



Representaciones Gráficas de Datos

Fabiola Medina Cabrera

Una representación gráfica es una manera de visualizar los datos cuantitativos. Las tablas estadísticas presentan la información de modo esquemático, mientras que los gráficos estadísticos muestran esa información de manera más expresiva. Cuando se hacen correctamente, las representaciones gráficas permiten, con solo verlas, entender la naturaleza de los datos, observar sus características e incluso formular conclusiones sobre el comportamiento de una población o muestra.

Para elaborar cualquier representación gráfica es necesario tomar en cuenta lo siguiente:

- El tipo de representación gráfica que alcance su objetivo con la mayor sencillez, será el más efectivo.
- Toda representación gráfica debe explicarse absolutamente por sí misma. Por lo tanto, deben estar correctamente indicados el título, el origen, las escalas y las claves explicativas.
- Tener indicado siempre las escalas y las unidades de medida.

Las representaciones gráficas (o gráficos estadísticos) más usuales son:

Diagrama de barras

Estos diagramas son especialmente útiles cuando se trata de relacionar datos cualitativos o cuantitativos de carácter discreto. Se usan para representar gráficamente series estadísticas de valores en un sistema de ejes cartesianos, de manera que en las abscisas se indica el valor de la variable estadística y en las ordenadas se señala su frecuencia absoluta. Resultan útiles cuando se requiere presentar a una variable en dos situaciones distintas.

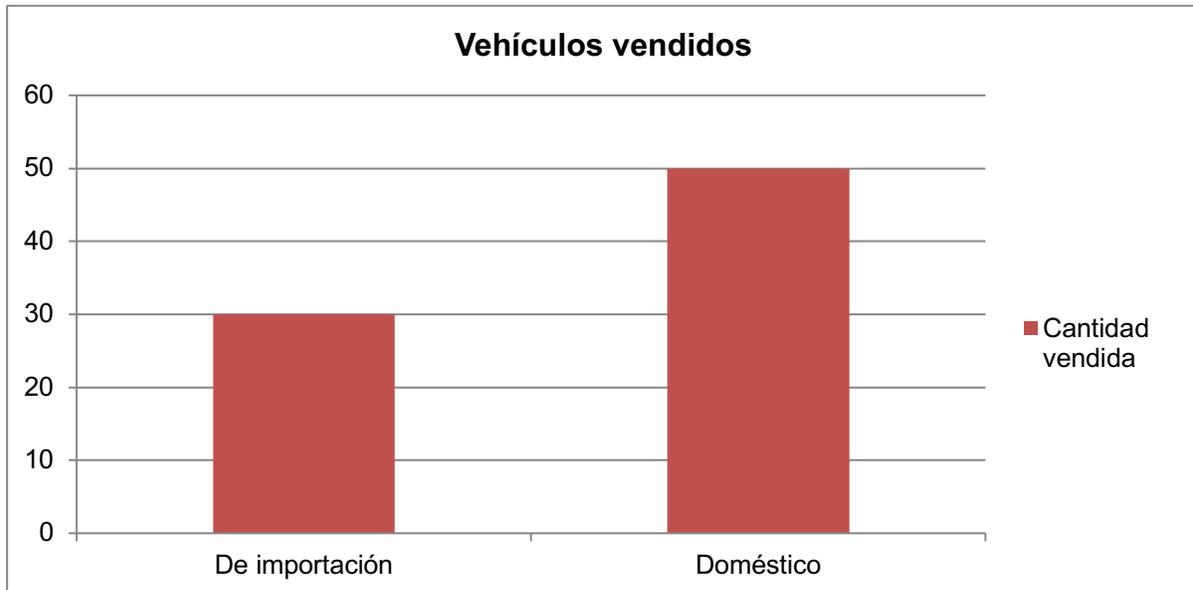
Utilizando los datos de una agencia de autos, tenemos la siguiente tabla:

Vehículos vendidos por una
agencia automotriz

Tipo de Vehículo	Cantidad vendida
De importación	30
Doméstico	50

La variable de interés es el tipo de vehículo y la cantidad de cada tipo de vehículos vendidos es la frecuencia de clase. Se representa el tipo de vehículo (doméstico o de importación) sobre el eje horizontal y la cantidad de cada artículo sobre el eje

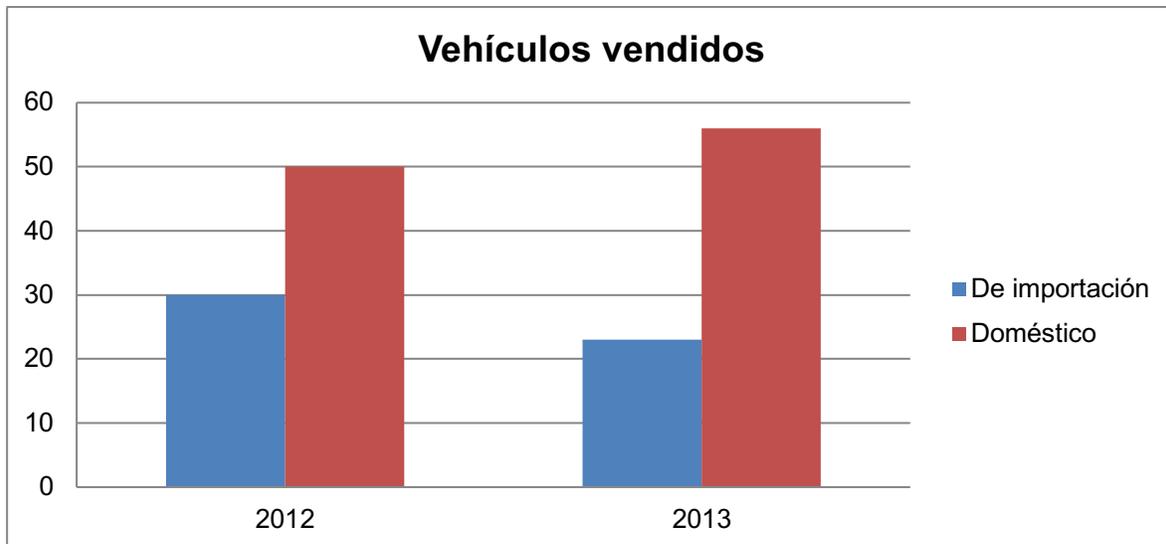
vertical. La altura de las barras o rectángulos, corresponde a la cantidad de vehículos vendidos de cada tipo.



Cuando se requiere representar a la misma variable en situaciones distintas, por ejemplo en momentos diferentes, la gráfica de barras queda:

Vehículos vendidos por una agencia
automotriz

Tipo de Vehículo	Cantidad vendida	
	2012	2013
De importación	30	23
Doméstico	50	56



Cuando se usan estas barras múltiples, caben dos alternativas; o bien las barras que representan los subgrupos se dibujan contiguas, o van separadas por un pequeño espacio. Si se emplea este último procedimiento el espacio entre grupo de barras. En todos los casos, cuando se usen barras múltiples será necesaria una clave para diferenciar los diversos subgrupos.

Las barras pueden ser representadas tanto horizontal como verticalmente; una norma para determinar la orientación de las barras, es que éstas se tracen verticalmente si ello no impide escribir debajo de la barra la clave correspondiente a esta; en caso contrario, se representarán horizontalmente. De esta manera, la clave puede ser leída sin hacer girar el diagrama. Independientemente de la posición de las barras, el origen de la escala debe coincidir con el cero, los intervalos deben progresar de forma regular o uniforme.

Histograma

Es un diagrama de barras especialmente recomendado para representar datos numéricos. La diferencia con las gráficas de barras es que en el histograma las barras van unidas, para indicar continuidad.

