Sprint Assignment 2

By: Nathan Ensley

1. 360-Degree View

Creating a 360-degree view of a problem means covering every aspect of the situation. When making a model, this means assembling unique data points that capture every part of the given situation. For example, the SEC conducted a survey assessing the gameday experience of football fans. For their survey, they included questions spanning every stage of a gameday, from when fans travel to the stadium to when they leave and head home. The survey also includes many questions about their experience inside the stadium, including their experience with staff, sign markings, stadium appearance, and seat quality. This wide array of questions helped the SEC create a dataset it could use to improve the fan experience.

Many have attempted to predict the March Madness tournament, where seemingly impossible upsets occur every year. In improving his predictions, particularly in identifying 2019 March Madness upsets, Brendon Bastean collected a set of 16 quantitative factors that he used to build his prediction model. These factors encompass every phase of a basketball game, including shooting, rebounding, defense, efficiency, and tempo. Bastean supplements these statistics with overall team performance, measuring the team's winning percentage and its strength of schedule. These game and performance metrics effectively give a 360-degree view of teams as they enter the March Madness tournament, improving predictions.

42 Analytics Video: SSAC25: Behind the Deal: The Analytics of Sports Sponsorship Valuation

In this video, the speaker explains how sports teams tackle the problem of valuing sponsorships. One sponsorship value tool is known as media equivalency. This metric measures the media exposure a stadium sponsor spot generates from being displayed inside the stadium and in the background of broadcast footage or photos in newspapers. The variables collected to calculate this media equivalency include the number of exposures, the number of impressions from people watching the game, the equivalent cost for other marketing, the exposure length, and the exposure quality of that given location. For example, if the sponsor spot is displayed in the corner of the broadcast, the exposure quality would be less than a sponsor spot in the screen's center.

For prospective sponsors, they examine even more variables when determining whether a sports sponsorship is worth the investment over marketing alternatives. In addition to media equivalency and overall impressions, sponsors look at the strength of that team's brand, their audience overlap with that team, and the expected impact on their business based on past ROI determinations. All of these variables help a sponsor capture a 360-degree view of a sports sponsorship beyond just the sponsorship cost and game attendance.

2. Making a Metric

Punters are commonly assessed based on their average punting yards; however, this does not account for other important attributes of punters, particularly the need to precisely place a punt deep inside their opponents' territory and prevent a significant return.

A. My composite metric is the Punter Rating, a rating of individual punters measuring their short-range and long-range punting ability.

Punter Rating = $100 \ * \ (0.4 \ * \ PPP \ (Yard \ Line \ \leq \ 60) \ + \ 0.6 \ * \ LPP \ (Yard \ Line \ > \ 61))$

The two individual parameters that make up punter rating, the Precision Punting Parameter (PPP) and the Long Punting Parameter (LPP), are indexed by the average percentage of plays in which a team takes a punt matching one of those scenarios, using historical NFL play data from 1999 - 2025. For example, the Precision Punting Parameter is multiplied by 0.4 because about 40% of the punts in that time period occurred 60 yards or closer to the team's opposing end zone.

Multiplying these decimal parameter values by 100 improves the readability of the metric and establishes 100 as an ideal rating, where a punter achieves 100% optimal outcomes in both short and long-range punting. However, punter rating can be higher than 100 and maxes out at 132 for especially optimal scenarios, as covered in the parameters shown below.

Precision Punting Parameter (PPP) =

$$\frac{\#\ of\ Punts\ Fair\ Caught,\ Downed,\ or\ No\ Return\ Inside\ 20 + 0.5 *\ \#\ of\ Punts\ Inside\ 5 - \#\ of\ 10{+}\ Yard\ Returns}{\#\ of\ Punts \leq 60\ Yard\ Line}$$

When a punt enters the endzone, the punt is considered a touchback, and the receiving team receives the ball at the 20-yard line, even though the punt traveled over 20 yards further. Thus, it is important for a punter to strategically place their punt the closer they are to their opponents' end zone. An ideal punt within 60 yards of an opponent's end zone would land inside the 20-yard line, as it would improve upon a full-legged kick into the end zone. Additionally, since punters are not prioritizing distance on precision punting attempts, they can increase the ball's hang time to neutralize the returner, preventing any potentially dangerous return from taking place. Thus, the ideal outcome for a punter would be a ball that ends inside the 20-yard line and has no return, commonly from the ball being downed or fair caught, which produces the following ideal outcome formula:

