Analysis of Mean Test Score

Anoushka Khayar, Gayatri Pant, and Subha Raut

Youngstown State University

STAT 6949

ABSTRACT

The main purpose of this paper is to determine the interaction between race and test preparation and their interaction on the mean score of the students. If there exists significant interaction, we simply use simple effects. Simple effects are comparisons of the cell means across levels of one factor for some or all levels of the other factor (Montgomery, 2021). If there doesn't exist interaction then, we use main effects. The main effects are comparisons of marginal means for one of the factors (Montgomery, 2021). Test preparation and Race are considered as the two fixed factors in our project. Test preparation consists of two categories (completed and none) whereas race consists of five levels (A, B, C, D, and E). There were three scores in the dataset: Reading scores, writing scores, and Math scores. We created a new variable and calculated the mean of three scores called "Mean" which represents the mean score of the students. The mean scores of the students were considered as the response variable.

A 2-factorial ANOVA test was performed to determine the interaction between race and test preparation and its effect on the mean score of students. We further conducted a multiple comparison test to meet the final conclusions. We analyzed the data using R statistical software throughout the procedure. The study found that there was no significant interaction between race and test on the mean score of the students (p-value =0.775>0.05). However, we can say that the race (p-value =2e-16<0.05) and the test(p-value=3.28e-07<0.05) individually have a significant effect on the mean score of the students. Thus, from Tukey's post hoc test, we concluded that there is a significant difference in mean groups between race and mean groups between tests.

Keyword: Race, Test preparation, interaction, main effect

Contents

References

1. Introduction

The results of a student's standardized high school tests have long been a deciding element in their college entrance. Higher test scores boost a student's chances of being accepted to school and of being eligible for merit-based scholarships. GPAs in high school are frequently viewed as a significant component, whereas test scores are viewed as equivalent. Because all students are graded on the same task, standardized test scores are thought to be more reliable. Various test preparation courses have been provided to high school students over the years in order to assist them improve their test scores. Despite the growth of test preparation classes, the question of whether they have a positive impact on a student's average score continues to be debated. Test preparation has led to a small boost in student's score. Research has demonstrated a favorable increase for students who took the ACT test a second time with test preparation, particularly for the ACT test. Various studies have also revealed a mixed result depending on the demographic being tested for the variation in test average score.

The purpose of this study is to see how the race and test preparation affect the mean score obtained by high school students. We investigated whether the completion of test preparation courses, is likely to affect the differences in average scores among different race groups. To do this, we examined students' average score of Reading, Writing, and Math score they obtained in their Standardized tests. We also examined whether certain race groups are more likely than others to obtained higher mean scores or whether these results are consistent across all race groups. For this study, we assumed that all the students completed the test preparation course.

1.1 Hypothesis

1.1.1 We wish to answer the following questions.

a. Do the factors, "Test Prep" and "Race" interact?
b. What effects do "Test Prep" and "Race" have on mean test scores?
c. Does completing a test prep result in higher mean score?

