

Student Crimtab Data Graphic Analysis

데이터테크전공 20173204 곽명빈

2020-10-06

Data Manipulation

```
load("./crimtab.RData")
```

산점도를 여러 유형으로 표현하기 위하여 필요한 패키지 설치

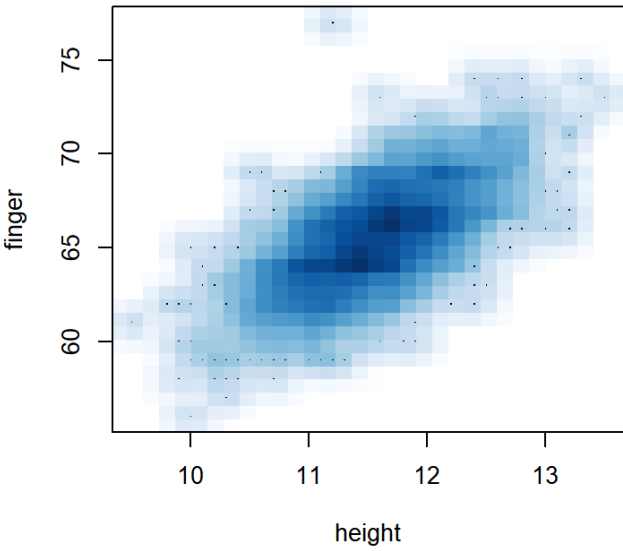
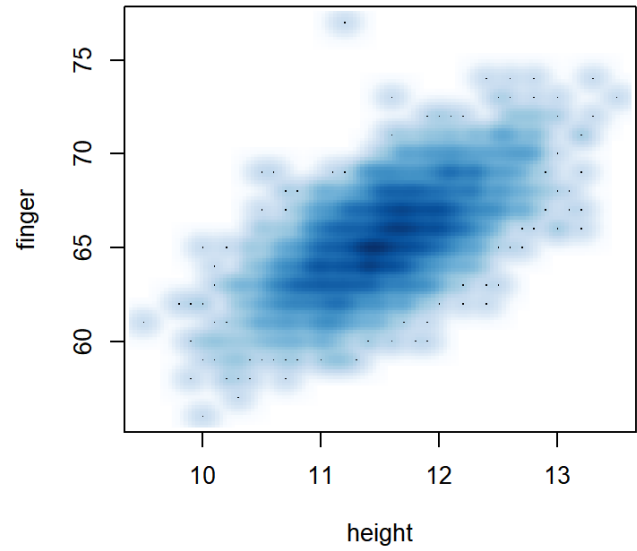
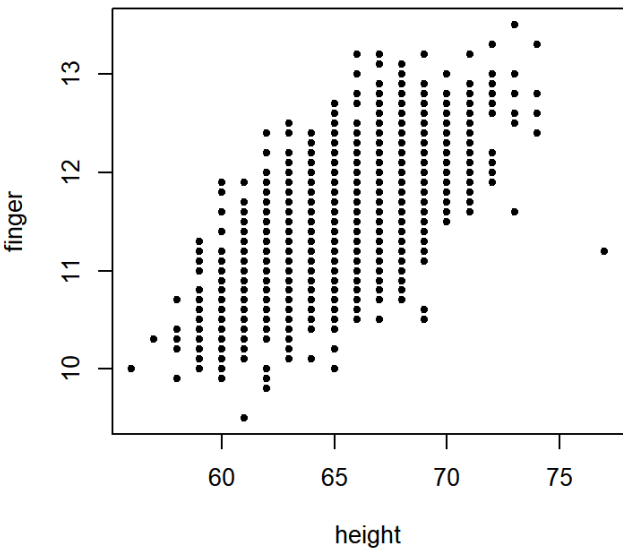
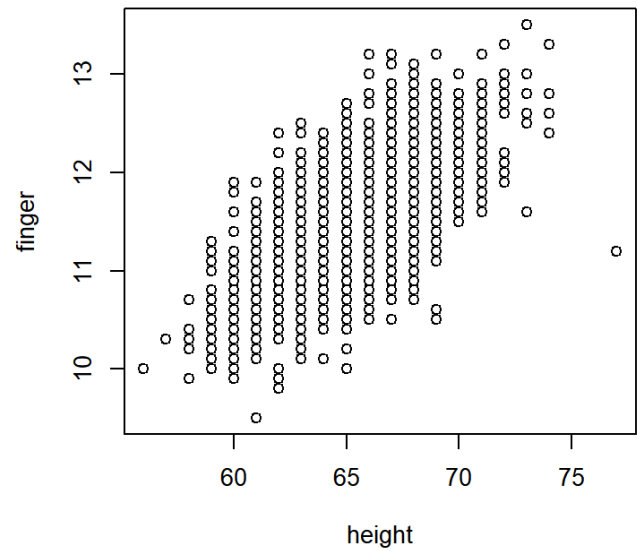
```
# install.packages("hexbin", repos = "https://cran.rstudio.com")  
library(hexbin)
```

