STOR 496 Final – UNC Women's Basketball Analytics

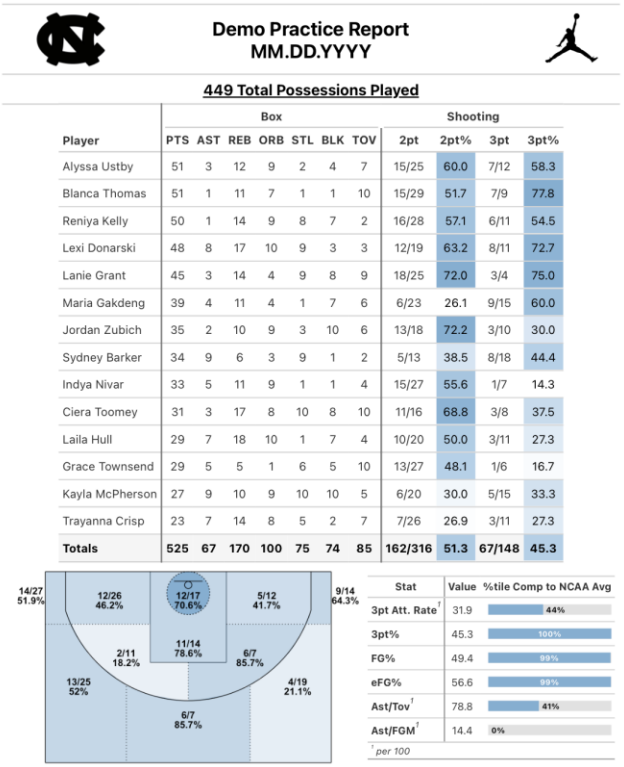
Gregory Lederer

Introduction

In this paper, I will discuss my role and some projects I did as a basketball analytics assistant with the UNC Women's Basketball team. Throughout the semester, I worked on a variety of projects, including developing a practice website to track player progress and team efficiency, building analytical scouting reports, and creating tools to evaluate our offensive sets and lineup combinations. Each of these tasks was designed to provide the coaching staff with valuable insights and data-driven recommendations to enhance team performance. The paper will discuss these projects and show some examples, highlighting how they contributed to tracking individual and team progress, optimizing game strategies, and supporting coaching decisions.

Practice Website

One of the initial main priorities I had was to develop and manage a practice website to help track player progression and team efficiency. I would be at each practice and track stats on two different macro-enabled excel sheets, one for shot tracking, and one for box stats. I then ran these sheets in R code to upload to my Shiny app, where all stats can easily be accessed. The R code utilized numerous packages such as *tidyverse* and *dplyr* to process the data, and *gt* and *ggplot* to visualize it. The website was crucial for tracking development over time by both an individual and the team. You can explore a demo of the website here.



