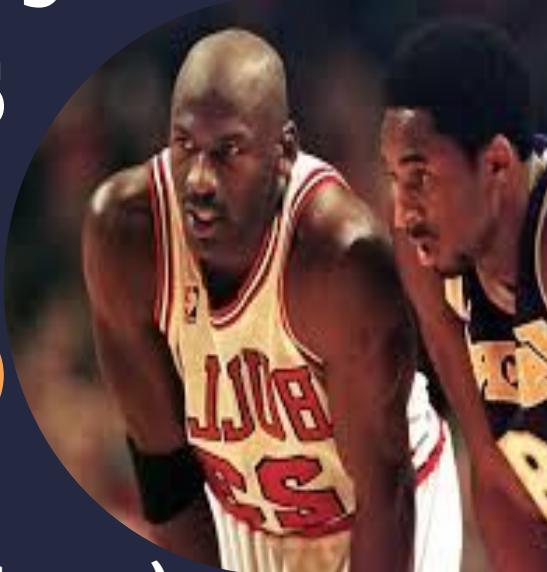


# Rookies of the year

## Vs MVP's



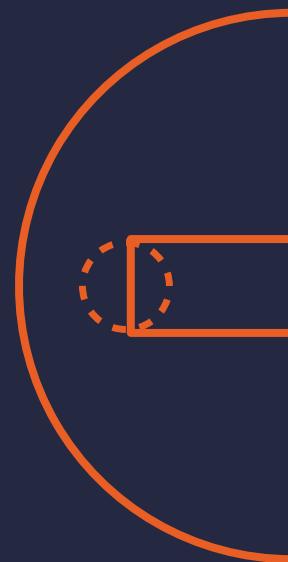
Player Salaries (2022-23 Season)

**Bryan Ramirez, Junyu Liu, Neil (Suritaneil) Sahota**



## Inspiration(The Why)

- **Basketball is fun and easy game to start playing at any level**
- **Watching basketball and tracking the stats is an extension.**
- **We played a lot of basketball growing up.**
- **We want see what are the most important skills when determining a player's worth.**



**This dataset looks at the salaries of all the player's and their key performance metrics in the NBA during the 2022 and 2023 season.**

## **—About Dataset**

# Dataset

## Dataset Dimensions

- (467 x 28)
- 467 NBA Players
- 26 Predictors
- 1 Response (Salary)
- 1 Column for Player Names

## Dealing with NAs

- Some columns had NA values (total of 42 NA values)
- Replaced NAs with the median value of column (as distribution was skewed)

# Variables:

**Variables we thought would be important.**

AST	Assists
PTS	Points scored
TRB	Total rebounds
X3P.	3 pointer percentage
FG.	Field goal percentage
STL	Steals

**Variables that were important.**

GP	Games played
GS	Games started
AST	Assists
TRB	Total Rebounds
PTS	Points scored
PPM	Points per minute

## Research Questions

Question 1



Can we predict the  
current Rookie of the  
Year salary?

Question 2



Can we predict the  
current MVP salary?

What about previous Rookie and  
MVP salaries?

# Predictor Multicollinearity

- Some predictors were highly correlated
  - After some research into dataset and found that some predictors were calculated based on other predictors
    - For example: Field Goals, 3 Pointers, 2 Pointers, and Free Throws include 3 related predictors
      - Each contain:
        - 1. Number of shots made by player
        - 2. Number of shots attempted by player
        - 3. Percentage of shots made by player

X3P	3 pointers made by player
X3PA	3 pointers shots attempted by player
X3P%	3 point percentage (X3P/X3PA)



