S4E5

UVM SAC

2023-11-14

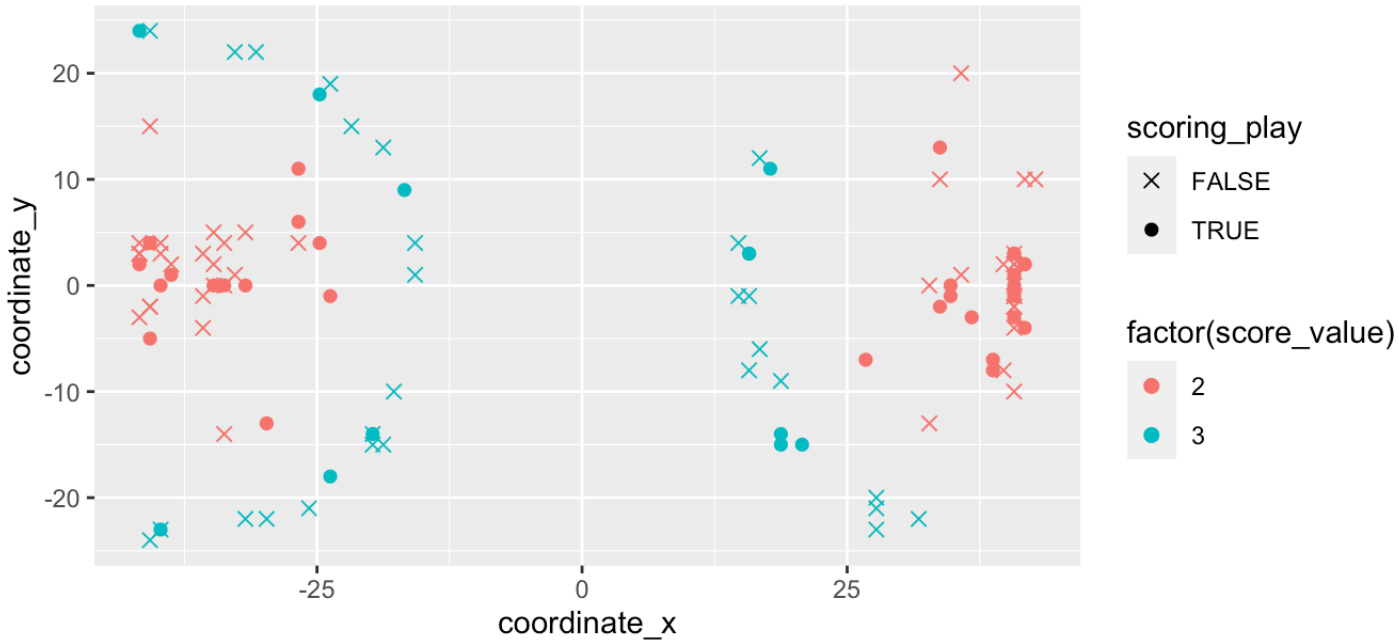
S4E5 - Basketball Data Visualization

In this weeks workshop, we're going to be working with R using an R markdown file, which combines markdown(formatted text like this one) and code chunks. Hopefully everyone can get their own R environment set up so that anyone can experiment with the data and make their own charts.

