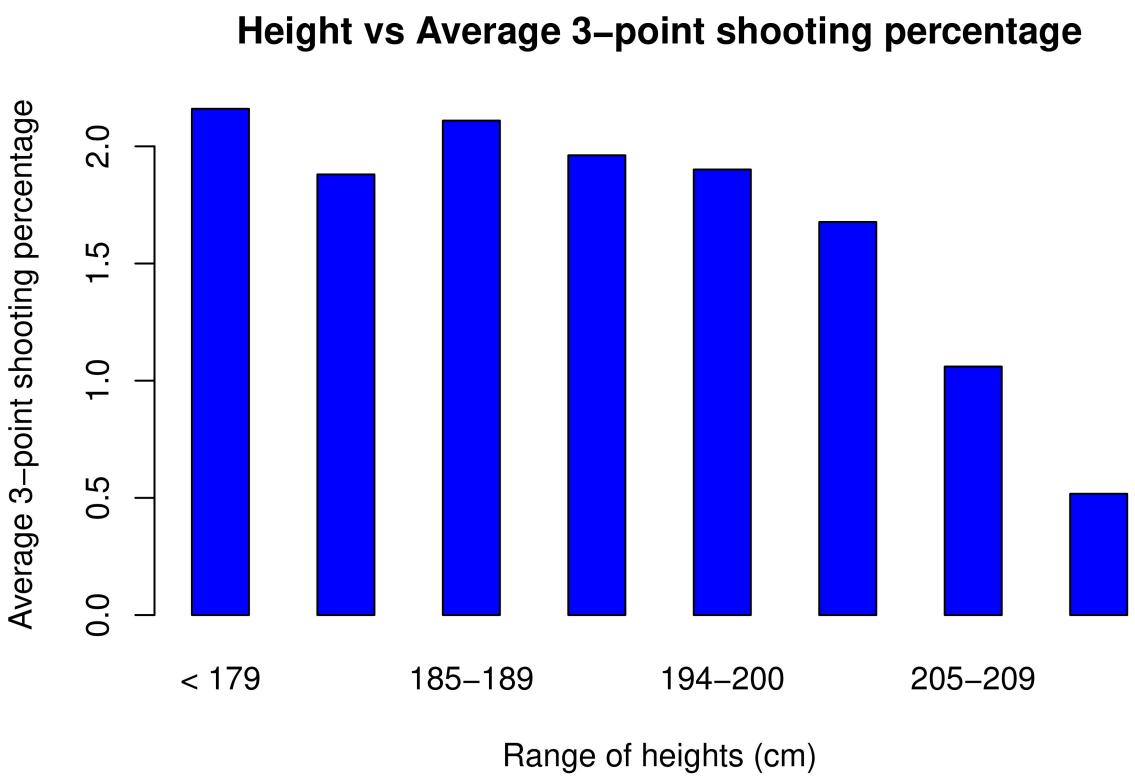
APPENDIX

```
heights<-vector()
tpp<-vector()
```

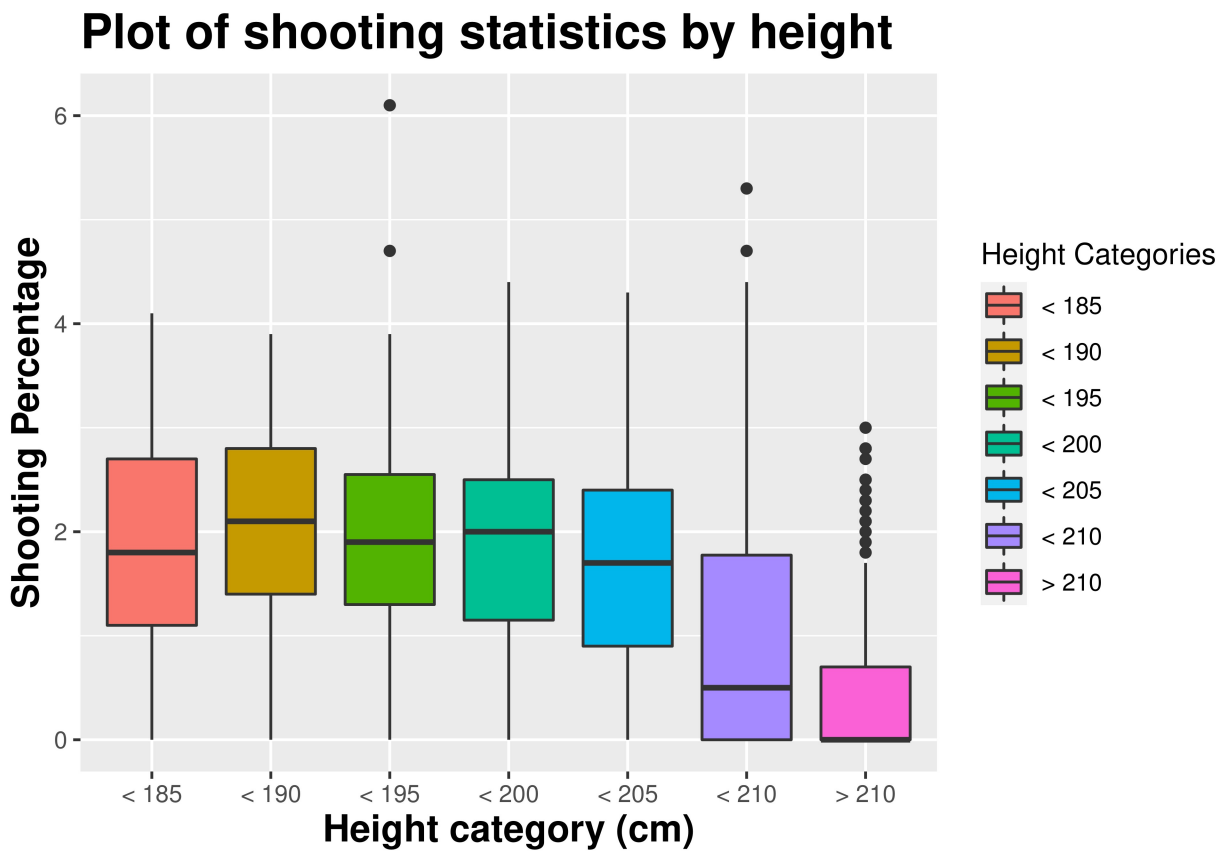
```
st<-read.csv("stats.csv")
n<-c(0, 0, 0, 0, 0, 0, 0, 0)
sum<-c(0, 0, 0, 0, 0, 0, 0, 0)
sheight<-c(0, 0, 0, 0, 0, 0, 0, 0)
less180<-vector()
less185<-vector()
less190<-vector()
less195<-vector()
less200<-vector()
less205<-vector()
less210<-vector()
more210<-vector()

for(row in 1:nrow(st)){
  height<-st[row, "height"]
  point<-st[row, "X3P"]

  heights[row] = st[row, "height"]
  tpp[row] = st[row, "X3P"]

  # 179-, 180 - 184, 185 - 189, 190 - 194, 194 - 200, 200 - 204, 205 - 209, 210+
  if(height < 180){
    n[1]<-n[1] + 1
    sum[1]<-sum[1]+point
    sheight[1]<-sheight[1]+height
    less180<-c(less180, point)
  }
  else if(height < 185){
    n[2]<-n[2] + 1
    sum[2]<-sum[2]+point
    sheight[2]<-sheight[2]+height
    less185<-c(less185, point)
  }
  else if(height < 190){
    n[3]<-n[3] + 1
    sum[3]<-sum[3]+point
    sheight[3]<-sheight[3]+height
    less190<-c(less190, point)
  }
  else if(height < 195){
    n[4]<-n[4] + 1
    sum[4]<-sum[4]+point
