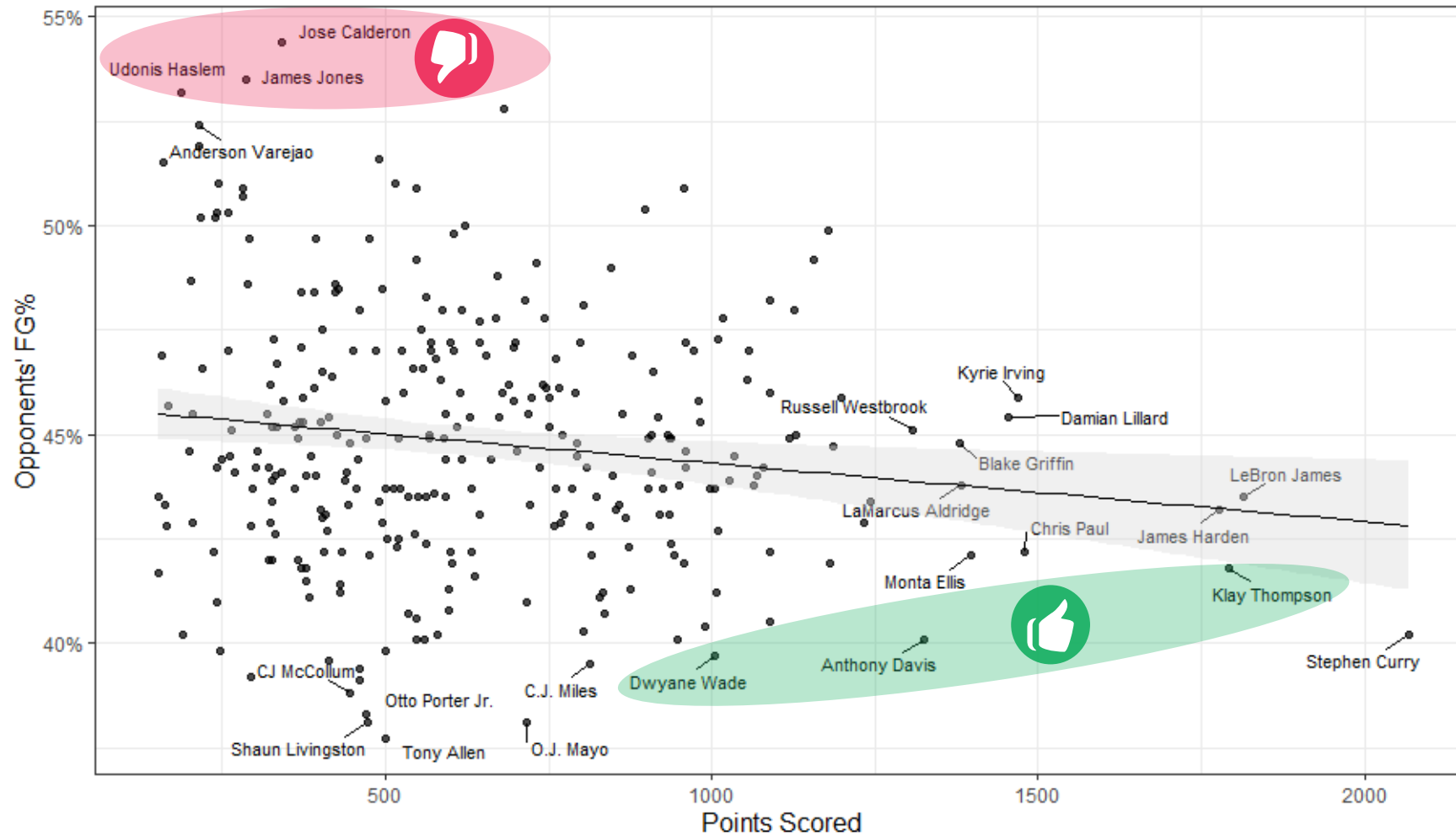


Two-Way Players



Evaluation of Two-Way Players

2014-15 NBA Season (Minimum: 300 shots guarded)



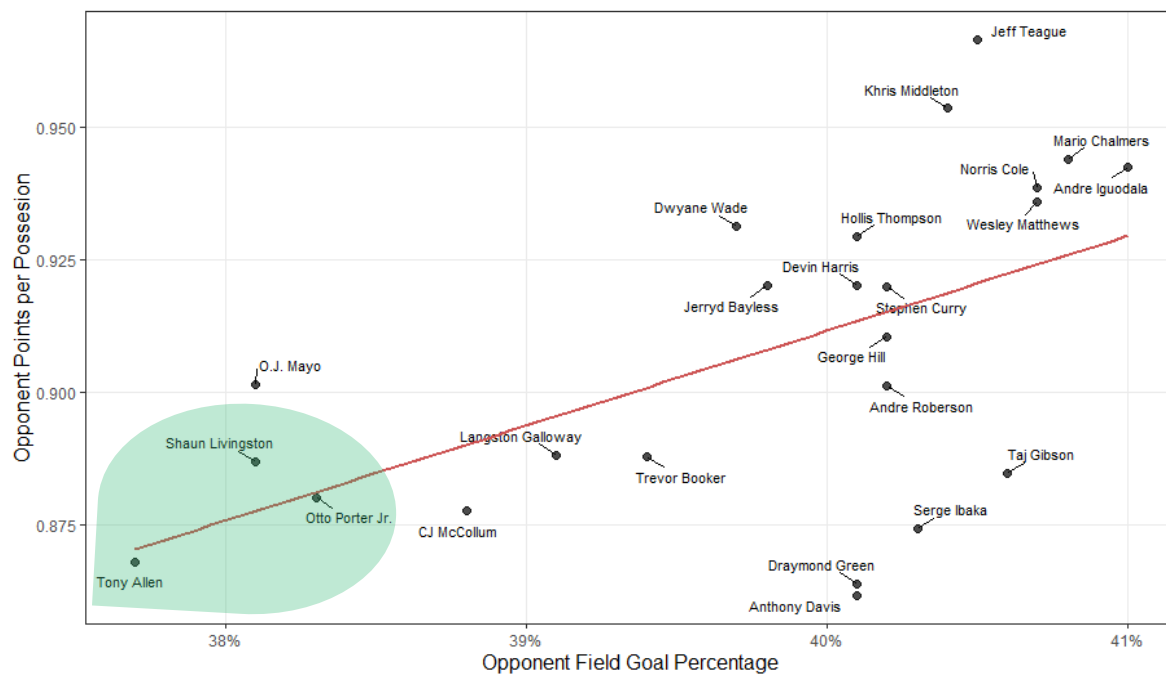


Best Defenders



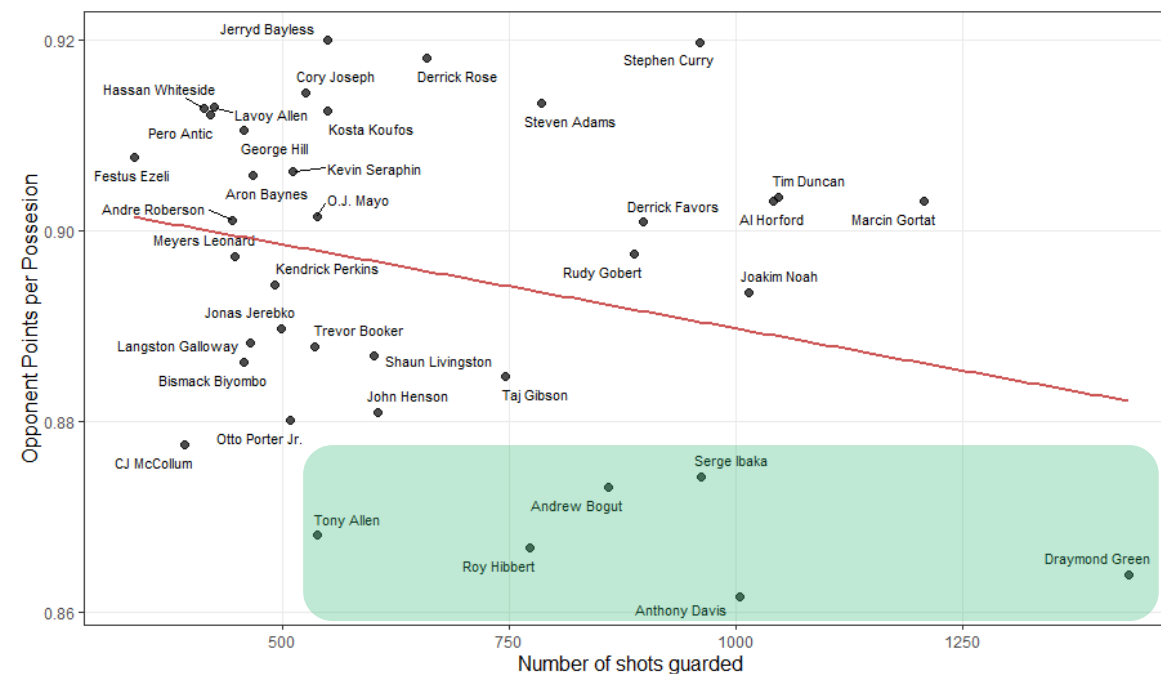
Top Defenders based on Opponents' FG%

2014-15 NBA Season (Minimum: 200 shots contested)



