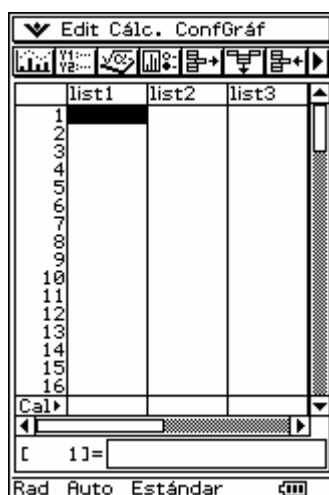


9. ESTADÍSTICA UNIDIMENSIONAL.













ARRANCAR LA APLICACIÓN ESTADÍSTICA

Para arrancar la aplicación Estadística hay que pulsar **Estadística** en el menú de aplicaciones; con esto se muestra la ventana del editor de listas



A continuación explicamos las operaciones que se pueden realizar usando los menús y los botones de la ventana anterior.

Abrir una lista existente	Edit - Abrir Lista	
Cerrar la lista seleccionada actualmente	Edit - Cerrar Lista	
Saltar a la línea 1 de la lista actual	Edit - Ir al principio	
Saltar a la línea siguiente de la última línea de la lista actual	Edit - Ir al final	
Ordenar los datos de una lista en orden ascendente	Edit - Ordenar (Asc.)	
Ordenar los datos de una lista en orden descendente	Edit - Ordenar (Desc.)	
Borrar una celda	Edit - Borrar - Celda	
Borrar todos los datos de una lista	Edit - Borrar - Columna	

Borrar una lista de la memoria	Edit - Borrar - Variable lista	
Insertar una celda en una lista	Edit - Insertar celda	
Convertir una expresión matemática a un valor		
Dibujar un gráfico estadístico		
Ver la ventana del editor de gráficos		
Aplicación Principal		
Ver el cuadro de diálogo de la ventana de visualización	▼ - Preferencias - Ventana vis.	
Ver el administrador de variables	▼ - Preferencias - Adm. de variable	
Cuadro de diálogo de configuración de gráficos estadísticos	ConfGráf . Opciones...	
Ver dos columnas en la ventana del editor de listas		
Ver tres columnas en la ventana del editor de listas		
Ver cuatro columnas en la ventana del editor de listas		

EL EDITOR DE LISTAS

El editor de listas es una herramienta para la creación y el mantenimiento de listas.

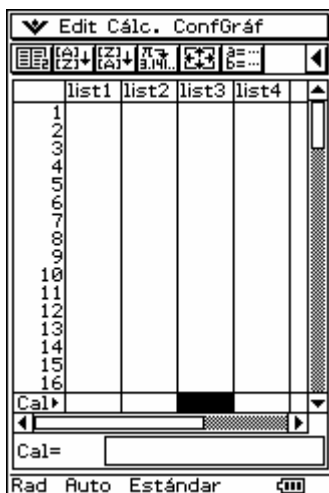
Una lista es un tipo de matriz que permite manipular múltiples datos como un grupo. Una lista tiene una columna y puede contener hasta 9999 filas. En la ventana del editor de listas pueden verse hasta 99 listas.

Las operaciones con listas se realizan usando la ventana del editor de listas, que aparece siempre que se arranca la aplicación Estadística.

Las listas son tratadas como variables y, al igual que las variables, son almacenadas en una carpeta en la memoria y pueden manipularse usando el

administrador de variables. Si se borra una lista de la pantalla, dicha lista todavía existe en la memoria como una variable y puede ser recuperada cuando sea necesario.

El nombre de lista aparece en la parte superior de cada lista. Los nombres de las variables de tipo lista pueden usarse dentro de las fórmulas de cálculo, como cualquier otro nombre de variable. La ventana del editor de listas inicial por defecto muestra seis listas (columnas) llamadas list1 a list6.



Se puede cambiar el nombre por defecto por el nombre que se desee. Para ello:

(1) En la ventana del editor de listas, toque la celda de nombre de lista en la parte superior de la lista a la que desea cambiar el nombre. Esto selecciona la celda de nombre de lista.

