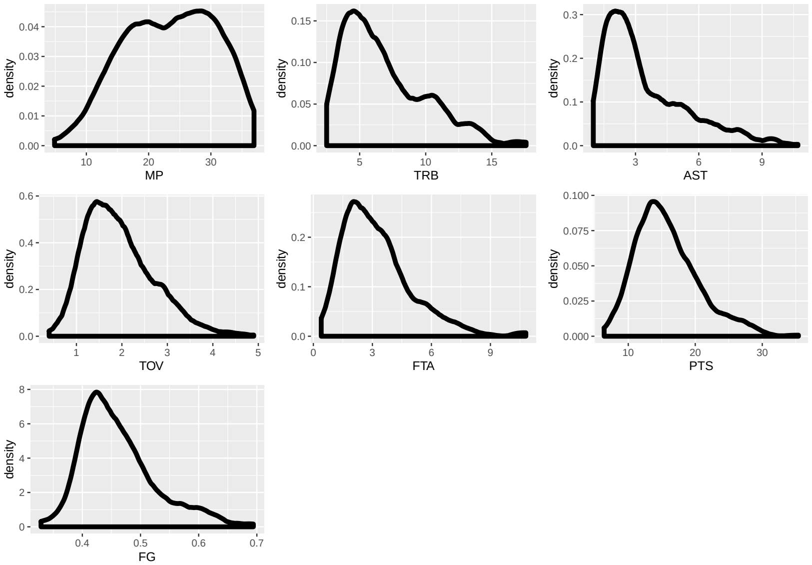
Predicting Minutes Played In The NBA

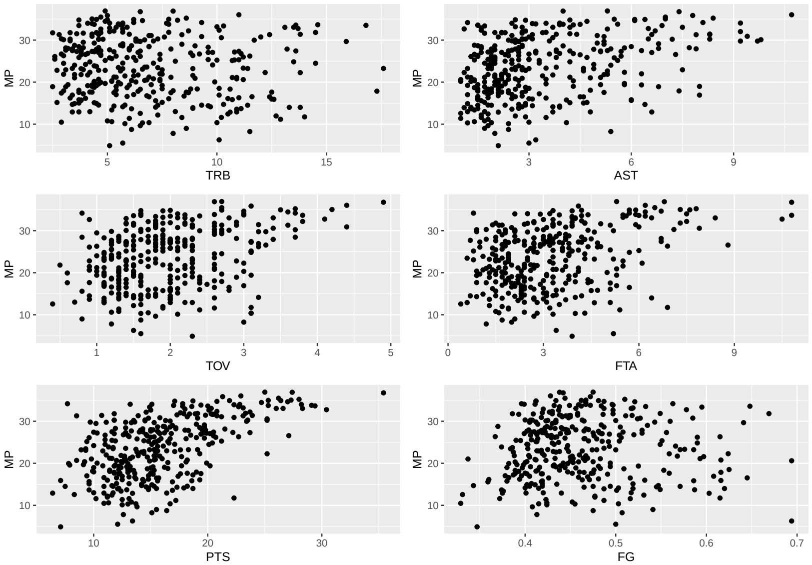
Background & Motivations

- We hope to predict minutes played through other types of basketball statistics.
- Minutes played is valuable to keep track of because more minutes played will offer a player more opportunities to contribute to the game.
- We looked at statistics from the 2018-2019 NBA season for players who played in at least 41 games, or half the season.
- We converted counting stats (eg rebounds) into per 36 minute rates. These rates can provide a good insight into a player's productivity without interference from the number of minutes played by the player.

