

Predicting Points Per Game in College Basketball

ACC & SEC Player Data 2017-2025

The Research Question



Can we predict a player's **points per game** using minutes played, assists, rebounds, 3-point %, field goal %, field goal attempts, position, and conference?

WHY IT MATTERS

Recruiting & Talent Evaluation

Helps coaches identify statistically which players are likely to score, reducing reliance on pure scouting.

Player Development

Understanding what drives scoring (volume vs. efficiency) informs how programs design practice and training.

Prior Research

Jones & Magel (2016): FG%, assists & turnovers explained >91% of NBA game point spreads

Taber et al. (2024): in-game stats were the strongest predictor of player efficiency in Div. I women's basketball

The Dataset



893

Player-Seasons

2017-25

Seasons Covered

Top 50

By Minutes Played

DATA SOURCES

ESPN Men's College Basketball: ACC & SEC per-game player stats, 18 columns, scraped 2017–2025 (aggregated from official game records)

Wikipedia NBA Draft pages (2017–2025): draft year, round, pick, and team; merged on player + season

KEY VARIABLES

Variable	Description
fga, fg%, 3p%	Shooting volume & efficiency
min	Average minutes per game
ast, reb	Assists & rebounds per game
pos, conference	Position (G/F/C), ACC (1) vs. SEC (0)
pts	Points per game

```
<class 'pandas.core.frame.DataFrame'>
```

```
Index: 893 entries, 0 to 899
```

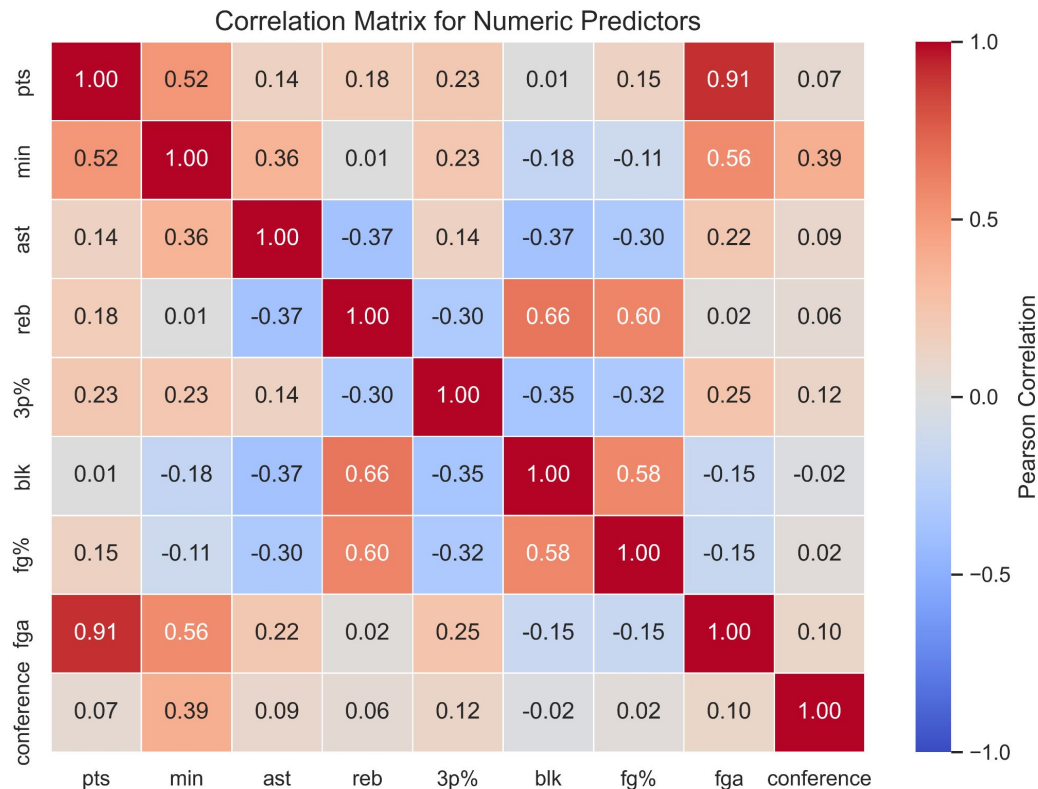
```
Data columns (total 21 columns):
```