(2) Introduzca hasta ocho bytes para el nombre de lista deseado, y luego presione **EXE**

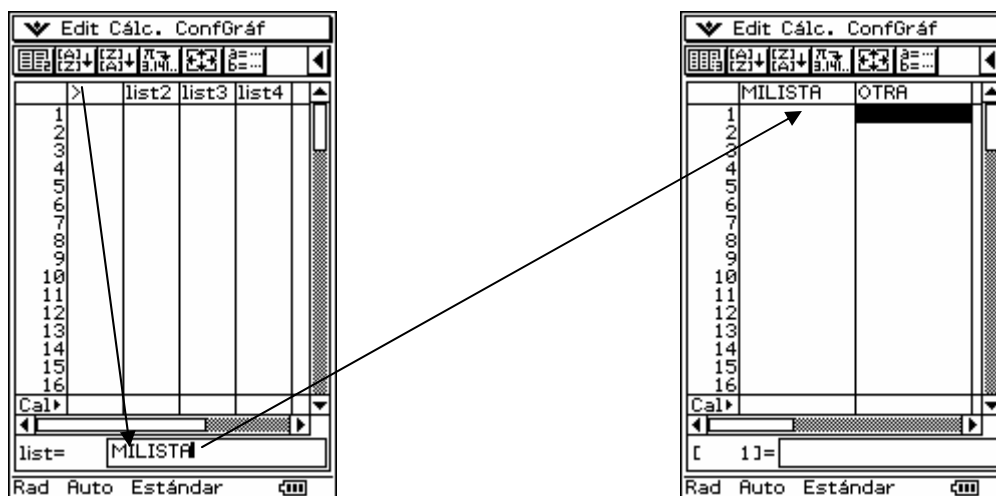
- No puede utilizar ninguna palabra reservada de la ClassPad como nombre de variable de tipo lista. Tampoco puede especificar un nombre de lista que ya esté siendo utilizado por otra lista.

Consejos:

- Si introduce un nombre de lista ya utilizado por otra lista, al tocar **EXE** se muestra el contenido de esa lista. Los datos de la lista existente reemplazan a los datos que ha introducido en la ventana del editor de listas.

- Si introduce un nombre de lista sin especificar una carpeta, el nombre de variable se almacena en la carpeta actual. Para almacenar el nombre de variable en otra carpeta, especifique el nombre de carpeta junto con el nombre de lista. Para

almacenar el nombre de variable de una lista llamada “a” en una carpeta llamada “abc”, por ejemplo, introduzca lo siguiente para el nombre de lista: abc\a.



Para abrir una lista existente, la manera más fácil es introducir el nombre de la lista en la celda de nombre de lista de una columna.

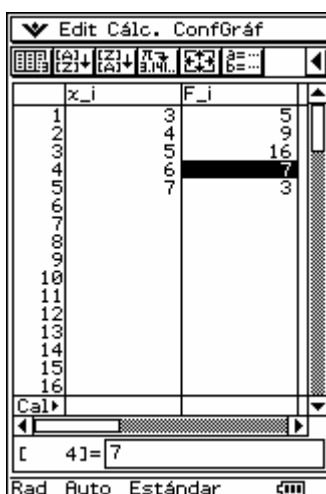
Consideremos el siguiente problema:

A la pregunta: *¿Cuántas personas forman tu hogar familiar?*, 40 personas respondieron esto:

5,5,4,7,4,3,5,5,3,4,6,4,6,5,6,4,6,5,5,5,4,5,7,6,5,5,4,3,5,3,5,6,7,4,5,4,3,5,6

Vamos a crear una tabla de datos y frecuencias.

En list1 vamos a colocar los valores de la variable, con lo cual vamos a renombrar por x_i ; en list2 vamos a colocar la frecuencia absoluta de cada valor de la variable y la llamamos F_i .