crimtab_bin 계산

```
crimtab_bin <- hexbin(crimtab_long_df$height,  
                      crimtab_long_df$finger,  
                      xbins = 50)
```

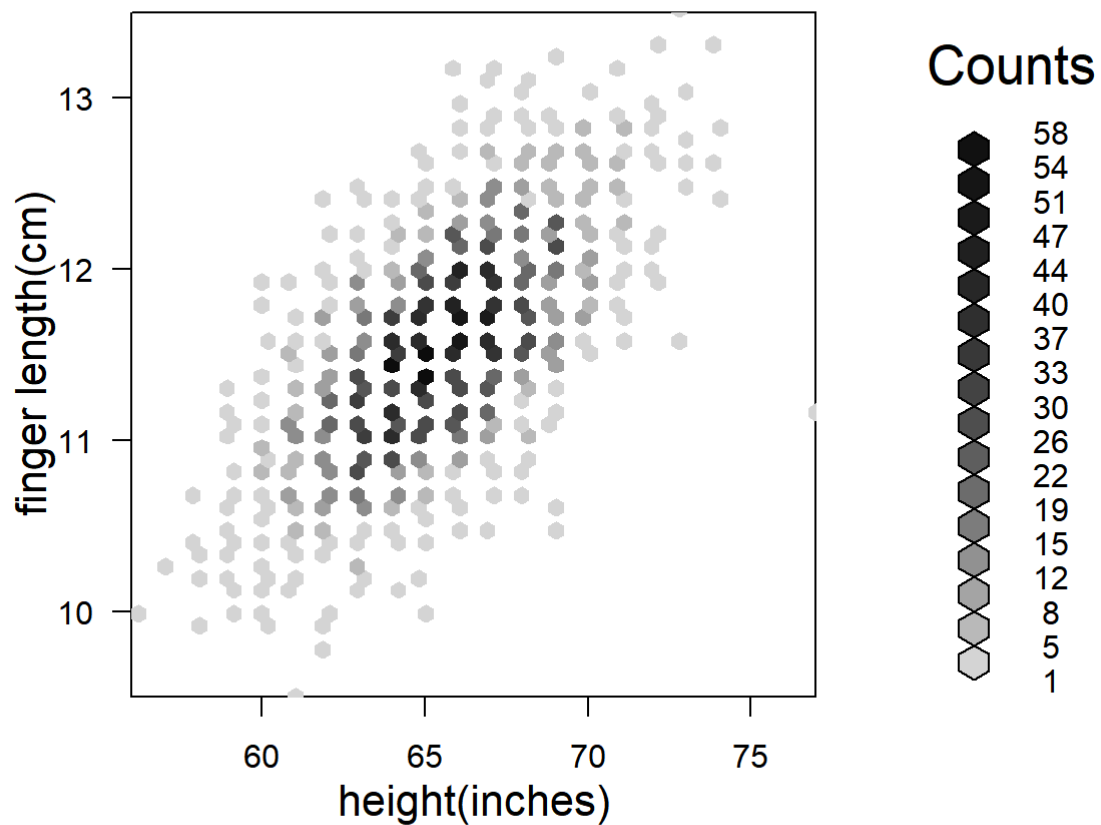
Plots

```
par(mfrow = c(2, 2))  
# plot(x = crimtab_long_df[, 2], y = crimtab_long_df[, "finger"])  
plot(crimtab_long_df[, 2:1])  
plot(crimtab_long_df[, 2:1],  
     pch = 20)  
# smoothScatter(crimtab_long_df[, "height"], crimtab_long_df[, "finger"], xlab = "height", ylab =  
# "finger")  
# smoothScatter(crimtab_long_df[, "height"], crimtab_long_df[, "finger"], nbin = 32, xlab = "height", ylab = "finger")  
smoothScatter(crimtab_long_df[,  
               xlab = "height",  
               ylab = "finger")  
smoothScatter(crimtab_long_df[,  
               nbin = 32,  
               xlab = "height",  
               ylab = "finger")
```



