
# Bust Basics

**Ethan Wang**

**Oct 3, 2023**

# Contents

# 1 Introduction

When it comes to the NBA, all anyone wants to talk about is the first round. And even within the 1st round, all the media buzz surrounds the lottery picks. I am not writing this to say that it shouldn't be that way, but I have always been curious and wanted to shed some light on the 2nd round of the NBA Draft.

Unless you are a mega basketball fan and know each team's roster top to bottom, there is likely a small abundance of roster spots filled by guys that you have never heard much about and pay no attention to unless they play for your team. Seeing as there are only 15 roster spots, give or take a few due to two-way and 10-day contracts, first round picks make up the vast majority of the NBA in any given year. This makes perfect sense. However, this does not mean that the NBA Draft's 2nd round inefficiencies should be understated in any way. From 2005 to 2019, only one-fifth of the 2nd round draft picks in that time period were still playing by 2019 (Kuo, 2019). From 1996 to 2015, 600 2nd rounders were drafted, and only two-thirds of them even saw the floor in an NBA game (Amar, 2016).

Those 2nd round picks are obviously made with the intention that those players will either, by some development magic, become a starters or at the very least fill a bench/rotational role. So how in the world do so many of these picks end up in the bust bin even though we have players like Fred VanVleet and Duncan Robinson making contributions on NBA Finals teams? We could simply attribute the inadequacy of the 2nd round to there not being enough talent available in a given draft class to fuel 60 picks worth of NBA talent. But I don't actually believe that to be true, so that is not what I plan on doing.

# 2 Datasets and the Subject of Analysis

In order to fuel my analysis, I needed to narrow down the scope of the NBA Draft because it takes players from all around the world playing basketball in many different leagues. The way I accomplished this was by establishing draftees of the NCAA Division 1 Men's Basketball League as the pool of players I would analyze. This way, I would not have to deal with the fact that some players are facing off against grown professionals overseas. Hence, I was able to scrape all my data off of Basketball Reference and NBA.com.

Narrowing the subject of analysis scope to only NCAA Division 1 constrained my sample size to players that both have available college statistics from one or more seasons and were drafted in the 2nd round. Sadly, any amazing international 2nd round picks, such as Nikola Jokic, will not be a part of the studied group. The first step in creating my dataset was to download the Basketball Reference statistics for every 2nd round draft pick's NBA career

within my studied time range. The time period contained in my data is 1989-2016. Reason being, in my opinion, the 2016 draft is the most recent that has had ample time to prove boom or bust. The next step was to use a web scraper to iteratively place each 2nd round pick's name into the College Basketball Reference search bar. If that player did not play college ball, their name wouldn't show up; if they did, the scraper would click on their profile and download their college stats.

After a period of data wrangling and merging the two sources of data, I produced the final dataset I used for my research during this project. There ended up being 537 draftees that were both drafted to an NBA team and had at least one season of college basketball under their belt. A decent amount of players were also filtered out due to the fact that their College Basketball Reference page was malfunctioning or they weren't in that database at all, so I had to forfeit those players as well. For these 537 players, I have all their college data in both season average and season total format and their NBA career statistics in both formats as well.

