



# Classifying & Evaluating NBA Players

Michael Hartnett

# Introduction/Background:

- NBA Season just completed first half of it's 82 game season marked by the All-Star Break festivities (held on March 7th 2021)
- Trade Deadline approaching (March 25th)
- Celtics Majority Owner, Wycliffe “Wyc” Grousbeck would like to solicit data driven analysis of their current roster, do they have a team composed of “playoff caliber” players?
- Celtics current record: 19-17 heading into the All-Star break

# Problem Statement

-Using both supervised and unsupervised machine learning, can I create a classification model to determine whether a current player and their statistics thus far this season is "playoff caliber"?

# Premodeling Methodology

- Acquire data and statistics for each player over previous 30 years:
  - Basic Stats: (points per game, rebounds, assists, blocks, steals fg%, 3pt%)
  - Advanced Stats: (player efficiency rating, true shooting %, box plus minus, win shares per 48)
- Establish player tiers (different tiers should have different expectations associated with being labeled playoff caliber)
- Defining playoff caliber: if the player's team made the playoffs that year they are denoted as playoff caliber

# Acquiring Data

-All data acquired through scraping different team tables found through basketball composite database site: <https://www.basketball-reference.com/>

## Advanced Stats Info:

-PER (player efficiency rating): Player Efficiency Rating is the overall rating of a player's per-minute statistical production. The league average is 15.00 every season.

-True Shooting %: what a player's shooting percentage would be if we accounted for free throws and 3-pointers.  $\text{True Shooting Percentage} = \text{Total points} / [(\text{FGA} + (0.44 \times \text{FTA}))]$

-BPM: uses a player's box score information, position, and the team's overall performance to estimate the player's contribution in **points above league average per 100 possessions played (league avg 0.0)**

-Win Shares Per 48: takes into account the various things a basketball player does to win or lose a game, and weighs them appropriately to provide a Win Share

# Player Tiers (2021)

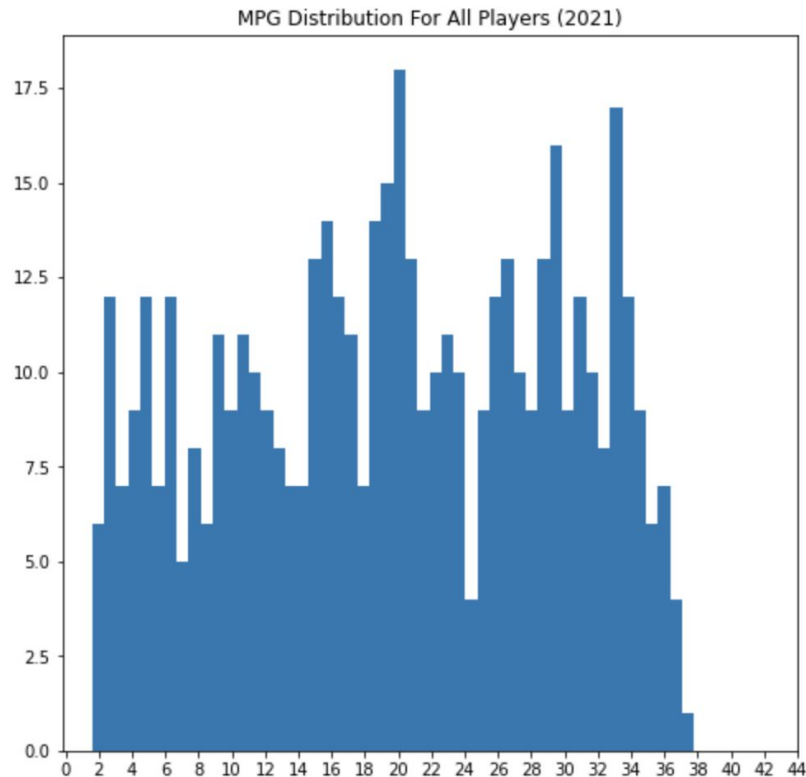
-**Major Players:** 30+ mpg, responsible for the bulk of production

