GUÍA DE PRÁCTICA

MEDIDAS DE TENDENCIA CENTRAL EN VARIABLE DISCRETA: DATOS DESAGRUPADOS Y AGRUPADOS CON MICROSOFT EXCEL

Diana María Robayo Botiva
Universidad Cooperativa de Colombia
Sede Villavicencio

Apropiación social del conocimiento
Generación de contenidos impresos
<https://repository.ucc.edu.co/handle/20.500.12494/7375>
N.º 01, febrero de 2020
doi: <https://doi.org/10.16925/gcgp.23>

NOTA LEGAL

El presente documento de trabajo ha sido incluido dentro de nuestro repositorio institucional como Apropiación social de conocimiento por solicitud del autor, con fines informativos, educativos o académicos. Asimismo, los argumentos, datos y análisis incluidos en el texto son responsabilidad absoluta del autor y no representan la opinión del Fondo Editorial o de la Universidad.

DISCLAIMER

This coursework paper has been uploaded to our institutional repository as Social Appropriation of Knowledge due to the request of the author. This document should be used for informational, educational or academic purposes only. Arguments, data and analysis included in this document represent authors' opinion not the Press or the University.

EDICIONES
Universidad Cooperativa
de Colombia

GUÍA DE PRÁCTICA

**MEDIDAS DE TENDENCIA
CENTRAL EN VARIABLE
DISCRETA: DATOS
DESAGRUPADOS Y
AGRUPADOS CON
MICROSOFT EXCEL**

Diana María Robayo Botiva

ACERCA DE LA AUTORA

Diana María Robayo Botiva, magíster en Comercio Electrónico del Instituto Tecnológico y de Estudios Superiores de Monterrey (México). Profesora de la Facultad de Ciencias Económicas, Administrativas y Contables de la Universidad Cooperativa de Colombia, sede Villavicencio, Colombia.

Correo-e: diana.robayo@campusucc.edu.co

CvLAC: http://scienti.colciencias.gov.co:8081/cv lac/visualizador/generarCurriculoCv.do?cod_rh=0000597864

Google Scholar: <https://scholar.google.com/citations?user=XL1poXcAAAAJ&hl=es>

CÓMO CITAR ESTE DOCUMENTO

Robayo-Botiva, D. M. (2020). *Medidas de tendencia central en variable discreta: datos desagrupados y agrupados con Microsoft Excel* (Generación de contenidos impresos N.º 1). Bogotá: Ediciones Universidad Cooperativa de Colombia. doi: <https://doi.org/10.16925/gcgp.23>

Este documento puede ser consultado, descargado o reproducido desde nuestro repositorio institucional (<http://repository.ucc.edu.co/handle/20.500.12494/7369>) para uso de sus contenidos, bajo la licencia de Creative Commons Reconocimiento-No Comercial-Sin Obra Derivada 4.0 Internacional. <http://creativecommons.org/licenses/by-nc-nd/4.0/>

TABLA DE CONTENIDO

INTRODUCCIÓN	7
PROPÓSITO	7
MARCO TEÓRICO	7
Media aritmética: $\bar{X}=\mu$	7
Media aritmética simple	7
Media aritmética ponderada	7
Cálculo de la media utilizando frecuencias relativas	7
Mediana (M_e)	8
Cálculo de datos sin agrupar u originales	8
Cálculo de datos agrupados o en tabla de frecuencias	8
Moda o valor modal (M_o)	8
OBJETIVOS	8
Materiales	8
DESCRIPCIÓN DE ACTIVIDADES Y PROCEDIMIENTOS DE LA PRÁCTICA	9
Ejercicio datos desagrupados	9
Variable discreta	9
Solución	9
Ejercicio datos agrupados	15
TRABAJO PARA LOS ESTUDIANTES	23
REFERENCIAS	24

