Websearch study

Aishat Alli

2023-04-29

1. This file describes a study in which participants were asked to find 100 distinct facts on the web using different search engines. The number of searches required and a subjective effort rating for each search engine were recorded. How many participants took part in this experiment?

```r
websearch2 = read.csv("websearch2.csv")
```

```r
websearch2$Subject = factor(websearch2$Subject) # convert to a nominal factor
websearch2$Engine = factor(websearch2$Engine)
websearch2$Order = factor(websearch2$Order) # convert to a nominal factor
summary(websearch2)
```

Answer: 30 participants took part in the experiment

```
##     Subject      Engine    Order     Searches          Effort    
##  1      : 2   Bing  :30   1:30   Min.   : 89.0   Min.   :1.00  
##  2      : 2   Google:30   2:30   1st Qu.:135.8   1st Qu.:2.00  
##  3      : 2                      Median :156.5   Median :4.00  
##  4      : 2                      Mean   :156.9   Mean   :3.90  
##  5      : 2                      3rd Qu.:175.2   3rd Qu.:5.25  
##  6      : 2                      Max.   :241.0   Max.   :7.00  
##  (Other):48
```

2. What was the average number of searches required for the search engine that had the greatest average overall?

```r
install.packages("plyr")
```

Answer: Bing search engine had the greatest average overall with average search of 165.97

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```r
library(plyr)
ddply(websearch2, ~ Engine, function(data) summary(data$Searches))
```

```
##   Engine Min. 1st Qu. Median     Mean 3rd Qu. Max.
## 1   Bing   89  144.25  164.0 165.9667  184.75  241
## 2 Google  108  131.00  147.5 147.8333  161.75  198
```

3. Conduct an order effect test on Searches using a paired-samples t-test, what is the p-value from such test?

```
install.packages("reshape2")
```

Answer: With the p-value being 0.7343, we don't have an order effect

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```
library(reshape2)
```

```
websearch2.wide.order = dcast(websearch2, Subject ~ Order, value.var="Searches") # go wide
t.test(websearch2.wide.order$"1", websearch2.wide.order$"2", paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  websearch2.wide.order$"1" and websearch2.wide.order$"2"
## t = 0.34273, df = 29, p-value = 0.7343
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -13.57786  19.04453
## sample estimates:
## mean difference
##        2.733333
```

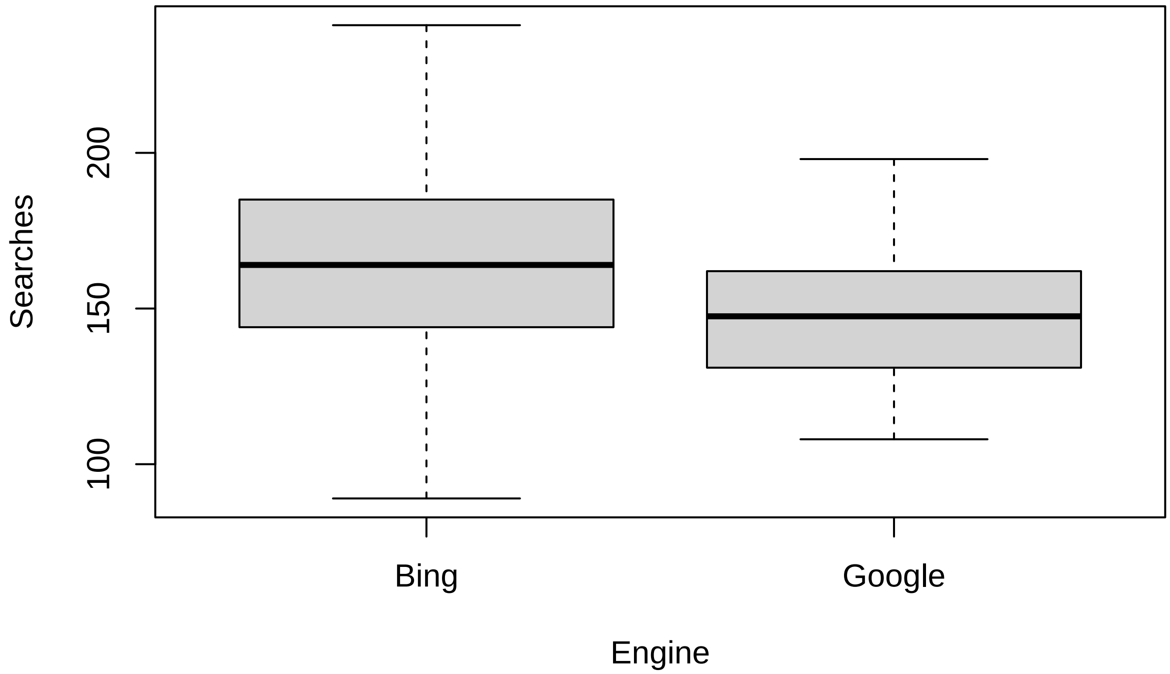
4. Conduct a paired-samples t-test on Searches by Engine. What is the absolute value of the t-statistic for such a test?

```
websearch2.wide.tech = dcast(websearch2, Subject ~ Engine, value.var="Searches") # go wide
t.test(websearch2.wide.tech$Bing, websearch2.wide.tech$Google, paired=TRUE)
```

