

ESTADÍSTICA DESCRIPTIVA CON EXCEL I

Óscar Machado Álvarez

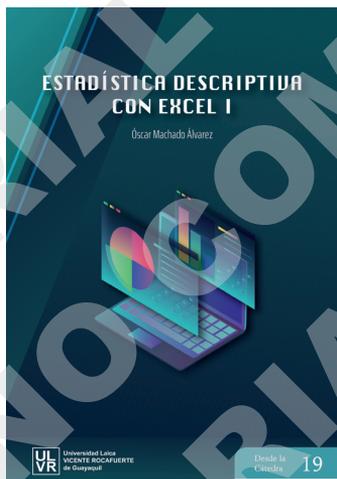


Universidad Laica
VICENTE ROCAFUERTE
de Guayaquil

Desde la
Cátedra

19





EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

Óscar Machado Álvarez

Estadística Descriptiva con Excel 1



Estadística Descriptiva con Excel 1.

Mg. Óscar Paul Machado Álvarez

El autor ejerce la calidad de docentes de la Universidad Laica VICENTE ROCAFUERTE de Guayaquil.

De esta edición:

Universidad Laica VICENTE ROCAFUERTE de Guayaquil, ULVR © 2021

Av. de las Américas #70, frente al Cuartel Modelo

Guayaquil, Ecuador

PBX: (00-593-4) 259-6500

www.ulvr.edu.ec



edilaica@ulvr.edu.ec

Av. de las Américas #70, frente al Cuartel Modelo

Guayaquil, Ecuador

PBX: (00-593-4) 259-6500, extensión 195

Estadística Descriptiva con Excel 1.

Primera Edición: 11 de febrero de 2021

ISBN: 978-9942-920-84-3

Derecho de autor: GYE-011900

Depósito Legal: GYE-000306

Tiraje: 100 ejemplares

**Desde la
Cátedra** **19**

Clasificación JEL:

C Métodos matemáticos y cuantitativos.

C4 Métodos econométricos y estadísticos.

C40 Generalidades.

Palabras clave:

Estadística, Matemáticas estadísticas, Inferencia estadística, Análisis estadístico.

Key words:

Statistics, Mathematical statistics, Statistical inference, Statistical analysis.

Diseño y diagramación: Econ. Patricia Navarrete Zavala / Inavarretez@ulvr.edu.ec

Portada: Dis. Gráf. Andrés Avilés Zavala / aavileszav@ulvr.edu.ec

El contenido de este libro puede ser utilizado, citando la fuente, de acuerdo a las Normas APA 7a. edición:

Machado, Ó. (2021). *Estadística Descriptiva con Excel 1* [Desde la Cátedra 19]. Editorial ULVR.

Consejo Editorial de la Universidad Laica VICENTE ROCAFUERTE de Guayaquil

Ph.D Aimara Rodríguez Fernández, *Rectora*

Ph.D Sonia Guerra Iglesias, *Vicerrectora Académica de Investigación, Grado y Posgrado*

Ph.D Rolando Villavicencio Santillán, *Vicerrector Administrativo*

Ph.D Rafael Iturralde Solórzano, *Decano de la Facultad de Administración*

Mg. Diana Almeida Aguilera, *Decana de la Facultad de Ciencias Sociales y Derecho*

Mg. Georgina Hinojosa Dazza, *Decana de la Facultad de Educación*

Mg. Alex Salvatierra Espinoza, *Decano de la Facultad de Ingeniería, Industria y Construcción*

Ing. Com. Alfredo Aguilar Hinojosa, *Director del Dpto. MKT y Relaciones Públicas*

Econ. Patricia Navarrete Zavala, *Coordinadora de la Editorial ULVR*

Queda rigurosamente prohibido, sin la autorización escrita de los titulares del Copyright, bajo las sanciones establecidas en leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© DRA imágenes tomadas de Internet.

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Agradecimiento

A Dios, por darme la fuerza para continuar en los momentos difíciles.

A la ULVR y sus autoridades, por permitirme expresar los conocimientos adquiridos en el devenir docente en la presente obra.



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Dedicatoria

A mi esposa Colo, por ser mi compañera en los buenos momentos y mi sostén y guía en los malos.

A mis hijos Emma y Óscar Jr., pues son la alegría de mi corazón y ese motor que me obliga a luchar diariamente.

A mi madre Teresa, quien me enseñó con amor mis primeras letras.

A mi padre William, quien me enseñó el valor del esfuerzo y la honestidad.

A mis hermanos, cuñados y suegros, quienes son parte fundamental de mi día a día.



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

Óscar Machado Álvarez

Estadística Descriptiva con Excel 1



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Presentación

A sabiendas de que la nueva generación de estudiantes y profesionales vive la era de la información a un ritmo vertiginoso, y consiente de la importancia que tiene la optimización del tiempo para los mismos, surgió la idea de elaborar un documento que les ayude a minimizar los tiempos de cálculo estadístico con la ayuda de Microsoft Excel, que es una herramienta que todos conocen pero que no explotan en todo su potencial.

Es por eso que el presente texto trata de explicar de manera sencilla y muy gráfica los cálculos más comunes en el análisis de los fenómenos estadísticos, con la finalidad de que el lector reduzca al mínimo los cálculos y omita errores que el proceso manual puede generar, ahorrando tiempo y alcanzando un análisis minucioso de los resultados finales. Espero que este libro les sea útil en su aprendizaje y que lo disfruten.

Óscar Machado Álvarez

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

Índice

Introducción	19
Antes de empezar	21
Capítulo 1	
Agrupación de Datos	
1.1 Número de clases óptimo	29
Ejercicios propuestos 1-3	32
1.2 Intervalo de clase	33
Ejercicios propuestos 4-12	38
1.3 Límites superiores e inferiores de clase	40
1.3.1 Límites de clase	40
Ejercicios propuestos 13-18	46
1.4 Frecuencia de clase	49
Ejercicios propuestos 19-23	53
1.5 Frecuencia acumulada	55
Ejercicios propuestos 24-26	57
1.6 Frecuencia relativa	58
Ejercicios propuestos 27-32	61
1.7 Frecuencia relativa acumulada	63
Ejercicios propuestos 33-37	65
1.8 Marca de clase	66
Ejercicios propuestos 38-40	71
1.9 Gráficas	73

1.9.1 Histograma de frecuencias	73
1.9.2 Polígono de frecuencias	79
1.9.3 Ojiva	82

Capítulo 2

Medidas de Tendencia Central

2.1 Generalidades	89
2.2 Media Aritmética	91
Ejercicios propuestos 41-43.....	93
2.3 Mediana.....	94
Ejercicios propuestos 44-47	95
2.4 Moda.....	96
Ejercicios propuestos 48-51	98
2.5 Media Geométrica	99
 Bibliografía.....	 102



Introducción

Consiente de los desafíos a los que debe enfrentarse un profesional moderno, además de la importancia de la reducción del tiempo de cálculo en los estudios estadísticos con el fin de priorizar el análisis; decidí realizar este texto para tratar sobre estadística con la ayuda de Microsoft Excel 2013.

Cabe resaltar que el presente trabajo no es un libro de estadística *per se*, ya que para consultarlo el lector debe tener conocimientos previos sobre la materia, más bien es un libro en el cual se encontrará como aplicar dichos conocimientos, con el auxilio de un software, que todo profesional actual está obligado a conocer.

La presente obra abordará la estadística descriptiva utilizando básicamente dos métodos, uno rápido y sencillo para que se tenga una idea general de los fenómenos estudiados, y otro mediante la utilización de fórmulas, funciones y herramientas de análisis que posee Excel como complemento, y que generalmente no se encuentran activados por defecto.

Los ejercicios y ejemplos propuestos serán –en lo posible– resultado de la experiencia del autor en los más de 10 años de docencia en la cátedra de Estadística aplicada a la Administración y Economía, además de los proyectos de investigación previos a la obtención del título de tercer nivel, de quienes quisieron colaborar con la elaboración del presente trabajo.

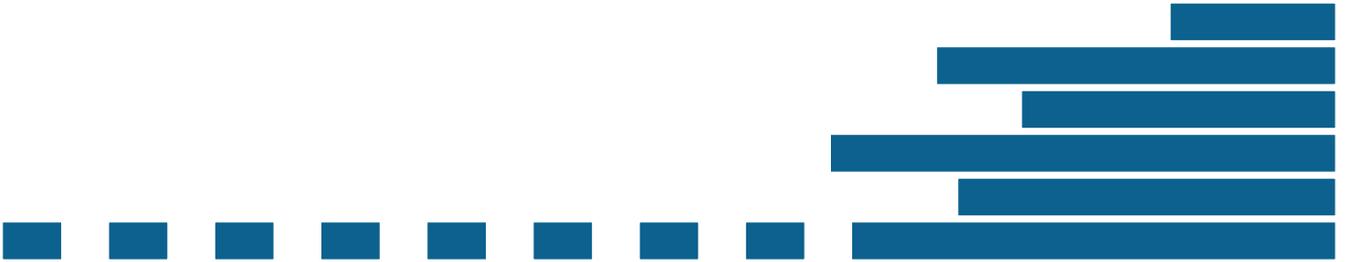
Para finalizar se insistirá en resolver los problemas más comunes a los que se enfrentan los estudiantes al momento de realizar cálculos estadísticos y más aún al momento de interpretarlos.

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Antes de empezar

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

Antes de empezar

Con el fin de estandarizar el conocimiento voy a hacer un par de precisiones:

Aunque versiones anteriores son muy parecidas y las fórmulas y funciones son prácticamente las mismas, se deja en claro que la versión con la que se trabajará este texto es la del 2013.

Se debe configurar la computadora para que reconozca de manera estándar todos los símbolos y comandos que utilizaremos de aquí en más, para ello vaya a Inicio, luego al Panel de Control y ubique las opciones de Reloj, Idioma y Región tal como lo muestra la figura I.



Figura I. Ajustar la configuración del equipo.

De *clik* en la opción sugerida y luego ubique nuevamente la opción Reloj, Idioma y Región y de *clik* en ésta (Figura II).

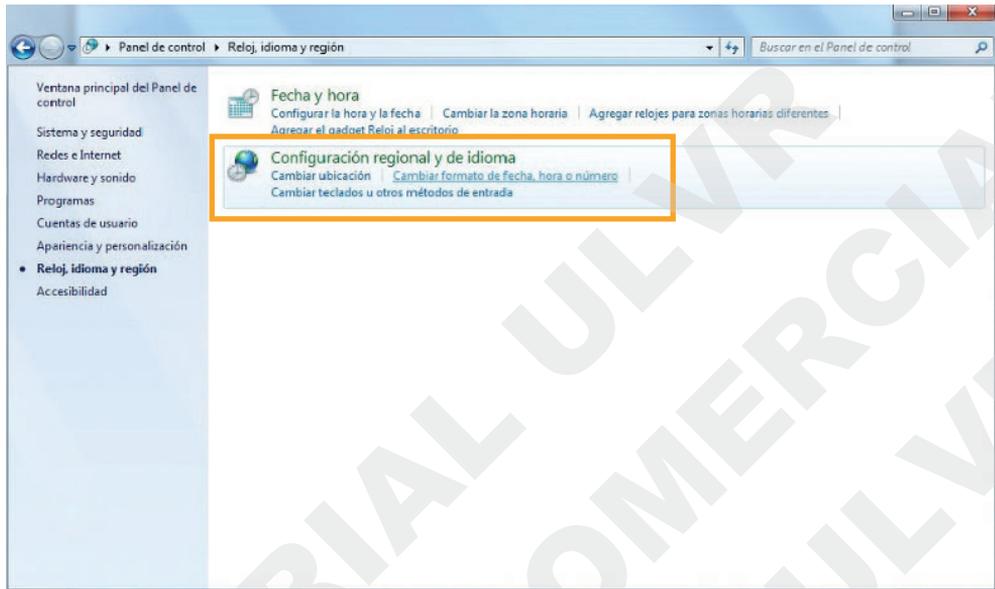


Figura II. Configuración regional y de idioma.

A continuación, de *clik* en **Cambiar formato de fecha, hora o número** y escoja en la ventana que se le abrirá la opción **Configuración adicional** (Figura III).

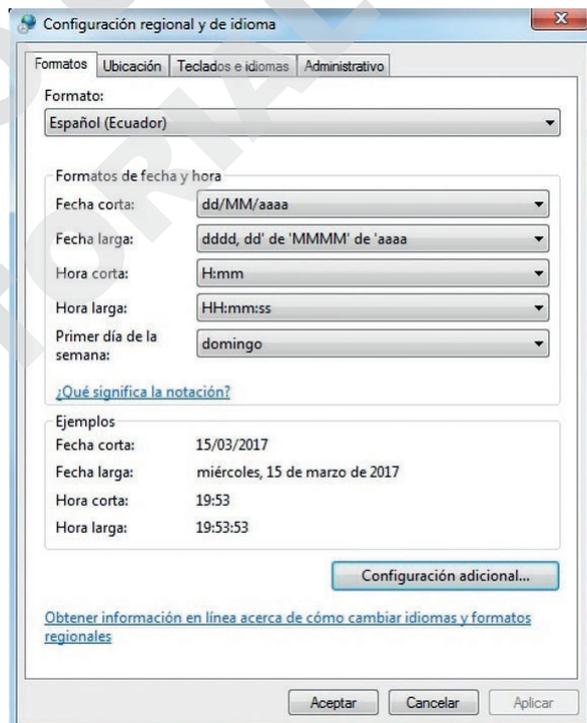


Figura III. Configuración adicional.

Como separador de decimales escoja la coma, y como separador de miles, el punto (Figura IV). Deje las demás opciones por defecto a menos que desee algo en específico y tenga conocimiento sobre configuraciones, ya que podría desconfigurar su máquina haciendo imposible obtener los resultados esperados. Si cree que cometió un error dele *clic* a Restablecer.

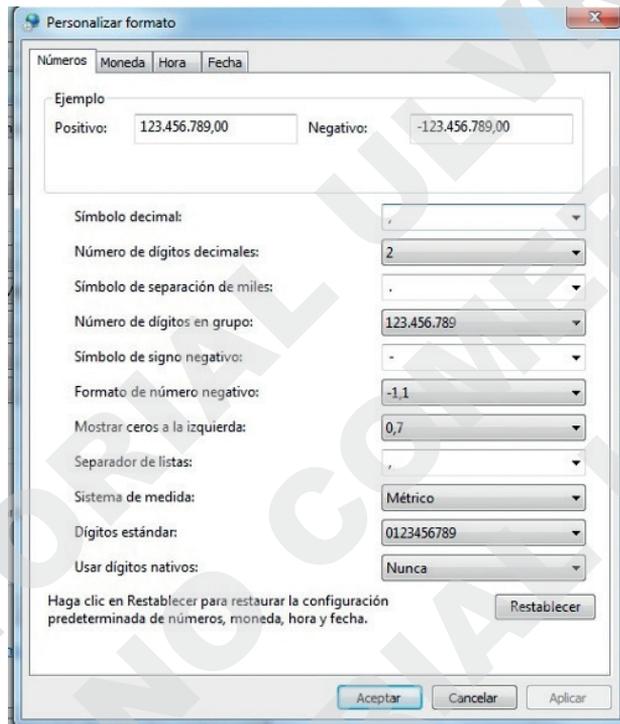


Figura IV. Configuración de separador de decimales y miles.

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Capítulo 1

Agrupación de Datos

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Capítulo 1

Agrupación de Datos

Cuando se decide hacer una investigación a fin de conocer cómo se comporta un fenómeno estadístico, el primer problema al que se enfrenta el investigador es analizar la gran cantidad de información que se recolecta, para ello se necesita organizar la información de manera que sea más fácil el análisis y la interpretación de los resultados. La forma de organizar esta información, cuando las variables son cuantitativas por supuesto, es a través de la utilización de tablas que agrupan los datos. Las tablas a las que se hace referencia se denominan *tablas de distribución de frecuencias*.

Una distribución de frecuencias es una “agrupación de datos cualitativos en clases mutuamente excluyentes, que muestra el número de observaciones que hay en cada clase” (Lind et al., 2012, p. 23). Las tablas de distribución de frecuencias agrupan gran cantidad de información disponible en una estructura de fácil comprensión. Cabe destacar que, aunque existen varias formas de agrupación de datos, el autor decide utilizar la que según su experiencia es la más sencilla.

Para diseñar correctamente una tabla de datos agrupados (nombre con el que también se conoce a la tabla de distribución de frecuencias) se seguirán 3 sencillos pasos que se desarrollarán en el ejemplo 1.

1.1 Número de clases óptimo

Existen varios métodos ampliamente difundidos para elegir el número de clases óptimo y aunque en este texto utilizaremos la regla de Sturges, se

recomienda utilizar el criterio del investigador para establecer el número de clases a utilizar, pues debido a que cada día se puede procesar mayor cantidad de información y la aplicación de esta fórmula podría generar un exceso de clases.

Sturges propone que para encontrar el número de clases adecuado (k_i) se debe aplicar la fórmula que se presenta en la ecuación 1, donde n es el tamaño de la muestra.

$$k_i = 1 + 1,322 \log n \quad (\text{Ecuación 1})$$

Ejemplo 1: Suponga que realizó una encuesta en la que se le consultó la edad a un grupo de mujeres que respondía a cierta característica de interés. El tamaño de la muestra n fue de 50. Si se quisiera agrupar las respuestas en una distribución de frecuencias, ¿cuál sería el tamaño óptimo de la muestra?

