Identifying The Most Impactful NBA Player in the 2020s

Connor Coniglio, Shaan Patel

Data Introduction

- The Dataset
  - “The Best NBA Players, According to Raptor” through FiveThirtyEight
- Key Metrics
  - Wins Above Replacement (WAR): Reflects players impact compared to replacement-level player
  - RAPTOR: Advanced offense, defense, and overall player contributions
- Data also accounts for regular/postseason differences

Project Overview

- Goal of Project
  - **Identify the most impactful NBA player of the 2020s**
- How?
  - Statistical Analysis of Key Metrics
  - Machine Learning-Based Ranking Model
  - Time-Series Analysis for Trend Direction
- Outcome
  - Determine the top performing superstars and provide insight on player trends and impact in this new generation of basketball

Data Preparation

```
df <- read_csv("modern_RAPTOR_by_player.csv")


#Filter data frame to only include 2020+ seasons
df <- df %>% filter(season >= 2020)

# Regular Season Data frame
regular_season <- df %>%
  mutate(war_total = NULL, war_playoffs = NULL)

# Playoff Data frame
playoffs <- df %>%
  filter(war_playoffs != 0) %>%
  select(player_name, season, war_playoffs, everything())
```

- Filter data to only include 2020 season and beyond
- Create playoff and regular season data frames

Statistical Analysis of Key Metrics

	Reg. Season	Playoffs
WAR Average	1.37	0.437
WAR SD	2.41	0.875
RAPTOR Mean	-1.72	-0.0842
RAPTOR SD	6.40	3.58

- Key Takeaways
  - Regular Season = higher deviation
  - Postseason = more competitive, narrower ranges and higher averages

Player Ranking Sorted by WAR Average

Top 10 Regular Season Players

	player_name	mean_war	mean_raptor
	<chr>	<dbl>	<dbl>
1	Nikola Jokic	15.9	9.63
2	Giannis Antetokounmpo	11.3	7.81
3	Rudy Gobert	11.2	6.77
4	Kawhi Leonard	10.7	8.73
5	Luka Doncic	9.77	6.19
6	Jayson Tatum	9.62	4.61
7	James Harden	9.53	6.06
8	Joel Embiid	9.08	6.73
9	LeBron James	8.67	6.36
10	Jimmy Butler	8.12	5.90

Top 10 Postseason Players

	player_name	mean_war_playoffs	mean_raptor_playoffs
	<chr>	<dbl>	<dbl>
1	Stephen Curry	4.78	6.80
2	Jamal Murray	4.38	3.67
3	LeBron James	3.55	7.17
4	Andrew Wiggins	3.34	2.14
5	Anthony Davis	3.25	5.57
6	Draymond Green	2.78	3.79
7	Kawhi Leonard	2.73	8.73
8	Giannis Antetokounmpo	2.68	7.81
9	Nicolas Batum	2.59	3.38
10	Jimmy Butler	2.53	5.90

