Predicting NBA All-Stars

Capstone Presentation

Lucas Winter
August 27, 2019

Agenda

- Problem Overview
- Data Collection & Cleaning
- EDA
- Modeling
- Results
- Conclusions
- Next Steps

Who Will Become an NBA All-Star?

Problem: Despite having more access to data and information than ever before, front offices and scouting departments of professional sports teams still continue to struggle to make the correct draft choices. Drafting a player is an inexact science, but having a way to predict future success of a player will make a teams draft decision that much easier and will give them a leg up on the competition.

Goal: To build a model that will predict the likelihood that a draft eligible College Basketball player will become an All-Star in the NBA.

Data Collection & Cleaning

- Data Sources: Sports-Reference.com & Bart Torvik
- Years: 2008 - 2019
- 55,939 individual college basketball seasons
  - Reduced to 581 single seasons
- Only Drafted Players
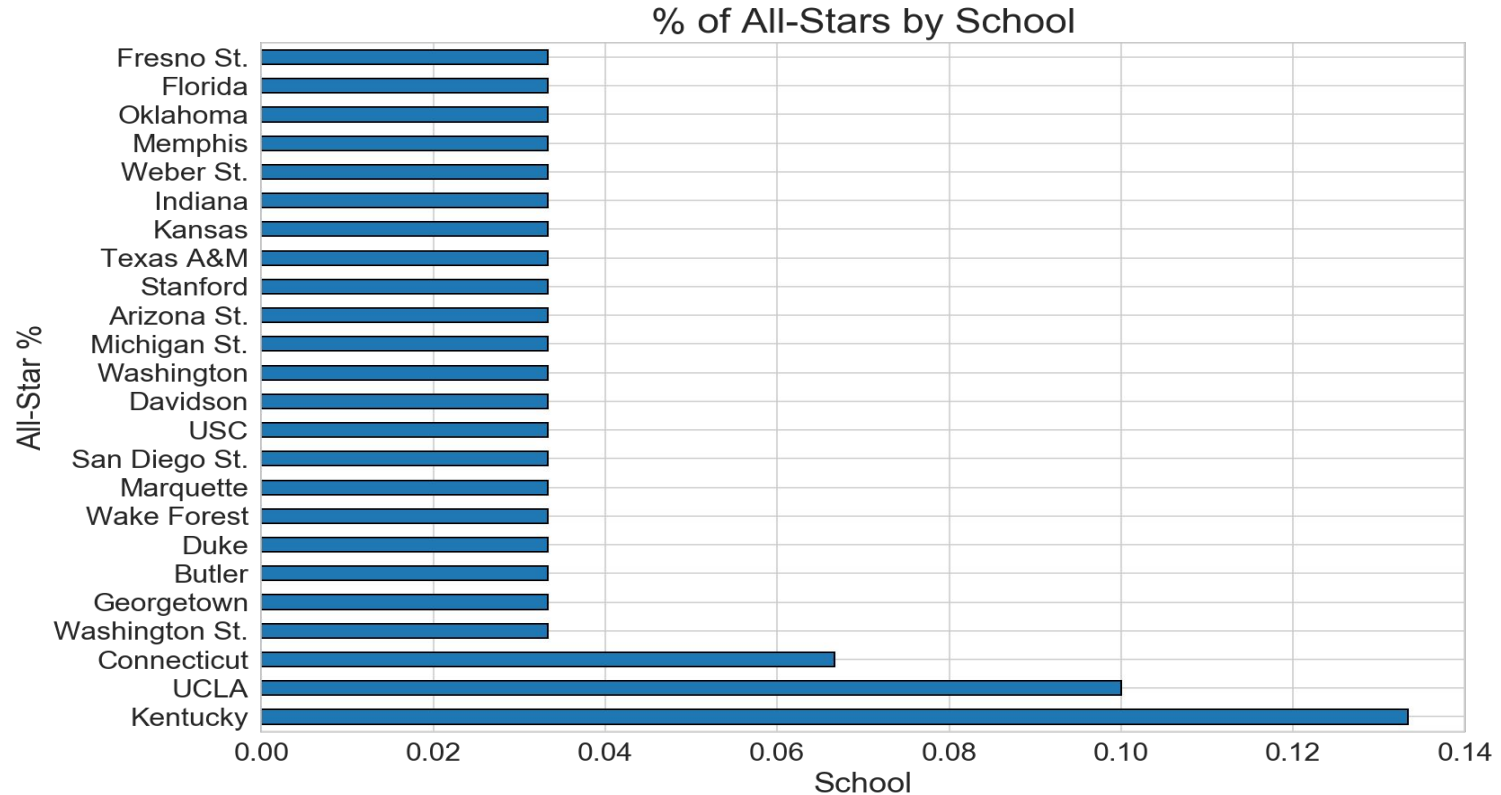- Removed International Players
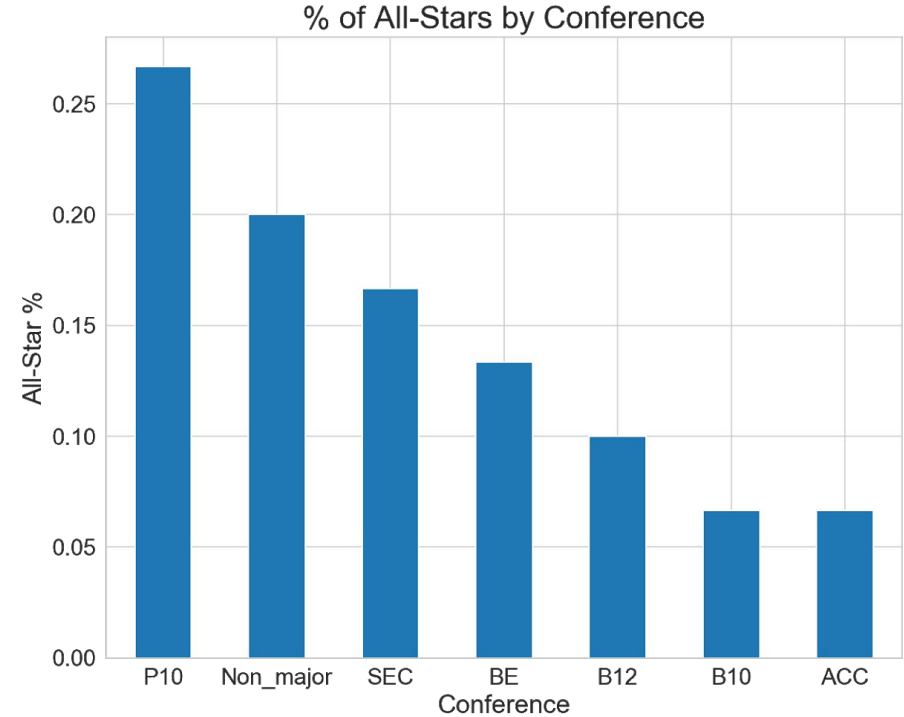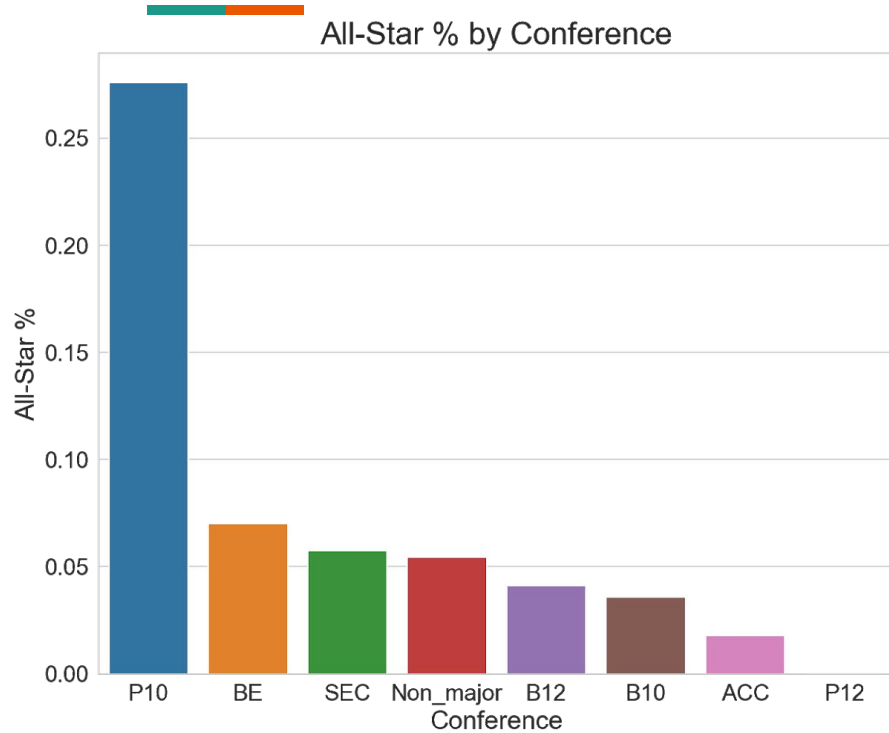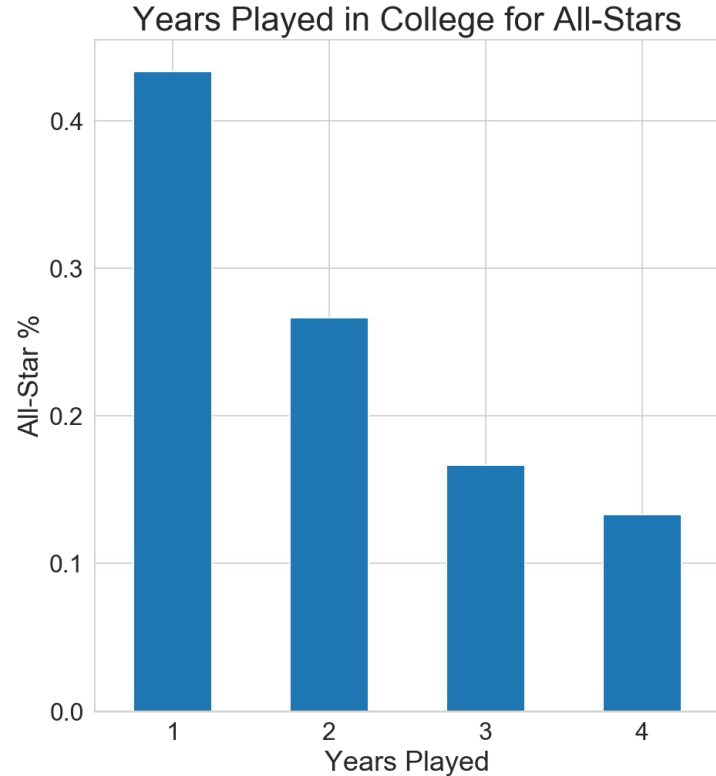
*30 New All-Stars since 2008

