HW9

Helped by : 106034054

Q1

Answer 4 question for each scenario

i. Would this scenario create systematic or random error (or both or neither)?
ii. Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?
iii. Will it increase or decrease our power to reject the null hypothesis?
iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

a. only collected data from a pool of young consumers, and missed many older customers who you suspect might use the product much less every day.
   i. Systematic error
   ii. diff
   iii. Yes, it will increase the power to reject the null hypothesis.
   iv. Type II error

b. 20 of the respondents should be removed from the data.
   i. Random error
   ii. n
   iii. No
   iv. Type II error.

c. Relaxing the C.I. from 95% to 90%
   i. A systematic error
   ii. Alpha
   iii. Yes , it will increase the power to reject the null hypothesis.
   iv. Type I error.

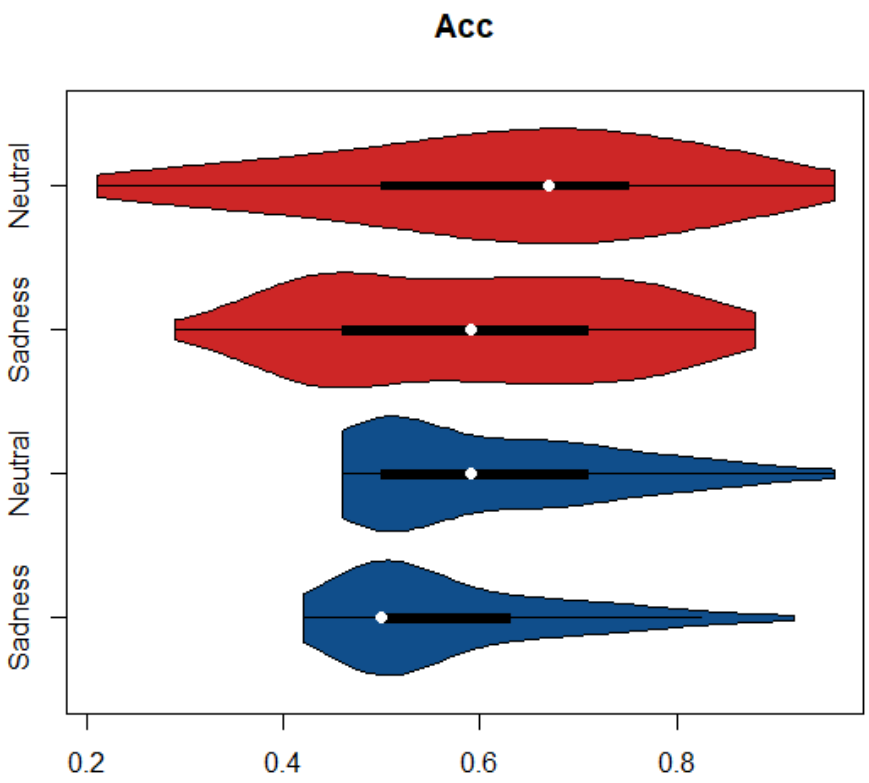
d. over-reports usage of older users.
   i. Both systematic error and random error
   ii. Diff and sd
   iii. Yes , it will decrease the power to reject the null hypothesis.
   iv. Type I error

Q2

a. Visualize the differences between blue-yellow accuracy (BY_ACC) and red-green accuracy (RG_ACC) for both the sad and neutral viewers (Emotion_Condition).

```
library(vioplot)
```

```
vioplot(BY_data[BY_data$Emotion_Condition == 'Sadness',]$ACC,
        BY_data[BY_data$Emotion_Condition == 'Neutral',]$ACC,
        RG_data[RG_data$Emotion_Condition == 'Sadness',]$ACC,
        RG_data[RG_data$Emotion_Condition == 'Neutral',]$ACC,
        main = 'Acc', names = c('Sadness','Neutral','Sadness','N
eutral'),
        col = c('dodgerblue4','dodgerblue4','firebrick3','fireb
rick3'),
        horizontal = TRUE)
```



