NBA HONORS SELECTIONS

Almona & Vedant

WHAT IS THE NBA?

- The National Basketball Association (NBA) is a professional basketball league in North America with 30 teams.
- The teams are divided into two conferences (west and east), playing 82 games in the regular season.
- At the end of each regular season, players may receive honors such as All-NBA, All-Defensive, and All-Rookie Team selections.

RESEARCH QUESTION

Can we predict NBA honors using regular season box score statistics?

WHY THIS PROJECT MATTERS

1. All-NBA honors influence supermax contract eligibility.

2. Recognition impacts player value and sports betting markets.

3. Understanding selection patterns can assist in front office decisions and analytics.

DATA SOURCE

Source: Kaggle dataset by Sumitro Datta

- https://www.kaggle.com/datasets/sumitrodatta/nba-aba-baa-stats

Scope: Includes NBA from 1947 - present, but we filtered for 2016-17 until 2023-24 excluding 2019-20 & 2020-21

DATA SOURCE

Variables:

- Games played (g), games started (gs), minutes played (mp), points per 100 possessions, assists, rebounds, steals, blocks, fouls, etc.

Why per 100 possessions stats (over per game or per 36 minutes)?

- Controls for pace and playing time.
- More standardized measure of efficiency across different eras and teams.

DATA SUMMARY

Eligibility Criteria:

- All-NBA/All-Defense: (as of 2023)
  - g >= 65 and mp >= 20 mins per game (except two games)
- All-Rookie: experience == 1

DATA SUMMARY

Descriptive Statistics:

- Correlation matrix to explore relationships.
- Players with more minutes, points, and efficiency tend to receive honors.

	g	gs	mp	fg_per_100_poss	fg_percent	x3p_per_100_poss	x3p_percent	x2p_per_100_poss	x2p_percent	ft_per_100
g	1.00	0.61	0.87	0.27	0.31	0.12	0.19	0.20	0.17	
gs	0.61	1.00	0.85	0.44	0.24	0.12	0.13	0.35	0.12	
mp	0.87	0.85	1.00	0.45	0.27	0.20	0.22	0.32	0.13	
fg_per_100_poss	0.27	0.44	0.45	1.00	0.47	0.20	0.14	0.85	0.31	
fg_percent	0.31	0.24	0.27	0.47	1.00	-0.26	0.03	0.59	0.74	
x3p_per_100_poss	0.12	0.12	0.20	0.20	-0.26	1.00	0.57	-0.34	-0.15	
x3p_percent	0.19	0.13	0.22	0.14	0.03	0.57	1.00	-0.17	-0.08	
x2p_per_100_poss	0.20	0.35	0.32	0.85	0.59	-0.34	-0.17	1.00	0.38	
x2p_percent	0.17	0.12	0.13	0.31	0.74	-0.15	-0.08	0.38	1.00	
ft_per_100_poss	0.12	0.33	0.31	0.62	0.18	-0.01	0.01	0.60	0.06	
ft_percent	0.20	0.18	0.25	0.17	-0.04	0.33	0.25	-0.01	-0.03	
orb_per_100_poss	-0.02	-0.03	-0.10	0.10	0.47	-0.53	-0.35	0.38	0.28	
drb_per_100_poss	0.08	0.14	0.09	0.26	0.38	-0.33	-0.22	0.43	0.27	
trb_per_100_poss	0.05	0.09	0.02	0.22	0.46	-0.45	-0.30	0.46	0.31	
ast_per_100_poss	0.08	0.22	0.23	0.26	-0.07	0.10	0.09	0.19	-0.11	

Showing 1 to 16 of 33 entries, 33 total columns

DATA SUMMARY

MODELING OVERVIEW

Approach:

- Logistic regression (GLM) for binary outcomes.
- Stepwise selection.
- Best subset regression using Cp criterion.

MODELING OVERVIEW

Separate Models Built For:

- All-NBA
- All-Defense
- All-Rookie

ALL-NBA MODEL

	estimate	p-value
(intercept)	-2.43	3.85×10^{-12}
mp	1.84×10^{-3}	0.000515
x3p_per_100_poss	2.50×10^{-1}	0.0955
x2p_per_100_poss	1.87×10^{-1}	0.0641
x2p_percent	6.92	0.0378
ft_percent	10.7	0.000294
trb_per_100_poss	1.23×10^{-1}	0.0376
ast_per_100_poss	1.48×10^{-1}	0.0128
blk_per_100_poss	4.91×10^{-1}	0.0406
shooting_foul_committed	-3.15×10^{-2}	0.00265
shooting_foul_drawn	9.74×10^{-3}	0.00588
playoffs	1.87	0.000277

ALL-NBA MODEL

Since 2-point scored and 2-point % are in the model, we tested whether or not adding an interaction between these terms would improve the model.

H_0 : The interaction term <u>does not</u> significantly improve prediction of All-NBA selection.

H_A : The interaction term significantly improves prediction of All-NBA selection.

```
> anova(test1, test2)
```

Output: p-value is 0.855, which is very high, so we fail to reject the null.

