

# Homework 2

Sangwon Chung

2019 9 30

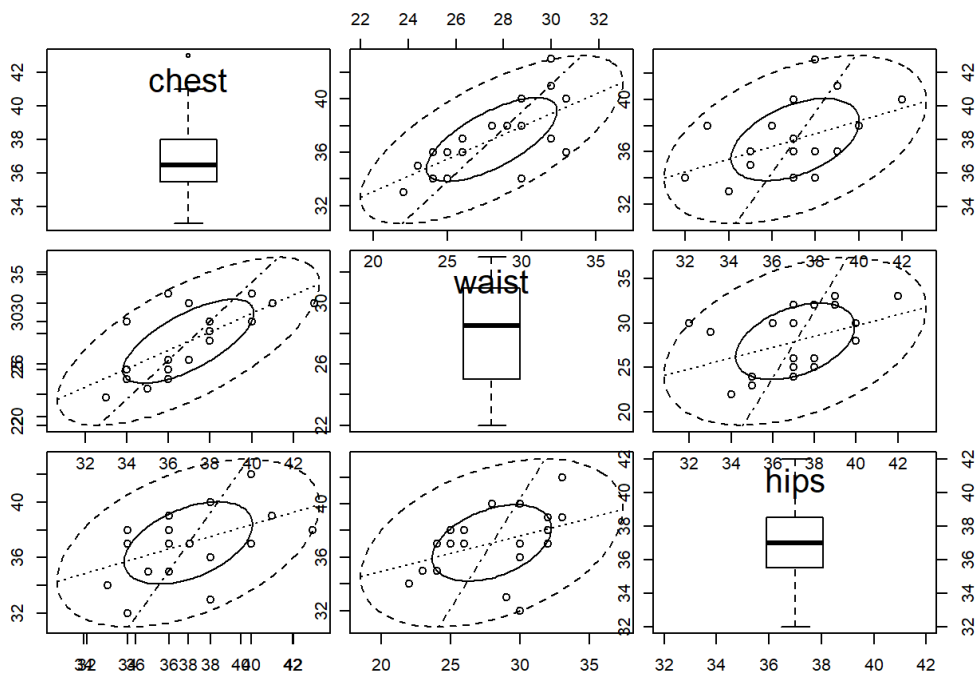
```
library(KernSmooth)
library(MVA)
library(dplyr)
library(ggplot2)
library(GGally)
library(ggpubr)
library(reshape)
```

```
Measure = read.delim("F:\\WDropbox\\WWDATA SET\\WWmeasuredata.txt", sep = " ")
```

## Question1

```
panel.bxp <- function(x, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 2, usr[3:4]))
  boxplot(x, add=TRUE)
}

pairs(Measure[,c("chest", "waist", "hips")],
      diag.panel = panel.bxp,
      panel = function(x,y){
        data = data.frame(cbind(x,y))
        par(new = TRUE)
        bv = bvbox(data)
      })
```



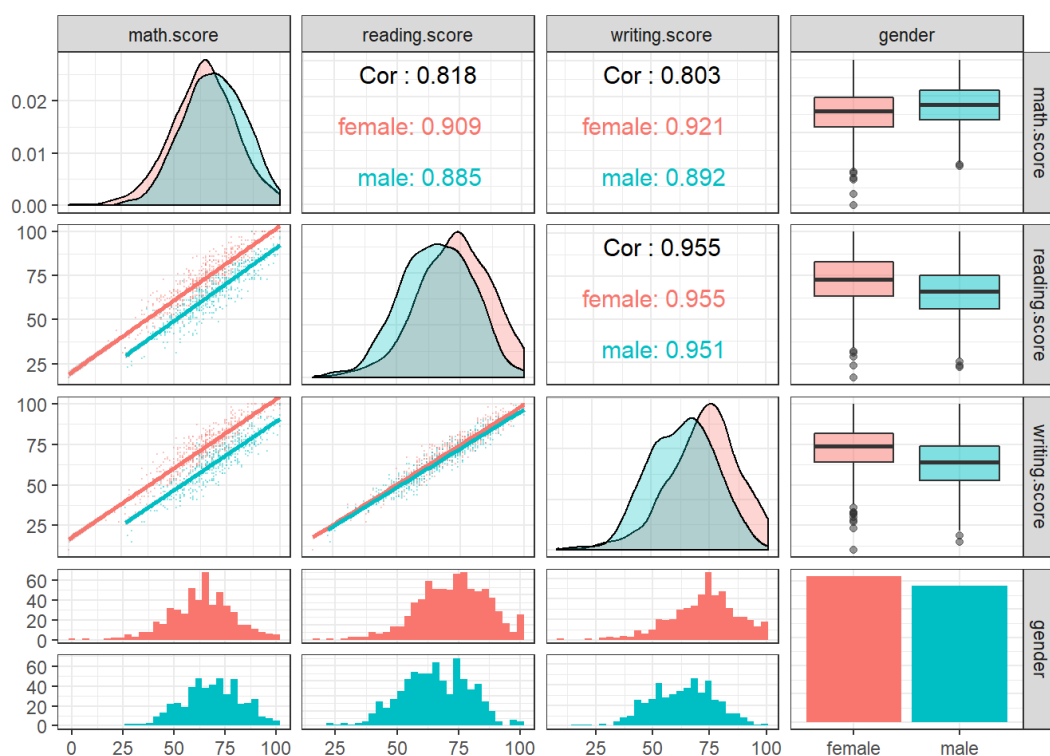
## Question2

```
Students = read.csv("F:\\WDropbox\\WWDATA SET\\WWstudents-performance-in-exams\\WWStudentsPerformance.csv")
```