# 3 How Bad is the Second Round Actually?

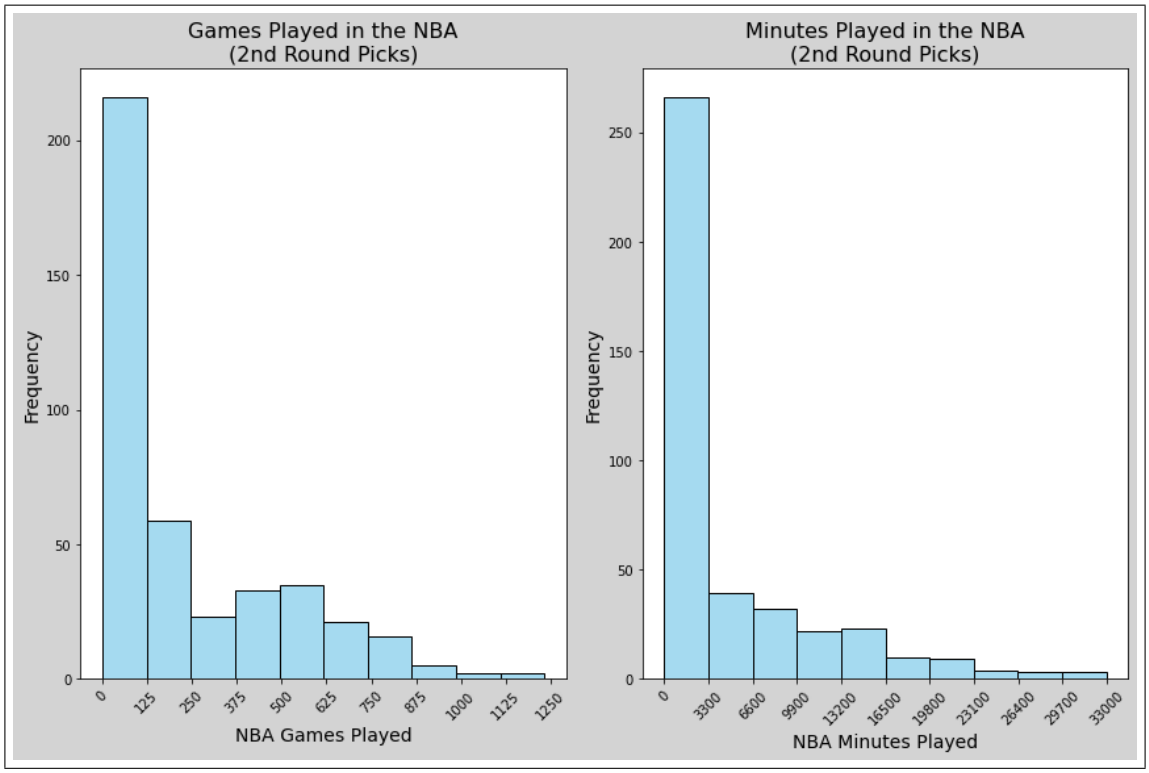


Figure 1: Career Games and Minutes Played

**Figure 1** depicts a distribution of my sample in terms of how many games and minutes they played in their NBA careers. As we can see, a little less than half of the sample falls between 0 and 125 games played, and around half of the sample has played between 0 and 3300 minutes. After all these years, the 2nd round does not even contain players that can last for more than one season, almost half the time. You might as well just keep that roster spot open in the hopes that some international phenom rears his head mid-season. **Table 1**

| 0 GP | 0 MP | Avg WS | Avg VORP |
|---|---|---|---|
| 104 | 105 | 6.663 | 0.925 |

Table 1: More Summary Statistics

is a depiction of how many players in this sample played zero games, how many played zero minutes, the average NBA career win shares of the dataset, and the average NBA career VORP. From a holistic perspective, it is clear that the productivity that comes out of the 2nd round is almost nonexistent.

When brainstorming how best to breakdown the anatomy of the 2nd round over the years using the data I have collected, my initial idea was to identify the standout performers drafted in the 2nd round and analyze what made them special. However, statistics are almost always more effective with a larger number of observations being utilized, and I realized I would waste much of the data that way as well. Therefore, I resolved to make this project a bust analysis.

# 4 Bust Analysis

## 4.1 A Definition of the Word "Bust"

When talking NBA, the term "bust" does not have a concrete definition, in my opinion. When a lottery pick ends up being a 12, 8 and 3 a night guy, people enjoy using the word bust to describe them simply because they did not end up an all-nba level talent. This is not the way I am using the term for the sake of my analysis.

The way I chose to define a bust for this project is any player that has not seen the NBA floor for at least 1000 minutes worth of game play. This is a very forgiving definition of a bust because 1000 minutes is not even a third of a season's worth of time. The reason I decided to keep the threshold relatively low is because I wanted the subjects of study to be either guys that never made it off the bench or guys that, when they did, it was evident they shouldn't be on the court. Any guy who was given a season or two worth of chances obviously showed promise in some way, even if productivity never came along with it. Maybe

they got injured, or maybe there was an unlucky trade deadline decision that moved them to the bench for good.

With my new "busts" dataset, I divided it up into guards, forwards, and centers. I used height cutoffs to do this where guards were any player 6'4" and below, forwards were players 6'5"-6'9", and centers were anyone above 6'9". The NBA has harbored positional fluidity for awhile, so height was the most objective way to classify position for the whole dataset.