2.Methodology

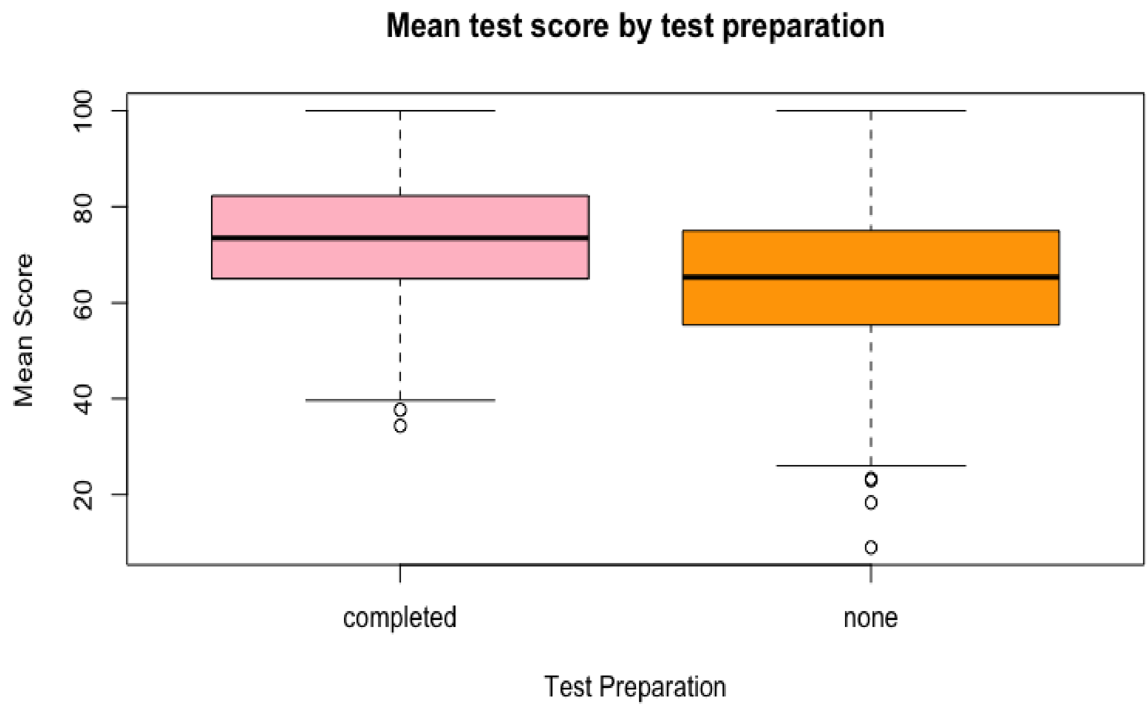
2.1 Data

The dataset was obtained from the Kaggle website which consist of data related to students form high school and their scores in Reading, Writing and Mathematics. There are total of 8 columns and 1000 rows in the original dataset. The columns are Gender, Race/Ethnicity, Parental education, Lunch, Test preparation, Math score, reading score, Writing score. The data didn't consist of any null values. We added a column "Mean" with the average scores from the 3 subject score column. The variables of interest in this study are race/ethnicity and test preparation. We will be performing two factor ANOVA method to determine the impact of race/ethnicity and test preparation on student's mean scores