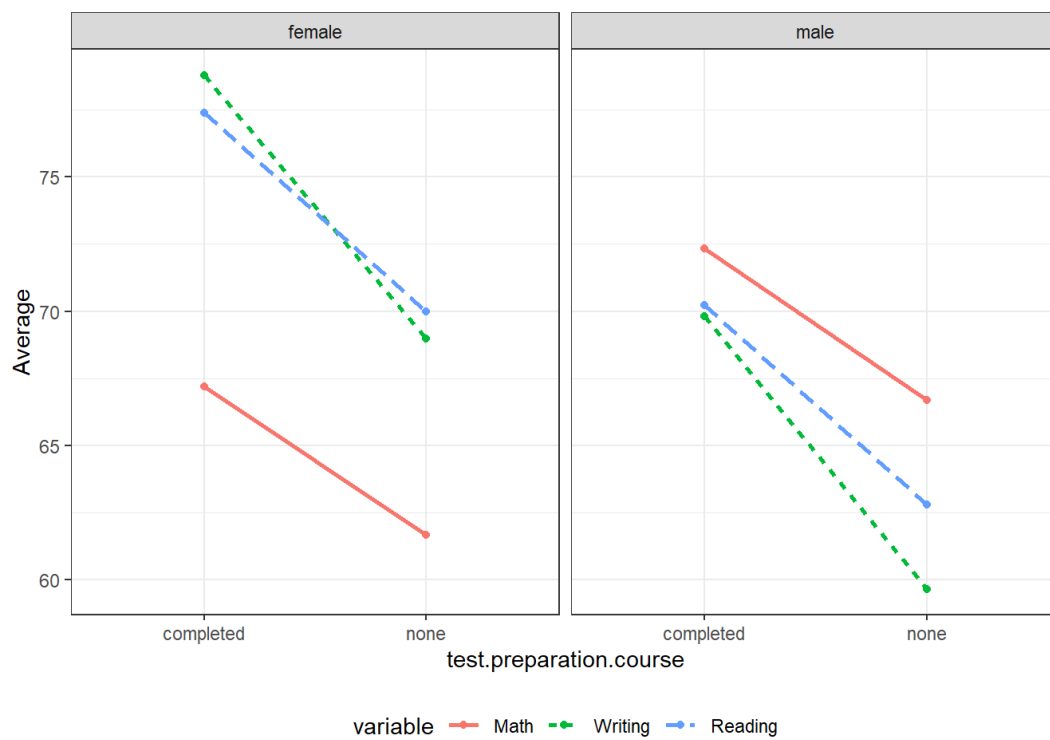
gender	race.ethnicity	parental.level.of.education	lunch	test.preparation.course	math.score	reading.score	writing.score
female	group B	bachelor's degree	standard	none	72	72	74
female	group C	some college	standard	completed	69	90	88
female	group B	master's degree	standard	none	90	95	93
male	group A	associate's degree	free/reduced	none	47	57	44
male	group C	some college	standard	none	76	78	75

