# UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

## CENTRO UNIVERSITARIO UAEM ZUMPANGO

LICENCIATURA EN TURISMO

UNIDAD DE APRENDIZAJE: ESTADISTICA

TEMA 1.5 : ESTADISTICA DESCRIPTIVA

M. EN C.  LUIS ENRIQUE KU MOO

FECHA: AGOSTO DE 2016

# UNIDAD DE APRENDIZAJE
## “Estadística”

## UNIDAD DE COMPETENCIA I:
“BASES CONCEPTUALES DE LA ESTADÍSTICA PARA EL ESTUDIO DEL TURISMO”

1. Relación entre la investigación y estadística en el contexto del turismo
2. Conceptos generales de estadística y análisis estadístico
3. Tipos de investigación: cuantitativa y cualitativa
4. Datos y variables
5. Estadística Descriptiva

# OBJETIVOS

Objetivos del área curricular o disciplinaria: Analizar y aplicar las diferentes perspectivas teórico-metodológicas de la investigación en ciencias sociales para abordar el estudio del turismo.

Objetivos de la unidad de aprendizaje: Aplicar los métodos y técnicas estadísticas para el análisis e interpretación de datos.

# INTRODUCCIÓN

La característica más importante que describe o resume un grupo de datos **es su posición.**

Los conjuntos de datos muestran una tendencia definida a agruparse o resumirse en torno a cierto punto, por lo que para cualquier conjunto particular de datos, es posible seleccionar **un valor típico para describir, representar o resumir todo el conjunto de datos.**

De acuerdo a la organización de los datos hay dos formas de estimar este tipo de medidas **ya sea para datos no agrupados o agrupados**

# MEDIA ARITMÉTICA O PROMEDIO

**La media aritmética** de un conjunto de datos es el cociente entre la suma de todos los datos y el número de estos.

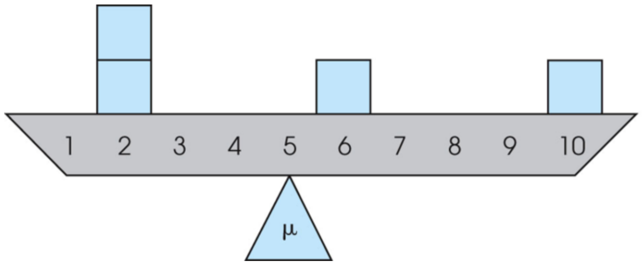
Paso 1. Se suman los datos individuales desde 1 hasta n o N

Paso 2. Se divide por el tamaño de la muestra o la población.

Fórmulas:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$ Media muestral

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$ Media poblacional



# MEDIA ARITMÉTICA

Ejemplo: Las ventas de la semana pasada fueron:

5, 6, 4, 7, 8, 4, 6

Hay 7 datos que suman 40

$$\text{Venta media} = \frac{5+6+4+7+8+4+6}{7} = \frac{40}{7} = 5.7$$

**media**


$$\frac{7+8+3+5+3+4+3+6+1+2+8}{11} =$$

$$= \frac{50}{11} = \boxed{4,5}$$

# MEDIA ARITMÉTICA: Datos agrupados

Paso 1. Definir las clases y determinar el punto medio de cada una

Paso 2. Determinar la frecuencia absoluta de cada clase

Paso 3. aplicar la siguiente fórmula.

$$\bar{x} = \frac{\sum M_i f_i}{n}$$

$M_i$= punto medio o marca de la clase

$f_i$ = frecuencia de la clase i

n = $\Sigma f_i$ = tamaño de la muestra

# MEDIA ARITMÉTICA: Ejemplo

| INTERVALO DE CLASE | MARCA DE CLASE | FRECUENCIA ABSOLUTA fi | FRECUENCIA ABSOLUTA ACUMULADA Fi | FRECUENCIA RELATIVA $F_i/n$ | FRECUENCIA RELATIVA ACUMULADA $F_I/n$ |
|---|---|---|---|---|---|
| 10 - 22 | 16 | 3 | 3 | 3/60 | 3/60 |
| 23 - 35 | 29 | 4 | 7 | 4/60 | 7/60 |
| 36 - 48 | 42 | 5 | 12 | 5/60 | 12/60 |
| 49 - 61 | 55 | 8 | 20 | 8/60 | 20/60 |
| 62 – 74 | 68 | 14 | 34 | 14/60 | 34/60 |
| 75 - 87 | 81 | 20 | 54 | 20/60 | 54/60 |
| 88 - 100 | 94 | 6 | 60 | 6/60 | 60/60 |
| | | 60 | | 60/60 | |

$$\bar{x} = \frac{\sum M_i f_i}{n} = \frac{(16 \times 3) + (29 \times 4) + (42 \times 5) + (55 \times 8) + (68 \times 14) + (81 \times 20) + (94 \times 6)}{60}$$

$$\bar{x} = \frac{3950}{60} = 65.833$$

# MEDIA PONDERADA

Esta medida surge como una forma de discriminar la importancia de cada dato. A través de un factor de ponderación se asigna el peso de cada dato dentro de la muestra o población.

