



---

---

## TABLAS Y GRÁFICAS

---

---

### 3.1 Introducción.

La recopilación de información es de gran importancia para analizar fenómenos o comportamientos en los campos educativo, social, económico, científico y político, entre otros. Igualmente es clave que los resultados numéricos de cualquier estudio se presenten de manera clara y concisa, de tal forma que la persona que lea la información obtenga rápidamente una sensación apropiada en torno a las características esenciales de los datos. Lo anterior es particularmente necesario cuando el conjunto de datos es muy grande. En realidad, una presentación efectiva de los datos usualmente revela características importantes tales como su rango de variación, el grado de simetría, qué tan concentrados o dispersos están los datos, en dónde se concentra la mayor parte de los datos, etcétera. En este capítulo se presentan algunas técnicas para graficar y tabular datos.

La manera de recopilar datos y presentar la información obtenida, de algún colectivo a estudiar, es una de las actividades que se desarrollan en el trabajo estadístico y como parte de esta labor es muy importante la forma en que los datos se organizan para su descripción, análisis, interpretación y publicación. Hay dos maneras básicas de presentar los datos que son de uso cotidiano en la estadística: las tablas y las gráficas. Es con base en estas dos formas que se pueden presentar y destacar diversas ideas que se desean expresar acerca de una información recopilada en forma de datos.

En este capítulo se hará una descripción de varias formas de organizar datos en forma de tablas, así como de presentar estos datos en diversas formas de representación gráfica. Además se dará cuenta de algunas características relevantes de la elaboración de tablas y gráficas de datos y distribuciones de frecuencias. Por ejemplo veremos, que el tipo de variable considerada en un análisis de datos, se relaciona e influye en la apropiada elección de la organización y presentación de los datos.

En la presentación se hablará de tres tipos de variables: en primer lugar, de variables categóricas o de atributos haciendo referencia a variables que pueden ubicarse en un nivel de medición nominal u ordinal, es decir, no numérico; en segundo lugar, de variables discretas para hacer referencia a las que corresponden a mediciones ordinales codificadas en forma numérica, o a las que provienen de conteos naturales sobre variables numéricas; y finalmente, de variables continuas para referirse a aquellas que se pueden medir en escalas de intervalo o de razón.

### 3.2 Tablas y distribuciones de frecuencias.

Buena parte de las tablas que se utilizan en estadísticas son conocidas como tablas de distribuciones de frecuencias. La tablas de frecuencia usualmente tienen dos columnas; una de ellas, la primera, muestra todos los posibles valores que asume la variable y la segunda, para cada valor de la variable, muestra el

número de veces que se presenta dicho valor, en el contexto en el que se está trabajando; tal número se denomina **frecuencia**.

Respecto a los tipos de tablas que se utilizan, se pueden mencionar al menos tres criterios para clasificarlas. De acuerdo al tipo de variables involucradas, se puede tener tablas de atributos, de variables discretas y de variables continuas. De acuerdo a la forma como se organizan las frecuencias se tienen: tablas de frecuencias absolutas, de frecuencias relativas, de frecuencias porcentuales, de frecuencias acumuladas absolutas, de frecuencias acumuladas relativas y de frecuencias acumuladas porcentuales. Finalmente, de acuerdo a la cantidad de variables consideradas se tienen tablas de una variable, tablas de dos variables y tablas multivariadas. En el segundo caso se pueden reconocer dos formas de presentación: en paralelo o en forma cruzada. Las tablas cruzadas también son llamadas como tablas de doble entrada, de contingencia o de correlación. En lo que sigue se presentará diversos esquemas y ejemplos de la manera de organizar tablas.

### 3.2.1 Tabla para una variable de atributo.

Este tipo de tabla se utiliza para representar información que provenga de una variable nominal. El arreglo tabular se presenta en dos columnas o filas, en una se indican las categorías o valores de la variable y en la otra la frecuencia. El esquema general es de esta forma:

Categorías de la variable A	Frecuencias absolutas
$A_1$	$f_1$
...	...
$A_k$	$f_k$
Total	N

Tabla 3.1. Esquema de tabla para una variable de atributo.

En este esquema  $A_i$  representa el nombre con el que se reconoce la categoría  $i$ ,  $f_i$  representa la frecuencia absoluta y N el número total de datos.

**Ejemplo 1.** En la Tabla 3.2 se presentan datos acerca del número de incendios forestales ocurridos en diferentes regiones españolas en el año 2004. Nótese que no se evidencia el uso de algún criterio para decidir sobre el orden en que se presentan las regiones. Por ejemplo, se podrían haber ordenado de manera ascendente de acuerdo a la frecuencia observada en cada región.

Región	Nº de Incendios
País Vasco	124
Cataluña	565
Galicia	10618
Andalucía	1055
Castilla La Mancha	1364
Navarra	261
Extremadura	1623
Madrid	379

Castilla y León	1942
Ceuta	3
Melilla	0
<b>Total</b>	<b>17.934</b>

Tabla 3.2. Número de incendios forestales ocurridos en diferentes regiones españolas en el año 2004.

### 3.2.1 Tabla para una variable discreta.

Este tipo de tabla es similar a la tabla que se usa para una variable de atributo. La diferencia con el caso anterior, es que los valores de la variable, es decir los  $x_i$ , son numéricos. El esquema general de este tipo de tabla se presenta en la Tabla 3.3

Valores de la variable	Frecuencias absolutas
$x_1$	$f_1$
$x_2$	$f_2$
...	...
$x_k$	$f_k$
Total	N

Tabla 3.3. Esquema de tabla de una distribución de frecuencias absolutas de una variable discreta.

**Ejemplo 2.** En la Tabla 3.4 se presenta la frecuencia del número de ausencias al trabajo en una empresa. Nótese que en este ejemplo se organizó la información en filas y no en columnas.

<b>Valor</b>	0	1	2	3	4	5	6	7	8	9
<b>Frecuencia</b>	12	8	5	4	5	8	0	5	2	1

Tabla 3.4. Frecuencias del número de ausencias al trabajo en una empresa

### 3.2.3 Tabla para una variable continua.

Cuando se quiere organizar la información de una variable continua o medida en una escala de intervalo o razón, puede ser necesario, sobre todo cuando la cantidad de datos es muy grande, reagrupar los valores en algunas pocas categorías llamadas intervalos de clase. En la práctica el número de estos intervalos suele oscilar entre cinco y veinte; sin embargo, no hay un criterio único o más recomendable para determinar el número de intervalos de clase. En general, el número de intervalos que se usen depende del contexto del que surgen los datos, de la experiencia de quien organiza la información, y/o de las intenciones de lo que quiera hacer notar el analista de datos. En la Tabla 3.5 se presenta un esquema general de este tipo de tabla.

Este tipo de tablas contiene elementos sobre los que es conveniente hacer algunos comentarios. La columna titulada “orden de clase”, es de uso opcional y se utiliza para numerar los intervalos de clase. En general, los intervalos de clase son una partición del conjunto en donde asume valores la variable. Esta partición es excluyente y exhaustiva, es decir, la unión de todos los intervalos es precisamente el conjunto de todos los valores que asume la variable y la intersección entre cualquier par de intervalos diferentes es vacía. Esto implica que un valor cualquiera de los datos siempre se puede clasificar de

manera única en alguno de los intervalos de clase que definen la partición. La Tabla 3.5 presenta un esquema general de este tipo de tablas. Aunque la convención para la utilización de los intervalos de clase en este esquema es la de utilizar un intervalo cerrado a la izquierda y abierto a la derecha, también se presentan variaciones a esta convención. Finalmente, la marca de clase es un valor representativo del intervalo de clase y corresponde a su punto medio.

Orden de Clase	Intervalo $[L_{i-1}, L_i)$	Marca de clase	Frecuencia absoluta
1	$[L_0, L_1)$	$m_1$	$f_1$
2	$[L_1, L_2)$	$m_2$	$f_2$
...	...	...	...
k	$[L_{k-1}, L_k]$	$m_k$	$f_k$
Total			N

Tabla 3.5. Esquema de tabla para una variable continua.

**Ejemplo 3.** En la Tabla 3.6 se presenta las temperaturas en grados centígrados reportadas por cien estaciones meteorológicas colombianas a las doce del día el 23 de febrero de 2006.

Intervalo	Marca de clase	Frecuencias absolutas
$[22,5; 25,5)$	24	45
$[25,5; 28,5)$	27	43
$[28,5; 31,5)$	30	11
$[31,5; 34,5]$	33	1
Total		100

Tabla 3.6. Temperaturas de cien estaciones meteorológicas de Colombia a las 12:00 el 23 de febrero de 2006.

**Ejemplo 4.** En la Tabla 3.7 se presenta un ejemplo hipotético de los resultados obtenidos por cincuenta estudiantes de una institución universitaria en un curso de estadística básica

Intervalo de Clase	Frecuencia $n_i$	Frecuencia Relativa ( $f_i$ )
30 – 39	4	0.08
40 – 59	14	0.28
60 – 69	12	0.24
70 - 89	16	0.32
90 – 100	4	0.08
Total	50	1

Tabla 3.7. Calificaciones en un examen obtenidas por cincuenta estudiantes de una institución universitaria

### 3.2.4 Tablas para diferentes tipos de frecuencias.

El hecho de que las frecuencias se puedan presentar en diferentes formas puede causar confusión a los inexpertos. Aunque no se acostumbra presentar, en trabajos prácticos, todos estos tipos de frecuencias de

manera simultánea, a manera de ilustración se presenta un esquema, en la Tabla 3.8, que recopila las diferentes formas de presentar las frecuencias. Por ejemplo, si  $N=50$ , una frecuencia absoluta de 25, corresponde a una frecuencia acumulada de 0.5 y a un porcentaje del 50%. Por otra parte, las frecuencias acumuladas lo que hacen es acumular el valor de las frecuencias absolutas, relativas o porcentuales.

Valores variable	Frecuencia absoluta	Frecuencia relativas	%	Frecuencia absoluta acumulada	Frecuencias relativa acumulada	% acumulado
$x_1$	$f_1$	$f_{r1}$	$p_1$	$F_1$	$F_{r1}$	$P_1$
$x_2$	$f_2$	$f_{r2}$	$p_2$	$F_2$	$F_{r2}$	$P_2$
...	...	...	...	...	...	...
$x_k$	$f_k$	$f_{rk}$	$p_k$	$F_k$	$F_{rk}$	$P_k$
Total	$N$	1	100			100

Tabla 3.8. Esquema de tabla con todas las versiones de tipos de frecuencias

### 3.2.5 Tablas de doble entrada.

Este tipo de tablas se utiliza con mucha frecuencia para organizar los resultados del cruce de frecuencias de dos variables. Además, cuando estas tablas se utilizan para analizar relaciones de dependencia entre las variables, se habla de tablas de contingencia. En estos casos es usual que en la última fila y columna se presenten los resultados de las frecuencias marginales correspondientes a cada variable.

