

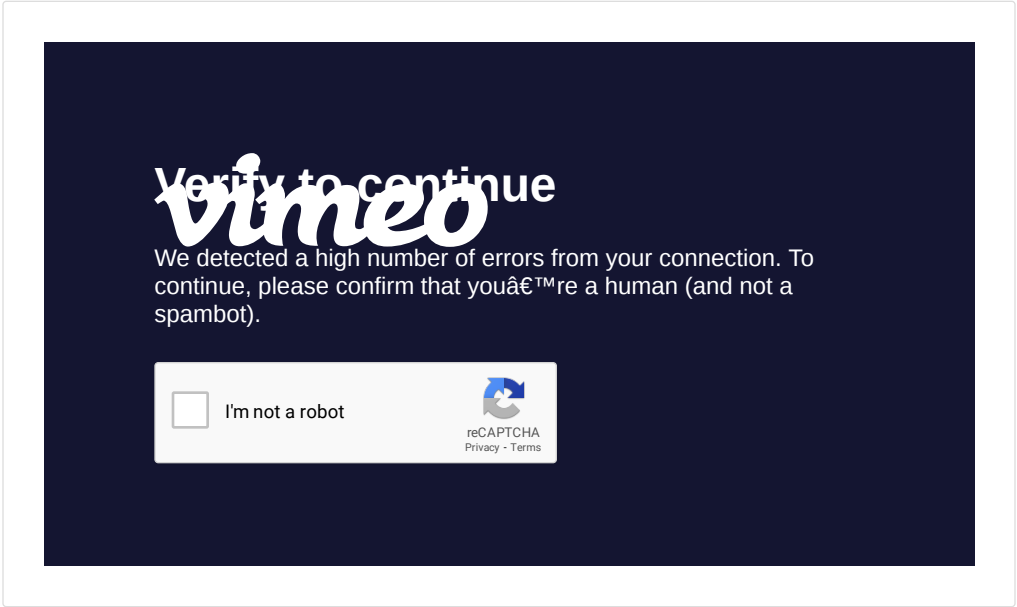
Módulo 4. Estadística



¿Tiene valor la información? ¿Da lo mismo contar o no con ella a la hora de tomar decisiones? A continuación, desarrollaremos sobre cómo recoger datos, ordenarlos, analizarlos y así obtener conclusiones. Sí, efectivamente conoceremos todo sobre la Estadística: los conceptos fundamentales y las distintas variables a tener en cuenta. Además, aprenderemos a interpretar y confeccionar distintos tipos de gráficos. Para ello, nos detendremos en los parámetros estadísticos moda, media y mediana.

- ≡ Video de inmersión
- ≡ Unidad 4.1 Métodos de recolección y resumen de datos
- ≡ Unidad 4.2 Estadística descriptiva
- ≡ Video de habilidades
- ≡ Cierre
- ≡ Referencias
- ≡ Descargá la lectura

Video de inmersión



Unidad 4.1 Métodos de recolección y resumen de datos

4.1.1. Población, elementos y variables

Las grandes masas de datos y números que nos proporcionan información, a veces pueden marear y aturdir. Necesitamos digerirlas y procesarlas de una manera accesible y de eso se ocupa la Estadística.

El uso de métodos estadísticos en diversas áreas implica la **recolección de información**. Estos métodos se utilizan con el objetivo de orientar acerca de qué cambios se deben realizar en el proceso para mejorar su calidad.

Un ejemplo podría ser

un estudio biomédico de un nuevo fármaco que reduce la hipertensión, 85% de los pacientes experimentaron alivio; aunque por lo general se reconoce que el medicamento anterior alivia a 80% de los pacientes que sufren hipertensión crónica. Sin embargo, el nuevo fármaco es más caro de elaborar y podría tener algunos efectos colaterales. ¿Se debería adoptar el nuevo medicamento? (Walpole, Myers y Myers, 2012, <https://aprenderly.com/doc/1101567/probabilidad-y-estadistica-para-ingenieria-y-ciencias?page=687>)

Los investigadores de la estadística han originado un gran número de métodos que permiten efectuar análisis de datos obtenidos, lo cual refleja la verdadera naturaleza de la ciencia que conocemos como **estadística inferencial** que permite **obtener conclusiones** (o inferencias) más allá de solo reportar datos (Walpole, Myers y Myers, 2012, <https://aprenderly.com/doc/1101567/probabilidad-y-estadistica-para-ingenieria-y-ciencias?page=687>).

En ocasiones solo se desea obtener alguna clase de resumen de un conjunto de datos representados en la muestra. En otras palabras, no se requiere estadística inferencial sino **estadística descriptiva**.

¿Para qué se emplea la estadística? Veamos algunos ejemplos:

1

Se desea tallar y pesar a los 900 niños de una escuela.

2

Se desea conocer la opinión de todas las mujeres argentinas, de edades entre 25 y 35 años, con respecto a un producto de belleza femenina.

3

Un fabricante de tarjetas para computadora desea eliminar defectos por un fallo en la fabricación.

En el primer caso, es muy fácil tallar y pesar a 900 niños. Los resultados obtenidos se tabularán y, si se quiere, se podrán presentar en un gráfico. En el segundo caso, resulta imposible conocer la opinión de millones de mujeres argentinas entre 25 y 35 años, no quedará más remedio que tomar un pequeño

conjunto de ellas y realizar la encuesta. En el último caso, las características que el fabricante de tarjetas de computadora desea eliminar podría ser un proceso muy largo si desea chequear una a una miles de tarjetas fabricadas. Un proceso de muestreo implicaría recolectar información de, por ejemplo, solo 50 tarjetas de computadora tomadas aleatoriamente durante el proceso y extender las conclusiones a la totalidad de tarjetas fabricadas.

Al conjunto de todos los **elementos** sobre los que se hace un estudio estadístico se lo denomina **población**.

A continuación, pensemos en los tres estudios estadísticos que analizamos anteriormente.

Tabla 1

Estudio estadístico	Población	Elementos	¿Requerirá analizar a toda la población?
1	900 niños de la escuela.	Cada uno de los niños de la escuela.	No
2	Todas las mujeres argentinas de edades entre 25 y 35 años.	Cada una de las mujeres argentinas de edades entre 25 y 35 años.	Sí
3	Todas las tarjetas de computadoras fabricadas durante cierto período.	Cada tarjeta de computadora fabricada durante cierto período.	Sí

Fuente: elaboración propia.

