

# Anatomy of a Made Shot

“Predicting Kobe Bryant’s Career Makes”

Ahmed El-Razi





# Career Snapshot

- 19 years
- 30,000+ shots
- 5 NBA Championships
- 3 Olympic Gold Medals



# Data

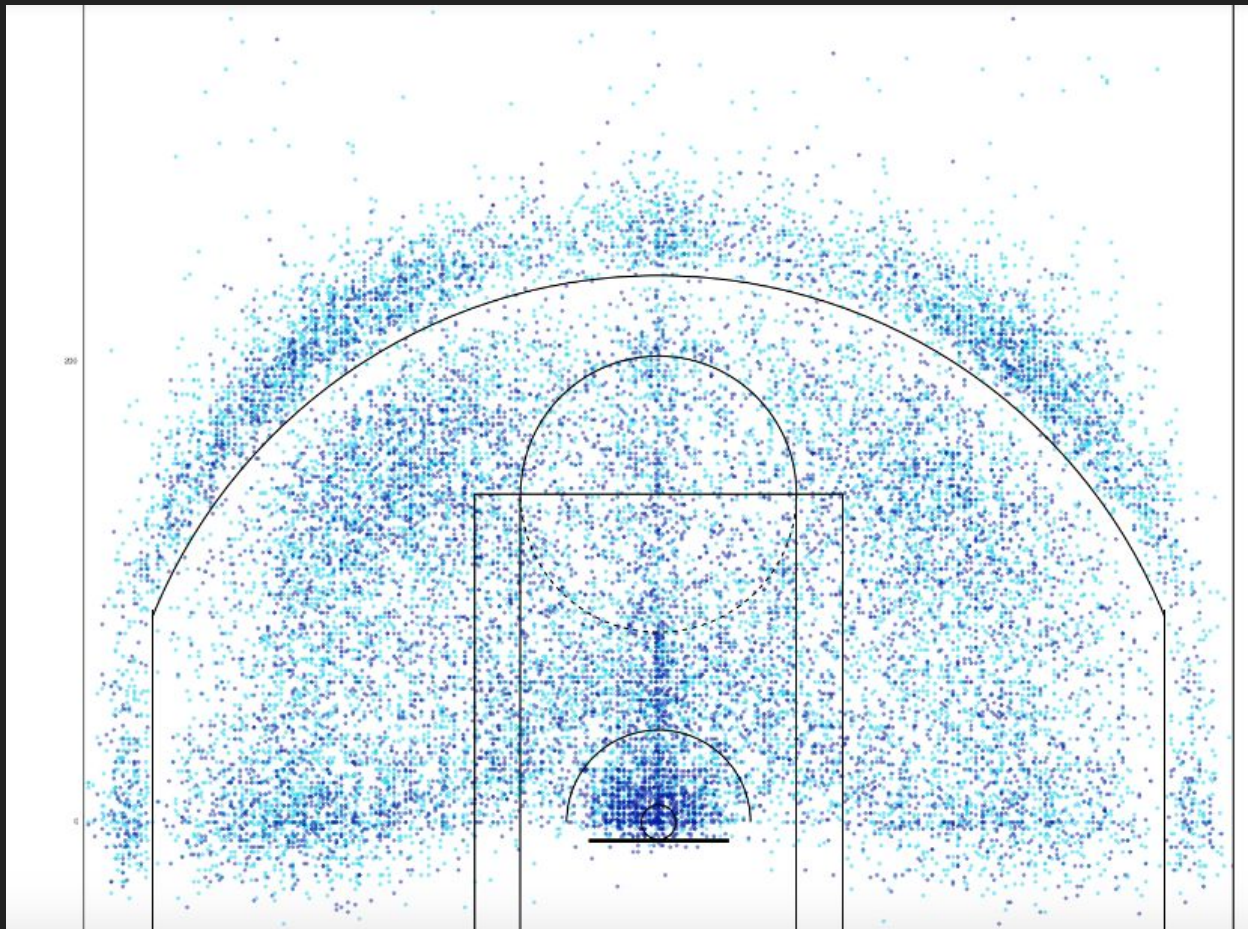
## Kaggle's dataset

- Career shots
  - Made/missed
  - Court location
  - Opponent
  - Distance from basket
- 30,697 observations, 25 features

The Kaggle logo, featuring the word "kaggle" in a lowercase, rounded, blue font. A small "TM" trademark symbol is located at the top right of the letter "e".

