
Work in Progress Presentation

— Abigail Weinstein, Abigail Brown,
and Michael Shepard —

Overview

Project Focus: NBA Regular Season Wins

We are investigating if certain reported gameplay and team statistics are correlated to more wins in a given season.

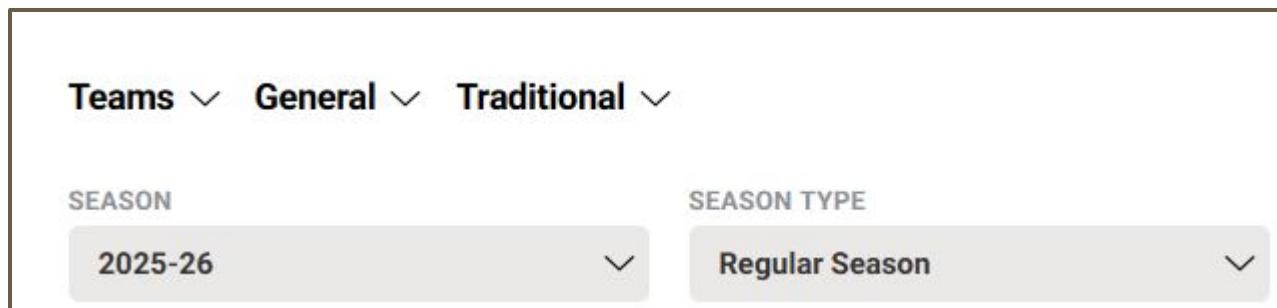
Specifically, we are observing how **attempted 3PT rate**, **average player age**, and **pace** of a team correlate to the number of wins earned over the regular season.

Data Collection

Source: National Basketball Association

Time Frame: 2015 - 2025

Unit of Interest: Team



Resulted in 11 separate data frames

Data Wrangling

Merged 11 season dataframes into one complete dataframe with:

- 11 years of regular season games
- 30 teams
- 29 variables

Rk	Team	Age	W	L	PW	PL	MOV	SOS	SRS	ORtg	DRtg	NRtg	Pace	FTr	3PAr	TS%
1	Golden State Warriors*	26.6	67	15	65	17	10.10	-0.09	10.01	111.6	101.4	+10.2	98.3	.239	.311	.571
2	Los Angeles Clippers*	28.8	56	26	58	24	6.59	0.22	6.80	112.4	105.5	+6.9	94.7	.303	.322	.565
3	San Antonio Spurs*	29.8	55	27	58	24	6.20	0.14	6.34	108.5	102.0	+6.5	93.8	.256	.269	.555
4	Atlanta Hawks*	27.8	60	22	56	26	5.43	-0.68	4.75	108.9	103.1	+5.8	93.9	.259	.321	.563
5	Cleveland Cavaliers*	26.9	53	29	53	29	4.48	-0.40	4.08	111.1	106.3	+4.8	92.3	.287	.334	.557
6	Portland Trail Blazers*	27.0	51	31	53	29	4.23	0.18	4.41	108.2	103.7	+4.5	94.2	.225	.316	.544
7	Houston Rockets*	27.6	56	26	50	32	3.44	0.38	3.82	107.0	103.4	+3.6	96.5	.312	.392	.548
8	Memphis Grizzlies*	29.6	55	27	50	32	3.24	0.38	3.62	105.7	102.2	+3.5	92.0	.276	.184	.531

```
32  
33 #merge dataframes  
34 allNBA <- rbind(  
35   data2015, data2016, data2017, data2018, data2019,  
36   data2020, data2021, data2022, data2023, data2024, data2025  
37 )  
38  
39 #save data frame  
40 write.csv(allNBA, file = "Data/allNBA.csv")  
41
```

Data Filtering

Selected eight primary variables:

- Team → identifies team data is associated with
- Year → identifies season data is associated with
- Rank → highlights overall team performance
- *Wins* → **outcome variable**
- Losses
- *3-Point Att.* → % of total points attempted that were 3-point shots, **predictor**
- *Age* → average team age during a given season, **predictor**
- *Pace* → number of possessions a team has per game related to playing time, **predictor**