    sheight[4]<-sheight[4]+height
```

```
    less195<-c(less195, point)
  }
  else if(height < 200){
    n[5]<-n[5] + 1
    sum[5]<-sum[5]+point
    sheight[5]<-sheight[5]+height
    less200<-c(less200, point)
  }
  else if(height < 204){
    n[6]<-n[6] + 1
    sum[6]<-sum[6]+point
    sheight[6]<-sheight[6]+height
    less205<-c(less205, point)
  }
  else if(height < 210){
    n[7]<-n[7] + 1
    sum[7]<-sum[7]+point
    sheight[7]<-sheight[7]+height
    less210<-c(less210, point)
  }
  else{
    n[8]<-n[8] + 1
    sum[8]<-sum[8] + point
    sheight[8]<-sheight[8]+height
    more210<-c(more210, point)
  }
}

avg<-c(sum[1]/n[1], sum[2]/n[2], sum[3]/n[3], sum[4]/n[4], sum[5]/n[5], sum[6]/n[6], sum[7]/n[7], sum[8
avgheight<-c(sheight[1]/n[1], sheight[2]/n[2], sheight[3]/n[3], sheight[4]/n[4], sheight[5]/n[5], sheigh

barplot(avg, xlab = "Range of heights (cm)", ylab = "Average 3-point shooting percentage",
        names.arg = c("< 179", "180-184", "185-189", "190-194", "194-200", "200-204", "205-209", "210 <
        main = "Height vs Average 3-point shooting percentage",
