PROBABILIDAD Y ESTADÍSTICA

ING. LAURA SUAREZ

PROGRAMA ANALÍTICO

1. Estadística descriptiva
2. Probabilidad
3. Variables aleatorias. Distribuciones de probabilidad
4. Inferencia estadística. Estimación de parámetros puntual
5. Estimación por intervalos de confianza
6. Pruebas de Hipótesis
7. Control estadístico de la calidad
8. Introducción al análisis de regresión
9. Análisis de varianza

VIAS DE COMUNICACIÓN

- Autogestión académica: calificaciones y condición académica.
- Aula virtual: clases asincrónicas, material de estudio, trabajos prácticos.
- suarezlaura296@gmail.com

CONCEPTOS

- Estadística: es la ciencia de recolectar, organizar, presentar, analizar e interpretar datos para ayudar en una toma de decisiones más efectiva.

- Probabilidad: busca datos numéricos más o menos objetivos para evaluar la posibilidad de un evento futuro, es decir, su grado de certidumbre o verisimilitud.

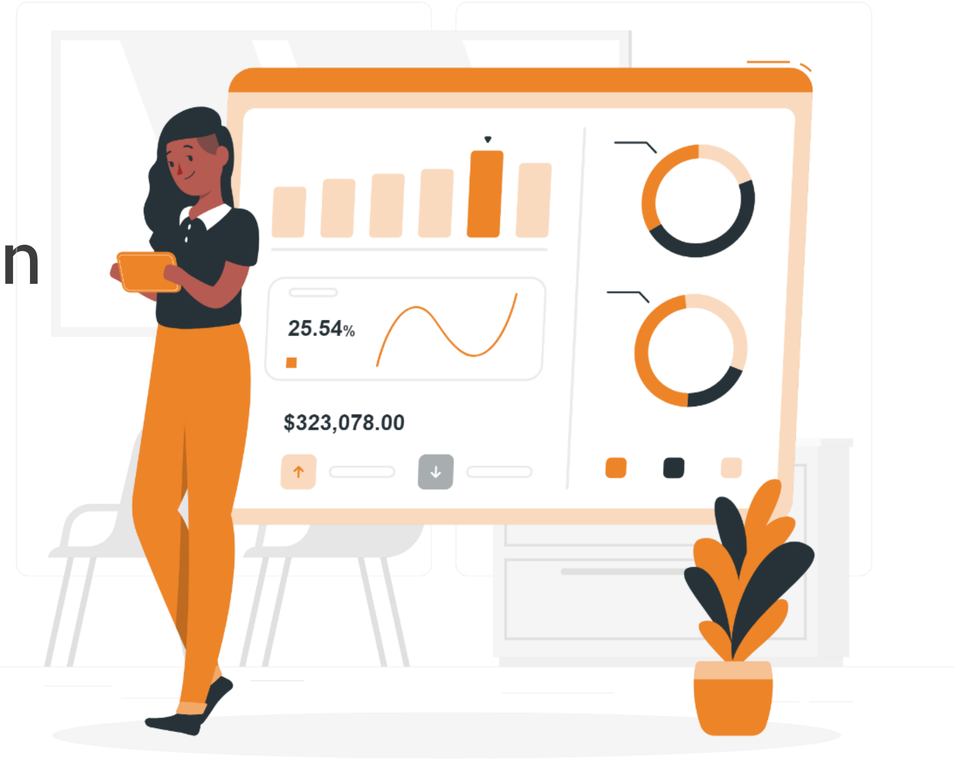
DATOS

- Datos: es toda información susceptible de ser sometida al análisis estadístico, es decir, que debe poder ser comparada, analizada e interpretada, razón por la cual se dice que los datos estadísticos presentan relaciones significativas.

- Fuentes de datos:
  1. Datos primarios: son los que recopila el propio investigador para resolver un problema particular.
  2. Datos secundarios: son aquellos que han sido reunidos y publicados por otras instituciones y se encuentran disponibles.

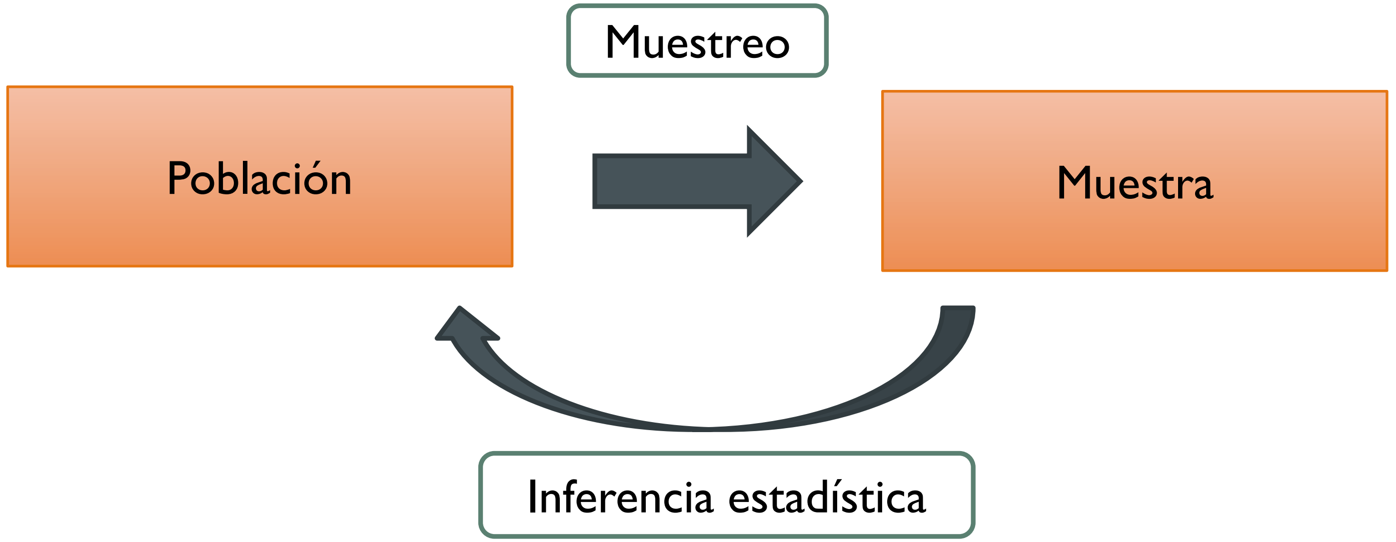
MÉTODO ESTADÍSTICO

- Formulación del problema
- Diseño del experimento o investigación
- Recolección de datos
- Procesamiento y análisis de datos
- Generalización e inferencia final



ESTADÍSTICA

- Estadística Descriptiva: se corresponde con la metodología utilizada para describir un conjunto de datos a través de métodos numéricos o gráficos. Por ejemplo, la evolución de las ventas en una empresa en un cierto periodo de tiempo
- Inferencia Estadística: incluye métodos que permiten hacer generalizaciones con respecto a la población en base a la información proporcionada por una muestra aleatoria, con un grado de incertidumbre cuantificable.



POBLACIÓN Y MUESTRA

- Población: es un conjunto completo de elementos o individuos de interés sobre los cuales se quiere realizar un estudio.

- Muestra: es una parte de la población a estudiar que sirve para representarla.

RAZONES PARA EL MUESTREO

- Costo
- Tiempo
- Pruebas destructivas

RAZONES PARA TRABAJAR CON LA POBLACIÓN

- Exactitud de los resultados
- Contar (censo)
- Sensibilidad de los resultados

DATOS

DATOS

- Datos: es toda información susceptible de ser sometida al análisis estadístico, es decir, que debe poder ser comparada, analizada e interpretada, razón por la cual se dice que los datos estadísticos presentan relaciones significativas.

- Fuentes de datos:

  1. Datos primarios: son los que recopila el propio investigador para resolver un problema particular. Métodos más conocidos para recolectar datos primarios: investigación mediante encuesta, experimental y observacional.

  2. Datos secundarios: son aquellos que han sido reunidos y publicados por otras instituciones y se encuentran disponibles. En general provenientes de organismos públicos tales como INDEC, BCRA, AFIP.

CLASIFICACIÓN SEGÚN SUS CARACTERÍSTICAS O SEGÚN LA VARIABLE

- Cualitativos: la característica bajo estudio es un atributo o categoría

- Cuantitativos: la característica bajo estudio es numérica
  - Discretos: surge de un proceso de conteo
  - Continuos: surge de un proceso de medición

CLASIFICACIÓN DE DATOS SEGÚN SU PRESENTACIÓN

- Simples: están presentados individualmente

  n = cantidad de datos= tamaño de la muestra

  N = cantidad de datos = tamaño de la población

22	32	15	25
31	25	17	25
14	25	17	21
19			

- Repetidos: están presentados en una tabla según las veces que se repiten cada uno

X	f
2	21
4	39
6	12
8	11
10	17

- Agrupados: están presentados en una tabla según a que clase pertenecen

Clase	f
de 2 a 4	15
de 4 a 6	12
de 6 a 8	19
de 8 a 10	4

DATOS AGRUPADOS

- Cuando los límites reales coinciden con los nominales

Clase	f	LRI	LRS	MC	c
de 2 a 4	15	2	4	3	2
de 4 a 6	12	4	6	5	2
de 6 a 8	19	6	8	7	2
de 8 a 10	4	8	10	9	2

$$Marca\ de\ clase = MC = \frac{(LRI + LRS)}{2}$$

$$c = ancho\ de\ clase = LRS - LRI$$

DATOS AGRUPADOS

- Cuando los límites reales NO coinciden con los nominales

$$Límite\ real\ inferior = LRI = \frac{Límite\ nominal\ inferior\ f_i + Límite\ nominal\ superior\ f_{i-1}}{2}$$

$$Límite\ real\ superior = LRS = \frac{Límite\ nominal\ inferior\ f_{i+1} + Límite\ nominal\ superior\ f_i}{2}$$

$$c = ancho\ de\ clase = LRS - LRI$$

Clase	f	LRI	LRS	c	MC
de 2,1 a 4	15	2,05	4,05	2	3,05
de 4,1 a 6	12	4,05	6,05	2	5,05
de 6,1 a 8	19	6,05	8,05	2	7,05
de 8,1 a 10	4	8,05	10,05	2	9,05

$$LRI = LRS_{fi} - c_{fi}$$

$$LRS = LRI_{fi} + c_{fi}$$

DISTRIBUCIONES DE FRECUENCIA

- Cuando trabajamos con datos repetidos o agrupados

Frecuencia absoluta

$$n = cantidad\ de\ datos = tamaño\ de\ la\ muestra = \sum f_i$$

$$N = cantidad\ de\ datos = tamaño\ de\ la\ población = \sum f_i$$

X	fi
2	21
4	39
6	12
8	11
10	17
n	100

DISTRIBUCIONES DE FRECUENCIA

Frecuencia acumulada

$$F_i = f_i + F_{i-1}$$

X	fi	Fi
2	21	21
4	39	60
6	12	72
8	11	83
10	17	100
n	100	

DISTRIBUCIONES DE FRECUENCIA

Frecuencia relativa

$$fr_i = \frac{f_i}{n}$$

X	fi	Fi	fri
2	21	21	0,21
4	39	60	0,39
6	12	72	0,12
8	11	83	0,11
10	17	100	0,17
n	100		1

DISTRIBUCIONES DE FRECUENCIA

Frecuencia acumulada relativa

$$fr_i = \frac{F_i}{n}$$

$$Fr_i = fr_i + Fr_{i-1}$$

X	fi	Fi	fri	Fri
2	21	21	0,21	0,21
4	39	60	0,39	0,60
6	12	72	0,12	0,72
8	11	83	0,11	0,83
10	17	100	0,17	1,00
n	100		1	

DISTRIBUCIONES DE FRECUENCIA

Frecuencia relativa porcentual

$$f\%_i = fr_i * 100$$

X	fi	Fi	fri	Fri	f%i
2	21	21	0,21	0,21	21
4	39	60	0,39	0,60	39
6	12	72	0,12	0,72	12
8	11	83	0,11	0,83	11
10	17	100	0,17	1,00	17
n	100		1		100