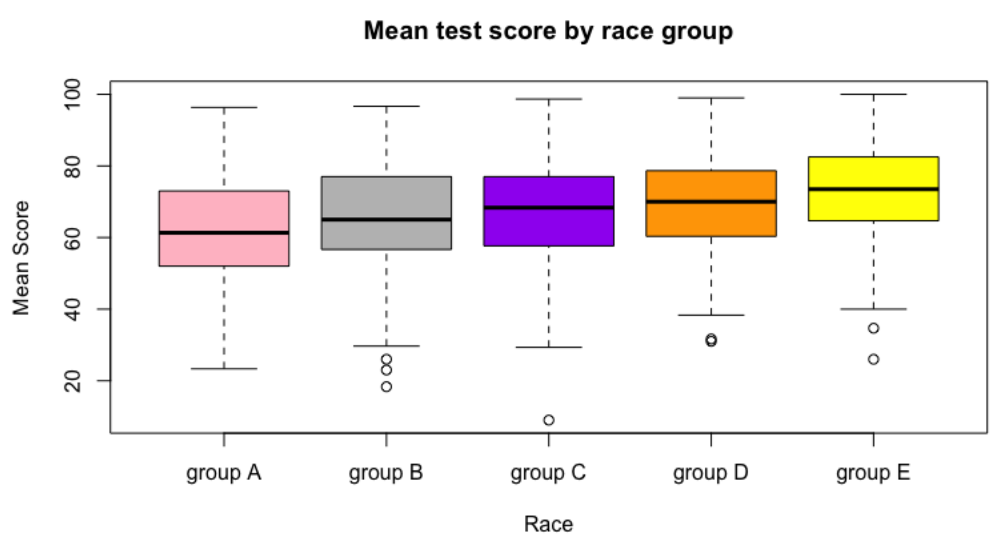
3.Descriptive analysis

3.1 Descriptive analysis for test

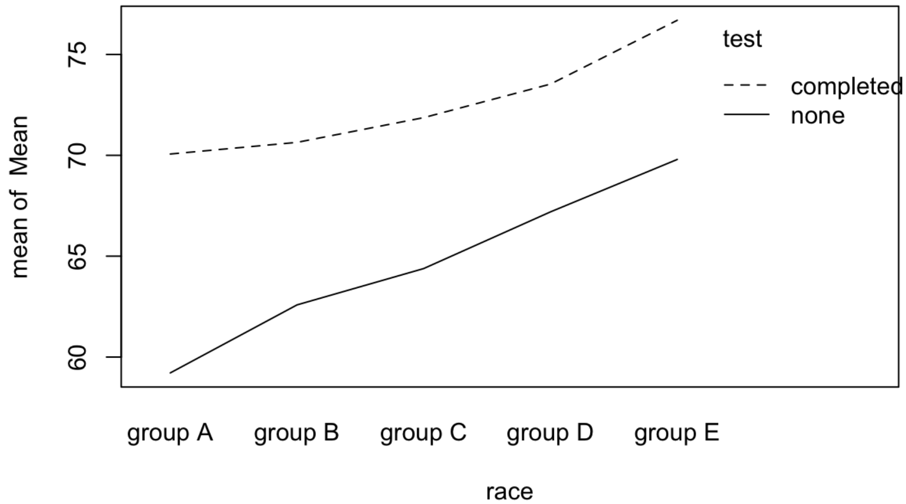
Test Preparation	Mean	Median	Standard deviation
Completed	72.67	73.50	
None	65.04	65.33	