## Pairs

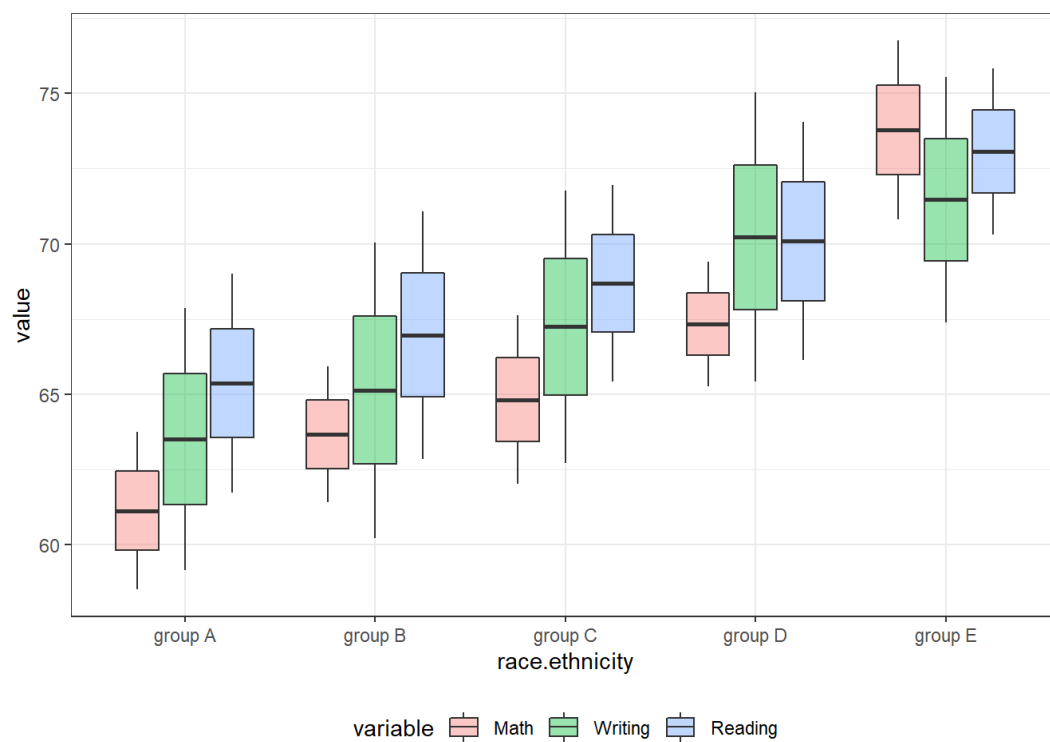
```
ggpairs(Students[,c("math.score", "reading.score", "writing.score", "gender")],
  mapping = aes(color = gender),
  diag = list(discrete="barDiag",
              continuous = wrap("densityDiag", alpha=0.3)),
  lower = list(continuous = wrap("smooth", alpha = 0.3, size=0.1)),
  upper = list(combo = wrap("box_no_facet", alpha=0.5),
              continuous = wrap("cor", size=4, alignPercent=0.8))) +
  theme_bw()
```



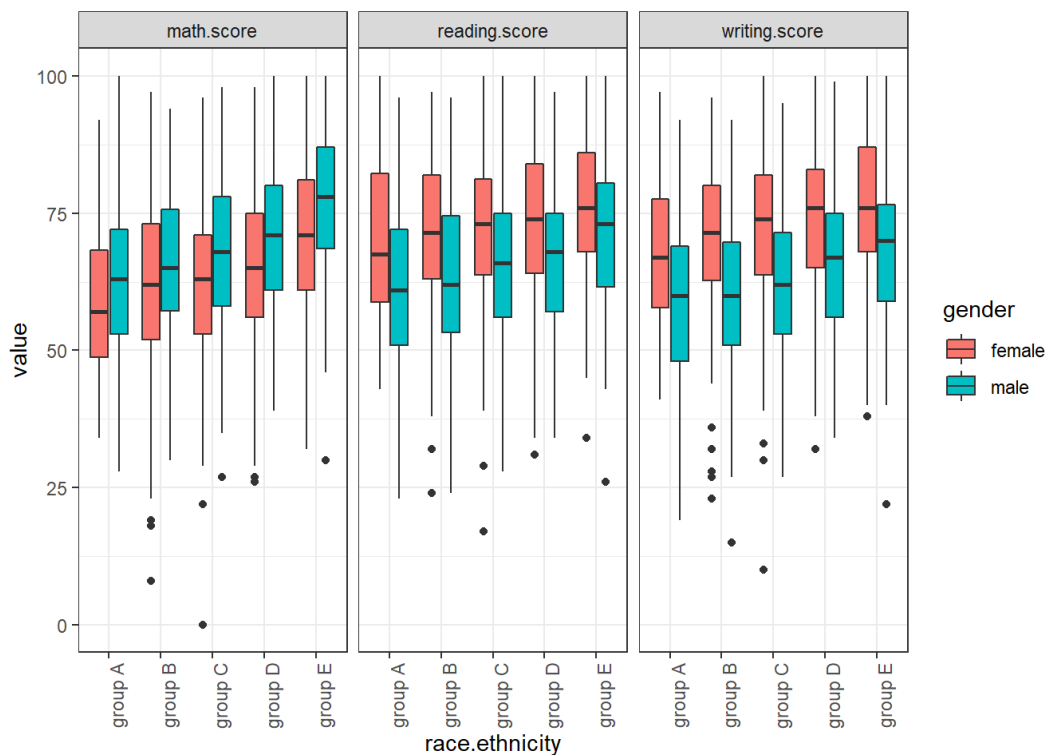
```
Students %>%
  group_by(test.preparation.course, gender) %>%
  summarise(Math = mean(math.score),
            Writing = mean(writing.score),
            Reading = mean(reading.score)) %>%
  as.data.frame() %>%
  melt(id = c("test.preparation.course", "gender")) %>%
  ggplot() +
  geom_point(aes(x = test.preparation.course, y = value, col = variable)) +
  geom_line(aes(x = test.preparation.course, y = value, col = variable, linetype = variable,
                group = variable), size = 1) +
  facet_wrap(~gender) +
  theme_bw() + ylab("Average") +
  theme(legend.position = "bottom")
```