Kernel Density Estimation



Scatterplots

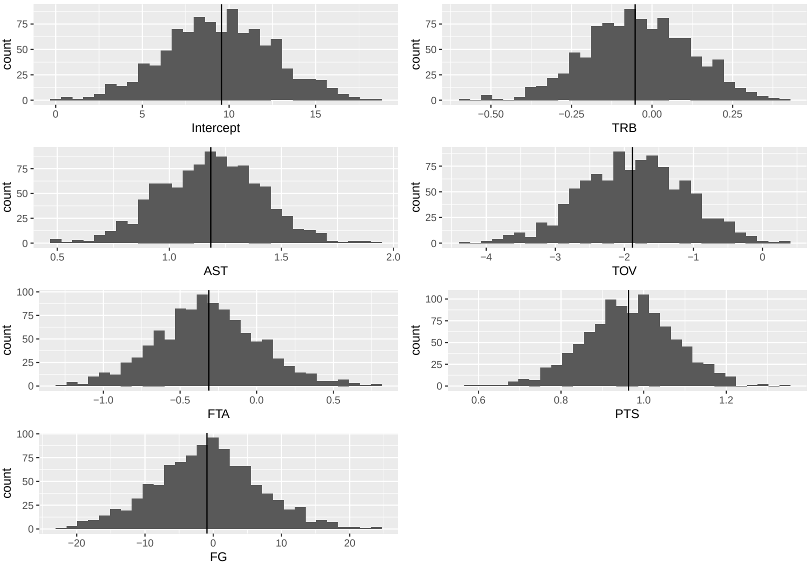


OLS Multiple Linear Regression

	Predictor	P-Value
(Intercept)	9.5726	0.0006
TRB	-0.0525	0.7443
AST	1.1850	0.0000
TOV	-1.8836	0.0152
FTA	-0.3123	0.3291
PTS	0.9633	0.0000
FG	-0.9236	0.8925

$MP = 9.573 - 0.052\ TRB + 1.185\ AST - 1.884\ TOV - 0.312\ FTA + 0.963\ PTS - 0.924\ FG$

Bootstrap



OLS Multiple Linear Regression

	Predictor	P-Value
(Intercept)	9.3395	0.0000
AST	1.2937	0.0000
TOV	-2.3198	0.0006
PTS	0.8941	0.0000

$$MP = 9.3395 + 1.2937\ \text{AST} - 2.3198\ \text{TOV} + 0.8941\ \text{PTS}$$

JHM Multiple Regression

	Predictor	P-Value
(Intercept)	9.2943	0.0014
TRB	-0.0898	0.5897
AST	1.1762	0.0000
TOV	-1.7645	0.0280
FTA	-0.3654	0.2701
PTS	0.9872	0.0000
FG	0.3410	0.9616

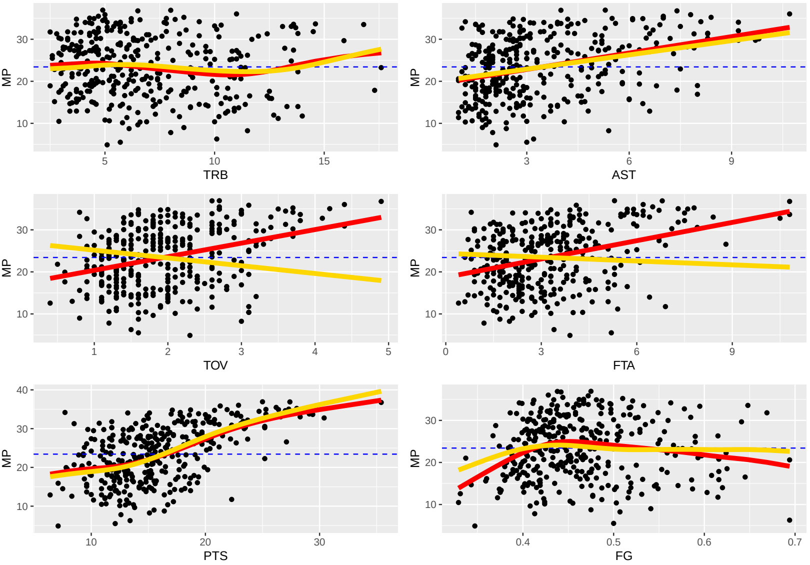
$$MP = 9.294 - 0.090\ TRB + 1.176\ AST - 1.764\ TOV - 0.365\ FTA + 0.987\ PTS + 0.341\ FG$$

JHM Multiple Regression

	Predictor	P-Value
(Intercept)	9.4239	0.0000
AST	1.2996	0.0000
TOV	-2.2720	0.0011
PTS	0.9083	0.0000