Top Defenders based on Opponents' Points per Possession

2014-15 NBA Season (Minimum: 200 shots contested)



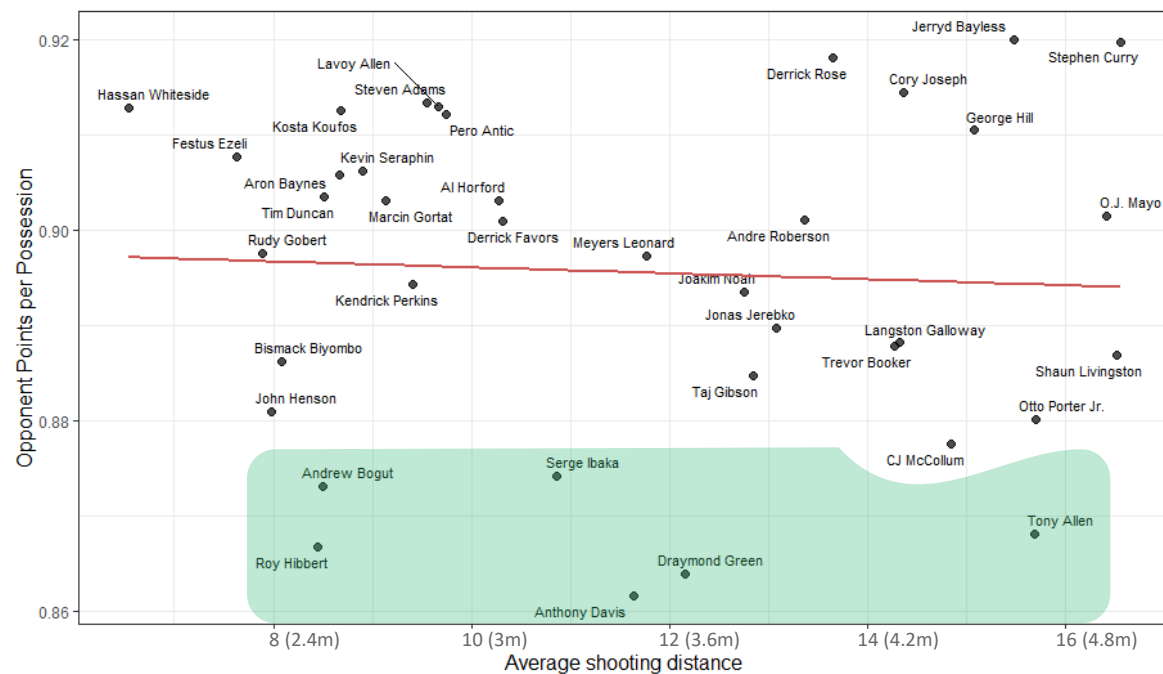


Best Defenders (Cont'd)



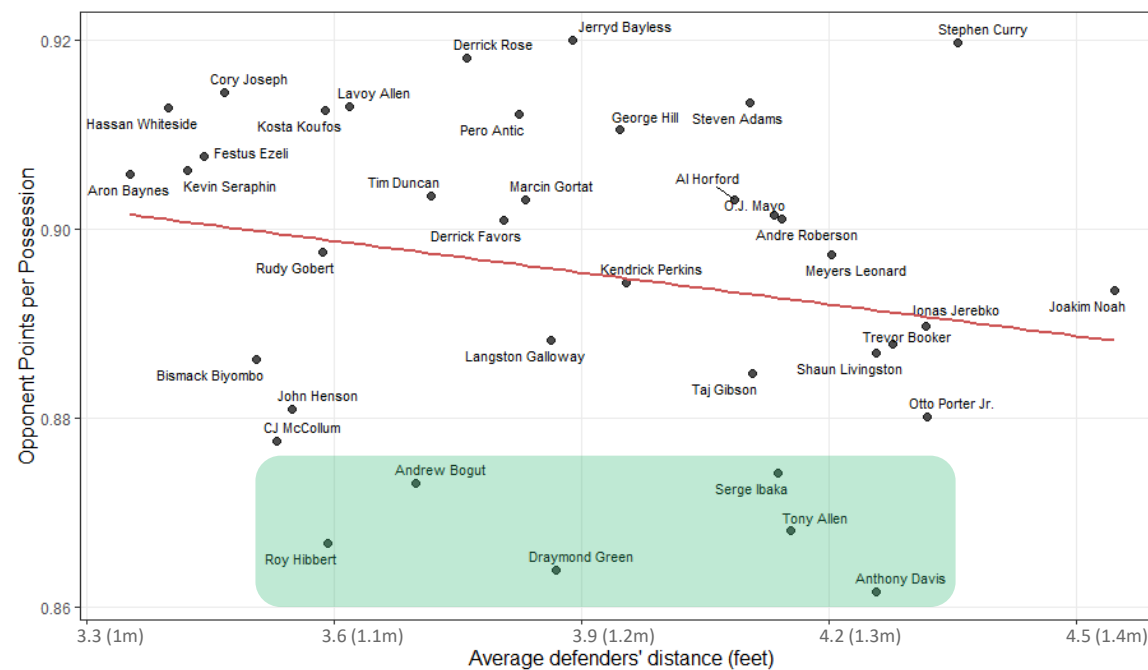
Top Defenders based on average shooting distance

2014-15 NBA Season (Minimum: 200 shots contested)



Top Defenders based on nearest defender's distance

2014-15 NBA Season (Minimum: 200 shots contested)