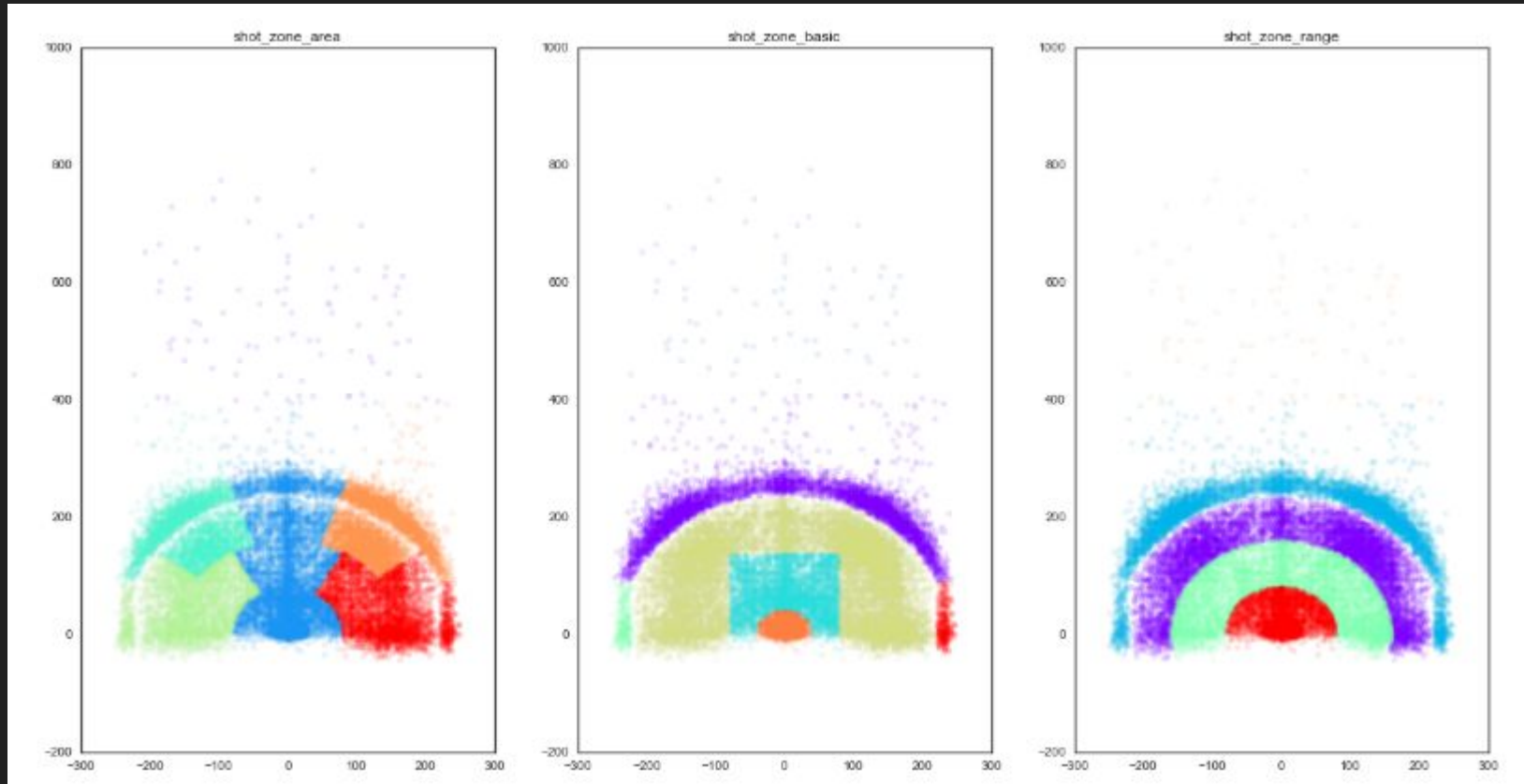
- **Objective: Build a model to predict which shots Kobe will make**

# Career Shot Chart





# Shot Zones



# Modeling

- Logistic regression with RFE
- Random Forest with GridSearch (1000 estimators)

## Features:

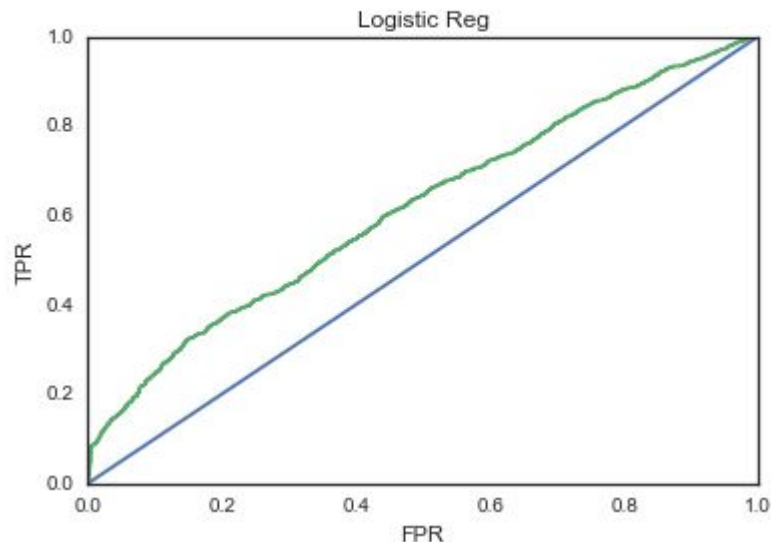
[shot type, shot distance, shot area, time remaining, opponent, period, game date, playoffs(binary) ]

