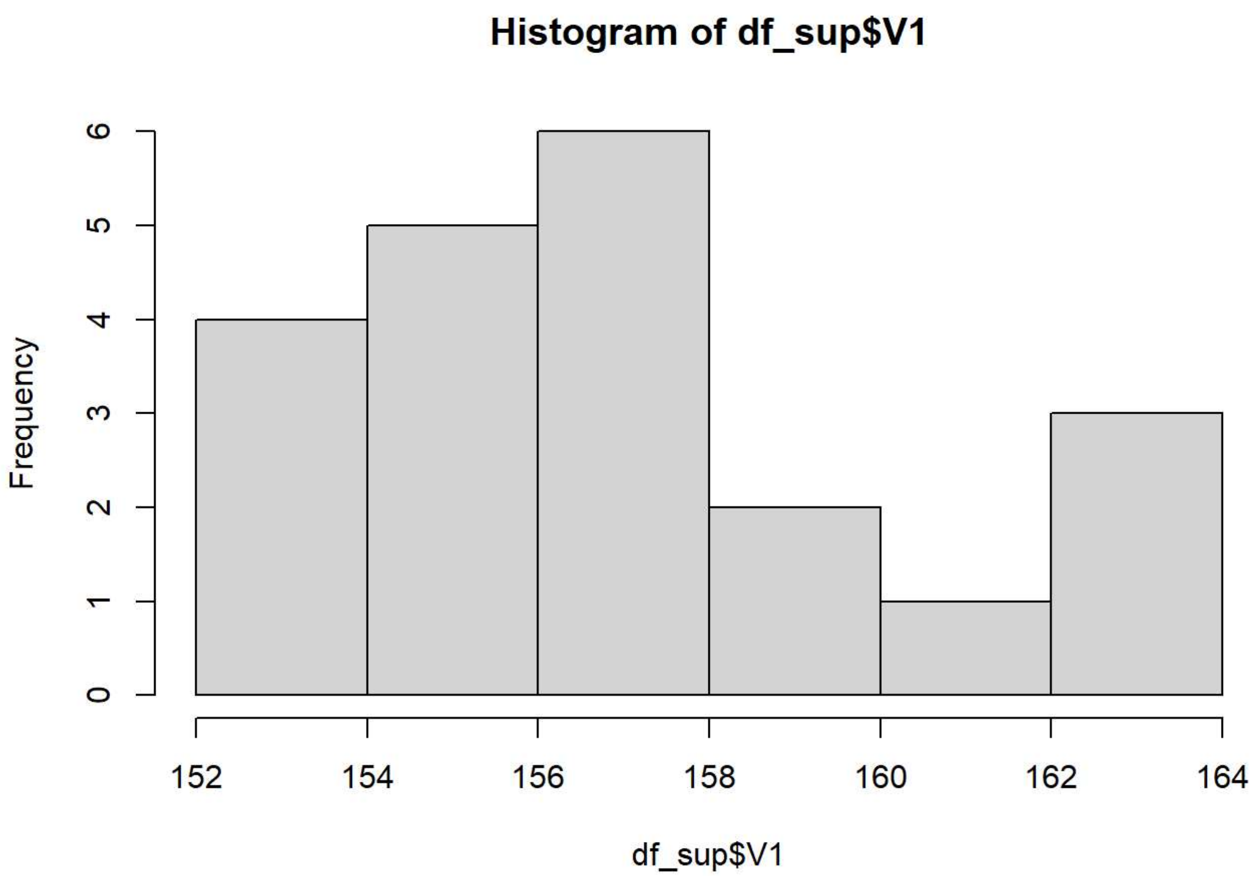
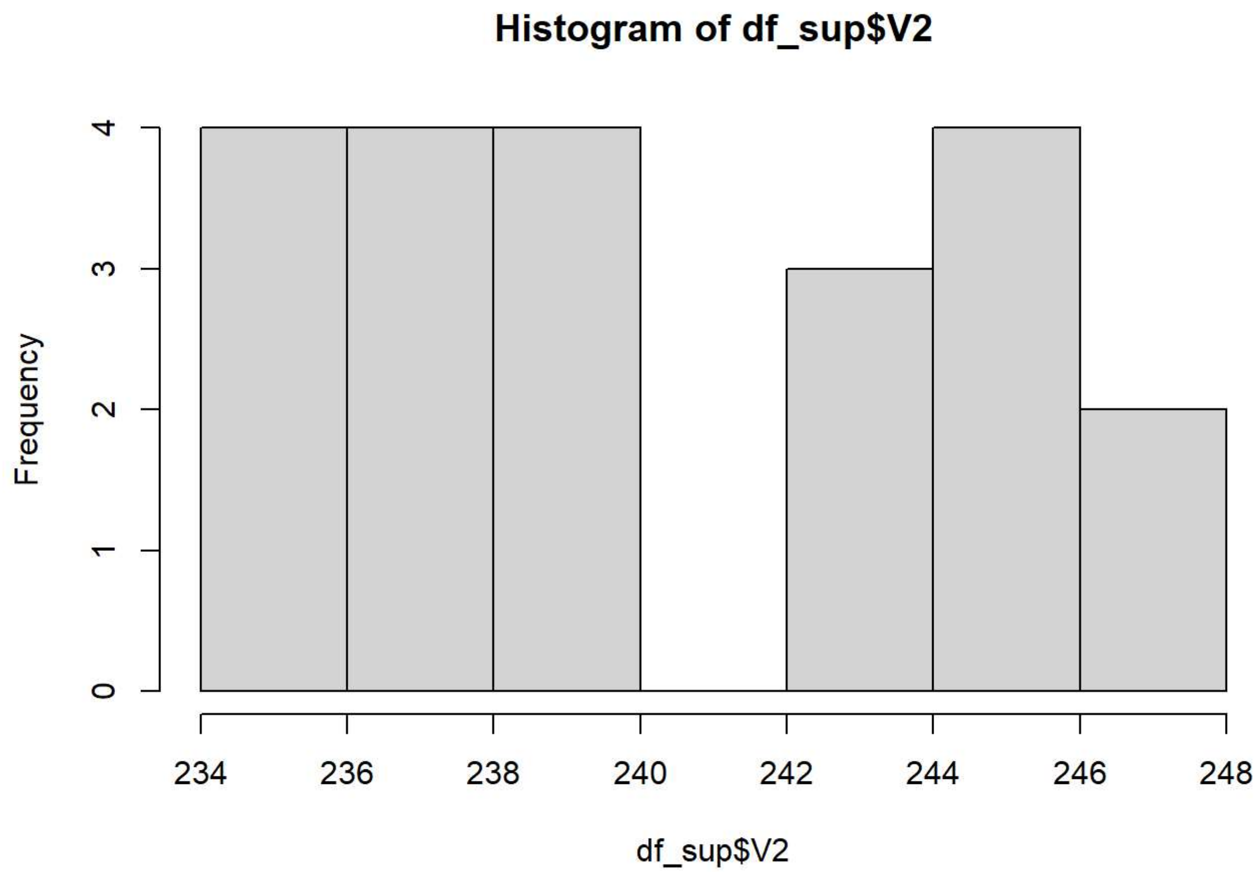
Guia 3.1 ejercicio 4

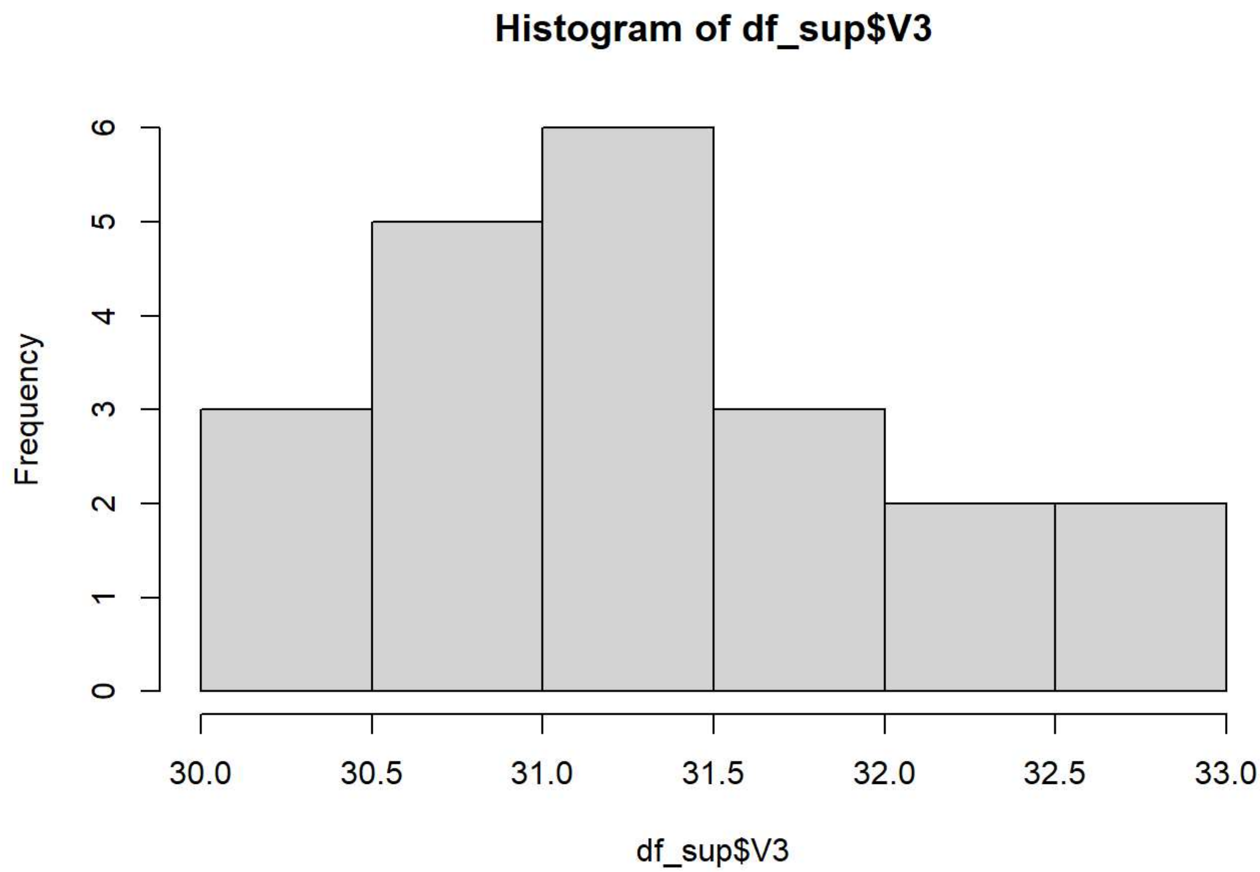
```
df_sup = read.table("C:\\Users\\Dell7400\\Documents\\Ale\\Facu\\Multivariado\\datos\\supervivien
tes.txt")
df_no_sup = read.table("C:\\Users\\Dell7400\\Documents\\Ale\\Facu\\Multivariado\\datos\\no_super
vivientes.txt")


# veamos primero la pinta de los datos

hist(df_sup$V1)
```