Ideal Outcome Formula = $\frac{\#\ of\ Punts\ Fair\ Caught,\ Downed,\ or\ No\ Return\ Inside\ 20}{\#\ of\ Punts \leq 60\ Yard\ Line}$

However, this formula, which measures the percentage of ideal outcomes, fails to weight the significant game importance of punts inside the 5-yard line or returns that are greater than 10 yards. As seen in Figure 1 below, the receiving team has negative points expected on first down plays that start within the 10-yard line as negative drive outcomes, like safeties and punts, increase in likelihood. This effect is magnified the closer the drive starts to the 1-yard line, where the team expects -0.7 points on that play. Thus, the PPP increases the weight for punts inside the 5-yard line by 50%.

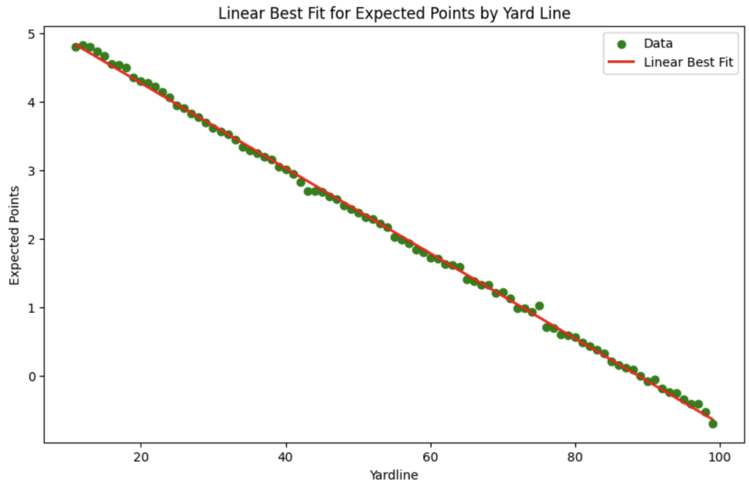

Figure 1. Expected Points by Yard Line (on 1st Down)

Additionally, a significant return of 10 yards or more increases the returning team's EPA by over 0.6 points. Punters can minimize long returns in precision punting situations by increasing the ball's hang time, allowing coverage teams to cover the punt sufficiently. If hang time is not long enough, returns can occur, as returners can safely receive the ball and maneuver around the opposition. Thus, a long return is penalized against the punter in PPP, as their punt's hang time can directly neutralize a returner's ability to return a punt.

Long Punting Parameter (LPP) = $\frac{\#\ of\ 48+\ Yard\ Punts\ +\ \#\ of\ 41+\ Net\ Yard\ Punts + 0.2*(60+\ Yard\ Punts)}{5\ *\ \#\ of\ Punts\ >=\ 60\ Yard\ Line}$

The other scenario punters enter is long-range punting. In this case, punters are pinned back deep in their zone and need to flip the field. Thus, contrary to precision punting, yardage is more important in long-range punting situations. As shown in Figure 2, the average yards per punt on a punt greater than 60 yards away from the endzone remains at a relatively stagnant 47.5 yards regardless of where that punt takes place. Thus, a 48+ yard punt would be above the expected yards that the punt would go. In the same manner, net punt yardage, which removes return yards from kicking yards, also remains relatively stable at 41 yards in long-punting situations.