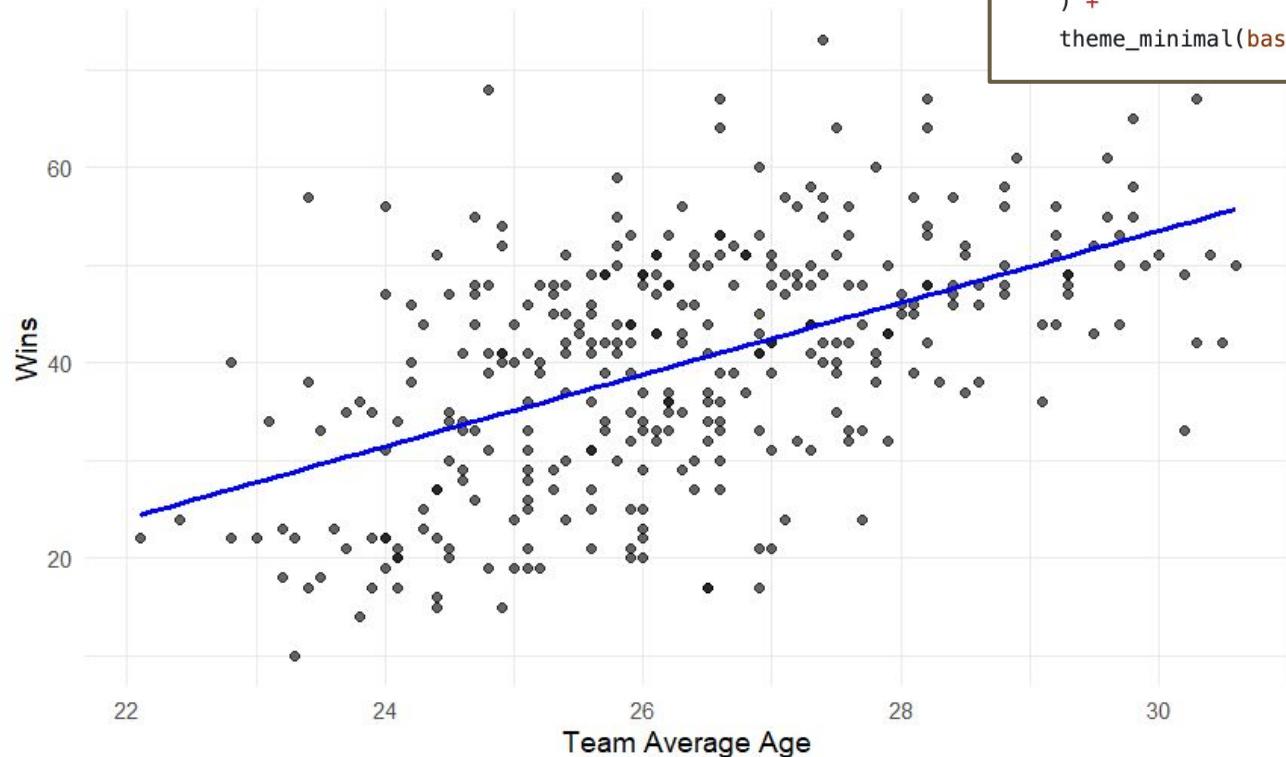
Data Summary

Summary Table, NBA Overall									
	Data Type	Description	Count	Min. Val.	Median Val.	Mean Val.	Max. Val.	Std. Dev.	Range
Team	String	Team Name	330	-	-	-	-	-	30
Team Rank	Ordered Num.	Rank in NBA	330	1	15.5	15.5	30	8.67	30
Year	Ordered Num.	Season Year	330	2015	2020	2020	2020	3.17	10
Wins	Numeric	# of wins	330	10	42	40	73	12.2	63
Losses	Numeric	# of losses	330	9	39	40	72	12	63
3-Point Att.	Numeric	% of Total Attempted Shots as Freethrows	330	0.179	0.362	0.358	0.536	0.0643	0.357
Pace	Numeric	Avg. Possessions per Game	330	90.4	98.2	98	105	2.68	14.7
Avg. Age	Numeric	Avg. Team Age	330	22.1	26.2	26.3	30.6	1.74	8.5

Team	Avg. Rank	Avg. Wins	Avg. Losses	Avg. 3-Point Attempt %	Avg. Pace	Avg. Age
Golden State Warriors	8.18	51.55	28.00	0.40	99.85	27.49
Los Angeles Clippers	8.73	48.45	31.73	0.36	97.83	28.75
San Antonio Spurs	14.18	41.36	38.73	0.32	97.86	27.18
Atlanta Hawks	17.45	38.64	41.09	0.37	99.30	26.00
Cleveland Cavaliers	13.45	43.45	36.09	0.38	96.47	26.54
Portland Trail Blazers	17.36	39.27	41.09	0.37	97.68	25.58
Milwaukee Bucks	10.91	48.00	32.27	0.36	99.03	27.42
Indiana Pacers	15.00	41.55	38.73	0.33	98.27	26.28
Utah Jazz	11.45	41.55	38.64	0.39	96.65	26.27
Boston Celtics	6.00	51.09	29.09	0.40	97.50	26.14
Phoenix Suns	17.36	36.82	43.45	0.34	98.89	26.33
Detroit Pistons	21.09	30.09	49.55	0.35	97.18	25.16
Miami Heat	14.18	43.00	37.27	0.37	95.73	27.35
Brooklyn Nets	19.73	34.45	45.73	0.37	98.22	26.76
Charlotte Hornets	21.27	32.55	47.00	0.37	97.38	25.60
Denver Nuggets	11.45	45.82	34.45	0.35	97.39	26.04
Sacramento Kings	20.73	35.09	45.09	0.34	98.57	26.34
Orlando Magic	21.91	32.18	48.09	0.34	97.42	24.71
Los Angeles Lakers	19.45	36.64	43.45	0.34	99.19	27.29
Minnesota Timberwolves	17.45	35.82	43.64	0.34	98.34	25.34