LISTA DE TABLAS

Tabla 1. Tabla datos cuantitativos discretos desagrupados	9
Tabla 2. Datos desagrupados organizados de menor a mayor	13
Tabla 3. Datos resaltados en azul para cálculo de mediana	13

LISTA DE FIGURAS

Figura 1. Datos en hoja de Microsoft Excel	9
Figura 2. Selección de la fórmula para calcular la media aritmética	10
Figura 3. Selección de datos para cálculo de la media aritmética	10
Figura 4. Resultado de la media aritmética	11
Figura 5. Selección de datos para cálculo de la mediana	11
Figura 6. Selección de grupos para cálculo de la mediana	12
Figura 7. Resultado de la mediana	12
Figura 8. Selección de fórmula para cálculo de moda	14
Figura 9. Selección de datos para el cálculo de la moda	14
Figura 10. Resultado de la moda	15
Figura 11. Elaboración de tabla de frecuencias para datos agrupados	15
Figura 12. Selección de fórmula para cálculo de la frecuencia absoluta	16
Figura 13. Selección de datos para cálculo de la frecuencia absoluta	16
Figura 14. Selección de grupo para cálculo de la frecuencia absoluta	17
Figura 15. Resultados de la frecuencia absoluta acumulada	17
Figura 16. Cálculo de la frecuencia relativa	18
Figura 17. Resultados de la frecuencia relativa	18
Figura 18. Cálculo de la frecuencia absoluta y relativa acumulada	19
Figura 19. Cálculo de $Y_i n_i$ para la media aritmética	20
Figura 20. Resultados de la columna $Y_i n_i$	20
Figura 21. Resultado de la media aritmética	21
Figura 22. Columna N_i resaltada para cálculo de mediana	22
Figura 23. Procedimiento para cálculo de la mediana	22

01 MEDIDAS DE TENDENCIA CENTRAL EN VARIABLE DISCRETA: DATOS DESAGRUPADOS Y AGRUPADOS CON MICROSOFT EXCEL

Diana María Robayo Botiva

Resumen

Paquetes como Microsoft Office facilitan el desarrollo de muchas tareas en distintos contextos, con lo que aumentan la productividad de los usuarios finales. Su sencillo manejo se debe a que cuentan con una GUI (interfaz gráfica de usuario) que incluye íconos, menús, barras de herramientas, etc., que hacen agradable su uso. Dentro del paquete de Microsoft Office, se encuentra Microsoft Excel que se utiliza en contabilidad y en finanzas, inventarios, ventas y en la producción o suministros, de modo que es susceptible de ser usado en cualquier área de las empresas. A su vez, sin importar su sector económico o tamaño las empresas utilizan este *software* debido a su flexibilidad y funcionalidad para el análisis de datos; por tanto, tiene lugar en la mayoría de procesos, y un conocimiento avanzado en su uso sugiere una ventaja competitiva (Telesup, 2017). Acorde a lo anterior, así como Excel es de gran utilidad en las empresas, también lo es en la cotidianidad. Por eso es pertinente que los estudiantes del área de ciencias económicas aprendan a utilizarla, de manera que conozcan todos sus beneficios, en especial en su campo disciplinar. Por ende, en esta guía se pretende que el estudiante utilice Excel en el desarrollo de ejercicios básicos de estadística descriptiva-variable discreta-datos desagrupados y agrupados, aplicando fórmulas de medidas de tendencia central, y que profundice en su uso en pro de su desarrollo profesional.

Palabras clave: fórmulas de Excel, media aritmética, mediana, moda.

INTRODUCCIÓN

En estadística descriptiva, en el contenido del programa se incluyen las medidas de tendencia central; por tanto, se plantea una práctica con la herramienta Microsoft Excel para que los estudiantes del programa de Administración de Empresas, además de ejercitar los temas vistos a través de ejercicios de contexto, puedan verificar los resultados haciendo uso de las fórmulas de Excel. Así mismo, la práctica genera un espacio para que el estudiante ahonde en esta herramienta ofimática y conozca todos sus usos, los cuales pueden ser aplicados en contextos reales.