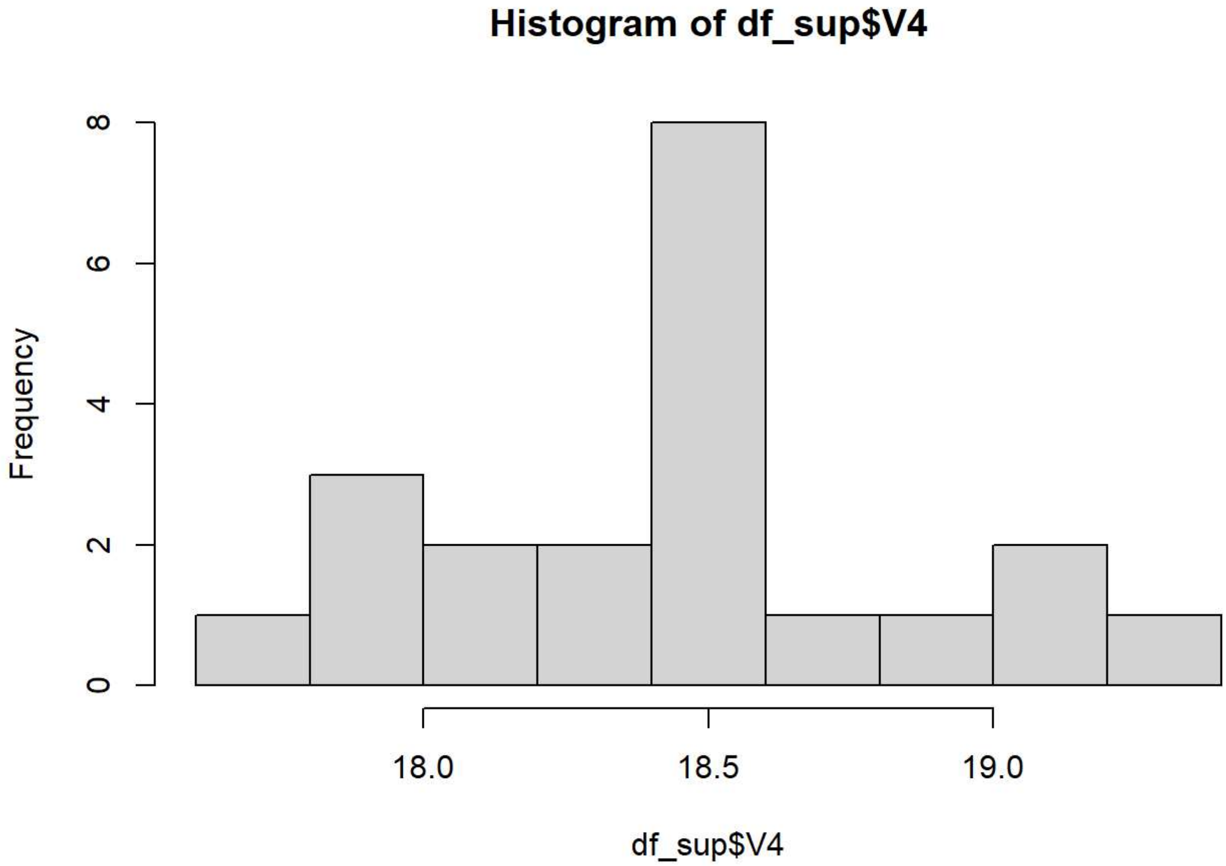

```
hist(df_sup$V2)
```



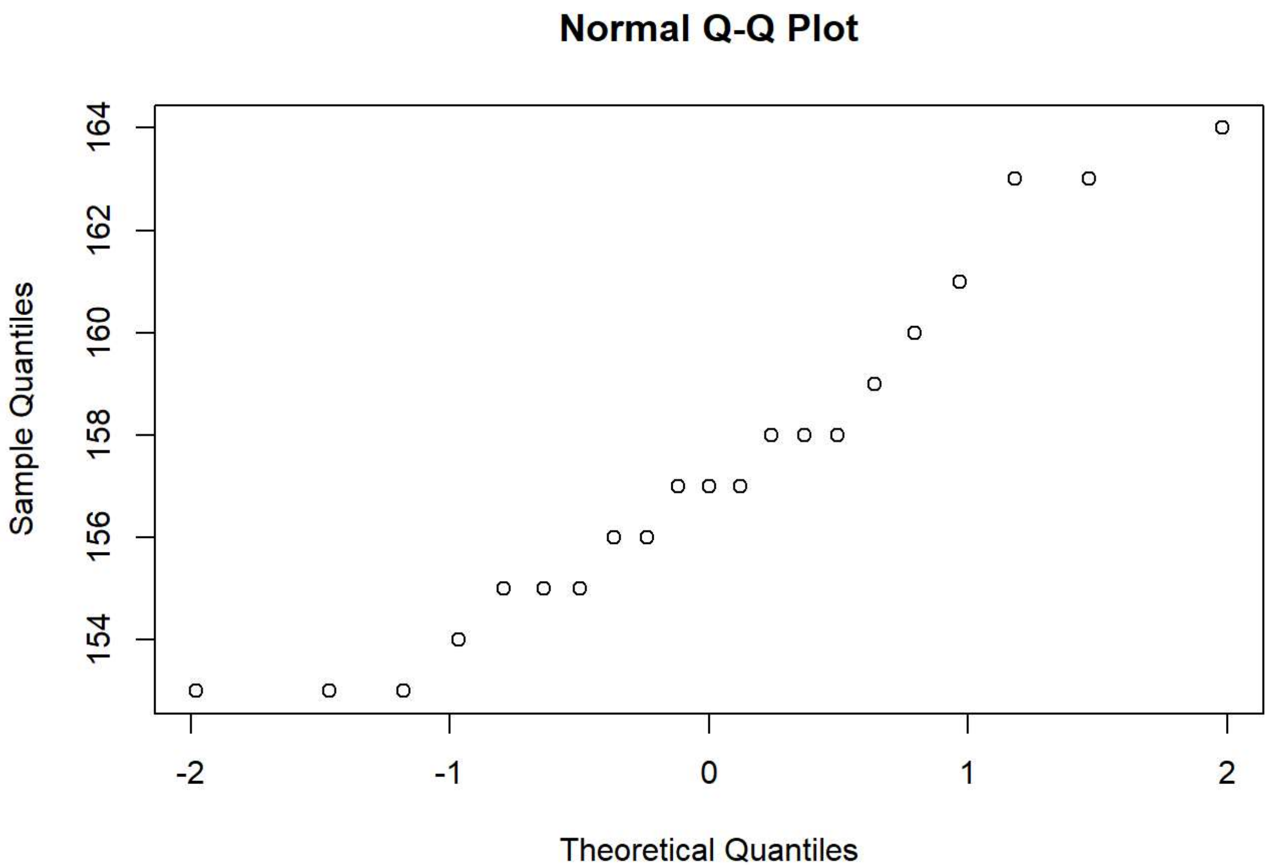

```
hist(df_sup$V3)
```



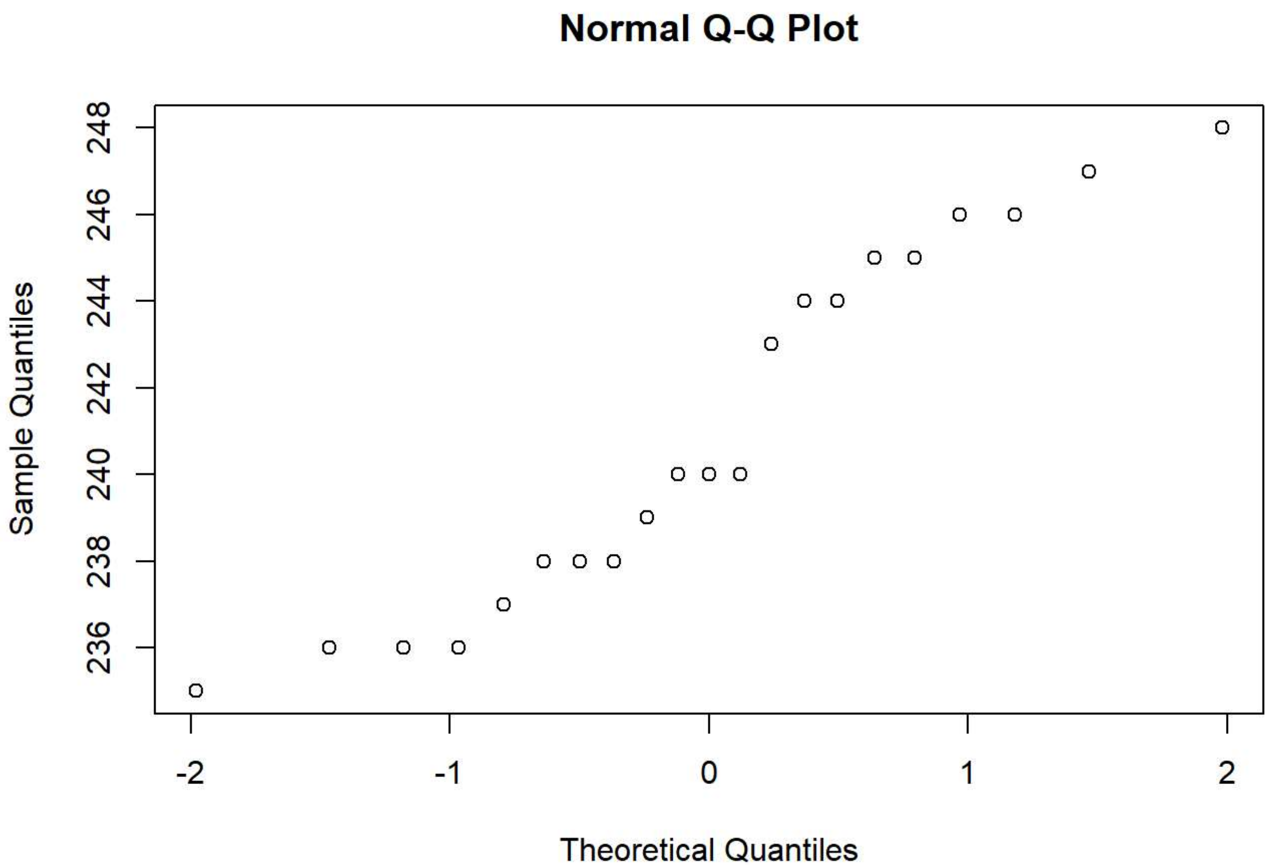

```
hist(df_sup$V4)
```



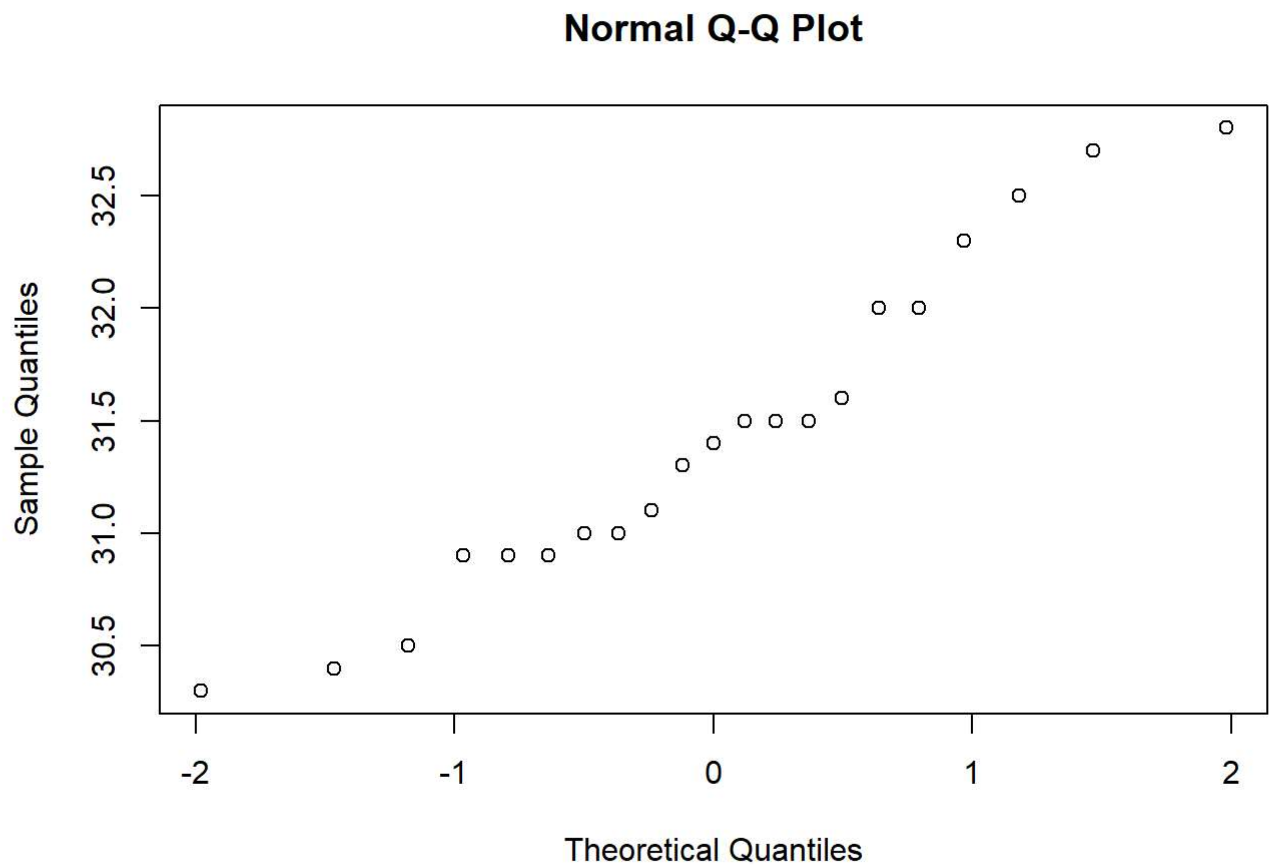

```
qqnorm(df_sup$V1)
```



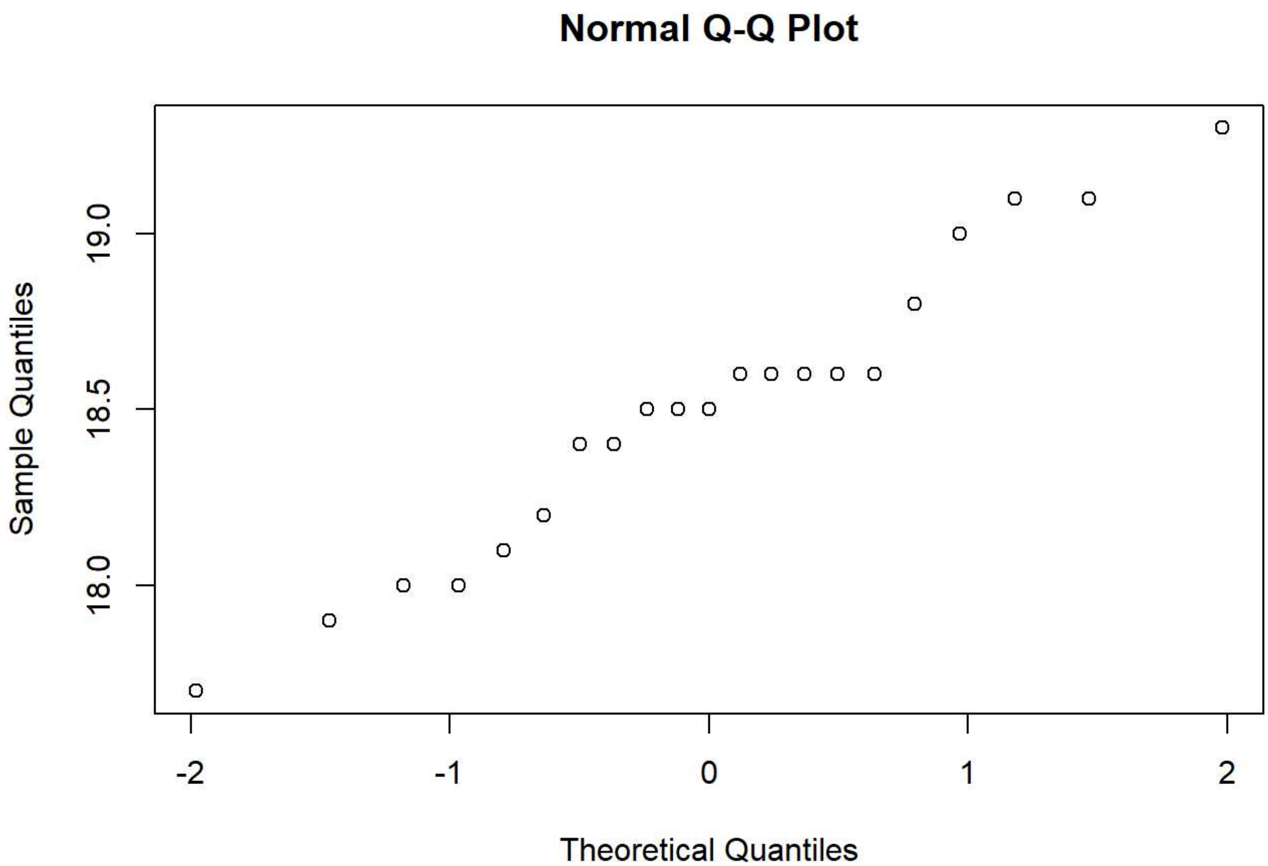

```
qqnorm(df_sup$V2)
```



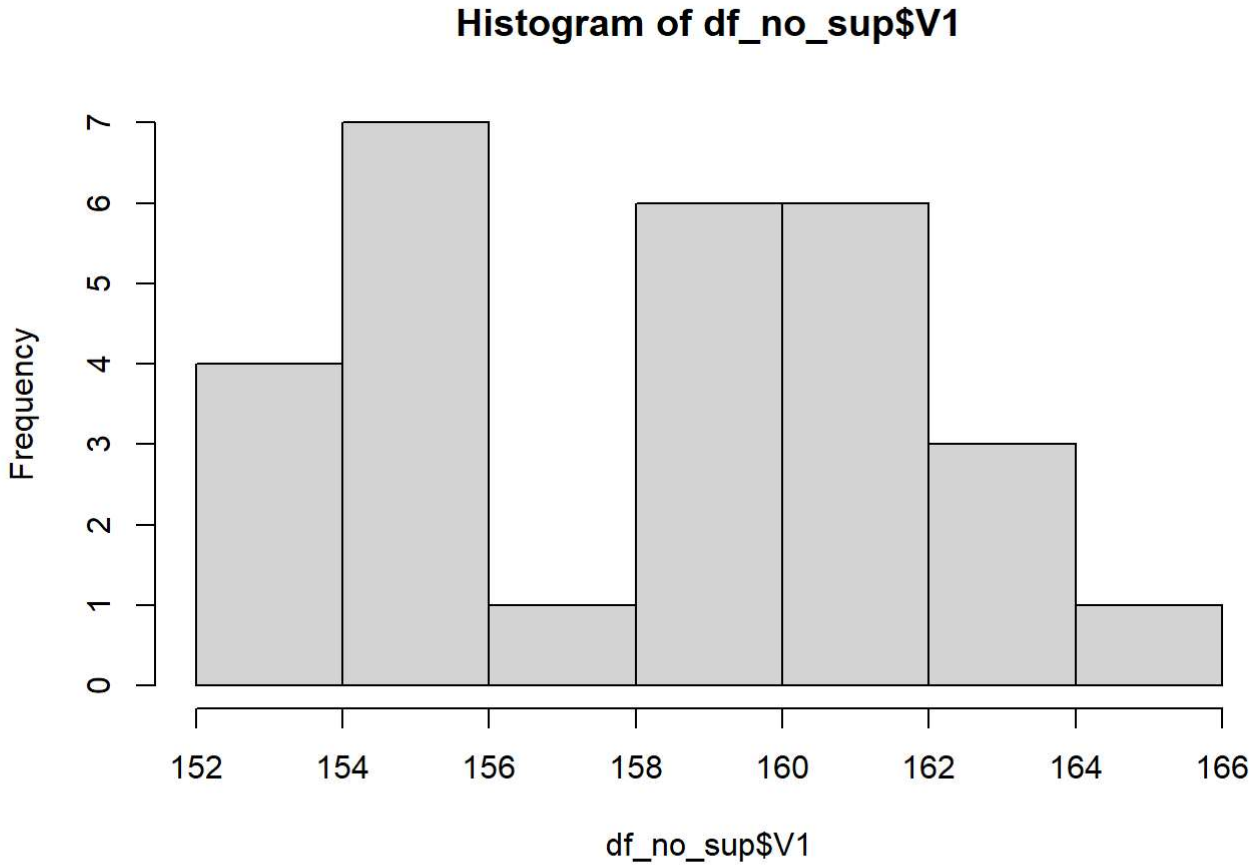

```
qqnorm(df_sup$V3)
```



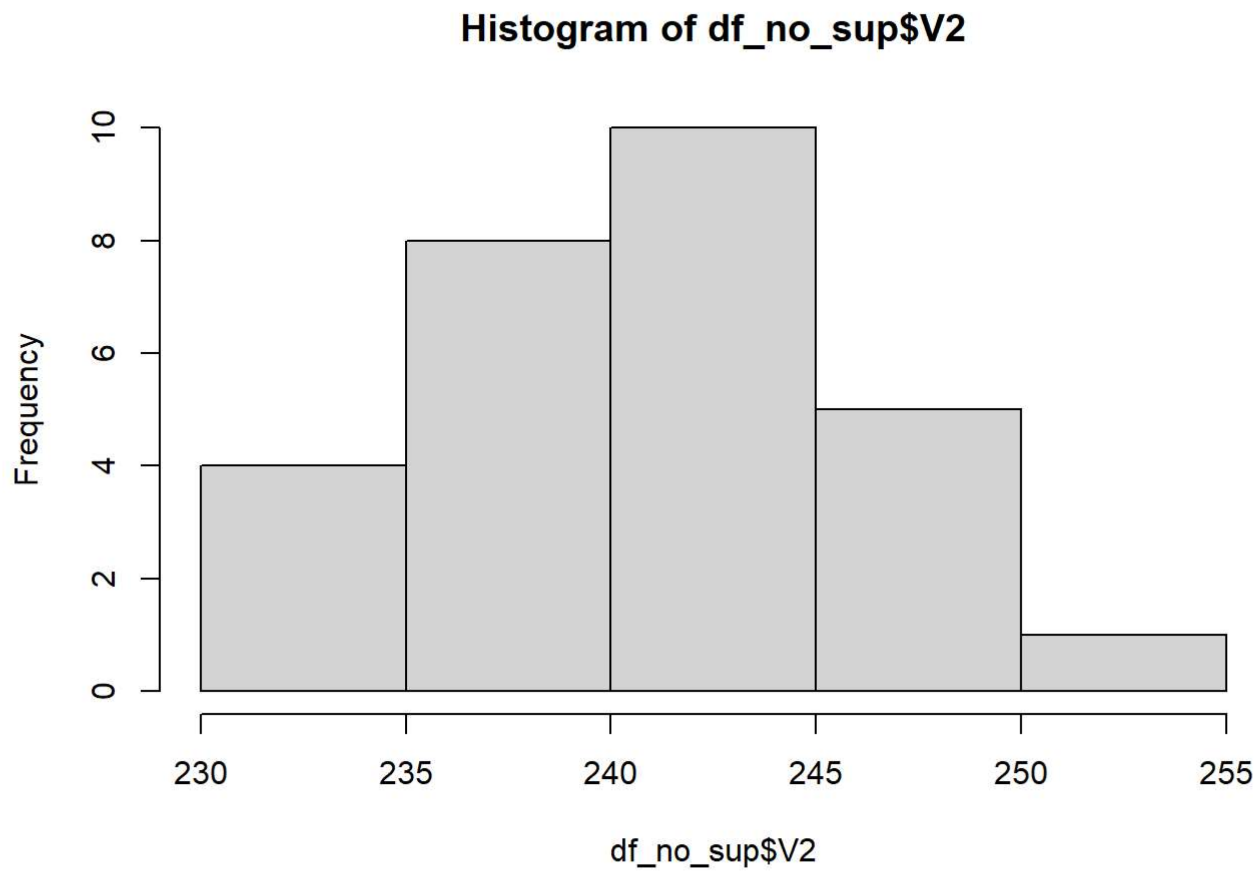

```
qqnorm(df_sup$V4)
```



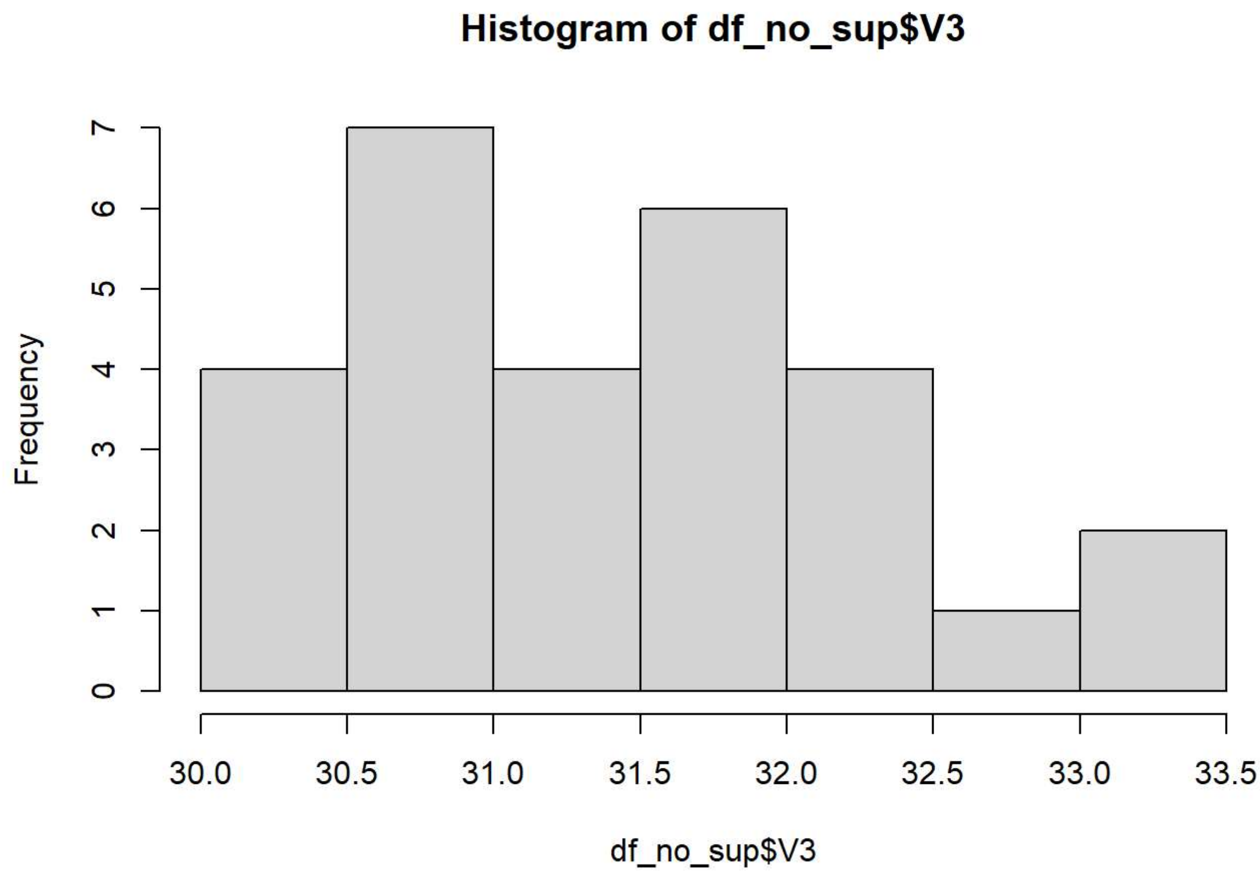

