Portfolio1

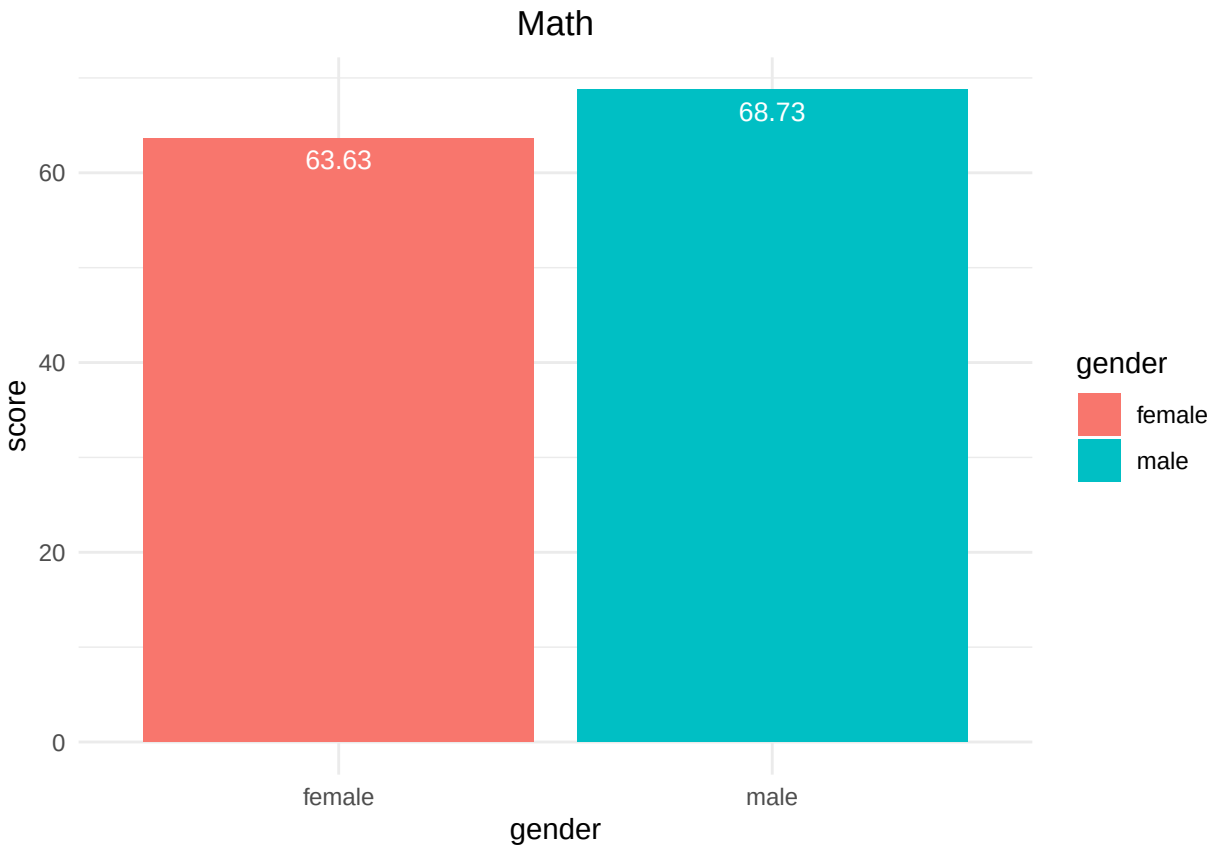
Negar

3/15/2022

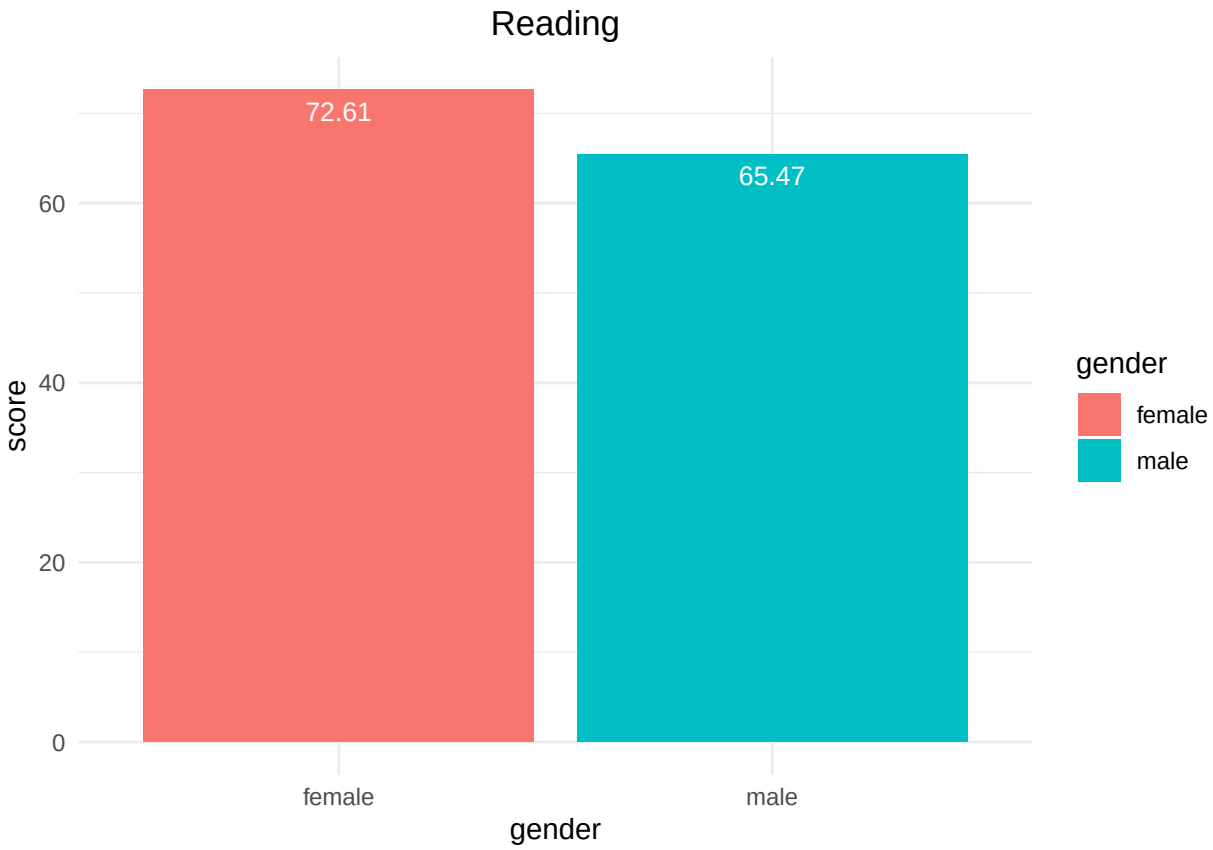
```
data = read.csv(here::here("Data","data","StudentsPerformance_renamed.csv"))
```