EDA

ALL-STARS BY SCHOOL

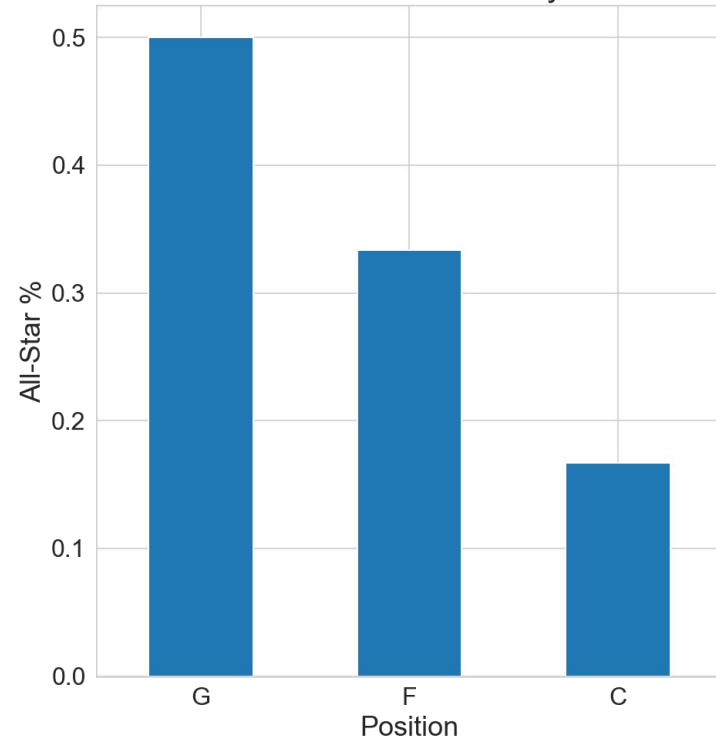% of All-Stars by School

ALL-STARS BY CONFERENCE

All-Star % by Conference

% of All-Stars by Conference

ALL-STARS BY YEARS PLAYED
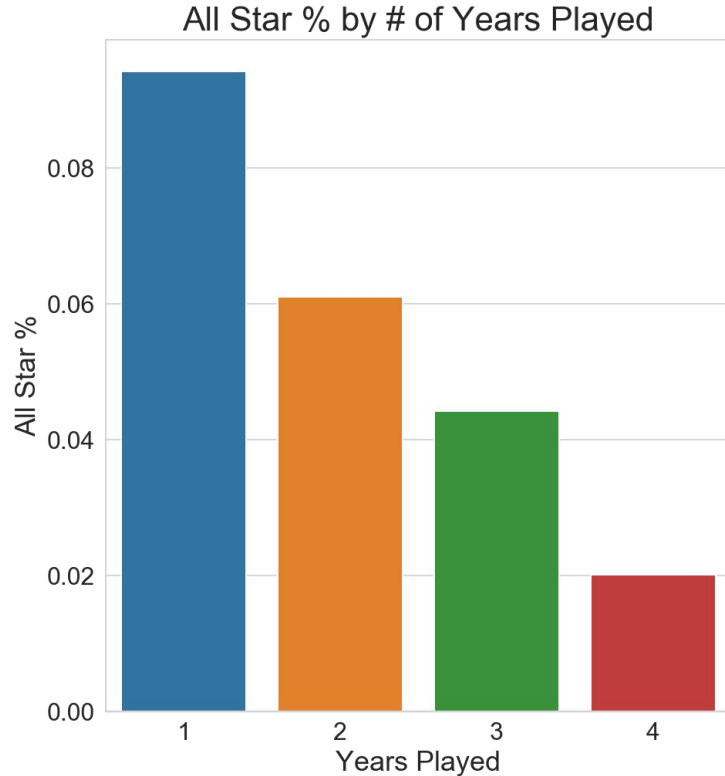
ALL-STARS BY POSITION
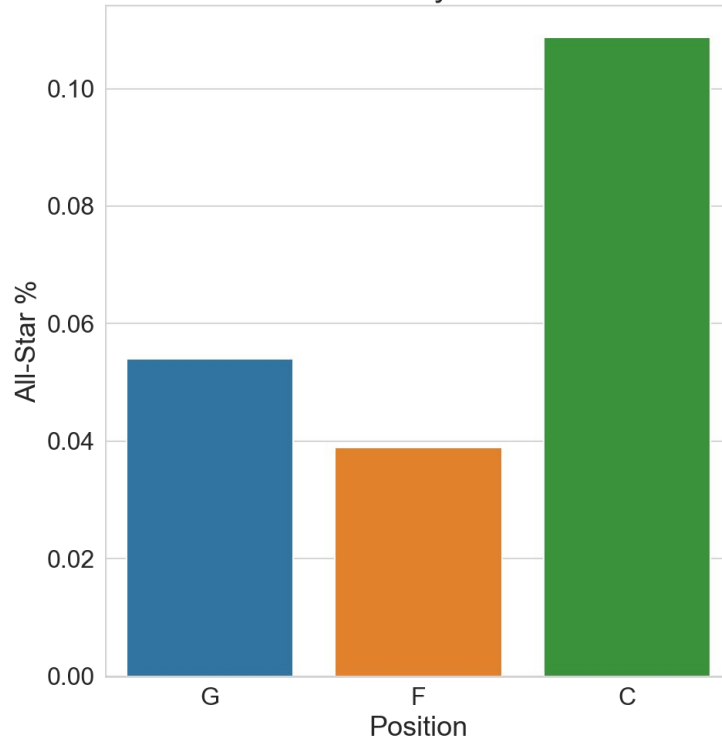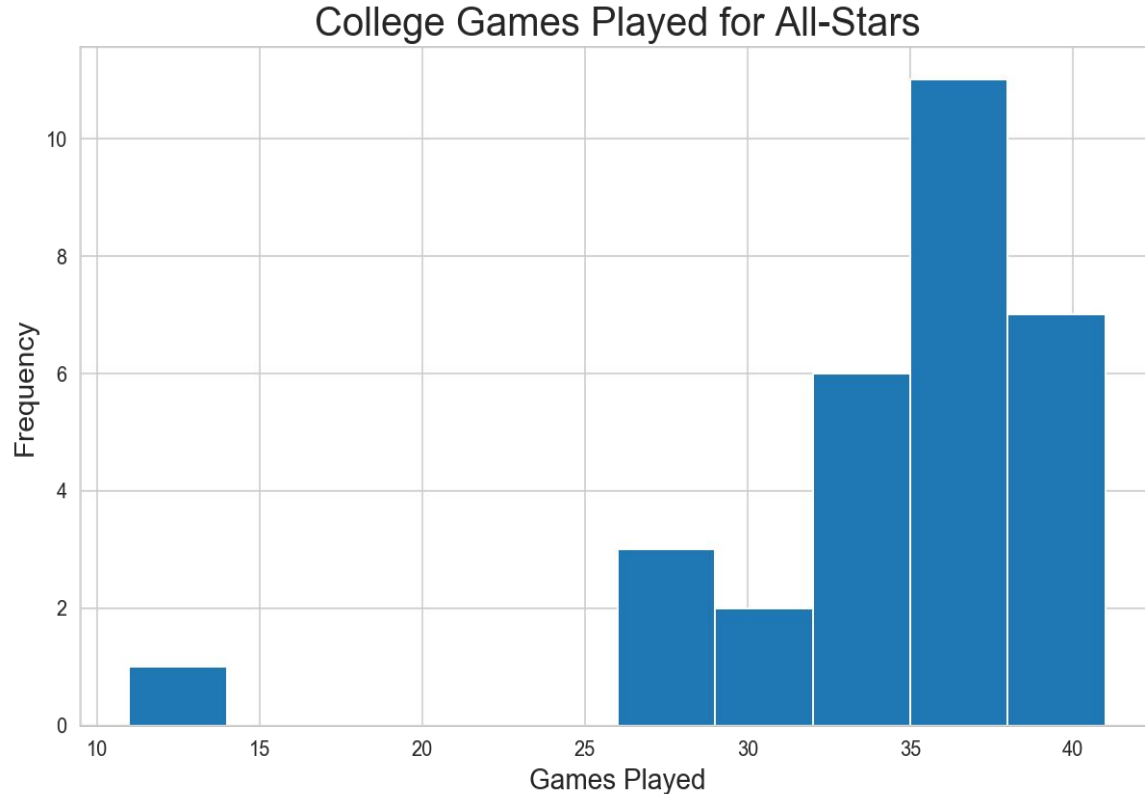
ALL-STARS BY GAMES PLAYED