```
hist(df_no_sup$V1)
```



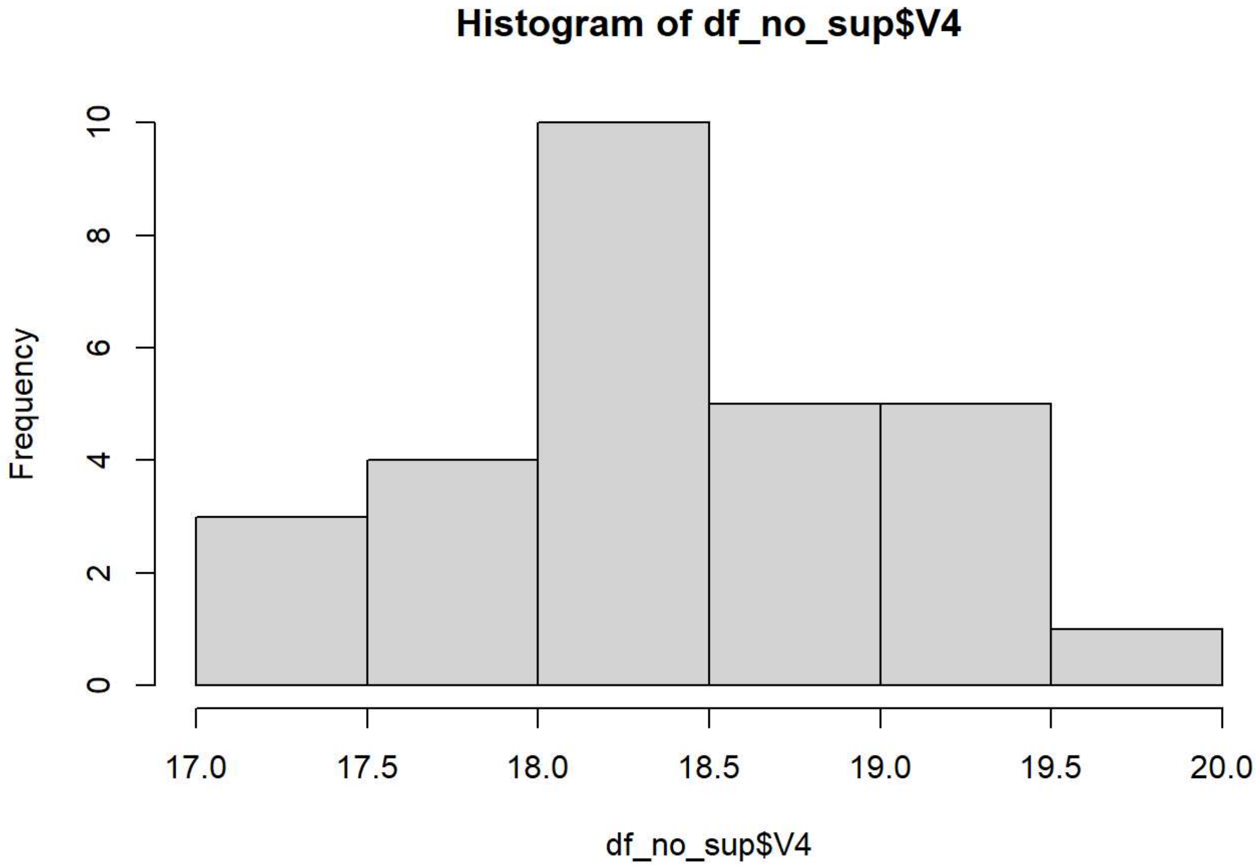

```
hist(df_no_sup$V2)
```



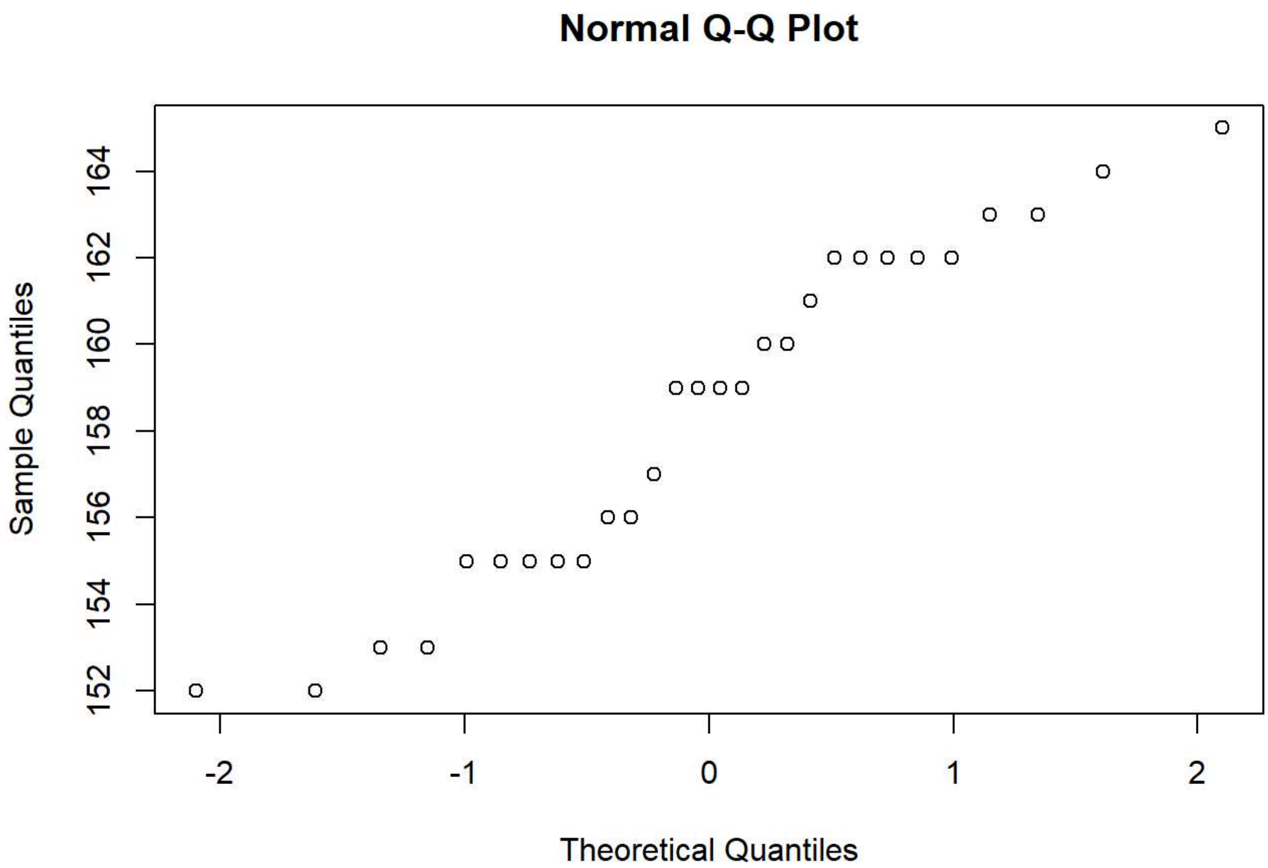

```
hist(df_no_sup$V3)
```



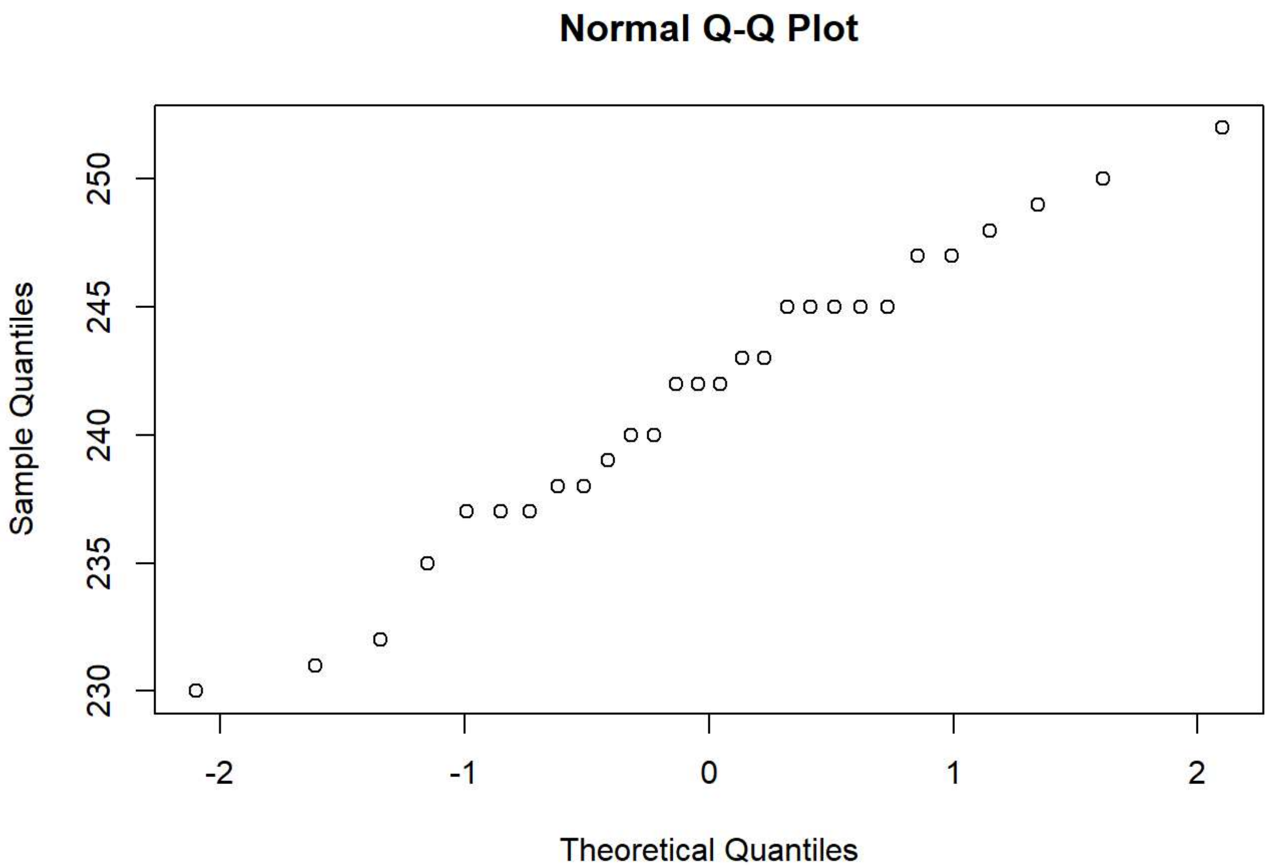

```
hist(df_no_sup$V4)
```



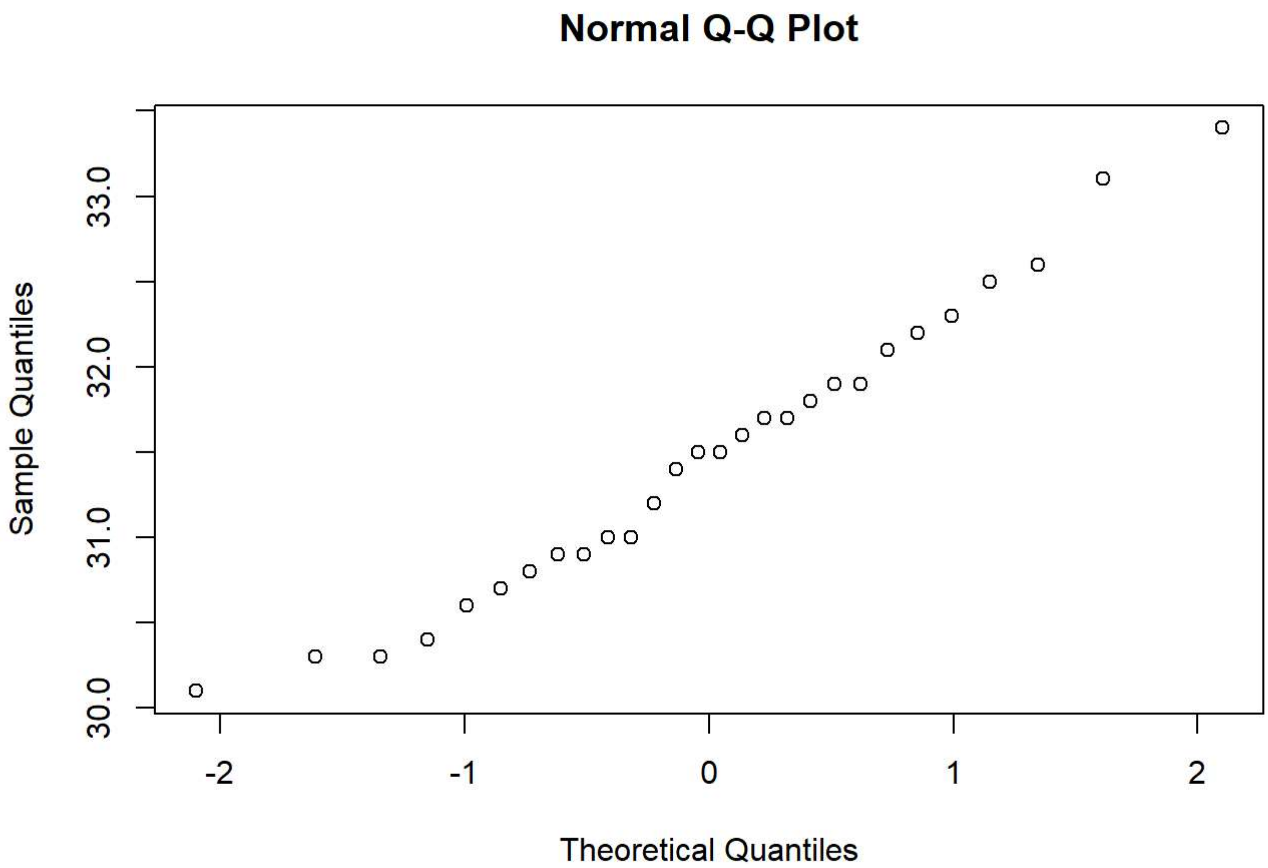

```
qqnorm(df_no_sup$V1)
```



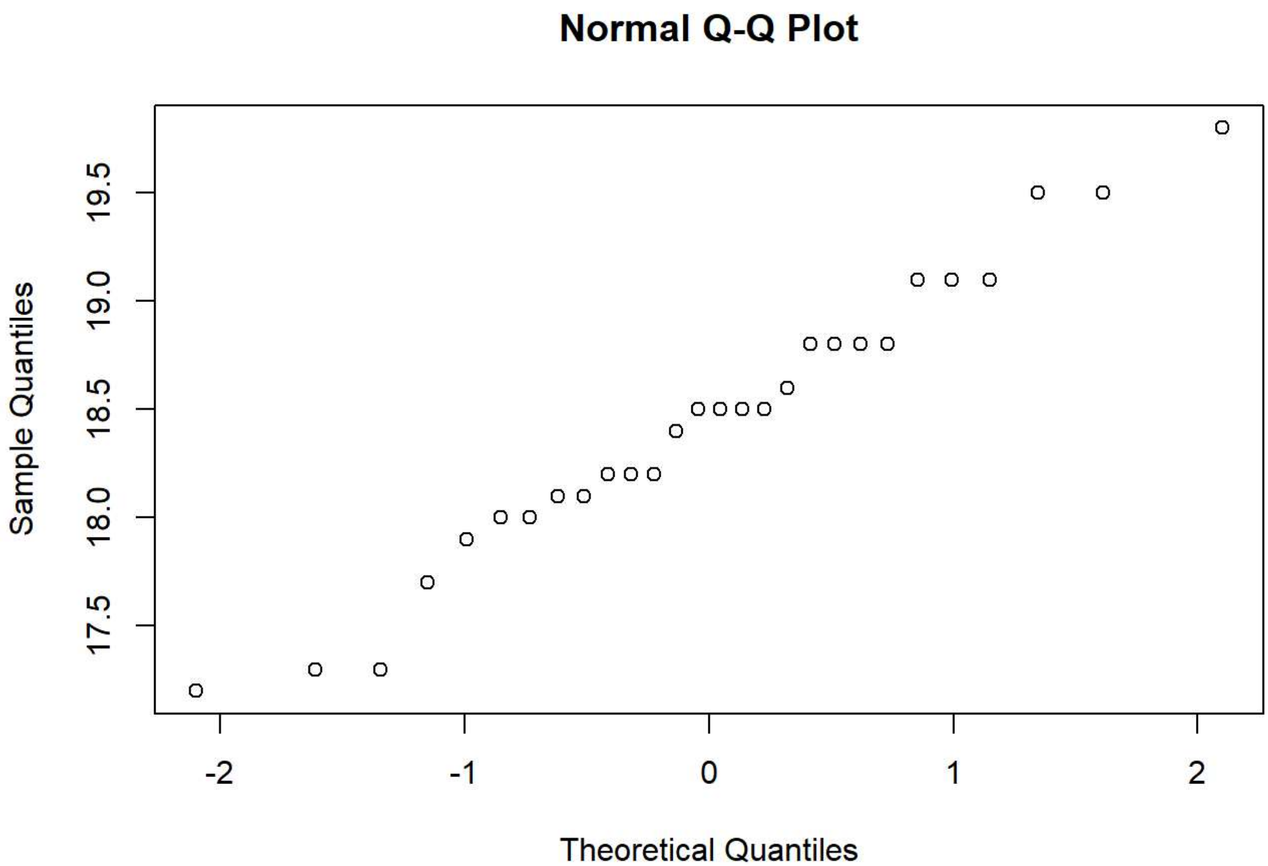

```
qqnorm(df_no_sup$V2)
```



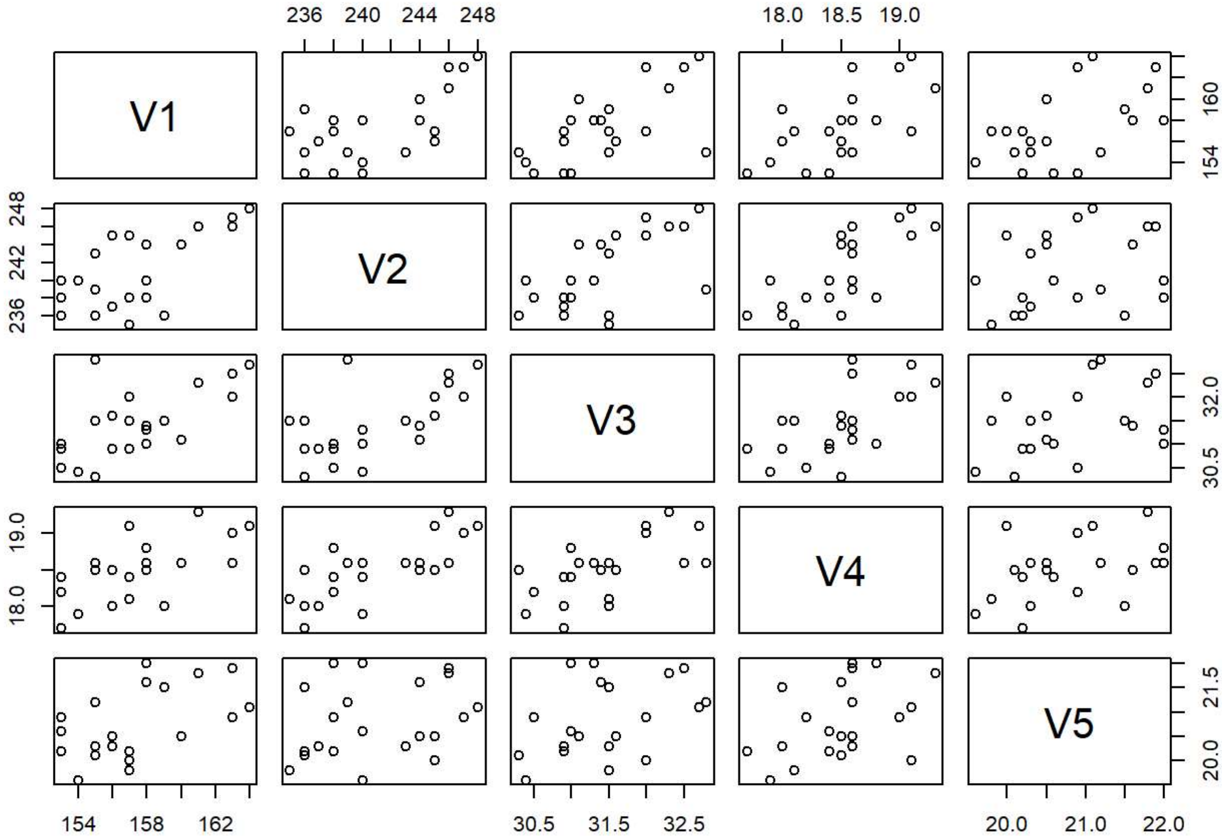

```
qqnorm(df_no_sup$V3)
```



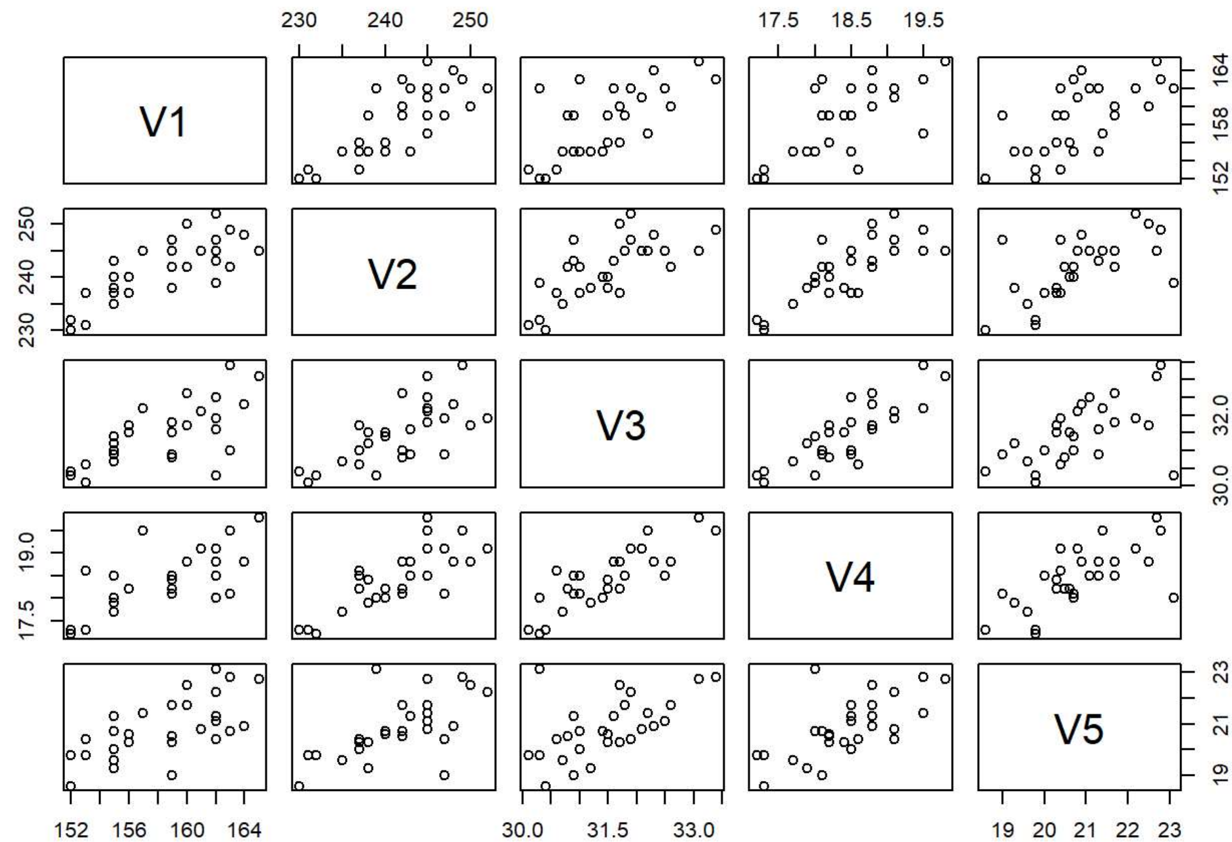

