NBA TREND ANALYSIS

EAS 503: Final Project

Contributors:

Achintya Pillai

James Hsu

1 Abstract

In the world today over 40% of the global games and sport market is overrun by legal betting. This number is highly likely going to rise due to an increase in the coming years due to an increase in betting social media applications such as Bovado and Bestbet. In recent years there have been more and more mobile and web applications which are catering for betting. A huge percentage of the betting takes place in the sport industry and the top 3 sports that are popular for the betting industry are Soccer, tennis and the NBA.

Our trend analysis can be very helpful for anyone trying to get into the betting market and equip a highly ill informed better with the correct and accurate information he need to make future predictions about a game.

This idea can be implemented across different domains of sports such as Soccer, NFL, UFC etc. and it could help a lot of people make accurate and educated bets.

2 Introduction

This project is intended to show a clear trend, relation and statistics about the national basketball association, the teams and players who play for them and get recruited for the teams. We have used a common dataset for the Basket ball association in north American which hold all the information for the different American basket leagues and their players, recruits and statistic numbers. We have queried the entire data set to obtain the information for the NBA league and the players since the 1990’s.

Basket ball is a sport that can not be analysis only by only into one component, the game is so diverse we have to make an educated decision by looking into many different aspects of the sport. Looking into a sport analytical research paper written by Manuel Janeira from the university of Porto and Jaime Sampaio, which can be found in the References [1], we have come to a conclusion that be should looking into the dynamics of athleticism, in games statistics such as rebounds, possession rates, turn over rates and most importantly improvements over time as a basis for our analysis and predictions as they seem to be the most significant predictors for our analysis and predictions.

3 Data

We have used a common dataset for the Basket ball association in north American which hold all the information for the different American basket leagues and their players, recruits and statistic numbers. We have queried the entire data set to obtain the information for the NBA league and the players since the 1990’s.

The data set consists for information such as player ids, player names, point per minute, rebounds per game, etc. It is a common data base and easily accessible by the public. If used, queried and analyzed correctly we can understand the data better and even make future prediction with a fairly high accuracy.

We have gone about this project with a few questions we intended to answer with our analysis, which are listed below:

1. How did the minutes over games change over the years?

2. Has the rebound per minute gotten better over the years?

3. How much has the sport of basket ball changed over the years?

4. How has the sport changed and what direction is it heading in?

5. Do different states in the United states have a consistent recruitment pattern or is it completely random?

6. Who are the best players of all time? Can we model our predictions just based on their statistics?

4 Model

We decided to analyze our data at three levels: trend analysis, in game statistic and recruitment patterns. These different levels allowed us to look for trends in layers, and draw conclusions from there. The trend analysis was completed using linear multi-mapped figures which show a clear tread over time. The correct games statistics were queried from the table and and modeled with the use of scatter, density and histograms to give us a better understanding of the process.

5 Analysis

5.1 Trend analysis over years

Below we can see a series of graphs that show us information over the years. The aspects of the game that graphed below are also some of the most important aspects of the game and are good ways to see how these have changed over a period of time and what we can expect to see for the upcoming years in the games. The trend analysis has been looked at from 1940 to 2010. From a lot of the graphs below you will see that the initial years, are usually trailing the 0 values, this is simply because there was not much of data collected during that period of time. We see a drop in general trends from 1988 and 1999 seasons because there was a NBA lockout due to contractual issues which was between the players and the team owner which lead to fewer games in season

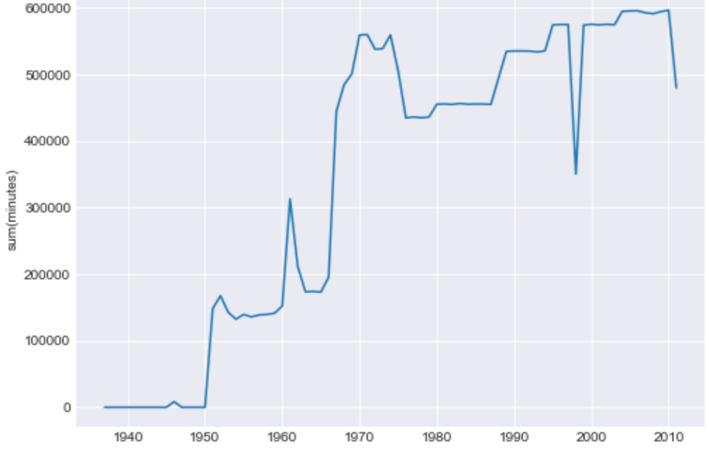
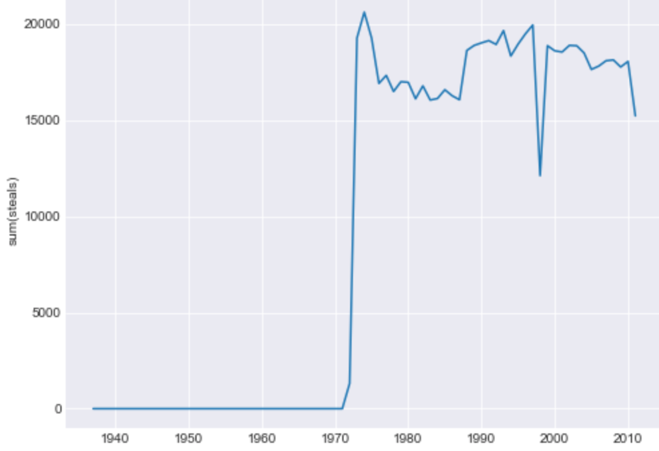