The package we're working with today is called hoopR

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.4.4     ✔ purrr   1.0.2
## ✔ tibble  3.2.1     ✔ dplyr   1.1.3
## ✔ tidyr   1.3.0     ✔ stringr 1.5.0
## ✔ readr   2.1.4     ✔ forcats 1.0.0
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE
```

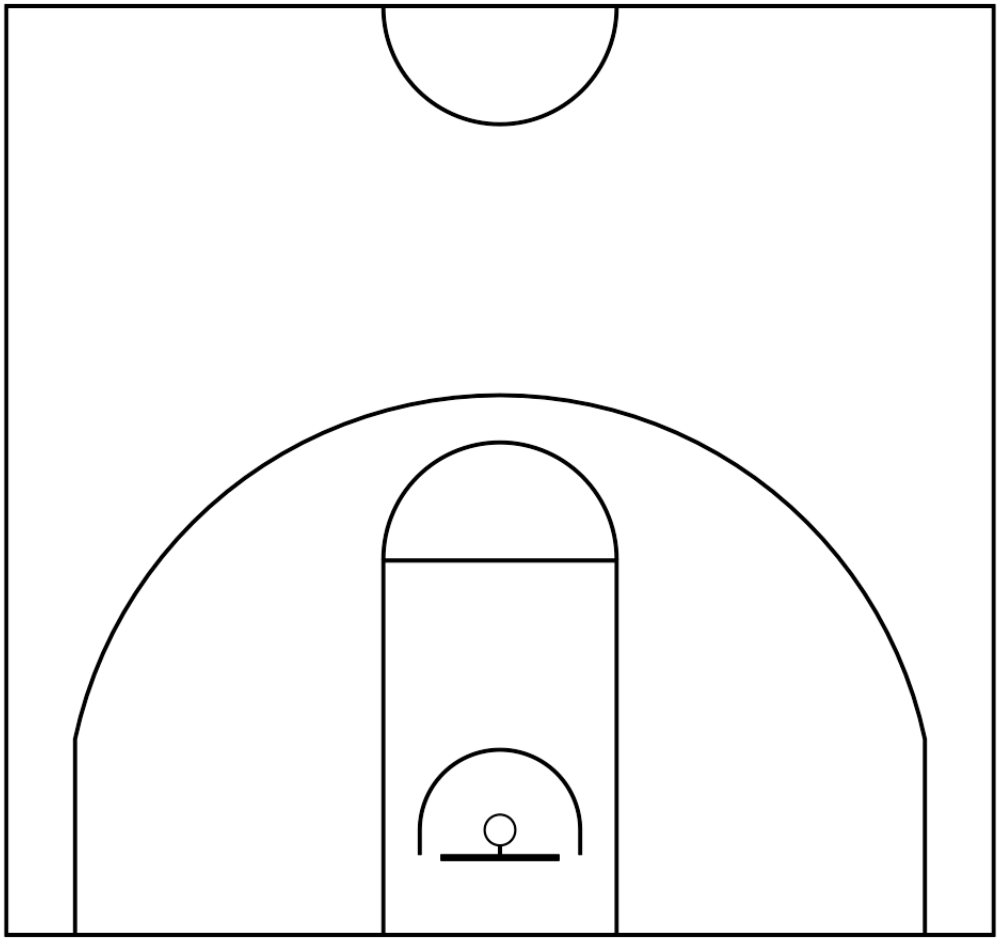
hoopR is an open source package that contains up to date NBA and NCAA play-by-play data



Lets try adding the lines on the court to better make sense of the data

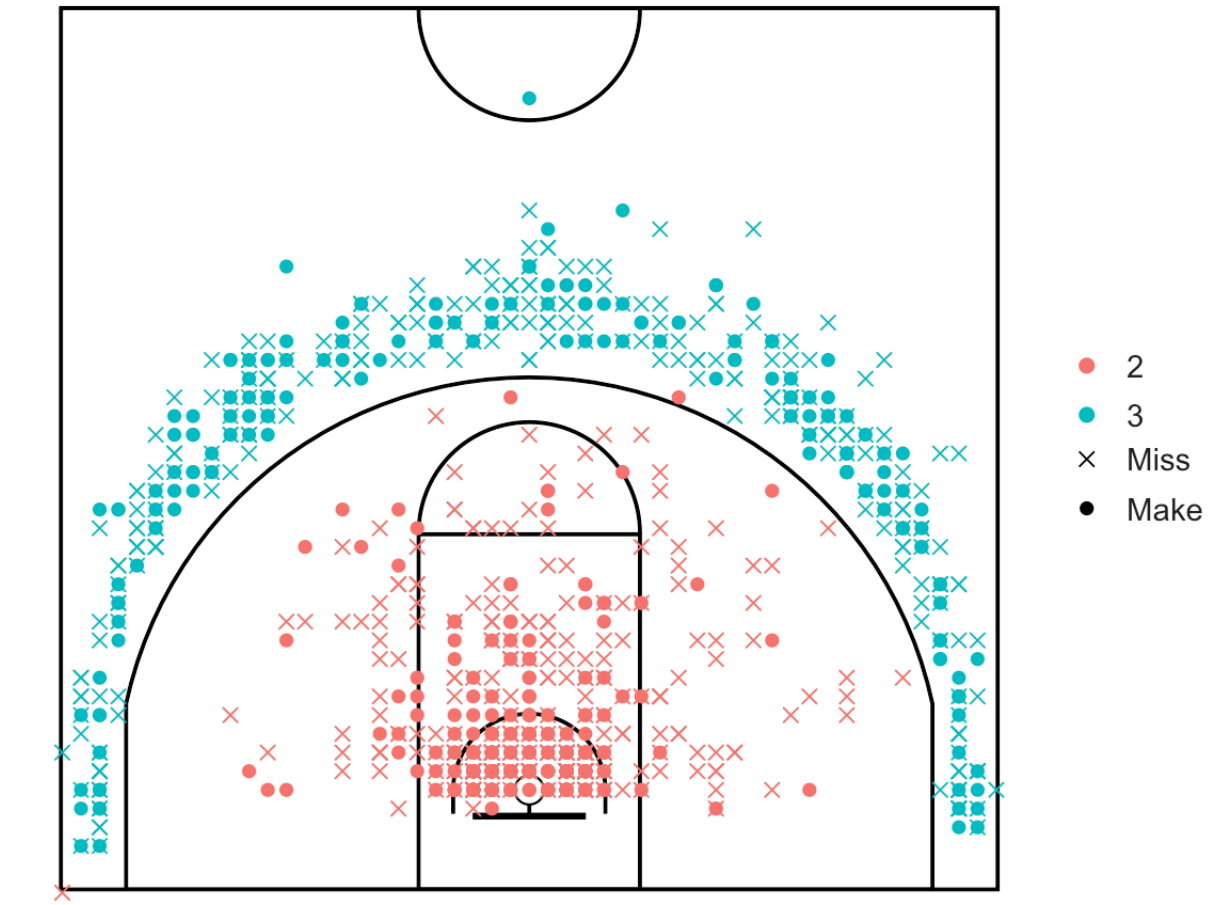
Below will be a lot of code that uses spacial objects to create to-scale lines for an NCAA court. The final result will be a theme that we can add to our shot chart graph that draws in the lines with only one line of code.

Implementing the court lines: We will have to make some changes to the graph to align with the theme. For the court lines, (0,0) is in the bottom left corner. In our previous graph (0, 0) is mid court. Also, with the court theme we only want to look at one team at a time. Also, for the theme, the x and y axes are going to be flipped.