```
qqnorm(df_no_sup$V4)
```



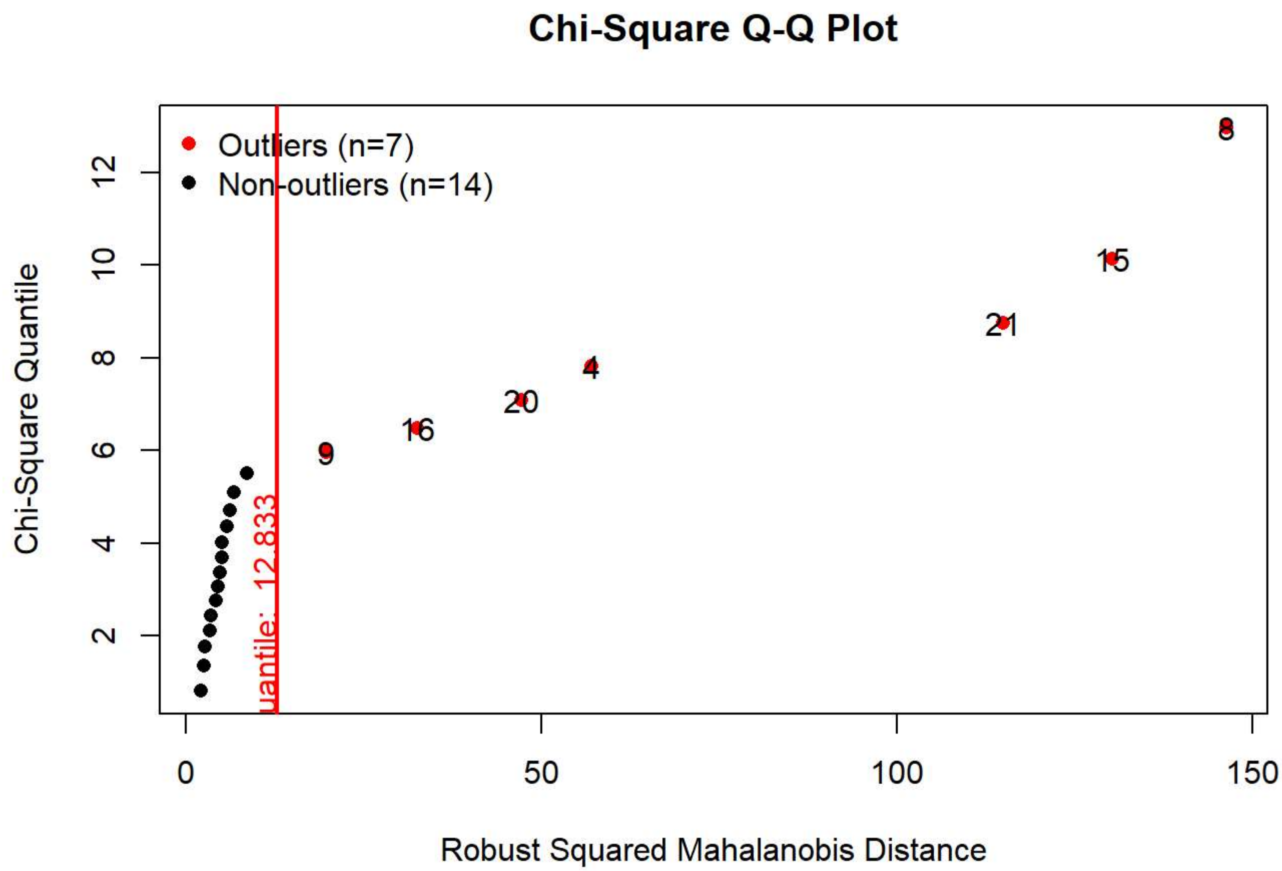

```
# En forma univariada pocas parecieran ser normales

# de todas formas, no hay que confiarse, son pocos datos


# Veamos un grafico bivariado

plot(df_sup)
```




```
# dependiendo el par de variables que vemaos, parecieran tener forma de elipse, posiblemente con mas observaciones


plot(df_no_sup)

# en este caso, la gran mayoria de par de variables tienen forma de elipse, al menos dos a dos parecieran ser normales