Average Player Age vs. Wins

Average Age vs Wins (2015–2025)



```
ggplot(AllNBA, aes(x = Age, y = W)) +  
  geom_point(alpha = 0.6) +  
  geom_smooth(method = "lm", se = FALSE, color = "blue") +  
  labs(  
    title = "Average Age vs Wins (2015–2025)",  
    x = "Team Average Age",  
    y = "Wins"  
) +  
  theme_minimal(base_size = 14)
```

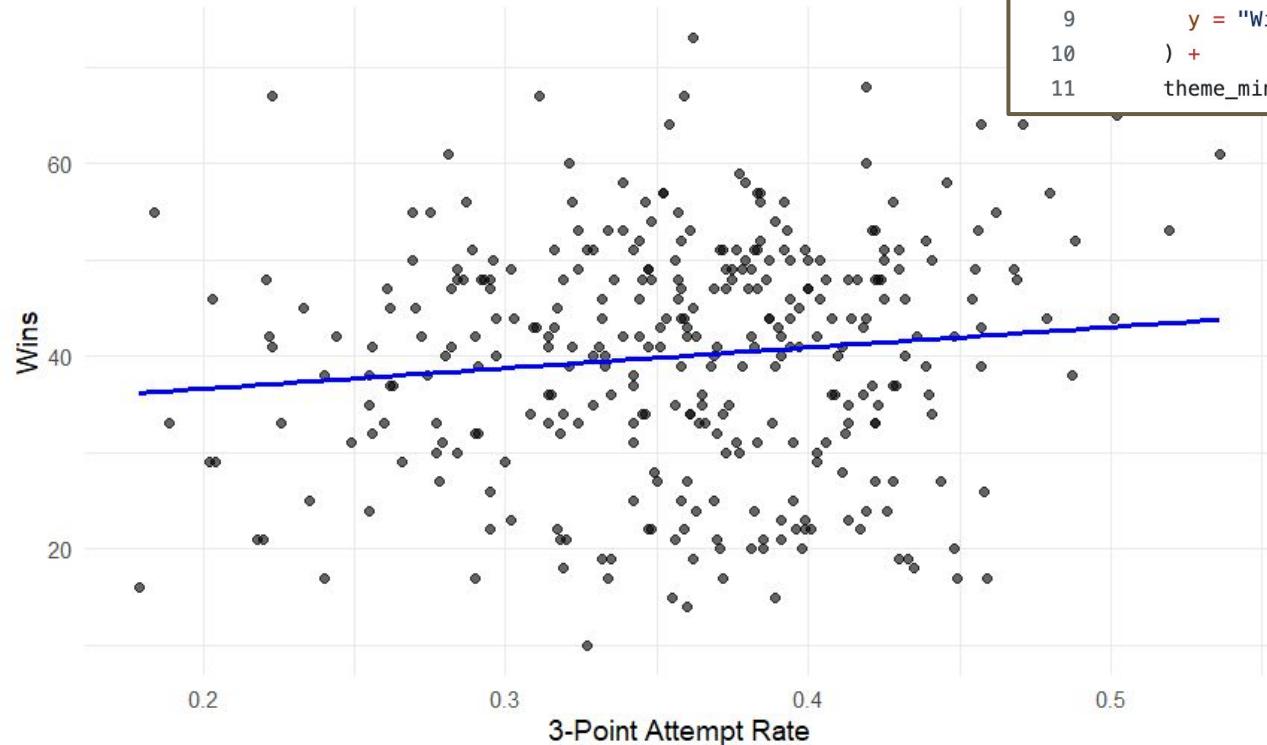
Avg Age:

$r = .5252$

Moderate,
Positive
Relationship

Attempted 3PT Rate vs. Wins

3P Attempt Rate vs Wins (2015–2025)



