Econometrics Final Project Report
Analyzing the Impact of Age and Performance Metrics on NBA Player

Introduction

My econometrics project was on analyzing how the Age of an NBA player affects their performance. My main goal was to see what the relationship was between a player's age and how well they played, particularly focusing on how many minutes they played and how many points they scored. I was doing this in the hopes of discovering the prime age of an NBA player.

Data Collection and Preparation

I started by Gathering data on all NBA players who participated in this current years regular season. I did not care if they played 82 games, or 5 minutes in one game, because I was unsure of where to cut it off. The dataset included variables such as Age, Points, MinutesPlayed, Field Goals Attempted (FGA), Three Pointers Attempted, Free Throws Attempted, Games Played, Assists, Rebounds, and Turnovers. After organizing the data in Excel, I imported it into Stata for analysis.

Model 1: Explaining Scoring (Points)

The first regression model used Points as the dependent variable. Independent variables included:

- Age
- AgeSquared (to capture non-linear effects and potential peak performance age)
- MinutesPlayed
- FGA, ThreePointersAttempted, FreeThrowsAttempted
- GamesPlayed

This model returned an R-squared of 0.9843, indicating that over 98% of the variation in points scored is explained by the model.

Key findings:

- Age had a positive and significant effect (coefficient = 0.326, $p < 0.001$), but AgeSquared was negative and significant (coefficient = -0.0058, $p < 0.001$), this confirmed a concave shape of how age would affect scoring.

- MinutesPlayed and GamesPlayed were significant, but interestingly, MinutesPlayed had a negative coefficient, likely due to multicollinearity or interaction effects.

- FGA and FreeThrowsAttempted had strong positive impacts, which is what I expected.

- ThreePointersAttempted had a negative coefficient, which suggests that players who rely too much on three's may have lower total point outputs, this is probably due to inefficiency or just because of different roles on the NBA court.

Model 2: Explaining Minutes Played
The second regression model focused on MinutesPlayed as the dependent variable, using the following independent variables:

- Age, AgeSquared

- Points, Assists, TotalRebounds

- FGA, FG%, FreeThrowsAttempted, Turnovers

- GamesPlayed

This model had an R-squared of 0.969.

Key findings:

- Age showed a positive and significant effect, while AgeSquared was marginally insignificant ($p = 0.054$), still suggesting diminishing returns as players get older.

- Points, Rebounds, and FreeThrowsAttempted all positively and significantly affected minutes played.

- Turnovers were positively associated with minutes (coefficient = 1.165), which possibly indicates that players who play more, naturally have more turnovers.

Conclusion

My regressions and analysis show that there certainly is a prime age rage on NBA players. After using derivatives to find the age where players score the most and when they play the most, I came to the conclusion that the Prime NBA age is around 31.25 years old.
The relationship between age and points scored confirms that players seem to peak in scoring at around age 30.75, and tend to receive the most playing time closer to age 31.75. This is likely due to the fact that NBA players aren't quite as good when they start due to inexperience, and eventually have a falloff due to athletic ability.

Regressions

```
. regress Points Age AgeSquared MinutesPlayed FGA ThreePointersAttempted GamesPlayed FreeThrowsAttempted
```

Source	SS	df	MS	Number of obs	=	644
				F(7, 636)	=	5776.95
Model	28862.2741	7	4123.18201	Prob > F	=	0.0000
Residual	453.932418	636	.713730218	R-squared	=	0.9845
				Adj R-squared	=	0.9843
Total	29316.2065	643	45.5928562	Root MSE	=	.84483

Points	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Age	.3260149	.0842363	3.87	0.000	.1606	.4914299
AgeSquared	-.0053847	.0015159	-3.55	0.000	-.0083615	-.0024078
MinutesPlayed	-.0031315	.0010875	-2.88	0.004	-.005267	-.0009959
FGA	1.115242	.0228239	48.86	0.000	1.070423	1.160061
ThreePointersAttempted	-.0621026	.028785	-2.16	0.031	-.1186278	-.0055774
GamesPlayed	.0102981	.001651	6.24	0.000	.007056	.0135402
FreeThrowsAttempted	.7318009	.0467028	15.67	0.000	.6400905	.8235112
_cons	-5.136055	1.148105	-4.47	0.000	-7.39059	-2.881521

```
. reg MinutesPlayed Age AgeSquared GamesPlayed Points Assists TotalRebounds FG FGA FreeThrowsAttempted Turnovers
```

Source	SS	df	MS	Number of obs	=	644
				F(10, 633)	=	2016.08
Model	618966.164	10	61896.6164	Prob > F	=	0.0000
Residual	19434.0283	633	30.7014665	R-squared	=	0.9696
				Adj R-squared	=	0.9691
Total	638400.192	643	992.846333	Root MSE	=	5.5409

MinutesPlayed	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Age	1.238433	.5603846	2.21	0.027	.1379952	2.338871
AgeSquared	-.0194588	.0100599	-1.93	0.054	-.0392136	.000296
GamesPlayed	.0912573	.0110946	8.23	0.000	.0694707	.1130439
Points	.1901188	.2611742	0.73	0.467	-.3227538	.7029914
Assists	-.0941398	.1150782	-0.82	0.414	-.320121	.1318413
TotalRebounds	-.1438011	.0417133	-3.45	0.001	-.2257143	-.0618879
FG	.4309503	.0125572	34.32	0.000	.4062915	.4556092
FGA	1.412669	.2972739	4.75	0.000	.8289071	1.996432
FreeThrowsAttempted	-.7701677	.3383728	-2.28	0.023	-1.434637	-.1056988
Turnovers	.165818	.0537843	3.08	0.002	.0602007	.2714353
_cons	-14.38516	7.67075	-1.88	0.061	-29.44836	.6780329