

OKC Thunder Data Analytics Project

Harry Svendsgaard

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.3
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
## filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
## intersect, setdiff, setequal, union
```

```
shots = read.csv("shots_data.csv")
```

```
# Assigns a zone to each shot, zone 3 is corner 3's, zone 2 is non-corner 3's, and zone 1 is all 2-pointers. Uses the distance formula to find non-corner 3's, since the x and y value can vary.  
shots <- mutate(shots,  
  zone = case_when(  
    abs(x) >= 22 & y <= 7.8 ~ 3,  
    y > 7.8 & sqrt(x^2 + y^2) >= 23.75 ~ 2,  
    TRUE ~ 1  
  ))
```

```
# Create boolean variables for each zone showing the attempts (if a shot was attempted in this zone or not) and whether or not the shot was made in this zone  
shots <- mutate(shots,  
  zone1att= case_when(zone == 1 ~ 1,  
    TRUE ~ 0),  
  zone2att= case_when(zone == 2 ~ 1,  
    TRUE ~ 0),  
  zone3att= case_when(zone == 3 ~ 1,  
    TRUE ~ 0),  
  zone1make= case_when(zone == 1 & fgmade == 1 ~ 1,  
    TRUE ~ 0),  
  zone2make= case_when(zone == 2 & fgmade == 1 ~ 1,  
    TRUE ~ 0),  
  zone3make= case_when(zone == 3 & fgmade == 1 ~ 1,  
    TRUE ~ 0))
```

Now, I will compute the shot distributions for each zone and each team.

```
totals <- shots %>%  
  group_by(team) %>%  
  summarise_at(c("zone1att", "zone2att", "zone3att", "zone1make", "zone2make", "zone3make"), su  
m)  
  
totals$zone1dis <- totals$zone1att / (totals$zone1att + totals$zone2att + totals$zone3att)  
  
totals$zone2dis <- totals$zone2att / (totals$zone1att + totals$zone2att + totals$zone3att)  
  
totals$zone3dis <- totals$zone3att / (totals$zone1att + totals$zone2att + totals$zone3att)
```

Now to find the effective field goal percentage for each team, in each zone.

```
totals$zone1eFG <- totals$zone1make / totals$zone1att  
totals$zone2eFG <- 1.5*totals$zone2make / totals$zone2att  
totals$zone3eFG <- 1.5*totals$zone3make / totals$zone3att
```

Write the data frame to a CSV file to submit with script

```
write.csv(totals, "C:\\Users\\harry\\Downloads\\ThunderProject\\ThunderProject.csv")
```