



Gaining Competitive Advantages in Fantasy Basketball

Jackson Au

Fantasy basketball is popular!

Fantasy sports is very popular due to its competitive nature and financial implications. Over the years, the community has only grown larger!



US-only, 18+ year old fantasy sports players

Year	Estimated Number of Players
2015	42 million
2016	42.8 million
2017	43.2 million
2019	45.9 million

PROBLEM STATEMENT

With the growing number of experienced players in fantasy basketball, gaining *any* slight advantage can help you win!

**Using historical data,
what competitive
advantages can we gain?**



Data utilized

- BasketballMonster Rankings (2015–2019)
- Historical NBA standings (2015-2019)
- Personal league data (2015-2019)



Exploratory Data Analysis

Player Name	Frequency
Josh Richardson	  
Kemba Walker	 
Marc Gasol	
Elfrid Payton	 
Goran Dragic	
Khris Middleton	 
Kyle Lowry	
Pau Gasol	

Common players on winning teams

	Rank	Y!Adp9	Value	g	m/g	p/g	3/g	r/g	a/g	s/g	b/g	fg%	fga/g	ft%	fta/g	to/g	
Rank	1	0.652382	-0.92498	-0.288287	-0.615155	-0.616684	-0.207787	-0.47587	-0.423108	-0.410678	-0.343628	-0.2155	-0.542053	-0.138555	-0.541224	-0.47787	
Y!Adp9	0.652382		1	-0.669149	0.191192	-0.729264	-0.748996	-0.291386	-0.476231	-0.511689	-0.399838	-0.263706	-0.0368222	-0.726969	-0.12397	-0.657027	-0.679002
Value	-0.92498	-0.669149		1	0.243861	0.607395	0.70728	0.289536	0.492005	0.435737	0.442399	0.358492	0.18776	0.61991	0.172572	0.630066	0.542566
g	-0.288287	0.191192	0.243861		1	-0.0917975	-0.123766	-0.017184	-0.117908	-0.00325432	-0.117359	-0.128608	0.00845838	-0.121598	-0.0617627	-0.114431	-0.0735887
m/g	-0.615155	-0.729264	0.607395	-0.0917975		1	0.76852	0.526195	0.245271	0.543135	0.482157	-0.0294193	-0.264772	0.796072	0.273205	0.591305	0.649606
p/g	-0.616684	-0.748996	0.70728	-0.123766	0.76852		1	0.508955	0.531437	0.2974	0.0594642	-0.0393243	0.969776	0.304206	0.856839	0.785763	
3/g	-0.207787	-0.291386	0.289536	-0.017184	0.526195	0.508955		1	-0.345646	0.311919	0.223191	-0.408496	-0.6282	0.577063	0.589711	0.183238	0.29793
r/g	-0.47587	-0.476231	0.492005	-0.117908	0.245271	0.335851	-0.345646		1	-0.00846589	0.0678945	0.630475	0.582531	0.220219	0.368466	0.49323	0.344193
a/g	-0.423108	-0.511689	0.435737	-0.00325432	0.543135	0.531437	0.311919	-0.00846589		1	0.522509	-0.244262	-0.236486	0.560942	0.198461	0.486596	0.779733
s/g	-0.410678	-0.399838	0.442399	-0.117359	0.482157	0.2974	0.223191	0.0678945	0.522509		1	-0.0273747	-0.255512	0.332954	-0.0179267	0.28108	0.397669
b/g	-0.343628	-0.263706	0.358492	-0.128608	-0.0294193	0.0594642	-0.408496	0.630475	-0.244262	-0.0273747		1	0.57204	-0.0515457	-0.42024	0.226994	0.0116334
fg%	-0.2155	-0.0368222	0.18776	0.00845838	-0.264772	-0.0393243	-0.6282	0.562531	-0.236486	-0.255512	0.57204		1	-0.208007	-0.49434	0.142618	-0.052303
fga/g	-0.542053	-0.726969	0.61991	-0.121598	0.796072	0.969776	0.577063	0.220219	0.560942	0.332954	-0.0515457	-0.208007		1	0.337578	0.755748	0.769834
ft%	-0.138555	-0.12397	0.172572	-0.0617627	0.273205	0.304206	0.589711	-0.368466	0.198461	-0.0179267	-0.42024	-0.49434	0.337578		1	0.0674559	0.0820268
fta/g	-0.541224	-0.657027	0.630066	-0.114431	0.591305	0.856839	0.183238	0.49323	0.486596	0.28108	0.226994	0.142618	0.755748	0.0674559		1	0.777254
to/g	-0.47787	-0.679002	0.542566	-0.0735887	0.649606	0.785763	0.29793	0.344193	0.779733	0.397669	0.0116334	-0.052303	0.769834	0.0820268	0.777254		1

