For my project, I'm analyzing how an NBA player's age affects their production on the court and how that differs by era/season. Even though a player's productivity can be measured with many variables, I'm interested in the player's average points per game just to keep it simple. I collected the data from a website called https://www.basketball-reference.com The site has all the stats from every player from any season in the NBA. After converting the CSV into a data frame, I cleaned up the data by removing rows that contain the column's name, removing players that average less than 10 minutes per game, adding a year column, and changing column names to make it easy for the reader to read. The data frame now contains only the things that will impact my visualization like the player's name, Position (The position they play in a game), current age, average minutes per game, and average points per game.

The visualization that I currently have is a strip plot made with plotly express. The strip plot is comparing age and average points per game. The points per game look sort of normally distributed.

From the topic proposal, the only thing I changed was the type of visualization I was going to use. Before I wanted to use a bar chart but now I decided on a strip plot because it can show a player's info when you hover over points.

I will keep the same strip plot but I will add more things to make the data look more readable. For example, I will try to differentiate the colors of data points and add a slider that will change the visualization look based on the year that the slider is set to. For example, if the slider is set to 1985-86 it will only show players that played in the season. A slide show might be used as well.