Part 1: Data Comprehension

1) Which NBA team(s) has drafted the most players who…

---

a) went to Duke and were drafted in or before the 2000 draft?

- Teams that drafted the most players 2 from Duke:
  ['DAL', 'MIN', 'PHO']

	TEAM	YEAR	PLAYER
0	DAL	1990	Phil Henderson
1	DAL	1995	Cherokee Parks
2	MIN	1992	Christian Laettner
3	MIN	1999	William Avery
4	PHO	1992	Brian Davis
5	PHO	1994	Antonio Lan

b) have a first name that begins with D and were drafted in an even year draft (1990, 1992, 1994, …)?

- Teams that drafted the most (7) D named players on even years from Duke:
  ['MIL', 'BOS', 'SEA']

	TEAM	YEAR	PLAYER
0	MIL	1998	Dirk Nowitzki
1	MIL	2002	Dan Gadzuric
2	MIL	2006	David Noel
3	MIL	2010	Darington Hobson
4	MIL	2012	Doron Lamb
5	MIL	2014	Damien Inglis
6	MIL	2018	Donte DiVincenzo
7	BOS	1990	Dee Brown
8	BOS	1992	Darren Morningstar
9	BOS	2002	Darius Songaila
10	BOS	2004	Delonte West
11	BOS	2016	Deyonta Davis
12	BOS	2016	Demetrius Jackson
13	BOS	2020	Desmond Bane
14	SEA	1992	Doug Christie
15	SEA	1994	Dontonio Wingfield
16	SEA	1996	Drew Barry
17	SEA	2000	Desmond Mason
18	SEA	2004	David Young
19	SEA	2006	Denham Brown
20	SEA	2008	DeVon Hardin

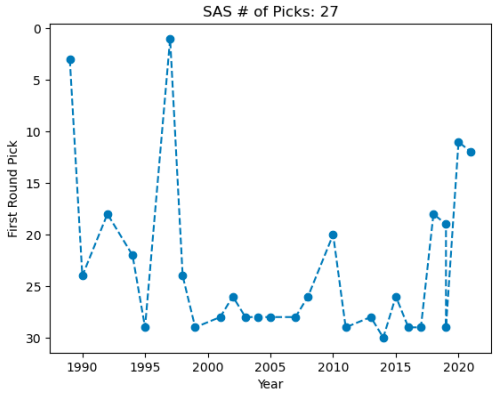
2) Describe the relationship between a team's first round pick slot in one year with their first-round pick slot in the subsequent year.

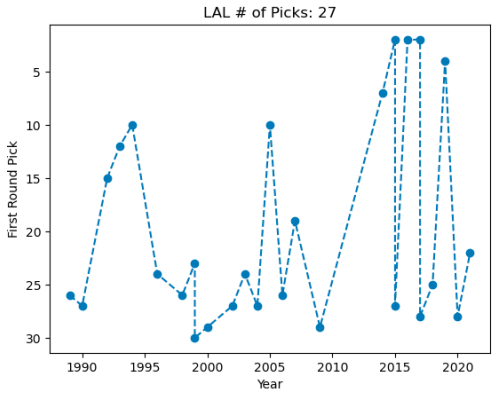
(In the analysis, I will not be taking into account trades, injuries, or rule changes that might have occurred throughout the previous season.)

There are three general trends for year to year first round pick slots.

i. Year #1: High Pick (Lottery Pick)
   Year #2: Low Pick
ii. Year #1: Low Pick
   Year #2: Low Pick
iii. Year #1: High Pick (Lottery Pick)
   Year #2: High Pick (Lottery Pick)
iv. *Year #1: Low Pick*
   Year #2: High Pick (Lottery Pick)

As an example, the San Antonio Spurs and the Los Angeles Lakers draft picks over the last 30 years.





The years after SAS had high draft picks (1989 and 1997) correlate with the first trend: High Pick/Low Pick. This signifies the team had great season success with the pick.

In the years in between 1990-1996, the team had low picks in the first round (correlating with the team' season success). This is an example of the second trend.

On the other-hand the Lakers have had a difficult decade where they had High Pick selections between 2014 – 2017. In this trend, the team does not see immediate impact by the selected prospect and subsequently are back in draft lottery contention.

The final trend Low Pick/High Pick, correlates with a season where the team undergoes a change. Usually, a star player's injury or trade. (SAS 1996-1997)

Part 2: Analytical Acumen

3) Analyze draft position value and team success/deficiencies compared to expectation.

a) Create a method for valuing each draft slot in the NBA Draft (picks 1 through 60 in most drafts).