-**Core Players:** 20-29.9 mpg, “6th man”, starters who split time

-**Rotational Players:** 10-19.0 mpg, spot minutes, younger or veteran players

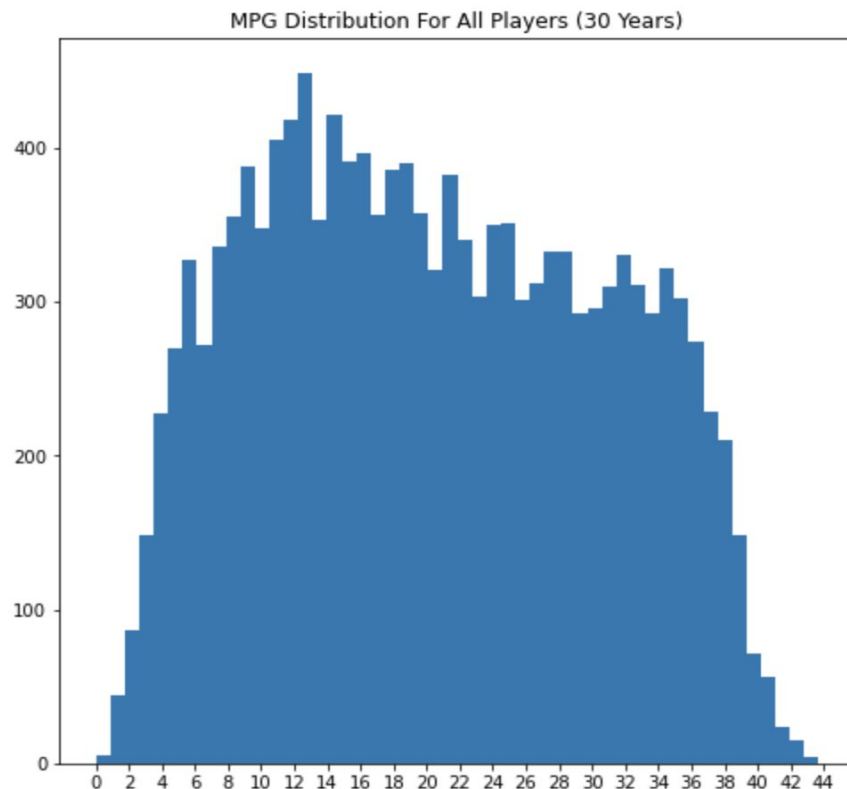
-**Reserve Players:** <10 mpg, situational players

```
Name: mpg, dtype: float64
count      494.000000
mean        19.751012
std         9.775469
min         1.600000
25%        11.525000
50%        19.900000
75%        28.375000
max         37.800000
Name: mpg, dtype: float64
```



# Player Tiers (All 30 Years)

count	13945.000000
mean	20.398745
std	10.168311
min	0.000000
25%	12.000000
50%	19.800000
75%	28.800000
max	43.700000

