Making Subjective NBA Awards Objective:
Predicting NBA End-of-Season Awards with Statistics

University of California, Berkeley
IEOR 142 - Introduction to Machine Learning and Data Analytics
Nilay Agarwalla, Raj Dasani, Bikram Khaira, Kenny Tran, Sarah Truong

Introduction + Motivation

The National Basketball Association (NBA) presents 12 annual awards to recognize the accomplishments of its teams, players, and coaches. Of those 12 awards, 6 individual awards are given to NBA basketball players based on their performance for the regular season. These awards are Most Valuable Player, Most Improved Player, Rookie of the Year, Defensive Player of the Year, and Sixth Man of the Year. The winner of each NBA award is selected by a panel of one hundred independent media members, where each panelist casts a vote for first to third place selections, with the exception of the Most Valuable Player Award where the voting panel members can cast a vote for first to fifth place selections. These awards are decided subjectively, with no clear procedure on how to weigh specific player statistics in place. For instance, several of these individual awards have been given to an individual player due to the team's overall season success, some have been given due to a player's best plays and statistics of the season, and some have been given due to a player's historical feat and ability to overcome strenuous challenges throughout the regular season. Thus, each panelist holds their own biases towards players from season to season, rarely taking an impartial approach with their voting.

The current, unstructured methods for deciding winners of NBA awards are not a consistent proxy for award distribution. Therefore, the aim of our project is to create a transparent and more uniform way to determine the end-of-season NBA awards using statistical models that are void of any individual biases. Using datasets that detail the individual statistics of each NBA player that has received at least one vote for these 6 individual awards, we attempt to answer the following questions: (1) What player statistics hold the most weight when our statistical models are predicting winners of each award between the years 2010 and 2020?, (2) Are voting members quantitatively consistent with how they determine their votes?, and (3) Can we build a predictive model to forecast future winners of these 6 individual NBA awards?

Data

We compiled and queried through datasets provided by public sources: basketball-reference.com and the official NBA.com website. With the help of Python packages and libraries, including Pandas, we cleaned and filtered the datasets and joined them to view each player's individual voting share for the 6 individual NBA awards with their basic and advanced player statistics from the years 2010 to 2020, to use as our predictors (X variables). In order to quantify player improvement for the Most Improved Player (MIP) for our models, we subtracted the current year's advanced and basic statistics from the previous year's advanced and basic statistics. This means that the data for Most Improved Player does not include statistics for the year 2009, and therefore the prediction of MIP for 2010 is void. Additionally, we applied the specific thresholds that the NBA has for each award, such as filtering to only include players who have played over 20 games for each season.

The way that NBA Awards Voting works is that a player receives points based on the number of votes they receive for an award. As mentioned earlier, each panelist casts a first, second, and third-place vote for Rookie of the Year, Sixth Man of the Year, Most Improved Player, and Defensive Player of the Year; a first-place vote is worth five points, a second-place vote is worth three points, and a third-place vote is worth one point. The Most Valuable Player Award, on the other hand, has each panelist cast a vote for first to fifth place selections; a first-place vote is worth ten points, a second-place vote is seven, a third-place vote is five, a fourth-place vote is three, and a fifth-place vote is one. Clearly, the number of votes cast for each award changes from year to year. For instance, a Defensive Player of the Year winner receiving 325 points in one year could mean that the same number of points they received the previous year was not enough for them to win the award. Therefore, we decided to collect voting shares that the players have received for the six awards. Voting shares are defined as the number of votes a player receives for an award divided by the maximum number of votes possible. These are our Y values.

Methods + Model Technique

With the data from 2010-2020, we decided to do a split of 8/3 years, with the training set being 2010 to 2017 and the test set to cover the years 2018 to 2020. This split will allow our models to analyze how consistent the predictions of the awards are across the decade.

To determine the success of the four models, we chose 3 key statistics to analyze each model. The first was R^2, the standard model stat. This was the main stat used to determine the best model within each type of model. After doing some initial analysis, it was noticed that R^2 had some key issues, most notably that it weighed eligible NBA players for each award the same. While this provided a holistic model, accounting for each player, it did not represent how well each model predicted the top players for the awards, and most importantly the winner.

