McDonald

Erika Martínez Meneses

2024-08-13

Lectura de Datos

```
file.choose()
```

```
## [1] "C:\\Users\\erika\\Documents\\Agos-Dic2024\\Estadística\\mc-
donalds-menu.csv"
```

```
library(readr)
data <- read_csv("C:\\Users\\erika\\Documents\\Agos-
Dic2024\\Estadística\\mc-donalds-menu.csv")
```

```
## Rows: 260 Columns: 24
## ── Column specification
────────────────────────────────────────────────────────
## Delimiter: ","
## chr  (3): Category, Item, Serving Size
## dbl (21): Calories, Calories from Fat, Total Fat, Total Fat (% Daily
Value),...
##
## ℹ Use `spec()` to retrieve the full column specification for this
data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet
this message.
```

```
calorias <- data$Calories
azucares <- data$Sugars
```

Exploración de Datos

```
head(data)
```

```
## # A tibble: 6 × 24
##   Category  Item         `Serving Size` Calories `Calories from Fat`
`Total Fat`
##   <chr>     <chr>        <chr>             <dbl>               <dbl>
<dbl>
## 1 Breakfast Egg McMuffin 4.8 oz (136 g)      300                 120
13
## 2 Breakfast Egg White D… 4.8 oz (135 g)      250                  70
8
## 3 Breakfast Sausage McM… 3.9 oz (111 g)      370                 200
23
```

```
## 4 Breakfast Sausage McM… 5.7 oz (161 g)      450                 250
28
## 5 Breakfast Sausage McM… 5.7 oz (161 g)      400                 210
23
## 6 Breakfast Steak & Egg… 6.5 oz (185 g)      430                 210
23
## # ℹ 18 more variables: `Total Fat (% Daily Value)` <dbl>,
## #   `Saturated Fat` <dbl>, `Saturated Fat (% Daily Value)` <dbl>,
## #   `Trans Fat` <dbl>, Cholesterol <dbl>, `Cholesterol (% Daily
Value)` <dbl>,
## #   Sodium <dbl>, `Sodium (% Daily Value)` <dbl>, Carbohydrates <dbl>,
## #   `Carbohydrates (% Daily Value)` <dbl>, `Dietary Fiber` <dbl>,
## #   `Dietary Fiber (% Daily Value)` <dbl>, Sugars <dbl>, Protein
<dbl>,
## #   `Vitamin A (% Daily Value)` <dbl>, `Vitamin C (% Daily Value)`
<dbl>, …
```

```r
summary(calorias)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0   210.0   340.0   368.3   500.0  1880.0
```

```r
summary(azucares)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    5.75   17.50   29.42   48.00  128.00
```

