

## STAT 451 Project 27 Proposal: Lines of Code:

```
nba_combine_data = pd.read_csv(r'C:\Users\kenne\OneDrive - UW-Madison\2023-2024 fall semester\Machine Learning\nba_combine_dataset.csv')
players_data = pd.read_csv(r'C:\Users\kenne\OneDrive - UW-Madison\2023-2024 fall semester\Machine Learning\Players.csv')
nba_combine_data.head(3)
```

Unnamed: 0	Player	Year	Draft pick	Height (No Shoes)	Height (With Shoes)	Wingspan	Standing reach	Vertical (Max)	Vertical (Max Reach)	Vertical (No Step)	Vertical (No Step Reach)	Weight	Body Fat	Hand (Length)	Hand (Width)	Bench	Agility	Sprint
0	Blake Griffin	2009	1.0	80.5	82.00	83.25	105.0	35.5	140.5	32.0	137.0	248.0	8.2	NaN	NaN	22.0	10.95	3.28
1	Terrence Williams	2009	11.0	77.0	78.25	81.00	103.5	37.0	140.5	30.5	134.0	213.0	5.1	NaN	NaN	9.0	11.15	3.18
2	Gerald Henderson	2009	12.0	76.0	77.00	82.25	102.5	35.0	137.5	31.5	134.0	215.0	4.4	NaN	NaN	8.0	11.17	3.14

```
players_data.head(3)
```

	_id	birthDate	birthPlace	career_AST	career_FG%	career_FG3%	career_FT%	career_G	career_PER	career_PTS	...	draft_pick	draft_round	draft_team	draft_year	height	highSchool
0	abdela01	24-Jun-68	Cairo, Egypt	0.3	50.2	0	70.1	256	13	5.7	...	25th overall	1st round	Portland Trail Blazers	1990	10-Jun	Bloomfield New Jersey
1	abdulza01	7-Apr-46	Brooklyn, New York	1.2	42.8	NaN	72.8	505	15.1	9.0	...	5th overall	1st round	Cincinnati Royals	1968	9-Jun	John Jay Brooklyn New York
2	abdulka01	16-Apr-47	New York, New York	3.6	55.9	5.6	72.1	1560	24.6	24.6	...	1st overall	1st round	Milwaukee Bucks	1969	2-Jul	Powder Mill in New York, New York

### After cleanup:

```
1 merged_df.head(3)
```

	Player	Height (No Shoes)	Wingspan	Standing reach	Vertical (Max Reach)	Weight	Body Fat	Agility	Sprint	draft_pick	career_PER	career_WS
0	Blake Griffin	80.5	83.25	105.0	140.5	248.0	8.2	10.95	3.28	1.0	22.2	75.3
1	Terrence Williams	77.0	81.00	103.5	140.5	213.0	5.1	11.15	3.18	11.0	11.5	-0.2
2	Gerald Henderson	76.0	82.25	102.5	137.5	215.0	4.4	11.17	3.14	64.0	13.4	32.1

```
1 merged_df.shape
```

```
(423, 12)
```

X features: Height (No shoes), Wingspan, Standing reach, Vertical (Max Reach), weight, Body Fat, Agility, Sprint

Y target variables: Career\_WS (Win Share- a measure that is assigned to players based on their offense, defense, and playing time. "A win share is worth one-third of a team win. If a team wins 60 games, there are 180 'Win Shares' to distribute among the players."(2)), Salaries

### Questions:

What player statistics are important for winning?

What player statistics are fiscally underrated?

Can we use NBA Combine numbers to accurately predict career statistics?

Methods: Feature Engineering, Feature Selection, Data Imputation, Binning, Decision Trees, Linear Regressor, Kernel Regressor, Model Performance Assessments, and Hyperparameter Tuning

Data Set Link: <https://www.kaggle.com/datasets/thedevastator/nba-draft-combine-measurement-data-from-2012-201>

