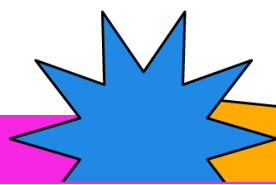
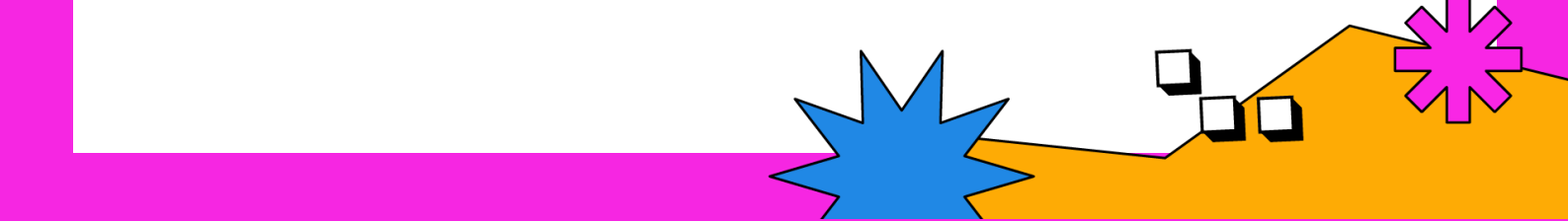


Algunas **orientaciones**, **fórmulas y tips** para usar en la resolución de ejercicios de **Estadística Descriptiva**





ÍNDICE

- 1** ¿Cómo armar **tablas de frecuencias** de datos **agrupados** con **intervalos de clase**?
 - 2** ¿Cómo calcular **las medidas de tendencia central** (MTC) con **tablas de frecuencias de datos agrupados** con **intervalos de clase**?
 - 3** ¿Cómo **calcular las medidas de tendencia central** (MTC)?
 - 4** ¿Cómo hacer **histogramas** en **Geogebra**?
- 



¿Cómo armar **tablas de frecuencias de datos agrupados con intervalos de clase?**



¡Hola!

A lo largo de este material encontrarán algunas **definiciones**, **fórmulas**, **orientaciones** y **tips** para resolver.



Lean la siguiente situación:

A Catalina y Santiago les encanta mirar series y películas. Están asociados a un grupo de fans de cine en Instagram: "Club de Cine MatShow". Quienes lo componen están en una franja etaria de 12 a 29 años.

El desafío fue averiguar: ***¿Qué cantidad de horas por día pasan mirando una película o serie?***

Armaron un **Formulario de Google** para recolectar la información. Y la compartieron al grupo de WhatApps del club.



Le agregaron otras preguntas, para conocer más a los cinéfilos/as

Pueden ver las respuestas en la planilla de cálculo

[EncuestaMatShow](#)



Es muy importante tener en cuenta la forma en que se realizan las **preguntas de las encuestas** y **cómo se recolectan los datos**. *¿Por qué?* De ello dependen las **conclusiones** que se puedan alcanzar.



Una **muestra de respuesta voluntaria** es aquella en la que los **propios encuestados deciden si la contestan** o no.

En este caso no es posible obtener conclusiones sobre la población, dado que no se tiene en cuenta la información que brinda la persona que no accedió a ella -ya sea por no utilizar el medio desde donde se la publicó o por no querer contestar, entre otros factores -. En cambio, si consideramos realizar encuestas en redes sociales, los que responden son usuarios/as de la red quedando fuera otros posibles encuestados.



Analicen la **variable cuantitativa continua**: “Cantidad de horas por día que pasan mirando una película o serie”

Catalina y Santiago, se hicieron estas preguntas:

- * ¿Cómo **ordenar los datos** obtenidos?
- * ¿Cuál es el **valor mínimo o máximo** de horas que un encuestado puede estar conectado en mirando una película o serie?
- * ¿Cuál será **el promedio** de la cantidad de horas que un encuestado pasa conectado mirando una película o serie?



Sigan los **pasos para responder** a las preguntas:

¿Cuál es el **valor mínimo o máximo** de horas que un encuestado puede estar conectado en mirando una película o serie?

Respuesta: Si observamos la planilla el mínimo valor es 1,5 hs.