Figure 1: Steals vs Years Figure 2: Minutes vs Years

Above we see graphs plotted for two important game play statistics, Minutes over time and steals over time.

From the general trend of the graphs we can see that the steals have generally been leveled. This is very natural because steals are highly advanced skills and need more time to rapidly increase. We see a drop in 1988 and 1999 seasons in steals because there was a NBA lockout due to contractual issues which was between the players and the team owner which lead to fewer games in season. The minutes over years show a increase in the in games minutes per game played over years. This shows that basketball is becoming even more popular over the years and hence we can say more money can go into this game which will be a good investment.

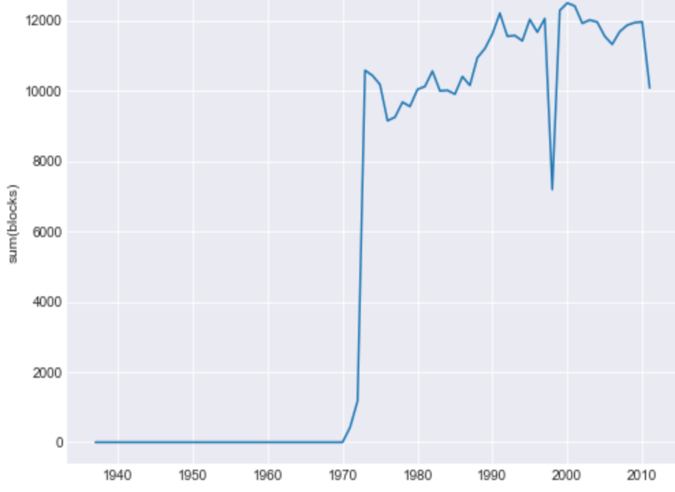
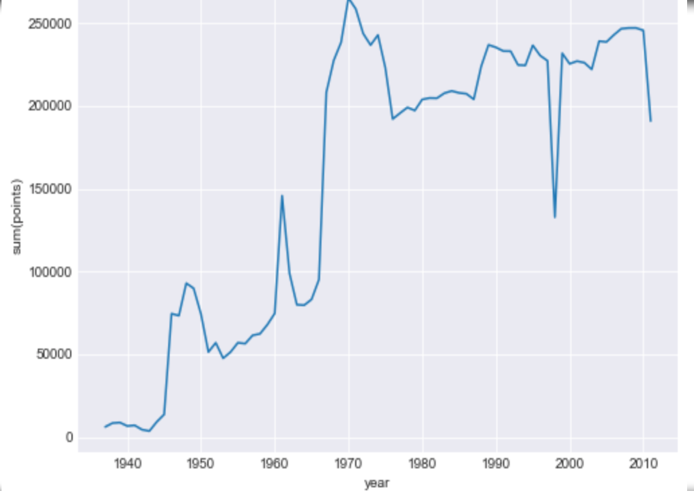


Figure 3: Points vs Years Figure 4: Blocks vs Years

From the graphs above we see a trend for points and blocks over the years. For points per game over the years we that there is a sharp increase in 1979 this seen mostly due to the invention of the 3 point line, which increased the total points accumulated as players could now score 3 point and 2 points in a games. Blocks per games is an important statistic because, the team that is better and has a higher block statistic would be reduces the other team’s ability to score.

