CBS
Colegio Bautista Shalom

Estadística Comercial
Sexto PCOC PFS
Primer Bimestre

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Contenidos

ESTADÍSTICA

- ✓ ¿POR QUÉ ESTUDIAMOS ESTADÍSTICA?
- ✓ CRITERIOS ESENCIALES EN LA ESTADÍSTICA.
- ✓ ORIGEN DE LA ESTADÍSTICA.
- ✓ TIPOS DE ESTADÍSTICA.
- ✓ CONCEPTOS BÁSICOS DE LA ESTADÍSTICA.
- ✓ GLOSARIO.
- ✓ ORDENAMIENTO DE LOS DATOS.
- ✓ TABULAR.
- ✓ DISTRIBUCIÓN.
- ✓ CLASE.
- ✓ FRECUENCIA.
- ✓ FRECUENCIA ABSOLUTA.
- ✓ FRECUENCIA RELATIVA.
- ✓ FRONTERAS DE CLASE.
- ✓ MARCA DE CLASE O PUNTO MEDIO.
- ✓ FRECUENCIA ACUMULADA Y FRECUENCIA RELATIVA ACUMULADA.
- ✓ REPRESENTACIÓN GRÁFICA DE DATOS.
  - GRAFICA DE BARRAS.
  - HISTOGRAMA.
  - GRAFICA BIDIRECCIONAL.
  - GRAFICA DE LÍNEAS O DIAGRAMA DE LÍNEAS.
  - POLÍGONO DE FRECUENCIAS.
  - GRAFICA O DIAGRAMA DE SECTORES.
  - PICTOGRAMA.
  - CARTOGRAMA.

NOTA: conforme vayas avanzando en tu aprendizaje, encontrarás ejercicios que debes resolver. Sigue las instrucciones de tu catedrático(a).

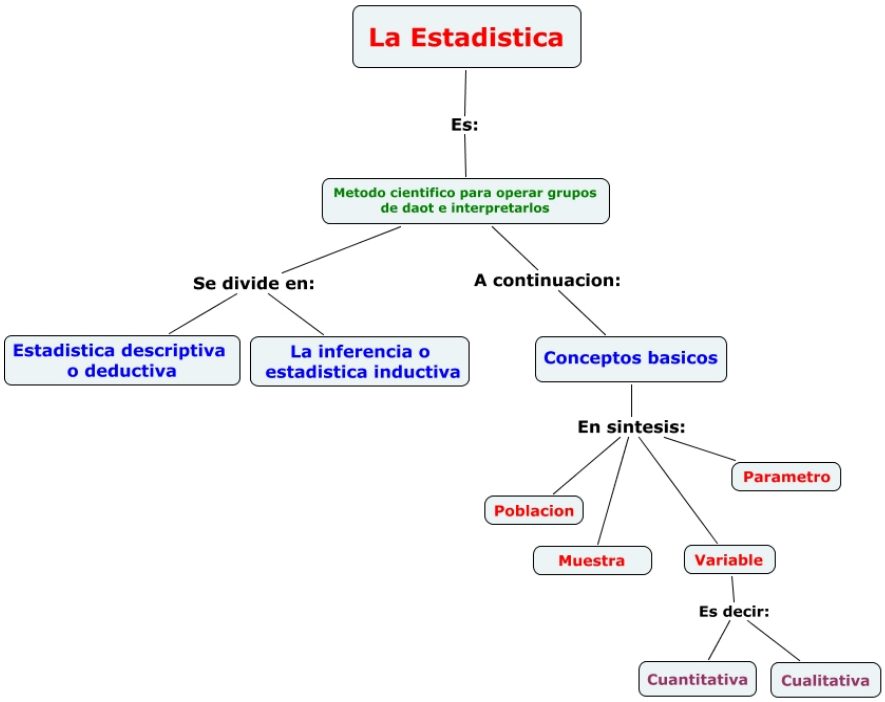
¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

ESTADÍSTICA

DEFINICIONES QUE PODEMOS DESTACAR:

1. Rama de la matemática que se ocupa de reunir, organizar, presentar, analizar e interpretar datos numéricos.
2. La Estadística es la ciencia cuyo objetivo es reunir una información cuantitativa concerniente a individuos, grupos, series de hechos… y deducir de ello gracias al análisis de estos datos unos significados precisos o unas previsiones para el futuro. La estadística, en general, es la ciencia que trata de la recopilación, organización presentación, análisis e interpretación de datos numéricos con el fin de realizar una toma de decisión más efectiva.
3. La ciencia que tiene por objeto aplicar las leyes de la cantidad a los hechos sociales para medir su intensidad, deducir las leyes que los rigen y hacer su predicción próxima.
4. La Estadística es la parte de las Matemáticas que se encarga del estudio de una determinada característica en una población, recogiendo los datos, organizándolos en tablas, representándolos gráficamente y analizándolos para sacar conclusiones de dicha población.
5. La estadística estudia los métodos científicos para recoger, organizar, resumir y analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas en tal análisis.

Podemos concluir que, la estadística se define como el arte y la ciencia de reunir datos, analizarlos, presentarlos e interpretarlos.



¿POR QUÉ ESTUDIAMOS ESTADÍSTICA?

Los conceptos y temas de la estadística se utilizan en la actualidad en un gran número de ocupaciones.

Las técnicas estadísticas constituyen una parte integral de las actividades de investigación en distintas áreas del saber humano.

La persona que comprenda los conceptos estadísticos y su metodología obtendrá mejor provecho de ellos.

El profesional, que entiende de estadística puede leer con inteligencia la literatura que sobre su campo de acción va apareciendo día con día.

Con frecuencia escuchamos en los medios de difusión comentarios como los siguientes:

- ✓ Se ha demostrado estadísticamente que el mayor porcentaje de las ventas de automóviles se registran en el primer trimestre del año.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

- ✓ La explotación de petróleo crudo en el último trimestre del año de 1993 ascendió a 285 millones de barriles, cuyo producto fue de 3698 millones de dólares.
- ✓ Estadísticamente no se ha demostrado que el huevo produce el colesterol en las personas que consumen mucho este producto.
- ✓ Se ha comprobado estadísticamente, que la pasta dental de mayor aceptación por el público es la que produce la fábrica Colgate-Palmolive.

Todas estas expresiones nos indican que la Estadística es una herramienta que ayuda a conocer la realidad. Sin embargo, también puede servir para distorsionar la verdad si no se tiene cuidado al usar los métodos estadísticos adecuadamente y si la interpretación de los resultados lo hacen incorrectamente.

Es famoso la frase que en cierta ocasión el ministro inglés Benjamin Disraeli dijo: Hay tres clases de mentiras que son: "Las simples, las malvadas y las de la Estadística".

Esta acusación hecha hace muchos años, ha llegado a convertirse en una descripción adecuada de algunos engaños que se pueden realizar mediante la Estadística.

Sucede algunas veces que en la presentación de resultados de alguna investigación poco seria, se utiliza la Estadística para dar la impresión de que todo ha sido realizado en forma científica y al comprobarlo, resulta que la información fue alterada con la finalidad de obtener un resultado favorable pero falso.

Se debe tener cuidado en los reportes estadísticos en no caer en una falacia; para evitarlo se debe tener una actitud crítica hacia cualquier escrito que leamos. Así mismo hacer un análisis cuidadoso que nos indique si el contenido y la interpretación son correctos o si se ha hecho un mal uso de la Estadística.

Para un análisis crítico de la información contenido en un reporte, debemos contestarnos las siguientes preguntas:

1) ¿Cómo es la muestra? La muestra debe ser representativa de la población en estudio.

2) ¿Qué se está midiendo y cómo? Es importante saber de qué manera se obtienen los datos y si las características de interés pueden ser medidas.
   - **a.** Ejemplo: Se desea saber qué tiempo en horas un alumno dedica al estudio, para ello el encuestador le hace la siguiente pregunta: Tú estudias dos horas diarias, ¿verdad? Esta forma de obtener la información es incorrecta toda vez que la respuesta es guiada por el encuestador y la mayoría van a ser afirmativas.

3) ¿Qué se está probando y cómo?

4) ¿Se está hablando siempre de lo mismo? Con frecuencia se usan datos reales para obtener conclusiones falsas. En algunos estudios las trampas de este tipo pueden estar en proceso de medición, veamos el siguiente ejemplo:
   - **a.** En una aldea se levantó un censo de todos los habitantes que no sabían leer para inscribirlos en los cursos de alfabetismo. El encuestador reportó 3000 analfabetos que fueron los que detectó, pero él no sabía que mucha de esta gente no querían aprender a leer y por lo tanto no se presentaron para su reporte.
   - **b.** En la misma aldea se anunció por la radio que toda persona que no supiera leer, el gobierno le otorgaría 100 quetzales mensuales de ayuda por lo tanto debían anotarse en la municipalidad las personas que cubrieran el requisito. A este anuncio acudieron 7,000 personas que no sabían leer. Esto nos muestra que la información no siempre va a ser verdadera y en casos como estos, depende del interés de la gente.

5) ¿Tiene sentido la información? Esta pregunta nos invita a analizar la información reportada y nosotros debemos ver si es congruente con el sentido común.
   - **a.** Ejemplo: Un estudio reporta que la cosecha de frijol está en relación directa con la precipitación pluvial, es decir, si llueve más, hay mayor cosecha. Esto es cierto en determinados límites toda vez que con demasiada lluvia se arruina la cosecha.

6) ¿Qué información falta? Existe información que no es estrictamente estadística pero es necesaria para poder captar la que nos interesa por estar relacionada con ella.
   - **a.** Ejemplo: En una librería se reportó que en tan sólo un mes se duplicaron las ventas, sin embargo, no se especificó que se refiere al mes de septiembre que fue el inicio de clases y además aumentó el costo de los libros debido a la demanda.

7) ¿Quién lo dice? Muchas de las estadísticas que dan los medios de difusión se basan en la publicidad.
   a. Ejemplo: Un anuncio de la televisión dice: "Kma su fábrica amiga en tan sólo un mes vendió 50,000 colchones; pero aún nos quedan 1,000 que se rematan a un precio muy bajo, venga y llévese el suyo. ¿Será cierta esta información, o solamente es un comercial sensacionalista?

CRITERIOS ESENCIALES EN LA ESTADÍSTICA

- ✓ **Validez:** medir lo que tiene que medir. Por ejemplo: para medir un peso debo usar una balanza y utilizar unidades de peso: libras, kilos… no debería usar un metro y usar medidas de longitud.
- ✓ **Pertinencia:** si tiene propósito o pertenece o corresponde a lo que se estudia. Ejemplo: si lo que se estudia está relacionado con el peso, lo que se estudia debe corresponder a estas unidades de medida.
- ✓ **Confiabilidad:** es el grado de precisión de lo estudiado. Ejemplo: no es correcto decir pesa más o menos 1 libra, confiable es decir: pesa 16 gramos exactos.