PROPÓSITO

Practicar las temáticas de medidas de tendencia central para datos desagrupados y agrupados en variable cuantitativa discreta del curso de estadística descriptiva, desarrollando ejercicios de contexto y utilizando la herramienta Microsoft Excel.

MARCO TEÓRICO

Las medidas de tendencia central, conocidas también como promedios, nos permiten determinar la posición de un valor respecto a un conjunto de datos, al cual consideramos representativo o típico para el total de las observaciones. Cuando son aplicadas a las características de las unidades de una muestra, a estas medidas se les denomina estimadores o estadígrafos. En cambio, cuando son aplicadas a las características de los elementos de una población, se les conoce como parámetros o valores estadísticos de la población. Según la naturaleza de los datos, la necesidad que se tiene o la característica que ellos representan, requieren la aplicación de un promedio especial dentro de los diferentes tipos que se expondrán a continuación. Para ello, deberán tenerse en cuenta algunas

reglas que serán básicas en la selección de estas herramientas de descripción (Martínez, 2012).

MEDIA ARITMÉTICA: $\bar{X} = \mu$

Es la **medida de posición** o **promedio** más conocida, la más utilizada y entendida por todos; dada su gran estabilidad, es la preferida en el muestreo y sus fórmulas admiten tratamiento algebraico. Su desventaja principal es ser muy sensible a cambios en sus valores u observaciones, también cuando algunos de sus valores extremos son demasiado grandes o pequeños (Martínez, 2012).

Media aritmética simple

Se trabaja con datos sin agrupar u originales, siendo (Martínez, 2012):

$$\bar{x} = \frac{\sum x_i}{n} \text{ donde } \bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

(minúsculas en las muestras)

$$\bar{X} = \frac{\sum X_i}{N} \text{ donde } \bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

(mayúsculas y letras griegas en la población)

Media aritmética ponderada

Se aplica en datos agrupados, es decir, aquellos que se encuentran organizados en una tabla de frecuencia, siendo las frecuencias absolutas sus ponderaciones (Martínez, 2012):

$$\bar{y} = \frac{\sum y_i n_i}{n} \text{ donde } \bar{y} = \frac{y_1 n_1 + y_2 n_2 + y_3 n_3 + \dots + y_m n_m}{n_1 + n_2 + \dots + n_m}$$

Cálculo de la media utilizando frecuencias relativas

Esta medida también se puede calcular utilizando como ponderaciones a las frecuencias relativas, con las cuales se obtendrá el mismo resultado (Martínez, 2012):

$$\bar{y} = \frac{\sum y_i n_i}{n} \text{ donde } \bar{y} = \frac{y_1 n_1 + y_2 n_2 + y_3 n_3 + \dots + y_m n_m}{n_1 + n_2 + \dots} \text{ se sabe } h_i$$

$$= \frac{n_i}{n} \text{ por lo que se tiene que}$$

$$\bar{y} = y_1 h_1 + y_2 h_2 + \dots + y_m h_m \Rightarrow \bar{y} = \sum y_i h_i$$

MEDIANA (M_E)

Al igual que la media, es considerada también como una medida de tendencia central. Su importancia es menor y sus fórmulas son rígidas ya que no admiten tratamiento algebraico; es por eso que se tendrán que aplicar seis procedimientos de cálculo diferente (que lo hace engorroso), como lo veremos a continuación:

Se define como “aquel valor de la variable que supera a no más de la mitad de las observaciones, al mismo tiempo, es superado por no más de la mitad de las observaciones” (Martínez, 2012); en otras palabras, se puede definir como el “valor central”. Se simboliza por M_E .