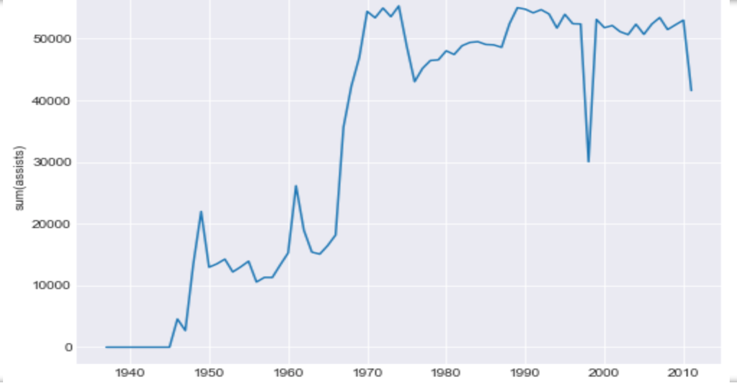
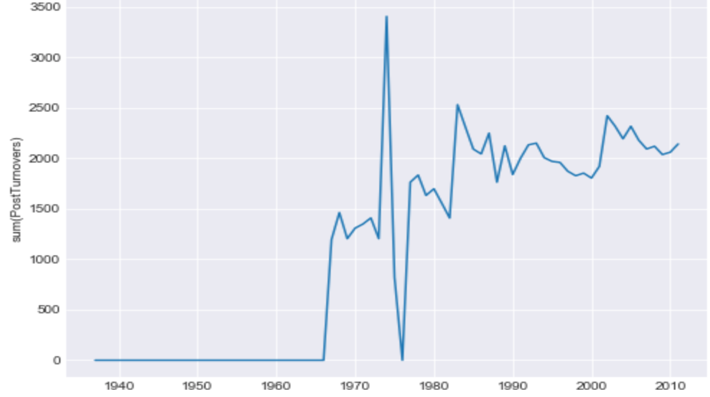


Figure 5: TurnOvers vs Years Figure 6: Assists vs Years

The Graphs above show us turnover and assists over the years. Turnovers is the rate at which the possession changes over the games and we see a general increase in that and assists is when you pass the ball at the last minute before you score to a team mate to increase the probability of the scoring.

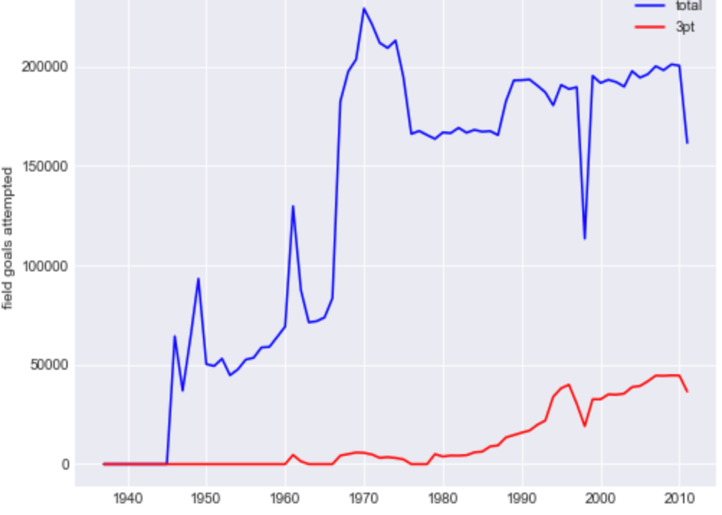
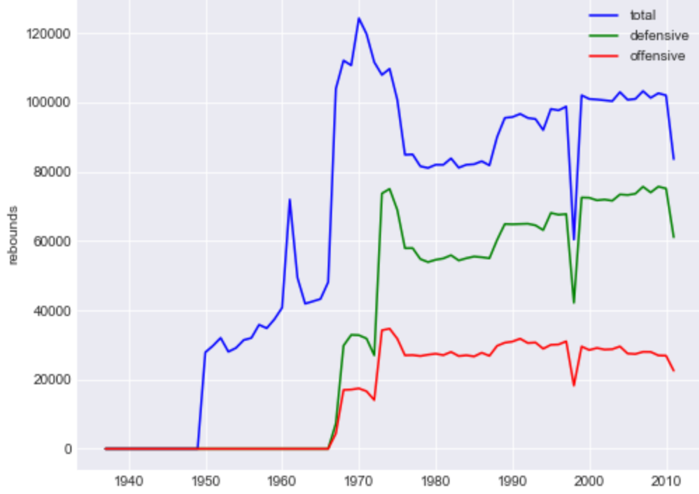
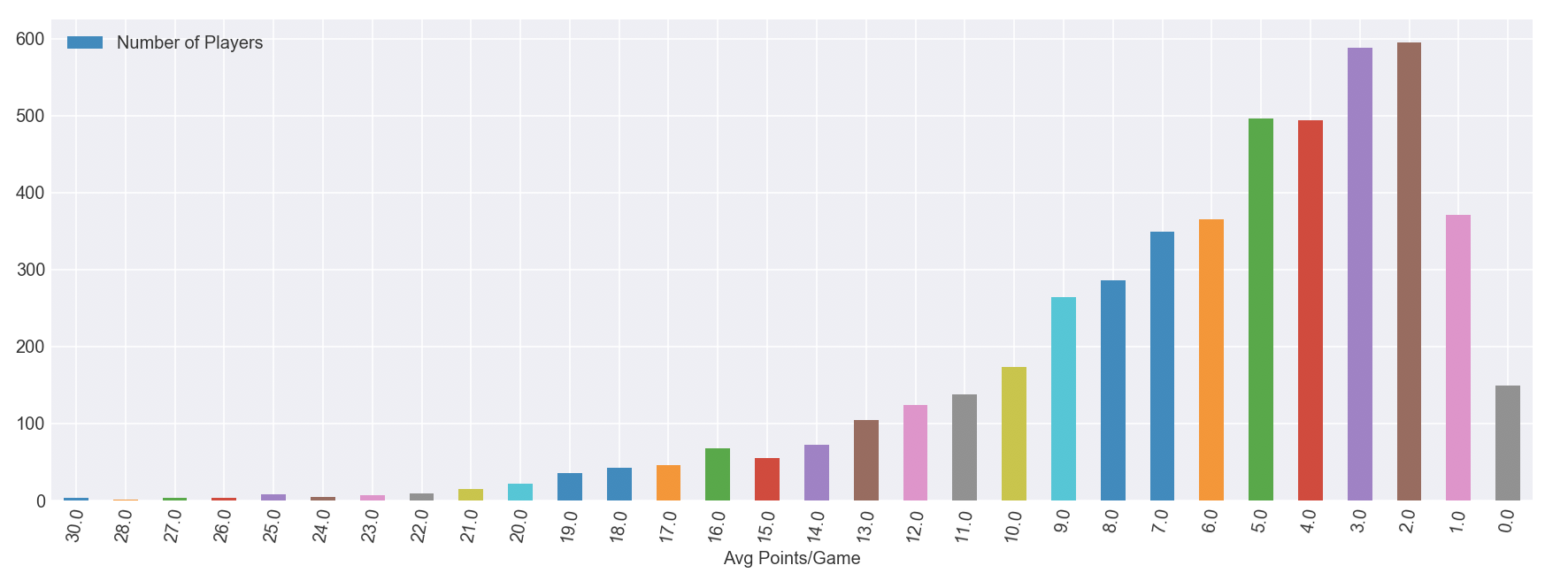


