

15

Estadística

Los recuentos estadísticos se remontan al origen de la historia. Existen documentos antiquísimos (papiros egipcios, tablillas de arcilla babilonias) donde hay constancia de censos de población y de bienes públicos.



© Grupo Anaya, S.A. Material fotocopiable autorizado.



Hace unos 2000 años, César Augusto ordenó que se realizara en todo el Imperio Romano una amplísima encuesta sobre habitantes, soldados, recursos de todo tipo y rentas públicas. Este es un antiguo antecedente de lo que actualmente se llama *estadística*. Pero no fue el primero.

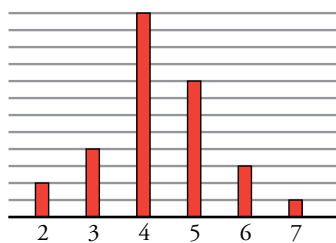
En la Biblia ya se recogen censos realizados por judíos. También los griegos, chinos e indios antiguos realizaron recuentos y encuestas.

Sin embargo, no es hasta el siglo XVII, en Europa, cuando la Estadística toma cuerpo como ciencia.

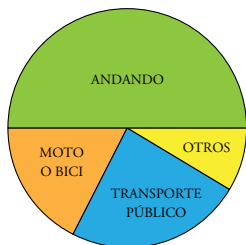
Nombre y apellidos: Fecha:

1 Proceso para realizar un estudio estadístico

¿Cuántas personas vivís en casa?



¿Cómo acudes a tu centro?



Ejemplos

- La marca del teléfono móvil es una variable cualitativa.
- El número de mensajes enviados es una variable cuantitativa.

A cada alumno de una clase se le han hecho las siguientes preguntas:

- ¿Cuántas personas vivís en tu casa?
- ¿Cómo acudes a tu centro de estudios?

Las respuestas dadas a cada una de estas preguntas es una **distribución estadística**.

Variable estadística

En cada una de las distribuciones anteriores se analiza una variable:

La primera variable es el *número de personas que viven en tu casa*. Es una *variable numérica (cuantitativa)*, pues los valores que puede tomar son números: 2, 3, 4, 5, 6, 7, ...

La segunda es *cómo acudes al centro docente*. El resultado puede ser: ANDANDO, TRANSPORTE PÚBLICO, MOTO-BICI, OTROS. Este tipo de variables se llama *cualitativa*.

Una **variable** se llama **cuantitativa** cuando toma valores numéricos, y **cualitativa**, cuando toma valores no numéricos.

Población y muestra


Los 30 alumnos de la clase anterior es el colectivo sobre el que estamos trabajando. A este colectivo objeto de estudio se le llama *población*.

Se llama **población** al conjunto de todos los elementos que se estudian. Cada uno de sus componentes se denomina **individuo**.

En ocasiones, la población es demasiado grande, por lo que conviene estudiar solo unos cuantos de sus individuos. A esta selección se le llama *muestra*.

Se denomina **muestra** a un subconjunto de la población elegido para realizar el estudio estadístico.

Piensa y practica

- Indica si cada una de las siguientes variables estadísticas es cuantitativa o cualitativa:
 - Equipo de fútbol preferido.
 - Edad.
 - Lugar de nacimiento.
 - Número de asignaturas suspendidas en la primera evaluación.
 - Asignaturas aprobadas en la segunda evaluación.
 - Número de viviendas que hay en tu calle.
 - Tiempo que tardas en correr los 100 m lisos.
-  Reconoce, en cada una de las siguientes situaciones, la población, la muestra y los individuos.
 - Una fábrica de bombillas quiere hacer un control de calidad. Para ello, analiza una bombilla de cada caja de 1 000.
 - Una farmacéutica visita a un médico de cada hospital para enseñarle sus nuevos productos.
 - Un agricultor recoge una naranja de cada uno de los árboles de su naranjal para comprobar la cantidad de zumo que puede obtenerse.
 - Tomo una golosina de cada cubo de la tienda.

158

Nombre y apellidos: Fecha:

Ejemplo

En este conjunto de datos:

3 0 2 5 2

3 1 2 1 1

0 5 2 1 3

Comprueba que:

$$f(0) = 2 \quad f(1) = 4$$

$$f(2) = 4 \quad f(3) = 3$$

$$f(4) = 0 \quad f(5) = 2$$

Frecuencia

En la distribución *cómo acudes a tu centro de enseñanza*, el número de individuos que utilizan el TRANSPORTE PÚBLICO es 7.