Machine Learning-Based Ranked Model

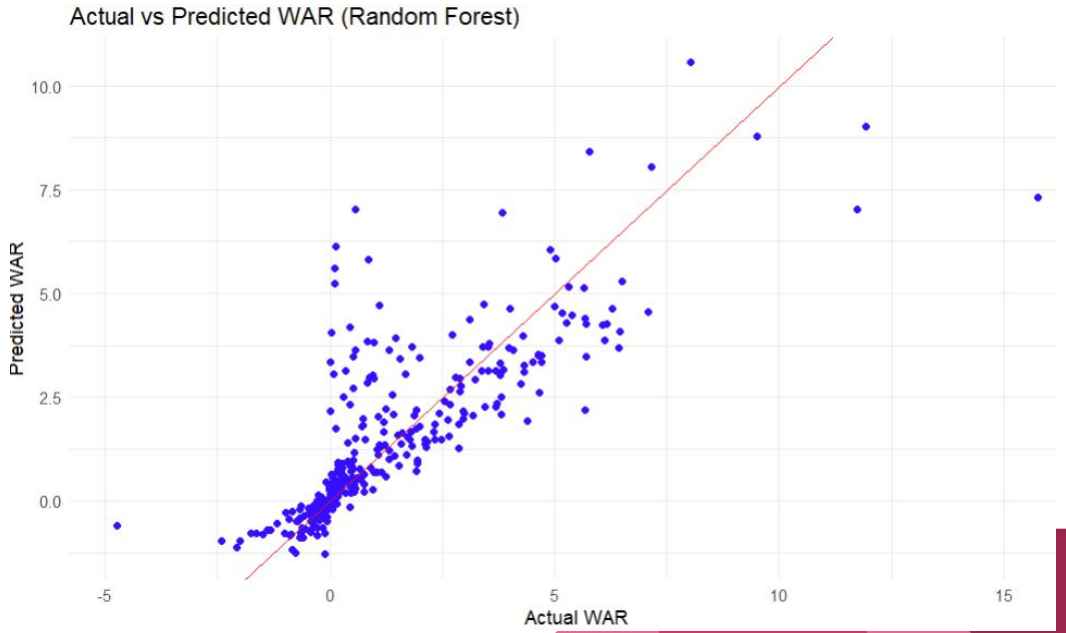
Training / Testing

- Dataset filtered to include seasons from 2020 onward.
- Features used: `raptor_box_offense`, `raptor_box_defense`, `raptor_box_total`, `raptor_onoff_offense`, `raptor_onoff_defense`, `raptor_onoff_total`, `raptor_total`, and `pace_impact`.
- Target variable: `war_reg_season` (Wins Above Replacement during the regular season).
- Dataset split:
  - **80% for training** (to build the model).
  - **20% for testing** (to evaluate model performance).
- Missing values dropped to ensure data integrity.

Machine Learning-Based Ranked Model

Evaluation

player_name	avg_predicted_war
<fct>	<dbl>
Giannis Antetokounmpo	10.6
Nikola Jokic	9.99
Luka Doncic	9.09
Rudy Gobert	8.68
LeBron James	8.59
Kawhi Leonard	8.23
Jayson Tatum	8.03
James Harden	8.01
Stephen Curry	7.81
Jimmy Butler	7.76



Machine Learning-Based Ranked Model

Tuning hyperparameters

2. Machine Learning: Random Forest

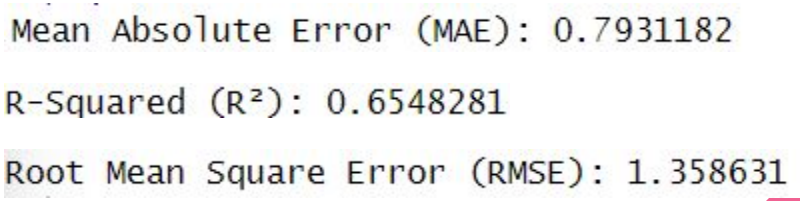
- **Feature Selection**:
  - Include derived metrics like `combined_metric` (WAR + RAPTOR) as an additional feature in the `ml_data` dataset. This could enhance the predictive power of your model.
- **Hyperparameter Tuning**:
  - You've included 200 trees (`ntree = 200`), but tuning other parameters like `mtry` (number of variables tried at each split) and `maxnodes` can improve model performance.

