

DATA VISUALIZATION OF THE “TRANSFER MARKET” IN THE COLLEGE BASKETBALL WORLD

XU LIAN
UNIVERSITY OF SAN FRANCISCO
MSDS 622 Final Project

A special thanks to Stephen Devlin for his guidance in this project.

Link to Github:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project

Introduction

In recent years, there is an ongoing trend that more coaches are into the “transfer market” in the NCAA Division I Men’s Basketball world. The motivation behind this phenomenon is based on the assumption that a transfer player is more seasoned than an average high school recruit. Coaches recruit such a type of players with an anticipation that the latter will give their teams an immediate boost.

This data visualization project attempts to analyze how players performed after they transferred to a new school. In addition, this work also serves as a foundation for a modeling project that attempts to estimate how certain players would project at certain levels in the future.

The dataset in this project comes from the four sources below:

- Transfer Dataset: This dataset contains all the records of players who transferred from 2012 to 2017. It is scraped from VerbalCommits.com.
- College Stats Dataset: This dataset contains various standard and advanced statistics of players who had played in any NCAA Men's Basketball Div. I game from 2012 to 2017. It is scraped from College Basketball Reference (<https://www.sports-reference.com/cbb/>).
- School Name Dataset: Since VerbalCommits.com and College Basketball Reference use different sets of school names (for example, it can be UTSA for one dataset and Texas-San Antonio for the other), it is necessary to have a dictionary of school names for merging purposes. The dataset is scraped from Wikipedia. In addition, geographic coordinates are obtained from Google Map API.
- Bucket Dataset: It is a usual procedure to divide all the NCAA Div. I program into three 'buckets': Low-Majors, Mid-Majors, and High-Majors. However, a basketball powerhouse like Gonzaga should not be considered as a Mid-Major-caliber team even though it is not affiliated with one of the Power Five Conferences (ACC, Big 12, Big Ten, Pac-12 and SEC). Hence, there is a need to introduce a better rating system. This dataset comes from the coaching staff of the University of San Francisco's Men's Basketball team, rating every program based on a 1-9 scale.

Table of Contents

5 | Summary of Data

Part I: General Information About Transfer Players

Part II: Using BPM to Examine Players Who “Transferred Up”

Part III: Examining Transfers’ Performance in Their New Schools

17 | Storyline

How We Use Historical Data to Project Transfers’ Future Performance

21 | Conclusion

22 | Appendix of Code

24 | Citations

Summary of Data

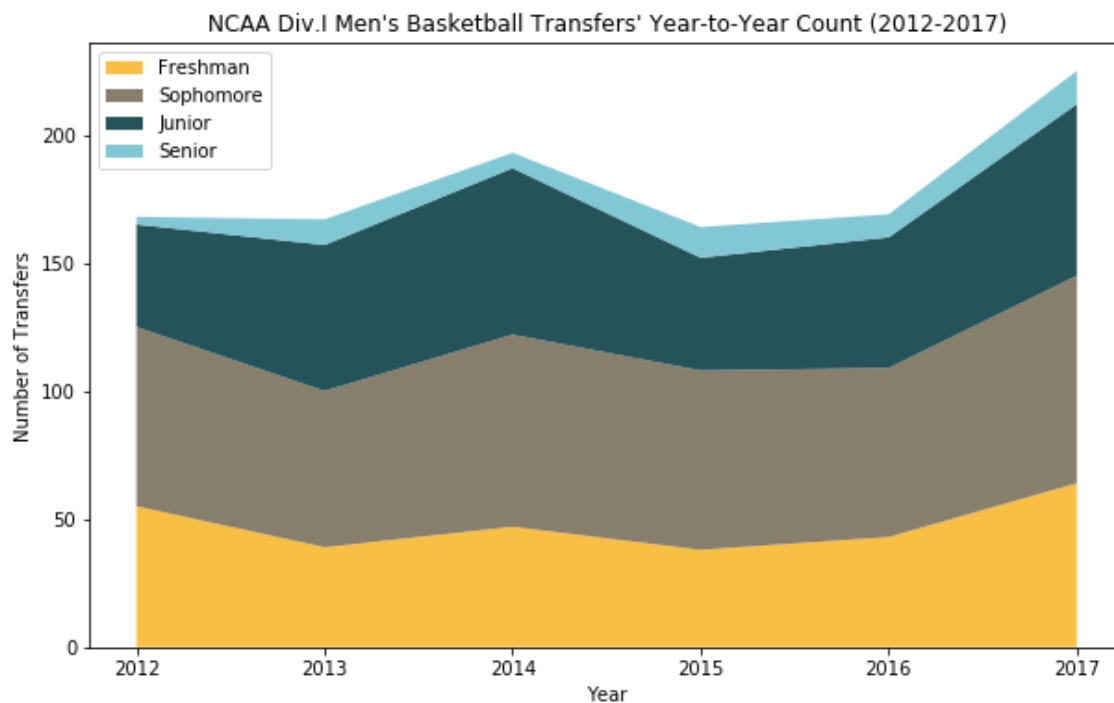
This project focuses on the players who transferred to other colleges from 2012 to 2017. After data processing, 1088 players are included.

The summary of data for this project consists of three parts.

- *Part I* gives general information regarding the transfer players.
- *Part II* focuses on the players who transferred up to a higher level and examines the quality of the data. For example, players who transferred up but walked on a High-Majors school should be excluded.
- *Part III* looks into players' performance with their new schools and introduces a new metric for the purpose of projecting a transfer player's performance in the future.

Part I: General Information About Transfer Players

The stacked area chart below shows the year-to-year trend of the “transfer market.” In total, the number of transfers increased over years, primarily due to the increase in the number of the transfers who are freshman and sophomore.