Figure 7: Rebounds vs time Figure 8: Field goal vs Time

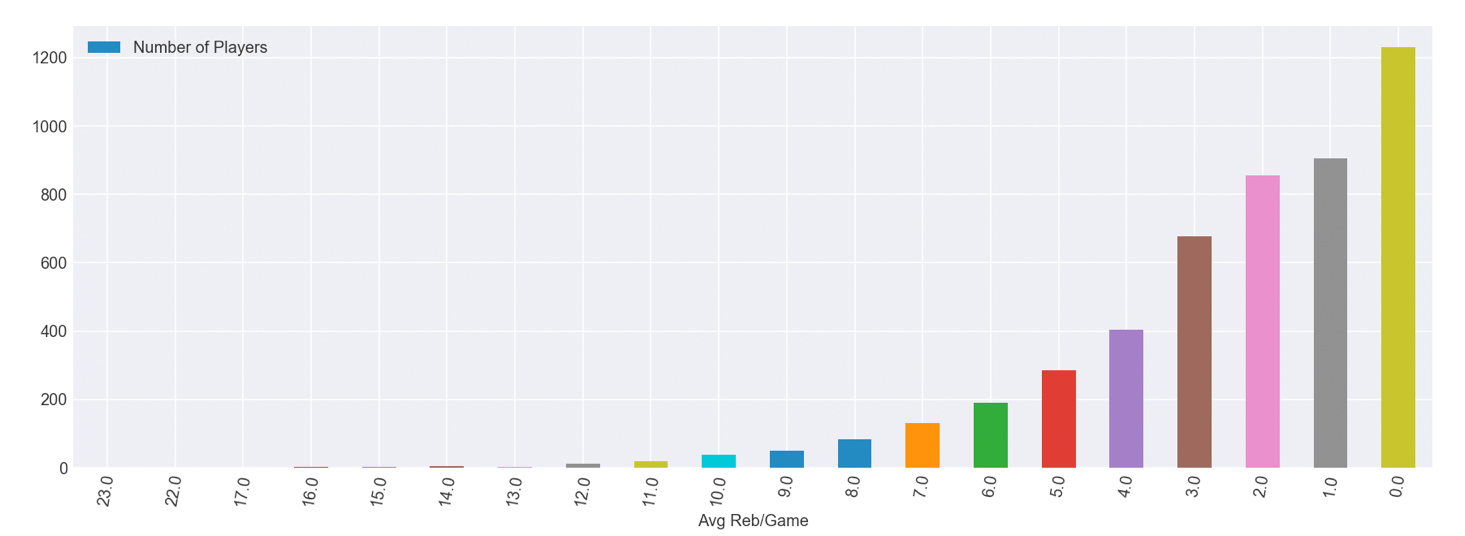
From the graphs above from figure 1 to 8 we see a general increase in the treads over the years. This is a very important statistic this shows us an increase in the pace of paly , this shows us that atheist have been getting faster and more skilled over the years. This is a important observation as this shows us that the the past statistics from the great players in a different era will not be a good predictor for the future players or bets because we can see from the graph that the pace of the game is always changing over time. Hence the statistics from different eras are always evolving. When betting if an individual uses the statistics of another good player to estimate the worth of his bet, this will not be a good idea as the era the other player is from will influence the decision.