With these concerns, a statistic that was focused more on what NBA fans most cared about, the top players for each award. This led to the creation of our own statistic, the Top 10 Mean Absolute Error of Predictions. It would be calculated by taking the top 10 players who received the highest voting share for each award and calculating the mean absolute error:

$$\sum_{i=1}^{10} \left| votingshare_{predicted_i} - votingshare_{actual_i} \right|$$

The third statistic made, and the most interpretable was simply a calculation of how many of the years the model predicted correctly.

The four models we chose for modeling were Linear Regression, Lasso Regression, Ridge Regression, and Random Forest. Linear Regression was expected to be the most interpretable model, Ridge and Lasso Regression was expected to be our most accurate model, and Random Forest was ideally supposed to be the middle between accuracy and interpretability. Linear Regression and Random Forest will also be paired to do Feature Engineering, in order to find the most important statistics for each award.

Linear Regression

The first model we decided to examine was Linear Regression, a linear approach to modeling the relationship between the basic and advanced NBA stats (listed in the Appendix) and the voting share of a player for an award. Via VIF (Variable Inflation Factor) (at a cutoff of 5) and then p-values (at a cutoff of 0.5), we then conducted Feature Engineering: VIF in order to help us quantify the severity of the multicollinearity that could potentially exist in our regression analysis and p-values to indicate the significance of the individual statistic to the voting share/model.

For example, our first award was MVP. Using the OLS feature of the StatsModels API, we were able to create a model training the decided independent variables, determined from implementing VIF, with the MVP voting share and then fitting the model. We also had to add a constant in order to ensure an intercept in the model. From there, we were able to view a summary of results which include R^2 values, coefficients, and p-values of the chosen stats and the intercept. The results for the MVP model and all the other awards models are below and in the appendix:

Lasso Regression and Ridge Regression

After the results of our linear regression model, we found two problems with the data. First, we realized that there were too many features, most of which had little to no effect on our response variable. Second, we realized that many of our variables exhibited multicollinearity. To address the first issue, a LASSO regression model was used as it performs variable selection by turning the coefficients of non-significant variables to zero, effectively removing them from the model. In addition, we also ran a

Ridge regression model to deal with the second issue as ridge regression deals with multicollinearity by increasing bias and reducing variance.

We standardized the independent variables and included interaction terms because we believed that the effects of certain variables may be dependent on others and we didn't want LASSO to remove them. For example, the variable Turnovers (TOV) may not be significant on its own but the interaction between Turnovers (TOV) and Assists (AST) hold significant value because players that average high amounts of turnovers and assists are players that control the basketball more, which would be influential on our response variable. We included interaction terms for the Ridge model as well. We then used cross-validation to find the best lambda for each model and trained a model with these lambda values.

Random Forest

Random forest regression calculates the average of all the predictions made by multiple but different regression decision trees, and because it decorrelates individual trees, the ensemble of decision trees brings an element of randomness, providing a nonlinear nature to our models that other linear algorithms inherently lack and preventing the issue of overfitting.

To set up the process of random forests, we did a grid search k-fold cross-validation with R^2 scoring over the list player advanced and basic statistics that we collected. Specifically, to set up our grid values, we tested odd features from 1 to 46 (46 is the number of statistics/columns that we have for each player in our dataset), and we controlled our minimum leaf samples to 5, set our estimators to 500 and splits to 5. Afterward, we specifically looked at the features that held the most weight in our models when predicting awards for each season. For example, our Most Valuable Player award has 2 dominant features/player statistics that were used to predict its voting share, and they were Value over Replacement Player (VORP) and Win Shares (WS).