Firstly, let's figure out the performance of each field for males and females.

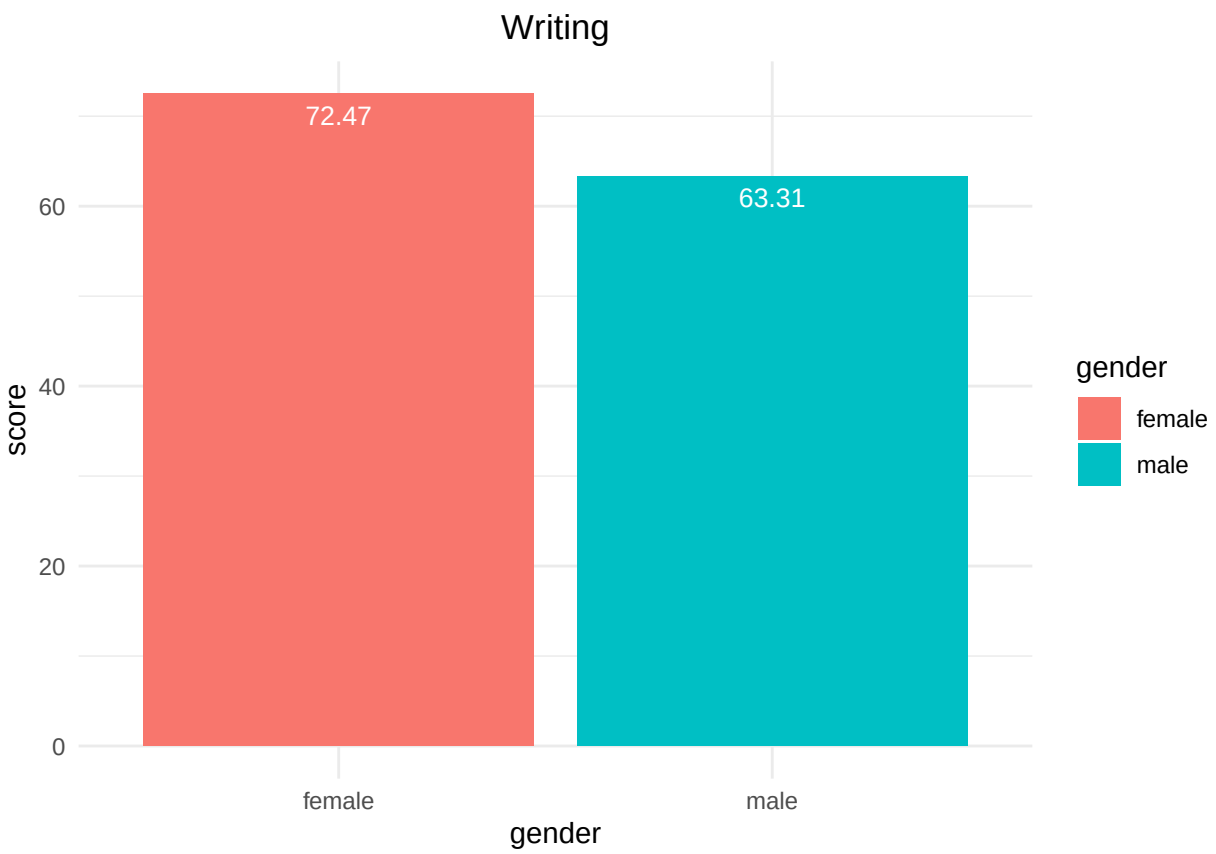
```
gender_performance <- data |> group_by(gender) |> summarise(
  math = mean(math),
  reading = mean(reading),
  writing = mean(writing))

plt1 <- ggplot(gender_performance, aes(x=gender, y=math, fill=gender)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(math, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Math")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "gender_math.tiff"), plt1,
       height = 6, width = 6)
plt1
```




```
plt2 <- ggplot(gender_performance, aes(x=gender, y=reading, fill=gender)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(reading, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Reading")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "gender_reading.tiff"), plt2,
       height = 6, width = 6)
plt2
```