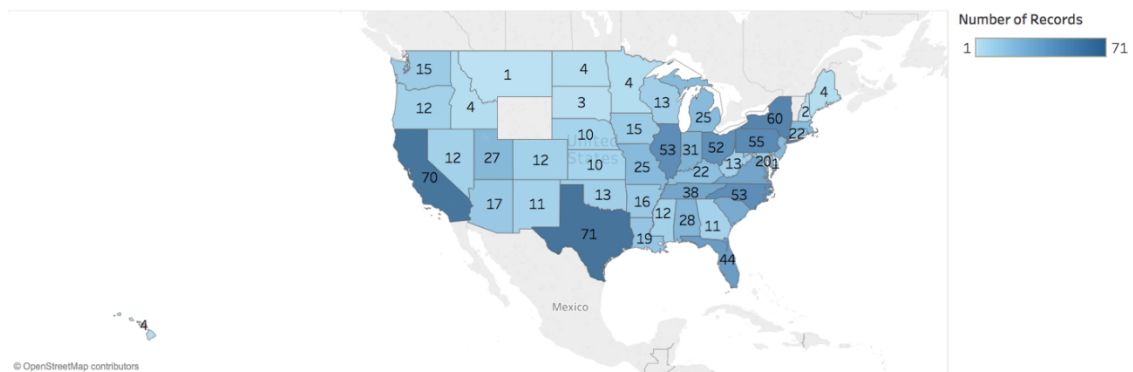


The choropleth map below illustrates the number of outgoing and incoming transfers by state.

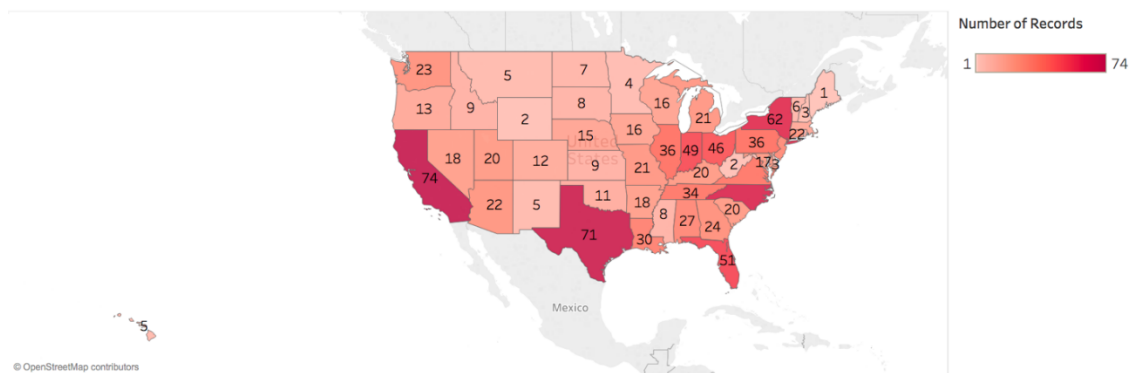
Unsurprisingly, California and Texas are the ones produced the most influx and outflux since both of these states have a large number of NCAA Div. I men's basketball programs.

Wyoming is the only state without a single outgoing transfer. This is not surprising, either. University of Wyoming is the only Div. I school in the state of Wyoming. If the program did not have an exit from 2012 to 2017, then there would be no outgoing transfers for the state.

Number of Outgoing Transfers by State (2012-2017)
NCAA Div.I Men's Basketball



Number of Incoming Transfers by State (2012-2017)
NCAA Div.I Men's Basketball

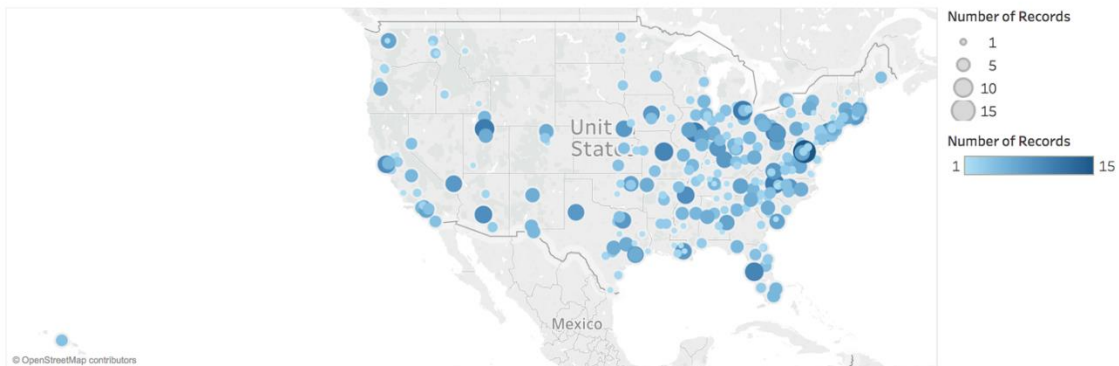


The bubble map below illustrates the number of outgoing and incoming transfers by school.

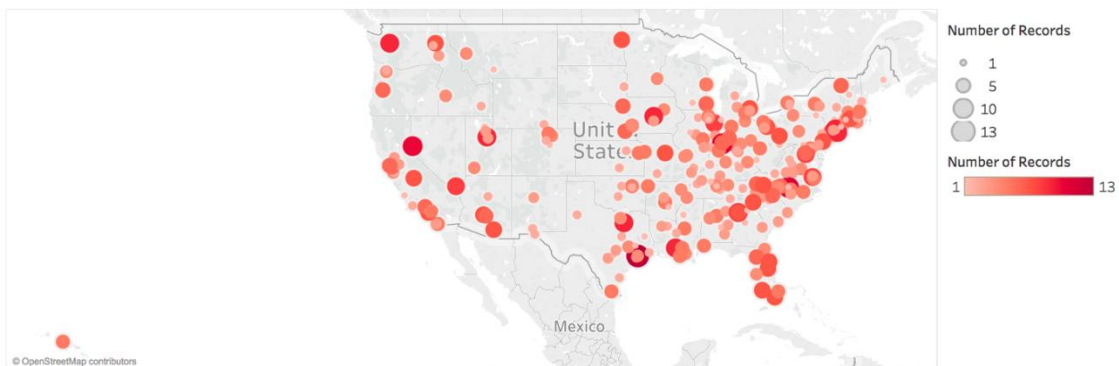
The school that lost the most players from 2012 to 2017 is George Washington. This makes sense since there have been reports that multiple players charged the then-head coach Mike Lonergan with "verbal and emotional abuse", ultimately leading to the school's decision to fire him in 2016. [1]

For the schools that had the most influx, there are a few interesting cases. Nevada had 10 transfers coming in during this period and saw them coming to fruition. In 2018, the Wolfpacks advanced to Sweet 16.