Most Important Variables Found From Each Model

NBA Award	Random Forest	Linear Regression	Ridge Regression	Lasso Regression
Most Valuable Player	VORP, WS	TS, X2PP, STL, X3PP, OWS	OWS*VORP, USG*VORP, TOVP*OBPM, TOVP*DBPM	BPM*VORP, DBPM*VORP, WS*VORP, OWS*VORP, TOV*VORP, TOV*BPM
Sixth Man of the Year	PTS, MP, FGM, FGA	TS, X2PP, X3PAR, X3PM, STL	FTP, AST, FGA*USG,FGP*FTR, X3PM*STL,X3PM*AST, X2PA*FTM	FGM*VORP, FGM*DWS, X3PM*FTM, X2PM*DWS, ORB*AST, ORB*X3PAR, STLP*TOVP
Defensive Player of the Year	DWS, DBPM, BLK	BLK, FTR, X3PAR	BPM*VORP, WS48*VORP, BLKP*BPM, STLP*WS48	DWS*DBPM, DWS*VORP, BLK*DBPM, BLK*DWS, STL*BLK,
Most Improved Player	PTS, FTM, FGM, X2PM, FTA	ORB, BLK, X3PM	WS*BPM, DWS*OBPM, DWS*WS, OWS*WS48, BLK*WS	PTS*OWS, PTS*USG, TRB*OWS,ORB*PTS, FTM*BLK, FTM*ORB
Rookie of the Year	PTS, FTM, FGA, FGM	TS, X2PP, FTP	FGM, FGA, DBPM,G*STL, GS*ORB, FGA*FTP, FGP*VORP	PTS*VORP, BLK*VORP,AST*VORP, AST*PTS, FTA*BPM, FTM*AST

Summary of Results + Discussion

Across our models, our R^2s had their fluctuations, and when looking at the Average Top 10 Player Errors, the best R^2 did not usually have the best average top 10 mean error. So, below is our best model for each award, prioritizing average top 10 player errors, and using OSR^2 and "Number of Accurate Predicted Winners," simply how many times the model predicted the winner correctly, as tiebreakers.

Best Model for Each Award

NBA Award	Best Model	Average Top 10 Player Errors	Number of Accurate Predicted Winners	OSR^2	2021 Predicted Winner
MVP	Random Forest	0.10	7/11	0.65	Nikola Jokic
ROY	Lasso Regression	0.10	10/11	0.80	LaMelo Ball
MIP	Random Forest	0.12	7/10	0.13	Nikola Jokic
6MOY	Random Forest	0.13	8/11	0.47	Jordan Clarkson
DPOY	Random Forest	0.11	7/11	0.37	Rudy Gobert

As seen above, our best model via average top 10 player errors was Random Forest, which was able to accurately predict 7 or 8 or the 11 years for each award. In terms of OSR^2 however, our lasso and ridge regressions, with their interaction terms, we're able to predict the more granular part of the data, and accurately provide 0s to more players who were not considered at all for the award.

The biggest takeaway we came across in our models was that most award predictions are dominated by just 2-6 stats, and they were usually advanced (Win Shares, Defensive Rating, etc). Placing these high weights on these key statistics seems to translate pretty closely to the stories or stereotypes of what the media depicts for each award. For example, many argue and believe that the MVP should be 'the best player on the best team,' and history has shown that. And in all of our models, Win Shares is a key predictor.

As a whole, there was some consistency across the types of stats each award model was predicting, and it yet again connects very much to the qualitative aspects of the awards. For MVP and DPOY, many of the most important stats, via feature engineering, seemed to be team-based stats, such as Win Shares and Defensive Win Shares, whereas for ROY, 6MOY, and MIP, basic yet direct impact stats such as Points, Free Throws Made, etc, were the key stats in each model.

In the appendix, you can see each model's predictions for each year!

Conclusion

Overall, our models found a lot of the key consistency across the awards but still were not able to necessarily capture the most qualitative part of the voting process, such as storyline or what many people in the media call the "eye test". A key example of this is the 2011 year where all our models predicted LeBron James to win the award, but Derrick Rose caught the heart of many fans across the NBA world as a young player who was on pace to be the next Bulls great after the dominance of Michael Jordan.

So while that is in fact the biggest question of our models and will probably never make it perfect, that is in fact the beauty of our models, where they will take each year equally, and produce as consistent results as possible. Whether that is right or not, is for the NBA world to decide.