Las características o propiedades que posee un determinado grupo de personas o elementos que queremos estudiar se denominan **variables estadísticas**. Existen variables de dos tipos:

- **Variables cuantitativas:** son las que se expresan mediante números. Estas pueden ser:
 - **Discretas:** si se pueden contar y se expresan con un número entero. Por ejemplo, cantidad de miembros de una familia, calificación numérica exacta de un examen, cantidad de productos con falla, número de pacientes infectados, etc.
 - **Continuas:** si se pueden medir y se expresan con un número real (peso o altura de una persona, precio de un producto, temperatura, etc.).
- **Variables cualitativas:** indican características o cualidades. Son las que no se pueden expresar mediante números (sexo, nacionalidad, nivel socioeconómico, estado civil, profesión, nivel de estudios, etc.).

Para obtener los datos necesarios de un estudio estadístico hace falta realizar una **encuesta**. Es muy importante que estas sean correctamente elaboradas.

4.1.2 Muestreo

Una **muestra** es un subconjunto de una población. El objetivo de tomar una muestra es reducir la cantidad de experiencias. Las propiedades que se obtengan se harán extensivas a toda la población. Una vez conocida la población, existen diferentes maneras de recoger una muestra, a esto se refiere el método de **muestreo**. En términos generales existen tres tipos de muestreos:

- 1

Muestreo probabilístico: todos los elementos de la población tienen la misma chance de integrar la muestra durante el proceso de selección.
- 2

Muestreo intencional: el propio investigador selecciona los elementos que mayor información puedan dar sobre la población en estudio.
- 3

Muestreo accidental: en este caso, el investigador incluye en su muestra los elementos de análisis que dispone en un momento dado.

Pensemos en algunos estudios estadísticos. Trataremos de determinar cuál es la población, sus elementos, si es necesario recoger una muestra, y si las variables son cuantitativas (discretas o continuas) o cualitativas.

Tabla 2

Estudio estadístico	Población	Elementos	¿Es necesario recoger una muestra?	V a r i a b l e	Tipo de variable	
					Cuantitativa (discreta o continua)	Cualitativa
Colores elegidos por los usuarios de automóviles en la ciudad de Córdoba.	Automóviles de la ciudad de Córdoba.	Cada automóvil de la ciudad de Córdoba.	Sí	C o l o r	-	•
Distancia del hogar a la planta de los empleados de una fábrica.	Empleados de la fábrica.	Cada empleado de la fábrica.	No	D i s t a	Continúa (las distancias recogidas serán valores reales, tales como 12,5 km, 500,5 m, etc.).	-

				n c i a s		
Marca de gaseosa preferida por los adolescentes argentinos.	Adolescentes argentinos.	Cada adolescente argentino.	Sí	M a r c a	-	•
Peso de los bebés al nacer en una maternidad.	Bebés recién nacidos de una maternidad.	Cada bebé recién nacido de una maternidad.	No	P e s o	Continúa (los pesos recogidos serán valores reales tales como 2,525 kg o 3,530 kg, etc.).	-
Número de personas que viven en los hogares de un barrio.	Hogares de un barrio.	Cada hogar de un barrio.	No	C a n t i d a d	Discreta (las cantidades obtenidas solo pueden ser números enteros positivos: 1, 2, 3, 4...).	-

Fuente: elaboración propia.

4.1.3. Distribución de frecuencias

Las notas obtenidas en un examen de un curso 2º año de un centro educativo fueron:

6	2	5	7	3	8	10	6
6	7	3	2	9	5	6	7
5	5	4	10	7	8	7	8
3	6	4	5	4	6	6	9
6	5	4	9	5	9	5	6

¿Cómo presentar estos datos para que su análisis resulte cómodo? La respuesta es sencilla. Presentaremos los datos ordenados y agrupados en una distribución de frecuencias. Las distribuciones de frecuencias son tablas que facilitan la lectura de los datos. A estas tablas las llamamos tablas de frecuencias.

Tabla 3

Notas	Frecuencia
1	0
2	2
3	3
4	4
5	8
6	9
7	5
8	3
9	4
10	2

Fuente: elaboración propia.

Esta es la tabla de frecuencias de las notas del examen. Leemos en la tabla que 2 estudiantes sacaron un 10, que 5 estudiantes sacaron un 7, etc., y decimos que la frecuencia de la nota 10 es 2, la de 7 es 5, etc.

Si nos interesa conocer la proporción de estudiantes que han obtenido una nota, bastará dividir la frecuencia de esa nota entre el total de estudiantes. Esa proporción, se llama frecuencia relativa. Si a cada uno de estos valores los multiplicamos por 100, obtendremos sus respectivos porcentajes.

Para distinguir la frecuencia relativa de la frecuencia, esta recibe también el nombre de frecuencia absoluta.

Tabla 4

Notas	Frecuencia absoluta	Frecuencia relativa	Porcentajes
-------	---------------------	---------------------	-------------

1	0	0	0
2	2	0,05	5
3	3	0,08	8
4	4	0,10	10
5	8	0,20	20
6	9	0,23	23
7	5	0.13	13
8	3	0,08	8
9	4	0,10	10
10	2	0,05	5
Total	40	1	100