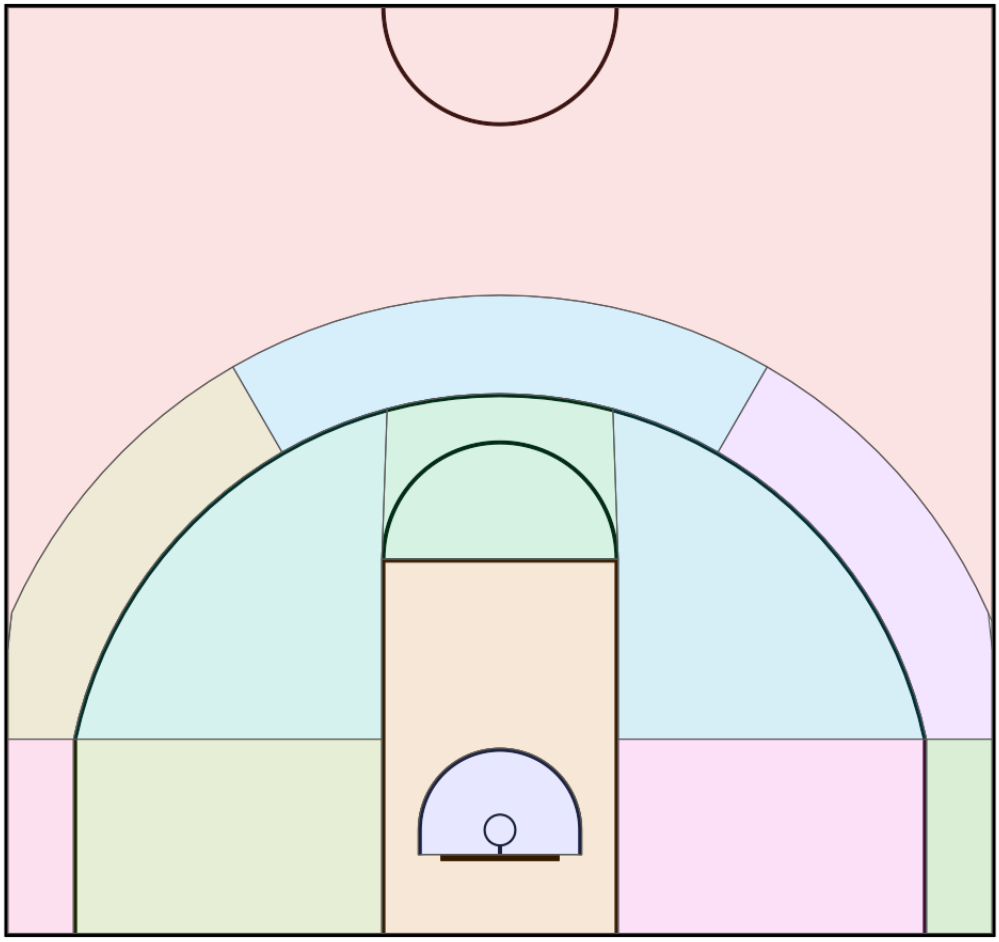


Adding shot charts to plot_court()

```
## Warning: Removed 1142 rows containing missing values (`geom_point()`).
```

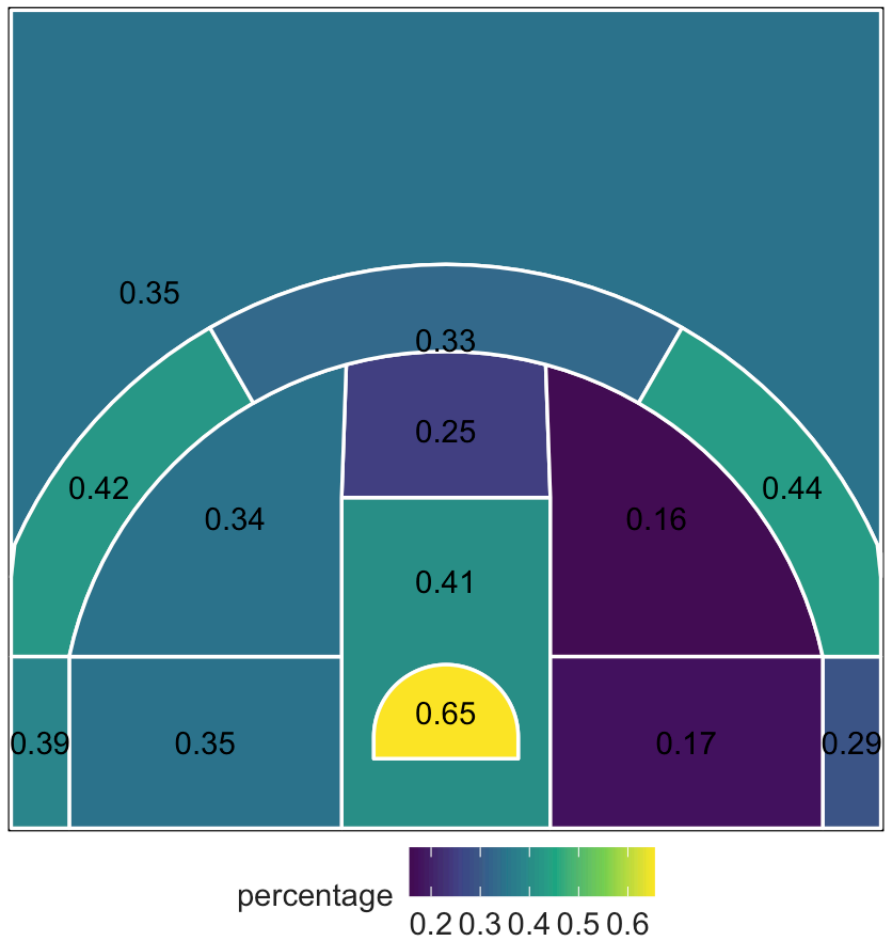


The data gets hard to read when you want to look at an entire season's worth of shots. To solve this we can try segmenting the court and look at shooting statistics for each zone.



