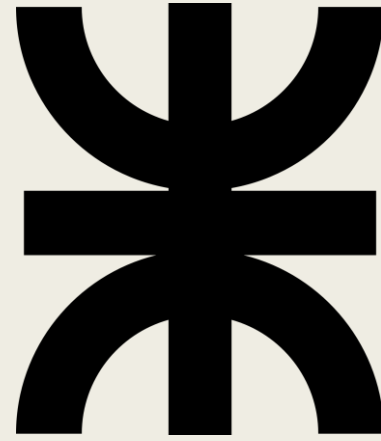


Estadística – Tecnicatura Universitaria en Programación



ESTADÍSTICA

Introducción

Ing. Sofía Pezzi



¿QUÉ ES LA ESTADÍSTICA



¿Qué es la estadística?

Ciencia y una rama de las matemáticas a través de la cual se recolecta, analiza, describe y estudia una serie de datos a fin de establecer comparaciones o variabilidades que permitan comprender un fenómeno en particular.

Aplicaciones de la estadística

- Método efectivo para describir con exactitud los valores de los datos económicos, políticos, sociales, psicológicos, biológicos y físicos, y sirve como herramienta para relacionar y analizar dichos datos.
- El trabajo del experto no consiste ya sólo en reunir y tabular los datos, sino sobre todo el proceso de interpretación de esa información.
- El desarrollo de la teoría de la probabilidad ha aumentado el alcance de las aplicaciones de la estadística. Muchos conjuntos de datos se pueden aproximar, con gran exactitud, utilizando determinadas distribuciones probabilísticas; los resultados de éstas se pueden utilizar para analizar datos estadísticos.



ESTADÍSTICA DESCRIPTIVA



¿Qué es la estadística descriptiva?

- La rama de la estadística que presenta técnicas para describir conjuntos de mediciones.
- La **estadística descriptiva** está formada por procedimientos empleados para resumir y describir las características importantes de un conjunto de mediciones.
- La **estadística inferencial** está formada por procedimientos empleados para sacar conclusiones acerca de características poblacionales, a partir de información contenida en una muestra sacada de esta población.

PROCEDIMIENTO:

1. Especifique las preguntas a contestar e identifique la población de interés.

El objetivo es determinar quién obtendrá más votos el día de la elección.

La población de interés es el conjunto de todos los votos en la elección presidencial.

2. Decida cómo seleccionar la muestra: procedimiento de muestreo.

¿La muestra es representativa de la población de interés?

Se podría buscar una muestra de votantes registrados en una única provincia, ¿esta muestra será representativa de todos los votantes?

¿La muestra es lo suficientemente grande para contestar las preguntas planteadas en el paso 1 sin perder tiempo y dinero en información adicional?

Un buen diseño de muestreo contestará las preguntas planteadas, con mínimo costo para el experimentador.

PROCEDIMIENTO:

3. Seleccione la muestra y analice la información muestral.

Método de análisis apropiado para extraerla.

4. Use la información del paso 3 para hacer una inferencia acerca de la población.

Es posible usar muchos procedimientos diferentes para hacer esta inferencia y algunos son mejores que otros.

Por ejemplo, podría haber 10 métodos diferentes para estimar la respuesta humana a un medicamento experimental, pero un procedimiento podría ser más preciso que los otros. Se debe usar el mejor procedimiento disponible para hacer inferencias.

5. Determine la confiabilidad de la inferencia.

Toda inferencia estadística debe incluir una medida de confiabilidad que dice cuánta confianza tiene usted en la inferencia.

VARIABLES Y DATOS

Una **variable** es una característica que cambia o varía con el tiempo y/o para diferentes personas u objetos bajo consideración.

EJEMPLO:

La temperatura corporal que varia



En una misma persona en distintos momentos del día

De una persona a otra.

La creencia religiosa.

La afiliación política.

Estatura, edad.

Cantidad de hijos.

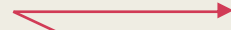

POBLACIÓN Y MUESTRA

EJEMPLO:

Para estudiar cuál es el candidato presidencial por el cual votarán los peruanos en las próximas elecciones, se toma una muestra de 3500 personas de todo el país. La pregunta es la siguiente, ¿por quién votará en las próximas elecciones presidenciales?

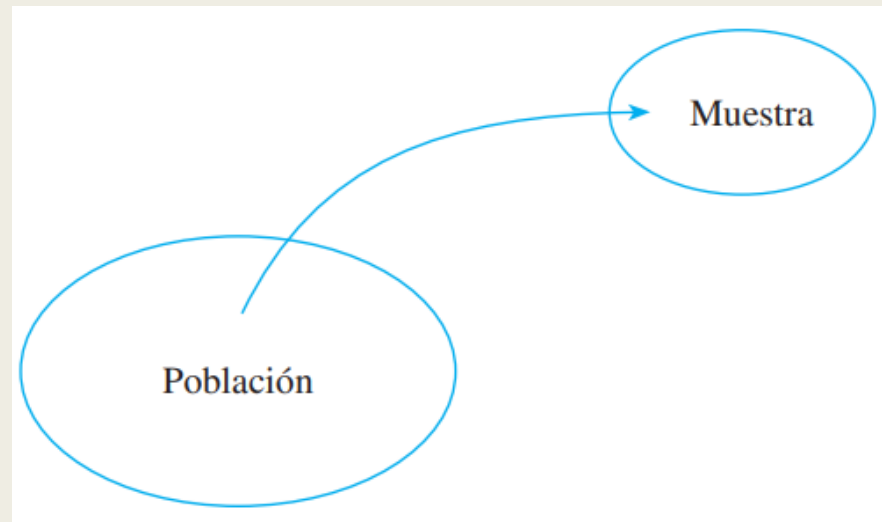
¿Qué es lo que queremos saber?

La población  Pero ésta puede ser difícil o imposible de enumerar.

Se elige una muestra  Tiene que ser bien representativa.
 Existen cálculos para definir el tamaño.

POBLACIÓN Y MUESTRA

Una **población** es el conjunto de mediciones de interés para el investigador.



Una **muestra** es un subconjunto de mediciones seleccionado de la población de interés.

POBLACIÓN Y MUESTRA

Una **unidad experimental** es el individuo u objeto en el que se mide una variable.

Resulta una sola medición o datos cuando una variable se mide en realidad en una unidad experimental.

EJEMPLO:

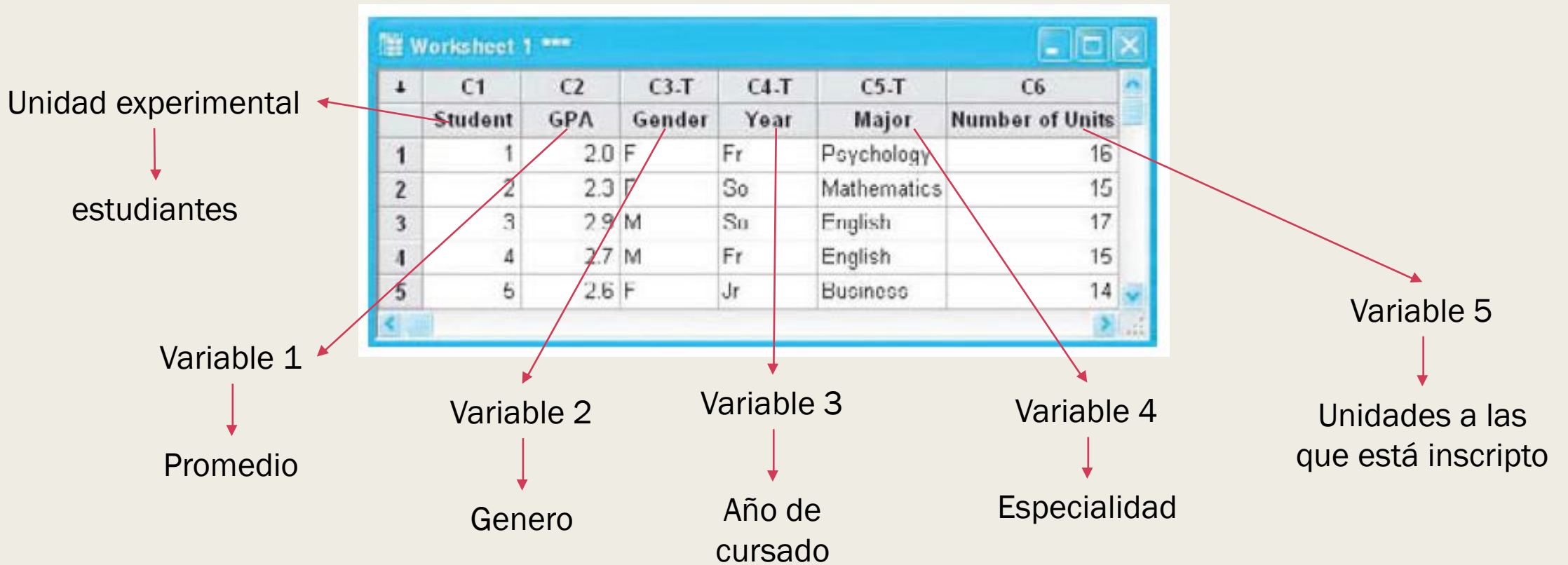
Para estudiar cuál es el candidato presidencial por el cual votarán los peruanos en las próximas elecciones, se toma una muestra de 3500 personas de todo el país. La pregunta es la siguiente, ¿por quién votará en las próximas elecciones presidenciales?

Se trata de describir o pronosticar el comportamiento de la población con base en información obtenida de una muestra representativa de esa población.

DATOS UNIVARIADOS Y BIVARIADOS

EJEMPLO:

De entre todos los alumnos de una gran universidad se selecciona un conjunto de cinco estudiantes y las mediciones se introducen en una hoja de cálculo:



DATOS UNIVARIADOS Y BIVARIADOS

Resultan **datos univariados** cuando se mide una sola variable en una sola unidad experimental.

EJEMPLO: Cuál candidato votarán los peruanos en las próximas elecciones.

Resultan **datos bivariados** cuando se miden dos variables en una sola unidad experimental.

EJEMPLO: Peso y altura de niños.

Resultan **datos multivariados** cuando se miden más de dos variables.

EJEMPLO: Ejemplo de los estudiantes.

EJERCITACIÓN

1. Identifique las unidades experimentales en los que se miden las variables siguientes:
 - a. Género de un estudiante.
 - b. Número de errores en un examen de medio semestre.
 - c. Edad de un paciente con cáncer.
 - d. Número de flores en una planta de azalea.
 - e. Color de un auto que entra a un estacionamiento.

TIPOS DE VARIABLES

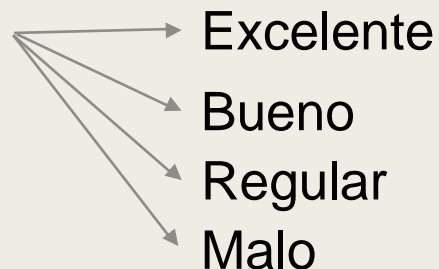
Las **variables cualitativas** miden una cualidad o característica en cada unidad experimental.

EJEMPLOS:

La afiliación
política



Clasificación de
gustos



TIPOS DE VARIABLES

Las **variables cuantitativas** miden una cantidad numérica en cada unidad experimental.

EJEMPLOS:

x = tasa preferencial de
interés

x = número de pasajeros en un vuelo de Argentina a Italia

x = peso de paquetes de encomienda

x = volumen de jugo de naranja en un vaso

TIPOS DE VARIABLES CUANTITATIVAS

Una **variable discreta** puede tomar sólo un número finito o contable de valores.

EJEMPLOS:

x = tasa preferencial de
interés

x = número de pasajeros en un vuelo de Argentina a Italia

x = peso de paquetes de encomienda

x = volumen de jugo de naranja en un vaso

Miembros en una
familia

Autos nuevos
vendidos

Productos
defectuosos

TIPOS DE VARIABLES CUANTITATIVAS

Una **variable continua** puede tomar infinitamente muchos valores correspondientes a los puntos en un intervalo de recta.

EJEMPLOS:

x = tasa preferencial de
interés

x = número de pasajeros en un vuelo de Argentina a Italia

x = peso de paquetes de encomienda

x = volumen de jugo de naranja en un vaso

EJERCITACIÓN

5 - Se seleccionan seis vehículos, de entre los que tienen permiso para estacionarse, y se registran los datos siguientes:

- ¿Cuáles son las unidades experimentales?
- ¿Cuáles son las variables que se miden? ¿Qué tipos de variables son?
- ¿Estos datos son univariados, bivariados o multivariados?

Vehículo	Tipo	Marca	¿Colectivo?	Distancia de viaje en una dirección (millas)	Antigüedad del vehículo (años)
1	Auto	Honda	No	23.6	6
2	Auto	Toyota	No	17.2	3
3	Camión	Toyota	No	10.1	4
4	Van	Dodge	Sí	31.7	2
5	Moto-cicleta	Harley-Davidson	No	25.5	1
6	Auto	Chevrolet	No	5.4	9

EJERCITACIÓN

6. Un conjunto de datos contiene las edades al fallecimiento de cada uno de los anteriores 38 presidentes de Estados Unidos ahora desaparecidos.
- a. ¿Este conjunto de mediciones es una población o una muestra?
 - b. ¿Cuál es la variable que se mide?
 - c. ¿La variable del inciso b) es cuantitativa o cualitativa?

EJERCITACIÓN

- Identifique cada una de las siguientes variables como cualitativas o cuantitativas:
 1. El uso más frecuente de su horno de microondas (recalentar, descongelar, calentar, otros)
 2. El número de consumidores que se niegan a contestar una encuesta por teléfono
 3. La puerta escogida por un ratón en un experimento de laberinto (A, B o C)
 4. El tiempo ganador para un caballo que corre en el Derby de Kentucky
 5. El número de niños en un grupo de quinto grado que leen al nivel de ese grado o mejor

EJERCITACIÓN

3. Identifique las siguientes variables cuantitativas como discretas o continuas:
 - a. Población en una región particular de un país.
 - b. Peso de periódicos recuperados para reciclar en un solo día.
 - c. Tiempo para completar un examen de sociología.
 - d. Número de consumidores en una encuesta de 1000 que consideran importante aplicar leyenda nutrimental en productos alimenticios.

RECOPIACIÓN DE DATOS

Una vez recolectados los datos, éstos pueden consolidarse y resumirse para mostrar la siguiente información:

- ¿Qué valores de la variable han sido medidos?
- ¿Con qué frecuencia se presenta cada uno de los valores?

Para ello se construye una **TABLA ESTADÍSTICA**

EJEMPLOS:

El número de estrellas de los hoteles de una ciudad viene dado por la siguiente serie:

2, 2, 3, 3, 3, 2, 2, 2, 2, 2, 3, 2, 1, 1, 1, 2, 2, 4, 1.