        space = 1.2, col = "blue")

less185<-c(less185, less180)

l185<-data.frame(group = "< 185", value = less185)
l190<-data.frame(group = "< 190", value = less190)
l195<-data.frame(group = "< 195", value = less195)
l200<-data.frame(group = "< 200", value = less200)
l205<-data.frame(group = "< 205", value = less205)
l210<-data.frame(group = "< 210", value = less210)
m210<-data.frame(group = "> 210", value = more210)

boxstuff<-rbind(l185, l190, l195, l200, l205, l210, m210)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.3
```




```
ggplot(boxstuff, aes(x=group, y=value, fill=group)) + geom_boxplot()  +
  xlab("Height category (cm)") +
  ylab("Shooting Percentage") +
  ggtitle("Plot of shooting statistics by height") +
  labs(fill = "Height Categories") +
  theme(plot.title = element_text(size = 18, face = "bold"),
        axis.title.x = element_text(size=14, face="bold"),
        axis.title.y = element_text(size=14, face="bold"))
```

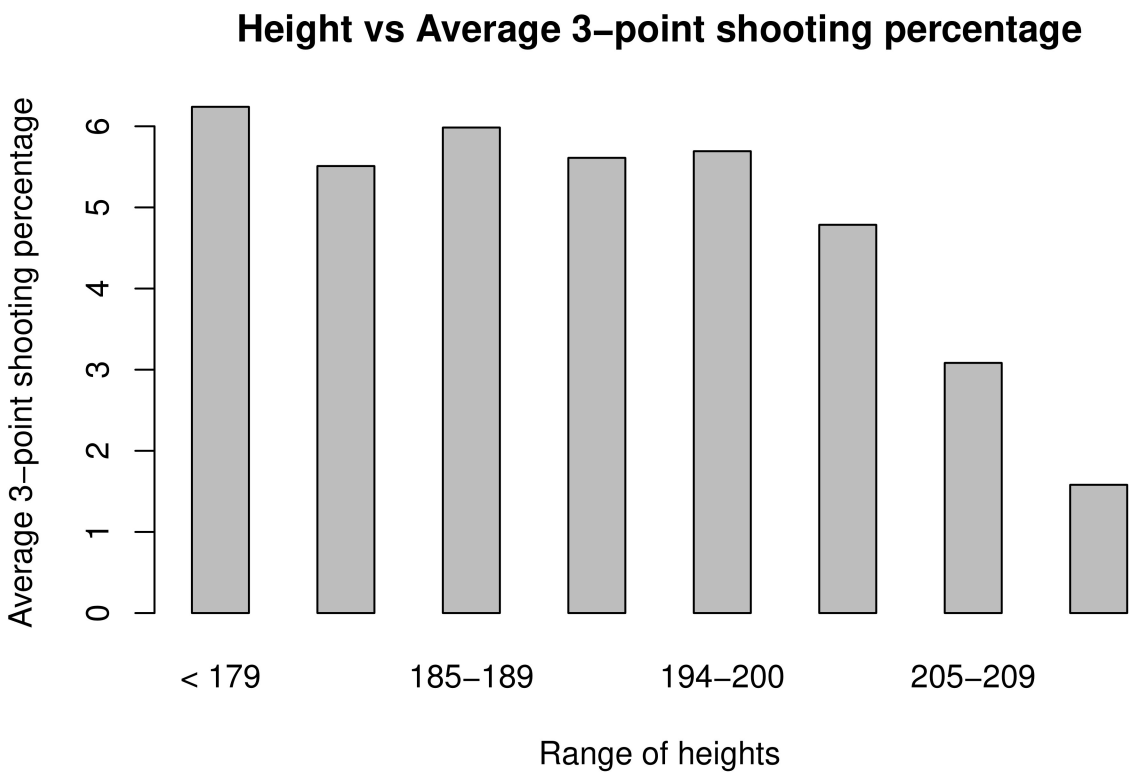



```
n1<-c(0, 0, 0, 0, 0, 0, 0, 0)
sum1<-c(0, 0, 0, 0, 0, 0, 0, 0)

for(row in 1:nrow(st)){