Number of Outgoing Transfers by School (2012-2017)
NCAA Div.I Men's Basketball



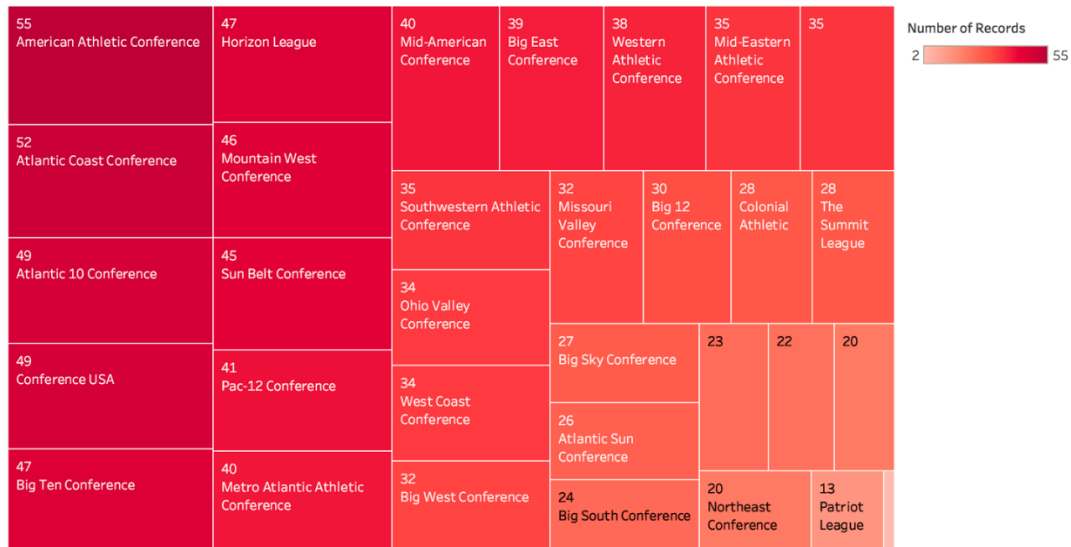
Number of Incoming Transfers by School (2012-2017)
NCAA Div.I Men's Basketball



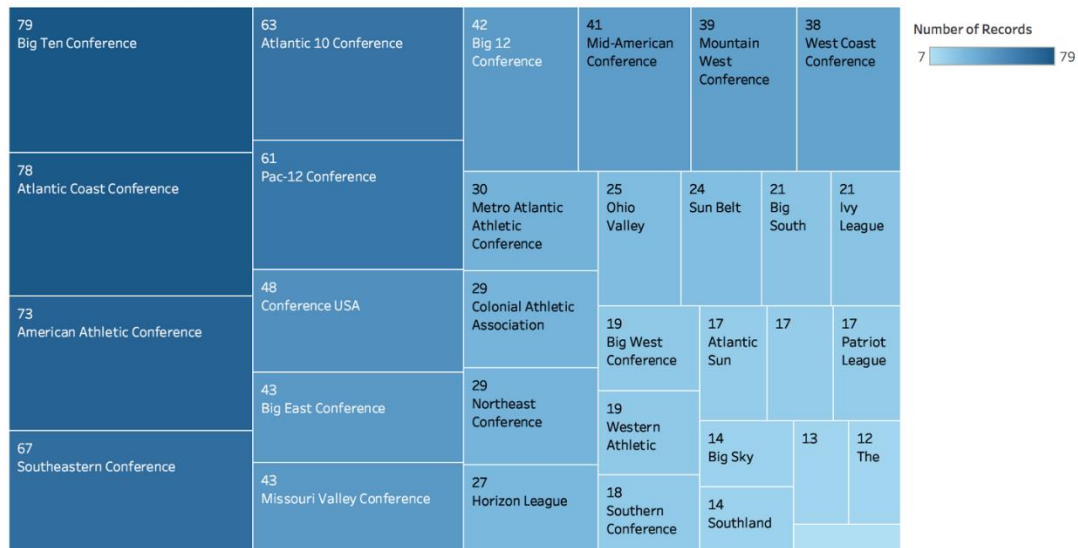
The tree maps below illustrate the number of outgoing and incoming transfers by conference.

It seems the conference with the most outflux also had the most influx, as ACC, AAC, A-10 and Big Ten went to top 5 in both of the tree maps.

Number of Incoming Transfers by Conference (2012-2017)
NCAA Div.I Men's Basketball