Lo expresamos así: $f(\text{TRANSPORTE PÚBLICO}) = 7$

Y se lee así: la frecuencia de TRANSPORTE PÚBLICO es 7.

El número de individuos correspondiente a un valor de la variable se llama **frecuencia** de ese valor.

Tablas de frecuencias

Una vez recogidos los datos correspondientes a una experiencia estadística, hay que tabularlos; es decir, hay que confeccionar con ellos una tabla en la que aparezcan, ordenadamente:

— Los *valores de la variable* que se está estudiando.

— El *número de individuos* de cada valor; es decir, su frecuencia.

Para hacer el recuento, se leen los datos uno a uno y se marca una señal en el correspondiente valor. Si las señales se agrupan de cinco en cinco, es más fácil contarlas.

Fijémonos, por ejemplo, en la distribución *número de personas que viven en tu casa*:

DATOS	RECuento	TABLA DE FRECUENCIAS																
4 5 4 7 6	2 → 2	<table border="1"> <thead> <tr> <th>VALORES</th> <th>FRECUENCIA</th> </tr> </thead> <tbody> <tr> <td>2</td> <td>2</td> </tr> <tr> <td>3</td> <td>4</td> </tr> <tr> <td>4</td> <td>12</td> </tr> <tr> <td>5</td> <td>8</td> </tr> <tr> <td>6</td> <td>3</td> </tr> <tr> <td>7</td> <td>1</td> </tr> <tr> <td></td> <td>30</td> </tr> </tbody> </table>	VALORES	FRECUENCIA	2	2	3	4	4	12	5	8	6	3	7	1		30
VALORES	FRECUENCIA																	
2	2																	
3	4																	
4	12																	
5	8																	
6	3																	
7	1																	
	30																	
2 4 6 3 5	3 → 4																	
4 6 5 5 3	4 → 12																	
4 4 4 5 3	5 → 8																	
2 5 4 4 5	6 → 3																	
4 5 4 3 4	7 → 1																	

En la **tabla de frecuencias**, cada valor tiene emparejada su frecuencia.

Piensa y practica

1. Se pregunta a 40 estudiantes qué prefieren hacer en su tiempo libre: deporte (D), leer (L), ver la tele (T), salir con amigos (S), jugar con videojuegos (V). Los resultados son:

S S D S V	S L S D T
L V S S L	D D S V L
D S S V S	D V D D V
V T S S D	L D T T L

Confecciona una tabla de frecuencias con los resultados obtenidos.

2. Se ha contabilizado el número de libros leídos en las vacaciones de verano por los 30 estudiantes de un curso. Estos son los resultados:

1 3 1 0 4	4 1 0 2 3
0 1 1 2 3	2 3 1 1 6
1 1 2 1 2	0 0 2 1 4

Realiza la correspondiente tabla de frecuencias.

3. Construye la tabla de frecuencias del ejemplo del margen de esta página.

3 Gráficos estadísticos

Las representaciones gráficas sirven para captar, de un solo golpe de vista, las características más sobresalientes de una distribución de datos.

Hay muchos tipos de representaciones gráficas. Vamos a ver las de uso más frecuente.

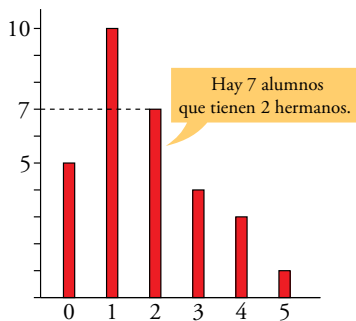
Diagrama de barras

Importante

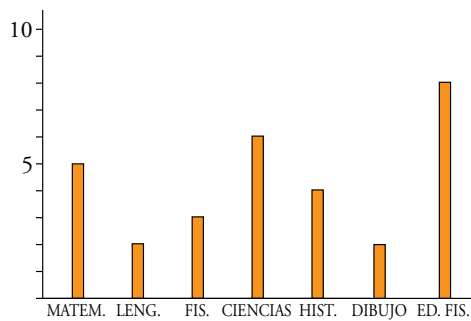
Las tablas de datos estadísticos y las representaciones gráficas son lenguajes, formas de dar información sumamente eficaces, pues lo que se dice mediante ellas “entra por los ojos”.

Procura mirar con atención las gráficas y las tablas para captar todo lo que te “dicen”.