  name<-st[row, "Player"]
  height<-st[row, "height"]
  point<-st[row, "X3PA"]
  # 179-, 180 - 184, 185 - 189, 190 - 194, 194 - 200, 200 - 204, 205 - 209, 210+


  if(height < 180){
    n1[1]<-n1[1] + 1
    sum1[1]<-sum1[1]+point
  }
  else if(height < 185){
    n1[2]<-n1[2] + 1
    sum1[2]<-sum1[2]+point
  }
  else if(height < 190){
    n1[3]<-n1[3] + 1
    sum1[3]<-sum1[3]+point
  }
  else if(height < 195){
    n1[4]<-n1[4] + 1
    sum1[4]<-sum1[4]+point
  }
```

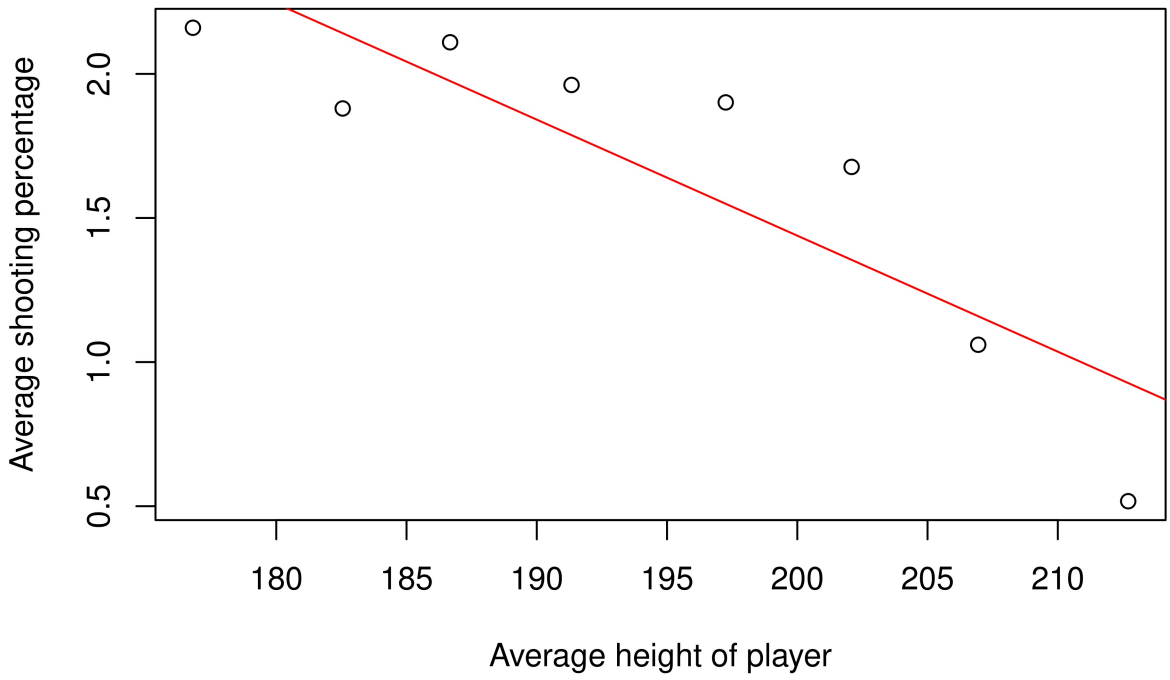
```
  else if(height < 200){
    n1[5]<-n1[5] + 1
    sum1[5]<-sum1[5]+point
  }
  else if(height < 204){
    n1[6]<-n1[6] + 1
    sum1[6]<-sum1[6]+point
  }
  else if(height < 210){
    n1[7]<-n1[7] + 1
    sum1[7]<-sum1[7]+point
  }
  else{
    n1[8]<-n1[8] + 1
    sum1[8]<-sum1[8] + point
  }
}

(avg1<-c(sum1[1]/n1[1], sum1[2]/n1[2], sum1[3]/n1[3], sum1[4]/n1[4], sum1[5]/n1[5], sum1[6]/n1[6], sum1
```

```
## [1] 6.240000 5.510000 5.984615 5.611111 5.692929 4.785593 3.083562 1.580992
```

```
barplot(avg1, xlab = "Range of heights", ylab = "Average 3-point shooting percentage",