#	Column	Non-Null Count	Dtype
0	position	893 non-null	object
1	gp	893 non-null	int64
2	min	893 non-null	float64
3	pts	893 non-null	float64
4	fgm	893 non-null	float64
5	fga	893 non-null	float64
6	fg%	893 non-null	float64
7	3pm	893 non-null	float64
8	3pa	893 non-null	float64
9	3p%	890 non-null	float64
10	ftm	893 non-null	float64
11	fta	893 non-null	float64
12	ft%	893 non-null	float64
13	reb	893 non-null	float64
14	ast	893 non-null	float64
15	stl	893 non-null	float64
16	blk	893 non-null	float64
17	to	893 non-null	float64
18	player	893 non-null	object
19	school	893 non-null	object
20	conference	893 non-null	int32

```
dtypes: float64(16), int32(1), int64(1), object(3)
```

```
memory usage: 150.0+ KB
```

Exploring the Data: Correlation Matrix



$$r = 0.91$$

pts $\leftrightarrow$ fga

fga

Strongest predictor of scoring ($r = 0.91$).
Shot volume drives points.

min

Correlated with pts ($r = 0.52$)
more minutes = more opportunity to score.

blk

Near-zero correlation with pts ($r = 0.01$)
excluded from the model.

Modeling Approach: Regression



01

Nearly identical performance

MLR matches Lasso and forward/backward selection within rounding.

02

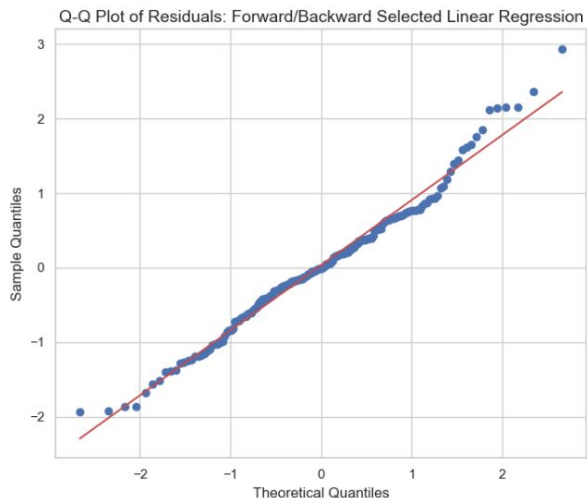
Interpretable coefficients

Coaches and analysts can directly read which variables drive scoring output.

03

Less tuning complexity

No hyperparameter search needed, simpler and more reliable in production.



Model Performance Table

	Model	RMSE	MAE	Test R ²
0	Baseline Mean Model	3.553	2.885	-0.016
1	Multiple Linear Regression	0.876	0.668	0.938
2	Lasso Regression	0.873	0.665	0.939
3	Forw/Back Selection Linear Regression	0.870	0.661	0.939
4	KNN Regression	1.282	1.026	0.868
5	Decision Tree Regression	1.101	0.860	0.902

Variance Inflation Factor (VIF) for Predictors

Feature	VIF
9 position_G	12.940
8 position_F	9.405
1 reb	3.012
7 blk	2.486
2 min	1.961
3 fg%	1.898
4 fga	1.609
5 ast	1.413
6 3p%	1.282
10 conference	1.209

Results: What Drives Scoring?



Forward & Backward Selection

	coef	std err	t	P> t
const	-9.0541	0.375	-24.146	0.000
fga	1.2665	0.015	86.281	0.000
fg%	0.1795	0.006	29.624	0.000
3p%	0.0367	0.004	9.175	0.000
conference	-0.3818	0.070	-5.464	0.000
ast	0.0714	0.028	2.580	0.010