	A	B	C	D	E
1		Género	Rango de Edades	Subtitulado	Cant tiempo por día (hs)
2	1	drama	16	si	1,25
3	2	acción	25	no	6
4	3	aventura	23	si	5,25
5	4	comedia	16	no	3,5
6	5	ciencia ficción	16	si	3,75
7	6	musical	15	no	4,25
8	7	aventura	23	si	5
9	8	aventura	17	no	4
10	9	ciencia ficción	22	si	3
11	10	drama	28	no	1,5
12	11	drama	29	no	2
13	12	acción	17	no	3,25
14	13	acción	16	si	1,25
15	14	ciencia ficción	15	si	2,25
16	15	comedia	23	no	3
17	16	comedia	22	no	4

18	17	romántica		21	no	5
19	18	aventura		19	si	1,25
20	19	aventura		27	si	6,25
21	20	romántica		27	no	6,5
22	21	ciencia ficción		29	no	3,5
23	22	acción		16	no	3
24	23	aventura		16	no	2
25	24	dibujo		12	si	1,5
26	25	acción		15	si	4
27	26	drama		15	no	5
28	27	romántica		28	si	3
29	28	aventura		27	si	6
30	29	ciencia ficción		26	no	1,25
31	30	romántica		25	no	1,25
32	31	dibujo		12	no	1,25
33	32	ciencia ficción		23	si	2
34	33	acción		27	si	2

Respuesta: Si observamos la planilla el máximo valor es 6,5 hs.

Como los datos son muchos y variados, se deberá pensar en otro modo diferente de organizarlos.

Por eso **usaremos las tablas de frecuencias agrupadas por intervalos de clases** en lugar de las tablas de frecuencia agrupadas por datos sencillos.



Esta guía indica **cómo usar las planillas de cálculo** y algunas sugerencias.

Las calculadoras también trabajan en modo estadística (SD). Pueden investigar cómo se usa en las de ustedes.



¿Qué pasos deben seguir?

➤ 1- Calculá el **Rango**

Tengan en cuenta el ejemplo, deberán calcular:

$$6,5 - 1,5 = 5 \quad \text{R} = D_m - d_m$$



El **rango** es la diferencia entre el dato mayor menos el dato menor.

➤ 2- Calculá la **cantidad de intervalos**

Muchas veces ya tenemos este valor. Cuando no contamos con ello, deben realizar la siguiente fórmula:

$$1 + 3,322 \cdot \log(n) \quad (\text{siendo } n \text{ el número de datos})$$



Un **intervalo de clase** se puede definir como todos los números reales que se encuentran entre dos. Pueden ser intervalos abiertos, cerrados, semiabiertos o semicerrados.

Utilicen los datos del ejemplo, reemplazando los valores:

$$1 + 3,322 \cdot \log(36) = 6.17$$

¿Qué significa este resultado?

Deberíamos armar **6 o 7 intervalos de clases**.

Muchos autores recomiendan que se **aproxime al menor valor e impar**. Entonces en este caso iríamos a **7 intervalos**.

➤ **3-** Calculen la amplitud

$$A = R/i$$

Utilicen los datos del ejemplo y reemplacen los valores:

$$A = 5/7$$

$$A = 0.72$$

➤ **4-** Armen la tabla:

Al símbolo que representa una clase, como 1,50-2,25 en la tabla, se le conoce como **intervalo de clase**.

A El primer intervalo debe comenzar siempre con el Dato menor, al que se denomina **extremo inferior** o límite inferior.

Li

Cant tiempo por día (hs)	fa
[1,5- 2,25)	11

B A ese Dato menor se le suma la Amplitud y de esa manera se encuentra el **extremo superior** o límite superior.

Ls

C Repitan el proceso.
Tengan en cuenta de comenzar cada intervalo con el extremo superior del intervalo anterior.


Cant tiempo por día (hs)	fa
[1,5- 2,25)	11
[2,25-2,97)	1
[2,97-3,69)	9
[3,69-4,41)	6
[4,41-5,13)	3
[5,13-5,85)	1
[5,85-6,57)	4
Total	35



Los términos clase e intervalo de clase se suelen usar indistintamente, aunque el intervalo de clase en realidad es un símbolo para la clase.