DISTRIBUCIONES DE FRECUENCIA

Frecuencia acumulada relativa porcentual

$$F\%_i = Fr_i * 100$$

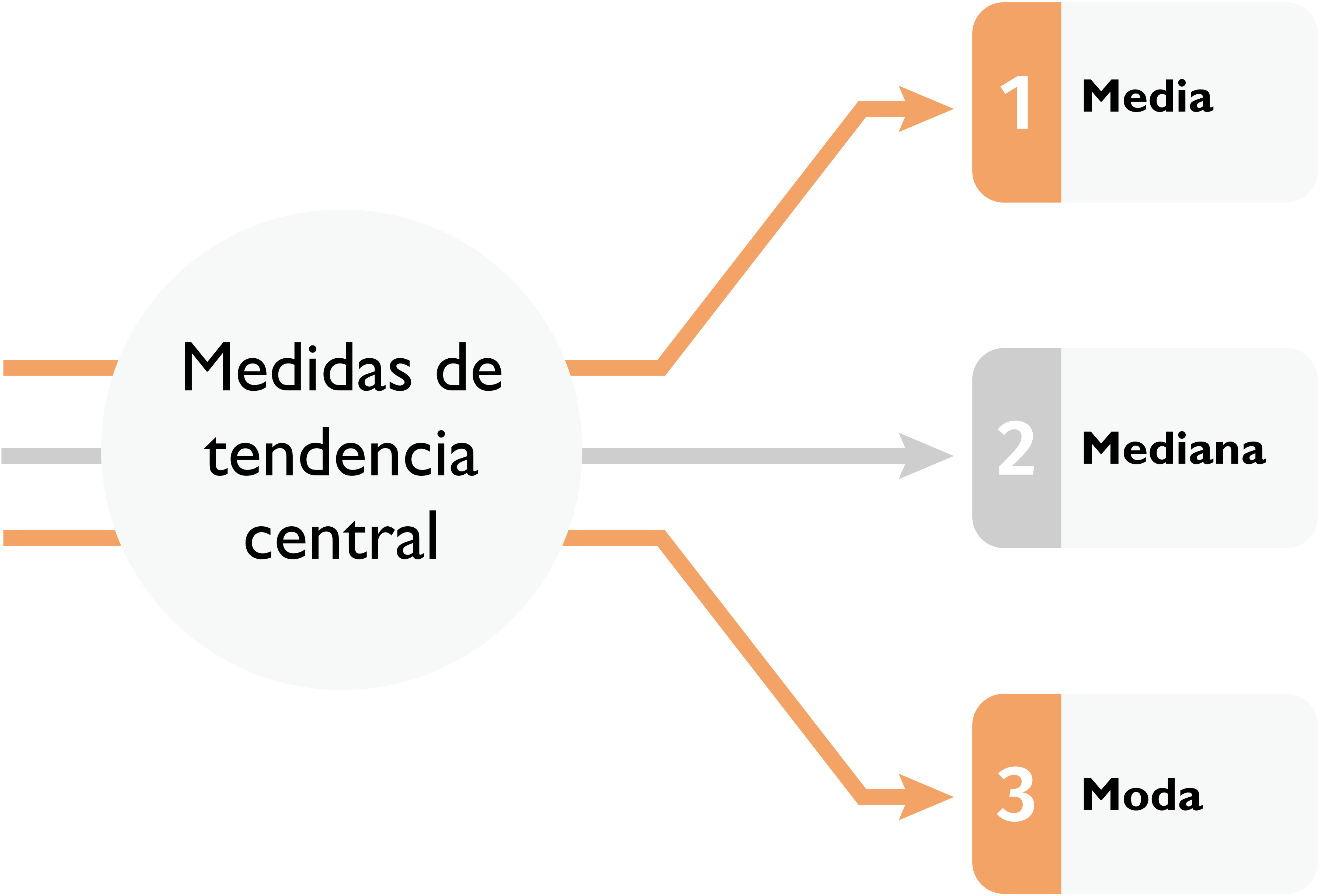
$$F\%_i = f\%_i + F\%_{i-1}$$

X	fi	Fi	fri	Fri	f%i	F%i
2	21	21	0,21	0,21	21	21
4	39	60	0,39	0,60	39	60
6	12	72	0,12	0,72	12	72
8	11	83	0,11	0,83	11	83
10	17	100	0,17	1,00	17	100
n	100		1		100	

MEDIDAS DESCRIPTIVAS

MEDIDAS DESCRIPTIVAS

- Medidas de tendencia central o posición
- Dispersión
- Forma


Medidas de tendencia central
1
Media
2
Mediana
3
Moda

MEDIA

- Media aritmética

Es el promedio aritmético de las observaciones

$$Media\ aritmetica = \frac{Suma\ de\ los\ datos}{Número\ de\ los\ datos}$$

$$\bar{X} = muestra$$

$$\mu = Población$$

FÓRMULAS MEDIA ARITMÉTICA

- Datos simples

$$\bar{X} = \frac{\sum x_i}{n}$$

- Datos repetidos

$$\bar{X} = \frac{\sum x_i . f_i}{n}$$

- Datos agrupados

$$\bar{X} = \frac{\sum MC_i . f_i}{n}$$

PROPIEDADES DE LA MEDIA

- La suma de las desviaciones respecto a la media aritmética, es igual a cero

$$\sum (x_i - \bar{x}) = 0$$

$$\sum (x_i - \bar{x}).f_i = 0$$

$$\sum (MC_i - \bar{x}).f_i = 0$$

PROPIEDADES DE LA MEDIA

- La suma de los cuadrados de las desviaciones es mínima cuando las desviaciones son obtenidas respecto de la media

$$\sum (x_i - \bar{x})^2 = Min$$

$$\sum (x_i - \bar{x})^2 . f_i = Min$$

$$\sum (MC_i - \bar{x})^2 . f_i = Min$$

PROPIEDADES DE LA MEDIA

- La media aritmética de una constante es igual a dicha constante

$$\bar{X}_{(k)} = k$$

- La media aritmética del producto de una constante por una variable es igual a la constante por la media de la variable

$$\bar{X}_{(y.k)} = k \,.\, \bar{X}_{(y)}$$

PROPIEDADES DE LA MEDIA

- La media aritmética de la suma de una variable mas una constante es igual a la media de la variable más la constante.

$$\bar{X}_{(y+k)} = k + \bar{X}_{(y)}$$

MEDIANA

Es el valor central de la variable

Cuando los datos están ordenados ya sea de menor a mayor o de mayor a menor es el valor de la variable que supera a no más de 50% de las observaciones y es superado por no más del 50% de las observaciones.

$$\tilde{X} = Muestra$$
$$Me = Población$$

17 19 $\boxed{21}$ 22 25

FÓRMULAS MEDIANA

- Datos simples
- Datos repetidos
- Datos agrupados

Primero se busca la posición del dato, luego el resultado de la mediana

n

IMPAR La posición es (n+1)/2

$$17 \quad 19 \quad \boxed{21} \quad 22 \quad 25 \qquad \tilde{X} = 21$$

PAR Hay dos posiciones: n/2 y n/2 + 1

Se saca el promedio de los valores

$$17 \quad 19 \quad \boxed{21} \quad \boxed{22} \quad 25 \quad 27 \qquad \tilde{X} = 21{,}5$$

FÓRMULAS MEDIANA

- Datos repetidos

Podemos usar la frecuencia acumulada para encontrar más fácil la posición

- Datos agrupados

Con la posición encontramos la clase acumulada, luego calculamos el valor de la mediana:

$$\tilde{X} = LRI + \frac{\frac{n}{2} - F_{i-1}}{f_i} . c$$

MODA

Es el valor mas se repite

Es el único estadístico que se puede utilizar con datos cualitativos

$\hat{x} = Muestra$

$Mo = Población$

14	15	17	17
19	21	22	25
25	25	25	31
32	32	38	40

FÓRMULAS MODA

- Datos simples — El dato que más se repite
- Datos repetidos — La frecuencia más alta
- Datos agrupados — Con la frecuencia más alta obtengo la clase modal. Luego cálculo la moda:

$$\hat{X} = LRI + \frac{d_1}{d_1 + d_2} \, .c \qquad \begin{aligned} d_1 &= f_i - f_{i-1} \\ d_2 &= f_i - f_{i+1} \end{aligned}$$

MEDIDAS DE DISPERSIÓN


Medidas de dispersión
1 Rango
2 Desviación media
3 Varianza
4 Desviación estándar

RANGO

- Rango o recorrido

Mide la amplitud de una distribución de frecuencias restándole al valor mayor el valor menor.

$$Rango = Máximo - Mínimo$$

EJEMPLOS

14 15 17 17 19 21 22 25 25 25 25 31 32

$$Rango = 32 - 14 = 18$$

X	f
2	21
4	39
6	12
8	11
10	17

$$Rango = 10 - 2 = 8$$

Clase	f	f acum	LRI	LRS	MC
de 2 a 4	15	15	2	4	3
de 4 a 6	12	27	4	6	5
de 6 a 8	19	46	6	8	7
de 8 a 10	4	50	8	10	9

$$Rango = 9 - 3 = 6$$

DESVIACIÓN MEDIA

Es la media aritmética de los valores absolutos de las desviaciones de los valores de la variable con respecto de la media

DESVIACIÓN MEDIA

- Datos simples

$$DM = \frac{\sum|x_i - \bar{x}|}{n}$$

- Datos repetidos

$$DM = \frac{\sum|x_i - \bar{x}| . f_i}{n}$$

- Datos agrupados

$$DM = \frac{\sum|MC_i - \bar{x}| . f_i}{n}$$

EJEMPLO

x	fi	$xi\,.\,fi$	$\lvert xi - \bar{x}\rvert$	$\lvert xi - \bar{x}\rvert\,.\,Fi$
2	21	42	3,28	68,88
4	39	156	1,28	49,92
6	12	72	0,72	8,64
8	11	88	2,72	29,92
10	17	170	4,72	80,24
	100	528		237,6

$\bar{x} = 5{,}28$ $\qquad DM = 2{,}376$

VARIANZA

Es la media aritmética de los cuadrados de las desviaciones de los valores de la variable con respecto de la media de la distribución

$$S^2 = Muestra$$
$$\sigma^2 = Población$$

VARIANZA

- Datos simples

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

- Datos repetidos

$$S^2 = \frac{\sum(x_i - \bar{x})^2 \, . \, f_i}{n-1}$$

- Datos agrupados

$$S^2 = \frac{\sum(MC_i - \bar{x})^2 \, . \, f_i}{n-1}$$

EJEMPLO

x	fi	$xi\,.\,fi$	$\lvert xi - \bar{x} \rvert$	$\lvert xi - x \rvert\,.\,Fi$	$(xi - \overline{x})^2$	$(xi - \overline{x})^2\,.\,fi$
2	21	42	3,28	68,88	10,7584	225,9264
4	39	156	1,28	49,92	1,6384	63,8976
6	12	72	0,72	8,64	0,5184	6,2208
8	11	88	2,72	29,92	7,3984	81,3824
10	17	170	4,72	80,24	22,2784	378,7328
	100	528		237,6		756,16

$\bar{x} = 5{,}28$ $\quad$ $DM = 2{,}376$ $\quad$ $S^2 = 7{,}638$

PROPIEDADES DE LA VARIANZA

- La varianza siempre es una cantidad no negativa $V(x) \geq 0$
- La varianza de una constante es cero $V(k) = 0$
- La varianza del producto de una constante por una variable es igual al cuadrado de la constante por la varianza de la variable $V(kx) = k^2 \,.V(x)$
- La varianza de la suma de una variable más una constante es igual a la varianza de la variable $V(x + k) = V(x)$

DESVIACIÓN ESTÁNDAR

La desviación típica o estándar es la raíz cuadrada con signo positivo de la varianza

$$S = Muestra$$

$$\sigma = Población$$

DESVIACIÓN ESTÁNDAR

- Datos simples

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}}$$

- Datos repetidos

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2 \, . \, f_i}{n - 1}}$$

- Datos agrupados

$$S = \sqrt{\frac{\sum(MC_i - \bar{x})^2 \, . \, f_i}{n - 1}}$$

EJEMPLO

x	fi	$xi\,.\,fi$	$\lvert xi - \bar{x}\rvert$	$\lvert xi - x\rvert\,.\,Fi$	$(xi - x)^2$	$(xi - x)^2\,.\,fi$
2	21	42	3,28	68,88	10,7584	225,9264
4	39	156	1,28	49,92	1,6384	63,8976
6	12	72	0,72	8,64	0,5184	6,2208
8	11	88	2,72	29,92	7,3984	81,3824
10	17	170	4,72	80,24	22,2784	378,7328
	100	528		237,6		756,16

$\bar{x} = 5{,}28$ $DM = 2{,}376$ $S^2 = 7{,}638$ $S = 2{,}764$

PROPIEDADES DE LA DESVIACIÓN ESTÁNDAR

- La desviación estándar siempre es un valor positivo $S(x) \geq 0$
- La desviación estándar de una constante es cero $S(k) = 0$
- Si todos los valores de la variable se multiplican por una misma constante, la desviación estándar queda multiplicada por el valor absoluto de dicha constante $S(kx) = k\,.\,S(x)$
- Si a todos los valores de una variable se le suma una misma constante la desviación estándar no varía $S(x+k) = S(x)$

Medidas descriptivas

- Medidas de tendencia central
  - Media — $\bar{X} = Muestra$; $\mu = Población$
  - Mediana — $\tilde{X} = Muestra$; $Me = Población$
  - Moda — $\hat{x} = Muestra$; $Mo = Población$
- Dispersión
  - Rango
  - Desviación media
  - Varianza — $S^2 = Muestra$; $\sigma^2 = Población$
  - Desviación estándar — $S = Muestra$; $\sigma = Población$
- Forma

MEDIDAS DE FORMA

MEDIDAS DE FORMA

- Asimetría
- Kurtosis

MOMENTOS

- Son generadores de fórmulas.
- Lo necesitamos para definir medidas de Asimetría y Kurtosis
- Los momentos de orden par serán positivos y los impares positivos o negativos