Here we see a p value of 0.01824. So it does look less than 0.05. We have a significanct difference and we can kind of visually confirm that by looking at the box plot between the Google and Bing. So it seems that Bing has more number of searches than Google

```
##
##  Paired t-test
##
## data:  websearch2.wide.tech$Bing and websearch2.wide.tech$Google
## t = 2.5021, df = 29, p-value = 0.01824
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##   3.310917 32.955750
## sample estimates:
## mean difference
##        18.13333
```

```
plot(Searches ~ Engine, data=websearch2) # confirm
```



#### 5. Conduct a nonparametric Wilcoxon signed-rank test on the Effort Likert-type ratings. Calculate an exact p-value. What is the p-value from such a test?

```
install.packages("coin")
```

Ans: We can see the p value is 0.5016. Meaning there is no significant difference between the Likert-type ratings.

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```
library(coin)
```

```
## Loading required package: survival
```

```
wilcoxsign_test(Effort ~ Engine | Subject, data=websearch2, distribution="exact")
```

```
##
##  Exact Wilcoxon-Pratt Signed-Rank Test
##
## data:  y by x (pos, neg)
##   stratified by block
## Z = 0.68343, p-value = 0.5016
## alternative hypothesis: true mu is not equal to 0
```

6. This file describes a study just like the one from websearch2.csv, except that now three search engines were used instead of two. Once again, the number of searches required and a subjective effort rating for each search engine were recorded. How many subjects took part in this new experiment?

```
websearch3 = read.csv("websearch3.csv")
```

```
websearch3$Subject = factor(websearch3$Subject) # convert to a nominal factor
websearch3$Engine = factor(websearch3$Engine)
websearch3$Order = factor(websearch3$Order) # convert to a nominal factor
summary(websearch3)
```

Ans: 30 subjects took part in the new experiment

```
##     Subject       Engine   Order     Searches          Effort
##  1      : 3   Bing  :30   1:30   Min.   : 92.0   Min.   :1.000
##  2      : 3   Google:30   2:30   1st Qu.:139.0   1st Qu.:3.000
##  3      : 3   Yahoo :30   3:30   Median :161.0   Median :4.000
##  4      : 3                      Mean   :161.6   Mean   :4.256
##  5      : 3                      3rd Qu.:181.8   3rd Qu.:6.000
##  6      : 3                      Max.   :236.0   Max.   :7.000
##  (Other):72
```

7. What was the average number of searches required for the search that had the greatest overall?

```
library(plyr)
ddply(websearch3, ~ Engine, function(data) summary(data$Searches))
```

Ans: Yahoo has the greatest average overall with average search of 172.4

```
##   Engine Min. 1st Qu. Median     Mean 3rd Qu. Max.
## 1   Bing  104  136.75  161.5 159.8333  180.75  214
## 2 Google   93  137.50  152.0 152.6667  168.50  200
## 3  Yahoo   92  141.75  169.5 172.4000  190.75  236
```

8. Conduct a repeated ANOVA to determine if there was an order effect on Searches. First determine whether there is a violation of sphericity. What is the value of Mauchly's W criterion? Among other things, to test for order effects.

```
library(ez)
m = ezANOVA(dv=Searches, within=Order, wid=Subject, type=3, data=websearch3)
m$Mauchly # p<.05 indicates a violation
```

Ans:Mauchly's W = 0.9416

```
##   Effect         W         p p<.05
## 2  Order 0.9416469 0.4309561
```

9. Interpret the result of Mauchly's test of sphericity, and then interpret the appropriate repeated measures ANOVA result. What is the p-value from the appropriate F-test?

Ans: We have a p value of 0.4309 which is greater than 0.05, meaning we do not have a violation of sphericity.

10. Conduct a repeated measures ANOVA on Searches by Engine. First determine whether there is a violation of sphericity. What is the value of Mauchly's W criterion? Among other things, to test for a significant main effects.