Población: Alumnos de una clase
Variable: Número de hermanos



Población: Alumnos de una clase
Variable: Asignatura preferida



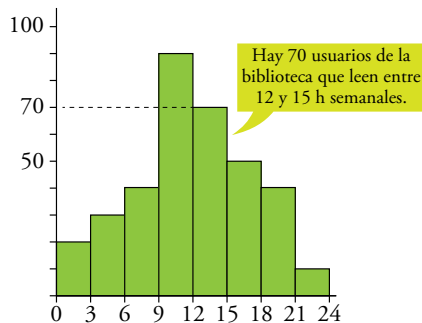
El **diagrama de barras** está formado por barras finas. Sirve para representar tablas de frecuencias de variables cualitativas, o bien cuantitativas que tomen pocos valores. Las alturas de las barras son proporcionales a las frecuencias correspondientes.

En la web

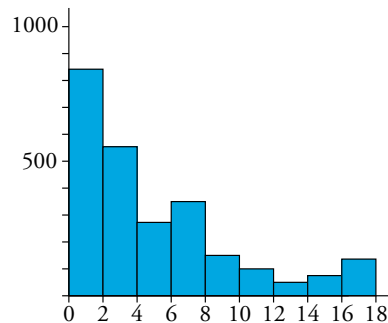
Interpreta un diagrama de barras.

Histograma

Población: Usuarios de una biblioteca
Variable: Tiempo (en h) de lectura semanal



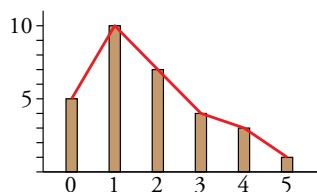
Población: Árboles de un parque
Variable: Altura (en m)



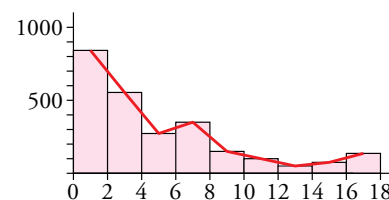
El **histograma** está formado por rectángulos anchos que se adosan unos a otros. Sirve para representar variables cuantitativas que tomen muchos valores diferentes. Las áreas de las barras son proporcionales a las frecuencias correspondientes.

Polígono de frecuencias

Población: Alumnos de una clase
Variable: Número de hermanos



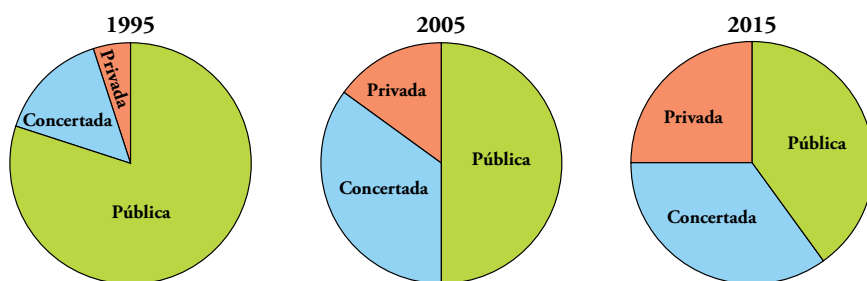
Población: Árboles de un parque
Variable: Altura (en m)



El **polígono de frecuencias** se utiliza para representar variables cuantitativas. Se construye uniendo los extremos de las barras o los puntos medios de los rectángulos de un histograma.

Diagrama de sectores

La evolución de la educación, según sea pública, concertada o privada, en una localidad se ve reflejada en los siguientes diagramas:



Utilidad

Los diagramas de sectores son muy útiles para ver la evolución de una misma variable.

Por ejemplo, podemos ver que en la localidad a la que se refieren los gráficos de la derecha el tipo de educación ha ido cambiando de pública a privada y concertada.

En la web

Interpreta un diagrama de sectores.

El **diagrama de sectores** sirve para representar variables de cualquier tipo. Cada sector representa un valor de la variable. El ángulo de cada sector es proporcional a la frecuencia correspondiente.

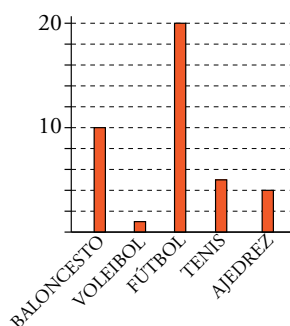
Piensa y practica

En la web

Practica con gráficos estadísticos.

1. Este diagrama de barras representa los deportes preferidos por los alumnos de una clase.

- ¿Cuál es el que más gusta? ¿Y el que menos?
- ¿Cuántos prefieren el tenis?
- ¿Cuántos alumnos hay en clase?