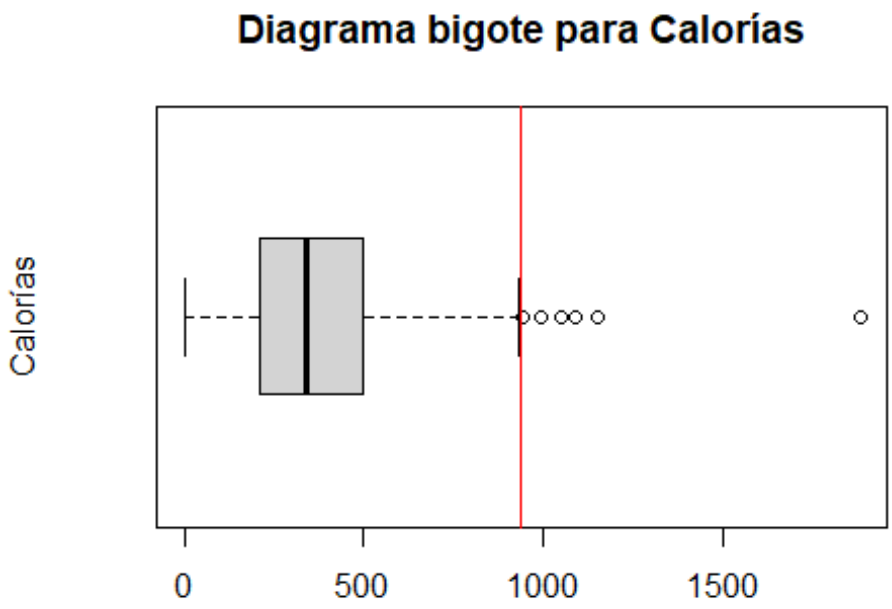
Datos atípicos

Diagrama de caja bigote

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

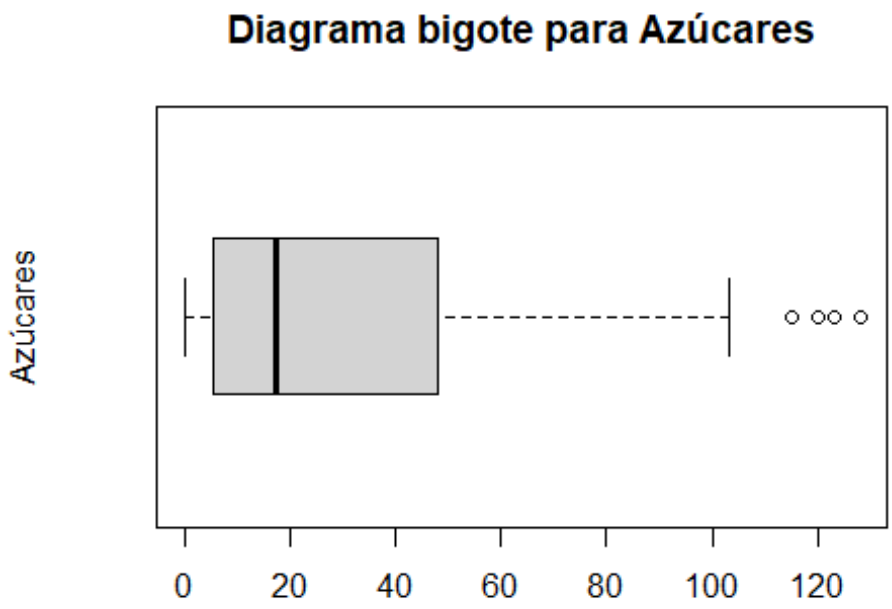
Calorías

```r
boxplot(calorias, horizontal = TRUE, main="Diagrama bigote para
Calorías", ylab="Calorías")
abline(v=q3+1.5*ri,col="red")
```



Azúcares

```
boxplot(azucares, horizontal = TRUE, main="Diagrama bigote para
Azúcares", ylab="Azúcares")
abline(v=q3+1.5*ri,col="red")
```



Podemos observar en los diagramas de caja que para ambas variables existen datos atípicos.

Rango intercuartílico y los cuartíles

El IQR mide la dispersión de la mitad central de los datos. Un IQR pequeño sugiere que los datos están más concentrados, mientras que un IQR grande indica mayor dispersión.

Calorías

```
q1_cal <- quantile(calorias, 0.25)
print("Q1")
```

```
## [1] "Q1"
```

```
q1_cal
```

```
## 25%
## 210
```

```
q3_cal <- quantile(calorias, 0.75)
print("Q3")
```

```
## [1] "Q3"
```

```
q3_cal
```

```
## 75%
## 500
```

```
iqr_cal <- q3_cal - q1_cal
print("IQR")
```

```
## [1] "IQR"
```

```
iqr_cal
```

```
## 75%
## 290
```

Azúcares

```
q1_sug <- quantile(azucares, 0.25)
print("Q1")
```

```
## [1] "Q1"
```

```
q1_sug
```

```
##  25%
## 5.75
```

```
q3_sug <- quantile(azucares, 0.75)
print("Q3")
```

```
## [1] "Q3"
```

```
q3_sug
```

```
## 75%
##  48
```

```
iqr_sug <- q3_sug - q1_sug
print("IQR")
```

```
## [1] "IQR"
```

```
iqr_sug
```

```
##   75%
## 42.25
```

Identifica la cota de 1.5 rangos intercuartílicos para datos atípicos, ¿hay datos atípicos de acuerdo con este criterio?

```
# Límite inferior y superior para datos atípicos (1.5 IQR)
lim_inf_cal <- q1_cal - 1.5 * iqr_cal
print("Límite inferior")
```

```
## [1] "Límite inferior"
```

```
lim_inf_cal
```

```
##  25%
## -225
```

```
lim_sup_cal <- q3_cal + 1.5 * iqr_cal
print("Límite superior")
```

```
## [1] "Límite superior"
```

```
lim_sup_cal
```

```
## 75%
## 935
```

```
outliers_cal_1.5iqr <- calorias[calorias < lim_inf_cal | calorias >
lim_sup_cal]
print("outliers")
```

```
## [1] "outliers"
```

```
outliers_cal_1.5iqr
```

```
## [1] 1090 1150  990 1050  940 1880
```

```
lim_inf_sug <- q1_sug - 1.5 * iqr_sug
print("Límite inferior")
```

```
## [1] "Límite inferior"
```

```
lim_inf_sug
```

```
##     25%
## -57.625
```

```
lim_sup_sug <- q3_sug + 1.5 * iqr_sug
print("Límite superior")
```

```
## [1] "Límite superior"
```

```
lim_sup_sug
```

```
##     75%
## 111.375
```

```
outliers_sug_1.5iqr <- azucares[azucares < lim_inf_sug | azucares >
lim_sup_sug]
print("outliers")
```

```
## [1] "outliers"
```

```
outliers_sug_1.5iqr
```

```
## [1] 123 120 115 128
```