# Hagamos un test de shapiro wilks para analizar la normalidad conjunta, tomando alfa 0.05

alfa = 0.05

library(MVN)
```



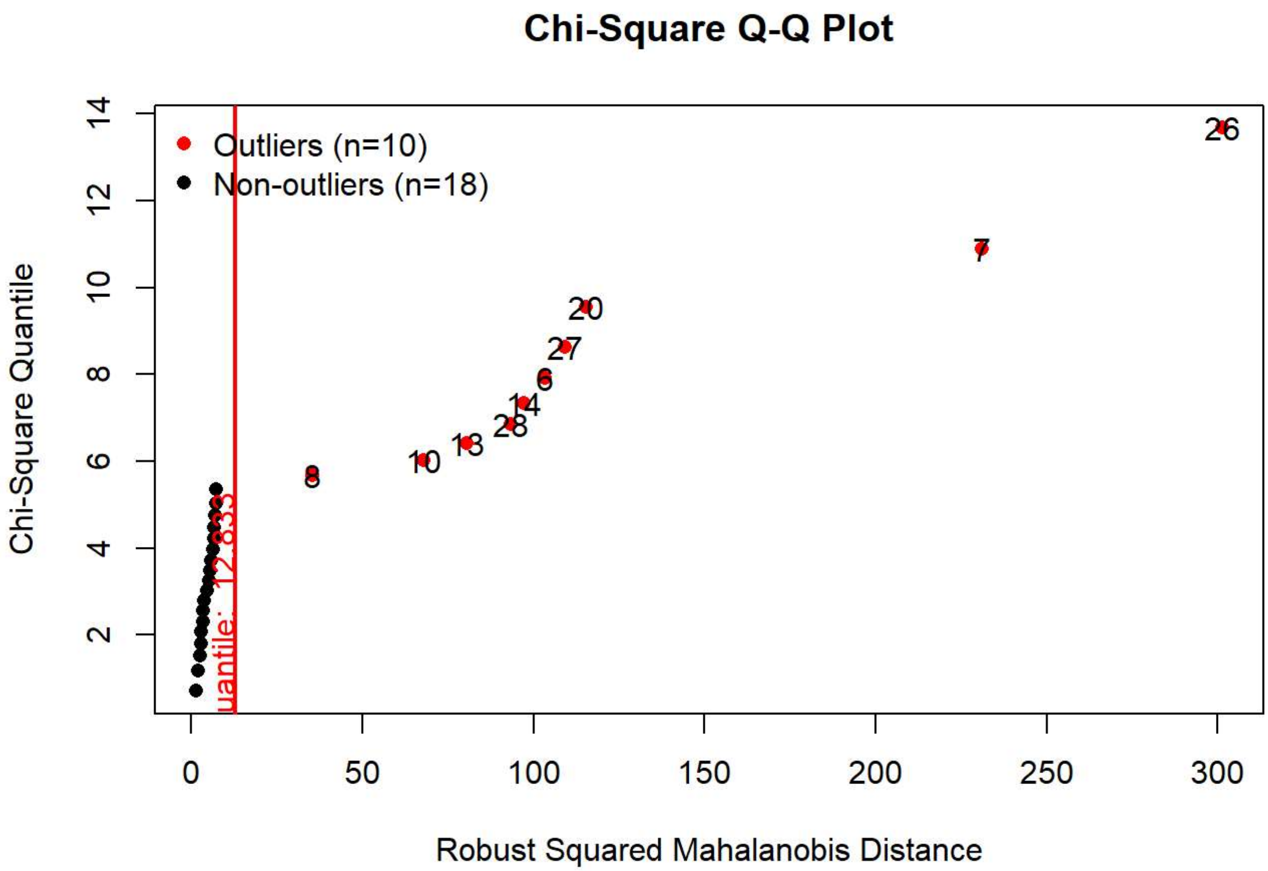

```
test_sup = mvn(df_sup, mvnTest = "hz", multivariateOutlierMethod = "quan")
```




```
pvalue_sup = test_sup$multivariateNormality[3]

ifelse(pvalue_sup<alfa,"rechazo, no hay normalidad","no rechazo, hay normalidad multivariada")
```

```
##      p value
## [1,] "no rechazo, hay normalidad multivariada"
```

```