MOMENTOS

1. **Momentos de una variable**

Se define como momento de orden "r" de una variable "x" a:

$$\bar{x}^r = \frac{\sum x_i^r}{n}$$

Casos particulares:

$$\bar{x}^0 = 1 \;\; ; \;\; \bar{x}^1 = \bar{x} = media$$

MOMENTOS

2. Momento respecto de un valor

Se define como momento de orden "r" de una variable "x" respecto de cierto valor k:

$$m_r = \frac{\sum (x_i - k)^r}{n}$$

Lo más frecuente es que k = media

Casos particulares:

$$m_0 = 1 \quad ; \quad m_1 = 0 \quad ; \quad m_2 = s^2$$

EJEMPLO: MOMENTO 5 RESPECTO DE LA MEDIA

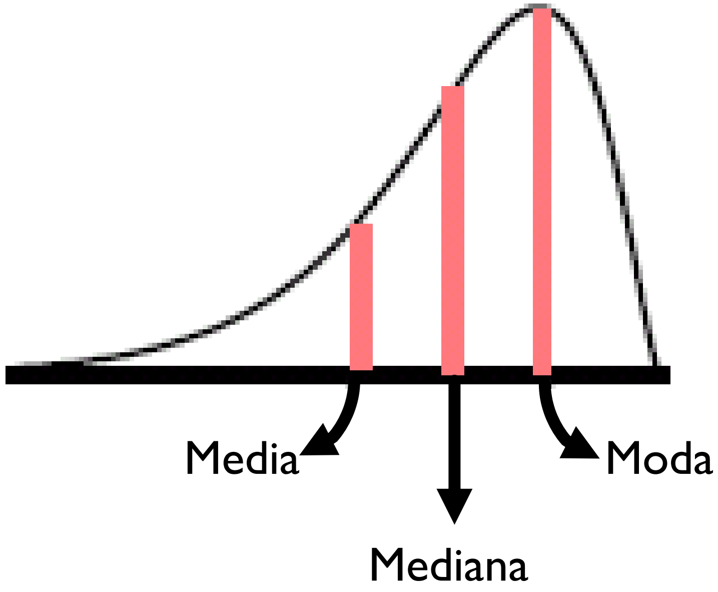
X	fi	$xi\,.\,fi$	$(xi\ -\bar{x})^2$	$(xi\ -\bar{x})^2.fi$	$(xi\ -\bar{x})^5$	$(xi\ -\bar{x})^5.fi$
2	21	42	10,7584	225,9264	-379,637599	-7972,38959
4	39	156	1,6384	63,8976	-3,43597384	-134,00298
6	12	72	0,5184	6,2208	0,19349176	2,32190116
8	11	88	7,3984	81,3824	148,882797	1637,71077
10	17	170	22,2784	378,7328	2342,66394	39825,287
	100	528		756,16		33358,9271

$$\bar{x} = 5{,}28$$

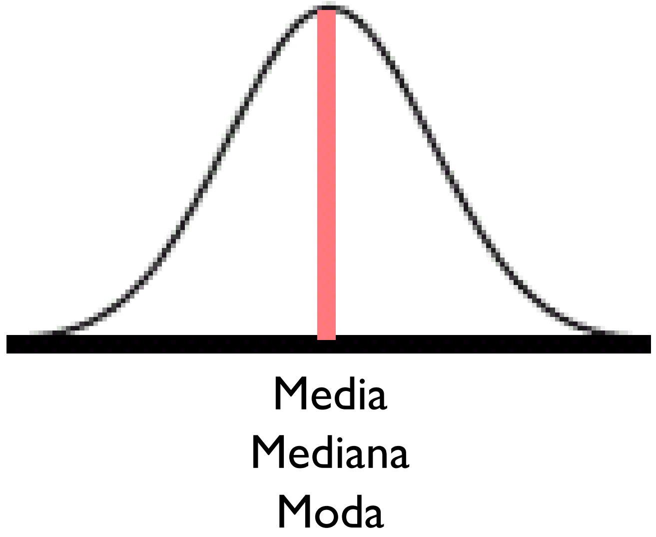
$$m_5 = \frac{\sum(xi\ -5{,}28)^5.fi}{100} = 333{,}59$$

ASIMETRÍA

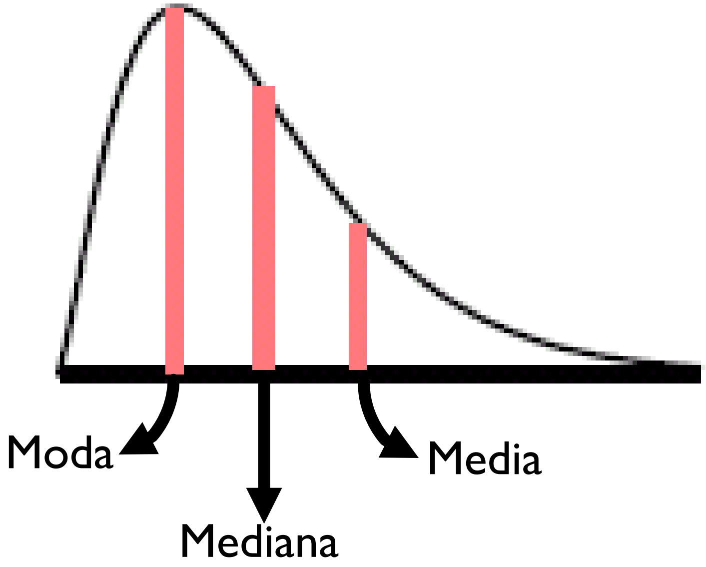
- También llamado sesgo indica si la curva formada por los valores de la serie presenta la misma forma a la izquierda y derecha de un valor central (media aritmética)



Asimetría negativa



Curva simétrica



Asimetría positiva

ASIMETRÍA

Coeficientes de Pearson

Se basan en relaciones entre las medidas de centralización

$$S_{k1} = \frac{\bar{x} - \hat{x}}{s} \qquad S_{k2} = \frac{3.(\bar{x} - \tilde{x})}{s}$$

- Si $S_k > 0$, la distribución es asimétrica positiva o a la derecha
- Si $S_k < 0$, la distribución es asimétrica negativa o a la izquierda
- Si $S_k = 0$, la distribución es simétrica (la media, la median y la moda son iguales)

EJEMPLO

X	fi	$xi\,.\,fi$	$(xi\ -\bar{x})^2$	$(xi\ -\bar{x})^2.\,fi$	Fi
2	21	42	10,7584	225,9264	21
4	39	156	1,6384	63,8976	60
6	12	72	0,5184	6,2208	72
8	11	88	7,3984	81,3824	83
10	17	170	22,2784	378,7328	100
	100	528		756,16	

$\bar{x} = 5{,}28$

$\hat{x} = 4$

$S^2 = 7{,}638$

$S = 2{,}764$

$\tilde{x} = 4$

$$S_{k1} = 0{,}46$$

$$S_{k2} = 1{,}39$$

ASIMETRÍA

Coeficiente de asimetría de Fisher

$$a_3 = \frac{m_3}{s^3} \qquad a_3 = \frac{\frac{\sum(x_i - \bar{x})^3}{n}}{s^3}$$

- Si $a_3 > 0$, la distribución es asimétrica positiva o a la derecha
- Si $a_3 < 0$, la distribución es asimétrica negativa o a la izquierda
- Si $a_3 = 0$, la distribución es simétrica (la media, la mediana y la moda son iguales)

EJEMPLO

X	fi	$xi \, . \, fi$	$(xi - \bar{x})^2$	$(xi - \bar{x})^2 . fi$	$(xi - \bar{x})^3$	$(xi - \bar{x})^3 . fi$
2	21	42	10,76	225,93	-35,29	-741,04
4	39	156	1,64	63,89	-2,09	-81,79
6	12	72	0,52	6,22	0,37	4,48
8	11	88	7,39	81,38	20,12	221,36
10	17	170	22,28	378,73	105,15	1787,62
	100	528		756,16		1190,63

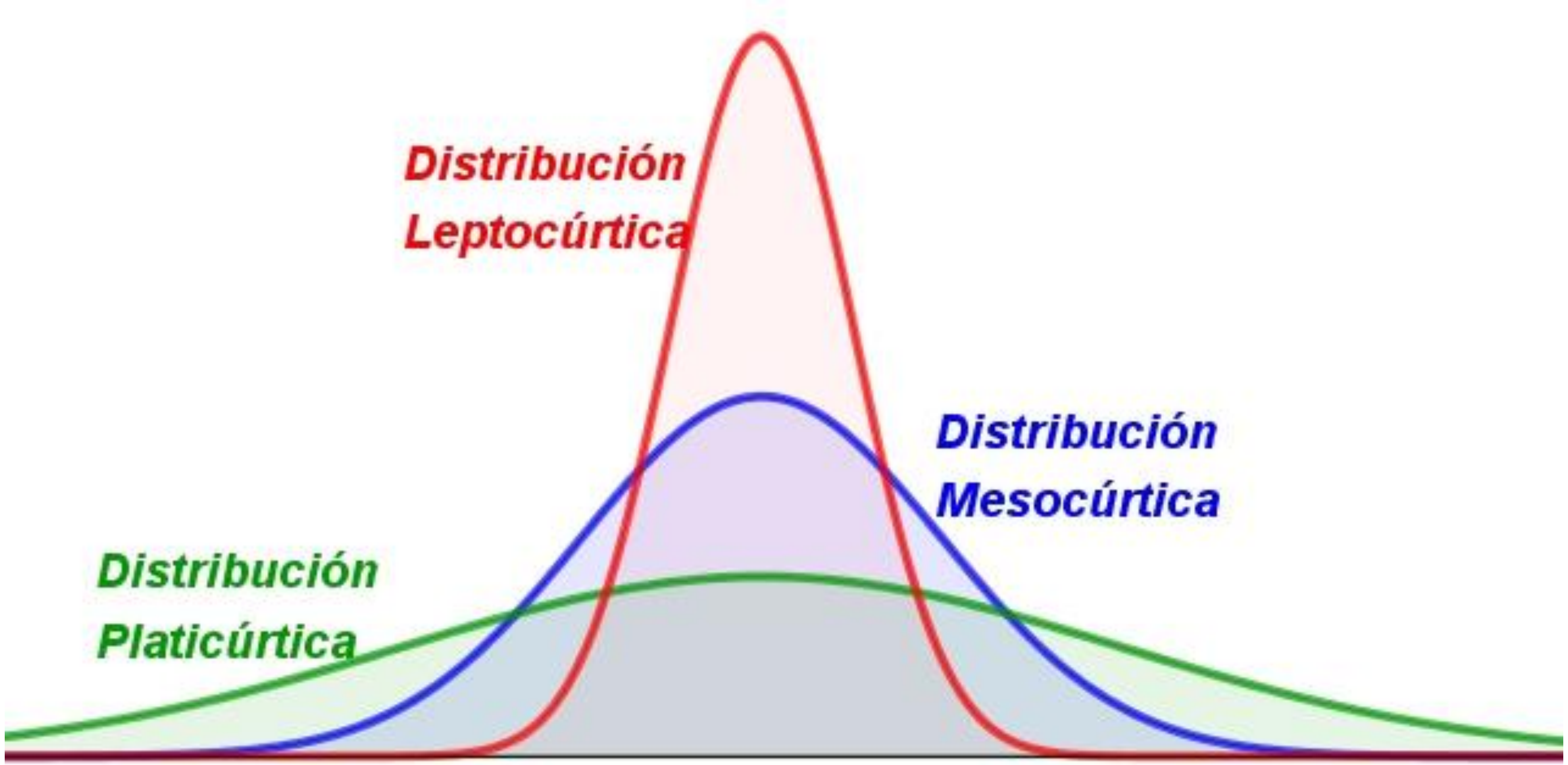
$\bar{x} = 5{,}28$

Calculado con **n-1** $S^2 = 7{,}638$ $S = 2{,}764$ $m_3 = 12{,}027$ $a_3 = 0{,}578$

Calculado con **n** $S^2 = 7{,}562$ $S = 2{,}749$ $m_3 = 11{,}906$ $a_3 = 0{,}564$

KURTOSIS

Analiza el grado de concentración que presentan los valores alrededor de la zona central de distribución



KURTOSIS

Coeficiente de Kurtosis

$$a_4 = \frac{m_4}{s^4} \qquad a_4 = \frac{\frac{\sum(x_i - \bar{x})^4}{n}}{s^4}$$

- Si $a_4 > 3$, la distribución es **leptocúrtica**, presenta un elevado grado de concentración alrededor de los valores centrales de la variable. Es más apuntada que la normal.
- Si $a_4 < 3$, la distribución es **platicúrtica**, presenta un reducido grado de concentración alrededor de los valores centrales de la variable. Es más aplastada que la normal
- Si $a_4 = 3$, la distribución **mesocúrtica**, presenta un grado de concentración medio alrededor de los valores centrales de la variable. Es igual que la normal

EJEMPLO

X	fi	$xi\,.\,fi$	$(xi - \bar{x})^2$	$(xi - \bar{x})^2.fi$	$(xi - \bar{x})^4$	$(xi - \bar{x})^4.fi$
2	21	42	10,76	225,93	115,743171	2430,60658
4	39	156	1,64	63,89	2,68435456	104,689828
6	12	72	0,52	6,22	0,26873856	3,22486272
8	11	88	7,39	81,38	54,7363226	602,099548
10	17	170	22,28	378,73	496,327107	8437,56081
	100	528		756,16		11578,1816

$\bar{x} = 5{,}28$

Calculado con **n-1** $S^2 = 7{,}638$ $S = 2{,}764$ $m_4 = 116{,}951$ $a_4 = 2{,}045$

Calculado con **n** $S^2 = 7{,}562$ $S = 2{,}749$ $m_4 = 115{,}782$ $a_4 = 1{,}985$

- Medidas descriptivas
  - Medidas de tendencia central
    - Media
    - Mediana
    - Moda
  - Dispersión
    - Rango
    - Desviación media
    - Varianza
    - Desviación estándar
  - Forma
    - Asimetría
      - Coeficientes de Pearson
      - Coeficiente de Fisher
    - Kurtosis
      - Coeficiente de Kurtosis

MEDIDAS DESCRIPTIVAS

RESOLUCIÓN DE EJERCICIOS

EJERCICIO 2

En un estudio de dos semanas sobre la productividad de los trabajadores de una empresa, se obtuvieron los siguientes datos sobre el número total de piezas aceptables que produjeron 50 operarios.