Demo Practice Report
MM.DD.YYYY

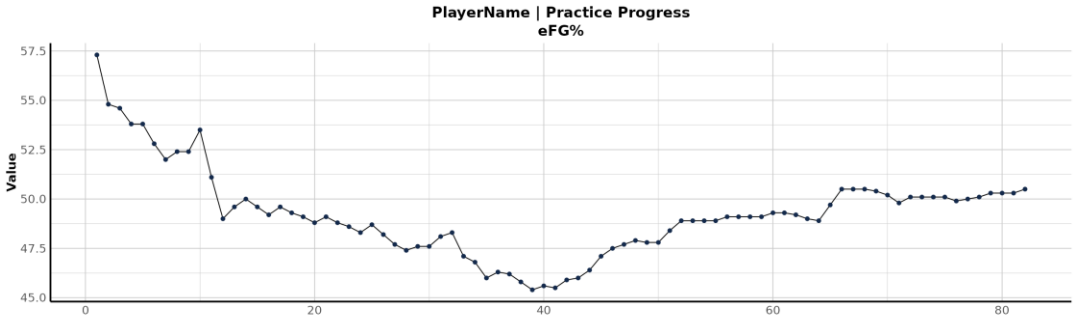
449 Total Possessions Played

Player	Box							Shooting			
	PTS	AST	REB	ORB	STL	BLK	TOV	2pt	2pt%	3pt	3pt%
Alyssa Ustby	51	3	12	9	2	4	7	15/25	60.0	7/12	58.3
Blanca Thomas	51	1	11	7	1	1	10	15/29	51.7	7/9	77.8
Reniya Kelly	50	1	14	9	8	7	2	16/28	57.1	6/11	54.5
Lexi Donarski	48	8	17	10	9	3	3	12/19	63.2	8/11	72.7
Lanie Grant	45	3	14	4	9	8	9	18/25	72.0	3/4	75.0
Maria Gakdeng	39	4	11	4	1	7	6	6/23	26.1	9/15	60.0
Jordan Zubich	35	2	10	9	3	10	6	13/18	72.2	3/10	30.0
Sydney Barker	34	9	6	3	9	1	2	5/13	38.5	8/18	44.4
Indya Nivar	33	5	11	9	1	1	4	15/27	55.6	1/7	14.3
Ciera Toomey	31	3	17	8	10	8	10	11/16	68.8	3/8	37.5
Laila Hull	29	7	18	10	1	7	4	10/20	50.0	3/11	27.3
Grace Townsend	29	5	5	1	6	5	10	13/27	48.1	1/6	16.7
Kayla McPherson	27	9	10	9	10	10	5	6/20	30.0	5/15	33.3
Trayanna Crisp	23	7	14	8	5	2	7	7/26	26.9	3/11	27.3
Totals	**525**	**67**	**170**	**100**	**75**	**74**	**85**	**162/316**	**51.3**	**67/148**	**45.3**

Stat	Value	%tile Comp to NCAA Avg
3pt Att. Rate†	31.9	44%
3pt%	45.3	100%
FG%	49.4	99%
eFG%	56.6	99%
Ast/Tov†	78.8	41%
Ast/FGM†	14.4	0%

† per 100

The practice report to the left is created after every practice and includes a box score and shooting statistics for each player. It's uploaded to the website for easy access, and I also compile weekly and monthly reports to compare performance over time. Coaches found it helpful to view both team and individual player shot charts, as it provided a clear breakdown of shooting efficiency and areas for improvement. The reports also feature the ability to compare team metrics to NCAA averages (seen as percentiles in bottom right visual.

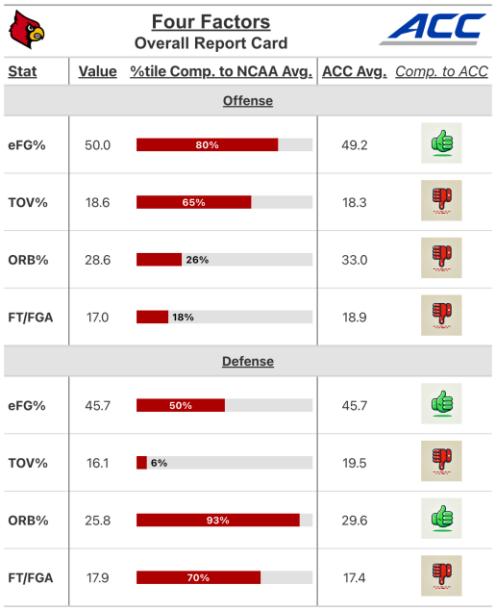


The line chart above tracks a player's eFG% over time, providing a simple way to see how their shooting performance changes throughout practices. It can also be used to track other metrics, like 3pt attempt rates, which gives a quick overview of shooting tendencies. These line charts are useful to show how a player is improving or what they need to continue to work more on. Coaches primarily looked at player box stats and team shot charts, with me bringing up sometimes how our rates are improving or getting worse, and how they compare to division 1 women's teams.

