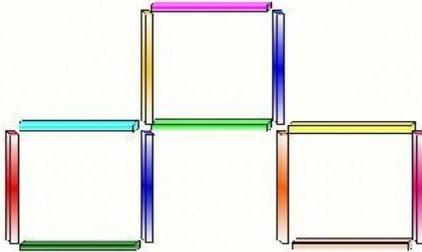


	INSTITUCIÓN EDUCATIVA JUAN DE DIOS CARVAJAL	Código FP 67
	FORMATO PARA ELABORACIÓN DE MÓDULOS	01-07-2020

MÓDULO No 2 FECHA: DE MARZO 15 AL 07 DE MAYO DE 2021			
Área o Asignatura: Estadística	Grados: 10.1 y 10.2	Intensidad Semanal:1 hora	Periodo: Primero
Docente(s): Paola Andrea Arenas Perez			
Fecha de entrega:22 de abril			
Contacto del/los docentes(s): WhatsApp 3168864320, paolaarenas@iejuandedioscarvajal.edu.co			
Horarios y forma de atención del/los docentes(s): WhatsApp desde las 8 am -4 pm de lunes a jueves, a través del correo institucional.			
Competencias a trabajar en este modulo			
Interpreta datos y hace uso crítico de la información.			
Interpreta gráficas y analiza la realidad.			
Introducción:			
<p>Los orígenes de la estadística se confunden con los de la humanidad. Sólo en tiempos recientes ha adquirido esta ciencia la categoría de disciplina relevante y de importancia práctica (Gutiérrez, 1994). Podría decirse que la historia de la estadística comienza, posiblemente, en la isla de Cerdeña donde existen algunos monumentos prehistóricos en cuyas piedras se encontraron ciertas muescas que permiten suponer un conteo del ganado cazado. La existencia de la estadística se debe a las necesidades de las sociedades y sus gobiernos de conocer hechos políticos, sociales y económicos en relación con el dominio, la riqueza y la producción (Carrasco, 2010).</p>			
<p>CONSTRUYE ESTA FIGURA DE 3 CUADRADOS</p>			
			
<p>MOVIENDO 3 PALILLOS HAS DE FORMAR 7 CUADRADOS</p>			

DISTRIBUCIONES DE FRECUENCIA AGRUPADA¹

Cuando se dispone de gran número de datos, es útil el distribuirlos en clases o categorías y determinar el número de individuos pertenecientes a cada clase, que es la frecuencia de clase. Una ordenación tabular de los datos en clases, reunidas las clases y con las frecuencias correspondientes a cada una, se conoce como una **Distribución de frecuencias o Tabla de frecuencias**.

La tabla de distribución de frecuencias se hace agrupando el conjunto de datos numéricos en clases o intervalos apropiados. Este procedimiento lo explicaremos mediante el desarrollo del siguiente ejercicio.

Los siguientes datos se recopilaron con el fin de determinar la edad de 50 estudiantes del grado 10^o de la sección nocturna de la IE Carvajal. Así los datos obtenidos fueron los siguientes:

15 16 17 18 19 15 20 18 20 17
 15 16 15 19 25 15 30 42 15 20
 15 16 19 20 16 15 16 20 20 42
 16 17 17 20 19 18 19 60 42 22
 19 19 25 17 25 31 20 25 30 42

Organizando este conjunto de datos en forma ascendente y haciendo el correspondiente recuento (número de veces que se repite cada valor) obtenemos el registro indicado a continuación. Observemos que el valor máximo es 60 y el menor es 15.

Edades de 50 estudiantes del grado 10^o de IE Carvajal sección nocturna.

Edad	Nº de veces que se repite
15 →	8
16 →	6
17 →	5
18 →	3
19 →	7
20 →	8
22 →	1
25 →	4
30 →	2
31 →	1
42 →	4
60 →	1

¹ <https://www.webcolegios.com/file/a73153.pdf>

En la anterior ordenación la variable X toma muchos valores diferentes y algunos de ellos tienen una frecuencia tan pequeña que no se justifica considerarlos por separado. Además, no podemos visualizar claramente las medidas de tendencia central y su cálculo se dificulta por la cantidad de operaciones que deben realizarse. Por tanto, es necesario agrupar los datos en clases o intervalos.