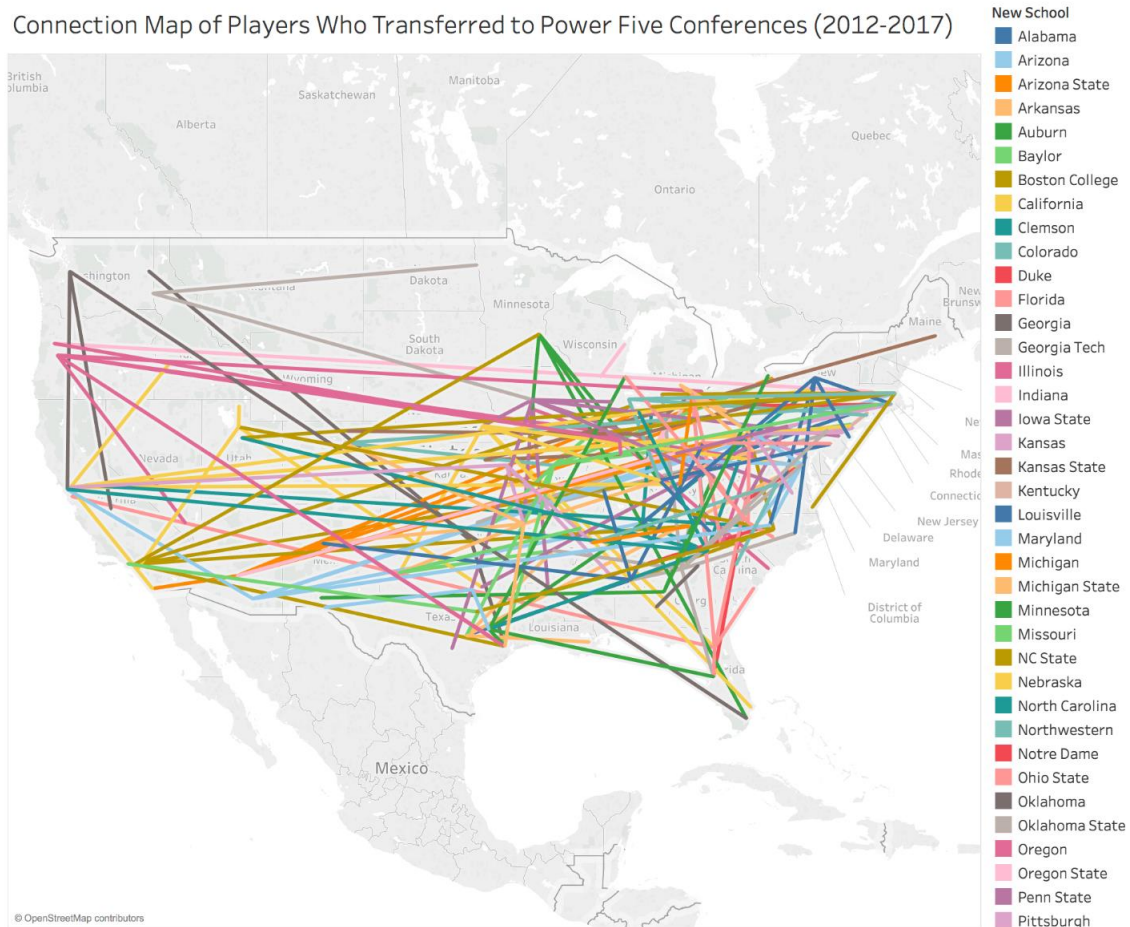


Number of Outgoing Transfers by Conference (2012-2017)
NCAA Div.I Men's Basketball

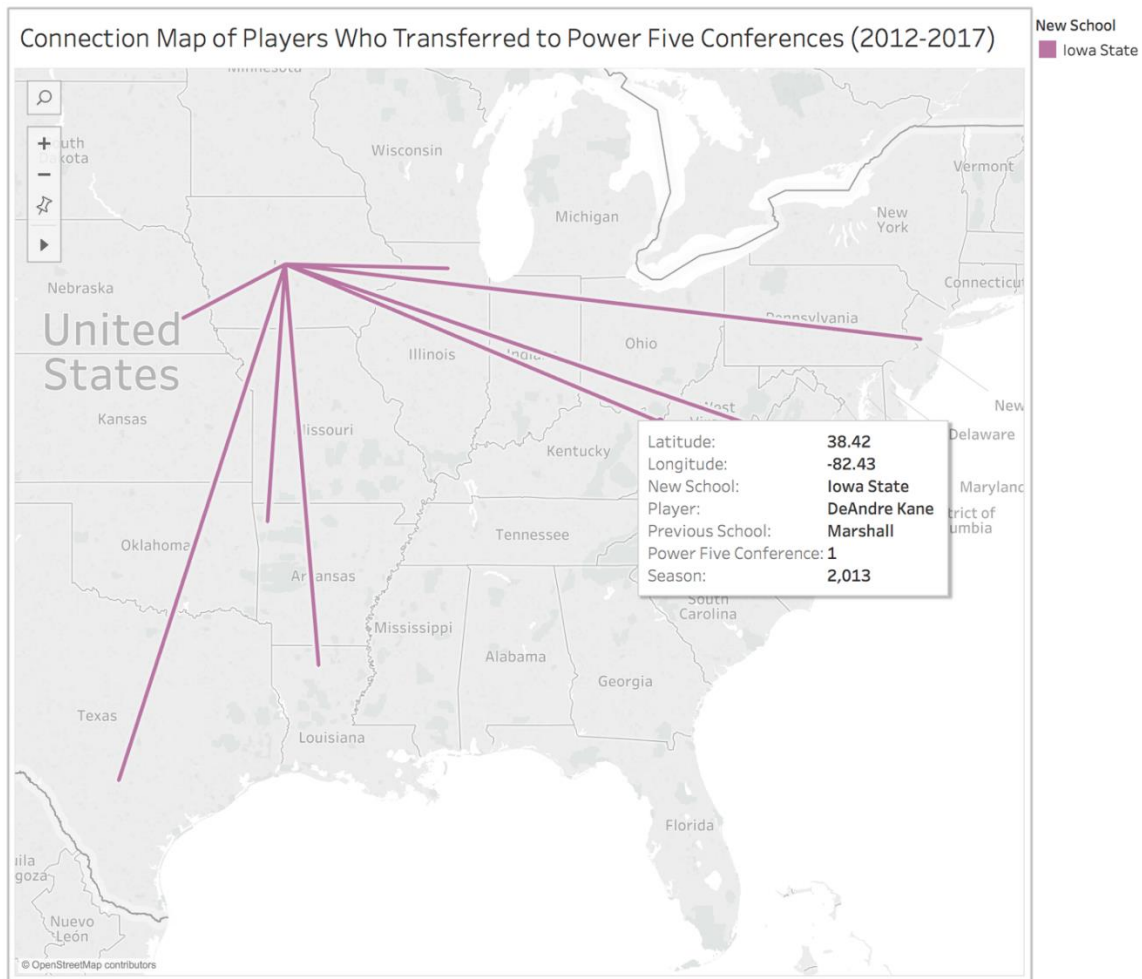


The connection map below shows all the records of players transferred to the 65 schools that are affiliated with the Power Five conferences (ACC, Big 12, Big Ten, Pac-12 and SEC) from 2012 to 2017.

Connection Map of Players Who Transferred to Power Five Conferences (2012-2017)