ORIGEN DE LA ESTADÍSTICA

El término alemán Statistik, introducido originalmente por Gottfried Achenwall en 1749, se refería al análisis de datos del Estado, es decir, la «ciencia del Estado». También se llamó aritmética política de acuerdo con la traducción literal del inglés. No fue hasta el siglo XIX cuando el término estadística adquirió el significado de recolectar y clasificar datos. Este concepto fue introducido por el militar británico sir John Sinclair (1754-1835).

En su origen, por tanto, la estadística estuvo asociada a los Estados o ciudades libres, para ser utilizados por el gobierno y cuerpos administrativos. La colección de datos acerca de estados y localidades continúa ampliamente a través de las instituciones de estadística ya sea nacionales e internacionales. En particular, los censos comenzaron a suministrar información regular acerca de la población de cada país. Así pues, los datos estadísticos se referían originalmente a los datos demográficos de una ciudad o Estado determinados.

Los libros bíblicos de Números y Crónicas incluyen en algunas partes trabajos de estadística. El primero contiene dos censos de la población de la Tierra de Israel y el segundo describe el bienestar material de las diversas tribus judías. Los griegos y los romanos realizaban censos cuya información se utilizaba para cobrar impuestos.

TIPOS DE ESTADÍSTICA

- ✓ **Estadística descriptiva o deductiva:**
  - Recolecta, agrupa y presenta los datos de una manera que permita describirlos de una manera fácil, comprensible y rápida.
  - Rama de la ciencia estadística que se encarga desde la recopilación, procesamiento y análisis de la información siendo sus conclusiones válidas sólo para el grupo analizado.
  - Realiza el estudio sobre la población completa, observando una característica de la misma y calculando unos parámetros que den información global de toda la población.
  - Es un proceso mediante el cual se recopila, organiza, presenta, analiza e interpreta datos de manera tal que describa fácil y rápidamente las características esenciales de dichos datos mediante el empleo de métodos gráficos, tabulares o numéricos.
  - La Estadística descriptiva se ocupa de tomar los datos de un conjunto, organizarlos en tablas o en representaciones gráficas y del cálculo de unos números que nos informen de manera global del conjunto estudiado.

- ✓ **Estadística inferencial o inductiva:**
  - Rama de la ciencia estadística que proporciona métodos y procedimientos que permiten obtener conclusiones para una población a partir del estudio de una o más muestras representativas.
  - Realiza el estudio descriptivo sobre un subconjunto de la población llamado muestra y, posteriormente, extiende los resultados obtenidos a toda la población.
  - Comprende los métodos y procedimientos que por medio de la inducción determina propiedades de una población estadística, a partir de una pequeña parte de la misma.
  - Parte de la estadística que comprende los métodos y procedimientos que por medio de la inducción determina propiedades de una población estadística, a partir de una pequeña parte de la misma.
  - La Estadística inferencial trata sobre la elaboración de conclusiones para la población, partiendo de los resultados de una muestra y del grado de fiabilidad de estas conclusiones.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

- En muchas situaciones se requiere información acerca de grupos grandes de elementos (individuos, empresas, votantes, hogares, productos, clientes, etc.). Pero, debido al tiempo, costo y a otras consideraciones, sólo es posible recolectar los datos de una pequeña parte de este grupo. Al grupo grande de elementos en un determinado estudio se le llama población y al grupo pequeño muestra.

CONCEPTOS BÁSICOS DE LA ESTADÍSTICA

Datos: algunas definiciones.

- ✓ Información sobre algo concreto que permite su conocimiento exacto o sirve para deducir las consecuencias derivadas de un hecho.
- ✓ Datos son hechos, informaciones o cifras que se recogen, analizan y resumen para su presentación e interpretación. A todos los datos reunidos para un determinado estudio se les llama conjunto de datos para el estudio.
- ✓ Es un valor particular.
- ✓ Características o números que son recolectados por observación. No son otra cosa que el producto de las observaciones efectuadas en las personas y objetos en los cuales se produce el fenómeno que queremos estudiar

A los datos reunidos de un determinado estudio se les llama conjunto de datos.

Los datos estadísticos pueden ser clasificados en cualitativos, cuantitativos, cronológicos y geográficos:

- ✓ **Datos cualitativos:** se refieren a la calidad y no a la cantidad de los datos. La calidad se refiere a la propiedad o conjunto de propiedades inherentes a algo, que permiten juzgar su valor. Ejemplo: Si deseamos clasificar los estudiantes que cursan la materia de estadística por su estado civil, observamos que pueden existir solteros, casados, divorciados, viudos.
- ✓ **Datos cuantitativos:** cuando los valores de los datos representan diferentes magnitudes, decimos que son datos cuantitativos. Ejemplo: Se clasifican los estudiantes de acuerdo a sus notas, observamos que los valores (nota) representan diferentes magnitudes.
- ✓ **Datos cronológicos:** cuando los valores de los datos varían en diferentes instantes o períodos de tiempo, los datos son reconocidos como cronológicos. Ejemplo: Al registrar los promedios de notas de los alumnos en los diferentes bimestres.
- ✓ **Datos geográficos:** cuando los datos están referidos a una localidad geográfica se dicen que son datos geográficos. Ejemplo: El número de estudiantes de educación superior en las distintas regiones del país.
- ✓ **Datos de sección transversal:** son los obtenidos en el mismo o aproximadamente el mismo momento (punto en el tiempo).
- ✓ **Datos de series de tiempo:** son datos obtenidos a lo largo de varios periodos.
- ✓ **Datos porcentuales:** cuando es necesario cuantificar una parte del total (100 * 20%) o determinar el valor relativo entre dos cantidades (cantidad A / cantidad B = C * 100)

Elementos o registro o casos: es el conjunto de mediciones o datos realizados. Llamado también **Individuo u Objeto:** es cualquier elemento que brinda información sobre el fenómeno que se estudia. Es cada elemento de la población.

Variables: es una característica de los elementos.

- ✓ **Variables unidimensionales:** sólo recogen información sobre una característica (por ejemplo: edad de los alumnos de una clase).
- ✓ **Variables bidimensionales:** recogen información sobre dos características de la población (por ejemplo: edad y altura de los alumnos de una clase).
- ✓ **Variables pluridimensionales:** recogen información sobre tres o más características (por ejemplo: edad, altura y peso de los alumnos de una clase).
- ✓ **Variable aleatoria:** es la que toma al azar los probables resultados de un experimento.
- ✓ **Variable dependiente:** es la que toma los valores correspondientes de un modelo matemático o que los toma debido a la influencia de otra variable independiente.
- ✓ **Cualitativas o atributos:** no se pueden medir numéricamente: bueno, malo, bonito, feo…
- ✓ **Cuantitativas:** tienen un valor numérico.
  - **Discretas:** valores enteros: …-2, -1, 0, 1, 2…
  - **Continuas:** puede tomar cualquier valor de un intervalo de la recta numérica real (valor decimal): …-2, -1.6, -1.5, -1.00001, 0, 1, 1.005, 1.7, 1.9, 2.01…

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

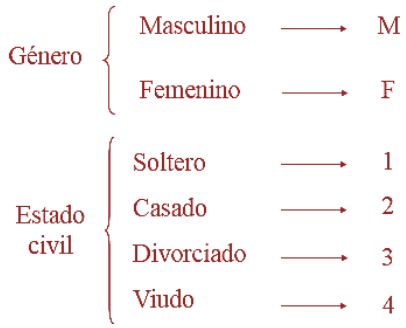
Escalas de medición de las variables: Las variables se pueden clasificar de acuerdo al nivel de medición.

- ✓ **Nominal:** se clasifica a los sujetos en categorías, mutuamente excluyentes y totalmente exhaustivas, tal que todos los sujetos clasificados en la misma categoría son equivalentes respecto a la variable que se está midiendo. Son aquellas en las que los códigos solo indican una diferencia en la categoría, clase, calidad o tipo.

  Sólo tiene sentido la relación de igualdad-desigualdad.

  Se pueden usar números, letras o símbolos para identificar a cada categoría de la variable.

  No se puede realizar ninguna operación aritmética en esta escala.



- ✓ **Ordinal:** se usa cuando es posible establecer una relación de orden entre las distintas categorías de la variable. Designan categorías, pero tienen la propiedad adicional de permitir la clasificar las categorías desde la mayor hasta la menor, de la mejor a la peor o de la primera a la última.

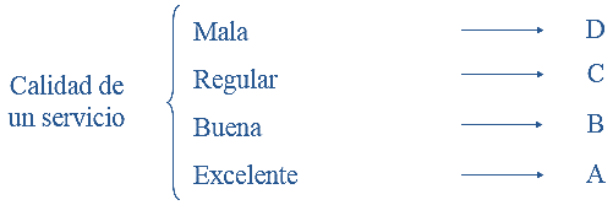
  Es decir, prevalece la relación de orden "mayor que" (>).

  Se pueden usar letras o números para identificar a cada categoría de la variable.

  Los números o letras usados deben reflejar el orden de las categorías.

  No se pueden realizar operaciones aritméticas entre los números asignados a las distintas categorías.

  Tales números solo reflejan una relación de orden.



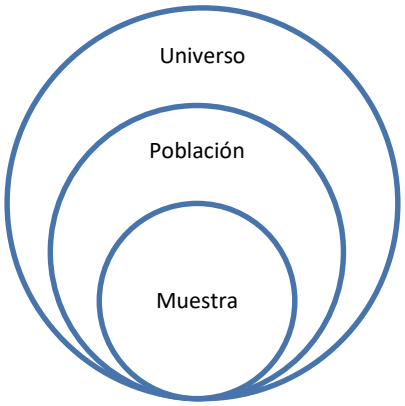
1. La temperatura de una ciudad medida en grados Fahrenheit o Celsius.
2. La altura de las ciudades usando como referencia el nivel del mar.
3. El rendimiento académico medido en una escala del 0 al 20.

- ✓ **De intervalo:** posee una unidad de medida constante y arbitraria.

  Posee un cero "arbitrario", es decir, no indica la ausencia de la característica que se está midiendo.

  Prevalece la relación de orden "mayor que" (>).

  Entre los valores de la variable solo es posible realizar la suma y la resta como operaciones aritméticas.



Observación: es el valor específico de una variable: es un dato.

Universo: es el conjunto de elemento que tiene una característica común, observable y capaz de ser medida. Ejemplo: todos los estudiantes de Guatemala.

Nota. Generalmente los términos población y universo tienden a considerarse sinónimos.