Abra una hoja de Excel y siga los siguientes pasos:

1 Ingrese por separado el tamaño de la muestra y las constantes, como se muestra en la Figura 1.

1	n	50,00
2	CONSTANTE 1	1,00
3	CONSTANTE 2	3,32
4	log n	
5	k_i	
6		

Figura 1

2 Para encontrar el logaritmo en base 10 del tamaño de la muestra escriba la fórmula (toda fórmula en Excel debe empezar con el símbolo = ó +) tal como lo muestra la Figura 2 =LOG10(B1). El resultado será el logaritmo en base 10 del tamaño de la muestra.

1	n	50,00
2	CONSTANTE 1	1,00
3	CONSTANTE 2	3,32
4	log n	=LOG10(B1)
5	k_i	
6		

Figura 2

3 Finalmente, en la celda B5 escriba la fórmula: $=B2+(B3*B4)$, tal como lo muestra la Figura 3. Es importante tomar en cuenta los paréntesis pues de no hacerlo podría obtenerse un número equivocado de clases.

1	n	50,00
2	CONSTANTE 1	1,00
3	CONSTANTE 2	3,32
4	log n	1,70
5	k_i	$=B2+(B3*B4)$
6		

Figura 3

Podría obtenerse el número adecuado de clase en menos pasos pero debe tomar en cuenta que se debe tener un conocimiento más profundo de las fórmulas en Excel, aunque no es nada difícil como se muestra en la Figura 4, $=1+(3,322*LOG10(B1))$

8	n	50,00
9	k_i	$=1+(3,322*LOG10(B1))$
10		

Figura 4

De cualquiera de las dos formas se obtiene un valor de 6,6439 por lo que se recomienda aproximar al inmediato superior, es decir, que el número adecuado de clases sería 7 para un tamaño de muestra de 50.

Ejercicios propuestos

1. Encuentre el número óptimo de clases para el ejemplo 1 suponiendo que el tamaño de la muestra es 100.
2. Se decide realizar un estudio para conocer los gustos y preferencias de 365 consumidores con relación a un nuevo plan de datos celulares.
3. Abra la hoja de cálculo **CAPÍTULO I.xls** y ubique la pestaña “**NÚMERO-CLASES**”, allí encontrará la fórmula para hacerlo rápidamente, una vez que ya haya comprendido la técnica.



Debido a la gran cantidad de datos que se manejan en la actualidad, el número de clases adecuado podría ser tan alto que se dificultaría la comprensión rápida de la distribución de los datos, por lo que se recomienda no exceder de 10 el número de clases, por ejemplo, si el tamaño de la muestra es de 4.000 elementos el tamaño de óptimo de clases (según Sturges) sería de 13 (compruébelo con el archivo de Excel **CAPÍTULO I.xls**). En este caso 13 clases son excesivas, a menos que desee hacer un análisis muy específico, lo recomendable sería un número de clases de entre 6 y 10, pero siempre tomando en cuenta el criterio de la persona que realiza el estudio.

1.2 Intervalo de clase

A éste se lo conoce también como anchura de clase y se obtiene con la fórmula de la ecuación 2.

$$i \geq \frac{H - L}{k} \quad (\text{Ecuación 2})$$

Siendo H el valor más alto observado y L el más bajo; mientras que k es el número de clases adecuado.

Suponga que para un programa de salud gubernamental se pesa a una muestra de 150 personas de la tercera edad (redondeando al entero más cercano y utilizando como unidad de medición libras) de un sector específico, los resultados se recogen en la tabla 1 (la información y cálculos los encontrará en el archivo de Excel llamado CAPÍTULO I, en la pestaña TABLA 1).

Tabla 1

208	247	108	158	249	190	234	135	217	205
203	123	151	207	188	139	186	235	169	100
189	119	142	125	157	180	196	169	227	166
134	235	195	141	215	123	153	213	221	155
249	176	189	156	148	154	211	234	177	239
119	115	219	201	236	143	112	111	140	224
223	212	134	156	232	148	160	110	136	197
233	159	144	162	213	122	152	104	208	240
104	232	173	187	113	237	158	147	192	224
143	211	104	213	113	146	137	195	158	158
103	230	105	181	201	229	119	117	247	152
231	135	125	215	232	165	156	189	140	138
162	156	117	239	203	218	204	133	178	152
115	174	239	172	129	222	143	113	122	125
191	225	247	197	144	103	221	206	222	215

El número de clases adecuado, utilizando el método aprendido anteriormente, es de 9 y se debe establecer el intervalo o anchura de clase (i), por lo que se procede a encontrar el valor más alto y más bajo observados abriendo una hoja de Excel e ingresando los datos desde la celda A1 (con fines de ejemplificar solamente, usted puede usar cualquier celda que desee) una vez ingresados los datos escriba la simbología de lo que está buscando y proceda a ingresar la información que ya conoce, como el tamaño de la muestra y el número de clases adecuado como lo muestra la Figura 5.

	A	B	C	D	E	F	G	H	I	J
1	208	247	108	158	249	190	234	135	217	205
2	203	123	151	207	188	139	186	235	169	100
3	189	119	142	125	157	180	196	169	227	166
4	134	235	195	141	215	123	153	213	221	155
5	255	176	189	156	148	154	211	234	177	239
6	119	115	219	201	236	143	112	111	140	224
7	223	212	134	156	232	148	160	110	136	197
8	233	159	144	162	213	122	152	104	208	240
9	104	232	173	187	113	237	158	147	192	224
10	143	211	104	213	113	146	137	195	158	158
11	103	230	105	181	201	229	119	117	247	152
12	231	135	125	215	232	165	156	189	140	138
13	162	156	117	239	203	218	204	133	178	152
14	115	174	239	172	129	222	143	113	122	125
15	191	225	247	197	144	103	221	206	222	215
16										
17										
18	n	150								
19	k	8								
20	H									
21	L									
22	i									

Figura 5

En la celda B20 se debe escribir la fórmula “ =MAX(” para encontrar el valor más alto observado y, cuando nos pida los números, elegimos todos los números de la tabla, es decir, los 150 datos de la forma en que se lo indica en la Figura 6.

El resultado es el dato más alto observado ó H.

Para encontrar el valor más bajo utilizar la fórmula “ =MIN(” i darle el mismo rango que elegimos para el máximo como lo indica la Figura 7.

	A	B	C	D	E	F	G	H	I	J	K
1	208	247	108	158	249	190	234	135	217	205	
2	203	123	151	207	188	139	186	235	169	100	
3	189	119	142	125	157	180	196	169	227	166	
4	134	235	195	141	215	123	153	213	221	155	
5	255	176	189	156	148	154	211	234	177	239	
6	119	115	219	201	236	143	112	111	140	224	
7	223	212	134	156	232	148	160	110	136	197	
8	233	159	144	162	213	122	152	104	208	240	
9	104	232	173	187	113	237	158	147	192	224	
10	143	211	104	213	113	146	137	195	158	158	
11	103	230	105	181	201	229	119	117	247	152	
12	231	135	125	215	232	165	156	189	140	138	
13	162	156	117	239	203	218	204	133	178	152	
14	115	174	239	172	129	222	143	113	122	125	
15	191	225	247	197	144	103	221	206	222	215	
16											
17											
18	n	150									
19	k	8									
20	H	=MAX(A1:J15)									
21	L										
22	i										

Figura 6

SUMA		=MIN(A1:J15)									
	A	B	C	D	E	F	G	H	I	J	K
1	208	247	108	158	249	190	234	135	217	205	
2	203	123	151	207	188	139	186	235	169	100	
3	189	119	142	125	157	180	196	169	227	166	
4	134	235	195	141	215	123	153	213	221	155	
5	255	176	189	156	148	154	211	234	177	239	
6	119	115	219	201	236	143	112	111	140	224	
7	223	212	134	156	232	148	160	110	136	197	
8	233	159	144	162	213	122	152	104	208	240	
9	104	232	173	187	113	237	158	147	192	224	
10	143	211	104	213	113	146	137	195	158	158	
11	103	230	105	181	201	229	119	117	247	152	
12	231	135	125	215	232	165	156	189	140	138	
13	162	156	117	239	203	218	204	133	178	152	
14	115	174	239	172	129	222	143	113	122	125	
15	191	225	247	197	144	103	221	206	222	215	
16											
17											
18	n	150									
19	k	8									
20	H	255									
21	L	=MIN(A1:J15)									
22	i										

Figura 7

Una vez establecido el valor máximo y el valor mínimo se debe calcular el intervalo i , también conocido como *amplitud de clase*, utilizando la fórmula de la ecuación 2.

Cabe destacar que el intervalo podríamos trabajarlo con el valor que nos da la operación incluyendo los decimales, pero se aconseja redondear al entero más cercano, e inclusive a la decena, centena, millar, etc, más cercano.

Observe la Figura 8 donde en la celda B22 se calcula el intervalo con dos decimales y, en la celda número B23 (Figura 9) se redondea al entero más cercano utilizando la fórmula “=REDONDEAR(B22;0)” eligiendo la celda B22 como primer criterio, es decir, el número que queremos redondear, separando los argumentos con el signo de puntuación “;” y finalmente tecleando el número cero para indicarle a la hoja de cálculo que redondee al entero más cercano.

	A	B	C	D	E	F	G	H	I	J	K
1	208	247	108	158	249	190	234	135	217	205	
2	203	123	151	207	188	139	186	235	169	100	
3	189	119	142	125	157	180	196	169	227	166	
4	134	235	195	141	215	123	153	213	221	155	
5	255	176	189	156	148	154	211	234	177	239	
6	119	115	219	201	236	143	112	111	140	224	
7	223	212	134	156	232	148	160	110	136	197	
8	233	159	144	162	213	122	152	104	208	240	
9	104	232	173	187	113	237	158	147	192	224	
10	143	211	104	213	113	146	137	195	158	158	
11	103	230	105	181	201	229	119	117	247	152	
12	231	135	125	215	232	165	156	189	140	138	
13	162	156	117	239	203	218	204	133	178	152	
14	115	174	239	172	129	222	143	113	122	125	
15	191	225	247	197	144	103	221	206	222	215	
16											
17											
18	n	150									
19	k	8									
20	H	255									
21	L	100									
22	i	=+(B20-B21)/B19									

Figura 8

17				
18	n	150		
19	k	8		
20	H	255		
21	L	100		
22	i	19,38		
23	i	=+REDONDEAR(B22;0)		
24				

Figura 9



Aunque se puede trabajar con intervalos no constantes entre clase y clase, o dejarlos abiertos, como por ejemplo mayores que (<) o menores que (>), es recomendable establecer un intervalo constante si se desea calcular luego valores para datos agrupados tales como media aritmética, mediana, varianza, desviación típica y otros; debido a que cuando se dejan los intervalos abiertos o no constantes, los resultados obtenidos pueden ser demasiado sesgados y no nos brindarían información que nos permita tomar lecturas acertadas ni decisiones efectivas.

Cuando se trata de cantidades grandes (cientos de miles o millones) es aconsejable expresar el intervalo en cantidades más pequeñas, dejando indicado las unidades de medición que se están tratando, por ejemplo, si el intervalo le da 123.000, sería mejor expresarlo como 123 (mil).

Ejercicios propuestos

4. Encuentre la anchura de clase para $H= 45$; $L=564$ y $k=5$.
5. Encuentre el intervalo para un grupo de datos cuyo mínimo es 12 y el máximo es 132 si deseamos construir una distribución de frecuencias de 4 clases.
6. Realice el cálculo para el ítem anterior suponiendo que se desean 6 clases.
7. Se desea estudiar el comportamiento de los ingresos de una empresa productora de detergente durante los últimos 5 años, para ello se recopiló información semanal cuyos resultados se muestran en la tabla del ejercicio 7, con esta información calcule un intervalo de clase suponiendo que se desea construir una distribución con 3 clases.

Tabla del ejercicio 7. Ingresos anuales de empresa productora de detergente (en dólares).

Mes/Año	2012	2013	2014	2015	2016
Enero	709.412,00	952.459,00	892.634,00	1.020.614,00	1.121.354,00
Febrero	1.205.280,00	1.219.132,00	1.106.904,00	1.035.590,00	1.228.739,00
Marzo	1.029.003,00	991.524,00	934.805,00	1.182.072,00	1.221.055,00
Abril	1.059.303,00	1.113.967,00	809.068,00	1.165.605,00	721.527,00
Mayo	670.936,00	1.188.259,00	719.164,00	733.209,00	1.206.161,00
Junio	1.166.351,00	1.203.975,00	1.103.703,00	1.078.869,00	844.925,00
Julio	999.712,00	777.397,00	1.068.755,00	1.064.469,00	1.000.618,00
Agosto	561.514,00	1.176.848,00	918.528,00	597.656,00	858.633,00
Septiembre	699.717,00	640.090,00	814.069,00	1.168.647,00	706.614,00
Octubre	1.186.643,00	775.975,00	1.231.344,00	978.577,00	882.438,00
Noviembre	923.893,00	1.031.730,00	782.504,00	1.116.149,00	957.474,00
Diciembre	1.068.500,00	1.229.371,00	860.436,00	780.006,00	1.103.950,00

8. Basados en las recomendaciones para el intervalo estime usted un número adecuado de clases y recalcule el intervalo.
9. ¿Por qué el intervalo del ejercicio 7 no sería adecuado?
10. Expresar el intervalo del ejercicio 7 en miles de dólares.

11. Exprese el intervalo del ejercicio 7 en millones de dólares.
12. Se quiere hacer un estudio que relaciona el peso de 140 estudiantes de primer semestre de la Universidad Laica VICENTE ROCAFUERTE de Guayaquil, y las actividades físicas que practican, junto con la frecuencia de las mismas. Se decidió por tal motivo pesar a los estudiantes utilizando una balanza digital de precisión cuya unidad de medición es en libras al segundo decimal redondeado. Los resultados se muestran en la tabla del ejercicio 12. Con esta información se requiere que encuentre el estudiante con menor peso, el que tiene mayor peso, que recomiende un número de clases adecuado y el intervalo de clase.

Tabla de datos del ejercicio 12.

251,45	238,45	285,80	182,10	245,88	166,69	162,91
191,81	249,86	243,33	189,82	253,95	257,77	169,75
128,78	254,18	259,97	142,06	142,86	175,88	223,89
275,03	234,95	246,34	235,01	291,24	153,87	261,83
123,93	295,92	257,31	271,30	177,16	137,21	123,09
131,52	248,53	203,96	265,66	266,72	265,49	147,14
189,90	214,05	208,72	182,47	217,41	140,57	232,48
129,46	256,74	188,72	247,12	258,82	125,29	173,27
237,70	298,05	233,75	141,45	181,57	235,69	159,39
156,52	209,97	248,32	297,48	210,66	243,50	121,53
267,26	293,24	201,82	273,51	159,04	164,67	238,31
145,26	122,02	248,57	165,18	122,58	145,42	222,22
244,35	276,73	192,57	251,19	173,90	165,54	145,79
167,90	234,47	125,28	208,88	159,23	138,75	185,95
194,20	187,46	291,17	154,32	292,28	238,11	207,11
183,73	172,45	259,05	201,85	141,30	289,18	126,61
294,29	284,47	288,06	204,63	286,10	129,89	271,18
262,17	268,38	135,97	222,49	199,72	133,24	223,70
256,63	186,55	125,39	181,21	299,55	295,17	145,30
191,94	149,20	120,46	179,33	151,42	189,31	159,06

1.3 Límites superiores e inferiores de clase

1.3.1 Límites de Clase

Son los valores menor y mayor que de encontrarse como datos en la muestra pertenecen a la clase en cuestión. (htt).

Para establecer el primer límite inferior (LI) es recomendable utilizar el mínimo, aunque también se podría redondear al entero cerrado inferior más cercano, por ejemplo en el caso del peso de las 150 personas de la tercera edad de la tabla 1, el límite inferior de la primera clase debería ser 100 debido a que el valor más bajo observado (L) es ese, como lo indica la Figura 10 que es la continuación del ejemplo que hemos venido utilizando (para un mejor entendimiento, los datos y cálculos los encontrará en el archivo de Excel llamado CAPÍTULO UNO, en la pestaña TABLA 1).

En el caso de los ingresos de la compañía productora de detergente de la tabla del ejercicio 7; el límite inferior de la primera clase debería ser \$561.514,00 porque ese es el valor (L) más bajo observado, aunque bien podía ser \$550.000,00 ó \$500.000,00 ó 550 mil ó 500 mil según se prefiera.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI
2	203	123	151	207	188	139	186	235	169	100		1ra.	=+B21
3	189	119	142	125	157	180	196	169	227	166		2da.	
4	134	235	195	141	215	123	153	213	221	155		3ra.	
5	255	176	189	156	148	154	211	234	177	239		4ta.	
6	119	115	219	201	236	143	112	111	140	224		5ta.	
7	223	212	134	156	232	148	160	110	136	197		6ta.	
8	233	159	144	162	213	122	152	104	208	240		7ma.	
9	104	232	173	187	113	237	158	147	192	224		8ava.	
10	143	211	104	213	113	146	137	195	158	158			
11	103	230	105	181	201	229	119	117	247	152			
12	231	135	125	215	232	165	156	189	140	138			
13	162	156	117	239	203	218	204	133	178	152			
14	115	174	239	172	129	222	143	113	122	125			
15	191	225	247	197	144	103	221	206	222	215			
16													
17													
18	n	150											
19	k	8											
20	H	255											
21	L	100											
22	i	19,38											
23	i	19											

Figura 10

Para encontrar el límite superior (LS) de la primera clase se debe sumar el intervalo o anchura de clase al primer límite inferior, por ejemplo, en el caso de del peso de los 150 personas de la tercera edad de la tabla 1, el primer límite inferior sería 100 y el primer límite superior sería 119 que resulta de sumar el primer LI al intervalo o anchura de clase, sin embargo para evitar confusiones a la hora de realizar el conteo de las frecuencias debemos establecer una regla que al principio podría parecer complicada pero que con la práctica se volverá un hábito.