Para agrupar este conjunto en clases o intervalos de datos debemos seguir los siguientes pasos:

Para agrupar este conjunto en clases o intervalos de datos debemos seguir los siguientes pasos:

1er paso: Rango o recorrido

Calculamos el rango o recorrido que representamos por R y que es la diferencia entre el $X_{\text{máx}}$ y el $X_{\text{mín}}$.

$$R = X_{\text{máx}} - X_{\text{mín}}$$

$$R = 60 - 15 = 45$$

Un rango de 45 años significa que la diferencia entre la mayor edad y la menor es 45 años.

2º Paso: Elección del número de intervalos

La *Regla de Sturges* nos propone que dadas N observaciones, el número k de intervalos viene dado por

$$k = 1 + 3.3 \log N$$

Donde N es el número de muestras observadas y k el número óptimo de clases o intervalos.

$$K = 1 + 3,3 \log 50 \Rightarrow k = 1 + 3,3(1,69) \Rightarrow k = 1 + 5,577 = 6,577 \approx 7 \text{ intervalos}$$

3er Paso: Amplitud de intervalos o clases

Repartimos el rango en clases o intervalos de la misma longitud o amplitud. Si a representa la amplitud de cada intervalo, entonces:

$$a = \frac{R}{k} \quad \rightarrow \quad a = \frac{45}{7} = 6,4$$

Cuando los datos sean valores enteros de la variable, entonces el cociente $\frac{R}{k}$ debe ser un número

entero. Si no ocurre que $\frac{R}{k}$ es entero, como en nuestro ejemplo ($\frac{R}{k} = 6,4$), debemos aproximar a al número entero más próximo por encima, es decir $a = 7$.

Si los datos de la muestra tienen cifras decimales, entonces debemos tomar una amplitud que tenga el mismo número de cifras decimales.

Así, si en un ejercicio encontramos que $\frac{R}{k} = 2,859$, entonces tomamos $a = 2,86$ si los datos tienen dos cifras decimales; $a = 2,9$ si los datos tienen una cifra decimal y $a = 3$ si los datos tienen son números enteros.

4º Paso: Límite de intervalos

Si $K = 7$ y $a = 7$, entonces el rango que vamos a repartir ya no es $R = 45$ sino $7 \times 7 = 49$. Este nuevo rango se representa por R_a y se llama rango ampliado.

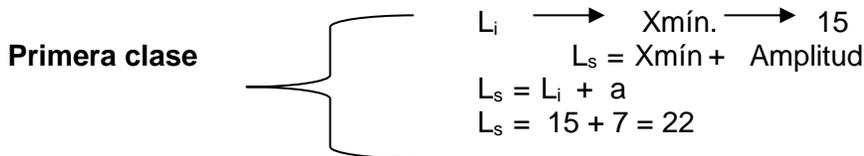
Si $R_a - R$ es la cantidad en que se amplía el rango, entonces en esta misma cantidad se debe ampliar el $X_{\text{máx}}$ o disminuir el $X_{\text{mín}}$ (o ambos) para que se cumpla:

$$R_a = X_{\text{máx}} - X_{\text{mín}}$$

En nuestro ejemplo: $49 = 64 - 15$ ó $49 = 60 - 11$ ó $49 = 62 - 13$

Si aumentamos el $X_{\text{máx}}$ en 4, entonces el $X_{\text{máx}} = 64$ es el límite superior del último intervalo.

Si al límite inferior $L_i = 15$ del primer intervalo se le adiciona la amplitud $a = 7$, el resultado $L_s = 15 + 7 = 22$ es el límite superior del primer intervalo, Así:



La primera clase está formada por todos los valores de x entre 15 y 22 años.

La segunda clase tiene como límite inferior el límite superior de la primera clase y como límite superior el inferior aumentado en la amplitud.

Segunda clase 22 - 29

Este procedimiento se repite hasta obtener un número ($K = 7$) de intervalos ya establecido que tiene a 64 como límite superior del último intervalo. (Ver tabla).

Clase	Intervalo	Marca de clase: X_i	Frecuencia absoluta: f_i	Frecuencia absoluta acumulada: F_i	Frecuencia relativa porcentual $\frac{f_i}{n} \cdot 100 = \%$	Frecuencia porcentual acumulada %	$X_i \cdot f_i$
1º	15 - 22	18,5	37	37	0,74=74%	74%	684.5
2º	22 - 29	25,5	5	42	0,10=10%	84%	127.5
3º	29 - 36	32,5	3	45	0,06=6%	90%	97.5
4º	36 - 43	39,5	4	49	0,08=8%	98%	158
5º	43 - 50	46,5	0	49	0,00=0%	98%	0
6º	50 - 57	53,5	0	49	0,00=0%	98%	0
7º	57 - 64	60,5	1	50	0,02=2%	100%	60.5
	Total		n=50		1,00=100%		

