

HW_5_Grebeniuk

Grebeniuk_Alena

12/2/2023

R Markdown

Exercise 1 . Pre-process the data so that you have the following four columns:

subj compatible correct rt

```
getwd()

## [1] "/Users/alénagrebenuk/Desktop/stats1_hw"

setwd("/Users/alénagrebenuk/Desktop/stats1_hw")
dat<-read.table("stroopdata.txt",header=FALSE)
head(dat)

##           V1      V2      V3 V4 V5 V6 V7   V8
## 1 training green  green  1 10  2  1 1243
## 2 training blue   blue  1 15  3  1 1391
## 3 training yellow green  0  2  4  2  979
## 4 training yellow yellow 1  1  4  1  894
## 5 training yellow  red   0  4  1  1 1169
## 6 training  red   green  0  6  1  2 1225

dat<-dat[,c(3,4,7,8)]
head(dat)

##           V3 V4 V7   V8
## 1 green  1  1 1243
## 2 blue  1  1 1391
## 3 green  0  2  979
## 4 yellow 1  1  894
## 5 red    0  1 1169
## 6 green  0  2 1225

colnames(dat)<-c("subj","compatible","correct","rt")
head(dat)

##    subj compatible correct  rt
## 1 green           1       1 1243
## 2 blue           1       1 1391
## 3 green           0       2  979
## 4 yellow          1       1  894
## 5 red            0       1 1169
## 6 green           0       2 1225
```

subj refers to subject id (factor)

```
dat$subj<-as.factor(dat$subj)
```

compatible is coded as 1 (compatible) or 0 (incompatible), also a factor

```
dat$compatible<-as.factor(dat$compatible)
```

correct (whether the response is correct or incorrect), this is an integer (numeric)

```
dat$correct<-as.integer(dat$correct)
```

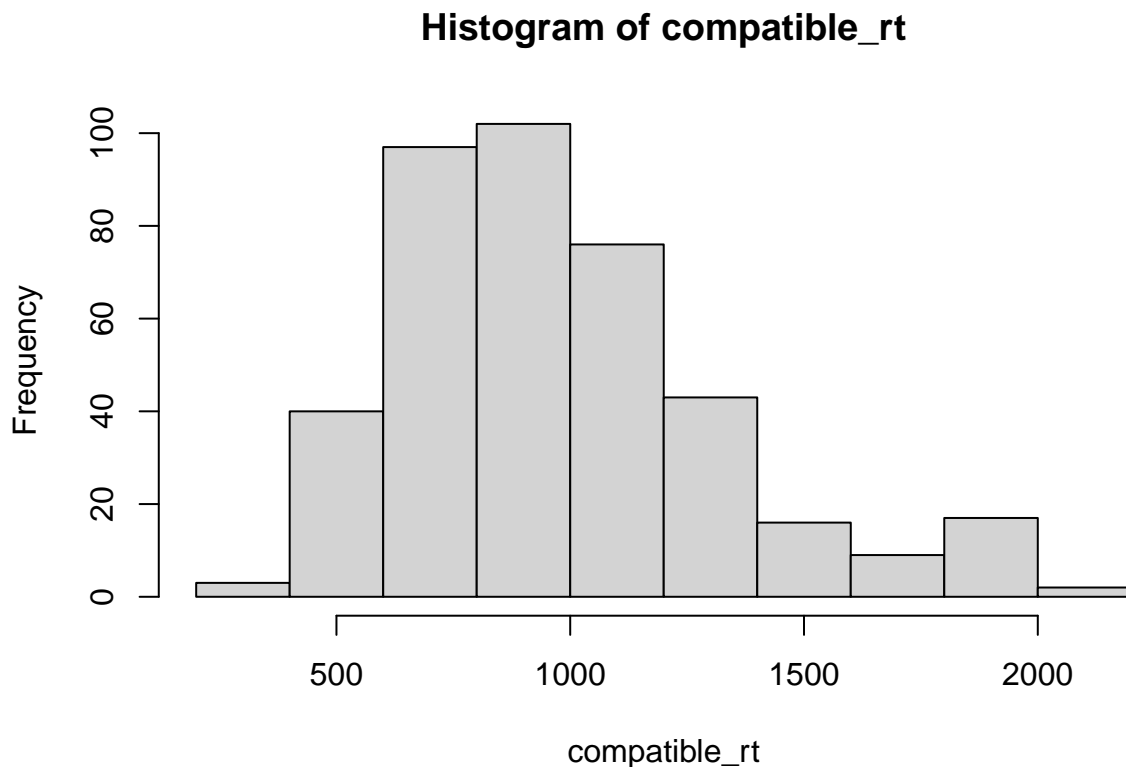
rt is the reaction time in milliseconds (also an integer, numeric)

```
dat$rt<-as.integer(dat$rt)
```

