NBA Players and Their True Value

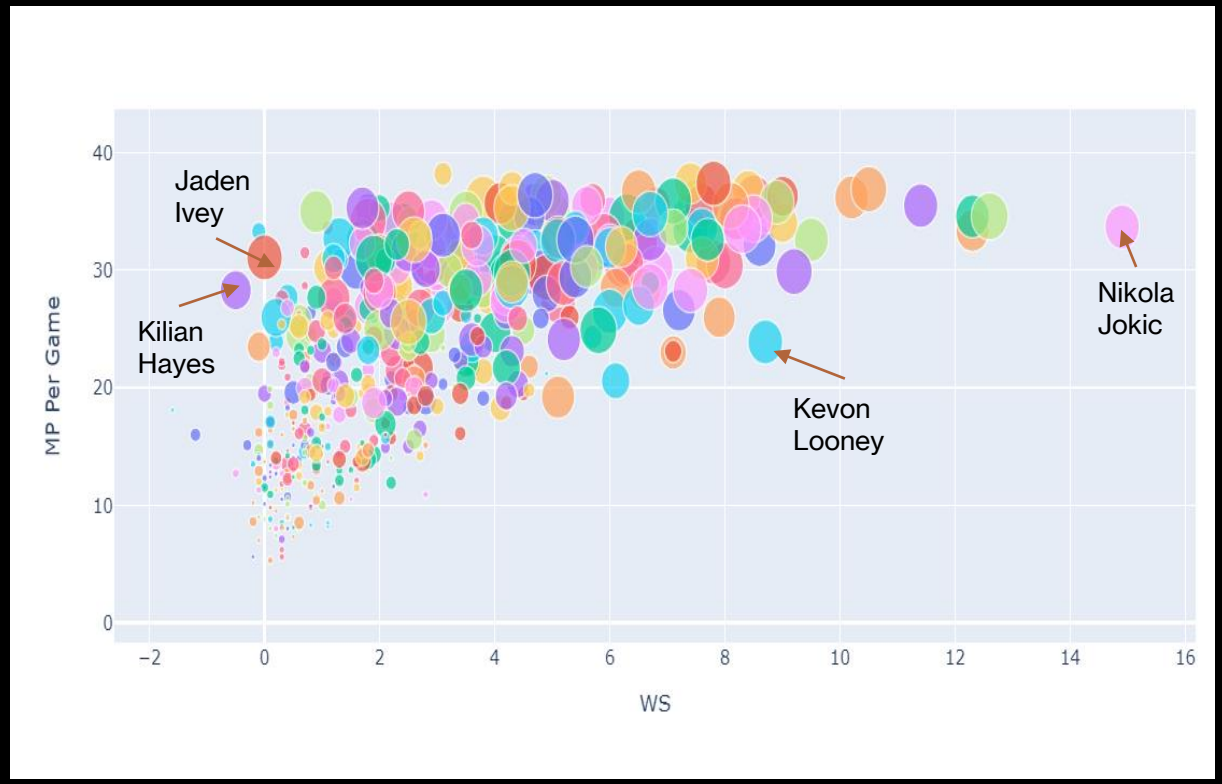
By: Garrett Atkinson, Jake Bonomi, & Ramon Henderson

Introduction

- The three of us are NBA analysts from Notre Dame
- We were brought in by a coalition of NBA front office workers, journalists, and league executives to help show which players are deserving of higher/lower financial compensation/recognition
  - In other words, figuring out their true “value”
- To dig into this project, we used a combination of two different data sets:
  - 1. A collection of 2022 player data from Pro Basketball Reference that captures total stats to advanced stats
  - 2. NBA contracts for players for the 2023-24 season and total Guaranteed Money
- We decided to address the following problems (these will each be displayed as the presentation goes on)

