

¿Qué es la estadística?

OBJETIVOS

Al concluir el capítulo, será capaz de:

1. Comprender la razón por la que estudia estadística.
2. Explicar los conceptos de *estadística descriptiva* y *estadística inferencial*.
3. Distinguir entre una *variable cualitativa* y una *variable cuantitativa*.
4. Describir la diferencia entre una *variable discreta* y una *variable continua*.
5. Distinguir entre los niveles de medición *nominal*, *ordinal*, *de intervalo* y *de razón*.



Usted se encuentra comprando un nuevo reproductor de música MP3, como el iPod de Apple. Los fabricantes indican la cantidad de canciones que almacena la memoria. Sin embargo, a usted le gustaría almacenar los musicales de Broadway, que duran más, por lo que le gustaría calcular cuántos musicales caben en su reproductor MP3. ¿Recogería información utilizando una muestra de una población? ¿Por qué razón? (véase ejercicio 8d) y objetivo 2).

Introducción

Hace más de cien años, H. G. Wells, escritor e historiador inglés, dijo que algún día el razonamiento cuantitativo sería tan importante para la gran mayoría de los ciudadanos como la capacidad de leer. No mencionó el área de los negocios, ya que la Revolución Industrial apenas iniciaba. No obstante, Wells tenía razón. Si bien la *experiencia en los negocios*, cierta *habilidad* para hacer *pronósticos razonados* y la *intuición* constituyen atributos fundamentales en los gerentes con éxito, los problemas que en la actualidad se presentan en los negocios tienden a ser demasiado complejos como para tomar decisiones sólo a partir de estos criterios.



Una de las herramientas utilizadas para tomar decisiones es la estadística. De la estadística no sólo se sirve la gente dedicada a los negocios; en nuestra vida cotidiana también aplicamos conceptos estadísticos. Por ejemplo, para comenzar el día, abra la regadera y deje correr el agua unos segundos. Enseguida moje su mano para percatarse si la temperatura es adecuada o *decidir* si abre más la llave del agua caliente o la del agua fría. Ahora suponga que está en una tienda comercial y quiere comprar una pizza congelada. Dos marcas tienen un puesto de promoción, y cada una le ofrece una pequeña rebanada. Después de probar, *decide* cuál comprar. En ambos ejemplos, usted toma la decisión y elige lo que hará, a partir de una muestra.

Las empresas enfrentan situaciones similares. Por ejemplo, Kellogg Company debe garantizar que la cantidad promedio de Raisin Bran en una caja de 25.5 gramos cumpla con la cantidad especificada en la etiqueta. Para hacerlo fija un peso *objetivo* un poco más alto que la cantidad que dice en la etiqueta. Las cajas se pesan después de llenarse. La báscula indica la distribución de los pesos del contenido por hora, así como la cantidad de cajas *desechadas* por no cumplir con las especificaciones de la etiqueta en el transcurso de dicha hora. El Departamento de Control de Calidad también selecciona de forma aleatoria muestras de la línea de producción y verifica la calidad del producto y el peso de la caja. Si es significativa la diferencia entre el peso promedio del producto y el peso objetivo o el porcentaje de cajas desechadas es muy alto, el proceso se ajusta.

Alan Greenspan, ex presidente del Departamento de la Reserva Federal de Estados Unidos, conoce y entiende la importancia de las herramientas y técnicas estadísticas para proporcionar información precisa y oportuna que sirva para hacer declaraciones públicas con la fuerza de movilizar mercados bursátiles globales e influir en la política. Al hablar frente al National Skills Summit, el doctor Greenspan dijo: "A los trabajadores se les debe preparar no sólo con conocimientos técnicos, sino también con la capacidad de crear, analizar y transformar la información, así como de relacionarse adecuadamente con otras personas. Es decir, deben ser capaces de separar los hechos de las opiniones y enseguida organizarlos en su forma más conveniente para analizar la información".

Como estudiante de administración o de economía, requerirá conocimientos básicos y habilidad para organizar, analizar y transformar datos, así como para presentar la información. En esta obra, aprenderá las técnicas y métodos estadísticos básicos que mejorarán su destreza para tomar buenas decisiones personales y de naturaleza administrativa.

¿Por qué se debe estudiar estadística?

Si revisa el plan de estudios de la universidad, se dará cuenta de que varios programas universitarios incluyen la estadística. ¿Por qué razón? ¿Cuáles son las diferencias entre los cursos de estadística que se imparten en la Facultad de Ingeniería, los Departamentos de Psicología o Sociología en la Escuela de Artes Liberales y la Facultad de Administración? La diferencia principal consiste en los ejemplos que se utilizan. El contenido del curso es el mismo. En la Facultad de Administración el interés son cuestiones como las utilidades, las horas de trabajo y los salarios. A los psicólogos les importan los resultados de las pruebas, y a los ingenieros la cantidad de unidades que fabrica determinada máquina. No obstante, en los tres casos, el interés se centra en el valor típico y la variación que experimentan los datos. También existe una diferencia en el nivel de

Ejemplos de por qué se estudia la estadística

los cálculos matemáticos requeridos. Un curso de estadística para ingenieros incluye el cálculo. Los cursos de estadística en las facultades de administración y pedagogía, por lo general, se imparten desde el punto de vista de las aplicaciones. Si usted ya estudió álgebra en la escuela secundaria, manejará adecuadamente la matemática que se emplea en el texto.

Entonces, ¿por qué se requiere la estadística en muchas empresas importantes? La primera razón consiste en que la información numérica prolifera por todas partes. Revise los periódicos (*USA Today*), revistas de noticias (*Time*, *Newsweek*, *U.S. News y World Report*), revistas de negocios (*BusinessWeek*, *Forbes*), revistas de interés general (*People*), revistas para mujeres (*Ladies*, *Home Journal* o *Elle*) o revistas deportivas (*Sports Illustrated*, *ESPN The Magazine*), y quedará abrumado con la cantidad de información numérica que contienen.

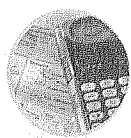
He aquí algunos ejemplos:

- En 2003 el ingreso familiar típico en Estados Unidos era de \$43 318. En el caso de las familias del noreste el ingreso típico era de \$46 742; en la región central de Estados Unidos de \$44 732; en el sur era de \$39 823 y en la región occidental de \$46 820. La información más reciente se puede localizar en la página <http://www.census.gov/hhes/income>.
- En julio de 2005, Boeing informó la entrega de 155 aeronaves para el periodo del 1 de enero de 2005 al 30 de junio de 2005. Esto representó un total de 113 naves Boeing 737 entregadas durante el periodo, y Southwest Airlines fue el comprador más importante con 22 aeronaves adquiridas. Verifique la información más reciente en la página de Boeing www.boeing.com, escriba *orders and deliveries* (órdenes y entregas) en el recuadro de búsqueda y, de la lista de posibles sitios de la red, seleccione el que ofrezca la información más reciente de órdenes y entregas. A la izquierda de esta página aparece una lista del mapa de ubicación de las órdenes, del cual puede elegir Current Year Deliveries.
- *USA Today* (www.usatoday.com) imprime *instantáneas*, que son el resultado de encuestas llevadas a cabo por diversas agencias de investigación, fundaciones y por el gobierno federal. Por ejemplo, muchos prefieren el correo electrónico en lugar del correo postal. Sin embargo, de acuerdo con una encuesta reciente, el Servicio Postal de Estados Unidos informa que 67% de los adultos señalan que el correo ordinario resulta más personal que el correo electrónico; 56% indica que les causa placer recibir el correo normal y 55% espera con ansias abrir el correo.