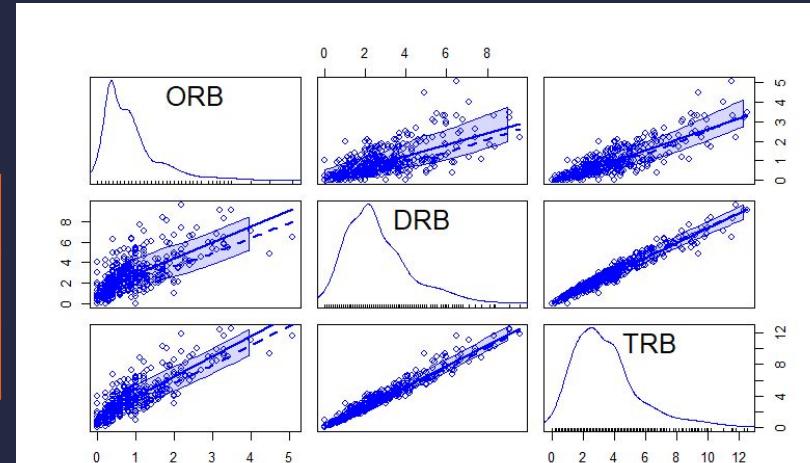
# Multicollinearity with Rebounds



Highly Correlated predictors

- Rebounds - player retrieving ball after missed shot
  - 3 related predictors for rebounds with high correlation, as  $TRB = ORB + DRB$
  - Remove ORBs and DRBs predictors in preference of TRBs

ORB	Offensive Rebounds by player
DRB	Defensive Rebounds by player
TRB	Total Rebounds by player



# eFG% (Effective Field Goal)

## Field Goal Percentage (FG%)

- Percentage of shots made from the court during gameplay, including two-pointer and three-pointer shots
  - FG% accounts for 3-pointer% and 2-pointer% in one value



## Effective Field Goal percentage (eFG%)

- Percentage of shots made from the court during gameplay, accounting for 3-pt shots being worth more than 2-pt shots
  - $eFG\% = (2\text{-pt FGs} + 1.5 \times 3\text{-pt FGs}) / FGA$
  - Higher weight for 3-pointer shots made, as it is worth more points.

# **Additional Predictors deemed unnecessary**



## **Age**

Age was taken out because of people like LeBron James, Steph Curry, and Kevin Durant.



## **\*Personal Fouls**

Salary will be consistent regardless of how many fouls, if the player is good enough. \* As long as it's not excessive.



## **\*Turnovers**

Salary will also be consistent for the same reason as personal fouls. \* As long as it's not excessive.

# Transformation

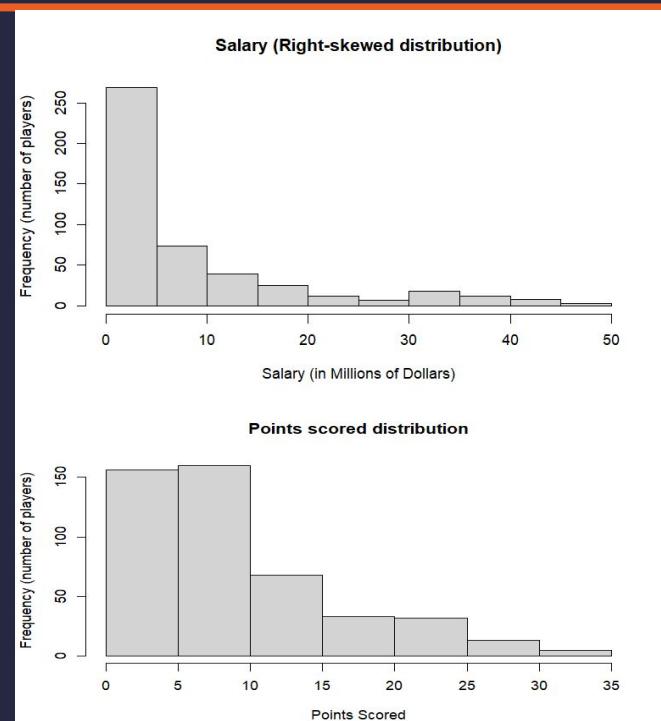
## Salary Transformation (on resp)

- As salary (response) was right skewed:
  - We utilized a log() transformation for salary in our MLR modeling



## Transformation (on pred)

- We utilized a log() transformation in the MVP model on the predictors GS, TRB, and PTS
- We utilized a log() transformation in the Rookie of the year model on the predictors GS, AST, and PPM