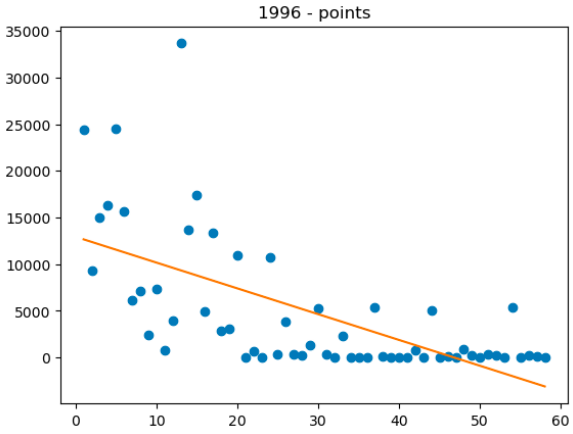
- For simplicity, I have ranked each draft spot on a log scale and I have manipulated the scale to be from 0-100. (This will be used for later comparison):

  2 - (np.log10(pick))) * 50
  (with 0 being the lowest possible ranking and 100 being the highest).

- In my code, I have isolated each year's draft and given a specific 'YEAR_PICK_RANKING' signifying the pick's value.

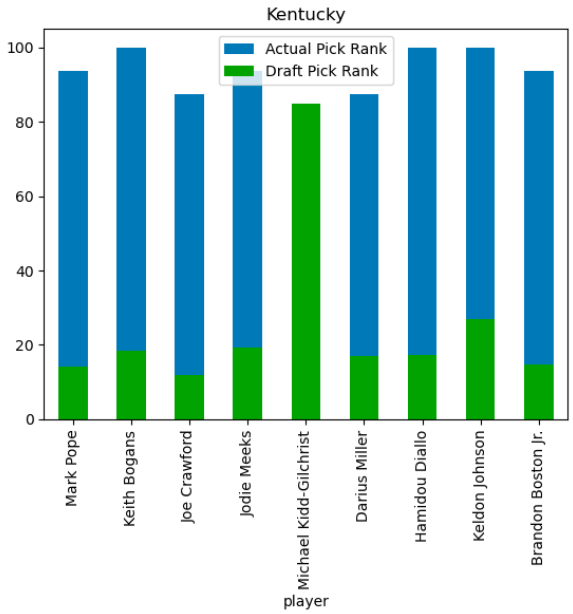
b) Conditional on the expected value of the draft positions, which NBA teams have over or underperformed the most when drafting during this time span. Which College Teams have had the players outperform expectations the most after entering the NBA?

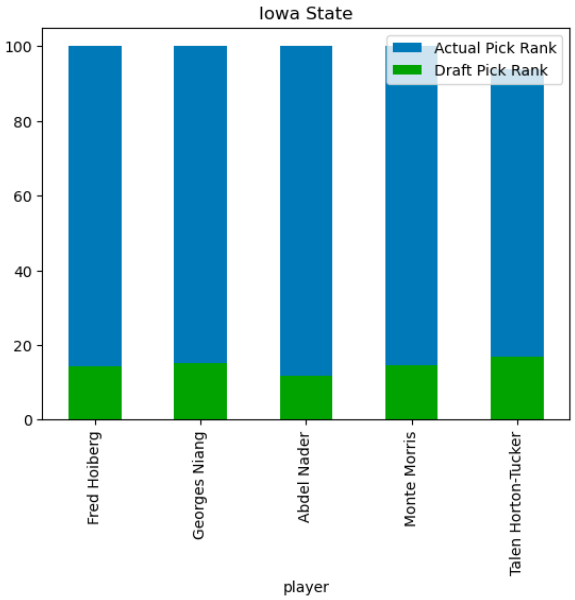
- To find the "over/under value" of a specific pick I have performed a simple linear regression on all stats for all players year by year. The expected value of the pick is on the orange line and the true value is the players actual statistical value.



- To find the players that outperformed their pick slot, I have given each player one point for every category they have overachieved their expected value (max 16 points = 16 statistical categories). This number was then normalized to 100 to compare with the expected value previously defined.

- The school that has over/under performed the most is Kentucky. Followed by, Iowa State.





c) Explain and present your findings with tables and visuals. What additional research areas would you focus on if given the opportunity to expand this study?

- Due to time constraints, I was unable to refine my approach. If given more time I would more accurately rank the different categories. For example, the weight of points is not equal to the weight of average minutes played.
- I would also like to dig deeper into removing the outliers from the linear regression. This would help find more players whose stats were not as high, but still had a large contribution compared to their expected contribution.