- Majority of All-Stars played between 30 and 40 games in their final college season.

- Kyrie Irving is the only All-Star to play less than 25 games.

- All-Stars tend to be healthy in college.

College Games Played for All-Stars

Do Traditional Stats Matter?

- Traditional basketball statistics are only slightly better for All-Stars in college than Non-All-Stars.

- Points per game only 1-point higher for All-Stars in college.

- More advanced metrics may be able to tell a better story.

all_star	0	1
PPG	15.34	16.24
new_2P_avg	0.52	0.53
new_3P_avg	0.36	0.35
new_FT_avg	0.72	0.72
AST_per	15.77	19.24
ORB_per	6.95	8.17

Feature Engineering - Bayes MAP Estimate

ADJUSTED STATISTICS

- FREE THROW %

- 2-POINT %
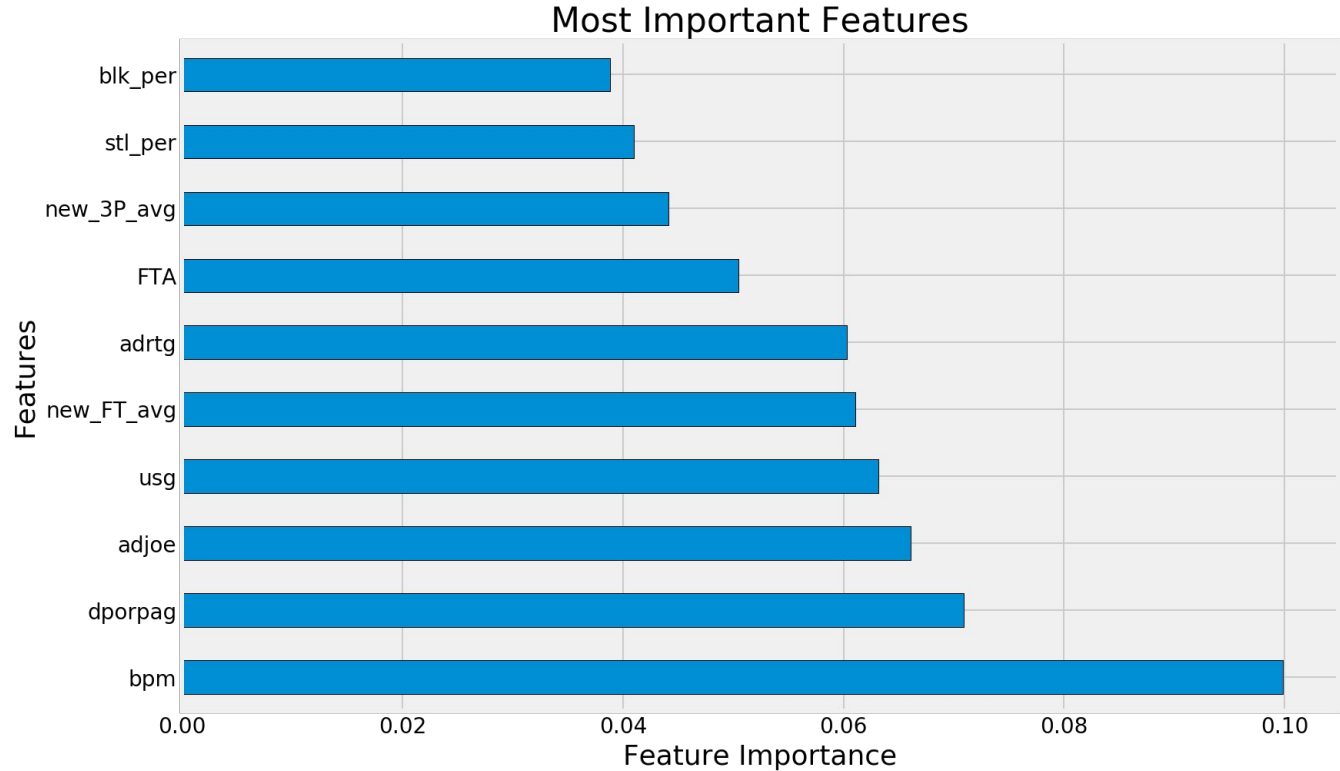
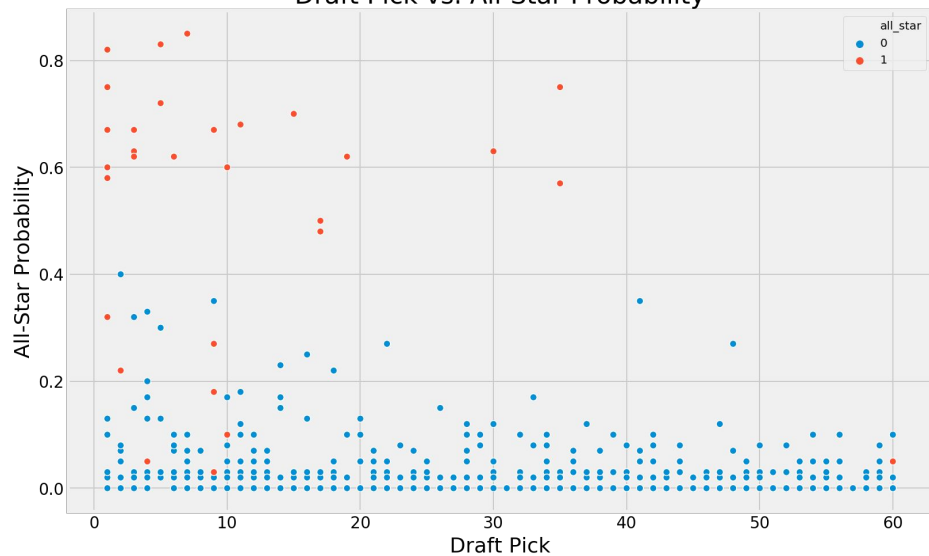- 3-POINT %