La fórmula para su cálculo es:

$$\bar{X}p = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

# MEDIA PONDERADA

## Ejemplo

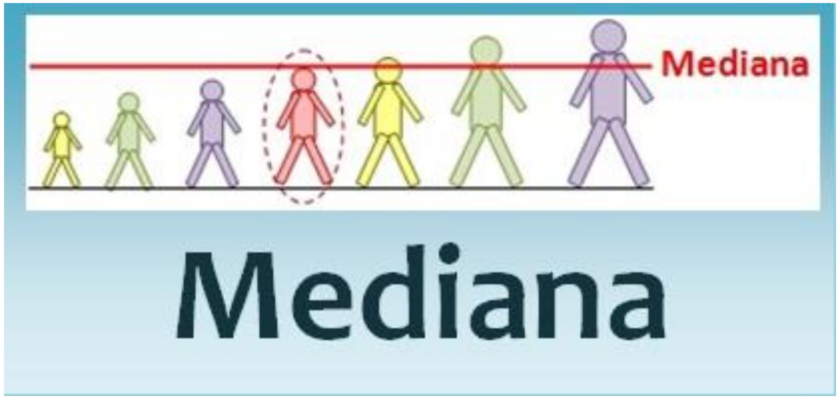
| x | w | wx |
|---|---|---|
| 9 | 0.65 | 5.85 |
| 7 | 0.20 | 1.40 |
| 8 | 0.15 | 1.20 |
| Total | 1.00 | 8.45 |

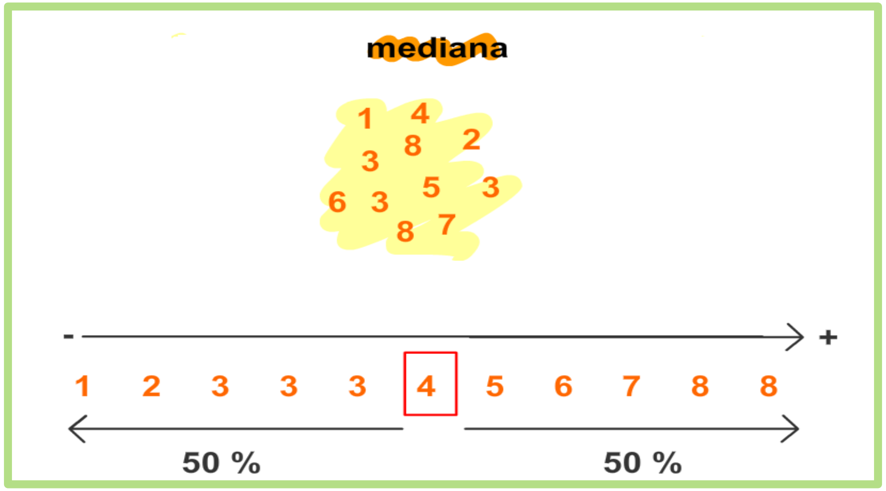
$$\bar{x} = \frac{\sum wx}{\sum w} = \frac{8.45}{1}$$

$$\bar{x} = 8.45$$

# MEDIANA

**La mediana** (me) de un conjunto de datos $x_1, x_2, \ldots, x_n$, es el valor $x_i$ que se encuentra en el punto medio o centro, cuando se ordenan los valores de menor a mayor.