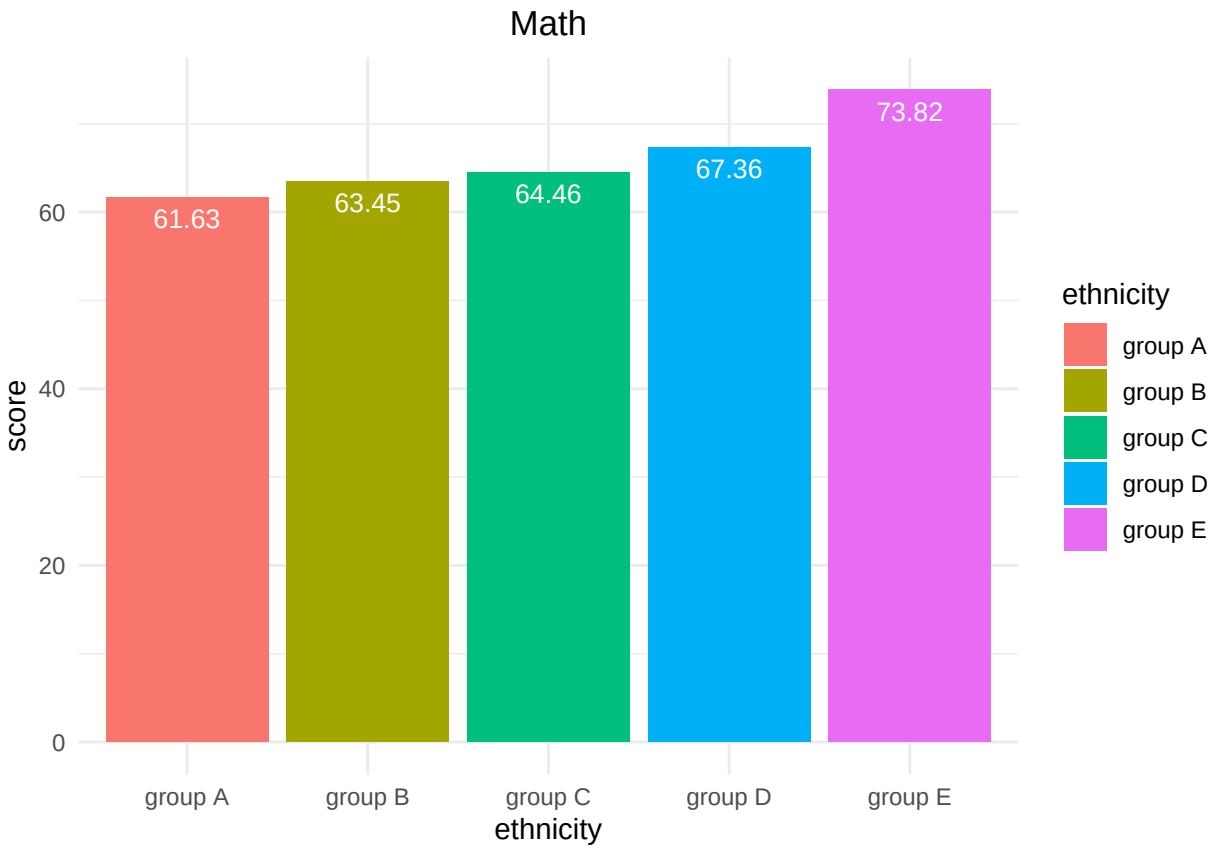



```
plt3 <- ggplot(gender_performance, aes(x=gender, y=writing, fill=gender)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(writing, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Writing")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "gender_writing.tiff"), plt3,
       height = 6, width = 6)
plt3
```