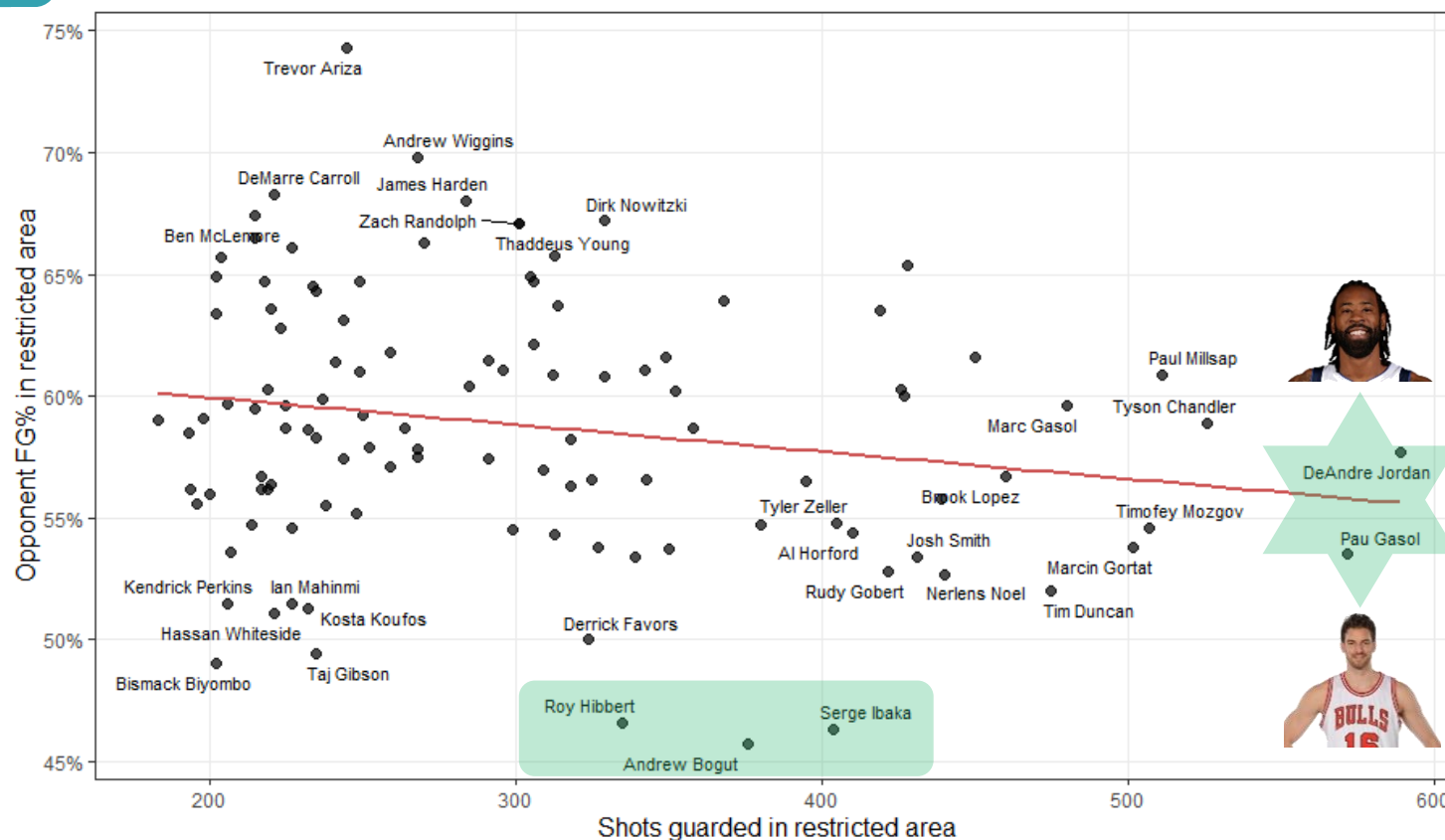


Best Rim Protectors



Top Rim Protectors

2014-15 NBA Season (Minimum: 150 shots contested)