65	36	49	84	79	56	38	43	67	36
43	78	37	40	68	72	55	62	32	82
88	50	60	56	57	46	39	57	73	65
59	48	76	74	70	51	40	75	56	45
35	62	58	49	76	38	53	43	42	51

a. Realice la tabla de distribución con las clases 30 a 39; 40 a 49; 50 a 59; 60 a 69; 70 a 79; 80 a 89
b. Calcular en base a la frecuencia calculada en el punto anterior: la media aritmética, mediana, moda, varianza, desviación típica, desviación media, primer y segundo coeficientes de Pearson, coeficiente de Kurtosis, segundo momento de la Variable y tercer momento respecto a la media
c. Confeccionar las distribuciones de frecuencias que conoce y un gráfico de ojiva.

EJERCICIO 2

Clase	f	LRI	LRS	MC	c
30 a 39	8	29,5	39,5	34,5	10
40 a 49	11	39,5	49,5	44,5	10
50 a 59	12	49,5	59,5	54,5	10
60 a 69	7	59,5	69,5	64,5	10
70 a 79	9	69,5	79,5	74,5	10
80 a 89	3	79,5	89,5	84,5	10
	50				

$$MC = \frac{(LRI + LRS)}{2}$$

$$c = LRS - LRI$$

$$LRI = \frac{Límite\ nominal\ inferior\ f_i + Límite\ nominal\ superior\ f_{i-1}}{2}$$

$$LRS = \frac{Límite\ nominal\ inferior\ f_{i+1} + Límite\ nominal\ superior\ f_i}{2}$$

EJERCICIO 2

Clase	f	LRI	LRS	MC	c	MCi . fi	f acum
30 a 39	8	29,5	39,5	34,5	10	276	8
40 a 49	11	39,5	49,5	44,5	10	489,5	19
50 a 59	12	49,5	59,5	54,5	10	654	31
60 a 69	7	59,5	69,5	64,5	10	451,5	38
70 a 79	9	69,5	79,5	74,5	10	670,5	47
80 a 89	3	79,5	89,5	84,5	10	253,5	50
	50					2795	

$$\bar{X} = \frac{\sum MC_i . f_i}{n} = \frac{2795}{50} = \boxed{55.9}$$

$$\tilde{X} = LRI + \frac{\frac{n}{2} - F_{i-1}}{f_i} . c = 49,5 + \frac{25 - 19}{12} . 10 = \boxed{54.5}$$

$$\hat{X} = LRI + \frac{d_1}{d_1 + d_2} . c = 49,5 + \frac{1}{1+5} .10 = \boxed{51,16}$$

$$d_1 = f_i - f_{i-1} = 12 - 11 = 1$$

$$d_2 = f_i - f_{i+1} = 12 - 7 = 5$$

EJERCICIO 2

Clase	f	LRI	LRS	MC	c	$\lvert MCi - \bar{x}\rvert . fi$	$(MCi - x)^2 . fi$
30 a 39	8	29,5	39,5	34,5	10	171,2	3663,68
40 a 49	11	39,5	49,5	44,5	10	125,4	1429,56
50 a 59	12	49,5	59,5	54,5	10	16,8	23,52
60 a 69	7	59,5	69,5	64,5	10	60,2	517,72
70 a 79	9	69,5	79,5	74,5	10	167,4	3113,64
80 a 89	3	79,5	89,5	84,5	10	85,8	2453,88
	50					626,8	11202

$$DM = \frac{\sum \lvert MC_i - \bar{x}\rvert \, . f_i}{n} = \frac{626{,}8}{50} = \boxed{12{,}54} \qquad S^2 = \frac{\sum (MC_i - \bar{x})^2 \, . f_i}{n-1} = \frac{11202}{49} = \boxed{228{,}61} \qquad S = \sqrt{\frac{\sum (MC_i - \bar{x})^2 \, . f_i}{n-1}} = \sqrt{228{,}61} = \boxed{15.12}$$

EJERCICIO 2

Clase	f	LRI	LRS	MC	c	$(MCi - x)^4 . fi$
30 a 39	8	29,5	39,5	34,5	10	1677818,89
40 a 49	11	39,5	49,5	44,5	10	185785,618
50 a 59	12	49,5	59,5	54,5	10	46,0992
60 a 69	7	59,5	69,5	64,5	10	38290,5712
70 a 79	9	69,5	79,5	74,5	10	1077194,89
80 a 89	3	79,5	89,5	84,5	10	2007175,68
	50					4986311,76

$$S_{k1} = \frac{\bar{x} - \hat{x}}{s} = \frac{55{,}9 - 51{,}16}{15{,}12} = \boxed{0{,}31}$$

$$S_{k2} = \frac{3.(\bar{x} - \tilde{x})}{s} = \frac{3.(55{,}9 - 54{,}5)}{15{,}12} = \boxed{0{,}28}$$

$$a_4 = \cancel{\frac{\frac{\sum(MC_i - \bar{x})^4 . fi}{n}}{s^4}}$$

$$a_4 = \frac{\frac{\sum(MC_i - \bar{x})^4 . fi}{n-1}}{s^4} = \frac{\frac{4986311{,}76}{49}}{15{,}12^4} = \boxed{1{,}94}$$

EJERCICIO 2

Clase	f	MC	MCi^2	$MCi^2 . fi$	$MCi - \bar{x}$	$(MCi - \bar{x})^3$	$(MCi - \bar{x})^3 . fi$
30 a 39	8	34,5	1190,25	9522	-21,4	-9800,344	-78402,752
40 a 49	11	44,5	1980,25	21782,75	-11,4	-1481,544	-16296,984
50 a 59	12	54,5	2970,25	35643	-1,4	-2,744	-32,928
60 a 69	7	64,5	4160,25	29121,75	8,6	636,056	4452,392
70 a 79	9	74,5	5550,25	49952,25	18,6	6434,856	57913,704
80 a 89	3	84,5	7140,25	21420,75	28,6	23393,656	70180,968
	50			167442,5			37814,4

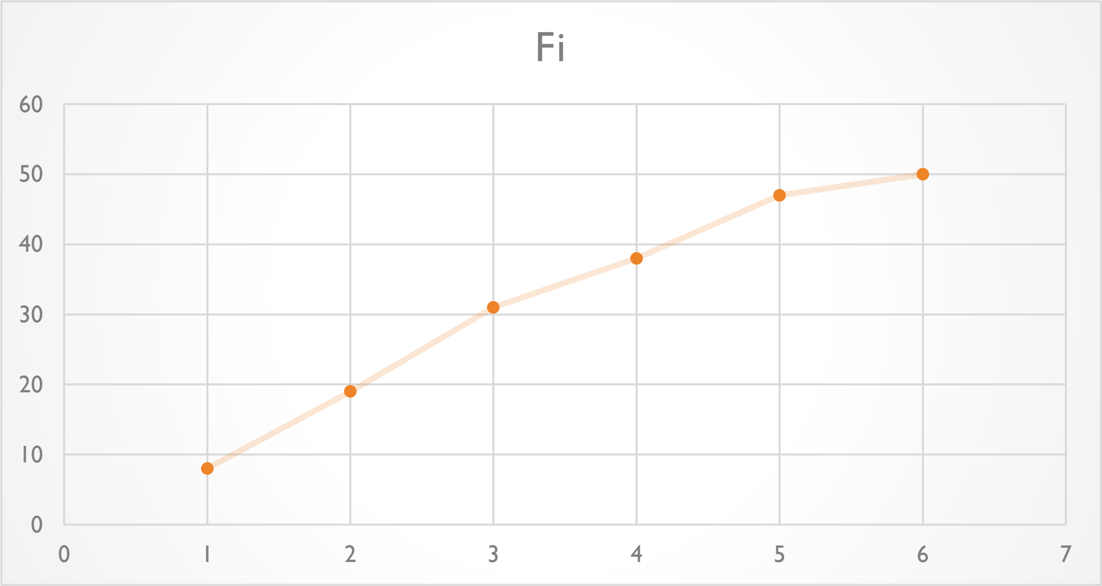
$$\bar{x}^r = \frac{\sum x_i^r}{n}$$

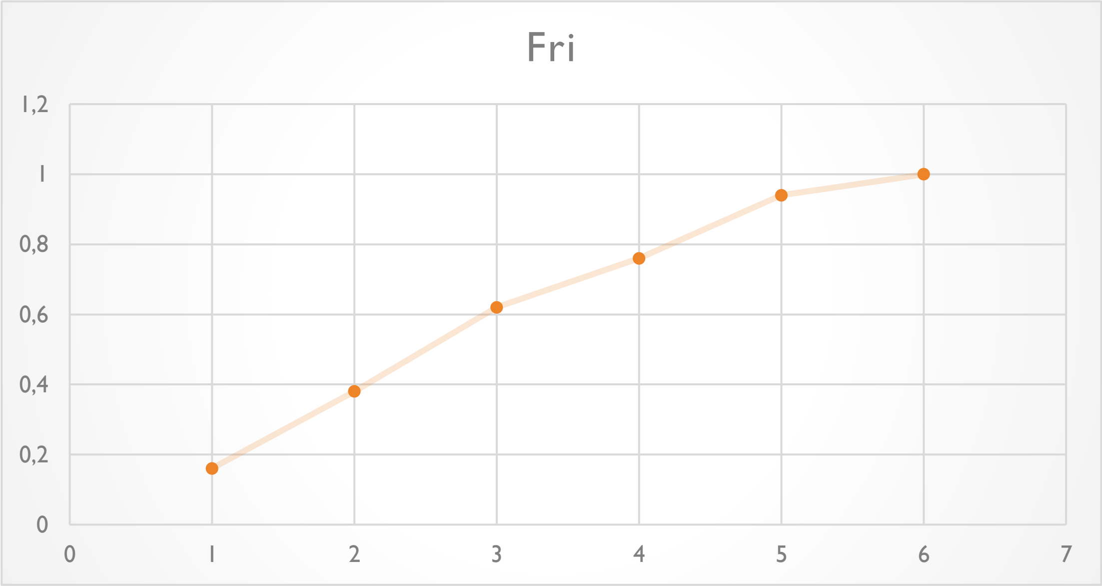
$$m_r = \frac{\sum (x_i - k)^r}{n}$$

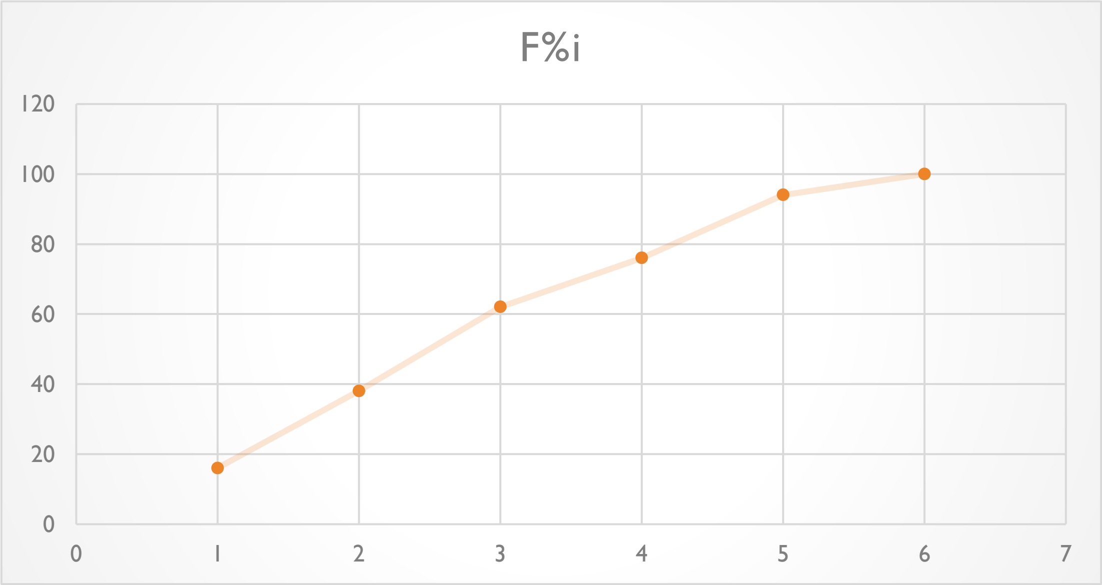
$$\bar{x}^2 = \frac{\sum MC_i^2 . fi}{n} = \frac{167442{,}5}{50} = \boxed{3348{,}85}$$

$$m_3 = \frac{\sum (MC_i - \bar{x})^3 . fi}{n} = \frac{37814{,}4}{50} = \boxed{756{,}29}$$