Cálculo de datos sin agrupar u originales

- Número impar de observaciones: cuando esta medida se aplica en datos originales o sin agrupar, lo primero que se debe hacer es organizarlos de menor a mayor o de mayor a menor. Para encontrar la posición del valor que corresponde a la mediana, se aplica $\frac{n+1}{2}$.
- Número par de observaciones: de la misma manera, se deben organizar los datos de menor a mayor o de mayor a menor; sin embargo, al aplicar $\frac{n+1}{2}$, se encontrarán dos valores en el centro de la serie, por lo cual la mediana será el promedio de ellos (Martínez, 2012).

Cálculo de datos agrupados o en tabla de frecuencias

A continuación, se relacionan las fórmulas que se aplican cuando la variable es discreta

y cuando es continua. En cada caso, se tienen dos procedimientos (Martínez, 2012):

Variable discreta:

$$\text{Cuando } N_{j-1} < \frac{n}{2} \text{ entonces la mediana será } M_E = y_j$$

$$\text{Cuando } N_{j-1} = \frac{n}{2} \text{ entonces la mediana será } M_E = \frac{y_j + y_{j+1}}{2}$$

Variable continua:

$$\text{Cuando } N_{j-1} < \frac{n}{2} \text{ entonces la mediana será } M_E = y'_{j-1} + c \left[\frac{\frac{n}{2} - N_{j-1}}{n_j} \right]$$

$$\text{Cuando } N_{j-1} = \frac{n}{2} \text{ entonces la mediana será } M_E = y'_{j-1}$$

MODA O VALOR MODAL (M_0)

Se define como “el valor de la variable que más se repite” o “aquel valor que presenta la máxima frecuencia”. Puede suceder que una distribución tenga dos modas, en este caso se dice que la distribución es bimodal, y en el caso de que haya más de dos modas, se dice que es plurimodal o multimodal. Es la única medida de posición que puede ser utilizada en atributos, es decir, cuando la característica es cualitativa. La moda, o valor modal, es la medida originalmente concebida como aquel valor de la variable que presenta el mayor número de observaciones, es decir, el valor de la variable que más veces se repite (Martínez, 2012).

OBJETIVOS

Interactuar con la aplicación Microsoft Excel para el cálculo de medidas de tendencia central en ejercicios de contexto, contemplando datos desagrupados y agrupados en variables cuantitativas discretas.

MATERIALES

- Equipos de cómputo con sistema operativo Windows.
- Software Microsoft Excel instalado.

DESCRIPCIÓN DE ACTIVIDADES Y PROCEDIMIENTOS DE LA PRÁCTICA

EJERCICIO DATOS DESAGRUPADOS

Variable discreta

Una encuesta realizada a 30 adolescentes para determinar la cantidad de veces que se conectan a internet arrojó los siguientes resultados:

TABLA 1

Tabla datos cuantitativos discretos desagrupados

1	5	3	5	1	7	5	3	2	4
4	2	3	6	6	6	5	3	2	4
5	3	2	2	4	3	6	1	4	6

Nota. Elaboración propia.

Con los datos anteriores, determine la media, la mediana y la moda en datos desagrupados.

Solución

Inicialmente, se puede evidenciar que los datos corresponden a variable cuantitativa discreta. Ahora, pasamos los datos a una hoja de cálculo de Microsoft Excel:

	B	C	D	E	F	G	H	I	J	K
4	1	5	3	5	1	7	5	3	2	4
5	4	2	3	6	6	6	5	3	2	4
6	5	3	2	2	4	3	6	1	4	6

FIGURA 1. Datos en hoja de Microsoft Excel. Elaboración propia.

Para calcular la media de datos desagrupados, utilizamos la fórmula promedio de la siguiente manera:

FIGURA 2. Selección de la fórmula para calcular la media aritmética. Elaboración propia.

En este caso, se tienen dos opciones: insertar los datos de forma manual, uno por uno, como indica la ecuación o seleccionando el rango de datos, de esta manera:

FIGURA 3. Selección de datos para cálculo de la media aritmética. Elaboración propia.