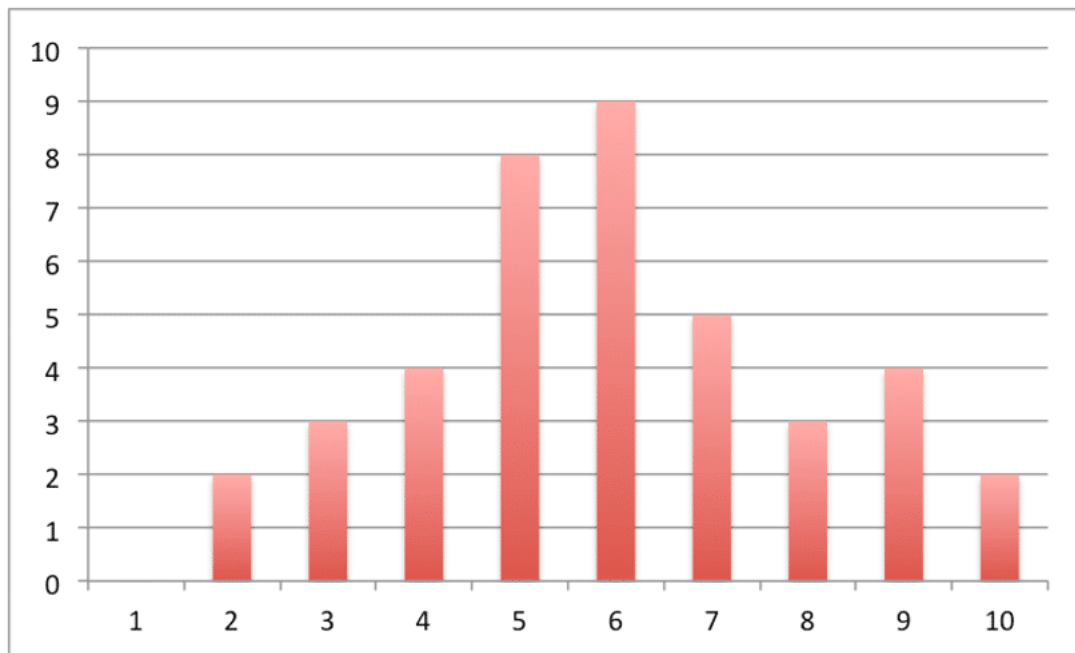
Fuente: elaboración propia.

4.1.4 Gráficos estadísticos

Una **gráfica** es la representación visual de los datos obtenidos en un trabajo estadístico. La finalidad de los gráficos estadísticos es que la información "entre por los ojos". Existen gráficos de distintos tipos, pero todos ellos resultan muy fáciles de interpretar, pues esa es, precisamente, su razón de ser. Los distintos tipos de gráficas se usan unas u otras, según el estudio que se realice. Las más importantes son: **diagrama de barras**, **histogramas**, **diagrama de sectores**, etc.

Representemos los datos del examen gráficamente por medio de un diagrama de barras.

Figura 1. Diagrama de barras



Fuente: elaboración propia.

La longitud de las barras es proporcional a las frecuencias. La tabla de frecuencias o el diagrama de barras permite analizar mejor una gran cantidad de datos, ya que estos se presentan ordenados y clasificados. El diagrama de barras se suele usar cuando la variable es discreta, como en este caso.

Histograma

Si quisiéramos hacer una tabla de frecuencia con los datos de las estaturas de los estudiantes de una clase, nos damos cuenta que sería inadecuado hacer una clasificación del estilo de la tabla anterior.

148	171	169	153	168	158	162	168
177	173	173	176	155	159	157	158
162	149	174	154	175	160	161	173
158	175	177	175	176	161	160	168
173	174	177	174	174	175	174	161

Como las estaturas son muy distintas, las agrupamos en intervalos.

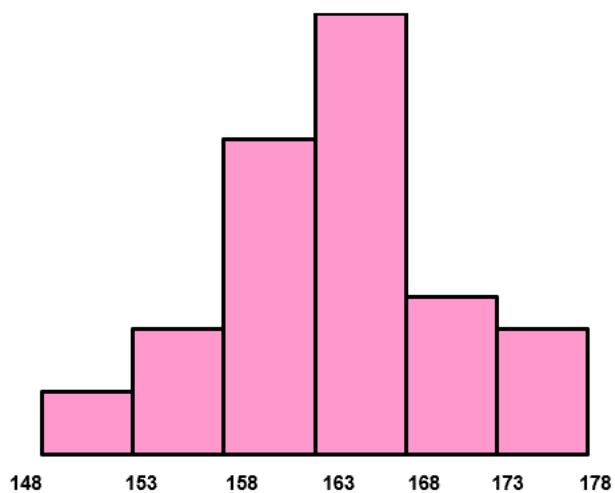
Tabla 5

Intervalo	Frecuencias
148 - 152	2

153 - 157	4
158 - 162	11
163 - 167	14
168 -172	5
173 - 177	4

Fuente: elaboración propia.

Figura 2



Fuente: elaboración propia.

Este diagrama se llama **histograma** y se suele usar cuando se agrupan los datos por intervalos. En cada uno de los intervalos se levanta una franja ancha como el intervalo y de la altura proporcional a la frecuencia.

Diagrama de sectores

Si quisiéramos clasificar las notas del examen del primer ejemplo en dos categorías, esto es, estudiantes aprobados y estudiantes desaprobados, al ser estas pocas categorías, resulta muy adecuado el **diagrama de sectores**.

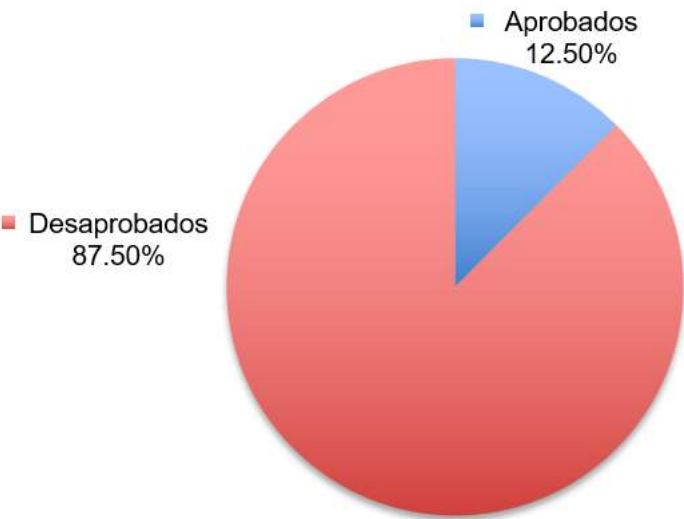
La nueva tabla de frecuencias quedaría de esta manera

Tabla 6

Exámenes	Frecuencia	Frecuencia relativa	Porcentaje
Aprobados	5	0,125	12,5
Desaprobados	35	0,875	87,5

Fuente: elaboración propia.