La regla en cuestión es al intervalo restarle 1 unidad (en este caso se resta una unidad por que los datos son inferiores a las 4 cifras y enteros, ya se explicará luego como proceder en otros casos), es decir, el intervalo entre clases sería 18; que resulta de restar 1 unidad al i que fue 19. Con esta operación la primera clase quedaría como $LI=100$; $LS=118$, como lo muestra la Figura 11 continuando con el archivo antes mencionado.

La unidad que se restó al intervalo de la primera clase se suma al LS de la primera clase para obtener el LI de la segunda clase (Figura 12), es decir que el LI de la segunda clase sería 119 y el LS 137 como lo muestra la Figura 13. Nótese que la celda B23 se encuentra entre símbolos de dólar (\$) debido a que queremos que esa celda siempre permanezca constante cuando arrastremos las fórmulas.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI	LS
2	203	123	151	207	188	139	186	235	169	100		1ra.	100	=+M2+(B23-1)
3	189	119	142	125	157	180	196	169	227	166		2da.		
4	134	235	195	141	215	123	153	213	221	155		3ra.		
5	255	176	189	156	148	154	211	234	177	239		4ta.		
6	119	115	219	201	236	143	112	111	140	224		5ta.		
7	223	212	134	156	232	148	160	110	136	197		6ta.		
8	233	159	144	162	213	122	152	104	208	240		7ma.		
9	104	232	173	187	113	237	158	147	192	224		8ava.		
10	143	211	104	213	113	146	137	195	158	158				
11	103	230	105	181	201	229	119	117	247	152				
12	231	135	125	215	232	165	156	189	140	138				
13	162	156	117	239	203	218	204	133	178	152				
14	115	174	239	172	129	222	143	113	122	125				
15	191	225	247	197	144	103	221	206	222	215				
16														
17														
18	n	150												
19	k	8												
20	H	255												
21	L	100												
22	i	19,38												
23	i	19												

Figura 11

K	L	M	N
	CLASES	LI	LS
	1ra.	100	118
	2da.	$=+N2+1$	
	3ra.		
	4ta.		
	5ta.		
	6ta.		
	7ma.		
	8ava.		

Figura 12

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI	LS	
2	203	123	151	207	188	139	186	235	169	100		1ra.	100	118	
3	189	119	142	125	157	180	196	169	227	166		2da.	119	$=+M3+(SBS23-1)$	
4	134	235	195	141	215	123	153	213	221	155		3ra.			
5	255	176	189	156	148	154	211	234	177	239		4ta.			
6	119	115	219	201	236	143	112	111	140	224		5ta.			
7	223	212	134	156	232	148	160	110	136	197		6ta.			
8	233	159	144	162	213	122	152	104	208	240		7ma.			
9	104	232	173	187	113	237	158	147	192	224		8ava.			
10	143	211	104	213	113	146	137	195	158	158					
11	103	230	105	181	201	229	119	117	247	152					
12	231	135	125	215	232	165	156	189	140	138					
13	162	156	117	239	203	218	204	133	178	152					
14	115	174	239	172	129	222	143	113	122	125					
15	191	225	247	197	144	103	221	206	222	215					
16															
17															
18	n	150													
19	k	8													
20	H	255													
21	L	100													
22	i	19,38													
23	i	19													

Figura 13

CLASES	LI	LS
1RA	100	118
2DA	119	$=+M3+(SBS23-1)$
3RA		
4TA		
5TA		
7MA		
8VA		

Figura 14

Para hallar los límites de las clases restantes se debe seleccionar los límites de la segunda clase, posicionarse en la esquina inferior y arrastrar las fórmulas como lo muestran las figuras 15 y 16.

CLASES	LI	LS
1ra.	100	118
2da.	119	137
3ra.		
4ta.		
5ta.		
6ta.		
7ma.		
8ava.		

Figura 15

L	M	N
CLASES	LI	LS
1ra.	100	118
2da.	119	137
3ra.	138	156
4ta.	157	175
5ta.	176	194
6ta.	195	213
7ma.	214	232
8ava.	233	251

Figura 16

Una vez hallados todos los límites se debe comprobar si la distribución se encuentra correctamente diseñada previo al conteo de datos, para ello debemos tomar en cuenta que una distribución se encuentra correctamente diseñada si el valor más bajo observado se encuentra incluido en la primera clase y el valor más alto en la última.

En el caso que se viene desarrollando (peso de 150 personas de la tercera edad), el valor más bajo observado es de 100 y efectivamente cae dentro de la primera clase, mientras que el valor más alto es de 255 y no se encuentra dentro de la última pues su valor excede al límite superior de la última clase por lo que la distribución estaría mal diseñada (cabe destacar que el ejemplo fue diseñado de esa manera para que el lector pueda superar estos inconvenientes utilizando los criterios que se expondrán en las siguientes líneas).

El problema se generó por que al calcular el número adecuado de clases el resultado fue de 8,23 y se decidió redondear a 8, cuando lo recomendable es que para el número de clases cualquier decimal que resulte luego del entero nos obliga a redondear al inmediato superior, es decir, para este caso el k debió ser 9, sin embargo nueve clases para 150 datos podría resultar excesivas por lo que se debe utilizar el criterio anotado en el tip 1.

Para corregir entonces el inconveniente se pueden tomar dos caminos:

1. Repetir el proceso descrito en esta sección utilizando como k un valor de 9.
2. Ajustar el intervalo de clase dejando el k sin variación.

Debido a que en esta sección ya se explicó cuál es el proceso y el lector ya se encuentra en capacidad de modificar tanto el intervalo como los límites de clase, se procederá a explicar cómo aplicar la segunda solución.

Para ajustar el intervalo de clase debe incrementar el mismo en una unidad paulatinamente hasta que observe que el último intervalo contenga al valor más alto observado, es decir, utilizando nuevamente el ejemplo del peso de las 150 personas, que puede ser suficiente con incrementar el intervalo a 20 ó continuar incrementándolo en una unidad hasta que se logre el objetivo.

Utilizando el archivo de Excel CAPÍTULO 1.xls y ubicándose en la pestaña TABLA 1 MODIFICADA, se puede modificar el intervalo hasta que la distribución se encuentre bien diseñada. Por ejemplo, si se modifica el intervalo a 20 (los límites se modifican automáticamente pues ya fueron formulados en la

sección anterior para que ocurra de esta manera) se puede apreciar que con eso es suficiente pues el valor más alto observado se encuentra contenido en la última clase tal como lo muestra la Figura 17.

Si aún no fuere suficiente el incremento descrito en el párrafo anterior usted puede utilizar el archivo e incrementar el intervalo hasta que se cumpla el objetivo de que la última clase diseñada contenga al valor más alto observado.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI	LS
2	203	123	151	207	188	139	186	235	169	100		1ra.	100	119
3	189	119	142	125	157	180	196	169	227	166		2da.	120	139
4	134	235	195	141	215	123	153	213	221	155		3ra.	140	159
5	249	176	189	156	148	154	211	234	177	239		4ta.	160	179
6	119	115	219	201	236	143	112	111	140	224		5ta.	180	199
7	223	212	134	156	232	148	160	110	136	197		6ta.	200	219
8	233	159	144	162	213	122	152	104	208	240		7ma.	220	239
9	104	232	173	187	113	237	158	147	192	224		8ava.	240	259
10	143	211	104	213	113	146	137	195	158	158				
11	103	230	105	181	201	229	119	117	247	152				
12	231	135	125	215	232	165	156	189	140	138				
13	162	156	117	239	203	218	204	133	178	152				
14	115	174	239	172	129	222	143	113	122	125				
15	191	225	247	197	144	103	221	206	222	215				
16														
17														
18	n	150												
19	k	8												
20	H	249												
21	L	100												
22	j	18,625												
23	i	20												
24														

Figura 17



Debido a la gran cantidad de información que se maneja actualmente, no se aconseja trabajar con un número alto de clases, sino más bien incrementar el intervalo cuidando que todas las observaciones se encuentren incluidas en las clases que se decidió tener para el análisis, más concretamente asegúrese que el número más bajo observado se encuentre incluido en la primera clase y el más alto en la última.

Ejercicios propuestos

13. Diseñar una distribución de frecuencias en una nueva hoja de cálculo, utilizando los datos de la tabla del ejercicio 7, bajo los siguientes criterios: $k=6$; $i=120.000$; LI (primera clase)=560.000.
14. Comparar sus resultados vs los que se encuentran en la pestaña “INGRESOS-EMPRESA DETERGENTE” de la hoja de cálculo CAPÍTULO I.xls (corrija su archivo en caso de ser necesario)
15. Diseñar una distribución de frecuencias en una nueva hoja de cálculo, utilizando los datos de la tabla del ejercicio 12, bajo los siguientes criterios: $k=5$; $i=35,82$; LI (primera clase)=120,46.
16. Comparar sus resultados vs los que se encuentran en la pestaña “PESO ESTUDIANTES” de la hoja de cálculo CAPÍTULO I.xls (corrija su archivo en caso de ser necesario)
17. Una compañía operadora de telefonía celular desea lanzar un nuevo plan de datos dirigido a público joven, para ello consulta cuanto estarían dispuestos a pagar por un plan con las características ofrecidas a un grupo de 306 jóvenes de entre 18 y 25 años; los resultados (en dólares americanos) se listan en la tabla del ejercicio 17. Con esta información se le pide que en una nueva hoja de cálculo diseñe una distribución de frecuencias utilizando $K=5$.
18. Las ventas diarias (en dólares americanos) de una estación de servicio 24/7 se recogen en la tabla del ejercicio 18; Con esta información se le pide que en una nueva hoja de cálculo diseñe una distribución de frecuencias utilizando $K=6$.

Tabla de datos del ejercicio 17.

21	11	12	13	15	16	27	16	28	18	21	27	20	20	16	28	21
19	20	28	15	13	26	12	26	11	29	11	20	16	21	19	12	15
19	24	26	15	21	29	14	28	19	12	10	18	27	15	23	10	10
13	17	10	24	23	17	13	12	18	27	29	21	16	27	21	30	26
17	13	10	30	19	30	20	11	23	18	20	18	10	28	10	23	20
11	24	11	12	17	26	30	19	22	10	16	17	27	25	17	15	16
18	21	30	17	10	28	17	24	26	25	29	15	26	22	19	16	29
16	30	12	21	28	10	16	23	24	13	16	16	28	30	23	17	16
10	13	10	24	21	16	21	29	10	24	16	11	24	22	28	22	16
12	26	22	24	19	25	17	14	11	29	13	26	24	15	29	19	13
19	17	23	25	18	30	11	26	13	19	12	23	18	18	22	20	12
16	24	27	25	13	29	11	27	26	23	24	11	12	13	29	27	24
25	26	11	30	19	11	21	12	17	23	15	20	22	30	26	20	15
10	29	22	13	11	21	23	21	14	19	21	25	16	19	19	30	21
30	17	24	16	30	18	11	11	29	30	11	20	20	22	26	22	11
29	30	19	10	15	11	21	28	27	14	23	22	25	11	24	21	23
24	10	16	27	21	21	25	15	15	12	11	22	29	17	24	27	11
25	14	14	21	28	28	29	15	26	19	24	25	25	17	16	14	24

Tabla de datos del ejercicio 18.

10.449	6.225	4.012	6.961	11.067	6.043	13.853	10.287	10.805	13.650	7.407	19.784	6.096
3.090	11.075	15.083	13.000	10.276	5.476	14.285	12.145	9.780	18.356	13.546	15.515	17.050
6.536	14.424	14.632	3.965	12.804	7.524	9.784	11.559	17.798	9.582	5.481	16.375	6.296
6.853	4.386	18.616	6.137	7.285	8.105	15.913	19.372	9.921	9.157	15.040	3.454	4.650
11.950	15.951	7.132	19.215	19.078	1.911	10.239	1.154	1.020	4.334	4.192	15.400	17.506
14.346	3.566	13.157	19.254	8.254	12.350	14.198	2.694	5.874	15.966	14.516	18.570	7.013
10.245	4.827	11.741	11.369	3.729	8.686	14.834	12.255	11.375	11.504	5.158	14.639	9.504
4.566	4.980	4.658	2.871	1.015	16.157	8.905	19.822	3.887	6.955	4.090	16.454	4.308
3.168	12.780	10.957	5.547	19.159	4.661	8.387	12.116	19.582	17.763	12.857	7.552	1.752
13.153	18.983	1.042	3.979	12.261	8.442	9.562	19.285	19.898	19.694	12.941	3.737	9.517
11.788	19.825	1.436	9.643	11.081	3.097	8.660	16.412	5.143	9.754	6.564	6.750	15.564
9.922	19.088	14.673	17.383	14.788	5.384	3.566	5.598	3.290	2.007	1.777	14.561	16.242
13.155	5.433	3.929	14.385	1.153	17.994	18.460	3.476	17.799	10.650	18.819	18.263	16.644
19.065	14.063	11.251	12.630	17.103	15.795	5.668	13.067	4.337	10.249	8.021	7.857	11.842
3.116	15.207	1.678	19.901	8.987	9.157	10.089	12.132	15.846	6.843	10.256	16.525	12.552
17.901	11.560	9.987	18.944	4.476	8.593	4.050	3.329	14.676	10.684	13.827	7.035	10.137
10.888	18.657	8.314	15.599	4.126	11.068	17.533	16.862	3.128	15.203	19.162	4.516	13.781
14.931	18.830	3.678	13.212	11.009	8.751	13.170	11.198	7.419	12.812	8.458	18.237	14.021
2.468	12.224	2.244	17.851	13.008	14.389	10.211	1.357	2.988	15.029	14.409	19.255	10.142
7.599	15.699	2.375	8.135	19.823	11.533	13.910	17.728	15.576	17.876	13.879	11.877	13.065
11.322	19.796	5.685	17.410	15.352	18.246	15.462	17.662	17.816	17.654	5.628	18.202	11.616
11.646	16.439	8.893	3.538	17.244	2.428	7.839	16.419	14.487	8.255	10.867	14.747	11.833
3.109	5.460	4.530	1.557	11.927	16.463	16.938	13.258	11.405	4.328	8.774	3.111	10.227
3.985	13.440	14.051	5.923	16.264	10.320	4.700	17.562	11.743	1.819	3.309	5.156	6.006
14.769	9.107	2.387	12.710	14.757	15.264	15.461	10.095	3.191	19.872	13.263	7.858	4.221
7.780	16.920	17.730	13.613	3.880	9.119	8.574	18.671	11.544	2.313	2.615	16.091	12.980
12.906	9.535	1.597	14.653	6.100	1.502	8.256	10.499	4.522	9.136	4.494	17.542	17.516
11.087	14.744	8.464	12.935	11.408	16.714	13.675	16.828	5.337	1.288	18.357	14.034	2.505

1.4 Frecuencia de clase

Una vez encontrados los límites, tanto inferiores como superiores, se debe realizar el conteo de valores que se encuentran dentro de cada clase; a este conteo se lo denomina *frecuencia de clase*.

Realizar este conteo es extremadamente difícil y toma demasiado tiempo cuando se trata de muchas observaciones, por eso se prefiere el uso de algún software especializado para realizar esta tarea, sin embargo muchos desconocen que Excel puede ser utilizado como una herramienta de auxilio para tales menesteres.

En este apartado se explicará paso a paso como realizar un conteo de frecuencias de manera rápida y precisa, ya que cuando existe gran cantidad de información, aunque se logre realizar el conteo manualmente, no se podría asegurar que no se cometió ningún error.

Para ejemplificar continuamos con el ejemplo de las 150 personas de la tercera edad (Tabla 1) donde ya se establecieron el número de clases adecuado (8), el intervalo (22) y se encontraron los LI y LS de cada clase y se encuentran en el archivo de Excel CAPÍTULO 1.xls., en la pestaña TABLA 1 MODIFICADA.

Ubique el archivo antes mencionado, elija la pestaña TABLA 1 MODIFICADA e ingrese una nueva columna denominada f que es la simbología generalmente utilizada para denominar la frecuencia de cada clase (que no es más que el número de observaciones que se encuentran dentro del rango de una clase específica), tal como lo muestra la Figura 18.

L	M	N	O
CLASES	LI	LS	f
1ra.	100	119	
2da.	120	139	
3ra.	140	159	
4ta.	160	179	
5ta.	180	199	
6ta.	200	219	
7ma.	220	239	
8ava.	240	259	

Figura 18

El siguiente paso es ubicarse en la celda O2 que pertenece a la primera clase para hallar la primera frecuencia.

Para hallar la frecuencia de la primera clase utilizaremos la función FRECUENCIA (al ingresar esta función se pedirán 2 argumentos, el primero corresponde a los datos observados y el segundo, separado por punto y coma, corresponde a los límites superiores de cada clase) de la siguiente manera: Escriba =FRECUENCIA(A1:J15;N2:N8) y teclee enter.

Se obtendrá como resultado 21, esto quiere decir que existen 21 observaciones que se encuentran dentro del rango de la primera clase, es decir, existen 21 adultos mayores que tienen un peso de entre 100 y 121 libras.

Cuando escriba la función, proporcione como primer argumento las celdas donde se encuentran ubicados los datos del fenómeno en estudio, en nuestro ejemplo, desde la celda A1 hasta la celda J15; luego separe los argumentos con la tecla punto y coma, para luego dar como segundo argumento los límites superiores de cada clase; en nuestro ejemplo desde la celda N2 a la celda N8, como lo muestra la tabla Figura 19.