```
install.packages("ez")
```

Ans: We have a Mauchly's W of 0.9420316

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```
library(ez)
```

```
m = ezANOVA(dv=Searches, within=Engine, wid=Subject, type=3, data=websearch3)
m$Mauchly # p<.05 indicates a violation
```

```
##   Effect         W         p p<.05
## 2 Engine 0.9420316 0.4334278
```

11. Interpret the result of Machly's test of sphericity, and then interpret the appropriate repeated measures ANOVA result.

Ans: We have a p value of 0.4334 which is greater than 0.05, meaning we do not have a violation of sphericity.

12. Conduct manual pairwise comparisons of Searches among levels of Engine using paired-samples t-tests and Holm's sequential Bonferroni procedure to correct for multiple comparisons. What is the smallest corrected p-value resulting from this set of tests?

```
library(reshape2)
websearch3.wide.engine = dcast(websearch3, Subject ~ Engine, value.var="Searches") # go wide
bi.ya = t.test(websearch3.wide.engine$Bing, websearch3.wide.engine$Yahoo, paired=TRUE)
go.bi = t.test(websearch3.wide.engine$Google, websearch3.wide.engine$Bing, paired=TRUE)
go.ya = t.test(websearch3.wide.engine$Google, websearch3.wide.engine$Yahoo, paired=TRUE)
p.adjust(c(bi.ya$p.value, go.bi$p.value, go.ya$p.value), method="holm")
```

Ans: The smallest corrected p-value resulting from this set of tests is 0.05066714

```
## [1] 0.37497103 0.37497103 0.05066714
```

13. Conduct a nonparametric Friedman test on the Effort Likert-type ratings. Calculate an asymtotic p-value, what is the Chi-square statistic from such a test?

```
library(coin)
friedman_test(Effort ~ Engine | Subject, data=websearch3, distribution="asymptotic")
```

Ans: chi-squared = 8.0182

```
##
##  Asymptotic Friedman Test
##
## data:  Effort by
##   Engine (Bing, Google, Yahoo)
##   stratified by Subject
## chi-squared = 8.0182, df = 2, p-value = 0.01815
```

14. Conduct manual pairwise comparisons of Effort among levels of Engine with Wilcoxon signed-rank tests, using Holm's sequential Bonferroni procedure to correct for multiple comparisons. What is the smallest corrected p-value resulting from this set of test?

```
library(plyr)
ddply(websearch3, ~ Engine, function(data) summary(data$Effort))
```