Clase	f	Fi	fri	Fri	f%i	F%i
30 a 39	8	8	0,16	0,16	16	16
40 a 49	11	19	0,22	0,38	22	38
50 a 59	12	31	0,24	0,62	24	62
60 a 69	7	38	0,14	0,76	14	76
70 a 79	9	47	0,18	0,94	18	94
80 a 89	3	50	0,06	1	6	100
	50					


Fi
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7


Fri
1,2
1
0,8
0,6
0,4
0,2
0
0
1
2
3
4
5
6
7


F%i
120
100
80
60
40
20
0
0
1
2
3
4
5
6
7

EJERCICIO 3

La siguiente tabla da la distribución de artículos defectuosos encontrados en 200 lotes de artículos manufacturados

No. de art. defect.	1	2	3	4	5	6	7
No. de lotes	10	25	58	65	20	16	6

a. Calcular: Media, mediana, moda, rango, desviación media, desviación típica y varianza, coeficientes de Pearson, coeficiente de Fisher y coeficiente de Kurtosis.
b. Construir cuatro distribuciones que Ud. conozca.
c. Construir un diagrama de Pareto y un círculo radiado

EJERCICIO 3

No. de art. defect.	No. de lotes	f acum	x.fi
1	10	10	10
2	25	35	50
3	58	93	174
4	65	158	260
5	20	178	100
6	16	194	96
7	6	200	42
	200		732

$$\bar{X} = \frac{\sum x_i . f_i}{n} = \frac{732}{200} = \boxed{3{,}66}$$

$$\tilde{X} = \boxed{4}$$

$$\hat{X} = \boxed{4}$$

EJERCICIO 3

No. de art. defect.	No. de lotes	x.fi	$\|xi - x\|.fi$	$(xi - x)^2.fi$
1	10	10	26,6	70,756
2	25	50	41,5	68,89
3	58	174	38,28	25,2648
4	65	260	22,1	7,514
5	20	100	26,8	35,912
6	16	96	37,44	87,6096
7	6	42	20,04	66,9336
	200	732	212,76	362,88

$$Rango = 7 - 1 = \boxed{6}$$

$$DM = \frac{\sum |x_i - \bar{x}| \, . \, f_i}{n} = \frac{212{,}76}{200} = \boxed{1{,}06}$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2 \, . \, f_i}{n - 1} = \frac{362{,}88}{199} = \boxed{1{,}82}$$

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 \, . \, f_i}{n - 1}} = \boxed{1{,}35}$$

EJERCICIO 3

No. de art. defect.	No. de lotes	$(xi-x)^3.fi$	$(xi-x)^4.fi$
1	10	-188,21	500,64
2	25	-114,36	189,83
3	58	-16,67	11,01
4	65	2,55	0,87
5	20	48,12	64,48
6	16	205,01	479,72
7	6	223,56	746,68
	200	159,99	1993,23

$$S_{k1} = \frac{\bar{x}-\hat{x}}{s} = \frac{3,66-4}{1,35} = \boxed{-0,25}$$

$$S_{k2} = \frac{3.(\bar{x}-\tilde{x})}{s} = \frac{3.(3,66-4)}{1,35} = \boxed{-0,76}$$

$$a_3 = \cancel{\frac{\frac{\sum(x_i-\bar{x})^3.fi}{n}}{s^3}}$$

$$a_3 = \frac{\frac{\sum(x_i-\bar{x})^3.fi}{n-1}}{s^3} = \frac{\frac{159,99}{199}}{1,35^3} = \boxed{0,33}$$

$$a_4 = \cancel{\frac{\frac{\sum(x_i-\bar{x})^4.fi}{n}}{s^4}}$$

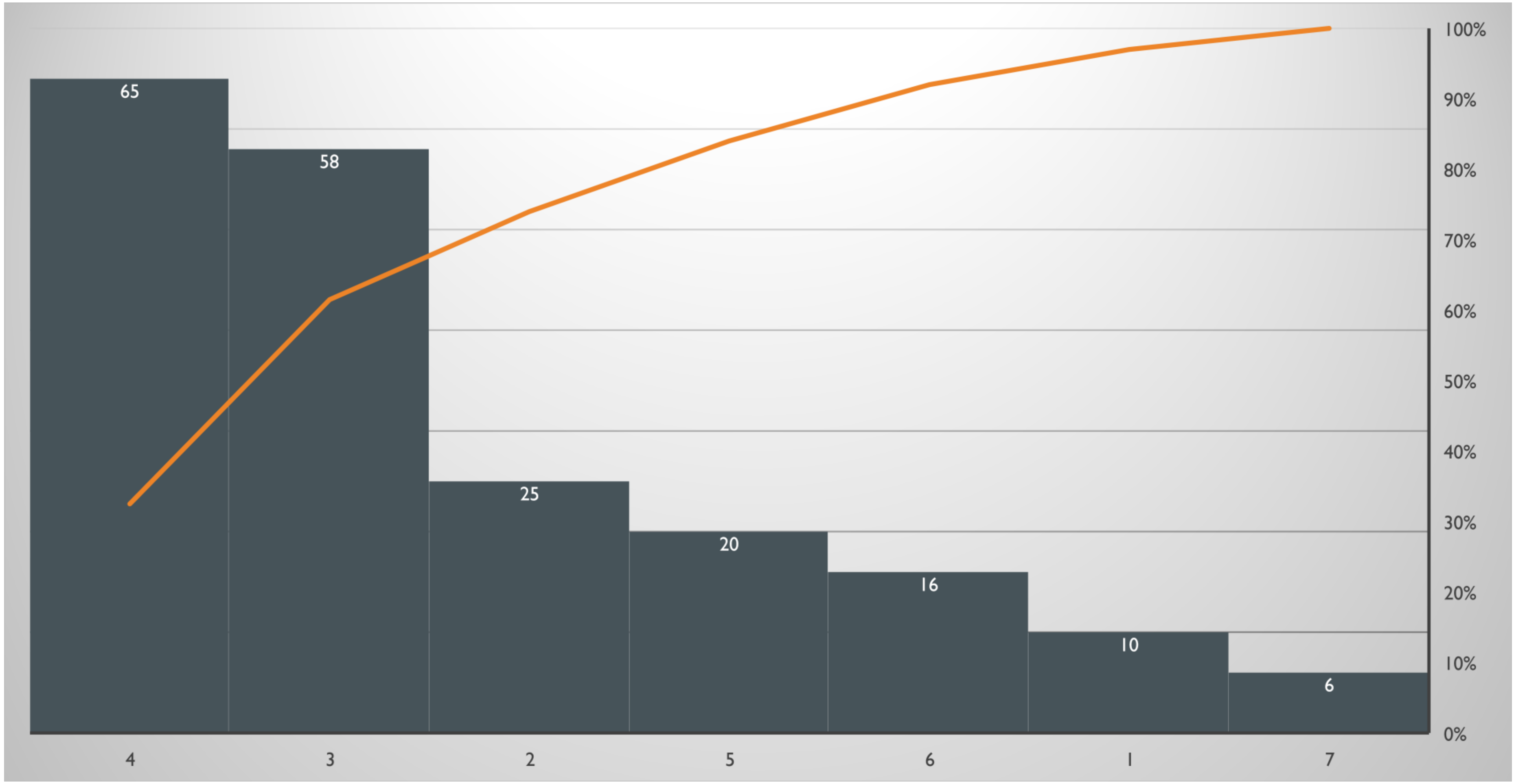
$$a_4 = \frac{\frac{\sum(x_i-\bar{x})^4.fi}{n-1}}{s^4} = \frac{\frac{1993,23}{199}}{1,35^4} = \boxed{3,02}$$

EJERCICIO 3

No. de art. defect.	No. de lotes	Fi	fri	Fri	f%i	F%i
1	10	10	0,05	0,05	5	5
2	25	35	0,125	0,175	12,5	17,5
3	58	93	0,29	0,465	29	46,5
4	65	158	0,325	0,79	32,5	79
5	20	178	0,1	0,89	10	89
6	16	194	0,08	0,97	8	97
7	6	200	0,03	1	3	100
	200					

EJERCICIO 3

No. de art. defect.	No. de lotes
4	65
3	58
2	25
5	20
6	16
1	10
7	6



EJERCICIO 4

La tabla muestra la distribución de diámetro de tornillos de una muestra de la producción del día

Diámetro en mm.	28	29	30	31	32	33	34
No. Tornillos	103	225	358	565	204	166	69

a. Calcular media aritmética, mediana, moda, varianza, desviación típica, desviación media, primer y segundo coeficientes de Pearson, coeficiente de Kurtosis, segundo momento de la Variable y tercer momento respecto a la media

a. Completar 3 distribuciones de frecuencias y confeccionar un histograma de frecuencias.

EJERCICIO 4

Diámetro en mm.	No. Tornillos	f acum	x.f
28	103	103	2884
29	225	328	6525
30	358	686	10740
31	565	1251	17515
32	204	1455	6528
33	166	1621	5478
34	69	1690	2346
	1690		52016

$$\bar{X} = \frac{\sum x_i . f_i}{n} = \frac{52016}{1690} = \boxed{30{,}78}$$

$$\tilde{X} = \boxed{31}$$

$$\hat{X} = \boxed{31}$$

EJERCICIO 4

Diámetro en mm.	No. Tornillos	xi.fi	$\lvert xi - x \rvert . fi$	$(xi - x)^2 . fi$
28	103	2884	286,21	795,28
29	225	6525	400,21	711,85
30	358	10740	278,77	217,08
31	565	17515	125,04	27,67
32	204	6528	249,15	304,28
33	166	5478	368,74	819,07
34	69	2346	222,27	716,00
	1690	52016	1930,37	3591,23

$$DM = \frac{\sum \lvert x_i - \bar{x} \rvert . f_i}{n} = \frac{1930,37}{1690} = \boxed{1,14}$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2 . f_i}{n - 1} = \frac{3591,23}{1689} = \boxed{2,12}$$

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 . f_i}{n - 1}} = \boxed{1,46}$$

EJERCICIO 4

Diámetro en mm.	No. Tornillos	$(xi-x)^3.fi$	$(xi-x)^4.fi$	$xi^2.fi$
28	103	-2209,84	6140,49	80752
29	225	-1266,16	2252,12	189225
30	358	-169,04	131,63	322200
31	565	6,12	1,36	542965
32	204	371,62	453,86	208896
33	166	1819,41	4041,46	180774
34	69	2306,45	7429,76	79764
	1690	858,56	20450,67	1604576

$$\bar{x}^2 = \frac{\sum x_i^2.fi}{n}$$

$$\bar{x}^2 = \frac{1604576}{1690} = \boxed{949,45}$$

$$S_{k1} = \frac{\bar{x}-\hat{x}}{s} = \frac{30,78-31}{1,46} = \boxed{-0,15}$$

$$S_{k2} = \frac{3.(\bar{x}-\tilde{x})}{s} = \frac{3.(30,78-31)}{1,46} = \boxed{-0,46}$$

$$a_3 = \frac{\frac{\sum(x_i-\bar{x})^3.fi}{n-1}}{s^3} = \frac{\frac{858,56}{1689}}{1,46^3} = \boxed{0,16}$$