5.2 Game play Analysis

At this level of analysis, we have looked into more details of the game that directly affect each game that is played. This will help in understanding specific information about the game that will help people bet for the correct numbers during the game.

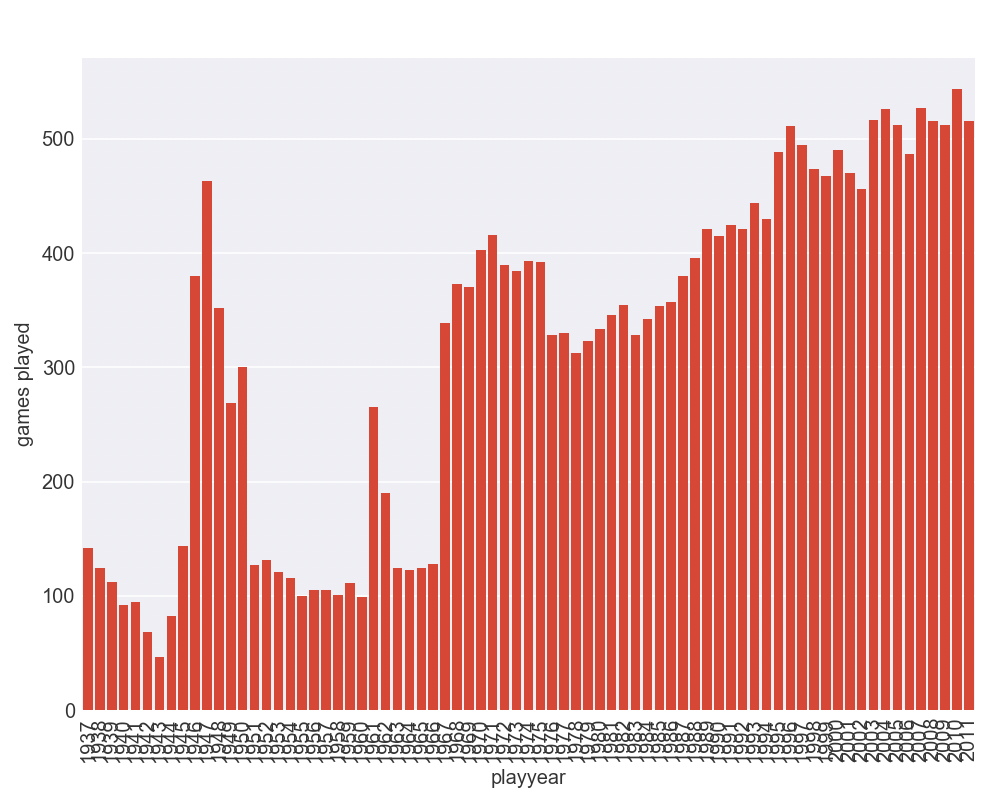


The graph highlights the average points of players per game. We can clearly see a trend in the graph, the number of average points and the number of players. The graph shows us that as the average points goes up the number of players actually goes down because only top tear athletics make an upwards of 12 points on average. The players with an average of 18, 19 and 20 points are considered to be very good and from the graph we can see that this is true because the number of players that make these top tear point are very few.



The graph above refers to the average rebound rates per game and how many players actually come under a certain bracket for rebound rates. Rebounds is actually a very important statistic to consider when deciding on a team because with a higher rebound rate this is always a chance for a team to gain more point per game as the game goes on. It is the frequency of ball captures after a attempt at scoring misses. Hence if a player has a better rebound rate he can counter the other team by going on the offensive and make a basket.

This graph above shows us that the number of players associated with a higher rebound average is actually very low and this is why we see the trend we see in the graph. Usually a good player has a average rebound rate of 11 or 10 per game. By looking at this statistic you should know when looking at statistic from a player or for a team if their rebound rates are exceptionally good, average or below average and make a better decision while choosing which team to bet on.