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

RECOPIACIÓN DE DATOS

EJEMPLOS:

El número de estrellas de los hoteles de una ciudad viene dado por la siguiente serie:

2, 2, 3, 3, 3, 2, 2, 2, 2, 2, 3, 2, 1, 1, 1, 2, 2, 4, 1.
3, 3, 4, 3, 4, 3, 1, 3, 4, 3, 3, 3, 2, 1, 3, 3, 3, 2, 3,

RECOPIACIÓN DE DATOS

La frecuencia es el número de mediciones en cada categoría

La frecuencia acumulada es el número de mediciones acumuladas

La frecuencia relativa es la proporción de mediciones en cada categoría

La frecuencia porcentual es la proporción de mediciones en cada categoría expresadas en %

RECOPIACIÓN DE DATOS

EJEMPLOS:

El número de estrellas de los hoteles de una ciudad viene dado por la siguiente serie:

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

N° de estrellas del hotel	Frecuencia absoluta	Fr acumulada	Fr relativa	Fr %
1	6	6	$6 / 38 = 0,1579$	15,79 %
2	12	$12 + 6 = 18$	$12 / 38 = 0,3158$	31,58 %
3	16	$16 + 18 = 34$	$16 / 38 = 0,4211$	42,11 %
4	4	$34 + 4 = 38$	$4 / 38 = 0,1052$	10,52 %
TOTAL	38	38	1	100 %

RECOPIACIÓN DE DATOS DISCRETOS

EJEMPLOS:

Las calificaciones de 50 alumnos en Matemáticas han sido las siguientes:

5, 2, 4, 9, 7, 4, 5, 6, 5, 7, 7, 5, 5, 2, 10, 5, 6, 5, 4, 5, 8, 8, 4, 0,
8, 4, 8, 6, 6, 3, 6, 7, 6, 6, 7, 6, 7, 3, 5, 6, 9, 6, 1, 4, 6, 3, 5, 5, 6, 7.

RECOPIACIÓN DE DATOS CONTINUOS

EJEMPLOS:

Los 40 alumnos de una clase han obtenido las siguientes puntuaciones, sobre 50, en un examen de Física.

3, 15, 24, 28, 33, 35, 38, 42, 23, 38, 36, 34, 29, 25, 17, 7, 34, 36, 39, 44, 31,
26, 20, 11, 13, 22, 27, 47, 39, 37, 34, 32, 35, 28, 38, 41, 48, 15, 32, 13.

X				
0 – 10				
11 – 20				
21 – 30				
31 – 40				
41 – 50				

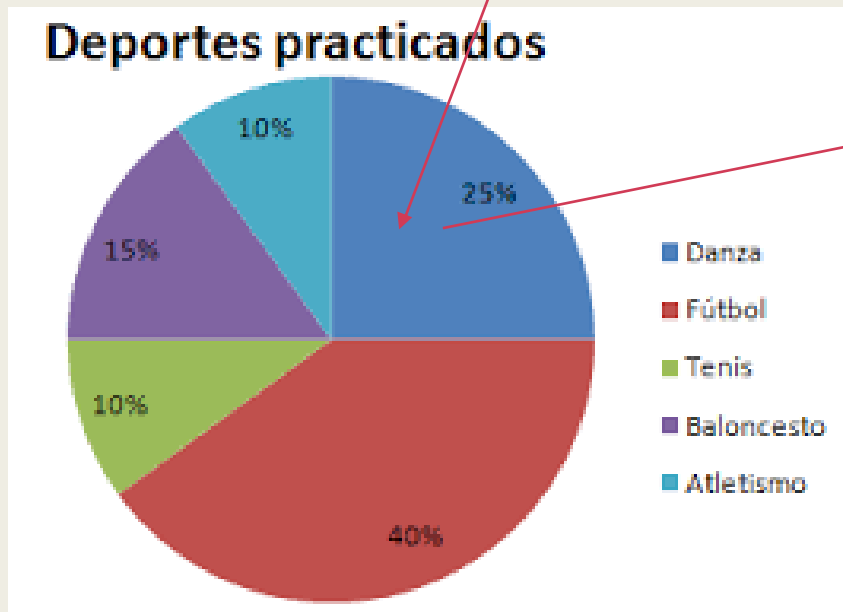


GRÁFICOS

Datos cualitativos – GRÁFICA CIRCULAR

Gráfico de PASTEL

Se divide en **SECTORES**.



Tamaño: Frecuencia relativa porcentual.

- Circulo entero → 360°
- Cada variable → Se calculan los $^{\circ}$

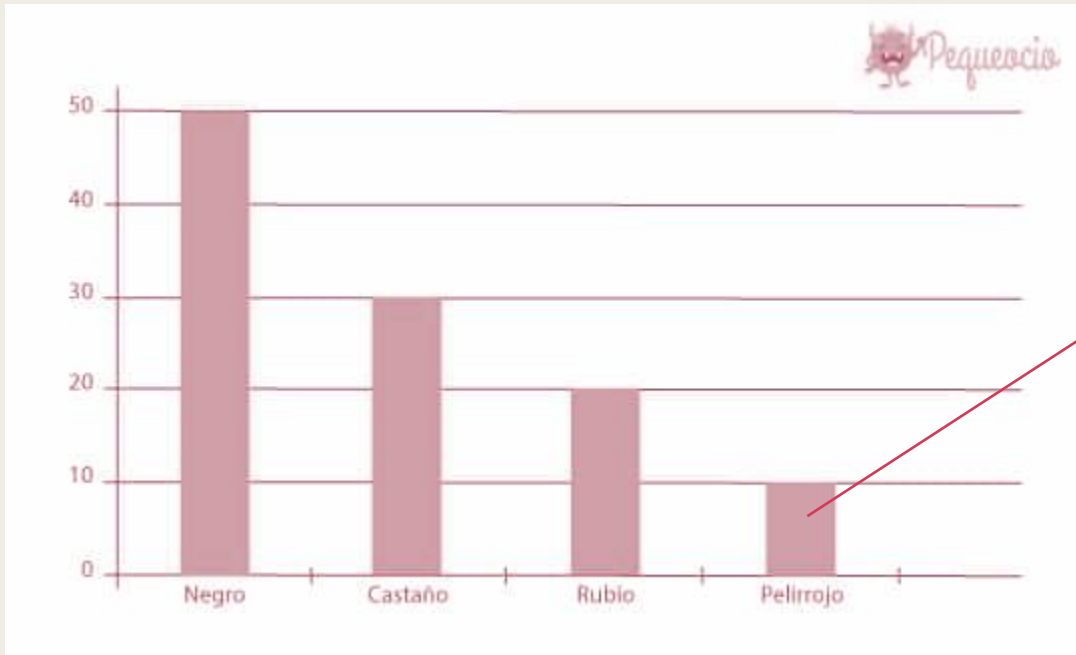
Datos cualitativos – DIAGRAMA CIRCULAR

X	Fr abs.	Fr acum	Fr rel	Fr %
0 – 10	2	2	0,050	5,0 %
11 – 20	6	8	0,150	15,0 %
21 – 30	10	18	0,250	25,0 %
31 – 40	17	35	0,425	42,5 %
41 – 50	5	40	0,125	12,5 %

Ángulo: Frecuencia relativa x 360

Datos cualitativos – GRÁFICA DE BARRAS

Las barras tienen la altura de las frecuencias.



Tamaño: Frecuencia Absoluta.

Destaca la frecuencia real de cada categoría

Gráfico de pastel → Muestra la relaciones de cada parte en el todo.

EJERCITACIÓN

8. Cincuenta personas se agrupan en cuatro categorías, A, B, C y D, y el número de personas que caen en cada categoría se muestra en la tabla:

- ¿Cuál es la unidad experimental?
- ¿Cuál es la variable que se mide? ¿Es cualitativa o cuantitativa?
- Construya una gráfica de pastel para describir los datos.
- Construya una gráfica de barras para describir los datos.
- ¿La forma de la gráfica de barras del inciso d) cambia, dependiendo del orden de presentación de las cuatro categorías? ¿Es importante el orden de presentación?
- ¿Qué proporción de las personas está en la categoría B, C o D?
- ¿Qué porcentaje de las personas no está en la categoría B?