Cuando el cociente $\frac{R}{k}$ es exacto y no hay necesidad de ampliar el rango, tanto el límite inferior del primer intervalo como el límite superior del último intervalo coinciden con los $x_{\text{mín.}}$ y $x_{\text{máx.}}$

5º paso: Marcas de clase: Como en cada intervalo podemos considerar infinitos valores reales de la variable x , debemos tomar uno de ellos que nos represente la clase y nos permita hacer gráficas y cálculos (como la media aritmética). A cada uno de estos valores se le llama **marca de clase** y su mejor representante es el punto medio del intervalo (o valor central).

Si x_i representa el punto medio del intervalo i -ésimo, entonces la primera marca de clase es:

$$x_1 = \frac{15 + 22}{2} = 18,5$$

Las otras marcas de clase se pueden obtener en forma similar, o sumando a la anterior la amplitud:

$$x_2 = 18,5 + 7 = 25,5$$

$$x_3 = 25,5 + 7 = 32,5$$

$$x_4 = 32,5 + 7 = 39,5$$

⋮

$$x_7 = 53,5 + 7 = 60,5$$

MEDIDAS DE TENDENCIA CENTRAL DE DATOS AGRUPADOS

Permiten un mejor análisis de los datos estadísticos. Las medidas de tendencia central son: Moda, Mediana y Media.

Aclaremos a través de un ejercicio cada uno de estos conceptos.

LA MEDIA: Se calcula sumando todos los productos de marca clase con la frecuencia absoluta respectiva y su resultado dividirlo por el número total de datos:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{n} = \frac{\sum_{i=1}^n x_i \cdot f_i}{n}$$

LA MODA: Para datos agrupados se encuentra en la mayor frecuencia absoluta de los intervalos, asimismo se puede determinar en el gráfico de barras pues la barra más alta equivale a la moda.

LA MEDIANA: en datos agrupados equivale al dato que se encuentra en toda la mitad, y para encontrarlo se necesita el método estudiado anteriormente para datos no agrupados; y luego se requiere de aplicar la siguiente fórmula.

$$Me = L_i + \frac{a \left[\frac{n}{2} - F_{i-1} \right]}{f_i} \quad \text{donde}$$

Me = Mediana

L_i = Límite inferior del intervalo donde se encuentra la mediana

a = amplitud

n = Número de datos

F_{i-1} = frecuencia acumulada anterior al intervalo que contiene la mediana

f_i = La frecuencia absoluta que corresponde al intervalo que contiene la mediana

Ejemplo: La tabla de datos correspondiente a las edades de 50 alumnos de la I.E. Carvajal sección nocturna está dada a continuación:

Clase	Intervalo	Marca de clase: X_i	Frecuencia absoluta: f_i	Frecuencia absoluta acumulada: F_i	Frecuencia relativa porcentual $\frac{f_i}{n} \cdot 100 = \%$	Frecuencia porcentual acumulada %	$X_i \cdot f_i$
1º	15 - 22	18,5	37	37	0,74=74%	74%	684.5
2º	22 - 29	25,5	5	42	0,10=10%	84%	127.5
3º	29 - 36	32,5	3	45	0,06=6%	90%	97.5
4º	36 - 43	39,5	4	49	0,08=8%	98%	158
5º	43 - 50	46,5	0	49	0,00=0%	98%	0
6º	50 - 57	53,5	0	49	0,00=0%	98%	0
7º	57 - 64	60,5	1	50	0,02=2%	100%	60.5
	Total		n=50		1,00=100%		

La moda es la edad de 18,5 años que es la marca de clase que corresponde a la mayor frecuencia absoluta 37

La mediana es la edad que se encuentra en el centro de todas las edades para calcularla utilizaremos la fórmula

$$Me = L_i + \frac{a \left[\frac{n}{2} - F_{i-1} \right]}{f_i}$$

Donde

M_e = Mediana, la cual se encuentra ubicada en el primer intervalo pues $\frac{n}{2} = \frac{50}{2} = 25$ y este dato se encuentra en la frecuencia acumulada del primer intervalo.