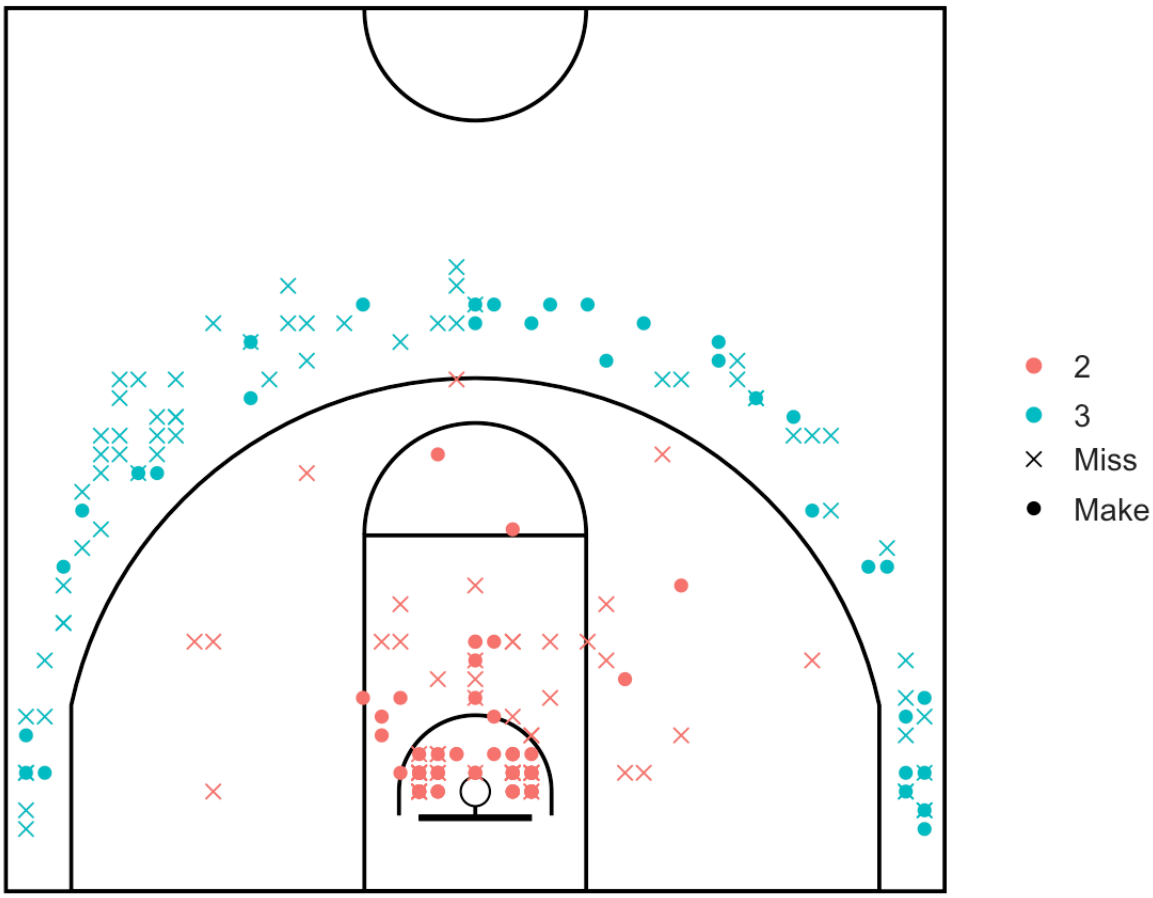
Lets take the shot data from uconn's games and see how they shot in each zone

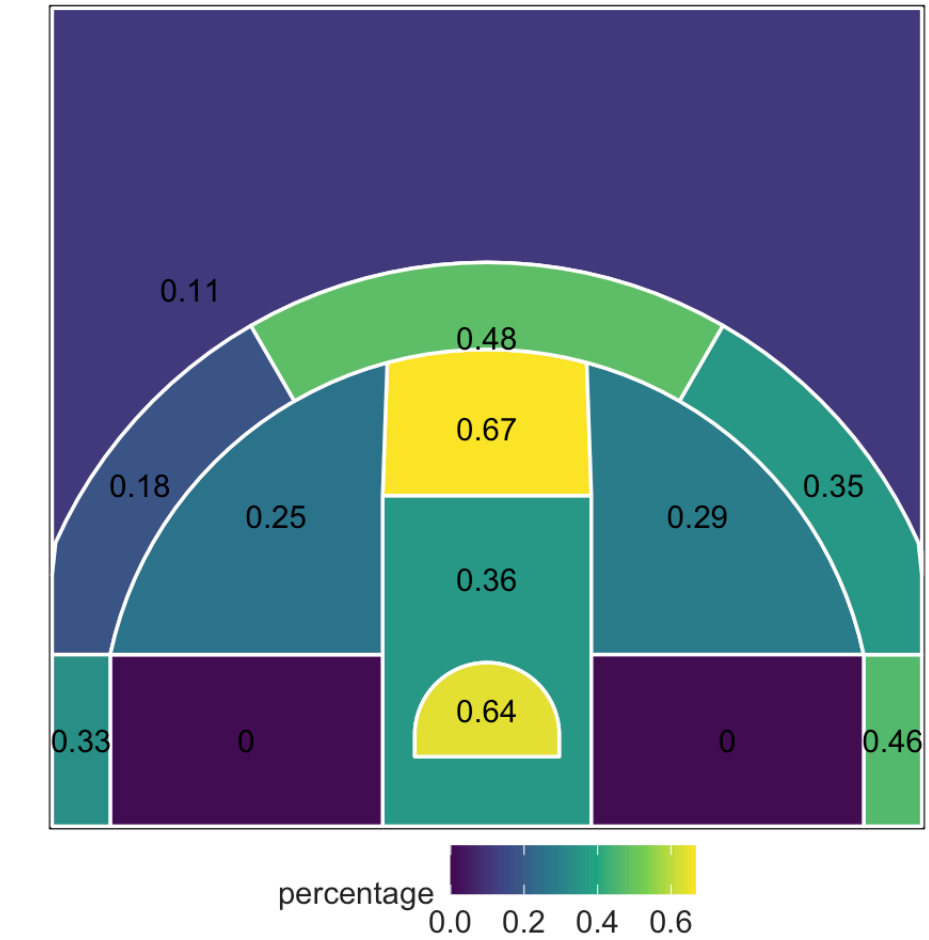
```
## # A tibble: 13 × 5
##    shot_zone_basic      num_shots num_makes percentage                    geom
##    <chr>                    <int>     <int>      <dbl>                  <POLYGON>
##  1 Deep 3                      57        20      0.351 ((0 14.43156, 0.2010777 …