Categoría	Frecuencia
A	11
B	14
C	20
D	5

EJERCITACIÓN

9. Un fabricante de jeans (pantalones vaqueros) tiene plantas en California, Arizona y Texas. Un grupo de 25 pares de jeans se selecciona al azar de entre la base de datos computarizada, registrándose el estado en el que se produce:

- ¿Cuál es la unidad experimental?
- ¿Cuál es la variable que se mide? ¿Es cualitativa o cuantitativa?
- Construya una gráfica de pastel para describir los datos.
- Construya una gráfica de barras para describir los datos.
- ¿Qué proporción de los jeans se hace en Texas?
- ¿Cuál estado produjo más jeans del grupo?
- Si se desea averiguar si las tres plantas produjeron iguales números de jeans, o si una produjo más jeans que las otras, ¿cómo se pueden usar las gráficas de las partes c y d para ayudar? ¿Qué conclusiones puede el lector sacar de estos datos?

CA	AZ	AZ	TX	CA
CA	CA	TX	TX	TX
AZ	AZ	CA	AZ	TX
CA	AZ	TX	TX	TX
CA	AZ	AZ	CA	CA

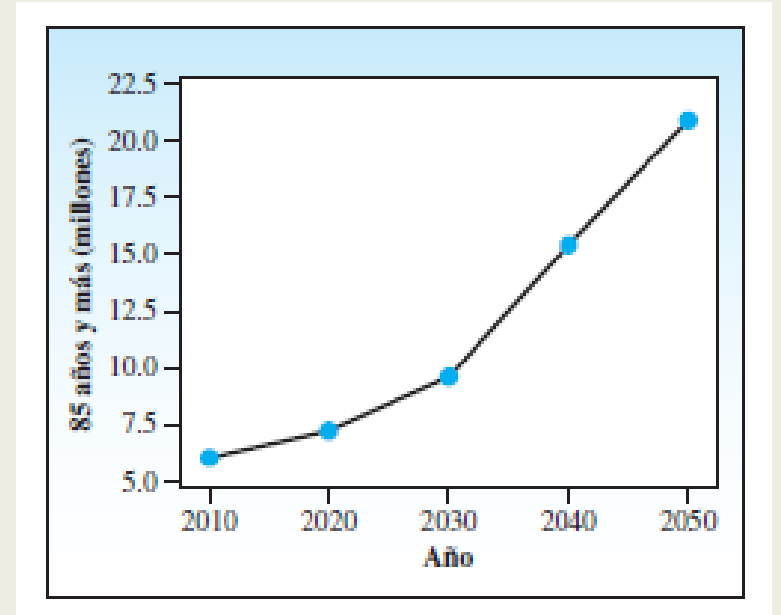
Datos cualitativos

GRAFICO DE BARAS

- <https://www.youtube.com/watch?v=wtZJwBzGVWc>
- <https://www.youtube.com/watch?v=Jb3cDEjjjIA>
- <https://www.youtube.com/watch?v=-LfsNnB4558>

Datos cuantitativos – GRÁFICA DE LÍNEA

- Variable → Eje Y
- Registrada en el tiempo → Eje X
- Conjunto de datos → Serie de tiempo



La idea es tratar de distinguir un patrón o tendencia que sea probable de continuar en el futuro y luego usar ese patrón para hacer predicciones precisas para el futuro inmediato.

Datos cuantitativos – GRÁFICA DE LÍNEA

- <https://es.tradingview.com/symbols/BTCUSD/>
- <https://www.cronista.com/MercadosOnline/moneda.html?id=ARSB>
- <https://www.smn.gob.ar/estadisticas>

Datos cuantitativos – GRÁFICA DE LÍNEA

EJEMPLOS:

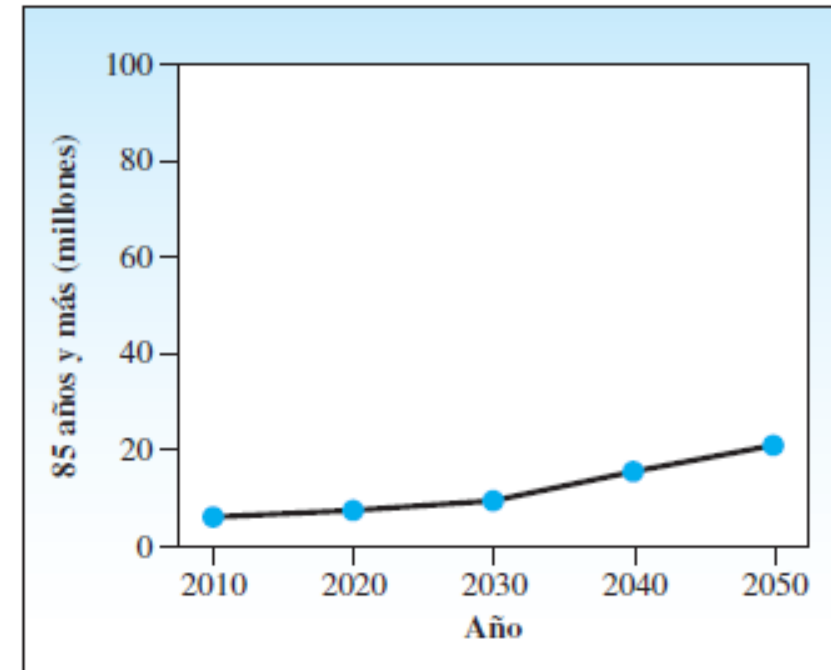
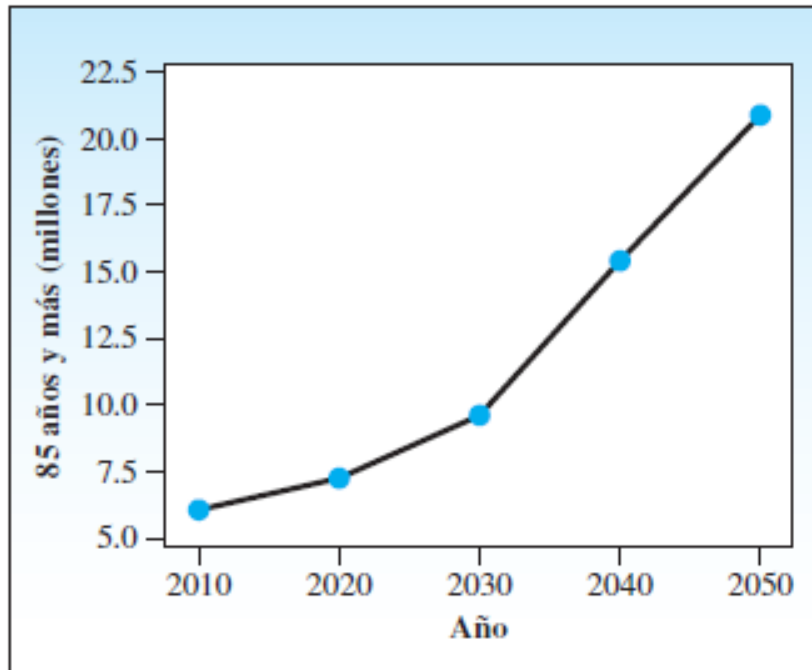
En el año 2025, el mayor de los “hijos de la explosión demográfica” (nacido en 1946) tendrá 79 años, y el mayor de los de la “Generación X” (nacido en 1965) estará a dos años de ser elegible para el Seguro Social. ¿Cómo afectará esto a las tendencias del consumidor en los siguientes 15 años? ¿Habrá suficientes fondos para los “hijos de la explosión demográfica” para recolectar prestaciones del Seguro Social? La Oficina de Censos de Estados Unidos da proyecciones para la parte de la población norteamericana que tendrá 85 años y más en los próximos años, como se muestra a continuación. Construya una gráfica de líneas para ilustrar los datos. ¿Cuál es el efecto de prolongar y contraer el eje vertical de la gráfica de línea?