Impact

Sports betting is a growing industry that has brought fans from all over the world across all professional sports together to make their best predictions on the outcome of games and awards. For those who want to anticipate winners and beat the odds, using statistical models can provide bettors with a more accurate set of results in contrast to their personal predicted winners based on their own sentiment. This also can pave the path for sportsbooks to display the statistical odds of game wins and award shares that are more accurate.

Certainly, the current voting system for NBA awards unveils an implicit bias currently being held by media members. Therefore, transforming NBA awards from being subjective to being objective can birth a new methodology that might be able to give these awards to more deserving NBA players. Determining winners of these 6 awards would be consistent for each season since they'd be purely numerical-based.

Finally, adopting statistical models to predict NBA awards provides opportunities to not only educate budding fans of professional basketball on the historical trends of each individual award but also pave a path for fans to help decide on how awards should be given. Fans can have a bigger role in the NBA awards process and alter the way they watch and analyze basketball games if the awards are decided objectively.

References

Datasets:

All data for 2020 (For all datasets, change the year in the URL to get different years):

1. Player Data:https://www.basketball-reference.com/leagues/NBA_2020_per_game.html
2. Voting Shares: https://www.basketball-reference.com/awards/awards_2020.html

Article Inspiration:

1. https://towardsdatascience.com/predicting-2020-21-nbas-most-valuable-player-using-machine-learning-24aaa869a740
2. https://towardsdatascience.com/using-data-science-to-predict-the-next-nba-mvp-30526e0443da
3. https://dribbleanalytics.blog/2019/04/ml-mvp-all-nba-predict-2019/

Appendix

Model's Success per Award

MVP:

Model / Stat	Predicted Winner? (/11)	Avg Top 10 Error	OSR^2
Linear Reg	6/11	0.22	0.29
Random Forest	7/11	0.10	0.65
Lasso Reg	5/11	0.15	0.61
Ridge Reg	7/11	0.13	0.67

ROY:

Model / Stat	Predicted Winner? (/11)	Avg Top 10 Error	OSR^2
Linear Reg	6/11	0.26	0.42
Random Forest	9/11	0.13	0.67
Lasso Reg	10/11	0.10	0.80
Ridge Reg	10/11	0.11	0.75

MIP:

Model / Stat	Predicted Winner? (/11)	Avg Top 10 Error	OSR^2
Linear Reg	1/10	0.14	0.035
Random Forest	7/10	0.12	0.13
Lasso Reg	2/10	0.15	0.07
Ridge Reg	0/10 :(	0.15	-0.02

6MOY:

Model / Stat	Predicted Winner? (/11)	Avg Top 10 Error	OSR^2
Linear Reg	5/11	0.17	0.10
Random Forest	8/11	0.13	0.47
Lasso Reg	7/11	0.14	0.52
Ridge Reg	8/11	0.14	0.54

DPOY:

Model / Stat	Predicted Winner? (/11)	Avg Top 10 Error	OSR^2
Linear Reg	3/11	0.16	0.12
Random Forest	7/11	0.11	0.37
Lasso Reg	6/11	0.12	0.35
Ridge Reg	6/11	0.13	0.15

Our Models' Assignment of each Award Winner:

Color Code:
Green: 3/4 or 4/4 models predicted correctly
Yellow: 2/4 or 1/4 models predicted correctly
Red: 0/4 models predicted correctly

MVP

	Linear Reg	**Random Forest**	**Lasso Reg**	**Ridge Reg**	**Actual Winner**
2010	LeBron James	LeBron James	LeBron James	LeBron James	LeBron James
2011	LeBron James	LeBron James	LeBron James	LeBron James	Derrick Rose
2012	LeBron James	LeBron James	LeBron James	LeBron James	LeBron James
2013	LeBron James	LeBron James	LeBron James	LeBron James	LeBron James
2014	Kevin Durant	Kevin Durant	Kevin Durant	Kevin Durant	Kevin Durant
2015	Kawhi Leonard	Stephen Curry	James Harden	Stephen Curry	Stephen Curry
2016	Stephen Curry	Stephen Curry	Stephen Curry	Stephen Curry	Stephen Curry
2017	James Harden	Russell Westbrook	Russell Westbrook	Russell Westbrook	Russell Westbrook
2018	James Harden	LeBron James	LeBron James	James Harden	James Harden
2019	James Harden	James Harden	James Harden	James Harden	Giannis Antetokounmpo
2020	James Harden	James Harden	James Harden	James Harden	Giannis Antetokounmpo
2021	Nikola Jokic	Nikola Jokic	Nikola Jokic	Nikola Jokic	TBD