Una segunda razón para inscribirse en un curso de estadística estriba en que las técnicas estadísticas se emplean para tomar decisiones que afectan la vida diaria. Es decir que éstas influyen en su bienestar. He aquí algunos ejemplos:

- Las compañías de seguros utilizan el análisis estadístico para establecer tarifas de seguros de casas, automóviles, de vida y de servicio médico. Las tablas disponibles contienen cálculos aproximados de que a una mujer de 20 años de edad le queden 60.25 años de vida; a una mujer de 87 años le queden 4.56 años de vida y a un hombre de 50 años 27.85. Las primas de seguros de vida se establecen con base en estos cálculos de expectativas de vida. Estas tablas se encuentran disponibles en www.ssa.gov/OACT/STATS/table4cb.html (este sitio acepta mayúsculas).
- La Agencia de Protección del Ambiente está interesada en la calidad del agua del lago Erie, entre otros. Con periodicidad toma muestras de agua para determinar el nivel de contaminación y mantener la norma de calidad.
- Los investigadores médicos estudian los índices de curación de enfermedades mediante la utilización de diferentes fármacos y diversos tratamientos. Por ejemplo, ¿cuál es el efecto que resulta de operar cierto tipo de lesión de rodilla o de aplicar terapia física? Si se ingiere una aspirina cada día, ¿se reduce el riesgo de un ataque al corazón?

Una tercera razón para inscribirse radica en que el conocimiento de sus métodos facilita la comprensión de la forma en que se toman decisiones y proporciona un entendimiento más claro de cómo le afectan.



Estadística en acción

Centre su atención en el título *Estadística en acción*. Lea con cuidado para obtener una idea de la amplia gama de aplicaciones de la estadística en la administración, economía, enfermería, cumplimiento de la ley, deportes y otras disciplinas.

- En 2005, *Forbes* publicó una lista de los estadounidenses más ricos. William Gates, fundador de Microsoft Corporation, es el hombre más rico. Su fortuna se calcula en 46 500 millones de dólares (www.forbes.com).
- En 2005 las cuatro compañías estadounidenses con mayores ingresos fueron ExxonMobil, General Motors, Ford y Chevron (www.industryweek.com).
- En Estados Unidos un típico estudiante graduado de la escuela secundaria gana 1.2 millones de dólares en el transcurso de su vida; un típico graduado universitario gana dos 2.1 millones de dólares y un típico posgraduado gana 2.5 millones de dólares (usgovinfo.about.com/library/weekly/aa072602a.htm).

Sin importar el empleo que haya elegido, usted encarará la necesidad de tomar decisiones en las que saber hacer un análisis de datos resultará de utilidad. Con el fin de tomar una decisión informada, será necesario llevar a cabo lo siguiente:

1. Determinar si existe información adecuada o si requiere información adicional.
2. Reunir información adicional, si se necesita, de manera que no se obtengan resultados erróneos.
3. Resumir los datos de manera útil e informativa.
4. Analizar la información disponible.
5. Obtener conclusiones y hacer inferencias al mismo tiempo que se evalúa el riesgo de tomar una decisión incorrecta.

Los métodos estadísticos expuestos en la obra le proporcionarán un esquema del proceso de toma de decisiones.

En suma, existen por lo menos tres razones para estudiar estadística: 1. Los datos proliferan por todas partes: 2. Las técnicas estadísticas se emplean en la toma de decisiones que influyen en su vida: 3. Sin importar la carrera que elija, tomará decisiones profesionales que incluyan datos. Una comprensión de los métodos estadísticos permite tomar decisiones con mayor eficacia.

¿Qué se entiende por estadística?

¿Cuál es la definición de *estadística*? Nos topamos con ella en el lenguaje cotidiano. En realidad, posee dos significados: en su acepción más común, la estadística se refiere a información numérica. Algunos ejemplos son el sueldo inicial de los graduados de universidad, el número de muertes provocadas por el alcoholismo el año pasado, el cambio en el promedio industrial Dow Jones de ayer a hoy y la cantidad de cuadrangulares conectados por los Chicago Cubs durante la temporada 2005. En estos ejemplos las estadísticas refieren un valor o un porcentaje. Otros ejemplos incluyen:

- El automóvil típico en Estados Unidos viaja 17 858 kilómetros al año; el autobús, 15 049 kilómetros al año y el camión, 22 433 kilómetros al año. En Canadá, la información correspondiente es de 16 687 kilómetros en el caso de los automóviles; de 31 895 en el caso de los autobuses y de 11 264.60 en el caso de los camiones.
- El tiempo promedio de espera para asesoría técnica es de 17 minutos.
- La longitud promedio del ciclo económico de negocios desde 1945 es de 61 meses.

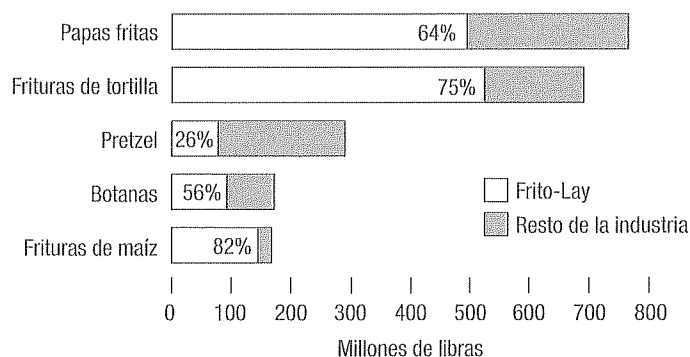
Todos éstos constituyen ejemplos de **estadísticas**. Una colección de información numérica recibe el nombre de **estadísticas**.

A menudo la información estadística se presenta en forma gráfica, la cual es útil porque capta la atención del lector e incluye una gran cantidad de información. Por ejemplo, la gráfica 1.1 muestra el volumen y las acciones de Frito-Lay respecto de las principales categorías de papas fritas y botanas en los supermercados de Estados Unidos. Es suficiente un vistazo para descubrir que se vendieron cerca de 800 millones de libras de papas fritas y que Frito-Lay vendió 64% del total. Observe, asimismo, que Frito-Lay posee 82% del mercado de frituras de maíz.

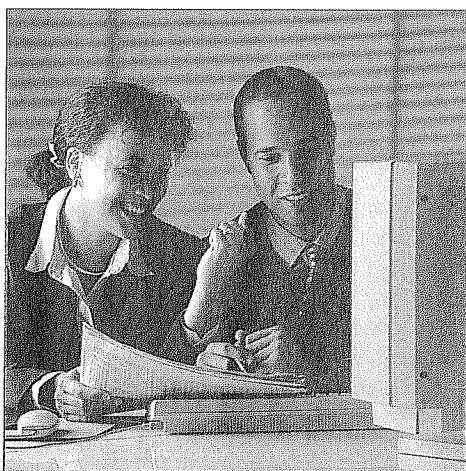
Como verá, la estadística tiene un significado mucho más amplio que la simple recolección y publicación de información numérica. Atienda a la siguiente definición de estadística:

ESTADÍSTICA Ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de propiciar la toma de decisiones más eficaz.

Como lo sugiere la definición, el primer paso en el estudio de un problema consiste en recoger datos relevantes. Éstos deben organizarse de alguna forma y, tal vez, representarse en una gráfica, como la gráfica 1.1. Sólo después de haber organizado los



GRÁFICA 1.1 Volumen y acciones de Frito-Lay en las principales categorías de botanas en los supermercados de Estados Unidos



datos es posible analizarlos e interpretarlos. He aquí algunos ejemplos de la necesidad de recoger datos.