Problem 1: Certain Players Deserve More/Less Minutes

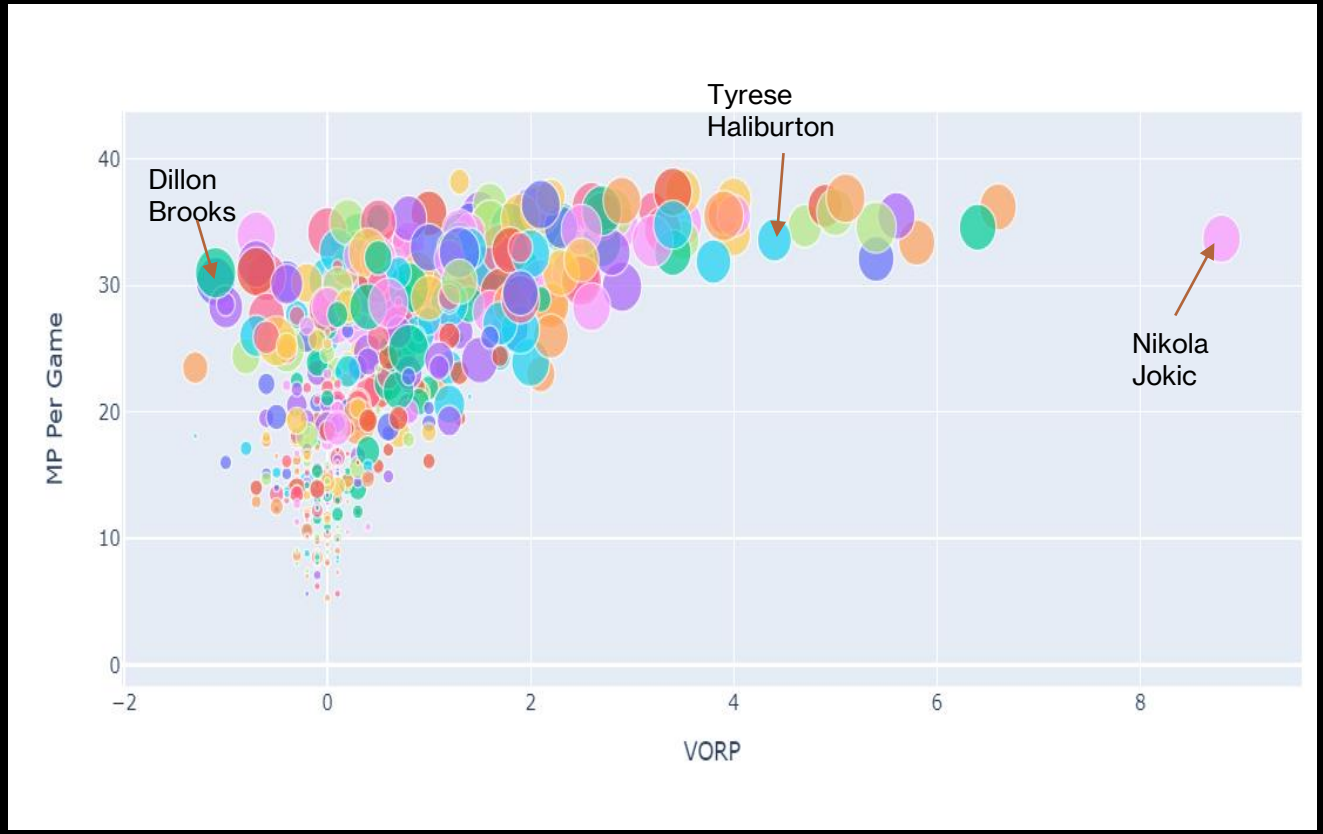
- To address this problem, we decided to focus on the relationship between certain advanced stats (WS, VORP, BPM) and the minutes per game they played.
  - Size of the bubbles was the games they started to account for players who played very few games but got a lot of minutes in the games they played
- First was Win Shares (WS), which divvies up team success to the individual players on the team.



- Based on the visual, we can see that Nikola Jokic deserves every bit of his minutes, while a player like Jaden Ivey, who also plays and starts frequently, should not have received as many minutes as he did
  - Kevon Looney can be argued to receive more minutes using this metric