ROY

	Linear Reg	**Random Forest**	**Lasso Reg**	**Ridge Reg**	**Actual Winner**
2010	Brandon	Tyreke Evans	Tyreke Evans	Tyreke Evans	Tyreke Evans

	Jennings				
2011	John Wall	Blake Griffin	Blake Griffin	Blake Griffin	Blake Griffin
2012	Kyrie Irving	Kyrie Irving	Kyrie Irving	Kyrie Irving	Kyrie Irving
2013	Damian Lillard	Damian Lillard	Damian Lillard	Damian Lillard	Damian Lillard
2014	Michael Carter-Williams	Michael Carter-Williams	Michael Carter-Williams	Michael Carter-William s	Michael Carter-Williams
2015	Andrew Wiggins	Andrew Wiggins	Andrew Wiggins	Andrew Wiggins	Andrew Wiggins
2016	Devin Booker	Karl-Anthony Towns	Karl-Anthony Towns	Karl-Anthony Towns	Karl-Anthony Towns
2017	Marquese Chriss	Joel Embiid	Joel Embiid	Joel Embiid	Malcolm Brogdon
2018	Ben Simmons	Donovan Mitchell	Ben Simmons	Ben Simmons	Ben Simmons
2019	Trae Young	Luka Doncic	Luka Doncic	Luka Doncic	Luka Doncic
2020	Ja Morant	Ja Morant	Ja Morant	Ja Morant	Ja Morant
2021	LaMelo Ball	Anthony Edwards	Lamelo Ball	Lamelo Ball	**TBD**

MIP

	Linear Reg	**Random Forest**	**Lasso Reg**	**Ridge Reg**	**Actual Winner**
2010	--	--	--	--	--
2011	Derrick Rose	Kevin Love	Dorell Wright	Troy Murphy	Kevin Love
2012	Nikola Pekovic	Jeremy Lin	Jeremy Lin	Lama Odom	Ryan Anderson
2013	Larry Sanders	Larry Sanders	James Harden	Stephen Curry	Paul George
2014	Deandre Jordan	Goran Dragic	Tony Wroten	Markieff Morris	Goran Dragic
2015	Rudy Gobert	Jimmy Butler	Shabazz Muhammad	Kevin Love	Jimmy Butler
2016	Kyle Lowery	CJ McCollum	CJ McCollum	Kyle Lowry	CJ McCollum
2017	Giannis Antetokounmpo	Giannis Antetokounmpo	Tim Hardaway Jr.	James Johnson	Giannis Antetokounmpo
2018	Tyreke Evans	Victor Oladipo	Kris Dunn	Isaiah Thomas	Victor Oladipo
2019	Nikola Vucevic	Pascal Siakam	Pascal Siakam	Tyreke Evans	Pascal Siakam
2020	Devonte' Graham	Devonte' Graham	Devonte'	Trae Young	Brandon Ingram