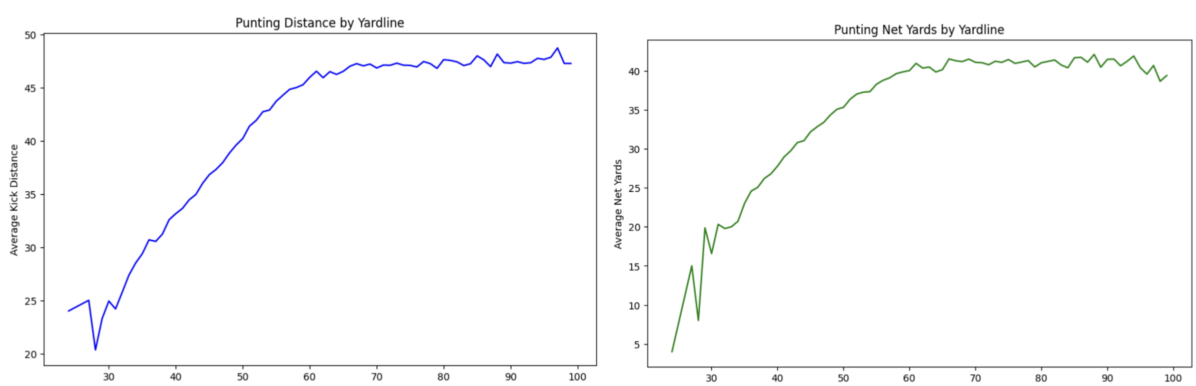

Figure 2. Average Kicking Distance by Yardline; Average Punting Net Yards by Yardline

Thus, the LPP helps to summarize the percentage of outcomes that are better than expected for punters in long punting situations. While a punter may not kick a punt over 47 yards, they may be increasing hang time to prevent the returner from getting a significant return. That's why a shorter punt may result in an ideal outcome for this parameter. Finally, a 20% boost is applied to punts longer than 60 yards, as these especially long punts do the job of significantly flipping the field, even if a returner may be able to return it.

C. The metric generally ranks the best punters of all time at the top. For example, the top two-ranked punters using Punter Rating are AJ Cole and Johnny Hecker. Cole has two all-pro selections and three pro bowl selections in his six-season career, while Hecker has six all-pro selections and four pro bowl selections in his thirteen-season career. Both All-Pro Selections and Pro Bowl Selections generally indicate a player is at the top of their respective position. Hecker was even named to the 2010s All-Decade team. All of the top five in Punter Rating have received at least one All-Pro selection in their careers, with Cameron Johnson at 6th being the highest-rated punter without an All-Pro selection.

	player_name	punter_rating
0	A.Cole	85.096347
1	J.Hekker	81.926962
2	M.Dickson	80.689300
3	L.Cooke	79.293588
4	T.Way	79.103846

Figure 3. Punter Rating by Player (Min. 250 Punts)

The Punter Rating generally favors the long-punting statistic, as expected, with eight of the top ten in LPP being in the top ten of Punter Rating. However, the Punter Rating still highlights precision punters, with Mitch Wishnowsky, the highest punter in PPP, being the 12th-highest-rated punter overall, even though his career average is just 45.6 yards, 3 yards shy of the league high of 48.6 yards. One significant surprise in the rating is that Shane Lechler, generally regarded as the best punter ever, with nine all-pro selections in his eighteen-year career, is rated just 32nd using the Punter Rating. Lechler does not even show up in the top ten of either LPP or PPP, with his precision punting significantly impacting his overall rating.

D. When charting punters with over 250 tracked punts, the punter rating has a positive correlation with career average punt yards. The highest punters in average yards score well with this statistic, as shown in Figure 4.

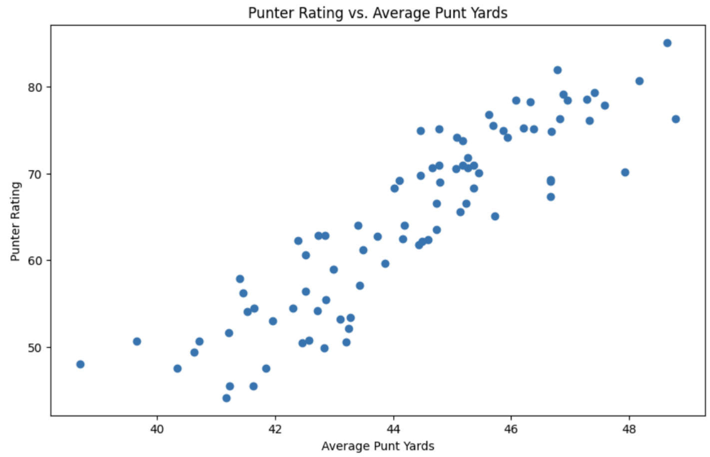

Figure 4. Punter Rating vs. Average Punt Yards

When looking at punters with the highest percentage of punts inside the 5-yard line, there appears to be no correlation with punter rating (Figure 5). This indicates that precision punting appears to have much less significance on punter rating than punt yardage does.