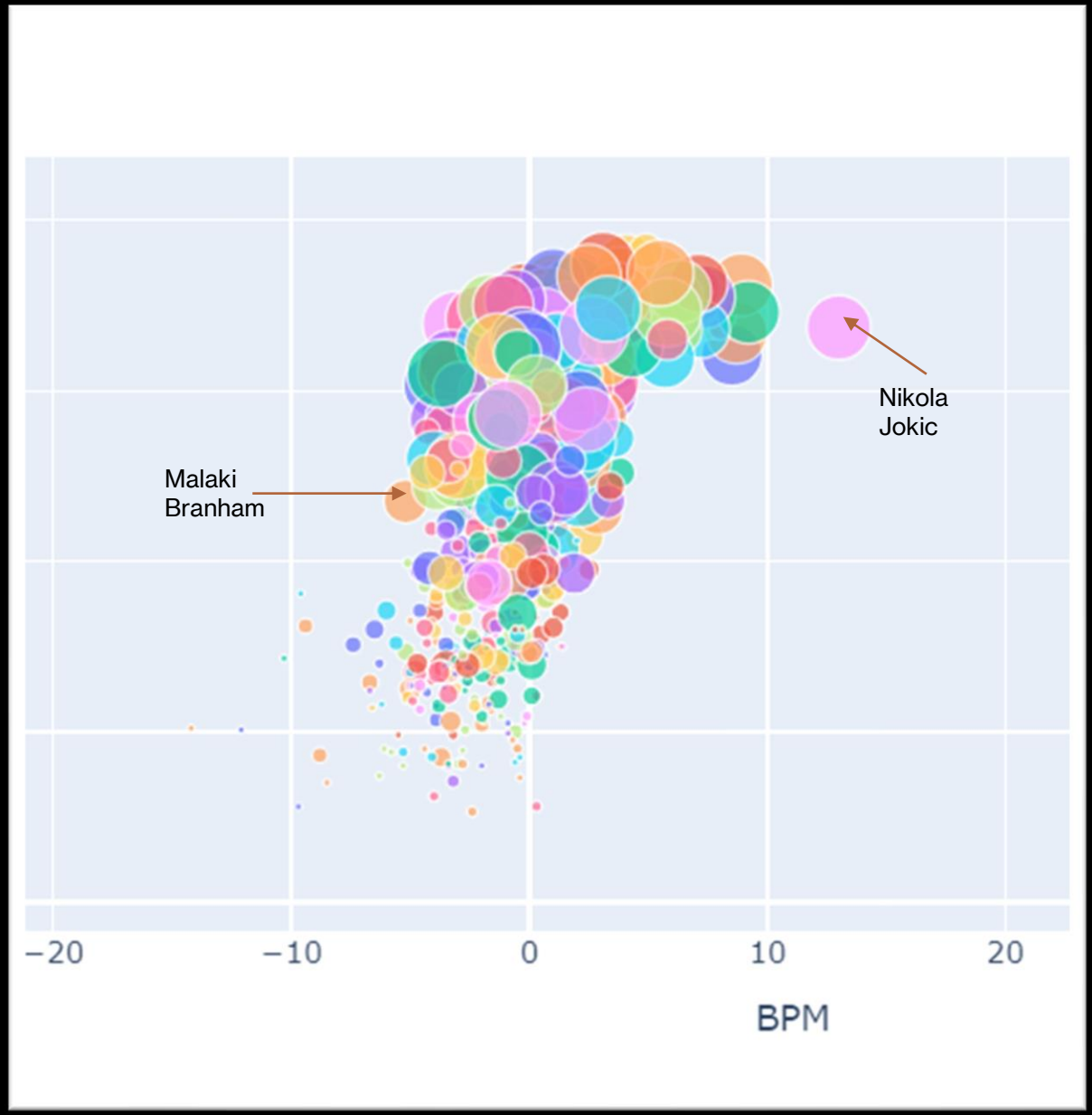
Problem 1: Continued

- Second was Value Over Replacement Player (VORP), using the same format as the previous visual.
  - VORP value means a player contribute that many points to their team in an NBA game, adjusted for playing time using box score stats
  - 10+ VORP is all-time stuff
- Once again, Jokic shows how valuable he is on the court, while a guy like Dillon Brooks actively hurts his team when he's on it, so he should have less minutes.



Problem 1: Continued

- Next is Box Plus/Minus (BPM), using same format
  - Similar stat to VORP, except it does not take playing time into account.
  - A BPM of 10 is phenomenal
- Once again, Nikola Jokic shows how dominant he is and shows how deserving he is of his minutes.
- Meanwhile, a player like Malaki Branham could use a lot of improvement



Problem 2: Main Box Score Stats Don't Always Indicate Player Impact

- We are defining the primary box score stats to be Points, Assists, Rebounds, Blocks, and Steals.
- We ran a linear regression for these stats in 4 different versions (season total, per game, per 36 minutes, per 100 possessions) to try and predict player VORP
  - We chose VORP because of its relationship with box score stats
- Decided to use the model that used season totals as the predictors, as it had the best fit, although the fit value shows reinforces the point of how basic box score stats aren't everything.

```
# Response Variable
Y = nba_2022_data['VORP']
```