Una vez hallada la primera frecuencia se debe arrastrar la fórmula hasta la celda O9 como lo muestra la Figura 20.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI	LS	f	
2	203	123	151	207	188	139	186	235	169	100		1ra.	100	119	=FRECUENCIA(A1:J15;N2:N9)	
3	189	119	142	125	157	180	196	169	227	166		2da.	120	139		
4	134	235	195	141	215	123	153	213	221	155		3ra.	140	159		
5	249	176	189	156	148	154	211	234	177	239		4ta.	160	179		
6	119	115	219	201	236	143	112	111	140	224		5ta.	180	199		
7	223	212	134	156	232	148	160	110	136	197		6ta.	200	219		
8	233	159	144	162	213	122	152	104	208	240		7ma.	220	239		
9	104	232	173	187	113	237	158	147	192	224		8ava.	240	259		
10	143	211	104	213	113	146	137	195	158	158						
11	103	230	105	181	201	229	119	117	247	152						
12	231	135	125	215	232	165	156	189	140	138						
13	162	156	117	239	203	218	204	133	178	152						
14	115	174	239	172	129	222	143	113	122	125						
15	191	225	247	197	144	103	221	206	222	215						
16																

Figura 19

CLASES	LI	LS	f
1ra.	100	119	21
2da.	120	139	36
3ra.	140	159	61
4ta.	160	179	68
5ta.	180	199	73
6ta.	200	219	81
7ma.	220	239	90
8ava.	240	259	85

Figura 20

Luego de arrastrar la fórmula no haga nada que dé como resultado la desección de las celdas que contengan la frecuencia, pues es necesario realizar ciertas precisiones que le permitirán corregir algunos aspectos que no concuerden con la respuesta correcta.

Si usted se ubica en la parte inferior derecha de la hoja de Excel (Figura 21) podrá notar que el recuento está correcto, pues son 8 frecuencias correspondientes a 8 clases, sin embargo la suma resulta 515, es decir la suma de las frecuencias es 515 de la frecuencia, cuando teóricamente se conoce que la sumatoria de la frecuencia ($\sum f$) debe ser igual al tamaño de la muestra o población según sea el caso (n ó N respectivamente).

Para que el resultado sea correcto en la hoja de Excel se debe transformar la columna de las frecuencias en una matriz (esto debido a un proceso matemático un poco complejo que no será analizado en este apartado pues lo que nos ocupa es encontrar las frecuencias correctas y no el procedimiento que el programa realiza para hallarlas).

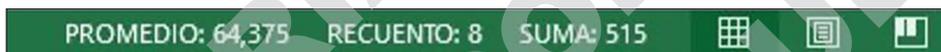


Figura 21

Una vez realizadas las aclaraciones se debe proceder a realizar un sencillo paso con el fin de encontrar las frecuencias correctas: Sin quitar la selección anterior (Figura 20) proceda a ubicarse en la barra de fórmulas (Figuras 22 y 23) y presione al mismo tiempo las teclas shift + ctrl + enter.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	208	247	108	158	249	190	234	135	217	205		CLASES	LI	LS	f
2	203	123	151	207	188	139	186	235	169	100		1ra.	100	119	=FRECUENCI
3	189	119	142	125	157	180	196	169	227	166		2da.	120	139	36
4	134	235	195	141	215	123	153	213	221	155		3ra.	140	159	61
5	249	176	189	156	148	154	211	234	177	239		4ta.	160	179	68
6	119	115	219	201	236	143	112	111	140	224		5ta.	180	199	73
7	223	212	134	156	232	148	160	110	136	197		6ta.	200	219	81
8	233	159	144	162	213	122	152	104	208	240		7ma.	220	239	90
9	104	232	173	187	113	237	158	147	192	224		8ava.	240	259	85
10	143	211	104	213	113	146	137	195	158	158					
11	103	230	105	181	201	229	119	117	247	152					
12	231	135	125	215	232	165	156	189	140	138					
13	162	156	117	239	203	218	204	133	178	152					
14	115	174	239	172	129	222	143	113	122	125					
15	191	225	247	197	144	103	221	206	222	215					
16															

Figura 22

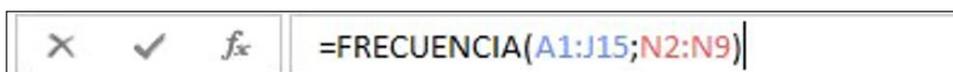


Figura 23

Note que al realizar el procedimiento antes descrito, la fórmula que se encuentra en la barra cambia y se transforma en una matriz, tal como lo muestra la Figura 24.



Figura 24

El resultado será la frecuencia de clases correcta para cada caso, es decir, el número de veces que las observaciones se encuentran dentro de cada clase respectiva (Figura 25).

De este modo se puede observar, por ejemplo, que existen 21 personas de la tercera edad que tienen un peso de entre 100 y 119 libras, 17 que tienen un peso de entre 120 y 139; 30, que cuentan con un peso de entre 140 y 159 libras, y así sucesivamente hasta la octava clase.

CLASES	LI	LS	f
1ra.	100	119	21
2da.	120	139	17
3ra.	140	159	30
4ta.	160	179	13
5ta.	180	199	16
6ta.	200	219	22
7ma.	220	239	25
8ava.	240	259	6

Figura 25

Si se ubica en la barra inferior derecha como en el proceso descrito anteriormente, podrá notar que ahora tanto el recuento como la sumatoria coinciden con lo correcto respectivamente; el recuento es 8 por que existen 8 clases mutuamente excluyentes y la suma 150 por que el número total de observaciones o tamaño de la muestra en este caso es de 150. Si a usted no le da como resultado esto, vuelva a repetir el procedimiento desde el principio pues algo no debe haberse hecho de forma incorrecta.

Ejercicios propuestos

19. Tomando en cuenta los datos del ejercicio 7, elabore una distribución de frecuencias encontrando: el número de clases adecuado, los límites de clase y la frecuencia absoluta; puede utilizar una hoja de cálculo nueva (recomendable para que ponga en práctica todos los puntos anteriores) o puede abrir el archivo “CAPÍTULO I.xls” y ubicarse en la pestaña “INGRESOS-EMPRESA DETERGENTE” y solo calcular la frecuencia absoluta.
20. Utilizando la información del ejercicio 12, elabore una distribución de frecuencias encontrando: el número de clases adecuado, los límites de clase y la frecuencia absoluta; puede utilizar una hoja de cálculo o puede abrir el archivo “CAPÍTULO I.xls” y ubicarse en la pestaña “PESO ESTUDIANTES” y solo calcular la frecuencia absoluta.
21. Tomando en cuenta los datos del ejercicio 17, elabore una distribución de frecuencias encontrando: el número de clases adecuado, los límites de clase y la frecuencia absoluta; puede utilizar una hoja de cálculo nueva o abrir el archivo “CAPÍTULO I.xls” y ubicarse en la pestaña “PLAN INTERNET” y solo calcular la frecuencia absoluta.
22. Utilizando la información del ejercicio 18, elabore una distribución de frecuencias encontrando: el número de clases adecuado, los límites de clase y la frecuencia absoluta; puede utilizar una hoja de cálculo o puede abrir el archivo “CAPÍTULO I.xls” y ubicarse en la pestaña “VENTAS E-S” y solo calcular la frecuencia absoluta.
23. Suponga que quiere establecer la relación entre la cantidad de galones de gasolina consumidos por un bus que brinda servicio de transportación escolar y los kilómetros recorridos durante 6 meses con el fin de establecer una nueva tarifa debido a que los costos de transportación se han incrementado pues han sucedido incrementos periódicos de los precios del combustible. La información de los kilómetros recorridos se encuentra en la tabla del ejercicio 23. Suponga que durante el tiempo de estudio los meses cuentan con 4 semanas y sólo se presta el servicio de transportación de lunes a viernes. Con esta información se pide que diseñe una distribución de frecuencias encontrando el número de clases adecuado, el intervalo de clase, los límites inferiores y superiores de clase y la frecuencia absoluta.

Tabla de datos del ejercicio 23.

58	50	71	52	39	83
15	84	40	30	28	94
41	70	33	96	36	54
18	58	35	2	74	91
44	70	42	19	80	43
62	38	10	15	65	38
20	42	112	9	4	105
45	38	63	36	88	12
20	20	34	98	72	70
112	102	32	105	120	70
24	90	55	53	58	73
101	79	92	90	55	102
97	88	45	67	88	69
44	85	3	42	42	118
42	17	29	58	69	20
78	92	108	69	75	69

1.5 Frecuencia acumulada

La frecuencia acumulada resulta de la suma sucesiva de las frecuencias inmediatas anteriores y sirve para comprender cuantas observaciones en total existen dentro de un rango de inicio y fin. Para ilustrarlo encontraremos la frecuencia acumulada para el ejercicio de las 150 personas de la tercera edad, donde la primera frecuencia acumulada siempre será la primera frecuencia absoluta (Figura 26).

CLASES	LI	LS	f	fa
1ra.	100	119	21	21 = 02
2da.	120	139	17	
3ra.	140	159	30	
4ta.	160	179	13	
5ta.	180	199	16	
6ta.	200	219	22	

Figura 26

Para leer este resultado debemos situarnos en el límite inferior de la primera clase y en el límite superior de la primera clase, entendiéndose entonces que 21 personas de la tercera edad tienen un peso de entre 100 y 119 libras.

Para hallar la segunda frecuencia acumulada se debe sumar la frecuencia acumulada de la primera clase más la frecuencia absoluta de la segunda clase, tal como indica la Figura 27.

CLASES	LI	LS	f	fa
1ra.	100	119	21	21
2da.	120	139	17	+P2+O3
3ra.	140	159	30	
4ta.	160	179	13	
5ta.	180	199	16	

Figura 27

El resultado será 38, lo que significa que en total 38 personas de la tercera edad tendrán entre 100 y 139 libras. Nótese que para la lectura se utilizó el límite inferior de la primera clase y el límite superior de la segunda clase.

Para encontrar las demás frecuencias acumuladas se debe, simplemente, arrastrar la fórmula tal como lo muestra la Figura 28.

CLASES	LI	LS	f	fa
1ra.	100	119	21	21
2da.	120	139	17	38
3ra.	140	159	30	68
4ta.	160	179	13	81
5ta.	180	199	16	97
6ta.	200	219	22	119
7ma.	220	239	25	144
8ava.	240	259	6	150

Figura 28

Nótese que la última frecuencia acumulada debe ser igual a n, es decir, la última frecuencia acumulada debe ser igual al tamaño de la muestra o de la población, según sea el caso (Figura 29).

	L	M	N	O	P
CLASES	LI	LS	f	fa	
1ra.	100	119	21	21	
2da.	120	139	17	38	
3ra.	140	159	30	68	
4ta.	160	179	13	81	
5ta.	180	199	16	97	
6ta.	200	219	22	119	
7ma.	220	239	25	144	
8ava.	240	259	6	150	
Sumatoria				150	

Figura 29

Ejercicios propuestos

24. Utilizando el archivo de Excel CAPÍTULO 1.xls y la pestaña “INGRESOS-EMPRESA DETERGENTE” encuentre e interprete la frecuencia acumulada (f_a).
25. Encuentre e interprete la f_a para el ejercicio del peso de los estudiantes.
26. A manera de práctica encuentre la f_a para todos los demás ejercicios propuestos.

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

1.6 Frecuencia relativa

La frecuencia relativa es la resultante de dividir cada una de las frecuencias absolutas para el tamaño de la muestra o población según sea el caso. Otra forma de entenderlo es el cociente entre cada una de las frecuencias absolutas y la sumatoria de las frecuencias absolutas.

$$fr = \frac{fi}{n} \quad \text{o} \quad fr = \frac{fi}{\sum fi} \quad (\text{Ecuación 3})$$

Tomando en cuenta una vez más el ejemplo del peso de las 150 personas de la tercera edad, dividimos la primera frecuencia para la sumatoria de las frecuencias recordando que para poder arrastrar las fórmulas, la celda que contenga a la sumatoria de las frecuencias debe permanecer fija, por lo tanto debe estar entre signos de dólares tal como lo muestra la figura 30.

O	P	Q
f	fa	fr
21	21	=O2/\$O\$10
17	38	
30	68	
13	81	
16	97	
22	119	
25	144	
6	150	
150		

Figura 30

Una vez obtenga el resultado dando enter, ya puede arrastrar la fórmula (asegurándose que la celda que contiene la sumatoria de las frecuencias se encuentre fija como en la Figura 30). El resultado se puede observar en la Figura 31.

CLASES	LI	LS	f	fa	fr
1ra.	100	119	21	21	0,14
2da.	120	139	17	38	0,1133333
3ra.	140	159	30	68	0,2
4ta.	160	179	13	81	0,0866667
5ta.	180	199	16	97	0,1066667
6ta.	200	219	22	119	0,1466667
7ma.	220	239	25	144	0,1666667
8ava.	240	259	6	150	0,04
Sumatoria			150		

Figura 31

Para expresar la columna de las frecuencias en porcentajes debe dejar seleccionados los datos obtenidos en el paso anterior y dirigirse a la parte superior en la cinta de opciones en la pestaña de inicio dando clic a la opción estilo porcentual, tal como lo muestran las Figuras 32 y 33.



Figura 32

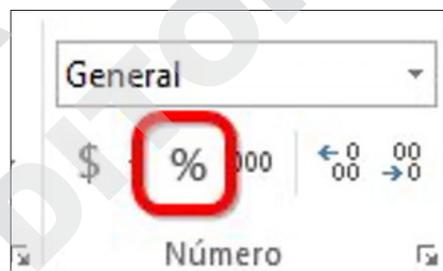


Figura 33

El resultado se muestra en la Figura 34. Si el número de decimales no resulta el mismo para todas las frecuencias relativas o si desea aumentar o disminuir el número de decimales para cada caso, puede utilizar los botones aumentar o disminuir decimales como se muestra en la Figura 35.

L	M	N	O	P	Q
CLASES	LI	LS	f	fa	fr
1ra.	100	119	21	21	14%
2da.	120	139	17	38	11%
3ra.	140	159	30	68	20%
4ta.	160	179	13	81	9%
5ta.	180	199	16	97	11%
6ta.	200	219	22	119	15%
7ma.	220	239	25	144	17%
8ava.	240	259	6	150	4%
Sumatoria			150		

Figura 34

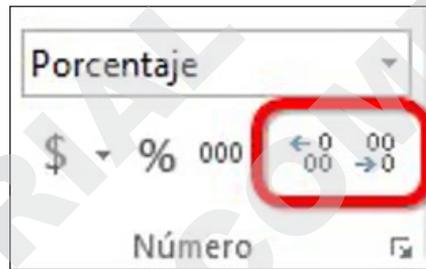


Figura 35

Ejercicios propuestos

27. Ubique la pestaña INGRESOS-EMPRESA DETERGENTE del archivo de Excel CAPÍTULO 1.xls y elabore la *fr* siguiendo los pasos descritos en el apartado anterior.
28. Encuentre la *fr* para el ejercicio 7, ubicándose en la pestaña INGRESOS-EMPRESA DETERGENTE del archivo de Excel CAPÍTULO 1.xls.
29. Encuentre la *fr* para el ejercicio 12, ubicándose en la pestaña PESO ESTUDIANTES del archivo de Excel CAPÍTULO 1.xls.
30. Encuentre la *fr* para el ejercicio 17, ubicándose en la pestaña PLAN INTERNET del archivo de Excel CAPÍTULO 1.xls.
31. Con la finalidad de establecer un precio referencial para un servicio de lubricación de vehículos a domicilio un par de emprendedores encuesta a 553 potenciales clientes consultándoles cuanto estarían dispuestos a pagar por el servicio, los resultados se listan en la tabla de datos y en la pestaña LUBRICACIÓN-DOMICILIO. Utilizando esta información elabore una distribución de frecuencias con $k=5$ llegando hasta la *fr*.
32. A manera de práctica complete todas las distribuciones de frecuencia propuestas hasta la frecuencia relativa.

Tabla de datos del ejercicio 31.

32	27	41	45	43	21	37	54	26	29	26	25	18	54
48	17	31	36	12	19	40	18	50	49	31	47	52	55
48	29	51	51	20	47	56	46	20	22	47	26	21	41
53	38	37	16	14	23	38	50	20	54	45	52	20	46
15	38	39	41	55	54	19	11	38	34	12	28	18	37
12	25	43	27	51	17	45	26	37	54	29	30	32	46
32	51	17	55	31	55	24	22	55	46	25	55	12	22
32	39	38	35	43	53	42	35	30	23	33	31	41	28
33	38	37	36	13	39	11	43	36	36	40	27	40	14
14	12	45	14	26	18	34	22	25	23	56	45	34	25
30	44	40	24	16	42	24	22	38	43	16	14	45	47
29	10	11	22	44	32	33	41	34	37	47	43	45	51
15	45	49	10	34	54	47	27	45	12	21	50	44	22
46	48	53	21	30	27	26	22	11	29	19	19	10	25
38	56	55	51	43	52	46	33	13	24	52	32	45	22
34	54	11	27	47	52	12	36	39	26	46	15	54	24
52	25	52	26	12	14	18	25	39	50	16	30	29	36
39	45	53	12	22	21	17	26	19	41	39	36	10	25
15	40	14	15	45	28	37	22	54	20	34	30	11	33
24	14	49	41	44	36	20	31	51	12	30	21	55	50
48	47	44	21	36	22	18	19	50	35	53	10	41	39
19	50	28	13	47	24	31	17	21	39	20	11	19	56
18	10	47	14	35	20	40	35	11	11	26	24	21	48
20	32	27	43	43	35	27	26	43	14	46	31	36	33
46	23	14	35	50	19	52	52	46	24	46	48	43	54
35	43	55	46	39	23	34	31	30	44	13	13	28	17
15	23	23	25	56	36	22	54	15	22	44	17	44	49
49	46	12	42	45	45	31	17	48	11	46	53	23	41
28	14	23	31	30	38	17	21	27	14	48	22	31	24
32	27	22	19	24	38	28	24	42	20	35	50	44	29
26	14	31	14	37	31	53	17	12	22	33	22	14	35
28	10	33	15	37	46	30	46	29	38	31	55	54	47
12	36	18	39	53	54	25	50	43	25	40	46	14	11
30	50	42	27	34	19	51	33	44	50	15	35	25	43
16	55	18	36	40	53	15	50	41	17	44	21	24	51
36	50	15	20	53	25	50	44	55	34	36	21	35	33
12	37	31	49	35	53	18	27	33	37	13	38	44	21
31	33	11	25	14	56	20	13	53	42	18	54	50	48
41	24	11	25	22	41	40	31	28	53	12	29	10	26
40	10	40	34	56	47	12	53	36	14	33	19	47	10

1.7 Frecuencia relativa acumulada

La frecuencia relativa acumulada resulta de la suma sucesiva de las frecuencias relativas inmediatas anteriores y sirve para comprender cuantas observaciones en total existen dentro de un rango de inicio y fin. Para ilustrarlo encontraremos la frecuencia relativa acumulada (*fra*) para el ejercicio de las 150 personas de la tercera edad, donde la primera *fra* siempre será la primera frecuencia relativa y las siguientes resultan de la suma sucesiva (Figura 36).

