TAREA 2
ACTIVIDAD 1

1. Calcular mediana y moda

a. Mediana

Se define a la mediana (*Ma*) como el valor que ocupa la posición central de un conjunto de observaciones, debiendo estar los datos previamente ordenados en sentido ascendente (o descendente) de magnitud.

Cantidad camas	Host/Resid. (f_i)
0-19	15
20-39	32
40-59	60
60-79	47
80-99	23
100-119	10
120-139	3
TOTAL	190

$$Ma = L_i + \frac{\frac{n}{2} - F_{a_{(i-1)}}}{fi} \cdot a$$

i	fi	Fa	li	n:2	ai
0-19	15	15	0		19
20-39	32	47	20		19
40-59	60	107	40		19
60-79	47	154	60	95	19
80-99	23	177	80		19
100-119	10	187	100		19
120-139	3	190	120		19

Ma =E6+G6*(F4-D5)/C6

Procedimiento:

1. Se anexa a la tabla original una columna conteniendo la frecuencia acumulada de cada intervalo.

2. Se calcula el punto medio de la población total, siendo esta última la sumatoria de todas las frecuencias (o, lo que es lo mismo, la cantidad total de objetos), n. En el intervalo que contenga ese valor medio estará ubicada la mediana.

En este caso, n = 190, $\frac{n}{2} = 95$

3. Se ubica el intervalo correspondiente. En este caso, se trata del intervalo entre 40 y 59, ya que su frecuencia acumulada es 107 y la frecuencia acumulada del intervalo anterior es 47, ergo, el número 95 está contenido en ese rango

4. Con estos datos se procede al cálculo, según la fórmula, donde:

- *Li* : límite inferior del intervalo de interés que contiene a la mediana, "*clase mediana*" (40)
- *n/2* : semisuma de las frecuencias acumuladas, punto medio de la serie de datos (95)
- *Fa(i-1)* : frecuencia acumulada del intervalo inmediato anterior a la *clase mediana* (47)
- *fi* : frecuencia absoluta de la *clase mediana* (60)
- *a*: amplitud de la *clase mediana* (19)

Esto arroja como resultado **$M_a = 55.2$**

b. Moda

La moda o modo (*Mo*) es el dato individual que más veces se repite en la serie. *Mo* será el valor más típico, más recurrente o bien, el que reúne la mayor frecuencia absoluta entre todos los valores (categorías) individuales observados en el conjunto de datos que se analiza.

Cantidad camas	Host/Resid. (f_i)
0-19	15
20-39	32
40-59	60
60-79	47
80-99	23
100-119	10
120-139	3
TOTAL	190

$$M_o = L_i + \frac{d_1}{d_1 + d_2} . a$$

i	fi	Li		
0-19	15	0	d1 =	28
20-39	32	20	d2 =	13
40-59	60	40		
60-79	47	60	d1+d2 =	41
80-99	23	80		
100-119	10	100	d1/(d1+d2) =	0.682927
120-139	3	119	a=	19
	Mo=	=K6+M9*M10		

Procedimiento:

1. Se determina la clase modal, es decir, el intervalo en el que se registra la mayor cantidad de repeticiones (40 – 59)
2. Se calculan d_1 : diferencia entre la frecuencia absoluta de la clase modal y la del intervalo inmediato anterior (28) , y d_2 : la misma diferencia con respecto al intervalo inmediato posterior (13)

3. Se establecen el límite inferior (40) y la amplitud de la clase modal (19)
4. Se procede al cálculo según la fórmula

El cálculo arroja un valor de ***M_o = 52.97***

ACTIVIDAD 2

a. Análisis de los cinco números.

Una forma aceptada y eficaz de integrar diferentes medidas descriptivas es la que se conoce como "el resumen de los cinco números".

Se procederá a aplicarlo a la ciudad de San Pedro, según los datos de la siguiente tabla:

- **$X_{mín}$**: el mínimo
- **Q_1**: el cuartil 1
- **M_a**: la mediana
- **Q_3**: el cuartil 3
- **$X_{máx}$**: el máximo

$X_{mín}$: **229**

Q_1: **744**

M_a : **1506.50**

Q_3: **2143.5**

$X_{máx}$: **270314**

Reg.	SAN PEDRO (1)		
	Superf. Total (has.)	Explot. Acum. (%)	Sup. Total Acum. (%)
1	229	10,0	0,080
2	744	20,0	0,338
3	744	30,0	0,596
4	744	40,0	0,855
5	1354	50,0	1,325
6	1659	60,0	1,901
7	1659	70,0	2,477
8	2628	80,0	3,399
9	7879	90,0	6,126
10	270314	100,0	100,000
Total	287954		

Cálculo de mediana, cuartil 1 y cuartil 3

Para establecer la mediana, dado que los datos están ordenados y expresados en forma atómica (no agrupados en rangos) basta con identificar el valor central de la serie de datos.

Debido a que la serie está integrada por un número par de datos, se procede a obtener el promedio de los dos valores centrales. En este caso, dichos valores son:

x_5 **: 1354** y x_6 **: 1659** $\Rightarrow M_a = \frac{1354+1659}{2}$ ***Ma = 1506.50***

Cuartiles:

$Q_1 = \frac{n}{4} \Rightarrow Q_1 = 2.5$

No existe la posición 2.5. El valor estará entre el valor de la posición 2 y el de la posición 3. Dado que tanto x_2 y x_3 son 744, el primer cuartil será directamente este valor

Q_1 = 744

$Q_3 = 3\frac{n}{4} \Rightarrow Q_3 = 7.5$

Nuevamente se obtiene una posición con decimales. El tercer cuartil estará entre los valores séptimo y octavo

Para obtener el valor exacto del cuartil se puede aplicar la fórmula

$$Q = x_i + d \cdot (x_{i+1} - x_i)$$

$x_i = 1659$ (x_7)

$x_{i+1} = 2628$ (x_8)

d = 0.5 (parte decimal obtenida en el cálculo inicial del cuartil)

Q_3 = 2143.5

b. Deciles y centiles

Se pide calcular el decil 7 y el centil 23.
Para ello usaremos las fórmulas respectivas:

$$D_k = k\frac{n}{10} \quad \text{y} \quad C_k = k\frac{n}{100}$$

En este caso particular, dado que contamos con diez datos, no sería necesario aplicar la fórmula para encontrar la posición que corresponde al decil 7, dado que directamente es el dato ubicado en la séptima posición. Puede verificarse:

k = 7

$D_7 = 7\frac{10}{10} \Rightarrow D_7 = 7$

Por lo tanto, el decil 7 será **$D_7 = x_7 = 1659$**

Para el cálculo del centil 23:

$C_{23} = 23\frac{10}{100} \Rightarrow C_{23} = 2.3$

El percentil 23 se ubicará entre el segundo y tercer dato de la serie. Dado que ambos tienen el mismo valor, no es necesario realizar cálculos adicionales.

$$\mathbf{C_{23} = 744}$$