## 4.2 Plot Analysis

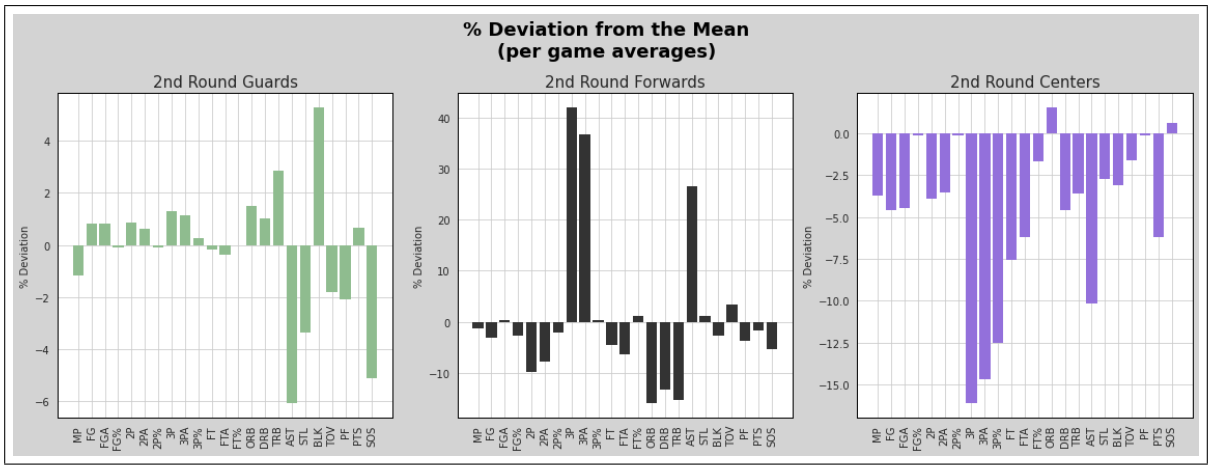


Figure 2: % Deviation from the Mean (career per game)

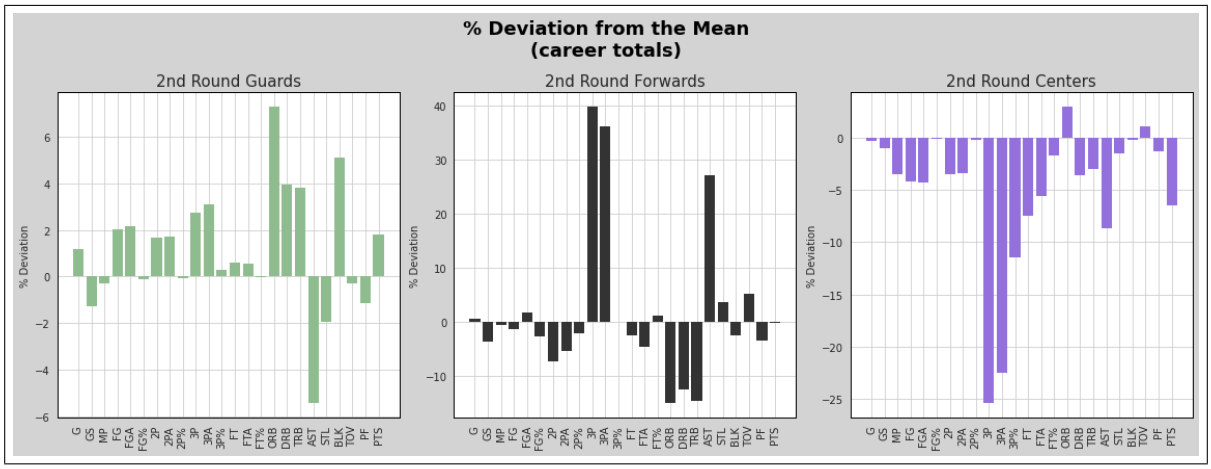


Figure 3: % Deviation from the Mean (career totals)

Shown above in **Figure 2** and **Figure 3** are the two main visualizations I will use for my analysis. For each player defined as a bust, I used their college statistics to create the aforementioned dataset. First and foremost, I want to discuss the statistic "% deviation

from the mean". The means used in this calculation are the means of the ENTIRE dataset, including both busts and non-busts for every statistic depicted. However, the players within the busts dataset are the only data points used to calculate the % deviation from the overall dataset's means. Therefore, % deviation from the mean represents how far, either negative or positive, the BUSTS collectively fall from the means of ALL 2nd round picks.

Now that we understand the significance of the statistics depicted on the plot, let's dive in. Right now, I will only discuss **Figure 2**. Let's start with guards. It is clear that minutes played per game, free throw %, and every statistic in between those two don't have much significance. Things like 2P% and 3P% are important statistics to pay attention to when drafting a guard, but since these are obvious necessities for an NBA prospect, it is likely every guard drafted met a certain threshold for these attributes. Things get interesting when we look at rebounds. For both offensive and team rebounds, the % deviation sits around +2%. It seems that rebounding aptitude is in a sense, overrated when analyzing a guard's college resume. Assists and steals are at huge negative deviations around -3% and -6%, respectively. It is obvious that playmaking and perimeter defense were undervalued when analyzing a guard's college averages. Interestingly enough, blocks lie on the opposite end of the spectrum at +5%, representing a huge overvaluation of the statistic. Strength of schedule is also at -5%, which is also very telling.

From an aggregate perspective, we can make some very informed assumptions about the inefficiencies when drafting 2nd round guards. When seeking players with potential, scouts most likely salivate when they see a big, strong guard who can rebound and impose his will on defense using athleticism. However, in reality, the guards who succeed are the ones who are fundamental playmakers with high IQ, maybe a little smaller and quicker with good hands. These assumptions are based on the fact that things like rebounding and blocks were clearly overvalued, but steals and assists were highly undervalued. Those facts create a pattern that shows a certain guard archetype. Also, the busts evidently had much weaker strengths of schedule. Power 5 guards ¿ other guards? Maybe.