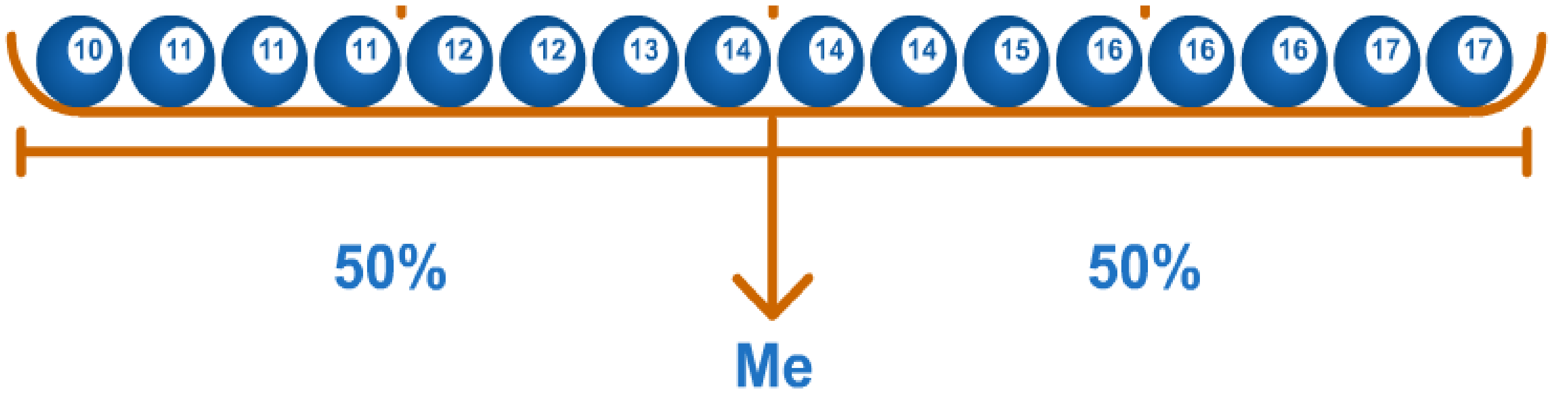
# MEDIANA

*Procedimiento de calculo:*

Paso 1.- Ordenar de menor a mayor los valores xi del conjunto de datos individuales, i = 1,2,…,n

Paso 2.- Aplicar la $\tilde{x} = \frac{n+1}{2}$ para localizar la posición.

Paso 3.- Identificar si n es impar o par sacar la media de los dos valores del centro.



# MEDIANA

Encontrar la mediana del siguiente conjunto de datos que corresponden al tiempo en segundos, requerido para marcar la compra de artículos en un supermercado que utiliza verificadores automáticos: 10, 15, 62, 53, 11, 38, 75, 112, 40, 22, 57.

Paso 1: Ordenar el conjunto de datos:

{ 10, 11, 15, 22, 38, 40, 53, 57, 62, 75, 112 }.

Paso 2: La mediana es $\tilde{x} = \frac{11+1}{2}$ = 6

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10, | 11, | 15, | 22, | 38, | 40, | 53, | 57, | 62, | 75, | 112 }. | Datos |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | Posición |

Paso 3: Identificar n = 11, impar

Mediana

Mediana2

# MEDIANA PARA DATOS AGRUPADOS

1.- Obtener la Clase de la Mediana, es decir, el primer intervalo que cumpla la condición: $F_j \geq n/2$

2.- Calcular la mediana con la siguiente ecuación:

$$\tilde{x} = Lm + \left[\frac{\frac{n}{2} - FA}{fm}\right] C$$

Donde

j = es el número del intervalo de clase que cumple la condición, j = 1, 2,…,k

$F_j$= Es la frecuencia acumulada del intervalo de clase j

Lm = Límite inferior del intervalo que corresponde a la clase mediana.

n = Total de datos

FA = Frecuencia acumulada absoluta de las clases anteriores a la clase mediana

fm = Frecuencia absoluta en la clase mediana.

C = Tamaño de clase. (amplitud o distancia del intervalo)

# MEDIANA PARA DATOS AGRUPADOS: Ejemplo

| INTERVALO DE CLASE | MARCA DE CLASE | FRECUENCIA ABSOLUTA fi | FRECUENCIA ABSOLUTA ACUMULADA Fi | FRECUENCIA RELATIVA $F_i$ /n | FRECUENCIA RELATIVA ACUMULADA $F_I$ /n |
|---|---|---|---|---|---|
| 10 - 22 | 16 | 3 | 3 | 3/60 | 3/60 |
| 23 - 35 | 29 | 4 | 7 | 4/60 | 7/60 |
| 36 - 48 | 42 | 5 | 12 | 5/60 | 12/60 |
| 49 - 61 | 55 | 8 | 20 | 8/60 | 20/60 |
| 62 – 74 | 68 | 14 | 34 | 14/60 | 34/60 |
| 75 - 87 | 81 | 20 | 54 | 20/60 | 54/60 |
| 88 - 100 | 94 | 6 | 60 | 6/60 | 60/60 |
| | | 60 | | 60/60 | |

1.- Obtener la Clase de la Mediana: n = 60 y n/2 = 30,
En el Quinto intervalo se cumple que: $F_5 = 34 \geq 30$.
El quinto intervalo se identificará como la clase mediana.

# MEDIANA PARA DATOS AGRUPADOS: Ejemplo

| No. | Intervalo | $f_i$ | $F_i$ |
|---|---|---|---|
| 4 | 49 -61 | 8 | 20 |
| 5 | 62 - 74 | 14 | 34 |

2.- Calcular la mediana:

Lm = 62 Limite inferior de la clase de la mediana

n = 60 Datos u observaciones

FA = $F_4$ = 20 Frec. acumul. absoluta inf.