			Graham		
2021	Michael Porter Jr.	Nikola Jokic	Jerami Grant	James Harden	TBD

6MOY

	Linear Reg	**Random Forest**	**Lasso Reg**	**Ridge Reg**	**Actual Winner**
2010	Jamal Crawford	Jamal Crawford	Jamal Crawford	Jamal Crawford	Jamal Crawford
2011	James Harden	Lamar Odom	Lamar Odom	Lamar Odom	Lamar Odom
2012	James Harden	James Harden	James Harden	James Harden	James Harden
2013	JR Smith	JR Smith	JR Smith	JR Smith	JR Smith
2014	Jamal Crawford	Jamal Crawford	Jamal Crawford	Jamal Crawford	Jamal Crawford
2015	Jamal Crawford	Jamal Crawford	Lou Williams	Jamal Crawford	Lou Williams
2016	Ryan Anderson	Ryan Anderson	Will Barton	Will Barton	Jamal Crawford
2017	Lou Williams	Lou Williams	Lou Williams	Lou Williams	Eric Gordon
2018	Lou Williams	Lou Williams	Lou Williams	Lou Williams	Lou Williams
2019	Terrence Ross	Lou Williams	Lou Williams	Lou Williams	Lou Williams
2020	Davis Bertans	Montrezl Harell	Lou Williams	Montrezl Harell	Montrezl Harell
2021	Jordan Clarkson	Jordan Clarkson	Jordan Clarkson	Jordan Clarkson	TBD

DPOY

	Linear Reg	**Random Forest**	**Lasso Reg**	**Ridge Reg**	**Actual Winner**
2010	Dwight Howard	Dwight Howard	Dwight Howard	Dwight Howard	Dwight Howard
2011	Dwight Howard	Dwight Howard	Dwight Howard	Dwight Howard	Dwight Howard
2012	Serge Ibaka	Dwight Howard	Dwight Howard	Dwight Howard	Tyson Chandler
2013	Tim Duncan	Marc Gasol	Marc Gasol	Marc Gasol	Marc Gasol
2014	DeAndre Jordan	Joakim Noah	Joakim Noah	Joakim Noah	Joakim Noah
2015	Anthony Davis	Kawhi Leonard	Kawhi Leonard	Kawhi Leonard	Kawhi Leonard

2016	Hassan Whiteside	Kawhi Leonard	Hassan Whiteside	Hassan Whiteside	Kawhi Leonard
2017	Rudy Gobert	Rudy Gobert	Rudy Gobert	Rudy Gobert	Draymond Green
2018	Anthony Davis	Andre Drummond	Anthony Davis	Anthony Davis	Rudy Gobert
2019	Rudy Gobert	Giannis Antetokounmpo	Giannis Antetokounmpo	Giannis Antetokounmpo	Rudy Gobert
2020	Anthony Davis	Giannis Antetokounmpo	Giannis Antetokounmpo	Giannis Antetokounmpo	Giannis Antetokounmpo
2021	Myles Turner	Rudy Gobert	Rudy Gobert	Rudy Gobert	TBD

Definition of Each Statistic and Feature:

All via Basketball-Reference:

AST - Assists

AST% - Assist Percentage (available since the 1964-65 season in the NBA); the formula is 100 * AST / (((MP / (Tm MP / 5)) * Tm FG) - FG). Assist percentage is an estimate of the percentage of teammate field goals a player assisted while he was on the floor.

Award Share - The formula is (award points) / (maximum number of award points). For example, in the 2002-03 MVP voting, Tim Duncan had 962 points out of a possible 1190. His MVP award share is 962 / 1190 = 0.81.

BLK - Blocks (available since the 1973-74 season in the NBA)

BLK% - Block Percentage (available since the 1973-74 season in the NBA); the formula is 100 * (BLK * (Tm MP / 5)) / (MP * (Opp FGA - Opp 3PA)). Block percentage is an estimate of the percentage of opponent two-point field goal attempts blocked by the player while he was on the floor.

BPM - Box Plus/Minus (available since the 1973-74 season in the NBA); a box score estimate of the points per 100 possessions that a player contributed above a league-average player, translated to an average team. Please see the article About Box Plus/Minus (BPM) for more information.

DPOY - Defensive Player of the Year

DRB - Defensive Rebounds (available since the 1973-74 season in the NBA)

DRB% - Defensive Rebound Percentage (available since the 1970-71 season in the NBA); the formula is 100 * (DRB * (Tm MP / 5)) / (MP * (Tm DRB + Opp ORB)). Defensive rebound percentage is an estimate of the percentage of available defensive rebounds a player grabbed while he was on the floor.

DRtg - Defensive Rating (available since the 1973-74 season in the NBA); for players and teams, it is points allowed per 100 possessions. This rating was developed by Dean Oliver, author of Basketball on Paper. Please see the article Calculating Individual Offensive and Defensive Ratings for more information.