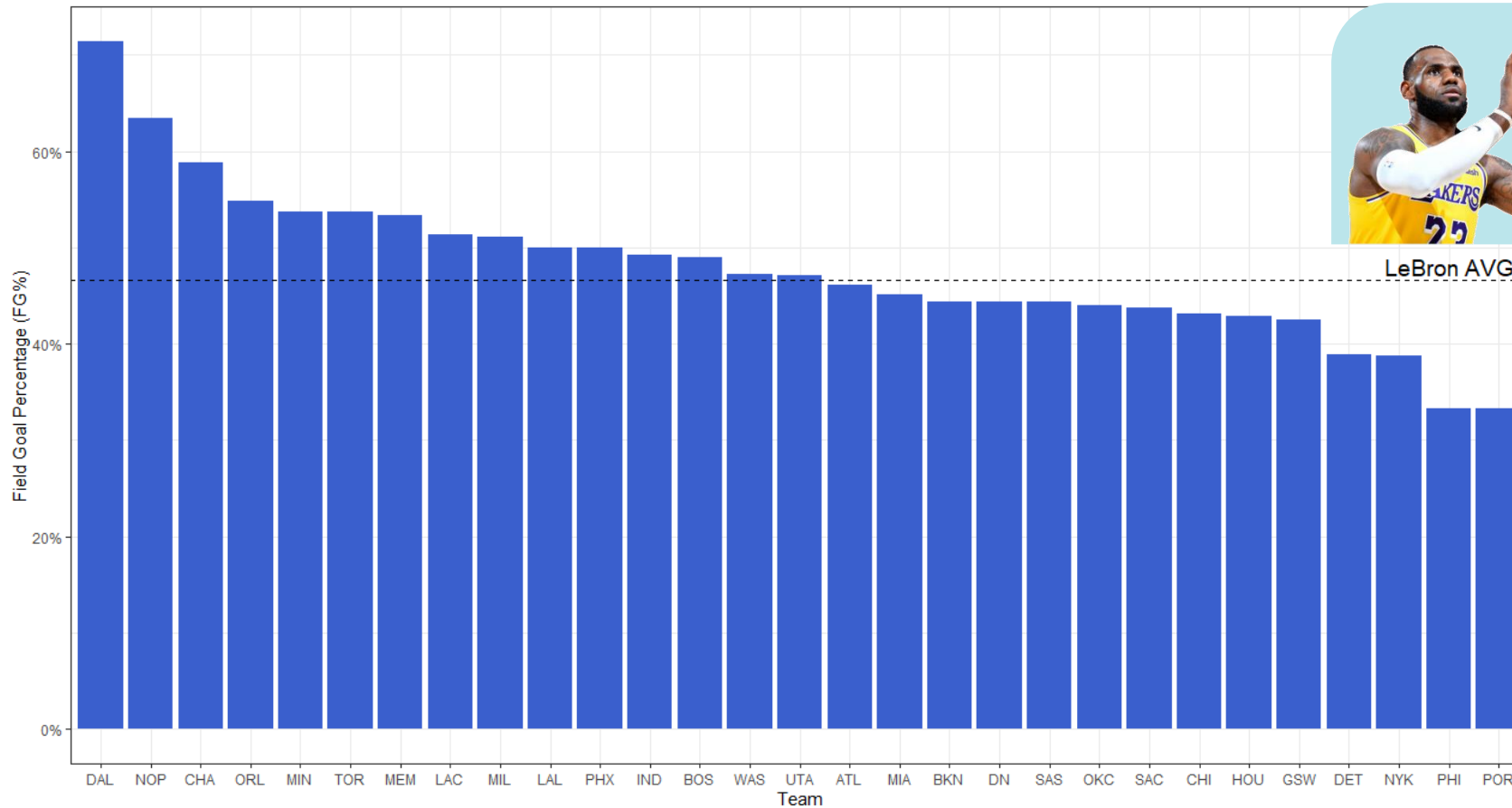


LeBron's FG% per Team



LeBron James' FG% per team

2014-15 NBA Season



LeBron AVG



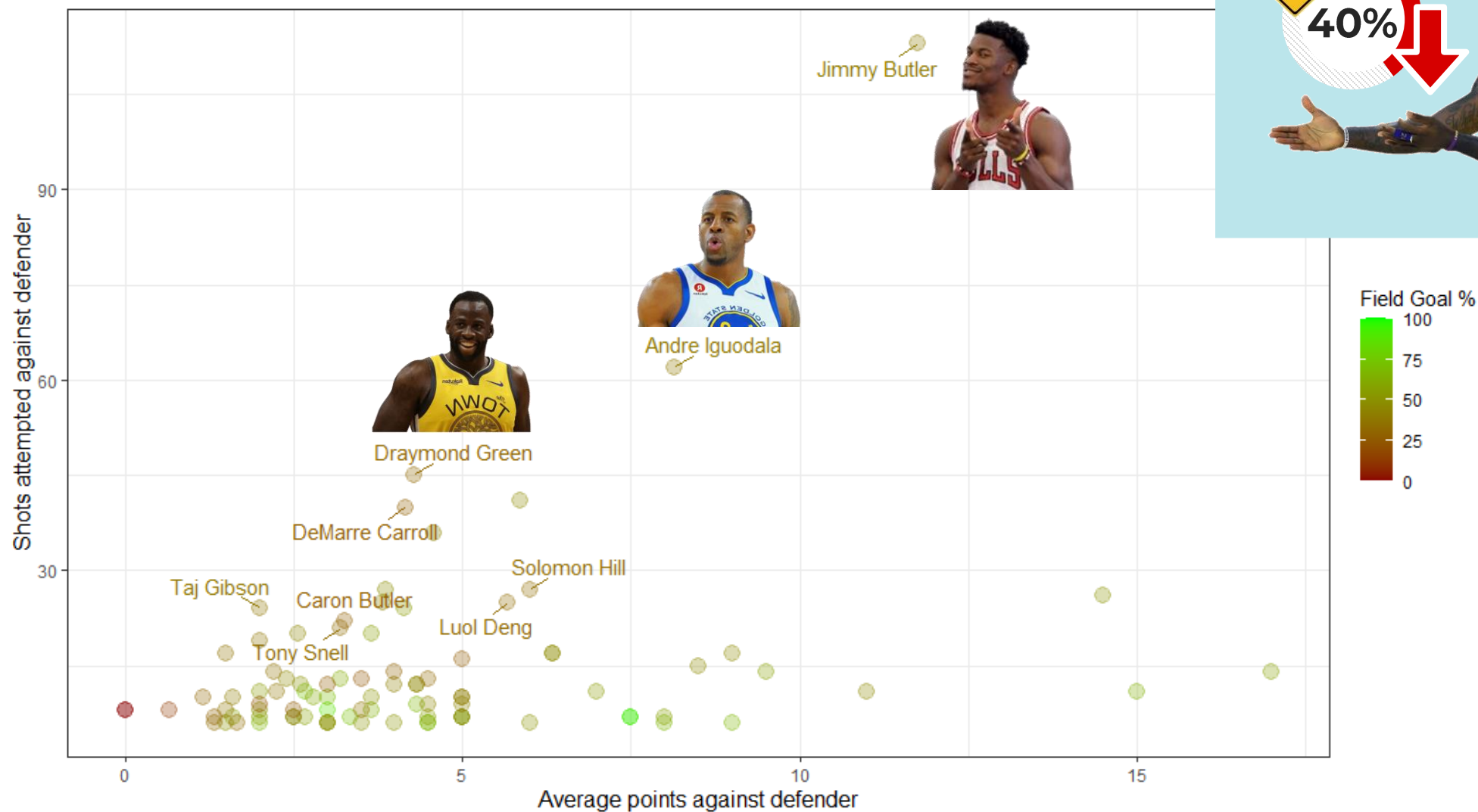


LeBron Stoppers



Top defenders against LeBron James

2014-15 NBA Season (Minimum: 5 shots attempted per defender)

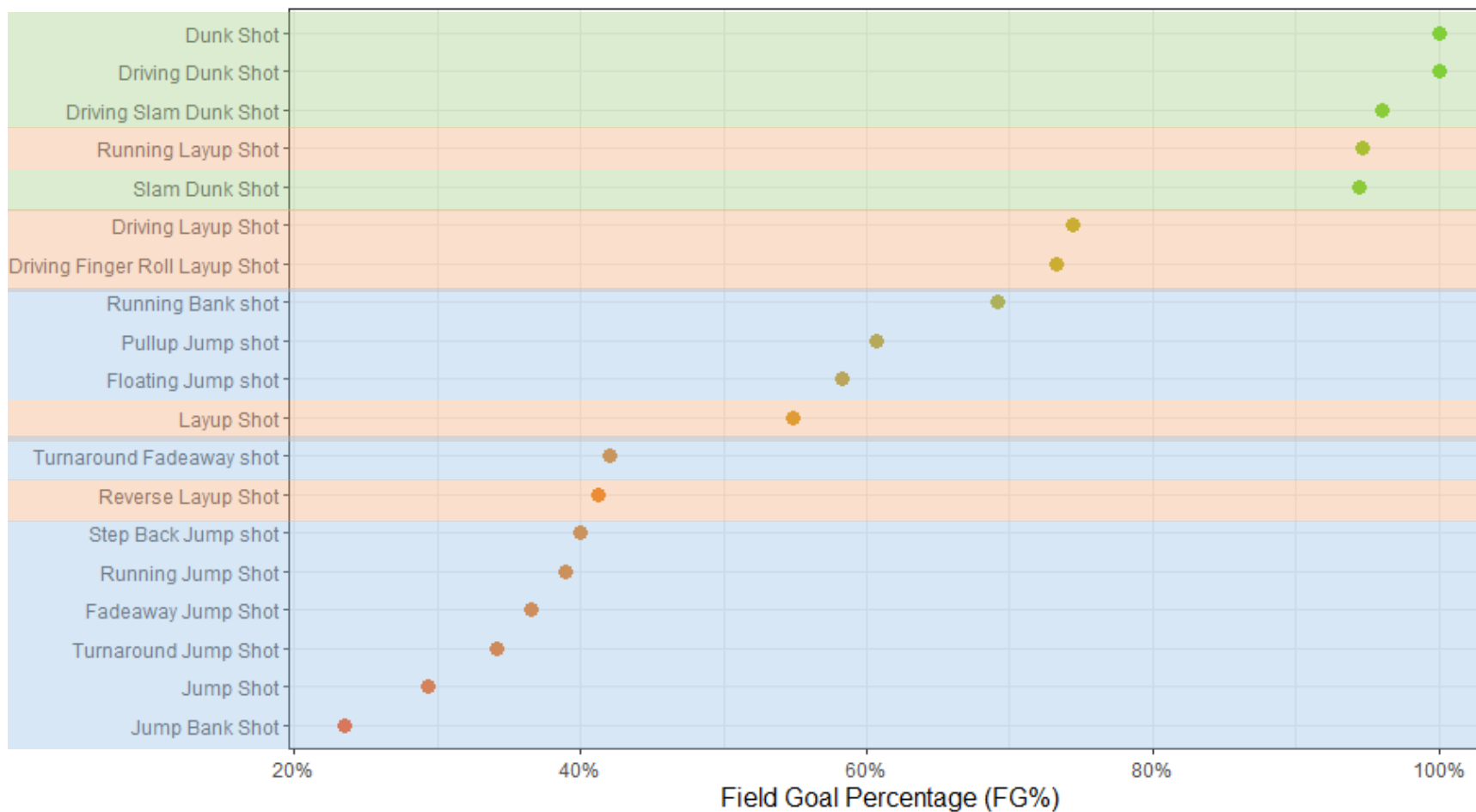




LeBron's FG% per Shot Type



LeBron James FG% per shot type
2014-15 NBA Season (Minimum: 15 shots attempts)