fm = $f_5$ = 14 Frec. Absoluta de la clase de la mediana

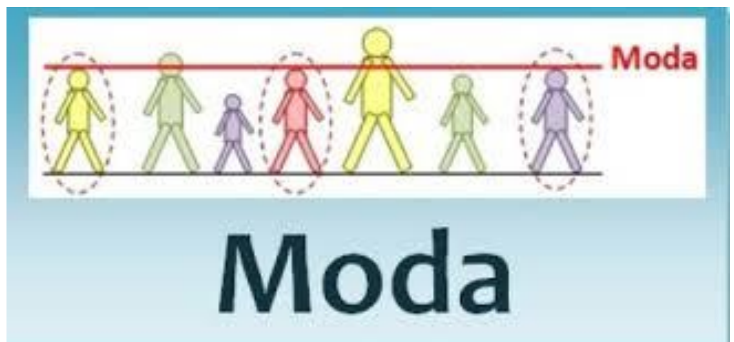
C = 13

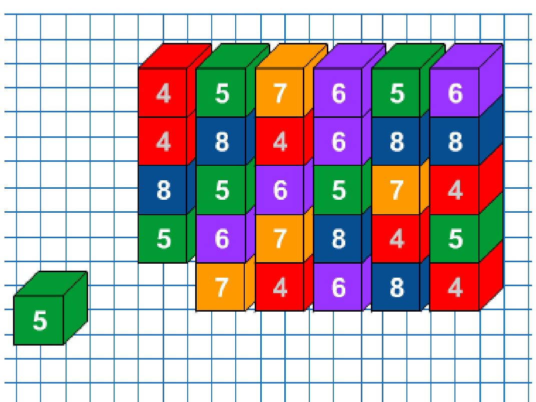
Me $= Lm + \left[\frac{\frac{n}{2} - FA}{fm}\right] C = 62 + \left[\frac{\frac{60}{2} - 20}{14}\right] (13)$ =71.286

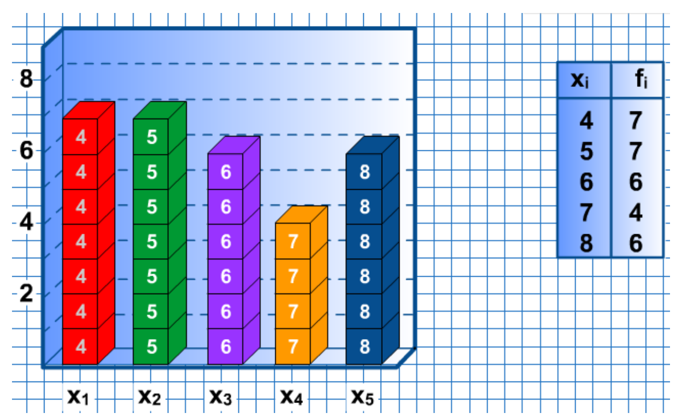
# MODA

**La moda** es la observación que se presenta con mayor frecuencia en la muestra o población.

Si los datos tienen una sola moda son unimodales, si tienen 2 son bimodales y así sucesivamente.



# MODA

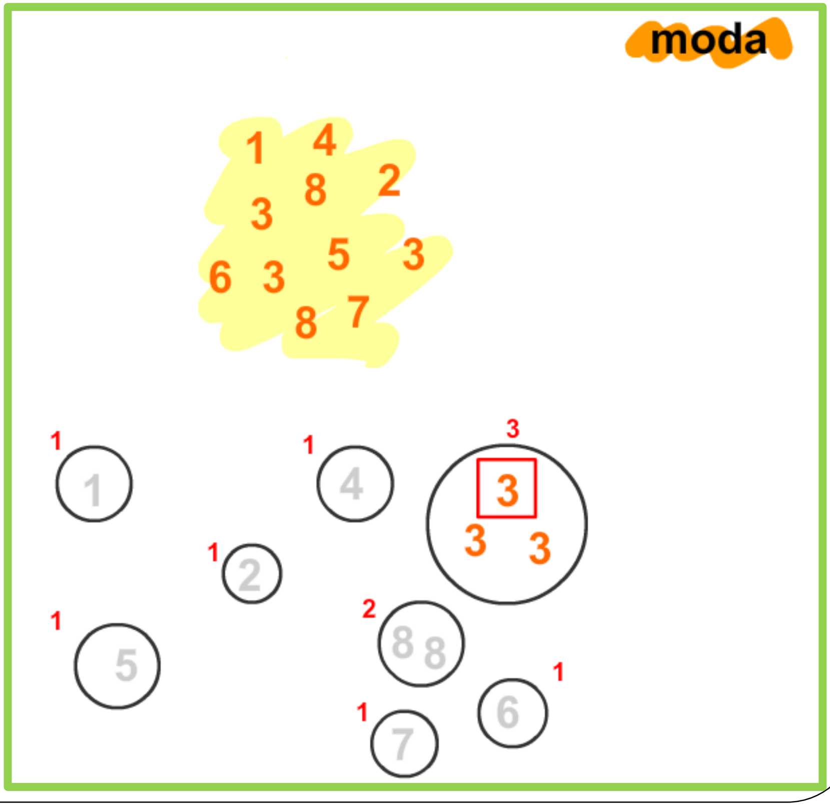
Procedimiento de calculo:

Paso 1. Ordenar los datos en forma ascendente

Paso 2. Identificar el o los datos con mayor frecuencia

Ejemplo: Sea 1, 2, 5, 1, 3, 2, 3, 7, 3, 6, 3, 4, y 3

Moda = 3



# MODA PARA DATOS AGRUPADOS

**Paso 1**. La Moda estará representada por la clase que posee la más alta frecuencia, denominándose **clase modal**.

**Paso 2**. El cálculo de la Moda se obtiene con la siguiente expresión:

$$MODA = Li + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] C$$

Donde:

$L_i$ = Límite inferior del intervalo de **la clase modal**.