- Los analistas dedicados a la investigación que trabajan para Merrill Lynch evalúan muchas facetas de determinadas acciones antes de hacer una recomendación de *compra* o *venta*. Recogen los datos de ventas anteriores de la compañía y calculan futuras ganancias. Antes de hacer recomendaciones, también consideran otros factores, como la demanda mundial prevista de los productos de la compañía, la fuerza de la competencia y el efecto del nuevo contrato con la administración sindical.
- El departamento de marketing de Colgate-Palmolive Co., fabricante de productos de limpieza, tiene la responsabilidad de hacer recomendaciones sobre la posible rentabilidad de un grupo de jabones faciales recién creados, con aromas frutales, como uva, naranja y piña. Antes de tomar la última decisión, los promotores de mercado examinarán el producto en diversos mercados. Es decir, los anunciarán y venderán en Topeka, Kansas y Tampa, Florida. A partir de los resultados de esta prueba de marketing en estas dos regiones, Colgate-Palmolive decidirá si vende o no los jabones en todo el país.
- El Gobierno está interesado en la situación actual y en el pronóstico de las tendencias económicas. Por lo que lleva a cabo una gran cantidad de encuestas para determinar la confianza del consumidor y el punto de vista de los administradores en lo que se refiere a ventas y producción para los siguientes doce meses. Los índices, como el índice de precios al consumidor (IPC), se elaboran cada mes para calcular la inflación. La información acerca de las ventas en tiendas departamentales, programas de vivienda, volumen de acciones y producción industrial son sólo algunos de los cientos de factores que se toman en cuenta al establecer la base de las proyecciones. Los bancos emplean estas proyecciones para determinar su tasa principal de préstamos, y el Departamento de la Reserva Federal las emplea para tomar decisiones sobre el nivel de control que aplicará al suministro de dinero.
- Los administradores deben tomar decisiones referentes a la calidad de sus productos o servicios. Por ejemplo, los consumidores se comunican con las compañías de software para solicitar asesoría técnica cuando no pueden resolver algún problema. El tiempo que un consumidor debe esperar para que un asesor técnico conteste la llamada constituye una medida de la calidad del servicio que se le brinda. Una compañía de software podría establecer un minuto como objetivo del tiempo representativo de respuesta. Entonces la compañía recabaría y analizaría los datos relativos al tiempo de respuesta. ¿Difiere el tiempo representativo de respuesta cierto día de la semana o parte de un día? Si los tiempos de respuesta se están creciendo, los administradores podrían tomar la decisión de aumentar la cantidad de asesores técnicos a ciertas horas del día o de la semana.

Tipos de estadística

Por lo general, el estudio de la estadística se divide en dos categorías: la estadística descriptiva y la estadística inferencial.

Estadística descriptiva

Es la ciencia que “recoge, organiza, presenta, analiza... datos”. Esta parte de la estadística recibe el nombre de **estadística descriptiva**.

ESTADÍSTICA DESCRIPTIVA Método para organizar, resumir y presentar datos de manera informativa.

Por ejemplo, el gobierno de Estados Unidos informa que en 1960, la población de este país fue de 179 323 000; en 1970, de 203 302 000; en 1980, de 226 542 000; en 1990, de 248 709 000 y en 2000 de 265 000 000. Esta información representa una estadística descriptiva. Se trata de estadística descriptiva si calcula el crecimiento porcentual de una década a otra. Sin embargo, *no* sería de naturaleza descriptiva si utilizara estos datos para calcular la población de Estados Unidos en el año 2010 o el crecimiento porcentual de 2000 a 2010. ¿Por qué? Dichas estadísticas no se están utilizando para hacer un resumen de poblaciones del pasado, sino para calcular poblaciones en el futuro. Los siguientes son ejemplos de estadística descriptiva.

- Hay un total de casi 68 859 kilómetros de carreteras interestatales en Estados Unidos. El sistema interestatal representa apenas 1% del total de carreteras de la nación, aunque alberga a más de 20% del tránsito. La más larga es la autopista I-90, que va de Boston a Seattle, una distancia de 4 957.32 kilómetros. La más corta es la I-878, localizada en Nueva York, cuya longitud es de 1.12 kilómetros. Alaska no cuenta con carreteras interestatales; Texas posee la mayor cantidad de kilómetros interestatales, 3 232, y Nueva York tiene la mayor parte de las rutas interestatales, 28 en total.
- De acuerdo con la Agencia de Estadística Laboral, en enero de 2006 el salario promedio por hora de los obreros era de \$17.73. Revise la información reciente sobre salarios y productividad de los trabajadores estadounidenses en la página de la Agencia de Estadística Laboral localizada en <http://www.bls.gov/home.htm>, seleccione Average Hourly Earnings.

Una masa de datos desorganizados —como el censo de población, los salarios semanales de miles de programadores de computadoras y las respuestas de 2000 votantes registrados para elegir presidente de Estados Unidos— resulta de poca utilidad. No obstante, las técnicas de la estadística descriptiva permiten organizar esta clase de datos y darles significado. Los datos se ordenan en una **distribución de frecuencia** (en el capítulo 2 se estudia este procedimiento). Se emplean diversas **clases de gráficas** para describir datos; en el capítulo 4 también se incluyen diversas formas básicas de gráficas.

Las medidas específicas de localización central, como la media, describen el valor central de un grupo de datos numéricos. Para describir la proximidad de un conjunto de datos en torno al promedio se emplean diversas medidas estadísticas. Estas medidas de tendencia central y dispersión se estudian en el capítulo 3.

Estadística inferencial

El segundo tipo es la **estadística inferencial**, también denominada **inferencia estadística**. El principal interés respecto de la estadística inferencial tiene que ver con encontrar algo relacionado con la población a partir de una muestra de dicha población. Por ejemplo, una encuesta reciente mostró que solamente 46% de los estudiantes del último grado de secundaria podían resolver problemas que incluyeran fracciones, decimales y porcentajes. Además, sólo 77% de los estudiantes de último año de secundaria pudo sumar correctamente el costo de una ensalada, una hamburguesa, unas papas fritas y un refresco de cola, que figuraban en el menú de un restaurante. Ya que éstas son

inferencias relacionadas con una población (todos los estudiantes de último grado de secundaria), basadas en datos de la muestra, se trata de estadística inferencial. Se podría considerar a la estadística inferencial como la *mejor conjetura* que es posible obtener del valor de una población sobre la base de la información de la muestra.

ESTADÍSTICA INFERENCIAL Métodos empleados para determinar una propiedad de una población con base en la información de una muestra.

Preste atención a las palabras *población* y *muestra* en la definición de estadística inferencial. Con frecuencia hacen referencia a la población que vive en Estados Unidos o a la población de 1 310 millones de habitantes de China. No obstante, en estadística, la palabra *población* posee un significado más amplio. Una **población** puede constar de *individuos* —como los estudiantes matriculados de la Universidad Estatal de Utah, los estudiantes de Contabilidad 201 o los presidentes de las compañías de Fortune 500—. También puede consistir en *objetos*, tales como las llantas Cobra G/T producidas en Cooper Tire and Rubber Company en la planta de Findlay, Ohio; las cuentas por cobrar al finalizar octubre por Lorraine Plastics, Inc.; o los reclamos de seguro de automóvil archivados durante el primer trimestre de 2006 en la Oficina Regional del Noreste de State Farm Insurance. Las *medidas* de interés podrían ser los resultados en el primer examen de los estudiantes de Contabilidad 201, el desgaste de la banda de rodamiento de las llantas Cooper, el monto en dólares de las notas por cobrar de Lorraine Plastics o la cantidad de reclamos de seguro de automóvil en State Farm. De esta manera, desde una perspectiva estadística una población no siempre tiene que ver con personas.

POBLACIÓN Conjunto de individuos u objetos de interés o medidas obtenidas a partir de todos los individuos u objetos de interés.

Con el objeto de inferir algo sobre una población, lo común es que tome una **muestra** de la población.

MUESTRA Porción o parte de la población de interés.

