groupwork6

Gafur Mammadov

2025-02-12

```
womenShots <- read.csv("~/Documents/math133/datasets/womenShots.csv")
summary(womenShots)
```
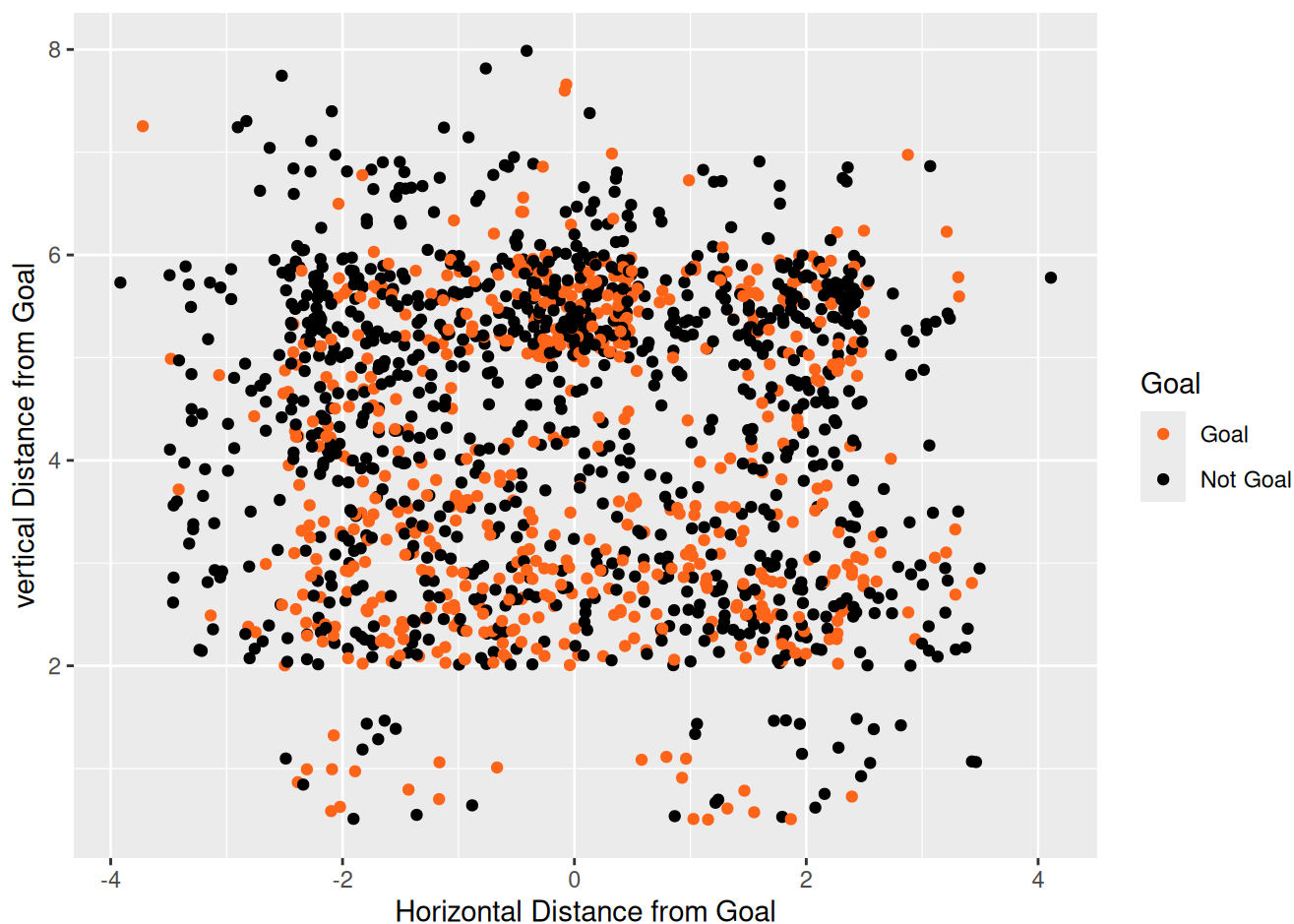
```
##       X.1              X            start.time          end.time
##  Min.   :   1.0   Min.   :   1.0   Length:1607        Length:1607
##  1st Qu.: 402.5   1st Qu.: 402.5   Class :character   Class :character
##  Median : 804.0   Median : 804.0   Mode  :character   Mode  :character
##  Mean   : 804.0   Mean   : 804.0
##  3rd Qu.:1205.5   3rd Qu.:1205.5
##  Max.   :1607.0   Max.   :1607.0
##
##    category         Nth.instance    X..descriptors   Offensive.Team
##  Length:1607        Min.   :   1.0   Min.   : 3.00   Length:1607
##  Class :character   1st Qu.: 402.5   1st Qu.:10.00   Class :character
##  Mode  :character   Median : 804.0   Median :11.00   Mode  :character
##                     Mean   : 804.0   Mean   :10.57
##                     3rd Qu.:1205.5   3rd Qu.:11.00
##                     Max.   :1607.0   Max.   :20.00
##
##     Position        Location          Tactic             Type
##  Min.   :  2.00   Length:1607        Length:1607        Length:1607
##  1st Qu.: 31.00   Class :character   Class :character   Class :character
##  Median : 38.00   Mode  :character   Mode  :character   Mode  :character
##  Mean   : 42.56
##  3rd Qu.: 61.00
##  Max.   :401.00
##
##      Hand              Result            Defense          Save.Diff
##  Length:1607        Length:1607        Length:1607        Min.   :1.000
##  Class :character   Class :character   Class :character   1st Qu.:1.000
##  Mode  :character   Mode  :character   Mode  :character   Median :2.000
##                                                           Mean   :2.409
##                                                           3rd Qu.:3.000
##                                                           Max.   :5.000
##
##    Shot.Diff        Athlete             x                 y
##  Min.   :1.000   Min.   : 2.000   Min.   :-4.00000   Min.   :1.00
##  1st Qu.:2.000   1st Qu.: 5.000   1st Qu.:-1.00000   1st Qu.:2.50
##  Median :3.000   Median :10.000   Median : 0.00000   Median :4.50
##  Mean   :3.049   Mean   : 9.219   Mean   :-0.08712   Mean   :4.28
##  3rd Qu.:4.000   3rd Qu.:12.000   3rd Qu.: 1.00000   3rd Qu.:5.50
##  Max.   :5.000   Max.   :20.000   Max.   : 4.00000   Max.   :7.50
##                  NA's   :753
##       xj               yj             distance          angle
##  Min.   :-3.91651   Min.   :0.5033   Min.   :1.414   Min.   :-1.10715
##  1st Qu.:-1.47059   1st Qu.:2.9334   1st Qu.:3.500   1st Qu.:-0.34877
##  Median :-0.12799   Median :4.5656   Median :4.924   Median : 0.00000
##  Mean   :-0.08145   Mean   :4.2785   Mean   :4.615   Mean   :-0.01262
##  3rd Qu.: 1.33195   3rd Qu.:5.5290   3rd Qu.:5.590   3rd Qu.: 0.34877
##  Max.   : 4.11045   Max.   :7.9869   Max.   :8.500   Max.   : 1.24905
##
##      Goal
##  Length:1607
##  Class :character
##  Mode  :character
```

```
##
##
##
##
```

```
womenShots %>% ggplot(aes(x=xj,y=yj,col=Goal))+geom_point()+
    scale_color_manual(values=c("#FF671D","black"))+
    labs(x="Horizontal Distance from Goal",y='vertical Distance from Goal')
```



```
set.seed(123)  # Ensure reproducibility
indices <- sample(1:nrow(womenShots), size = 0.7 * nrow(womenShots))
train_data <- womenShots[indices, ]
test_data <- womenShots[-indices, ]
```

```
goal_knn = knn(train_data[,c("xj","yj")],test_data[,c("xj","yj")],cl=train_data$Goal,k=1
0)
head(goal_knn)
```

```
## [1] Not Goal Goal     Not Goal Not Goal Not Goal Not Goal
## Levels: Goal Not Goal
```

```
sum(goal_knn!=test_data$Goal)/length(test_data$Goal)
```

```
## [1] 0.4120083
```