En el eje de las abscisas de un histograma se representa la variable. En el eje de las ordenadas se representan las densidades de frecuencia.

En el caso del histograma cada intervalo se representa sobre el eje de las abscisas (X), este será la base del rectángulo que se dibuja sobre él con altura igual o proporcional a su frecuencia absoluta. Como los intervalos son consecutivos, los rectángulos quedan uno junto a otro. En caso de que los intervalos tengan la misma amplitud, en lugar de utilizar las densidades de frecuencia, se pueden utilizar las frecuencias absolutas.



Las reglas generales para formar distribuciones de frecuencia, y por lo tanto, para construir histogramas son las siguientes:

- a) Ordenar los datos
- b) Determinar el **Rango**, que es la diferencia entre el mayor y el menor de los números.

Por ejemplo, si se recopiló un conjunto de estaturas de 100 estudiantes, y la mayor es de 188 cm y la menor es de 152 cm; el rango es:
 $188 - 152 = 36$.

- c) Delimitar los intervalos de clase o categorías, para formularlos es necesario en primer término dividir el rango entre el número de intervalos que se considere adecuado (generalmente entre cinco y veinte, dependiendo de los datos); y de acuerdo con el valor obteniendo se determinan los límites para cada uno de los intervalos.

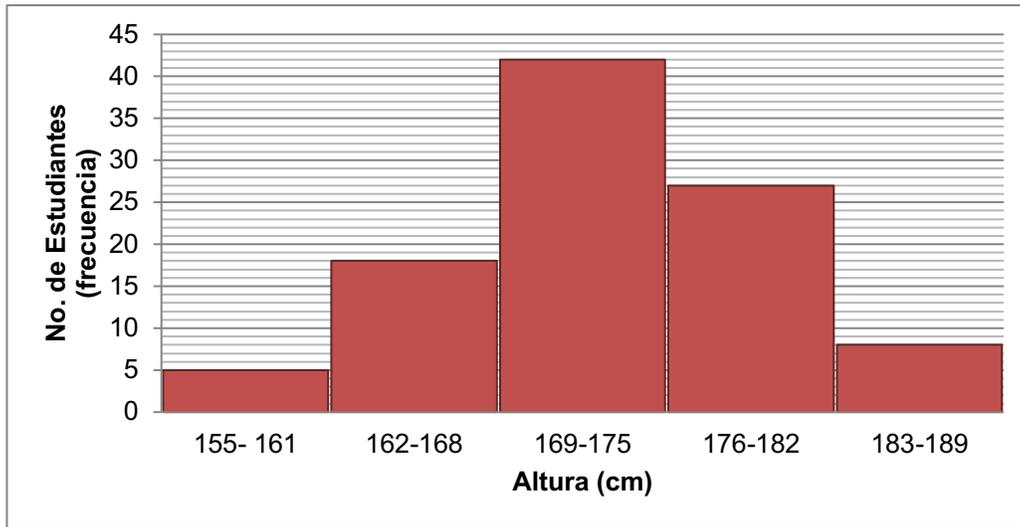
Continuando con el mismo ejemplo, si se quieren distribuir las estaturas de los estudiantes en 5 categorías o intervalos de clase: $36 \div 5 = 7.2 \approx 7$, cada categoría o intervalo de clase debe abarcar 7 valores:

- d) Determinar las Frecuencias de Clase, o sea, el número de individuos que pertenece a cada categoría o intervalo de clase. En el ejemplo se puede observar que 5 estudiantes tienen una estatura entre 155 y 161, al intervalo de clase se 155 -161 le corresponde una frecuencia de 5, así sucesivamente para cada intervalo.
- e) Construir una tabla de frecuencias, es decir, una representación tabular que reúna todos los datos correspondientes a los intervalos de clase y las frecuencias correspondientes:

Tabla de frecuencias

Intervalo de clase	Altura	Número de estudiantes
1	155- 161	5
2	162-168	18
3	169-175	42
4	176-182	27
5	183-189	8

f) Construir el histograma de acuerdo con los datos de la tabla de frecuencia:



Polígono de Frecuencias:

Se utilizan para variables estadísticas cuantitativas, discretas o continuas.