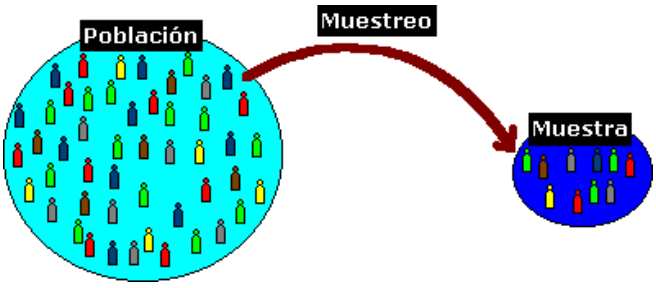
Población: conjunto de todas las mediciones u observaciones hechas sobre una o varias de las características de los elementos del universo. Conjunto de individuos. Población estadística es un conjunto de personas, entidades u objetos del cual se quiere saber algo que nos interesa para tomar una determinación acertada. Población es un conjunto definido, limitado y accesible del universo que forma el referente para la elección de la muestra. Es el grupo al que se intenta generalizar los resultados". Ejemplo: los estudiantes del Colegio Bautista Shalom.

- ✓ **Población finita:** es el conjunto compuesto por una cantidad limitada de elementos, como el número de especies, el número de estudiantes, el número de obreros.
- ✓ **Población infinita:** es la que tiene un número extremadamente grande de componentes, como el conjunto de estrellas.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

- ✓ **Población real:** es todo el grupo de elementos concretos, como las personas que en Guatemala se dedican a actividades artísticas.
- ✓ **Población hipotética:** es el conjunto de situaciones posibles imaginables en que puede presentarse un suceso, como por ejemplo las formas de reaccionar de una persona ante una catástrofe.
- ✓ **Población estable:** es aquella en que sus calores o cualidades no presentan variaciones, o éstas, por pequeñas que sean, son despreciables, como la rotación de la tierra o la velocidad de la luz.
- ✓ **Población inestable:** es la que contienen los valores en constante cambio. Prácticamente la totalidad de las poblaciones corresponden a este tipo. El cambio de los valores se presenta en el tiempo o en el espacio.
- ✓ **Población aleatoria:** es la que presenta cambios en sus calores debidos al azar, sin que exista una causa aparente, como las variaciones en el contenido del producto.
- ✓ **Población dependiente:** es la que cambia sus valores debido a una causa determinada y medida. La dependencia puede ser total, como las variaciones obtenidas en una función matemática, la regresión lineal, por ejemplo. La dependencia es parcial cuando la causa influye en la variable dependiente en una proporción menor a la total, por ejemplo, el incremento en las ventas proveniente de un mayor gasto publicitario.
- ✓ **Población binomial:** es aquella en la que se busca la presencia o ausencia de una característica, por ejemplo, la presencia de ozono en el aire.
- ✓ **Población polinomial**: es la que tiene varias características que deben ser definidas, medidas o estimadas, como la obediencia, la inteligencia y la edad de los alumnos de postgrado.

Muestra: subconjunto obtenido de la población. Una muestra es un conjunto de medidas u observaciones tomadas a partir de una población dada; es un subconjunto de la población. Las muestras se obtienen con la intención de inferir propiedades de la totalidad de la población, para lo cual deben ser representativas de la misma. Las muestras se toman debido a que no es factible desde el punto de vista económico usar a toda la población. En algunos casos es imposible recolectar todas las posibles observaciones en la población. Ejemplo: los estudiantes de 4 y 5 bachillerato del Colegio Bautista Shalom.



El muestreo: En un universo de trabajo en donde se desea aplicar un análisis estadístico, cuando el muestreo cubre a todos los elementos de la población., se realiza un censo. En muchos de los casos, la realización de un censo no es posible por ser muy costoso, muy extenso o que la muestra se destruya como resultado del análisis. En tales oportunidades se debe practicar un análisis muestral. La muestra es una parte seleccionada de la población que deberá ser representativa, es decir, reflejar adecuadamente las características que deseamos analizar en el conjunto en estudio.

Se pueden realizar diferentes tipos de muestreo, que quedan clasificados en dos grandes grupos: probabilísticos y no probabilísticos.

- ✓ **Muestreo no probabilístico:** por su parte, los elementos de la muestra se seleccionan siguiendo criterios determinados siempre procurando la representatividad de la muestra.
- ✓ **Muestreo probabilístico:** todos los individuos o elementos de la población tienen la misma probabilidad de ser incluidos en la muestra extraída, asegurándonos la representatividad de la misma.

  - **Muestreo probabilístico aleatorio simple:** cuando todos los elementos de la población tienen la misma probabilidad de ser seleccionados en la muestra y esta probabilidad es conocida. Este tipo de muestreo es más recomendable, pero resulta mucho más difícil de llevarse a cabo y, por lo tanto, es más costoso. Para seleccionar una muestra de este tipo se requiere tener en forma de lista todos los elementos que integran la población investigada y utilizar tablas de números aleatorios. **Ejemplo:** A un grupo de 100 personas se les numera de uno a cien y se depositan en una urna 100 bolitas a su vez numeradas de uno a cien. Para obtener una muestra aleatoria simple de 20 elementos, tendríamos que sacar 20 bolitas numeradas de la urna que nos seleccionarán en forma completamente al azar a los 20 elementos escogidos para que opinen sobre un nuevo producto.
  - **Muestreo probabilístico aleatorio sistemático:** el cual es susceptible de ser más preciso que el muestreo aleatorio simple. Se elige un primer elemento del universo y luego se van escogiendo otros elementos igualmente espaciados a partir del primero. Consiste en dividir la población en n estratos, compuestos por las primeras K unidades, las segundas k unidades y así sucesivamente.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Ejemplo: a partir de una lista de 100 establecimientos de comestibles, deseamos seleccionar una muestra probabilística de 20 tiendas. La forma de hacerlo sería:

- dividir 100 entre 20 para obtener 5, que es un proceso sistemático
- extraer un número al azar entre 1 y 5. Supóngase que es el número 2 el cual corresponde al primer elemento seleccionado.
- Se incluyen en la muestra de establecimientos numerados: 2, 7, 12, 17, 22… 97.

Ejemplo general:

Universo: los estudiantes de Guatemala.
Población: los estudiantes del Colegio Bautista Shalom.
Muestra: los estudiantes de 4 y 5 bachillerato que pesen entre 120 y 160 libras entre 16 y 18 años y que midan entre 1.50 y 1.65 metros.

Estudiante	Sexo	Edad	Peso	Altura	Zona donde vive	Grado	Ingresos	Variables
1	F	16	120	1.55	18	4	Q500	
2	M	17	130	1.6	6	5	Q500	
3	M	18	140	159	18	5	Q500	Registro
4	M	17	125	1.62	6	5	Q500	
5	F	16	160	1.63	2	4	Q500	

Observación o dato

Variables: estudiante, sexo, edad, peso, altura, zona, grado, ingresos…

Glosario:

- ✓ **Estadística:** el arte y la ciencia de recolectar, analizar, presentar e interpretar datos.
- ✓ **Datos:** los hechos y las cifras que se recolectan, analizan y resumen para su presentación e interpretación.
- ✓ **Conjunto de datos:** todos los datos recolectados en un estudio determinado.
- ✓ **Elementos:** Entidades sobre las que se recolectan los datos.
- ✓ **Variable:** una característica que interesa de un elemento.
- ✓ **Observación:** el conjunto de mediciones obtenidas de un elemento determinado.
- ✓ **Escala nominal:** escala de medición de una variable cuando los datos son etiquetas o nombres que se emplean para identificar un atributo de un elemento. Los datos nominales pueden ser no numéricos o numéricos.
- ✓ **Escala ordinal:** escala de medición de una variable cuando los datos presentan las propiedades de los datos nominales y el orden o jerarquía de los datos tiene sentido. Los datos ordinales pueden ser no numéricos o numéricos.
- ✓ **Escala de intervalo:** escala de medición de una variable cuando los datos presentan las propiedades de los datos ordinales y los intervalos entre valores se expresan en términos de una unidad o medida fija. Los datos de intervalo siempre son numéricos.
- ✓ **Escala de razón:** escala de medición de una variable cuando los datos presentan todas las propiedades de los datos de intervalo y la razón entre dos valores tiene sentido. Los datos de razón siempre son numéricos.
- ✓ **Datos cualitativos:** etiquetas o nombres utilizados para identificar un atributo de cada elemento.
- ✓ Los datos cualitativos usan las escalas de medición nominal y ordinal y pueden ser no numéricos o numéricos.
- ✓ **Datos cuantitativos:** valores numéricos que indican cuánto o cuántos de algo. Los datos cuantitativos se obtienen mediante la escala de intervalo o de razón.
- ✓ **Variable cualitativa:** una variable con datos cualitativos.
- ✓ **Variable cuantitativa:** una variable con datos cuantitativos.
- ✓ **Datos de sección transversal:** datos recolectados en el mismo o aproximadamente en el mismo momento.
- ✓ **Datos de series de tiempo:** datos recolectados a lo largo de varios periodos de tiempo.
- ✓ **Estadística descriptiva:** resúmenes tabulares, gráficos o numéricos de datos.
- ✓ **Población:** conjunto de todos los elementos que interesan en un estudio determinado.
- ✓ **Muestra:** un subconjunto de la población.
- ✓ **Censo:** un estudio para recolectar los datos de toda la población.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

- ✓ **Encuesta muestral:** un estudio para recolectar los datos de una muestra.
- ✓ **Inferencia estadística:** el proceso de emplear los datos obtenidos de una muestra para hacer estimaciones probar hipótesis acerca de las características de la población.

ACTIVIDAD DE EVALUACIÓN 1

1. La revista Condé Nast Traveler realiza una encuesta anual entre sus suscriptores con objeto de determinar los mejores alojamientos del mundo. En la tabla siguiente se presenta una muestra de nueve hoteles europeos (Condé Nast Traveler, enero de 2000).

Nombre del lugar	País	Precio de la habitación	Número de habitaciones	Calificación general
Graveteye Manor	Inglaterra	$$	18	83.6
Villa d'Este	Italia	$$$$	166	86.3
Hotel Prem	Alemania	$	54	77.8
Hotel d'Europe	Francia	$$	47	76.8
Palace Luzern	Suiza	$$	326	80.9
Royal Crescent Hotel	Inglaterra	$$$	45	73.7
Hotel Sacher	Austria	$$$	120	85.5
Duc de Bourgogne	Bélgica	$	10	76.9
Villa Gallici	Francia	$$	22	90.6

Fuente: Condé Nast Traveler, enero de 2000.

Los precios de una habitación doble estándar van de $(precio más bajo) a $$$$(precio más alto). La calificación general corresponde a la evaluación de habitaciones, servicio, restaurante, ubicación/atmósfera y áreas públicas; cuanta más alta sea la calificación general, mayor es el nivel de satisfacción.

a) ¿Cuántos elementos hay en este conjunto de datos? 9
b) ¿Cuántas variables hay en este conjunto de datos? 5
c) ¿Cuáles variables son cualitativas? Nombre, país y precio de la habitación
d) ¿Cuáles variables son cuantitativas?: Número de habitaciones y calificación general.
e) ¿Cuál es el número promedio de habitaciones en los nueve hoteles? La cantidad promedio de habitaciones: 808/9 = 89.78 o aproximadamente 90 habitaciones
f) Calcule la calificación general promedio. El promedio de las puntuaciones = 732.1/9 = 81.34
g) ¿Qué porcentaje de los hoteles se encuentra en Inglaterra? 2/9 = 0.222 * 100 = 22.22%
h) ¿En qué porcentaje de los hoteles el precio de la habitación es de $$? 4/9 = 0.4444 * 100 = 44.44 %

2. Los equipos de sonido todo en uno, llamados minicomponentes, cuentan con sintonizador AM/FM, casetera doble, cargador para un disco compacto con bocinas separadas. En la tabla siguiente se muestran los precios de menudeo, calidad de sonido, capacidad para discos compactos, sensibilidad y selectividad de la sintonización y cantidad de caseteras en los artículos de una muestra de 10 minicomponentes (Consumer Report Buying Guide 2002).