        names.arg = c("< 179", "180-184", "185-189", "190-194", "194-200", "200-204", "205-209", "210 <
        main = "Height vs Average 3-point shooting percentage",
        space = 1.2)
```



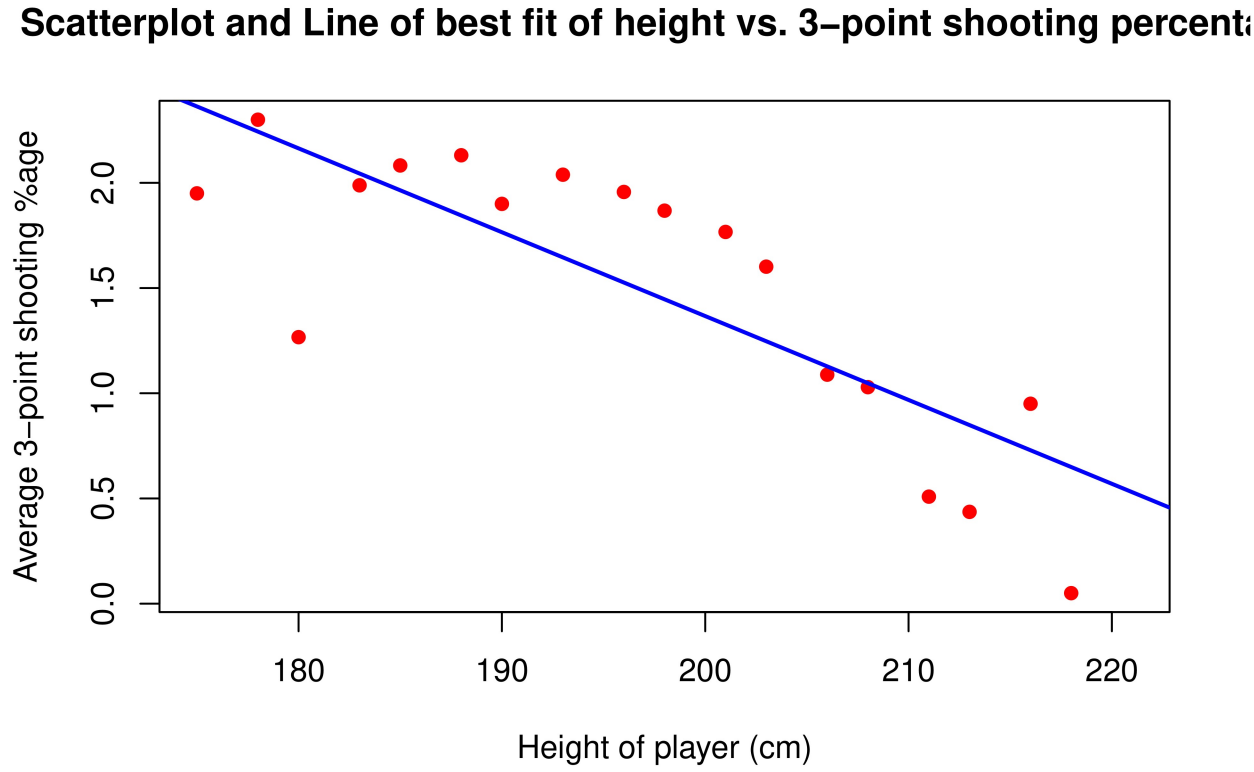

```
#plot(heights, tpp)
#pred = lm(tpp ~ heights)
#abline(pred, col = "blue")

plot(avgheight, avg, xlab = "Average height of player", ylab = "Average shooting percentage")
predAvg = lm(avg ~ avgheight)
abline(predAvg, col = "red")
```




```
#plot(heights, resid(pred))

#summary(pred)
summary(predAvg)
```

```
##
## Call:
## lm(formula = avg ~ avgheight)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41000 -0.22473  0.01781  0.21113  0.35172
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.489630   1.824529   5.201  0.00201 **
## avgheight   -0.040255   0.009362  -4.300  0.00509 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3075 on 6 degrees of freedom
## Multiple R-squared:  0.755,  Adjusted R-squared:  0.7142
## F-statistic: 18.49 on 1 and 6 DF,  p-value: 0.005094
```

```r
numb<-(1:(max(st$height)-min(st$height)))*0