DWS - Defensive Win Shares; please see the article Calculating Win Shares for more information.

eFG% - Effective Field Goal Percentage; the formula is (FG + 0.5 * 3P) / FGA. This statistic adjusts for the fact that a 3-point field goal is worth one more point than a 2-point field goal. For example, suppose Player A goes 4 for 10 with 2 threes, while Player B goes 5 for 10 with 0 threes. Each player would have 10 points from field goals, and thus would have the same effective field goal percentage (50%).

FG - Field Goals (includes both 2-point field goals and 3-point field goals)

FG% - Field Goal Percentage; the formula is FG / FGA.

FGA - Field Goal Attempts (includes both 2-point field goal attempts and 3-point field goal attempts)

FT - Free Throws
FT% - Free Throw Percentage; the formula is FT / FTA.
FTA - Free Throw Attempts
G - Games
GS - Games Started (available since the 1982 season
MVP - Most Valuable Player
MP - Minutes Played (available since the 1951-52 season)
MOV - Margin of Victory; the formula is PTS - Opp PTS.
ORtg - Offensive Rating (available since the 1977-78 season in the NBA); for players, it is points produced per 100 possessions, while for teams it is points scored per 100 possessions. This rating was developed by Dean Oliver, author of Basketball on Paper. Please see the article Calculating Individual Offensive and Defensive Ratings for more information.
Opp - Opponent
ORB - Offensive Rebounds (available since the 1973-74 season in the NBA
Player - Registered name of NBA Player
PProd - Points Produced; Dean Oliver's measure of offensive points produced. Please see the article Calculating Individual Offensive and Defensive Ratings for more information.
PTS - Points
ROY - Rookie of the Year
SMOY - Sixth Man of the Year
STL - Steals (available since the 1973-74 season in the NBA)
STL% - Steal Percentage (available since the 1973-74 season in the NBA); the formula is 100 * (STL * (Tm MP / 5)) / (MP * Opp Poss). Steal Percentage is an estimate of the percentage of opponent possessions that end with a steal by the player while he was on the floor.
TOV - Turnovers (available since the 1977-78 season in the NBA)
TOV% - Turnover Percentage (available since the 1977-78 season in the NBA); the formula is 100 * TOV / (FGA + 0.44 * FTA + TOV). Turnover percentage is an estimate of turnovers per 100 plays.
TRB - Total Rebounds (available since the 1950-51 season)
TRB% - Total Rebound Percentage (available since the 1970-71 season in the NBA); the formula is 100 * (TRB * (Tm MP / 5)) / (MP * (Tm TRB + Opp TRB)). Total rebound percentage is an estimate of the percentage of available rebounds a player grabbed while he was on the floor.
TS% - True Shooting Percentage; the formula is PTS / (2 * TSA). True shooting percentage is a measure of shooting efficiency that takes into account field goals, 3-point field goals, and free throws.
TSA - True Shooting Attempts; the formula is FGA + 0.44 * FTA
Usg% - Usage Percentage (available since the 1977-78 season in the NBA); the formula is 100 * ((FGA + 0.44 * FTA + TOV) * (Tm MP / 5)) / (MP * (Tm FGA + 0.44 * Tm FTA + Tm TOV)). Usage percentage is an estimate of the percentage of team plays used by a player while he was on the floor.
VORP - Value Over Replacement Player (available since the 1973-74 season in the NBA); a box score estimate of the points per 100 TEAM possessions that a player contributed above a replacement-level (-2.0) player, translated to an average team and prorated to an 82-game season. Multiply by 2.70 to convert to wins over replacement. Please see the article About Box Plus/Minus (BPM) for more information.
WS - Win Shares; an estimate of the number of wins contributed by a player. Please see the article Calculating Win Shares for more information.
WS/48 - Win Shares Per 48 Minutes (available since the 1951-52 season in the NBA); an estimate of the number of wins contributed by the player per 48 minutes (league average is approximately 0.100). Please see the article Calculating Win Shares for more information.