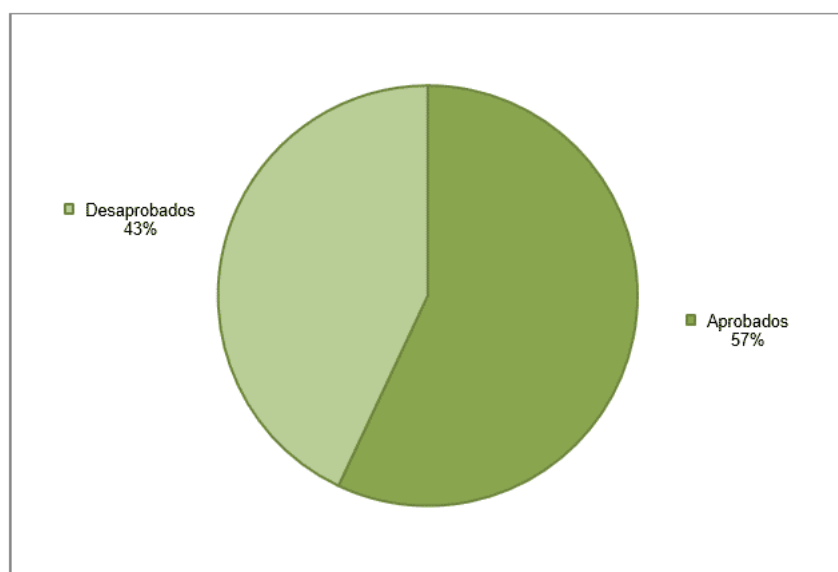
Figura 3



Fuente: elaboración propia.

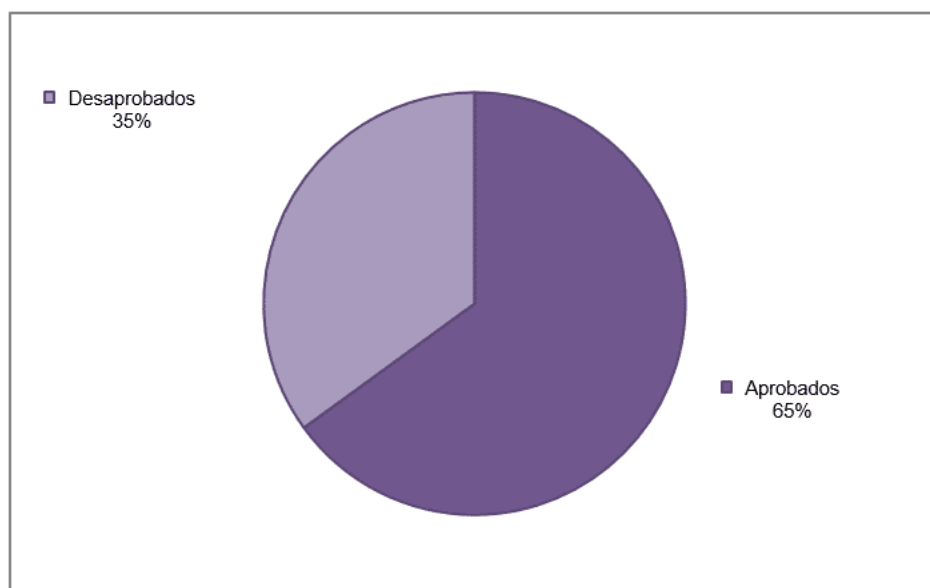
El diagrama de sectores es adecuado para representar varias situaciones y establecer comparaciones. Por ejemplo, para un centro educativo, poder contrastar año a año el porcentaje de estudiantes aprobados y desaprobados en alguna materia.

Figura 4: Porcentaje de estudiantes aprobados y desaprobados – Año 2018



Fuente: elaboración propia.

Figura 5: Porcentaje de estudiantes aprobados y desaprobados – Año 2019



Fuente: elaboración propia.

En resumen



A las características o propiedades que posee un determinado grupo de personas o a los elementos que queremos estudiar se los denomina **variable estadística**. Las hay de dos tipos:

- **Variables cuantitativas:** son las que se expresan mediante números.
- **Variables cualitativas:** son las que no se pueden expresar mediante números.



Al conjunto de todos los **elementos** sobre los que se hace un estudio estadístico se le denomina **población**. A una representación significativa de esta se la denomina **muestra**.



Para obtener los datos necesarios de un estudio estadístico hace falta realizar una **encuesta**. Es muy importante que estén bien elaboradas.



Cuando se tienen todos los datos hay que agruparlos, preferiblemente en tablas. De este modo, se obtienen las **frecuencias** de cada uno de los valores. Hay varios tipos de frecuencias.



Una **gráfica** es la representación visual de los datos obtenidos en un trabajo estadístico. Así, de una ojeada podemos ver, con facilidad, los datos obtenidos.



Hay muchos tipos de gráficas. Se usan unas u otras según el estudio que se realiza. Las más importantes son: **diagrama de barras**, **histogramas**, **diagramas de sectores**, **pictogramas**, etc.

Unidad 4.2 Estadística descriptiva

Parámetros estadísticos

Las tablas de frecuencias son una forma organizada de dar toda (o casi toda) la información de los datos que disponemos. "Con las gráficas estadísticas se pierde algo de información, pero el mensaje 'entra por los ojos'" (Carmesi Matematic, s.f., <https://carmesimatic.webcindario.com/estadisticamat.htm>). A la hora de poder hacer referencias o comparaciones de una distribución con otras es necesario el uso de parámetros estadísticos.

Parámetros estadísticos. Medidas de centralización o de tendencia central



Moda o modo: el valor que se presenta con mayor frecuencia.



Media: es el cálculo del **promedio**.