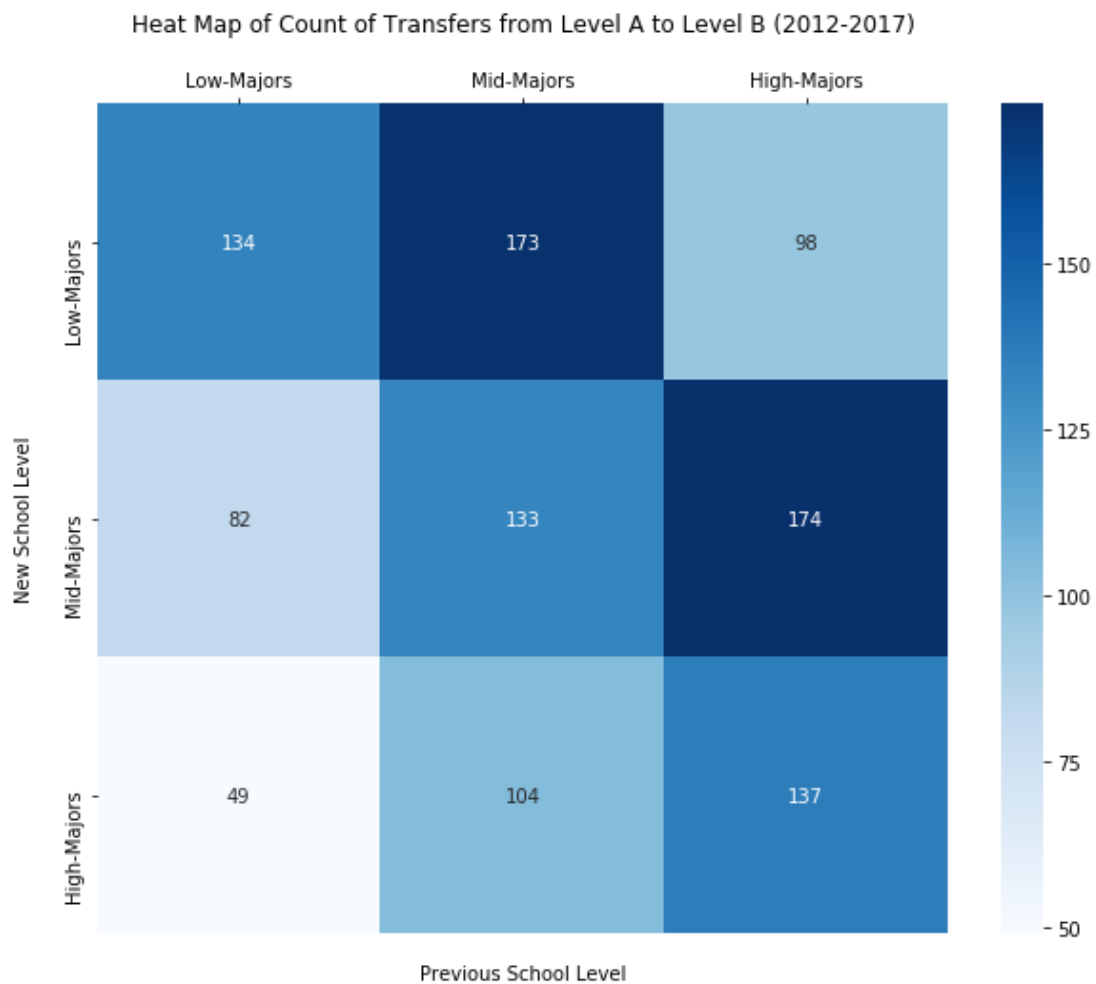
This connection map could serve as an interaction plot that allows people to check out the transfer information of the schools which they are interested in. Below is an example of Iowa State, a program that is known for recruiting highly-profiled transfers. [2]



The heat map below incorporates the ratings regarding program strength by the University of San Francisco's coaching staff. All of the NCAA Div. I men's basketball programs are classified into three wide levels: Low-Majors (1-3), Mid-Majors (4-6), and High-Majors (7-9).

According to the heat map, the cases of players transferring from High-Majors to Mid-Majors and those transferring from Mid-Majors to Low-Majors are the most populous.

However, for the following parts, we instead focused on the cases of players transferring to a higher competition level.

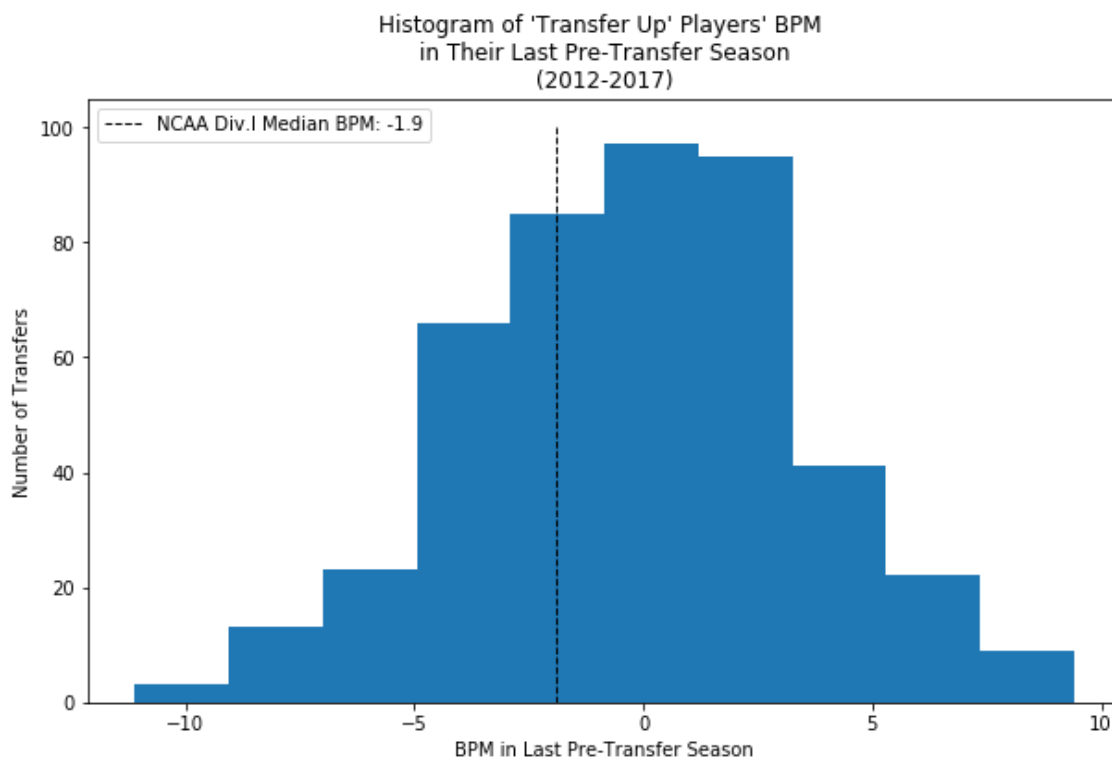


Part II: Using BPM to Examine Players Who “Transferred Up”

Box Plus-Minus (BPM) is a popular “one-number-tells-all” metric that attempts to summarize a player’s contribution or value to a team. For this part, it is used to examine the quality of the transfers.