3.2 Descriptive analysis for race



3.3 Interaction Plot



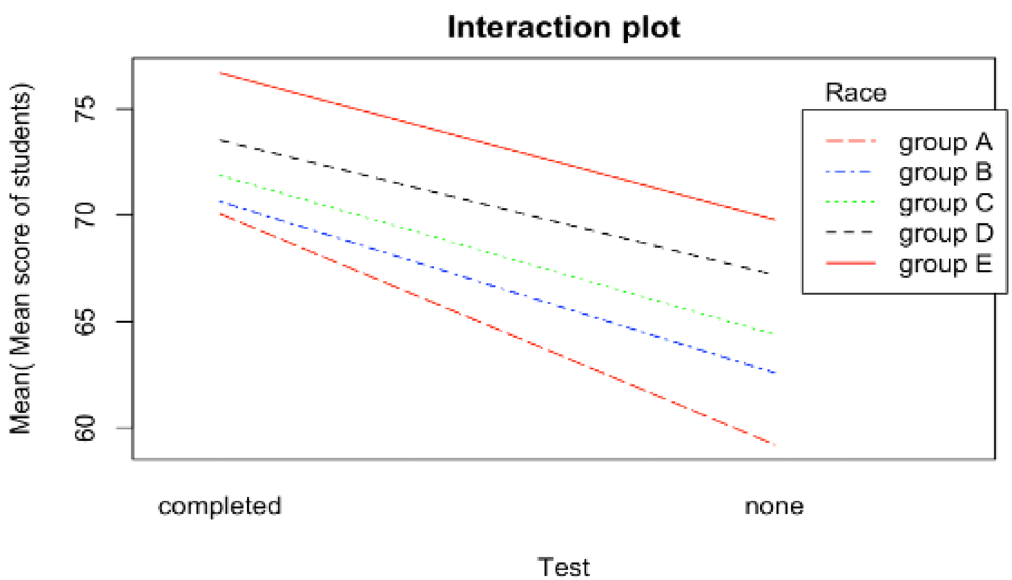


Figure 3.3: Interaction plot of the Variables

3.4 Main effects of test and race

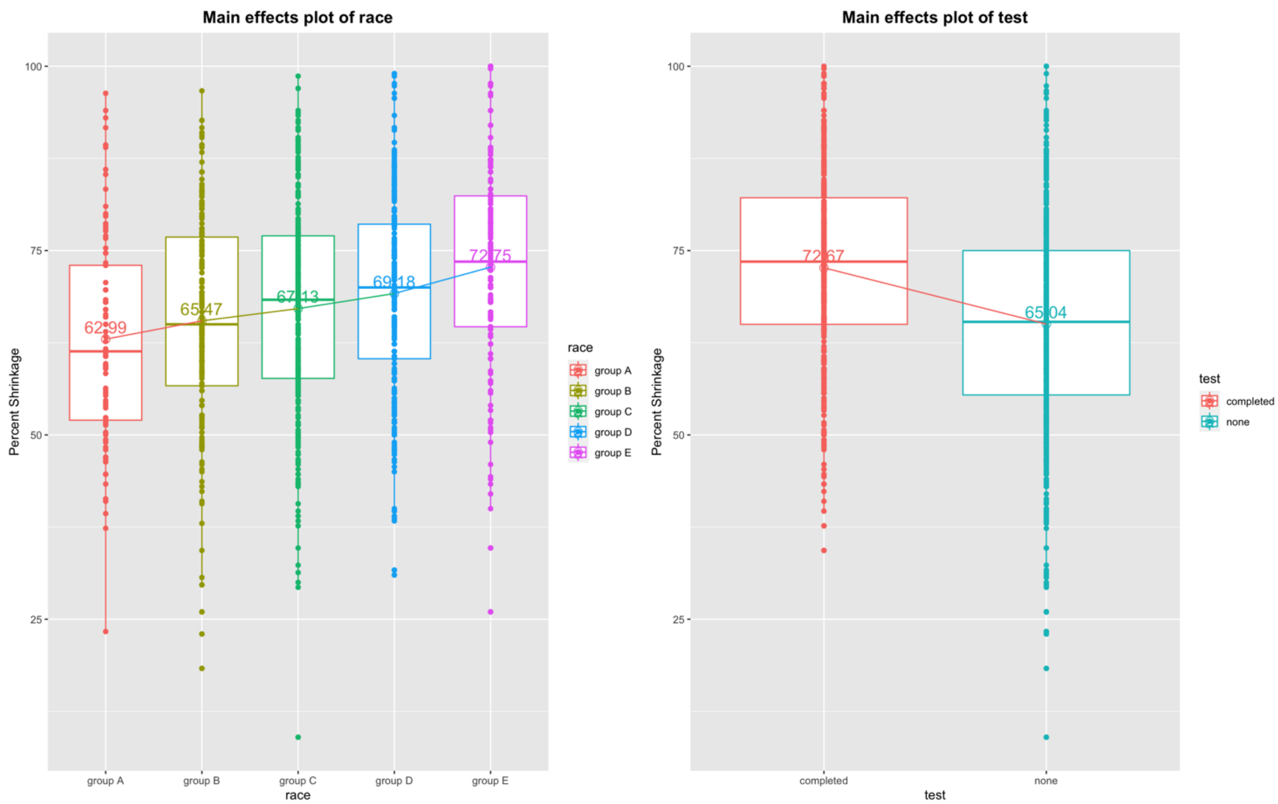


Figure 3.4: Main effects plot of race and test

4. Analysis

4.1 A Two-Factor Fixed Effects Model

Five levels of race were used in this experiment: A,B,C,D,E. Also, two levels of test were used; completed, non-compleceted. Race and test both were considered as our fixed variable. So, we fit a two-factor fixed effects model. The model is:

$Y_{ijk} = \mu_i + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$ Where, $i = 1,\ldots,a$, $j = 1,\ldots,b$, $k = 1,\ldots,n$

- μ is the overall mean.
- τ_i is the effect of the ith level of factor A.
- β_j is the effect of the jth level of factor B.
- $(\tau\beta)_{ij}$ is the effect of interaction between race and test. We know that if interaction exists, we cannot rely on the p-values for the race or for the test prep.
- $\varepsilon_{ijk} \sim$ iid $N(0,\sigma^2)$.

4.2 The F-test

Testing whether factors A(test) and B(race) interact:

$H_0 : (\tau\beta)_{ij} = 0$ for all i, j

H_1 : Atleast one $(\tau\beta)_{ij} \neq 0$

If there is no significant interaction, then we test the equality of effects of test and race.