Razones por las que se toman muestras

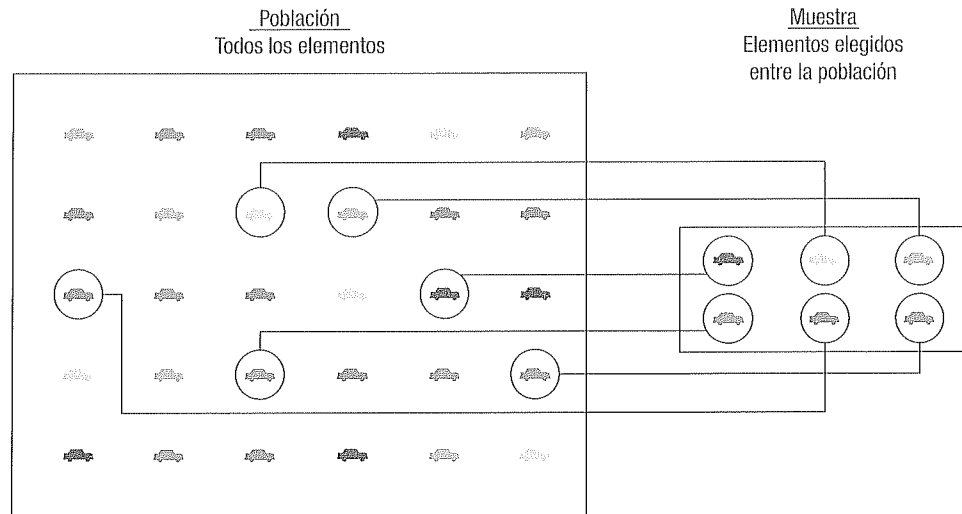
¿Por qué tomar una muestra en lugar de estudiar a cada miembro de la población? Una muestra de votantes registrados se hace necesaria en virtud de los costos prohibitivos de ponerse en contacto con millones de electores antes de una elección. Las pruebas en el trigo acerca de la humedad que lo destruye, hacen imprescindible la toma de una muestra. Si los catadores de vino probaran todo el vino, no quedaría una gota para vender. En la práctica resulta imposible que unos cuantos biólogos marinos capturen y rastreen a todas las focas en el océano. (Éstas y otras razones para tomar muestras se estudian en el capítulo 8.)

La toma de muestras para aprender algo sobre una población es de uso frecuente en administración, agricultura, política y acciones de gobierno, según lo muestran los siguientes ejemplos:

- Las cadenas de televisión hacen un monitoreo continuo de la popularidad de sus programas contratando a Nielsen y a otras organizaciones con el fin de que éstas tomen muestras sobre las preferencias de los teleespectadores. Por ejemplo, en una muestra de 800 personas que ven el televisor a la hora de mayor audiencia, 320, o 40%, señaló que vio *CSI (Crime Scene Investigation)* la semana pasada. Estos índices de audiencia se emplean para establecer tarifas de publicidad o para suspender programas.
- Gamous and Associates, una firma de contadores públicos, realiza una auditoría a Pronto Printing Company. Para comenzar, la firma contable elige una muestra aleatoria de 100 facturas y verifica la exactitud de cada factura. Por lo menos hay un error en cinco facturas; por consiguiente, la firma de contadores calcula que 5% de la población de facturas contiene un error por lo menos.
- Una muestra aleatoria de 1 260 graduados de marketing de escuelas que imparten la carrera en cuatro años mostró que su sueldo inicial promedio era de \$42 694. Por

tanto, se estima que el sueldo inicial promedio de todos los graduados de contabilidad de instituciones que imparten la carrera en cuatro años es de \$42 694.

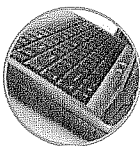
La relación entre una muestra y una población se presenta abajo. Por ejemplo, desea calcular los kilómetros promedio por litro de los vehículos SUV (*sport utility vehicles*). Se eligen seis SUV de la población. Se emplea la cantidad promedio de KPL (kilómetros por litro) de los seis para calcular la cantidad de KPL en el caso de la población.



Le recomendamos que realice el ejercicio de autoevaluación.

Enseguida aparece un ejercicio de autoevaluación. Estos ejercicios se encuentran intercalados en cada capítulo. Someten a prueba su comprensión del material precedente. La respuesta y método de solución aparecen al final del capítulo. La respuesta a la siguiente autoevaluación se encuentra en la página 19. El lector debe intentar resolverlos y después comparar su respuesta.

Autoevaluación 1.1



Las respuestas se localizan al final del capítulo.

La empresa de publicidad con sede en Atlanta, Brandon and Associates, solicitó a una muestra de 1 960 consumidores que probaran un platillo con pollo recién elaborado por Boston Market. De las 1 960 personas de la muestra, 1176 dijeron que comprarían el alimento si se comercializaba.

- ¿Qué podría informar Brandon and Associates a Boston Market respecto de la aceptación en la población del platillo de pollo?
- ¿Es un ejemplo de estadística descriptiva o estadística inferencial? Explique su respuesta.

Tipos de variables

Existen dos tipos básicos de variables: 1) cualitativas y 2) cuantitativas (véase gráfica 1.2). Cuando la característica que se estudia es de naturaleza no numérica, recibe el nombre de **variable cualitativa** o **atributo**. Algunos ejemplos de variables cualitativas son el género, la filiación religiosa, tipo de automóvil que se posee, estado de nacimiento y color de ojos. Cuando los datos son de naturaleza cualitativa, importa la cantidad o proporción que caen dentro de cada categoría. Por ejemplo, ¿qué porcentaje de la población tiene ojos azules? ¿Cuántos católicos o cuántos protestantes hay en Estados Unidos? ¿Qué porcentaje del total de automóviles vendidos el mes pasado eran SUV? Los datos cualitativos se resumen en tablas o gráficas de barras (capítulo 2).

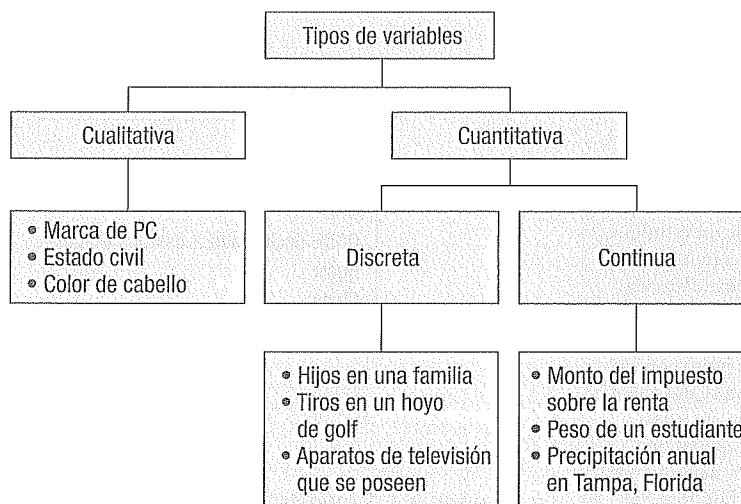
Cuando la variable que se estudia aparece en forma numérica, la variable se denomina **variable cuantitativa**. Ejemplos de variables cuantitativas son el saldo en su cuenta de cheques, las edades de los presidentes de la compañía, la vida de la batería de un automóvil —aproximadamente 42 meses— y el número de hijos que hay en una familia.

Variable cualitativa

Variable cuantitativa

Las variables cuantitativas pueden ser discretas o continuas. Las **variables discretas** adoptan sólo ciertos valores y existen *vacíos* entre ellos. Ejemplos de variables discretas son el número de camas en una casa (1, 2, 3, 4, etc.); el número de automóviles que en una hora usan la Salida 25, carretera I-4, en Florida, cerca del Walt Disney World (326, 421, etc.) y el número de estudiantes en cada sección de un curso de estadística (25 en la sección A, 42 en la sección B y 18 en la sección C). Aquí se cuenta, por ejemplo, el número de automóviles que arriban a la Salida 25, carretera I-4, y el número de estudiantes de estadística en cada sección. Observe que en una casa hay 3 o 4 camas, pero no 3.56. Por consiguiente, existe un *vacío* entre los valores posibles. Las variables discretas son el resultado de una relación numérica.

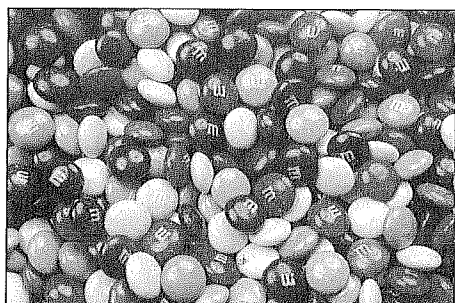
Las observaciones de una **variable continua** toman cualquier valor dentro de un intervalo específico. Ejemplos de variables continuas son la presión del aire en una llanta y el peso de un cargamento de tomates. Otros ejemplos son la cantidad de cereal con pasas que contiene una caja y la duración de los vuelos de Orlando a San Diego. El promedio de puntos al graduarse (PPG) constituye una variable continua. Podría expresar el PPG de determinado estudiante como 3.2576952. Se acostumbra redondear a 3 lugares decimales (3.258). Por lo general las variables continuas son el resultado de mediciones.