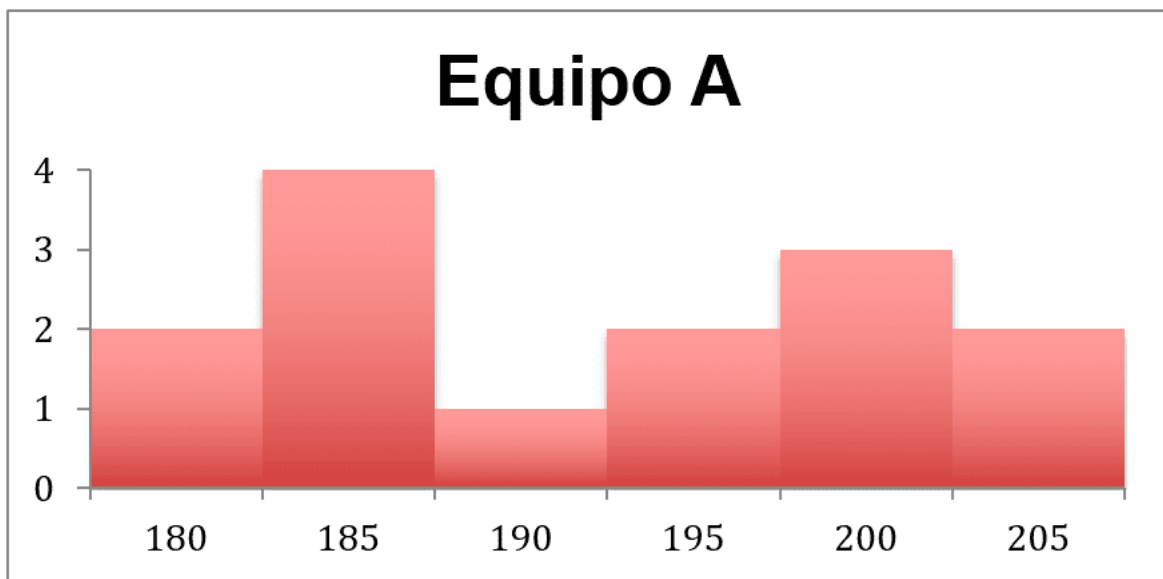


Mediana: es el valor del individuo o elemento que ocuparía el lugar central si se colocaran los elementos ordenados de menor a mayor.

4.2.1 Media

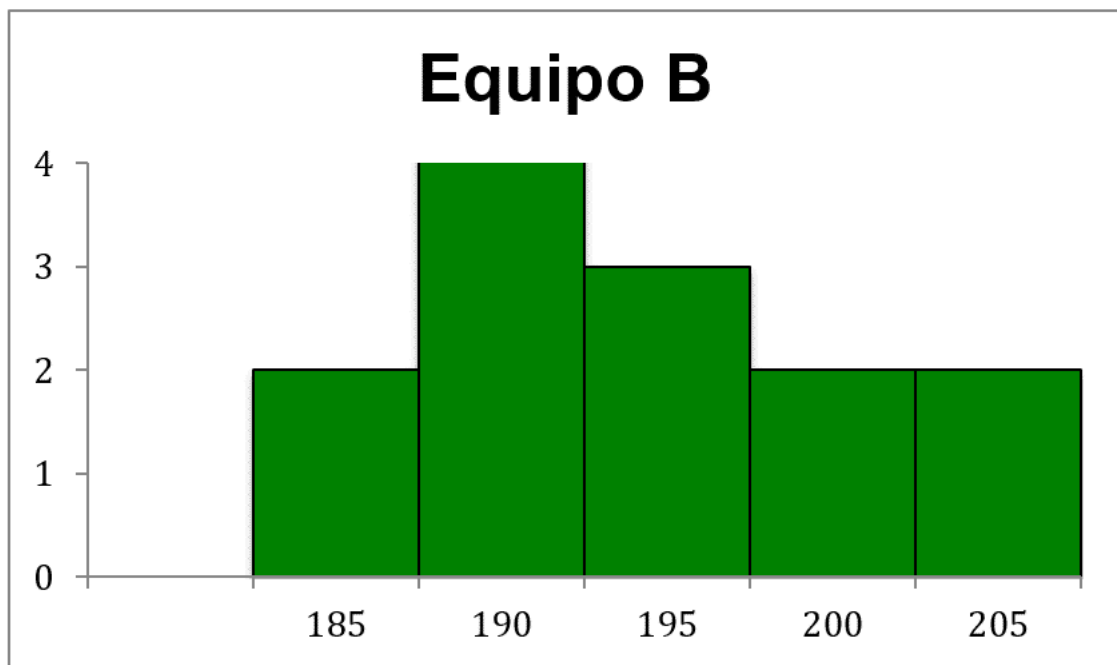
Se quiere hacer un estudio comparativo sobre las estaturas de los jugadores que integran diferentes equipos de básquet. Las estaturas de los jugadores de básquet de los equipos A y B se resumen en las siguientes gráficas.

Figura 6



Fuente: elaboración propia.

Figura 7



Fuente: elaboración propia.

Comparar a partir de estas gráficas la altura de dos equipos es posible. Pero si tuviésemos que comparar las alturas de 11 equipos de básquet sería un trabajo muy difícil.

¿Qué valor numérico podría representar a cada una de las distribuciones? El valor ideal para representar una distribución de datos es la **media o promedio**.

Veamos ahora las tablas de frecuencias de los jugadores de básquet de los equipos A y B que muestran las alturas de sus 14 jugadores.

Tabla 7: Equipo A

Alturas (en cm)	Jugadores Equipo A
183	1
184	1
187	1
188	1
189	2
193	1
196	1
199	1
202	1
203	1
204	1
206	1
207	1

Fuente: elaboración propia.

Tabla 8: Equipo B

Alturas (en cm)	Jugadores Equipo B

Observá que ampliamos la tabla, de modo que en la última columna tengamos el producto de la primera por la segunda. De esta manera, evitamos escribir 2 veces la calificación 2, 3 veces la calificación 3, etc.

Tabla 9

Notas x	Frecuencia absoluta f	$x.f$
1	0	0
2	2	4
3	3	9
4	4	16
5	8	40
6	9	54
7	5	35
8	3	24
9	4	36
10	2	20
Total	40	238

Fuente: elaboración propia.

La suma de todos estos productos lo dividiremos por el total de estudiantes.

$$\bar{x} = \frac{238}{40} = 5,95$$

La nota media del examen es 5,95.

La fórmula general para calcular la media es:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 \dots + x_n}{n}$$

Si los datos vienen agrupados en una tabla de frecuencias, la fórmula general es:

$$\bar{x} = \frac{f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + \dots + f_n \cdot x_n}{n}$$

4.2.2 Mediana

La mediana es el valor del elemento "que ocuparía el lugar central si se colocaran ordenados de menor a mayor" (Carmesi Matematic, s.f., <https://carmesimatic.webcindario.com/estadisticamat.htm>).