- RFE singled out **shot type** and **shot area**

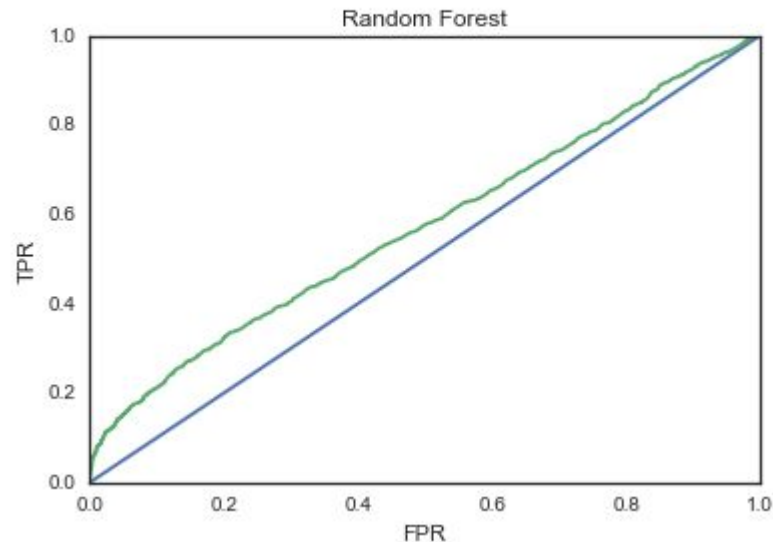


# Logistic vs. Random Forest (AUC)

fpr 0.148822269807  
tpr 0.320359281437  
auc: 0.615266866084



fpr 0.256502426838  
tpr 0.371685201027  
auc: 0.571223337968



# Logistic vs. Random Forest (Precision/Recall)

Training Data Accuracy: 0.62

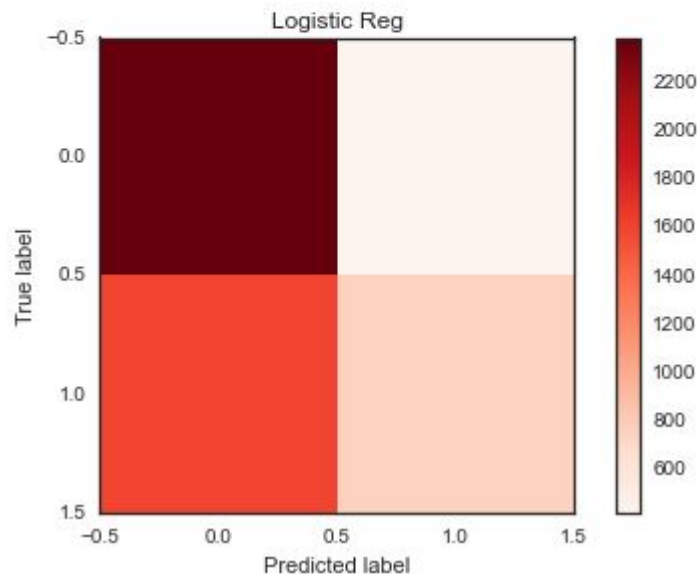
Test Data Accuracy: 0.61

[[2385 417]

[1589 749]]

Precision: 0.60

Recall: 0.85



Training Data Accuracy: 1.00

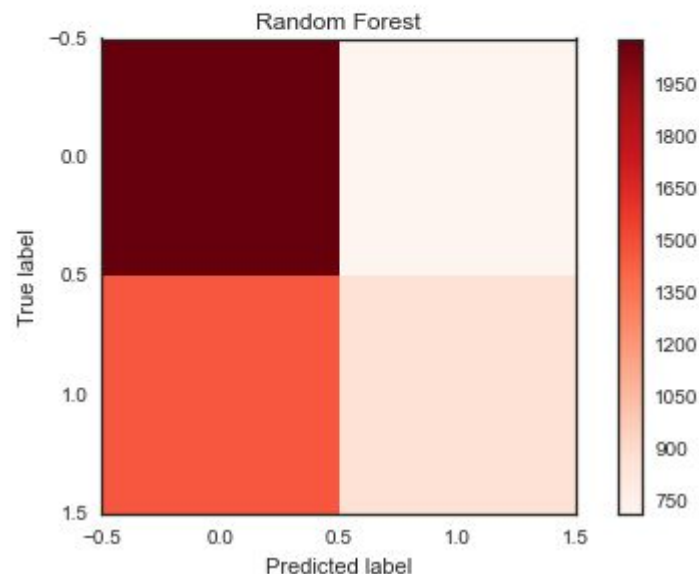
Test Data Accuracy: 0.57

[[2085 717]

[1470 868]]

Precision: 0.59

Recall: 0.74



# Conclusions

- Logistic model > Random Forest
  - Accuracy
  - Precision
  - Recall
- Overall predictive power: moderate
- Feature engineering to improve model

# Acknowledgements

- Savvas Tjortjoglou (code for “Career Shot Chart”)
- Dixhom (code for “Shot Zones”)

Thank you.

**I DON'T ALWAYS PASS BUT WHEN I DO**