GRÁFICA 1.2 Resumen de los tipos de variables

Niveles de medición

Los datos se clasifican por niveles de medición. El nivel de medición de los datos rige los cálculos que se llevan a cabo con el fin de resumir y presentar los datos. También determina las pruebas estadísticas que se deben realizar. Por ejemplo, en una bolsa de M&M hay lunetas de seis diferentes colores. Suponga que asigna el 1 al café, el 2 al amarillo, el 3 al azul, el 4 al naranja, el 5 al verde y el 6 al rojo. Sume la cantidad de lunetas que hay en una bolsa, la divide entre el número de lunetas e informa que el color promedio es 3.56. ¿Significa que el color promedio es azul o anaranjado? Desde luego que no. Otro ejemplo, en la pista de una escuela secundaria hay ocho competidores para la carrera de 400 metros. Para indicar el orden en que llegan a la meta dice que la media es de 4.5. ¿Qué revela este promedio? ¡Nada! En ambos casos, no se empleó adecuadamente el nivel de medición.



De hecho, existen cuatro niveles de medición: nominal, ordinal, de intervalo y de razón. La medición más baja, o más primaria, corresponde al nivel nominal. La más alta, o el nivel que proporciona la mayor información relacionada con la observación, es la medición de razón.

Datos de nivel nominal

En el caso del **nivel nominal** de medición, las observaciones acerca de una variable cualitativa sólo se clasifican y cuentan. No existe una forma particular para ordenar las etiquetas. La clasificación de los seis colores de las lunetas de chocolate de leche M&M constituye un ejemplo del nivel nominal de medición. Simplemente se clasifican las lunetas por color. No existe un orden natural. Es decir, no presenta primero las lunetas cafés, las anaranjadas o las de cualquier color. El género representa otro ejemplo del nivel nominal de medición. Suponga que hace un conteo de los estudiantes que entran a un partido de fútbol con credencial e informa cuántos son hombres y cuántas mujeres. Podría presentar primero a los hombres o a las mujeres. Para el nivel nominal, la medición consiste en contar. La tabla 1.1 muestra un análisis de las fuentes de suministro mundial de petróleo. La variable de interés se refiere al país o región. Se trata de una variable de nivel nominal porque registra la información de acuerdo con la fuente de suministro del petróleo y no existe orden natural. No se confunda por el hecho de que la variable se resume informando la cantidad de barriles producidos por día.



Estadística en acción

¿Dónde tiene sus orígenes la estadística? En 1662 John Graunt publicó el artículo "Natural and Political Observations Made upon Bills of Mortality". Las observaciones del autor eran el resultado de un estudio y análisis de una publicación religiosa semanal llamada *Bill of Mortality*, la cual incluía nacimientos, bautizos y muertes junto con sus causas. Graunt se dio cuenta de que *Bills of Mortality* representaba apenas una fracción de los nacimientos y muertes en Londres. Sin embargo, utilizó los datos para llegar a conclusiones relativas al impacto de las enfermedades, como la peste, en la población. Su lógica constituye un ejemplo de inferencia estadística. Su análisis e interpretación de los datos marcan el inicio de la estadística.

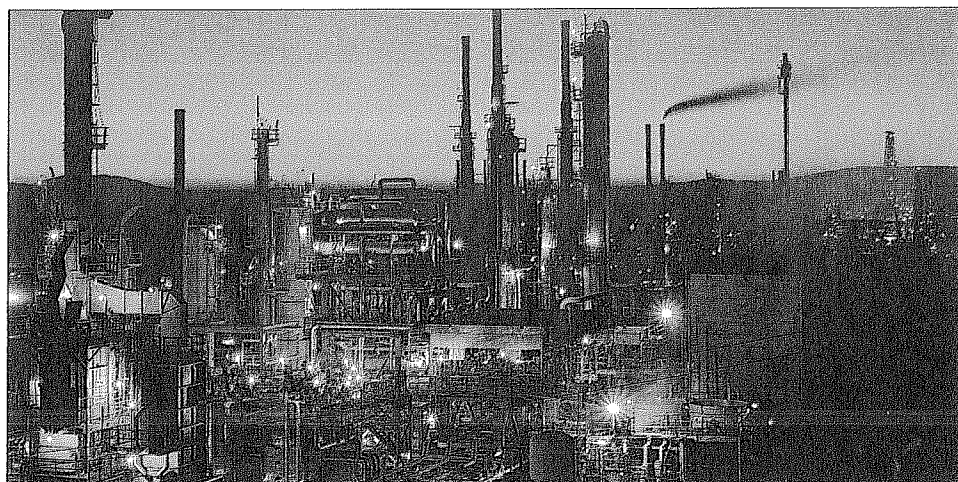
TABLA 1.1 Suministro mundial de petróleo para 2004

Fuente	Millones de barriles diarios	Porcentaje
OPEP	32.91	39.7
OCDE (incluyendo a Estados Unidos)*	22.76	27.4
Rusia	11.33	13.7
China	3.62	4.4
Otra	12.35	14.9
	82.97	100.1

*El promedio diario en Estados Unidos es de 8.69 millones de barriles, o 10.5% del total.

La tabla 1.1 muestra el rasgo esencial de la escala nominal de medición: no existe un orden particular en las categorías.

Con el fin de procesar datos referentes a la producción de petróleo, al género, al empleo por industria, etc., a menudo las categorías se codifican con los números 1, 2, 3,



etcétera: el 1 representa a la OPEP; el 2, a la OCDE, por ejemplo. Esto facilita el cálculo con la ayuda de la computadora. Sin embargo, aunque ha asignado números a las diversas categorías, esto no le autoriza a realizar operaciones con los números. Por ejemplo, $1 + 2$ no es igual a 3, es decir que OPEP + OCDE no es igual a Rusia. En resumen, los datos de nivel nominal poseen las siguientes propiedades:

1. Las categorías de datos se encuentran representadas por etiquetas o nombres.
2. Aun cuando las etiquetas se codifiquen con números, las categorías de datos no tienen ningún orden lógico.

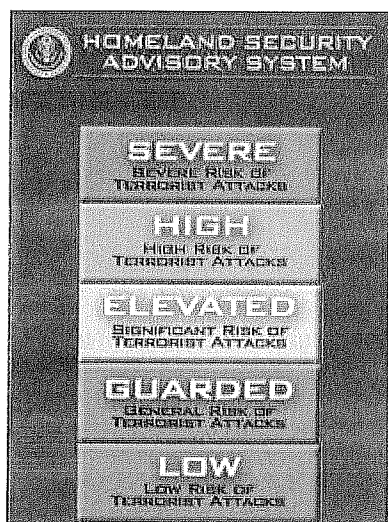
Datos de nivel ordinal

El nivel inmediato superior de datos es el **nivel ordinal**. La tabla 1.2 contiene las calificaciones que los alumnos del profesor James Bruner le otorgaron después de un curso de introducción a las finanzas. Cada estudiante de la clase respondió la pregunta: "En términos generales, ¿cómo calificas al profesor del curso?" La calificación variable ilustra el uso de la escala ordinal de medición. Una calificación es *más alta* o *mejor*, que la siguiente: *superior* es mejor que *bueno*, *bueno* es mejor que *promedio*, etc. Sin embargo, no es posible distinguir la magnitud de las diferencias entre los grupos. ¿La diferencia entre *superior* y *bueno* es la misma que entre *malo* e *inferior*? No es posible afirmarlo. Si sustituye 5 por *superior* y 4 por *bueno*, concluirá que la calificación *superior* es mejor que la calificación *bueno*, pero si añade una calificación de *superior* y una de *bueno* no espere que el resultado tenga significado. Además, no debe concluir que la calificación de *bueno* (calificación de 4) sea necesariamente dos veces más alta que *malo* (calificación de 2). Sólo tendrá claro que la calificación *bueno* es mejor que la calificación *malo*; no en qué grado es mejor calificación.