Un intervalo de clase que, por lo menos teóricamente, no tenga indicado el límite de clase superior o el límite de clase inferior, se conoce como **intervalo de clase abierto**.



2 ¿Cómo calcular **las medidas de tendencia central (MTC)** con tablas de frecuencias de datos agrupados con intervalos de clase?





Las **MTC** son utilizadas también para presentar resultados de datos muy extensos, como los estudios que incluyen grandes poblaciones, por ejemplo.

Gracias a estas medidas, también se pueden proyectar límites o valores hacia los que tiende a inclinarse la variable que estás analizando.

Si la variable es cuantitativa continua, la media se calcula usando un “representante” de cada intervalo llamado marca de clase.

➤ 1- Calculen la **marca de clase**



La **marca de clase** es el punto medio del intervalo de clase y se obtiene sumando los límites de clase inferior y superior, dividiendo entre 2.



¿Qué pasos deben seguir?

Observen la tercera columna

Con la calculadora **verifiquen** los resultados

Sigan el ejemplo

Cant tiempo por día (hs)	fa	Marca de clase
[1,5- 2,25)	11	$(1,5+2,25)/2=1,875$
[2,25-2,97)	1	2,61
[2,97-3,69)	9	3,33
[3,69-4,41)	6	4,05
[4,41-5,13)	3	4,77
[5,13-5,85)	1	5,49
[5,85-6,57)	4	6,21
Total	35	

➤ **2-** Pasos para hacer el cálculo de la **media aritmética** en una planilla de cálculo:

$$\sum_{i=1}^k \frac{x_i * f_i}{n} = \bar{x}$$

Dato:
fi=fa



El mismo proceso se puede hacer con calculadora y en papel.

Primero: Se le indica a la planilla de cálculo la siguiente fórmula: **fx=C3*D3**.



Deben tener en cuenta donde se encuentra ubicada la columna para realizar este cálculo, en base a ello se debe realizar la fórmula, indicando la columna y fila para multiplicar.

Queda así:

Cant tiempo por día (hs)	fa	Marca de clase	xi*fa
[1,5- 2,25)	11	(1,5+2,25)/2=1,875	
[2,25-2,97)	1		2,61
[2,97-3,69)	9		3,33
[3,69-4,41)	6		4,05
[4,41-5,13)	3		4,77
[5,13-5,85)	1		5,49
[5,85-6,57)	4		6,21
Total	35		0

AUTOCOMPLETAR

Autocompletar sugerido

Ctrl+Enter para autocompletar.

Mostrar fórmula

✓ ✕

Segundo: Si aparece el cartel de **autocompletar**, se puede dar enter en la tilde y repetir la fórmula.

Tercero: Se suma la columna.

Cant tiempo por día (hs)	fa	Marca de clase	xi*fa
[1,5- 2,25)	11	$(1,5+2,25)/2=1,875$	20,625
[2,25-2,97)	1	2,61	2,61
[2,97-3,69)	9	3,33	29,97
[3,69-4,41)	6	4,05	24,3
[4,41-5,13)	3	4,77	14,31
[5,13-5,85)	1	5,49	5,49
[5,85-6,57)	4	6,21	24,84
Total	35		122,145

Cuarto: se aplica la fórmula de la media aritmética

$$122,145/35= 3,489$$

$$\sum_{i=1}^k \frac{x_i * f_i}{n} = \bar{x}$$



➤ **3-** Pasos para hacer el cálculo del **intervalo modal**:



En un estudio estadístico, el dato que más se repite –es decir, el de mayor frecuencia– es **la moda**. La moda no tiene por qué ser una sola. Si la mayor frecuencia se repite, entonces habrá más de una.

Cuando la variable es **cuantitativa continua**, por lo general no se tienen todos los datos a la vista y lo importante es en qué intervalo hay más datos. Por eso se habla del **intervalo modal** en vez de la moda.