Año	2010	2020	2030	2040	2050
85 o más (millones)	6.1	7.3	9.6	15.4	20.9

Datos cuantitativos – GRÁFICA DE LÍNEA

EJEMPLOS:

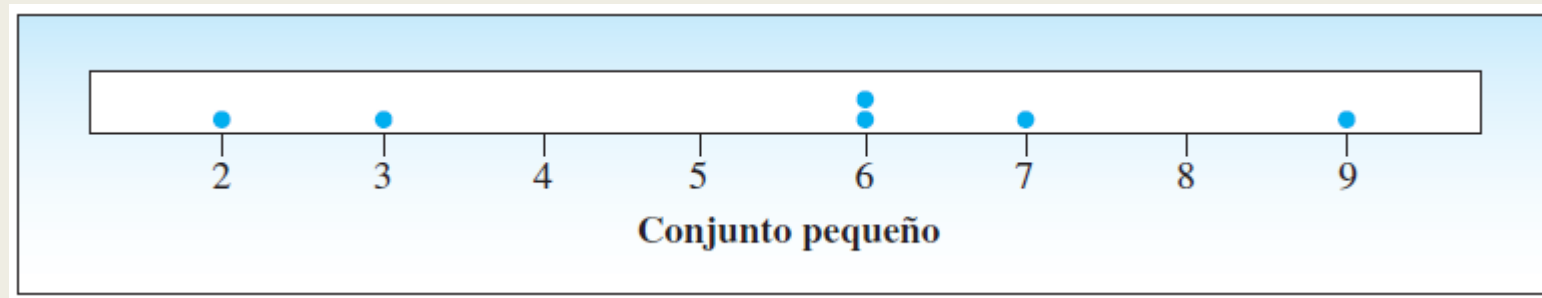
Año	2010	2020	2030	2040	2050
85 o más (millones)	6.1	7.3	9.6	15.4	20.9



Datos cuantitativos – GRÁFICA DE PUNTOS

- Gráfica más sencilla para datos cuantitativos.
- Frecuencia variable → Tantos puntos como datos.
- Variable → Eje X
- Conjunto de datos → Serie de tiempo

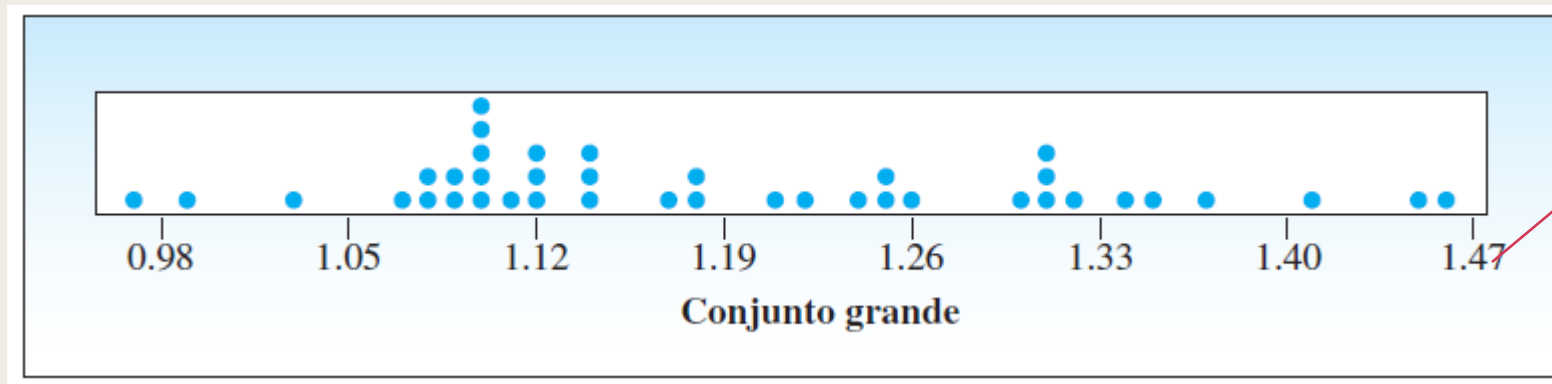
Variables discretas:



Datos cuantitativos – GRÁFICA DE PUNTOS

- Gráfica más sencilla para datos cuantitativos.
- Frecuencia variable —————> Tantos puntos como datos.
- Variable —————> Eje X
- Conjunto de datos —————> Serie de tiempo

Variables continuas:

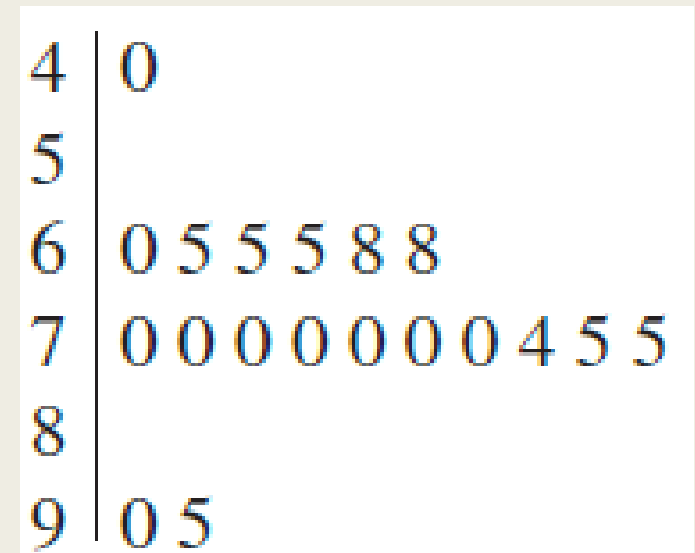


Se arma como cualquier gráfica a escala

Datos cuantitativos

GRÁFICA DE TALLO Y HOJAS

- Gráfica sencilla para exhibir la distribución de un conjunto de datos cuantitativos.
- Se divide cada observación entre las unidades y las decenas.
- Izquierda → Tallo
- Derecha → Hojas



Datos cuantitativos

GRÁFICA DE TALLO Y HOJAS

EJEMPLOS:

En la tabla se muestra una lista de precios (en dólares) de 19 marcas de zapatos deportivos.

Construya una gráfica de tallo y hoja para mostrar la distribución de los datos.

90	70	70	70	75	70
65	68	60	74	70	95
75	70	68	65	40	65
70					

Datos cuantitativos

GRÁFICA DE TALLO Y HOJAS

- Cuando hay pocos tallos y muchas hojas

- Se hacen las hojas en dos líneas.

Las hojas 0-4 en la primera línea y las hojas 5-9 en la segunda

- Se hacen las hojas en cinco líneas.

Las hojas 0-1, 2-3, 4-5, 6-7 y 8-9 en las cinco líneas,

- Si se gira el gráfico, se apilan los datos y nos queda una “gráfica de puntos”.
- Las hojas pueden ser decimales.