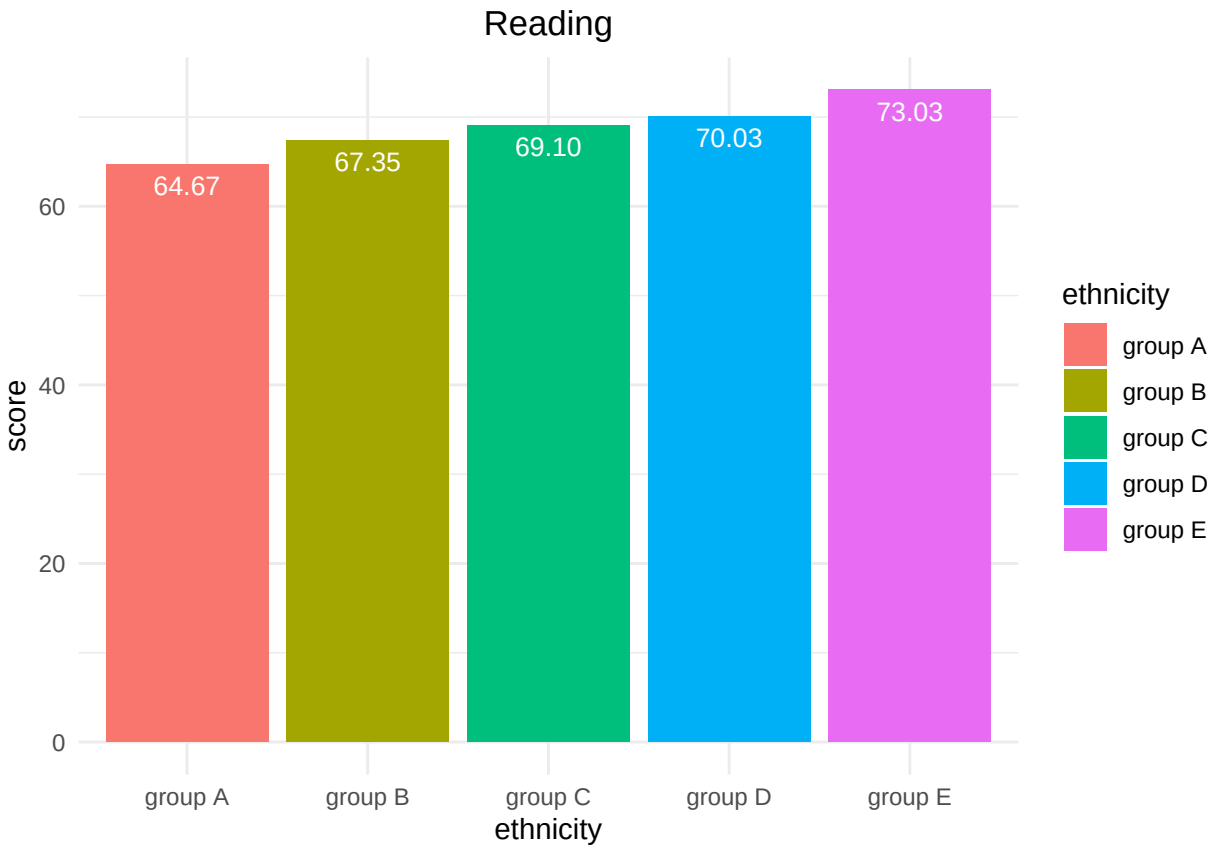


We can see that male has better performance on math field, but worse on reading and writing. Secondly, see the performance of ethnicity.

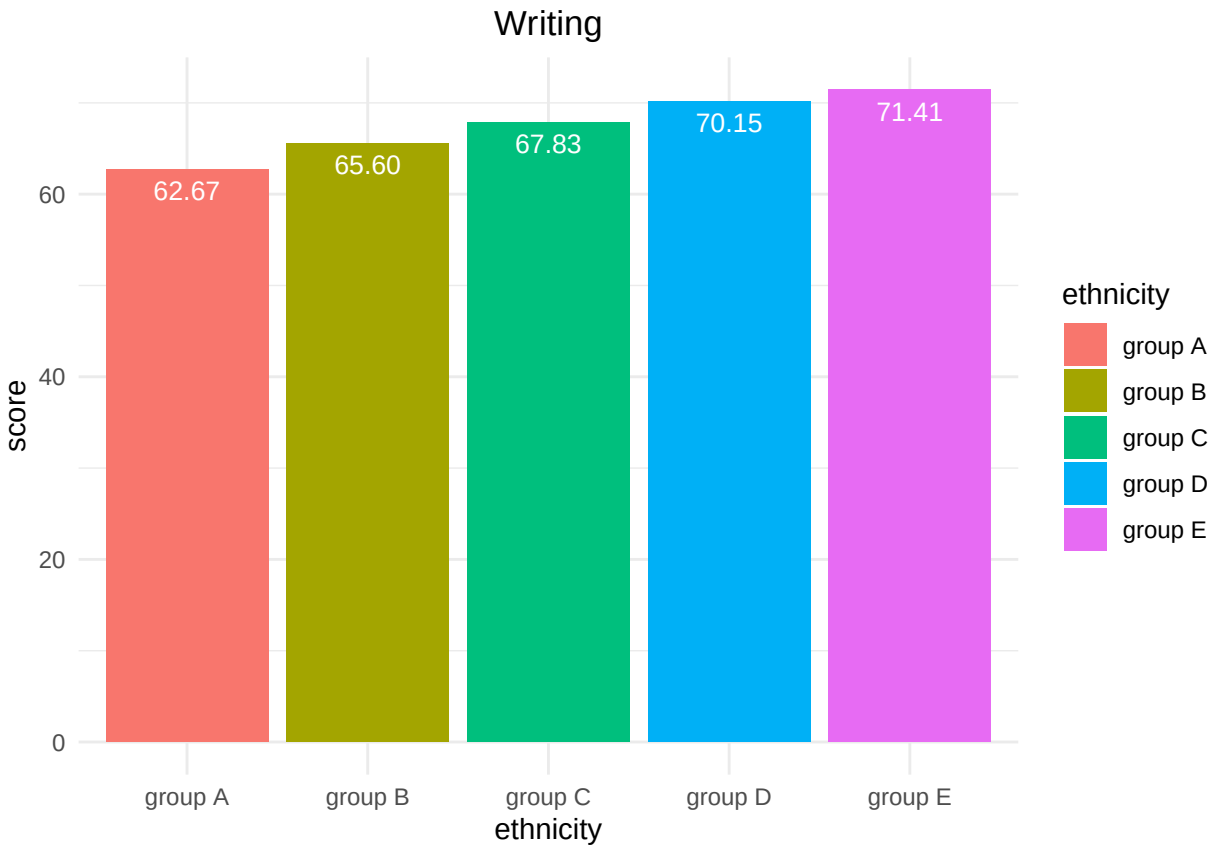
```
ethnicity_performance <- data |> group_by(ethnicity) |> summarise(
  math = mean(math),
  reading = mean(reading),
  writing = mean(writing))

plt1_1 <- ggplot(ethnicity_performance, aes(x=ethnicity, y=math, fill=ethnicity)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(math, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Math")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "ethnicity_math.tiff"), plt1_1,
       height = 6, width = 6)
plt1_1
```




```
plt2_2 <- ggplot(ethnicity_performance, aes(x=ethnicity, y=reading, fill=ethnicity)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(reading, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Reading")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "ethnicity_reading.tiff"), plt2_2,
       height = 6, width = 6)
plt2_2
```