• Testing the equality of the effects of factor A(test):

$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$

H_1 : At least one $\tau_i \neq 0$

• Testing the equality of the effects of factor B(race):

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$

H_1 : At least one $\beta_b \neq 0$

4.3 The results from the two-factor ANOVA are listed below

Table 4.1: The results from the two-factor ANOVA are listed below

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
race	4	7164	1791	9.707	1.06e-07 ***
test	1	12925	12925	70.054	< 2e-16 ***
race:test	4	329	82	0.446	0.775
Residuals	990	182651	184		

Table (4.1) shows the ANOVA findings in conjunction with the interaction and main effect plots. We then looked to determine if there was a substantial relationship between the test and race. The null hypothesis is not rejected because the p-value is 0.775. This means we'll argue that there isn't any interaction and won't include it in our model. We can check the equivalence of the effects of test and the equality of effects from race because there is no interaction. Race has a p-value of 1.06e-07. Because this is less than 0.05, we will reject the null hypothesis and adopt the alternative hypothesis, which states that the impact race was not zero. The test has a p-value less than 2e-16 so we will reject the null hypothesis and support the alternative hypothesis, that the test is not equal to zero, because this is nearly equal to zero than 0.05. In our model, we'll additionally consider the race and test. This means that both the race and test have an impact on the Mean.

In the next section, we run the Tukey's test. The purpose of Tukey's test is to figure out which groups in our sample are differ. It is, however, a number that represents the distance between groups, to compare every mean with every other mean.

4.4 Tukey test for test and race

We used a Tukey test to compare the means from the test and race.

4.4.1 Tukey test for test

Table 4.2: Tukey test for test

Variable Name	diff	lwr	upr	p adj
group B-group A	2.475912	-2.2920516	7.243875	0.6154456
group C-group A	4.139152	-0.3106668	8.588971	0.0823759
group D-group A	6.186880	1.6326966	10.741063	0.0020177
group E-group A	9.759872	4.7276319	14.792111	0.0000014
group C-group B	1.663240	-1.7384075	5.064888	0.6685986
group D-group B	3.710968	0.1738924	7.248044	0.0342884
group E-group B	7.283960	3.1495003	11.418420	0.0000168
group D-group C	2.047728	-1.0471583	5.142614	0.3693360

group E-group C 5.620720 1.8575864 9.383853 0.0004631

group E-group D 3.572992 -0.3129920 7.458975 0.0886021

From the Table (4.2) we got the confidence interval for the mean of group B-group A is [-2.2920516, 7.243875]. Because this interval includes zero and the associated p-value is 0.6154456, we can say that there is no significant difference in the means between group B and group A.

The confidence interval for group C-group A is [-0.3106668, 8.588971]. Because this interval includes zero and the associated p-value is 0.0823759, we can say that there is no significant difference in the means between group C and group A.

The confidence interval for group D-group A is [1.6326966, 10.741063]. This interval does not include zero and the associated p-value is 0.0020177 which is less than 0.05 alpha level, we can say there is significant difference in means between group D and group A.

The confidence interval for the mean of group E-group A is [4.7276319, 14.792111]. Because this interval does not include zero and the associated p-value is 0.0000014, we can say that there is significant difference in the means between group E and group A.

The confidence interval for group C-group B is [-1.7384075, 5.064888]. Because this interval includes zero and the associated p-value is 0.6685986, we can say that there is no significant difference in the means between group C and group B.

The confidence interval for group D-group B is [0.1738924, 7.248044]. Because this interval does not include zero and the associated p-value is 0.0342884, we can say that there is significant difference in the means between group D and group B.

The confidence interval for group E-group B is [3.1495003, 11.418420]. This interval does not include zero and the associated p-value is 0.0000168 which is less than 0.05 alpha level, we can say there is significant difference in means between group E and group B.

The confidence interval for the mean of group D-group C is [-1.0471583, 5.142614]. Because this interval includes zero and the associated p-value is 0.369336, we can say that there is no significant difference in the means between group D and group C.

The confidence interval for group E-group C is [1.8575864, 9.383853]. Because this interval does not include zero and the associated p-value is 0.0004631, we can say that there is a significant difference in the means between group E and group C.

The confidence interval for group E-group D is [-0.3129920, 7.458975]. Because this interval includes zero and the associated p-value is 0.0886021, we can say that there is no significant difference in the means between group E and group D.

4.4.2 Tukey test for race

Table 4.3: Tukey test for race

Variable Name	diff	lwr	upr	p adj
none-completed	-7.478328	-9.236508	-5.720148	0

The confidence interval for the mean of none−completed is [-9.236508 -5.720148]. Because this interval does not include zero and the associated p-value is 0, we can say that there is significant difference in the means.