Free Throw % (Bayes)
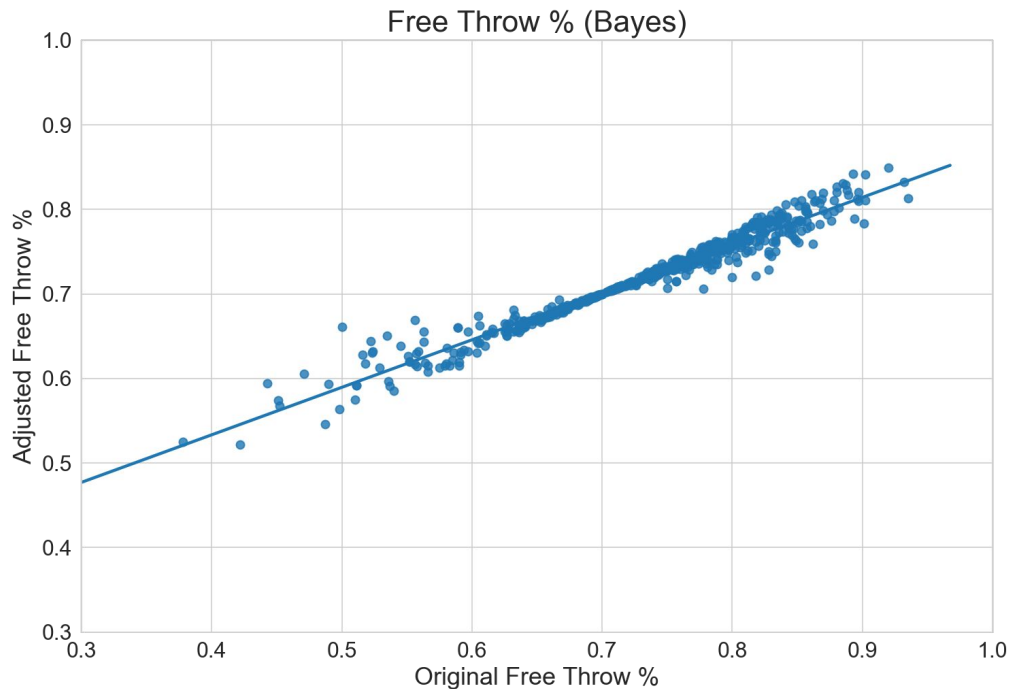
Model

Random Forest

ACCURACY

- Train Score: 99.5%
- Test Score : 93.9%

- Provided interpretable probabilities.