Correlation Matrix – All variables

Exploratory Data Analysis - Takeaways

Player Name	Frequency
Josh Richardson	  
Kemba Walker	
Marc Gasol	
Elfrid Payton	
Goran Dragic	
Khris Middleton	
Kyle Lowry	
Pau Gasol	

Common players on winning teams (out of 8 rosters)

Player Name	Avg cost (across 4 years)
Josh Richardson	12
Kemba Walker	29
Marc Gasol	25.8
Elfrid Payton	6
Goran Dragic	15.3
Khris Middleton	15.5
Kyle Lowry	32.5
Pau Gasol	14.3

None of these players are “superstars” and are inexpensive to draft: 6 – 32 dollar range

Exploratory Data Analysis - Takeaways

	Rank	Y!Adp9	Value	g	m/g	p/g	3/g	r/g	a/g	s/g	b/g	fg%	fga/g	ft%	fta/g	to/g
Rank	1	0.652382	-0.92498	-0.288287	-0.615155	-0.616684	-0.207787	-0.47587	-0.423108	-0.410678	-0.343628	-0.2155	-0.542053	-0.138555	-0.541224	-0.47787
Y!Adp9	0.652382	1	-0.669149	0.191192	-0.729264	-0.748996	-0.291386	-0.476231	-0.511689	-0.399838	-0.263706	-0.0368222	-0.726969	-0.12397	-0.657027	-0.679002
Value	-0.92498	-0.669149	1	0.243861	0.607395	0.70728	0.289536	0.492005	0.435737	0.442399	0.358492	0.18776	0.61991	0.172572	0.630066	0.542566
g	-0.288287	0.191192	0.243861	1	-0.0917975	-0.123766	-0.017184	-0.117908	-0.00325432	-0.117359	-0.128608	0.00845838	-0.121598	-0.0617627	-0.114431	-0.0735887
m/g	-0.615155	-0.729264	0.607395	-0.0917975	1	0.76852	0.526195	0.245271	0.543135	0.482157	-0.0294193	-0.264772	0.796072	0.273205	0.591305	0.649606
p/g	-0.616684	-0.748996	0.70728	-0.123766	0.76852	1	0.508955	0.335851	0.531437	0.2974	0.0594642	-0.0393243	0.969776	0.304206	0.856839	0.785763
3/g	-0.207787	-0.291386	0.289536	-0.017184	0.526195	0.508955	1	-0.345646	0.311919	0.223191	-0.408496	-0.6282	0.577063	0.589711	0.183238	0.29793
r/g	-0.47587	-0.476231	0.492005	-0.117908	0.245271	0.335851	-0.345646	1	-0.00846589	0.0678945	0.630475	0.562531	0.220219	-0.368466	0.49323	0.344193
a/g	-0.423108	-0.511689	0.435737	-0.00325432	0.543135	0.531437	0.311919	-0.00846589	1	0.522509	-0.244262	-0.236486	0.560942	0.198461	0.485696	0.779733
s/g	-0.410678	-0.399838	0.442399	-0.117359	0.482157	0.2974	0.223191	0.0678945	0.522509	1	-0.0273747	-0.255512	0.332954	-0.0179267	0.28108	0.397669
b/g	-0.343628	-0.263706	0.358492	-0.128608	-0.0294193	0.0594642	-0.408496	0.630475	-0.244262	-0.0273747	1	0.57204	-0.0515457	-0.42024	0.226994	0.0116334
fg%	-0.2155	-0.0368222	0.18776	0.00845838	-0.264772	-0.0393243	-0.6282	0.562531	-0.236486	-0.255512	0.57204	1	-0.208007	-0.49434	0.142618	-0.052303
fga/g	-0.542053	-0.726969	0.61991	-0.121598	0.796072	0.969776	0.577063	0.220219	0.560942	0.332954	-0.0515457	-0.208007	1	0.337578	0.755748	0.769834
ft%	-0.138555	-0.12397	0.172572	-0.0617627	0.273205	0.304206	0.589711	-0.368466	0.198461	-0.0179267	-0.42024	-0.49434	0.337578	1	0.0674559	0.0820268
fta/g	-0.541224	-0.657027	0.630066	-0.114431	0.591305	0.856839	0.183238	0.49323	0.485696	0.28108	0.226994	0.142618	0.755748	0.0674559	1	0.777254
to/g	-0.47787	-0.679002	0.542566	-0.0735887	0.649606	0.785763	0.29793	0.344193	0.779733	0.397669	0.0116334	-0.052303	0.769834	0.0820268	0.777254	1