Marca y modelo	Precio ($)	Calidad de sonido	Capacidad para CD	Sintonización FM	Caseteras
Aiwa NSX-AJ800	250	Buena	3	Regular	2
JVC FS-SD1000	500	Buena	1	Muy buena	0
JVC MX-G50	200	Muy buena	3	Excelente	2
Panasonic SC-PM11	170	Regular	5	Muy buena	1
RCA RS 1283	170	Buena	3	Mala	0
Sharp CD-BA2600	150	Buena	3	Buena	2
Sony CHC-CL1	300	Muy buena	3	Muy buena	1
Sony MHC-NX1	500	Buena	5	Excelente	2
Yamaha GX-505	400	Muy buena	3	Excelente	1
Yamaha MCR-E100	500	Muy buena	1	Excelente	0

a) ¿Cuántos elementos contiene este conjunto de datos? 10
b) ¿Cuál es la población? Todas las marcas.
c) Calcule el precio promedio en la muestra. 314.
d) Con los resultados del inciso c, estime el precio promedio para la población. 314

3. La Columbia House vende discos compactos a los miembros de su club de venta por correo. En una

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

encuesta sobre música se les pidió a los nuevos miembros del club que llenaran un cuestionario con 11 preguntas. Algunas de las preguntas eran las siguientes. Responde si los datos que se obtienen con cada pregunta son cualitativos o cuantitativos.

a) ¿Cuántos discos compactos has comprado en los últimos 12 meses? Cuantitativo.
b) ¿Eres miembro de algún club de venta de libros por correo (Sí o No)? Cualitativo.
c) ¿Cuál es tu edad? Cuantitativo.
d) Incluyéndote a ti, de cuántas personas (adultos y niños) consta tu familia. Cuantitativo.
e) ¿Qué tipo de música te interesa comprar? Se presentaban quince categorías entre las que se encontraban rock pesado, rock ligero, música contemporánea para adultos, rap y rancheras. Cualitativo.

4. La empresa Gallup realizó una encuesta telefónica empleando una muestra aleatoria nacional compuesta de 1005 adultos de 18 años o más. En la encuesta se les preguntó a los participantes "Cómo considera que es su salud física en este momento" (www.gallup.com, 7 de febrero de 2002)". Las respuestas podían ser Excelente, Buena, Regular o Ninguna opinión.

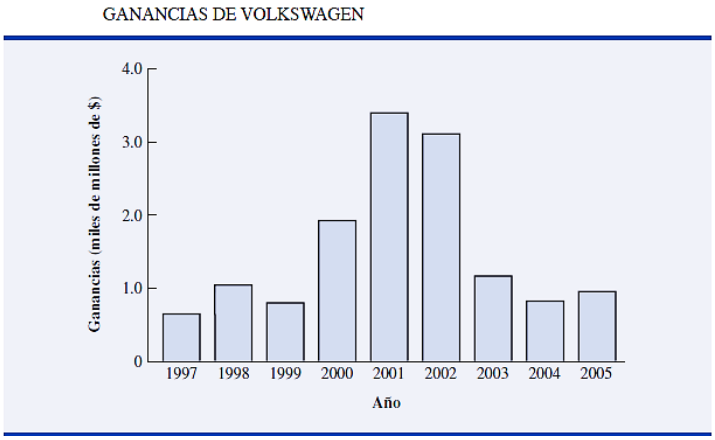
a) ¿Cuál es el tamaño de la muestra de esta investigación? 1005
b) ¿Son estos datos cualitativos o cuantitativos? Cualitativos.
c) ¿Sería conveniente usar promedios o porcentajes para resumir los datos de estas preguntas? Porcentajes.
d) De las personas que respondieron, 29% dijo que su salud era excelente. ¿Cuántos fueron los individuos que dieron esta respuesta? 1005 * 29% = 291.45

5. La Oficina de Visitantes a Hawái recolecta datos de los visitantes. Entre las 16 preguntas hechas a los pasajeros de un vuelo de llegada en junio de 2003 estaban las siguientes:

- ✓ Este viaje a Hawái es mi 1o., 2o., 3o., 4o…
- ✓ La principal razón de este viaje es: (10 categorías para escoger entre las que se encontraban vacaciones, luna de miel, una convención).
- ✓ Dónde voy a alojarme: (11 categorías entre las que se encontraban hotel, departamento, parientes, acampar).
- ✓ Total de días en Hawái

a) ¿Cuál es la población que se estudia? Todas las personas que visitan Hawái.
b) ¿El uso de un cuestionario es una buena manera de tener información de los pasajeros en los vuelos de llegada? Si.
c) Diga de cada una de las cuatro preguntas si los datos que suministra son cualitativos o cuantitativos. Cuantitativas: 1 y 4. Cualitativas: 2 y 3.

6. En la figura siguiente se presenta una gráfica de barras que resume las ganancias de Volkswagen de los años 1997 a 2005 (BusinessWeek, 26 de diciembre de 2005).



¿Estos son datos cualitativos o cuantitativos? Las ganancias en miles de millones son datos cuantitativos.
a) ¿Son datos de series de tiempo o datos de sección transversal? De serie de tiempo de 1997 a 2005
b) ¿Cuál es la variable que interesa en la gráfica? La ganancia
c) Comente la tendencia en las ganancias de Volkswagen a lo largo del tiempo. De 1997 a 1999 las ganancias

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

son bajas. En el año 2000 empiezan a subir y luego bajan en el 2002.

d) ¿Qué advertencia sugiere esta gráfica acerca de la proyección de datos como los de las ganancias de Volkswagen hacia el futuro? Que la ganancia tiende a permanecer estable con relación a los años 2003 y 2004.

ORDENAMIENTO DE LOS DATOS

Generalmente los datos se obtienen de observaciones con instrumentos de recolección de datos estructurados adecuadamente a la ocasión o situación; entre las que podemos mencionar la encuesta, los censos poblacionales y la entrevista. Los datos que se obtienen normalmente aparecen en forma desordenada, por consiguiente es necesario ordenado y se puede hacer de dos maneras

- ✓ Ascendente: en este caso se busca el más pequeño o de menor valor y se ordena de menor a mayor.
- ✓ Descendente: en este caso los valores so deben de colocarse de mayor a menor.

Ejemplo: a continuación se te presentan los resultados de 25 alumnos de la clase de matemática.

80	60	55	75	85
90	50	55	45	90
50	75	60	55	65
55	50	70	65	70
90	90	95	85	60

Ordenados ascendentemente quedarían así: 45, 50, 50, 50, 55, 55, 55, 55, 60, 60, 60, 65, 65, 70, 70, 75, 75, 80, 85, 85, 90, 90, 90, 90, 95.

Ordenados descendentemente quedarían así: 95, 90, 90, 90, 90, 85, 85, 80, 75, 75, 70, 70, 65, 65, 60, 60, 60, 55, 55, 55, 55, 50, 50, 50, 45.

Cuando el número de datos es grande, es conveniente organizar los datos numéricos en un arreglo ordenado para ayudar a comprender la información.

- ✓ **Tabular:** es la representación estructurada o arreglo ordenado, en forma de tabla, de toda la información que se ha recogido sobre la variable que se estudia.
- ✓ **Distribución:** es el acto de distribuir. Distribuir es dividir algo entre varias personas, designando lo que a cada una corresponde, según voluntad, conveniencia, regla o derecho. En estadística es la amplitud entre un límite menor y uno mayor (intervalo) en la cual se encuentran divididos los datos. Se suelen utilizar intervalos impares: 5, 7, 9, 11… y no deben tener menos de 5 ni más de 15. Para crear la distribución es necesario utilizar el sentido común.
- ✓ **Clase:** una clase es un rango, grupo o intervalo de datos numéricos. Se agrupan los datos en intervalos que tengan la misma amplitud denominados clases. A cada clase se le asigna su frecuencia correspondiente.
  - **Límites de la clase:** cada clase está delimitada por el límite inferior de la clase y el límite superior de la clase.
  - **Amplitud de la clase:** la amplitud de la clase es la diferencia entre el límite superior e inferior de la clase.
  - **Marca de clase:** la marca de clase es el punto medio de cada intervalo y es el valor que representa a todo el intervalo para el cálculo de algunos parámetros.

FRECUENCIA DE DATOS

- ✓ **Frecuencia:** frecuencia es una repetición de un hecho o un suceso. Es también el número de veces que se repite un proceso periódico en un intervalo de tiempo determinado. Número de oscilaciones, vibraciones u ondas por unidad de tiempo en cualquier fenómeno periódico. La cantidad de veces que se repite un dato. Es el número de elementos comprendidos en un intervalo con una distribución determinada. En Estadística, frecuencia es el número de veces que el valor de una variable se repite. Se distinguen dos tipos principales de frecuencia: relativa y absoluta.
  - **Frecuencia absoluta:** la frecuencia absoluta es el número de veces que se repite un hecho en un experimento o un estudio. Se suele representar de la siguiente forma: n_i.
  - **Frecuencia relativa:** es el resultado de la división entre el valor de la frecuencia absoluta (n_i) y el tamaño de la muestra (N). Se suele representar de esta forma: f_i. Puede aparecer de forma decimal, como fracción o como un porcentaje.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

- ✓ **Distribución de Frecuencia:** una distribución de frecuencia es un resumen tabular de datos que muestra el número (frecuencia) de elementos en cada una de las diferentes clases o intervalos diferentes (que no se sobreponen). Es la representación estructurada, en forma de tabla, de toda la información que se ha recogido sobre la variable que se estudia. La utilidad estriba en que permite comprender los datos, lo que no se logra fácilmente con la simple observación de éstos en su forma desorganizada original.

Ejemplo: clasificar los siguientes sabores de comida:

Salado	Dulce	Acido	Amargo	Dulce	Acido	Amargo	Salado	Dulce	Dulce
Acido	Dulce	Salado	Amargo	Amargo	Salado	Dulce	Dulce	Acido	Salado
Salado	Salado	Acido	Salado	Dulce	Acido	Amargo	Acido	Salado	Dulce

1. Se cuentan los elementos que forman el conjunto de datos: El total da 30

Salado	*Dulce*	Acido	Amargo	*Dulce*	Acido	Amargo	Salado	*Dulce*	*Dulce*
Acido	*Dulce*	Salado	Amargo	Amargo	Salado	*Dulce*	*Dulce*	Acido	Salado
Salado	Salado	Acido	Salado	*Dulce*	Acido	Amargo	Acido	Salado	*Dulce*

2. Se identifica cada uno de los elementos, para el caso de este ejemplo se utilizará un color para cada sabor de comida.