El promedio es entonces:

FIGURA 4. Resultado de la media aritmética. Elaboración propia.

Es decir, en promedio los adolescentes se conectan a internet 3,8 veces en el día, lo cual es verificable utilizando la fórmula de la media aritmética simple, de la siguiente forma:

$$\bar{x} = \frac{1+4+5+5+2+3+3+3+2+5+6+2+1+6+4+7+6+3+5+5+6+3+3+1+2+2+4+4+4+6}{30}$$

$$\bar{x} = \frac{113}{30} = 3,7666$$

Para el cálculo de la mediana en datos desagrupados, se utiliza la fórmula de la mediana de la siguiente manera:

FIGURA 5. Selección de datos para cálculo de la mediana. Elaboración propia.

En este caso, se tienen dos opciones: insertar los datos de forma manual, uno por uno, o seleccionando el rango de datos, de esta manera:

FIGURA 6. Selección de grupos para cálculo de la mediana. Elaboración propia.

Luego la mediana es:

FIGURA 7. Resultado de la mediana. Elaboración propia.

Es decir, la mediana es de cuatro veces en el día, lo cual indica que es el dato que se encuentra en la mitad de los 30 datos. Esto es verificable organizando primero los datos de menor a mayor:

TABLA 2

Datos desagrupados organizados de menor a mayor

1	1	1	2	2	2	2	2	3	3
3	3	3	3	4	4	4	4	4	5
5	5	5	5	6	6	6	6	6	7

Nota. Elaboración propia.

Luego aplicamos la ecuación de $\frac{n+1}{2} = \frac{30+1}{2} = 15,5$. Esto nos indica que la mediana está entre la decimoquinta y la decimosexta posición:

TABLA 3

Datos resaltados en azul para cálculo de mediana

1	1	1	2	2	2	2	2	3	3
3	3	3	3	4	4	4	4	4	5
5	5	5	5	6	6	6	6	6	7

Nota. Elaboración propia.

Es decir, la mediana sería el promedio de los números resaltados en verde oscuro:

$$M_e = \frac{4+4}{2} = 4$$

Resultado igual al arrojado por Microsoft Excel.

Finalmente, para la moda en datos desagrupados se utiliza la fórmula denominada moda, de la siguiente manera:

FIGURA 8. Selección de fórmula para cálculo de moda. Elaboración propia.

En ese caso, se utiliza MODA.UNO para que nos arroje un solo dato que es el que más se repite; sin embargo, en otros ejercicios se puede presentar que sea bimodal o plurimodal, por lo que la fórmula sería diferente. Igualmente, se tienen dos opciones: insertar los datos de forma manual, uno por uno, como indica la ecuación o seleccionando el rango de datos, de esta manera:

FIGURA 9. Selección de datos para el cálculo de la moda. Elaboración propia.

Por lo que la moda sería:

FIGURA 10. Resultado de la moda. Elaboración propia.

La moda es 3, el cual es el valor que más se repite, es decir, la cantidad de veces más frecuente de conexión a internet es tres veces en el día.

Ahora bien, esto es verificable contando la cantidad de veces de cada dato del rango de datos del ejercicio. Al hacer esto, se concluye que el dato que más se repite es el 3, por lo que $M_o = 3$.

EJERCICIO DATOS AGRUPADOS

Para solucionar el ejercicio anterior como datos agrupados, primero se debe realizar una tabla en Excel como se relaciona a continuación:

FIGURA 11. Elaboración de tabla de frecuencias para datos agrupados. Elaboración propia.

Donde y_i es la variable cuantitativa discreta, n_i es la frecuencia absoluta, h_i es la frecuencia relativa, N_i es la frecuencia absoluta acumulada y H_i es la frecuencia relativa acumulada. Para determinar la frecuencia absoluta, se utiliza la fórmula de frecuencia de la siguiente manera:

ejercicio literatura gris - Excel

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda Buscar

Portapapeles Fuente Alineación Número Estilos

SUMA X ✓ fx =FRECUENCIA(

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3												
4		1	5	3	5	1	7	5	3	2	4	
5		4	2	3	6	6	6	5	3	2	4	
6		5	3	2	2	4	3	6	1	4	6	
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												

Hoja1

Introducir

FIGURA 12. Selección de fórmula para cálculo de la frecuencia absoluta. Elaboración propia.