✅ So: adding the interaction <u>does not significantly improve the model's fit.</u>

ALL-DEFENSE MODEL

	estimate	p-value
(intercept)	9.90	3.85×10^{-12}
gs	0.055	4.31×10^{-5}
posPF	-21.52	6.00×10^{-6}
posPG	-17.09	0.000121
posSF	-14.60	0.000157
posSG	-18.66	1.06×10^{-5}
drb_per_100_poss	-0.80	0.003364
stl_per_100_poss	1.47	1.68×10^{-7}
pf_per_100_poss	-2.12	0.00110
posPF:drb_per_100_poss	1.06	0.000225
posSG:blk_per_100_poss	3.83	0.00141
posSF:pf_per_100_poss	2.16	0.0931
drb_per_100_poss:blk_per_100_poss	0.44	0.000109

Other interactions: blk_per_100_poss*shooting_foul_committed, blk_per_100_poss*offensive_foul_drawn

ALL-ROOKIE MODEL

	estimate	p-value
(intercept)	-6.85	1.33×10^{-10}
mp	0.00168	2.74×10^{-11}
fg_per_100_poss	0.280	0.00150
playoffs	0.821	0.0365
blk_per_100_poss	1.88	0.00737
pf_per_100_poss	0.0538	0.748
blk_per_100_poss:pf_per_100_poss	-0.290	0.0311

Insights:

- Minutes and scoring per 100 possessions are strong indicators.
- Rookies who contribute defensively and avoid fouling are more likely to be selected.

ALL-ROOKIE MODEL

Main Term: `pf_per_100_poss`

- **Estimate:** `+0.054;` **p-value:** `0.748` (not statistically significant)
- **Interpretation:** When `blk_per_100_poss = 0`, an increase in personal fouls per 100 possessions is associated with a **slight increase** in the odds of making All-Rookie — **but this effect is statistically meaningless** on its own.

Interaction Term: `blk_per_100_poss:pf_per_100_poss`

- **Estimate:** `-0.290;` **p-value:** `0.031` (statistically significant)
- **Interpretation:** The **positive effect of blocks on All-Rookie odds decreases** as personal fouls increase. In other words:

  "Blocks are good, but not if you're fouling a lot."

This suggests that efficiency matters — rookies who can block shots without fouling are more likely to be rewarded.

2025 SEASON PREDICTIONS: ALL-NBA

Player Name	Team	%Prob	Player Name	Team	%Prob	Player Name	Team	%Prob
Nikola Jokić	DEN	85.65	Tyler Herro	MIA	65.72	Tyrese Haliburton	IND	38.02
Shai Gilgeous-Alexander	OKC	81.33	LeBron James	LAL	64.86	Karl-Anthony Towns	NYK	23.88
Giannis Antetokounmpo	MIL	78.93	Anthony Edwards	MIN	63.18	Jalen Brunson	NYK	22.54
James Harden	LAC	65.79	Stephen Curry	GSW	45.15	Darius Garland	CLE	19.85
Jayson Tatum	BOS	65.53	Cade Cunningham	DET	39.17	Austin Reaves	LAL	17.68

MVP Finalists

2025 SEASON PREDICTIONS: ALL-DEFENSE

Player Name	Team	%Prob	Player Name	Team	%Prob
Toumani Camara	POR	96.40	Evan Mobley	CLE	50.61
Dyson Daniels	ATL	80.41	Nikola Jokić	DEN	44.00
Derrick White	BOS	60.00	Draymond Green	GSW	39.42
Bam Adebayo	MIA	58.95	Giannis Antetokounmpo	MIL	34.65
Jalen Williams	OKC	51.43	Jarrett Allen	CLE	32.35

Defensive Player Of the Year Finalists

2025 SEASON PREDICTIONS: ALL-ROOKIE

Player Name	Position	Team	%Prob
Kel'el Ware	C	MIA	72.50
Alex Sarr	C	WAS	69.27
Yves Missi	C	NOP	62.72
Stephon Castle	PG	SAS	51.19
Matas Buzelis	SF	CHI	43.80

Player Name	Position	Team	%Prob
Zaccharie Risacher	SF	ATL	42.61
Jaylen Wells	SG	MEM	39.85
Branden Carlson	C	OKC	39.49
Bub Carrington	PG	WAS	36.11
Zach Edey	C	MEM	34.76

Rookie Of the Year Finalists

OUTCOME: ALL-ROOKIE

ANNOUNCED TUESDAY, MAY 20 @2PM ET

Player Name	Position	Team	%Prob	Est. Rank
Stephon Castle	PG	SAS	51.19	4
Zach Edey	C	MEM	34.76	10
Zaccharie Risacher	SF	ATL	42.61	6
Alex Sarr	C	WAS	69.27	2
Jaylen Wells	SG	MEM	39.85	7
1st-Team				
Matas Buzelis	SF	CHI	43.80	5
Bub Carrington	PG	WAS	36.11	9
Donovan Clingan	C	POR	5.34	43
Yves Missi	C	NOP	62.72	3
Kel'el Ware	C	MIA	72.50	1
2nd-Team				

Rookie of the Year Winner

Did not make our Top-10

*Branden Carlson, ranked 8th in our model, did not make the team

LIMITATIONS & ASSUMPTIONS

- **Limitations**:
    - No advanced metrics (e.g., BPM, RAPTOR)
    - Voter biases
    - Some player/team context omitted (e.g., team record, media coverage)
- **Assumptions**:
    - Selection criteria consistent across years
    - Variables accurately reported and standardized

CONCLUSION

- **Summary**:
  - Logistic regression can reasonably predict honors selections.
  - Defensive selections more complex due to interactions.
  - Rookie selection influenced by playing time and efficiency.

- **Implications**:
  - Teams can use this to identify rising stars or undervalued contributors.
  - Sports media and bettors may apply similar models.

THANK YOU. QA?