Problem set 3

S670 Fall 2019

Upload a HTMLor PDF document with your code, graphs, and write-up to Canvas by 11:59 pm, Thursday 19th September.

Four variables in the data set `NBA2017-18.csv` are:

- `PTS`: points
- `TRB`: rebounds
- `MP`: minutes played
- `Age`: player's age in years.

Among other things, we might wish to study whether players who score a lot of points better or worse at rebounding than players who don't score many points.

1. Plot rebounds as a function of points, and add a regression line. Is there a strong relationship?
2. Plot *rebounds per minute played* as a function of *points per minute played*. Is there a strong relationship?
3. Can you find a way of showing (graphically or mathematically) that there really is a weak but meaningful relationship between points per minute played and rebounds per minute played? Is this relationship positive or negative?

Now let's look at the relationship between age and minutes played.

4. Plot minutes played as a function of age, and add TWO different curves (not lines.) Describe, in words, how you think the average minutes played among NBA players changes with age.
5. Plot minutes played as a function of age, and this time add a regression line. What does this line tell you?