Positional NBA FG% Comparison

1. The standard deviation of field goal percentage is 0.04. The coefficient of variation is 0.093, or 9.3%.
2. Based on the histogram that includes all players, the data is bimodal. When all NBA players in the data set are considered, there are two clear peaks in the histogram. This suggests that there are multiple subclasses in the entire data set. When histograms are created for each position, it is clear that the positional data is more unimodal. The histograms for guards, power forwards, and small forwards all contain just one obvious peak in the data. For centers, there are two or three peaks, which suggests that there may be even more subclasses for centers. When all players are included in the data, the distribution is right-skewed. The right tail is longer than the left, and the histogram shows that a solid amount of the data is bunched up on the left. This means the mode is less than the median, which is then less than the mean.
3. Player 0053 had the highest FG% relative to their position, shooting 61.5% from the field. This tops the average FG% for centers (52.3%) by 9.2%. The most efficient guard was Player 0068, who shot 51.9% from the field. This beats the average FG% for guards (44.7%) by 7.2%. The most efficient power forward was Player 0022, who shot 54.3% from the field. This tops the average FG% for power forwards (49.5%) by 4.8%. The most efficient small forward was Player 0001, who shot 52% from the field. This tops the average FG% for small forwards (45.9%) by 6.1%.
4. A great way to compare scoring efficiency across different positions is to consider the shot selection of each player. Obviously, not all shots are created equally, meaning some shots are more difficult than others. It might be best to include 3-point % as a variable to consider when discussing scoring efficiency. 3-pointers are obviously more difficult than most 2-pointers, and smaller players (guards) typically take more 3-pointers than bigger players (centers). Additionally, it could be beneficial to consider the shooting zones on the basketball court. For example, a 3-point shot from the far corner of the court is tougher than a 3-point shot from the top of the key. There are many ways to make an even better scoring efficiency comparison, as these are just a few ideas.

Entropy in the NFL

1. The Buffalo Bills had the most variation in their run/pass play calling on 3rd and 4th and 2 in the NFL in 2024. The Bills' standardized entropy was 99.9%, which is near perfect variation in run/pass play calling in this scenario. The Los Angeles Rams were very close to this mark.
2. The team with the least variation was the Minnesota Vikings. The standardized entropy for the Vikings actually can't be calculated in Excel because the software returns 'undefined' for the natural log of 0. This 0 represents the percentage of running plays that the Vikings ran in these scenarios. This obviously means they did not do any of these run plays, and instead they opted to always pass on 3rd and 4th and 2 in 2024. This means there is no variation at all for the Vikings.
3. From the scatter plot, there is a slight positive relationship between standardized entropy and first down conversion percentage. The equation for the line of best fit from the simple linear regression is $y = 0.1029x + 0.4609$. This means that an increase by 1% in standardized entropy for 3rd and 4th and 2 is predicted to improve first down conversion rate by .1029%. This is obviously not a huge number, and the r-squared for this regression is just 0.0698. These stats indicate that there is a slight positive trend, but it is not that strong.
4. There's a lot of information needed to determine how important entropy is for success rate. First, I would want to see the context of each 3rd and 4th and 2 scenario that this PFF data set provides. Where on the field are these plays happening? Is the game still in reach for both

teams? What time during the game are these plays happening? Are both teams fully healthy during these plays? Next, I would want to know more about the specific formation and play design that the offense is running. There is an obvious scenario (Eagles QB sneak) where the strength of a specific play outweighs the importance of entropy. For teams that do have high standardized entropy and high success rate, I would see how strong their offense is at both passing and running the ball. If the offense is strong at both, then the unpredictability of the high entropy is not necessarily the reason for the high success rate. All of these points give more of an idea as to whether or not entropy is the reason for success rate versus some other factors instead.

5. To assess the optimal entropy for a given team, I would first look at the strengths and weaknesses of the current roster. If a team had a good short-yardage running back and an offensive line that could run block well, then I would consider having a more run-heavy approach, which would decrease standardized entropy. This could be a team like the Baltimore Ravens or the Philadelphia Eagles for last year. If my team had a strong receiving core and an accurate quarterback, I would consider having a more pass-heavy approach, which would also decrease standardized entropy. This could be a team like the Cincinnati Bengals or the Dallas Cowboys for this year. If my offense was balanced in terms of roster construction, I would maintain a balanced run/pass approach, which would lead to a high standardized entropy. Past my team's roster, I would analyze the strengths and weaknesses of my opponent's defense as well. If my opponent's defensive line and linebacker core struggled to stop the run, I would consider having a run-heavy approach. If my opponent's secondary struggled in short-yardage passing, I would consider having a pass-heavy approach in this scenario. If I was facing a balanced defense, I would play to my offense's strengths or maintain a balanced play calling plan. Roster construction is just one factor I would consider, as there are many others.