Scouting Reports

Another key aspect of my role was building detailed analytical scouting reports for almost every game. These reports were tailored to the specific preferences of the assistant coach

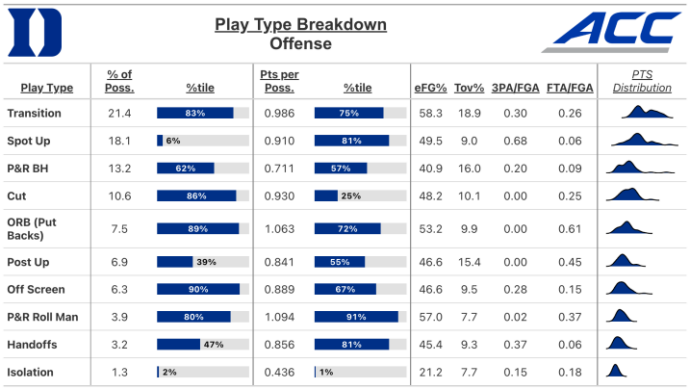
assigned to scout that opponent. I met with the coaches regularly to discuss their numbers and adjust the reports based on their needs. Each scouting report began with a summary of the opponent's team statistics, focusing on the four factors of basketball efficiency. From there, I provided a deeper dive into their offensive and defensive schemes, followed by an analysis of their top 4-5 players' tendencies. I also addressed specific questions the coaches had about the opponent, ensuring they had the insights they needed to strategize effectively. In addition, I prepared extra notes for each meeting to ensure coaches were fully equipped with all relevant information. You can view a sample scouting report here.



Four Factors
Overall Report Card

Stat	Value	%tile Comp. to NCAA Avg.	ACC Avg.	*Comp. to ACC*
Offense				
eFG%	50.0	80%	49.2	
TOV%	18.6	65%	18.3	
ORB%	28.6	26%	33.0	
FT/FGA	17.0	18%	18.9	
Defense				
eFG%	45.7	50%	45.7	
TOV%	16.1	6%	19.5	
ORB%	25.8	93%	29.6	
FT/FGA	17.9	70%	17.4	

The visual to the left is an example of a simple table shown in the reports. The data is seamlessly pulled from Sports Reference and compares the Four Factors of the team to the rest of the D1 women's teams, and the ACC average.



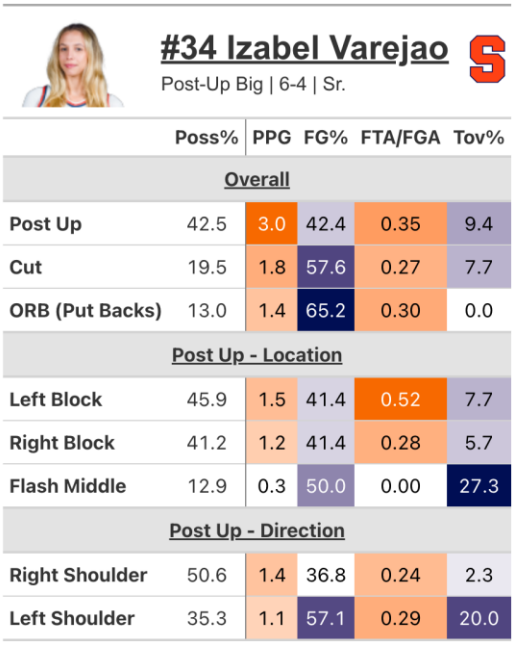
Play Type Breakdown
Offense

Play Type	% of Poss.	%tile	Pts per Poss.	%tile	eFG%	Tov%	3PA/FGA	FTA/FGA	*PTS Distribution*
Transition	21.4	83%	0.986	75%	58.3	18.9	0.30	0.26	
Spot Up	18.1	6%	0.910	81%	49.5	9.0	0.68	0.06	
P&R BH	13.2	62%	0.711	57%	40.9	16.0	0.20	0.09	
Cut	10.6	86%	0.930	25%	48.2	10.1	0.00	0.25	
ORB (Put Backs)	7.5	89%	1.063	72%	53.2	9.9	0.00	0.61	
Post Up	6.9	39%	0.841	55%	46.6	15.4	0.00	0.45	
Off Screen	6.3	90%	0.889	67%	46.6	9.5	0.28	0.15	
P&R Roll Man	3.9	80%	1.094	91%	57.0	7.7	0.02	0.37	
Handoffs	3.2	47%	0.856	81%	45.4	9.3	0.37	0.06	
Isolation	1.3	2%	0.436	1%	21.2	7.7	0.15	0.18	

The visual above and to the right (Duke's Play Type Breakdown), shows their offensive plays, uses of each, and efficiency compared to the league. This visual, used both for looking at their offense and defense, helps our coaches understand what they run the most and are best at. As you can see above, Duke is a fast-paced team operating their offense in transition more than 80% of other Division 1 teams, however they do not shoot as many spot up shots as most other teams while being very efficient though. As you can see on the right most column as well their scoring from spot up shots varies a lot, I then discuss that variation to see the effect it has on winning.