CLASES	LI	LS	f	fa	fr	fra
1ra.	100	119	21	21	14%	+Q2
2da.	120	139	17	38	11%	
3ra.	140	159	30	68	20%	
4ta.	160	179	13	81	9%	
5ta.	180	199	16	97	11%	
6ta.	200	219	22	119	15%	
7ma.	220	239	25	144	17%	
8ava.	240	259	6	150	4%	
Sumatoria			150			

Figura 36

La segunda *fra* resulta de sumar la primera *fra* + la *fr* de la segunda clase, para este ejemplo resulta ser 25% (Figura 37) lo cual quiere decir que existe un 25% de adultos mayores que tienen un peso de entre 100 y 139 libras. Una vez hallada la segunda *fra* ya puede arrastrar la fórmula hasta el final de la serie, tomando en cuenta que la última frecuencia relativa acumulada debe ser igual al 100% (Figura 38). Si la última *fra* resulta ser diferente de 100% repita el proceso pues algo debe haber salido mal.

CLASES	LI	LS	f	fa	fr	fra
1ra.	100	119	21	21	14%	14%
2da.	120	139	17	38	11%	+R2+Q3
3ra.	140	159	30	68	20%	
4ta.	160	179	13	81	9%	
5ta.	180	199	16	97	11%	
6ta.	200	219	22	119	15%	
7ma.	220	239	25	144	17%	
8ava.	240	259	6	150	4%	
Sumatoria			150			

Figura 37

fra
14%
25%
45%
54%
65%
79%
96%
100%

Figura 38

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

Ejercicios propuestos

33. Ubique la pestaña INGRESOS-EMPRESA DETERGENTE del archivo de Excel CAPÍTULO 1.xls y elabore la *fra* siguiendo los pasos descritos en el apartado anterior.
34. Encuentre la *fra* para el ejercicio 7, ubicándose en la pestaña INGRESOS-EMPRESA DETERGENTE del archivo de Excel CAPÍTULO 1.xls.
35. Encuentre la *fra* para el ejercicio 12, ubicándose en la pestaña PESO ESTUDIANTES del archivo de Excel CAPÍTULO 1.xls.
36. Encuentre la *fra* para el ejercicio 17, ubicándose en la pestaña PLAN INTERNET del archivo de Excel CAPÍTULO 1.xls.
37. A manera de práctica complete todas las distribuciones de frecuencia propuestas hasta la frecuencia relativa acumulada.

1.8 Marca de clase

La marca de clase es el punto medio de cada clase y resulta de la suma de los límites tanto inferiores como superiores de cada clase dividido para 2, se podría decir también que la marca de clase es el promedio de cada clase. La fórmula para la marca de clase (mc) se puede observar en la ecuación 4.

$$mc = \frac{LI_i + LS_i}{2} \quad (\text{Ecuación 4})$$

Excel facilita mucho el cálculo de la mc , en este apartado utilizaremos dos métodos, uno es mediante la utilización de una fórmula sencilla y otra a través de la utilización de la *función promedio*.

Para el primer método utilizaremos una vez más el ejemplo de las 150 personas de la tercera edad que se encuentra en la pestaña TABLA 1 MODIFICADA del archivo de Excel CAPÍTULO I.xls. Sume el límite inferior de la primera clase con el límite superior de la misma, una vez realizado esto divídalo para dos, tomando en cuenta que la primera operación debe estar entre paréntesis para que Excel comprenda que debe primero realizar la suma y luego dividir esta suma para dos tal como lo muestra la Figura 39.

CLASES	LI	LS	f	fa	fr	fra	mc
1ra.	100	119	21	21	14%	14%	=(M2+N2)/2
2da.	120	139	17	38	11%	25%	
3ra.	140	159	30	68	20%	45%	
4ta.	160	179	13	81	9%	54%	
5ta.	180	199	16	97	11%	65%	
6ta.	200	219	22	119	15%	79%	
7ma.	220	239	25	144	17%	96%	
8ava.	240	259	6	150	4%	100%	
Sumatoria			150				

Figura 39

Una vez realizada esta operación ya puede arrastrar fórmula o copiar las fórmulas utilizando el botón rellenar (de la pestaña inicio) y escogiendo rellenar hacia abajo tal como lo muestran las Figuras 40 a 43.

	mc
%	109,5
%	
%	
%	
%	
%	
%	
%	

Figura 40



Figura 41

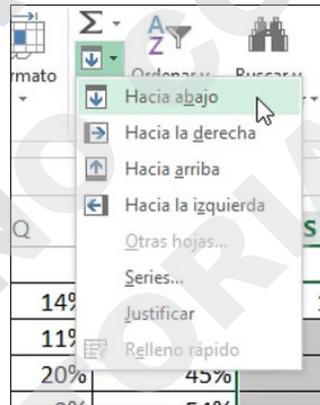


Figura 42

CLASES	LI	LS	f	fa	fr	fra	mc
1ra.	100	119	21	21	14%	14%	109,5
2da.	120	139	17	38	11%	25%	129,5
3ra.	140	159	30	68	20%	45%	149,5
4ta.	160	179	13	81	9%	54%	169,5
5ta.	180	199	16	97	11%	65%	189,5
6ta.	200	219	22	119	15%	79%	209,5
7ma.	220	239	25	144	17%	96%	229,5
8ava.	240	259	6	150	4%	100%	249,5
Sumatoria			150				

Figura 43

El resultado se lee de la siguiente forma: existen 21 personas de la tercera edad que tienen un peso promedio de 109,5 libras; 17 que tienen un peso medio de 129,5 libras, etc.

Ahora bien, en párrafos anteriores se mencionó que se podían utilizar dos métodos por lo que a continuación se explicará cómo utilizar la función promedio.

Para el segundo método, ubíquese en el mismo archivo y borre lo realizado anteriormente (sólo la columna de la marca de clase) y colocándose en la primera celda escriba “=PROMEDIO(“; como primer argumento escoja el límite inferior de la primera clase y como segundo argumento el límite superior de la segunda clase, el proceso y resultado de la utilización de la función mencionada se encuentra en las Figuras 44 y 45.

L	M	N	O	P	Q	R	S	T
CLASES	Li	Ls	f	fa	fi	fra	mc	
1ra.	100	119	21	21	14%	14%	=PROMEDIO(M2;N2)	
2da.	120	139	17	38	11%	25%		
3ra.	140	159	30	68	20%	45%		
4ta.	160	179	13	81	9%	54%		
5ta.	180	199	16	97	11%	65%		
6ta.	200	219	22	119	15%	79%		
7ma.	220	239	25	144	17%	96%		
8ava.	240	259	6	150	4%	100%		
Sumatoria			150					

Figura 44

=PROMEDIO(M2;N2)

Figura 45

El resultado obtenido mediante este método debe ser igual que el obtenido por el anteriormente descrito. Para completar las marcas de clase puede arrastrar la fórmula, rellenar hacia abajo o copiar y pegar las fórmulas utilizando el botón de pegado especial de la pestaña inicio tal como se ilustra en las Figuras 46 a 48.

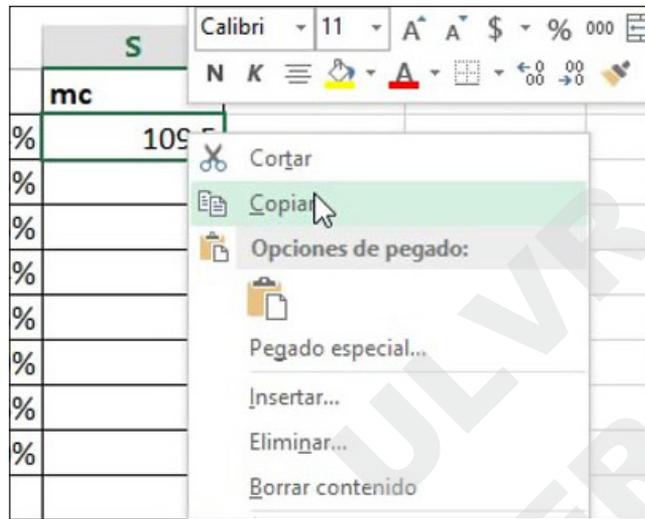


Figura 46

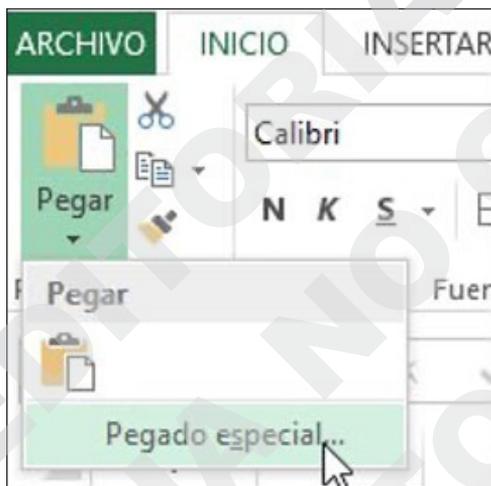


Figura 47

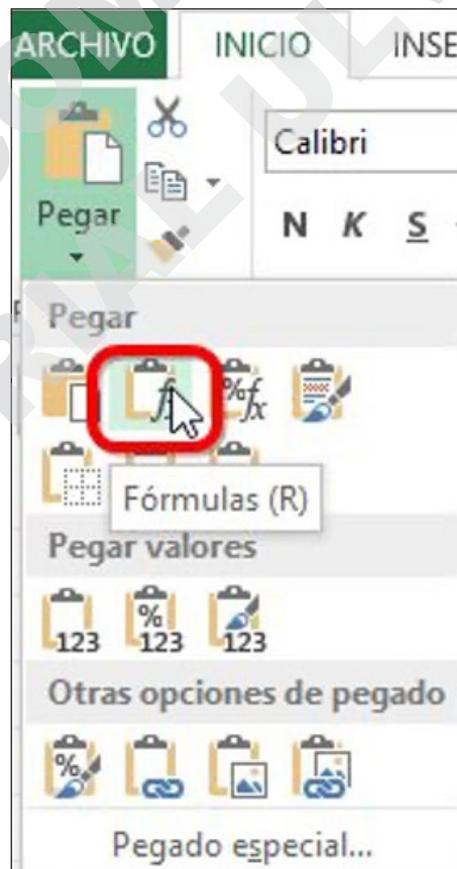


Figura 48



La marca de clase sirve para conocer el punto medio de cada clase y de esta forma analizar el número de observaciones en el promedio de cada clase, siempre y cuando el intervalo de clase siempre sea el mismo para cada clase, es decir, siempre y cuando el intervalo sea constante. También sirve para hallar el valor típico, que no es nada más que la marca de clase de la clase que tiene mayor frecuencia.

Ejemplo 2: Imagine que la tabla del ejemplo, representa una distribución de frecuencia particular.

Clases		F
5	11	2
12	18	6
19	25	25
26	32	13
33	39	9
40	46	1

Se podría decir que el valor típico de esta distribución es de 22, porque es el punto medio de la segunda clase (que se encuentra resaltado en amarillo).

Ejercicios propuestos

38. Encuentre las marcas de clase para la tabla del ejemplo 2 (archivo de Excel CAPÍTULO I.xls pestaña TABLA 8) y verifique que el valor típico es de 22.
39. Encuentre las marcas de clase para los ejercicios propuestos en las tablas 9, 10, 11 y 12 del archivo de Excel CAPÍTULO I.xls.
40. La tabla 13 muestra una distribución de frecuencias del consumo mensual en bebidas alcohólicas de un grupo de 250 ejecutivos de alto rango. Utilizando esta información y los métodos de cálculo anteriores, encuentre el consumo típico en bebidas alcohólicas de esas familias (archivo de Excel CAPÍTULO I.xls pestaña TABLA 13).

Tabla 9.

Clases		F	Mc
12,5	22,5	12	
22,6	32,6	13	
32,7	42,7	7	
42,8	52,8	61	
52,9	62,9	5	
63,0	73,0	36	
73,1	83,1	20	

Tabla 10.

Clases		F	Mc
1.200	1.500	12	
1.501	1.801	13	
1.802	2.102	15	
2.103	2.403	35	
2.404	2.704	23	
2.705	3.005	12	
3.006	3.306	6	

Tabla 11.

Clases (millones de \$)		F	Mc
1,2	3,2	1	
3,3	5,3	5	
5,4	7,4	8	
7,5	9,5	2	
9,6	11,6	3	

Tabla 12.

Clases (toneladas)		F	Mc
0,12	0,32	23	
0,42	0,62	18	
0,72	0,92	14	
1,02	1,22	56	
1,32	1,52	54	
1,62	1,82	36	
1,92	2,12	20	

Tabla 13.

Consumo de alcohol (\$)		F	Mc
50,0	80,0	23	
81,0	111,0	25	
112,0	142,0	43	
143,0	173,0	85	
174,0	204,0	34	
205,0	235,0	36	
236,0	266,0	4	

1.9 Gráficas

Con el fin de que la comprensión de un fenómeno estadístico sea más fácil se acude a las gráficas pues estas tienen un impacto más visual ante quien desea interpretar una tabla de distribución de frecuencias; en este apartado aprenderemos a elaborar 3 tipos de gráficas comúnmente utilizadas en estadística, las cuales son: Histogramas de frecuencias (gráfica de barras), polígono de frecuencias (gráfica de líneas) y ojivas (gráfica incremental).

1.9.1 Histograma de frecuencias

Un histograma de frecuencias es una gráfica de barras que sirve para representar visualmente una distribución de frecuencias, continuando con el ejemplo de las 150 personas de la tercera edad, abra el archivo CAPÍTULO I.xls ubicando la pestaña TABLA 1 MODIFICADA, seleccione la columna de la frecuencia tal como lo muestra la Figura 49.

CLASES	LI	LS	f	fa	fr	fra	mc
1ra.	100	119	21	21	14%	14%	109,5
2da.	120	139	17	38	11%	25%	
3ra.	140	159	30	68	20%	45%	
4ta.	160	179	13	81	9%	54%	
5ta.	180	199	16	97	11%	65%	
6ta.	200	219	22	119	15%	79%	
7ma.	220	239	25	144	17%	96%	
8ava.	240	259	6	150	4%	100%	
Sumatoria			150				

Figura 49

Seguidamente ubique la pestaña insertar y ubique el ítem “insertar gráfico de columnas” y elija cualquiera de las opciones que se despliegan, en este caso se eligió el más básico (Figuras 50 a 52).

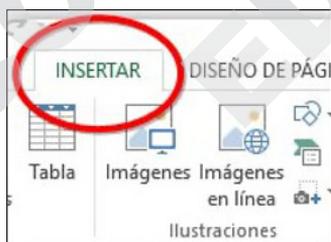


Figura 50



Figura 51

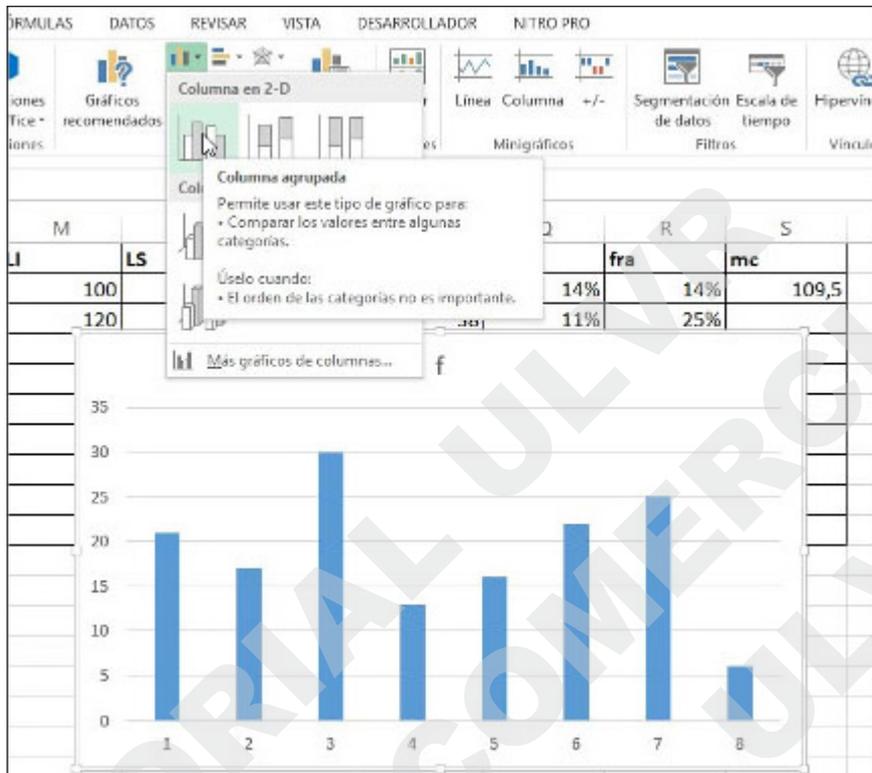


Figura 52

Una vez realizados estos sencillos pasos usted ya tendrá como resultado un histograma de frecuencias que le ayudará comprender el fenómeno estudiado de forma gráfica. Luego debe editar el gráfico para que contenga toda la información necesaria y que de esta manera cualquiera que quiera leer el gráfico pueda entenderlo ya que debe recordar que no se grafica para entenderlo uno mismo sino que se lo hace para que cualquier persona lo entienda.