L_i = Límite inferior del intervalo que contiene la mediana, es decir, 15

a = amplitud = 7 porque $22 - 15 = 7$

n = Número de datos = 50

F_{i-1} = frecuencia acumulada anterior al intervalo que contiene la mediana = 0 porque antes de este primer intervalo no hay datos acumulados.

f_i = La frecuencia absoluta que corresponde al intervalo que contiene la mediana = 37

$$Me = 15 + \frac{7 \left[\frac{50}{2} - 0 \right]}{37}$$

$$Me = 15 + \frac{7[25]}{37}$$

Por tanto: $Me = 15 + \frac{175}{37}$

$$Me = 15 + 4,73$$

$$Me = 19,73$$

Lo cual significa que la **mediana** o la edad que se encuentra en el centro de los datos recolectados es 19.

HISTOGRAMA DE FRECUENCIAS²

Un histograma es una representación gráfica de una variable en forma de barras, teniendo en cuenta que la superficie de cada barra es proporcional a la frecuencia de los valores representados. Un histograma nos permite ver cómo se distribuyen los valores de la variable en estudio.

Usamos los histogramas cuando analizamos [variables continuas](#), o cuando trabajamos con variables discretas que toman un gran número de valores y son agrupadas en intervalos. Cuando tenemos variables cualitativas, se emplean los [diagramas de barras](#).

¿Cómo construir un histograma?

Partimos de una [tabla de frecuencias con datos agrupados](#), y seguimos los siguientes pasos:

1. En el eje horizontal (X), colocamos los límites de clase. Opcionalmente, puedes colocar las marcas de clase.
2. En el eje vertical (Y), colocamos las frecuencias. Se suele tomar la frecuencia absoluta, pero también se puede trabajar con la frecuencia relativa o con la frecuencia porcentual.
3. Dibujamos las barras de cada clase, teniendo en cuenta que la altura de cada barra es igual a la frecuencia.

Ejemplo 1

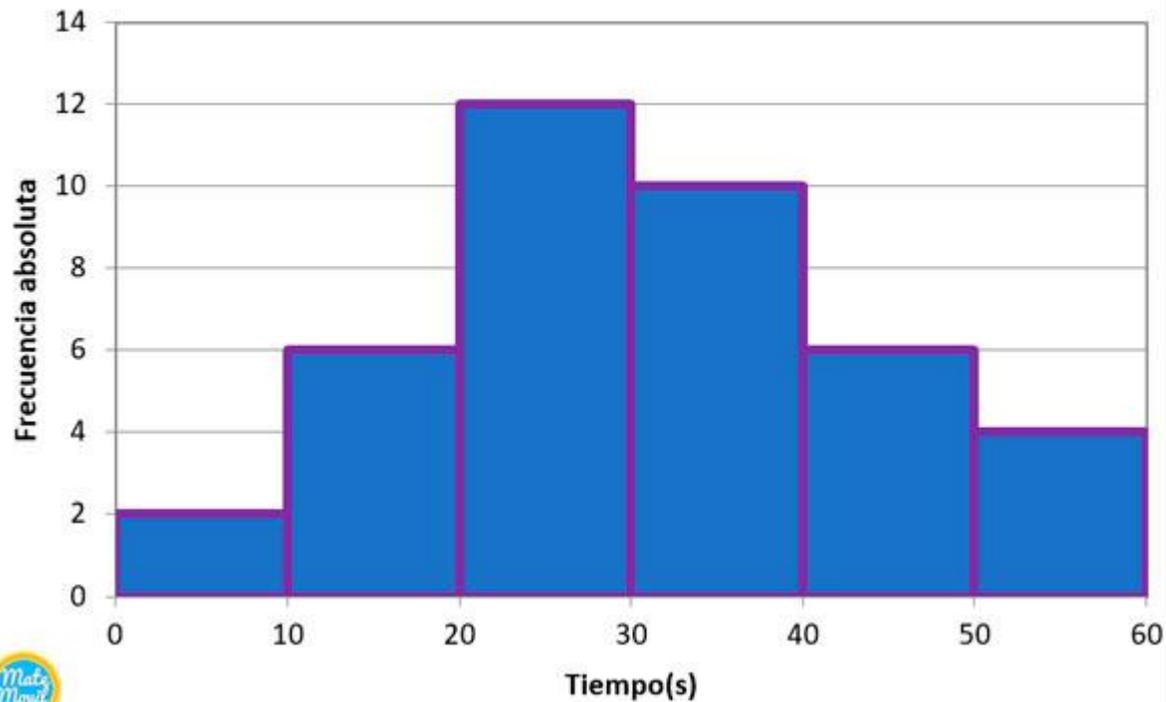
Se registran los tiempos de las llamadas recibidas en un call center, y se obtiene la siguiente tabla de frecuencias con datos agrupados. Construir un histograma de frecuencias.