Machine Learning-Based Ranked Model

Insights and Rankings

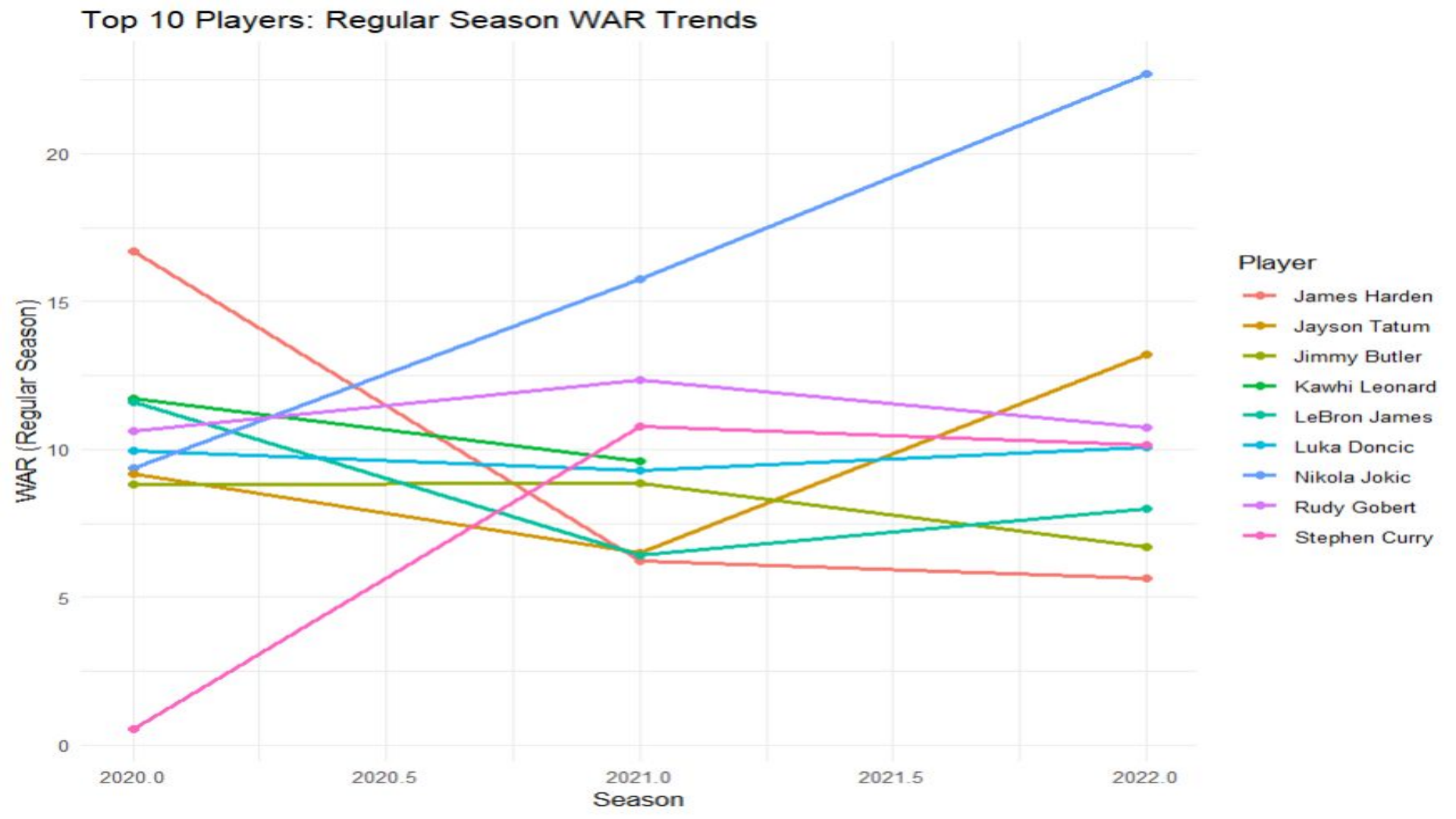
Content

- **Objective**: Summarize key insights and trends from the analysis.
- **Key Points**:
  - Top Players (Regular Season):
    - Highlight top-ranked players from the `player_ranking_ml` results.
  - Trend Analysis:
    - Discuss how top players like Nikola Jokic and Giannis Antetokounmpo performed consistently across seasons.
    - Mention observed discrepancies (e.g., regular-season stars underperforming in playoffs).
  - Player Comparison:
    - Comparison of regular season vs playoff performances.



Time Series Analysis on Trend Direction

- For our time series analysis, we wanted to see how key players have evolved their game over the past couple of years (2020, 2021, 2022)
  - NOTE: We only focused on regular season statistics for this section. Playoff stats were inconclusive due to many players missing the postseason and it being such a small scope
- We used the top 10 players predicted in our ML model
  - Allows us to see trends for these players in this time period


Top 10 Players: Regular Season WAR Trends
WAR (Regular Season)
0
5
10
15
20
2020.0
2020.5
2021.0
2021.5
2022.0
Season
Player
James Harden
Jayson Tatum
Jimmy Butler
Kawhi Leonard
LeBron James
Luka Doncic
Nikola Jokic
Rudy Gobert
Stephen Curry

So, Who is the best player?

Based on our analysis, we can confirm that the best player of the 2020 decade thus far in the NBA is **Nikola Jokic!**

- In hindsight, Jokic has won a MVP in 2021, 2022, and 2024, so this makes sense
- This dataset only extends to 2022, so it is interesting to see that Jokic was already well on his way to stardom and is currently continuing his journey

Why Jokic?

- Statistical Analysis
  - Jokic is far and away the best regular season player (based on both mean WAR && Raptor)
- Machine learning-Based Raking
  - The model only explains ~65% of the variance
  - Mean Absolute Error (MAE): 0.7931182

Giannis Antetokounmpo	10.6
Nikola Jokic	9.99

- Trend Direction Time Series Analysis
  - In the line graph, Jokic starts around the middle of the pack, but consistently shoots up until he is in a league of his own in 2022

Our final decision was between Jokic and Giannis, although Giannis was predicted be the better player with our model, Jokic's well roundedness and high marks in every category allowed him to be more impactful in this timeframe, and ultimately, our decision for the better player.

Thank you!