$\Delta_1$ = Frecuencia de la clase modal menos la frecuencia de la clase inmediata anterior.

$\Delta_2$ = Frecuencia de la clase modal menos la frecuencia de la clase inmediata posterior.

C = Tamaño de la clase.

# MODA PARA DATOS AGRUPADOS

| No. | Intervalo | $f_i$ | $F_i$ |
|---|---|---|---|
| 5 | 62 -74 | 14 | 34 |
| 6 | 75 - 87 | 20 | 54 |
| 7 | 88 -100 | 6 | 60 |

**Clase modal**: 6, intervalo (75 – 87) ya que

$f_6$ = 20 (la mayor frecuencia absoluta)

$L_I = 75$

$\Delta_1 = f_6 - f_5 = 20 - 14 = 6$

$\Delta_2 = f_6 - f_7 = 20 - 6 = 14$

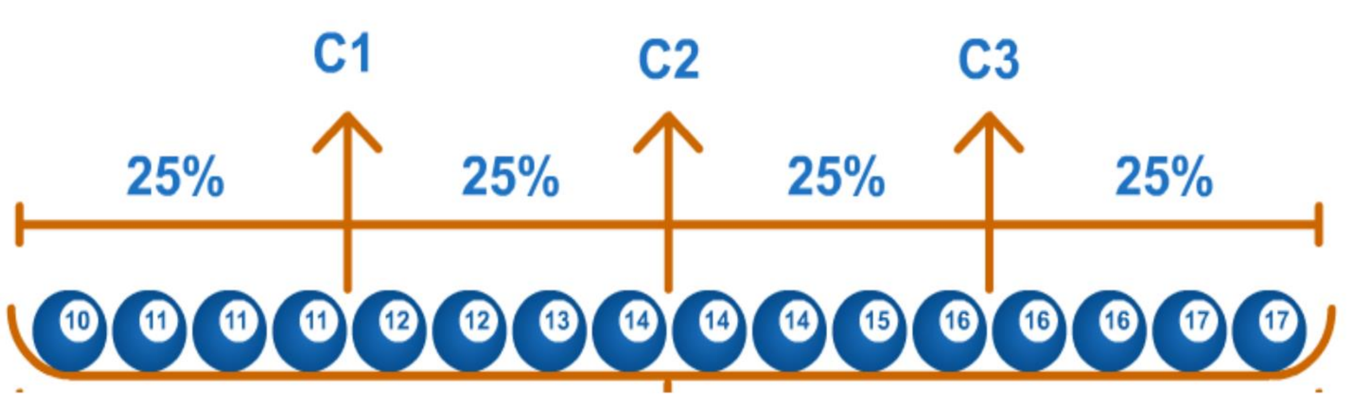
$C = 13$

$$MODA = Li + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] C = 75 + \left[\frac{6}{6 + 14}\right] 13 = 78.9$$

# OTRAS MEDIDAS DE TENDENCIA CENTRAL

**Cuartil**. Un cuartil es una medida de posición que permite determinar que valor de un grupo de datos es de tal forma que sólo cierto porcentaje del total de datos está por debajo de dicho valor.

$$Q_i = \left(\frac{n+1}{4}\right)\mathrm{i},\ \ \mathrm{donde\ i} = 1, 2\ \mathrm{ó}\ 3$$