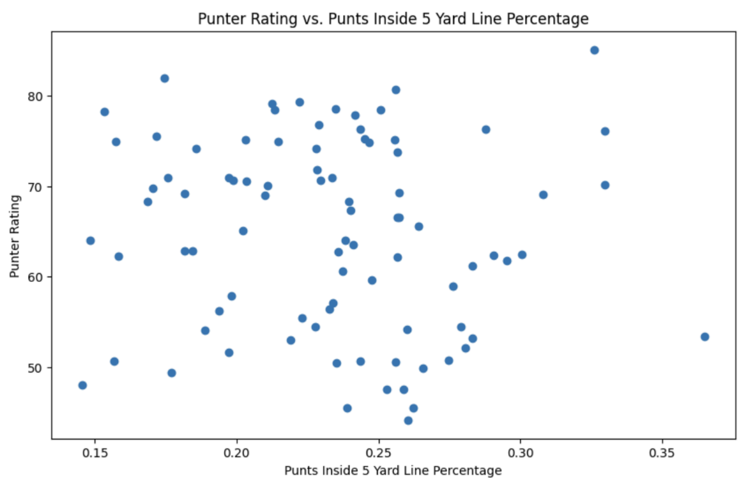

Figure 5. Punter Rating vs. Punts Inside 5 Yard Line %

Now, taking a look at some team-based statistics from the 2024 season, I compared how their primary punters' punter rating compared to different performance outcomes. For example, the opposing team's points seemed to have a decreasing trend as a team's punter rating increased.

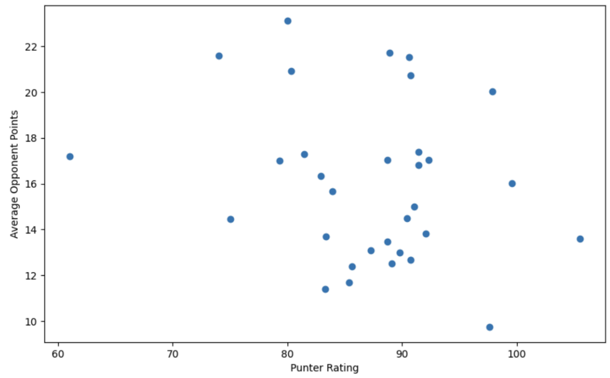

Figure 6. Punter Rating vs. Average Opposing Points (2024 Season)

There does not appear to be a significant correlation between punter rating and winning %. While all of the best teams have above-average punters according to Punter Rating, the trend is not consistent across all winning percentages.

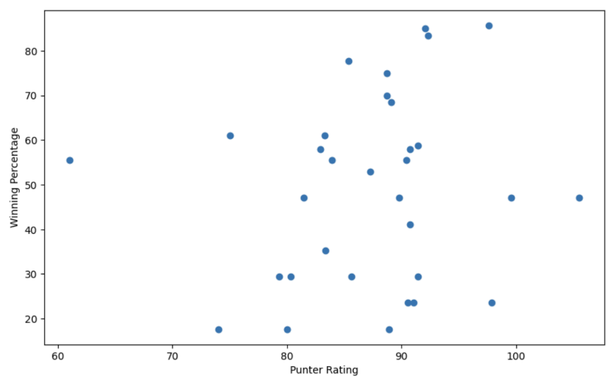

Figure 7. Punter Rating vs Winning % (2024 Season)

Finally, while Punter Rating attempts to capture the best precision punters, punter rating does not have a correlation with the number of safeties scored. A safety occurs so rarely that even a punter consistently punting within the 5-yard line may never force the opposing team into a safety.