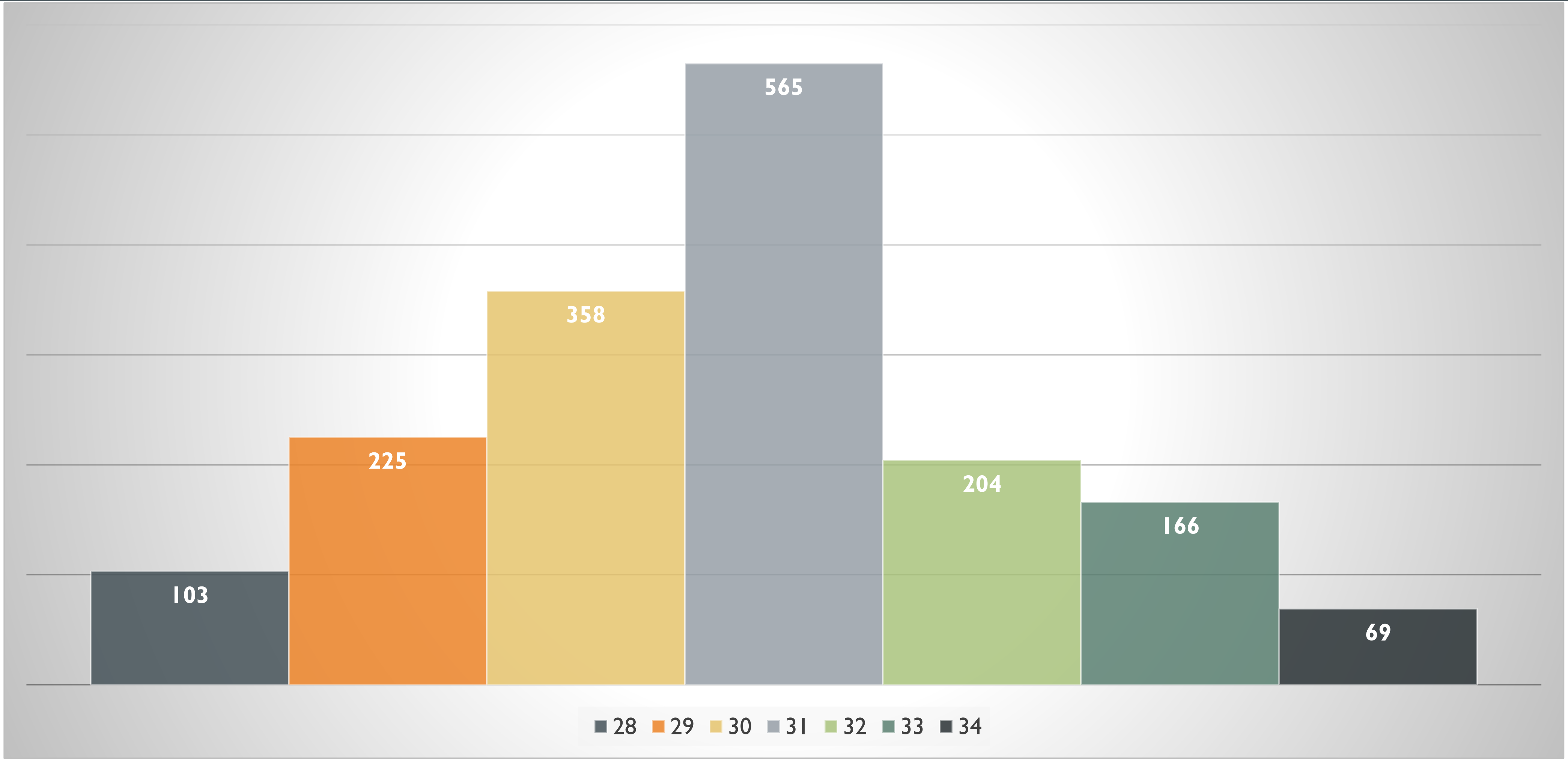
$$a_4 = \frac{\frac{\sum(x_i-\bar{x})^4.fi}{n-1}}{s^4} = \frac{\frac{20450,67}{1689}}{1,46^4} = \boxed{2,68}$$

$$m_3 = \frac{\sum(x_i-\bar{x})^3.fi}{n-1} = \frac{858,56}{1689} = \boxed{0,51}$$

EJERCICIO 4

Diámetro en mm.	No. Tornillos	Fi	fri	Fri
28	103	103	0,06	0,06
29	225	328	0,13	0,19
30	358	686	0,21	0,41
31	565	1251	0,33	0,74
32	204	1455	0,12	0,86
33	166	1621	0,10	0,96
34	69	1690	0,04	1
	1690			

EJERCICIO 4



EJERCICIO 10

Con los siguientes datos muestrales calcule la media, la mediana, la moda y el coeficiente de asimetría de Fisher (explique el significado del valor obtenido):

xi	6	8	10	13	18
fi	4	10	7	4	3

xi	fi	f acum	x.f	(xi-x)2 fi	(xi-x)3 fi
6	4	4	24	64	-256
8	10	14	80	40	-80
10	7	21	70	0	0
13	4	25	52	36	108
18	3	28	54	192	1536
	28		280	332	1308

$$\bar{X} = \frac{\sum x_i . f_i}{n} = \frac{280}{28} = \boxed{10{,}29}$$

$$\tilde{X} = \boxed{9}$$

$$\hat{X} = \boxed{8}$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2 . f_i}{n-1} = \frac{332}{27} = 12{,}29$$

$$S = 3{,}51$$

$$a_3 = \frac{\frac{\sum (x_i - \bar{x})^3 . fi}{n-1}}{s^3} = \frac{\frac{1308}{27}}{3{,}51^3} = \boxed{1{,}12}$$

MEDIDAS DESCRIPTIVAS

MEDIDAS DE POSICIÓN

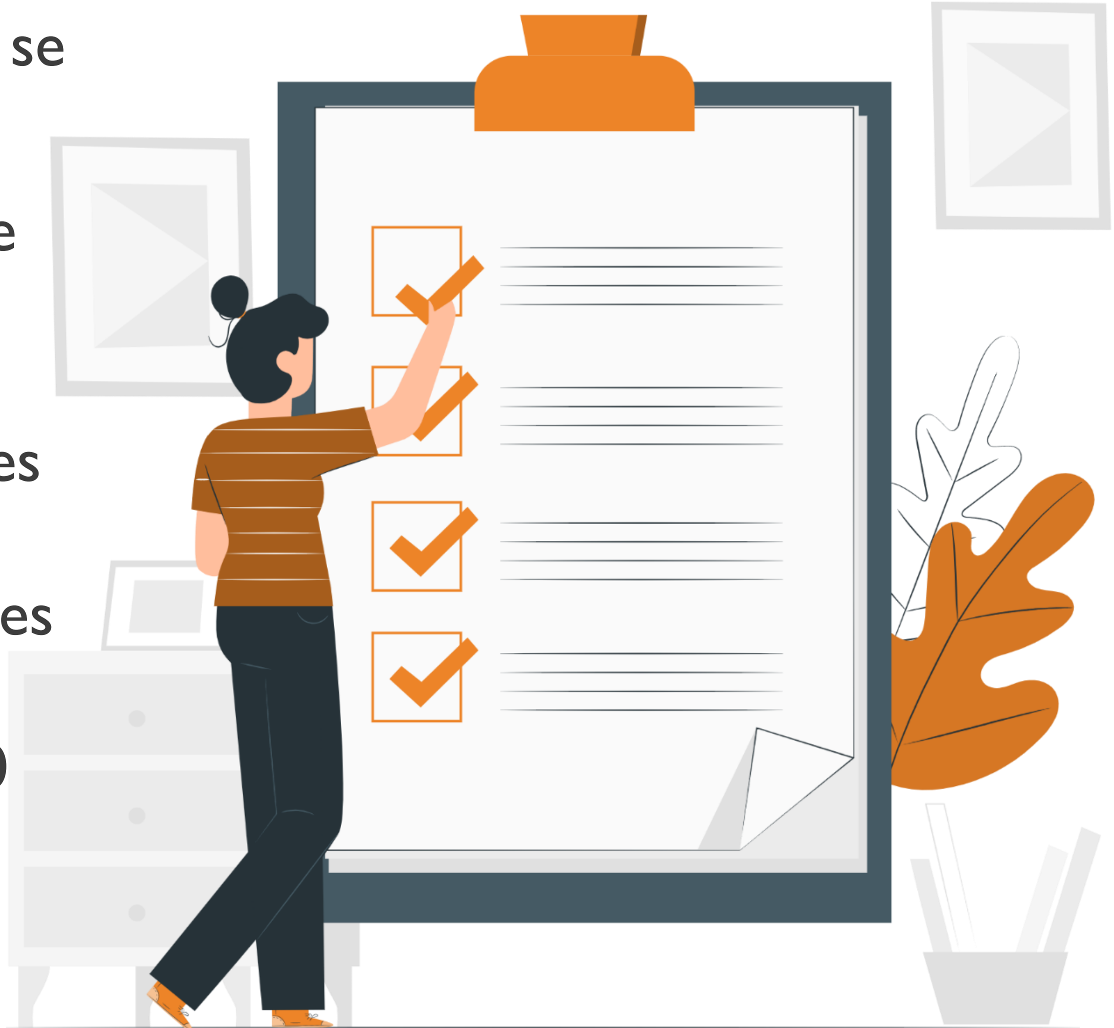
PERCENTILES

El percentil p es un valor tal que por lo menos p por ciento de las observaciones son menores o iguales que este valor y por lo menos $(100 - p)$ por ciento de las observaciones son mayores o iguales que este valor

EJEMPLO

Las puntuaciones en los exámenes de admisión se suelen dar en términos de percentiles.

Por ejemplo, suponga que un estudiante obtiene 54 puntos en la parte verbal del examen de admisión. Esto no dice mucho acerca de este estudiante en relación con los demás estudiantes que realizaron el examen. Sin embargo, si esta puntuación corresponde al percentil 70, entonces 70% de los estudiantes obtuvieron una puntuación menor a la de dicho estudiante y 30 de los estudiantes obtuvieron una puntuación mayor



PERCENTILES

1. **Ordenar** los datos de menor a mayor (colocar los datos en orden ascendente).
2. **Calcular** el índice *i* donde *p* es el percentil deseado y *n* es el tamaño de muestra.

$$i = \left(\frac{p}{100}\right).n$$

3. **Determinar** el percentil

A) Si i no es un número entero, debe redondearlo. El primer entero mayor que i denota la posición del percentil p.

B) Si i es un número entero, el percentil p es el promedio de los valores en las posiciones i e i + 1.

PERCENTILES

- Para datos agrupados:

$$P = LRI + \frac{i - F_{i-1}}{f_i} \, . \, c$$

EJEMPLO

- Los datos siguientes corresponden a sueldos mensuales iniciales en una muestra de 12 egresados de la carrera de administración. Determine el percentil 85.

3450	3490
3550	3730
3650	3540
3480	3925
3355	3520
3310	3480

EJEMPLO

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

$$i = \left(\frac{p}{100}\right).n$$

$$i = \left(\frac{85}{100}\right).12$$

$$i = 10{,}2$$

Como *i* no es un número entero, se debe *redondear.* La posición del percentil 85 es el primer entero mayor que 10.2, es la posición 11

Entonces el percentil 85 es el dato en la posición 11, o sea 3730.

EJEMPLO

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

Si calculo el percentil 50, se obtiene:

$$i = \left(\frac{p}{100}\right).n$$

$$i = \left(\frac{50}{100}\right).12$$

$$i = 6$$

Como *i* es un número entero, de acuerdo el percentil 50 es el promedio de los valores de los datos que se encuentran en las posiciones seis y siete; de manera que el percentil
50 es:

$$\frac{(3490\ +\ 3520)}{2} = 3505$$

El *percentil 50 coincide con la mediana.*

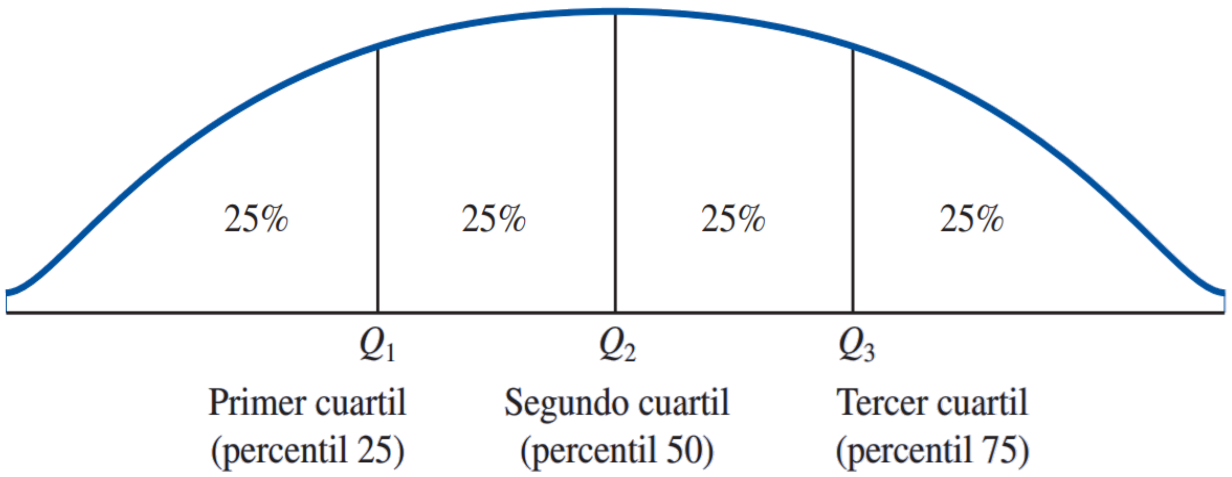
CUARTILES

Con frecuencia es conveniente dividir los datos en cuatro partes; así, cada parte contiene una cuarta parte o 25% de las observaciones. A los puntos de división se les conoce como cuartiles y están definidos como sigue:

Q1 = primer cuartil, o percentil 25

Q2 = segundo cuartil, o percentil 50

Q3 = tercer cuartil, o percentil 75



EJEMPLO

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

Q2, el segundo cuartil (la mediana), es 3505.