**Ejemplo 5.** La Tabla 3.9 presenta los resultados de lanzar un par de dados distinguibles (por ejemplo un dado rojo y uno azul) al lanzarlos en 120 ocasiones. Además, el ejemplo ilustra una manera manual y práctica de hacer conteos de frecuencia.





































		Resultados dado-2						Total
		1	2	3	4	5	6	
Resultados dado-1	1	 4	 2	 2	 5	 3	 5	21
	2	 4	 2	 2	 1	 7	 6	22
	3	 3	 3	 4	 4	 1	 2	17
	4	 7	 1	 3	 4	 2	 6	23
	5	 2	 2	 3	 2	 4	 4	17
	6	 5	 2	 2	 2	 3	 6	20
Total		25	12	16	18	20	29	120

Tabla 3.9. Tabla de doble entrada donde han registrado los resultados del lanzamiento de un par de dados distinguibles en 120 ocasiones

**Ejemplo 6.** La Tabla 3.10 presenta los resultados de la intención de separación de 66 parejas heterosexuales cruzada con el tipo de unión bajo la que conviven. Nótese que en este ejemplo, se puede establecer si el tipo de unión está relacionada con la intención de separación en una relación de pareja.

		Tipo de unión			Total general
		Civil	Libre	Religiosa	
Intención de separación	no	7	17	12	36
	si	10	11	9	30
	Total general	17	28	21	66

Tabla 3.10. Tabla de doble entrada donde han registrado los resultados del lanzamiento de un par de dados distinguibles en 120 ocasiones

### 3.2.6 Comentarios acerca del uso de tablas.

Las tablas de atributos son el tipo de tablas que se utilizan más a menudo para organizar la información de atributos, es decir, de variables cualitativas de nivel nominal u ordinal. Este tipo de tablas se puede organizar en orden ascendente o descendente de acuerdo a la frecuencia de los valores o se acuerdo a algún orden impuesto al atributo como por ejemplo el orden alfabético. También se pueden dar casos en que las categorías de organización contengan a su vez subcategorías sobre las cuales también se realizan conteos.

En cuanto a las tablas de variables discretas son las que se utilizan más a menudo para organizar la información de variables numéricas discretas o cuantitativas ordinales. No es recomendable su uso cuando existen demasiados valores diferentes de la variable de observación. Este tipo de tablas, usualmente se presentan en orden ascendente de acuerdo a los valores de la variable.

Las tablas para variables continuas son las que se utilizan más a menudo para organizar la información de variables cuantitativas que se miden en un nivel de intervalo o de razón. En general se debe sopesar, para determinar el número de intervalos de clase a contemplar, el error generado por el agrupamiento utilizado con las dificultades de cálculo posterior, especialmente si no se tiene a la mano un computador para los cálculos. La construcción de este tipo de tablas exige la generación de criterios de partición para realizar el proceso de agrupamiento por el cual se generan los intervalos de clase; en general, es preferible que la amplitud de las clases sea constante, aunque en ocasiones pueden verse mejor algunas características de los datos si se utilizan clases de diferente amplitud. Este tipo de tablas puede generar problemas de cálculo de estadísticas, como la media aritmética, o de representación gráfica cuando los intervalos extremos no se definen de manera acotada.

Finalmente, respecto a las tablas de doble entrada, se puede decir que son de uso frecuente cuando se consideran dos variables de tipo categórico, aunque también las hay para organizar el cruce de dos variables continuas. Este tipo de tablas es llamado como tabla de contingencia, cuando se utilizan para analizar la relación de dependencia entre las dos variables categóricas, o como tabla de correlación cuando las variables consideradas son continuas. En este tipo de tablas se debe tener que las categorías de clasificación sean excluyentes y exhaustivas.

### 3.3 Gráficos.

La forma en que los datos se organizan para su descripción, análisis, interpretación y publicación es de importancia fundamental en la estadística. La representación gráfica es una manera de presentar los datos y que permite destacar algunas ideas que se desean expresar. Al igual que en el caso del uso de representaciones tabulares, las representaciones gráficas son de tipos muy variados, según se trate de variables unidimensionales, bidimensionales, multidimensionales o de conjuntos de datos observados a intervalos regulares de tiempo (como en el caso de las series de tiempo o cronológicas), de atributos, etcétera. Para la selección apropiada de un tipo de gráfico se debe considerar el tipo de variable y la cantidad de clases o categorías. Aquí se presentarán los de mayor uso.

#### 3.3.1 Columnas.

Un diagrama de bloques es una gráfica que se emplea para representar la distribución de una variable discreta o nominal. Consta de una serie de rectángulos, cada uno de los cuales representa una categoría de la variable. Las bases de los rectángulos están sobre una misma recta y se nombran con los valores que toma la variable cuya distribución se quiere representar. Las bases de todos los rectángulos tienen la misma longitud y la altura de cada uno de ellos es proporcional al número de observaciones de la muestra que están incluidas en cada clase. Los rectángulos que conforman la gráfica están separados entre sí para indicar que entre uno y otro valor de la variable no hay más valores.

**Ejemplo 7.** En la Figura 3.1 se muestra la calificación promedio asignado a un grupo de veinticinco trabajadores en cuanto a sus actitudes de disponibilidad hacia el trabajo, iniciativa para realizar labores y capacidad de cuestionamiento de su labor en una empresa de producción de flores.

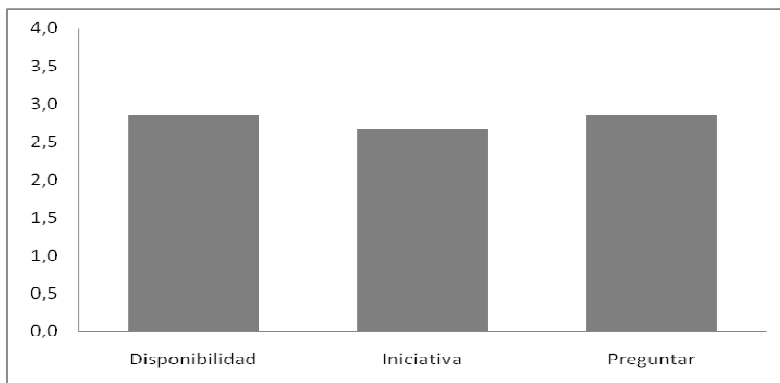


Figura 3.1. Calificación promedio asignada por veinticinco trabajadores a su actitud en términos de su disponibilidad, iniciativa y capacidad de cuestionamiento de su trabajo.

#### 3.3.2 Circular.

Un diagrama circular es una gráfica que se emplea para representar la distribución de una variable categórica. Para construirlo se utiliza un círculo: se divide en tantos sectores como categorías tenga la variable. El tamaño de cada sector (o sea del ángulo central correspondiente) debe ser proporcional al número de observaciones de la muestra que están incluidas en cada clase. En otras palabras, el gráfico debe presentar el aporte de cada valor con respecto al total.

**Ejemplo 8.** La Figura 3.2 muestra la distribución de las inscripciones de estudiantes a diferentes carreras en una universidad. Para mayor claridad, también se incluye la tabla de donde proviene la información.

Programas	Diseño	Electrónica	Matemáticas	Química	Biología	Física	Total
Inscritos	250	150	350	200	150	100	1200
Porcentaje	21%	13%	29%	17%	13%	8%	100%

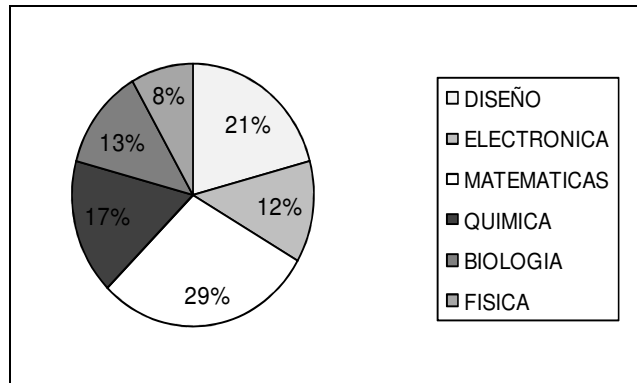


Figura 3.2. Distribuciones de las inscripciones de los estudiantes a diferentes en una Universidad.

### 3.3.3 Puntos.

Un diagrama de puntos es una gráfica que se emplea para dar una idea aproximada de la forma de la distribución de una variable cuantitativa discreta. Sobre una misma recta (usualmente horizontal) se disponen en orden ascendente los posibles valores de la variable y encima de cada uno de esos valores se anotan tantos puntos como veces se repita el valor.

**Ejemplo 9.** Cuarenta estudiantes universitarios participaron en un estudio acerca del efecto del sueño sobre las puntuaciones en los exámenes. La Figura 3.3 presenta los resultados de 20 de los estudiantes que estuvieron voluntariamente despiertos estudiando toda la noche anterior al examen (grupo experimental que no durmió). En la Figura 3.4 se presentan los resultados de los otros 20 estudiantes (grupo control) que se acostaron a las 11 p.m. la noche anterior al examen. Las puntuaciones en el examen se muestran en los siguientes gráficos. Cada punto representa la puntuación de un estudiante particular. Por ejemplo, los dos puntos encima del número 80 en la figura 3.4, indican que dos estudiantes en el grupo control tuvieron una puntuación de 80 en el examen.

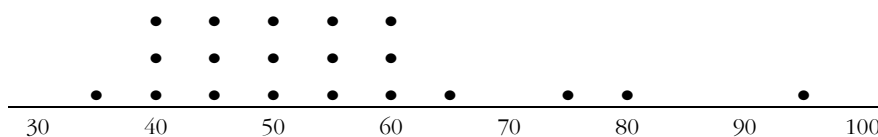


Figura 3.3. Puntuaciones del grupo experimental.



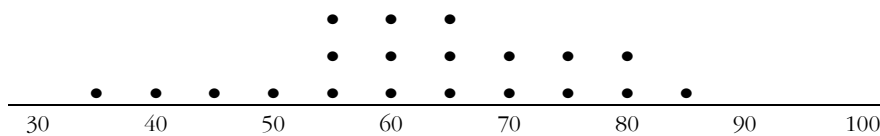


Figura 3.4. Puntuaciones del grupo control.

### 3.3.4 Barras.

Un diagrama de barras es una gráfica que se emplea para representar la distribución de una variable cuantitativa discreta o nominal. Sobre una misma recta (usualmente vertical) se disponen en orden ascendente los posibles valores de la variable y al lado de cada uno de esos valores se trazan segmentos de recta cuya longitud es proporcional a la frecuencia de cada valor de la variable. En realidad este tipo de gráfico se puede ver como equivalente al de columnas, salvo por la disposición de las barras en forma horizontal. Por lo tanto, las alturas de las barras deben ser iguales y el largo de las mismas debe ser proporcional al número de observaciones de la muestra que están incluidas en cada clase.