Observá la fa el intervalo que tiene mayor cantidad de datos es **[1,5-2,25]**

Cant tiempo por día (hs)	fa
[1,5- 2,25)	11
[2,25-2,97)	1
[2,97-3,69)	9
[3,69-4,41)	6
[4,41-5,13)	3
[5,13-5,85)	1
[5,85-6,57)	4
Total	35

➤ **4-** Pasos para hacer el cálculo del intervalo donde se encuentra **la mediana**:

Primero se debe calcular la mitad de los datos: $\frac{n}{2}$, para encontrar el intervalo de clase que está a la mitad.

Segundo marquen ese renglón e intervalo en el cuál aparecen la mitad de los datos

Busquen en la frecuencia acumulada el valor que lo contenga.

Volvemos al ejemplo inicial:

$$\frac{n}{2} = 35/2 = 17,5$$

Lo podemos ubicar en 21.

Cant tiempo por día (hs)	fa	Marca de clase	Fa	xi*fa
[1,5- 2,25)	11	1,875	11	20,625
[2,25-2,97)	1	2,61	12	2,61
[2,97-3,69)	9	3,33	21	29,97
[3,69-4,41)	6	4,05	27	24,3
[4,41-5,13)	3	4,77	30	14,31
[5,13-5,85)	1	5,49	31	5,49
[5,85-6,57)	4	6,21	35	24,84
Total	35			122,145

Tercero deben completar la fórmula para hallar la mediana. Luego suponemos que los valores en esa clase se distribuyen de manera uniforme y podemos resolver.

$$Me = L_i + \frac{\frac{n}{2} - F_{a-1}}{f_a} \cdot A$$

f_a : frecuencia de la clase de la mediana

L_i : es el límite inferior del intervalo de la clase de la mediana

A: amplitud

F_{a-1} : es la frecuencia acumulada de la clase el intervalo anterior al de la clase de la mediana

Resolvemos:

Se reemplaza en la fórmula. Se deben tener en cuenta los datos

Cant tiempo por día (hs)	fa	Marca de clase	Fa	xi*fa
[1,5- 2,25)	11	1,875	11	20,625
[2,25-2,97)	1	2,61	12	2,61
[2,97-3,69)	9	3,33	21	29,97
[3,69-4,41)	6	4,05	27	24,3
[4,41-5,13)	3	4,77	30	14,31
[5,13-5,85)	1	5,49	31	5,49
[5,85-6,57)	4	6,21	35	24,84
Total	35			122,145

fa: 9

Li= 2,97

A= 0,72

Fa-1=12

n/2= 17,5

$$Me = 2,97 + \frac{17,5-12}{9} \cdot 0,72$$

Me= 3,41



La mediana y la moda son medidas de tendencia central que por sus propiedades destacan los valores individuales de un colectivo; en cambio, **la media aritmética**, al promediar todos los valores igualando en un justo reparto todas las observaciones, suprime sus individualidades.



3

**¿Cómo calcular las
medidas de tendencia
central (MTC) ?**



Las **MTC** también son utilizadas para presentar resultados de datos muy extensos, como los estudios que incluyen grandes poblaciones, por ejemplo.

Gracias a estas medidas, también se pueden proyectar límites o valores hacia los que tiende a inclinarse la variable que estás analizando.

En **Estadística** existen tres medidas llamadas “**medidas de tendencia central**”:

- 1. El promedio o media aritmética**
- 2. La moda.**
- 3. La mediana**

Se denominan **medidas de tendencia central** porque los valores que se obtienen son valores representativos que resumen un conjunto de datos, mostrando la tendencia hacia el centro mismo.

Se puede decir que las medidas de tendencia central **nos sirven para identificar cuáles son las tendencias** en un conjunto de datos o **hacia dónde se inclina o agrupa más la información.**



EL PROMEDIO O MEDIA ARITMÉTICA

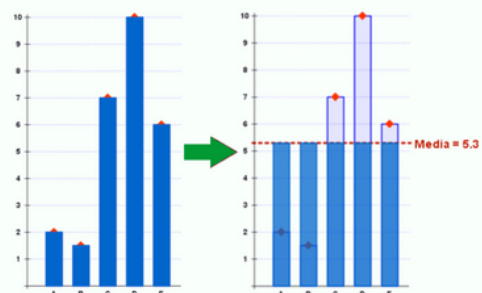
¿Qué es?