4.5 Model Adequacy

4.5.1 Normal Quantile-Quantile plot.

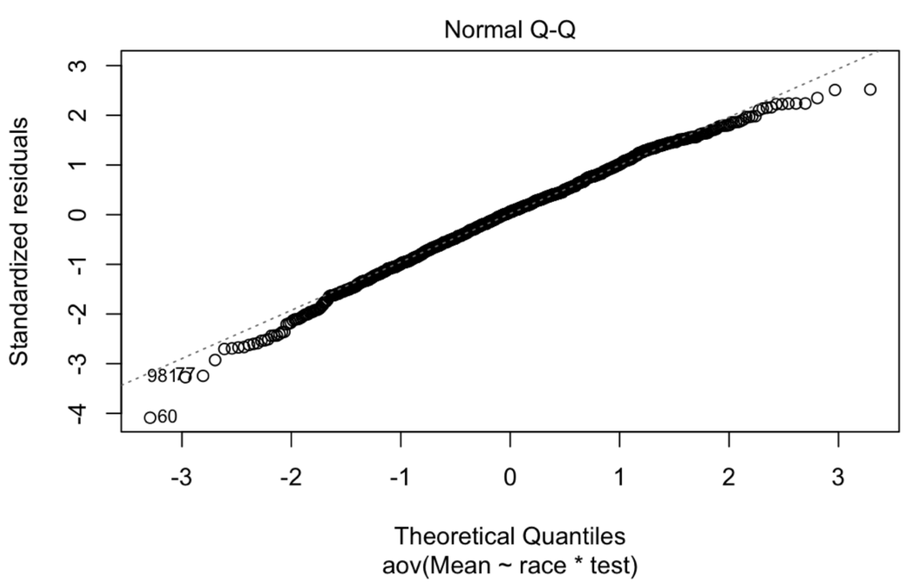


Figure 4.1: Normal Quantile-Quantile plot

With reference to the Figure (4.1) above it can be depicted that the normality assumption seems to be verified, a good number of points are along the line, also we will run a shapiro test to see if we end up with same result.

Shapiro-Wilk normality test data: W = 0.99369, p-value = 0.0003106

From the shapiro test we got the p-value 0.0003106, since the p-value is less than 0.05 it gives statistical evidence that the normality assumption is not satisfied.

4.6 Constant variance assumption

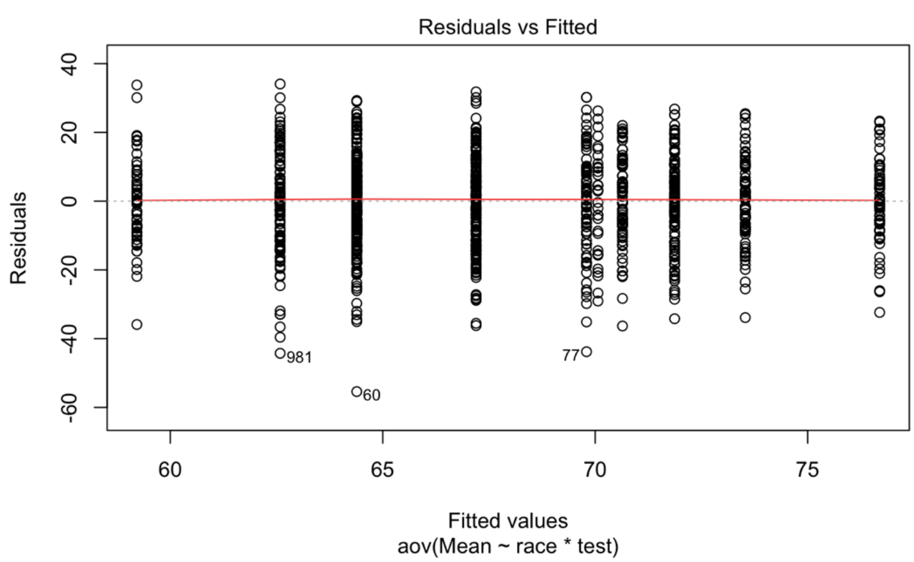


Figure 4.2: Constant variance plot

5. Conclusion and recommendation

```
---

title: "Untitled"

author: "Anoushka Khayar"

date: "4/18/2022"

output:

  word_document: default

  html_document: default

---