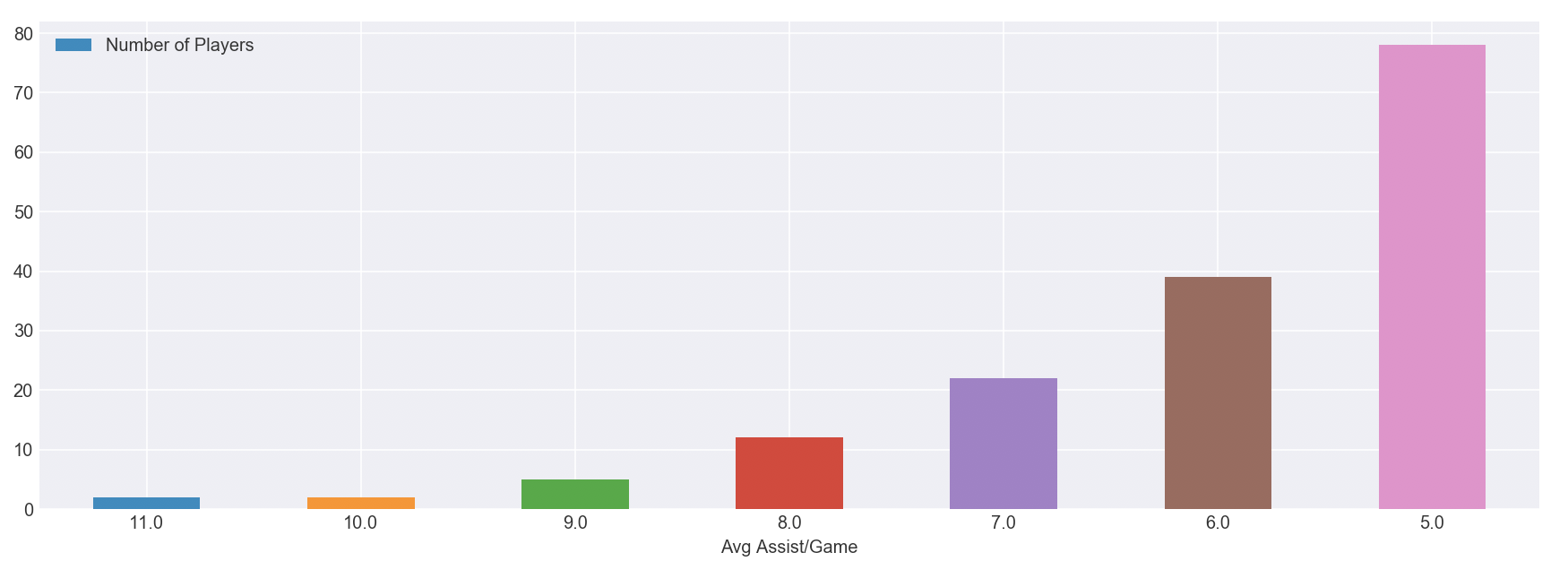
The graph above is a little deviation from the previous graphs we have been looking at as it shows us the total number of games played over the years. I have showed this just to emphasize on the fact that NBA is a good sport and is only going to get more popular over time as we can see a steady increase in games played over the years, excluding the dips we see in the graph which is caused by many factors due to new rules and lockouts over the years, but disregarding those aspects we have clearly seen and upward trend in the popularity of the

sport as we can see that the game plays have increased. This also tells us the rate at which players are getting better and the level that new recruits need to be at when they are signed.



The graph above shows a list of best players we have queried over the years. Can we look into the statistics of these players to make prediction of other players?

As I have mentioned above, we can not. This is because players such as Michel Jordan can not be compared to upcoming and currently well know players such as LeBron James, because they are from two completely different eras. The level of competition, averages and general rates and regulations have changed to a point where a comparison between Michael Jordan and LeBron James will not be reasonable for betting purposes. The players must be from the same age and era to make a reasonable decision on which team or player or statistic you choose to put your money on.



The graph above is a very underrated graph in our generation because, people think it does not affect the game to a point that will cause a win or loss. That might be true when we only look at the average assist of one

player, but when we look into the average assists of multiple players from one team we can dramatically increase our chances of winning.

We can see that the graph tells us that the higher the average assists are the lower the number of players that can actually achieve the average number.

