Anova testing Accuracy

Hypothesis Testing for 8 interaction technique's Accuracy:

```
#Inputting Data: Analysis Excel sheet has "Accuracy" data of 8 interaction techniques
library(openxlsx)
library(stats)
library(gridExtra)
library(grid)
data <- read.xlsx("Analysis.xlsx",colNames = FALSE, sheet="Accuracy")
#Transpose
Accuracy<- c(data[,1], data[,2], data[,3], data[,4], data[,5],
                 data[,6], data[,7], data[,8])
techniques <- c(rep('technique1', 8),rep('technique2', 8),rep('technique3', 8),
                rep('technique4', 8),rep('technique5', 8),rep('technique6', 8),
                rep('technique7', 8),rep('technique8', 8))
combine <- cbind(Accuracy, techniques)
```

Computing Mean and Standard Deviation of 8 technique's accuracy

```
mean1<- mean(data$X1)
sd1 = sqrt(var(data$X1))
mean2<- mean(data$X2)
sd2 = sqrt(var(data$X2))
mean3<- mean(data$X3)
sd3 = sqrt(var(data$X3))
mean4<- mean(data$X4)
sd4 = sqrt(var(data$X4))
mean5<- mean(data$X5)
sd5 = sqrt(var(data$X5))
mean6<- mean(data$X6)
sd6 = sqrt(var(data$X6))
mean7<- mean(data$X7)
sd7 = sqrt(var(data$X7))
mean8<- mean(data$X8)
sd8 = sqrt(var(data$X8))

# Print in table
r1 <- c("Technique id", "Mean", "Standard Deviation")
r2 <- c("1", round(mean1, 2), round(sd1, 2))
r3 <- c("2", round(mean2, 2), round(sd2, 2))
r4 <- c("3", round(mean3, 2), round(sd3, 2))
r5 <- c("4", round(mean4, 2), round(sd4, 2))
r6 <- c("5", round(mean5, 2), round(sd5, 2))
```

```r
r7 <- c("6", round(mean6, 2), round(sd6, 2))
r8 <- c("7", round(mean7, 2), round(sd7, 2))
r9 <- c("8", round(mean8, 2), round(sd8, 2))
tab <- rbind(r1, r2, r3, r4, r5, r6, r7,r8,r9)
grid.table(tab, rows = NULL)
```

Technique id	Mean	Standard Deviation
1	0.73	0.28
2	0.53	0.34
3	0.56	0.18
4	0.45	0.17
5	0.65	0.29
6	0.29	0.13
7	0.23	0.08
8	0.21	0.07

Anova Tests : Mean Accuracy

Study Populations: We have 8 participants in total

Population 1:

Population 2:

Research hypotheses: Not all the means of the techniques are equal.

Null hypotheses: All the mean of the techniques are equal

```r
results <- aov(Accuracy ~ techniques)
anova(results)
```

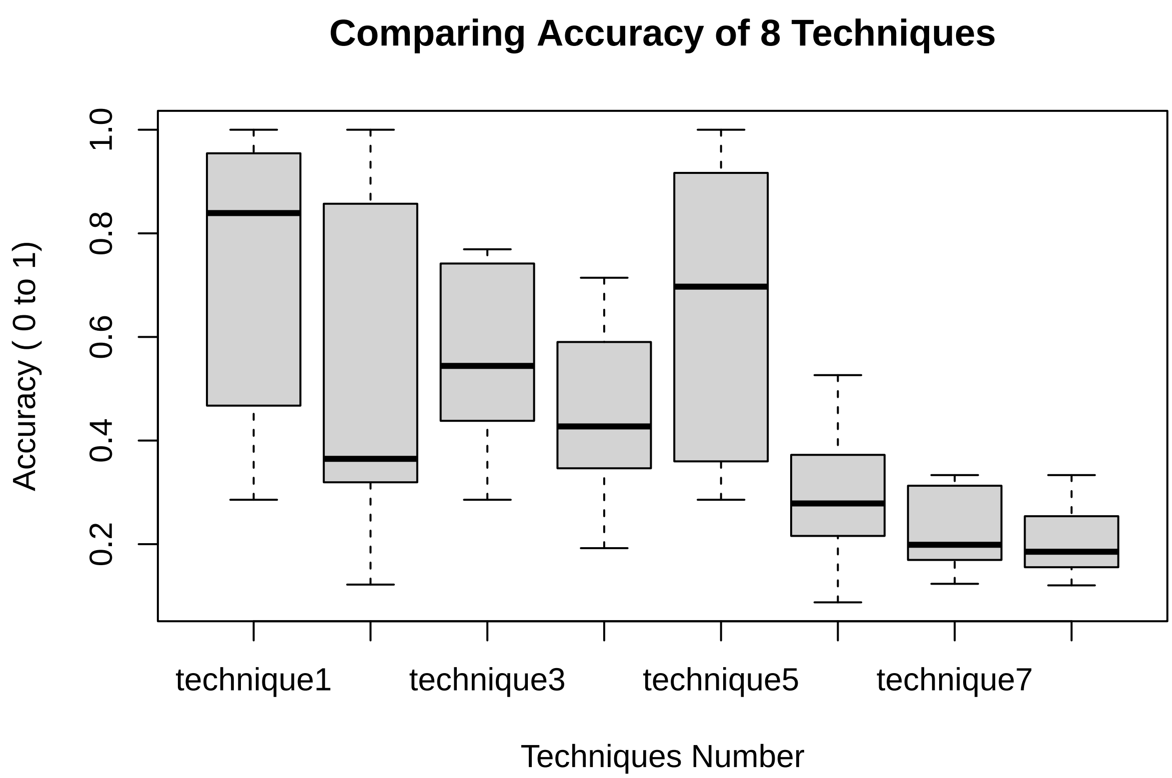
```
## Analysis of Variance Table
##
## Response: Accuracy
##             Df Sum Sq  Mean Sq F value    Pr(>F)
## techniques  7 2.1589 0.308410  6.7109 8.857e-06 ***
```

```
## Residuals  56 2.5736 0.045957
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The decision is to reject the null hypothesis, as the p-value is < 0.05. There is sufficient evidence to see that the mean values for all techniques are not all equal. At least two of these techniques has mean values that were different from each other.

Creating side by side box plots