Exercise 2

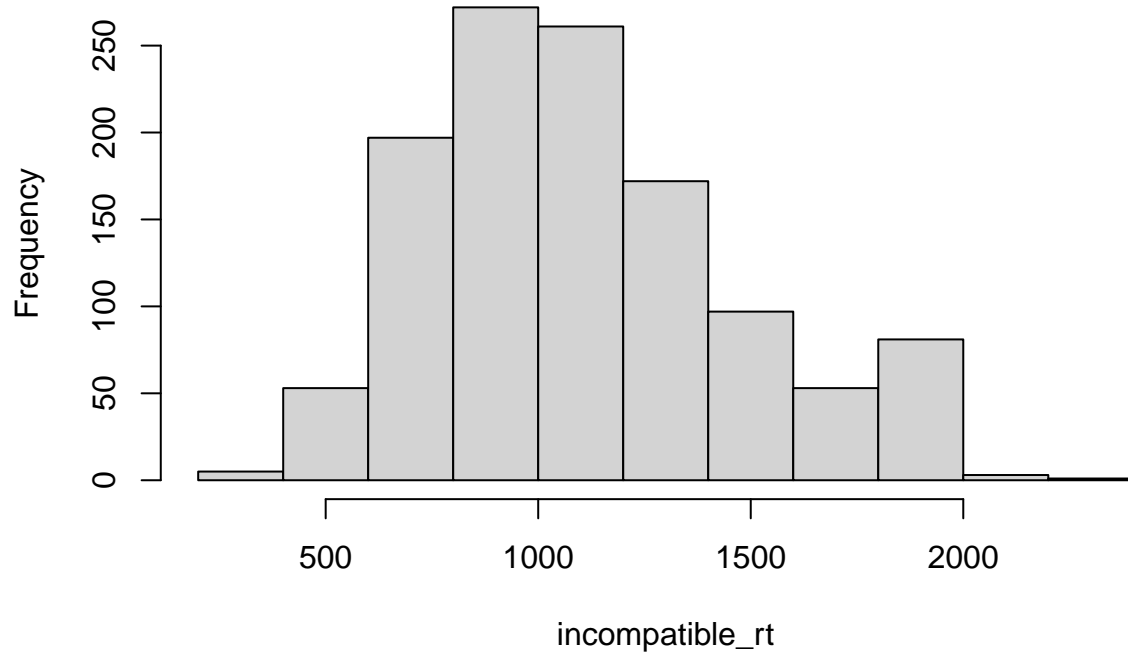
Plot the distribution of reaction times for the compatible and incompatible conditions using histograms and boxplots.

```
compatible_rt<-subset(dat,compatible==1)$rt  
incompatible_rt<-subset(dat, compatible==0)$rt  
hist(compatible_rt)
```