Para ejecutarla, se deben seleccionar las celdas donde se ubicarán los resultados, es decir, en la columna D de la fila 9 a la 15. Luego seleccionamos todos los datos (B4:K6):

ejercicio literatura gris - Excel

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda Buscar

Portapapeles Fuente Alineación Número Estilos

B4 X ✓ fx =FRECUENCIA(B4:K6;

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3												
4		1	5	3	5	1	7	5	3	2	4	
5		4	2	3	6	6	6	5	3	2	4	
6		5	3	2	2	4	3	6	1	4	6	
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												

Hoja1

Introducir

FIGURA 13. Selección de datos para cálculo de la frecuencia absoluta. Elaboración propia.

Se escribe punto y coma, luego se seleccionan los grupos (C9:C15):

FIGURA 14. Selección de grupo para cálculo de la frecuencia absoluta. Elaboración propia.

Para que el resultado sea correcto, se deben presionar las teclas Ctrl + Shift + Enter al mismo tiempo:

FIGURA 15. Resultados de la frecuencia absoluta acumulada. Elaboración propia.

De la tabla de frecuencias, se puede identificar que tres adolescentes se conectan una vez en el día; cinco adolescentes, dos veces en el día, y así sucesivamente.

Para calcular la frecuencia relativa, se utiliza la ecuación $h_i = \frac{n_i}{n}$, donde n_i es la frecuencia absoluta y n es la muestra, que para este caso son 30 adolescentes, por tanto:

	B	C	D	E	F	G	H	I	J	K	L
1											
2											
3											
4	1	5	3	5	1	7	5	3	2	4	
5	4	2	3	6	6	6	5	3	2	4	
6	5	3	2	2	4	3	6	1	4	6	
7											
8											
9		Yi	ni	hi	Ni	Hi					
10		1		=D9/30							
11		2									
12		3									
13		4									
14		5									
15		6									
16		7									
17											
18											
19											
20											
21											
22											

FIGURA 16. Cálculo de la frecuencia relativa. Elaboración propia.

Donde D9 hace referencia a 3, que es el valor cuando Y_i es 1 y se divide en 30, que es la muestra.

	B	C	D	E	F	G	H	I	J	K	L
1											
2											
3											
4	1	5	3	5	1	7	5	3	2	4	
5	4	2	3	6	6	6	5	3	2	4	
6	5	3	2	2	4	3	6	1	4	6	
7											
8											
9		Yi	ni	hi	Ni	Hi					
10		1		0,10							
11		2		0,17							
12		3		0,20							
13		4		0,17							
14		5		0,17							
15		6		0,17							
16		7		0,03							
17		Total	30	1							
18											
19											
20											
21											
22											

FIGURA 17. Resultados de la frecuencia relativa. Elaboración propia.

Se ajusta h_i a 2 decimales. La sumatoria de toda la columna h_i , debe ser igual a 1, que equivale al 100% de los datos, es decir, los 30 adolescentes. Para el cálculo de la frecuencia absoluta acumulada, se suman los valores de n_i :

$$N_1 = n_1$$

$$N_2 = n_1 + n_2$$

$$N_3 = n_1 + n_2 + n_3 \text{ y así sucesivamente.}$$

En el caso de la frecuencia relativa acumulada, es la suma de los valores de h_i , es decir:

$$H_1 = h_1$$

$$H_2 = h_1 + h_2$$

$$H_3 = h_1 + h_2 + h_3 \text{ y así sucesivamente.}$$

Por lo que, en Excel, se reflejaría de la siguiente manera:

Y _i	n _i	h _i	N _i	H _i
1	3	0,10	3	0,10
2	5	0,17	8	0,27
3	6	0,20	14	0,47
4	5	0,17	19	0,63
5	5	0,17	24	0,80
6	5	0,17	29	0,97
7	1	0,03	30	1,00
Total	30	1		

FIGURA 18. Cálculo de la frecuencia absoluta y relativa acumulada. Elaboración propia.