Salado
Dulce
Acido
Amargo

3. Se va a formar una tabla (tabular) con cada uno de los sabores, cada sabor se coloca en un renglón (clase) diferente, por ejemplo:

4. Se agregan dos columnas a la derecha de los sabores y un renglón en la parte superior:

Salado		
Dulce		
Acido		
Amargo		

5. En el primer renglón coloque Frecuencia de clase y Frecuencia relativa como se muestra a continuación:

	Frecuencia de clase	Frecuencia relativa
Salado		
Dulce		
Acido		
Amargo		

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

6. Para llenar la columna de Frecuencia de clase cuente en la primera tabla las veces que aparece la palabra Salado y escriba el resultado. Se debe repetir este proceso para las palabras Dulce, Acido y Amargo.

	Frecuencia de clase	Frecuencia relativa
Salado	9	
Dulce	9	
Acido	7	
Amargo	5	

7. Para llenar la columna de Frecuencia relativa repita el número obtenido en la Frecuencia de clase y divídalo entre el número de elementos como se muestra en la tabla siguiente:

	Frecuencia de clase: n_i.	Frecuencia relativa: $f_i = n_i / N$
Salado	9	9 / 30
Dulce	9	9 / 30
Acido	7	7 / 30
Amargo	5	5 / 30

Ejemplo:

18	19	20	24	26	28	17	19	23	16	21	30	23	27	29	15	18	26

Tamaño de la muestra: N = 18

Ordenar ascendentemente:

15	16	17	18	18	19	19	20	21	23	23	24	26	26	27	28	29	30

Rango: En una muestra o población, el rango es la distancia entre el dato mayor y el dato menor, se calcula restando ambos datos.

Rango = Dato Mayor – Dato Menor

Rango: 30-15=15

Intervalo: se sugiere 5.

Se puede calcular por una formula llamada Regla de Sturges: 1+3.332*logaritmo de N.

En este caso:

1+3.332 * logaritmo de 18
1+3.332 * 1255 = 1+ 4.18 = 5.18

Siempre se aproxima al número impar entero inferior = 5

Amplitud: rango/ intervalo. Amplitud: 15/5 = 3

Al definir los intervalos es necesario verificar si los datos menor y mayor quedan incluidos, en este caso el 15 y el 30 están incluidos. Si el número menor o el mayor no están incluidos los intervalos no están bien definidos y habrá que definirlos de nuevo. Luego se cuenta cuantos datos hay en el intervalo sin incluir los datos que sean iguales al límite superior.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Intervalo 1	15	18	3	3/18 = 0.17 = 17%
Intervalo 2	18	21	5	5/18 = 0.27 = 27%
Intervalo 3	21	24	3	3/18 = 0.17 = 17%
Intervalo 4	24	27	3	3.18 = 0.17 = 17%
Intervalo 5	27	30	4	4.18 = 0.22 = 22%
Total			18	1.00 100%

EJERCICIO 1.

1.1. Con la tabla siguiente efectuar las operaciones siguientes:

18.60	13.80	10.40	15.00	16.00	22.10	16.20	36.10	11.60	7.80
22.60	17.90	25.30	32.80	16.60	13.60	8.50	23.70	14.20	22.90
17.70	26.30	9.20	24.90	17.90	26.50	26.60	16.50	18.10	24.80
16.60	32.30	14.00	11.60	20.00	33.80	15.80	15.20	24.00	16.40
24.10	23.20	17.30	10.50	15.00	20.20	20.20	17.30	16.60	16.90
22.00	23.90	24.00	22.20	21.80	12.20	22.00	9.60	8.00	20.40
17.20	18.30	13.00	10.60	17.20	8.90	16.80	14.20	15.70	8.00
17.70	16.10	17.80	11.60	10.40	13.60	8.40	12.60	8.10	11.60
21.10	20.50	19.80	24.80	9.70	25.10	31.80	24.90	20.00	17.60

Ordenar:

7.80	7.80	8.00	8.10	8.40	8.50	8.50	9.20	9.60	9.70
10.40	10.40	10.40	10.50	11.60	11.60	11.60	11.60	12.20	12.60
13.60	13.60	13.60	13.80	13.80	14.20	14.20	15.00	15.00	15.00
16.00	16.00	16.10	16.20	16.20	16.50	16.60	16.60	16.60	16.90
17.30	17.30	17.60	17.70	17.70	17.80	17.90	17.90	17.90	18.10
18.60	18.60	19.80	20.00	20.20	20.20	20.40	20.50	21.10	21.80
22.00	22.00	22.10	22.10	22.20	22.60	22.60	22.90	22.90	23.20
23.70	23.70	23.90	24.00	24.10	24.80	24.80	24.90	24.90	25.10
25.30	25.30	26.30	26.50	26.60	31.80	32.80	32.80	36.10	36.10

Determinar N: 90

Determinar Rango: 36.10 – 7.80 = 28.3

Intervalo sugerido: 7

Regla de Sturges:
1+3.332 * logaritmo de 90
1+3.332 * 1.954 = 1+ 6.5115 = 7.512

Siempre se aproxima al número impar entero inferior = 7
Intervalos: 7.

Amplitud: 28.3/7 = 4.04 = 4

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Intervalo 1	7.7	11.7	18	18/90 = 0.2 = 20%
Intervalo 2	11.7	15.7	13	
Intervalo 3	15.7	19.7	24	

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Intervalo 4	19.7	23.7	17	
Intervalo 5	23.7	27.7	13	
Intervalo 6	27.7	31.7	0	
Intervalo 7	31.7	35.7	4	
Intervalo 8	35.7	39.7	1	
Total			90	

1.2. Con los datos siguientes: 55, 51, 60, 56, 64, 56, 63, 63, 61, 57, 63, 50, 49, 70, 72, 54, 48, 53, 58, 66, 68, 45, 74, 65, 58, 61, 62, 59, 64, 57, 63, 52, 67

Ordenar:

Determinar N:

Determinar Rango:

Intervalo sugerido:

Regla de Sturges:

Amplitud:

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Total				

1.3. Con los datos siguientes:

5	8	10	14	15	23	25	28	29	30
30	35	36	40	43	44	44	45	45	46
46	46	47	48	48	49	49	49	49	50
50	50	50	51	51	51	52	53	55	57
60	66	70	72	75	75	84	84	88	94

Ordenar:

Determinar N:

Determinar Rango:

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Intervalo sugerido:

Regla de Sturges:

Amplitud:

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Total				

ACTIVIDAD DE EVALUACIÓN 02.

Con los datos siguientes:

8	8	7	5	4	3	4	7	5	9
3	4	7	6	5	7	3	5	4	5
3	9	7	6	8	4	6	9	7	8
3	4	9	5						

Ordenar:

Determinar N:

Determinar Rango:

Intervalo sugerido:

Regla de Sturges:

Amplitud:

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Total				

Con los datos siguientes:

18	25	15	30	22	27	15	18	21	19
28	27	16	19	18	30	17	16	20	15
14	19	23	27	17	14	16	20	21	28

Ordenar:

Determinar N:

Determinar Rango:

Intervalo sugerido:

Regla de Sturges:

Amplitud:

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Total				

Con los datos siguientes:

30	38	36	35	29	28	30	35	40	48
50	20	25	56	30	27	29	46	41	31
31	39	28	36	37	52	44	49	52	56
58	40	39	38	40	27	24	30	32	35
38	26	25	24	60	55	48	37	31	30
22	20	24	26	23	22	28	27	48	31

Ordenar:

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Determinar N:

Determinar Rango:

Intervalo sugerido: 7

Regla de Sturges:

Amplitud:

Frecuencia relativa expresada de manera decimal, como fracción y como un porcentaje.

	Límite inferior	Límite superior	Frecuencias	Frecuencia relativa
Total				

FRONTERAS DE CLASE

La frontera de clase o límite de clase es el límite menor y mayor del intervalo, el límite menor se incluye como parte del intervalo, pero el mayor no forma parte de él.

En el ejemplo siguiente la distribución de frecuencias está formada por 8 intervalos que empiezan en 14 y terminan en 70.

Si los datos de los intervalos son enteros, se busca el número más cercano al límite superior.

PRECIO POR PLATO	RESTAURANTES	Fronteras de clase
14 pero menos de 21	1	14 , 21
21 pero menos de 28	5	21 , 28
28 pero menos de 35	7	28 , 35
35 pero menos de 42	16	35 , 42
42 pero menos de 49	10	42 , 49
49 pero menos de 56	9	49 , 56
56 pero menos de 63	1	56 , 63
63 pero menos de 70	1	63 , 70
Total	50	

Si el intervalo es "De 10 a menos de 20" y los datos de la encuesta tienen dos decimales, entonces al límite superior se le resta 0.01 y se obtiene 19.99 la frontera real superior.

Si el intervalo es "De 10 a menos de 20" y los datos tienen 1 decimal, entonces la frontera superior se le resta 0.1 y se obtiene 19.9 como frontera superior

La frontera inferior siempre es la misma.

En el ejemplo las fronteras reales son:

PRECIO POR PLATO	RESTAURANTES	Fronteras reales de clase
14 pero menos de 21	1	14 , 20.9
21 pero menos de 28	5	21 , 27.9
28 pero menos de 35	7	28 , 34.9
35 pero menos de 42	16	35 , 41.9
42 pero menos de 49	10	42 , 48.9
49 pero menos de 56	9	49 , 55.9
56 pero menos de 63	1	56 , 62.9
63 pero menos de 70	1	63 , 69.9
Total	50	

Marca de clase o punto medio

Es el punto medio de un intervalo de clase, se calcula sumando sus fronteras y dividiendo el resultado entre 2. Es decir, el intervalo es dividido a la mitad.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

$$MC = \frac{FronteraInferior + FronteraSuperior}{2}$$

En el ejemplo anterior la marca de clase se calcula así:

PRECIO POR PLATO	RESTAURANTES	Marca de Clase
14 pero menos de 21	1	$\frac{14 + 21}{2} = 17.50$
21 pero menos de 28	5	$\frac{21 + 28}{2} = 24.50$
28 pero menos de 35	7	$\frac{28 + 35}{2} = 31.50$
35 pero menos de 42	16	$\frac{35 + 42}{2} = 38.50$
42 pero menos de 49	10	$\frac{42 + 49}{2} = 45.50$
49 pero menos de 56	9	$\frac{49 + 56}{2} = 52.50$
56 pero menos de 63	1	$\frac{56 + 63}{2} = 59.50$
63 pero menos de 70	1	$\frac{63 + 70}{2} = 66.50$
Total	50	

Frecuencia acumulada y frecuencia relativa acumulada: es la suma parcial de cada intervalo, permite hacer observaciones sobre grupos de intervalos que están por debajo de él. Para calcular la frecuencia acumulada se puede seguir el siguiente proceso:

- ✓ Se suman todas las frecuencias para obtener el tamaño de la muestra.
- ✓ El primer intervalo acumulado es igual a la frecuencia del primer intervalo
- ✓ Se suma la frecuencia de un intervalo con todas las frecuencias de los intervalos anteriores.
- ✓ La frecuencia acumulada de cada intervalo nunca es menor que el valor del intervalo anterior.
- ✓ El último intervalo debe ser igual al tamaño de la muestra (la suma de todas las frecuencias).

