Python A | YSP 2023

Shawn Ma, Shiv Patel

June 14th, 2023

NBA Project Code Analysis

In [1]

* Import our packages necessary for data collection and analysis

In [2]

* Import and transform CSVs of NBA Playoff Statistics into data frames while cherry-picking specific statistics we plan to look at

In [3]

* Transform the data frame information about all 8 statistics (name, mpg, gp, ppg, apg, rpg, 3pa, drtg) into lists that represent every year. This makes the data much easier to handle and cross reference to compare different years

In [4]

* Clean the lists of data to only include players who have been in the NBA Playoffs for the 2019-2020 season and at least 1 additional season in any year from 2020-2023. Then, track those player’s specific statistics in new independent stat lists as for later analysis (done per year)

In [5]

* Create a function called “checksig” that can take any statistic of list and compare it to the control data to determine whether the data/statistical references are statistically significant or not. This lets us know where the differences of whatever statistic we are comparing to the COVID season is truly an anomaly

In [6]

* Create a function called “findDiff” that can return the difference of any 2 given data sets (different years) in whatever statistic we will test.

In [7]

* Create a function called “makeHist” that can create/print a histogram given any sample data (in Dataframe), statistic, and column. This makes it very easy to quicky visualize trends between statistics to identify trends from a birds eye view.

In [8]

* Create a function called “scatterplot” that takes various values to create a visually pleasing scatterplot as well as a trend line/line of best fit as another way to visualize trends between statistics of different years

In [9]

* Create a function called “findPlayer” that sorts players based off a dataset with various statistics in increasing improvement to identify star players

In [10]

* Create a function called “findConsis” that finds the most consistent player for any statistic in any year.

In [11-20]

* These are various uses of the functions I just listed to cover almost every combination of years, statistics, and players so we can begin to identify major trends within the COVID and non-COVID years.