2. Estas son las notas en matemáticas de un grupo de alumnos en las tres evaluaciones del año:



Explica cómo han evolucionado.

3. Representa con un diagrama de barras los datos del ejercicio 5 del epígrafe anterior.

4 Parámetros estadísticos

Los primeros parámetros que vamos a ver aquí (media, mediana y moda) son de **centralización**, pues designan valores en torno a los cuales se distribuyen los datos.

Media solo para variables cuantitativas

Es claro que solo es posible obtener la media de una variable cuantitativa. Imagina tener que hallar la media de esta distribución:

A, A, A, B, B, A, C, A, A, B, C

Media

Cuatro amigos han ido a pescar. Están son las truchas capturadas por cada uno:

2, 1, 4, 5

La media de las truchas pescadas es: $\frac{2+1+4+5}{4} = \frac{12}{4} = 3$

Esto significa que si juntan todos los peces y los reparten por igual, a cada uno le tocan 3 truchas.

La **media** de varias cantidades es la suma de todas las cantidades dividida por el número de ellas. También se conoce como promedio.

Mediana

Las alturas de tres jirafas son 4 m, 5 m y 6 m. Al colocarlas ordenadas según sus alturas, decimos que la de en medio es la mediana.

Se llama **mediana** de un conjunto de datos numéricos al que ocupa el valor central una vez ordenados.

Para calcularla, ordenamos las cantidades de menor a mayor y elegimos la de en medio. Si hay un número par de datos, la mediana es el promedio de los dos valores centrales.

Ejemplos

- 4, 4, 7, 7, 8, 8, 10 → MEDIANA = 7
- 1, 5, 6, 6, 9, 10 → MEDIANA = 6
- 1, 5, 6, 9, 9, 10 → MEDIANA = $\frac{6+9}{2} = 7,5$

Moda

Cuando algo se lleva mucho, se dice que está de moda.

La moda es lo que más se lleva.

En una distribución estadística, se llama **moda** al dato que aparece con más frecuencia. Se puede obtener la moda de una variable cuantitativa o cualitativa.

Una distribución puede tener varias modas. Si hay dos datos que tienen la frecuencia más alta, la distribución será bimodal; si son tres datos, la distribución es trimodal, y así sucesivamente.

Ejemplos

- 1, 2, 2, 2, 3, 4, 5, 5 → La moda es 2.
- 1, 1, 2, 3, 4, 4, 5, 6 → Las modas son 1 y 4. Distribución bimodal.

Hay distintos parámetros que sirven para cuantificar cómo de dispersos están los datos. Se les llama **parámetros de dispersión**. Pero aquí solo vamos a estudiar uno de ellos: *el recorrido*.

Ejemplos

- 1, 2, 2, 2, 3, 4, 5, 5
El recorrido es $5 - 1 = 4$.
- 0, 0, 0, 0, 0, 1, 1, 1
El recorrido es $1 - 0 = 1$.
- 5, 5, 5, 5, 5, 5, 5, 5
El recorrido es $5 - 5 = 0$.

Recorrido o rango

El recorrido nos muestra la longitud del tramo en el que se encuentran los datos.

Se denomina **recorrido** de un conjunto de datos a la diferencia entre el dato mayor y el menor.

$$\text{RECORRIDO} = \text{Valor máximo} - \text{Valor mínimo}$$

Ejercicios resueltos

1. Hallar la media, la mediana, la moda y el recorrido de esta distribución:

4, 6, 8, 8, 9, 10, 10, 10

• $\text{MEDIA} = \frac{4+6+8+8+9+10+10+10}{8} = 8,125$

- Los valores están ordenados. Los dos datos centrales son 8 y 9. Por tanto, la mediana es el promedio de ellos; $\text{MEDIANA} = 8,5$.
- El valor que está más veces es el 10; $\text{MODA} = 10$.
- El valor máximo menos el mínimo es $10 - 4 = 6$; $\text{RECORRIDO} = 6$.

2. Hallar la media, la mediana, la moda y el recorrido de esta distribución:

I, V, P, O, V, O, P, P, I, V

Cuando la variables es cualitativa, no se puede hallar ni la media, ni la mediana, ni el recorrido. La moda es P.

$$\text{MODA} = \text{P}$$

3. Hallar la media, la mediana, la moda y el recorrido de esta distribución de notas de un examen:

2 1 5 7 4 3 8 7 9 5
5 6 4 9 2 4 6 7 7 8
4 5 5

• $\text{MEDIA} = \frac{2+1+5+7+4+3+\dots+4+5+5}{8} = 5,35$

- Se ordenan los valores:
1 2 2 3 4 4 4 4 5 5 5 5 5 6 6 7 7 7 7 8 8 9 9
- El valor central corresponde a 5; $\text{MEDIANA} = 5$.
- El valor que está más veces es el 5; $\text{MODA} = 5$.
- El valor máximo menos el mínimo es $9 - 1 = 8$; $\text{RECORRIDO} = 8$.