Para editar debe al menos insertar un título en la gráfica y colocar las etiquetas en los ejes, también se recomienda insertar las frecuencias en las columnas, a continuación se editará la gráfica de la Figura 52.

Para ello ubíquese en la pestaña PESO PERSONAS TERCERA EDAD del archivo de Excel CAPÍTULO 1.xls. Note que es el mismo ejercicio ya trabajado pero ya editado para una presentación formal, además cada columna está plenamente identificada y con sus respectivas unidades de medición lo cual es muy importante al momento de editar el gráfico.

En la columna A se utilizó un conector “&” que permite realizar cadenas entre texto y números (Figura 53). Las columnas B, C, D y E muestran la frecuencia, la frecuencia acumulada, la frecuencia relativa y la frecuencia relativa acumulada respectivamente (Figura 54).



Figura 53

A	B	C	D	E
Peso 150 personas 3ra. Edad (libras)	f	fa	fr (%)	fra (%)
100 - 119	21	21	14,00	14,00
120 - 139	17	38	11,33	25,33
140 - 159	30	68	20,00	45,33
160 - 179	13	81	8,67	54,00
180 - 199	16	97	10,67	64,67
200 - 219	22	119	14,67	79,33
220 - 239	25	144	16,67	96,00
240 - 259	6	150	4,00	100,00
	150		100,00	

Figura 54

Ahora bien, siguiendo el proceso antes descrito inserte un gráfico para proceder a la edición.

Dando clic derecho en la gráfica que insertó ubique SELECCIONAR DATOS y elija esa opción (Figura 54), se le abrirá un cuadro de diálogo (Figura 55), aquí usted podrá insertar –entre otras cosas– el nombre que le asignará a la gráfica así como corregir los valores asignados al eje de las abscisas.

Comenzaremos por asignar el título de la gráfica: en el lado izquierdo elija MODIFICAR (Figura 56) y en el cuadro de diálogo (Figura 57) elija la celda A1, que es donde se encuentra escrito el título que le dará a la gráfica (Figura 58); dele clic en ACEPTAR para volver al cuadro de diálogo anterior, en donde se ubicará ahora en el lado derecho (Figura 59), aquí elegirá EDITAR y a continuación en el cuadro de diálogo elegirá los valores de la columna A (Figura 60).

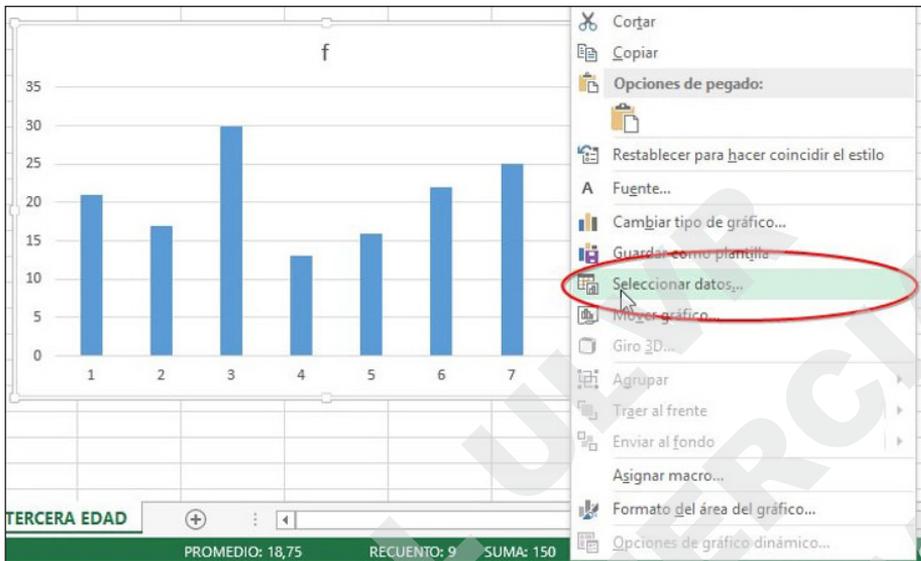


Figura 55



Figura 56

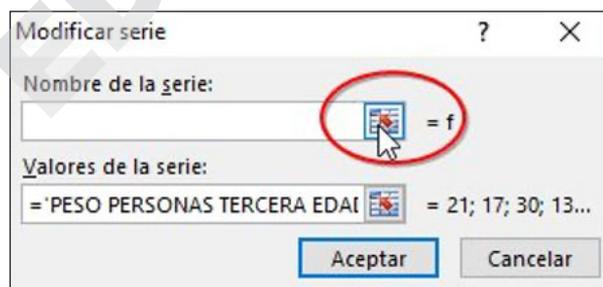


Figura 57

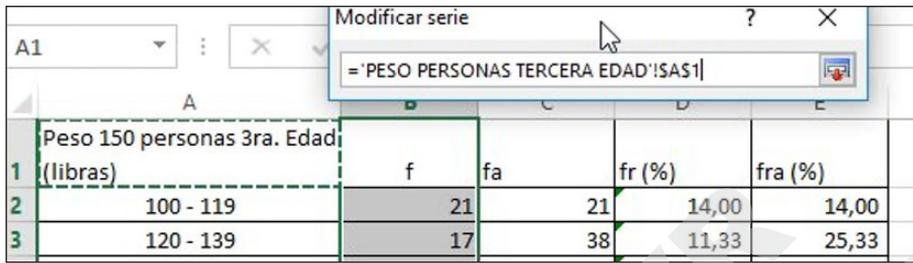


Figura 58

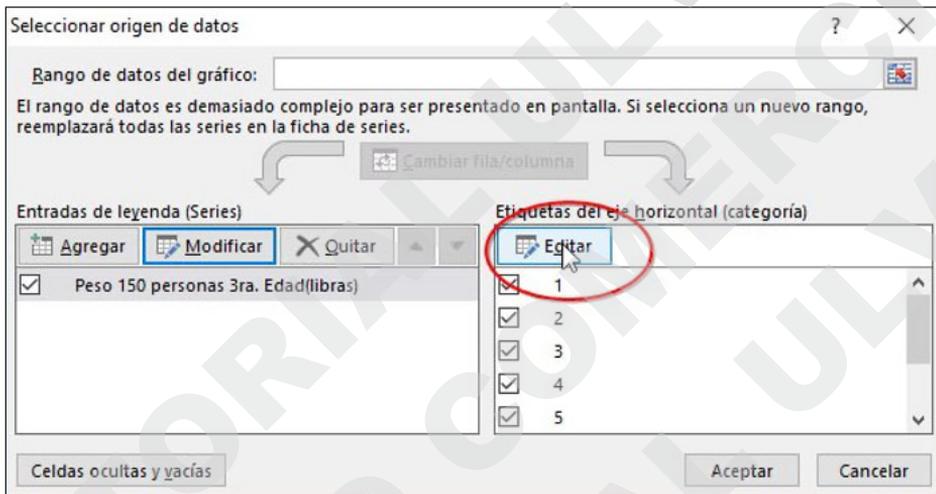


Figura 59

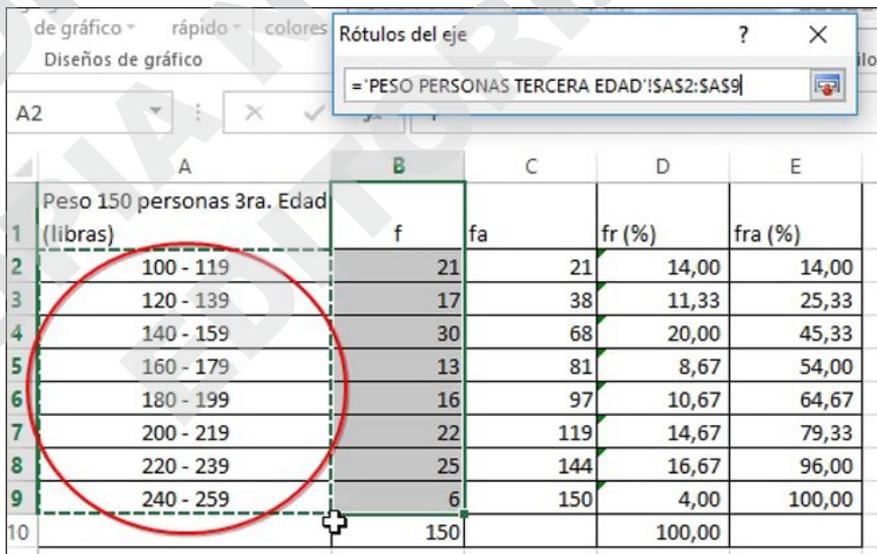


Figura 60

El resultado se puede apreciar en la Figura 61. Ahora bien, todo gráfico debe contener la mayor cantidad de información posible de tal modo que el lector pueda comprender de manera fácil lo que éstos quieren expresar; por lo que se deben agregar etiquetas a las barras que permitan observar gráfica y numéricamente el comportamiento del fenómeno estudiado por lo que dando click derecho sobre el gráfico, se desplegarán opciones de edición que le ayudarán a incorporar toda la información que usted considere relevante.

El resultado general se puede apreciar en la Figura 62, tomando en cuenta que se puede editar el gráfico de tal forma que pueda incluir los títulos de los ejes o en lugar de las frecuencias absolutas incluir las frecuencias relativas; inclusive incluir ambas (f y fa). Puede personalizar los colores de las barras, etc.

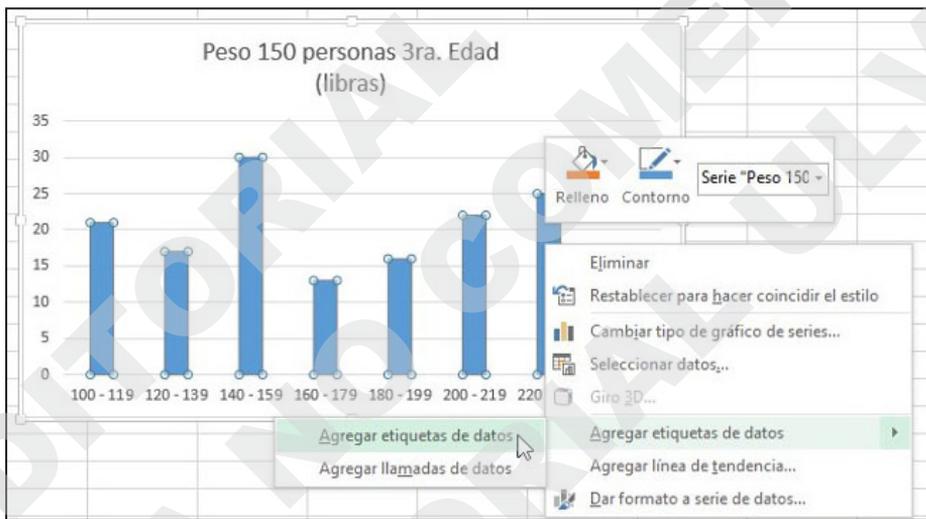


Figura 61

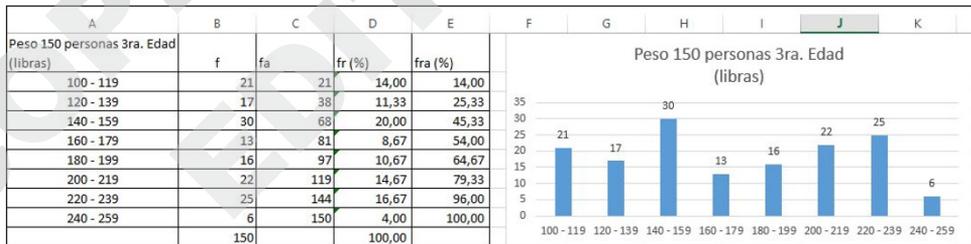


Figura 62



Aunque se debe hacer hincapié en el hecho de que los gráficos deben tener la mayor cantidad de información posible para que sean de fácil comprensión, se debe mantener lo más plano que se pueda para que el resultado final proyecte profesionalismo; procure evitar utilizar formatos en 3D o colores que ocasionen un ruido visual (que no resulte agradable a la vista o no parezca profesional).

1.9.2 Polígono de frecuencias

Un polígono de frecuencias es una gráfica lineal que sirve para representar visualmente una distribución de frecuencias.

Continuando con el ejemplo de las 150 personas de la tercera edad, abra el archivo CAPÍTULO I.xls ubicando la pestaña TABLA 1 MODIFICADA, seleccione la columna de la frecuencia, seguidamente ubique la pestaña insertar y ubique el ítem INSERTAR GRÁFICO DE LÍNEAS, y elija la opción de LÍNEA 2D, el resultado será un histograma básico que usted debe editar de tal forma que contenga toda la información necesaria para una fácil lectura.

Para editar la gráfica siga los mismos pasos que utilizó en el apartado anterior recordando que para un resultado más profesional debe utilizar –en lo posible– figuras plana y con colores que no ocasionen ruido visual. Todo el proceso a seguir se muestra en las Figuras 63 a 70.