Identifica la cota de 3 desviaciones estándar alrededor de la media, ¿hay datos atípicos de acuerdo con este criterio?

```
# Límite inferior y superior para datos atípicos (3 desviaciones
estándar)
mean_cal <- mean(calorias)
print("Media")
```

```
## [1] "Media"
```

```
mean_cal
```

```
## [1] 368.2692
```

```
sd_cal <- sd(calorias)
print("Desviación estándar")
```

```
## [1] "Desviación estándar"
```

```
sd_cal
```

```
## [1] 240.2699
```

```
outliers_cal_3sd <- calorias[calorias < mean_cal - 3 * sd_cal | calorias
> mean_cal + 3 * sd_cal]
print("Outliers")
```

```
## [1] "Outliers"
```

```
outliers_cal_3sd
```

```
## [1] 1090 1150 1880
```

```
mean_sug <- mean(azucares)
print("Media")
```

```
## [1] "Media"
```

```
mean_sug
```

```
## [1] 29.42308
```

```
sd_sug <- sd(azucares)
print("Desviación estandar")
```

```
## [1] "Desviación estandar"
```

```
sd_sug
```

```
## [1] 28.6798
```

```
outliers_sug_3sd <- azucares[azucares < mean_sug - 3 * sd_sug | azucares
> mean_sug + 3 * sd_sug]
print("Outliers")
```

```
## [1] "Outliers"

outliers_sug_3sd

## [1] 123 120 128
```

Podemos observar que a través de los diferentes métodos que se encuentran datos fuera de los límites, datos atípicos y comunmente estos datos serían eliminados, sin embargo, en el contexto de nuestro problema estos datos son relevantes para mantener el análisis realista, ya que los datos son representativos del menú, existen postre extremadamente azucarado y hamburguesas muy alta en calorías por lo que no considero conveniente eliminar estos datos.

