

ESTALMAT 2007-2008

Sesión 20: Estadística Descriptiva (Analizando y visualizando datos)

Antonio Moreno, Patricia Román

PRIMERA PARTE

ACTIVIDAD 1.

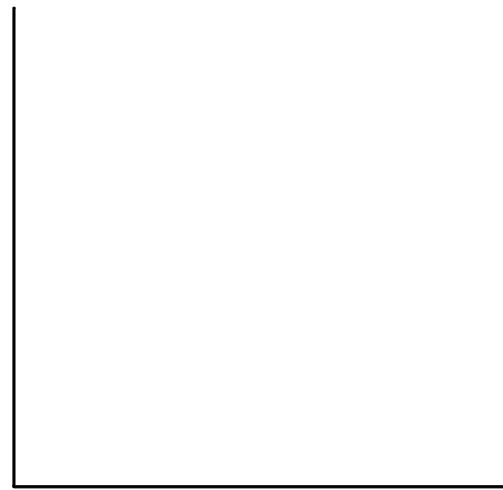
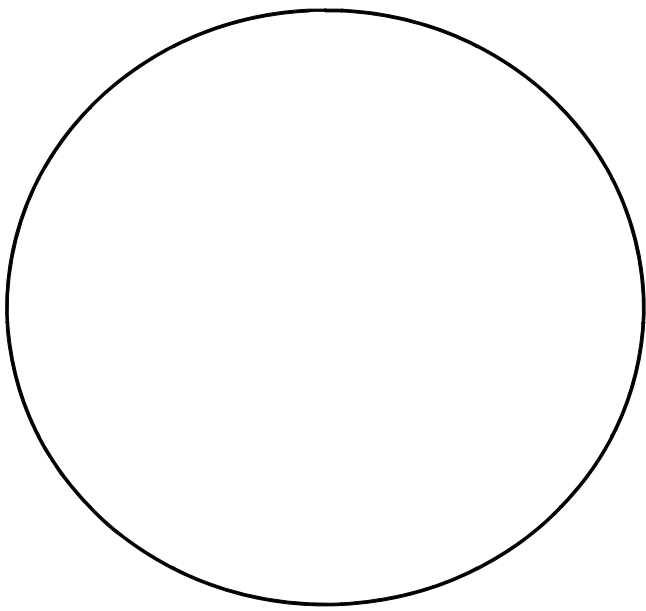
Indica el tipo de variable estadística de cada uno de estos cuatro estudios estadísticos, efectuados por una cadena de televisión:

- a) Programa preferido
- b) Tiempo medio diario de sintonización
- c) Opinión sobre la cantidad de tiempo dedicada a informativos
- d) Número de programas musicales que les gustaría que se emitiera

ACTIVIDAD 2. AGRUPANDO DATOS

Clasifica las variables estadísticas que aparecen en la tabla.

Elabora una tabla de distribución de frecuencias para la variable OJOS.
Representa gráficamente los datos.



Elabora una tabla de distribución de frecuencias para la variable CALZADO.
Representa gráficamente los datos



Elabora una tabla de distribución de frecuencias para la variable ALTURA.
Representa gráficamente los datos.



Nº	Sexo	Ojos	Pelo	Herm.	Peso	Altura	Calzado	Mes	Año
1	M	verd.	rubio	3	61	163	36	nov	78
2	V	verd.	cast.	4	52	164	39	enero	76
3	V	cast.	cast.	4	64	182	42	mar	77
4	M	cast.	cast.	3	53	164	37	julio	78
5	M	cast.	cast.	2	53	168	38	nov	78
6	V	cast.	cast.	0	59	164	40	oct	77
7	M	cast.	cast.	3	49	153	37	sept	78
8	V	cast.	negro	3	70	171	42	ago	78
9	V	azules	cast.	2	50	167	40	dic	77
10	V	azules	rubio	3	60	160	41	ago	78
11	V	cast.	cast.	1	62	173	41	mayo	76
12	V	cast.	cast.	4	72	180	43	abril	76
13	M	cast.	cast.	2	64	163	38	sept	76
14	V	cast.	negro	2	75	180	43	junio	75
15	V	verd.	cast.	2	75	170	42	dic	77
16	V	cast.	negro	1	70	180	44	oct	77
17	M	azules	rubio	3	54	165	38	mar	78
18	M	cast.	cast.	3	51	152	35	nov	76
19	V	verd.	cast.	3	60	175	44	julio	78
20	V	cast.	negro	3	50	155	39	mar	78
21	V	cast.	negro	2	54	160	39	mar	78
22	M	azules	rubio	2	59	170	40	febr	78
23	V	verd.	cast.	2	55	168	42	mayo	78
24	M	cast.	cast.	2	40	159	35	abril	78
25	M	cast.	cast.	3	60	165	36	nov	78
26	V	cast.	cast.	4	63	170	40	ago	77
27	M	cast.	cast.	1	51	165	37	ago	78
28	V	cast.	cast.	3	47	156	38	nov	78
29	M	azules	cast.	3	48	162	36	febr	78
30	V	cast.	negro	3	53	173	42	nov	78

ACTIVIDAD 3. DIAGRAMA DE TALLO Y HOJAS



A. Usa un diagrama de tallo y hojas para comparar los resultados de un examen (sobre cien puntos) de dos clases

Clase 1: 80, 62, 53, 76, 76, 31, 59, 78, 84, 66, 71, 50, 79, 69, 87, 64, 56, 65, 58, 78, 75, 60, 51.