```{r setup, include=FALSE}

knitr::opts_chunk$set(echo = TRUE)

```



```{r library}

library(readxl)
```
```

```
mydata0 <- read_excel("~/Desktop/StudentsPerformance.xlsx")

View(mydata0)

mydata0$race = as.factor(mydata0$race)

mydata0$test = as.factor(mydata0$test)

mydata0$gender = as.factor(mydata0$gender)
```

```{r }
### Data display and summary


## marginal and cell means

tapply(mydata0$Mean,mydata0$race,mean, data=mydata0) # marginal mean for fabric

tapply(mydata0$Mean,mydata0$test,mean, data=mydata0)   # marginal mean for temperature

tapply(mydata0$Mean, list(mydata0$race,mydata0$test), mean) #cell mean


p_fabric1 = ggplot(mydata0,aes(x=race,y=Mean,color=race)) +

  geom_boxplot() +   # add a boxplot

  geom_point() +     # add a scatterplot

  stat_summary(fun=mean, geom="point", shape=21, size=3) + # add means

  stat_summary(fun=mean, geom="line", aes(group=1))  + # join means to draw a mean line plot

  stat_summary(aes(label=round(..y..,2)), fun=mean, geom="text", size=5,
```

```
            vjust = -.5) +
  labs(title ="Main effects plot of race",
       x = "race", y="Percent Shrinkage") +   # add title and axis labels
  theme(plot.title = element_text(size=14, hjust=.5,face="bold"))

p_temp1 = ggplot(mydata0,aes(x=test,y=Mean,color=test)) +
  geom_boxplot() +   # add a boxplot
  geom_point() +     # add a scatterplot
  stat_summary(fun=mean, geom="point", shape=21, size=3) + # add means
  stat_summary(fun=mean, geom="line", aes(group=1))  + # join means to draw a mean line plot
  stat_summary(aes(label=round(..y..,2)), fun=mean, geom="text", size=5,
            vjust = -.5) +
  labs(title ="Main effects plot of test",
       x = "test", y="Percent Shrinkage") +   # add title and axis labels
  theme(plot.title = element_text(size=14, hjust=.5,face="bold"))



library("gridExtra")
grid.arrange(p_fabric1, p_temp1, ncol=2)
## interaction plot
interaction.plot(x.factor=mydata0$test,     # variable to plot on x-axis
                 trace.factor=mydata0$race, # variable to specify "traces"
```

```
        response=mydata0$Mean,     # variable to plot on y-axis
        ylab="Mean",
        xlab="Test",
        main="Interaction plot",
        col=c("red","blue","green","black"),
        trace.label="Race",      # label for legend
        legend=T,
        leg.bty="o",  # put a box around the legend
        fixed=T
)

## conditional plot

coplot(Mean ~ race | test, show.given=T,
    xlab=c("Race","Test"),
    panel=panel.smooth,rows=1,data=mydata0)

coplot(Mean ~ test | race, show.given=T,
    xlab=c("Test","Race"),
    panel=panel.smooth,rows=1,data=mydata0)
```

###

```
### ANOVA

```{r ANOVA}

results=aov(Mean~test*race, data=mydata0)

##options(show.signif.stars=F)

summary(results)


# Compute tables of results from an aov model fit

model.tables(results,type="means",se=T)

```

###

model adequacy checking

```{r adr}

## normality

qqnorm(results$res)

qqline(results$res)

shapiro.test(results$res)


# constant variance

plot(results$fitted,results$res, xlab="Fitted",ylab="Residuals")
```
```

```
library("car")

# Levene's test does not work here - within-group variability is 0.

# see explanation below

leveneTest(Mean~test*race, data=mydata0)




# Plot of residuals by test

plot(as.numeric(mydata0$test),results$residuals,

    xlab="Test", ylab="Residuals")

leveneTest(Mean~test,data=mydata0)


# Plot of residuals by fabric

plot(as.numeric(mydata0$race),results$residuals,

    xlab="Race", ylab="Residuals")

leveneTest(Mean~race,data=mydata0)

```
```
```{r tuket}

## test main effects if no interaction

# main effects of test

## TukeyHSD(results)

TukeyHSD(results, "test")

TukeyHSD(results, "race")
```
```

```
```