```r
boxplot(Accuracy~techniques,main = "Comparing Accuracy of 8 Techniques",
        xlab = "Techniques Number", ylab = "Accuracy ( 0 to 1)")
```



The boxplots show that the differences in performance in terms of accuracy for techniques (1,6), (1, 7) and (1,8) due to how little of the boxplots "overlap". However, techniques 6,7,8 overlap significantly, so there will likely be some similarities. Additionally, techniques 1,2,3,4 and 5 also overlap quite a bit, indicating that we may find some similarities in those techniques as well. This is supported by the mean of techniques 1 and 3 being similar to each other, as well as the mean of the techniques 1 and 2 and 1 and 5 having similar means. It is also evident that the techniques (8,3), (6,5), (7,5) and (8,5) have more of a difference in their means.

Tukey HSD (Honest Significant Difference) post hoc comparision

A pairwise comparison technique that uses the Studentized range distribution to construct simultaneous confidence intervals for differences of all pairs of means. Studentization means dividing a mean value by its standard error. We are computing the Tukey HSD with confidence level of 0.95.

```
TukeyHSD(results, conf.level=0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Accuracy ~ techniques)
##
## $techniques
##                               diff        lwr          upr     p adj
## technique2-technique1 -0.20041054 -0.5378671  0.137046026 0.5767352
## technique3-technique1 -0.16312777 -0.5005843  0.174328797 0.7924608
## technique4-technique1 -0.27165801 -0.6091146  0.065798563 0.2027301
## technique5-technique1 -0.07186677 -0.4093233  0.265589803 0.9974577
## technique6-technique1 -0.43249030 -0.7699469 -0.095033733 0.0039106
## technique7-technique1 -0.49859589 -0.8360525 -0.161139323 0.0005165
## technique8-technique1 -0.52055202 -0.8580086 -0.183095452 0.0002552
## technique3-technique2  0.03728277 -0.3001738  0.374739340 0.9999666
## technique4-technique2 -0.07124746 -0.4087040  0.266209106 0.9975927
## technique5-technique2  0.12854378 -0.2089128  0.466000346 0.9288064
## technique6-technique2 -0.23207976 -0.5695363  0.105376810 0.3879534
## technique7-technique2 -0.29818535 -0.6356419  0.039271220 0.1204659
## technique8-technique2 -0.32014148 -0.6575980  0.017315091 0.0747836
## technique4-technique3 -0.10853023 -0.4459868  0.228926334 0.9706245
## technique5-technique3  0.09126101 -0.2461956  0.428717575 0.9890584
## technique6-technique3 -0.26936253 -0.6068191  0.068094039 0.2114199
## technique7-technique3 -0.33546812 -0.6729247  0.001988449 0.0524247
## technique8-technique3 -0.35742425 -0.6948808 -0.019967680 0.0306076
## technique5-technique4  0.19979124 -0.1376653  0.537247809 0.5805553
## technique6-technique4 -0.16083230 -0.4982889  0.176624273 0.8039265
## technique7-technique4 -0.22693789 -0.5643945  0.110518683 0.4169605
## technique8-technique4 -0.24889401 -0.5863506  0.088562554 0.3003669
## technique6-technique5 -0.36062354 -0.6980801 -0.023166967 0.0282231
## technique7-technique5 -0.42672913 -0.7641857 -0.089272557 0.0046266
## technique8-technique5 -0.44868525 -0.7861418 -0.111228686 0.0024188
## technique7-technique6 -0.06610559 -0.4035622  0.271350979 0.9985044
## technique8-technique6 -0.08806172 -0.4255183  0.249394850 0.9911454
## technique8-technique7 -0.02195613 -0.3594127  0.315500440 0.9999991
```

As we can see from the table above, techniques (6,1) has $p(=0.0039106) < 0.05$, techniques (7,1) has $p(=0.0005165) < 0.05$, techniques (8,1) has $p(=0.0002552) < 0.05$. Also, it is evident that techniques(8,3) has $p(=0.0306076) < 0.05$, techniques (6,5) has $p(=0.0282231) < 0.05$, techniques (7,5) $p(=0.0046266) < 0.05$ and techniques (8,5) $p(=0.0024188) < 0.05$ SO we can conclude that technique 1 has different mean than technique 6,7 and 8.Also, technique 5 has different mean than technique 6,7 and 8. Therefore, technique 1 and 5 is significantly more accurate than technique 6,7,8.