5.3 Points per game analysis.

Figure (a): Log points over min Figure (b): Average Points

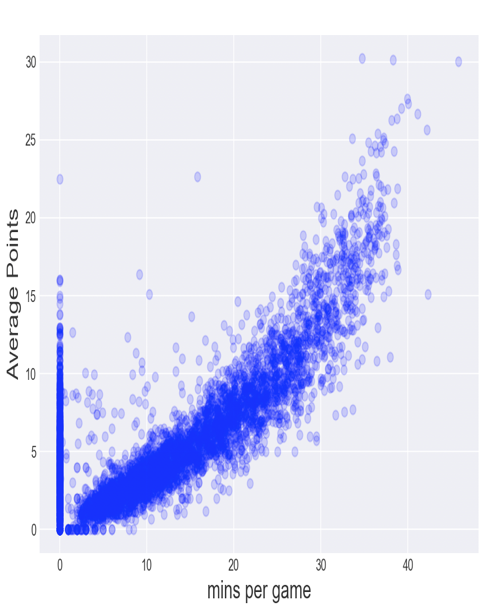
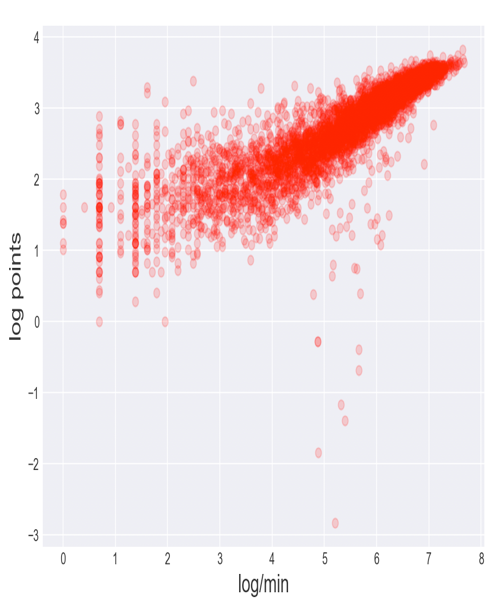
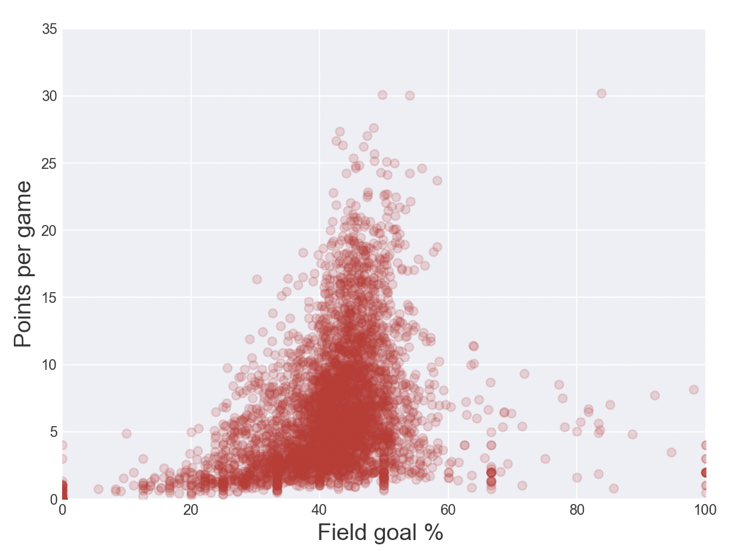