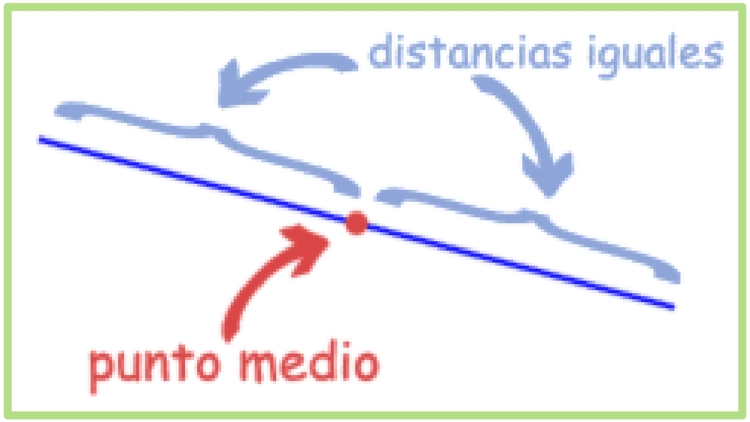



Cuartil

# OTRAS MEDIDAS DE TENDENCIA CENTRAL

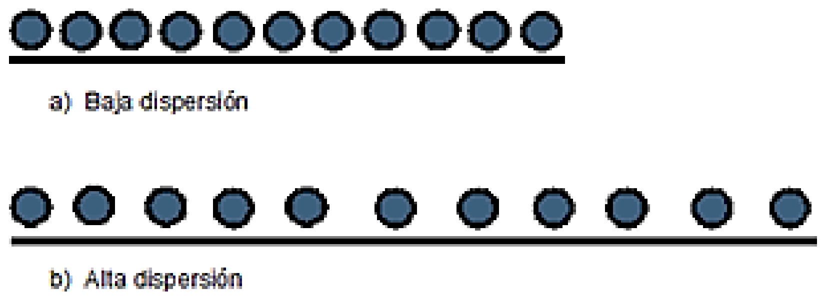
**Punto medio** o Rango Medio. Es el punto medio entre el máximo y el mínimo valor observado RM= (DM + dm) / 2.

Ejemplo: Si los datos son 3, 5, 7, 12, 9, 8. El Rango medio sería: RM = ( 12+3) / 2= 7.5



# Medidas de dispersión, variación o variabilidad.

Las medidas de dispersión sirven para expresar la cantidad de difusión o dispersión que hay entre los números, indican que tan alejados o dispersos se encuentran unos de otros.




Las medidas de dispersión son :

- Rango.
- Desviación media
- Varianza.
- Desviación Estándar.
- Coeficiente de variación.

# Medidas de dispersión: Rango.

Rango. Es la diferencia entre el máximo y el mínimo valor observado. R = (DM - dm).

Ejemplo: Si los datos son 3, 5, 7, 12, 9, 8.

R = ( 12 - 3)  = 9

**Ejercicio.** Encuentre el rango de los siguientes datos:

6.0, 5.0, 11.0, 33.0, 4.0, 5.0, 80.0, 18.0, 35.0, 17.0, 23.0, 4.0, 14.0, 11.0, 9.0, 9.0, 8.0, 4.0, 20.0, 5.0, 8.9, 21.0, 9.2, 3.0, 2.0, 0.3

# Rango para datos agrupados.

Limite superior de la última clase menos el límite inferior de la primera clase.

| INTERVALO DE CLASE | MARCA DE CLASE | FRECUENCIA ABSOLUTA fi | FRECUENCIA ABSOLUTA ACUMULADA Fi | FRECUENCIA RELATIVA $F_i/n$ | FRECUENCIA RELATIVA ACUMULADA $F_I/n$ |
|---|---|---|---|---|---|
| 10 - 22 | 16 | 3 | 3 | 3/60 | 3/60 |
| 23 - 35 | 29 | 4 | 7 | 4/60 | 7/60 |
| 36 - 48 | 42 | 5 | 12 | 5/60 | 12/60 |
| 49 - 61 | 55 | 8 | 20 | 8/60 | 20/60 |
| 62 – 74 | 68 | 14 | 34 | 14/60 | 34/60 |
| 75 - 87 | 81 | 20 | 54 | 20/60 | 54/60 |
| 88 - 100 | 94 | 6 | 60 | 6/60 | 60/60 |
| | | 60 | | 60/60 | |

Rango= 100-10 = 90

# Medidas de dispersión: Desviación media.

**Desviación media:** Media aritmética de los valores absolutos de las desviaciones respecto a la media aritmética. $dm = \frac{\sum|x_i - \bar{x}|}{n}$ o $DM = \frac{\sum|x_i - \mu|}{N}$

Ejemplo: 2, 2, 4, 4, 4, 5, 6, 7, 8, 8,

| Datos $(x_i)$ | Desviaciones $(x_i - \bar{x})$ | Valor absoluto $|(x_i - \bar{x})|$ |
|---|---|---|
| 2 | 2 – 5 = -3 | 3 |
| 2 | 2 – 5 = -3 | 3 |
| 4 | 4 – 5 = -1 | 1 |
| 4 | 4 – 5 = -1 | 1 |
| 4 | 4 – 5 = -1 | 1 |
| 5 | 5 – 5 = 0 | 0 |
| 6 | 6 – 5 = 1 | 1 |
| 7 | 7 – 5 = 2 | 2 |
| 8 | 8 – 5 = 3 | 3 |
| 8 | 8 – 5 = 3 | 3 |
| n = 10 | Suma = 0 | Suma = 18 |

$Media = 5$

$$DM = \frac{\sum|x_i - \mu|}{N} = \frac{18}{10}$$

$$DM = 1.8$$

# Medidas de dispersión: Varianza.

La varianza. Es la media aritmética de las desviaciones cuadráticas respecto a la media.

