

Does age matter in the NBA?

Suren S. Bhakta

Introduction

- **Motivation:** Growing up, basketball has always been my favorite sport. HOWEVER, this project stems larger than just basketball. I wanted to see if there is a true peak in athletic performance, specifically ages 26-30 in the NBA.
- Question:** Is there a difference in Points per game in the NBA by age groups **19-25, 26-30, and 30+**.
- **Data collection:** I collected my data by importing a spreadsheet of all players points per game from the season 20-21 into excel. Following gaining my random sample, I imported the scoring outputs of each and every player and averaged it out for whatever intervals they played over.
 - After shortening my list to fit the criteria, I used a random number generator to "pick" a player from the adjusted sample
 - My sample size was 75.
 - Outliers I removed were players that only played in one of the age groups
 - <https://www.basketball-reference.com/>

Method

Method: For my research, I felt that a “single factor Anova test assuming unequal variances” to determine whether there is a difference or not. I felt this to be the most justifiable method due to the presence of one numerical variable, points per game (PPG), being manipulated by three different categorical variables, age.

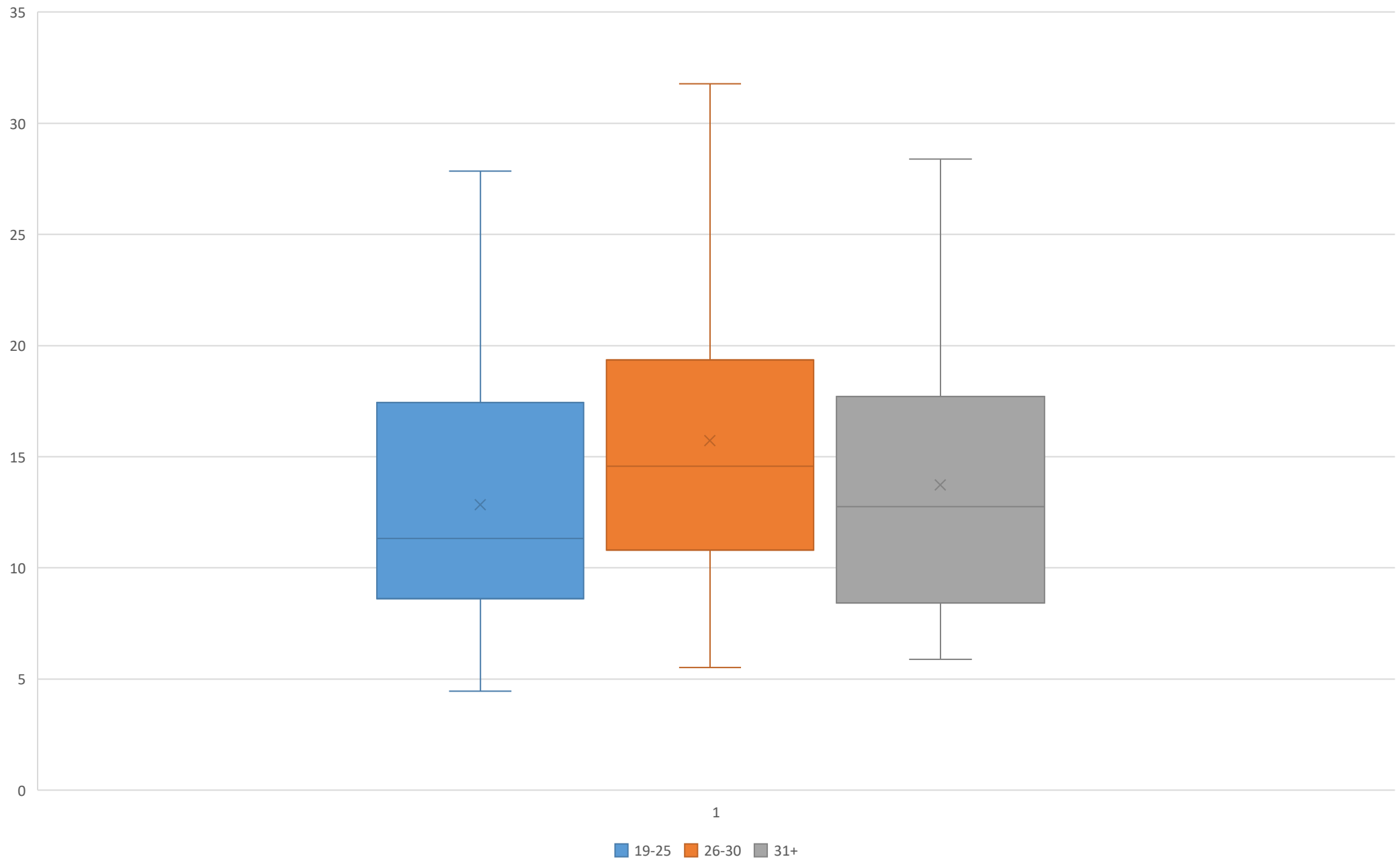
Assumptions/Conditions: For a single factor Anova test, the assumptions I checked were Random sample, Independent samples (groups), Normal distribution or large sample (25+) and homoscedasticity.