newsum<-(1:(max(st$height)-min(st$height)))*0
for(row in 1:nrow(st)){
  height<-st[row, "height"]
  point<-st[row, "X3P"]

  newsum[height - 174] = newsum[height - 174] + point
  numb[height - 174] = numb[height - 174] + 1
}
x<-c(175:max(st$height))
newavg<-newsum/numb
plot(x, newavg, xlab = "Height of player (cm)", ylab = "Average 3-point shooting %age", main = "Scatterp
modelthis<-lm(newavg ~ x)
abline(modelthis, col = "blue", lwd = 2)
```




```r
summary(modelthis)
```

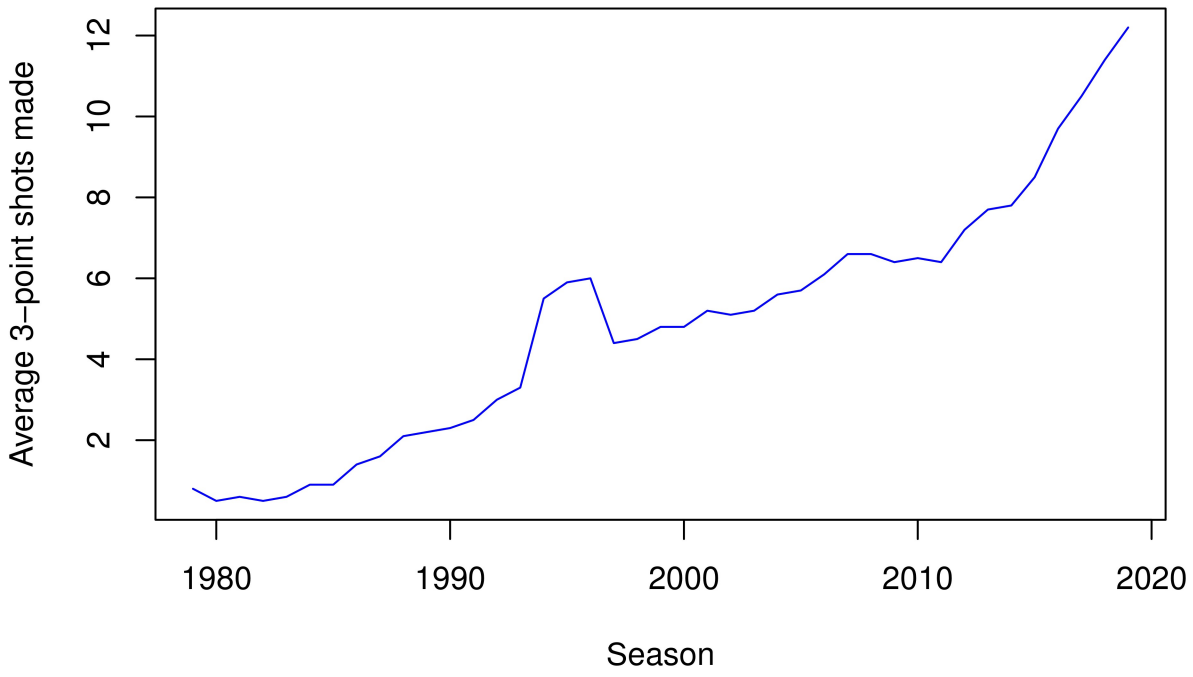
```
##
## Call:
## lm(formula = newavg ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.89713 -0.32317  0.08729  0.33734  0.43982
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.337662   1.445122   6.462 7.86e-06 ***
## x           -0.039855   0.007327  -5.439 5.46e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4099 on 16 degrees of freedom
##   (29 observations deleted due to missingness)
## Multiple R-squared:  0.649,  Adjusted R-squared:  0.6271
## F-statistic: 29.58 on 1 and 16 DF,  p-value: 5.46e-05
```

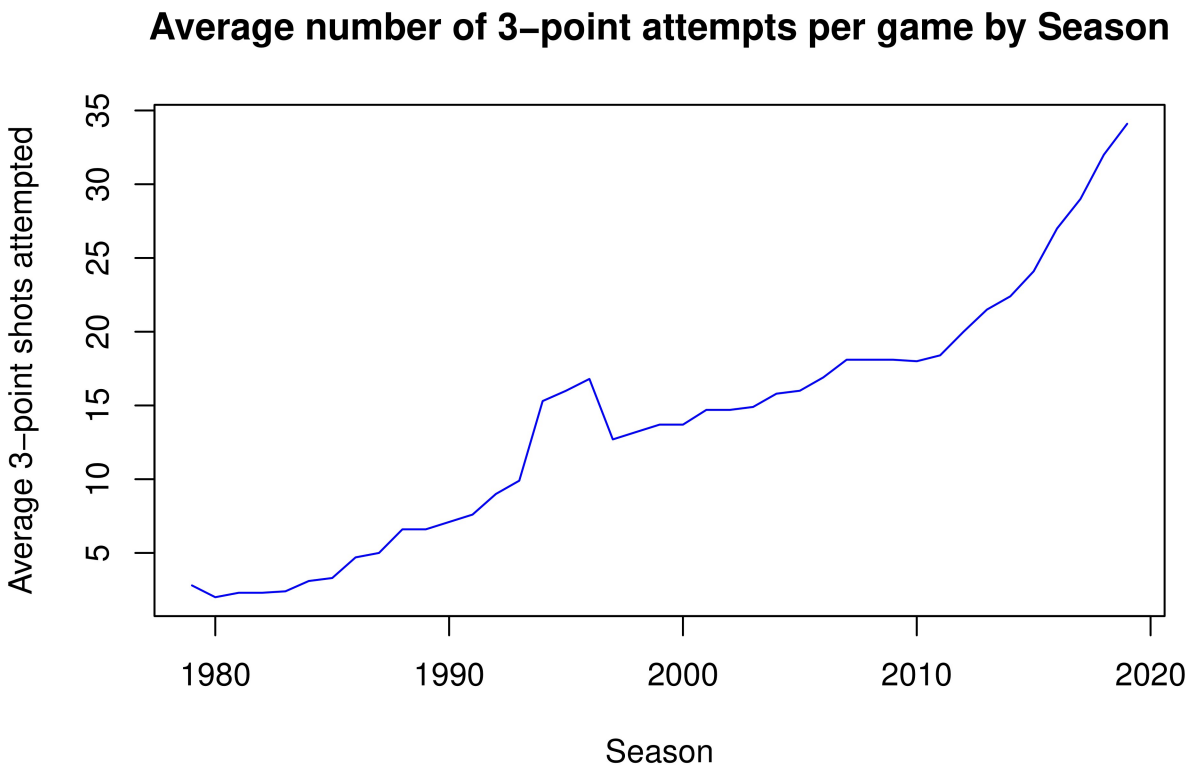
