

Tarea 2

A01275465 Carol Arrieta Moreno

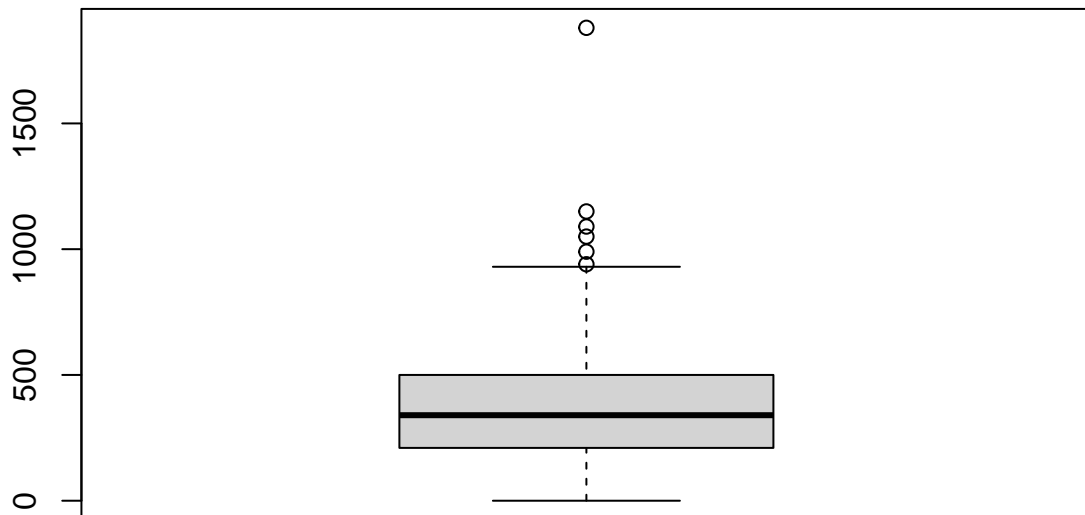
2023-08-17

1. Baja el archivo de trabajo: datos de McDonald

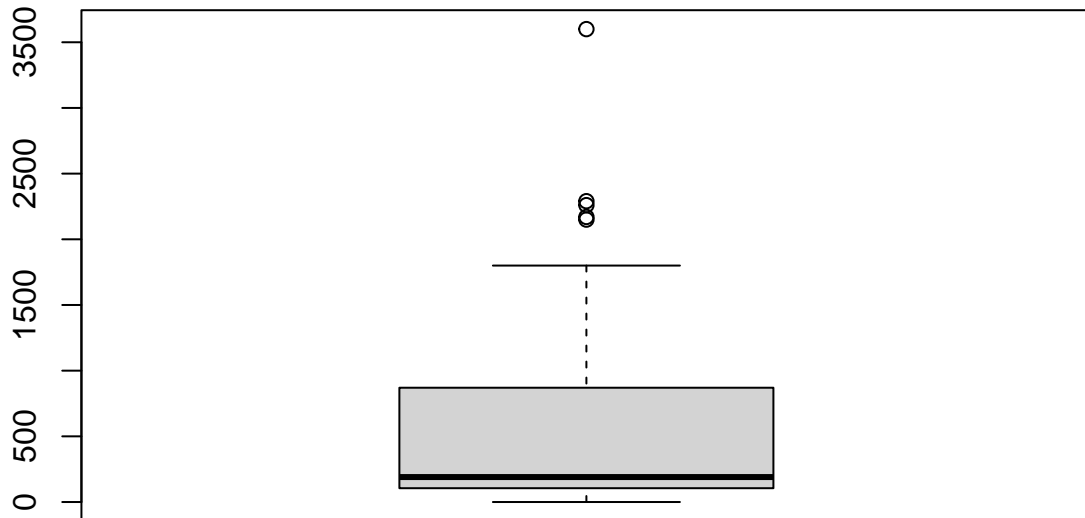
```
M=read.csv("mc-donalds-menu-1.csv") #leer la base de datos
```

2. Analiza 2 de las siguientes variables en cuanto a sus datos atípicos:

```
calorias = M[4]  
Sodio = M[13]  
boxplot(calorias)
```



```
boxplot(Sodio)
```