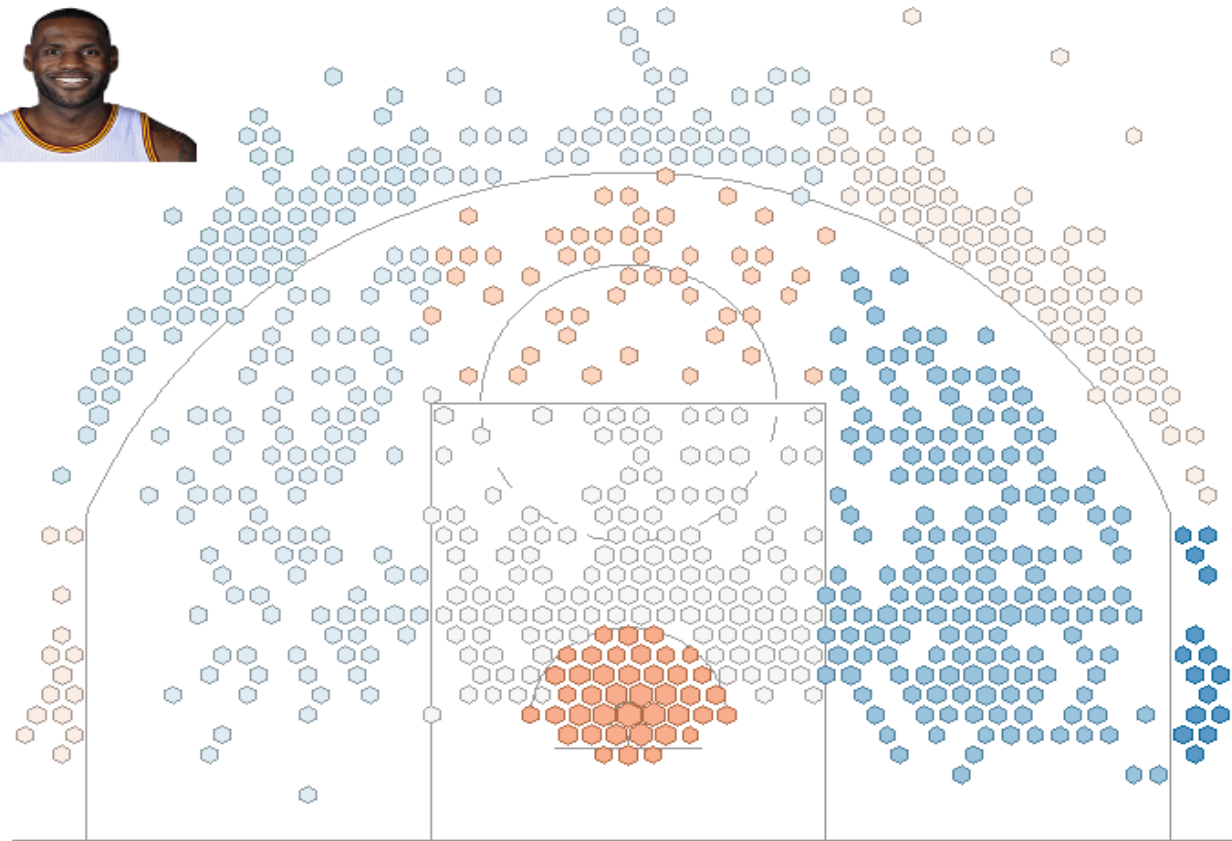




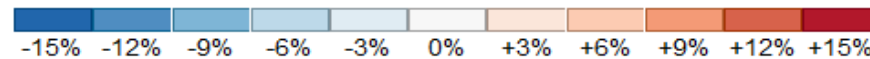
LeBron's Shot Chart



Evaluation of LeBron James' shot efficiency 2014-15 NBA Season



Field Goal Percentage vs. League Average



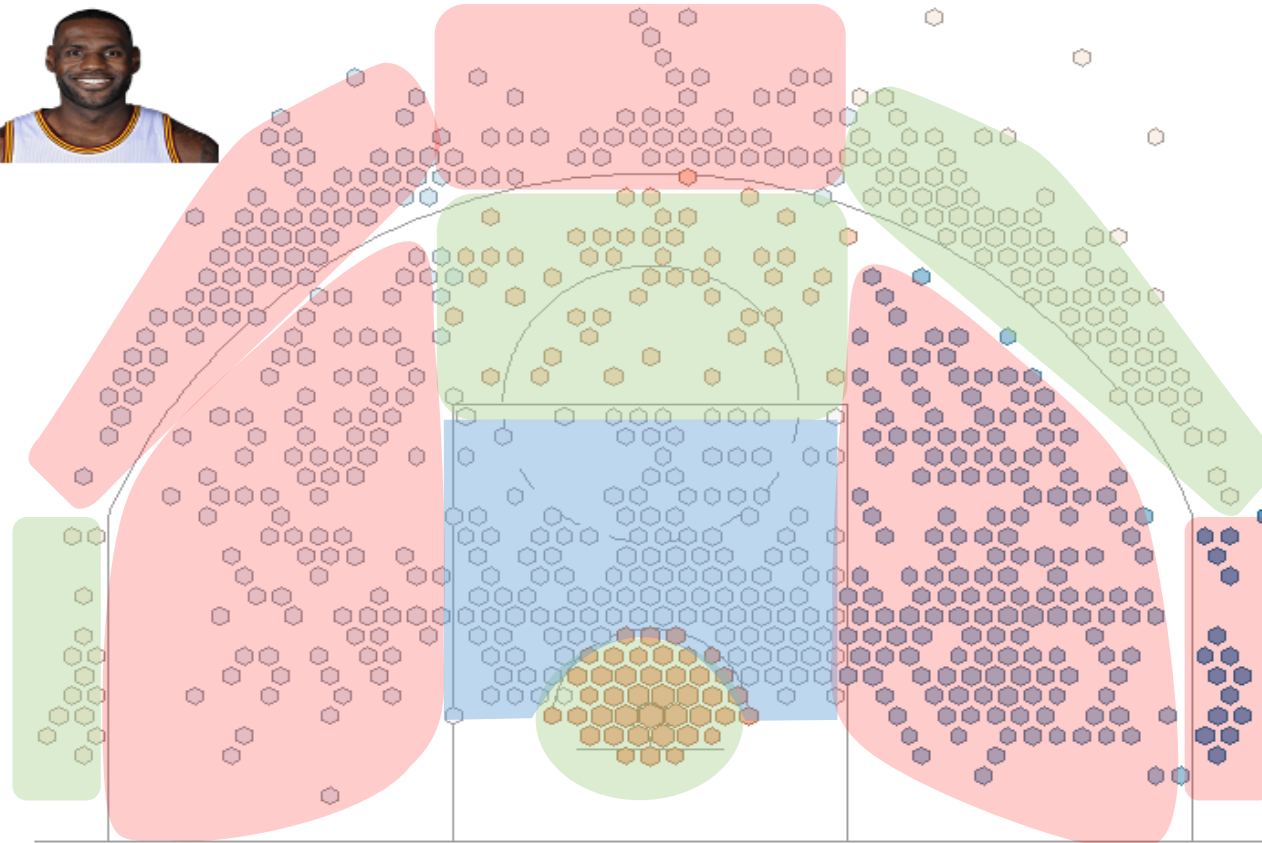


LeBron's Shot Chart



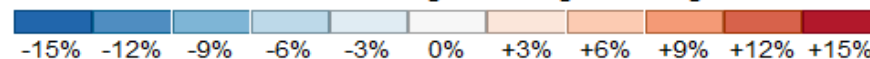
Evaluation of LeBron James' shot efficiency

2014-15 NBA Season



-  Better
-  Worse
-  Average

Field Goal Percentage vs. League Average





AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

04 Methodology

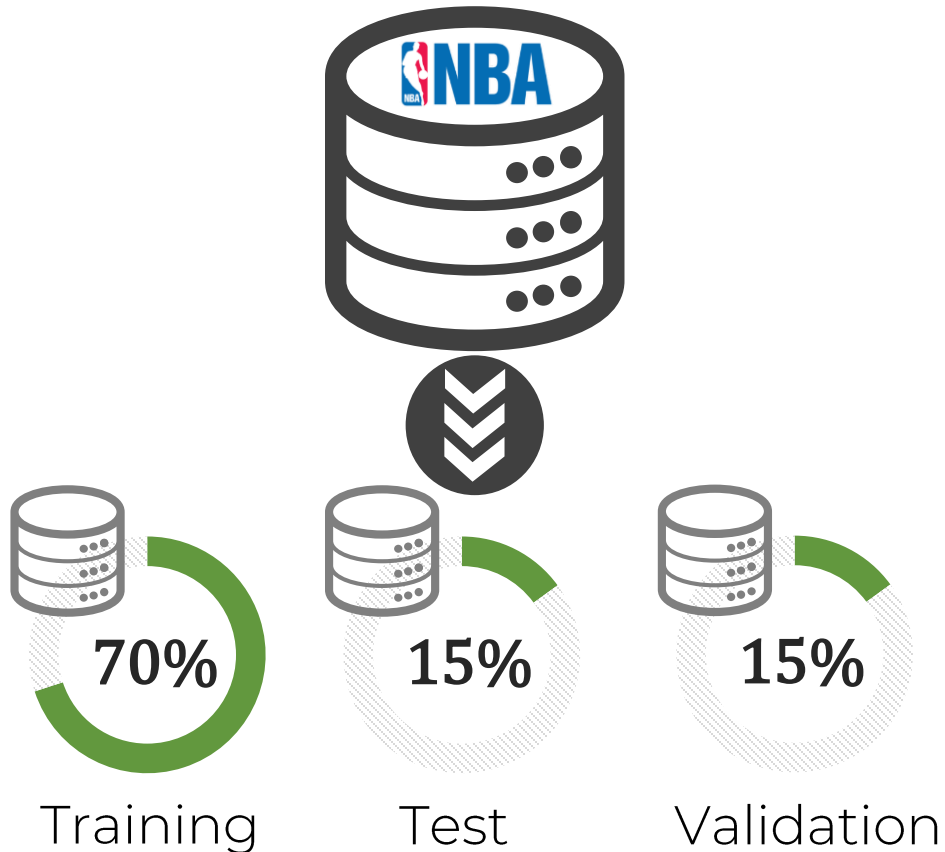
05 Analysis Variants

06 Discussion and Conclusions

Methodology



➤ Data Splitting



➤ Performance Metrics

		Predicted	
		FALSE	TRUE
Actual	FALSE	True Negative (TN)	False Positive (FP)
	TRUE	False Negative (FN)	True Positive (TP)