Piensa y practica

1. En la cola del autobús hay siete personas, cuyas edades son:

15, 12, 23, 30, 71, 55 y 12

Calcula la media, la mediana, la moda y el recorrido de las edades.

2. Las notas del último examen de matemáticas de seis amigos son:

9, 5, 10, 5, 7, 6

Halla la media, la mediana, la moda y el recorrido de las notas.

3. ¿Verdadero o falso?

- a) En una distribución cuya variable es cualitativa no se puede calcular ni la media, ni la mediana, ni el rango.
- b) La media de las edades de Ana, su madre y su abuela es 40. Si la abuela de Ana tiene 60 años, Ana debe tener 20 años.
- c) La mediana de las puntuaciones de once equipos es 40. Si el 5.º clasificado tiene 42 puntos, el 7.º tendrá 38.

Nombre y apellidos: Fecha:

Ejercicios y problemas

Variables estadísticas y frecuencias

- Indica si es cualitativa o cuantitativa cada una de las variables por las que se pregunta.
 - ¿Cuántos hermanos sois en casa?
 - ¿Qué medio de transporte prefieres?
 - ¿Qué deporte prefieres practicar?
 - ¿Qué edad tienes?
- Indica en cada caso la población, la muestra y los individuos:
 - Se quiere estudiar las migraciones anuales de las ballenas del océano índico. Para ello, se colocan radiotransmisores en 30 de estos cetáceos.
 - Se quiere saber qué opinan los mayores de edad sobre las nuevas iniciativas del ayuntamiento. Para ello se entrevista a 100 personas elegidas al azar.
 - Para saber qué opinión se tiene de una nueva sección de una revista, se ha preguntado telefónicamente a 50 lectores.

- Elabora una tabla con las frecuencias correspondientes a este recuento.

Soleado	
Sol y nubes	
Nublado	
Lluvia fina	
Lluvia torrencial	
Nieve	

- Se ha preguntado a 50 personas qué tipo de película prefieren. Las respuestas son: Romántica (RO), Terror (TE), Comedia (CO), Aventuras (AV), *Thriller* (TH), Acción (AC), Otras (O). Dados los resultados, haz un recuento y elabora una tabla de frecuencias:

RO, TE, RO, TH, AC	AV, O, AC, TH, TE
TE, RO, CO, CO, TH	RO, TH, AC, O, AC
AC, AV, TE, O, TH	TE, AV, TH, AC, AC
TE, RO, RO, AV, TH	TH, AC, AC, CO, TE
RO, TE, RO, AV, TH	AC, O, TH, RO, RO

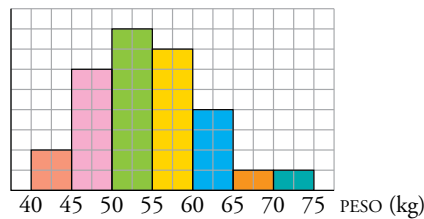
- Se ha pasado una prueba tipo test a 30 estudiantes. Estos son los errores cometidos por cada uno:

2 3 2 1 0	1 2 2 1 3
0 3 2 4 0	1 1 2 0 2
2 3 2 0 0	1 0 4 2 1

Realiza la tabla de frecuencias.

Gráficos estadísticos

- El peso de los alumnos de una clase viene reflejado en el siguiente histograma:



Hay un solo alumno que pesa más de 70 kg.

- ¿De qué color es la barra donde se ubica un alumno de 57 kg?
 - ¿Cuántos alumnos pesan entre 60 kg y 65 kg?
 - ¿Cuántos alumnos pesan más de 50 kg?
 - ¿Cuántos alumnos hay en clase?
- Este diagrama de sectores representa la distribución de los 24 estudiantes de una clase de 1.º de ESO, según se queden o no a comer en el colegio:



- ¿Qué fracción de los alumnos se quedan a comer?
 - ¿Qué porcentaje no se queda nunca?
 - ¿Qué tanto por ciento se queda a veces?
- Dibuja el diagrama de barras y el polígono de frecuencias correspondiente al ejercicio 3.
 - Haz lo mismo para el ejercicio 4.
 - Repítelo para el ejercicio 5.
 - Se ha pasado una encuesta a 40 empresarios sobre el número de continentes que han visitado:

1 1 2 1 2	1 2 4 1 3
2 1 1 2 3	2 2 3 4 1
1 3 1 4 1	1 2 1 1 1
1 1 2 1 1	2 1 4 4 3

- Construye una tabla de frecuencias.
- Representa los resultados en un diagrama de barras.