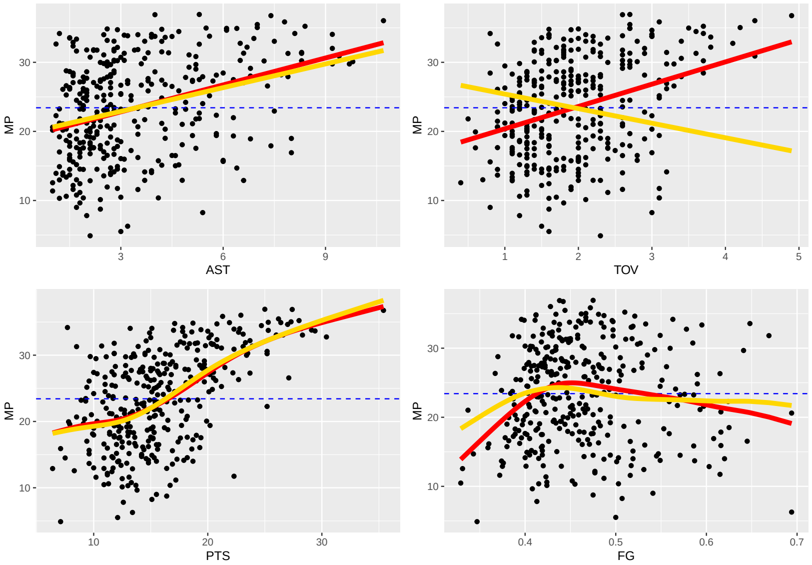
$$MP = 9.4239 + 1.2996 \text{ AST} - 2.272 \text{ TOV} + 0.9083 \text{ PTS}$$

Generalized Additive Model



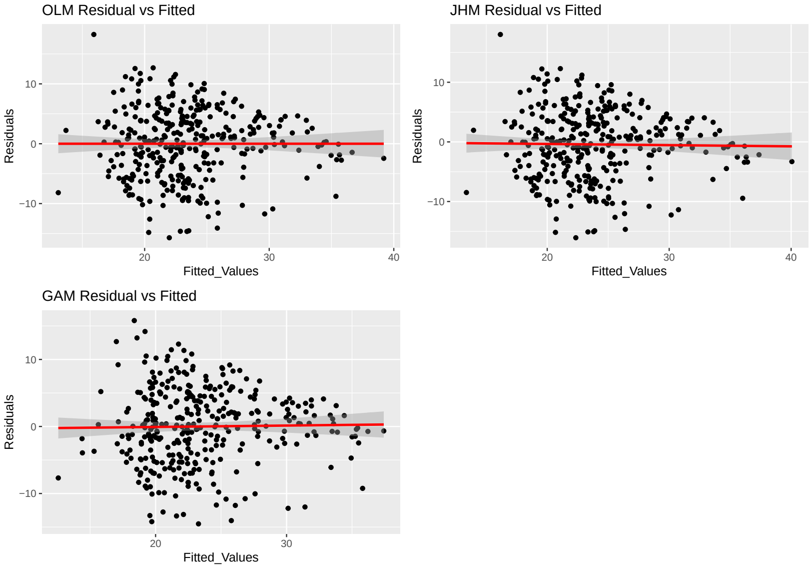
AIC: 2221.5334353

Generalized Additive Model



AIC: 2219.0000079

Residual Plots



Kolmogorov Smirnov Test

OLS	0.6180
JHM	0.5157
GAM	0.5157

Not enough evidence to say the residual distributions stray from Normal.

Model Fit CV (Both Models)

	cv.rsq	cv.adjrsq	cv.propL1
OLS	0.3365	0.3248	0.2344
JHM	0.3346	0.3229	0.2371
GAM	0.3422	0.3126	0.2481

	cv.rsq	cv.adjrsq	cv.propL1
OLS	0.3421	0.3364	0.2356
JHM	0.3381	0.3323	0.2356
GAM	0.3447	0.3253	0.2512

Model Fit and CV (Reduced Models)

	rsq	adjrsq	propL1
OLS	0.3586	0.3530	0.2446
JHM	0.3550	0.3494	0.2473
GAM	0.3961	0.3782	0.2787

	cv.rsq	cv.adjrsq	cv.propL1
OLS	0.3447	0.3390	0.2357
JHM	0.3401	0.3344	0.2360
GAM	0.3491	0.3299	0.2532

Takeaways

- Raw box score numbers do not do a good job of predicting the number of minutes played
- The models tested seem to have similar values for R^2, $AdjR^2$, and $L1_{prop}$.
- Future studies could examine different models for different positions, the usage of other stats, or different filtering conditions