Una vez concluida la introducción de datos, va a ser la máquina la que realice el trabajo; nosotros nos limitaremos a reflexionar sobre lo que hay que hacer, lo que significa cada cosa y a dar órdenes para la obtención de las demás columnas. Ahora bastará con situarse encima de la leyenda "list3" y pensar en lo que vamos a hacer. En este caso vamos a generar la columna de las frecuencias absolutas acumuladas de cada valor de la variable, que es el número total de individuos para los que la variable toma valores menores o iguales que ese valor. A esa lista la llamamos F_{Ai} . Para ello, procedemos así:

En la ventana del editor de listas, seleccionamos la celda "Cal" de la lista donde queremos introducir los resultados de cálculo; en nuestro caso es la lista que hemos llamado F_{Ai} .

En el cuadro "Cal=", introducimos la fórmula (utilizando el teclado virtual **cat**, o directamente) **cuml(F_i)**; presionando **EXE** los valores aparecerán en la nueva lista.

z_i	F_i	F_{Ai}
1	3	5
2	4	9
3	5	16
4	6	30
5	7	37
6	7	40
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

Cal= cuml(F_i)

Por ejemplo, la tercera fila indica que el dato 5 aparece con una frecuencia acumulada de 30, es decir, hay 30 personas que tienen 5 o menos miembros en su familia.

En la cuarta lista vamos a crear la columna de frecuencias relativas de cada valor de la variable, que es el cociente de dividir su frecuencia absoluta por el número total de individuos, o sea $F_i/40$. A esa columna le llamamos f_i

F_Ai	f_i	list5
1	5	0.125
2	14	0.225
3	30	0.4
4	37	0.175
5	40	0.075
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

Cal= F_i/40

La primera fila, por ejemplo, indica que el dato 3 aparece con una frecuencia relativa de 0.125, es decir, 0.125 individuos de dada 1 tiene 3 miembros en su familia. Pero puede resultar más real expresar lo anterior en porcentajes, así que en lista 5 vamos a generar la columna de las frecuencias relativas pero expresadas en porcentajes; con el teclado virtual o directamente escribimos en la celda de cal **percent(f_i)**

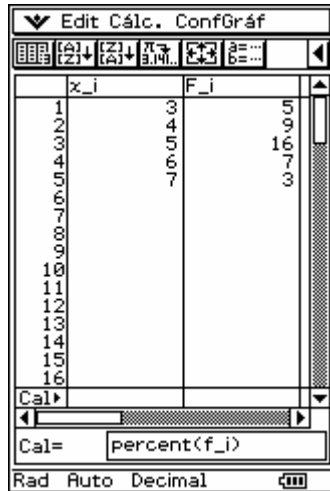
F_Ai	f_i	list5	
1	5	0.125	12.5
2	14	0.225	22.5
3	30	0.4	40
4	37	0.175	17.5
5	40	0.075	7.5
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			

Cal= percent(f_i)

Por ejemplo, en la fila 1 el dato 3 aparece con una frecuencia relativa del 12.5%

CÁLCULOS ESTADÍSTICOS

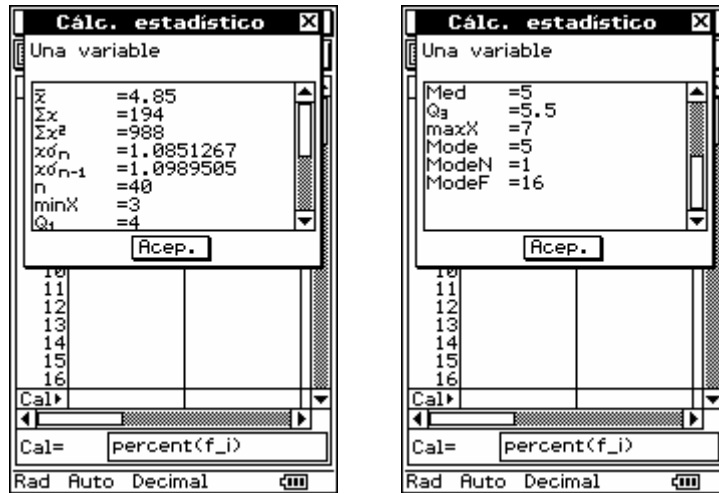
Vamos a obtener los valores de las diferentes medidas de centralización y de dispersión. Observemos que realmente solo nos interesan las dos primeras listas, o sea, los valores de la variable y los de su frecuencia absoluta