✓ 0.0s

```
# Total Boxscore Stats
X = nba_2022_data[['PTS', 'AST', 'TRB', 'STL', 'BLK']]
```

✓ 0.0s

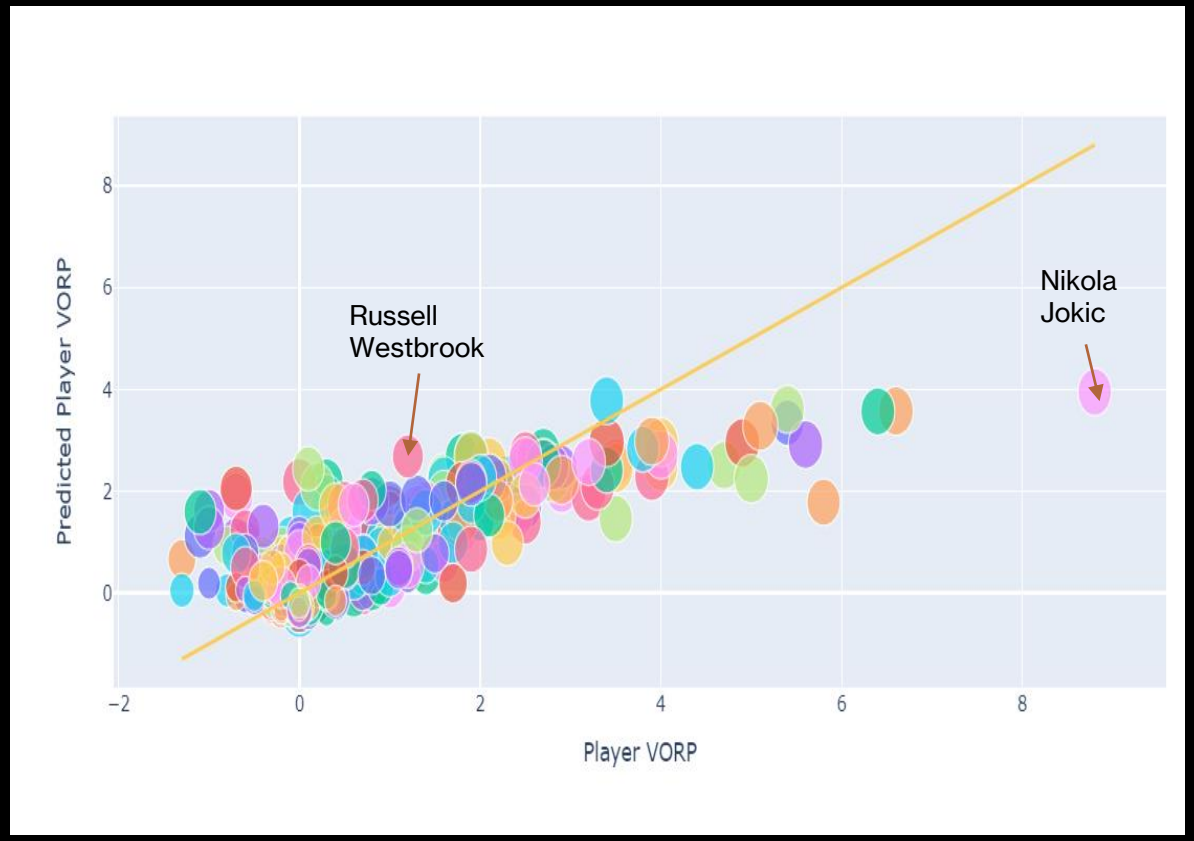
```
reg1 = LinearRegression().fit(X, Y)
reg1.score(X, Y)
```

✓ 0.1s

```
0.6171862015579404
```

Problem 2: Continued

- We then used the regression to create the model to the right.
  - Below the line = better VORP than predicted
  - Above the line = worse VORP than predicted
  - Also set the size to minutes per game further showcase who needs more minutes.
- Nikola Jokic shows himself to be the most valuable player again, as his impact on the court goes well beyond his box score stats, as do other superstars
- Russell Westbrook is a case where his stats don't reflect his true impact, but in a negative way