Clase 2: 71, 68, 56, 79, 73, 51, 48, 83, 64, 58, 75, 45, 91, 80, 59, 34, 55, 73, 81, 62, 64, 69.

a) ¿Qué clase ha realizado mejor el examen?

b) ¿Qué clase parece menos variable en términos de habilidades de los alumnos?

B. Los siguientes son los primeros sueldos (en cientos de euros) de licenciados que realizan trabajos técnicos:

15; 11; 9; 12; 7,5; 8; 10; 16; 11; 12; 10; 9; 10; 14; 9,5;
13; 8,5; 8; 8,5; 12; 11; 9; 11; 7; 10; 13.



Dibuja un diagrama de tallos y hojas para los datos.
Comenta su forma. ¿Es lo que habrías esperado?

ACTIVIDAD 4. EXPERIMENTOS CON ANIMALES



Se ha inyectado el bacilo de la tuberculosis a 72 cobayas para un experimento médico. Los investigadores anotaban el tiempo de supervivencia. El primer animal murió después de 43 días. El más fuerte vivió 598 días. La lista completa es:

43	45	53	56	56	57	58	66	67	73
74	79	80	80	81	81	81	82	83	83
84	88	89	91	91	92	92	97	99	99
100	100	101	102	102	102	103	104	107	108
109	113	114	118	121	123	126	128	137	138
139	144	145	147	156	162	174	178	179	184
191	198	211	214	243	249	329	380	403	511
522	598								

Un diagrama de "tallo y hojas" no es una buena herramienta para representar estos datos.

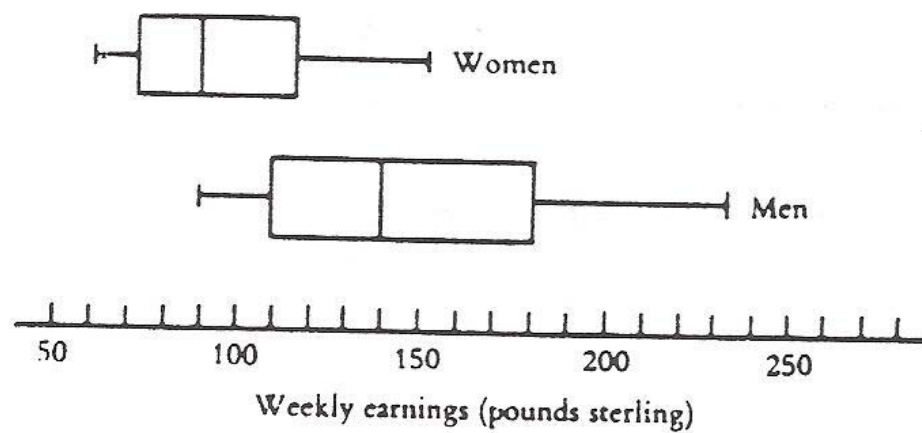
1. ¿Por qué?
2. ¿Crees que la media te proporcionará una buena información de los datos?
3. Calcula la mediana de los números de la tabla anterior
4. ¿Cuál es la mediana de la primera mitad de los datos? ¿Cuál es la mediana de la segunda mitad? ¿Qué representan estos datos?

La mediana de la primera mitad de los datos se llama primer cuartil
La mediana de la segunda mitad de los datos se llama tercer cuartil.

5. ¿Cuál es el segundo cuartil?
6. Dibuja un diagrama de caja para representar estos datos

ACTIVIDAD 5. SALARIOS SEMANALES

El siguiente diagrama de cajas representa los salarios semanales de hombres y mujeres en Inglaterra.



"Casi el 50% de todos los hombres ganan más que el máximo salario semanal de las mujeres"

"Todos los hombres ganan más que el 50% de las mujeres peor pagadas"

a) ¿Son estas afirmaciones consistentes con el diagrama de cajas?

b) Busca en el centro (50%) de ambos grupos. ¿Qué representan los límites en cada caso?

c) ¿Cuál ha sido el uso del gráfico?

ESTALMAT 2007-2008

Sesión 20: Estadística Descriptiva (Analizando y visualizando datos)

Antonio Moreno, Patricia Román

SEGUNDA PARTE: ANÁLISIS CRÍTICO DE LA INFORMACIÓN PROPORCIONADA POR ESTADÍSTICAS

ACTIVIDAD 1. Realizar un debate sobre la manipulación que se puede llevar a cabo en la toma de datos, realización de encuestas, etc. Casos reales.

ACTIVIDAD 2. Comentar las siguientes afirmaciones para detectar conclusiones erróneas y estudios mal planteados:

- El 45 % de los fallecidos en accidentes de tráfico no llevaban puesto el cinturón de seguridad.