TABLA 1.2 Calificaciones a un profesor de finanzas

Calificación	Frecuencia
Superior	6
Bueno	28
Promedio	25
Malo	12
Inferior	3

Otro ejemplo de datos de nivel ordinal es el *Homeland Security Advisory System*. El Departamento de Seguridad Nacional publica información relativa al riesgo de que las autoridades federal, estatal y local, así como los estadounidenses, sean víctimas de ataques terroristas. A la izquierda aparecen los primeros cinco niveles de riesgo, que van del más bajo al más alto y se incluye una descripción y códigos de colores.



Éste es un ejemplo de la escala ordinal, ya que conoce el orden o los grados de los niveles de riesgo —el naranja es superior al amarillo—, aunque la diferencia en cuanto a riesgo no es necesariamente la misma. En otras palabras, la diferencia en cuanto al nivel de riesgo entre el amarillo y el naranja no es la misma que la existente entre el verde y el azul. Consulte los niveles actuales de riesgo y conozca más sobre los diversos niveles en la siguiente dirección: www.whitehouse.gov/homeland.

En resumen, las propiedades del nivel ordinal de los datos son las siguientes:

1. Las clasificaciones de los datos se encuentran representadas por conjuntos de etiquetas o nombres (alto, medio, bajo), las cuales tienen valores relativos.
2. En consecuencia, los valores relativos de los datos se pueden clasificar u ordenar.

Datos de nivel de intervalo

El **nivel de intervalo** de medición es el nivel inmediato superior. Incluye todas las características del nivel ordinal, pero, además, la diferencia entre valores constituye una magnitud constante. Un ejemplo de nivel de intervalo de medición es la temperatura. Suponga que las temperaturas altas durante tres días consecutivos de invierno en Boston son de 28, 31 y 20 grados Fahrenheit. Estas temperaturas se clasifican fácilmente, aunque, además, es posible determinar la diferencia entre ellas, gracias a que un grado Fahrenheit representa una unidad de medición constante. Diferencias iguales entre dos temperaturas son las mismas, sin importar su posición en la escala. Es decir, la diferencia entre 10 y 15 grados Fahrenheit es de 5; la diferencia entre 50 y 55 grados también es de 5. Es importante destacar que 0 es un punto más en la escala. No representa la ausencia de estado. Cero grados Fahrenheit no representa la ausencia de calor, sino sencillamente el hecho de que hace frío. De hecho, 0 grados Fahrenheit equivale aproximadamente a -18 grados en la escala Celsius.

Otro ejemplo de escala de intervalo de medición consiste en las tallas de ropa para dama. Enseguida se muestran datos referentes a diversas medidas de una prenda de una mujer caucásica típica.

Talla	Busto (pulgadas)	Cintura (pulgadas)	Cadera (pulgadas)
8	32	24	35
10	34	26	37
12	36	28	39
14	38	30	41
16	40	32	43
18	42	34	45
20	44	36	47
22	46	38	49
24	48	40	51
26	50	42	53
28	52	44	55

¿Por qué razón la *talla* es una medición de intervalo? Observe que conforme la talla cambia 2 unidades (de la talla 10 a la 12, o de la talla 24 a la 26), cada medida aumenta 2 pulgadas. En otras palabras, los intervalos son los mismos.

No existe un punto cero natural que represente una talla. Una prenda *talla cero* no está hecha de *cero* material. Más bien, se trata de una prenda con 24 pulgadas de busto, 16 pulgadas de cintura y 27 de cadera. Además, las razones no tienen significado alguno. Si divide una talla 28 entre una talla 14, no obtiene la misma respuesta que si divide una talla 20 entre una 10. Ninguna razón es igual a dos, como sugeriría el número de *talla*. En resumen, si las distancias entre los números tienen sentido, aunque las razones no, entonces tiene una escala de intervalo de medición.

Las propiedades de los datos de nivel de intervalo son las siguientes:

1. Las clasificaciones de datos se ordenan de acuerdo con el grado que posea de la característica en cuestión.
2. Diferencias iguales en la característica representan diferencias iguales en las mediciones.

Datos de nivel de razón

Todos los datos cuantitativos son registrados en el nivel de razón de la medición. El **nivel de razón** es el *más alto*. Posee todas las características del nivel de intervalo, aunque, además, el punto 0 tiene sentido y la razón entre dos números es significativa. Ejemplos de la escala de razón de medición incluyen salarios, unidades de producción, peso, cambios en los precios de las acciones, la distancia entre sucursales y la altura. El dinero ilustra bien el caso. Si tiene cero dólares, entonces no tiene dinero. El peso constituye otro ejemplo. Si el cuadrante de la escala de un dispositivo correctamente calibrado se ubica en 0, entonces hay una ausencia total de peso. La razón entre dos

números también resulta significativa. Si Jim gana \$40 000 anuales vendiendo seguros y Rob gana \$80 000 al año vendiendo automóviles, entonces Rob gana el doble de lo que gana Jim.

La tabla 1.3 ilustra el uso de la escala de razón de medición, muestra los ingresos de cuatro parejas de padre e hijo.

TABLA 1.3 Combinaciones de ingresos de padre e hijo

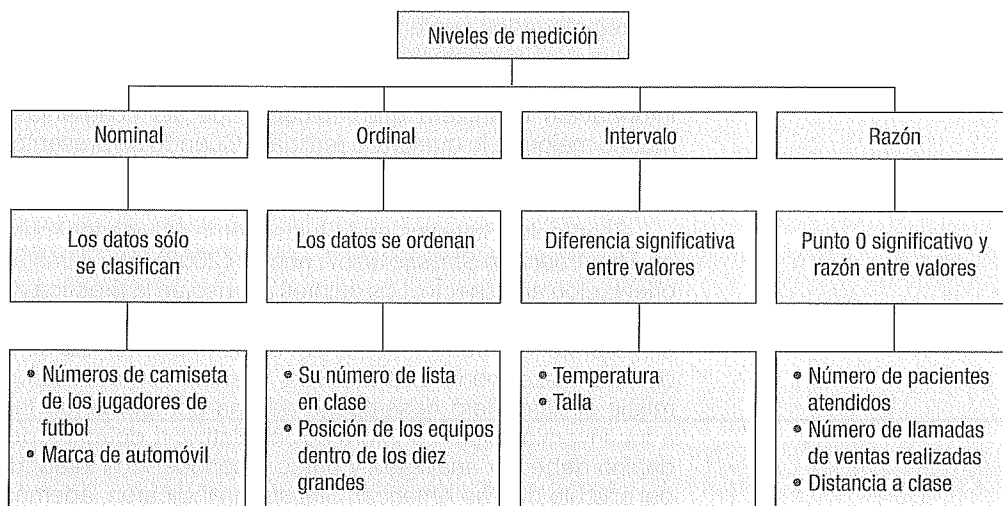
Nombre	Padre	Hijo
Lahey	\$80 000	\$ 40 000
Nale	90 000	30 000
Rho	60 000	120 000
Steele	75 000	130 000

Observe que Lahey, padre, gana el doble de lo que gana su hijo. En la familia de Rho, el hijo percibe el doble de ingresos que su padre.

En resumen, las propiedades de los datos de nivel de intervalo son las siguientes:

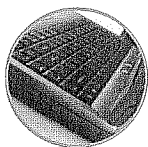
1. Las clasificaciones de datos se ordenan de acuerdo con la cantidad de características que poseen.
2. Diferencias iguales en la característica representan diferencias iguales en los números asignados a las clasificaciones.
3. El punto cero representa la ausencia de características y la razón entre dos números es significativa.

La gráfica 1.3 resume las principales características de los diversos niveles de medición.



GRÁFICA 1.3 Resumen de las características de los niveles de medición

Autoevaluación 1.2



¿Cuál es el nivel de medición que reflejan los siguientes datos?

- a) La edad de cada persona en una muestra de 50 adultos que escuchan una de las 1 230 estaciones de radio que transmiten entrevistas en Estados Unidos es:

35	29	41	34	44	46	42	42	37	47
30	36	41	39	44	39	43	43	44	40
47	37	41	27	33	33	39	38	43	22
44	39	35	35	41	42	37	42	38	43
35	37	38	43	40	48	42	31	51	34

- b) En una encuesta de 200 propietarios de automóviles de lujo, 100 eran de California, 50 de Nueva York, 30 de Illinois y 20 de Ohio.