Fórmulas: $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$ o $\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$

| Datos $(x_i)$ | Desviaciones $(x_i - \bar{x})$ | Cuadrados $(x_i - \bar{x})^2$ |
|---|---|---|
| 2 | 2 – 5 = -3 | 9 |
| 2 | 2 – 5 = -3 | 9 |
| 4 | 4 – 5 = -1 | 1 |
| 4 | 4 – 5 = -1 | 1 |
| 4 | 4 – 5 = -1 | 1 |
| 5 | 5 – 5 = 0 | 0 |
| 6 | 6 – 5 = 1 | 1 |
| 7 | 7 – 5 = 2 | 4 |
| 8 | 8 – 5 = 3 | 9 |
| 8 | 8 – 5 = 3 | 9 |
| n = 10 | Suma = 0 | Suma = 44 |

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{44}{9} = 4.89$$

o

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N} = \frac{44}{10} = 4.4$$

# Varianza para datos agrupados.

La fórmula de la varianza para datos agrupados usada como estimador de la varianza poblacional es:

$$s^2 = \frac{\sum f(x-\bar{x})^2}{n-1} = \frac{\sum f x^2}{n-1} - \frac{n\bar{x}^2}{n-1}$$

$$\sigma^2 = \frac{\sum f(x_i-\mu)^2}{N} = \frac{\sum f x^2}{N} - \mu^2$$

Donde:

*f* = frecuencia de clase

*x* = Es el punto medio de la clase.

μ = media poblacional

$\bar{x}$= Media muestral

# Varianza para datos agrupados.

$\bar{x} = 65.8$

$$s^2 = \frac{\sum f x^2}{n-1} - \frac{n\bar{x}^2}{n-1}$$

$$= \frac{286124}{59} - \frac{60(65.833^2)}{59}$$

= 4849.56 - 4407.44 = 442.12

$$\sigma^2 = \frac{\sum f x^2}{N} - \mu^2$$

$$\sigma^2 = \frac{286124}{60} - 65.833^2$$

= 4768.73 - 4333.98 = 434.75

| INTERVALO DE CLASE | MARCA | FREC. | $x^2$ | $fx^2$ |
|---|---|---|---|---|
| 10 - 22 | 16 | 3 | 256 | 768 |
| 23 - 35 | 29 | 4 | 841 | 3364 |
| 36 - 48 | 42 | 5 | 1764 | 8820 |
| 49 - 61 | 55 | 8 | 3025 | 24200 |
| 62 – 74 | 68 | 14 | 4624 | 64736 |
| 75 - 87 | 81 | 20 | 6561 | 131220 |
| 88 - 100 | 94 | 6 | 8836 | 53016 |
| | | 60 | | 286124 |

# Medidas de dispersión: Desviación estándar.

**Desviación Estándar:** Es la raíz cuadrada de la varianza.

$$\sqrt{s^2} = \sqrt{\frac{\sum(x_i-\overline{x})^2}{n-1}} = \sqrt{\frac{44}{9}} = \sqrt{4.89} = 2.211$$

$$\sqrt{\sigma^2} = \sqrt{\frac{\sum(x_i-\mu)^2}{N}} = \sqrt{\frac{44}{10}} = \sqrt{4.4} = 2.098$$

# Medidas de dispersión: Coeficiente de variación.

El coeficiente de variación es la razón de la desviación estándar a la media aritmética, expresada como porcentaje:

$$CV = \frac{s}{\bar{x}}(100\%)$$

Ejemplo:

$$CV = \frac{s}{\bar{x}}(100\%) = \frac{2.21}{5}(100\%) = \frac{221.1}{5}\% = 44.22\%$$

$$CV = \frac{\sigma}{\mu}(100\%) = \frac{2.098}{5}(100\%) = \frac{209.8}{5}\% = 41.96\ \%$$