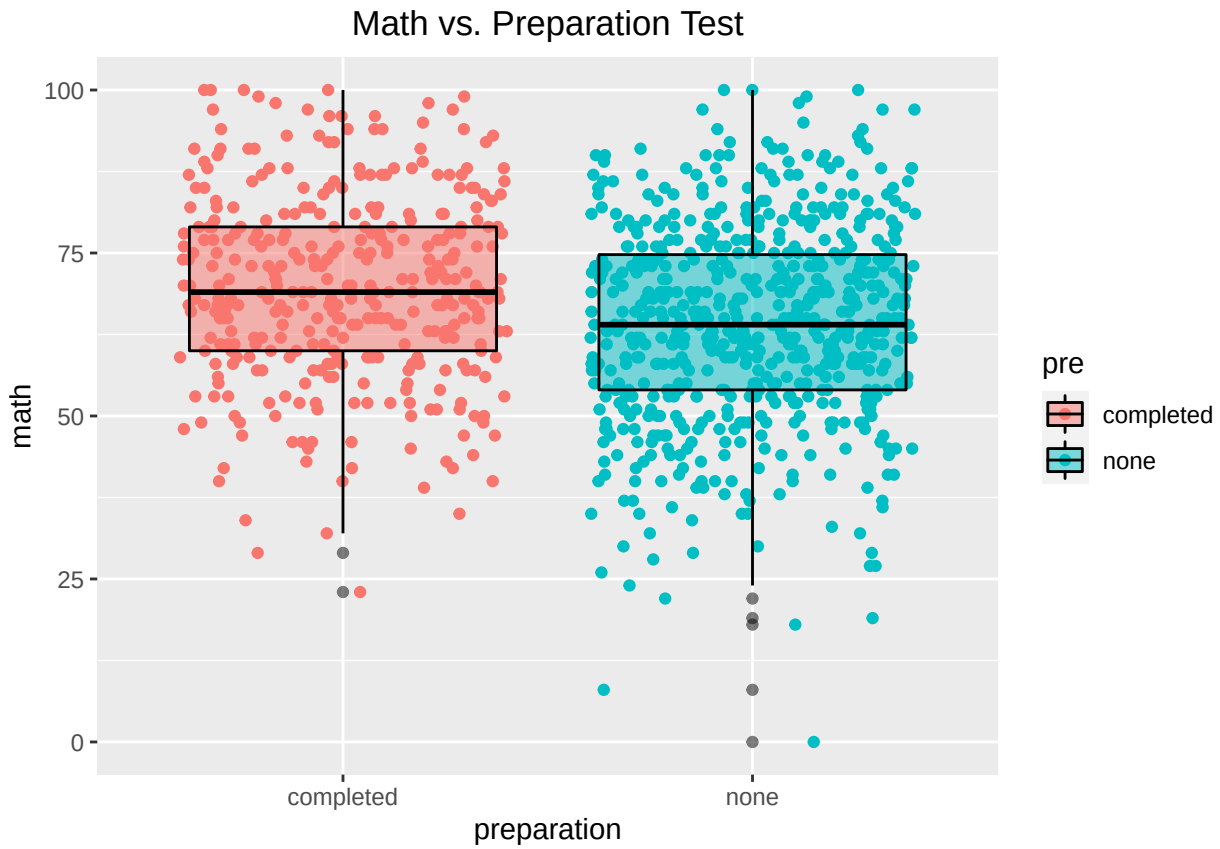



```
plt3_3 <- ggplot(ethnicity_performance, aes(x=ethnicity, y=writing, fill=ethnicity)) +
  geom_bar(stat="identity")+
  geom_text(aes(label=formattable(writing, digits = 2, format = "f")),
            vjust=1.6, color="white", size=3.5)+
  theme_minimal()+
  ylab("score")+
  ggtitle("Writing")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "ethnicity_writing.tiff"), plt3_3,
       height = 6, width = 6)
plt3_3
```

