Assignment-8

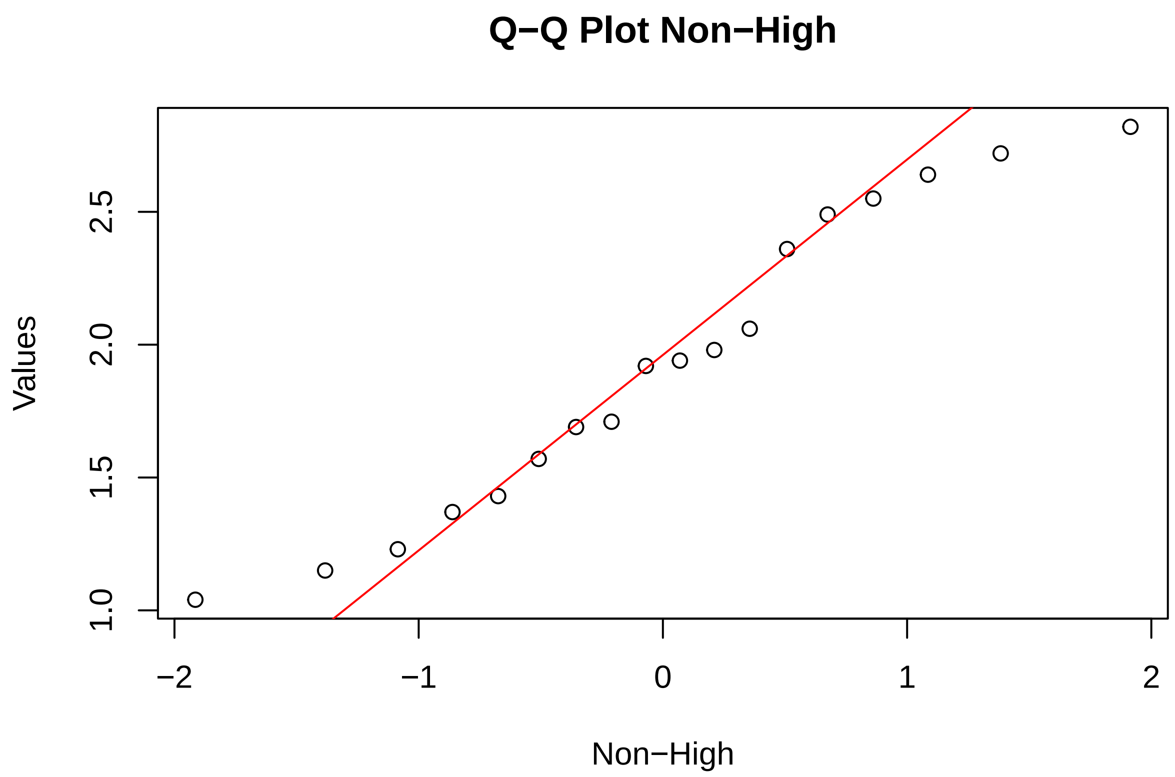
Ashish Verma

2024-03-27

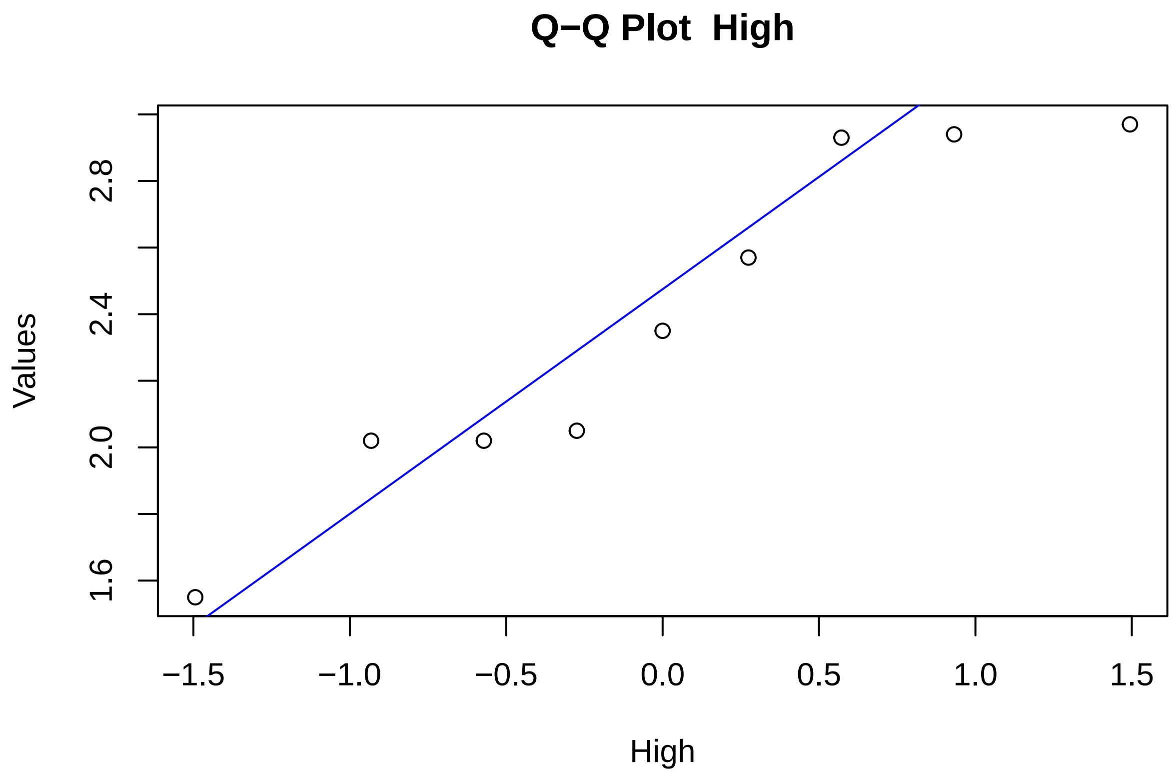
Section 9.2 Problem #22

Solution(a)

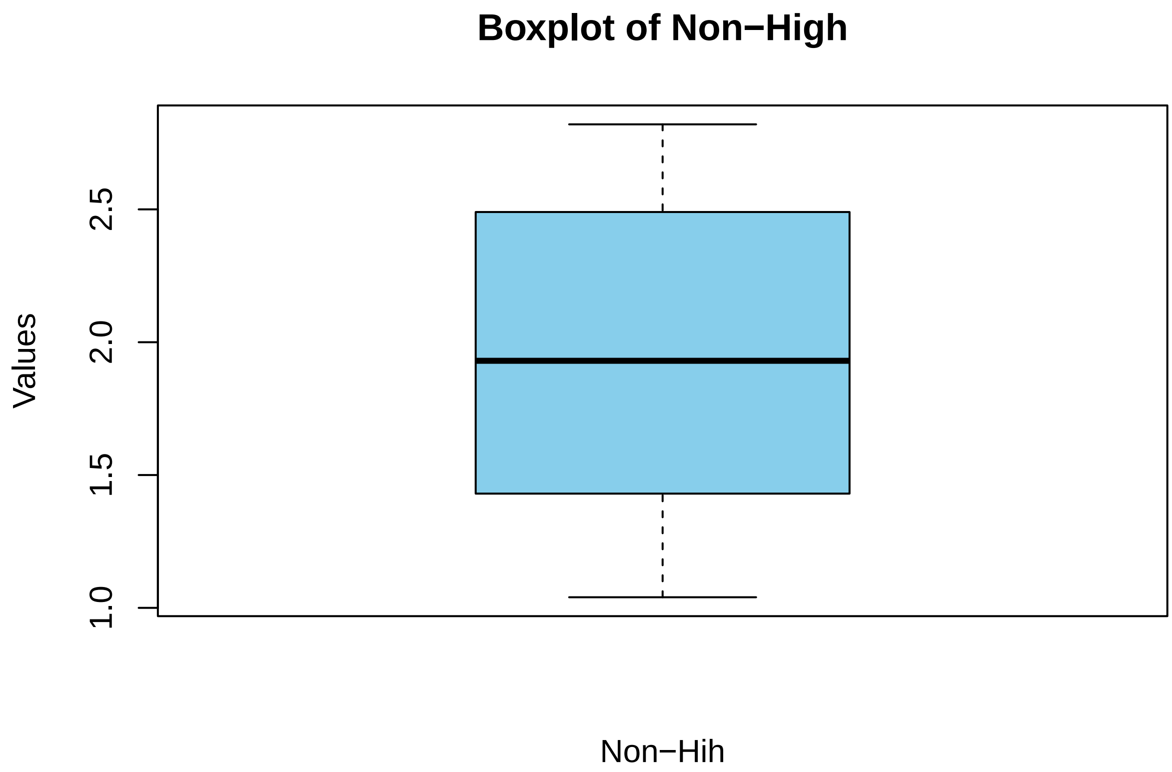
```
x<-c(1.04,1.15,1.23,1.69,1.92,1.98,2.36,2.49,2.72,1.37,1.43,1.57,1.71,1.94,2.06,2.55,2.64,2.82) ##Non-h
y<-c(1.55,2.02,2.02,2.05,2.35,2.57,2.93,2.94,2.97) ##High
qqnorm(x,main = "Q-Q Plot Non-High", xlab = "Non-High", ylab = "Values")
qqline(x,col = "red")
```