Quitar los datos atípicos

A pesar de que se ha decidido no eliminar los datos atípicos, a continuación se muestra el proceso que seguiría para eliminar dichos datos.

```
X1= data[calorias<q3+1.5*ri,]  #En la matriz M, quitar datos más allá de
3 rangos intercuartílicos arriba de q3 de la variable X
summary(X1)

##    Category             Item           Serving Size          Calories
##  Length:254         Length:254         Length:254         Min.   :
0.0
##  Class :character   Class :character   Class :character   1st
Qu.:202.5
##  Mode  :character   Mode  :character   Mode  :character   Median
:335.0
##                                                           Mean
:349.0
##                                                           3rd
Qu.:480.0
##                                                           Max.
:930.0
##  Calories from Fat   Total Fat      Total Fat (% Daily Value)
Saturated Fat
##  Min.   :  0.0     Min.   : 0.000   Min.   : 0.00            Min.   :
0.000
##  1st Qu.: 12.5     1st Qu.: 1.625   1st Qu.: 2.25            1st Qu.:
1.000
##  Median :100.0     Median :11.000   Median :17.00            Median :
5.000
##  Mean   :116.3     Mean   :12.969   Mean   :19.97            Mean   :
5.752
##  3rd Qu.:197.5     3rd Qu.:22.000   3rd Qu.:33.00            3rd Qu.:
9.000
##  Max.   :470.0     Max.   :52.000   Max.   :80.00            Max.
:20.000
##  Saturated Fat (% Daily Value)   Trans Fat       Cholesterol
##  Min.   :  0.00                Min.   :0.0000   Min.   :  0.00
```

```
##  1st Qu.:  4.00                1st Qu.:0.0000   1st Qu.:  5.00
##  Median : 24.00                Median :0.0000   Median : 30.00
##  Mean   : 28.68                Mean   :0.2047   Mean   : 49.70
##  3rd Qu.: 45.00                3rd Qu.:0.0000   3rd Qu.: 63.75
##  Max.   :102.00                Max.   :2.5000   Max.   :555.00
##  Cholesterol (% Daily Value)     Sodium       Sodium (% Daily Value)
##  Min.   :  0.00              Min.   :   0.0   Min.   : 0.00
##  1st Qu.:  2.00              1st Qu.:  92.5   1st Qu.: 4.00
##  Median : 11.00              Median : 185.0   Median : 8.00
##  Mean   : 16.64              Mean   : 451.3   Mean   :18.82
##  3rd Qu.: 21.00              3rd Qu.: 817.5   3rd Qu.:34.00
##  Max.   :185.00              Max.   :1720.0   Max.   :72.00
##  Carbohydrates    Carbohydrates (% Daily Value) Dietary Fiber
##  Min.   :  0.00   Min.   : 0.00                 Min.   :0.000
##  1st Qu.: 30.00   1st Qu.:10.00                 1st Qu.:0.000
##  Median : 43.50   Median :14.50                 Median :1.000
##  Mean   : 45.99   Mean   :15.33                 Mean   :1.531
##  3rd Qu.: 58.00   3rd Qu.:19.00                 3rd Qu.:2.000
##  Max.   :141.00   Max.   :47.00                 Max.   :7.000
##  Dietary Fiber (% Daily Value)     Sugars          Protein
##  Min.   : 0.000                Min.   :  0.00   Min.   : 0.00
##  1st Qu.: 0.000                1st Qu.:  6.00   1st Qu.: 4.00
##  Median : 5.000                Median : 19.00   Median :12.00
##  Mean   : 6.142                Mean   : 29.84   Mean   :12.58
##  3rd Qu.:10.000                3rd Qu.: 48.00   3rd Qu.:18.00
##  Max.   :28.000                Max.   :128.00   Max.   :48.00
##  Vitamin A (% Daily Value) Vitamin C (% Daily Value) Calcium (% Daily
Value)
##  Min.   :  0.00            Min.   :  0.000           Min.   : 0.00
##  1st Qu.:  2.00            1st Qu.:  0.000           1st Qu.: 6.00
##  Median :  8.00            Median :  0.000           Median :20.00
##  Mean   : 13.61            Mean   :  8.614           Mean   :21.01
##  3rd Qu.: 15.00            3rd Qu.:  4.000           3rd Qu.:30.00
##  Max.   :170.00            Max.   :240.000           Max.   :70.00
##  Iron (% Daily Value)
##  Min.   : 0.000
##  1st Qu.: 0.000
##  Median : 4.000
##  Mean   : 7.228
##  3rd Qu.:15.000
##  Max.   :35.000
```

```
X1= data[azucares<q3+1.5*ri,]  #En la matriz M, quitar datos más allá de
3 rangos intercuartílicos arriba de q3 de la variable X
summary(X1)
```

```
##    Category            Item           Serving Size          Calories
##  Length:260         Length:260         Length:260         Min.   :
0.0
##  Class :character   Class :character   Class :character   1st Qu.:
```

```
210.0
##  Mode  :character   Mode  :character   Mode  :character   Median :
340.0
##                                                          Mean   :
368.3
##                                                          3rd Qu.:
500.0
##                                                          Max.
:1880.0
##  Calories from Fat   Total Fat       Total Fat (% Daily Value)
Saturated Fat
##  Min.   :   0.0    Min.   :  0.000   Min.   :  0.00            Min.
: 0.000
##  1st Qu.:  20.0    1st Qu.:  2.375   1st Qu.:  3.75            1st
Qu.: 1.000
##  Median : 100.0    Median : 11.000   Median : 17.00            Median
: 5.000
##  Mean   : 127.1    Mean   : 14.165   Mean   : 21.82            Mean
: 6.008
##  3rd Qu.: 200.0    3rd Qu.: 22.250   3rd Qu.: 35.00            3rd
Qu.:10.000
##  Max.   :1060.0    Max.   :118.000   Max.   :182.00            Max.
:20.000
##  Saturated Fat (% Daily Value)   Trans Fat       Cholesterol
##  Min.   :  0.00                Min.   :0.0000   Min.   :  0.00
##  1st Qu.:  4.75                1st Qu.:0.0000   1st Qu.:  5.00
##  Median : 24.00                Median :0.0000   Median : 35.00
##  Mean   : 29.97                Mean   :0.2038   Mean   : 54.94
##  3rd Qu.: 48.00                3rd Qu.:0.0000   3rd Qu.: 65.00
##  Max.   :102.00                Max.   :2.5000   Max.   :575.00
##  Cholesterol (% Daily Value)     Sodium       Sodium (% Daily Value)
##  Min.   :  0.00              Min.   :   0.0   Min.   :  0.00
##  1st Qu.:  2.00              1st Qu.: 107.5   1st Qu.:  4.75
##  Median : 11.00              Median : 190.0   Median :  8.00
##  Mean   : 18.39              Mean   : 495.8   Mean   : 20.68
##  3rd Qu.: 21.25              3rd Qu.: 865.0   3rd Qu.: 36.25
##  Max.   :192.00              Max.   :3600.0   Max.   :150.00
##  Carbohydrates    Carbohydrates (% Daily Value) Dietary Fiber
##  Min.   :  0.00   Min.   : 0.00                 Min.   :0.000
##  1st Qu.: 30.00   1st Qu.:10.00                 1st Qu.:0.000
##  Median : 44.00   Median :15.00                 Median :1.000
##  Mean   : 47.35   Mean   :15.78                 Mean   :1.631
##  3rd Qu.: 60.00   3rd Qu.:20.00                 3rd Qu.:3.000
##  Max.   :141.00   Max.   :47.00                 Max.   :7.000
##  Dietary Fiber (% Daily Value)     Sugars          Protein
##  Min.   : 0.000                Min.   :  0.00   Min.   : 0.00
##  1st Qu.: 0.000                1st Qu.:  5.75   1st Qu.: 4.00
##  Median : 5.000                Median : 17.50   Median :12.00
##  Mean   : 6.531                Mean   : 29.42   Mean   :13.34
##  3rd Qu.:10.000                3rd Qu.: 48.00   3rd Qu.:19.00
```

```
##  Max.   :28.000                Max.   :128.00   Max.   :87.00
##  Vitamin A (% Daily Value) Vitamin C (% Daily Value) Calcium (% Daily
Value)
##  Min.   :  0.00            Min.   :  0.000          Min.   : 0.00
##  1st Qu.:  2.00            1st Qu.:  0.000          1st Qu.: 6.00
##  Median :  8.00            Median :  0.000          Median :20.00
##  Mean   : 13.43            Mean   :  8.535          Mean   :20.97
##  3rd Qu.: 15.00            3rd Qu.:  4.000          3rd Qu.:30.00
##  Max.   :170.00            Max.   :240.000          Max.   :70.00
##  Iron (% Daily Value)
##  Min.   : 0.000
##  1st Qu.: 0.000
##  Median : 4.000
##  Mean   : 7.735
##  3rd Qu.:15.000
##  Max.   :40.000
```

Prueba de normalidad univariada

H_{0} = La muestra proviene de una distribución normal h_{1} = La muestra no proviene de una distribución normal

```
library(nortest)
ad.test(calorias)

##
##  Anderson-Darling normality test
##
## data:  calorias
## A = 2.5088, p-value = 2.369e-06
```

El valor p es extremadamente pequeño (mucho menor que 0.05), lo que significa que hay suficiente evidencia para rechazar la hipótesis nula. Esto indica que la muestra de calorías no proviene de una distribución normal.

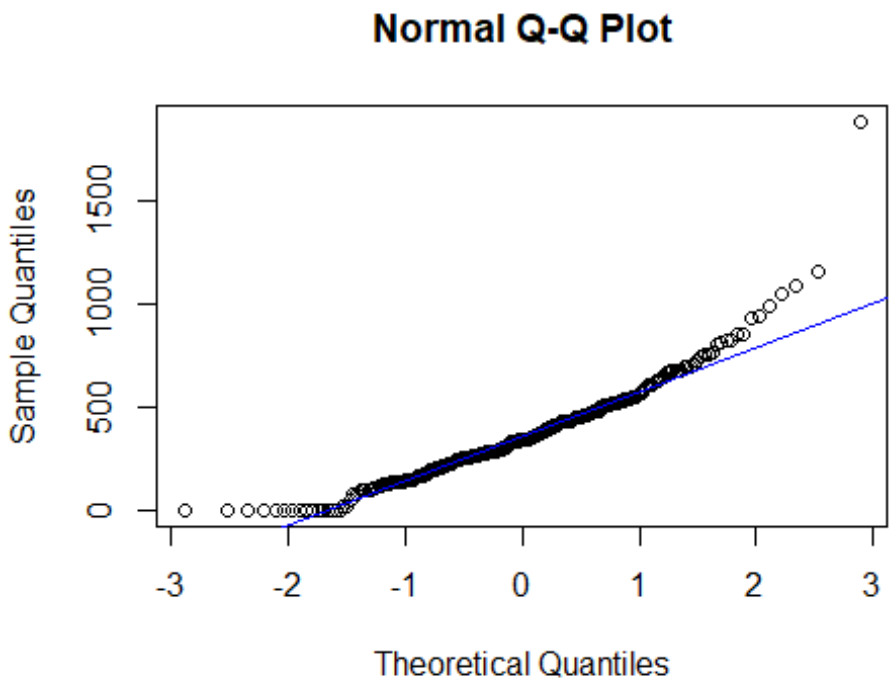
```
library(nortest)
ad.test(azucares)

##
##  Anderson-Darling normality test
##
## data:  azucares
## A = 9.9899, p-value < 2.2e-16
```

En el caso del azucar, el valor p es aún más pequeño. Esto refuerza aún más la evidencia en contra de la hipótesis nula, sugiriendo que la muestra de azúcares no proviene de una distribución normal.

QQPlot

Calorías

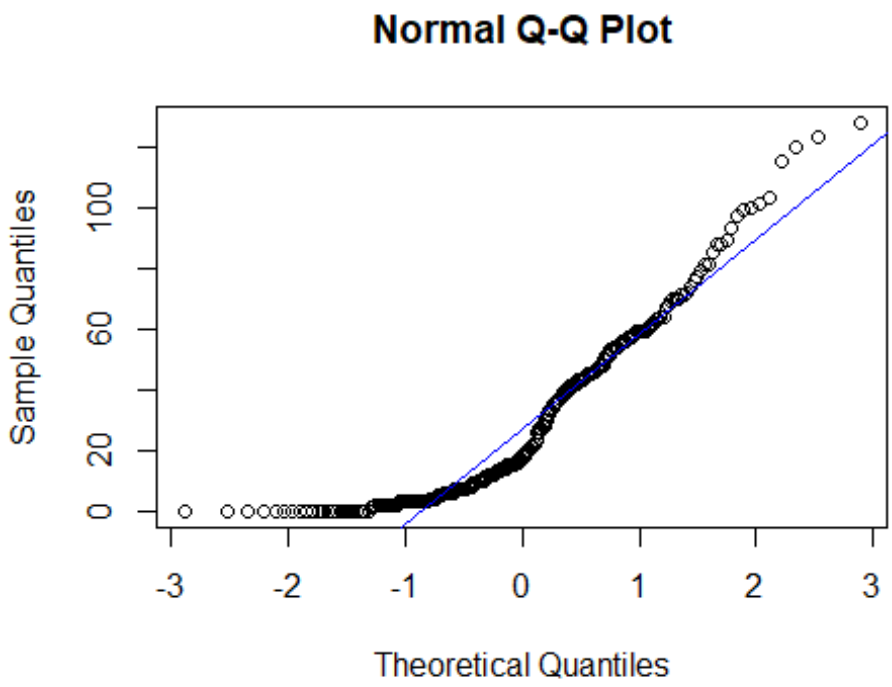
```
qqnorm(calorias)
qqline(calorias, col = "blue")
```



Cuando los puntos en el QQ plot siguen aproximadamente la línea recta, los datos son aproximadamente normales, en el caso de las calorías en su mayoría siguen la línea, sin embargo, los puntos en la cola izquierda están por debajo de la línea teórica, lo que indica que hay menos valores extremadamente bajos de calorías de lo que se esperaría en una distribución normal, y en la superior los puntos se alejan considerablemente de la línea, con varios puntos muy por encima de la línea. Esto indica la presencia de valores extremadamente altos en la distribución de calorías que no serían esperados si los datos fueran normalmente distribuidos.

Azúcares

```
qqnorm(azucares)
qqline(azucares, col = "blue")
```



En el caso de los azucares podemos observar que pasa un comportamiento muy similar en la cola superior e inferior, sin embargo aquí es más marcado y la mayoría de los puntos se alejan considerablemente de la línea.

Coeficiente de sesgo y el coeficiente de curtosis

Sesgo: Indica la asimetría de la distribución. Un sesgo cercano a 0 indica simetría; valores positivos indican una cola larga a la derecha, y negativos a la izquierda.

Curtosis: Mide la "agudeza" de la distribución. Un valor de curtosis cercano a 3 indica una distribución normal (mesocúrtica), valores mayores indican distribuciones con colas más pesadas (leptocúrtica), y menores indican colas ligeras (platicúrtica).

Calorías

```
library(moments)
sesgo_cal <- skewness(calorias)
print("Sesgo")
```

```
## [1] "Sesgo"
```

```
sesgo_cal
```

```
## [1] 1.444105
```

```
curtosis_cal <- kurtosis(calorias)
print("Curtosis")
```

```
## [1] "Curtosis"
```

```
curtosis_cal
```

```
## [1] 8.645274
```

En el caso de las calorías los valores indican que la distribución de las calorías está sesgada hacia la derecha. Esto significa que hay una cola más larga en el lado derecho de la distribución, lo que sugiere que existen valores más altos de calorías que están menos representados. Y por el lado de la curtosis es bastante alta, lo que indica que la distribución tiene colas pesadas (leptocúrtica). Esto significa que hay más valores extremos (o datos atípicos) en ambas colas en comparación con una distribución normal

Azúcares

```
sesgo_sug <- skewness(azucares)
print("Sesgo")
```

```
## [1] "Sesgo"
```

```
sesgo_sug
```

```
## [1] 1.025977
```

```
curtosis_sug <- kurtosis(azucares)
print("Curtosis")
```

```
## [1] "Curtosis"
```

```
curtosis_sug
```

```
## [1] 3.487744
```

En el caso de los azucares los valores también indican una distribución sesgada hacia la derecha, aunque el sesgo es menor que en el caso de las calorías. Esto sugiere que, aunque hay valores altos de azúcares, son menos extremos que los valores altos de calorías, mientras que la curtosis es ligeramente superior a la de una distribución normal, lo que indica que la distribución tiene colas ligeramente más pesadas. Esto sugiere la presencia de algunos valores atípicos pero es mucho menor que las calorías y se aproxima bastante a 3.

Comparación de media, mediana y rango medio

Calorías

```
media_cal <- mean(calorias)
print("Media")
```

```
## [1] "Media"
```

```
media_cal
```

```
## [1] 368.2692
```

```
mediana_cal <- median(calorias)
print("Mediana")
```

```
## [1] "Mediana"
```

```
mediana_cal
```

```
## [1] 340
```

```
rango_medio_cal <- (q3_cal + q1_cal) / 2
print("Rango")
```

```
## [1] "Rango"
```

```
rango_medio_cal
```

```
## 75%
## 355
```

Azúcares

```
media_sug <- mean(azucares)
print("Media")
```

```
## [1] "Media"
```

```
media_sug
```

```
## [1] 29.42308
```

```
mediana_sug <- median(azucares)
print("Mediana")
```

```
## [1] "Mediana"
```

```
mediana_sug
```

```
## [1] 17.5
```

```
rango_medio_sug <- (q3_sug + q1_sug) / 2
print("Rango")
```

```
## [1] "Rango"
```

```
rango_medio_sug
```