##  2 In The Paint (Non-R…       138        56      0.406 ((19 19, 31 19, 31 0, 19…
##  3 Left 3                      95        40      0.421 ((0 9.863118, 0 14.43156…
##  4 Left Block                  23         8      0.348 ((19 0, 19 9.863118, 3.3…
##  5 Left Corner 3               31         9      0.290 ((46.66 0, 46.66 9.86311…
##  6 Mid-Range Center            20         5      0.25  ((19 19, 19.26825 26.641…
##  7 Mid-Range Left              29        10      0.345 ((3.34 9.863118, 3.6088 …
##  8 Mid-Range Right             32         5      0.156 ((46.66 9.863118, 46.659…
##  9 Middle 3                   120        40      0.333 ((11.4271 28.75895, 12.6…
## 10 Restricted Area            379       248      0.654 ((20.85 4, 20.85 5.25, 2…
## 11 Right 3                    110        48      0.436 ((50 9.863118, 50 15.805…
## 12 Right Block                 23         4      0.174 ((31 0, 31 9.863118, 46.…
## 13 Right Corner 3              31        12      0.387 ((0 0, 0 9.863118, 3.34 …
```



Putting it all together

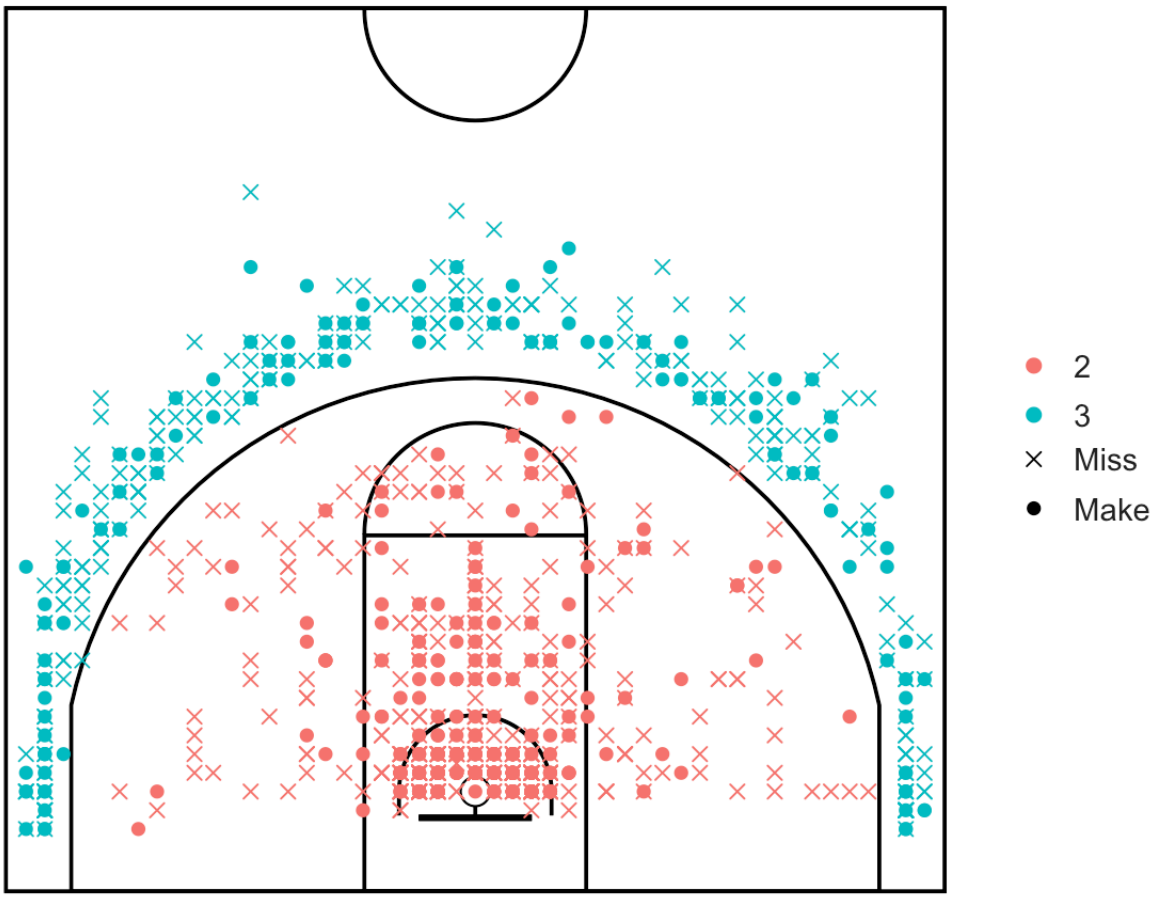
2023 Vermont Catamounts Shot Chart

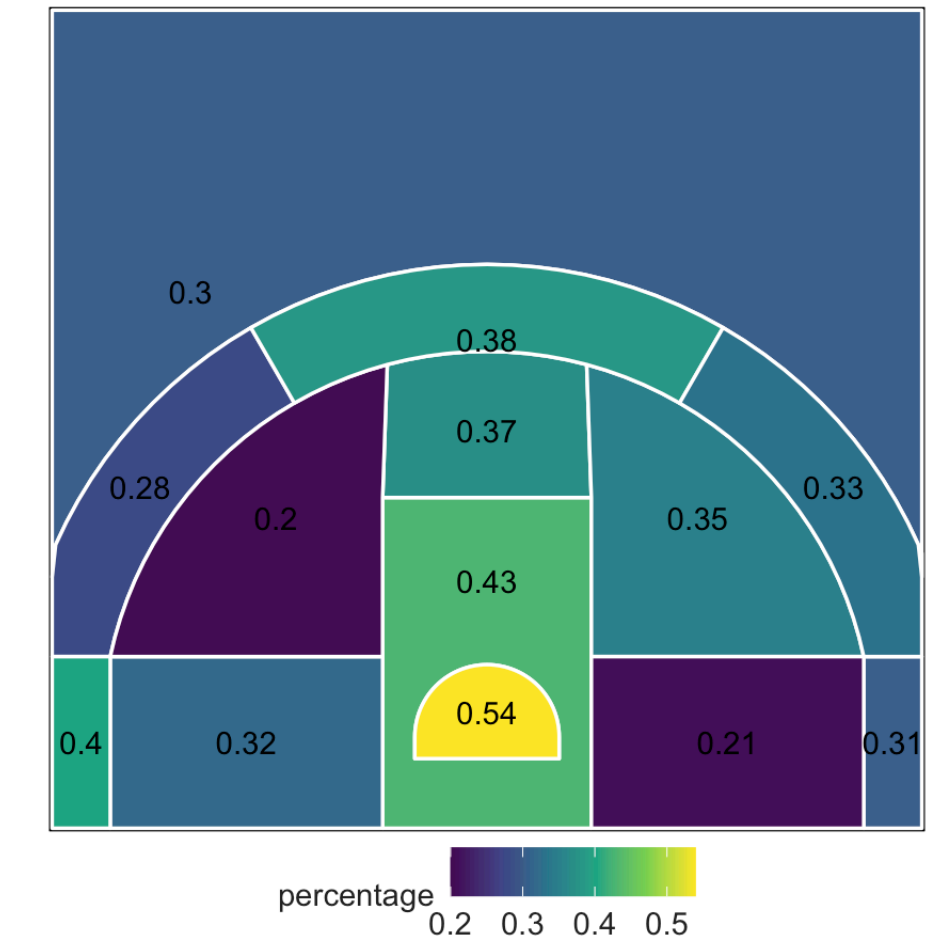




Try it yourself!!

2023 Illinois Shot Chart




2023 Illinois Shooting by Zone
0.3
0.38
0.37
0.28
0.2
0.33
0.35
0.43
0.54
0.4
0.32
0.21
0.31
percentage
0.2 0.3 0.4 0.5