KEY INSIGHTS

fga (coef = 1.2665) dominates; shot volume is the #1 driver of scoring

fg% (coef = 0.1795); efficiency matters but less than volume

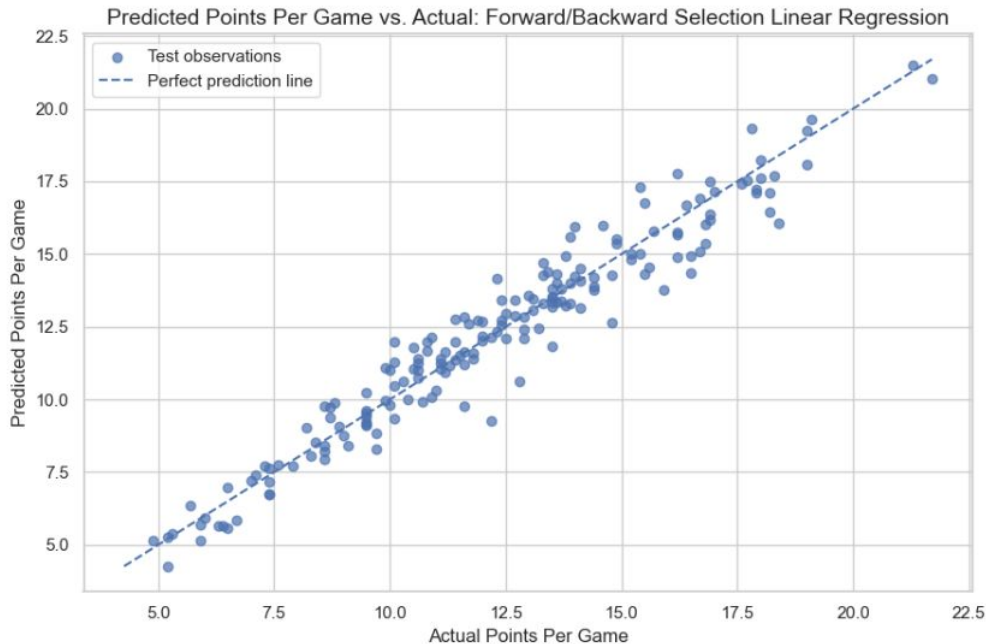
3p% and ast, have small but positive effects

Feature Selection Predictors

	Forward Selection	Backward Selection
0	fga	ast
1	fg%	3p%
2	3p%	fg%
3	conference	fga
4	ast	conference

$$PPG = -9.054 + 1.267(FGA) + 0.180(FG\%) + 0.037(3P\%) - 0.382(Conference) + 0.071(AST)$$

Results: Predicted vs. Actual



0.870

Test RMSE

0.661

Test MAE

0.939

Test R^2

3.553

Baseline RMSE

Conclusion & Takeaways



Box score statistics alone explain 93.9% of the variation in ACC and SEC college scoring output, shot volume and efficiency are what matter most

FINDINGS	LIMITATIONS	FUTURE DIRECTIONS
fga & fg% are the dominant predictors of scoring	Missing pace/tempo context, raw stats are system-dependent, less data on centers	Use possession-adjusted or per-100-possession stats
Conference has a small but real effect (ACC-0.43 pts)	No strength-of-schedule adjustment	Incorporate team-level features (pace, coaching system)
MLR outperforms KNN and Decision Tree	High VIF on position variables limits those coefficients	Extend to transfer portal predictions

References



ESPN. (2025). 2017–25 Men's College Basketball ACC & SEC Stat Leaders. ESPN Internet Ventures.

Jones, E. S., & Magel, R. C. (2016). Predicting Outcomes of NBA Basketball Games. *Journal of Advance Research in Business, Management and Accounting*.

Taber, et al. (2024). Machine learning to predict performance in Division I women's basketball.

Coma-Ribes, M., et al. (2026). Identifying Individual Playing Roles in Professional and Semi-Professional Basketball. *Retos*.

Wikipedia contributors. (2025). NBA Draft (2017–2025). Wikipedia, The Free Encyclopedia.

OpenAI. ChatGPT. Used for assistance with visualizations and modeling.