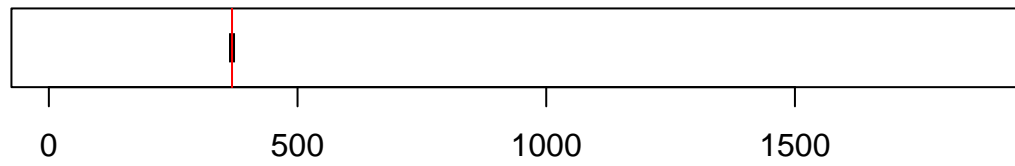
3. Analiza 2 de las siguientes variables en cuanto a sus datos normalidad:

```
X = mean(calorias[["Calories"]])
q1=quantile(X,0.25) #Cuantil 1 de la variable X
q3=quantile(X,0.75)
ri= q3-q1 #Rango intercuartílico de X
par(mfrow=c(2,1)) #Matriz de gráficos de 2x1
y1=0
y2= max(calorias[["Calories"]])
boxplot(X,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red") #línea vertical en el límite de los datos atípicos o extremos
X1= M[M$X<q3+1.5*ri,c("X")] #En la matriz M, quitar datos más allá de 3 rangos intercuartílicos arriba
summary(X1)
```

```
## Length Class Mode
##      0  NULL  NULL
```

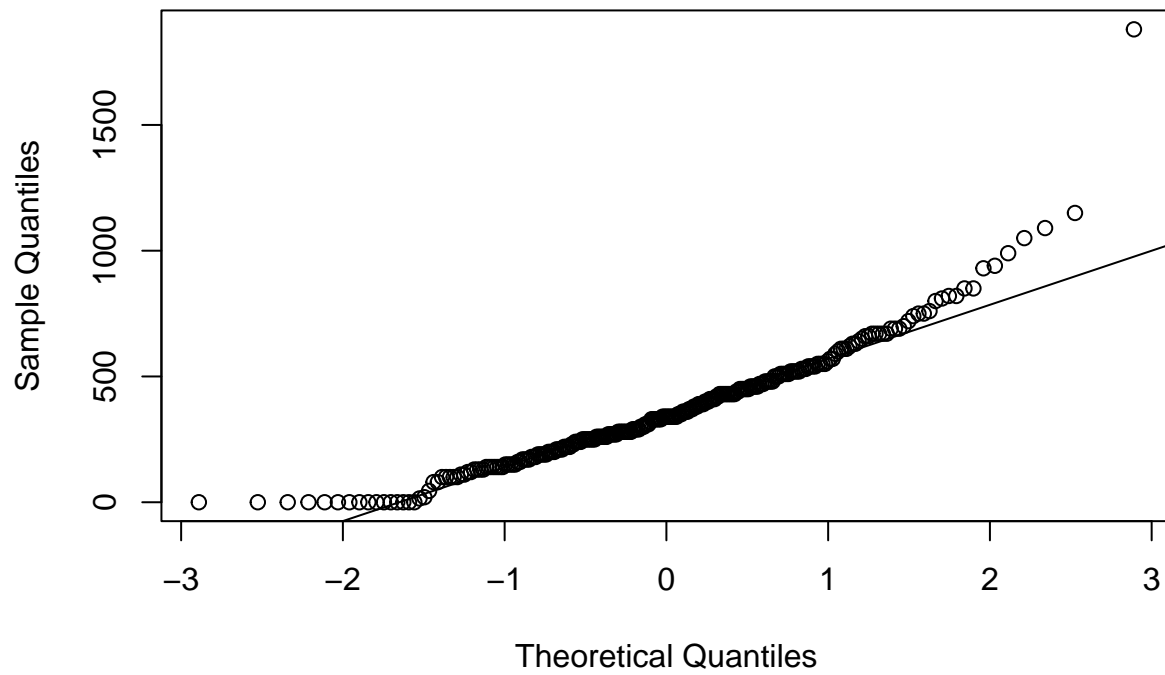
```
summary(X)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    368.3   368.3   368.3   368.3   368.3   368.3
```