$$i = \left(\frac{p}{100}\right).n = \left(\frac{25}{100}\right).12 = 3$$

Como i es un entero el primer cuartil, o el percentil 25, es el promedio del tercer y cuarto valores de los datos:

$$Q1 = \frac{(3450 + 3480)}{2} = 3465$$

$$i = \left(\frac{p}{100}\right).n = \left(\frac{75}{100}\right).12 = 9$$

Como i es un entero el tercer cuartil, o el percentil 75, es el promedio del noveno y décimo valores de los datos:

$$Q3 = \frac{(3550 + 3650)}{2} = 3600$$

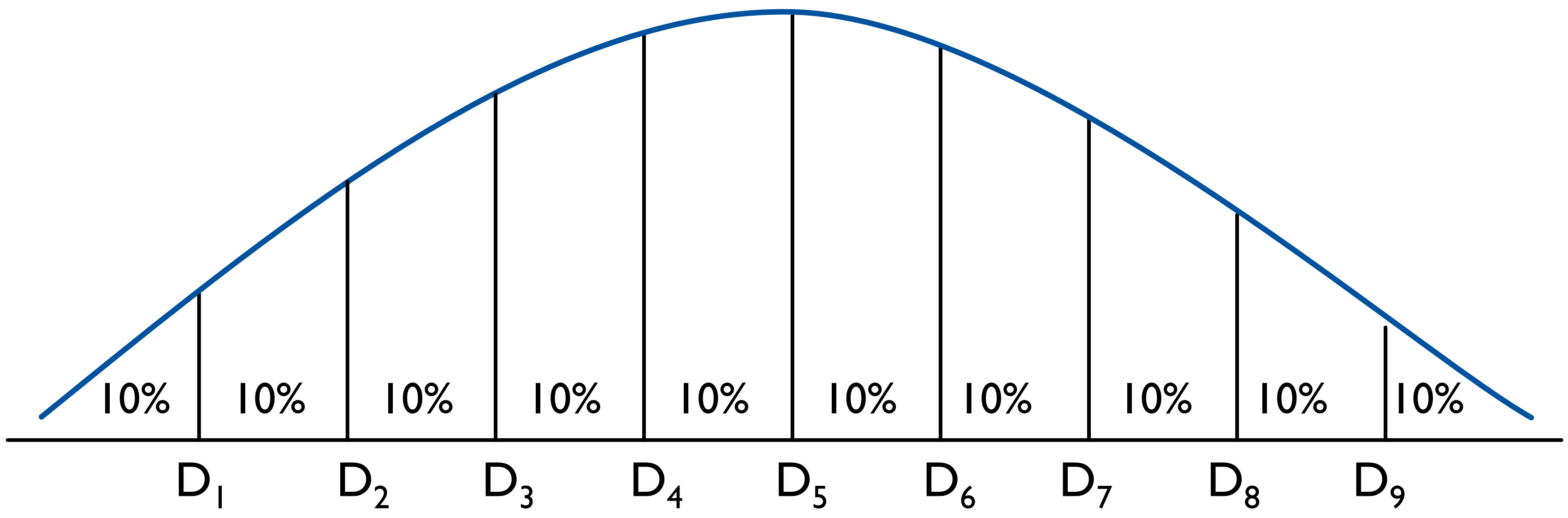
EJEMPLO

3310	3355	3450	3480	3480	3490	3520	3540	3550	3650	3730	3925

$Q_1 = 3465$ $\qquad$ $Q_2 = 3505$ (Mediana) $\qquad$ $Q_3 = 3600$

DECILES

- Son los valores de la variable que dividen un conjunto ordenado de datos en diez partes iguales



MEDIDAS DE DISPERSIÓN

RANGO INTERCUARTÍLICO

Una medida que no es afectada por los valores extremos es el rango intercuartílico (RIC). Esta medida de variabilidad es la diferencia entre el tercer cuartil Q3 y el primer cuartil Q1. En otras palabras, el rango intercuartílico es el rango en que se encuentra el 50% central de los datos.

$$IQR = Q3 - Q1$$

COEFICIENTE DE VARIACIÓN

El coeficiente de variación es la relación entre la desviación típica de una muestra y su media