Veamos un ejemplo: en una oficina se preguntó a los empleados cuántos hijos tiene cada uno. Los resultados fueron los siguientes:

1 2 1 0 5 3 4

Para calcular la mediana, es decir, el valor central de estos datos los ordenaremos de menor a mayor, tal como se muestra a continuación:

0 1 1 2 3 4 5

Éste es el valor que ocupa el lugar de la mediana.

Denotaremos a la mediana con el símbolo M_e . En este caso, $M_e = 2$.

Si el número de datos es par, la mediana será la media o el promedio de los dos valores centrales.

Se hizo la misma pregunta en otro sector de la empresa y los resultados fueron los siguientes:

1 2 4 2 6 1 4 4

Ordenamos los datos de menor a mayor:

1 1 2 2 4 4 4 6

Estos valores ocupan el lugar central.

Calculamos el promedio de estos dos valores:

$$\frac{2 + 4}{2} = \frac{6}{2} = 3$$

El valor de la mediana es 3.

4.2.3 Moda o modo

En una distribución de frecuencia, al valor más frecuente se le llama modo o moda (lo que más se repite está de moda).

Veamos tres ejemplos.

Figura 8



Fuente: elaboración propia.

Figura 9



Fuente: elaboración propia.

Tabla 10

Notas del examen final	Frecuencia absoluta
1	0
2	2
3	3
4	4
5	8
6	9

7	5
8	3
9	4
10	2

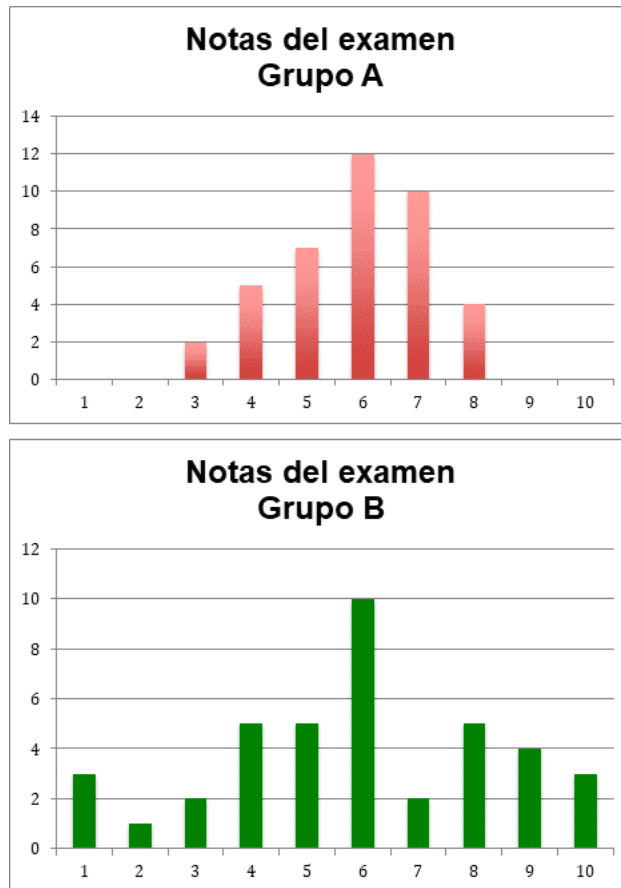
Fuente: elaboración propia.

Resulta sencillo observar el valor de la moda en los tres ejemplos anteriores. En el gráfico de barras que representa las edades de los estudiantes de un curso de fotografía, la moda la ocupa la edad 29; en la encuesta representada en el gráfico de sectores, la moda es realizar actividad física durante el fin de semana; finalmente, en la tabla de frecuencias con las notas de un examen, la moda es la calificación 6.

4.2.4 Otros parámetros estadísticos: la varianza y la desviación estándar

Los diagramas de barras muestran las notas de un examen en dos grupos A y B. La nota media o promedio en ambos grupos es la misma: 5,9.

Figura 10



Fuente: elaboración propia.

A pesar de que los grupos A y B tienen la misma media, sus distribuciones son muy diferentes. Las notas del grupo A son más parejas, mientras que las notas del grupo B son más extremas.

Por esto, necesitamos, además de la media, otro parámetro que mida cómo están dispersos los datos con relación a la media.

Veamos el concepto en un ejemplo práctico. Para esto trataremos de averiguar lo separados que están de la media los datos de la distribución. Para ello, vamos a medir la separación (la diferencia) entre cada dato y la media. A partir de estas diferencias, siguiendo los pasos que veremos abajo, encontraremos la **desviación estándar**.

Hacemos los cálculos para los exámenes del grupo A.

1

A cada valor le restamos la media.

2

Elevamos al cuadrado las diferencias obtenidas.

3

Multiplicamos los cuadrados por la frecuencia.

Tabla 11

x	f	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f$
1	0	$1 - 5,9 = -4,9$	$(-4,9)^2 = 24,01$	$24,01 \cdot 0 = 0$
2	0	$2 - 5,9 = -3,9$	$(-3,9)^2 = 15,21$	$15,21 \cdot 0 = 0$
3	2	$3 - 5,9 = -2,9$	$(-2,9)^2 = 8,41$	$8,41 \cdot 2 = 16,82$
4	5	$4 - 5,9 = -1,9$	$(-1,9)^2 = 3,61$	$3,61 \cdot 5 = 18,05$
5	7	$5 - 5,9 = -0,9$	$(-0,9)^2 = 0,81$	$0,81 \cdot 7 = 5,67$
6	12	$6 - 5,9 = 0,1$	$(0,1)^2 = 0,01$	$0,01 \cdot 12 = 0,12$
7	10	$7 - 5,9 = 1,1$	$(1,1)^2 = 1,21$	$1,21 \cdot 10 = 12,1$
8	4	$8 - 5,9 = 2,1$	$(2,1)^2 = 4,41$	$4,41 \cdot 4 = 17,64$
9	0	$9 - 5,9 = 3,1$	$(3,1)^2 = 9,61$	$9,61 \cdot 0 = 0$
10	0	$10 - 5,9 = 4,1$	$(4,1)^2 = 16,81$	$16,81 \cdot 0 = 0$
Total	40			70,4

Fuente: elaboración propia.

4

Sumamos los cuadrados multiplicados por la frecuencia y lo dividimos entre el total de exámenes.

$$\frac{70,4}{40} = 1,76$$

El valor obtenido se llama **varianza** de la distribución.