Para el cálculo de la media aritmética, se utiliza la ecuación de la media aritmética ponderada, $\bar{y} = \frac{\sum y_i n_i}{n}$, por lo que en Excel se calcula primero la columna $Y_i n_i$:

FIGURA 19. Cálculo de Y_{pi} para la media aritmética. Elaboración propia.

Luego, se suma y se divide en la cantidad de datos, utilizando la fórmula de suma, es decir:

FIGURA 20. Resultados de la columna Y_{pi} . Elaboración propia.

Dando el siguiente resultado:

The screenshot shows an Excel spreadsheet with the following data and calculations:

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3												
4		1	5	3	5	1	7	5	3	2	4	
5		4	2	3	6	6	6	5	3	2	4	
6		5	3	2	2	4	3	6	1	4	6	
7												
			Yi	ni	hi	Ni	Hi	yini				
9			1		3		3	0,10			3	
10			2		5		8	0,17			10	
11			3		6		14	0,20			18	
12			4		5		19	0,17			20	
13			5		5		24	0,17			25	
14			6		5		29	0,17			30	
15			7		1		30	0,03			7	
16			Total		30		1				3,76666667	
17											Media	

FIGURA 21. Resultado de la media aritmética. Elaboración propia.

Este resultado es igual al resultado de datos desagrupados, ya que es una variable discreta. También se puede corroborar utilizando la fórmula de la siguiente manera:

$$\bar{y} = \frac{(1*3)+(2*5)+(3*6)+(4*5)+(5*5)+(6*5)+(7*1)}{30} = 3,76666$$

Para el caso de la mediana, como son datos agrupados, se aplica directamente la ecuación, teniendo presente la columna de la frecuencia absoluta acumulada:

The screenshot shows an Excel spreadsheet with two tables. The first table is a 6x10 grid of data. The second table is a summary table with 7 rows and 7 columns.

	Yi	ni	hi	Ni	Hi	yini
1	1	3	0,10	3	0,10	3
2	2	5	0,17	8	0,27	10
3	3	6	0,20	14	0,47	18
4	4	5	0,17	19	0,63	20
5	5	5	0,17	24	0,80	25
6	6	5	0,17	29	0,97	30
7	7	1	0,03	30	1,00	7
Total		30	1			3,7666667
						Media

FIGURA 22. Columna N_i resaltada para cálculo de mediana. Elaboración propia.

En este caso, primero calculamos $n/2$, lo cual sería $(30/2)=15$, este valor se busca en la columna N_i , y como no se encuentra, entonces el inmediatamente anterior se denota como N_{j-1} y el inmediatamente superior, como N_j .

The screenshot shows the same Excel spreadsheet as Figure 22, but with additional annotations. The formula bar shows $=F10+D11$. Arrows point from the text Y_i to the cell containing 3 in the first row of the summary table, and from the text N_{j-1} to the cell containing 14 in the third row of the summary table. Another arrow points from the text N_j to the cell containing 19 in the fourth row of the summary table.

	Yi	ni	hi	Ni	Hi	yini
1	1	3	0,10	3	0,10	3
2	2	5	0,17	8	0,27	10
3	3	6	0,20	14	0,47	18
4	4	5	0,17	19	0,63	20
5	5	5	0,17	24	0,80	25
6	6	5	0,17	29	0,97	30
7	7	1	0,03	30	1,00	7
Total		30	1			3,7666667
						Media

FIGURA 23. Procedimiento para cálculo de la mediana. Elaboración propia.