The red violin plots mean the data from red-green color-axis, the blue ones are from blue-yellow color-axis.

b. Run a t-test (traditional) to check if there is a significant difference in blue-yellow accuracy between sad and neutral participants at 95% confidence.

```
t.test(BY_data[BY_data$Emotion_Condition == 'Sadness',]$ACC,
       BY_data[BY_data$Emotion_Condition == 'Neutral',]$ACC,
       conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  BY_data[BY_data$Emotion_Condition == "Sadness", ]$AC
C and BY_data[BY_data$Emotion_Condition == "Neutral", ]$ACC
## t = -2.0435, df = 125.61, p-value = 0.04309
## alternative hypothesis: true difference in means is not equa
l to 0
## 95 percent confidence interval:
##  -0.086308149 -0.001384159
## sample estimates:
## mean of x mean of y
## 0.5690769 0.6129231
```

From the p-value, 0.04309< 0.05 so that there is no significant difference in blue-yellow accuracy between sad and neutral participants.

c. Run a t-test (traditional) to check if there is a significant difference in red-green accuracy between sad and neutral participants at 95% confidence.

```
t.test(RG_data[RG_data$Emotion_Condition == 'Sadness',]$ACC,
       RG_data[RG_data$Emotion_Condition == 'Neutral',]$ACC,
       conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  RG_data[RG_data$Emotion_Condition == "Sadness", ]$AC
C and RG_data[RG_data$Emotion_Condition == "Neutral", ]$ACC
## t = -0.87491, df = 121.98, p-value = 0.3833
## alternative hypothesis: true difference in means is not equa
l to 0
## 95 percent confidence interval:
##  -0.08432635  0.03263405
## sample estimates:
## mean of x mean of y
## 0.5944615 0.6203077
```

From the p-value, 0.3833 > 0.05 so that there is significant difference in red-green accuracy between sad and neutral participants.

d. Do the above t-tests support a claim that there is an interaction between emotion and color axis?

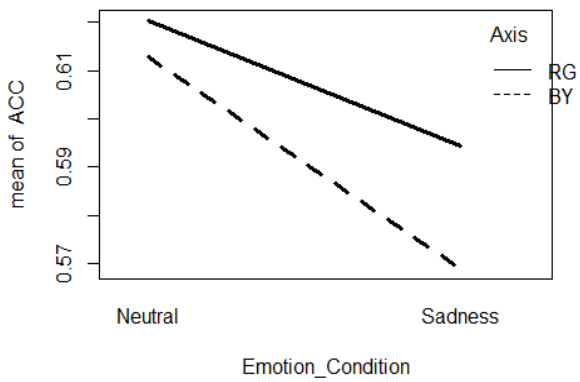
No. It is hard to find out that is there has any interaction between.

e. Running a factorial design ANOVA, are any of these three factors (emotion/color-axis/interaction) possibly influencing color perception accuracy at any meaningful level of confidence?

```
all_data <- rbind(BY_data, RG_data)
summary(aov(formula = ACC ~ Axis + Emotion_Condition + Axis:Emotion_C
ondition, data=all_data))
```

```
##                          Df  Sum Sq Mean Sq F value Pr(>F)
## Axis                      1  0.017 0.01745   0.806 0.3703
## Emotion_Condition         1  0.079 0.07893   3.644 0.0574 .
## Axis:Emotion_Condition    1  0.005 0.00526   0.243 0.6224
## Residuals               256  5.545 0.02166
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
with(all_data,
     interaction.plot(
       x.factor = Emotion_Condition,
       trace.factor =  Axis,
       response = ACC,
       lwd = 3
     ))
```



From the probability on the Axis, emotion condition and the interaction between there are all above 0.05, which are significant different. Also, the interaction seems influencing a lot in this case. By the interaction plot, the distance between two lines change that means there are interaction between.