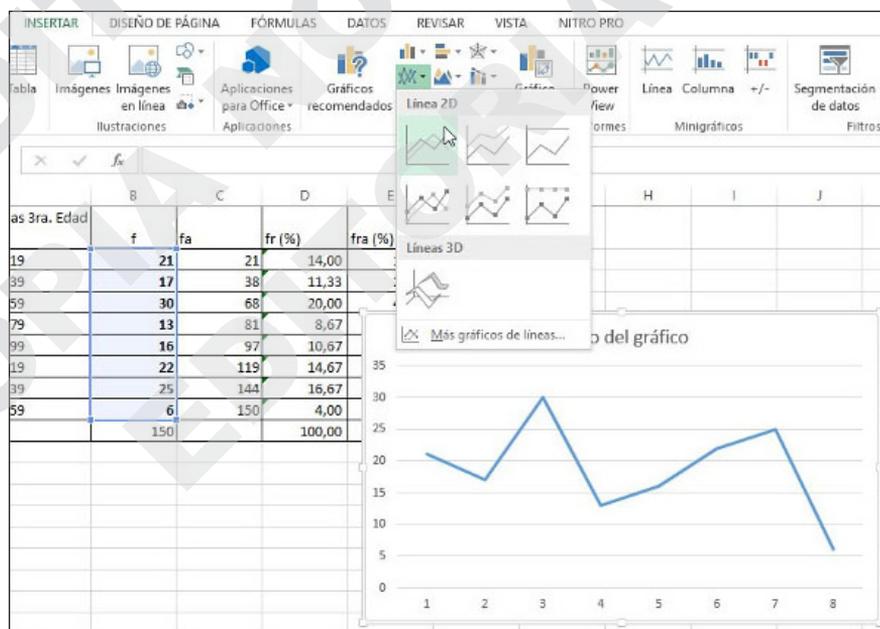


Figura 63

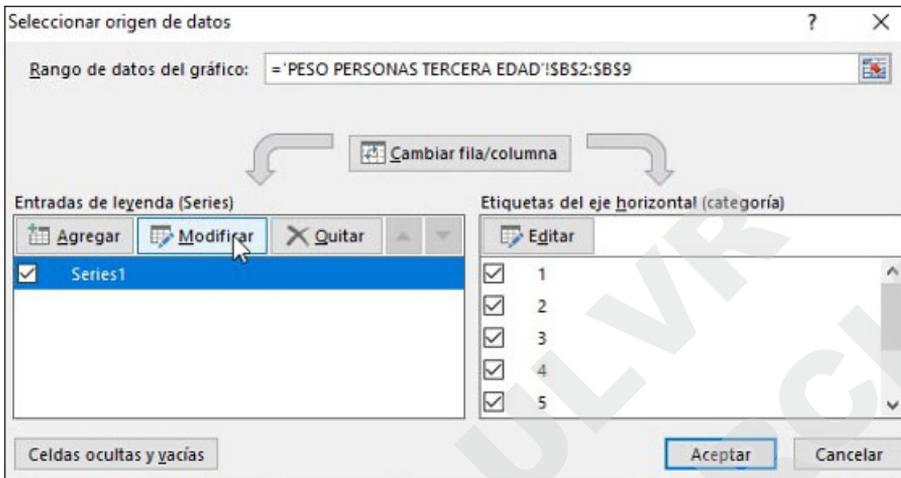


Figura 64

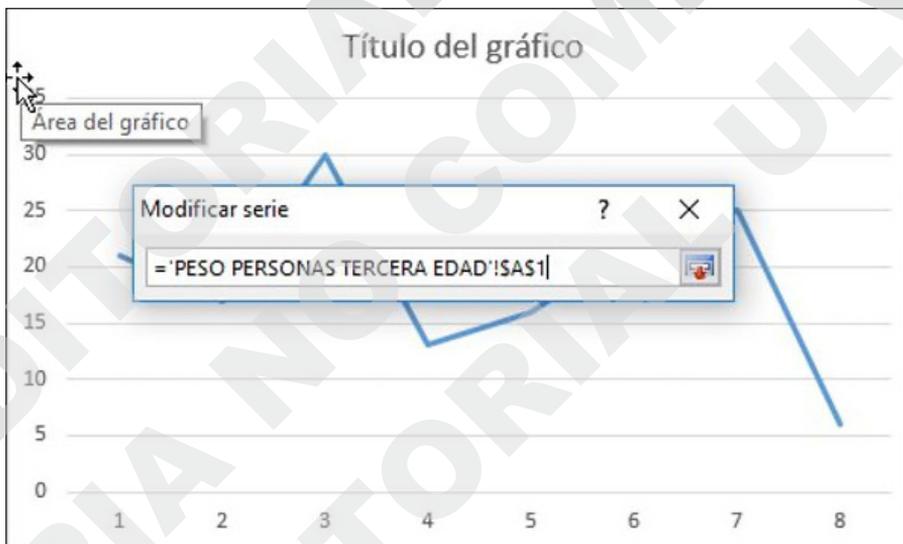


Figura 65

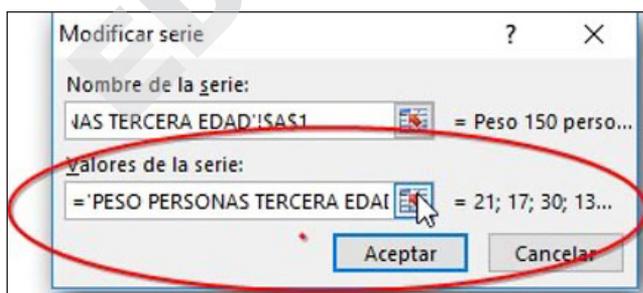


Figura 66

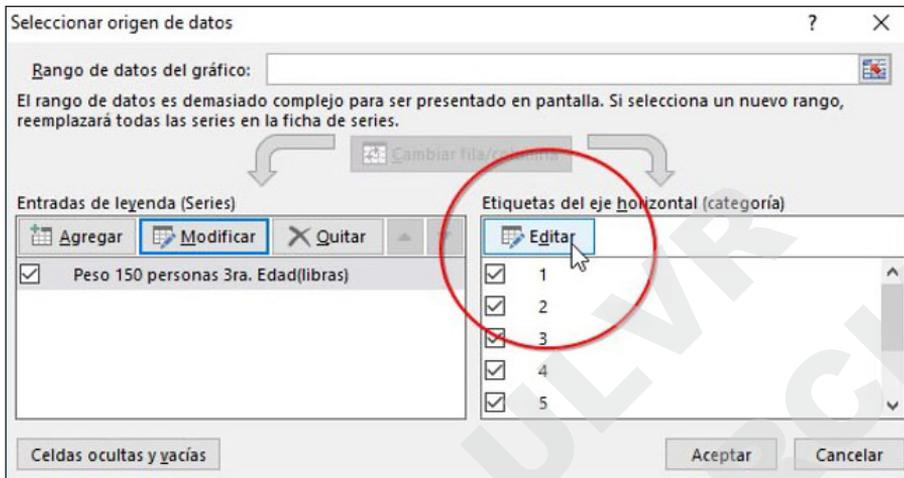


Figura 67

Peso 150 personas 3ra. Edad (libras)	f	f	f	f	f
100 - 119	21	38	11,33	25,33	35
120 - 139	17	68	20,00	45,33	30
140 - 159	30	81	8,67	54,00	25
160 - 179	13	97	10,67	64,67	20
180 - 199	16	119	14,67	79,33	15
200 - 219	22	144	16,67	96,00	10
220 - 239	25	150	4,00	100,00	5
240 - 259	6				
	150		100,00		

Rotulos del eje

RCERA EDAD!\$A\$2:\$A\$9

Figura 68

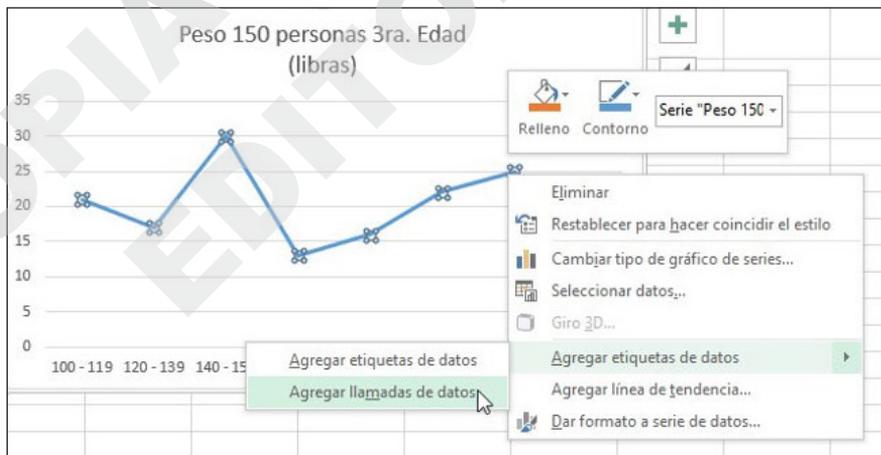


Figura 69



Figura 70

1.9.3 Ojiva

Una ojiva muestra el incremento que sufre la frecuencia cuando se acumula, por lo que básicamente es incremental, esto se debe a que se utiliza la frecuencia absoluta acumulada para representarla por lo que siempre será creciente.

Continuando con el ejemplo de las 150 personas de la tercera edad, abra el archivo CAPÍTULO I.xls ubicando la pestaña TABLA 1 MODIFICADA, seleccione la columna de la frecuencia acumulada (fa), seguidamente ubique la pestaña INSERTAR y ubique el ítem INSERTAR GRÁFICO DE LÍNEAS, y elija la opción de LÍNEA 2D (Figuras 71 y 72), el resultado será una ojiva básica que usted debe editar de tal forma que contenga toda la información necesaria para una fácil lectura.

Peso 150 personas 3ra. Edad (libras)	f	fa	fr (%)	fra (%)
100 - 119	21	21	14,00	14,00
120 - 139	17	38	11,33	25,33
140 - 159	30	68	20,00	45,33
160 - 179	13	81	8,67	54,00
180 - 199	16	97	10,67	64,67
200 - 219	22	119	14,67	79,33
220 - 239	25	144	16,67	96,00
240 - 259	6	150	4,00	100,00
	150		100,00	

Figura 71

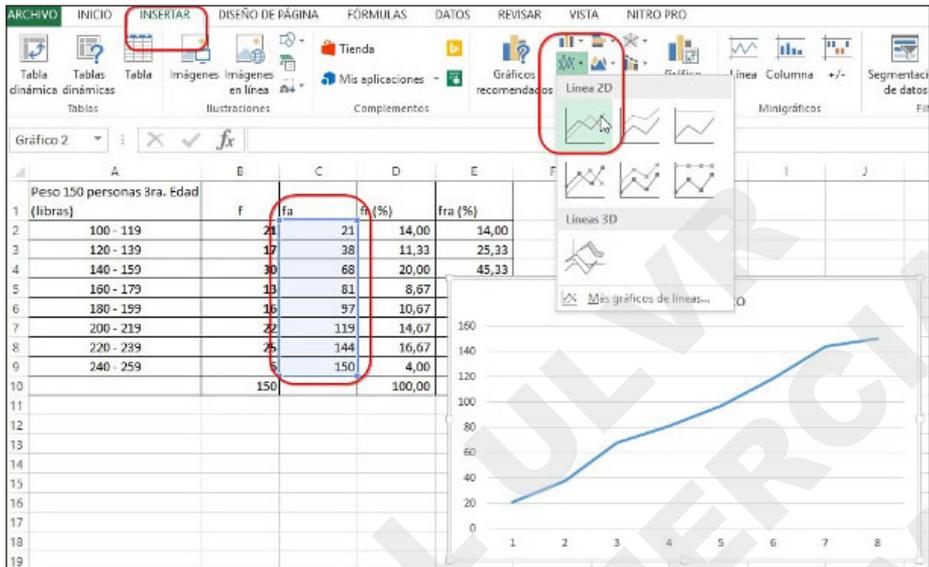


Figura 72

Para editar el gráfico debe dar click derecho sobre éste y al desplegarse el menú elegir SELECCIONAR DATOS tal como lo muestra la Figura 73.

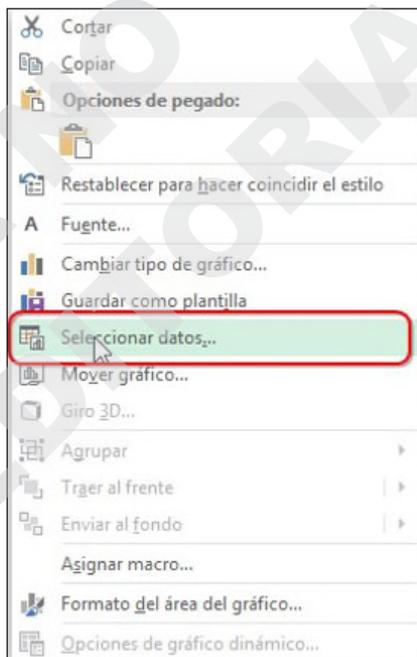


Figura 73

Una vez realizada la acción anterior debe elegir MODIFICAR sobre el lado derecho en la casilla ENTRADAS DE LEYENDA (Series), a diferencia de gráficas anteriores no elija la celda donde se encuentra el título de la tabla sino escriba manualmente la palabra OJIVA tal como lo ilustra la Figura 74.

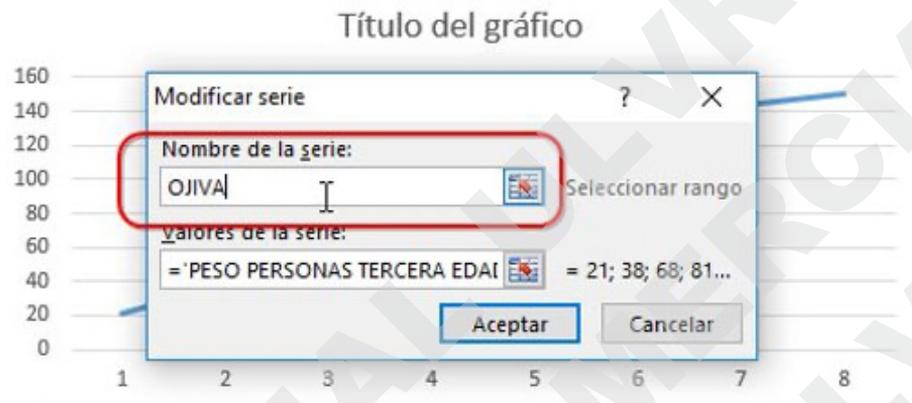


Figura 74

Adicionalmente, en la pestaña DISEÑO, puede elegir el formato que más se adapte a sus necesidades, siempre recordando las recomendaciones hechas en apartados anteriores sobre la sobriedad de los mismos; en este caso solo por cuestiones de ejemplificar se eligió el cuarto modelo, contado de izquierda a derecha, pero se pudo haber elegido cualquier otro ya que eso depende de sus gustos y preferencias (Figura 75).

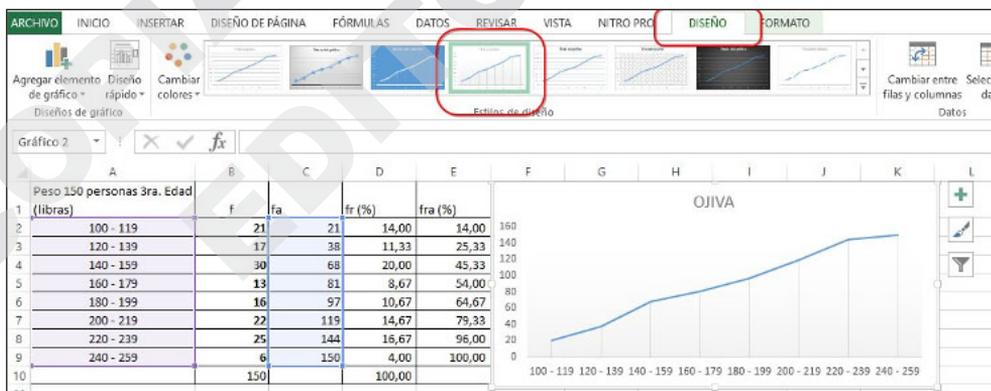


Figura 75

Dando click derecho sobre la ojiva se despliega un menú donde debe elegir AGREGAR ETIQUETAS DE DATOS y nuevamente AGREGAR ETIQUETA DE DATOS en el submenú que se despliega a continuación con el fin de incluir cada una de las frecuencias acumuladas en la gráfica. Luego si desea puede arrastrar cada valor hacia el lugar donde usted crea que visualmente sea más agradable (ver Figuras 76 y 77).

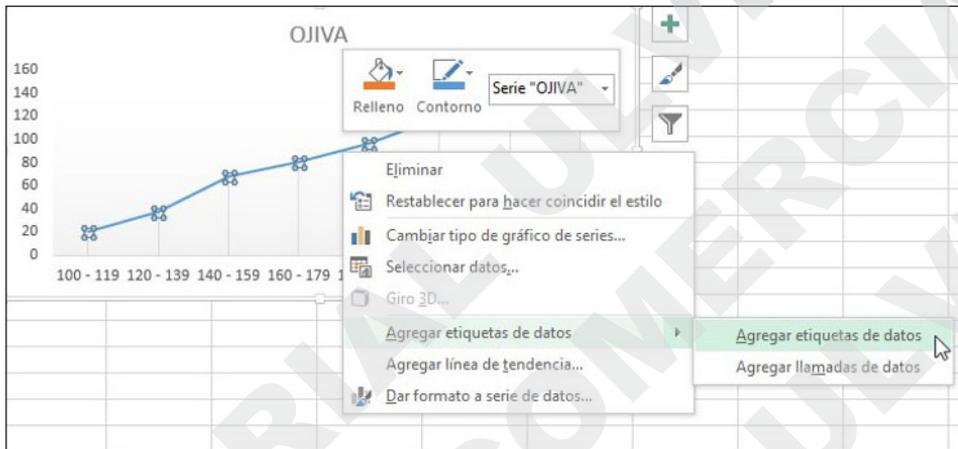


Figura 76

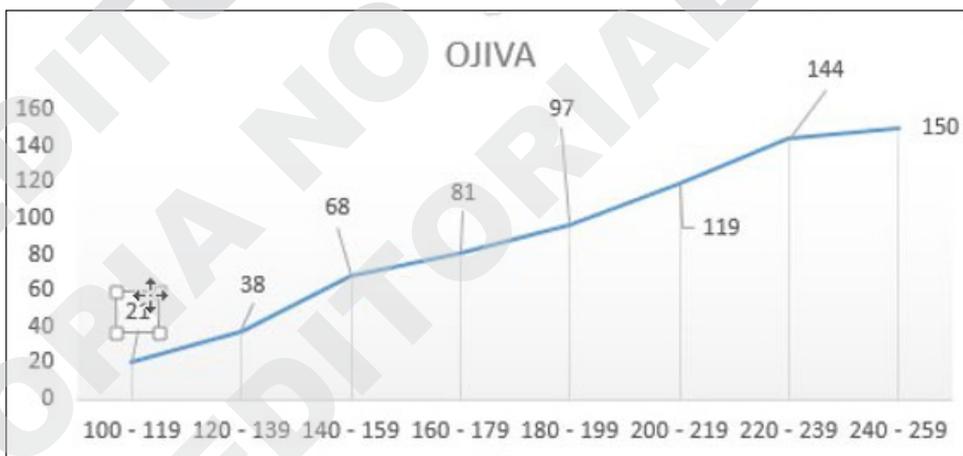


Figura 77

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Capítulo 2

Medidas de Tendencia Central

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR



Capítulo 2

Medidas de Tendencia Central

2.1 Generalidades

Las medidas de tendencia central agrupan a las observaciones de un fenómeno de manera tal, que le dan al investigador una idea general de cómo se comportan; por esto, es muy común que tanto el estudiante como el profesional acudan a las medidas de tendencia central para representar el comportamiento aproximado de dichos fenómenos.

Aunque existen varias medidas de tendencia central, en el presente trabajo se abordarán las que, a criterio del autor, son las más importantes: media aritmética, mediana, moda, y un caso especial que es la media geométrica, la que se utilizará bajo las circunstancias que se detallan más adelante.

Cabe destacar que en esta obra no se abordarán las medidas de tendencia central para datos agrupados, puesto que para el autor y debido a la gran capacidad de procesamiento de Excel, éstas no tienen razón de ser, y paso a explicar el por qué.

A través de la experiencia, he podido notar que quienes utilizan medidas de tendencia central para representar algún fenómeno, creen que deben aplicar todas al mismo tiempo, cuando esto no es necesariamente lo correcto, debido a que dependiendo de lo que se investigue, la elección correcta puede variar en gran medida una de otra.

Se aconseja que quienes deseen interpretar el comportamiento aproximado de un fenómeno elijan la que más se ajuste al estudio del mismo. La media

aritmética es la medida de tendencia central por excelencia, así que no es raro que al momento de generalizar un fenómeno se la utilice para generalizar un aspecto relevante, sin embargo es muy sensible ante observaciones atípicas, es decir, observaciones extremadamente altas o bajas; en esos casos, se acude a la mediana; y sólo si se quiere tener una idea rápida y muy general del fenómeno, a la moda.

Por ende antes de elegir cuál utilizar, se deben conocer las características de la información a fin de hacer un uso correcto de la medida de tendencia central más adecuada, tomando en cuenta que se debe evitar lo más posible reportar las tres al mismo tiempo para describir un fenómeno.

2.2 Media Aritmética

La Media Aritmética es la medida de tendencia central por excelencia, y no es más que la suma de todas las observaciones dividida para el total de dichas observaciones.