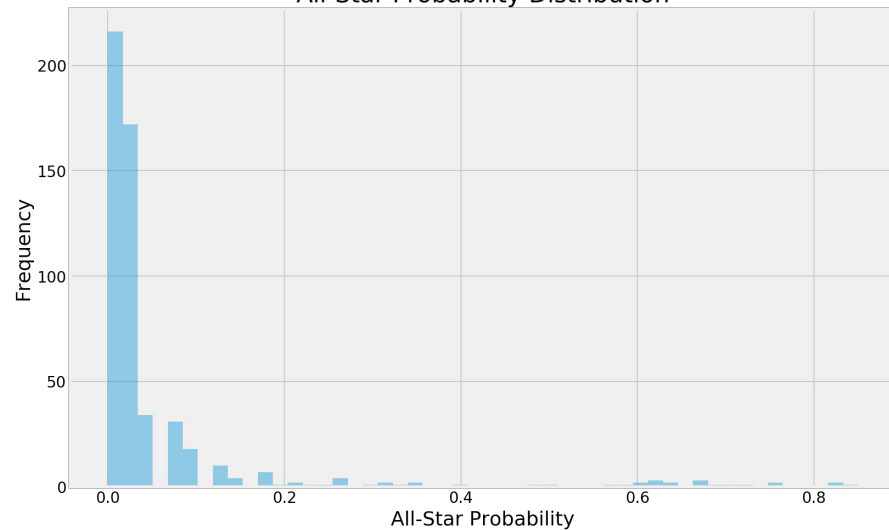- Efficiency metrics reign supreme.
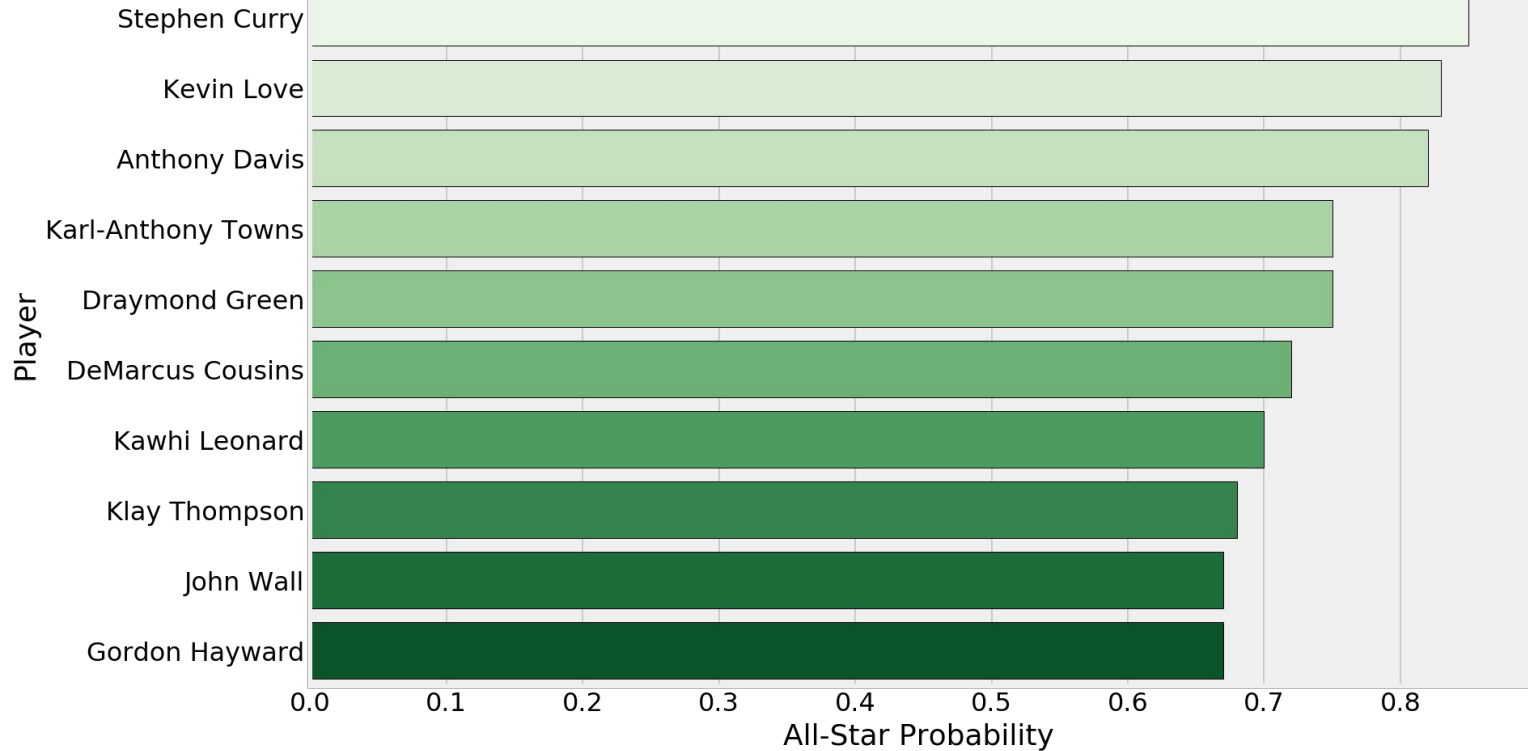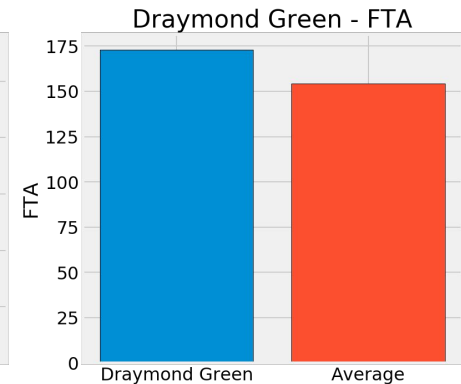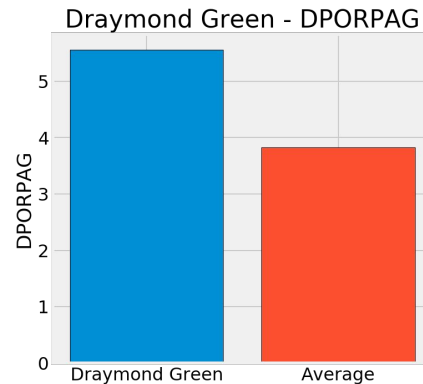