**Ejemplo 10.** En la Figura 3.5 se muestran los resultados de una investigación en la que una muestra de 107 profesores respondió una prueba acerca de la interpretación y el razonamiento con probabilidades.

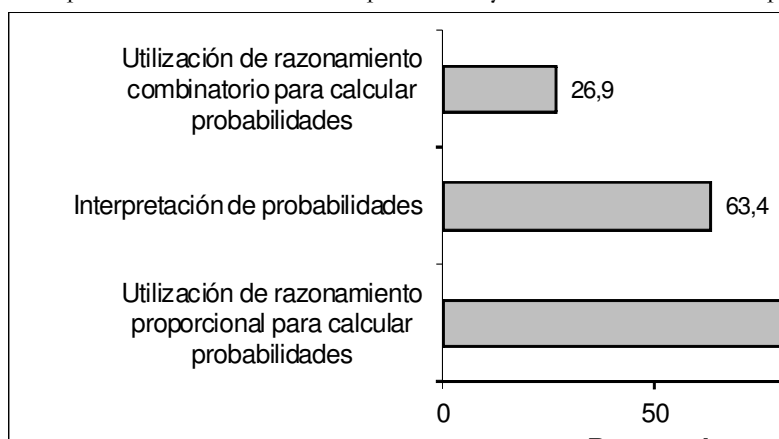


Figura 3.5. Porcentaje de aciertos de profesores en una prueba acerca de interpretación y razonamiento con probabilidades.

### 3.3.5 Histograma.

Un histograma es una gráfica que se emplea para representar la distribución de una variable cuantitativa continua. Está constituida por rectángulos ubicados sobre una misma recta. Cada uno de los grupos en que se clasifica la variable está representado por la base de un rectángulo; y la altura del mismo es proporcional a la frecuencia del correspondiente grupo de valores. Además tales rectángulos son adyacentes. Aunque el histograma es una forma típica de visualizar el comportamiento de una variable continua, también se puede utilizar para una variable discreta cuando tiene un número muy elevado de valores. En su elaboración, para elegir la cantidad apropiada de bloques se toma la parte entera de la cantidad de datos, y para determinar la longitud de la base de los rectángulos (longitud de clase) se divide la diferencia de los datos extremos por la cantidad de bloques.

**Ejemplo 11.** El histograma que se presenta en la Figura 3.6, representa el número de visitas que ha tenido una página Web de Internet, dependiendo de la hora de la visita. Nótese que en este ejemplo las frecuencias se han indicado, encima de los bloques que conforman el histograma.

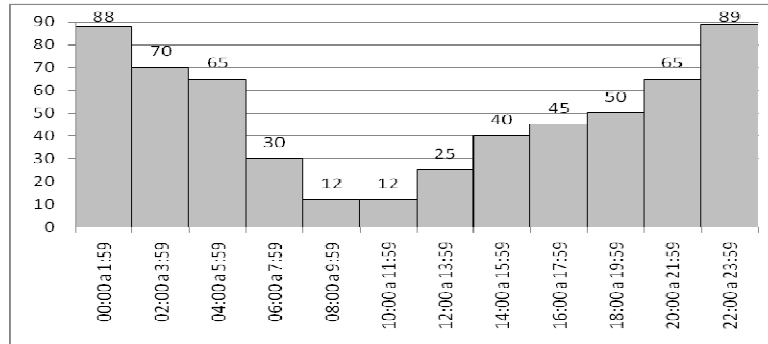


Figura 3.6. Porcentaje de aciertos de profesores en una prueba acerca de interpretación y razonamiento con probabilidades

### 3.3.6 Polígono de frecuencias.

Un polígono de frecuencias es un tipo de gráfico que suaviza el contorno de un histograma. Es una forma geométrica obtenida de segmentos de recta que unen los puntos medios de los intervalos de clase adyacentes del histograma. Los polígonos de frecuencia son especialmente útiles para comparar varias distribuciones en donde la superposición de histogramas causaría confusión.

**Ejemplo 12.** El polígono de frecuencias que se muestra en la Figura 3.7 corresponde a la suavización que se realizó al histograma presentado en la Figura 3.6.

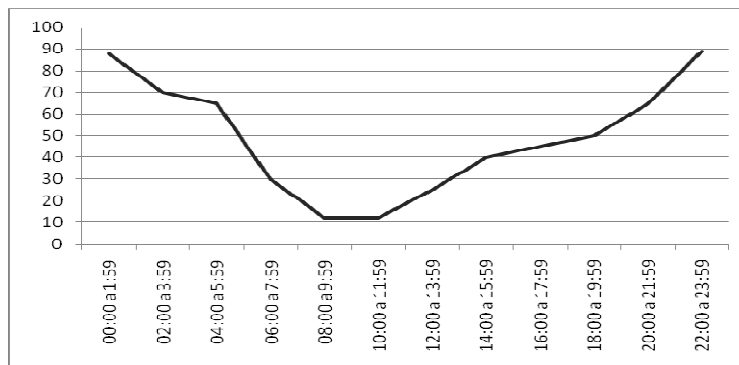


Figura 3.7. Suavización de la gráfica de la Figura 3.5.

### 3.3.7 Tallo y hojas.

El gráfico de tallo y hojas es uno de los métodos básicos del análisis exploratorio de datos que aplica a variables numéricas. La idea de la construcción del mismo se basa en definir un tallo de valores que represente el primer o primeros dígitos del cuerpo de datos, y a lado de cada uno de estos valores se escriben los siguientes dígitos de cada número, que semejan las hojas que se desprenden de un tallo. Una ventaja de este gráfico es la de permitir ver la forma de la distribución de los datos preservando toda o buena parte de la información numérica.

**Ejemplo 13.** En la Figura 3.8 se presenta el gráfico de tallo y hojas correspondiente a las notas de veinte estudiantes en un examen. También se incluye el conjunto de datos para que se verifique la construcción del gráfico.

78	93	61	100	70	83	88	74	97	72
66	73	76	81	83	64	91	70	77	86

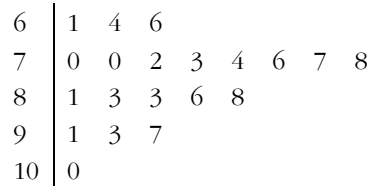


Figura 3.8. Gráfico de tallos y hojas de las calificaciones de los 20 estudiantes.

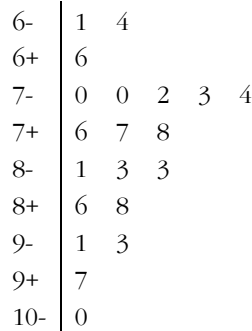


Figura 3.9. Gráfico de tallos y hojas extendido de las calificaciones de los 20 estudiantes.

Por otra parte, en la Figura 3.9 se ilustra una manera de “abrir” el gráfico de tallo y hojas presentado en la Figura 3.8. La convención utilizada consiste en usar el signo “-” para colocar las “hojas” con valores entre 0 y 4 y el signo “+” para colocar las “hojas” entre 5 y 9.

### 3.3.8 Gráfico de caja (Boxplot).

El gráfico de caja, conocido en la literatura inglesa como **Boxplot**, proviene del análisis exploratorio de datos y es muy apropiado para mostrar el comportamiento de los datos cuando interesa presentarlos estratificados por alguna variable cualitativa. Para su construcción es necesario calcular los cuartiles ( $Q_1$ ,  $Q_2$  y  $Q_3$ ), el rango intercuartílico (RQ) y los límites que determinan la zona de anomalías ( $Q_1 - 1,5 \times RQ$  y  $Q_3 + 1,5 \times RQ$ ).

**Ejemplo 14.** En la Figura 3.10 se presentan los resultados obtenidos en grado décimo en un examen final de matemáticas en cuatro cursos dirigidos por diferentes profesores. Se puede observar, por ejemplo, que en el curso 1, hubo un estudiante con un puntaje que sobresale respecto a los demás estudiantes de su curso, mientras que en el curso 2, se identifica a un estudiante que obtuvo un puntaje significativamente menor que los puntajes obtenidos por los demás estudiantes del curso. Asimismo, en cuanto a la homogeneidad de los resultados obtenidos, se puede observar que el curso 4 es el más heterogéneo de los cuatro cursos.

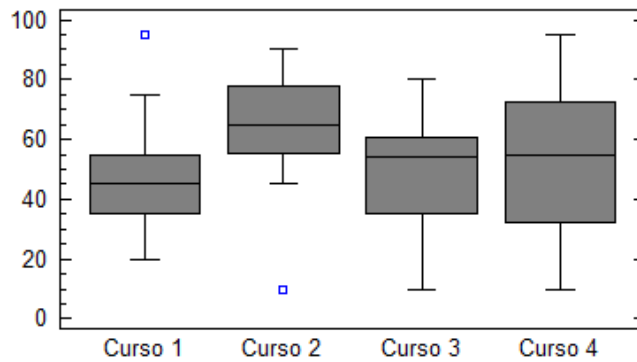


Figura 3.10. Resultados en examen de matemáticas de los estudiantes de cuatro cursos, del mismo nivel académico, en un colegio

**Ejemplo 15.** La Figura 3.11 ilustra la construcción del gráfico de caja para los 20 datos ordenados de la siguiente tabla:

Orden:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Datos:	5	7	8	9	9	9	9	9	9	9	10	10	11	11	11	11	14	15	16	18

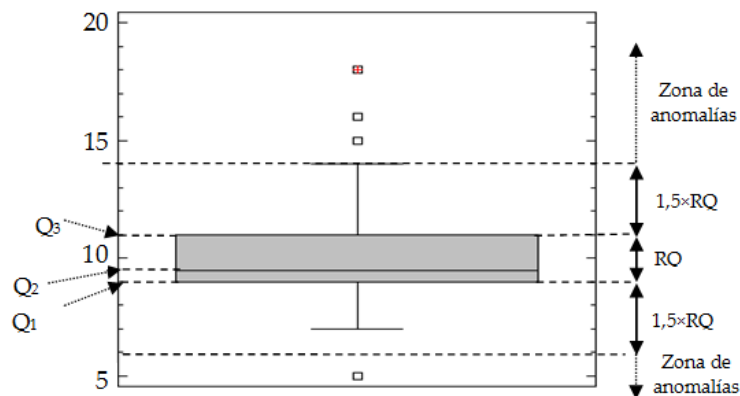
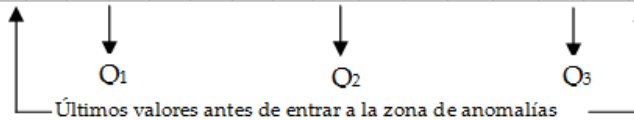


Figura 3.11. Gráfico de caja donde  $Q_1$  es el primer cuartil,  $Q_2$  es la mediana,  $Q_3$  es el tercer cuartil y  $RQ$  es el rango intercuartílico ( $Q_3 - Q_1$ )

### 3.3.9 Dispersión (Nube de puntos).

Este gráfico muestra los puntos asociados con los datos, en un plano bidimensional y se usa para representar los pares de valores (nube de puntos) de una distribución bivalente. Cada dato, representado por un punto, nos indica una sola unidad de observación sobre las cuales han sido hechas dos mediciones,  $x$  y  $y$ . Los valores de cada una de las mediciones son llevadas a escala sobre los ejes  $x$  y  $y$

respectivamente. Los gráficos de dispersión son útiles para analizar la correlación entre variables cuantitativas, ya que la nube de puntos que se genera permite identificar tendencias o ausencias de correlación.