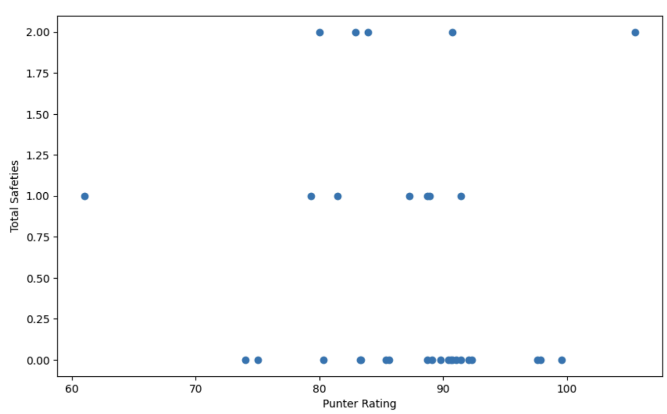

Figure 8. Punter Rating vs Total Safeties (2024 Season)

3. Visualization Analysis

Figure 9 attempts to visualize where different drivers are gaining or losing time on track against other drivers in qualifying. This graphic does include a lengthy caption, helping to provide specific context to the audience in helping them understand the graphic. Additionally, color is used minimally, only to highlight the two main drivers for comparison. Other drivers' times are sorted to the back by decreasing that line's capacity. With the captions on the lines themselves, the viewer can directly understand what caused the rapid time difference in a particular sector of the track, producing that "Aha" moment that Knaflic emphasizes. While it's interesting to see the other three drivers' times, it may be preferable to remove those lines on the graphic, with Knaflic emphasizing to "remove the non-essential." Finally, without a label on the yellow, blue, and white lines, it is difficult for the audience to know what driver those lines represent without looking at the caption above. Thus, to improve accessibility, having another label next to those colored lines may be preferable.

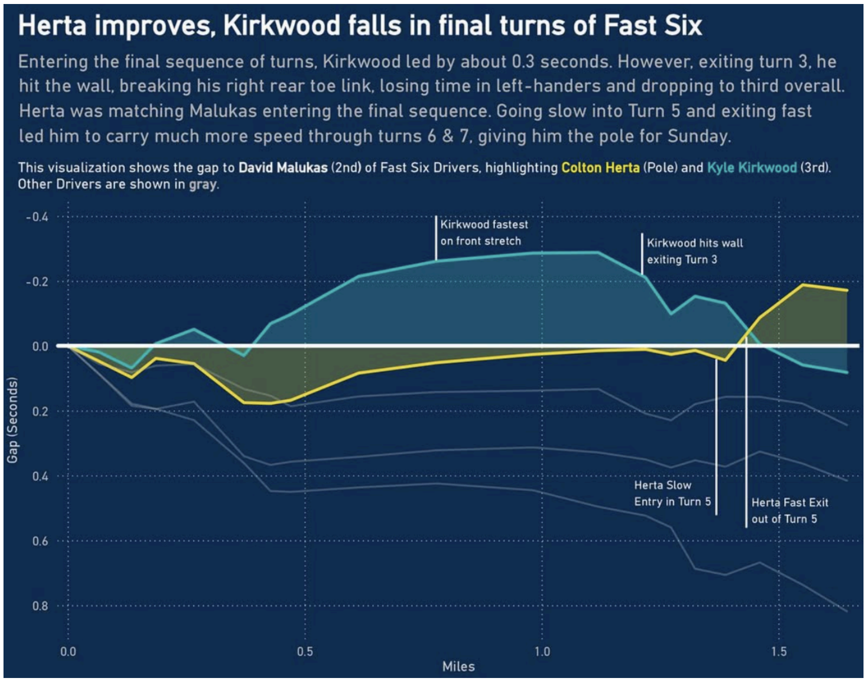

Figure 9. Fast Six Time Gaps over Lap

Figure 10 is an example of a visualization is a graph that falls short of the principles that Knaflic laid out. For one, the graphic includes excessive labels for some of the teams on the graphic. This overcomplicates the graphic and makes it difficult to know where to look. Additionally, the graphic clutters the view by including labels in each of the quadrants that don't add pertinent information for understanding the graphic. Words are important for accessibility, said Knaflic. In this example, there is no title and small axis labels, making it inaccessible. It is also very difficult to know where to specifically look. There are teams dotted everywhere on the graphic, making it difficult for the reader to know what is important and demonstrating an "aha" moment.