Es un **concepto matemático** usado en estadística. También llamada promedio que se obtiene sumando todos los datos y dividiéndolos por el total de datos. Como el resultado se obtiene a partir de la totalidad de los datos disponibles, no es necesario que estén ordenados o agrupados.



La **media aritmética se destaca** por representar el reparto equitativo, el equilibrio, la equidad. Es el valor que tendrían los datos, si todos ellos fueran iguales. O también, el valor que correspondería a cada uno de los datos de la distribución si su suma total se repartiera por igual.

En este ejemplo podemos ver cómo las observaciones mayores se equilibran con las observaciones menores observando cómo actúa la media aritmética como punto de equilibrio.



¿Cómo se calcula?

Cuando se tiene gran cantidad de datos se usa **una pequeña fórmula** para calcular el promedio o media aritmética.

$$\sum_{i=1}^k \frac{x_i * f_i}{n} = \bar{x}$$

Este cálculo también se puede hacer utilizando las herramientas que nos facilitan las hojas o planillas de cálculo. Para recordar estas funcionalidades, pueden recorrer nuevamente el Tutorial Hojas y Planillas de Cálculo.



LA MEDIANA

¿Qué es?

La mediana es otra medida de tendencia central. **Es el valor central del conjunto de datos ordenados de menor a mayor que deja la misma cantidad de datos a cada lado.**

¿Cómo se calcula?

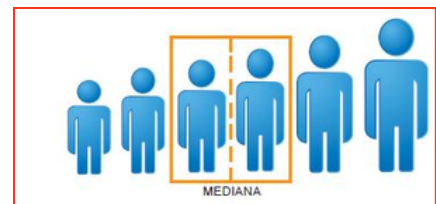
Para hacer este cálculo se puede proceder de diferentes formas según la cantidad de datos con los que contamos. Estos cálculos pueden realizarse de forma manual o si cantidad de datos es demasiada podemos ayudarnos con la fórmula que incluyen las planillas de cálculo.

Si la cantidad **de datos es chica**, es posible obtener la mediana ordenando los datos de menor a mayor. La mediana será el dato que deje a izquierda y derecha de él la misma cantidad de datos.

Si el número de datos es **impar**, la mediana es el valor de la variable que está en la posición $(n+1)/2$.



Si el número de datos es **impar**, la mediana es el valor de la variable que está en la posición $(n+1)/2$.



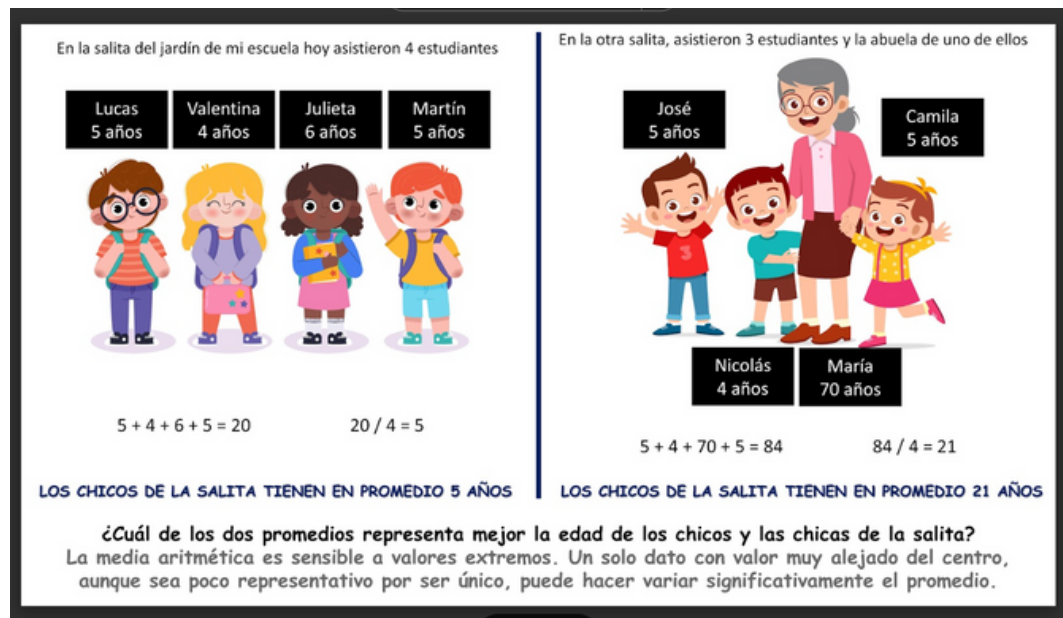


IMPORTANTE

El valor de la media aritmética se puede ver muy afectado por la presencia de un valor extremo, aunque la mediana no es tan sensible a esto.

Un valor extremo es un valor que está muy alejado de la gran mayoría de los datos.

Un ejemplo:



De este modo, una **desventaja** de la media aritmética es que se ve afectada en gran medida por la presencia de valores atípicos o raros, aumentando o disminuyendo su valor según sean las características del mismo.



LA MODA

¿Qué es?

Es el número que más se repite.

¿Cómo se calcula?

Compartimos un ejemplo para datos sin agrupar.

8-14-9-12-15-11-12-8-10-14-12-9-12

Observarvamos la lista de datos e identificamos la cifra que más aparece.

Si seleccionaron el 12 es correcto

La **moda es 12**, porque se repite 5 veces, mientras que el catorce, que es otra cifra que aparece mucho, solo está tres veces.



Si hubiera dos datos que se repiten la misma cantidad de veces, se denomina **bimodal**.

Resumen para calcular la media, la mediana y la moda

- 👉 Organizar el conjuntos de datos.
- 👉 Para calcular **la media** se debe sumar todos los datos y dividirlos por la cantidad de datos que tengas.
- 👉 Para **la mediana**, se ordenan los datos de menor a mayor o viceversa y se elige justo el número que se encuentra en la mitad de todos.
- 👉 Para **la moda**, descubre el dato que más se repite.



4

**¿Cómo hacer
histogramas en
Geogebra?**

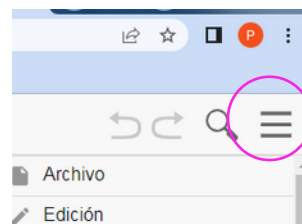
Para realizar **los histogramas** y que no sean planilla de cálculo de libro office calc o excel **se puede utilizar geogebra.**

Ingresen

GeoGebra Classic
Free online apps bundle from GeoGebra: get graphing, geometry, algebra, 3D, statistics, probability, all in one tool!
[geogebra.org](https://www.geogebra.org)



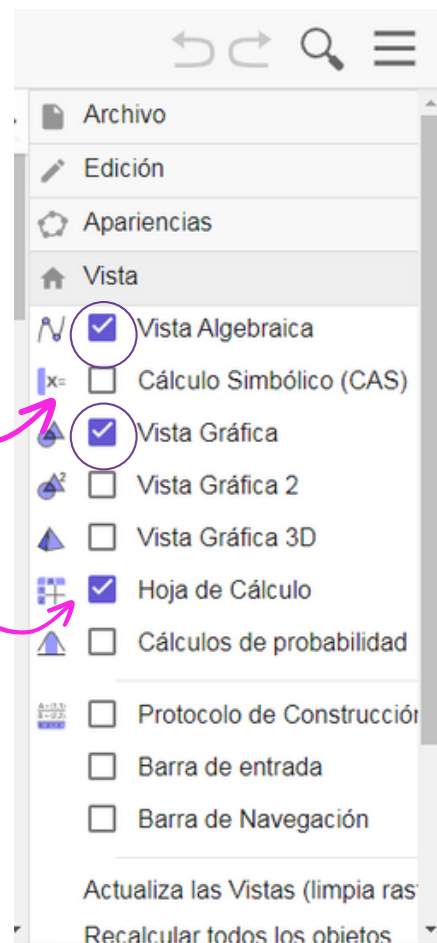
Primero: Desplieguen el menú, en el margen derecho arriba



Se visualizan estas **opciones:**

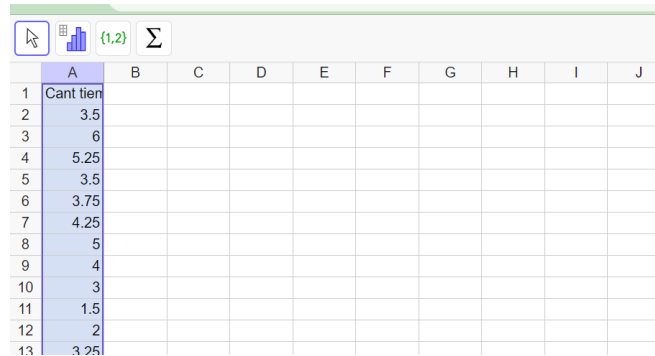
Destilden: **"Vista Gráfica"** y **"Vista Algebraica"**

Tilden **"Hoja de Cálculo"**



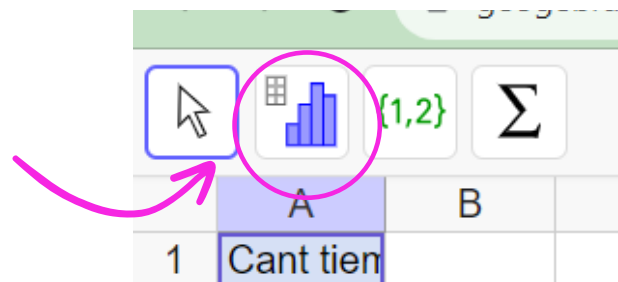
Segundo: Copien la columna de datos "Cantidad de tiempo por día que pasan mirando una película o serie"

Tercero: Peguen en la primera columna de la hoja de cálculo de GeoGebra.

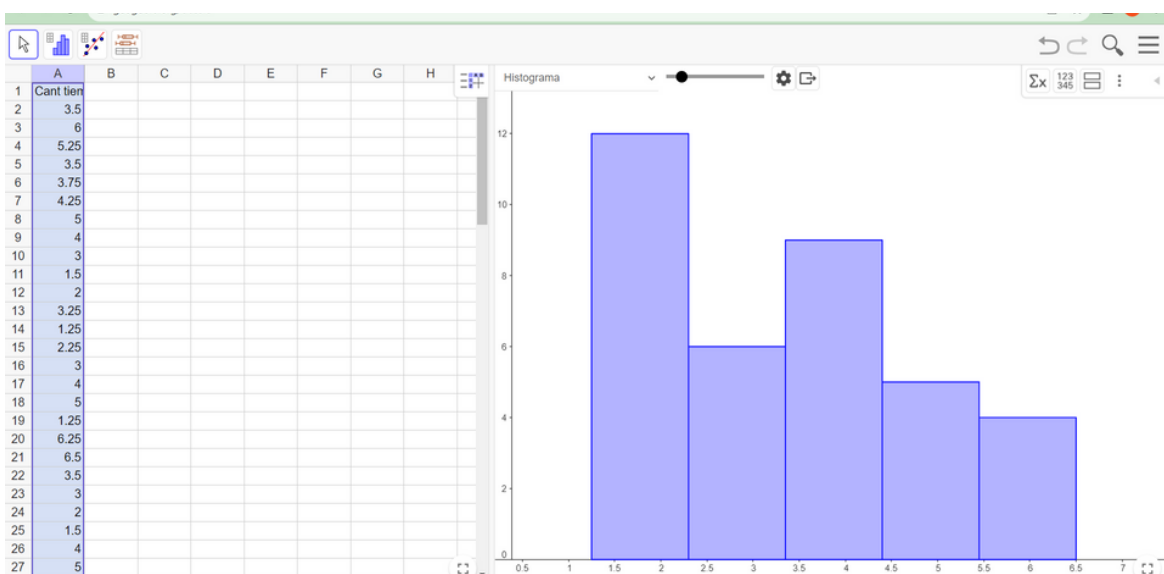


	A	B	C	D	E	F	G	H	I	J
1	Cant tien									
2	3.5									
3	6									
4	5.25									
5	3.5									
6	3.75									
7	4.25									
8	5									
9	4									
10	3									
11	1.5									
12	2									
13	3.25									

Cuarto: Seleccionen el ícono del gráfico, en "Análisis de una variable"



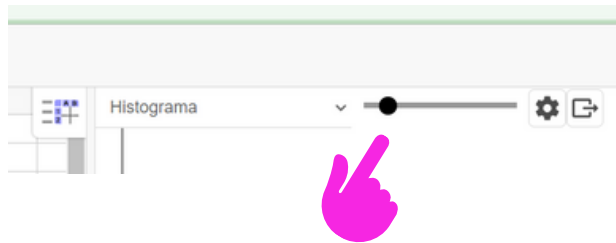
Aparecerá la imagen gráfica y se crea un histograma:



Es decir que **GeoGebra** arma de forma **automática** los intervalos de clases.



Aunque sea automático, **se pueden modificar** con la **función deslizable** y cambiar la amplitud del intervalo.

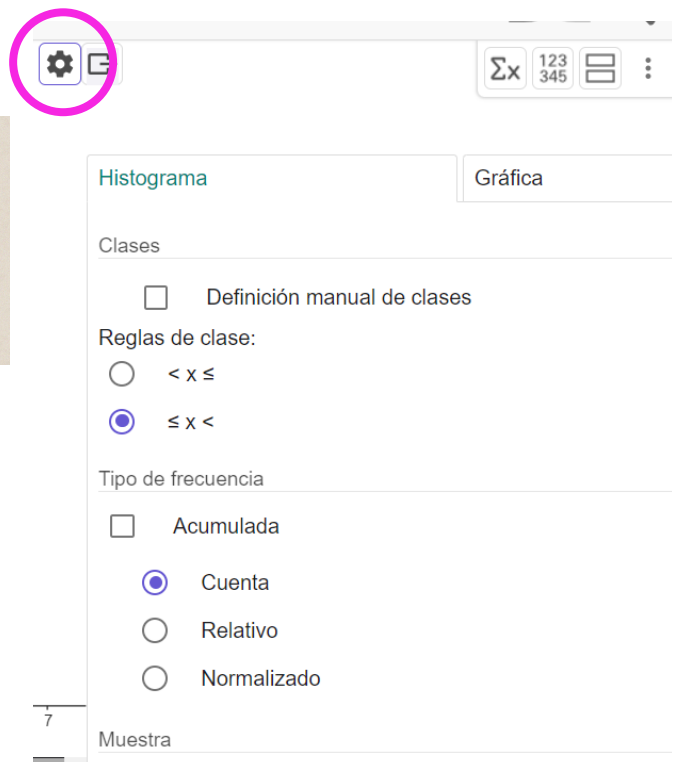


Mové el deslizador y cambia la amplitud.



También se puede cambiar la **configuración**. Haciendo clic en el **botón de la ruedita** se despliega la siguiente ventana.

Con distintas opciones, las cuales se pueden ir *tildando* o *destildando*.



Dentro de las opciones podemos ver la tabla de frecuencia.

Seleccionen **"Tabla de frecuencia"**

Histograma Gráfica

Clases

Definición manual de clases

Reglas de clase:

$< x \leq$

$\leq x <$

Tipo de frecuencia

Acumulada

Cuenta

Relativo

Normalizado

Muestra

Histograma

Tabla de frecuencias

Aparece de este modo:

Intervalo	Frecuencia
1.25 - 2.3	12
2.3 - 3.35	6
3.35 - 4.4	9
4.4 - 5.45	5
5.45 - 6.5	4



En el caso que quisieramos que GeoGebra, nos haga **otros cálculos**, como los de medida de tendencia central (por ejemplo), deben clicar en "mostrar estadísticas" y aparece:

Σx 123 345

Estadísticas	
n	36
Media	3.3958
σ	1.5594
s	1.5816
Σx	122.25
Σx^2	502.6875
Mín	1.25
Q1	2
Mediana	3.375
Q3	4.625
Máx	6.5



CONSEJO GENERAL DE EDUCACIÓN
Gobierno de Entre Ríos

@prender

→ EL PORTAL DEL SISTEMA
EDUCATIVO DE **ENTRE RÍOS**

Algunas **orientaciones, fórmulas y tips** para
usar en la resolución de ejercicios de
Estadística Descriptiva