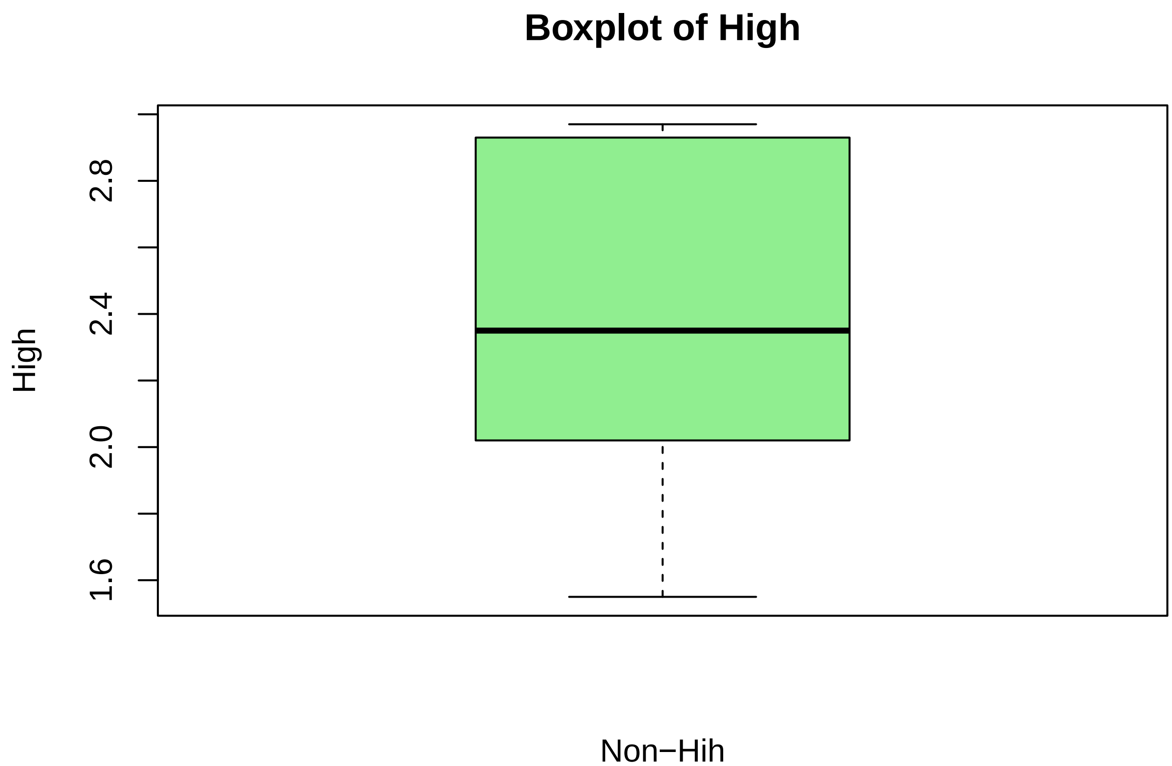

```
qqnorm(y,main = "Q-Q Plot  High", xlab = "High", ylab = "Values")
qqline(y,col = "blue")
```




```r
boxplot(x,
        main = "Boxplot of Non-High",  # Title of the plot
        xlab = "Non-Hih",             # Label for x-axis
        ylab = "Values",             # Label for y-axis
        col = "skyblue",  # Colors for the boxes
        outline = TRUE)              # Show outliers
```




```
boxplot(y,
        main = "Boxplot of High",  # Title of the plot
        xlab = "Non-Hih",              # Label for x-axis
        ylab = "High",              # Label for y-axis
        col = "lightgreen",  # Colors for the boxes
        outline = TRUE)              # Show outliers
```



