

EJEMPLO DE MANEJO ESTADÍSTICO DE UN GRUPO DE DATOS.

INSTRUCCIONES: Este es un ejercicio para que practiques pero también para que te autoevalúes, por lo que lo recomendable es que primero resuelvas todo el ejemplo y después veas los resultados, para compararlos con los tuyos.

Para el siguiente grupo de Datos; obtener:

- Media, Moda y Mediana
- Varianza, Desviación Estándar y Coeficiente de Variación
- 1er Cuartil, 6to Decil y 71° Percentil.
- Gráfica de Barras, Polígono de Frecuencias y Ojiva Ascendente
- Sesgo y Curtosis

Los datos son: 12, 24, 35, 42, 10, 9, 8, 31, 27, 33, 24, 15, 17, 8, 22, 31, 43, 51, 13, 14, 28, 33, 25, 14, 7, 22, 30, 31, 45, 37, 43, 12, 29, 30, 45, 47, 52, 31, 33, 20, 19, 15, 9, 14, 17, 32, 51, 40, 15, 32.

2) Organización de los Datos

Los datos se ordenan de menor a mayor y se cuentan para obtener su frecuencia. En este caso se uso un Diagrama de Árbol.

Decenas	Unidades	No. de Datos
0	9, 8, 8, 7, 9	5
1	2, 0, 5, 7, 3, 4, 4, 2, 9, 5, 4, 7, 5	13
2	4, 7, 4, 2, 8, 5, 2, 9, 0	9
3	5, 1, 3, 1, 3, 0, 1, 7, 0, 1, 3, 2, 2	13
4	2, 3, 5, 3, 5, 7, 0	7
5	1, 2, 1	3
	Total	50

Al analizar los datos ordenados, se obtiene:

Dato Menor=7 Dato Mayor=52 No. de Datos=50

Como son 50 datos, estos deben agruparse en clases:

$R = 52 - 7 = 45$ $K = 50 = 7.07$ $c = 45 / 7.07 = 6.36$

Entonces $c = 7$; por lo que los datos se agruparan en 7 clases o intervalos de ancho 7.

Intervalo	f	x
7 - 13	9	10
14 - 20	10	17
21 - 27	6	24
28 - 34	13	31
35 - 41	3	38
42 - 48	6	45
49 - 55	3	52
$\Sigma f = n =$	50	

3) Procesamiento de los Datos:

Una vez obtenidos, los intervalos, su frecuencia y marca de clase, se construye la tabla de distribución de frecuencias y se realizan las operaciones necesarias: $\bar{X} = 1347 / 50 = 26.94$

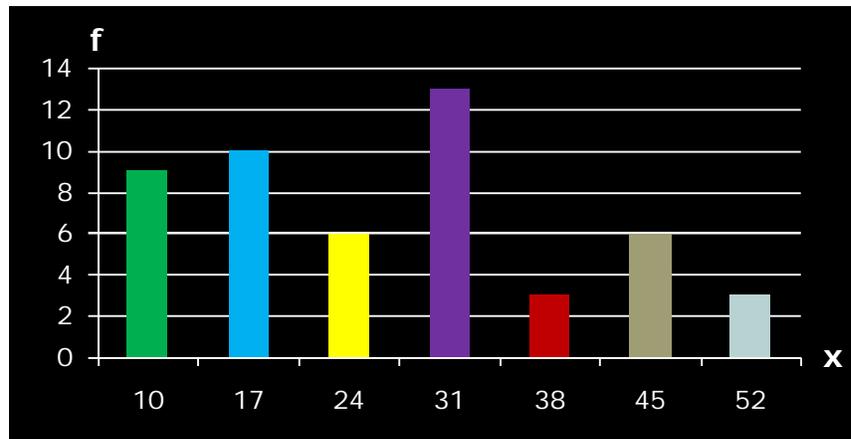
Intervalo	f	x	xf	fa	$x-\bar{X}$	$f(x-\bar{X})^2$	$f(x-\bar{X})^3$	$f(x-\bar{X})^4$
7 - 13	9	10	90	9	-16.94	2582.6724	-43750.4705	741132.97
14 - 20	10	17	170	19	-9.94	988.036	-9821.07784	97621.5137
21 - 27	6	24	144	25	-2.94	51.8616	-152.473104	448.270926
28 - 34	13	31	403	38	4.06	214.2868	870.004408	3532.2179
35 - 41	3	38	114	41	11.06	366.9708	4058.69705	44889.1894
42 - 48	6	45	270	47	18.06	1956.9816	35343.0877	638296.164
49 - 55	3	52	156	50	25.06	1884.0108	47213.3106	1183165.56
Sf = n=	50	$\Sigma xf =$	1347		$\Sigma =$	8044.82	33761.0784	2709085.89

Con los resultados obtenidos, se aplican las fórmulas necesarias y se obtienen los resultados solicitados:

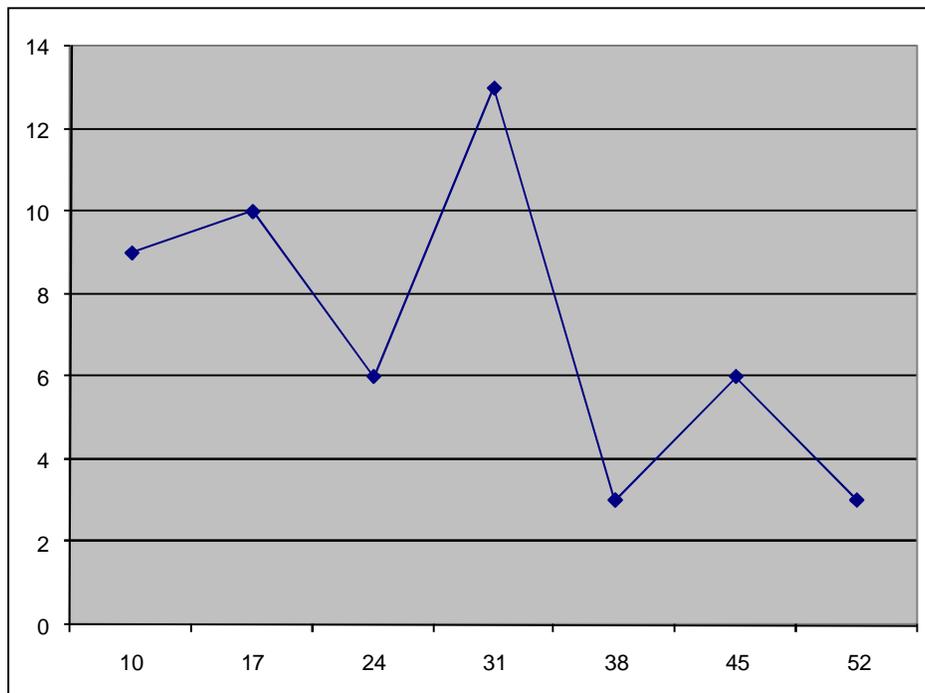
- Media $\bar{X} = 1347/50 = 26.94$
- Mediana $Me = [(20.5 + (25-19)/6)]7 = 27.5$
- Moda $Mo = 27.5 + [7/(7+10)]7 = 30.38$
- Varianza $S^2 = 8044.82 / (50-1) = 164.18$
- Desviación Estándar $S = 164.18 = 12.81$
- Coeficiente de Variación $CV = 12.81 / 26.94 = 0.47$
- 1er Cuartil $Q_1 = 13.5 + [((50/4)-1)/10]7 = 15.95$
- 6to Decil $D_6 = 27.5 + [((50/10)-6)/13]7 = 30.19$
- 71vo Percentil $P_{71} = 27.5 + [((50/100)-71)/13]7 = 33.15$
- Sesgo $a = 33761.07 / (12.81)^3 \times 50 = 0.32$
- Curtosis $g = [2709085.89 / (12.81)^4 \times 50] - 3 = -0.98$

4) Presentación de los Datos

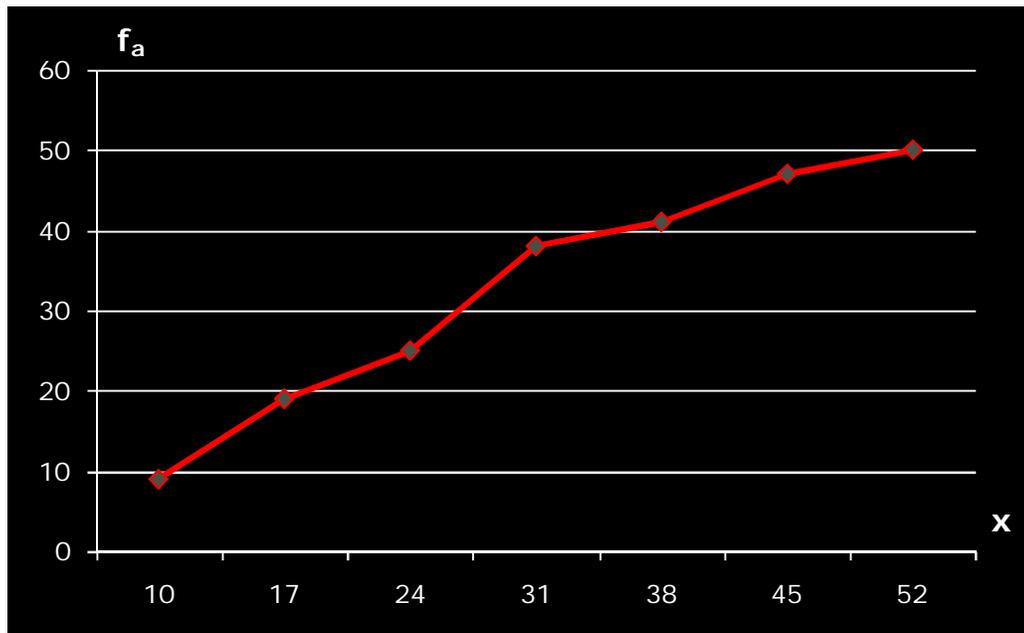
Gráfica de Barras



Polígono de Frecuencias



Ojiva Ascendente



5) Análisis de los Resultados y Conclusiones

Analizando los resultados obtenidos, se puede concluir que:

- De acuerdo a las medidas de tendencia central, el valor representativo de los datos está entre 26.9 y 30.3.
- De acuerdo a las medidas de dispersión, los datos están casi en 13 unidades dispersos alrededor de la media (es decir la mayoría de los datos está entre 14.13 y 40.75).
- Según las medidas de posición la cuarta parte de los datos están antes del valor 15.95 y casi las tres cuartas partes (71%) están antes del 33.15.
- Según lo visto en las gráficas, el sesgo y la curtosis, los datos están un poco sesgados a la derecha (es decir, hay un poco más de datos a la derecha de la media que a la izquierda, porque g fue positiva) y la curva es baja (el conjunto de datos es heterogéneo, porque g fue negativa).