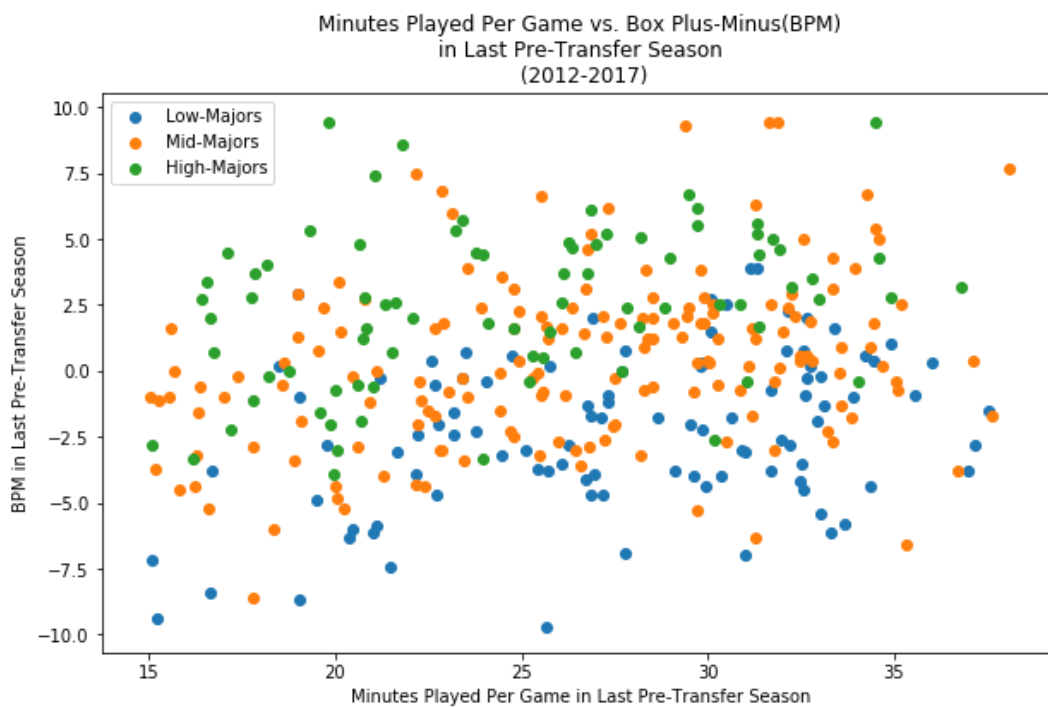
In general, when people see a player transferring from a lower level to a higher one, they would tend to assume that player has made enough strides to play at a bigger stage. However, the histogram below suggests that there are some cases with a vastly different profile.

From 2012 to 2017, the median BPM of NCAA Div. I Men’s Basketball is around -1.9. It is interesting to see a good portion of the ‘transfer up’ cases ended up on the left side of -1.9 BPM. Moreover, some of them even got into the sub-ten territory.



The scatterplot below takes a deeper look at the “transfer up” cohort. It includes all the players who “transferred up” and played at least 15 minutes in their last pre-transfer season.

From the plot, most of the players with an extremely poor BPM values come from Low-Majors schools. In addition, the majority did not log a lot of minutes, either. Therefore, it is reasonable to suspect that these players ‘transferred up’ to walk-on.



The table below lists four players that transferred from Low-Majors to High-Majors with a super low BPM in their last pre-transfer seasons. After searching their situations individually, it turns out all of these players walked-on in their new teams.

	Team	Player	MP	BPM	Season	prev_school	new_school	prev_school_wide_bucket	new_school_wide_bucket	Type
462	robert-morris	Braden Burke	487.0	-9.4	2017	Robert Morris	Michigan State	Low-Majors	High-Majors	sit out
696	lipscomb	Carter Sanderson	360.0	-8.6	2014	Lipscomb	Syracuse	Low-Majors	High-Majors	graduate
3056	chicago-state	Kurt Karis	294.0	-8.4	2015	Chicago State	USC	Low-Majors	High-Majors	sit out
2345	houston-baptist	Jimmy Barton	516.0	-8.4	2014	Houston Baptist	Missouri	Low-Majors	High-Majors	sit out

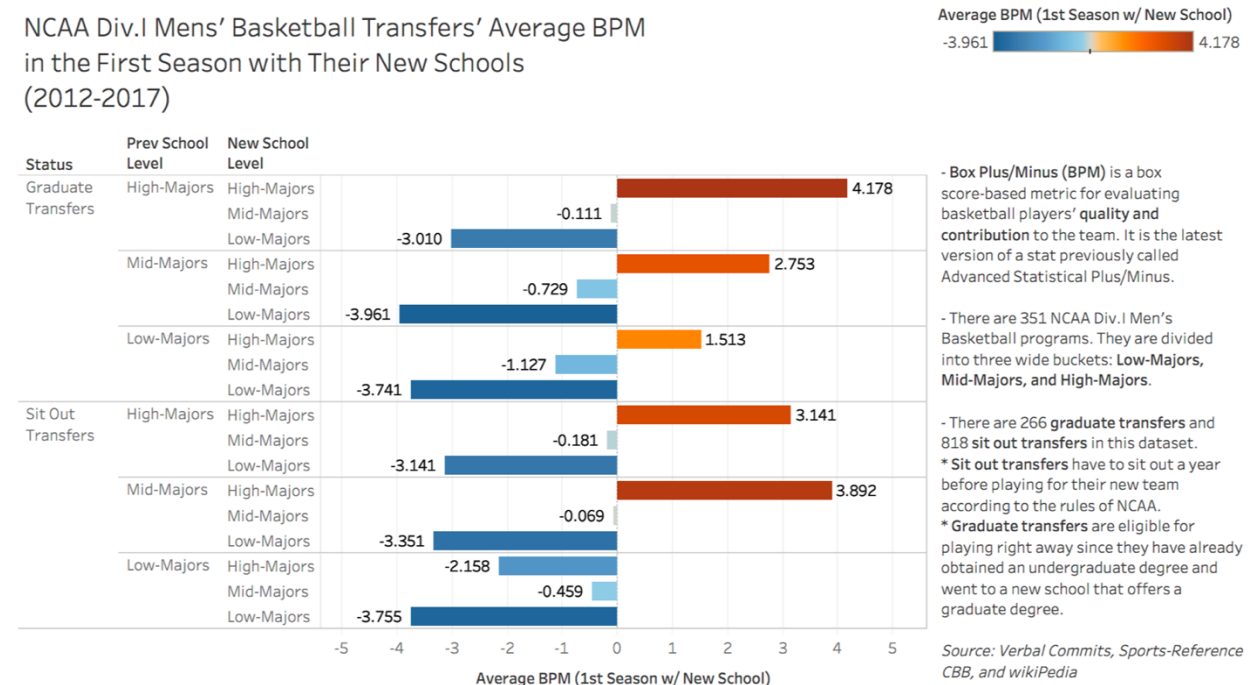
Part III: Examining Transfers' Performance in Their New Schools

The barplot below takes a look at how players performed in their first post-transfer season.