# Predictor Modifications

The reason we did this was because a lot of the predictors had high correlation with other predictors (including MP and PTS).  
**(Multicollinearity- VIF went down)**

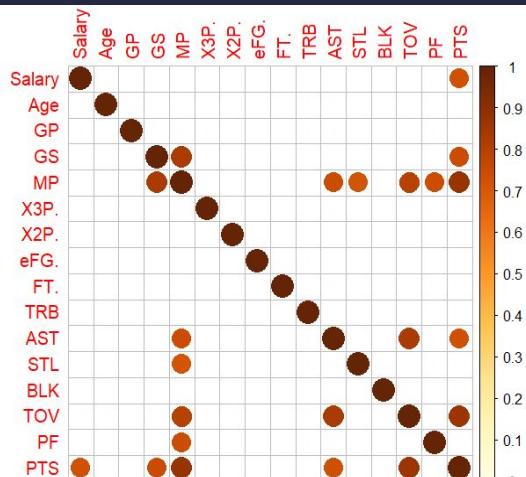
**GP \* MP = Total  
Minutes played**



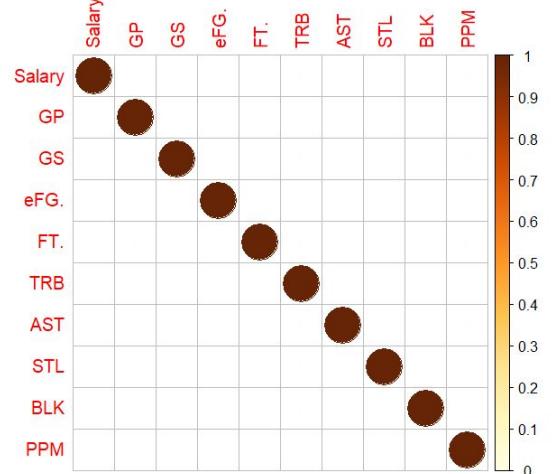
**PTS / MP = Points  
scored per Minute**

# Correlation Plot ( > 0.7)

**Before**



**After**



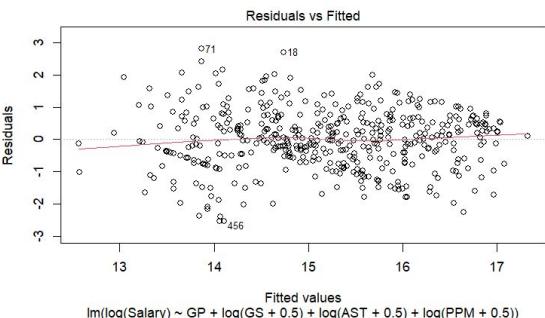
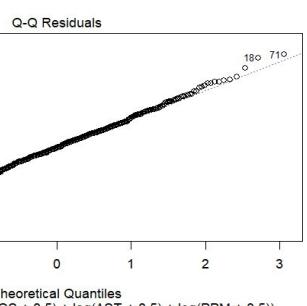
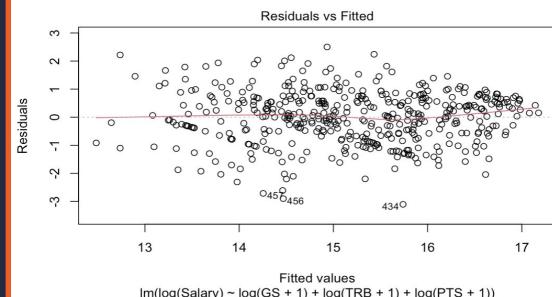
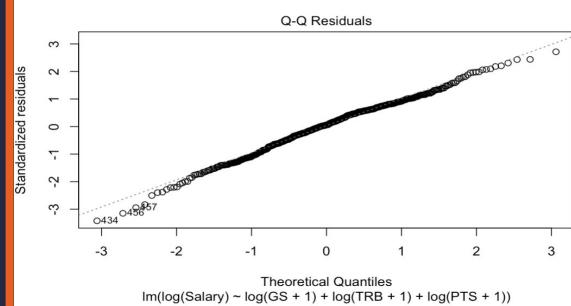
# Diagnostics Plots