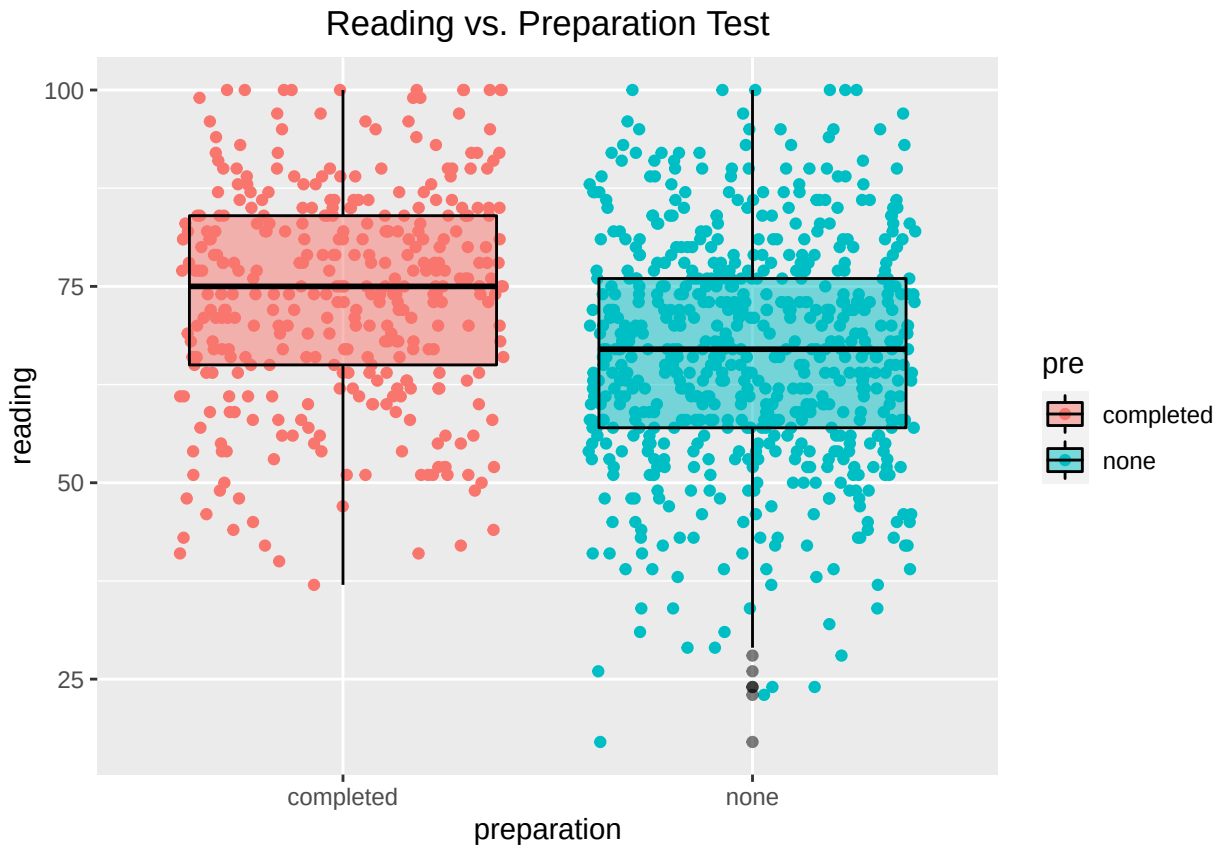


Group E, obviously, has best performance for all the fields, and group A is the worst. Then, let's see the result of score and test preparation.

```
plt1_1_1 <- ggplot(data) +
  aes(x = pre,
      y = math,
      fill = pre,
      color = pre) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  xlab("preparation")+
  ggtitle("Math vs. Preparation Test")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "preparation_math.tiff"), plt1_1_1,
       height = 6, width = 6)
plt1_1_1
```