**Ejemplo 16.** En la Figura 3.12 se ilustra el uso de este tipo de gráfico.

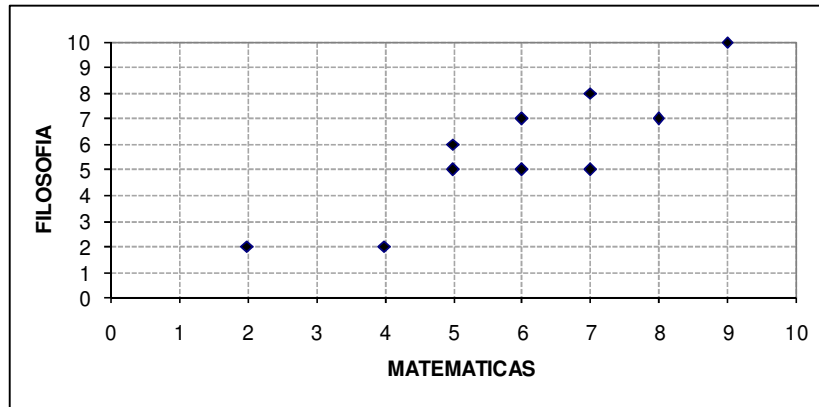


Figura 3.12. Resultados obtenidos en matemáticas y estadística por diez estudiantes.

### 3.3.10 Pareto.

El gráfico de Pareto permite visualizar e identificar las principales razones que generan o se atribuyen a un evento que se desee estudiar. En particular, se utiliza en la supervisión de control de calidad cuando, por ejemplo, se desea determinar las razones que provocan una mala calidad en la elaboración del algún producto textil. Nótese además, que este no es propiamente un gráfico de frecuencias pues la suma de las ocurrencias de los eventos es, en general, más alta que el total de los objetos o productos que se revisan.

**Ejemplo 17.** La Figura 3.13 muestra las razones atribuidas por treinta consumidores para devolver las chaquetas vendidas en un almacén.

En el gráfico de Pareto la suma de las ocurrencias es mucho mayor que 30, que es el número de chaquetas revisadas. Para su elaboración lo que se hace es contar el número de ocurrencias que se atribuyen al fenómeno en cuestión y luego se muestran las ocurrencias en columnas ordenadas de mayor a menor.

### 3.3.11 Pictogramas.

Los pictogramas usualmente consisten en dibujos icónicos cuyo tamaño suele representar cantidades o proporciones de la variable o variables de las que se está informando.

**Ejemplo 18.** En la Figura 3.14 se muestra un ejemplo de la manera como se utiliza el pictograma.

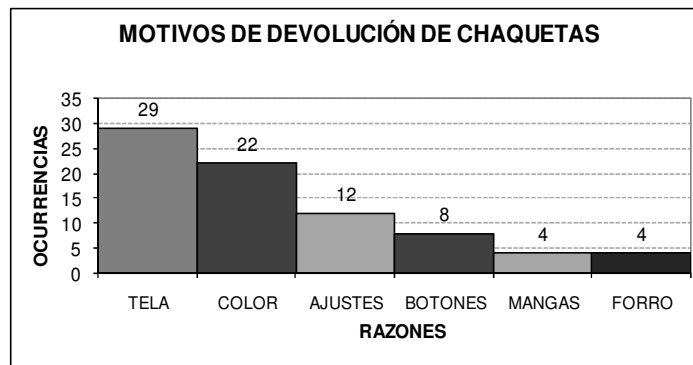


Figura 3.13. Razones para la devolución de chaquetas en un almacén.



Figura 3.14. Población de E.U. de 1930 a 1990

### 3.3.12 Perfiles.

Los gráficos de perfiles se pueden considerar como una extensión de los gráficos de columnas o de barras. Se usan para comparar resultados atribuibles a los efectos de controlar u observar dos variables.

**Ejemplo 19.** En la Figura 3.15 y en la tabla que se presenta a continuación, se comparan los promedios de efecto de sueño (en horas) de tres tipos de drogas somníferas en diferentes dosificaciones en una muestra de 60 pacientes voluntarios.

Dosis en miligramos	Droga 1	Droga 2	Droga 3
95	5	2	1
96	6	2	1
97	7	3	2
98	8	5	3
99	8	6	2

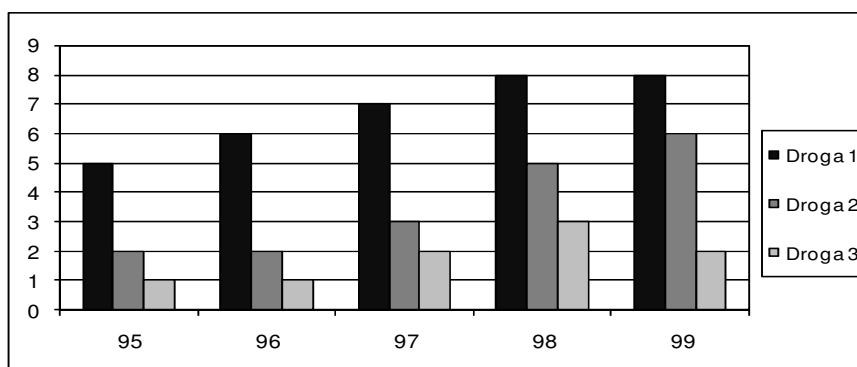


Figura 3.15. Comparación del efecto promedios de sueño (en horas) de tres tipos de drogas somníferas en diferentes dosis, en una muestra de 60 pacientes voluntarios.

En la Figura 3.16, se representan los datos de la tabla utilizando un diagrama de perfiles con barras horizontales.

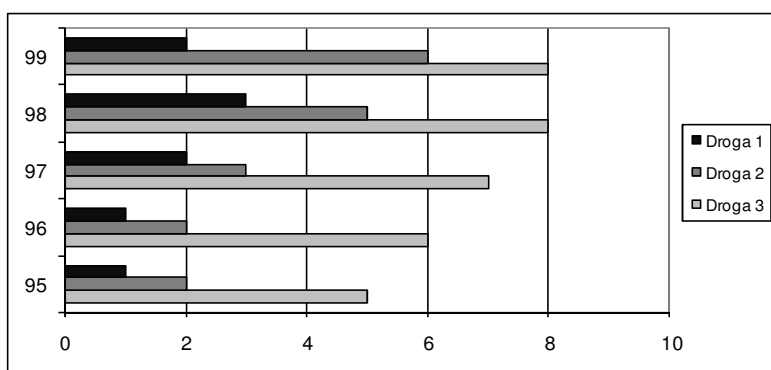


Figura 3.16. Versión en presentación de barras de los datos de la gráfica de la Figura 3.15.

### 3.3.13 Columnas o barras apiladas.

Una variación de la gráfica de perfiles la aporta los gráficos de columnas o de barras apiladas, que permiten comparar entre categorías el aporte de cada valor al total. Como se verá en los ejemplos este tipo de gráficas se puede presentar en versiones absolutas o porcentuales.

**Ejemplo 20.** Los datos presentados en la tabla siguiente muestran el número de autos, camionetas y camperos vendidos durante los meses de diciembre de los años 1995 a 1999 en un concesionario de venta de carros. En las Figuras 3.17 y 3.18, se muestra la versión de apilamiento absoluta en columnas y barras respectivamente.

Año	Autos	Camionetas	Camperos
95	5	2	1
96	6	2	1
97	7	3	2
98	8	5	3
99	8	6	2

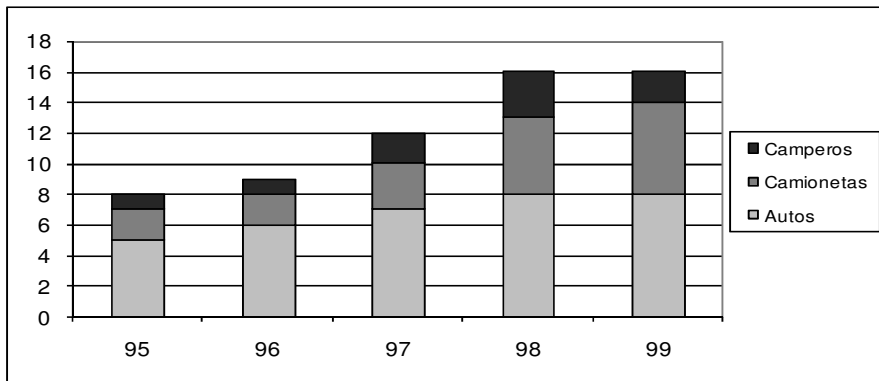


Figura 3.17. Número de autos, camionetas y camperos vendidos durante los meses de diciembre de los años 1995 a 1999 en un concesionario de venta de carros.

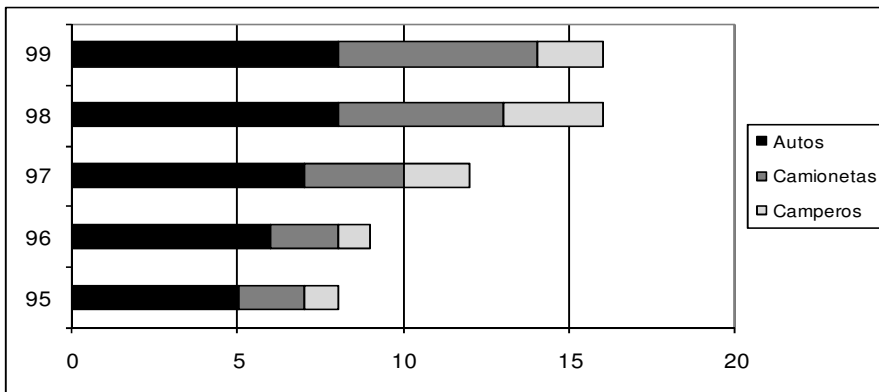


Figura 3.18. Versión presentada en barras de la Figura 3.17. Observe que en este caso se ha hecho un manejo menos preciso de la escala numérica horizontal.