Presenting these tables took some explaining at first, especially the meaning of percentiles (the bars), and the scoring distribution. What helped the most was me just writing it down on a post-it and attaching it to the page. For example, for the table on Louisville I would write "Louisville's offensive eFG% is better than 80% of other division 1 women's

basketball teams." Early on I made sure they knew that for all bars it was better to be closer to 100 than 0, even for metrics like TOV% where having a high TOV% is bad. The same went for the play type breakdowns for the points per possession defensive metrics, as for percentage of possessions the closer the percentile was to 100, always meant more of the possessions they were seeing compared to other division 1 teams. The points distribution column they did not draw from much, that was more of for me to show inconsistencies in how much they score/allow from those play types. For example, I would say "As you can see their points in transition vary a lot, in wins they average 18 points per game, and in losses they average 12 points per game, never scoring 18 or more points in transition when they lose."



#34 Izabel Varejao

Post-Up Big | 6-4 | Sr.

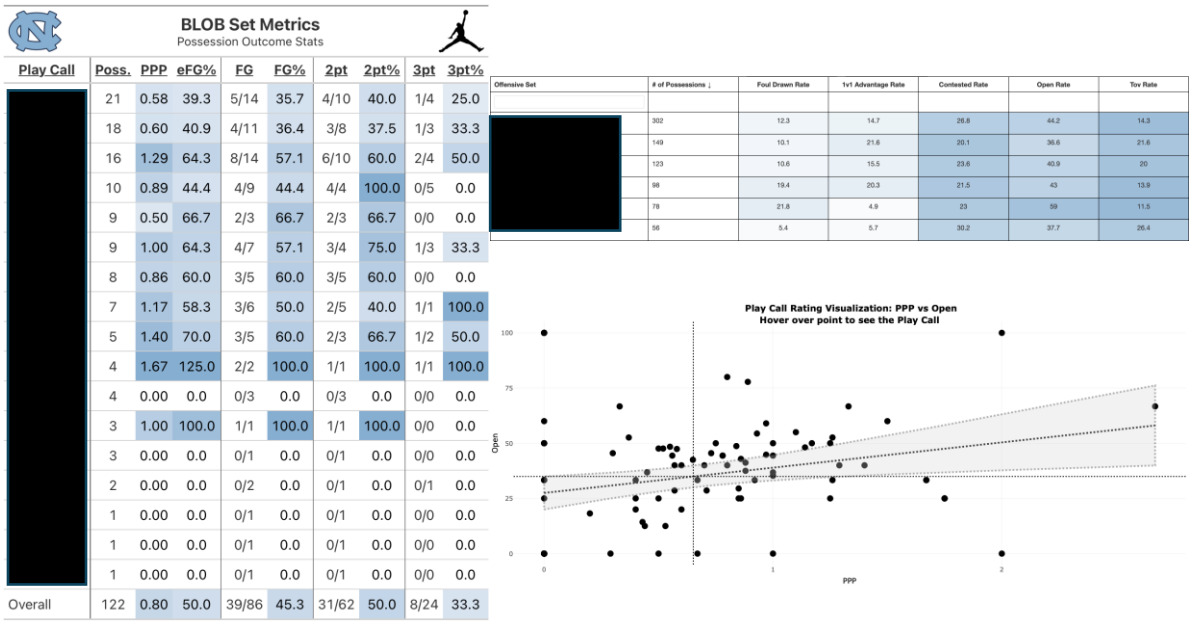
	Poss%	PPG	FG%	FTA/FGA	Tov%
Overall					
Post Up	42.5	3.0	42.4	0.35	9.4
Cut	19.5	1.8	57.6	0.27	7.7
ORB (Put Backs)	13.0	1.4	65.2	0.30	0.0
Post Up - Location					
Left Block	45.9	1.5	41.4	0.52	7.7
Right Block	41.2	1.2	41.4	0.28	5.7
Flash Middle	12.9	0.3	50.0	0.00	27.3
Post Up - Direction					
Right Shoulder	50.6	1.4	36.8	0.24	2.3
Left Shoulder	35.3	1.1	57.1	0.29	20.0

