Overview

For this assignment you will be working with a [dataset of NBA basketball data]().
You should use Python to process this data, calculate values, and produce graphs. Then, prepare a PDF document to report on your findings and upload it to I-Learn. Rather than submitting a python program, you will have a PDF that shows snippets of code, followed by the resulting graphs, or discussion of what you discovered.

Because the focus of this assignment is the result of your work, rather than creating a general purpose program, your code is not required to be as elegant as in other assignments. But keep in mind that using good style will help you keep things organized.

In addition, you are welcome to use the interactive console for elements of this if you prefer, rather than having to put all of the code in a Python program to submit in the end. But keep in mind that at the end, you'll need to put the code into your PDF that you used for the various elements of your assignment.

TOOLS

Please note that there are *many* ways to accomplish these tasks. You will need to search for examples and solutions online. *This is by design*. Part of learning how to work with data science libraries is learning how to discover and sort through various options. As you do so, try to keep an eye out for best practices, etc. For example, you shouldn't iterate through each row of the Pandas dataframe to compute an average manually. It has a built-in function to do this.
There are lots of great visualization libraries out there. You are welcome to use any that you would like ([here is an overview]() of some of the most popular ones). As a place to start, I would recommend using Seaborn. It is relatively straightforward and

can produce professional looking graphs without too much effort.

Please walk through [this Pandas tutorial](#), as it will walk step by step through the process of many of the elements you need to complete for this assignment.

Requirements

PART I - SPECIFIC ASSIGNMENTS

These requirements are intended to be straightforward in the expectations. It may still require research on your part to learn how to accomplish each one, but their shouldn't be ambiguity in the task itself. Please complete each one and report on it in your PDF document.

1. Calculate the mean and median number of points scored. (In other words, each row is the amount of points a player scored during a particular season. Calculate the median of these values. The result of this is that we have the median number of points players score each season.)

2. Determine the highest number of points recorded in a single season. Identify who scored those points and the year they did so.

3. Produce a boxplot that shows the distribution of total points, total assists, and total rebounds (each of these three is a separate box plot, but they can be on the same scale and in the same graphic).

4. Produce a plot that shows how the number of points scored has changed over time by showing the median of points scored per year, over time. The x-axis is the year and the y-axis is the median number of points among all players for that year.

PART II - COME UP WITH SUPPORTING EVIDENCE

These requirements still ask you to address a specific question, but require you to exercise a little bit of creativity in exactly what to produce.

1. Some players score a lot of points because they attempt a lot of shots. Among players that have scored a lot of points, are there

some that are much more efficient (points per attempt) than others?

2. It seems like some players may excel in one statistical category, but produce very little in other areas. Are there any players that are exceptional across many categories?

3. Much has been said about the rise of the three-point shot in recent years. It seems that players are shooting and making more three-point shots than ever. Recognizing that this dataset doesn't contain the very most recent data, do you see a trend of more three-point shots either across the league or among certain groups of players? Is there a point at which popularity increased dramatically?

PART III - SHOW CREATIVITY

These requirements expect you to come up with your own insight.

1. Many sports analysts argue about which player is the GOAT (the Greatest Of All Time). Based on this data, who would you say is the GOAT? Provide evidence to back up your decision.

2. The biographical data in this dataset contains information about home towns, home states, and home countries for these players. Can you find anything interesting about players who came from a similar location?

3. Find something else in this dataset that you consider interesting. Produce a graph to communicate your insight.

Submission

Upload to I-Learn a PDF file with your findings and the code used to produce your results.