```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.0.3
```

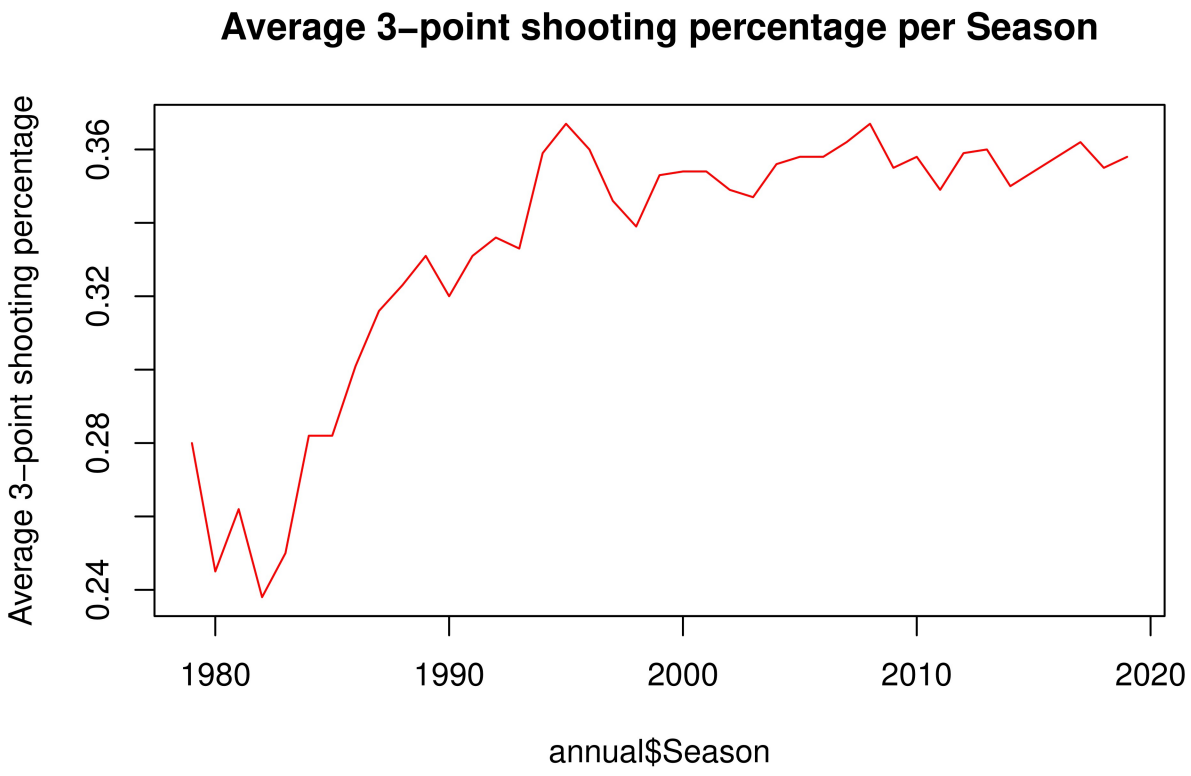
```
annual<-read_excel("annual.xlsx")
plot(annual$Season, annual$TP, xlab = "Season", ylab = "Average 3-point shots made", col = "blue", type
```




```
plot(annual$Season, annual$TPA, xlab = "Season", ylab = "Average 3-point shots attempted", type = "l",
```




```
plot(annual$Season, annual$TPp, ylab = "Average 3-point shooting percentage", type = "l", col = "red", r
```




```
ress<-aov(value~group, data = boxstuff)
summary(ress)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## group         6  208.1   34.68   32.07 <2e-16 ***
## Residuals   653  706.2    1.08
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(ress)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = value ~ group, data = boxstuff)
##
## $group
##                     diff        lwr         upr     p adj
## < 190-< 185  0.17361538 -0.5749434  0.92217421 0.9933556
## < 195-< 185  0.02561616 -0.6628387  0.71407105 0.9999998
## < 200-< 185 -0.03498990 -0.7234448  0.65346499 0.9999990
