The Data

The dataset I'm using is NBA (professional basketball) player stats from the 2020-21 regular season. The data is taken from ESPN. Only players who played in at least half of the games (41 games) and whose average play time was at least a third of a game (20 minutes) for that season are included in the dataset. There are 218 players that meet that criteria. There are 23 variables in the dataset, of which 12 will be used. They are Player Name, Team, Position, and the rest are key stats that we will look at.

The Goal

The goal of this project is to see which positions (PG, SG, SF, PF, C) are better at which stats. The stats are Points, Shooting %, 3 Point %, Free Throw %, Rebounds, Assists, Steals, Blocks, and Turnovers. I want to find out if certain positions are overall better at certain stats, or if there is an even/random distribution with some stats.

Tasks

The tasks for this project are as follows:
Allow NBA fans to explore player stats
Compare positions in one particular stat
Compare 2 stats to each other while also showing position
Find trends with particular stats
Allow user to select and explore desired/interested stats

Link to Visual

https://github.com/Mike-Prodo/NBA-Visualization

Key Elements

The key elements used in the visualization are as follows:
Histogram: compares two stats to each other
Dropdown selection for x and y variable: allows the user to select the stats they want
Position type as color variable: allows the user to see which positions are better at which stats
Filter position by clicking on position in legend: makes it easier to see how selected positions perform
Hover over data point to see information (tooltip): allows the user to see specific information (Player, Team, Position) of an individual data point
Pan and zoom: enables the user to look more closely at the data

Evaluation Approach

My target population for this project is people who like the NBA and want to break down player stats. The people recruited for the evaluation process were friends who watch and follow the NBA.
The evaluation procedure technique I used was thinkalouds. I wanted the people who evaluated my visualization to record their thoughts as they were interacting with it. I wanted to get their immediate feedback.
The overall feedback I received from them was positive. In general, they really liked the visualization. They talked about specific elements of the visualization they liked and elements they didn't enjoy or were confused at which I will discuss further in the next section.

Synthesis of Findings

Visualizing the data in a histogram worked well. Most of my evaluators liked that they could select the two variables they wanted, as well as select the same variable for the x and y axis to look at just one variable. Everyone liked that they could hover over a data point and see the individual player information. Being able to click on a particular position in the legend and see that position in the histogram was very helpful. People liked that they could zoom further in and see the data points more clearly. However, some feedback I received was that there were a lot of data points and in some situations it looked a bit cluttered and was difficult to see trends. There were 218 data points so in the future one thing I could change is trim the data down to have 20 players for each position to get 100 players total. Then there is less clutter, and it would be easier to see individual points.