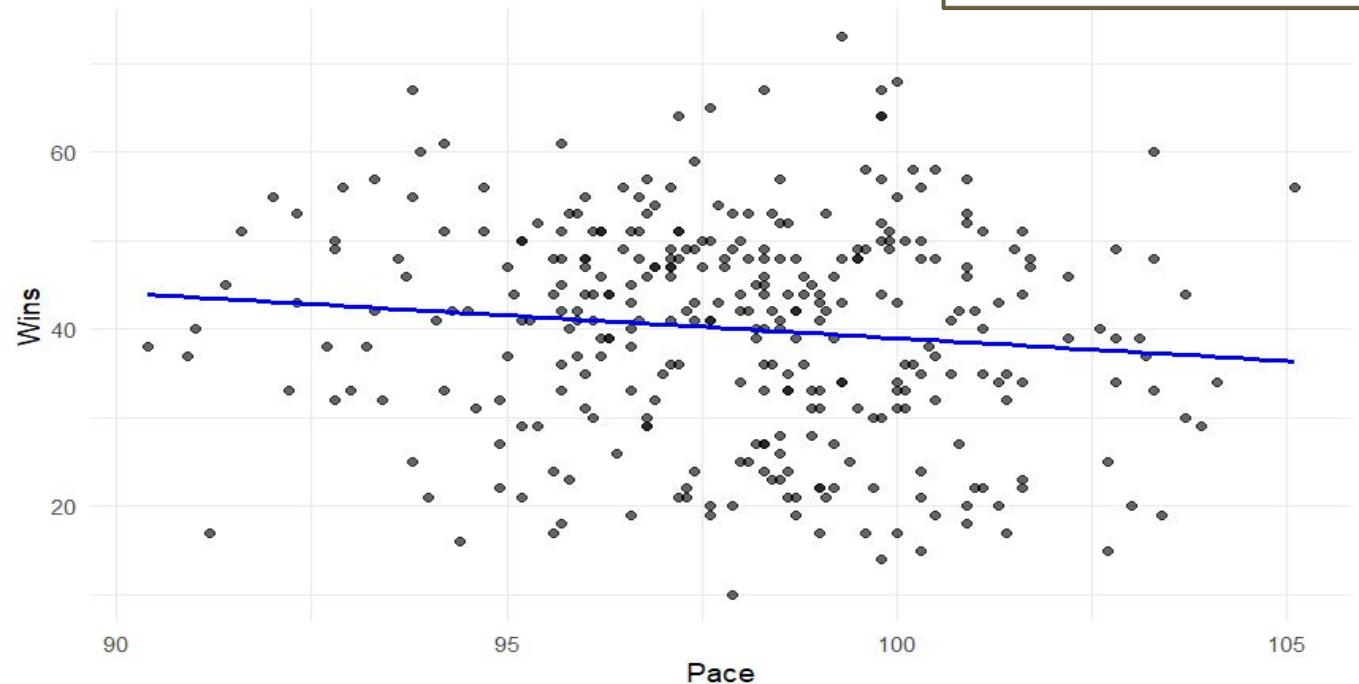
```
1 library(ggplot2)
2
3 ggplot(AllNBA, aes(x = ThreePAR, y = W)) +
4   geom_point(alpha = 0.6) +
5   geom_smooth(method = "lm", se = FALSE, color = "blue") +
6   labs(
7     title = "3P Attempt Rate vs Wins (2015–2025)",
8     x = "3-Point Attempt Rate",
9     y = "Wins"
10   ) +
11   theme_minimal(base_size = 14)
```

Rate of
Attempted 3PT:

 $r = .1131$
Weak, Positive
Relationship

Pace vs. Wins

Pace vs Wins (2015–2025)



```
23 ggplot(AllNBA, aes(x = Pace, y = W)) +  
24   geom_point(alpha = 0.6) +  
25   geom_smooth(method = "lm", se = FALSE, color = "blue") +  
26   labs(  
27     title = "Pace vs Wins (2015–2025)",  
28     x = "Pace",  
29     y = "Wins"  
30   ) +  
31   theme_minimal(base_size = 14)
```

Pace:

$r = -.1133$
**Weak, Negative
Relationship**

Linear Regression Model

Model 1			
	Coef.	p-Value	Std. Error
(Intercept)	-29.878	0.266	26.829
Age	3.538 ^{**}	< 0.001	0.340
ThreePointAtt.	24.172 [*]	0.020	10.331
Pace	-0.326	0.199	0.253
Num.Obs.	330		
R2	0.288		

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Key Findings:

- Age has significant impact on team wins
- 3-Point Attempts have a significant impact on team wins

Model is not well fit → What would fit better?

Moving Forward

- Evaluate statistical significance of each factor coefficient
- Apply our predictive model determining team success during the regular season
- Possibly evaluate other variables we have available in the data set
- Refine visualizations

Thank You!