Para una variable discreta, el polígono de frecuencias se obtiene uniendo por una poligonal, los extremos superiores de las barras.

Para una variable continua, el polígono de frecuencias se obtiene uniendo por una poligonal los puntos medios de la base superior de los polígonos del histograma.

Las escalas utilizadas para representar los polígonos de frecuencias influyen mucho por el impacto visual de los mismos.

En los polígonos de frecuencia es posible representar más de una variable en el mismo gráfico.

Se construye trazando una línea que una los puntos medios de los techos de los rectángulos o barras de un histograma; así tenemos el polígono de frecuencia del ejemplo anterior:

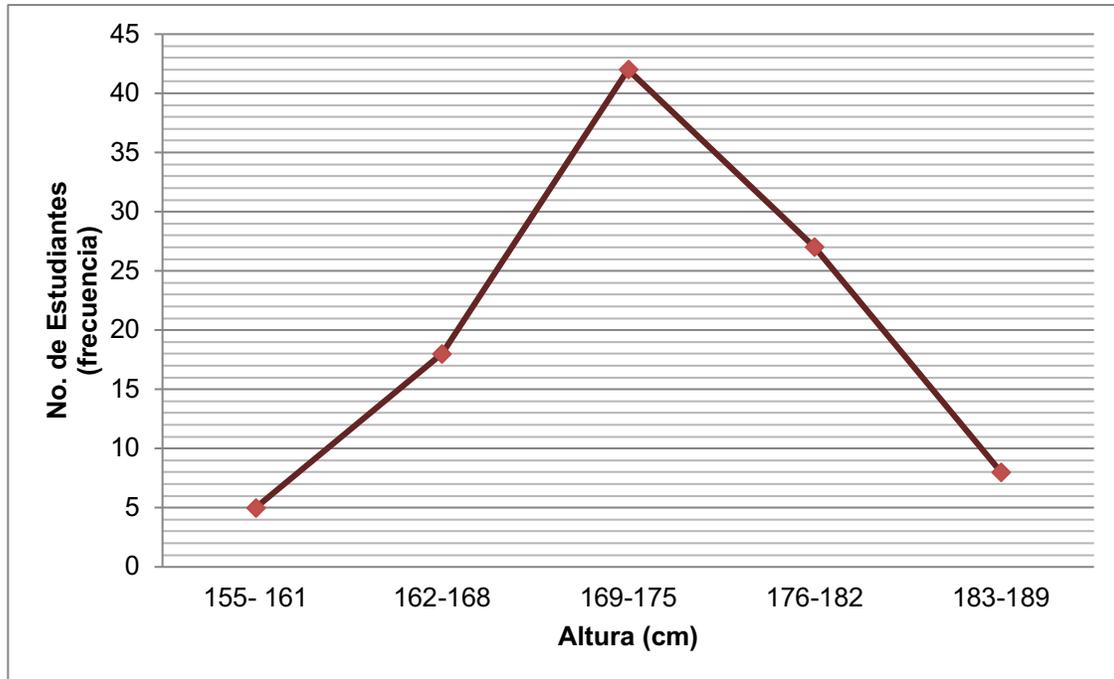


Diagrama de Sectores (o gráfica de pastel).

En los diagramas de sectores, también llamados circulares o de pastel, se muestra el valor de la frecuencia de la variable señalada como un sector circular dentro de un círculo completo. Por ello, resultan útiles particularmente para mostrar comparaciones entre datos, sobre todo en forma de frecuencias relativas de las variables expresadas en forma de porcentaje.