1.There are no outliers in the data.

2.The data is almost normal distributed

3.The median value of High is greater than Non-high.

Soltuion(b)

Yes its reasonable to use T test for the given data set,sample size of M<40 and sample size of N<40 and population variances of both Non-high and High are unknown.

```
m<-length(x)
n<-length(y)
cat("The length of M",m)
```

```
## The length of M 18
```

```
cat("The length of N",n)
```

```
## The length of N 9
```

Solution(c)

```r
xbar<-mean(x)
ybar<-mean(y)
samplex<-sd(x)
sampley<-sd(y)
m<-length(x)
n<-length(y)
alpha<-0.01
sample_diff<-sqrt((samplex^2)/m+(sampley^2)/n)
ttesting<-(xbar - ybar)/sample_diff
cat("T-test results",ttesting)
```

```
## T-test results -2.092321
```

```r
numerator<-(samplex^2/m+sampley^2/n)^2
denominator1<-(samplex^2/m)^2/(m -1)
denominator2<-(sampley^2/n)^2/(n -1)
degree_of_freedom<-ceiling(numerator/(denominator1+denominator2))
p_value<-pt(ttesting,df=degree_of_freedom)
cat("P-value results",p_value)
```

```
## P-value results 0.02542456
```

```r
if (p_value < alpha) {
cat("Reject the null hypothesis.\n")
} else {
 cat("Fail to reject the null hypothesis.\n")
}
```

```
## Fail to reject the null hypothesis.
```

Section 9.2 Problem #28

Solution:

```r
x_yf<-c(29, 34, 33, 27, 28, 32, 31, 34, 32, 27)
y_of<-c(18, 15, 23, 13, 12)
mu_diff<-10
m<-length(x_yf)
n<-length(y_of)
xbar<-mean(x_yf)
ybar<-mean(y_of)
samplex<-sd(x_yf)
sampley<-sd(y_of)
alpha<-0.1

sample_diff<-sqrt((samplex^2)/m+(sampley^2)/n)
ttesting<-((xbar - ybar)-mu_diff)/sample_diff
cat("T-test result",ttesting)
```

```
## T-test result 2.076432
```

```r
numerator<-(samplex^2/m+sampley^2/n)^2
denominator1<-(samplex^2/m)^2/(m -1)
denominator2<-(sampley^2/n)^2/(n -1)
degree_of_freedom<-ceiling(numerator/(denominator1+denominator2))
#qtesting<-qt(1 -alpha,df=degree_of_freedom)
p_value<-pt(ttesting,df=degree_of_freedom,lower.tail = FALSE)
cat("P-value result",p_value)
```

```
## P-value result 0.04157212
```

```r
if (p_value <alpha) {
cat("Reject the null hypothesis.There is evidence that suggest that true average maximum lean angle for older females is more than 10 degree smaller\n")
} else {
 cat("Fail to reject the null hypothesis.There is no evidence  suggest that true average maximum lean angle for older females is more than 10 degree smaller\n")
}
```

```
## Reject the null hypothesis.There is evidence that suggest that true average maximum lean
## angle for older females is more than 10 degree smaller
```