Para ver los resultados procedemos así:

- En la barra de menús, tocamos [Cálculo] y luego [Una variable].
- En el cuadro de diálogo que aparece especificamos el nombre [ListaX], que en nuestro caso es main/x_i y seleccionamos la opción [Frec] asignando main/F_i. (La palabra main aparece para indicarnos que la variable está guardada en la carpeta principal).
- Pulsamos OK





Aparece el cuadro de diálogo de cálculos estadísticos con los resultados que vamos a comentar a continuación.

- La media \bar{x} de los miembros de la familia es de 4.85 miembros.
- $\sum x$. Si sumamos todos los miembros de las familias nos daría 194
- $\sum x^2$. La suma de los cuadrados de los miembros de la familia es 988.
- $x\sigma_n$. La desviación típica del número de miembros de la familia tomados como población es 1.08512672.
- $x\sigma_{n-1}$ La desviación típica del número de miembros de la familia tomados como muestra es 1.09895054.
- n. El número de familias es de 40.
- MinX. El menor número de miembros en una familia es 3.
- Q_1 . Primer cuartil. Las familias de 4 miembros tiene el 25% de la distribución con un número de miembros menor o igual que ellas.
- Med. Mediana. Una familia de 5 miembros deja a cada lado el mismo número de datos.
- Q_3 . Tercer cuartil. Las familias de 5.5 miembros tiene el 75% de la distribución con un número de miembros menor o igual que ellas.
- maxX. El mayor número de miembros de una familia es 7.
- Mode. Moda. El número de miembros más frecuentes en una familia es 5.
- ModeN. Número de elementos iguales a la moda de los datos, que en nuestro caso es 1.
- ModeF. Frecuencia de la moda de los datos, o sea, 5 tiene una frecuencia absoluta de 16.

REPRESENTACIÓN GRÁFICA DE DATOS ESTADÍSTICOS DE UNA VARIABLE.

Antes de dibujar un gráfico estadístico, primero es necesario configurar su “configuración de gráfico estadístico” usando el menú [ConfGráf].

La configuración de gráficos estadísticos le permite configurar los parámetros para controlar el tipo de gráfico, las listas que contienen los datos de un gráfico, el tipo de marcadores que se usarán y otras opciones. Se pueden almacenar hasta nueve configuraciones de gráfico estadístico en la memoria, llamadas Gráfico Estadís. 1, Gráfico Estadís. 2, y así sucesivamente, para ser recuperadas posteriormente

Al pulsar [ConfGráf.] en la barra de menús de la ventana del editor de listas, aparece un menú como el que sigue

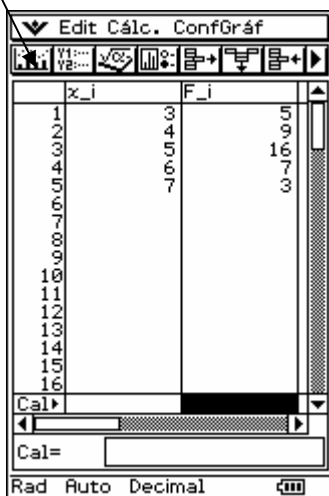


Si ahora pulsamos [Opciones...] aparecerá el cuadro de diálogo de configuración de los gráficos estadísticos.