The player tendency table to the right is used to look at how individuals are being used most in the offense, and what they are the best and worst at. We give these tendencies to our practice guys so they can then play with similar styles and tendencies, so our girls get used to guarding that type of player. For example, the athlete featured on the right operates in the post primarily, where she typically plants herself on a block to get the ball, then is far more efficient shooting when she is turning over her left shoulder. This is useful as our coaches then know to tell our girls to force her to shoot over her right shoulder.

All data for the reports was collected from either Synergy, an advanced website with in-depth analytics on every level of basketball from AAU to the NBA, or Sports Reference. The data from Sports Reference was seamlessly pulled into R using the *htmltools* package. Pulling data for Synergy was not as easy as there we do not have an API at the moment. Data was pulled by sometimes downloading the data to an excel sheet, or manually inputting it, next season I plan to experiment more and talk with our Synergy representative about ways to pull this data easier. All additional notes I had for reports came from other trends I was seeing on Synergy, Sports Reference, or Her Hoops Stats. All visuals were made using the *gt* package and this source for the percentile bars and were processed using the *tidyverse* and *dplyr* package.

Offense Set Metrics

One of the key projects I worked on was developing a website to track the efficiency of our offensive sets. An assistant coach approached me with the question of wanting to know, by March, which plays we were efficient at and which ones we needed to improve. Using the data he provided after each game; I built a Shiny app that tracks various metrics for each offensive set. The app allows us to see how many times each play was run, along with shooting and scoring efficiency, foul rates, turnover rates, and more. It also includes the ability to group plays together, such as comparing Horns sets to Floppy sets, and further allows comparisons within those sets themselves. This tool provided a comprehensive look at our offensive performance, making it easier to identify strengths and areas for improvement.



BLOB Set Metrics
Possession Outcome Stats

Play Call	Poss.	PPP	eFG%	FG	FG%	2pt	2pt%	3pt	3pt%
[redacted]	21	0.58	39.3	5/14	35.7	4/10	40.0	1/4	25.0
	18	0.60	40.9	4/11	36.4	3/8	37.5	1/3	33.3
	16	1.29	64.3	8/14	57.1	6/10	60.0	2/4	50.0
	10	0.89	44.4	4/9	44.4	4/4	100.0	0/5	0.0
	9	0.50	66.7	2/3	66.7	2/3	66.7	0/0	0.0
	9	1.00	64.3	4/7	57.1	3/4	75.0	1/3	33.3
	8	0.86	60.0	3/5	60.0	3/5	60.0	0/0	0.0
	7	1.17	58.3	3/6	50.0	2/5	40.0	1/1	100.0
	5	1.40	70.0	3/5	60.0	2/3	66.7	1/2	50.0
	4	1.67	125.0	2/2	100.0	1/1	100.0	1/1	100.0
	4	0.00	0.0	0/3	0.0	0/3	0.0	0/0	0.0
	3	1.00	100.0	1/1	100.0	1/1	100.0	0/0	0.0
	3	0.00	0.0	0/1	0.0	0/1	0.0	0/0	0.0
	2	0.00	0.0	0/2	0.0	0/1	0.0	0/1	0.0
	1	0.00	0.0	0/1	0.0	0/1	0.0	0/0	0.0
	1	0.00	0.0	0/1	0.0	0/1	0.0	0/0	0.0
	1	0.00	0.0	0/1	0.0	0/1	0.0	0/0	0.0
Overall	122	0.80	50.0	39/86	45.3	31/62	50.0	8/24	33.3