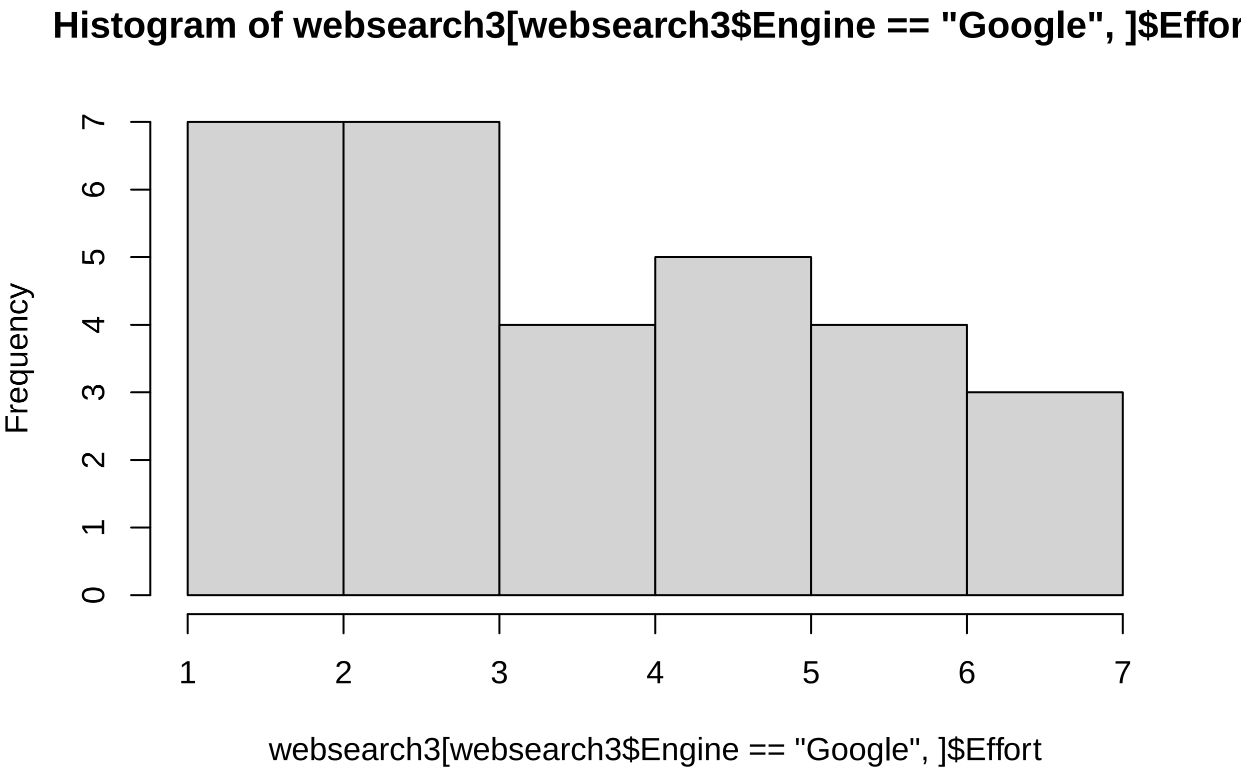
Ans: The smallest corrected p-value = 0.02332207

```
##   Engine Min. 1st Qu. Median     Mean 3rd Qu. Max.
## 1   Bing    1       3      4 3.900000       5    7
## 2 Google    1       3      4 3.866667       5    7
## 3  Yahoo    3       4      5 5.000000       6    7
```

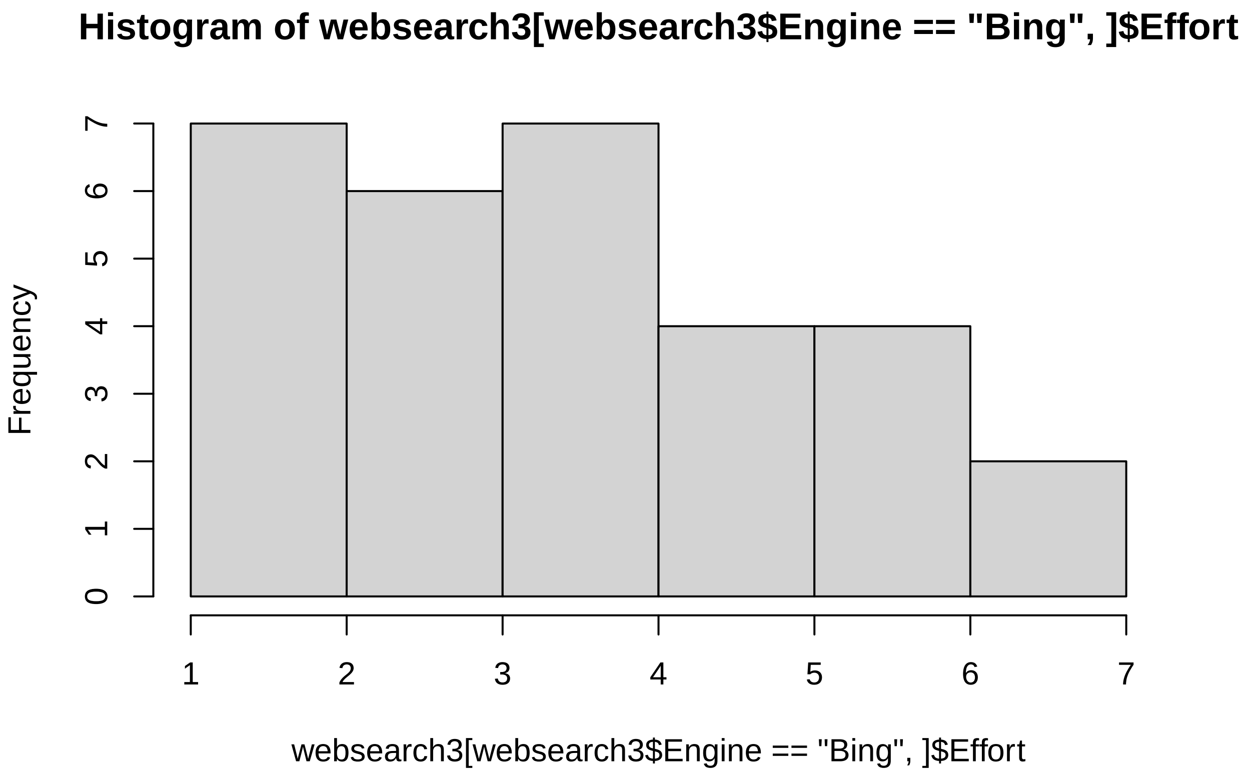
```
ddply(websearch3, ~ Engine, summarise, Effort.mean=mean(Effort), Effort.sd=sd(Effort))
```

```
##   Engine Effort.mean Effort.sd
## 1   Bing    3.900000  1.626293
## 2 Google    3.866667  1.925032
## 3  Yahoo    5.000000  1.389617
```