In general, players that transferred to a High-Majors school tend to achieve a higher BPM. This makes sense since they made enough strides in the lower levels that enticed interests from Power Conference schools.

On the other hand, players who 'transfer down' generally don't produce or perform, even though the level of competition went lower.

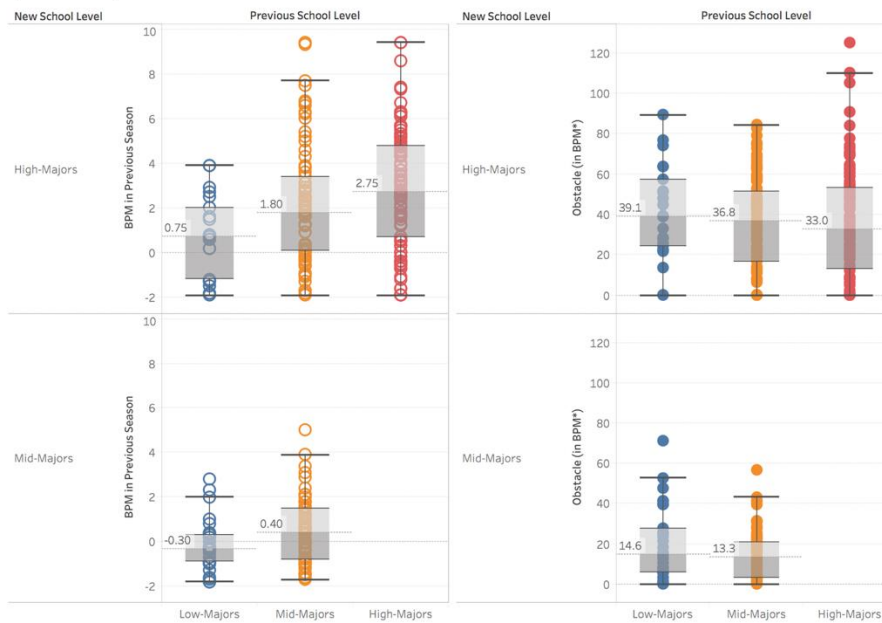
One interesting observation is that graduate transfers who went from Low-Majors to High-Majors perform much better than those similar-typed sit out transfers. One of the factors could be years of experience.



Suppose a player transfers from a low competition level to a higher one. He would probably to take a back seat in his new team due to the presence of more established players.

The boxplot below attempts to quantifies and visualizes the “depth”. Details are given in the **Storyline** section.

Obstacles for NCAA Men’s Basketball Players
Who Transferred Up
(2012-2017)



Box Plus-Minus

Box Plus/Minus (BPM) is a box score-based metric for evaluating basketball players' quality and contribution to the team. It is the latest version of a stat previously called Advanced Statistical Plus/Minus.

* Mechanism of Calculating the Obstacle

Obstacle = Total BPM of players who had a higher BPM than the transfer in the latter's new school in previous season. For example, Transfer A had a 0.5 BPM in his last season with his previous school. In that season, Player 1, 2, 3 were the only three players in the Transfer A's new school who had a higher BPM. These three players had 1, 1.5 and 2 BPM, respectively. Therefore, Transfer A's "obstacle" score is 4.5 BPM.

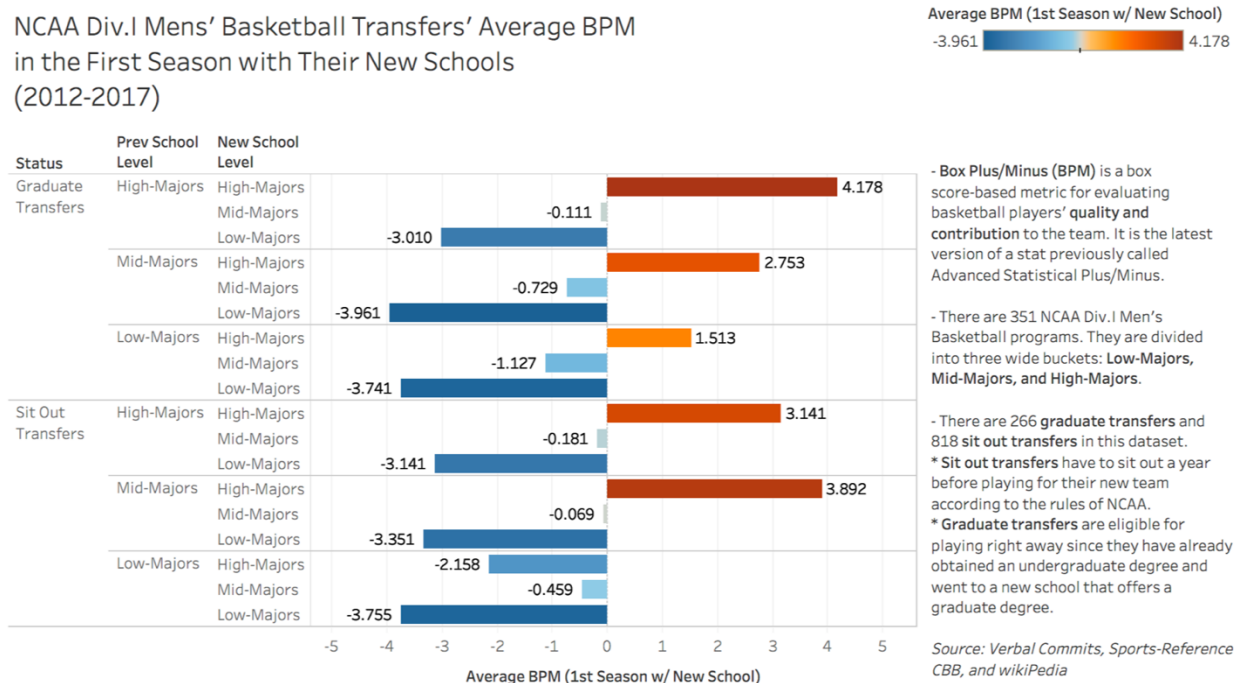
Storyline

How We Use Historical Data to Project Transfers' Future Performance

One party that can be benefitted from the analysis on historical data of the transfers, are the coaches. A model that is able to project a potential transfer's future value can help their decision makings.