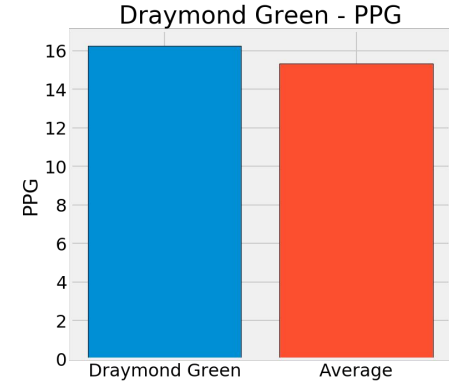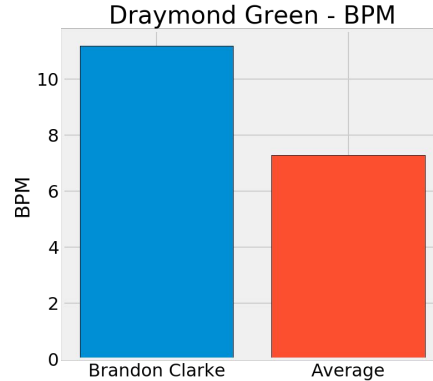Most Important Features

2008-2018

Top 10 All-Star Probabilities

Draymond Green

- Drafted 35th overall

- 3-Time All-Star
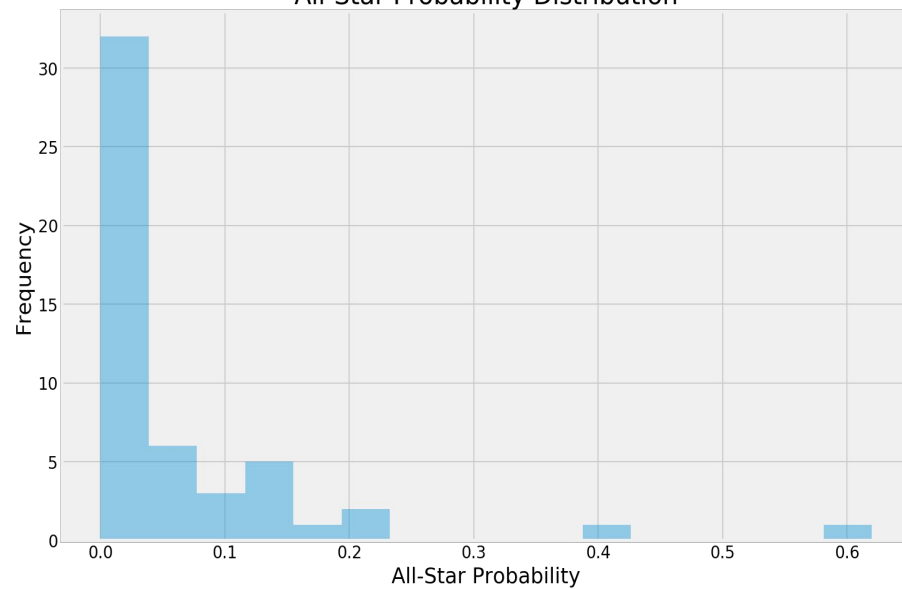
- 3-Time NBA Champion

- What did front offices miss?

2019
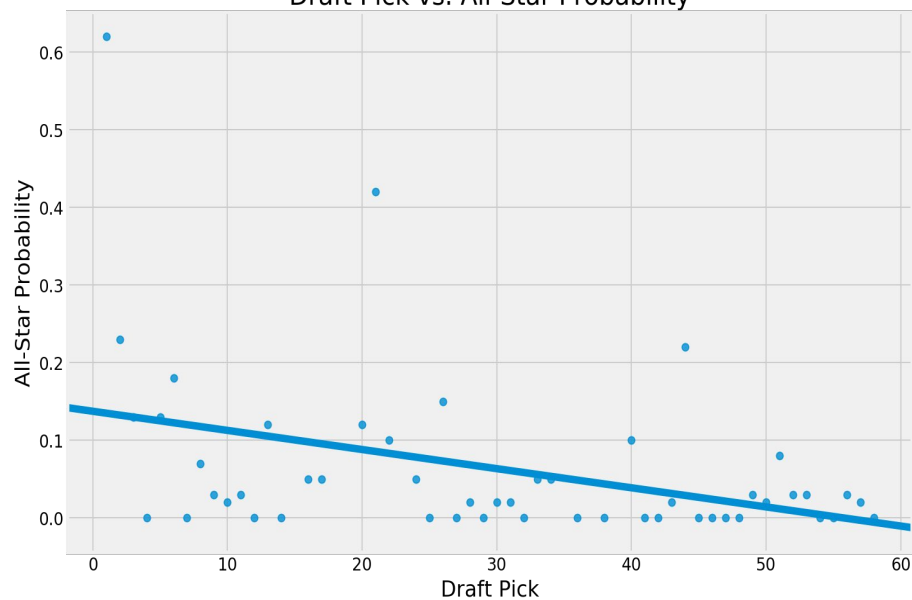
Highest Probabilities

- Zion Williamson had the highest probability of becoming an All-Star

- Mix of early round draft picks and late round draft picks

- 25% of players with greater than 5% of making the All-Star game

| player_name	school	pick	all_star_prob
Zion Williamson	Duke	1.0	0.62
Brandon Clarke	Gonzaga	21.0	0.42
Ja Morant	Murray St.	2.0	0.23
Bol Bol	Oregon	44.0	0.22
Jarrett Culver	Texas Tech	6.0	0.18
Dylan Windler	Belmont	26.0	0.15
R.J. Barrett	Duke	3.0	0.13
Darius Garland	Vanderbilt	5.0	0.13
Matisse Thybulle	Washington	20.0	0.12
Tyler Herro	Kentucky	13.0	0.12
Justin James	Wyoming	40.0	0.10
Grant Williams	Tennessee	22.0	0.10
Tremont Waters	LSU	51.0	0.08
Jaxson Hayes	Texas	8.0	0.07