Tiempo de llamadas	Marcas de clase	Frecuencia absoluta	Frecuencia acumulada	Frecuencia porcentual
[0 - 10)	5	2	2	5%
[10 - 20)	15	6	8	15%
[20 - 30)	25	12	20	30%
[30 - 40)	35	10	30	25%
[40 - 50)	45	6	36	15%
[50 - 60]	55	4	40	10%
Total		40		100%

Solución:

² <https://matemovil.com/histogramas-ejemplos-y-ejercicios/>

Histograma de tiempos de llamadas



Recuerda que si vas a trabajar con una variable cualitativa o variable discreta que asume pocos valores, deberás usar un [diagrama de barras](#) y no un histograma.

Polígono de frecuencias

Es un gráfico que se forma uniendo los puntos medios de la parte superior de las barras mediante segmentos de recta. El polígono de frecuencias es de mucha utilidad cuando se representa más de una serie en una misma gráfica.

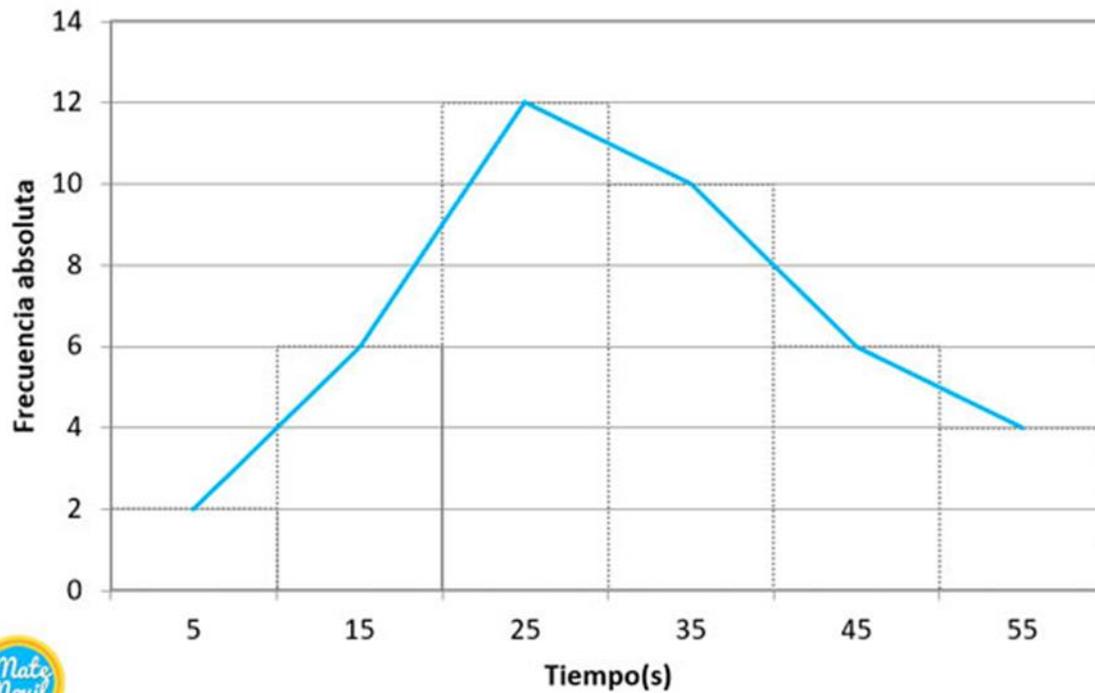
Los polígonos de frecuencias se trazan tomando en cuenta las marcas de clase de cada barra.

Ejemplo 2

A partir del histograma del ejemplo anterior, construir el polígono de frecuencias.

Solución:

Polígono de frecuencias de tiempos de llamadas



Taller:

Ejercicios:

1. La tabla dada a continuación muestra la información sobre el número de casos de urgencias atendidos diariamente en un hospital durante un trimestre. Hallar la moda, mediana y media aritmética de la demanda del servicio de urgencias en ese hospital.

X_i	f_i	F_i	% Acumulado	$X_i \cdot f_i$
15	3	3	3,33	45
18	4	7	7,78	72
19	10	17	18,89	190
21	16	33	36,67	336
22	12	45	50,00	264
25	12	57	63,33	300
28	16	73	81,11	448
31	8	81	90,00	248
35	7	88	97,78	245
40	2	90	100,00	80
Total	N = 90			2228

2. A una reunión asisten 6 personas con edades de 15, 16, 18, 20, 12 y 14 años. ¿Cuál es la media aritmética? ¿Cuál es la mediana? ¿Cuál de estos valores es más representativo? ¿Por qué?