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
X1	n1	n1	f1 = n1 / n	f1
X2	n2	n1 + n2	f2 = n2 / n	f1 + f2
Xn-1	nn-1	n1 + n2 +..+ nn-1	fn-1 = nn-1 / n	f1 + f2 +..+fn-1
Xn	nn	S n	fn = nn / n	S f

PRECIO POR PLATO	RESTAURANTES	Frecuencia acumulada
14 pero menos de 21	1	1 = 1
21 pero menos de 28	5	1 + 5 = 6
28 pero menos de 35	7	1 + 5 + 7 = 13
35 pero menos de 42	16	1 + 5 + 7 +16 = 29
42 pero menos de 49	10	1 + 5 + 7 +16 + 10 = 39
49 pero menos de 56	9	1 + 5 + 7 +16 + 10 + 9 = 48
56 pero menos de 63	1	1 + 5 + 7 +16 + 10 + 9 + 1 = 49
63 pero menos de 70	1	1 + 5 + 7 +16 + 10 + 9 + 1 + 1 = 50
Total	50	

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Ejemplo: se han calculado las cuatro primeras líneas, es necesario concluir la tabla:

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
1.20	1	1	1/30 = 0.03 = 3.33%	3.3%
1.21	4	5	4/30 = 0.133 = 13.3%	16.6%
1.22	4	9	4/30 = 0.133 = 13.3%	30.0%
1.23	2	11	2/30 = 0.066 = 6.6%	36.6%
1.24	1			
1.25	2			
1.26	3			
1.27	3			
1.28	4			
1.29	3			
1.30	3			
Total	30			

EJERCICIO 2. **Con la tabla siguiente efectuar las operaciones siguientes:**

216	269	235	235	224	249	244	240	252	210
256	228	233	239	253	237	221	273	229	236
215	260	230	228	230	255	245	244	230	231
240	222	223							

Ordenar:

Determinar N:

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

				Frecuencias		Frecuencias acumulada	
	Límite inferior	Límite superior	MC_i	X_i	f_i	N_i	F_i
Intervalo 1							
Intervalo 2							
Intervalo 3							
Intervalo 4							
Intervalo 5							
Intervalo 6							
Intervalo 7							
Intervalo 8							
Total							

Con la tabla siguiente efectuar las operaciones siguientes:

40	40	44	48	50	50	53	54	58	60
60	60	61	61	64	68	69	70	74	77
78	79	80	80	85	89	90	91	91	92

Determinar N:

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

Con la tabla siguiente efectuar las operaciones siguientes:

5	40	45	35	20	15	50	45	25	15
20	55	35	45	15	10	5	50	55	60
55	35	45	65	60	45	15	5	5	25
30	45	50	15	10	5	10	5	5	50
55	15	10	50	20	25	35	45	55	60

Ordenar:

Determinar N:

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

ACTIVIDAD DE EVALUACIÓN 02

Con la tabla siguiente efectuar las operaciones siguientes:

5	40	45	35	20	15	50	45	25	15
20	55	35	45	15	10	5	50	55	60
55	35	45	65	60	45	15	5	5	25
30	45	50	15	10	5	10	5	5	50
55	15	10	50	20	25	35	45	55	60

Ordenar:

Determinar N:
Determinar Rango:
Intervalo por la Regla de Sturges:
Amplitud:
Frecuencia absoluta = X_i
Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.
Frecuencia acumulada: absoluta N_i y relativa F_i
Marca de clase = MC_i

				Frecuencias		Frecuencias acumulada	
	Límite inferior	Límite superior	MC_i	X_i	f_i	N_i	F_i
Intervalo 1							
Intervalo 2							
Intervalo 3							
Intervalo 4							
Intervalo 5							
Intervalo 6							
Intervalo 7							
Intervalo 8							
Total							

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Con la tabla siguiente efectuar las operaciones siguientes:

14	21	23	21	16	19	22	25	16	16
24	24	25	19	16	19	18	19	21	12
16	17	18	23	25	20	23	16	20	19
24	26	15	22	24	20	22	24	22	20

Ordenar:

Determinar N:

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

				Frecuencias		Frecuencias acumulada	
	Límite inferior	Límite superior	MC_i	X_i	f_i	N_i	F_i
Intervalo 1							
Intervalo 2							
Intervalo 3							
Intervalo 4							
Intervalo 5							
Intervalo 6							
Intervalo 7							
Intervalo 8							
Total							

El personal de un consultorio analiza los tiempos de espera de los pacientes que requieren servicio de emergencia. Los datos siguientes son los tiempos de espera en minutos recolectados a lo largo de un mes.

2	5	10	12	4	4	5	17	11	8
9	8	12	21	6	8	7	13	18	3

Determinar N:

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

				Frecuencias		Frecuencias acumulada	
	Límite inferior	Límite superior	MC_i	X_i	f_i	N_i	F_i
Total							

¿Cuál es la proporción de los pacientes que requieren servicio de emergencia y esperan 9 minutos o menos?

Las velocidades de N vehículos fueron medidas por la Policía Municipal de Transito y generaron la tabla siguiente:

27	23	22	38	43	24	35	26	28	20
18	25	23	22	63	31	30	41	45	29
43	27	29	28	27	25	29	28	24	37
28	18	29	26	33	25	27	25	34	32
36	22	33	32	21	23	24	18	48	23
15	38	26	23	21					

El policía que tomo las velocidades tomo las velocidades en millas por lo cual es necesario convertirlas a kilómetros para determinar cuántos violaron el límite de velocidad de 80 kilómetros por hora. Una milla es equivalente a 1.60934 kilómetros. ¿Cuántos violaron el límite de los 80 kilómetros y cuál es el porcentaje? ¿Cuál es el rango de velocidad más utilizado?

Ordenar:

Determinar N:

Determinar Rango:

Intervalo por la Regla de Sturges:

Amplitud:

Frecuencia absoluta = X_i

Frecuencia relativa f_i expresada de manera decimal, como fracción y como un porcentaje.

Frecuencia acumulada: absoluta N_i y relativa F_i

Marca de clase = MC_i

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

				Frecuencias		Frecuencias acumulada	
	Límite inferior	Límite superior	MC_i	X_i	f_i	N_i	F_i
Total							

Las siguientes datos son de los estudiantes de una clase en la que un observador, durante el tiempo que ha durado una sesión de clase de 2 horas, ha anotado el número de veces que ha participado cada uno de los estudiantes dirigiéndose a todo el grupo en voz alta.

¿Cuántos estudiantes son en total?

¿Qué porcentaje y cantidad de estudiantes participaron en menos de 2 ocasiones en la sesión de clase?

¿Qué porcentaje y cantidad de estudiantes participaron 5 veces? ¿Cuántos son?

2	2	3	0	3	1	8	0	3	9
1	1	0	4	0	2	9	5	0	1
9	8								

¿Qué porcentaje y cantidad de estudiantes participaron 4 veces o menos?

¿Qué porcentaje y cantidad de estudiantes participaron más de 4 veces?

¿Qué proporción de estudiantes participaron al menos una vez?

¿Qué porcentaje y cantidad de estudiantes participaron entre 2 y 5 veces, ambas inclusive?

¿Qué porcentaje y cantidad de estudiantes participaron 8 o 9 veces?

REPRESENTACIÓN GRÁFICA DE DATOS

"Una imagen dice más que mil palabras", esta frase explica la importancia de presentar los datos en forma gráfica.

Un gráfico o representación gráfica es un tipo de representación de datos, generalmente numéricos, mediante recursos gráficos, para que se manifieste visualmente la relación matemática o correlación estadística que guardan entre sí.

La representación gráfica de datos es un procedimiento técnico para la facilitación de interpretación de resultados.

En estadística denominamos gráficos a aquellas imágenes que, combinando la utilización De sombreado, colores, puntos, líneas, símbolos, números, texto y un sistema de referencia (coordenadas), permiten presentar información cuantitativa.

La utilidad de los gráficos es doble, ya que pueden servir no sólo como sustituto a las tablas, sino que también constituyen por sí mismos una poderosa herramienta para el análisis de los datos, siendo en ocasiones el medio más efectivo no sólo para describir y resumir la información, sino también para analizarla.

Un gráfico estadístico es una representación visual de una serie de datos estadísticos.

Es una herramienta muy eficaz y debe tener las características siguientes:

- ✓ Captar la atención del lector.
- ✓ Debe transmitir en forma inmediata una idea general: de fácil interpretación.
- ✓ Presentar la información de forma sencilla, clara y precisa.
- ✓ No debe inducir al error.
- ✓ Facilitar la comparación de datos y destacar las tendencias y las diferencias.
- ✓ No debe contener más líneas o símbolos que los que el ojo pueda seguir cómodamente.
- ✓ Todo gráfico debe explicarse por sí mismo; por eso deben colocarse claramente el título, escalas, origen y leyendas.

Las gráficas se construyen a partir de los datos siguientes:

Cualitativos: se refieren a cualidades o modalidades que no pueden expresarse numéricamente. Pueden ser:

- ✓ **Ordinales:** si siguen un orden o secuencia. Ejemplo: el abecedario, los meses del año…
- ✓ **Categóricos:** si no siguen ningún orden. Ejemplo: el estado civil de las personas: solteros, casados, viudos, divorciados y separados, sexo…

Cuantitativos: se refieren a cantidades o valores numéricos. Pueden ser:

- ✓ **Discretos:** si toman valores enteros (0, 1, 2, 3...). Ejemplo: el número de hijos, el número de alumnos de una clase…
- ✓ **Continuos:** si pueden tomar cualquier valor dentro de un intervalo. Ejemplo: la estatura o el peso de las personas.

Existe una gran variedad de gráficos y la selección apropiada de algunos de ellos para la representación de la información dependerá, entre otras cosas, del tipo de datos, y la preferencia e interés del que presenta la gráfica.

GRAFICA DE BARRAS

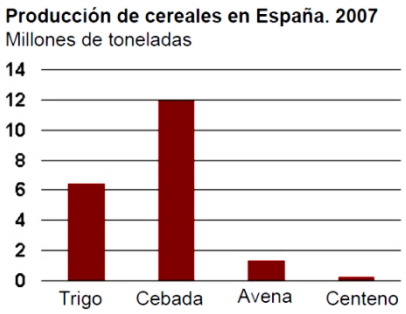
Se recomienda para representar series cronológicas, datos cualitativos (ordinales o nominales) y en general para datos donde exista algún orden.