Before building a projection model, there is a variable selection process. For this project, a natural option for the response variable would be BPM in transfers' first season with their new school, since BPM is one of the best metrics that summarizes a player's contribution to a team.

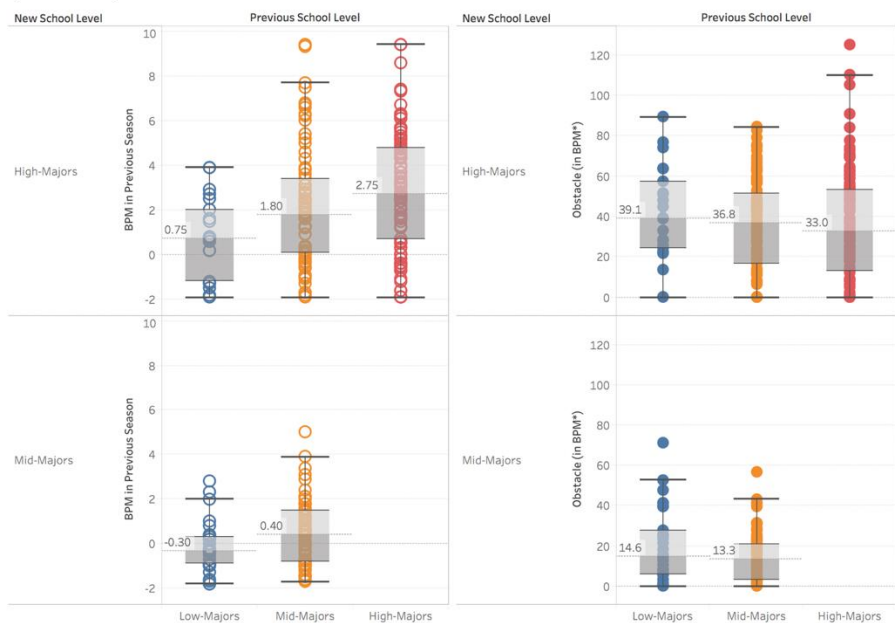
Let's revisit the barplot. It provides some valuable information. For example, players that "transferred up" to High-Majors are expected to yield a high BPM, while those transferring to Mid-Majors and Low-Majors are expected to have a subzero BPM.



However, such information could lead to flawed results in modeling. Suppose I am a coach of a Mid-Majors school. I want to recruit a player that has the caliber to transfer into a High-Majors school. If I use the information in the bar plot above, does that mean the targeted player would perform worse on my team than how he performs for a High-Majors school? Therefore, we need some metrics to counter this problem. Specifically, we need a metric that can serve as a penalty.

Therefore, in the boxplot below, I introduce a metric called “obstacle”. It is the sum of BPM of players that had a higher BPM in the previous season than the transfer player. The idea behind the metric is, if you transfer into a High-Majors program like Duke, then you would face more competition in the team. However, if you transfer into a Mid-Majors program instead, there would be fewer guys who are in front of you in the team’s depth chart. By having less in-team competition, you would probably play better than how you perform for a High-Majors school.

Obstacles for NCAA Men’s Basketball Players
Who Transferred Up
(2012-2017)



Box Plus-Minus

Box Plus/Minus (BPM) is a box score-based metric for evaluating basketball players' quality and contribution to the team. It is the latest version of a stat previously called Advanced Statistical Plus/Minus.

* Mechanism of Calculating the Obstacle

Obstacle = Total BPM of players who had a higher BPM than the transfer in the latter's new school in previous season. For example, Transfer A had a 0.5 BPM in his last season with his previous school. In that season, Player 1, 2, 3 were the only three players in the Transfer A's new school who had a higher BPM. These three players had 1, 1.5 and 2 BPM, respectively. Therefore, Transfer A's 'obstacle' score is 4.5 BPM.

Since this plot only considers the “transfer up” cases, there are no cases regarding High-Majors players transferring into a Mid-Majors school.

The left side focuses on the individual performance. Players transferred into High-Majors generally had a higher BPM in their last pre-transfer seasons than those who transferred to Mid-Majors, on average.

The “obstacle” metric comes in the right side. Players transferred to High-Majors programs would get more “obstacle”, like 20+ BPM, than those who transferred to Mid-Majors.

Therefore, the plot confirms the assumption stated previously, and we can use this “obstacle” feature for the modeling.

In addition, this “obstacle” metric is also independent. For example, Florida State and Duke are both High-Majors but with different strength. When we project a potential transfer into Duke, the latter would get a different “obstacle” score from the one where we project him into FSU. Hence, the “obstacle” metric would allow us to make projections on a school-by-school basis.

Conclusion

College basketball coaches invest more effort in the “transfer market” than ever, because they believe in seasoned college players’ abilities to produce right away. A model that projects a transfer’s value in future could help coaches to make a better judgement on their potential targets.

Nonetheless, since high-caliber players usually go to High-Majors program and yield high production, it creates difficulties when projecting such a type of player into the Mid-Majors.

This project introduces “obstacle”, a metric that serves as a penalty variable. By using data visualization, the metric shows its ability to help project a High-Majors-caliber player to a higher BPM if he elects to transfer to a Mid-Majors program, delivering a quality attempt to the challenge.

Appendix of Code

Histogram:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Histogram.ipynb

Barplot:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Barplot.ipynb

Boxplot:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Boxplot.ipynb

Scatterplot:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Scatterplot.ipynb

Bubble Map:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Bubble_Map.ipynb

Choropleth Map:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Chloropleth_Map.ipynb

Connection Map:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Connection_Map.ipynb

Heat Map:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Heat_Map.ipynb

Stacked Area Chart:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Stacked_Area_Chart.ipynb

Tree Map:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Tree_Maps.ipynb

Interactive Plot:

https://github.com/xulianrenzoku/College_Div_I_Mens_Basketball_Transfer_Project/blob/master/Code/Connection_Map.twb

Citations

- [1] Kilgore, Adam (July 21, 2016). "GW basketball players report coach's 'verbal and emotional abuse'; many fled school". *The Washington Post*. Retrieved July 21, 2016.
- [2] NCB - Iowa State coach Fred Hoiberg has the Cyclones rolling thanks to transfers - *ESPN The Magazine* - ESPN