Ejercicios

Al final del libro se encuentran las respuestas a los ejercicios impares.

- ¿Cuál es el nivel de medición de cada una de las siguientes variables?
 - Coeficientes intelectuales de los estudiantes.
 - La distancia que viajan los estudiantes para llegar a clases.
 - Las calificaciones de los estudiantes en el primer examen de estadística.
 - Una clasificación de estudiantes por fecha de nacimiento.
 - Una clasificación de estudiantes que cursan primero, segundo, tercero o último grado.
 - Número de horas que los alumnos estudian a la semana.
- ¿Cuál es el nivel de medición de los siguientes artículos relacionados con el negocio de los periódicos?
 - El número de periódicos vendidos todos los domingos durante 2006.
 - Los diferentes departamentos, como edición, publicidad, deportes, etcétera.
 - Un resumen del número de periódicos vendidos por condado.
 - Cantidad de años que cada empleado ha laborado en el periódico.
- Localice en la última edición de *USA Today* o en el periódico de la localidad ejemplos de cada nivel de medición. Redacte un breve resumen de lo que descubra.
- En los siguientes casos determine si el grupo representa una muestra o una población.
 - Los participantes en el estudio de un nuevo fármaco contra el colesterol.
 - Los conductores que recibieron una multa por exceso de velocidad en la ciudad de Kansas el último mes.
 - Beneficiarios del programa de asistencia social en Cook County (Chicago), Illinois.
 - Las 30 acciones que forman parte del promedio industrial Dow Jones.

Ética y estadística

Al seguir de cerca los sucesos de Enron, Tyco, HealthSouth, WorldCom y otros desastres relacionados con empresas, los estudiantes de administración necesitan comprender que estos acontecimientos se debieron a la interpretación equivocada de los datos administrativos y financieros. En cada caso, el personal comunicó a los inversionistas información financiera que indicaba que las compañías se estaban desempeñando mucho mejor de lo que era la realidad. Cuando se presentó la información verdadera, las compañías tenían un valor muy inferior al que se anunciaba. El resultado fue que muchos inversionistas perdieron todo o casi todo el dinero que invirtieron en estas compañías.

El artículo "Statistics and Ethics: Some Advice for Young Statisticians", que apareció en *The American Statistician* 57, núm. 1 (2003) (www.amstat.org/profession), proporciona orientación al respecto. Los autores aconsejan la práctica de la estadística con integridad y honestidad, e instan a "hacer lo correcto" cuando se recoja, organice, resuma, analice e interprete información numérica. La contribución real de la estadística a la sociedad es de naturaleza moral. Los analistas financieros necesitan proporcionar información que refleje el verdadero desempeño de una compañía, de tal manera que no desorienten a los inversionistas. La información relativa a defectos de un producto que puede ser dañino debe ser analizada y darse a conocer con integridad y honestidad. Los autores del artículo de *The American Statistician* indicaron, además, que cuando se practique la estadística, es necesario mantener "un punto de vista independiente y con principios".

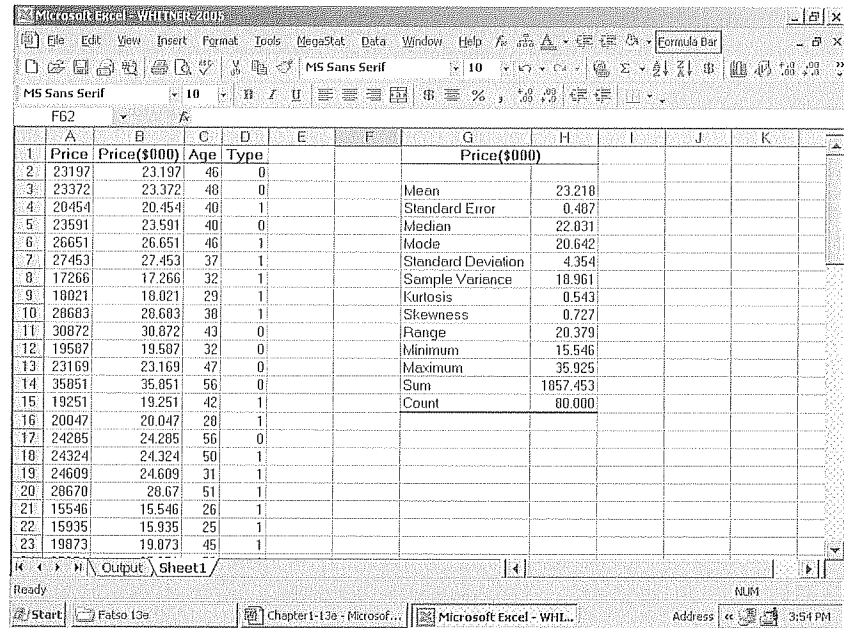
Conforme el lector avance, atenderá a cuestiones éticas relacionadas con la recopilación, análisis, presentación e interpretación de información estadística. Es de esperarse, asimismo, que conforme el lector aprenda más estadística, se convierta en un consumidor crítico. Por ejemplo, pondrá en tela de juicio un informe basado en datos que no representan fielmente a la población, otro que no contenga estadísticas relevantes, uno que incluya una elección incorrecta de medidas estadísticas o una presentación de datos tendenciosa en un intento deliberado por desorientar o tergiversar los hechos.

Aplicaciones de la computadora

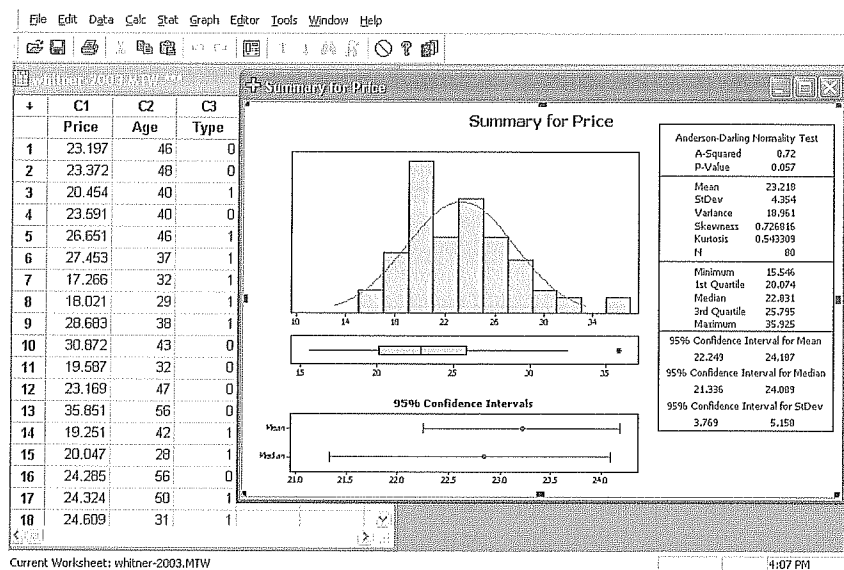
En la actualidad las computadoras están disponibles en la mayoría de las escuelas de formación profesional y universidades. Las hojas de cálculo, como Microsoft Excel, y los paquetes de software de estadística, como MINITAB, se encuentran disponibles en la mayoría de los laboratorios de computadoras. El paquete Microsoft Excel viene incluido con muchas computadoras domésticas. En el texto se emplea tanto Excel como MINITAB

para las aplicaciones. También se utiliza un complemento de Excel llamado MegaStat, que proporciona a Excel la capacidad para generar informes estadísticos adicionales.