En uno de los ejes se posicionan las distintas categorías o modalidades de la variable cualitativa o discreta (en el ejemplo, el tipo de cereal) y en el otro el valor o frecuencia de cada categoría en una determinada escala (en el ejemplo, la producción en millones de toneladas de granos).

Producción Agrícola 2007

Cereales	Miles de toneladas
Cebada	11.945
Trigo	6.436
Avena	4.310
Centeno	261

Fuente: Ministerio de Agricultura, Alimentación y Medio Ambiente



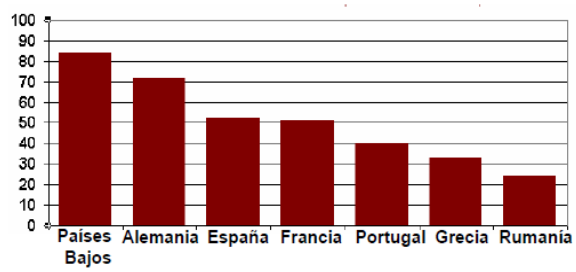
La orientación del gráfico puede ser:

- ✓ **Vertical:** las distintas categorías están situadas en el eje horizontal y las barras de frecuencias crecen verticalmente.
- ✓ **Horizontal:** las categorías se sitúan en el eje vertical y las barras crecen horizontalmente. Suelen usarse cuando hay muchas categorías o sus nombres son demasiado largos.

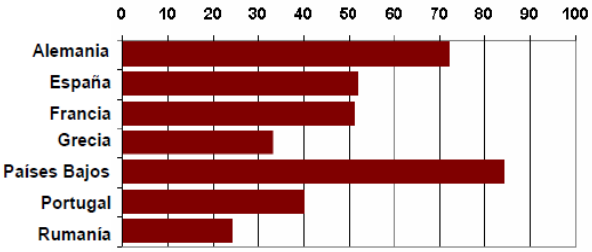
Las categorías pueden ordenarse alfabéticamente facilitando su búsqueda o por sus frecuencias facilitando la comparación de los datos. Veamos el siguiente ejemplo del porcentaje habitantes usuarios de internet del año 2007 por países.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Orientación vertical y orden por frecuencias

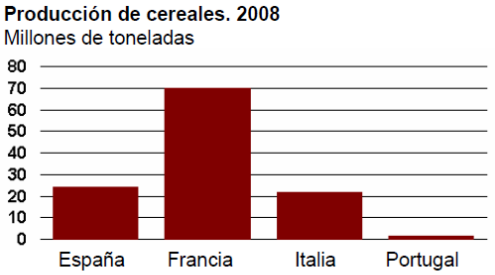


Orientación horizontal y orden alfabético

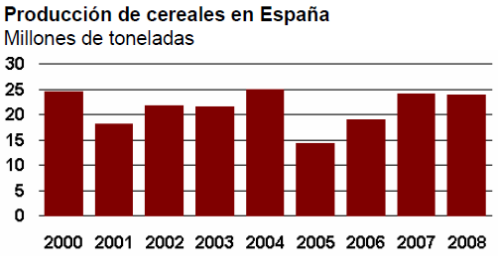


Se suelen usar para:

- ✓ Ver la evolución en el tiempo de una magnitud concreta.

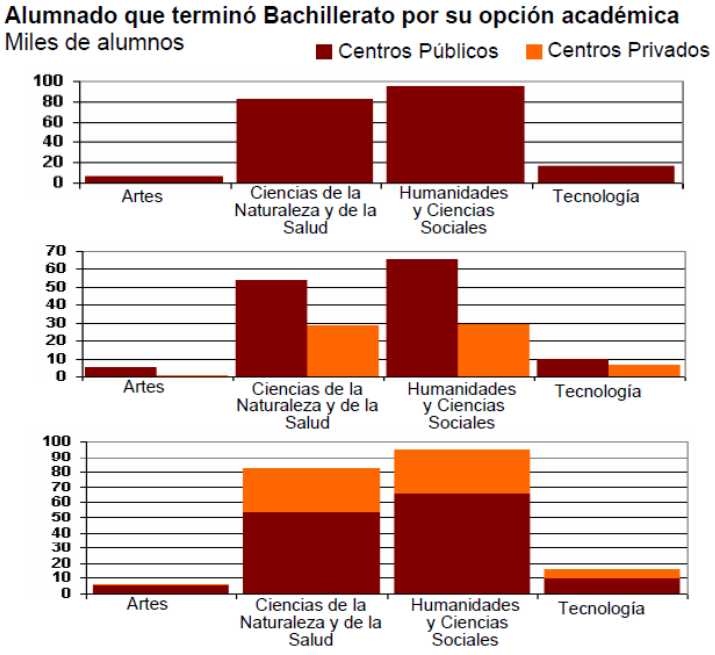


- ✓ Comparar magnitudes de varias categorías.



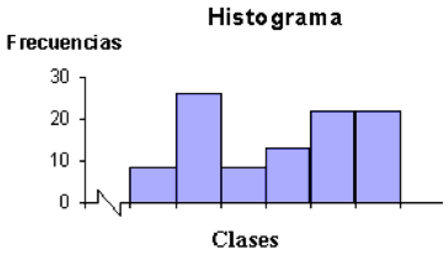
Tipos de gráficos de barras:

- ✓ Sencillo: Contiene una única serie de datos.
- ✓ Agrupado: Contiene varias series de datos y cada una se representa por un tipo de barra de un mismo color o textura.
- ✓ Apilado: Contiene varias series de datos. La barra se divide en segmentos de diferentes colores o texturas y cada uno de ellos representa una serie.



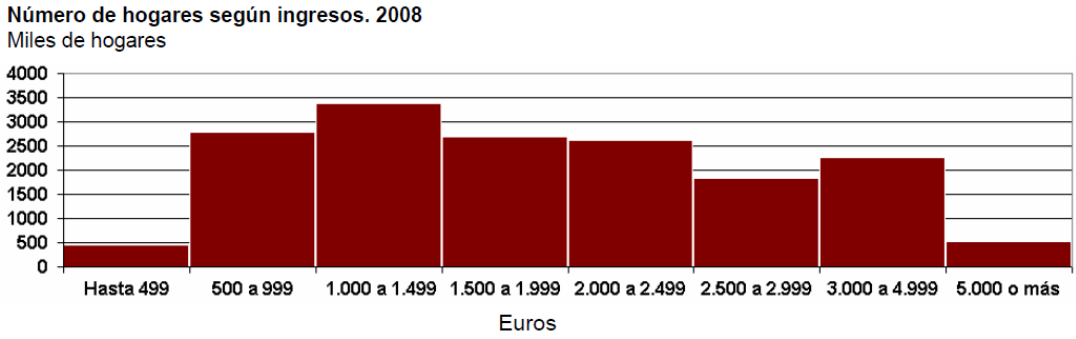
HISTOGRAMA

Se usa para representar las frecuencias de una variable cuantitativa continua. El histograma es el gráfico adecuado para ilustrar el comportamiento de los datos agrupados en intervalos de clases, es un gráfico de barras compuesto por varios rectángulos adyacentes, que



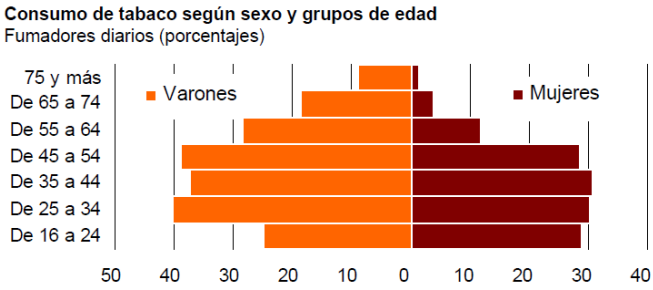
¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

representan a la tabla de distribución de frecuencias de cierta variable cuantitativa. En el eje horizontal se marcan los intervalos (límites de clases), y cada intervalo es la base de cada rectángulo; en el eje vertical se marcan las alturas de los rectángulos la cual viene dada por las frecuencias absolutas simples o relativas. No existe separación entre las barras.



GRAFICA BIDIRECCIONAL

Tiene orientación horizontal y contiene dos series de datos cuyas barras de frecuencias crecen en sentidos opuestos.



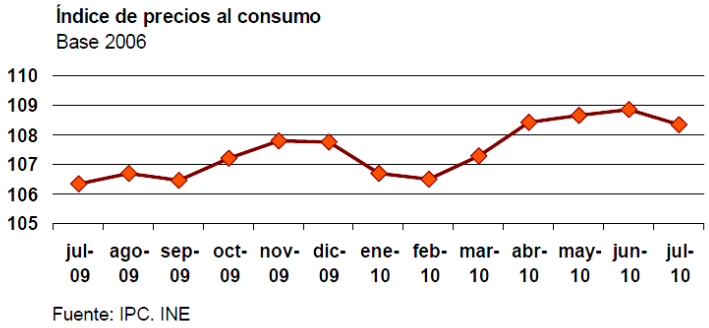
GRAFICA DE LÍNEAS O DIAGRAMA DE LÍNEAS

Un gráfico de líneas es una representación gráfica en un eje cartesiano de la relación que existe entre dos variables reflejando con claridad los cambios producidos.

En cada eje se representa cada una de las variables cuya relación se quiere observar (en el ejemplo, en el eje horizontal los meses y en el eje vertical la media mensual del IPC correspondiente a esos meses).

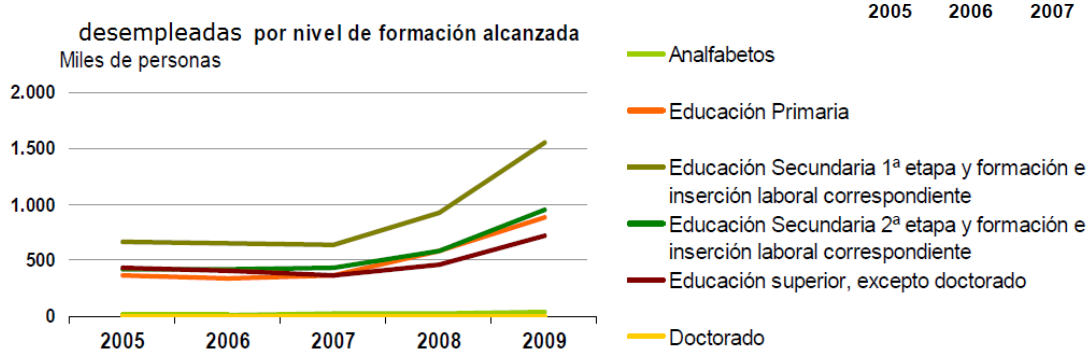
IPC 2009 - 2010
Base 2006

	IPC
jul-09	106,3
ago-09	106,7
sep-09	106,4
oct-09	107,2
nov-09	107,8
dic-09	107,8
ene-10	106,7
feb-10	106,5
mar-10	107,3
abr-10	108,4
may-10	108,7
jun-10	108,9
jul-10	108,4



Fuente: IPC. INE

Se suelen usar para presentar tendencias temporales. En el eje horizontal se ha de posicionar la variable que indica las unidades de tiempo y en el vertical se introduce la escala de la variable cuya variación en el tiempo queremos ver. Pueden aparecer varias variables para compararlas.