test_no_sup = mvn(df_no_sup, mvnTest = "hz", multivariateOutlierMethod = "quan")
```




```
pvalue_no_sup = test_no_sup$multivariateNormality[3]

ifelse(pvalue_no_sup<alfa,"rechazo, no hay normalidad","no rechazo, hay normalidad multivariada"
)
```

```
##      p value
## [1,] "no rechazo, hay normalidad multivariada"
```

```
# Al parecer hay normalidad multivariada en ambas muestras, procedemos a construir los estadisti
cos de Hotelling suponiendo

#- Normalidad en los vectores aleatorios e independencia entre muestas
#- Igualdad de matriz de varianzas y covarianzas


n1 = nrow(df_sup)
n2 = nrow(df_no_sup)


df_1_medias = apply(df_sup,2,mean)
df_2_medias = apply(df_no_sup,2,mean)

resta = df_1_medias-df_2_medias

s1 = cov(df_sup)
s2 = cov(df_no_sup)

s = ((n1-1)*s1+(n2-1)*s2)/(n1+n2-2)

# construyo el valor del estadistico To_2

To_2 = ((n1*n2)/(n1+n2))*t(resta)%*%solve(s)%*%resta

p = ncol(df_sup)


Fo= ((n1+n2-p-1)/((n1+n2-2)*p)) * To_2

#busco el fractil de la t student cn p,n1+n2-p-1 grados de libertad

gl1 = p
gl2 = n1+n2-p-1

Fcritico = qf(0.99, gl1, gl2, lower.tail = T, log.p = F)

ifelse(Fo>Fcritico,"Rechazo Ho, no hay igualdad entre los vectores de medias","No Rechazo Ho, ha
y igualdad entre los vectores de medias")
```

```
##      [,1]
## [1,] "No Rechazo Ho, hay igualdad entre los vectores de medias"
```

```
#b

#La combinación lineal del componente de medias es donde se alcanza el supremo, entonces


solve(s)%*%resta
```

```
##           [,1]
## V1 -0.15532570
## V2 -0.02649058
## V3 -0.09285760
## V4  1.03247387
## V5  0.06932512
```