## < 205-< 185 -0.25888136 -0.9360694  0.41830665 0.9185651
## < 210-< 185 -0.87572603 -1.5414648 -0.20998722 0.0021152
## > 210-< 185 -1.41864463 -2.0943634 -0.74292585 0.0000000
## < 195-< 190 -0.14799922 -0.6747697  0.37877129 0.9816919
```

```
## < 200-< 190 -0.20860528 -0.7353758  0.31816523 0.9047701
## < 205-< 190 -0.43249674 -0.9444544  0.07946091 0.1614268
## < 210-< 190 -1.04934141 -1.5460558 -0.55262702 0.0000000
## > 210-< 190 -1.59226001 -2.1022727 -1.08224737 0.0000000
## < 200-< 195 -0.06060606 -0.4977754  0.37656323 0.9996329
## < 205-< 195 -0.28449752 -0.7036997  0.13470465 0.4110124
## < 210-< 195 -0.90134219 -1.3017857 -0.50089867 0.0000000
## > 210-< 195 -1.44426079 -1.8610854 -1.02743622 0.0000000
## < 205-< 200 -0.22389146 -0.6430936  0.19531071 0.6955246
## < 210-< 200 -0.84073613 -1.2411796 -0.44029261 0.0000000
## > 210-< 200 -1.38365473 -1.8004793 -0.96683016 0.0000000
## < 210-< 205 -0.61684467 -0.9975919 -0.23609742 0.0000421
## > 210-< 205 -1.15976327 -1.5577032 -0.76182335 0.0000000
## > 210-< 210 -0.54291860 -0.9210465 -0.16479066 0.0004944
```