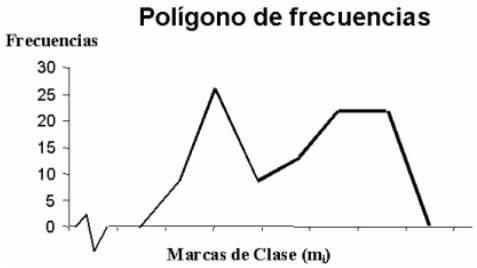


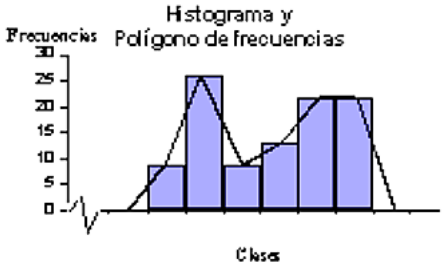
¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

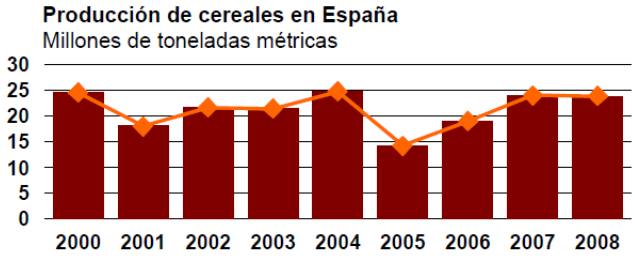
POLÍGONO DE FRECUENCIAS

Si se unen los puntos medios de las bases superiores de las barras en los gráficos de barra se obtiene el polígono de frecuencias.

En el eje horizontal se marcan las marcas de clase de cada intervalo y para cada una de estas marcas se marcan las alturas en el eje vertical, las cuales vienen dadas por las frecuencias respectivas (absolutas simple o relativas).







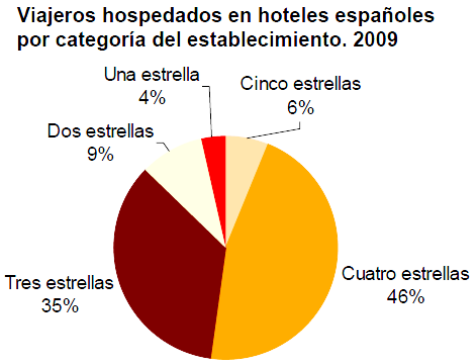
GRAFICA O DIAGRAMA DE SECTORES

Llamado también grafico de torta, es el más indicado para datos cualitativos nominales y ordinales. Es una representación circular de las frecuencias relativas de una variable cualitativa o discreta que permite, de una manera sencilla y rápida, su comparación. Consiste en dividir el círculo en tantos sectores como categorías tenga la variable y donde a cada sector le corresponde un área proporcional a la frecuencia absoluta o relativa asociada con la modalidad que representa. Son útiles cuando las categorías son pocas.

El círculo representa la totalidad que se quiere observar (en el ejemplo, total de viajeros hospedados en hoteles) y cada porción, llamadas sectores, representan la proporción de cada categoría de la variable (en el ejemplo, tipo de hotel) respecto el total. Suele expresarse en porcentajes.

Alojamientos Turísticos. 2009

Categoría	Número de viajeros
Total	69.152.754
Cinco estrellas	4.216.253
Cuatro estrellas	31.960.442
Tres estrellas	24.079.125
Dos estrellas	6.331.715
Una estrella	2.565.219



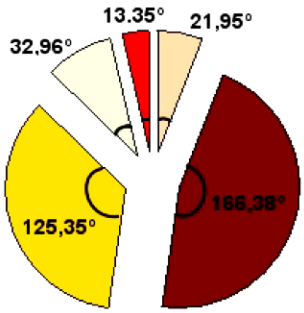
Obtención de los **ángulos** de cada sector:

Categoría	Frecuencia absoluta	Frecuencia relativa	Porcentaje (%)
Cinco estrellas	4.216.253	0,06	6
Cuatro estrellas	31.960.442	0,46	46
Tres estrellas	24.079.125	0,35	35
Dos estrellas	6.331.715	0,09	9
Una estrella	2.565.219	0,04	4
Total	69.152.754	1,00	100

$$\text{Frecuencia relativa} = \frac{\text{Frecuencia absoluta}}{\text{Suma de todas las frecuencias absolutas}}$$

$$\text{Porcentaje} = \text{Frecuencia relativa} \times 100$$

$$\text{Ángulo} = \text{Frecuencia relativa} \times 360 \text{ (grados que tiene una circunferencia)}$$



¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

PICTOGRAMA

Se usan para hacer más llamativa la representación.

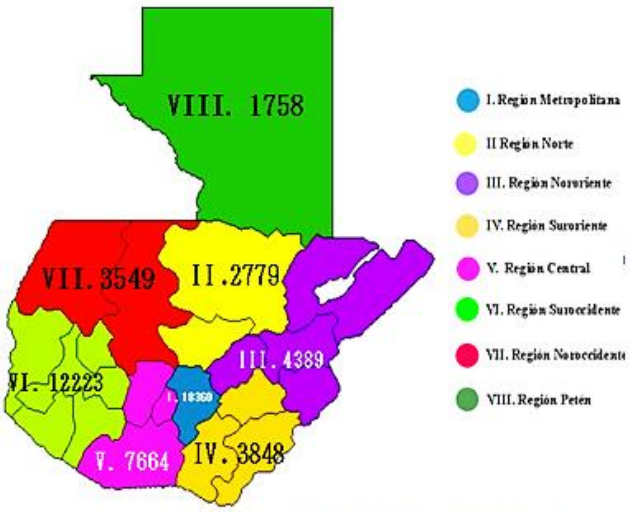
Aceites de cocina	Medicamentos	Envases de plástico, metálicos	Pilas y pequeñas baterías	Papel y cartón	Vidrio
24	69,5	71,8	72,3	74,5	75,3

Es un gráfico de barras horizontal, para graficar las frecuencias, se usan dibujos, símbolos o figuras alusivos al tema de estudio. Cada dibujo representa un número o cantidad específica de unidades, por lo tanto, debe repetirse tantas veces como sea necesario para reflejar una magnitud determinada.

CARTOGRAMA

Un cartograma es un mapa en el que se presentan datos estadísticos por regiones bien poniendo el número o coloreando las distintas zonas en función del dato que representan.

Mapa regional con el número de denuncias de violencia intrafamiliar por región durante el año 2011.

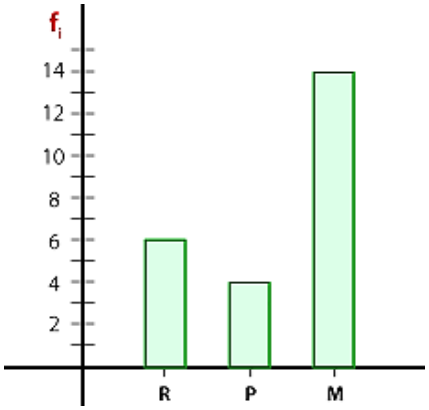


Fuente: Área de Transparencia GAM, Organismo Judicial.

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)

Ejemplos:

- ✓ El siguiente diagrama de barras indica el color de pelo de los alumnos de la clase de Mario. Completa la tabla con las frecuencias absolutas correspondientes a cada color y responde las siguientes preguntas:

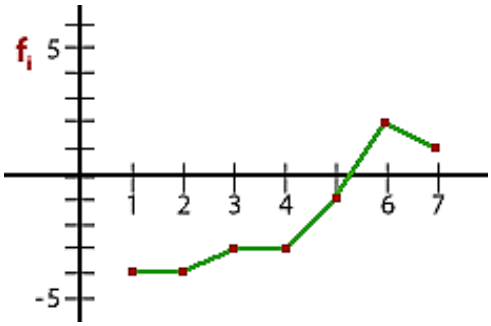


¿Cuántos estudiantes hay en total? 24
¿Qué tipo de pelo predomina en la clase? Moreno
¿Cuántos estudiantes de color de pelo rubio hay? 6
¿Cuántos estudiantes de color de pelo pelirrojo hay? 4
¿Cuántos estudiantes de color de pelo moreno hay? 14

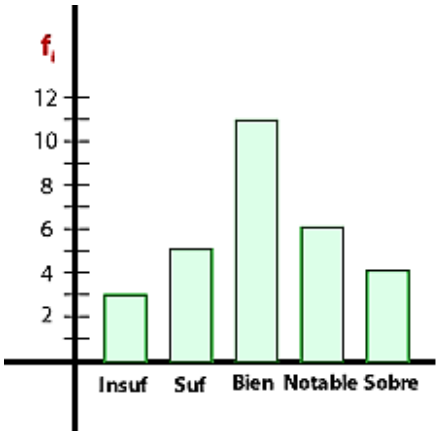
- ✓ El siguiente polígono de frecuencia muestra el promedio de temperatura diaria en una ciudad a lo largo los siete día de una semana.

Indicar la temperatura de cada día:

Día	Temperatura
1	-4
2	-4
3	-3
4	-3
5	-1
6	+2
7	+1



- ✓ El siguiente diagrama de barras muestra las notas de los alumnos de una clase de una clase.



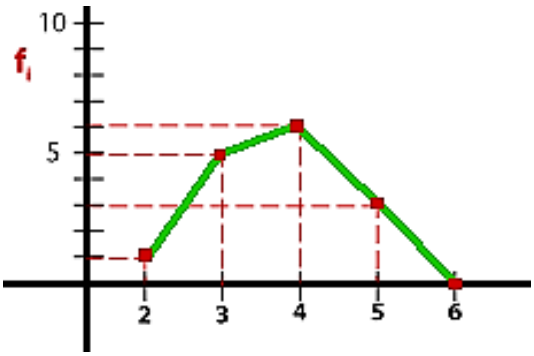
Cuantos alumnos hay con las notas:

Nota	Cantidad
Insuficiente	3
Suficiente	5
Bien	11
Notable	6
Sobresaliente	4

¿Cuál es la nota más común?

¿Cuántos estudiantes hay?

- ✓ Los siguientes valores indican el número de comidas al día que hace un grupo de quince amigos: 3, 4, 2, 3, 4, 3, 4, 5, 4, 5, 3, 4, 5, 3, 4.



Cuantas personas toman el número de comidas indicado:

Número de comidas	Cantidad
2	1
3	5
4	6
5	3
6	0

¡Den gracias al SEÑOR porque él es bueno; su gran amor perdura para siempre! Salmos 107:1 (NVI)