The checks for each assumption goes as follows:

- Random Sample: see **introduction** slide
- Independent groups: no players were or groups were repeated
- Normal Distribution: see counts for each grouping (**slide 5**)
- Homoscedasticity: Ensuring variances were not 4 times larger than another group (**slide 5**)

Group Comparison

Points Per Game (PPG) by Age Range



Results

	19-25	26-30	31+
Mean	12.84	15.72	13.73
Variance	31.2	38.194	36.45
Minimum	4.45	5.52	5.89
Maximum	27.84	31.78	28.4
Count	74	75	67

ANOVA F-Table						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	323.19	2	161.60	4.58	0.011	3.04
Within Groups	7509.75	213	35.26			
Total	7832.94	215				

Post Hoc Test Results

With a f-stat of 4.58 which was more extreme than our f-crit of 3.04, post-hoc testing was necessary. Post hoc test was done with an adjusted alpha of $\alpha = 0.017$

	19-25	26-30
df	146	
t Stat	-2.9837891	
P(T<=t) two-tail	0.00333827	
t Critical two-tail	1.97634565	

	19-25	31+
df	135	
t Stat	-0.9048693	
P(T<=t) two-tail	0.36714663	
t Critical two-tail	1.97769228	

	26-30	31+
df	139	
t Stat	1.93864624	
P(T<=t) two-tail	0.05457084	
t Critical two-tail	1.97717772	

Conclusion

H₀: $\mu_{19-25} = \mu_{26-30} = \mu_{31+}$

H_a: One of the groups that differs by age does not equal the other in average points per game

Following the check of the assumptions that the groups were randomly sampled (**slide 2**), along with independent and normal sampling (**slide 5**), and homoscedasticity (**slide 5**), we were able to do a single factor anova test and concluded that there is a difference in average points per game over the three different age groups. We can conclude this because our f-stat of 4.58 was more extreme than our f-critical of 3.04. This is reassured since our p-value of 0.011 is less than our alpha of 0.05.

With our f-stat being more extreme than our f-critical, at least one of the groups differed. As a result, post-hoc testing was performed with an adjusted alpha value of 0.017 with three comparisons. After conducting post hoc testing, it was found that the **19-25** and **26-30** age group differed significantly ($t(74)=-2.98$, $p=0.0033$). However, neither the **19-25/31+** comparison, ($t(67)=-0.90$, $p=.37$), nor the **26-30/31+** comparison ($t(67)=1.94$, $p=0.05$) significantly differed in average points per game.

Limitations

One limitation that comes to mind comes from the 31+ age group. Some of the players who qualified for the 31+ age group may have not played long enough to signify a drop-off in points per game average.

Another limitation that comes to mind is injuries during one's career. Injuries can cause a number of different abnormalities to points per game. For some players, they are never able to get back to who they once were, especially if it were an injury in their early twenties. However, for others an injury gives them time to recover and prologues their career and efficiency.

Implications/Future Directions

Answer both of these questions:

- 1) An implication from this study is that there is a peak in someone's career in the NBA. Although this could be related to peak body performance, but there are other factors that can go into someone's points per game output such as experience or change of team
- 2) If I were to conduct a study under the same premise, I would conduct it under the same sample. However, I would put an emphasis on external factors that would impact an individuals point per game and make sure to research each player to see if there were other factors besides age that would impact the scoring output.