- Las estadísticas muestran que casi todos los accidentes de circulación se producen entre vehículos que ruedan a velocidad moderada. Muy pocos ocurren a más de 150 Km. por hora. ¿Significa esto que resulta más seguro conducir a gran velocidad?
- Si las estadísticas mostrarán que la mortalidad por tuberculosis es mayor en Segovia que en las demás provincias, ¿significa esto que el clima segoviano favorece el contagio tuberculoso?
- Un reciente estudio psicopedagógico ha mostrado que los niños de pie grande saben leer mejor que los de pie pequeño. ¿Permitirá el tamaño del pie medir la capacidad de lectura de los niños?
- Suele decirse que casi todos los accidentes de automóvil ocurren cerca de casa. ¿Significa esto que viajar por carretera, a muchos kilómetros de nuestra ciudad, es menos peligroso que callejear por nuestro barrio?
- Según las estadísticas, los diestros viven más que los zurdos. ¿Será eso cierto?

ACTIVIDAD 3. Hasta cierto punto no es exagerado afirmar que hay escasas medidas estadísticas que superen a los porcentajes en la frecuencia en que se cometen equivocaciones honestas o interpretaciones malintencionadas. Analiza las siguientes cuestiones:

- ¿Proporcionan las siguientes frases información para afirmar un excesivo aumento del número de alumnas casadas con profesores o del número de homicidios en una ciudad?
 - El 33.3 % de las alumnas matriculadas en una universidad se han casado con miembros del profesorado.
 - Los homicidios en una ciudad han aumentado en un 60 % con respecto al año anterior.

- "Un comerciante, tras realizar un cursillo de Marketing, y decidido a incrementar sus ventas a la vez que mantenía el margen de beneficio, aumentó los precios de sus productos un 25%, ofreciendo posteriormente a sus clientes un descuento del 25 %". ¿Es correcta esa estrategia comercial, en cuanto a dicho margen de beneficio?

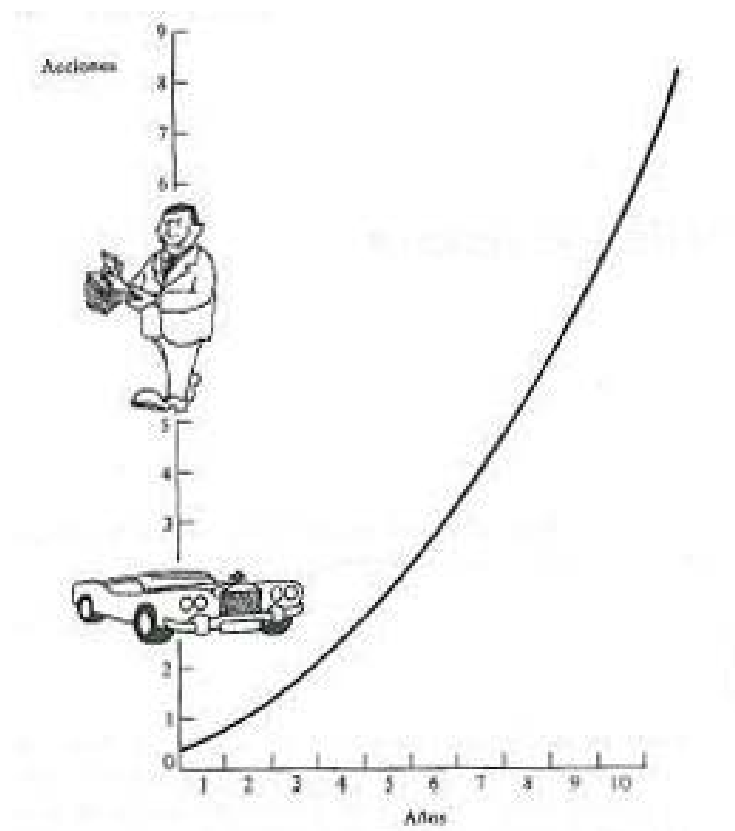
- "El Gobierno propone subir un determinado tipo de impuesto del 4 % al 5 %, y afirma que la subida es pequeña, de tan sólo el 1 %. La Oposición, por su parte, dice que se trata de una subida enorme, del 25 %". ¿Pueden tener razón ambos, Oposición y Gobierno?