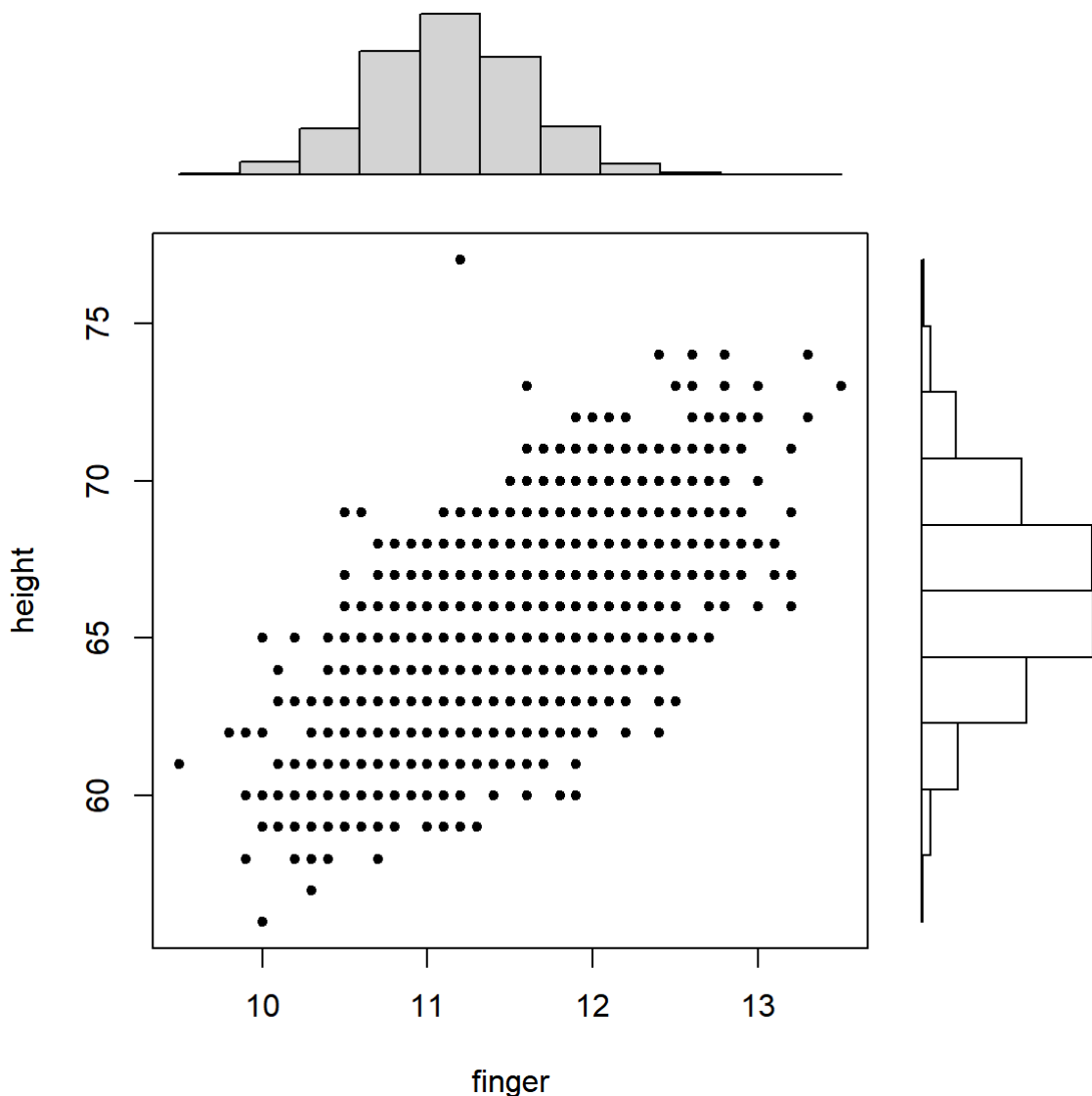
Plot crimtab_bin

```
par(mfrow = c(1, 1))  
plot(crimtab_bin,  
     xlab = "height(inches)",  
     ylab = "finger length(cm)")
```