- **Ejemplo 3:** Suponga que quiere saber cuánto gastó en promedio en entradas al estadio del equipo de su preferencia durante el presente año, para ello revisa una muestra de las contraseñas de 10 partidos a los que acudió y los anota por separado; los resultados se recogen en la tabla 14 (la información y cálculos los encontrará en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 14).

5	8	18	6	8	8	10	15	15	20
---	---	----	---	---	---	----	----	----	----

- Para conocer el promedio de forma rápida solo basta seleccionar los valores y observar el resultado rápido mostrado en la parte inferior derecha de la hoja de cálculo, donde además del promedio se muestra el recuento de los datos y la sumatoria total (Figura 78).

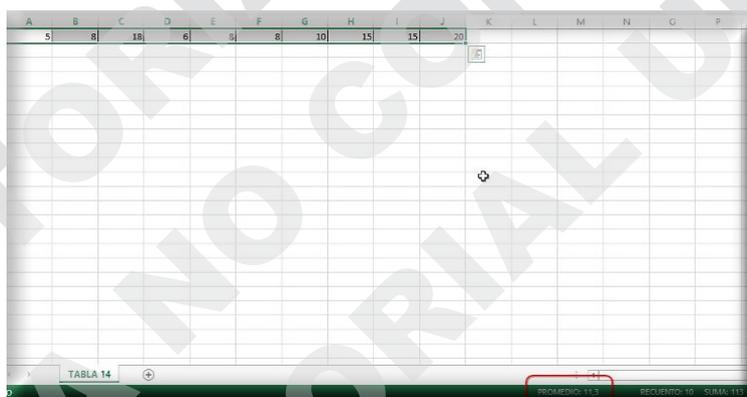


Figura 78

Ahora bien, si lo que se desea es dejar el resultado indicado en la hoja de cálculo, existen 2 formas de hacerlo:

La primera forma emplea una fórmula que suma todos los valores y el resultado lo divide para el tamaño de la muestra. Para la primera parte de la fórmula se usa la función SUMA y luego se divide para el total de los datos observados.

Todos los cálculos que se realizan en Excel empiezan con el ingreso del condicional IGUAL (=); sin embargo, existe una manera aún más fácil que consiste en escribir el condicional MAS (+) que resulta equivalente al igual al

momento de comenzar un cálculo, se destaca que utilizar el condicional (+) porque en las mayoría de las computadoras para escribir el símbolo (=) se debe utilizar la tecla SHIFT. El cálculo de la media y su fórmula se muestran en la Figura 79.

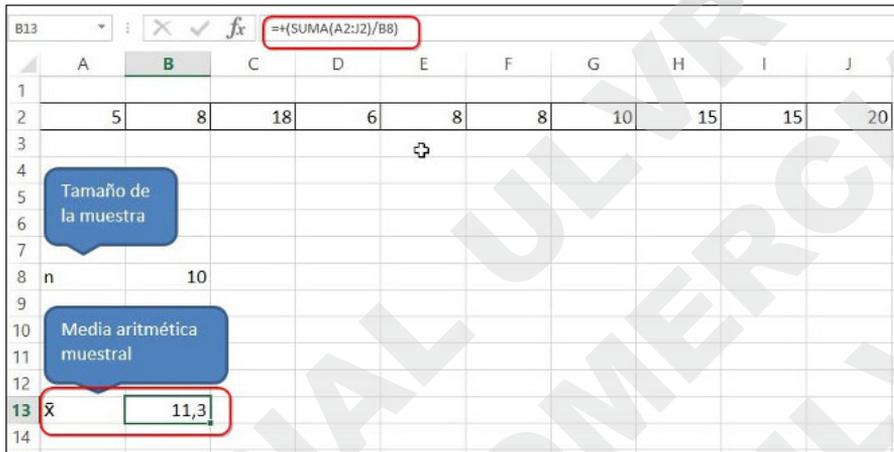


Figura 79

En la celda B13 se observa el resultado del cálculo de la media aritmética (\bar{x}) y en la celda B8 el tamaño de la muestra (n); cabe destacar que cuando el número de observaciones es demasiado grande, se puede utilizar la función CONTAR, seleccionar el rango de datos deseado y el resultado será el conteo de las observaciones. Para hallar \bar{x} se inició con la función SUMA, luego se seleccionó todos los datos desde A2 hasta J2, y finalmente se lo dividió para el tamaño de la muestra que se encuentra en la celda B8. La fórmula completa se muestra en la Figura 80.

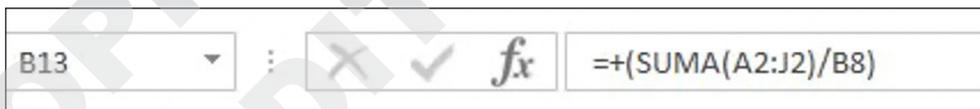


Figura 80

Ejercicios propuestos

41. Suponga que se le designa la tarea de realizar un estudio de mercado para un grupo objetivo específico, el tamaño de la muestra (n) es de 544. La consulta fue acerca de cuántos están dispuestos a pagar por un artículo en particular. Las respuestas se encuentran en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 15. ¿Cuál es valor promedio que están dispuestos a pagar los consultados?
42. El peso en libras de 50 pacientes de una clínica de estética fue controlado durante una semana. Los resultados se encuentran en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 16. ¿Cuál es el peso promedio de los pacientes?
43. El consumo de energía de 6.500 familias de una comunidad fue redondeado al KW más cercano, y registrado en una base de datos con la finalidad de proponer un plan de energía limpia. Los resultados se encuentran en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 17. ¿Cuál es el consumo medio de energía de las familias?

2.3 Mediana

Cuando los valores observados no son homogéneos, la media aritmética puede no ser representativa para el conjunto de observaciones, mucho más cuando existen valores extremos (valores muy altos o muy bajos), también conocidos como atípicos. Entonces lo adecuado es utilizar la mediana, tomando en cuenta que la mediana solo debe utilizarse bajo estas circunstancias puesto que ésta solo toma en cuenta un valor central cuando el número de observaciones es impar, y dos valores centrales cuando el número de observaciones es par.

Para encontrar la mediana se debe empezar con el símbolo (+) para luego escribir la función mediana y seleccionar los datos. El resultado será el valor mediano, tal como lo muestran las Figuras 81 a 84.

	310	657	540	734	73	64	698	492	590	640
MEDIANA										

Figura 81

	A	B	C	D	E	F	G	H
	310	657	540	734	73	64	698	492
MEDIANA		+MED						
		<ul style="list-style-type: none"> ⊗ MEDIA.ACOTADA ⊗ MEDIA.ARMO ⊗ MEDIA.GEOM ⊗ MEDIANA 						

Devuelve la mediana o el número central de un conjunto de números

Figura 82

B3										
	A	B	C	D	E	F	G	H	I	J
1	310	657	540	734	73	64	698	492	590	640
2										
3	MEDIANA	+MEDIANA(A1:J1)								
4										

Figura 83

B3										
	A	B	C	D	E	F	G	H	I	J
1	310	657	540	734	73	64	698	492	590	640
2										
3	MEDIANA	565								

Figura 84

Ejercicios propuestos

44. Encuentre la mediana en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 14.
45. Encuentre la mediana en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 15.
46. Encuentre la mediana en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 16.
47. Encuentre la mediana en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 17.

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

2.4 Moda

La moda es tal vez la medida de tendencia central menos utilizada pues solo sirve para tener una idea muy general de lo que está sucediendo dentro del fenómeno estudiado debido a la característica de la misma (es simplemente el valor que más se repite).

La moda no existe si hay valores que se no se repiten, y si se observan varios valores con la misma frecuencia, entonces existen varias modas, por lo que en ambos casos no es de utilidad.

La mejor manera de conseguir encontrar la moda es seleccionando un grupo de celdas donde aparecerán las posibles modas (ya que no se sabe con antelación si existe una o varias modas cuando se trata de un gran número de observaciones) en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 19, tal como lo muestra la Figura 85.

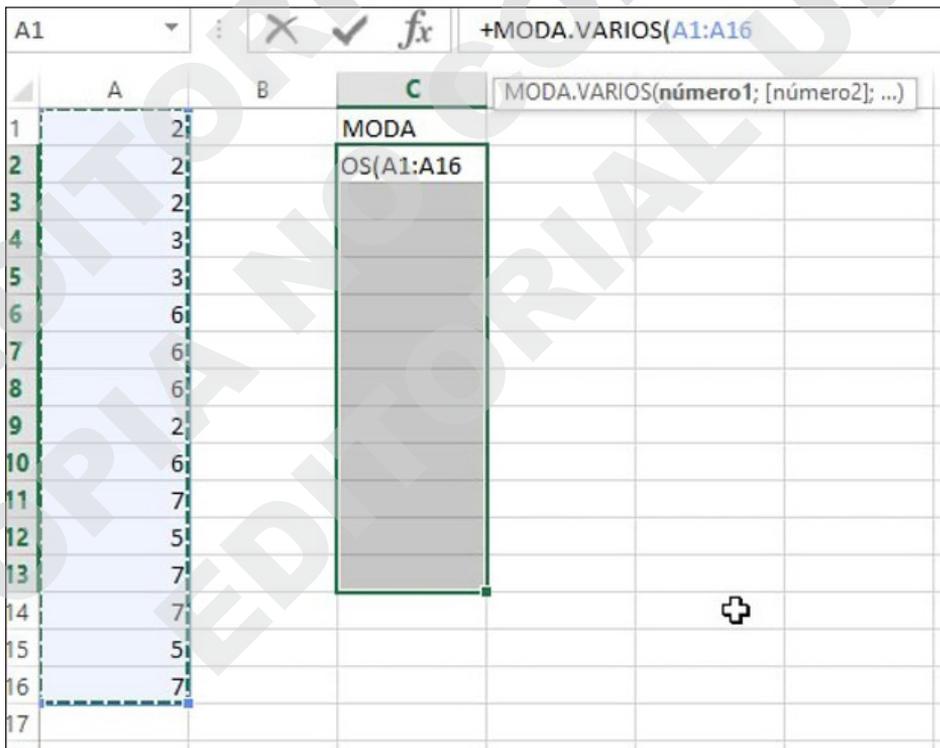


Figura 85

Luego se escribe en la barra de fórmulas la función “MODA. VARIOS” seleccionando a continuación el rango de observaciones; acto seguido (este paso es muy importante ya que transforma la fórmula en matricial) se debe presionar al mismo tiempo las teclas CTRL+SHIFT+ENTER. El resultado será una o varias modas dependiendo el caso (Figuras 86 y 87).

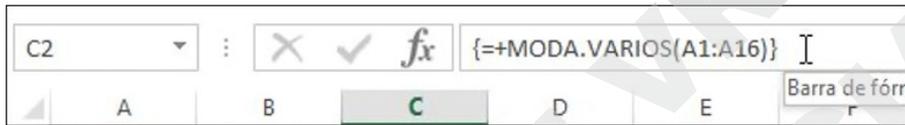


Figura 86

The image shows the Excel spreadsheet with the results of the `MODA.VARIOS` function. The results are displayed in column C, starting from row 1. The first cell (C1) contains the text 'MODA'. The subsequent cells (C2 to C13) contain numerical values: 2, 6, 7, and then a series of #N/A errors. The data in column A is as follows:

	A	B	C	D	E
1			MODA		
2	2		2		
3	2		6		
4	3		7		
5	3		#N/A		
6	6		#N/A		
7	6		#N/A		
8	6		#N/A		
9	2		#N/A		
10	6		#N/A		
11	7		#N/A		
12	5		#N/A		
13	7		#N/A		
14	7				
15	5				
16	7				

Figura 87

Ejercicios propuestos

48. Encuentre la moda en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 14.
49. Encuentre la moda en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 15.
50. Encuentre la moda en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 16.
51. Encuentre la moda en el archivo de Excel llamado CAPÍTULO II, en la pestaña TABLA 17.



Cuando reporte medidas de tendencia central con el fin de describir fenómenos, no es necesario reportar las tres (media aritmética, mediana y moda), inclusive podría resultar confuso si las tres tienen valores muy diferentes. Lo correcto sería utilizar en primer lugar la media aritmética ya que esta es la medida de tendencia central por excelencia; y luego, en el caso de que las observaciones no sean homogéneas, utilizar la mediana; finalmente, si lo que desea es hacer una descripción muy general y solo así, utilizar la moda.

2.5 Media Geométrica

La media geométrica es un el caso especial, pues se utiliza sólo cuando lo que se quiere promediar se encuentra expresado en valores relativos, es decir, porcentajes, es por eso que es muy útil al momento de promediar índices o valores porcentuales.

Para el cálculo de la MG (media geométrica) se debe utilizar la función “MEDIA.GEOM” tal como lo muestran la Figuras 88 a 93.

The screenshot shows an Excel spreadsheet with the following data in column A:

	A	B	C
1	32%		
2	36%		
3	35%		
4	64%		
5	26%		
6			
7	MG		
8			
9			

The formula bar for cell B7 shows the function f_x being entered.

Figura 88

The screenshot shows the same Excel spreadsheet as Figure 88, but with the function list open for the formula bar. The list includes:

- MEDIA.ACOTADA
- MEDIA.ARMO
- MEDIA.GEOM
- MEDIANA

The description for the selected function, MEDIA.GEOM, is: "Devuelve la media geométrica de una matriz o rango de datos numéricos positivos".

Figura 89

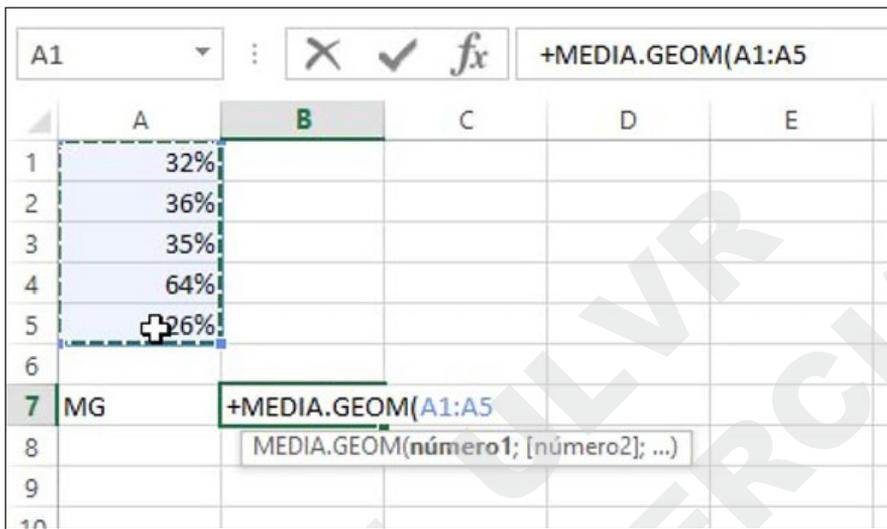


Figura 90

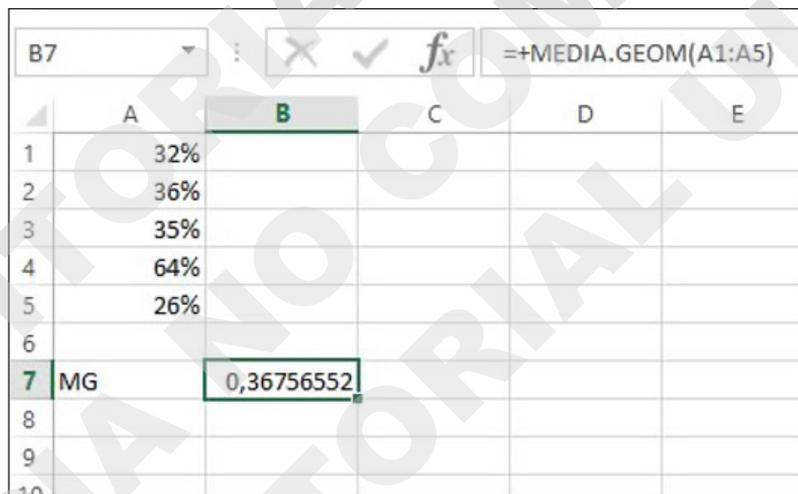


Figura 91



Figura 92

		B7 : \times \checkmark f_x =+MEDIA.GEOM(A1:A5)				
	A	B	C	D	E	
1	32%					
2	36%					
3	35%					
4	64%					
5	26%					
6						
7	MG	37%				
8						
9						

Figura 93

EDITORIAL ULVR
 COPIA NO COMERCIAL
 EDITORIAL ULVR



Bibliografía

Lind, D., Marchal, S. y Wathen, S. (2012). *Estadística aplicada a los negocios y la economía* (15ª ed.). McGrawHill Educación.

Universidad Nacional Autónoma de México. (s. f.). Tablas de Frecuencia. *Probabilidad y Estadística FI UNAM*. <http://probabilidadyestadisticafiunam.blogspot.com/p/blog-page.html>

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR

EDITORIAL ULVR
COPIA NO COMERCIAL
EDITORIAL ULVR





Estadística descriptiva con Excel 1, en palabras de su autor, no se trata de un libro de estadística por sí mismo sino una herramienta didáctica para que estudiantes y profesionales puedan realizar cálculos estadísticos, más ágiles y confiables, utilizando Excel.

A través de ejemplos desarrollados, el autor guía al lector paso a paso en la aplicación de las herramientas de Excel para realizar los cálculos más utilizados en la Estadística, como son, la media aritmética, la mediana, la media geométrica, así como graficar los resultados usando los distintos formatos de presentación con que cuenta Excel.

Complementando al texto, 51 ejercicios se adjuntan en un archivo de Excel para que el lector desarrolle las competencias propuestas en el libro, de tal manera que el autor se convierte en un tutor que interactúa con él y lo lleva a aplicar, repasar y afianzar el conocimiento que se adquiere a medida que se avanza en el dominio de las técnicas propuestas.