**Ejemplo 21.** Respecto a los mismos datos considerados en el ejemplo anterior, en las Figuras 3.19 y 3.20, se muestra la versión de apilamiento porcentual en columnas y barras respectivamente.

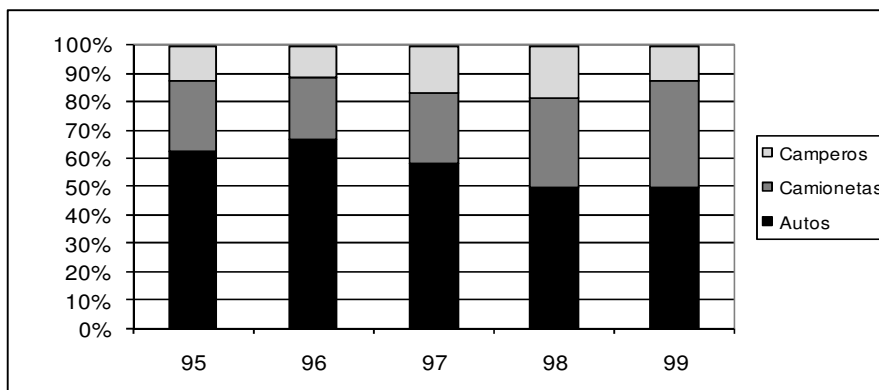


Figura 3.19. Porcentaje de autos, camionetas y camperos vendidos durante los meses de diciembre de los años 1995 a 1999 en un concesionario de venta de carros.



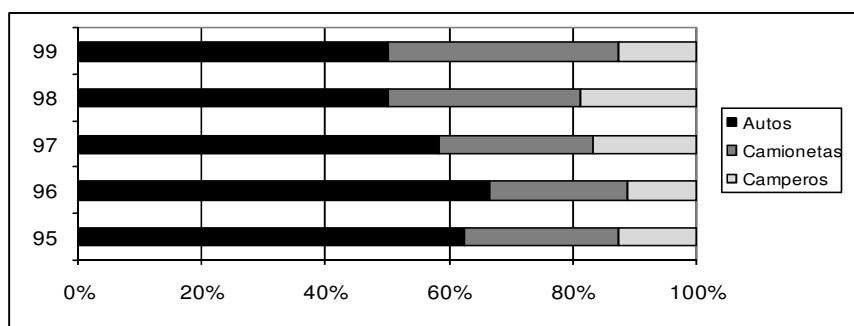


Figura 3.20. Versión en barras de la Figura 3.18.

### 3.3.14 Líneas.

Este tipo de gráficos se utiliza para presentar tendencias a lo largo del tiempo o entre categorías. Cuando una de las variables es el tiempo y la otra es por ejemplo, las ventas, el gráfico también se conoce con el nombre de series de tiempo. Por otra parte, y al igual que en los casos de los gráficos de columnas y barras, en los gráficos de líneas se pueden elaborar apilamientos absolutos y porcentuales.

**Ejemplo 22.** Los datos de los ejemplos presentados en el apartado sobre gráficos de columnas o barras apiladas también se pueden utilizar para representar las diferentes modalidades de los gráficos de líneas.

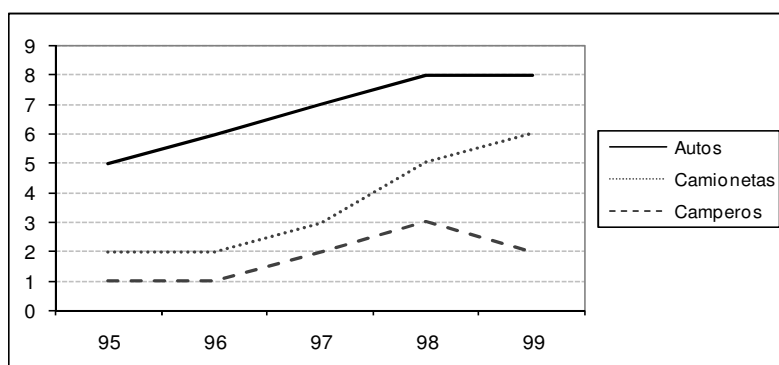


Figura 3.21. Número de autos, camionetas y camperos vendidos durante los meses de diciembre de los años 1995 a 1999 .

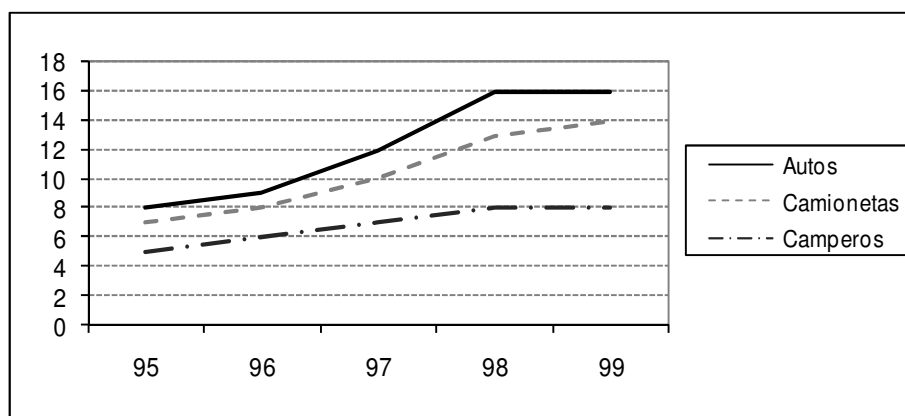


Figura 3.22. Versión apilada absoluta de gráfico de líneas del número de autos, camionetas y camperos vendidos durante diciembre en los años 1995 a 1999.

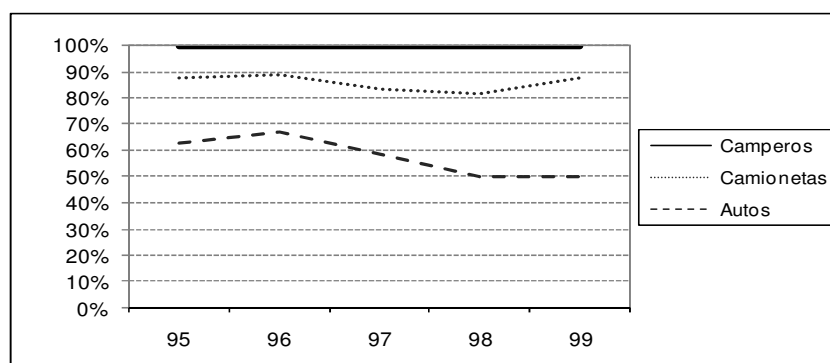


Figura 3.23. Versión apilada porcentual de gráfico de líneas del número de autos, camionetas y camperos vendidos durante diciembre en los años 1995 a 1999.

En la Figura 3.21 se muestra una versión sin apilar del número de autos, camionetas y camperos vendidos durante los meses de diciembre en los años 1995 a 1999 en un concesionario de venta de carros, mientras que en las Figuras 3.22 y 3.23, se muestran las versiones apiladas absolutas y porcentuales respectivamente de los mismos datos.

### 3.3.15 Áreas.

En este tipo de gráfico, como en los gráficos de línea, también se presenta la tendencia de los valores de los datos a lo largo del tiempo o entre categorías.

**Ejemplo 23.** El presupuesto de inversión del gobierno en miles de millones de pesos en las áreas de educación y ciencia y tecnología se presenta en la siguiente tabla.

Años	Educación	Ciencia y tecnología	Totales
2001	250	200	450
2002	200	150	350
2003	250	180	430
2004	270	150	420
2005	220	200	420
2006	350	250	600

Obsérvese que en la Figura 3.24 se presentan los presupuestos sin realizar apilamientos, mientras que en la Figura 3.25 y 3.26 respectivamente se utilizan los apilamientos absolutos y porcentuales.

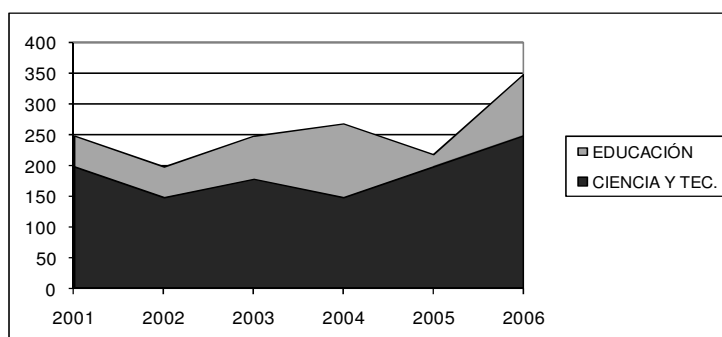


Figura 3.24. Presupuesto de inversión de los años 2001 a 2006, en educación y ciencia y tecnología (versión sin apilamientos).

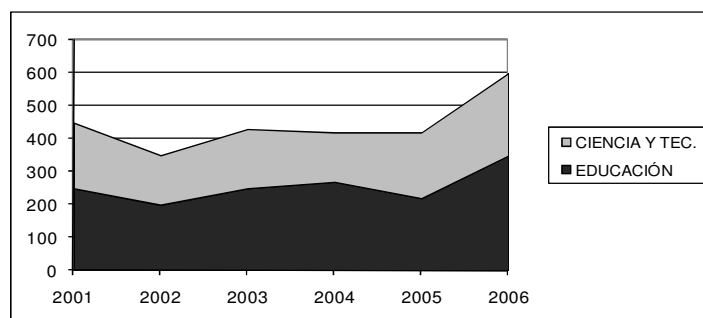


Figura 3.25. Presupuesto de inversión de los años 2001 a 2006, en educación y ciencia y tecnología (versión con apilamientos absolutos).

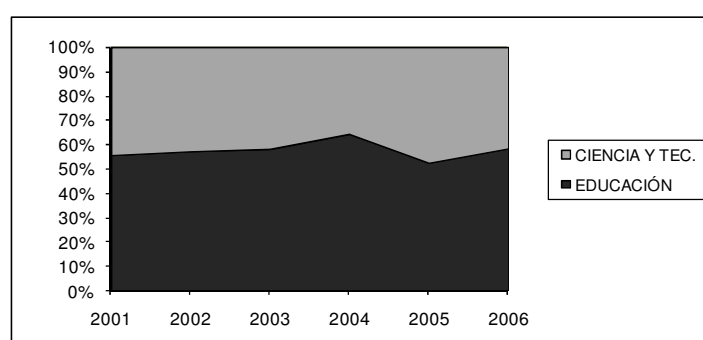


Figura 3.26. Presupuesto de inversión de los años 2001 a 2006, en educación y ciencia y tecnología (versión con apilamientos porcentuales).

### 3.3.16 Columnas en tres dimensiones.

Los gráficos columnas en tres dimensiones sirven para representar la información que proviene de tablas de doble entrada ya que con base en ellos se pueden comparar frecuencias entre categorías bivariantes. En realidad, este tipo de gráfico puede verse como extensión del histograma aplicado a dos variables o como variación de los gráficos de columnas (barras), o de perfiles para dos variables.