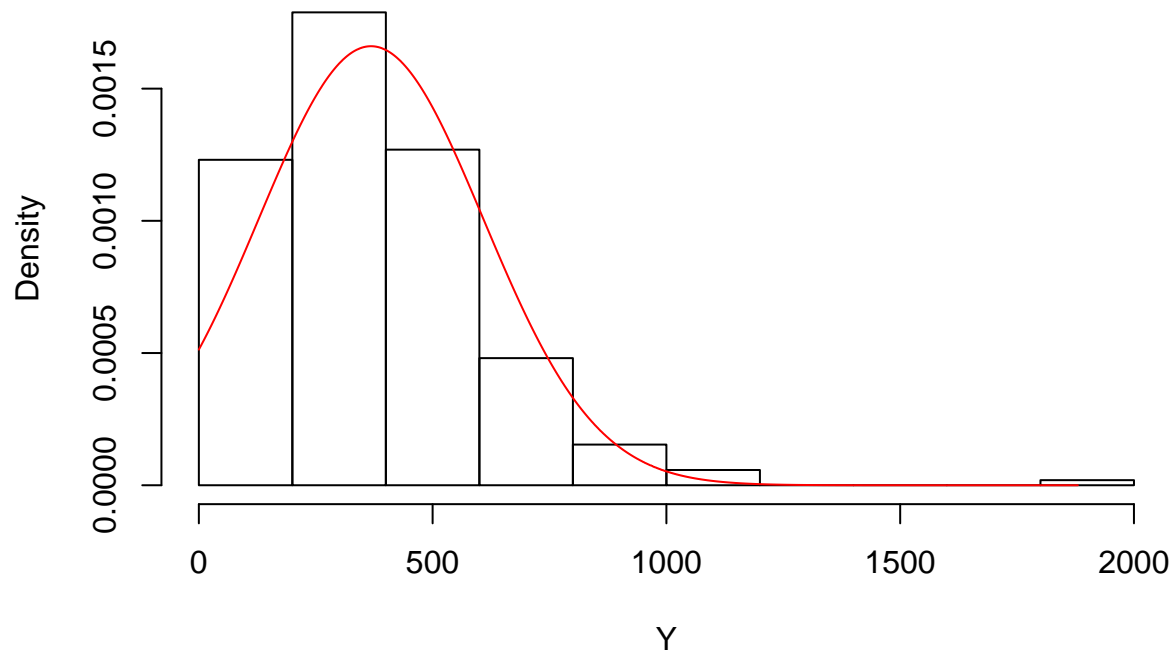
```
Y=calorias[["Calories"]]  
qqnorm(Y)  
qqline(Y)
```

Normal Q-Q Plot



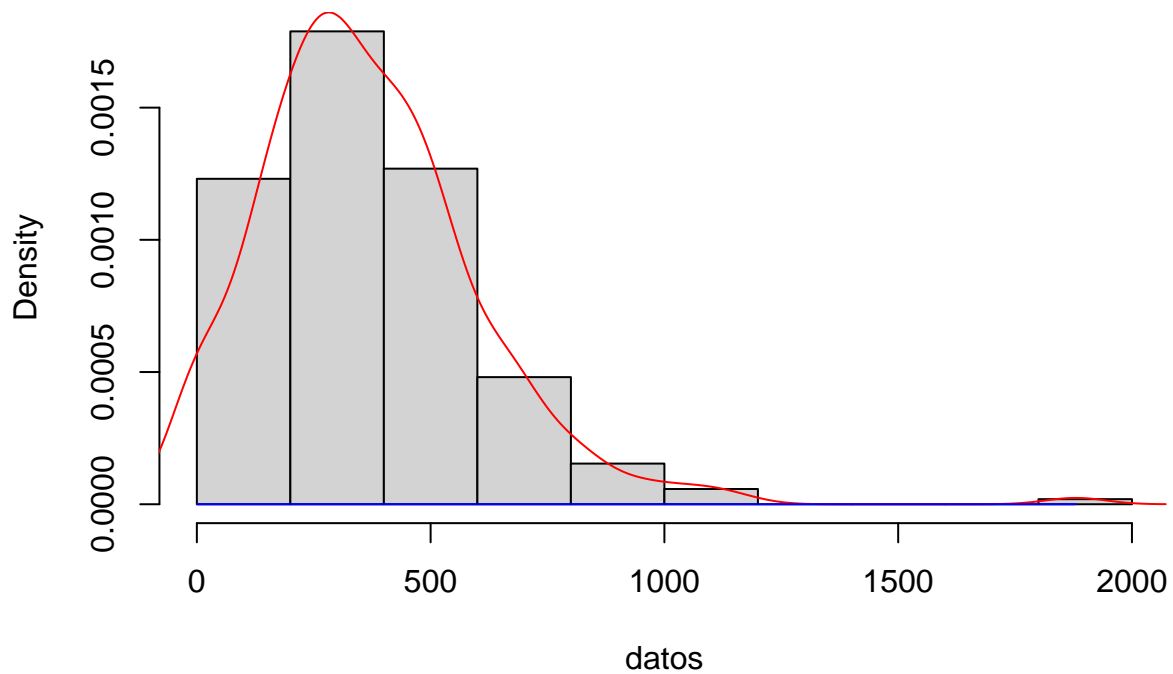
```
hist(Y,prob=TRUE,col=0)
x=seq(min(Y),max(Y),0.1)
y=dnorm(x,mean(Y),sd(Y))
lines(x,y,col="red")
```