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

➤ Accuracy Benchmark

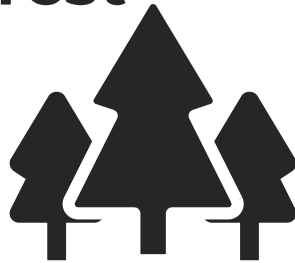


Model Building

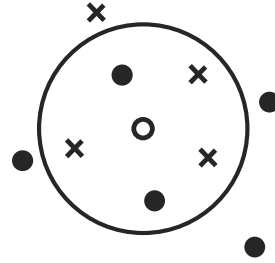


**Decision
Tree**

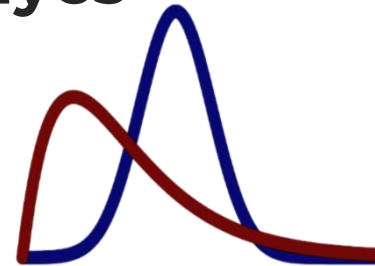
Random
Forest



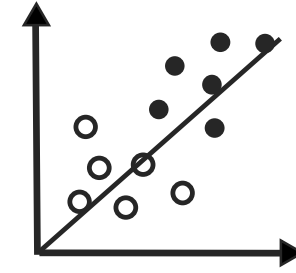
**k-Nearest
Neighbors**



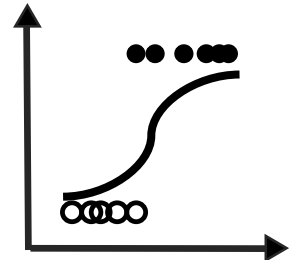
Naïve
Bayes



Linear
Probability



Logistic
Regression



Model Building

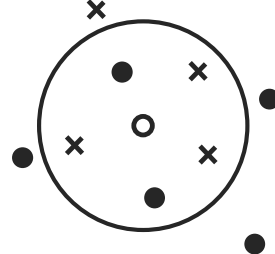


**Decision
Tree**

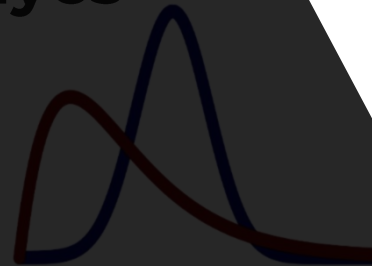
Random
Forest



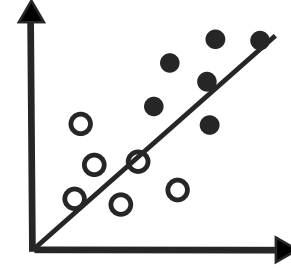
**k-Nearest
Neighbors**



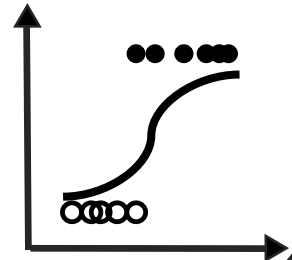
Naïve
Bayes



Linear
Probability



Logistic
Regression



Model Building



**Feature
Selection**

Random
Forest



k-Nea
Neigh



**Hyper-parameter
Tuning**

Support
Vector
Machines

Lin
Proba

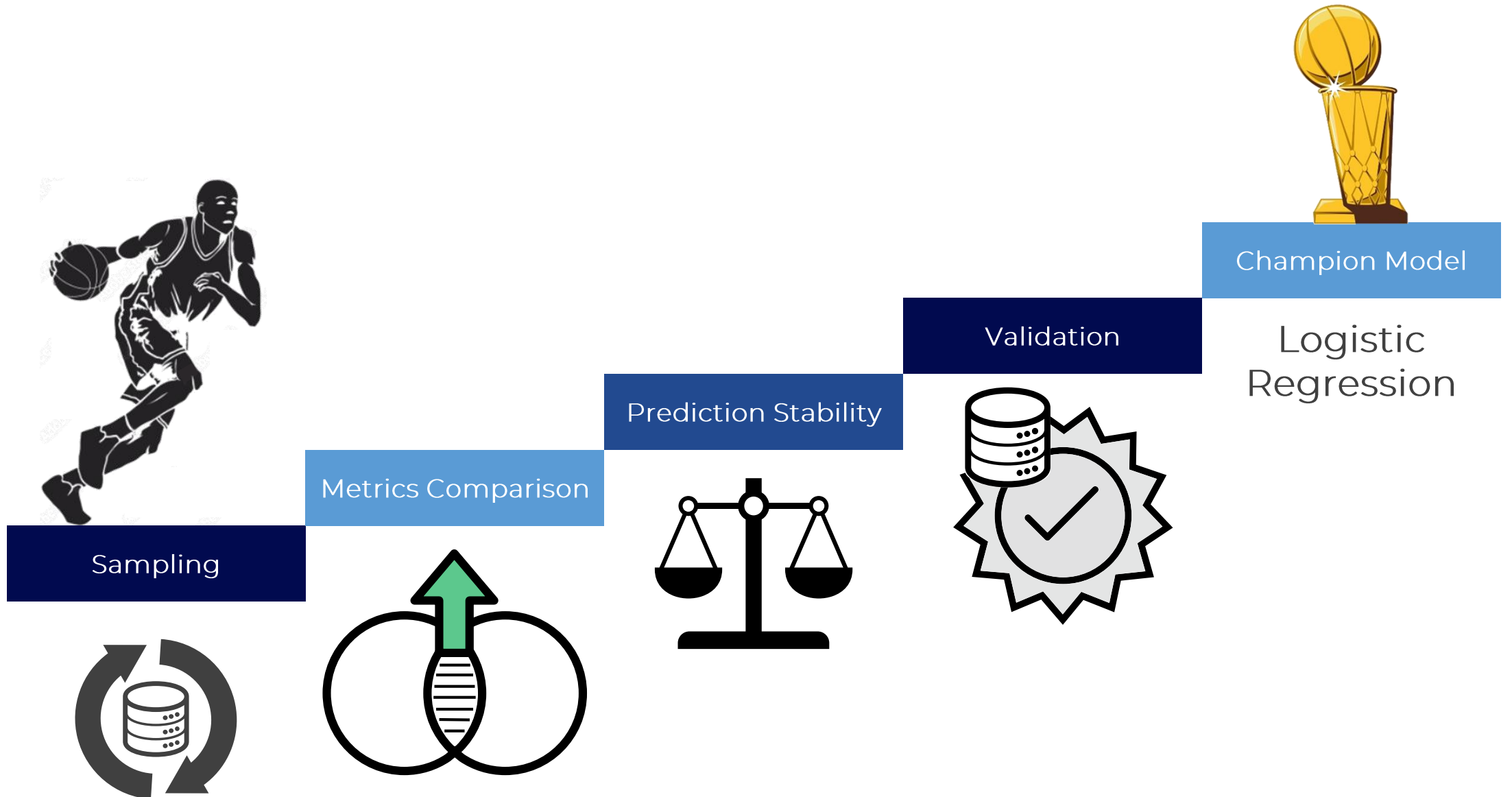


**Feature
Importance**

Decision
Tree



Model Evaluation





AGENDA

01 Introduction

02 Data Preparation

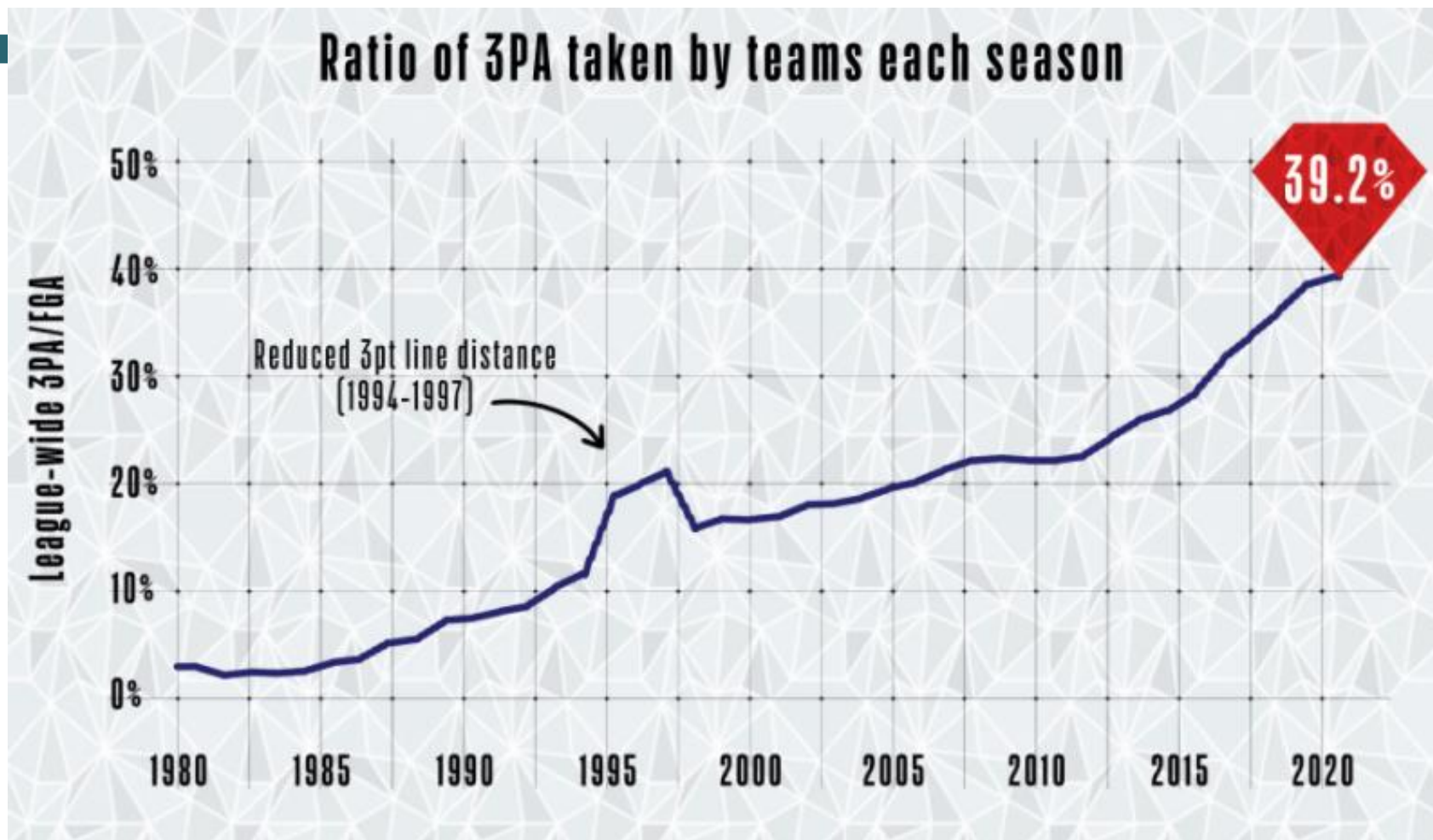