La raíz cuadrada de la varianza se llama **desviación estándar** y se lo simboliza con la letra griega σ (sigma).

En este caso, $\sigma = \sqrt{1,76} \approx 1,33$ es la desviación estándar de la distribución de exámenes del grupo A.

Vamos a calcular ahora la desviación estándar de la distribución de exámenes del grupo B.

Repetimos el procedimiento anterior.

Tabla 12

x	f	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f$
1	3	$1 - 5,9 = -4,9$	$(-4,9)^2 = 24,01$	$24,01 \cdot 3 = 72,03$
2	1	$2 - 5,9 = -3,9$	$(-3,9)^2 = 15,21$	$15,21 \cdot 1 = 15,21$
3	2	$3 - 5,9 = -2,9$	$(-2,9)^2 = 8,41$	$8,41 \cdot 2 = 16,82$
4	5	$4 - 5,9 = -1,9$	$(-1,9)^2 = 3,61$	$3,61 \cdot 5 = 18,05$
5	5	$5 - 5,9 = -0,9$	$(-0,9)^2 = 0,81$	$0,81 \cdot 5 = 4,05$
6	10	$6 - 5,9 = 0,1$	$(0,1)^2 = 0,01$	$0,01 \cdot 10 = 0,01$
7	2	$7 - 5,9 = 1,1$	$(1,1)^2 = 1,21$	$1,21 \cdot 2 = 2,42$
8	5	$8 - 5,9 = 2,1$	$(2,1)^2 = 4,41$	$4,41 \cdot 5 = 22,05$
9	4	$9 - 5,9 = 3,1$	$(3,1)^2 = 9,61$	$9,61 \cdot 4 = 38,44$
10	3	$10 - 5,9 = 4,1$	$(4,1)^2 = 16,81$	$16,81 \cdot 3 = 50,43$
Total	40			239,51

Fuente: elaboración propia.

$$\frac{239,51}{40} = 5,99$$

El valor obtenido se llama **varianza** (σ) de la distribución.

En este caso, $\sigma = \sqrt{5,99} \approx 2,45$ es la desviación estándar de la distribución de exámenes del grupo B.

La desviación estándar en B es mayor que en A, esto significa que las notas de los exámenes del grupo B están más separadas de la media que las del equipo A, como se observan en los diagramas de barras.

La **fórmula** general para calcular la **desviación estándar** es:

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{N}}$$

$x_1, x_2, x_3, \dots, x_n$: cada uno de los valores que toma x .

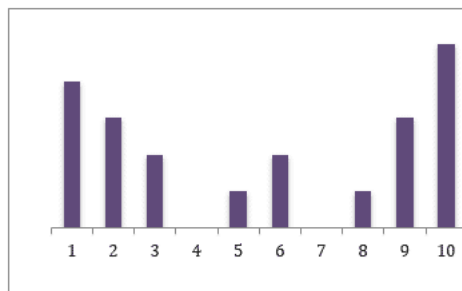
\bar{x} : media

N : es la suma de las frecuencias.

La media, \bar{x} , de una distribución ya sabés lo que significa. La desviación estándar, σ , mide la dispersión de los valores y es un valor mayor cuanto más alejados estén de la media.

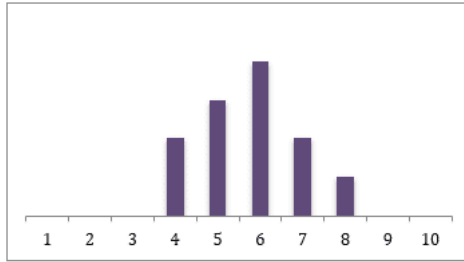
Veamos otro ejemplo. Las siguientes distribuciones tienen la misma media 5, aproximadamente. Sin embargo, sus desviaciones estándar son 1, 2, y 3. ¿Cuál será la desviación estándar de cada una?

Figura 11: Gráfica A



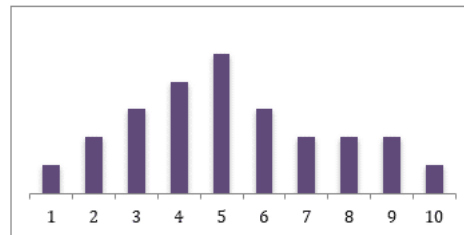
Fuente: elaboración propia.

Figura 12: Gráfica B



Fuente: elaboración propia.

Figura 13: Gráfica C



Fuente: elaboración propia.

Observemos las gráficas y determinemos cuál de ella tiene los valores menos alejados respecto a la

$$\bar{x} = 5.$$

De las tres gráficas, la que tiene los valores de x más agrupados alrededor de la media es la gráfica b. La gráfica C, tiene valores alejados de la media, pero tiene concentrada una gran cantidad de valores alrededor de ella. Por último, la gráfica A es la que tiene los valores más dispersos y pocos valores cercanos a

$$\bar{x} = 5.$$

Según lo observado, la desviación estándar que le correspondería a cada distribución sería:

Gráfica B, $\sigma=1$.

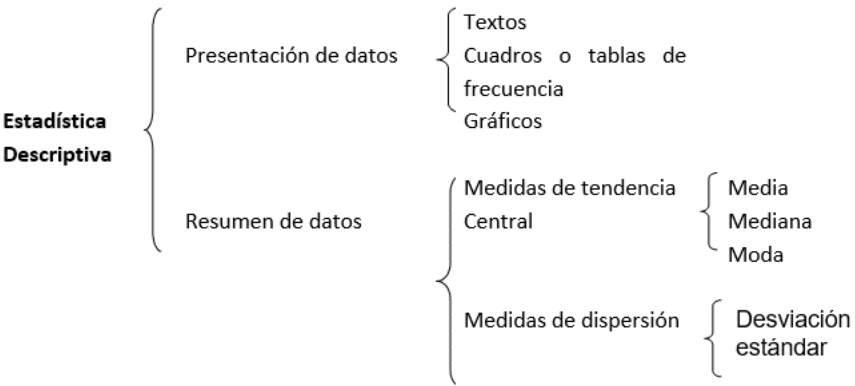
Gráfica C, $\sigma=2$.

Gráfica A, $\sigma=3$.

En conclusión, los ejemplos que hemos observado nos llevan a entender que solo el conocimiento de la media aritmética no es suficiente para caracterizar a un conjunto de datos.