Offensive Set	# of Possessions ↓	Foul Drawn Rate	1v1 Advantage Rate	Contested Rate	Open Rate	Tov Rate
[redacted]	302	12.3	14.7	26.8	44.2	14.3
	149	10.1	21.6	20.1	36.6	21.6
	123	10.6	15.5	23.6	40.9	20
	98	19.4	20.3	21.5	43	13.9
	78	21.8	4.9	23	59	11.5
	56	5.4	5.7	30.2	37.7	26.4

The three visuals above are examples of some of the visuals that are seen on the website. The visual on the left is an example of comparing all BLOB (baseline out of bounds) sets to each other. The visual on the top right is an example of comparing grouped sets to each other (Horns, Floppy, Chin, etc.) and looking at the rates we are fouled, get an advantage, shoot a contested or open shot, and turn the ball over. The visual on the bottom right shows the relationship between out PPP (points per possession) and our open shots for all play types, when using in the shiny app if you were to hover over a point you would be able to see what set it is, and the number of times we have run it along with its PPP and open rate values.

The visuals used on the Shiny App were made using the *gt*, *ggplot*, and *reactable* packages, with the plot using the *plotly* feature in Shiny Apps so it is interactive. The data was read in through the *readxl* package and processed using the *tidyverse* and *dplyr* packages. The assistants that used this website were very good at making sense of it. They found the

tables straight forward as they were the ones providing me with the data so they had expectations of what would be there. After explaining what open rate, contested rate, and more was once, they were able to make use of it. For example, if our Horns sets had an open rate of 44.4, I would say, "44.4% of our Horns sets we got an open shot," and they were then able to make sense of the rest through that. The plots where they could select the x and y axis, I recommended to keep PPP on the x as that's what was most important to draw from for these sets, they understood well after I explained each metric and how the lines/shaded regions show the relationship it has on the efficiency of the set.

Miscellaneous Project – Lineup and Player Combination Report

One of the miscellaneous projects I took on was evaluating our overall offensive efficiency by analyzing different player combinations, particularly focusing on our starting lineup. I noticed that one of our players was negatively impacting our offensive efficiency, and after diving deeper into the data, it became clear that some player combinations were performing much better than others. I started the project on my own and then shared it with one of our coaches to present a more detailed analysis. I first made sure to show which player statistics had the highest correlation to team winning percentage by using data from the past three and a half seasons and tailored the report to highlight those metrics. This allowed me to pinpoint which combinations of players were most effective on the court and which ones might be hindering our offensive flow.

This specific project used data from Sports Reference, Synergy, Her Hoops Stats, and CBB Analytics. Data was either pulled directly into R using the *htmltools* function, copy and pasted into Excel, or just directly typed into R. All visuals were primarily made through the *ggplot* and *gt* functions. At first discussing and explaining some of the visuals were difficult, but after going over it a couple times and changing my wording the coaches were able to understand. For example, when discussing individual correlation, the phrase "25% of the variability in team winning percentage can be explained by eFG%," was very helpful for them. Other than that, all other plots and metrics were straight forward, and the report included a glossary to answer all their questions on the meaning of some metrics and how they are calculated.

In the future I plan to continue to do projects like this on a smaller, more consistent scale, showing these analyses regularly to help our coaches maintain overall team efficiency. By making these evaluations on a more frequent basis, I hope to provide ongoing insights that can aid in making more informed decisions about rotations, player pairings, and in-game strategies. This will allow the coaching staff to make adjustments throughout the season,

ensuring we're maximizing our offensive potential and making data-driven decisions to improve team performance.

Future Plans

Looking ahead, I plan to continue working with the UNC Women's Basketball team next season, building on what I've established this year. During the off-season, I aim to improve and streamline the scouting process, making it faster and more seamless to get detailed reports to coaches. I will also focus on adding to the practice website by adding new features. Additionally, I plan to improve the halftime reports that coaches receive during games, incorporating more visuals and insights to support in-game decision-making. These efforts will help ensure that we continue to make data-driven improvements across all aspects of team performance.

Code Samples

Some examples of R code used this season can be found at the following GitHub link. Examples include the code for the website and offensive sets shiny apps, analytical scouting reports (from earlier in the season), and the code used to make in game shot charts.