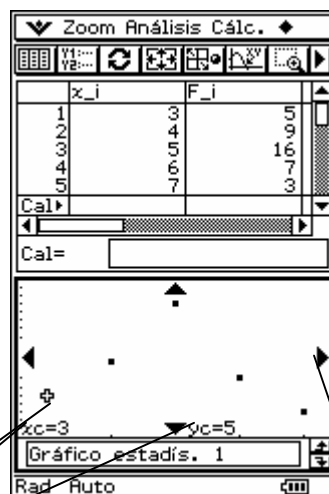
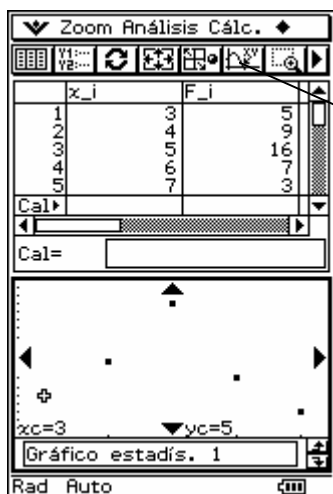


Hay una lengüeta para cada configuración de gráfico estadístico, llamada Gráfico Estadís.1 a Gráfico Estadíst.9. Tocando los botones de selección se puede ver la configuración de gráfico estadístico que se quiere cambiar, y pulsando [Def.] se aplican las opciones elegidas. Así, en la pantalla anterior tenemos un gráfico de **Nube de Puntos (Disper.)** que queremos se dibuje (Dibujo: On); le hemos indicado que los datos de la variable están en x_i y sus frecuencias en F_i , y con una marca (Marca) de punto grande (Punto Gr.).

Ahora pulsamos aquí

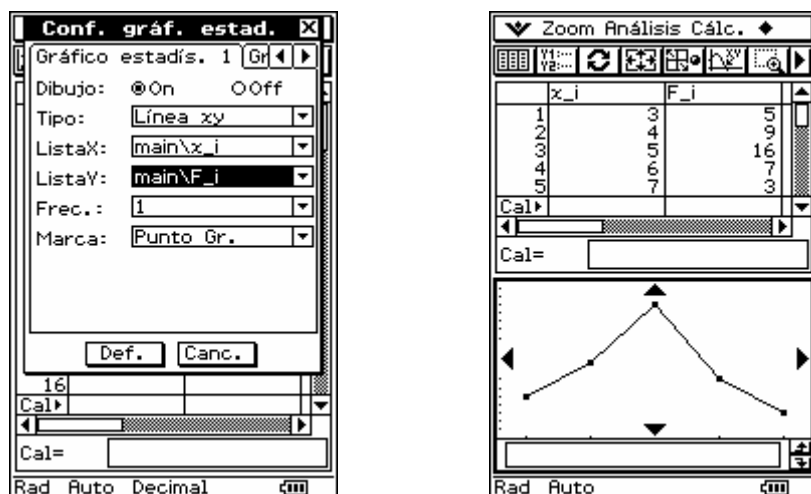


y nos aparecerá el gráfico

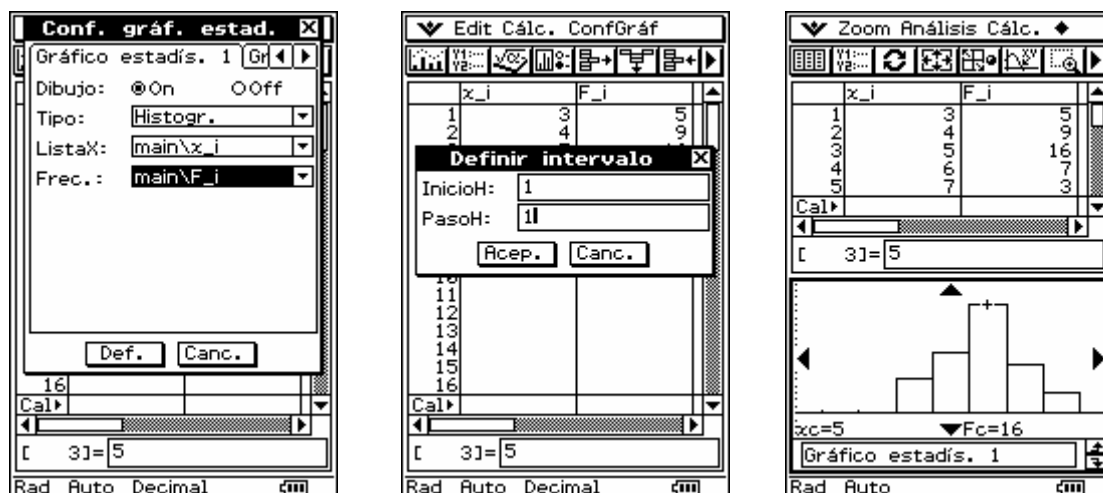


Aparece marcado el punto y sus valores (mover con flechas).