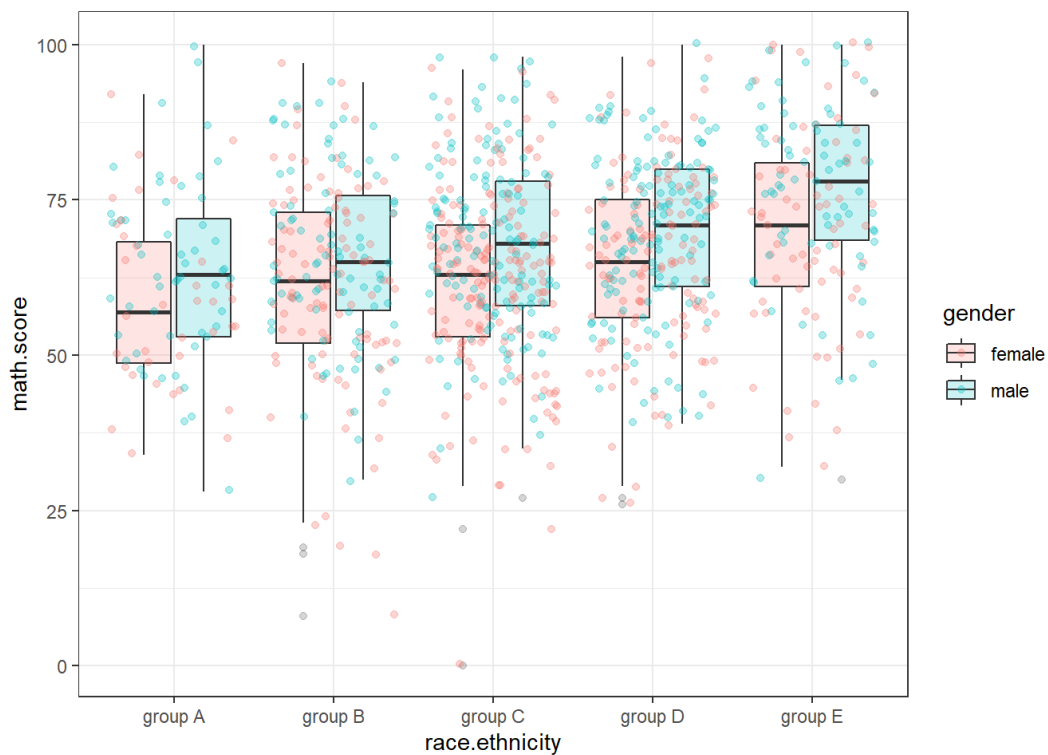
```
Students %>%
  group_by(race.ethnicity,gender) %>%
  summarise(Math = mean(math.score),
            Writing = mean(writing.score),
            Reading = mean(reading.score)) %>%
  as.data.frame() %>%
  melt(id.vars = c("race.ethnicity","gender")) %>%
  ggplot() +
  geom_boxplot(aes(x = race.ethnicity, y= value, fill = variable), alpha = 0.4) +
  theme_bw() +
  theme(legend.position = "bottom")
```



```
Students[,c("race.ethnicity", "gender", "math.score", "reading.score", "writing.score")] %>%
  melt(id.vars = c("race.ethnicity", "gender")) %>%
  ggplot() +
    geom_boxplot(aes(x = race.ethnicity, y = value, fill = gender)) +
    facet_wrap(~ variable) +
    theme_bw() +
    theme(axis.text.x = element_text(angle = 90))
```



```
# 수학점수
ggplot(Students) +
  geom_boxplot(aes(x = race.ethnicity, y = math.score, fill = gender), alpha = 0.2) +
  geom_jitter(aes(x = race.ethnicity, y = math.score, col = gender), alpha = 0.3) +
  theme_bw()
```



```

plotA = ggplot(Students) +
  geom_density(aes(x = math.score, fill = lunch), alpha = 0.4) +
  theme_bw()

plotB = ggplot(Students) +
  geom_density(aes(x = reading.score, fill = lunch), alpha = 0.4) +
  theme_bw()

plotC = ggplot(Students) +
  geom_density(aes(x = writing.score, fill = lunch), alpha = 0.4) +
  theme_bw()

D = ggarrange(plotA, plotB, plotC, ncol = 1)

annotate_figure(D, top = text_grob("Score Group by lunch"))

```