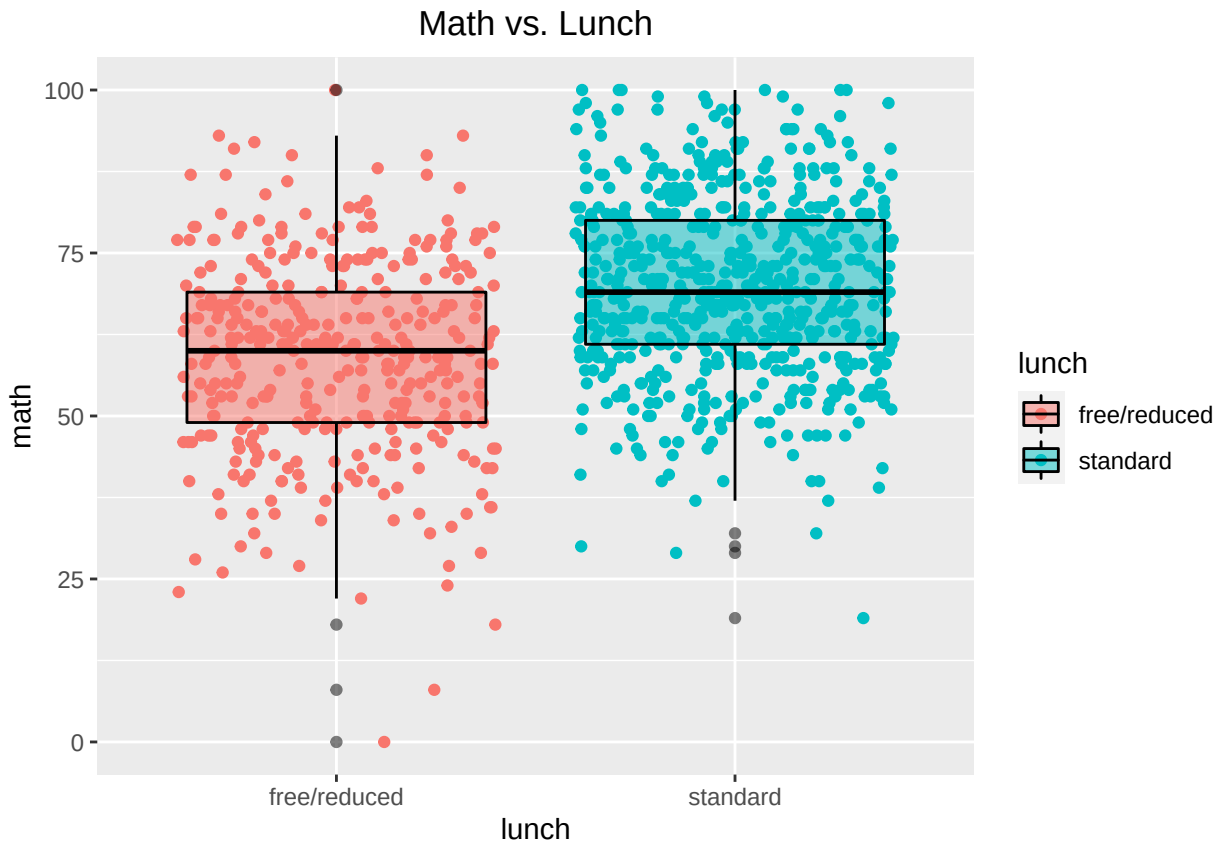

```
plt2_2_2 <- ggplot(data) +
  aes(x = pre,
      y = reading,
      fill = pre,
      color = pre) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  xlab("preparation")+
  ggtitle("Reading vs. Preparation Test")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "preparation_reading.tiff"), plt2_2_2,
       height = 6, width = 6)
plt2_2_2
```



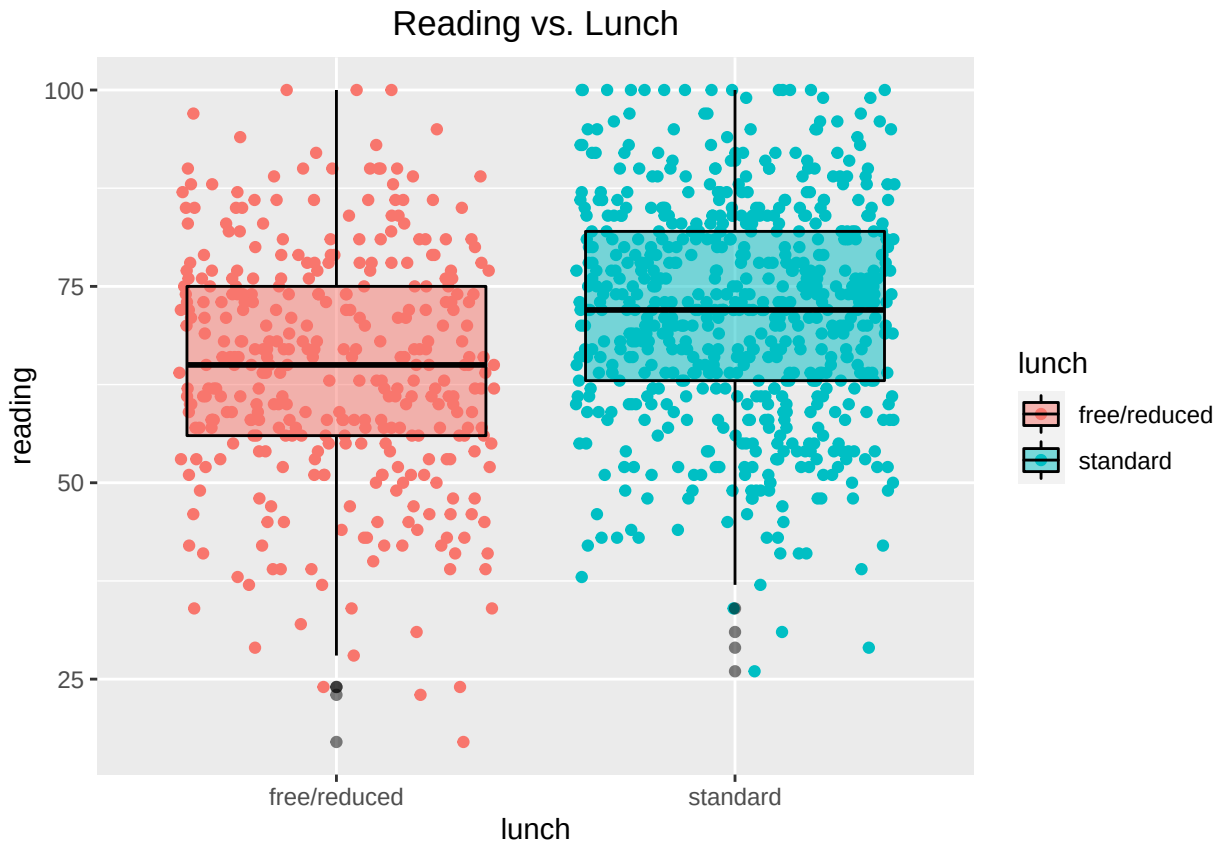

```
plt3_3_3 <- ggplot(data) +
  aes(x = pre,
      y = writing,
      fill = pre,
      color = pre) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  xlab("preparation")+
  ggtitle("Writing vs. Preparation Test")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "preparation_writing.tiff"), plt3_3_3,
       height = 6, width = 6)
plt3_3_3
```

The score distribution got narrower if students complete the preparation before test, and also we can see that the average of the score is better.