$$C.V = \frac{S}{\bar{x}}$$

EJEMPLO

3310	3355	3450	3480	3480	3490	3520	3540	3550	3650	3730	3925

$Q_1 = 3465$ $\quad$ $Q_2 = 3505$ (Mediana) $\quad$ $Q_3 = 3600$

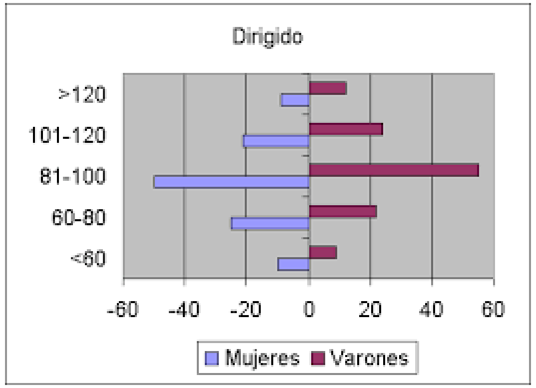
$$IQR = Q3 - Q1$$

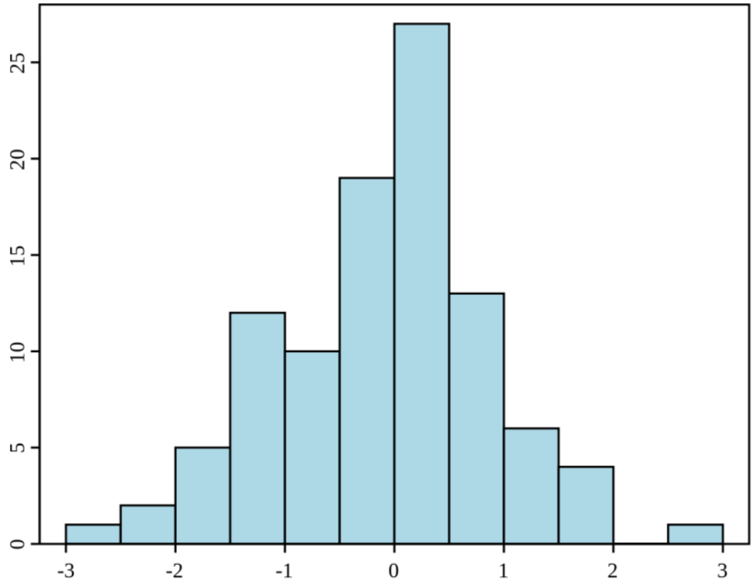
$$IQR = 3600 - 3465$$

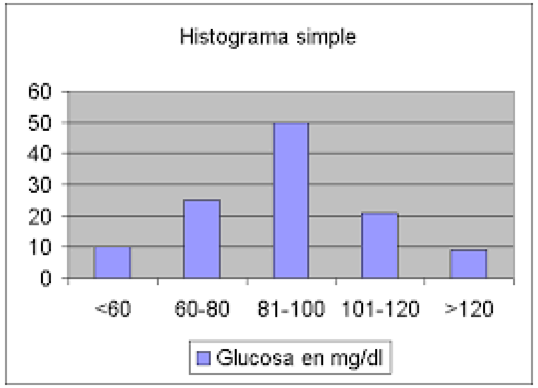
GRÁFICOS

GRÁFICOS

- Histogramas de frecuencias

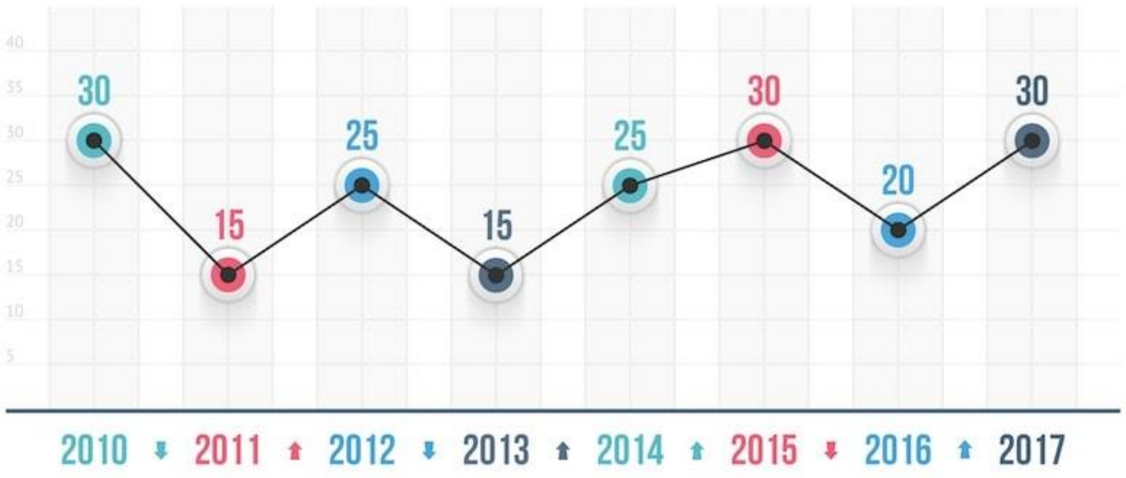


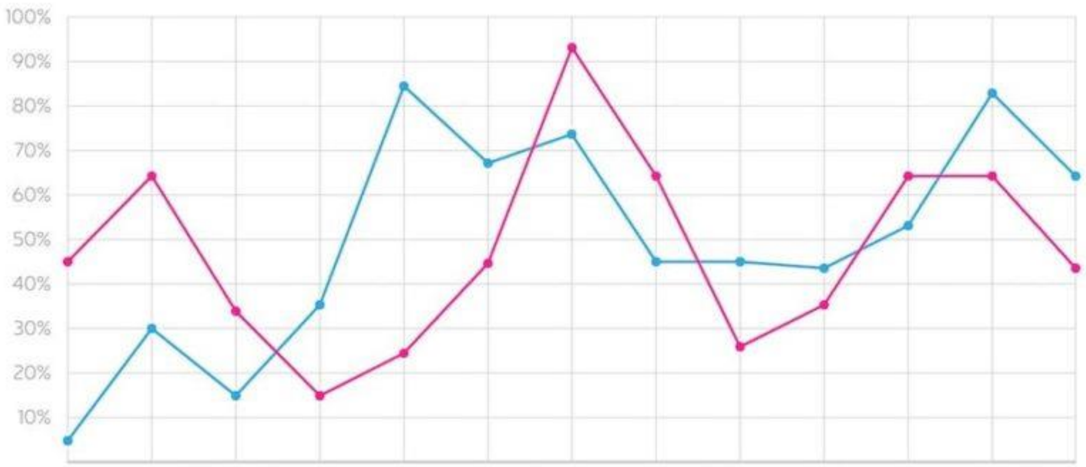




GRÁFICOS

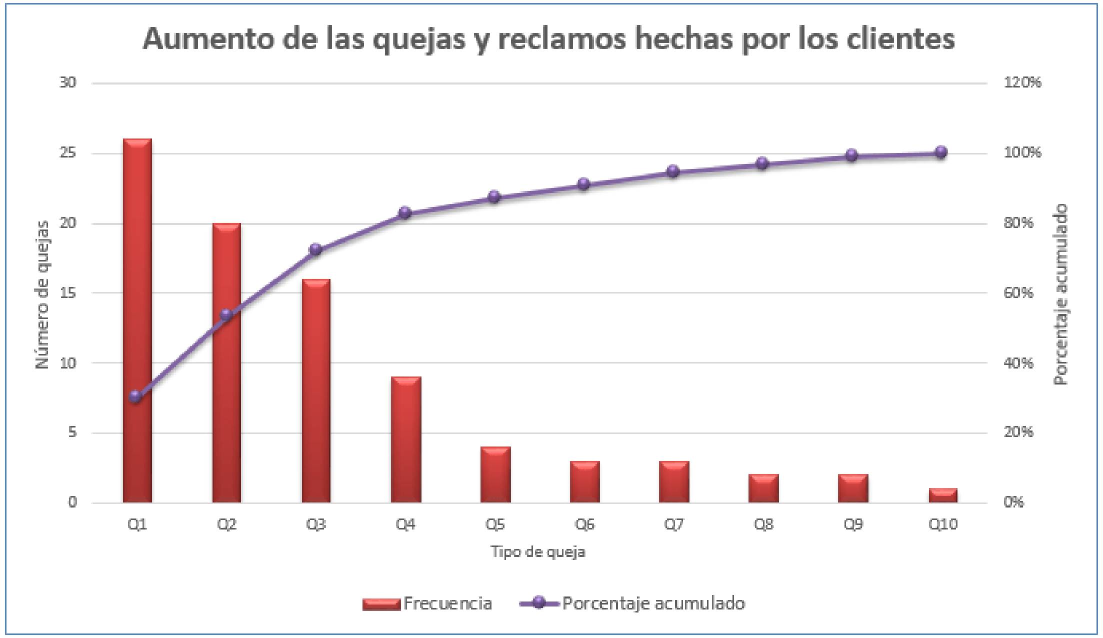
- Polígonos de frecuencia

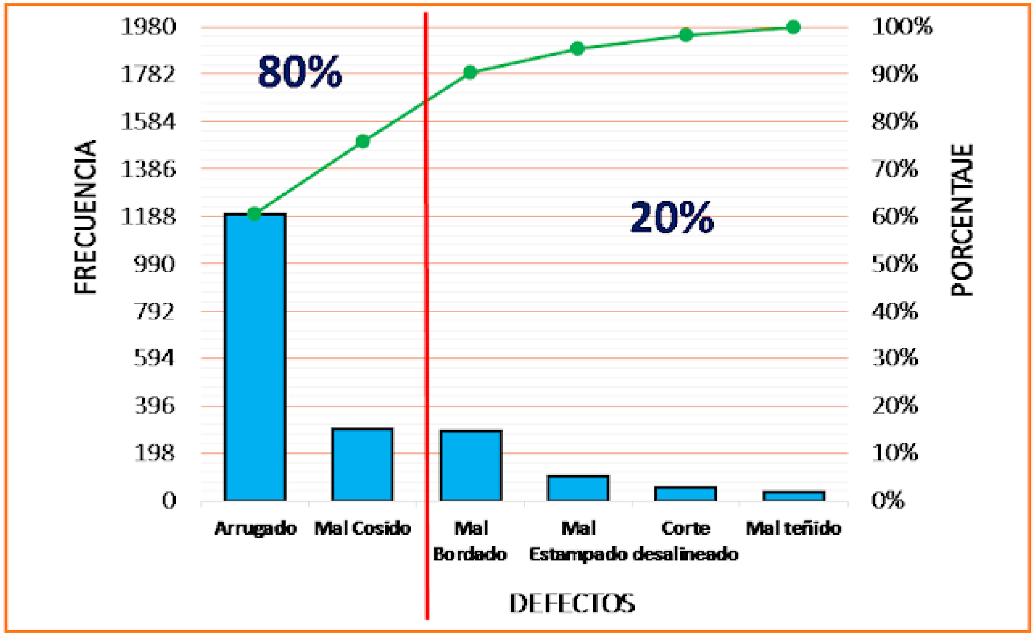




GRÁFICOS

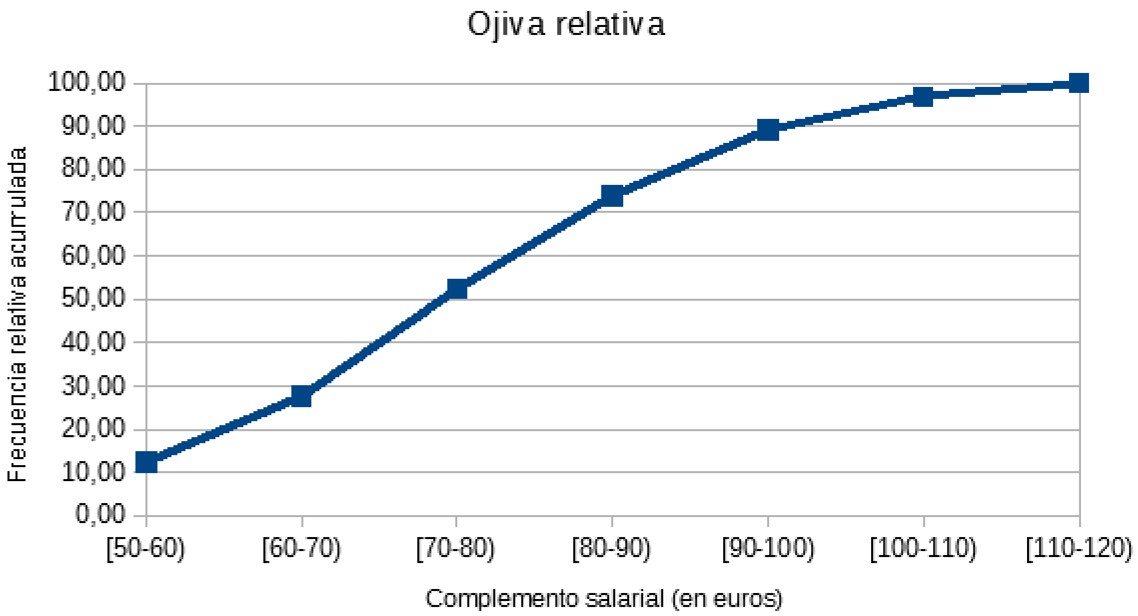
- Diagrama de Pareto

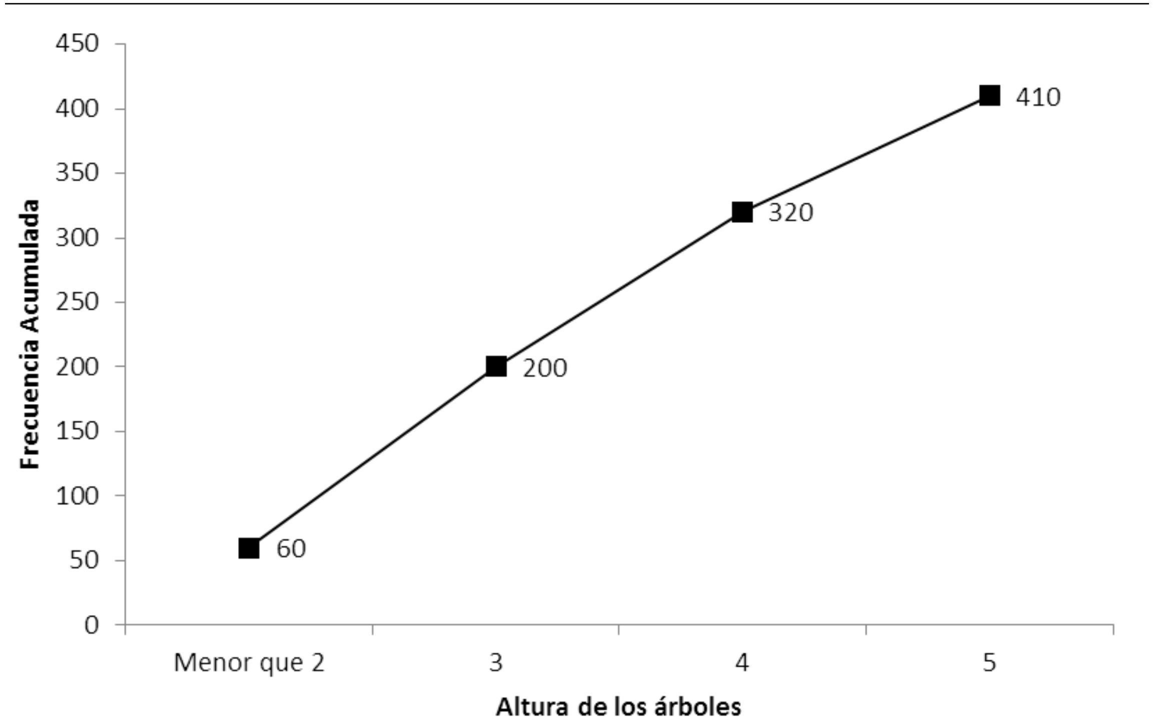




GRÁFICOS

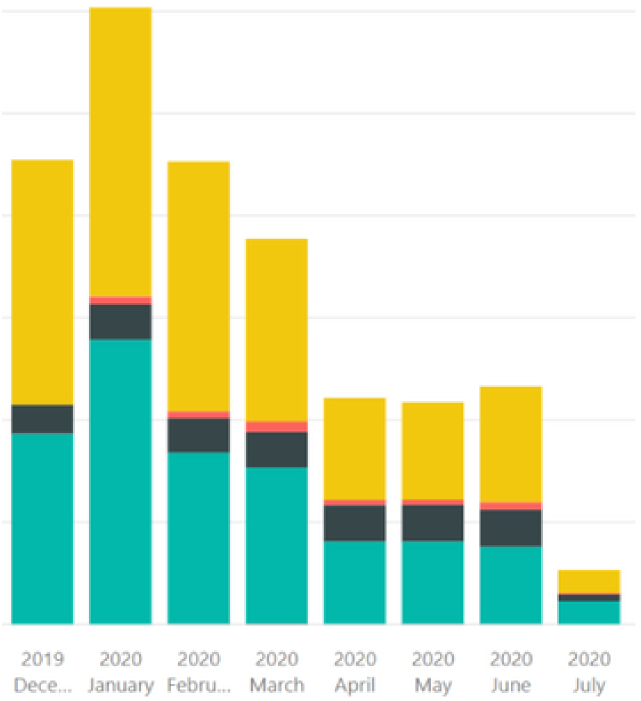
- Gráfico de Ojiva

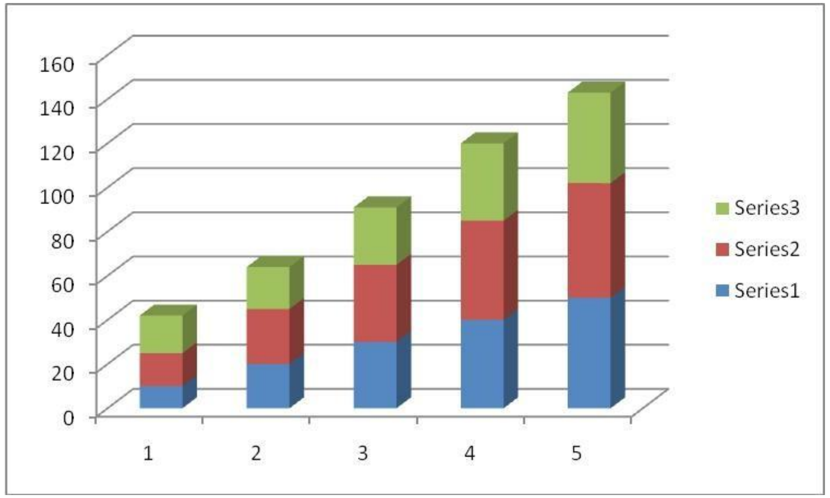


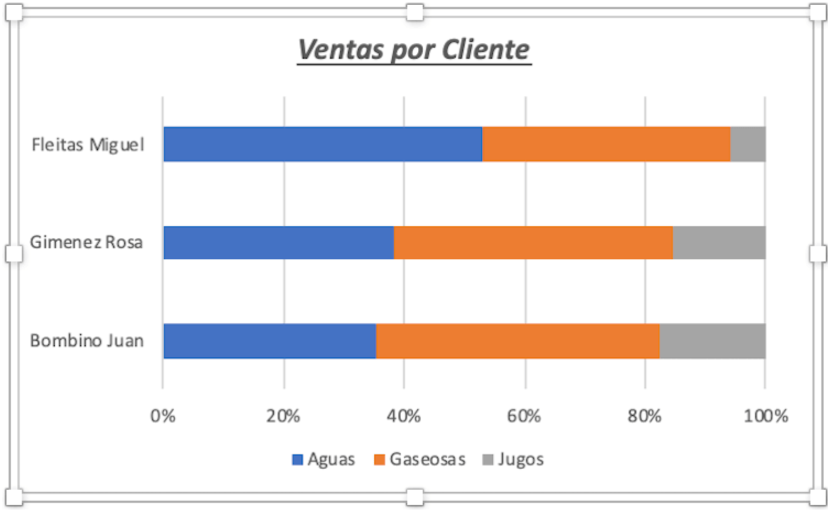


GRÁFICOS

- Barra porcentual

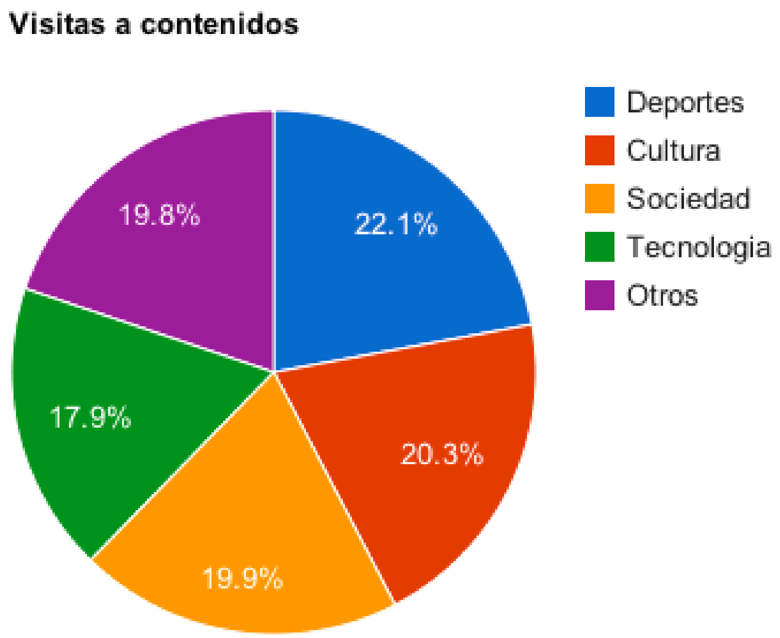


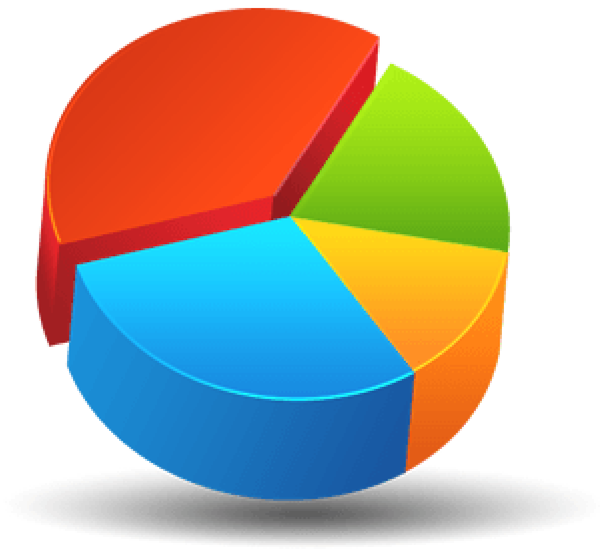




GRÁFICOS

- Circulo radiado





GRÁFICOS

- Gráfico de cajas