03 2014-15 NBA Season Insights

04 Methodology

05 Analysis Variants

06 Discussion and Conclusions

3-Pointer Proliferation



Efficiency



Spacing



Analysis Variants





AGENDA

01 Introduction

02 Data Preparation

03 2014-15 NBA Season Insights

04 Methodology

05 Analysis Variants

06 Discussion and Conclusions



“

**“All good things must
come to an end.”**

— Geoffrey Chaucer



Discussion and Conclusions



Findings



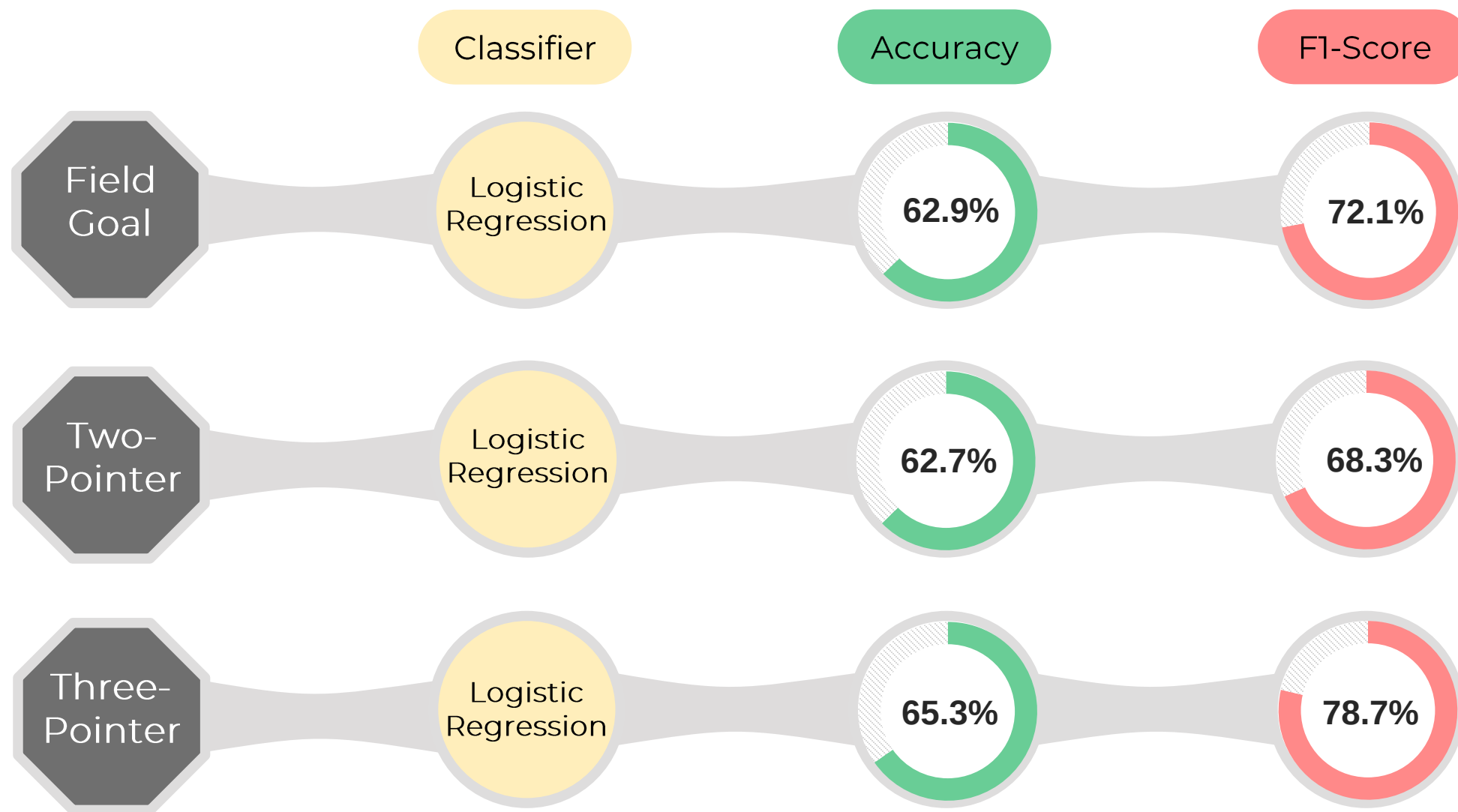
Applications



Future Work

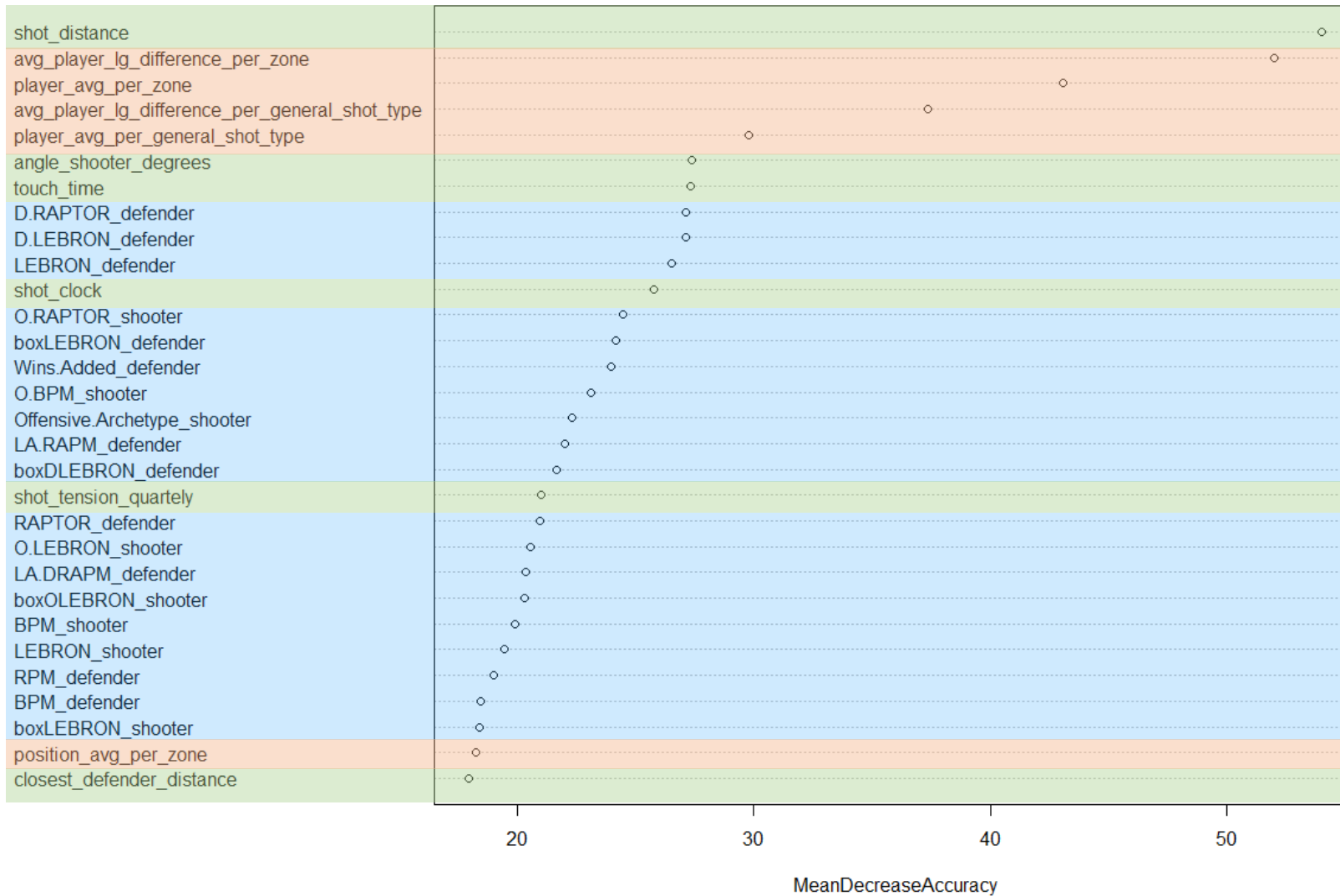


Results Overview



All Field Goals

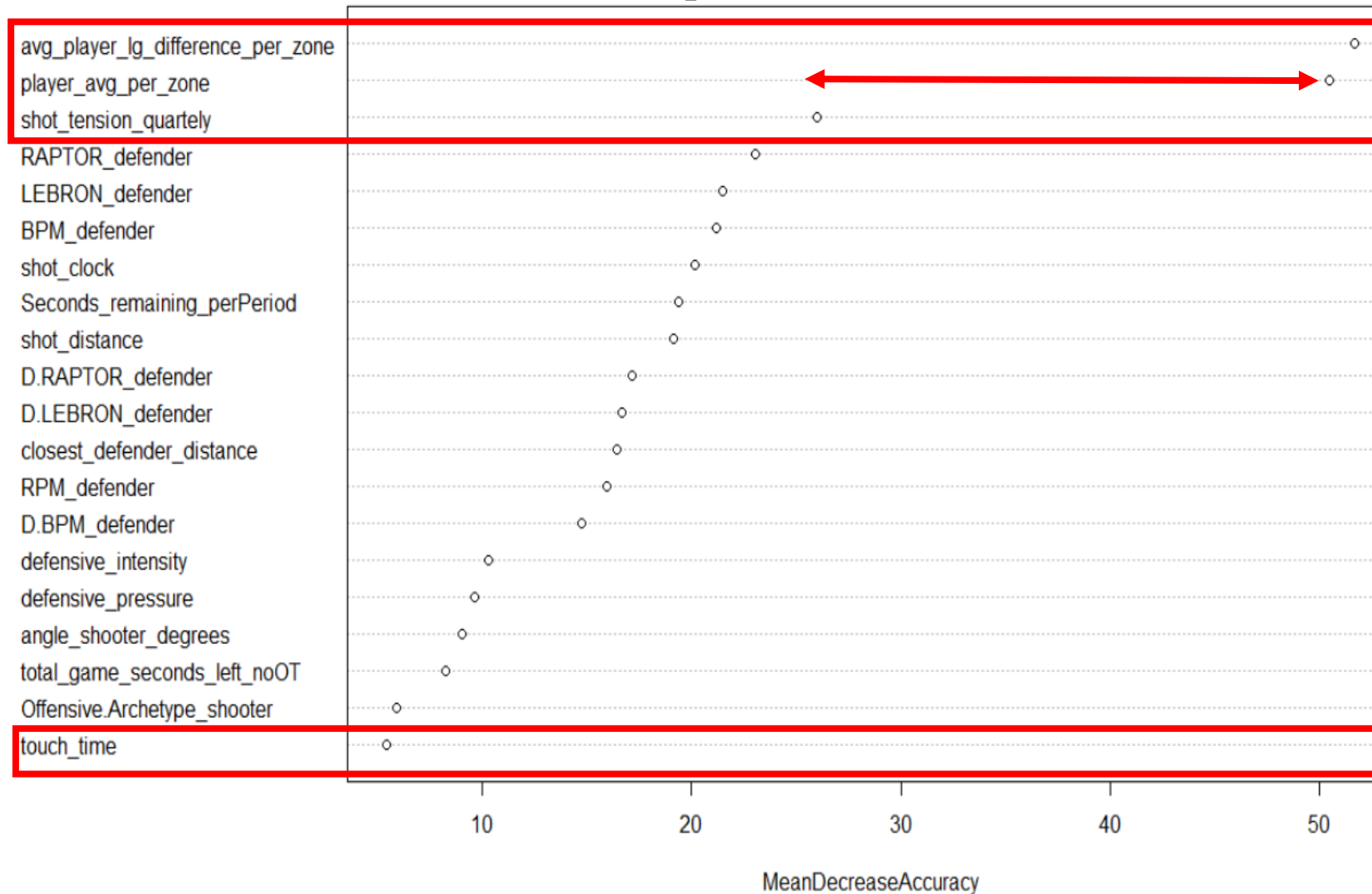
Feature Importance



-  In-game features
-  Seasonal FG% stats
-  Player Metrics

Three-Pointers

Feature Importance





Two-Pointers



Shooting Zone



Defensive Pressure



Possession Type



Shooting Ability



Shot Tension



Player Archetype



Shot Angle



Dribbling



Other Indicators



Season Stats

Applications

In-game
strategy



Roster
construction



Training
Regimes



Player
Benchmark