ADP and Value seems highly correlated with RANK (our target), so we may want to avoid using those when modeling.

Can we utilize data to help
us draft higher-ranked
players, *consistently*?

Model [baseline]

Our baseline model was a Logistic Regression implementation with some validation techniques:

MODEL	ACCURACY
Linear Regression w/ train-test split	50.0%
Linear Regression w/ k-fold	40.23%

Our baseline model is below average

Model [improved!]

In this model, I find additional data that can help our predictability. As shown below, our accuracy for prediction goes up!

MODEL	ACCURACY
Random Forest w/ train-test split	64.25%

Best Model [model tuning + feature engineering]

After finding model tuning and feature engineering, I created a much better model than our baseline

MODEL	ACCURACY
Random Forest with feature engineering + model tuning	82.06%

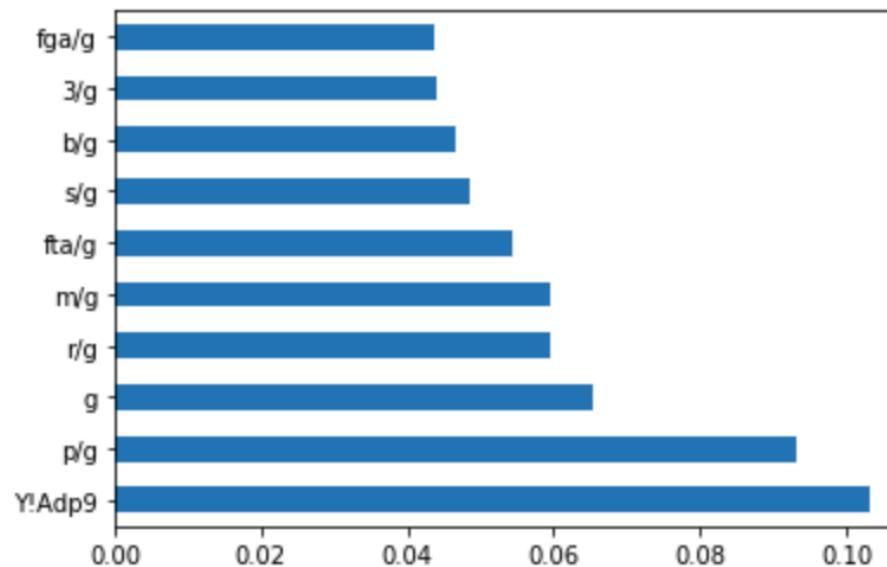
*With model tuning and feature engineering, I
can make better use of my data.*

Model Comparison

MODEL	ACCURACY
Linear Regression (Baseline)	50%
Random Forest (more data)	64.25%
Random Forest (model tuning)	82.06%

Diving Deep into our model

Using Recursive Feature Elimination, we determine which *attributes* were most important in our final model:



FINAL CONCLUSIONS

Let's sum up the matter with some final conclusions about the presentation.

Auction Drafts

Going for a stars-and-scrubs approach will have less room for error. Aim for players in the \$6-30 range.

Important Features

Assists did not land in the top-10 most important features. Value can be found outside of this often overhyped category.

Potential breakouts

Look for players that may be given the opportunity to play more. Minutes Per Game were a key indicator of a high ranked player.

NBA parity

In the past, it was easy to avoid resting players. However, with the NBA approaching parity, there is evidence of high-ranked players on playoff teams.

Thank you!