- "En un paquete de cereales, que se encuentra a la venta en la actualidad en las grandes superficies, se puede leer lo siguiente: Premiamos tu fidelidad con un paquete más grande en el que

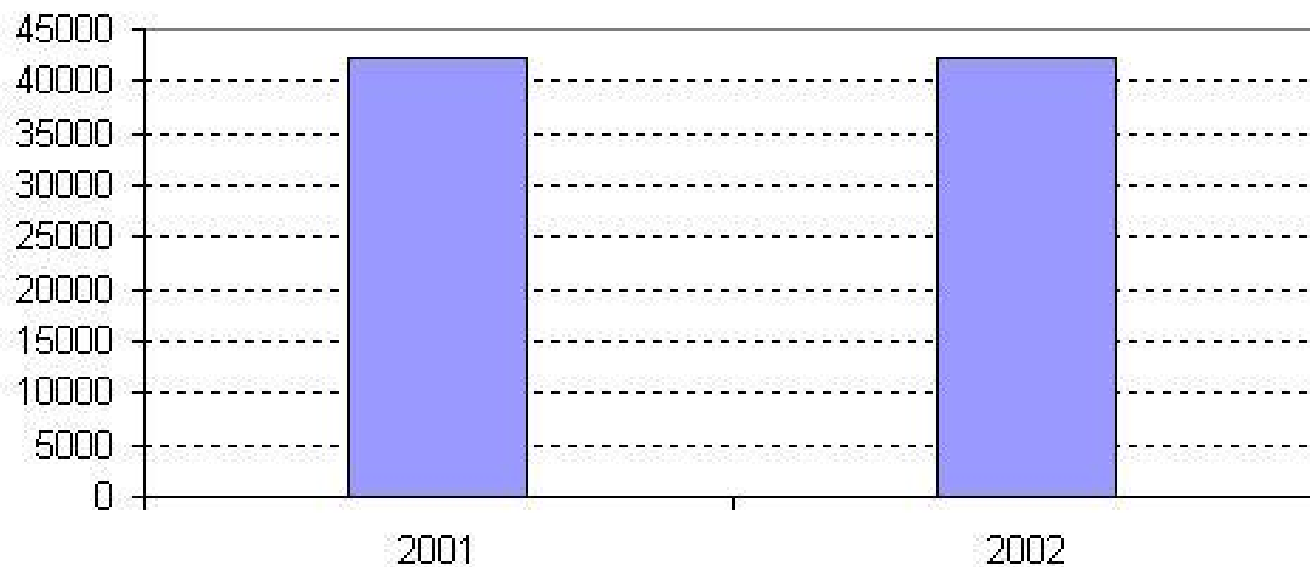
encontrarás un 20 % más de cereales por el mismo precio: de cada cinco desayunos, uno es gratis... ¡no dejes pasar esta oportunidad!".
¿Qué opinas acerca de la veracidad de lo afirmado en la publicidad?