Player Comparisons
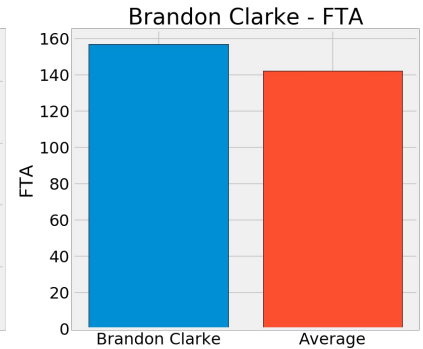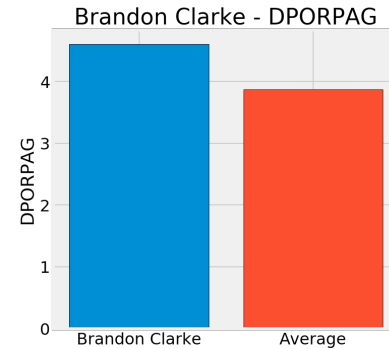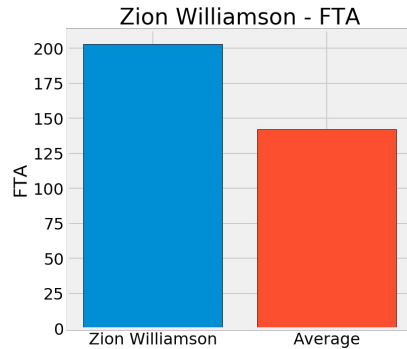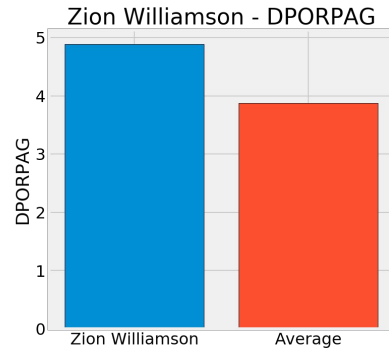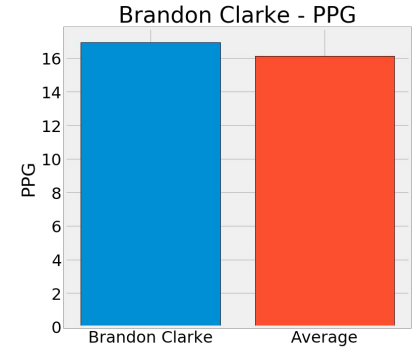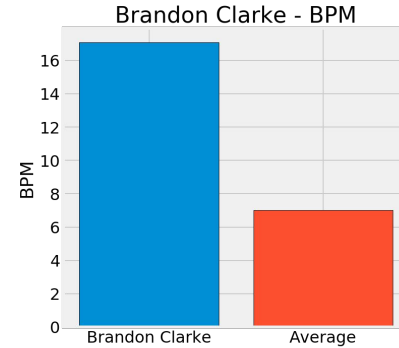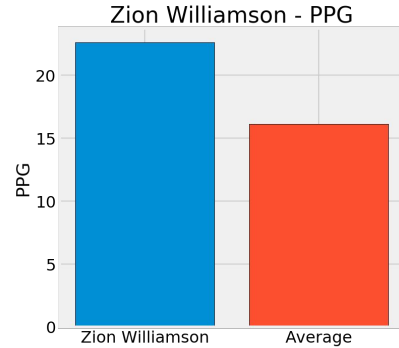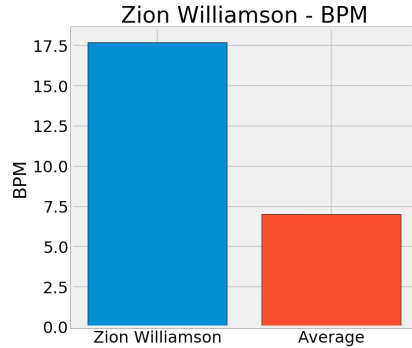
Conclusion

Utilizing a random forest model, we were able to predict correctly if a player had become an All-Star at some point in their career 99% of the time on the training set and 94% on the test set.

This included players such as Draymond Green who was drafted in the second round and Kyrie Irving who only played 11 games in his college career.

When testing my model on unseen data, the 2019 drafted players, it predicted Zion Williamson to have the highest probability of becoming an All-Star. The player with the second highest probability was Brandon Clarke from Gonzaga. He happened to be selected 21st overall which is not a range where All-Stars are typically taken.

Next Steps

- Collect more data
    - Data going back further than 2008.
    - New features:
        - High school recruiting rank
        - Number of wins in college
        - Athletic measurements from the NBA combine.
    - International players and high school players

- Quantify the mental aspects of the game.
    - Player Aptitude

- Predict draft picks and draft position.

Resources

- Basketball-Reference.com
- Bart Torvik
- MIT Sloan Sports Analytics Conference