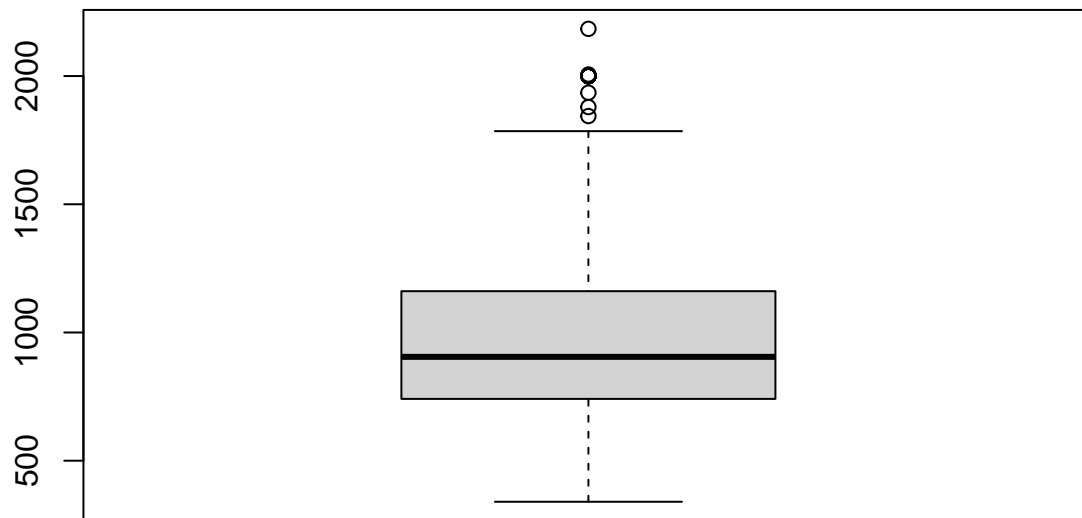


```
hist(incompatible_rt)
```

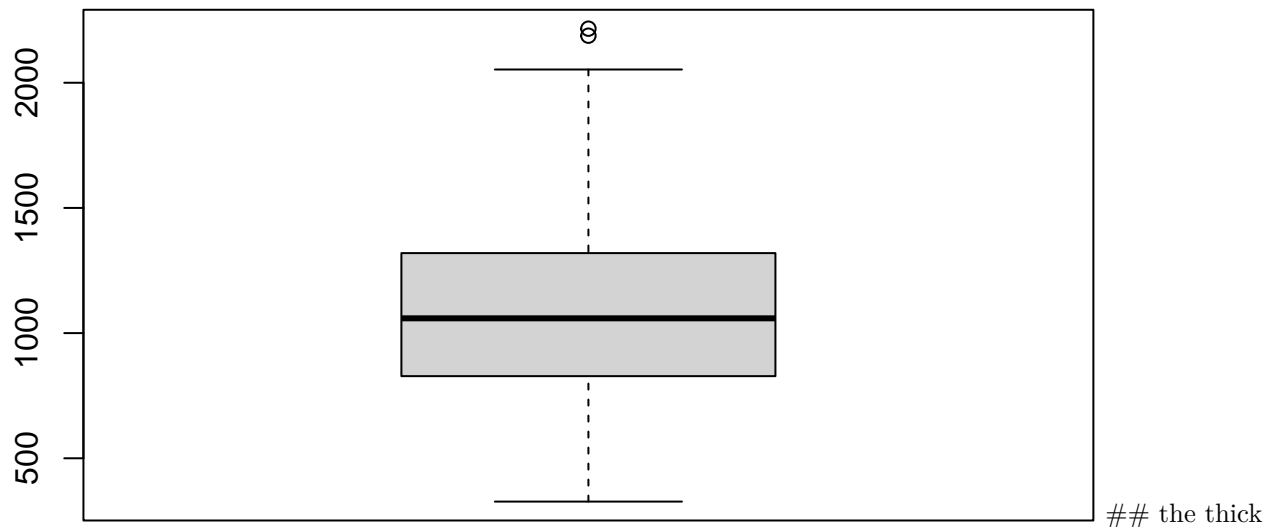
Histogram of incompatible_rt



```
boxplot(compatible_rt)
```



```
boxplot(incompatible_rt)
```



the thick horizontal line in the box is MEDIAN - (the middle value of the dataset). ## what are the ends of the whiskers? -THOSE ARE ENDS OF OBSERVED DATA POINTS - (Minimum and maximum values within a certain range). ## and what do the top and bottom edges of the box represent? - the Bottom is the First Quartile (Q1)(the median of the lower half of the data) ## the Top Edge is the Third Quartile (Q3)(the median of the upper half of the data.)

Exercise 3

What is the mean difference between the incongruent and congruent conditions (across all subjects)?

```
##r
mean(subset(dat,compatible==1)$rt)
## [1] 982.6296
mean(subset(dat,compatible==0)$rt)
## [1] 1108.183
```

Exercise 4

For each subject, what are the means differences between the incongruent and congruent conditions?

```
with(dat,tapply(rt, IND=list(subj,compatible),mean))

##           0           1
## blue   1101.026  937.5000
## green   1130.951 1034.3654
## red     1087.438  966.7800
## yellow 1111.794  989.7327
```

Exercise 5

What do you conclude from the differences in means? Can we conclude that incongruent conditions have longer reaction times than congruent conditions?

we can conclude that incongruent conditions have longer reaction times than congruent conditions. Since the difference in mean showed that $1108.1833 > 982.6296$.