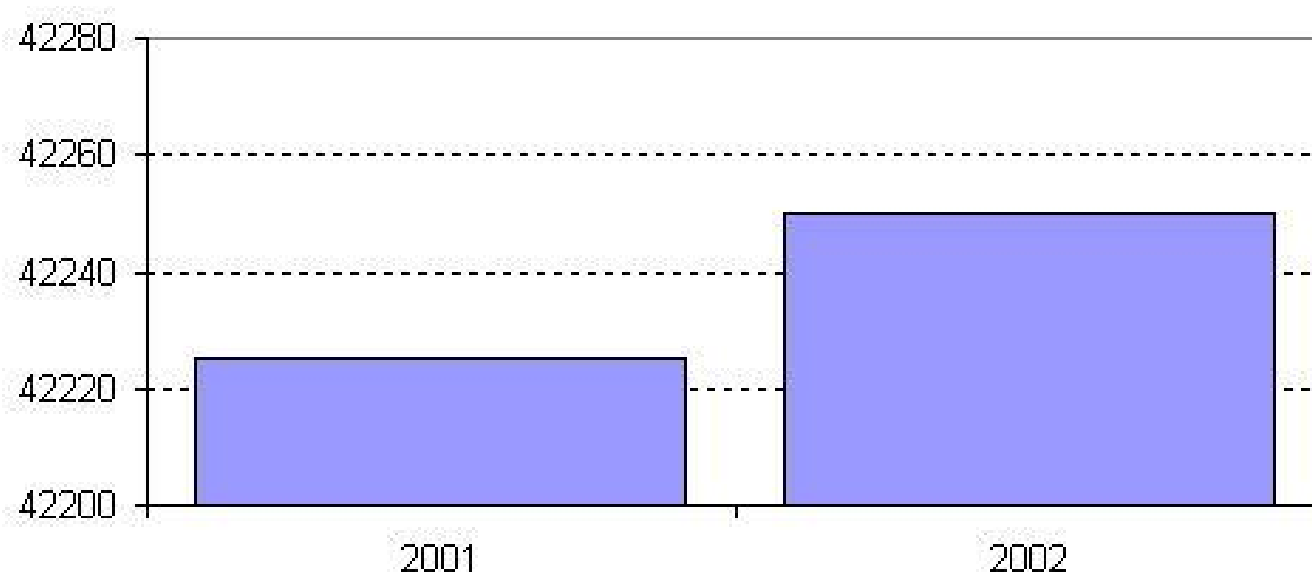
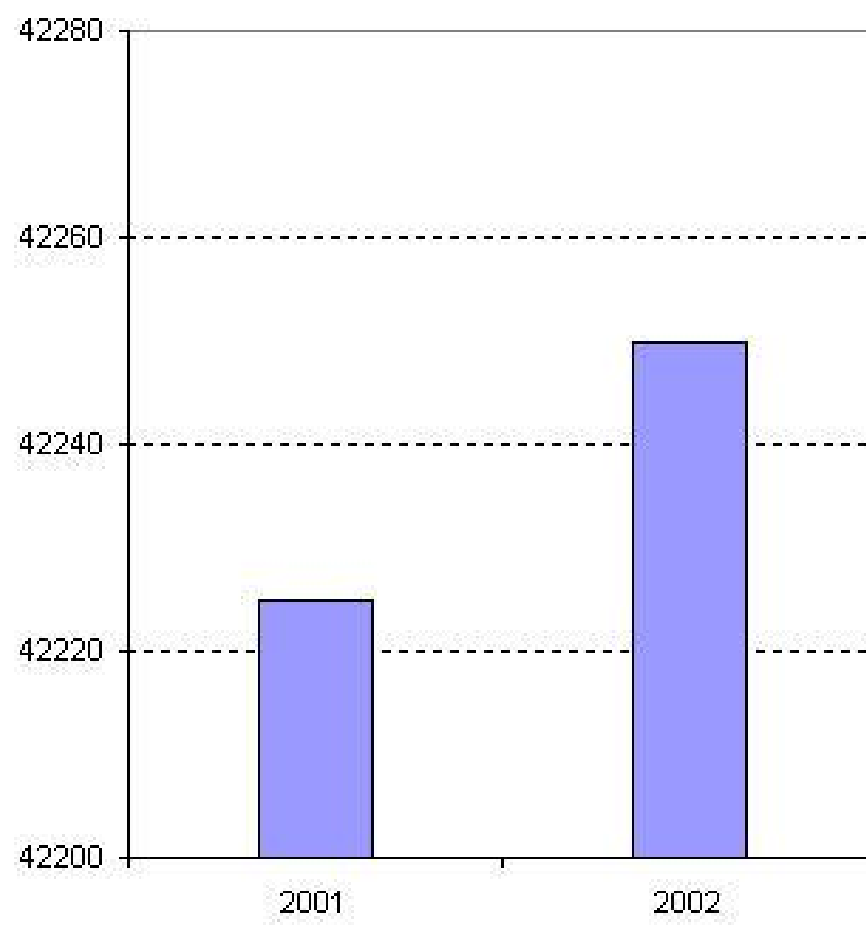
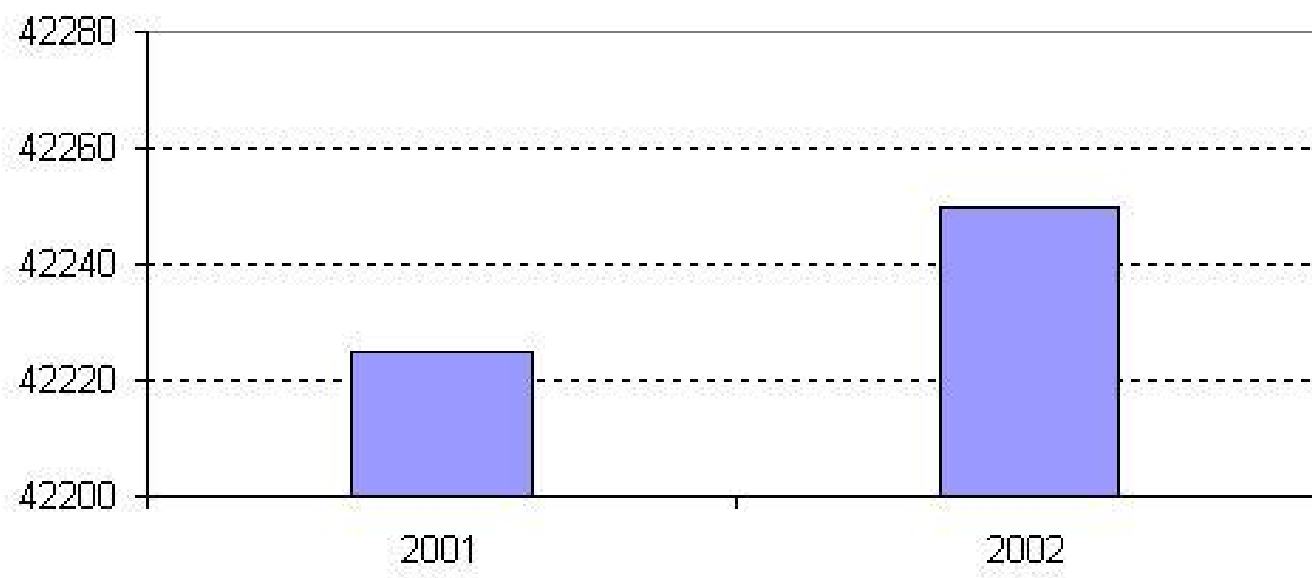
ACTIVIDAD 4: Gráficas engañosas. Analizar las siguientes gráficas y discutir formas de manipular la información mediante gráficas.