Let's move on to forwards. Very polarizing. +40% deviation for 3 pointers made and 3 point attempts? Wow. This is very interesting because we all know those forwards who can fill it up from behind the line have changed the league. This tells me one thing: we must remember that I am analyzing the 2nd round. When looking for a franchise guy, a 6'7" forward who can score at all three levels is the perfect candidate. But what about when looking for a rotational SF or PF that can defend and provide useful bench minutes? Maybe we should look away from 3 point aptitude in that case. It is not likely that your 2nd round forward pick will be taking many 3 point shots away from your main shot-takers anyway so it actually makes perfect sense that this was far overvalued when drafting 2nd round forwards.

For the exact same reason, assists were also extremely overvalued. What 2nd round forward is going to have the ball in his hand? It is almost poetic that all three rebounding categories were plainly undervalued in 2nd round forwards.

So what did we learn about the forwards. We should maybe not be fooled by a guy who can hit step-back 3s if he can't rebound, defend, and do all the things a role player should do well. In the second round, that is.

As for centers, I don't have much analysis to provide. This is because it seems that centers who ended up as busts were just unmistakably worse in every statistical category during college than those who succeeded. Interpret that how you will.

I would now like to discuss **Figure 3** a little. In my opinion, for the sake of bust analysis, **Figure 2** is simply a more objective and better way to make assumptions. Per game averages over a player's college career consider everything from how he was right out of the gate to how he progressed, all the way to his final year in college. Totals, on the other hand, are not as useful because players did not all play the same amount of games or minutes. Whether an injury stopped them or they got drafted after 1-2 years vs. 3-4 years, no one plays the exact same amount of basketball in college. However, that does not mean this plot is useless. For the most part, the direction of deviation remains the same when moving from **Figure 2** to **Figure 3**, but the magnitude is what changes. For guards and centers, deviations seem to get more extreme, but for forwards, they get a little less extreme. Another, more general assumption can be made from this. And that is that for forwards, more college experience might counteract the extent to which overvaluing and undervaluing statistics play a role. The opposite is the case for guards and centers. More college experience means the totals hold more weight for a given overvalued or undervalued trait of a given player.

# 5 Conclusion

I fully understand that none of the data research done in this project is sufficient to make conclusive statements due to the lack of causal findings. However, my dataset contains the majority of second round picks from the 27 most recent drafts that have been given enough time to pan out. This is more than enough to attribute significance to the trends within the data.

Not only is the sample size sufficient, but the findings are quite intuitive as well. Between both guards and forwards, the data showed that valuable traits for each position were exactly what they should be. Guards that make their success through finding the open man usually tend to pan out, and guards that were highly coveted in draft rooms because of their size and athletic ability tend to fall short of the mark. Forwards that can rebound and provide

utility on the defensive end are much more likely to make good on their roster spot than forwards that were drafted for their three point prowess. Based on the data, the center position remains a mystery to me because it seems obvious that these guys would end up as busts if they were below average in every single statistical category except for offensive rebounding. Once again though, it could be a similar issue to the forwards. If a center gets a block or two a game that he sends into the third row of the bleachers or seems like a lethal, above-the-rim lob threat, that might fool some teams into drafting him.

To summarize the sentiment of my findings in this project, I would say that at times, teams may lose sight of what is actually still available when the 2nd round comes around. At the end of the day, only five guys can be on the court at a time for a given team, and that does not leave much room for could be potential to shine. Billions of people play basketball around the world, and less than 600 guys can play it at the very highest level. The craziest part is that among those 600 guys, only a handful have the ability to dominate on a nightly basis, and having the ability doesn't even mean they will. Hence, it makes sense to me that after the 30 best guys are drafted, it is exceptionally rare that a team will find a bonafide star in the making during the 2nd round. Maybe it's time for the 2nd round of the NBA draft to be a place where you put your faith in a player that lives and breathes fundamentals. That guy is much more likely to live up to his projection of "solid rotational piece", than the other guy is to live up to his projection of "longshot second coming of Kevin Durant".

-To see my Python code for this project: https://github.com/ethwang17/BustBasics.git
-Code is in the "nba_scraper.py" and "nba_draft.py" files

# Works Cited

1. Kuo, N. (2019, March 2). Trash or treasure? an analysis of NBA Second Round Picks. https://sportsanalytics.berkeley.edu/articles/trash-or-treasure.html

2. Amar, A. (2016, June 28). NBA draft 2016: Historically, how many second rounders become rotationnbsp;players?. SLC Dunk. https://www.slcdunk.com/nba-draft/2016/6/27/12048960/nba-draft-2016-historically-how-many-second-rounders-become-rotation