Jeremy Dang

ID: 131177953

Date: August 10, 2024

Final Report

Introduction

National Basketball League (NBA) players generally live a larger life than the general population with being much more involved with the public and matching massively larger salaries. The goal of this research was to see if NBA players have similar motivations, namely temperature, tax, and cost of living, to other occupations when choosing a new team in free agency or when traded. Some of the questions were aided by consulting others to see what are their motivations when changing jobs and moving cities, primarily when looking at taxes and cost of living. The cities of the previous and new teams were collected into a database, and the resulting annual average temperature for those cities, along with the tax rate and per capita income for those states were analyzed and compared.

Learning Experiences

Undergoing this project taught me that research is not an easy thing to do but that when you find the answers to your questions, it is a very satisfying feeling but that you can always try to learn more. I also got to see the evolution of the league as players got more bargaining power, decided where they wanted to go, and formed teams of their own choosing rather than having their contract traded, like in the formative years of the league, or that their next team found that

they fit there better, which was much more common later on. All data is on https://github.com/danceswithme/nba_analysis.

Major Tasks and Assignments

The primary tasks in this project were to develop the pool of players that fit my criteria, then determine which aspects to focus on for my analysis, pulling the data from various resources, analyzing them in notebooks and on maps, and drawing conclusions from those data.

I first utilized basketball-reference.com to pull the list of basketball players that fit my criteria, which was inserted below. With the full list, I then only had players who changed teams either via free agency or through trade and had chosen their next team. This was far easier to determine for free agency, which was obvious, and in trades, for those that happened more recently as a benefit of the 24/7 news cycle is that all parts of a player's trade requests were constantly recorded, especially if they were more well-known, so it was a lot easier to determine if they chose their next destination. For older players, most articles would only just say that they were traded, and usually it was a lack of fit on their old team, a better fit for their new team, or a combination of those two. This list of players was gradually culled until the final list was transferred onto MongoDB for better storage and usage with other programs.

Star Criteria	4 All Star Game Appearances	2 All-NBA	1 MVP	1 FMVP	3 DPOY	NBA 75 Team
Amount of Players	209	190	42 (DR J won both, 6 + Dr J ABA)	34	4	76

Next, the annual temperatures for the cities and the tax rate (converted into tiers) and per capita incomes for the states were pulled from various sources. Annual temperature was pulled from open-meteo.com, an open source weather website that allows for API pulls without

requiring keys. Tax rate came from Tax Foundation and statcan.gc.ca was used for Ontario's tax rate. The tax rate was converted into tiers of none, low, moderate, and high because small differences in percentages were unlikely to make a difference in a person's decision. The per capita income was pulled as comma separated values from the United States Regional Economic Analysis Project (REAP) and from canada.ca for Ontario.

After having all the data collected into MongoDB, I analyzed them in a variety of manners. I first did statistical analysis and other graphing in Jupyter Notebook and found that using probability, none of the three had a significant difference when a player changed teams of their own choice. This is where having the tax rate in tiers may have hurt my analysis as there were not enough samples or deltas for analysis while temperature had a p-value of 0.4848 and PCI had a p-value of 0.3747, although both were negative, stating that the previous city and state, respectively had a larger value than the new one. This is contrary to the common belief that players choose teams for warmer weather or lower taxes, but I did not include lesser players who may have chosen differently. I also created frequency graphs to look at the most common destinations, and found that in general, the most popular destinations were the historically greatest teams, Los Angeles Lakers, Boston Celtics, Philadelphia 76ers, but sometimes also fit what the general population liked; that players like cities that are lively and would be somewhere they would want to hang out when not playing games. I also binned sets of years into eras: starting to 1980 which was considered near the beginning of true free agency, and more importantly, was when Magic Johnson and Larry Bird helped save the league and grow its popularity massively, followed by 1999 which was when the league started to grow more worldwide as its superstars were no longer just Americans, then in 2015 which was the start of the player empowerment era, where players were able to force trades to a team of their own

choice, also creating superteams with players that they wanted to play with, up to the present day. For the most part, Los Angeles was the most popular destination for those eras, with the exception of 1980-1999 for which there were no massively popular destinations and actually Atlanta was the most popular but only had 7 while other cities had 6. The era before had limitations as players usually had their contracts sold to the next team so mainly players went to the best teams of that era, the Lakers, Celtics, and 76ers. Los Angeles became much more popular after 2015, which also may have to do with the fact that the Los Angeles Clippers became a much more successful and popular team, especially after Steve Ballmer purchased the team and invested much of his own money into making them more than just the little brother of the Lakers. I also mapped the data in QGIS and the destination popularity bubble showed more of the same while the temperature mapping mainly just showed that a surprisingly popular destination early on was the San Diego Clippers prior to them moving to Los Angeles in the late 70s. I think the temperature not having much of an impact had to do with some of the greatest franchises being in relatively cold weather, the Boston Celtics, Chicago Bulls, and Philadelphia 76ers, to name some. I think it is because if players wanted warm weather they could do so during the summer in the offseason and would rather be on teams that give them the best chance for personal glory, be it statistics, fame, endorsements, or team success, ultimately winning the championship.

Research Challenges

The biggest challenges I faced were teaching myself how to pull data using API requests, using a MongoDB rather than relying on recording everything into Excel, general Python and Jupyter lab programs, and figuring out how to put everything together. Another challenge was in consistency of data as I made a multitude of mistakes including not using the true per capita

income, which REAP included and had been adjusted for inflation, which would've made analysis between generations easier, and also didn't account for the conversion rate between currencies. Also using just states and not cities meant the analysis did not always make sense, especially when there were several basketball teams within the same state and it was not too uncommon that a player only changed cities but not states. I was not able to conclude much from my research but that was from other limiting factors, as I wanted to limit my analysis to the non-basketball factors in a city when there are many others that could've been included but I did not from ethical or other reasons, those are proximity to home, desired teammates or coaches, or contract, amongst others. I wanted to respect data privacy by only pulling from public information, keep my analysis mostly objective, and focus on the non-basketball reasons that a player could change teams.

Conclusion

Going through this project was time-consuming but did show me the benefits of challenging oneself and being placed in a situation where one feels "dumb" in. This doesn't mean that the research is above your head and you don't get anything out of it, it means you are delving deeper into something that you don't know well but would like to know more of and it can help you grow as you then can't always rely on an outside source. Additionally, I came out of this wanting to grow my skills more in the data analytics sphere, along with trying to improve my project in the future by collecting data from more than just star players and also cleaning up my data and making sure everything is good and would be easily readable and usable by another person.