Figure (c): Field goals

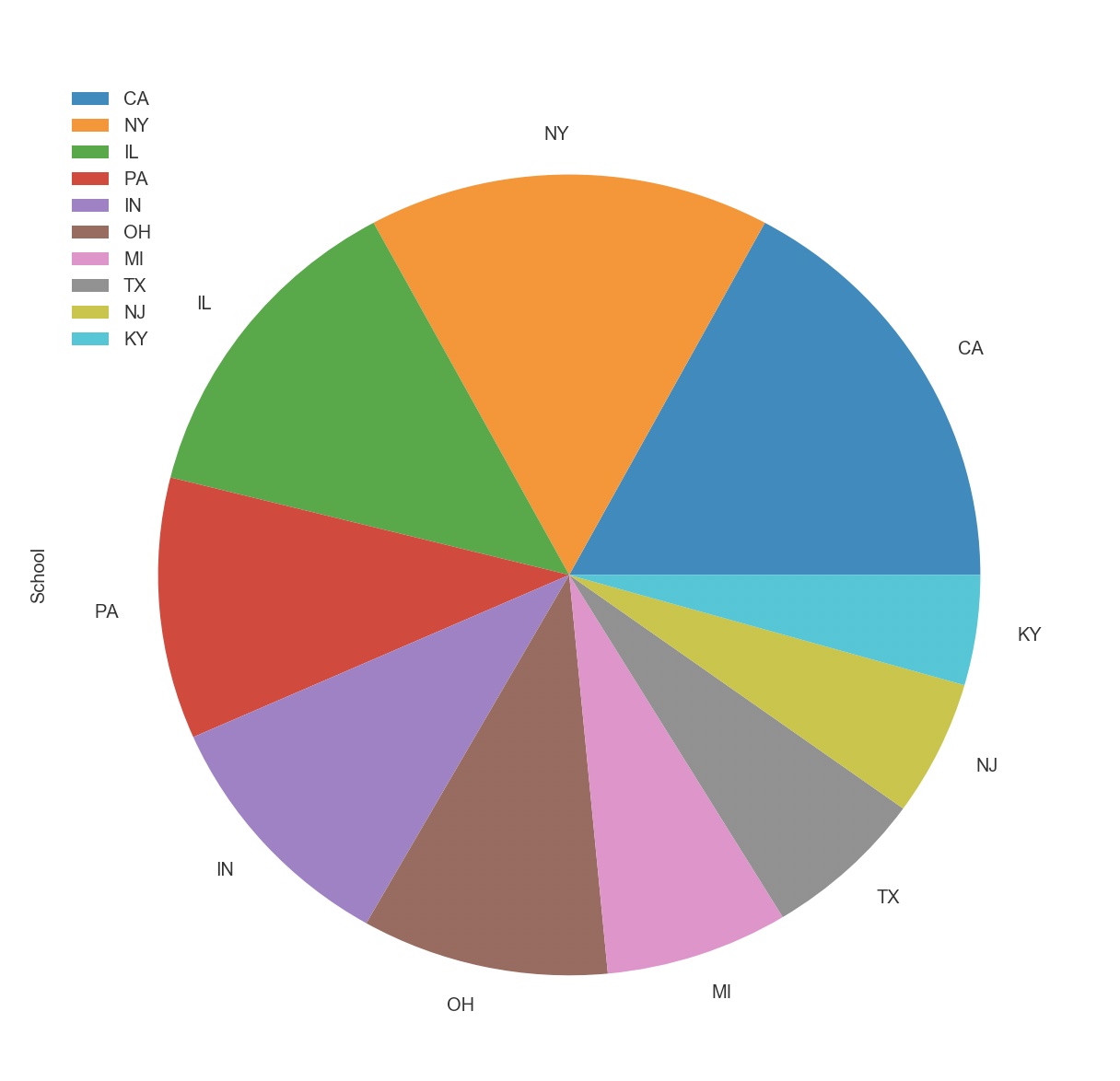


The three graphs above (a), (b) and (c) show us the point based trends we can use to interpret the game.

As you can see from graph (a) and (b) there are an exponential increase relationship between minutes per game and average points per game for the players. This relationship is really well presented in graph (b) where there are an increase curve. As the the minutes per game increases the average points per game increases. Also we can see from the graph we see the average points that are scored over a particular time interval in the game. We use an exponential graph to show this because this is a continuous process and can be measured over a time

interval. The results should be a good measure for us to judge what our team should be scoring at any particular minute in the game to be considered to be playing at a good level.

In graph (c) it shows the trend of current NBA scoring distribution. We can see this is a normal distribution. The reason is that players who tend to have higher Field goal percentages are center or power forward. So they do not tend to score a lot but since they specialize in under the basketball scoring they have a higher percentages of field goal. (Field Goal percentage= field goal made/field gold attempted) majority of player average about 35-45 field goal percentages. And very few players have averaged more than 25 points per game. Players like LeBron James, Kobe Bryant , Michael Jordan, Wilt Chamberlain have all more than 40 percent field goal percentage and averaged from 25 to 30 points per game and these players are considered the best basketball players that have ever lived.



**This graph shows us the top ten states where the NBA players are from. Can you can see from the graph majority of the players are from New York, California and Illinois. This graph can tell us that if you are a basketball player you have a better chance of getting into the NBA if you are from one of these ten states. The reason behind that NY,CA and IL has more NBA players is because that the high school basketball program within these top 3 states are very well developed. Also within NY and CA, there are many poverty places such as Brooklyn(New York), Bronx(New York)and Compton(LA) . Younger generations from these places are more motivated to get into NBA in order to escape from the cycle of poverty .**

6 Results and Conclusion

From our research we have learned a lot about basketball, trends in the games, player statistics, Average analysis of different systems of the game, the different components and part that would affect the game in the long run. How the trends can help us make a prediction in the future and it also showed us good ways on understanding how to look into statistic to make a informed and educated bet, which would yield with a higher probability of you actually making money.

From the information we have been given, we can see that, now an indicidual should know the average numbers, trends, points per game, points over minutes, rebound averages, best players, best teams and frequently chosen state recruits.

Knowing these average number trends and game stats, as I have listed above is a very powerful tool for a regular betting client. We can also target the younger generation that might be influenced by social media and get into the sport to direct them with the accurate information. This will help mitigate the risk for people losing more money or making uneducated betting decisions.

7 Future Research Direction

For the future this analysis can be made much better and more powerful if we conducted the same form on analysis on other sports such as soccer and tennis as with this we can form a database for new betting individuals of a trend and game play analysis of many sports so they will always have a general idea of what they should focus on while betting.

8 References

[1] - Manuel Janeira from the university of Porto and Jaime Sampaio. Retrieved April, 2003, from https://www.researchgate.net/publication/233623969\_Statistical\_analyses\_of\_basketball\_team\_performance\_Understanding\_teams'\_wins\_and\_losses\_according\_to\_a\_different\_index\_of\_ball\_possessions