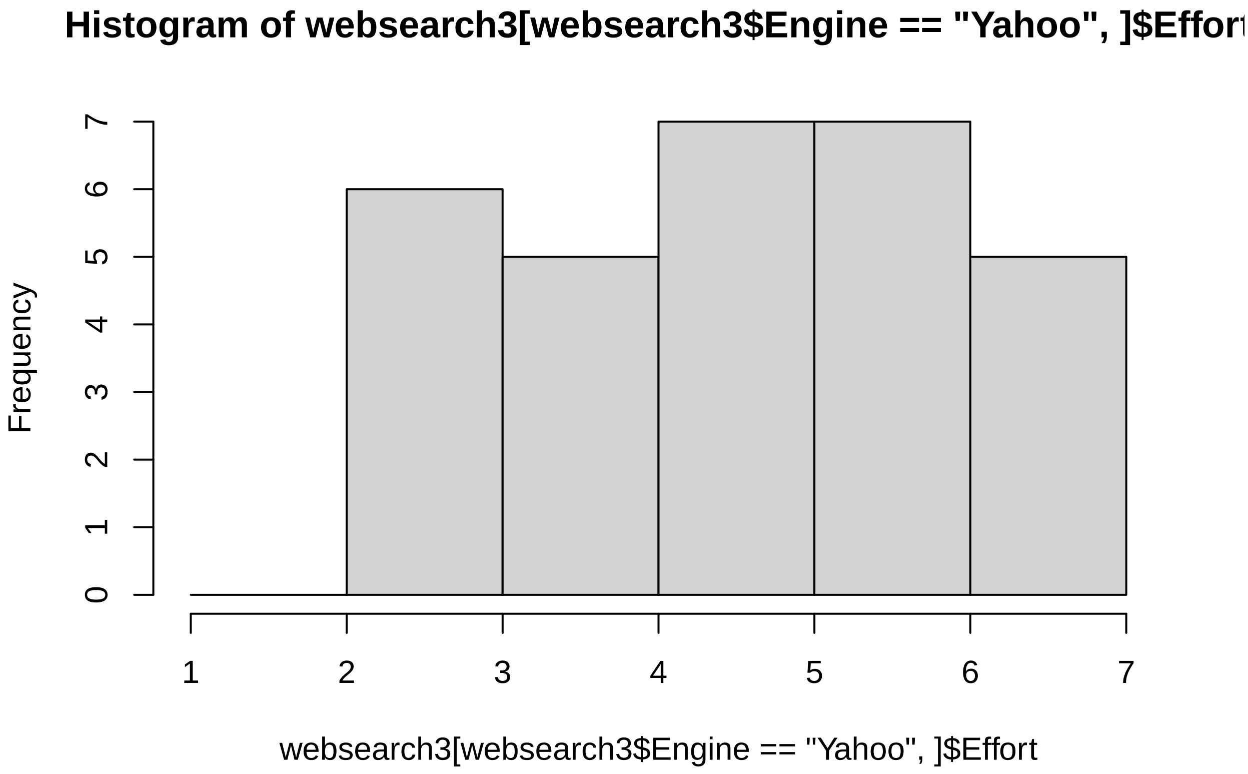
```
hist(websearch3[websearch3$Engine == "Google",]$Effort, breaks=c(1:7), xlim=c(1,7))
```