- Players at the bottom-left of the graph that are beyond the line are players that arguably should have more minutes

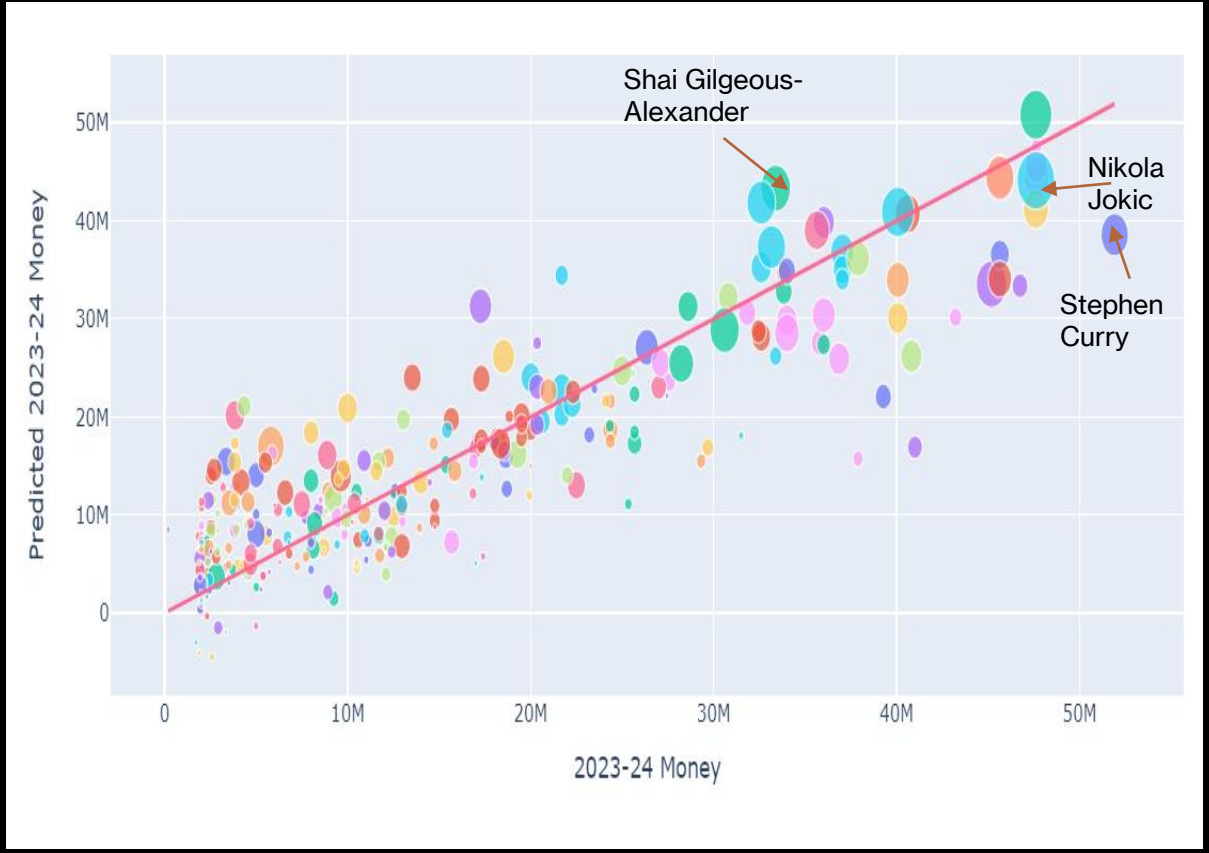
Problem 3: To Be Paid, Or Not To Be Paid?

- Some of the recent extensions signed by NBA players have been record-breaking contracts, particularly for stars
  - This obviously matters because it effects how a roster can be constructed
- Merged this data set with NBA contract data for the 2023-24 season and total Guaranteed money, as we thought it would be interesting to look at if past play warranted the money being paid

```
merged_nba = pd.merge(nba_2022_data, nba_contracts, on="Player")
merged_nba = merged_nba.drop(columns = ['-9999', 'Rk_y', 'Tm_y'])
merged_nba.head()
```