Future Work



Multiple
Seasons



Data-tracking
Enhancement



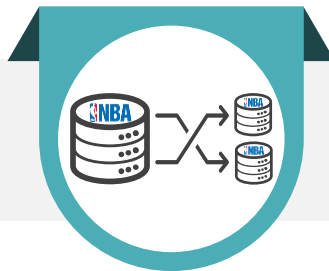
Arm
Raised



Foul
Trouble



Player
Fatigue



Data
Splitting



More
Classifiers



Analysis
Variants



Resources



Books & Articles:

Bocskocsky, A. et al., 2014. Heat Check: New Evidence on the Hot Hand in Basketball

Franks, A. et al., 2015. Characterizing the Spatial Structure of Defensive Skill in Professional Basketball

Keshri, S. K., 2019. Essays in Basketball Analytics

Liu, H. et al., 2019. Predicting the outcome of a basketball shot

Molnar, C., 2019. Interpretable machine learning

Partnow, S., 2021. The Midrange Theory

Zuccolotto, P. et al., 2020. Basketball Data Science with Applications in R

Data Sources:

[Basketball-Reference.com](https://basketball-reference.com)

[Bball-index.com](https://bball-index.com)

[NBA.com](https://nba.com)

[NBAsavant.com](https://nbasavant.com)

[NBAstuffer.com](https://nbastuffer.com)

Package hoopR

Package nbastatR

[Statsperform.com](https://statsperform.com)



Any
Questions?