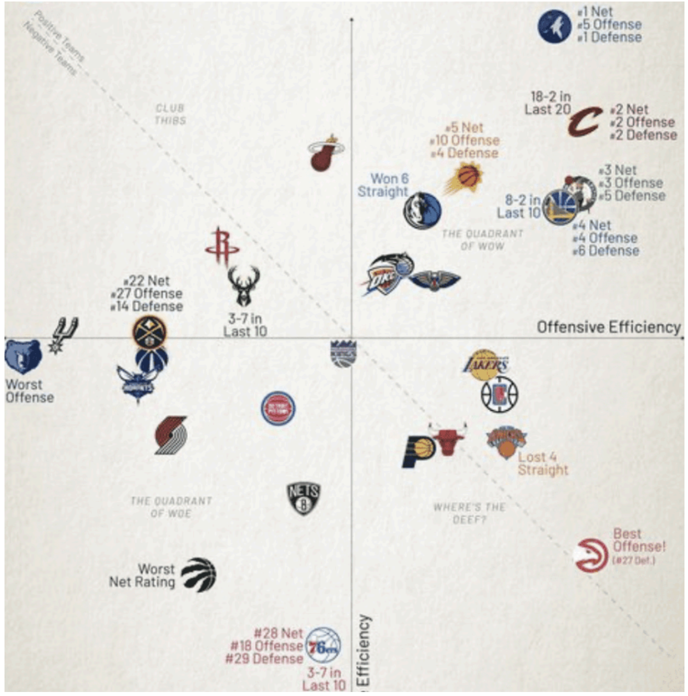

Figure 10. Offensive Efficiency vs. Defensive Efficiency (NBA)

4. Visualizations

For the IU Indianapolis women's basketball team, I used practice data to show how their players were performing in terms of shot quality and shot accuracy. The X Axis shows Shot Rubric Average while the Y Axis shows Effective Field Goal %. Since the audience, the coaches, would not understand the meaning of Shot Rubric Average, I changed that axis to Shot Quality in the final plot. While Knafflic recommended simplifying the graph and not using color, I used color to indicate what position each player played. This helps viewers as they can now tell that Kamara, Neveah, and Kaylin way outperform the other guards in shot accuracy.

Overall, this graph helps indicate who is taking good shots and how well they are capitalizing on those opportunities. The further right on the graph, the better shot quality that player has. The further up on the graph, the higher that player's shooting percentage is. This graph helps the team quickly identify positive and negative outliers. For example, Kamara doesn't take great shots, but she can make these difficult shots. Also, Sydney takes great shots, but underperforms her expected shooting percentage for those shots.