산점도와 함께 주변분포 표시

```
par(mar = c(4, 4, 1, 1))
par(fig = c(0, 0.8, 0, 0.8))
plot(crimtab_long_df, pch = 20)
par(fig = c(0, 0.8, 0.68, 1), new = TRUE)
# hist(crimtab_long_df[, "height"], axes = FALSE, ann = FALSE)
hist(crimtab_long_df$height, axes = FALSE, ann = FALSE)
par(fig = c(0.68, 1, 0, 0.8), new = TRUE)
# barplot(table(cut(crimtab_long_df[, "finger"], breaks = 10))), space = 0, col = "white", horiz
= TRUE, axes = FALSE, axisnames = FALSE)
barplot(table(cut(crimtab_long_df$finger, breaks = 10)),
        space = 0, col = "white", horiz = TRUE, axes = FALSE, axisnames = FALSE)
```

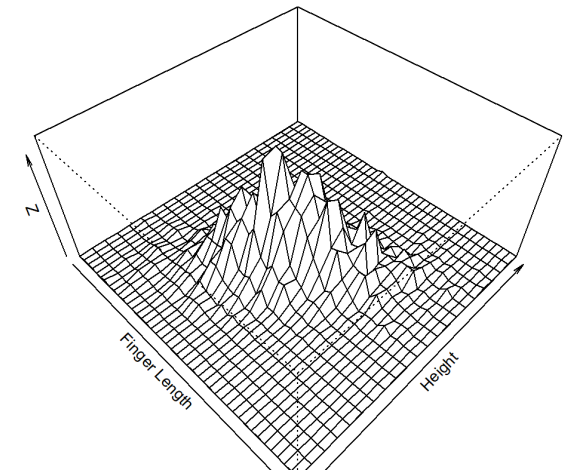
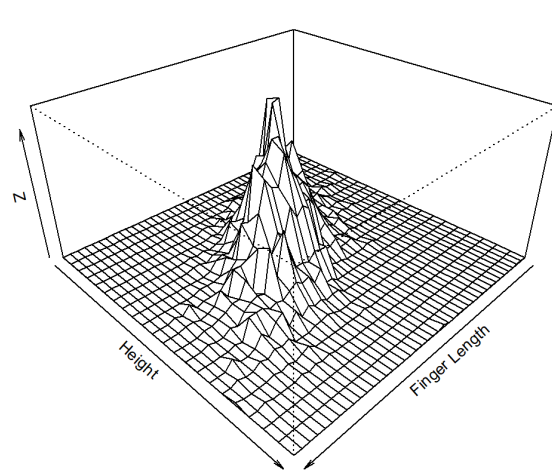
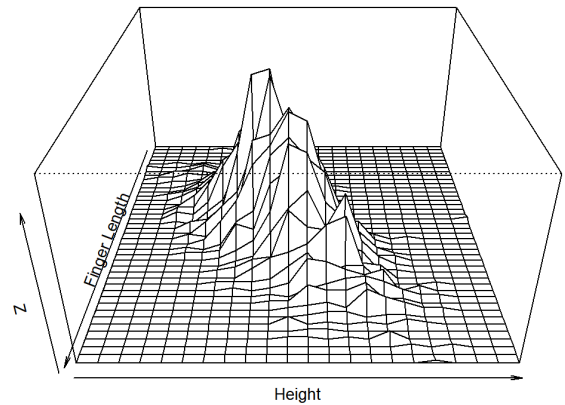
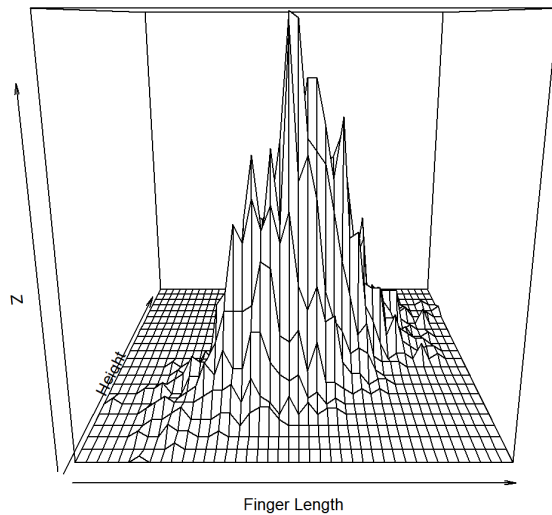


```
par(fig = c(0, 1, 0, 1))
par(mar = c(5, 4, 1, 1) + 0.1)
```