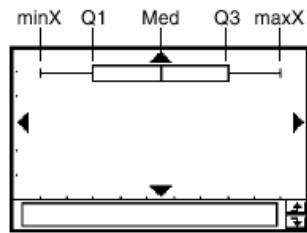
Ahora vamos a dibujar el **polígono de frecuencias absolutas (Línea xy)**



Para dibujar el **diagrama de barras (Histogr.)**



Los **diagramas de cajas Med(MedBox)**, o **diagramas de la media en recuadro** o, a veces más conocidos, **diagramas de cajas y bigotes**, son representaciones gráficas de una distribución estadística unidimensional que reflejan directamente 5 parámetros, e indirectamente el rango. También dan una idea de la simetría, el sesgo y la dispersión de los datos de la distribución permitiendo contrastar conjuntos de datos diferentes de una misma variable.



En nuestro caso

Conf. gráf. estad.

Gráfico estadís. 1 Gr

Dibujo: On Off

Tipo: CajaMed

ListaX: main\x_i

Frec.: main\F_i

Def. Canc.

16

Cal

[2]=4

Rad Auto Decimal

Zoom Análisis Cál.

x_i	F_i
1	3
2	4
3	16
4	7
5	3

Cal

[2]=4

Rad Auto

Zoom Análisis Cál.

x_i	F_i
1	3
2	4
3	16
4	7
5	3

Cal

[2]=4

minX=3

Gráfico estadís. 1

Rad Auto

Zoom Análisis Cál.

x_i	F_i
1	3
2	4
3	16
4	7
5	3

Cal

[2]=4

Q₁=4

Gráfico estadís. 1

Rad Auto

Zoom Análisis Cál.

x_i	F_i
1	3
2	4
3	16
4	7
5	3

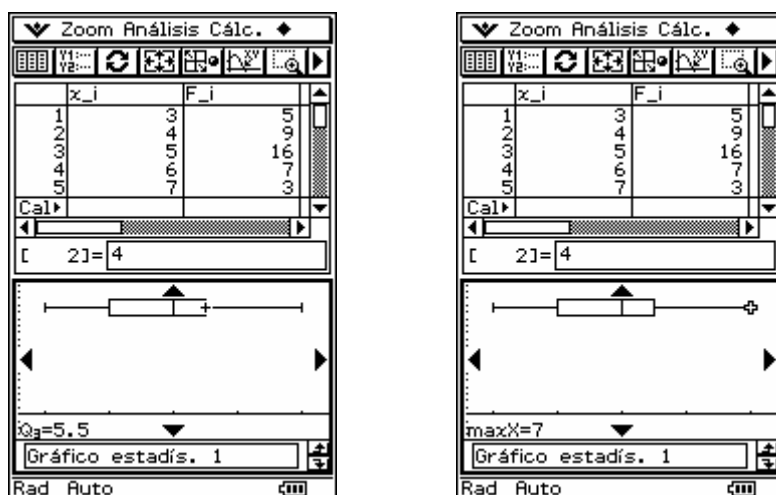
Cal

[2]=4

Med=5

Gráfico estadís. 1

Rad Auto



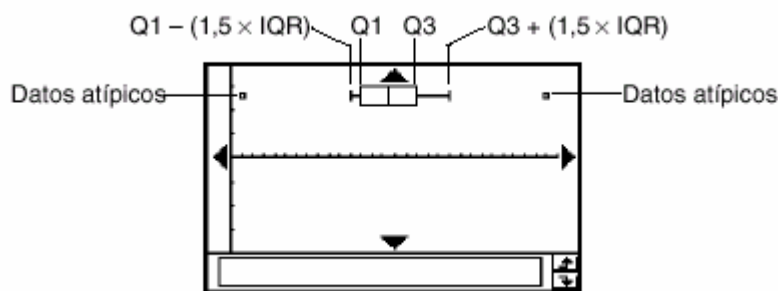
Existe una variación del diagrama anterior llamado **diagrama de cajas modificado (ModBox)**

