**Question 1.**
**R code**

```
# Make a new wide format dataframe `staiData` containing only "participant", "group", and three sets of stai_scores at different timepoint
staiData<-labdata[c("participant","group","timepoint","stai_score")]
staiData <- reshape(staiData, direction="wide", idvar="participant",timevar="timepoint")
staiData <- subset(staiData, select=-c(`group.week 4`,`group.week 8`))
names(staiData) <- c("participant","group","stai_baseline","stai_week4","stai_week8")

# Decide whether their anxiety level improved from baseline to after the intervention for every participant
change <- logical()
for (i in 1:(nrow(staiData))){
  if (staiData$group[i]=="A"){          # if the participant is in group A
    change[[i]] <- (staiData$stai_baseline[i]-staiData$stai_week4[i])>=3      # compare data obtained from baseline and week 4
  }
  if (staiData$group[i]=="B"){          # if the participant is in group B
    change[[i]] <- (staiData$stai_baseline[i]-staiData$stai_week8[i])>=3      # compare data obtained from baseline and week 8
  }
}

# Make a table of change
table(change)
```

**Table**

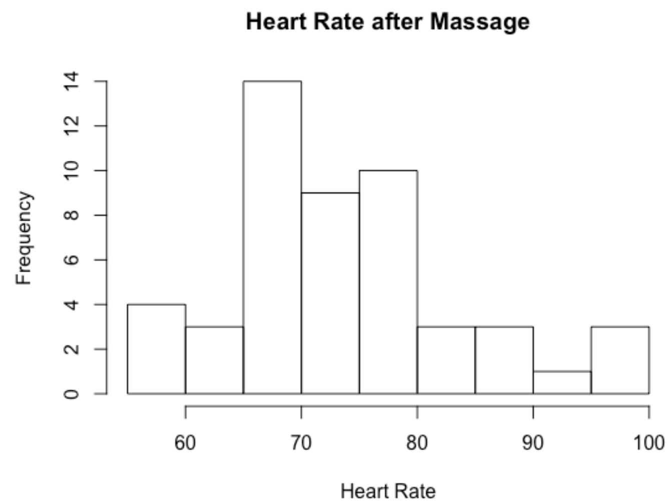
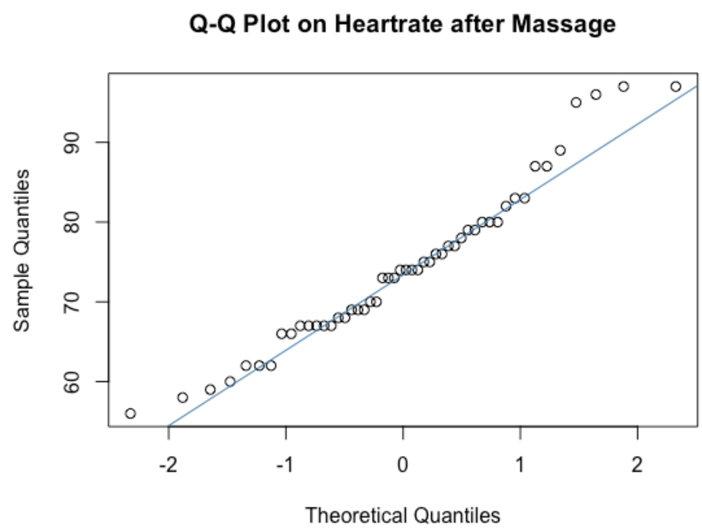
```
FALSE  TRUE
   22    29
```

**Question 2.**
**a) inference procedure**
t confidence interval for the population mean

b)
Note: simply copy and paste (or recreate) the template table given below as many times as needed, i.e. once for each condition of the model.

| | |
|---|---|
| **Condition** | The data is a SRS from a large population |
| **Valid?** | Yes |
| **Justification** | The data is obtained from a repeated measures design with randomization on the order of participation in two treatments which means it is a SRS. The desired population is all Japanese adults which is large enough. |
| **R code and associated output (if applicable)** | N/A |

| | |
|---|---|
| **Condition** | Observations from the population has a normal distribution with mean μ and standard deviation σ |
| **Valid?** | Yes |
| **Justification** | A histogram and a Q-Q plot are drawn based on the heart rate data obtained from the adults who have been using aromatherapy foot massage for an extended period of time and a straight Q-Q line is added. The deviations from the line are minimal which indicates normal distribution. Besides, the sample is large enough (>50), so although the histogram is right skewed, the population is robust to deviations from the Q-Q line. |
| **R code and associated output (if applicable)** | labdata$heartrate <- as.integer(labdata$heartrate) # convert the heartrate to integer from factors<br>heartrateAfter = labdata$heartrate[(labdata$group=="A"&labdata$timepoint=="week 4")\|(labdata$group=="B"&labdata$timepoint=="week 8")]<br>heartrateAfter <- heartrateAfter[!is.na(heartrateAfter)] # make a subset containing the heartrate data obtained after inventation for both group A and B<br>hist(heartrateAfter, main="Heart Rate after Massage", xlab="Heart Rate", ylab="Frequency") # make a histogram with the heartrate data<br>qqnorm(heartrateAfter, main="Q-Q Plot on Heartrate after Massage") # make a Q-Q plot with the heartrate data<br>qqline(heartrateAfter, col="steelblue") # add a Q-Q line to the Q-Q plot |

| | |
|---|---|
| | Heart Rate after Massage
Frequency
0
2
4
6
8
10
12
14
60
70
80
90
100
Heart RateQ-Q Plot on Heartrate after Massage
Sample Quantiles
60
70
80
90
-2
-1
0
1
2
Theoretical Quantiles |

**Question 3.**

a) **inference procedure (copy & paste from Question 2a):**
t confidence interval for the population mean

b) **R code and associated Output**
R Code:
labdata$heartrate <- as.integer(labdata$heartrate)     # convert the heartrate to integer from factors
heartrateAfter = labdata$heartrate[(labdata$group=="A"&labdata$timepoint=="week 4")|(labdata$group=="B"&labdata$timepoint=="week 8")]  # make a subset containing the heartrate data obtained after invention for both group A and B
heartrateAfter <- heartrateAfter[!is.na(heartrateAfter)]        # Remove the NA in the list
mean = mean(heartrateAfter)             # Calculate the mean
sd <- sqrt(var(heartrateAfter))           # Calculate the standard deviation
mean   # print the mean
sd        # print the standard deviation

Output :
[1] 74.24
[1] 10.06015

c) **conclusion**
The mean heartrate for adults who have been using aromatherapy foot massage for an extended period of time is 74.24 ± 10.06 BPM.

## Question 4.

**Description:**

As the research question indicates, the independent variable is whether one is engaged in aromatherapy foot massage, and the response variable is the blood pressure. The null hypothesis would be engaging in aromatherapy foot massage does not have change the blood pressure of healthy adults. The alternative hypothesis would be engaging in aromatherapy foot massage could make the blood pressure of healthy adults closer to normal range.

To determine the definition of “a positive therapeutic impact”, we will find the normal range (as the “healthy blood pressure”) of the systolic and diastolic blood pressure of Japanese adults. After calculating the average systolic and diastolic blood pressure (=(sysBP1+sysBP2)/2 and =(diaBP1+diaBP2)/2, respectively) for each participant at each timepoint, we will calculate the difference between the average blood pressure and the healthy blood pressure range. The mean difference calculated at the baseline and after the intervention (timepoint==”week 4” for group A, timepoint==”week 8” for group B) will be compared using two-sample t procedure.