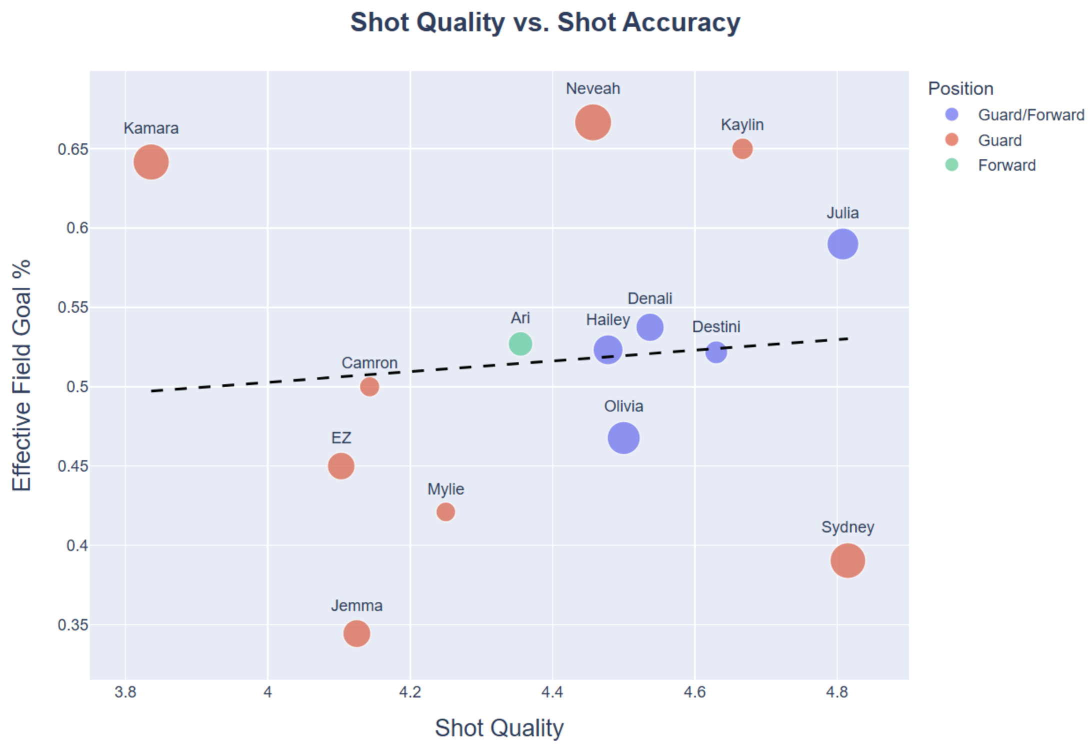

Figure 11. Shot Quality vs. Shot Accuracy (for IU Indy Basketball)

The next plot I produced for the IU Indianapolis coaches shows assists vs. turnovers by player. The coaches' goal was to improve their assist-turnover rate from last year, which was negative for every single player. In this plot, I used color to indicate what areas were ideal. It's good for players to have more assists than turnovers. Thus, I used green for that region while using red for players who had more turnovers than assists. The dotted line indicates a perfectly equal assist-turnover rate, which would be a neutral outcome. I also did not use color for position as before, since there is already a lot of color present in this diagram.

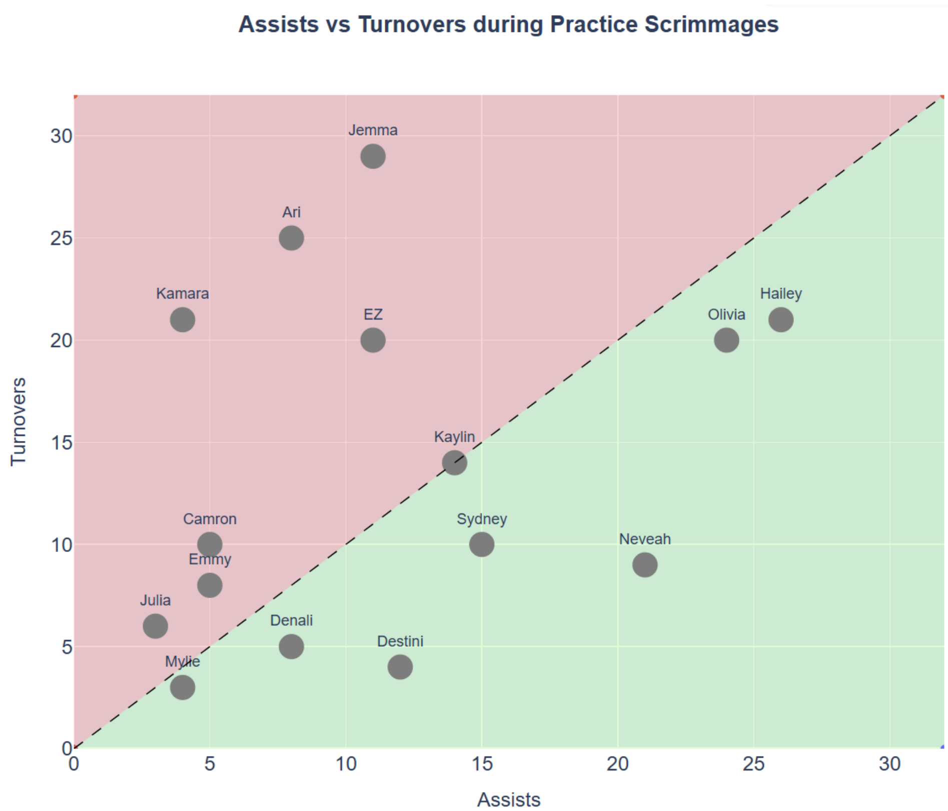

Figure 12. Assists vs. Turnover in Practice