A) Evolución de las acciones compradas por una compañía

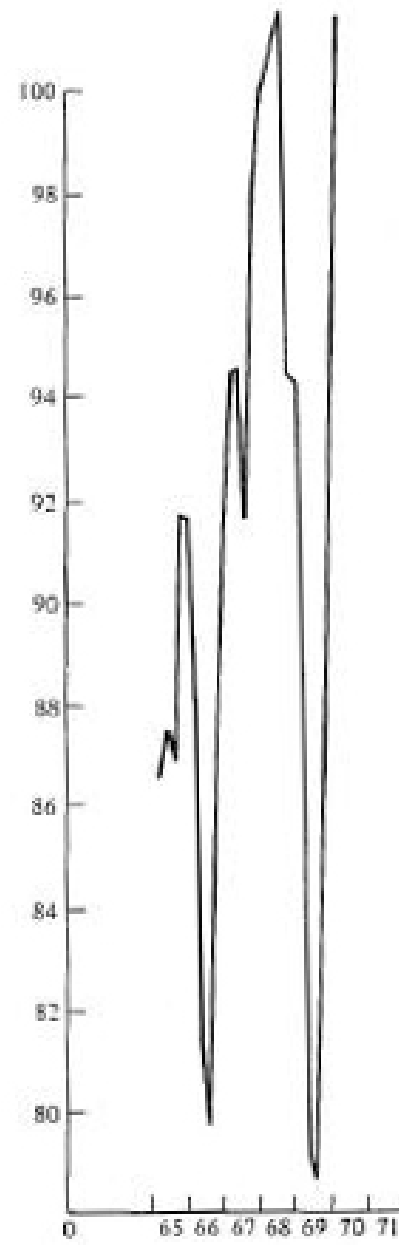
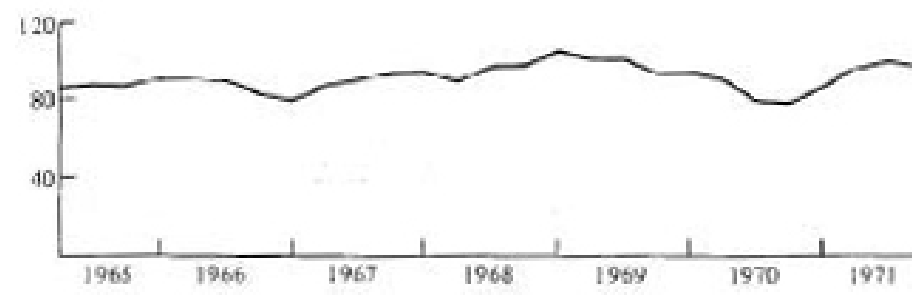


B) Cantidad dedicada a salarios en una empresa en los años 2001 y 2002





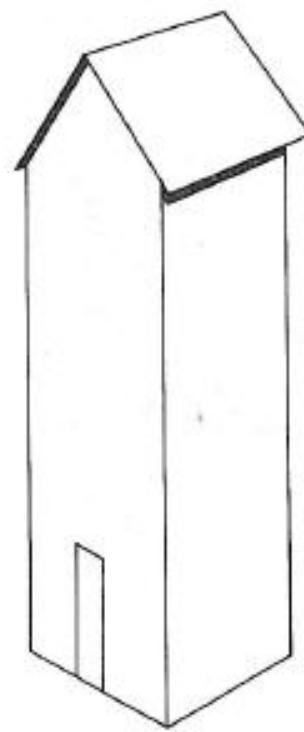
C) Cotizaciones de una determinada acción a lo largo de los años



D) Número de viviendas de nueva construcción en una determinada región. Primer año: 7200. Segundo año: 30400.