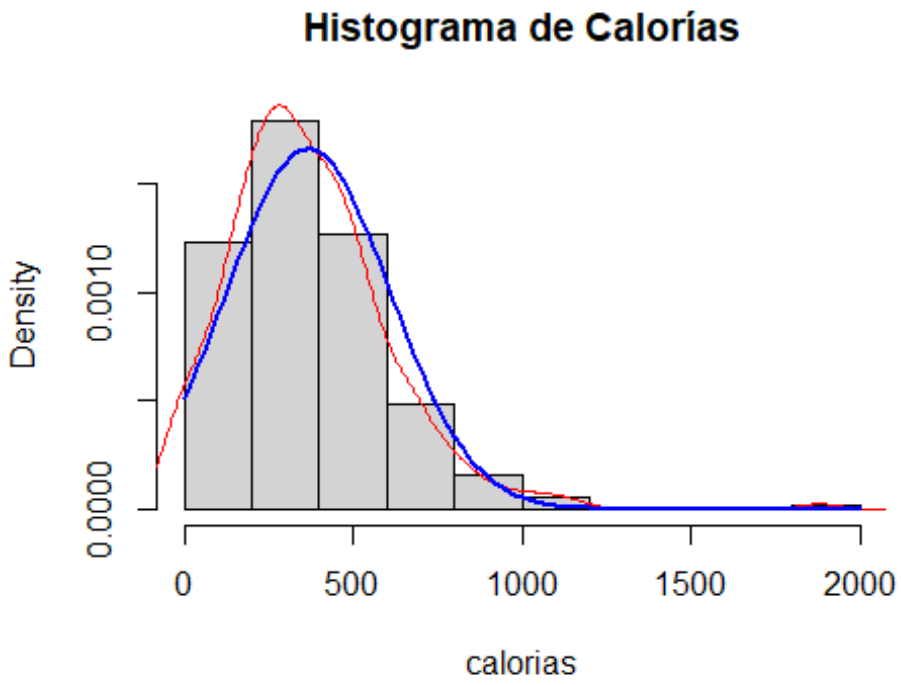
```
##    75%
## 26.875
```

Los valores de Azucar son mucho menores que los valores de calorías

Histograma y su distribución teórica de probabilidad

Calorías

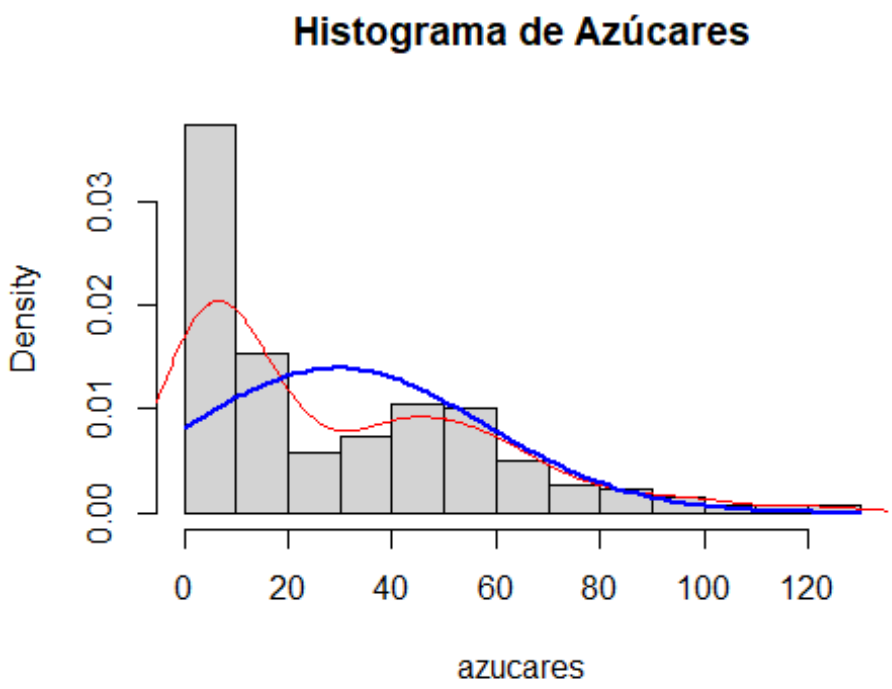
```
hist(calorias, freq=FALSE, main="Histograma de Calorías")
lines(density(calorias), col="red")
curve(dnorm(x, mean=mean(calorias), sd=sd(calorias)), add=TRUE,
col="blue", lwd=2)
```



El histograma de calorías muestra una distribución claramente asimétrica hacia la derecha. Esto significa que hay una mayor concentración de productos con valores de calorías más bajos (alrededor de 400 a 600 calorías), con algunos productos que tienen valores mucho más altos, que llegan hasta aproximadamente 2000 calorías. La curva roja representa la densidad empírica observada, mientras que la curva azul es la densidad teórica de una distribución normal ajustada a los datos La diferencia entre la curva empírica (roja) y la teórica (azul) resalta la falta de ajuste de los datos a una distribución normal. Es evidente que la distribución empírica es más concentrada en la parte izquierda con una cola larga a la derecha.

Azucares

```
hist(azucares, freq=FALSE, main="Histograma de Azúcares")
lines(density(azucares), col="red")
curve(dnorm(x, mean=mean(azucares), sd=sd(azucares)), add=TRUE,
col="blue", lwd=2)
```



El histograma de azúcares también muestra una distribución asimétrica hacia la derecha, aunque menos pronunciada que en el caso de las calorías. Hay una concentración de productos con bajos niveles de azúcar (entre 0 y 20 gramos), pero también existen algunos productos con cantidades significativamente mayores de azúcar, hasta más de 120 gramos. La curva de densidad empírica (roja) nuevamente muestra una forma que no se ajusta bien a la curva de densidad normal (azul), especialmente en las colas de la distribución.La curva teórica subestima la densidad en la parte central y sobreestima en las colas, indicando que los datos de azúcares también tienen colas más pesadas de lo que se esperaría en una distribución normal.

Identifica cómo influyen los datos atípicos en la normalidad de los datos

En el caso de las calorías, la clara desviación de la normalidad y la presencia de una cola larga hacia valores altos sugieren que la eliminación o tratamiento de datos atípicos podría ser necesario si la normalidad es importante para un análisis posterior.Mientras que para los azúcares, aunque la distribución también se desvía de la normalidad, es menos extrema, por lo que se podría optar por mantener los datos atípicos si representan características importantes del menú, o transformarlos si la normalidad es crítica para el análisis. Sin embargo para ambos casos decidí conservar los valores atípicos porque como ya mencioné anteriormente son representativos del menú.