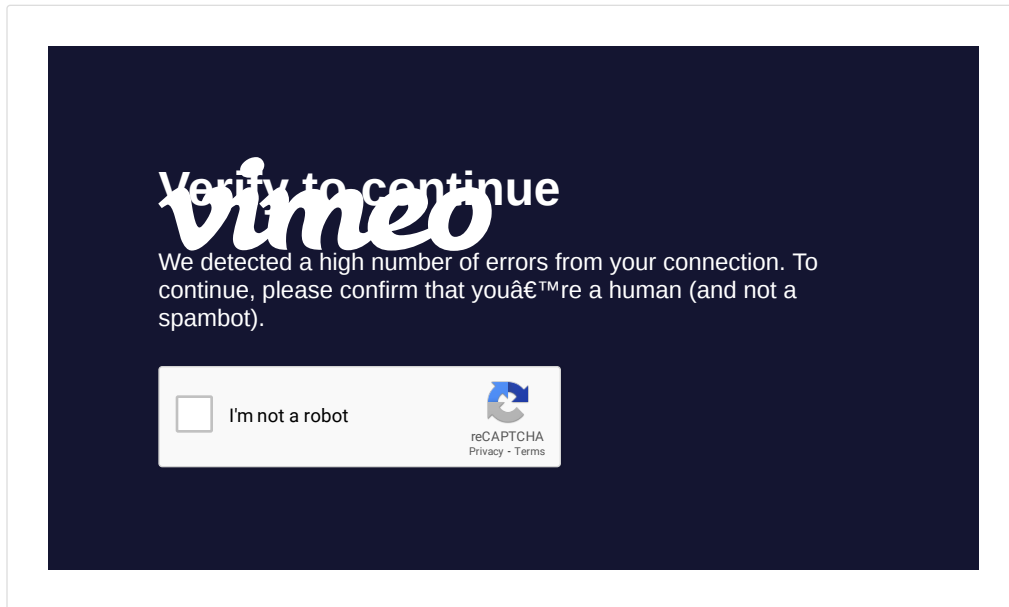
En resumen

Figura: 14



Fuente: elaboración propia.

Video de habilidades



¿Cuáles de las siguientes afirmaciones son verdaderas?

- ☐ La media es un parámetro estadístico.
- ☐ Para calcular la media se suman todos los datos.
- ☐ Calcular la media es calcular el promedio.
- ☐ La media es el valor ideal para representar una distribución de datos.

SUBMIT

¿Cuáles de las siguientes afirmaciones son verdaderas?

- ☐ La mediana es un parámetro estadístico.
- ☐ Para calcular la mediana primero se ordenan los datos de mayor a menor.
- ☐ Si el número de datos es par, una vez ordenados los datos se calcula el promedio de los dos centrales.
- ☐ Para calcular la mediana primero se ordenan los datos de menor a mayor. La mediana es el valor central.
- ☐ El símbolo que representa a la mediana es M_0 .

SUBMIT

¿Cuáles de las siguientes afirmaciones son verdaderas?

- ☐ La moda o modo es un parámetro estadístico.
- ☐ La moda es el valor que se repite con menor frecuencia.
- ☐ Sea el conjunto $\{1, 3, 5, 5, 1, 7, 5, 2, 5, 3, 2, 1\}$ la moda es 1.
- ☐ Sea el conjunto $\{1, 3, 5, 7, 9\}$ la moda no existe.
- ☐ La moda es el valor que se repite con mayor frecuencia.

SUBMIT

Las siguientes son las edades de un grupo de estudio. Calculá la media, la mediana y la moda de estos datos.

18	19	21	18	19	21	24	18	22
----	----	----	----	----	----	----	----	----

☐

Media = 20

☐

Moda = 21

☐

Media = 18

☐

Mediana = 19

☐

Mediana = 18

☐

Moda = 18

SUBMIT

¿Cuáles de las siguientes afirmaciones son verdaderas?

☐

En el video, las edades del grupo A son menos dispersas respecto a la media, por lo tanto, su desviación estándar es menor a la desviación estándar del grupo B.

☐

Cuanto menos alejados están los datos de la media, más grande es la desviación estándar.

☐

Los edades del grupo B son parejas y cercanas al valor de la media.

☐

El grupo B no tiene edades extremas.

☐

Cuanto más alejados están los datos de la media, más grande es la desviación estándar.

SUBMIT

Cierre

La importancia de la Estadística

Permite analizar y procesar grandes masas de datos y números, de manera de obtener información para tomar decisiones, u obtener conclusiones.

Población

Se denomina así al conjunto de todos los elementos sobre los que se hace un estudio estadístico. Una muestra es un subconjunto de una población.

Los dos tipos de variables estadísticas

- Variables cuantitativas: son las que se expresan mediante números. Pueden ser discretas o continuas.
- Variables cualitativas: indican características o cualidades. Son las que no se pueden expresar mediante números (sexo, nacionalidad, nivel socioeconómico, estado civil, profesión, nivel de estudios, etc.).

Distribuciones de frecuencias

Son tablas que facilitan la lectura de los datos. A estas tablas las llamamos tablas de frecuencias.

Gráficos estadísticos

Son representaciones visuales de los datos obtenidos en un trabajo estadístico. La finalidad de los mismos es que la información “entre por los ojos”. Existen gráficos de distintos tipos, pero todos ellos resultan muy fáciles de interpretar, pues esa es, precisamente, su razón de ser. Los distintos tipos de gráficas se usan unas u otras, según el estudio que se realice. Las más importantes son: diagrama de barras, histogramas, diagrama de sectores, etc.

Parámetros estadísticos

- Moda o modo: es el valor que se presenta con mayor frecuencia
- Media: es el cálculo del promedio.

- Mediana: es el valor del individuo o elemento que ocuparía el lugar central si se colocaran los elementos ordenados de menor a mayor.

Referencias

Carmesi Matematic (s.f.). Elementos de probabilidad y estadística. Recuperado de <https://carmesimatic.webcindario.com/estadisticamat.htm>

Walpole, R.; Myers, V. y Myers, S. (2012). Probabilidad y estadística para ingeniería y ciencias. Recuperado de <https://aprenderly.com/doc/1101567/probabilidad-y-estadistica-para-ingenieria-y-ciencias?page=687>