**Ejemplo 24.** La venta de carros en diciembre de 2008 (autos, camionetas y camperos) en tres concesionarios diferentes se puede volver a utilizar para ilustrar el uso de este tipo de gráficos. Los datos que se representan en la Figura 3.27 se organizaron en la siguiente tabla:

	C1	C2	C3
Autos	2	3	4
Camionetas	1	5	3
Camperos	4	2	1

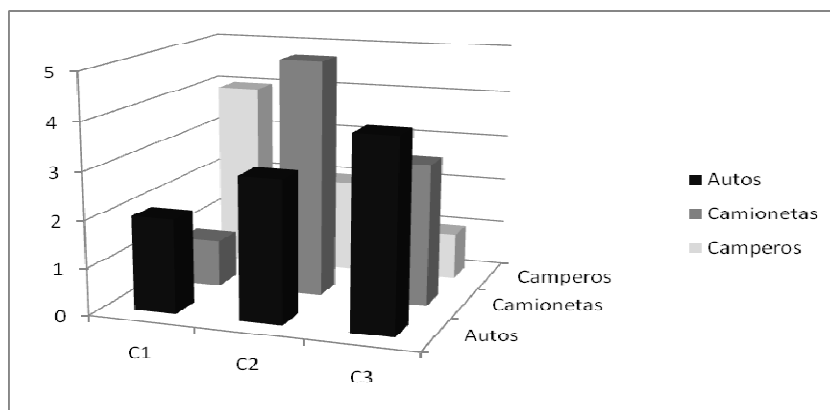


Figura 3.27. Número de carros vendidos en tres concesionarios diferentes en diciembre de 2008

### 3.3.17 Cotizaciones.

Los gráficos de cotizaciones, como lo indica su nombre, se utilizan en el mercado bursátil para analizar las fluctuaciones de los precios de las acciones.

**Ejemplo 25.** Las Figuras 3.28, 3.29 y 3.30 ilustran la utilización de los gráficos de cotizaciones, en tres modalidades. Los datos que se tomaron para elaborar estos gráficos aparecen en la siguiente tabla y corresponden al comportamiento de una acción particular. Las columnas corresponden al día de la semana, volumen de acciones transadas en el día, precio de apertura en el día, máximo precio alcanzado durante el día, mínimo valor alcanzado durante el día y precio de cierre en el día.

Día	Volumen	Apertura	Máximo	Mínimo	Cierre
1	200	55	58	52	56
2	250	53	60	52	54
3	300	59	62	56	60
4	250	54	56	54	55
5	300	56	54	57	55
6	150	58	56	59	57
7	200	56	60	52	58

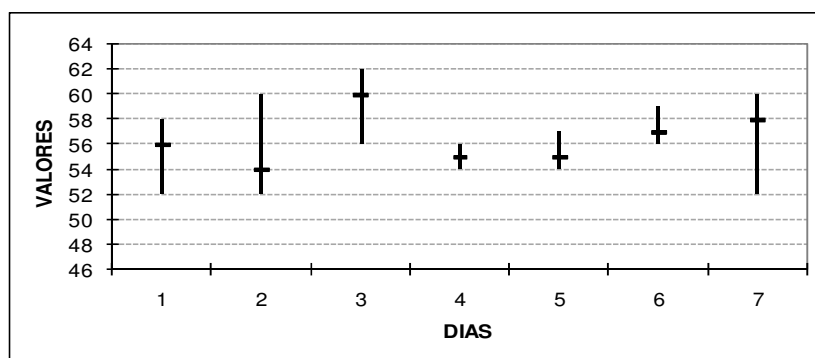


Figura 3.28. Fluctuaciones de los valores en miles de pesos de una acción durante una semana. Se indican valores extremos alcanzados y el precio de cierre con un rectángulo.

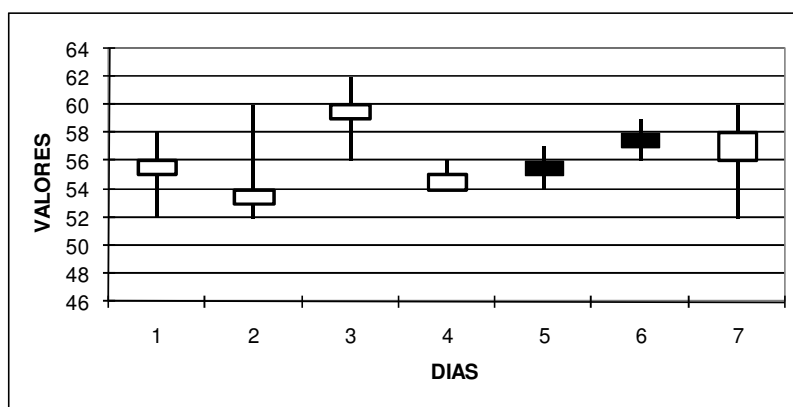


Figura 3.29. Fluctuaciones de los valores (en miles de pesos) de una acción durante siete días. Se indican los valores máximos y mínimos alcanzados y los precios de apertura y cierre de la acción se representan con un rectángulo. El rectángulo es negro cuando el precio de apertura es menor que el de cierre.

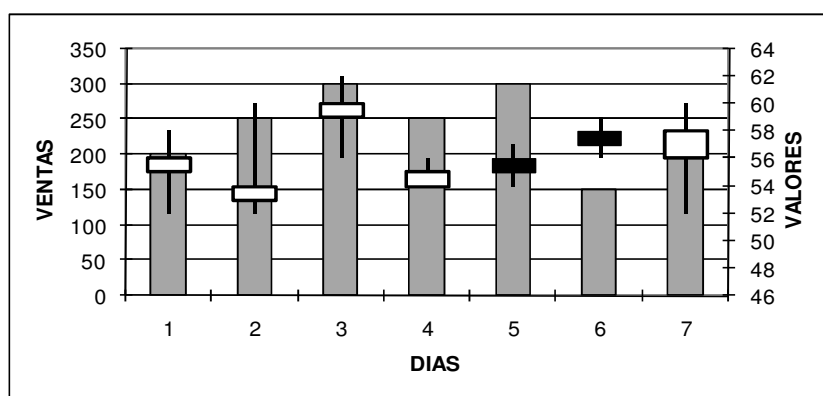


Figura 3.30. Fluctuaciones de los valores (en miles de pesos) de una acción durante siete días. Este gráfico contiene la información del gráfico anterior, pero además, incluye la información de monto total de las ventas en millones de pesos.

### 3.3.18 Ábaco de probabilidad.

Este tipo de gráfico se utiliza para representar tablas de frecuencias relativas asociados a una variable discreta o para representar una tabla de probabilidades de variables aleatorias discretas.

**Ejemplo 26.** Un juego consiste en lanzar 3 dados distinguibles y apostarle a la aparición de cierto número. Si resulta una vez el “6” el jugador se gana un punto, si resulta dos veces el “6” se gana dos puntos, si resulta tres veces el “6” se gana tres puntos y si no sale el “6” se pierde un punto.

Los posibles resultados de este juego son 216 ternas: 111, 112, 121, 211, 113, 131, 311, ..., 666.

La variable aleatoria es el puntaje ganado por el jugador, la cual toma los valores -1, 1, 2 y 3, los cuales dependen de la cantidad de veces que aparece el “6”.

En la siguiente tabla aparecen las posibles ganancias y la cantidad de resultados para los cuales se obtendrían las respectivas ganancias:

Cantidad de "6"	Ganancia (Puntaje)	Cantidad de resultados	Frecuencias relativas
0	-1	125	125/216
1	1	75	75/216
2	2	15	15/216
3	3	1	1/216

El gráfico que se muestra en la Figura 3.31, similar a un diagrama de barras, se conoce como ábaco de probabilidad.

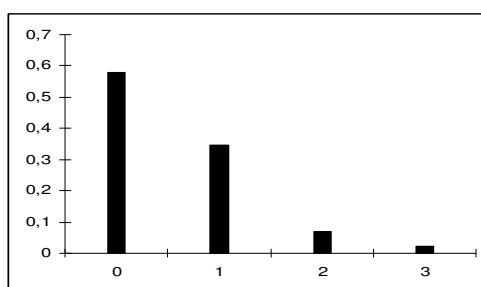


Figura 3.31. Probabilidades de ganancia de un jugador de dados..

**Ejemplo 27.** Un estudiante presenta un examen que contiene 4 preguntas de Falso y Verdadero. Las dos primeras preguntas son de Estadística y las dos últimas son de Probabilidad.  $X$  es la cantidad de veces que puede marcar Verdadero en las dos primeras preguntas, y  $Y$  es la cantidad de veces que puede marcar Verdadero en todo el examen.

Las preguntas  $P_1$ ,  $P_2$ ,  $P_3$  y  $P_4$  del examen se pueden responder de las siguientes maneras:

No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$P_1$	V	V	V	V	F	V	F	V	F	V	F	V	F	F	F	F
$P_2$	V	V	V	F	V	V	F	F	V	F	V	F	V	F	F	F
$P_3$	V	V	F	V	V	F	V	F	V	V	F	F	F	V	F	F
$P_4$	V	F	V	V	V	F	V	V	F	F	V	F	F	F	V	F

La tabla anterior nos permite realizar los conteos para las variables  $X$  e  $Y$  que se muestran en la siguiente tabla:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$X$	2	2	2	1	1	2	0	1	1	1	1	1	1	0	0	0
$Y$	4	3	3	3	3	2	2	2	2	2	2	1	1	1	1	0

Se puede ver que hay 16 maneras de responder el examen. En el primer caso, la variable  $X$  toma el valor 2 y la variable  $Y$  toma el valor 4. Con base en esta tabla se obtiene la siguiente tabla de frecuencias, y con base en esta tabla se obtiene el ábaco para las dos variables que se muestra en la Figura 3.32.

		Y				
		0	1	2	3	4
X	0	1	2	1	0	0
	1	0	2	4	2	0
	2	0	0	1	2	1

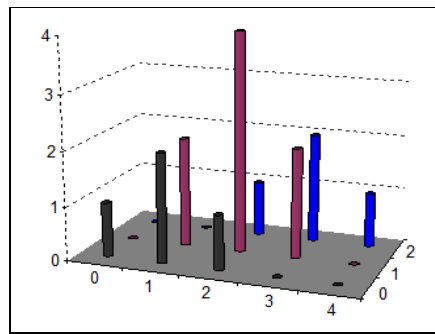


Figura 3.32. Ábaco para las variables X,Y.

### 3.3.19 Comentarios acerca del uso de gráficos.

El gráfico apropiado para representar una información se elige teniendo en cuenta la variable estadística, la tabla y la cantidad de datos. La finalidad de las gráficas es visualizar mejor la información. Por ello, la mejor gráfica suele ser la más simple. Al presentar gráficas siempre se debe buscar sencillez y claridad, de tal manera que éstas sean comprensibles sin la ayuda de las descripciones de textos. Toda gráfica debe tener título e indicar con claridad el contenido de la misma. Las columnas, barras y líneas en los diversos gráficos de este tipo deben preservar en lo posible el nivel de frecuencia cero para no generar gráficos engañosos.

Las gráficas no se deben ver como un sustituto de tablas u otros medios de representación estadística, más bien son un complemento del análisis de datos. En realidad, la gráfica es considerada como el principal medio de expresión de la estadística por diferentes razones: es más llamativa y sugestiva, deja en la memoria una expresión más duradera que las tablas o el sólo texto y requiere de un menor tiempo de lectura.

### 3.4 Ejercicios.

- La persona encargada de administrar un taller de carros, quiere determinar los motivos que inducen con más frecuencia a que los clientes ingresen su automóvil para revisión y mantenimiento. Después de recoger la información va a presentarla en un gráfico.
  - ¿Cuál es la variable que se va a medir? ¿De qué tipo es?
  - ¿Qué tipo de diagrama es el más adecuado? ¿Por qué?
- Un estudiante de derecho quiere determinar la proporción de estudiantes que hay en cada una de las carreras que conforman la facultad de Humanidades para corroborar su afirmación de que los estudiantes de derecho son los que deciden en una votación pues son la mayoría. Después de que haya recogido la información va a presentarla en un gráfico.
  - ¿Cuál es la variable que se va a medir? ¿De qué tipo es?
  - ¿Qué tipo de diagrama es el más adecuado? ¿Por qué?
- El administrador de un supermercado está interesado en determinar si es necesario instalar más cajas registradoras en el almacén para darle una atención más rápida a la comunidad. Para el efecto, la persona encargada de hacer la investigación toma, un día cualquiera, una muestra aleatoria de 50 compradores y anota el tiempo que cada uno de ellos gasta haciendo cola para pagar sus compras. Después de que haya recogido la información va a presentarla en un gráfico.

- a) ¿Cuál es la variable que se va a medir? ¿De qué tipo es?
- b) ¿Qué tipo de diagrama es el más adecuado? ¿Por qué?
4. En el curso de estadística se han realizado tres parciales. El profesor del curso cree que, en general, sus alumnos van bien pues la mayoría de ellos han aprobado los tres parciales. El profesor quiere presentar la información correspondiente en un diagrama.
- a) ¿Cuál es la variable que se va a medir? ¿De qué tipo es?
- b) ¿Qué tipo de diagrama es el más adecuado? ¿Por qué?
5. El lenguaje es la herramienta propia del hombre para comunicarse con los otros miembros de una sociedad. Con el fin de comprobar el cumplimiento de la función comunicativa de su periódico en toda la sociedad cucuteña, el director de un destacado diario de la capital de Santander del Norte, adelantó una investigación que pretendía medir el alcance de la información contenida en el periódico con respecto al grado de comprensión que de ésta lograba el lector. Para tal efecto, se tomó una muestra de cien suscriptores del diario a quienes se pidió calificar tres tipos de escritos según el grado de dificultad de comprensión que presentaban. A continuación se especifica un poco más el estudio a través de cuatro tablas: la primera presenta la muestra estratificada por nivel socioeconómico de los suscriptores del periódico, la segunda indica los tipos de textos sometidos a calificación, la tercera indica la escala de calificaciones para el grado de dificultad de comprensión del texto y la última tabla presenta los resultados, de la calificación general a los tres textos, obtenidos según el nivel socioeconómico.

<b>Alto</b>	20	<b>A</b>	Temas de interés general	<b>+2</b>	Muy difícil
<b>Medio alto</b>	30	<b>B</b>	Temas sociales y políticos	<b>+1</b>	Difícil
<b>Medio bajo</b>	20	<b>C</b>	Temas especializados	<b>0</b>	No muy difícil
<b>Bajo</b>	30			<b>-1</b>	Fácil
				<b>-2</b>	Muy fácil

<b>Alto</b>	0	+1	+1	0	-1	-1	-1	+2	0	0
	0	+1	+1	-1	-1	-1	0	0	+1	-1
<b>Medio alto</b>	-1	-2	+2	+2	0	0	+1	+1	+1	0
	-1	-2	-2	0	0	+1	+1	+2	-1	0
<b>Medio bajo</b>	-2	-1	0	0	0	+1	0	0	0	+1
	-1	0	+1	0	0	+1	-1	-2	+2	+1
<b>Bajo</b>	0	0	+1	+1	0	-1	-1	-1	-1	0
	0	0	+1	-1	-2	+2	+2	0	0	0
<b>Bajo</b>	+1	+1	+1	+2	0	0	0	-1	0	0
	+1	+1	+1	+2	0	0	+1	+1	0	0

- a) ¿Cuál es el problema de estudio?
- b) ¿Cuál es la muestra de estudio?
- c) ¿Cuáles son las variables y de qué tipo son?



- d) Elabore una tabla de frecuencias para las calificaciones dadas y represente la información con un gráfico apropiado.
- e) Elabore tablas de frecuencias y represente en gráficos apropiados las calificaciones obtenidas según los cuatro niveles socioeconómicos considerados en el conjunto de datos presentado.
- f) Compare los cuatro gráficos y concluya sobre la relación entre el nivel socioeconómico y el grado de comprensión.
- g) ¿Puede concluir algo sobre la diferencia entre grado de comprensión y el tipo de texto presentado? Explique su respuesta.
6. La siguiente tabla muestra el número de hombres y mujeres en los diferentes departamentos de una empresa. Contabilidad (C), Estadística (E), Servicios generales (G), Tesorería (T), Sistemas (S), Ventas (V) y Dirección (D).

	C	E	G	T	S	V	D
Hombres	30	35	35	20	25	40	10
Mujeres	40	20	30	25	30	50	20

- a) Proponga un gráfico apropiado para representar el número de mujeres por departamento.
- b) Proponga un gráfico apropiado para representar el número de mujeres y hombres por departamento.
- c) ¿En qué departamento hay mayor número de empleados? ¿Cuál gráfica muestra más claramente este resultado?
- d) ¿En qué departamento hay mayor diferencia entre el número de hombres y mujeres? ¿Cuál gráfica muestra claramente este resultado?
- e) Para observar en qué departamento hay mayor porcentaje de mujeres, ¿qué tipo de gráfico se puede utilizar?
7. La tabla que se muestra más abajo se refiere al comercio de Colombia en el grupo Andino entre los años 1975 y 1979. Datos en millones de pesos. Proponga un gráfico en el que pueda comparar las exportaciones y las importaciones por año. ¿Tendría sentido presentar los datos en un gráfico de columnas o barras porcentuales? Explique.

Año	Exportaciones	Importaciones
1975	80	55
1976	90	40
1977	90	55
1978	140	80
1979	160	90

8. En la siguiente tabla se muestran las razones por las cuales 30 computadores configurados en la universidad se encuentran en el taller de mantenimiento. Las partes se compran a diferentes proveedores. Las razones por las cuales cada computador entró al taller se detallan en la siguiente tabla:

Nº	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
1	x			x	x		x			x
2			x		x		x		x	
3		x	x		x	x		x	x	
4	x		x				x		x	x
5						x				
6		x	x		x			x		
7					x		x		x	
8	x		x		x			x	x	x
9						x			x	
10		x	x				x		x	
11					x	x		x		
12				x			x			x
13		x	x		x	x		x		
14					x		x		x	
15	x	x		x			x		x	x

Nº	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
16	x		x			x				x
17		x		x	x			x	x	
18			x	x	x					
19		x	x		x			x	x	
20	x			x	x				x	x
21		x	x				x			
22		x	x			x				
23			x		x			x	x	
24	x			x	x				x	x
25		x	x						x	
26					x	x				
27	x			x	x			x		x
28		x	x				x		x	
29				x	x				x	
30	x		x				x			x

Donde las razones son:

R1 = Actualizar sistema operativo	R2 = Cambio de batería de la bios
R3 = Molesta la tarjeta de sonido	R4 = Molesta la tarjeta de video
R5 = Se desconfigura el teclado	R6 = Dañada la tarjeta de red
R7 = Escasa memoria	R8 = Daño del disco duro
R9 = Daño en unidad de CD	R10 = Daño en la unidad 3 ½ A

- a) Represente gráficamente esta información.
- b) ¿Sugiera en qué caso se deben tomar medidas correctivas?
9. El siguiente conjunto de datos corresponde al consumo en metros cúbicos de agua de 81 residencias de una ciudadela durante el segundo semestre del año 2006.

40,00	40,30	40,60	40,90	41,20	41,50	41,80	42,10	42,40
40,25	40,70	41,15	41,60	42,05	42,50	42,95	43,40	43,85
40,50	41,10	41,70	42,30	42,90	43,50	44,10	44,70	45,30
40,75	41,50	42,25	43,00	43,75	44,50	45,25	46,00	46,75
41,00	41,90	42,80	43,70	44,60	45,50	46,40	47,30	48,20
41,25	42,30	43,35	44,40	45,45	46,50	47,55	48,60	49,65
41,50	42,70	43,90	45,10	46,30	47,50	48,70	49,90	51,10
41,75	43,10	44,45	45,80	47,15	48,50	49,85	51,20	52,55
42,00	43,50	45,00	46,50	48,00	49,50	51,00	52,50	54,00

- a) Elabore tablas de frecuencias relativas y acumuladas.
- b) Represente de tres formas diferentes el conjunto de datos.

10. En un zocriadero destinado a la cría de chigüiros para exportación se ha descuidado la alimentación de los animales y se ha presentado un desarrollo inesperado en estos. Se han clasificado los animales en diez grupos, teniendo en cuenta el peso en kilogramos. La siguiente tabla muestra la cantidad de animales en cada categoría de pesos:

<b>Pesos 35.00 - 85.00</b>	<b>Cantidad de animales</b>	<b>Porcentaje de animales</b>	<b>Porcentaje acumulado</b>	<b>Marca de la categoría</b>
35.00 - 40.00	20			
40.10 - 45.00	25			
45.10 - 50.00	30			
50.10 - 55.00	10			
55.10 - 60.00	15			
60.10 - 65.00	20			
65.10 - 70.00	25			
70.10 - 75.00	35			
75.10 - 80.00	10			
80.10 - 85.00	10			

- a) Complete la tabla.
  - b) Estime cuántos animales pesan menos de 70.00 kilos.
  - c) Estime cuántos animales pesan menos de 58.00 kilos.
11. Indague como se elaboran los gráficos de: anillos, burbujas, radial y superficie. Proponga cuatro conjuntos de datos con los cuales tenga sentido la utilización de estos diagramas.
12. En el análisis multivariado de datos existen algunos métodos de representación gráfica como los dispersogramas, las curvas de Andrews y los rostros de Chernov. Investigue cual es la finalidad de estos gráficos y proponga conjuntos de datos para los que tenga sentido su utilización.
13. La siguiente tabla presenta la información correspondiente a cantidad de estudiantes matriculados en un postgrado de Estadística en una universidad del estado, durante el periodo 1995 - 2004. H=(Cantidad de hombres), M=(Cantidad de mujeres).