Parámetros estadísticos

10. Calcula la media, la mediana, la moda y el recorrido de las siguientes distribuciones:

- a) 2, 4, 4, 41, 17, 13, 24
- b) 1, 3, 8, 9, 4, 1, 1, 7, 10, 10
- c) 1, 3, 5, 4, 2, 8, 9, 6, 10, 6
- d) 1, 2, 3, 4, 3, 2, 1

11. Halla la media, la mediana, la moda y el recorrido de los datos recogidos en los anteriores ejercicios 5 y 9.

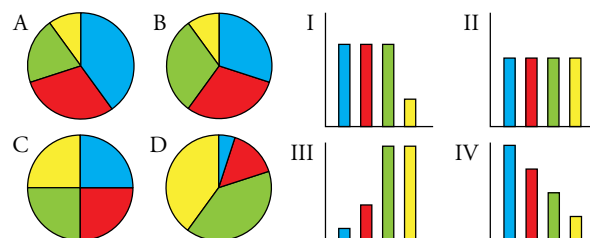
12. Lanzamos un dado 40 veces. Estos son los resultados obtenidos:

3 5 1 2 5	5 3 4 6 2
4 3 6 4 1	6 4 2 6 1
4 3 5 6 2	1 5 6 6 2
4 2 3 2 6	5 4 1 6 1

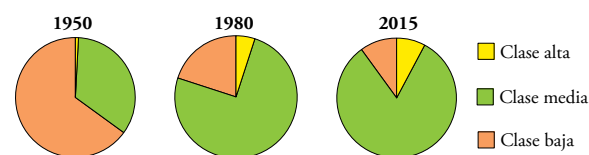
Calcula la media, la mediana, la moda y el recorrido de la distribución.

Resuelve problemas

13. Asocia a cada diagrama de sectores su correspondiente diagrama de barras:



14. La evolución de las clases sociales de un cierto país viene dada por estos tres diagramas de sectores:



- a) Explica cómo han evolucionado las clases sociales en el país a lo largo de estos últimos 65 años.
- b) ¿Crees que es positivo lo que ha ocurrido? Razonalo.

Autoevaluación

1. Indica cuáles son variables cualitativas y cuáles cuantitativas:

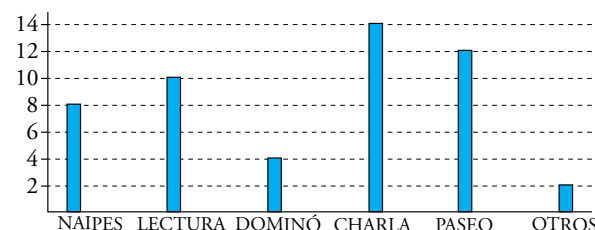
- a) Color de zapatos o zapatillas.
- b) Resultado de un partido en la quiniela (1, X, 2).
- c) Tiempo en recorrer cierta distancia.

2. Los 40 componentes del equipo de tiro con arco realizan una competición. Estos son los resultados del número de dianas que ha conseguido cada uno:

3 2 5 2 0	2 5 3 2 2
2 1 2 3 4	4 3 5 2 1
2 3 2 1 4	5 2 2 3 1
2 3 0 3 0	2 0 2 3 5

- a) Construye una tabla de frecuencias.
- b) Representa los datos en un diagrama de barras.
- c) Calcula la media, la mediana, la moda y el recorrido.

3. Este diagrama de barras muestra lo que más les gusta hacer a un grupo de jubilados en su tiempo libre:



- a) ¿Qué es lo que más prefieren hacer? ¿Y lo que menos?
- b) ¿Cuántos prefieren jugar a los naipes?
- c) ¿Cuántos prefieren dedicarse a la lectura?
- d) ¿Cuántos jubilados han sido encuestados?
- e) ¿Qué porcentaje de ellos prefiere salir de paseo?

© Grupo Anaya, S. A. Material fotocopiable autorizado.

Nombre y apellidos: Fecha: