

# **Instituto Politécnico Nacional**

## **Escuela Superior de Computo**

**Programación para la ciencia de datos.**

**Cristal Karina Galindo Durán**

**Practica 5:**  
**Estadística Descriptiva**

**Vianey Maravilla Pérez**

**3AM1**

**Unidad temática a la que corresponde la práctica.**

II. Análisis exploratorio de Datos.

**Objetivo:** Realizar scripts en Lenguaje R que permitan implementar diferentes estimadores estadísticos de centralización, posición y dispersión.

**INTRODUCCIÓN:**

La **Estadística** es el área de las matemáticas que estudia la variabilidad, colección, organización análisis, interpretación, y presentación de los datos, así como el proceso aleatorio que los genera siguiendo las distribuciones de probabilidad. La estadística está conformada por dos áreas: Estadística, Inferencial y Descriptiva.

Por su parte, la **Estadística Descriptiva** es el conjunto de técnicas y procedimientos que ayudan a describir, mostrar y resumir la información de un conjunto de datos.

En esta actividad se incluyen un conjunto de ejercicios que le permiten al discente poner en práctica conceptos sobre los diferentes estimadores estadísticos en Lenguaje R.

1. Los siguientes datos representan los pesos en gramos de contenido de 16 cajas de cereal que se seleccionaron al azar de un proceso de llenado con el propósito de verificar medidas de tendencia central, posición (cuartiles y deciles); así como, de dispersión obténgalas e interprete.

506	508	499	503	504	510	497	512
514	505	493	496	506	502	509	496

### **Consideraciones:**

Para poder emplear el programa adecuadamente necesitamos saber como sacar manualmente las medidas de tendencia central, posición, puesto que en el programa se pedirá no usar las funciones y hacerlas con un método propio para así poner a prueba nuestras habilidades.

### **Procedimiento:**

Se logró ejecutar el programa de una manera eficaz, creando función por función para que nos pudiera calcular correctamente cada una de las operaciones requeridas, así como también se les hizo la llamada a cada función en la función principal.

## Programa:

```
1 "Los siguientes datos representan los pesos en gramos de contenido de 16 cajas de cereal
2 que se seleccionaron al azul de un proceso de llenado con el propósito de verificar medidas
3 de tendencia central, posición (cuartiles y deciles); así como, de dispersión, obténgalas e interprete
4 Hecho por: Maravilla Pérez Vianey 3AM1"
5
6 #Definimos nuestro vector sin ordenar para luego entonces ordenarlos
7
8 totalcereal<-c(506, 508, 499, 503, 504, 510, 497, 512, 514, 505, 493, 496, 506, 502, 509, 496)
9
10 cereal<-sort(totalcereal)
11
12 #Comenzamos a calcular lo estipulado en la descripción
13
14 PrincipalCereal<- function()
15 {
16   mediaA()
17   mediana()
18   moda()
19   Cuartil1()
20   Cuartil2()
21   Cuartil3()
22   Var()
23   DE()
24 }
25
26 #Calculamos la Media Aritmetica
27
28 mediaA<-function()
29 {
30
31   p<- sum(totalcereal)
32   n<- length(totalcereal)
33   prom<- p/n
34   cat("\nLa media Aritmetica es:", prom)
35 }
36
37
38 #Calculamos la mediana
39
40 mediana<-function()
41 {
42
43   n<- length(totalcereal)
44   o<-sort(totalcereal)
45   ifelse((n %% 2) == 1, o[(n+1)/2], mean(o[n/2+0:1]))
46 }
47
48 cat("\nLa mediana es:", mediana())
49
50
51 #Calculamos la moda
52
53
54 moda<-function()
55 {
56   names(which.max(table(cereal)))
57 }
58
59 cat("\nLa moda es:", moda())
60
61
62 #Calculamos los cuartiles
63
64 Cuartil1<-function()
65 {
66
67   r<-length(totalcereal)
68   p<-totalcereal[(r + 1) / 4]
69   cat("\nEl primer cuartil es:", p)
70 }
71
72
73
```

```

73
74
75
76 Cuartil2<-function()
77 {
78   r<-length(totalcereal)
79   p2<-median(totalcereal)
80   cat("\nEl segundo cuartil es:", p2)
81 }
82
83
84
85 Cuartil3<-function()
86 {
87   r<-length(totalcereal)
88   p3<-totalcereal[(3/4)*(r+1)]
89   cat("\nEl tercer cuartil es:", p3)
90 }
91
92
93
94
95 #Calculamos la varianza
96
97 Var<- function()
98 {
99
100   med <- mean(totalcereal)
101   r <- length(totalcereal)
102   suma<-(r-med)^2
103   sum<- ((suma)/(r-1))
104   cat("\nLa varianza es de:", sum)
105 }
106
107
108
109 #Calculamos la Desviación Estandar
109 #Calculamos la Desviación Estandar
110
111 DE<-function()
112 {
113
114   med<- mean(totalcereal)
115   r<- length(totalcereal)
116   sqrt(sum((totalcereal - med)^2)/(r - 1))
117
118 }
119
120 cat("\nLa desviación estandar es:", DE())
121
122
123
124
125
126
127
128
129
130
131
132
133
134

```

## RESULTADOS:

```
La desviación estandar es: 6.20215
> totalcereal
[1] 506 508 499 503 504 510 497 512 514 505 493 496
[13] 506 502 509 496
> cereal
[1] 493 496 496 497 499 502 503 504 505 506 506 508
[13] 509 510 512 514
> PrincipalCereal()

La media Aritmetica es: 503.75
El primer cuartil es: 503
El segundo cuartil es: 504.5
El tercer cuartil es: 496
La varianza es de: 15860[1] 6.20215
> DE()
[1] 6.20215
> |
```

## Conclusión:

Durante el proceso de esta práctica, se logró poner a prueba los conocimientos básicos tanto en el Lenguaje R, como en Probabilidad y Estadística.

En esta práctica se presentaron diferentes dificultades, una de ellas personales, y por ese motivo no pude crear correctamente el segundo problema, los ejercicios de las prácticas requieren un poco más de tiempo, aún así envío el procedimiento del primer problema que pude hacer.