

Our dataset came from NBA.com. We decided to take the top 60 ranked players, which are ranked by the number of points they scored last season, and analyze them based on eight different variables. We included field goal percentage (FG%), which is the percentage of shots (both 2 and 3 points) a player made, percentage of three point shots made (%3P), percentage of free throws made (%FT), average number of rebounds in a game, average number of assists in a game, average number of steals in a game, and average number of points in a game.

Field goal percentage was measured by the NBA by counting the number of 2 or 3 point baskets made by each player, and dividing it by the total number that player shot. Three-point percentage measures the number of three point baskets made divided by the number attempted by each player. Percentage of free throws is measured by counting the number of free throw baskets made divided by the number attempted. The average number of rebounds, assists, steals, and points are each measured by counting the number that occur each game and finding the average for each over the season.

Each row represents a player and each observation in the row is based on our variables. We wanted our main focus of the project would be to determine which of these factors has the greatest effect on wins in a season, so we created a second file that included team names and the number of wins they had that season. We plan to use coding to merge the two datasets in order to come to our conclusions.