Se utiliza para todo tipo de variable estadística, cuantitativa o cualitativa. Consiste en dibujar sectores sobre un círculo, siendo la amplitud de los sectores proporcional a su frecuencia absoluta, cada sector se rellena con un color diferente.

El cálculo de la amplitud en grados sexagesimales del sector correspondiente se realiza así:

$$\text{ángulo} = (\text{frecuencia relativa}) * 360$$

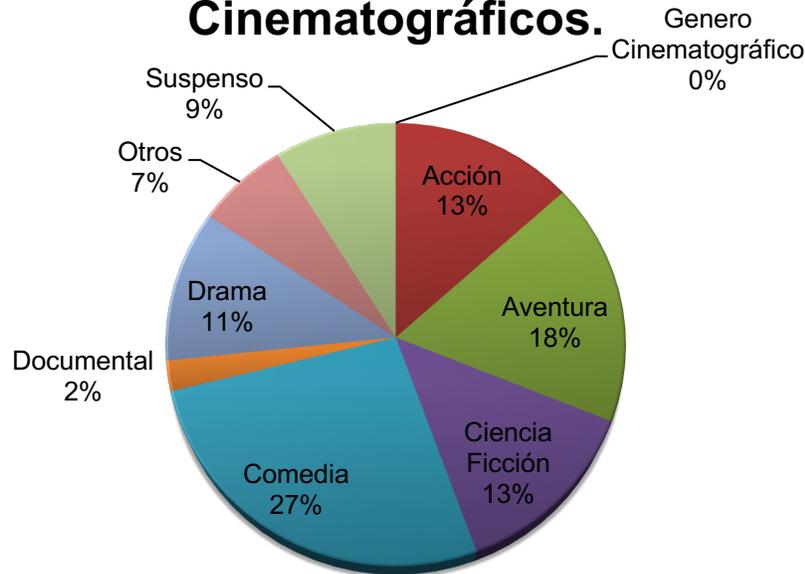
Por ejemplo, tomando la frecuencia con que se proyectan cierto género de películas en las salas de cine de una ciudad determinada.

Género Cinematográfico	
Drama	5

Comedia	12
Documental	1
Aventura	8
Ciencia Ficción	6
Acción	6
Suspense	4

42

Distribución de Géneros Cinematográficos.



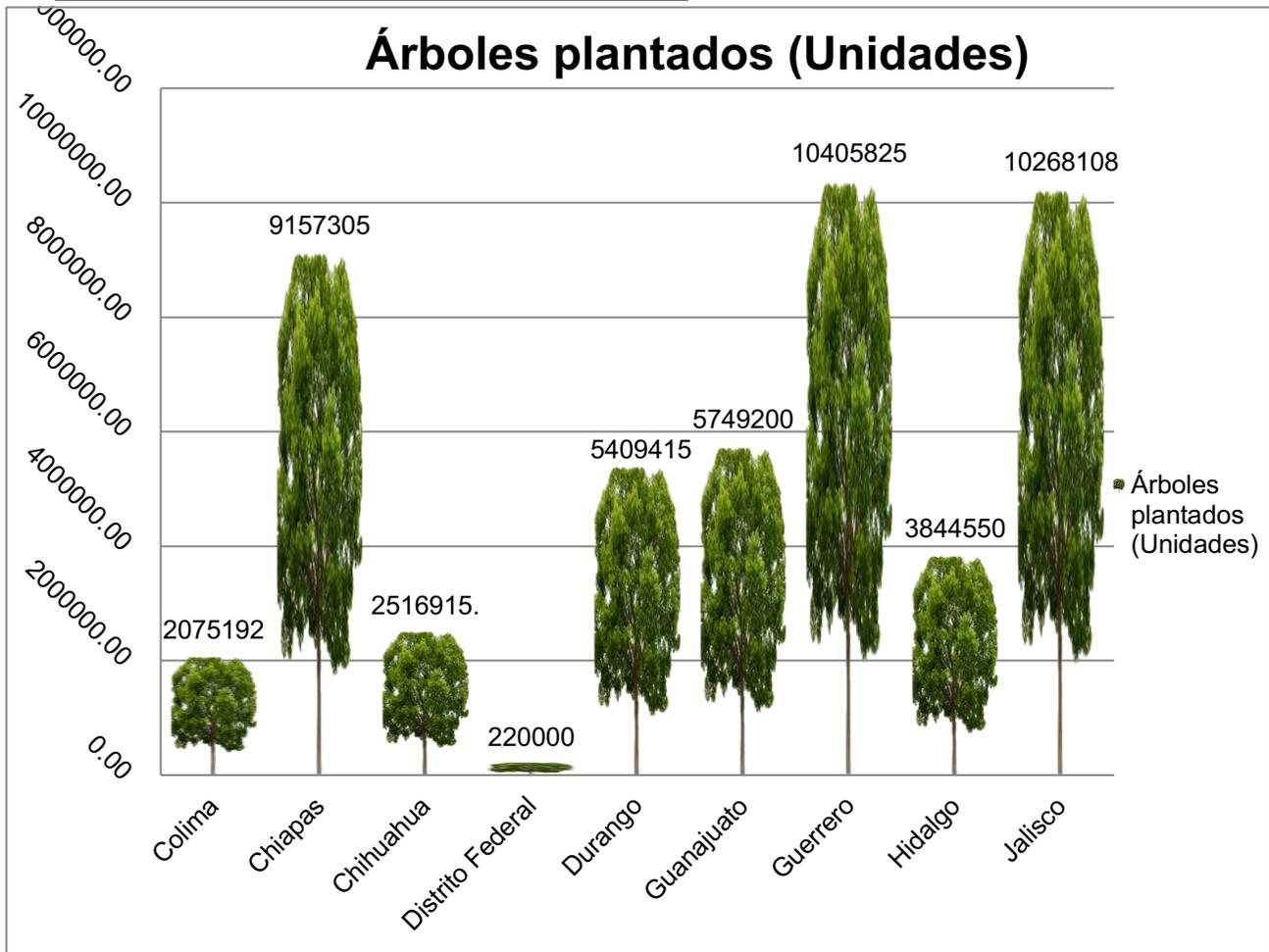
Pictogramas

Para aligerar la presentación de datos estadísticos, con frecuencia se recurre a imágenes pictóricas representativas del valor de las variables. Dos formas comunes de expresión gráfica de los datos son:

- Los pictogramas, que muestran diagramas figurativos con figuras o motivos que aluden a la distribución estadística analizada (por ejemplo, una imagen antropomórfica para indicar tamaños, alturas u otros).

Entidad federativa	Árboles plantados (Unidades)
Estados Unidos Mexicanos	110'559,240
Colima	2'075,192
Chiapas	9'157,305

Chihuahua	2'516,915
Distrito Federal	220000
Durango	5'409,415
Guanajuato	5'749,200
Guerrero	10'405,825
Hidalgo	3'844,550
Jalisco	10'268,108



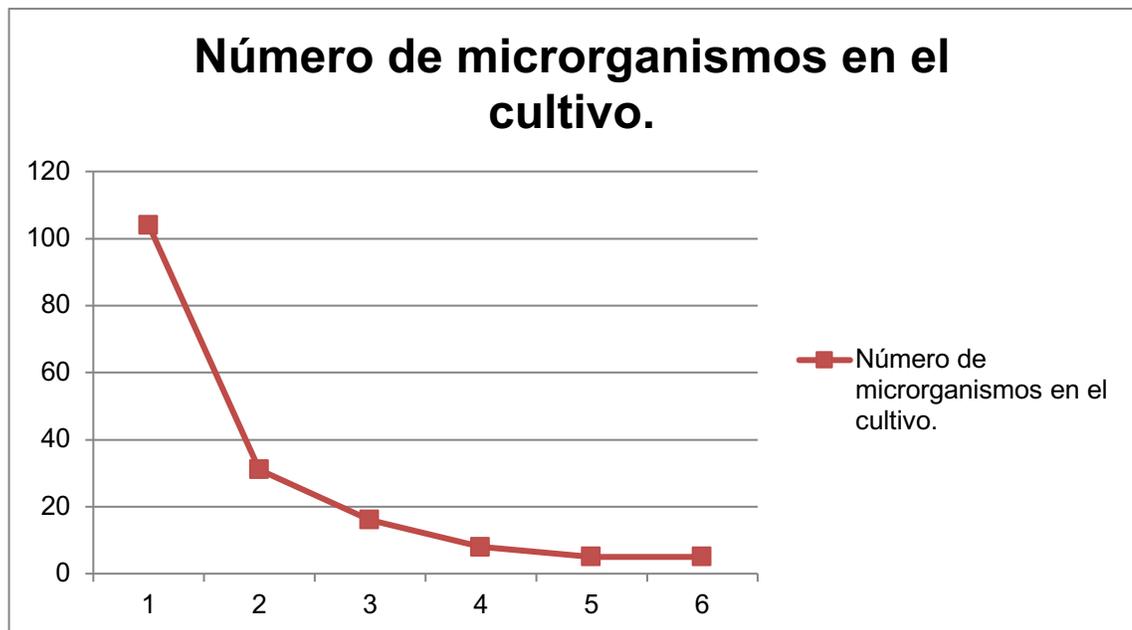
Fuente propia, elaborada con datos de: INEGI. Anuarios Estadísticos de los Estados, 2010. Con base en datos de CONAFOR. Ags, Ags., 2011.

Gráfica de línea

Es la representación gráfica que relaciona dos (o más) variables en un sistema de coordenadas. Este tipo de gráficas resumen una gran cantidad de datos, son útiles porque pueden mostrar relaciones entre factores, y por tanto de ellas se deriva una gran cantidad de información.

Por ejemplo, supongamos que liberamos un número de microorganismos, en un recipiente de cultivo, durante las siguientes 6 horas tomamos nota del número de microorganismos que se hay. Los datos correspondientes se presentan en la siguiente tabla:

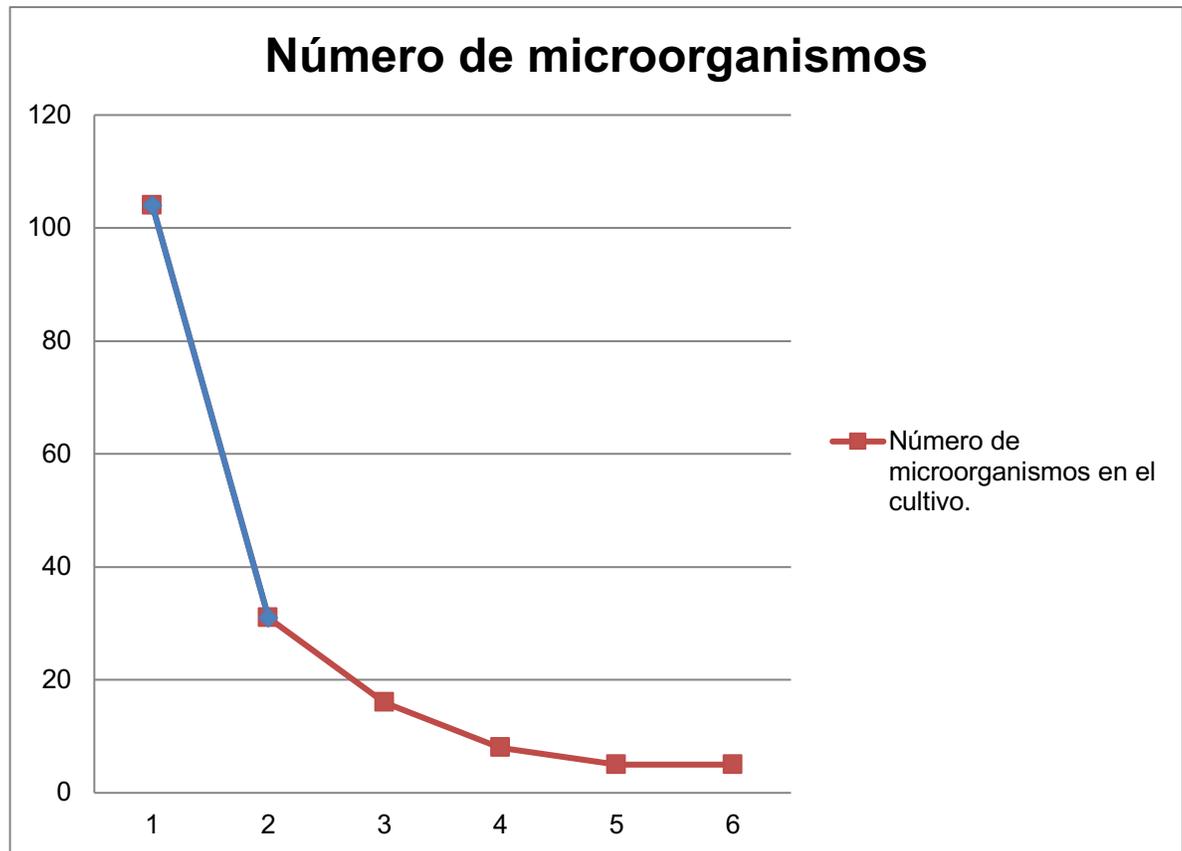
Número de microorganismos en el cultivo.	
Número de horas después de la liberación.	Número de microorganismos en el cultivo.
1	104
2	31
3	16
4	8
5	5
6	5



En la gráfica se puede ver que existe una relación entre el tiempo que ha pasado desde que se liberaron y el número de microorganismos dentro del cultivo, encontramos que el número de individuos disminuye gradualmente a medida que pasa el tiempo. Puesto que la línea no desciende en línea recta, sino que forma una curva de inclinación descendente, la gráfica provee algo más de información; la inclinación de la línea entre la primera y la segunda hora indica que un número mayor de organismos disminuyen durante la primera hora, en comparación con cualquier otro intervalo observado. En otras palabras, a partir de la gráfica podemos saber algo sobre la velocidad con la cual disminuye el número

individuos; la inclinación de la línea da una medida cuantitativa de esta velocidad, si se trata matemáticamente.

Por otro lado, en la gráfica puede observarse que no medimos el número de microorganismos entre las horas 1 y 2, pero lo hubiéramos hecho, habríamos esperado un número alrededor de 50, es decir, en una línea directa entre 104 y 31. Tomando como base esta suposición, nosotros trazamos una línea en la gráfica conectando estos dos datos.



Al trazar esta línea en la gráfica, tenemos un ejemplo de **interpolación**; que quiere decir, determinar una unidad no medida en base a la tendencia que muestra cada partida de datos a cada lado. No hay garantía de que esto sea verdad para cualquier caso individual, sin embargo, la experiencia ha mostrado que cuando se tiene una muestra de datos bastante grande, la interpolación es un proceso válido.

Bibliografía

Johnson, R., & Kuby, P. (2012). *Estadística elemental*. México: Cengage Learning.

Lind, D. A., Marchal, W. G., & Wathen, S. A. (2008). *Estadística Aplicada a los negocios y la economía*. México: McGrawHill.



Moore S., D. (2000). *Estadística aplicada básica*. Barcelona, España: W. H. Freeman and Company.

Ross, S. (2005). *Introducción a la Estadística*. Barcelona: Reverté.