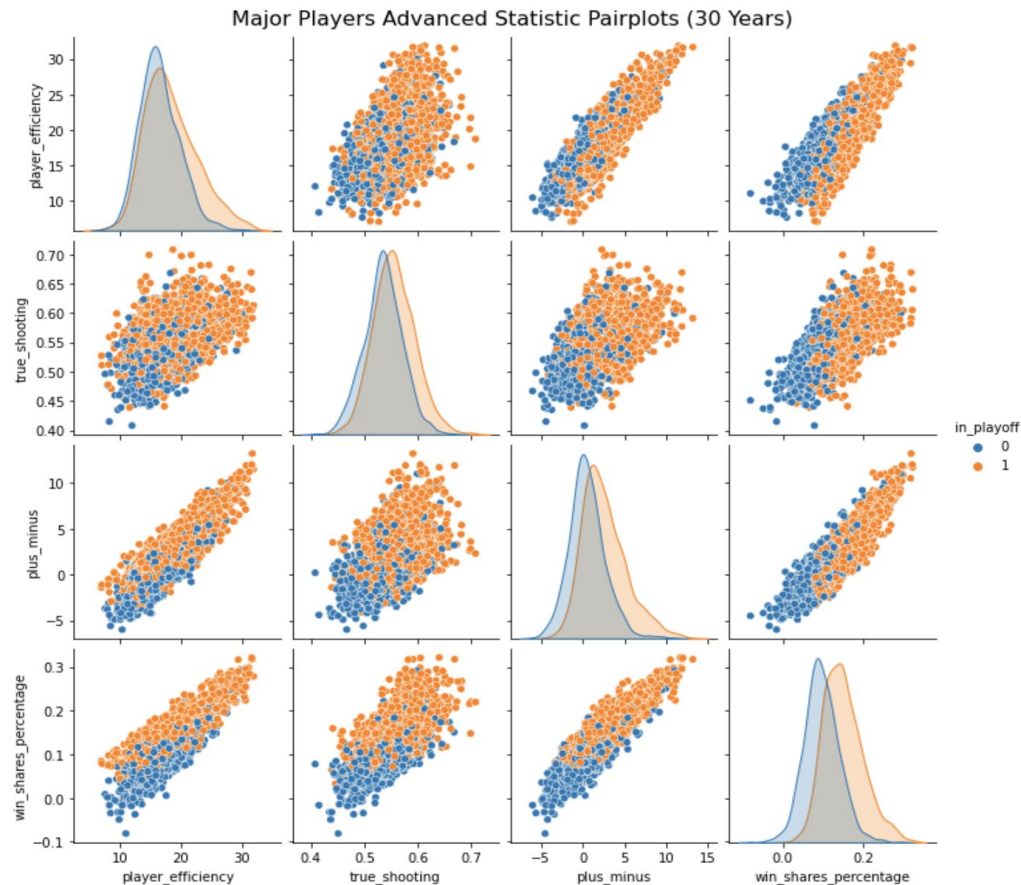




# EDA Statistics: Advanced Metrics

-Advanced statistics w/  
'in\_playoffs' hue shows we should  
be able to build a classification  
model effectively

-Sidenote: Displaying high degrees  
of multicollinearity, but not  
surprising since all cuts of  
impact/contributions similarly  
based on player performance.  
(shouldn't affect our model  
predictive power)





# KMeans Clustering: By Decade

Decade Clusters for All Major Players (over 30 years)

labels	0	2	1
year	1995.017889	2005.635936	2015.368597
games_played	71.981216	72.460019	68.576837
mpg	34.900537	35.110536	33.401002
fg%	0.470748	0.458312	0.463898
3pfg%	0.257787	0.294197	0.319150
rebounds	6.229964	6.016839	5.948552
assists	3.838104	3.661430	3.936303
steals	1.259213	1.129633	1.134521
blocks	0.731485	0.675165	0.626392
points_per_game	16.435242	16.512700	17.187751
player_efficiency	17.138283	17.453998	18.018708
true_shooting	0.541538	0.541243	0.556043
plus_minus	1.470572	1.534807	1.836971
win_shares_percentage	0.122292	0.123344	0.126518
in_playoff	0.567979	0.511759	0.526726

Decade Clusters for All Major "Playoff Caliber" Players (over 30 years)

labels	2	0	1
year	1994.462069	2004.438757	2014.958095
games_played	73.074138	74.341865	71.215238
mpg	35.020862	35.281901	33.762286
fg%	0.482962	0.462371	0.474617
3pfg%	0.261128	0.293892	0.322697
rebounds	6.496034	6.137294	6.198095
assists	3.942759	3.823949	4.122857
steals	1.298621	1.170567	1.170667
blocks	0.786552	0.726874	0.677524
points_per_game	16.729655	16.789031	17.787619
player_efficiency	17.973966	18.130713	19.064571
true_shooting	0.554362	0.544912	0.566998
plus_minus	2.445172	2.363803	2.873524
win_shares_percentage	0.148347	0.144976	0.152910
in_playoff	1.000000	1.000000	1.000000

# Modeling

- Tried out a handful of classification models (KNN, RandomForest, Bagging etc..) most effective models for each tier either LogisticRegression or SVM
- Tier 1: Major Players **SVM** Train Acc: 82.6% / Test Acc: 81.3%
- Tier 2: Core Players **LogReg** Train Acc: 78.5% / Test Acc: 76.1%
- Tier 3: Rotational Players **LogReg** Train Acc: 74.1% / Test Acc: 72.0%
- Tier 4: Reserve Players **SVM** Train Acc: 74.1% / Test Acc: 63.7%