MVP

01

X

02



ROO  
KIE

03

X

04

## 2 Final Models

**The MVP model was  
better at  
predicting MVP's salary.**

$$\log(\text{salary}) \sim \log(\text{GS}) + \\ \log(\text{TRB}) + \log(\text{PTS})$$

**The Rookie model was  
better at predicting  
rookie's salary.**

$$\log(\text{salary}) \sim \text{GS} + \log(\text{GP}) \\ + \log(\text{AST}) + \log(\text{PPM})$$

**The reason why the MVP model is better than the rookie model and vice versa is because MVP's are usually given to players who put up really high stats that rookies are not usually able to produce yet.**

# Model Interpretation(coefficients)

**MVP**

$\log(\text{salary}) \sim$   
 $\log(\text{GS}) + \log(\text{TRB}) +$   
 $\log(\text{PTS})$



For every 1 % increase in Games Started the predicted salary is expected to increase by about 0.19% when holding all other predictors constant.

For every 1 % increase in Rebounds the predicted salary is expected to increase by about 0.23% when holding all other predictors constant.

For every 1 % increase in Points scored the predicted salary is expected to increase by about 0.95% when holding all other predictors constant.

# Model Interpretation(coefficients)



## Rookie

For every 1% increase in GS, the predicted salary is expected to increase by 0.18%, holding all other predictors constant

For every 1% increase in AST, the predicted salary is expected to increase by 0.64%, holding all other predictors constant

$$\log(\text{salary}) \sim \text{GS} + \log(\text{GP}) + \log(\text{AST}) + \log(\text{PPM})$$



For every additional Game Played, the predicted salary is expected to increase by 1.08%, holding all other predictors constant

For every 1% increase in PPM, the predicted salary is expected to increase by 1.19%, holding all other predictors constant.

## **Players that were removed to help with salary prediction**

**10 day contract(Late season signing)**

**Shaquille Harrison  
Skylar Mays  
Stanley Umude**

**Exhibit 10 contract(1 year contract)**

**Jacob Gilyard  
RaiQuan Gray**

**Two Way Contracts(Both NBA and G-League)**

**Mac McClung,Gabe York,  
Justin Minaya, Jay Scrubb,  
Jay Huff,Lindell Wigginton**

**Standard contract (very low salary)**

**Jeenathan Williams  
Kobi Simmons**

**GP GS AST PPM**

**2019 - 2020**

**GS PTS TRB**



- **Predicted salary**
  - -\$20,972,839
- **Actual salary**
  - -\$~~8,730,240~~ (salary cap)

**Ja Morant  
(Rookie of the Year)**

- **Predicted salary**
  - -\$27,344,063
  -
- **Actual salary**
  - -\$25,842,697

**Giannis Antetokounmpo  
(MVP)**

**GP GS AST PPM**

**2022 - 2023**

**GS PTS TRB**



- **Predicted salary**
  - -\$15,350,401
- **Actual salary**
  - -\$11,055,120

**Paolo Banchero  
(Rookie of the Year)**



- **Predicted salary**
  - - 28,835,563
- **Actual salary**
  - - 33,616,770

**Joel Embiid (MVP)**

**GP GS AST PPM**

**2023- 2024**

**GS PTS TRB**



- **Predicted salary**
  - -\$12,330,470
- **Actual salary**
  - -\$12,160,680

**Victor Wembanyama**  
**The Alien**  
**(Rookie of the Year)**



- **Predicted salary**
  - - 24,994,091
- **Actual salary**
  - -47,607,350

**Nikola Jokic**  
**Joker**  
**(MVP)**

# Final Thoughts



- **Basketball is a very complicated sport, especially when you are trying to determine what star players should be paid.**
- **We also think that the new mvp this year should still be joker(Nikola Jokic) based on pure stats alone.**