Histogram of Y



```
datos=calorias[["Calories"]]  
hist(datos,freq=FALSE)  
lines(density(datos),col="red")  
curve(dnorm(x,mean=mean(datos,sd=sd(datos))), from=y1, to=y2, add=TRUE,col="blue",lwd=1)
```

Histogram of datos

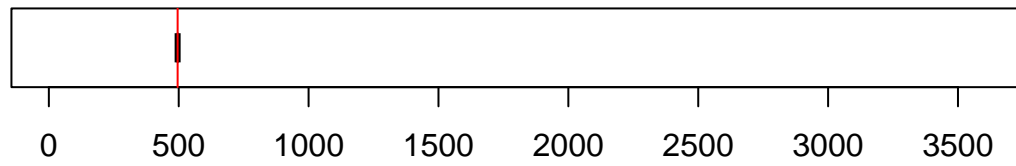


```
X = mean(Sodio[["Sodium"]])
q1=quantile(X,0.25) #Cuantil 1 de la variable X
q3=quantile(X,0.75)
ri= q3-q1 #Rango intercuartílico de X
par(mfrow=c(2,1)) #Matriz de gráficos de 2x1
y1=0
y2= max(Sodio[["Sodium"]])
boxplot(X,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red") #línea vertical en el límite de los datos atípicos o extremos
X1= M[M$X<q3+1.5*ri,c("X")] #En la matriz M, quitar datos más allá de 3 rangos intercuartílicos arriba
summary(X1)
```

```
## Length Class Mode
##      0  NULL  NULL
```

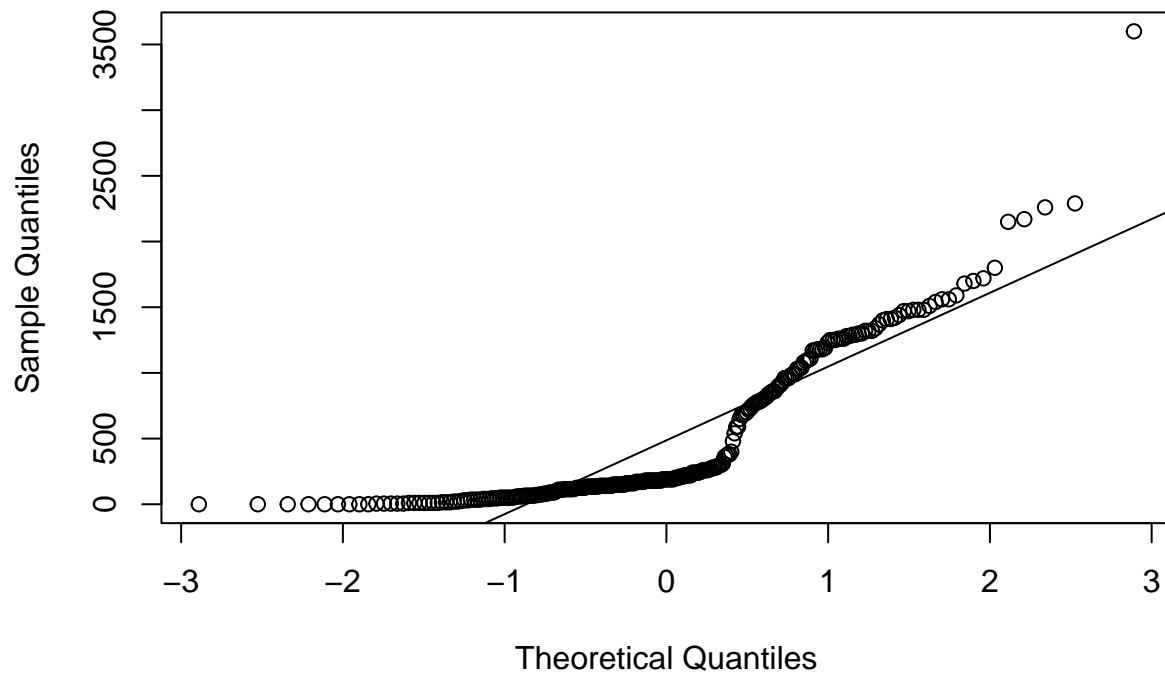
```
summary(X)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    495.8  495.8   495.8   495.8  495.8   495.8
```