Section 9.2 Problem #32

Solution(a)

```r
m<-28
n<-16
xbar<-801
ybar<-780
xsd<-117
ysd<-72
alpha<-1 - 0.99
sample_diff<-sqrt((xsd^2)/m+(ysd^2)/n)

numerator<-(xsd^2/m+ysd^2/n)^2
denominator1<-(xsd^2/m)^2/(m -1)
denominator2<-(ysd^2/n)^2/(n -1)
degree_of_freedom<-ceiling(numerator/(denominator1+denominator2))
ttesting_critical<-qt(1 - (alpha/2),df=degree_of_freedom)
lower_interval<-(xbar -ybar)-ttesting_critical*sample_diff
upper_interval<-(xbar -ybar)+ttesting_critical*sample_diff
cat("The 99% confidence interval is given by",lower_interval,upper_interval)
```

```
## The 99% confidence interval is given by -55.92531 97.92531
```

Solution(b)

```r
alpha_new<-0.05
m<-28
n<-16
xbar<-801
ybar<-780
xsd<-117
ysd<-72
sample_diff<-sqrt((xsd^2)/m+(ysd^2)/n)
ttesting<-((xbar - ybar))/sample_diff
cat("T-test result",ttesting)
```

```
## T-test result 0.7365507
```

```r
numerator<-(xsd^2/m+ysd^2/n)^2
denominator1<-(xsd^2/m)^2/(m -1)
denominator2<-(ysd^2/n)^2/(n -1)
degree_of_freedom<-ceiling(numerator/(denominator1+denominator2))
p_value<-pt(ttesting,df=degree_of_freedom,lower.tail = FALSE)
cat("P-value",p_value)
```

```
## P-value 0.232745
```

```r
if (p_value < alpha_new) {
cat("Reject the null hypothesis.\n")
} else {
 cat("Fail to reject the null hypothesis.No sufficient evidence to support the claim.\n")
}
```

```
## Fail to reject the null hypothesis.No sufficient evidence to support the claim.
```

Part2

Solution

mu1= the population mean satisfaction level for graduates from State School S1. mu2= the population mean satisfaction level for graduates from State School S2. The null hypothesis (H0) and the alternative hypothesis (H1) can be stated as follows: H0: mu1=mu2 (There is no significant difference in the mean satisfaction levels between the two schools.) H1:mu1!=mu2 (There is a significant difference in the mean satisfaction levels between the two schools.)

```r
# Sample data
school_s1 <- c(69, 75, 79, 80, 81, 82, 86, 89, 91, 94, 97)
school_s2 <- c(59, 62, 66, 70, 70, 75, 76, 77, 78, 79, 81, 84, 85, 86, 94)
mean_s1 <- mean(school_s1)
mean_s2 <- mean(school_s2)
sd_s1 <- sd(school_s1)
sd_s2 <- sd(school_s2)
n1 <- length(school_s1)
n2 <- length(school_s2)
```

```r
# Calculate pooled standard deviation
sp <- sqrt(((n1 - 1) * sd_s1^2 + (n2 - 1) * sd_s2^2) / (n1 + n2 - 2))

# Calculate t-test statistic
t_stat <- (mean_s1 - mean_s2) / (sp * sqrt(1/n1 + 1/n2))

# Degrees of freedom
df <- n1 + n2 - 2

# Calculate p-value
p_value <- 2 * pt(-abs(t_stat), df)

# Alpha level
alpha <- 0.01

# Print results
cat("t-test statistic:", t_stat, "\n")
```

```
## t-test statistic: 2.160625
```

```r
cat("Degrees of freedom:", df, "\n")
```

```
## Degrees of freedom: 24
```

```r
cat("p-value:", p_value, "\n")
```

```
## p-value: 0.04092527
```

```r
# Check if the difference is significant at alpha = 1%
if (p_value < alpha) {
  cat("The difference between the means is significant at the 1% level.\n")
} else {
  cat("The difference between the means is not significant at the 1% level.\n")
}
```

```
## The difference between the means is not significant at the 1% level.
```