WELLS FARGO		UNC	
PTS	FG	PTS	FG
1 FOSTER	10-18	1 FOSTER	10-18
2 FOSTER	10-18	2 FOSTER	10-18
3 FOSTER	10-18	3 FOSTER	10-18
4 FOSTER	10-18	4 FOSTER	10-18
5 FOSTER	10-18	5 FOSTER	10-18
6 FOSTER	10-18	6 FOSTER	10-18
7 FOSTER	10-18	7 FOSTER	10-18
8 FOSTER	10-18	8 FOSTER	10-18
9 FOSTER	10-18	9 FOSTER	10-18
10 FOSTER	10-18	10 FOSTER	10-18
11 FOSTER	10-18	11 FOSTER	10-18
12 FOSTER	10-18	12 FOSTER	10-18
13 FOSTER	10-18	13 FOSTER	10-18
14 FOSTER	10-18	14 FOSTER	10-18
15 FOSTER	10-18	15 FOSTER	10-18
16 FOSTER	10-18	16 FOSTER	10-18
17 FOSTER	10-18	17 FOSTER	10-18
18 FOSTER	10-18	18 FOSTER	10-18
19 FOSTER	10-18	19 FOSTER	10-18
20 FOSTER	10-18	20 FOSTER	10-18
21 FOSTER	10-18	21 FOSTER	10-18
22 FOSTER	10-18	22 FOSTER	10-18
23 FOSTER	10-18	23 FOSTER	10-18
24 FOSTER	10-18	24 FOSTER	10-18
25 FOSTER	10-18	25 FOSTER	10-18
26 FOSTER	10-18	26 FOSTER	10-18
27 FOSTER	10-18	27 FOSTER	10-18
28 FOSTER	10-18	28 FOSTER	10-18
29 FOSTER	10-18	29 FOSTER	10-18
30 FOSTER	10-18	30 FOSTER	10-18
31 FOSTER	10-18	31 FOSTER	10-18
32 FOSTER	10-18	32 FOSTER	10-18
33 FOSTER	10-18	33 FOSTER	10-18
34 FOSTER	10-18	34 FOSTER	10-18
35 FOSTER	10-18	35 FOSTER	10-18
36 FOSTER	10-18	36 FOSTER	10-18
37 FOSTER	10-18	37 FOSTER	10-18
38 FOSTER	10-18	38 FOSTER	10-18
39 FOSTER	10-18	39 FOSTER	10-18
40 FOSTER	10-18	40 FOSTER	10-18
41 FOSTER	10-18	41 FOSTER	10-18
42 FOSTER	10-18	42 FOSTER	10-18
43 FOSTER	10-18	43 FOSTER	10-18
44 FOSTER	10-18	44 FOSTER	10-18
45 FOSTER	10-18	45 FOSTER	10-18
46 FOSTER	10-18	46 FOSTER	10-18
47 FOSTER	10-18	47 FOSTER	10-18
48 FOSTER	10-18	48 FOSTER	10-18
49 FOSTER	10-18	49 FOSTER	10-18
50 FOSTER	10-18	50 FOSTER	10-18
51 FOSTER	10-18	51 FOSTER	10-18
52 FOSTER	10-18	52 FOSTER	10-18
53 FOSTER	10-18	53 FOSTER	10-18
54 FOSTER	10-18	54 FOSTER	10-18
55 FOSTER	10-18	55 FOSTER	10-18
56 FOSTER	10-18	56 FOSTER	10-18
57 FOSTER	10-18	57 FOSTER	10-18
58 FOSTER	10-18	58 FOSTER	10-18
59 FOSTER	10-18	59 FOSTER	10-18
60 FOSTER	10-18	60 FOSTER	10-18
61 FOSTER	10-18	61 FOSTER	10-18
62 FOSTER	10-18	62 FOSTER	10-18
63 FOSTER	10-18	63 FOSTER	10-18
64 FOSTER	10-18	64 FOSTER	10-18
65 FOSTER	10-18	65 FOSTER	10-18
66 FOSTER	10-18	66 FOSTER	10-18
67 FOSTER	10-18	67 FOSTER	10-18
68 FOSTER	10-18	68 FOSTER	10-18
69 FOSTER	10-18	69 FOSTER	10-18
70 FOSTER	10-18	70 FOSTER	10-18
71 FOSTER	10-18	71 FOSTER	10-18
72 FOSTER	10-18	72 FOSTER	10-18
73 FOSTER	10-18	73 FOSTER	10-18
74 FOSTER	10-18	74 FOSTER	10-18
75 FOSTER	10-18	75 FOSTER	10-18
76 FOSTER	10-18	76 FOSTER	10-18
77 FOSTER	10-18	77 FOSTER	10-18
78 FOSTER	10-18	78 FOSTER	10-18
79 FOSTER	10-18	79 FOSTER	10-18
80 FOSTER	10-18	80 FOSTER	10-18
81 FOSTER	10-18	81 FOSTER	10-18
82 FOSTER	10-18	82 FOSTER	10-18
83 FOSTER	10-18	83 FOSTER	10-18
84 FOSTER	10-18	84 FOSTER	10-18
85 FOSTER	10-18	85 FOSTER	10-18
86 FOSTER	10-18	86 FOSTER	10-18
87 FOSTER	10-18	87 FOSTER	10-18
88 FOSTER	10-18	88 FOSTER	10-18
89 FOSTER	10-18	89 FOSTER	10-18
90 FOSTER	10-18	90 FOSTER	10-18
91 FOSTER	10-18	91 FOSTER	10-18
92 FOSTER	10-18	92 FOSTER	10-18
93 FOSTER	10-18	93 FOSTER	10-18
94 FOSTER	10-18	94 FOSTER	10-18
95 FOSTER	10-18	95 FOSTER	10-18
96 FOSTER	10-18	96 FOSTER	10-18
97 FOSTER	10-18	97 FOSTER	10-18
98 FOSTER	10-18	98 FOSTER	10-18
99 FOSTER	10-18	99 FOSTER	10-18
100 FOSTER	10-18	100 FOSTER	10-18

COUNT 10 68 2ND 2.0 68