Problem 3: 2023-24 Regression

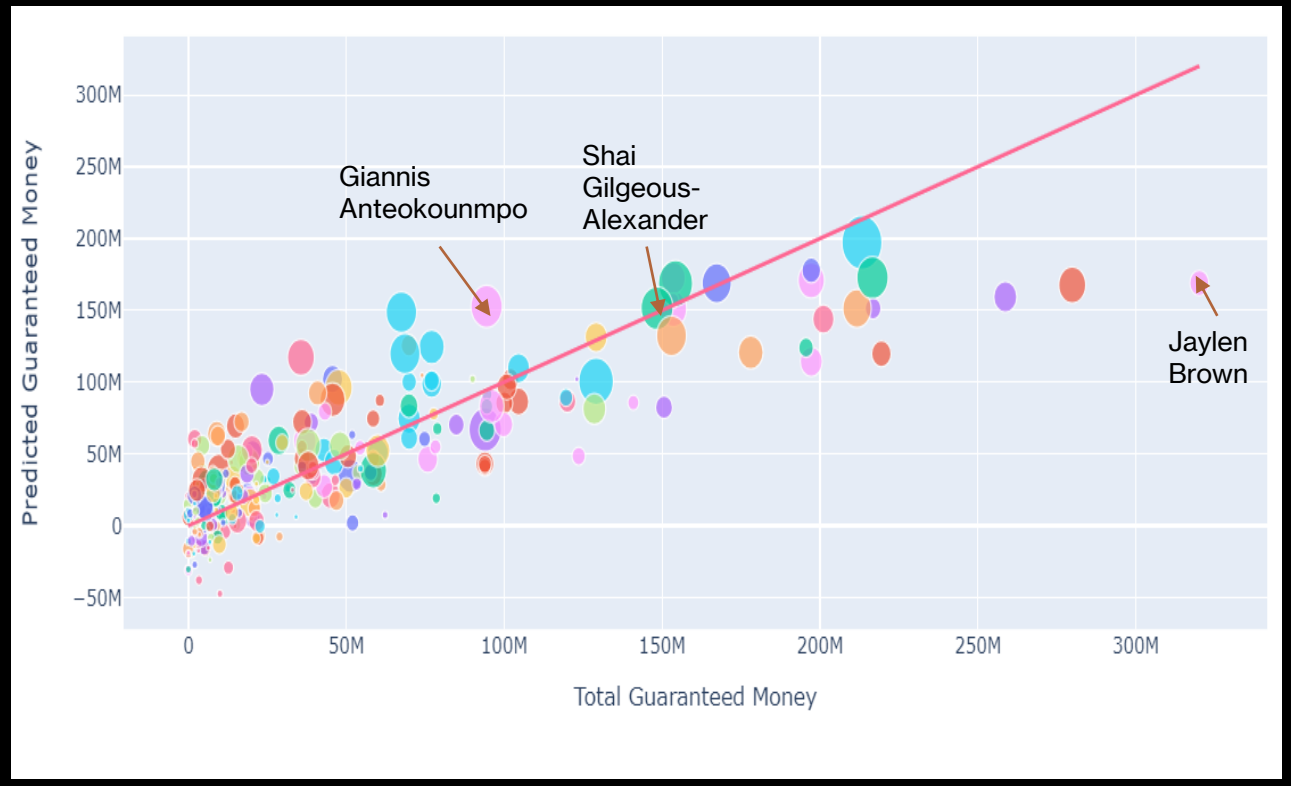
- Used all numerical columns in the player data set to predict their performance and how much money they would have earned for this year.
- For the use of the visual, I filtered out all players who had a negative VORP value so that we can see if high-impact players are, for the most part, earning their money.
- Model fit = 0.818



- This visual shows that players like Stephen Curry earn more money than their 2022 production says they deserve, players like Shai Gilgeous-Alexander are being underpaid, while players like Nikola Jokic are being fairly compensated for their performance.

Problem 3 Guaranteed Regression

- Same exact process as previous model except Guaranteed money is the response.
- Fit = 0.708
- Players like Giannis = deserve more
- Players like J. Brown = deserve less
- Players like SGA = fairly compensated in terms of guaranteed money



Conclusion

- After going through our results, our recommendations for resolving our problem are:
  1. Using advanced stats, figure out who is deserving of more-or-less minutes
  2. Realize that the primary box score stats that everyone uses is not the end-all be-all
  3. Try to pay players as appropriately as possible so that no gross overpays occur
- What would have improved our findings: finding easily accessible data that uses tracking data would allow us to go more in depth in terms of certain shots that players may take, minute-by-minute data, etc.