# Model Intuition: Do the Results Make Sense?

-Compared our classification results to the 1st place team of the Western Conference: the 27-9 Utah Jazz and last place team: the 8-29 Minnesota Timberwolves

-Jazz: 14 Playoff Caliber, 2 Non-Playoff Caliber

-TWolves: 2 Playoff Caliber, 14 Non-Playoff Caliber

# Celtics Analysis: Major Players

- 3 Non Playoff Caliber, 1 Playoff Caliber
- “...Asked if he thinks, as currently constituted, the Celtics are good enough to win a title, he replied ‘I do not.’ (excerpt from interview with Celtics GM Danny Ainge)
- As a Celtics fan, these model results & interview made me cry :’(



Image source: <https://images.app.goo.gl/umbQBWJvTCV5Z7JA9>

Interview source: <https://985thesportshub.com/2021/02/18/danny-ainge-says-celtics-do-not-have-a-championship-roster/>

# Celtics Analysis: Core Players

- 1 out of 3 core players classified as playoff caliber: Daniel Theis
- Rookie Peyton Pritchard and veteran big man, Tristan Thompson, non-playoff caliber
- Theis, a Center, has shown he can hit 3s at a 37% rate this half--valuable in today's game



Image source: <https://images.app.goo.gl/TRyA5U815Sdk9YTh7>

# Celtics Analysis: Rotational Players

-3 of the 7 rotational players classify as “playoff caliber”: Rob Williams, Semi Ojeleye, Carsen Edwards

-Player to Watch: Rob Williams--first round of modeling he was not “playoff caliber” but after a strong week of play model changed his classification



Image source: <https://images.app.goo.gl/sJhV6VbCrcw5K96P8>

# Celtics Analysis: Reserve Players

- Both Celtic reserves qualified as playoff caliber
- While reserves certainly add value, not the tier we want with the higher percentage of playoff caliber players
- Tacko Fall is 7'5" tall.



Image source: <https://images.app.goo.gl/d9LV81nVY9ZTDblV6>



# Recommendations

- Major Players: Despite Celts stars not faring well with model, I suggest keeping current stars. Kemba is trending upwards coming off knee injury, Tatum & Brown 2 young stars with tremendous potential
- Core Players: Explore trade options here, either big man can go (depending on return), Pritchard worth keeping, high percentage from beyond the arc, also only a rookie
- Rotational Players: Keep Robert Williams, rest of the players can be worth packaging in a trade to acquire a potential playoff caliber core player here.
- Reserve Players: Keep fan favorite Tacko Fall, as mentioned he's 7'5" tall!

# Next Steps/Model Reflection

- Model imperfections: presence of multicollinearity in our features makes it difficult to analyze individual features
- Also we trained with a full season worth of data, while we are predicting solely with only a half season, so model does not incorporate for potential 2nd half improvements
- Difficult to quantify the effect of statistics having a great player on a bad team, his metrics may be hampered by poor support around him. (might qualify as playoff caliber on one team, but not on the other)

# Questions & Additional Celtics Articles

-Kemba Trade Rumors:

<https://www.masslive.com/celtics/2021/02/kemba-walker-trade-rumors-celtics-tried-like-hell-to-trade-guard-danny-ainge-knew-knee-wasnt-right-report.html>

-The Two Jays (Tatum & Brown):

[https://www.espn.com/nba/story/\\_/id/30984749/why-jayson-tatum-jaylen-brown-enough-boston-celtics-season](https://www.espn.com/nba/story/_/id/30984749/why-jayson-tatum-jaylen-brown-enough-boston-celtics-season)

<https://www.celticsblog.com/2021/3/11/22321297/the-evolution-of-jayson-tatums-shot-making-ability-boston-celtics-nba>

-Danny Ainge on a Potential Mid-Season Trade:

<https://nesn.com/2021/03/boston-celtics-trade-rumors-danny-ainge-tpe/>