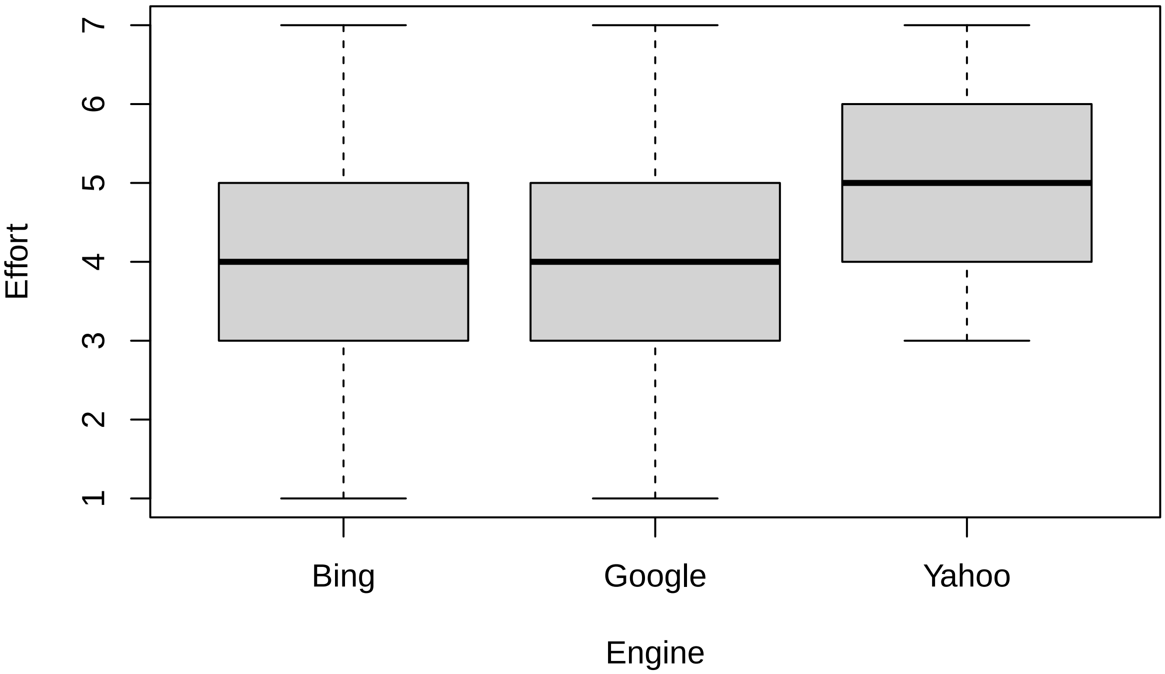
```
hist(websearch3[websearch3$Engine == "Bing",]$Effort, breaks=c(1:7), xlim=c(1,7))
```



```
hist(websearch3[websearch3$Engine == "Yahoo",]$Effort, breaks=c(1:7), xlim=c(1,7)) # new one
```



```
plot(Effort ~ Engine, data=websearch3) # boxplot
```



```
library(reshape2)
websearch3.wide.effort = dcast(websearch3, Subject ~ Engine, value.var="Effort") # go wide
bi.ya = t.test(websearch3.wide.effort$Bing, websearch3.wide.effort$Yahoo, paired=TRUE)
go.bi = t.test(websearch3.wide.effort$Google, websearch3.wide.effort$Bing, paired=TRUE)
go.ya = t.test(websearch3.wide.effort$Google, websearch3.wide.effort$Yahoo, paired=TRUE)
p.adjust(c(bi.ya$p.value, go.bi$p.value, go.ya$p.value), method="holm")
```

```
## [1] 0.02332207 0.94518326 0.03130003
```

Conclusions from the analysis

1. For the first study in which participants used Bing and Google, there was no order effect, meaning that the orders in which the search engines were used among the participants did not influence their answers.
2. We confirmed that Bing has more number of searches than Google, although there was no significant difference between the likert type ratings of the two search engines.
3. For the second study just like the first one except that three search engines were used instead of two, Yahoo has the greatest average overall.
4. We do not have a validation of sphericity on searches by the three search engines, meaning that the f-statistic calculated is valid and can be used to determine statistical significance.
5. For the comparisons of searches among levels of search engine, there was no significant difference.
6. For the comparisons of effort among levels of searches, with confidence, we could detect a significant difference between Bing and Yahoo, likewise Google and Yahoo. Although, there was no significance difference when we compared Google and Bing.
7. Lastly, Yahoo had the most number of searches and also required the least effort.