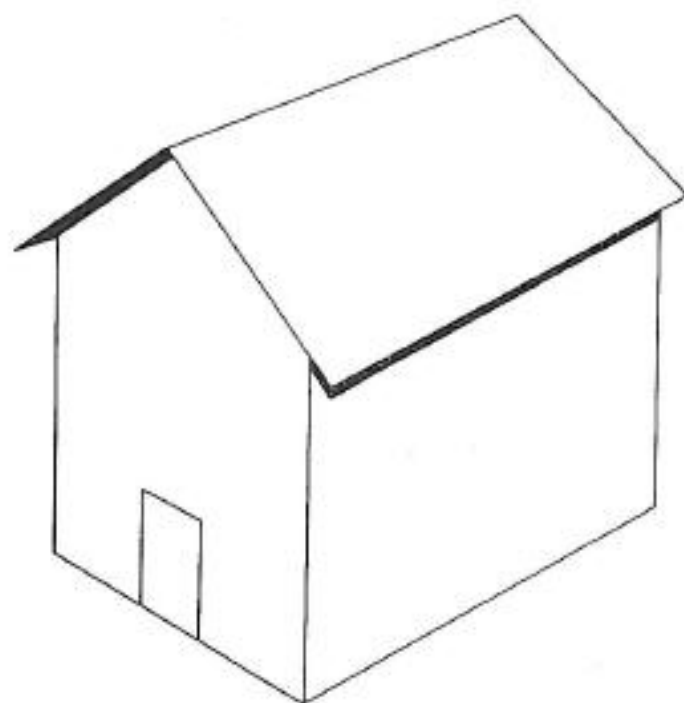
Primer año



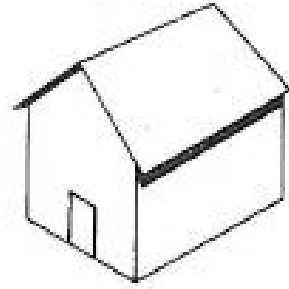
Segundo año



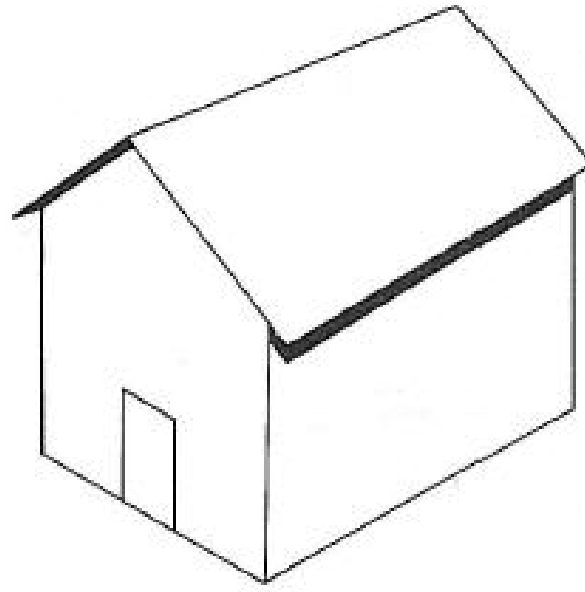
Primer año



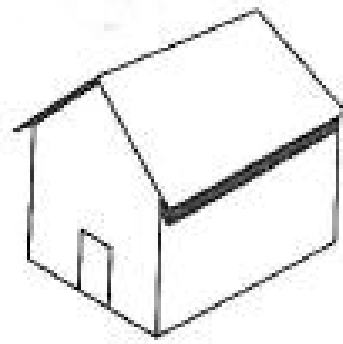
Segundo año



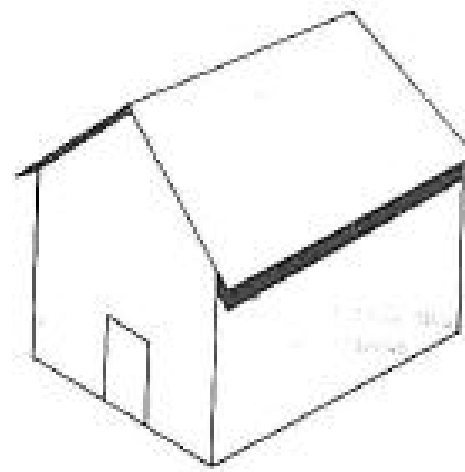
Primer año



Segundo año

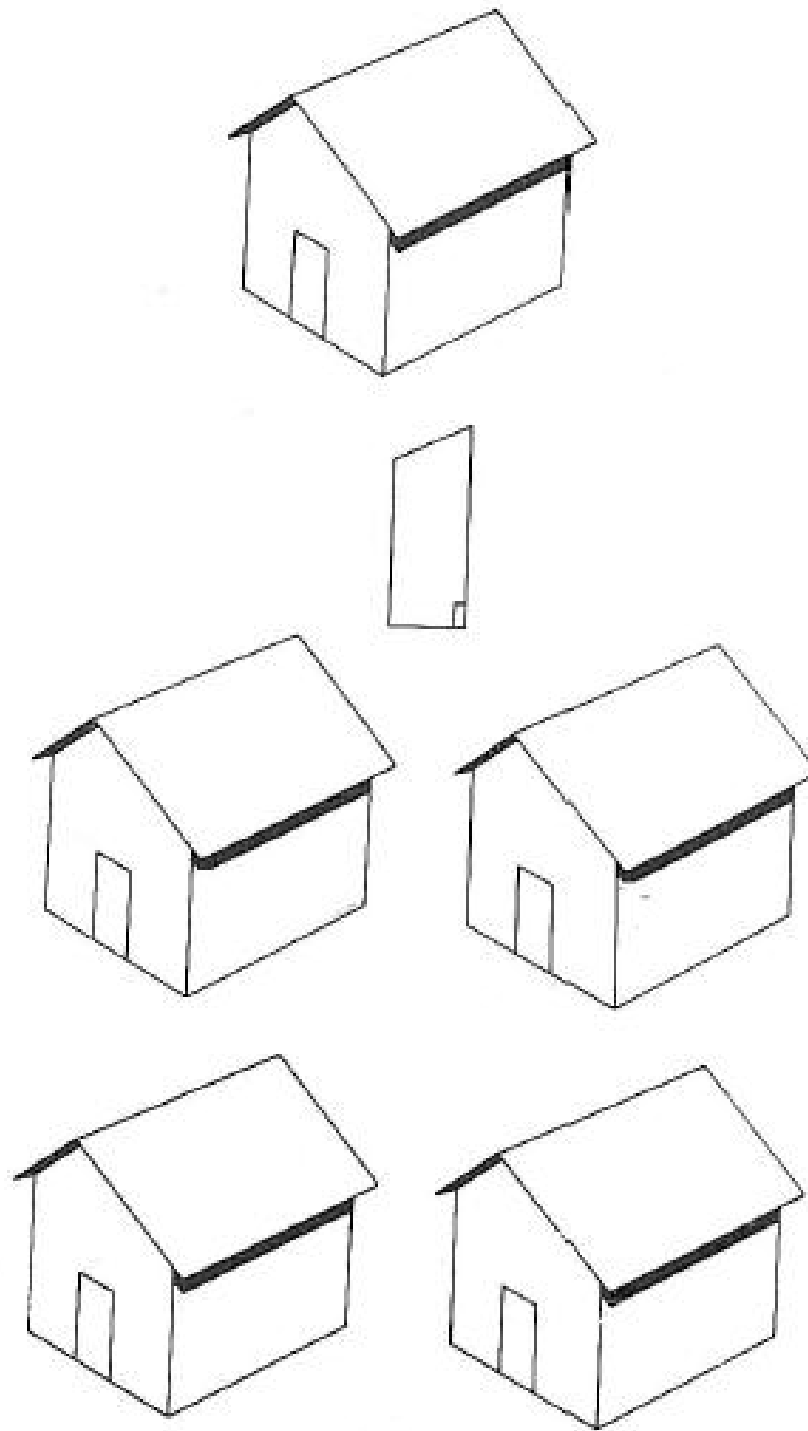


Primer año



Segundo año

Primer año



Segundo año

ACTIVIDAD 5: Análisis de las distintas medidas de tendencia central: influencia en ellas de valores extremos.

- Supongamos que preguntamos a cinco personas cuántas horas ven de televisión a la semana y los datos son: 5, 7, 6, 3 y 6. A continuación añadimos un dato más de una persona que ve la televisión 120 horas a la semana. Calcular en ambos casos la media, la mediana y la moda y analizar los resultados.

- Calcular la media, mediana y moda en los siguientes casos, asociados a datos de las cantidades que gastaron los suscriptores de una determinada revista en regalos:

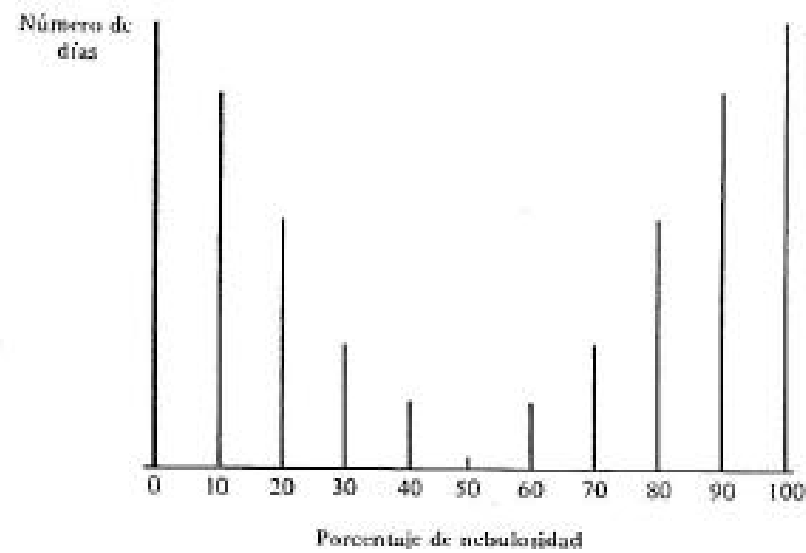
Suscriptor	Gasto
Pérez	48
Gómez	48
García	50
Martínez	60
Sánchez	70

Suscriptor	Gasto
Pérez	48
Gómez	48
García	50
Martínez	60
Sánchez	70
Vázquez	300

Suscriptor	Gasto
Pérez	48
Gómez	48
García	50
Martínez	60
Sánchez	70
Vázquez	300
Suárez	2224

- Comentar ejemplos de casos reales.

ACTIVIDAD 6: Análisis de las distintas medidas de tendencia central: representatividad de las mismas. Realizar dicho análisis sobre los siguientes datos de porcentaje de nebulosidad de una ciudad.



ACTIVIDAD 7: Analiza el siguiente ejemplo de la típica broma estadística desde el punto de vista de la información proporcionada únicamente por las medidas de tendencia central (media, mediana y moda).

Situaciones planteadas:

- Yo me como dos pollo y tu ninguno.
- Yo me como un pollo y medio y tu medio pollo.
- Cada uno nos comemos un pollo.

ACTIVIDAD 8. Análisis de medidas de dispersión en datos dispares.
¿Qué datos presentan una mayor dispersión: la longitud de los tornillos fabricados diariamente por una máquina (que tienen media 2 cm y desviación típica 1 cm) o los salarios de un grupo de recién graduados (que tienen media 1000 euros y desviación típica 100)?