Año	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
H	15	18	20	25	20	20	18	10	22	24
M	25	25	22	15	15	17	22	35	25	20

Para la información organizada en la tabla, elabore un diagrama de perfiles, un diagrama de líneas y un diagrama de áreas.

14. La información que se presenta en la siguiente tabla corresponde a 20 pacientes atendidos en un consultorio pediátrico durante la primera semana del mes de junio. Se ha recabado la siguiente información: X=Edad (en meses), Y=Peso (en gramos), Z=Talla (en centímetros).

Paciente	X = Edad	Y = Peso	Z = Talla
Bebé 01	3	3500	55
Bebé 02	4	4500	60
Bebé 03	5	6000	60
Bebé 04	4	5000	55
Bebé 05	5	4500	60
Bebé 06	5	7500	65
Bebé 07	5	6500	65
Bebé 08	6	6000	65
Bebé 09	6	7000	55
Bebé 10	6	6500	70
Bebé 11	8	7500	70
Bebé 12	8	8000	75
Bebé 13	9	10000	70
Bebé 14	9	8500	80
Bebé 15	11	10000	85
Bebé 16	1	4000	45
Bebé 17	1	3500	45
Bebé 18	11	10500	90
Bebé 19	12	9500	85
Bebé 20	7	8000	70

- a) Elabore un diagrama de dispersión para las variables Y y Z.  
b) Elabore un diagrama de columnas 3D para las variables X y Z.  
c) Elabore un diagrama de perfiles para las variables X, Y y Z.
15. Una caja de compensación ha establecido 10 categorías para clasificar a sus afiliados de acuerdo a sus salarios, de tal manera que un afiliado queda clasificado en la categoría  $C_n$  si su salario se encuentra en el intervalo  $[535.000n, 535.000(n+1))$ , donde  $n=1,2,3,\dots,10$ . La cantidad de empleados en cada categoría es:

$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$C_{10}$
250	350	500	700	700	750	400	400	200	250

- a) ¿Qué porcentaje de afiliados tiene salarios entre \$2.500.000 y \$3.500.000?

- b) ¿Qué porcentaje de afiliados tiene salarios superiores o iguales a \$3.000.000?  
 c) Elabore un gráfico adecuado para representar la información de la caja de compensación.
16. Un estudiante contesta al azar las 10 preguntas de un examen, en donde cada pregunta tiene cuatro opciones de respuesta, pero solo una opción es la respuesta correcta. Elabore una tabla de frecuencias relativas en donde se clasifiquen todas las posibles formas de contestar el examen de acuerdo a la variable aleatoria  $X$  definida como “Cantidad de aciertos que puede obtener el estudiante”.
17. En un instituto que cuenta con 12 especialistas (3 fisiólogos, 4 genetistas, 5 inmunólogos) se seleccionará aleatoriamente un grupo de 4 para realizar una investigación en un laboratorio extranjero. Elabore una tabla de frecuencias relativas en donde los grupos se clasifiquen los posibles grupos de acuerdo a la cantidad de fisiólogos que queden en los grupos.
18. Una moneda que no tiene Cara y Sello sino 2 y 3 se lanza 10 veces. Se anota la sucesión de los 10 números resultantes (2's y 3's) y se suman estos 10 números. Elabore una tabla de frecuencias relativas en donde se clasifiquen los posibles resultados de acuerdo a la suma de los 10 números resultantes.
19. Se extraen simultáneamente 5 cartas de una baraja de 52 cartas y se observa la cantidad de Ases resultantes en la muestra extraída. Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras de acuerdo a la cantidad de Ases posibles en las muestras.
20. En un recipiente se encuentran 5 objetos: 3 buenos y 2 defectuosos. Se extrae una muestra de dos objetos, uno tras otro sin reemplazo. Sea  $X$  el número de objetos defectuosos en la primera extracción, y  $Y$  el número de objetos defectuosos en la segunda extracción.
- a) Elabore una tabla de frecuencias absolutas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $X$  y  $Y$ .
- b) Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $X$  y  $Y$ .
21. Una compañía ofrece dos seguros a sus clientes: seguro de automóvil y seguro de vivienda. Las pólizas para seguro de automóvil son de \$ 100, \$ 200 y \$ 300; mientras que las pólizas para seguro de vivienda son de \$ 50, \$ 150 y \$ 250. Sea  $X$  la cantidad pagada por la póliza de automóvil, y  $Y$  la cantidad pagada por la póliza de vivienda. La siguiente tabla muestra la cantidad de personas en los diferentes pares de seguros. Hallar la tabla de frecuencias relativas.

(X,Y)		Y		
		50	150	250
X	100	30	40	50
	200	60	20	20
	300	60	40	80

22. Se lanza 5 veces una moneda y se anotan las figuras que resultan (Cara, Sello). Sea  $X$  el número de caras en los dos primeros lanzamientos, y  $Y$  el número total de caras en los cinco lanzamientos. Elabore una tabla de frecuencias relativas en donde se clasifiquen los posibles resultados en los cinco lanzamientos de acuerdo a las variables  $X$  y  $Y$ .

23. Un dado balanceado se lanza dos veces y se anotan los números obtenidos. Sea  $X$  es el número de cuatros, y  $Y$  es el número de cincos, que se obtienen en dos lanzamientos del dado. Elabore una tabla de frecuencias relativas en donde se clasifiquen los posibles resultados en los dos lanzamientos de acuerdo a las variables  $X$  y  $Y$ .
24. Se quiere conformar una brigada de salud de 10 profesionales para atender una emergencia en una población que sufrió las inundaciones de dos ríos que la atraviesan. La secretaría de salud de la región dispone de 12 médicos, 15 enfermeras, 6 gastroenterólogos y 7 epidemiólogos. Se define  $X$  como el número de médicos en la brigada, y  $Y$  como el número de enfermeras en la brigada. Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles brigadas de acuerdo a las variables  $X$  y  $Y$ .
25. Una moneda que no tiene Cara y Sello sino 2 y 4 se lanza cuatro veces. Se anota el número que resulta a medida que se lanza la moneda. Sea  $X$  la suma acumulada en los dos primeros lanzamientos, y  $Y$  la suma acumulada en los cuatro lanzamientos. Elabore una tabla de frecuencias relativas en donde se clasifiquen los posibles resultados en los cuatro lanzamientos de acuerdo a las variables  $X$  y  $Y$ .
26. Un almacén tiene en existencia 30 componentes de cierto tipo, 8 de los cuales fueron proporcionados por el proveedor 1, 10 por el proveedor 2 y 12 por el proveedor 3. Se va a seleccionar una muestra de 6 componentes. Sea  $X$  el número de componentes seleccionados del proveedor 1,  $Y$  el número de componentes seleccionados del proveedor 2, y  $Z$  el número de componentes seleccionados del proveedor 3.
- Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $X$  y  $Y$ .
  - Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $X$  y  $Z$ .
  - Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $Y$  y  $Z$ .
27. Se seleccionan 4 repuestos para un bolígrafo de una caja que contiene 3 repuestos azules, 4 rojos y 5 negros. Supóngase que  $X$  representa el número de repuestos azules en la muestra seleccionada, y  $Y$  el número de repuestos rojos en la muestra seleccionada. Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles muestras de acuerdo a las variables  $X$  y  $Y$ .
28. En un grupo de 9 ejecutivos de una empresa hay 4 que son casados, 3 solteros y 2 divorciados. Se debe hacer una selección 3 ejecutivos. Sea  $X$  el número de ejecutivos casados en la selección,  $Y$  el número de ejecutivos solteros en la selección. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen las posibles selecciones de acuerdo a las variables  $X$  y  $Y$ .
29. Un dado balanceado se lanza 5 veces y se anotan los números obtenidos. Sea  $X$  es el número de cuatros en los tres primeros lanzamientos, y  $Y$  es el número de cuatros en los cinco lanzamientos. Elabore una tabla de frecuencias relativas en donde se clasifiquen los posibles resultados en los cinco lanzamientos de acuerdo a las variables  $X$  y  $Y$ .
30. Una urna contiene 6 bolas de colores diferentes: dos numeradas con el número 3 y cuatro numeradas con el número 6. Se seleccionan dos bolas simultáneamente. Sea  $X$  el menor de los números

seleccionados, y  $Y$  el mayor de los números seleccionados. Elabore una tabla de frecuencias relativas en donde se clasifiquen las posibles selecciones de acuerdo a las variables  $X$  y  $Y$ .

31. Una urna contiene 5 bolas de colores diferentes: dos numeradas con el número 2 y tres numeradas con el número 3. Se seleccionan tres bolas simultáneamente. Sea  $X$  el menor de los números seleccionados, y  $Y$  el mayor de los números seleccionados. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen las posibles selecciones de acuerdo a las variables  $X$  y  $Y$ .
32. Se lanza cuatro veces un dado y se anota el número que resulta en cada lanzamiento. Sea  $X$  la cantidad de números pares en los dos primeros lanzamientos, y  $Y$  la cantidad de números pares en los cuatro lanzamientos. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen los posibles resultados de acuerdo a las dos variables definidas.
33. De una baraja de 52 cartas se extraen 5 cartas, una tras otra sin reposición, y se va anotando el valor de la carta. Sea  $X$  la cantidad de Ases acumulados al finalizar la tercera extracción, y  $Y$  la cantidad de Ases acumulados al finalizar la quinta extracción. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen los posibles resultados de las cinco extracciones de acuerdo a las variables  $X$  y  $Y$ .
34. Se sacan simultáneamente 4 cartas de las 16 cartas de las letras J, Q, K y A, que tiene una baraja de 52 cartas. Sea  $X$  el número de Jotas, y  $Y$  el número de Ases que resultan en la muestra seleccionada. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen las posibles muestras resultantes de acuerdo a las variables  $X$  y  $Y$ .
35. De una baraja de 52 cartas se seleccionan simultáneamente 5 cartas, se anota valor y palo de las cartas seleccionadas, y no se reponen. Luego se seleccionan simultáneamente otras tres cartas de las 47 cartas restantes, se anota valor y palo de las cartas seleccionadas, y no se reponen. Sea  $X$  la cantidad de cartas de tréboles en la primera selección, y  $Y$  la cantidad de cartas de tréboles en la segunda selección. Elabore las tablas de frecuencias absolutas y relativas en donde se clasifiquen los posibles resultados de las dos selecciones hechas de acuerdo a las variables  $X$  y  $Y$ .