Datos cuantitativos

GRÁFICA DE TALLO Y HOJAS

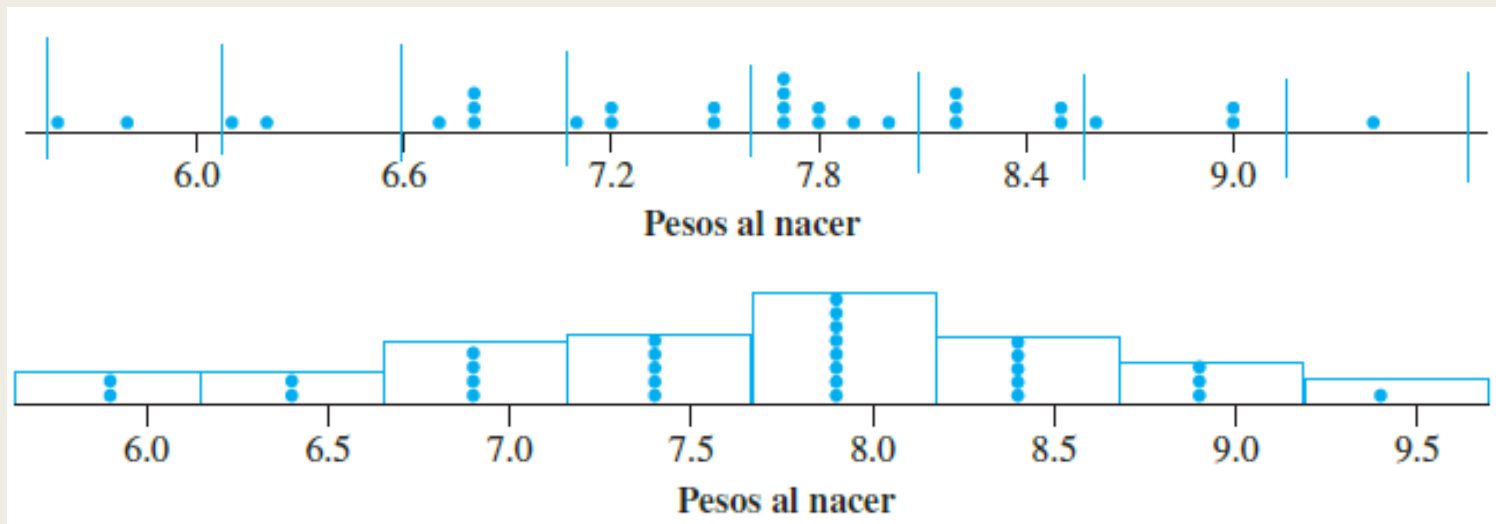
EJEMPLOS:

Los datos de la tabla son los pesos de 30 bebés de gestación completa al momento de nacer, nacidos en un hospital metropolitano y registrados al décimo de libra más cercano. Construya una gráfica de tallo y hoja para mostrar la distribución de los datos.

7.2	7.8	6.8	6.2	8.2
8.0	8.2	5.6	8.6	7.1
8.2	7.7	7.5	7.2	7.7
5.8	6.8	6.8	8.5	7.5
6.1	7.9	9.4	9.0	7.8
8.5	9.0	7.7	6.7	7.7

Datos cuantitativos – HISTOGRAMA

- Similar al gráfico de barra —————> Datos cuantitativos
- Se grafican los datos como una gráfica de puntos
 - Se dividen los intervalos en subintervalos
 - Se colocan los puntos correspondientes
 - Se coloca una línea sobre cada intervalo



Datos cuantitativos – HISTOGRAMA

Gráfica de barras en la que la altura de la barra muestra “con qué frecuencia” (medida como proporción o frecuencia relativa) las mediciones caen en una clase o subintervalo particular.

- Subintervalos → Eje X
- Frecuencia → Eje Y
- Se divide en intervalos de igual longitud.

Tamaño de muestra	25	50	100	200	500
Número de clases	6	7	8	9	10

EJEMPLOS:

$$9,4 - 5,6 = 3,8 \quad \xrightarrow{\text{Se dividen en 8 intervalos}}$$

$$3,8 / 8 = 0,475$$

↓
Peso más grande y más chico.

↘
Se redondea a 0,5

Datos cuantitativos

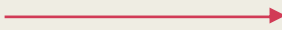
HISTOGRAMA

EJEMPLOS:

Los datos de la tabla son los pesos de 30 bebés de gestación completa al momento de nacer, nacidos en un hospital metropolitano y registrados al décimo de libra más cercano. Construya una gráfica de tallo y hoja para mostrar la distribución de los datos.

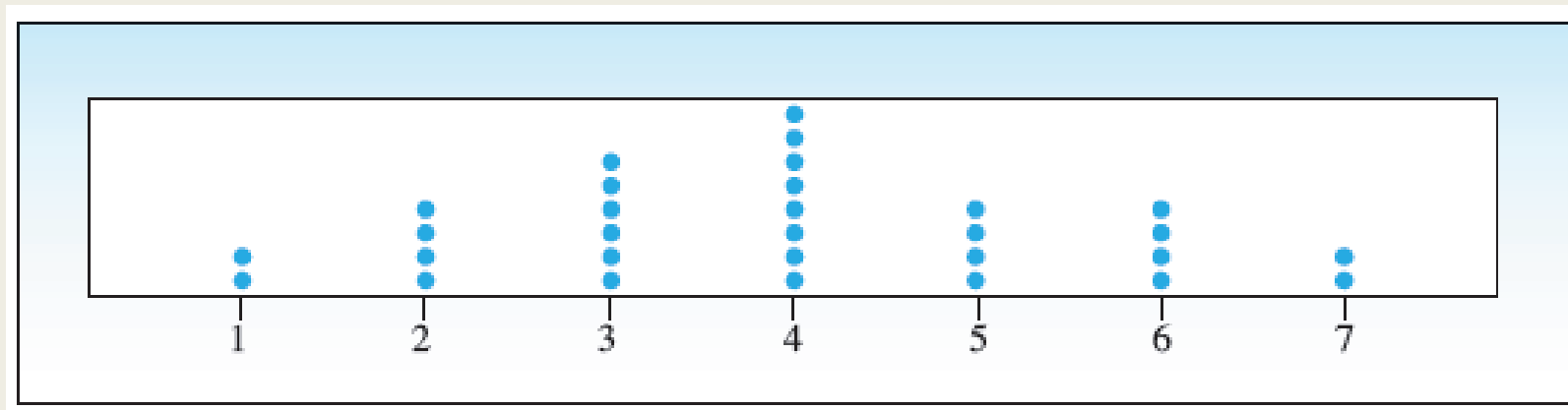
7.2	7.8	6.8	6.2	8.2
8.0	8.2	5.6	8.6	7.1
8.2	7.7	7.5	7.2	7.7
5.8	6.8	6.8	8.5	7.5
6.1	7.9	9.4	9.0	7.8
8.5	9.0	7.7	6.7	7.7

INTERPRETACIÓN DE GRÁFICOS

- Una vez creada la grafica  Se siguen los siguientes pasos:
 1. Verificar las escalas horizontales y verticales.
 2. Examinar el lugar de la distribución de datos
 - a) ¿Donde está el centro de la distribución?
 - b) Si se comparan dos distribuciones: ¿están centradas en el mismo lugar?
 3. Examinar la forma de la distribución
 - a) ¿La distribución tiene un pico?
 - b) ¿En qué medición se encuentra ese pico?
 4. Buscar mediciones poco comunes o resultados atípicos.
 - a) Pueden no ser representativos.

DISTRIBUCIONES

Una distribución es **simétrica** si los lados izquierdo y derecho de la distribución, cuando se divide en el valor medio, forman imágenes espejo.



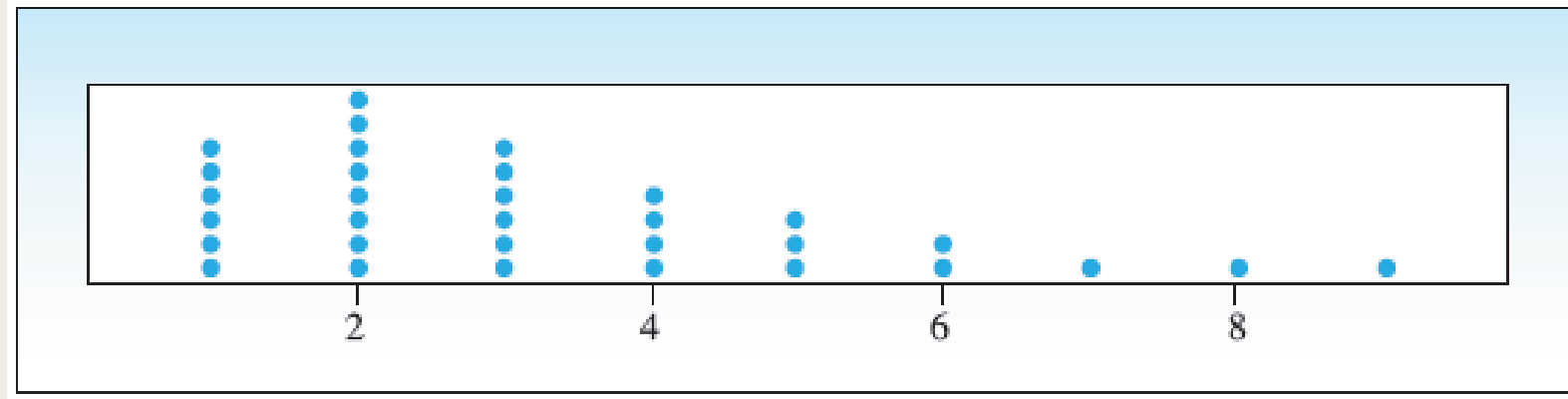
Distribución *relativamente simétrica* con un solo pico situado en $x = 4$.