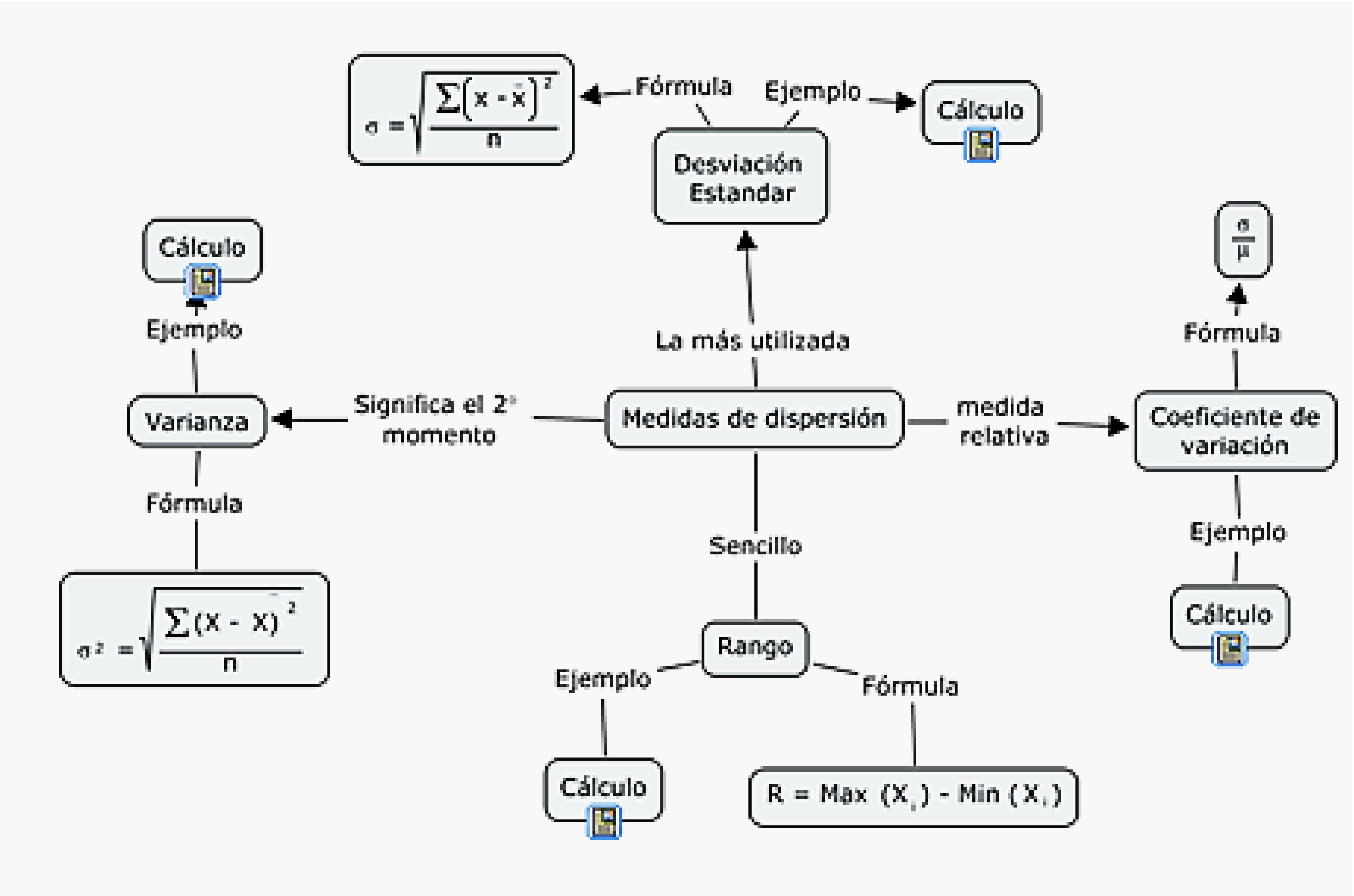
# Hacer un Resumen de medidas de Variación

Mencionar las medidas de dispersión

Conceptos de las medidas

Formulas.

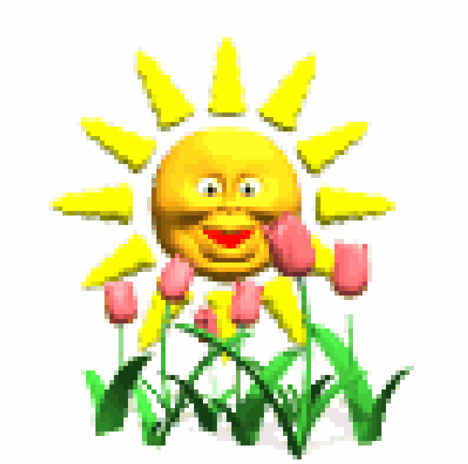
# Ejemplo de mapa conceptual



# Ejercicio sobre medidas de Variación

Porcentaje de ocupación promedio diaria 2007-2012

| Años | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|
| Lunes | 61.74 | 61.96 | 48.93 | 57.07 | 61.33 | 55.97 |
| Martes | 70.18 | 68.86 | 56.22 | 65.84 | 70.22 | 66.27 |
| Miércoles | 73.66 | 72.84 | 58.56 | 70.04 | 75.30 | 71.20 |
| Jueves | 69.70 | 68.37 | 54.59 | 64.65 | 71.13 | 61.80 |
| Viernes | 50.19 | 50.93 | 40.01 | 49.15 | 55.15 | 46.96 |
| Sábado | 43.38 | 44.70 | 35.12 | 41.99 | 49.20 | 41.14 |
| Domingo | 45.35 | 45.08 | 35.60 | 42.64 | 46.77 | 42.87 |



FIN DE LA PRESENTACION