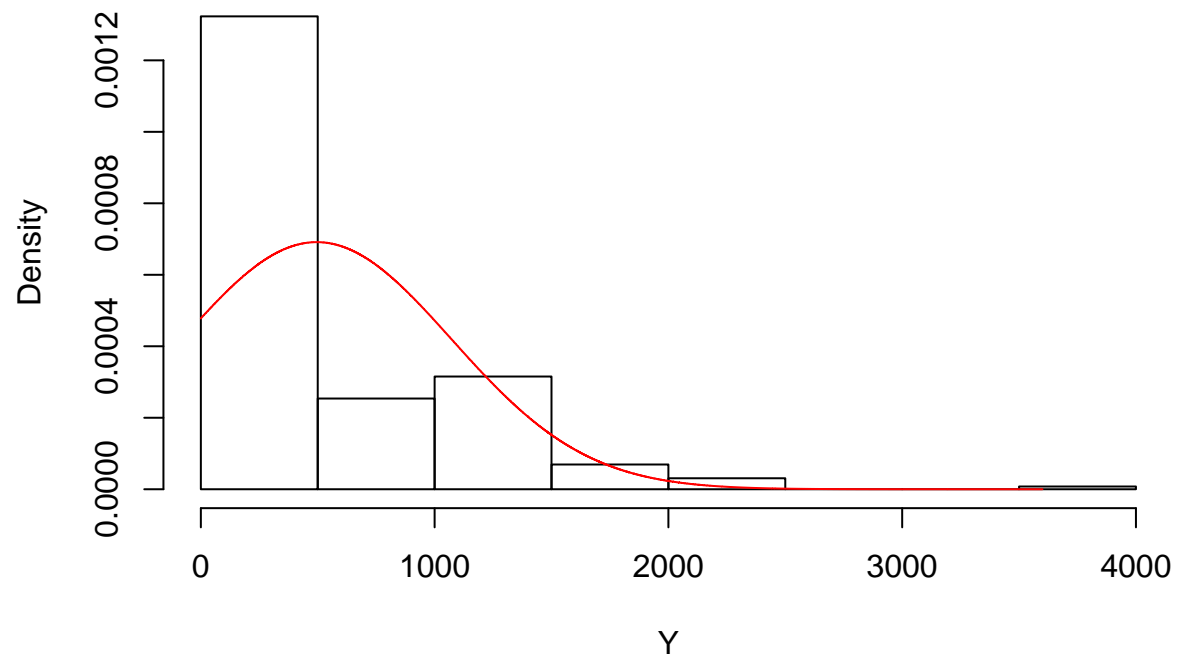
```
Y=Sodio[["Sodium"]]  
qqnorm(Y)  
qqline(Y)
```

Normal Q-Q Plot



```
hist(Y,prob=TRUE,col=0)
x=seq(min(Y),max(Y),0.1)
y=dnorm(x,mean(Y),sd(Y))
lines(x,y,col="red")
```


Histogram of Y



```
datos=Sodio[["Sodium"]]  
hist(datos,freq=FALSE)  
lines(density(datos),col="red")  
curve(dnorm(x,mean=mean(datos,sd=sd(datos))), from=y1, to=y2, add=TRUE,col="blue",lwd=1)
```

Histogram of datos