Si se dobla la página en este pico, las mitades izquierda y derecha *casi* serían imágenes espejo.

Que está inclinado, torcido

DISTRIBUCIONES

Una distribución está **sesgada a la derecha** si una proporción más grande de las mediciones se encuentra a la derecha del valor pico.

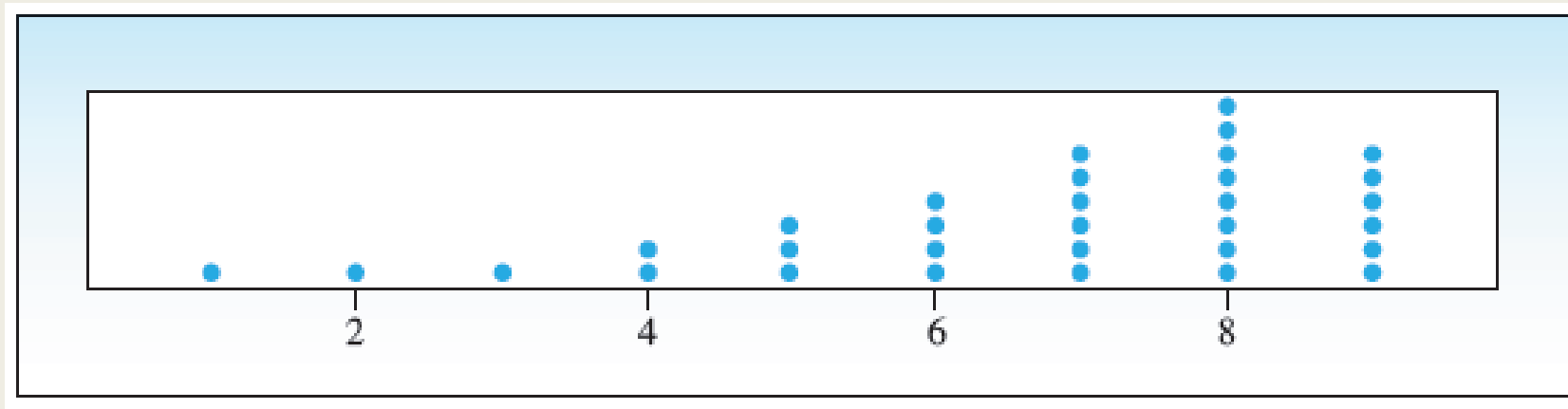


Tiene una larga “cola derecha”, lo cual significa que hay unas pocas observaciones extraordinariamente grandes.

Si se dobla la página en el pico, estaría en el lado derecho una proporción más grande de mediciones que en el izquierdo.

DISTRIBUCIONES

Una distribución está **sesgada a la izquierda** si una proporción mayor de las mediciones está a la izquierda del valor pico.

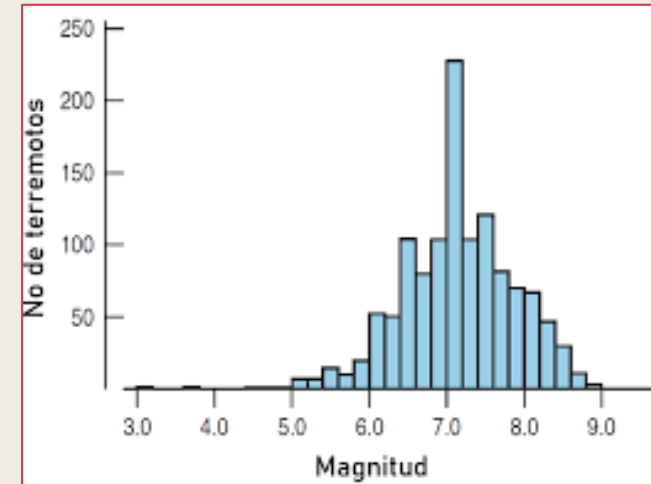


La tercera gráfica de puntos con una larga “cola a la izquierda”, por lo que tiene pocas mediciones extraordinariamente chicas.

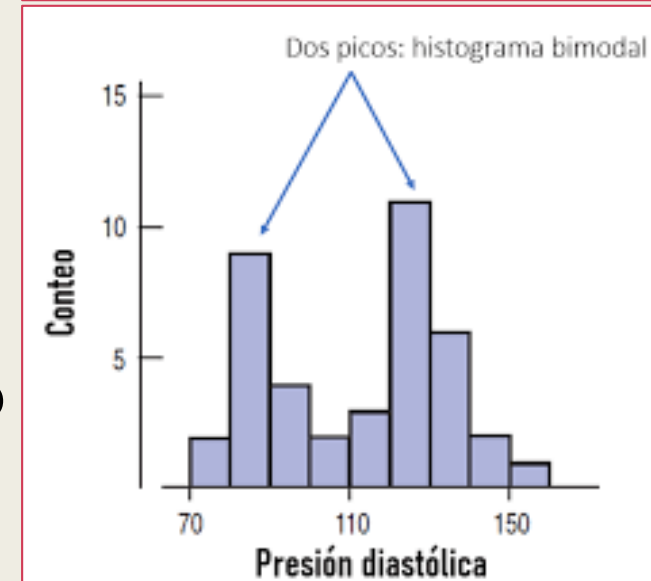
Si se dobla la página en el pico, estaría en el lado izquierdo una proporción más grande de mediciones que en el derecho.

DISTRIBUCIONES

Una distribución es **unimodal** si tiene un pico.



Una distribución **bimodal** tiene dos picos.



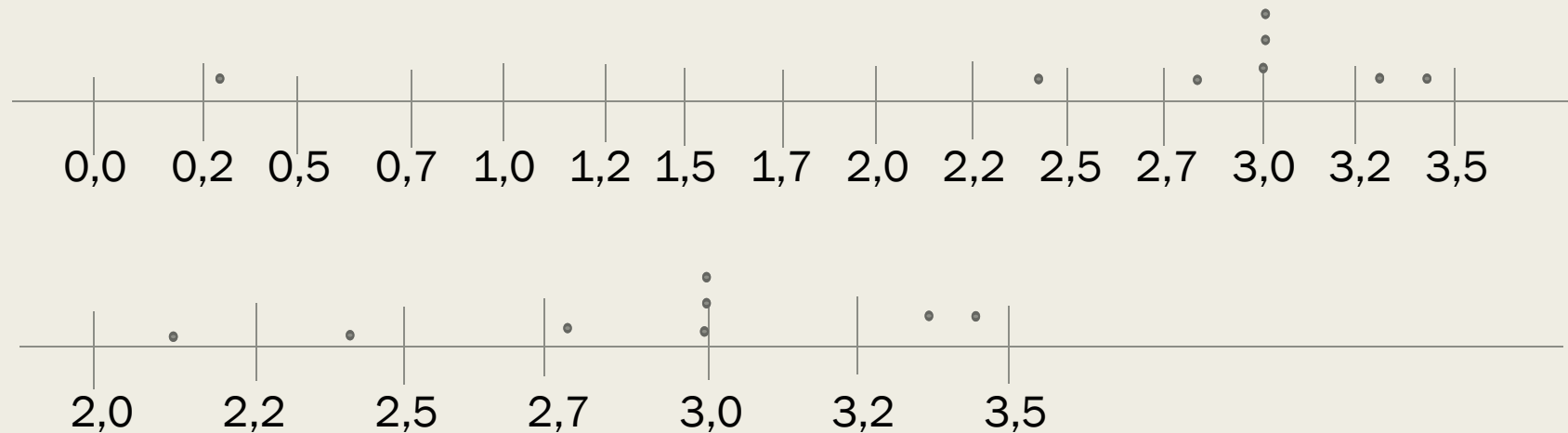
Las distribuciones bimodales representan a veces una combinación de dos poblaciones diferentes del conjunto de datos.