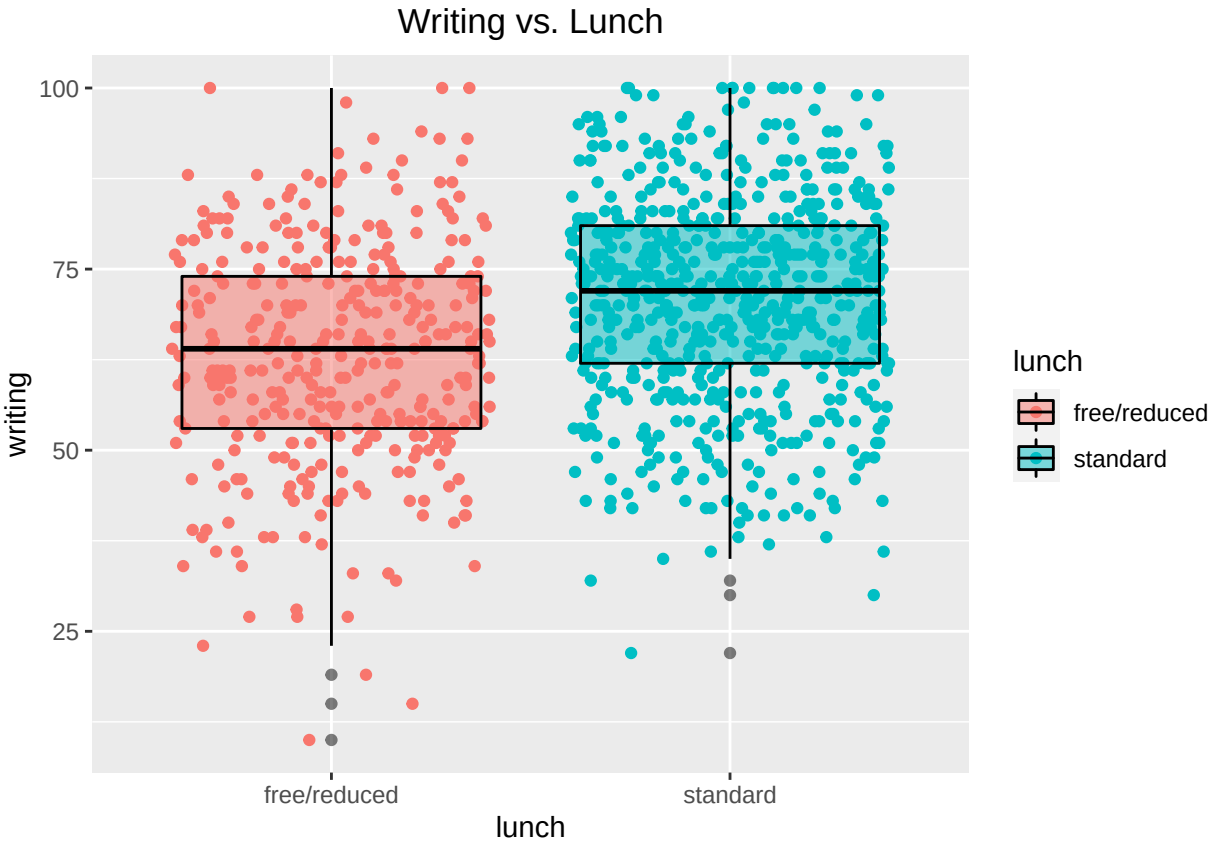
```
plt1_1_1_1 <- ggplot(data) +
  aes(x = lunch,
      y = math,
      fill = lunch,
      color = lunch) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  ggtitle("Math vs. Lunch")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "lunch_math.tiff"), plt1_1_1_1,
       height = 6, width = 6)
plt1_1_1_1
```




```
plt2_2_2_2 <- ggplot(data) +
  aes(x = lunch,
      y = reading,
      fill = lunch,
      color = lunch) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  ggtitle("Reading vs. Lunch")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "lunch_reading.tiff"), plt2_2_2_2,
       height = 6, width = 6)
plt2_2_2_2
```




```
plt3_3_3_3 <- ggplot(data) +
  aes(x = lunch,
      y = writing,
      fill = lunch,
      color = lunch) +
  geom_jitter(height = 0,
              width = .4) +
  geom_boxplot(color = "black",
               alpha = .5)+
  ggtitle("Writing vs. Lunch")+
  ggeasy::easy_center_title()
ggsave(here::here("Output", "Figures", "lunch_writing.tiff"), plt3_3_3_3,
       height = 6, width = 6)
plt3_3_3_3
```



As we expect, students are easier to get better score once they eat standardly.