persp()

`persp()` 를 활용하면 다양한 각도에서 3차원 겨냥도를 그려볼 수 있음. x 축은 행, y 축은 열에 펼쳐진 격자를 0에서 1까지로 조정. θ 와 ϕ 는 박스를 돌려보는 각도이고, `expand` 는 박스 높이의 상대적인 비율임. x 축과 y 축의 라벨 이외에는 디폴트값을 적용시킨 겨냥도와 적절히 조정한 겨냥도를 비교해 볼 것,

```
par(mfrow = c(2, 2))
persp(crimtab_2,
      xlab = "Finger Length",
      ylab = "Height")
persp(crimtab_2,
      xlab = "Finger Length",
      ylab = "Height",
      theta = 90,
      phi = 30,
      expand = 0.5,
      scale = TRUE)
persp(crimtab_2,
      xlab = "Finger Length",
      ylab = "Height",
      theta = 135,
      phi = 30,
      expand = 0.5,
      scale = TRUE)
persp(crimtab_2,
      xlab = "Finger Length",
      ylab = "Height",
      theta = 45,
      phi = 45,
      expand = 0.5,
      scale = TRUE)
```



```
par(mfrow = c(1, 1))
```

Comments

이번시간에는 “Student” : william. S. Gosset이 3000명의 죄수를 바탕으로 한 키와 손가락길이의 관계에대한 측정치에 대해 배울수 있었습니다. 두개를 살펴보니 가운데 밀집되어 있는 것을 확인 할 수있었습니다.

as.data.frame을 이용해서 변환해주는 것이 중요하다는 것을 알게 되었습니다. 산점도를 사용하여crimtab_bin을 시각화를 하였습니다. 기존의 데이터로만 보다가 산점도로 표현하니, 두 변수의 관계에 대해 더 잘 이해 할 수 있는 시간이였습니다. new=true를통해 겹쳐그리게 되어 쉽게 이해할수 있게 되었습니다. 또한persp()를통해 겨냥도 그리는 법을 알게 되었습니다. 3차원으로 보니 2차원 자료에비해 더 알아보기 쉬운 느낌을 받을수 있었습니다.