DISTRIBUCIONES - ERRORES

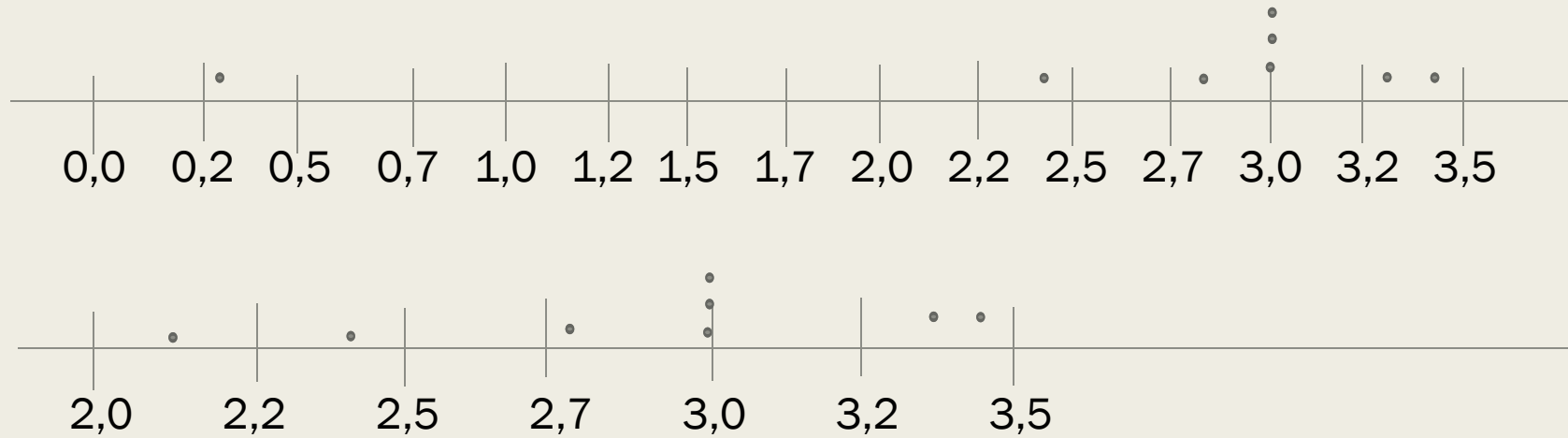
EJEMPLOS:

Un asistente administrativo del departamento de atletismo de una universidad local está observando los promedios de calificaciones de ocho miembros del equipo femenino de voley. El asistente introduce los promedios en la base de datos pero por accidente coloca mal el punto decimal de la última entrada.

2.8 - 3.0 - 3.0 - 3.3 - 2.4 - 3.4 - 3.0 - .21



DISTRIBUCIONES - ERRORES



Cuando se comparen gráficas creadas para dos conjuntos de datos, se deben **comparar** sus **escalas** de medición, **ubicaciones** y **formas**, y buscar mediciones poco comunes o **resultados atípicos**.

Dan información muy valiosa que no debe ser soslayada.

No siempre son causados por errores o introducción errónea de datos.

Si las escalas difieren en mucho, debe tenerse cuidado al hacer comparaciones o ¡sacar conclusiones que pudieran ser imprecisas!

EJERCITACIÓN

- 11.** Construya una gráfica de tallo y hoja para estas 50 mediciones:
- a. Use la gráfica de tallo y hoja para hallar la observación mínima.
 - b. Encuentre la octava y novena observaciones más grandes.
 - c. Si desea hacer un histograma de frecuencia relativa: ¿Aproximadamente cuántos intervalos de clase debe usar?
 - d. Suponga que usted decide usar clases que empiezan en 1,6 con ancho de clase de 0,5 (es decir, 1,6 a 2,1...) Construya el histograma de frecuencia relativa para los datos.
 - e. ¿Qué fracción de las mediciones es menos a 5,1?
 - f. ¿Qué fracción de las mediciones es mayor a 3,6?

3.1	4.9	2.8	3.6	2.5	4.5	3.5	3.7	4.1	4.9
2.9	2.1	3.5	4.0	3.7	2.7	4.0	4.4	3.7	4.2
3.8	6.2	2.5	2.9	2.8	5.1	1.8	5.6	2.2	3.4
2.5	3.6	5.1	4.8	1.6	3.6	6.1	4.7	3.9	3.9
4.3	5.7	3.7	4.6	4.0	5.6	4.9	4.2	3.1	3.9

EJERCITACIÓN

10. Las cuatro ramas de las fuerzas armadas en Estados Unidos son muy diferentes en su formación con respecto a las distribuciones de género, raza y edad. La tabla siguiente muestra el desglose racial de los miembros del Ejército y la Fuerza Aérea de Estados Unidos.

- Defina la variable que se ha medido en esta tabla
- ¿La variable es cuantitativa o cualitativa?
- ¿Qué representan los números?
- Construya una gráfica de pastel para describir el desglose racial en el Ejército de Estados Unidos.
- Construya una gráfica de barras para describir el desglose de razas en la Fuerza Aérea de Estados Unidos.
- ¿Qué porcentaje de los miembros del Ejército de Estados Unidos son minorías, es decir, no blancos?
¿Cuál es este porcentaje en la Fuerza Aérea de Estados Unidos?

	Ejército	Fuerza Aérea
Blanco	58.4%	75.5%
Negro	26.3%	16.2%
Latino	8.9%	5.0%
Otro	6.4%	3.3%

EJERCITACIÓN

12. Una variable discreta puede tomar sólo los valores de 0, 1 o 2. Un conjunto de 20 mediciones en esta variable se muestra:

- a. Construya un histograma de frecuencia relativa para los datos.
- b. ¿Qué proporción de las mediciones es mayor a 1?
- c. ¿Qué proporción de las mediciones es menor a 2?
- d. Si una medición se selecciona al azar de entre las 20 mediciones mostradas, ¿cuál es la probabilidad de que sea un 2?
- e. Describa la forma de la distribución. ¿Ve algunos resultados atípicos?

1	2	1	0	2
2	1	1	0	0
2	2	1	1	0
0	1	2	1	1

EJERCITACIÓN

13. El tiempo (en meses) entre el inicio de una enfermedad en particular y su recurrencia se registró para $n = 50$ pacientes:

- Construya un histograma de frecuencia relativa para los datos.
- ¿Describiría usted la forma como aproximadamente simétrica, sesgada a la derecha o sesgada a la izquierda?
- Dé la fracción de tiempos de recurrencia menores o iguales a 10 meses.

2.1	4.4	2.7	32.3	9.9	9.0	2.0	6.6	3.9	1.6
14.7	9.6	16.7	7.4	8.2	19.2	6.9	4.3	3.3	1.2
4.1	18.4	.2	6.1	13.5	7.4	.2	8.3	.3	1.3
14.1	1.0	2.4	2.4	18.0	8.7	24.0	1.4	8.2	5.8
1.6	3.5	11.4	18.0	26.7	3.7	12.6	23.1	5.6	.4