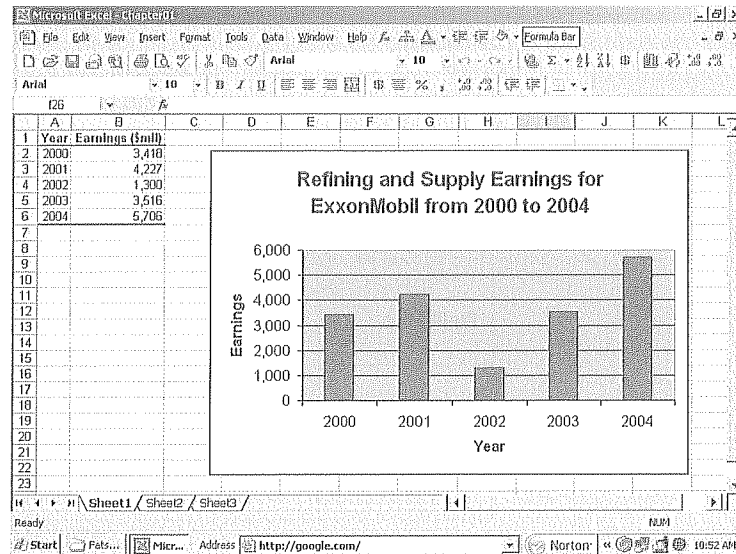
El siguiente ejemplo muestra la aplicación de las computadoras en el análisis estadístico. En los capítulos 2, 3 y 4 aparecen los métodos para resumir y describir datos. Un ejemplo utilizado en dichos capítulos se refiere al precio, expresado en miles de dólares, de 80 vehículos vendidos el mes pasado en Whitner Autoplex. La siguiente presentación de Excel revela, entre otras cosas: 1) Ochenta vehículos se vendieron el mes pasado. 2) El precio medio (promedio) de venta fue de \$23 218. 3) Los precios de venta iban de un mínimo de \$15 546 a un máximo de \$35 925.



La siguiente página se toma del sistema MINITAB, contiene mucha de la misma información.



Si hubiera empleado una calculadora para llegar a estas medidas y otras que se necesitan para analizar plenamente los precios de venta, hubiera requerido horas de cálculos. Además, la posibilidad de cometer un error aritmético es alta cuando se maneja una gran cantidad de valores. Por otra parte, los paquetes de software de estadística y las hojas de cálculo proporcionan información exacta en segundos.



Redacte un breve informe con un análisis de las utilidades de ExxonMobil durante dicho periodo. ¿Se incrementaron las utilidades o disminuyeron?

ejercicios.com



En los siguientes ejercicios se hace uso de la World Wide Web, una fuente de información rica y en crecimiento. Debido a la naturaleza cambiante y de la continua revisión de los sitios web, es posible que se encuentren diferentes menús y que las direcciones exactas, o URL, cambien. Cuando visite una página, hay que prepararse para buscar el vínculo.

15. Suponga que recién abrió una cuenta en Ameritrade, Inc., un corredor de bolsa en línea. Usted decide comprar acciones, ya sea de Johnson & Johnson (una compañía farmacéutica) o de PepsiCo (empresa matriz de Pepsi y Frito-Lay). Si desea hacer una comparación entre las dos compañías, visite la página <http://finance.yahoo.com> y, en el espacio que dice **Get Quotes**, escriba las letras JNJ y PEP, que son los respectivos símbolos de las compañías. Haga clic en **Go** para obtener información reciente sobre el precio de venta de las dos acciones. A la derecha de esta información, dé clic en **More** y enseguida en **Analyst Opinion**. Aquí hay información de unos analistas accionarios que evaluaron las acciones. Los corredores de bolsa califican la acción con 1, si se trata de una buena compra, y con 5 si se trata de una buena venta. ¿Qué nivel de medición corresponde a esta información? ¿Qué acciones se recomiendan?

Ejercicios de la base de datos

16. Regrese a los datos de Real Estate que aparecen en el texto, que incluyen información sobre casas vendidas en la zona de Denver, Colorado, el año pasado. Considere las siguientes variables: precio de venta, número de recámaras, ubicación y distancia al centro de la ciudad.
 - a) De las variables, ¿cuáles son cualitativas y cuáles cuantitativas?
 - b) Determine el nivel de medición de cada una de las variables.
17. Consulte los datos Baseball 2005, que contienen información de los treinta equipos de las Ligas Mayores de Béisbol para la temporada 2005. Considere las siguientes variables: número de victorias, salario del equipo, asistencia durante la temporada, si el equipo jugó los partidos como anfitrión sobre césped, pasto sintético o superficie artificial, así como el número de carreras anotadas.
 - a) ¿Cuáles de estas variables son cuantitativas y cuáles cualitativas?
 - b) Determine el nivel de medición de cada una de las variables.
18. Vaya a los datos Wage, que incluyen información de los salarios anuales de una muestra de 100 trabajadores. También incluye variables sobre la industria, años de educación y género de cada trabajador.
 - a) ¿Cuáles de las doce variables son cuantitativas y cuáles cualitativas?
 - b) Determine el nivel de medición de cada variable.

19. Consulte los datos CIA, que incluyen información demográfica y económica sobre 46 países.
- a) ¿Qué variables son cuantitativas y cuáles cualitativas?
 - b) Determine el nivel de medición de cada variable.

Capítulo 1 Respuestas a las autoevaluaciones



- 1.1 a) Sobre la base de la muestra de 1 960 consumidores, estimamos que, si lo comercializa, 60% de los consumidores comprará el platillo de pollo $(1\ 176/1\ 960) \times 100 = 60\%$.
- b) Estadística inferencial, ya que se empleó una muestra para llegar a una conclusión relativa a la reacción de los consumidores de la población en caso de que se comercializara el platillo de pollo.
- 1.2 a) La edad es una variable de escala de razón. Una persona de 40 años tiene el doble de edad que una de 20.
- b) Escala nominal. Podría ordenar indistintamente los estados.

2

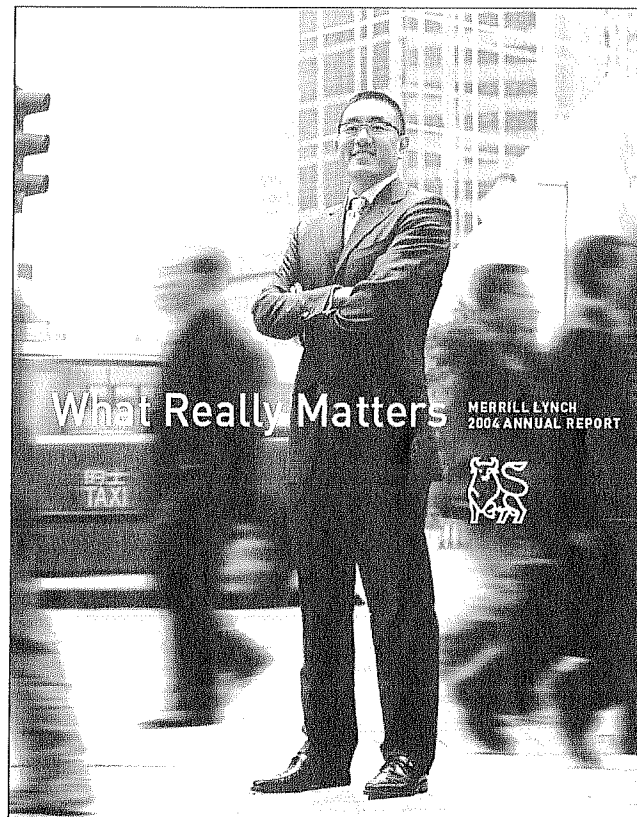
OBJETIVOS

Al concluir el capítulo, será capaz de:

1. Organizar los datos cualitativos en una *tabla de frecuencias*.
2. Representar una tabla de frecuencias como una *gráfica de barras* o una *gráfica de pastel*.
3. Organizar datos cuantitativos en una *distribución de frecuencias*.
4. Representar una distribución de frecuencias de datos cuantitativos por medio de *histogramas*, *polígonos de frecuencia* y *polígonos de frecuencias acumuladas*.

Descripción de datos:

Tablas de frecuencias, distribuciones de frecuencias y su representación gráfica



Merrill Lynch recién concluyó el estudio de una cartera de inversiones en línea para una muestra de clientes. Elabore un histograma con los datos de los 70 participantes en el estudio (véase ejercicio 39 y objetivo 4).