El tiempo en segundos registrado por un grupo de 40 atletas en los 100 metros planos, presenta el siguiente conjunto de datos estadísticos numéricos:

2. Los tiempos en minutos empleados por un grupo de atletas en recorrer 15 Km. Están representados en la siguiente tabla. Calcula el tiempo promedio empleado por los atletas.

Tiempo X_i	Frecuencia Absoluta f_i	$X_i \cdot f_i$
120	2	
130	5	
135	4	
180	7	
200	10	
215	8	
230	4	

- Calcula la mediana y la moda en los ejercicios anteriores.
- Calcula la mediana de los números: 15, 6, 3, 8, 10.
- Calcula la mediana de los números: 3, 6, 7, 10, 15, 18.

3. Con los datos dados a continuación llene la tabla de frecuencias.

Las notas obtenidas por los alumnos de 10^o grado en estadística fueron:
 3 alumnos obtuvieron 30; 6 alumnos obtuvieron 40; 9 alumnos obtuvieron 50; 10 alumnos obtuvieron 60; 7 obtuvieron 70; 5 obtuvieron 80, 2 obtuvieron 90; 3 obtuvieron 100.

- ❖ Con los datos anteriores, completa la tabla.
- ❖ Calcula la media aritmética o nota promedio obtenida por los alumnos
- ❖ Halla la Moda y la Mediana.

Notas = X_i	Frecuencia Absoluta = f_i	Frecuencia Absoluta Acumulada = F_i	Frecuencia Relativa o Porcentual = %	Frecuencia Porcentual Acumulada	$X_i \cdot f_i$
30					
40					
50					
60					
70					
80					
90					
100		45		100%	
Total	45		1,00 = 100%		

Promedio = $\bar{x} = \frac{\quad}{45}$

Moda =

Mediana =

Ejercicio: La siguiente tabla muestra las 178 calificaciones obtenidas por los estudiantes del grado décimo en la prueba final de matemáticas (escala de 1 a 10).

Intervalos	Marca de clase: X_i	Frecuencia absoluta: f_i	Frecuencia absoluta acumulada: F_i	Frecuencia porcentual acumulada %	$X_i \cdot f_i$
0,5-1,5		3			
1,5-2,5		7			
2,5-3,5		20			
3,5-4,5		25			
4,5-5,5		30			
5,5-6,5		40			
6,5-7,5		22			
7,5-8,5		20			
8,5-9,5		8			
9,5-10,5		3			

1. Completar la tabla de distribución de frecuencias
2. Construir un histograma, un polígono de frecuencias y una ojiva
3. Hallar la moda, mediana y media.

Evaluación:

Solucionar todo el taller del módulo en los respectivos cuadernos y a mano, colocando un título alusivo al principio de la solución y numerando cada punto. Desarrolle cada punto con la estructura ejercicio – solución – respuesta. Se envían todo con sus debidos procedimientos.

Usted deberá presentar las fotografías o escaneos de dicho trabajo, para lo cual usted marcará cada archivo o página de forma consecutiva al orden de los ejercicios del taller. NOTA: Si le queda fácil presentar el archivo en PDF, sería lo más adecuado, mas no es una obligación. (Se puede usar camscanner)

PROCESO DE AUTOEVALUACIÓN DE TRABAJO CON MÓDULOS					
EJE A EVALUAR	CRITERIO DE EVALUACIÓN	SUPERIOR	ALTO	BÁSICO	BAJO
Responsabilidad e Interés	Leí toda la información del módulo.				
	Realicé las actividades teniendo en cuenta un horario				
	Realicé las actividades completas y en el tiempo establecido				
	Seguí atentamente las instrucciones de cada actividad				
	Hice preguntas a un adulto cuando no comprendí algo				
Académico	Estudie todos los temas propuestos				
	Desarrolle cada una de las actividades propuestas				
	Hice uso de la estética para la realización de trabajos				
Sobre el módulo	El contenido fue interesante y entretenido				
	Los temas incluidos me aportaron como persona				
	El contenido fue de fácil comprensión				
Mi opinión sobre el trabajo realizado y sugerencias:					

Bibliografía:

¹ <https://www.webcolegios.com/file/a73153.pdf>

² <https://matemovil.com/histogramas-ejemplos-y-ejercicios/>