El diagrama de cajas modificado (ModBox) utiliza la “regla $1,5 \times IQR$ ” para definir los límites para los “datos atípicos (outliers)”, que son valores que no siguen el mismo patrón que el resto de los datos.

La sigla “IQR” significa “rango intercuartil”, que es la longitud de la caja del gráfico CajaMed descrito antes. El rango intercuartil IQR se calcula de la manera siguiente:

$$IQR = Q3 \text{ (tercer cuartil)} - Q1 \text{ (primer cuartil)}.$$

Cualquier valor que esté $1,5 \times IQR$ por debajo del primer cuartil o $1,5 \times IQR$ por encima del tercer cuartil, es un dato atípico. Los datos atípicos se indican con círculos.



En nuestro ejemplo no hay datos atípicos, con lo que si intentamos dibujar este diagrama nos quedaría el mismo que el de cajas y bigotes.

ACTIVIDADES

1. A unas oposiciones de auxiliar administrativo se han presentado un total de 1000 aspirantes. Éstos han realizado un primer ejercicio que se ha calificado en una escala entera de 0 a 100 puntos. Los resultados se muestran en la siguiente tabla:

Intervalo	x	f
[40,45)	42.5	1
[45,50)	47.5	2
[50,55)	52.5	10
[55,60)	57.5	46
[60,65)	62.5	120
[65,70)	67.5	202
[70,75)	72.5	242
[75,80)	77.5	205
[80,85)	82.5	113
[85,90)	87.5	45
[90,95)	92.5	12
[95,100)	97.5	2

Construir la tabla de frecuencias tal como se ha hecho a lo largo del tema y calcular los valores de las medidas de centralización y de dispersión.

Ensayar todos los gráficos de una distribución unidimensional con estos datos

2. Una zapatería ha vendido en un día 25 pares de zapatos de caballero de las siguientes tallas:

40	41	40	42	40
41	43	43	45	44
45	42	40	41	45
45	44	40	44	41
42	43	43	46	41

Construya una tabla de frecuencias y halle la media aritmética, la moda, la mediana, la desviación media, la varianza y la desviación típica.

Represente los datos en un diagrama adecuado.

3. El coeficiente de variación de una distribución, CV , es el cociente entre la desviación típica y la media aritmética:

La siguiente tabla muestra las calificaciones obtenidas por Paco y Eva en diez controles de matemáticas:

Notas de Paco	4	5	5	4	6	7	8	9	3	9
Notas de Eva	6	6	5	6	7	7	6	5	7	5

- a) Halle sus medias y desviaciones típicas. ¿Quién es más regular?
- b) Construya los polígonos de frecuencias de ambas distribuciones. ¿Cuál de ellas se ajusta a la de una distribución normal?
4. Las puntuaciones obtenidas por 30 alumnos de 4º de ESO en una prueba de inteligencia han sido

100	102	98	95	92	105	121	110	84	87
94	99	98	112	123	145	116	93	89	85
86	97	114	127	103	104	135	128	109	110

Agrupe los datos en intervalos de clase y construya su tabla de frecuencias absolutas, relativas y porcentuales.

Realice algunos gráficos adecuados.

5. De la distribución de una variable sabemos que el recorrido es 60, y la distribución está dividida en seis intervalos de amplitud constante. Las frecuencias de cada intervalo son, por orden, 7, 11, 15, 10, 5, 2. La media aritmética es $35,2$. Una vez hallada la distribución (su tabla de frecuencias), represéntela gráficamente, y hallar los parámetros de centralización y dispersión.