Luego, como $N_{j-1} < \frac{n}{2}$, entonces $M_e = y_j$, es decir, $N_{j-1} < 15$, $M_e = 4$, lo cual concuerda con la solución de datos desagrupados.

Para hallar la moda, que es el dato que más se repite, con la tabla de frecuencia se puede identificar rápidamente que el dato es 3, por ende, $M_o = 3$.

TRABAJO PARA LOS ESTUDIANTES

- Los siguientes son los datos que corresponden al número de retrasos (en la mañana) por parte de los empleados de una empresa:

2	5	0	0	6	3	5	0	0	2	7	4	4	2	3	8	4	3	3	1
3	2	3	2	4	2	2	3	1	3	2	0	0	1	4	2	2	0	3	5

- Determine la media, la mediana y la moda con datos desagrupados utilizando Microsoft Excel. Interprete los resultados.
 - Elabore la tabla de frecuencia utilizando Microsoft Excel y determine la media, la mediana y la moda. (Martínez, 2012)
- Una encuesta realizada a 30 fumadores para determinar el número de cigarrillos que encienden (fuman) en un día corriente arrojó los siguientes resultados:

3	7	5	10	8	4	5	8	10	8	8	4	5	3	10	5
7	10	8	5	5	12	8	4	4	3	5	8	12	10		

- Determine la media, la mediana y la moda con datos desagrupados utili-

zando Microsoft Excel. Interprete los resultados.

- Elabore la tabla de frecuencia utilizando Microsoft Excel y determine la media, la mediana y la moda. (Martínez, 2012)
- En un estudio realizado sobre la aplicación de tatuajes en el cuerpo, a estudiantes universitarios se les consultaron aspectos como los que se ilustran en la tabla:

Género	Edad (años)	Actividad laboral	N.º de tatuajes	Parte del cuerpo tatuada
Masculino	19	No	2	Espalda
Femenino	20	Sí	1	Pierna
Masculino	26	Sí	1	Brazo
Masculino	27	Sí	4	Hombros
Masculino	26	Sí	1	Brazo
Femenino	20	Sí	2	Espalda
Masculino	17	No	1	Pierna
Masculino	27	Sí	2	Brazo
Femenino	17	No	1	Pie
Femenino	18	No	1	Pierna
Masculino	17	Sí	5	Brazo
Femenino	19	No	1	Hombros
Femenino	20	No	2	Hombros
Femenino	17	No	4	Espalda
Masculino	22	Sí	1	Espalda
Femenino	17	No	2	Pierna
Femenino	18	No	1	Cuello
Masculino	20	No	4	Brazo
Masculino	18	No	1	Brazo
Masculino	24	No	1	Pierna

Clasifique cada una de las variables (género, edad, actividad laboral, número de tatuajes, parte del cuerpo tatuada) según su tipo (cualitativas, cuantitativas discretas o continuas). Para la variable “número de tatuajes”:

- a. Determine la media, la mediana y la moda con datos desagrupados utilizando Microsoft Excel. Interprete los resultados.
 - b. Elabore la tabla de frecuencia utilizando Microsoft Excel y determine la media, la mediana y la moda. (Hernández, 2016)
4. En una empresa manufacturera se realizó un control de calidad sobre la producción.

Fueron seleccionados 20 productos y se les identificó el número de imperfectos que presentaban. Los resultados fueron los siguientes:

1 0 2 3 2 0 2 1 1 4

1 1 2 4 0 3 3 1 4 1

- a. Determine la media, la mediana y la moda con datos desagrupados utilizando Microsoft Excel. Interprete los resultados.
- b. Elabore la tabla de frecuencia utilizando Microsoft Excel y determine la media, la mediana y la moda. (Hernández, 2016)

REFERENCIAS

- Hernández, G. J. (2016). *Elementos básicos de estadística descriptiva para el análisis de datos*. Fondo Editorial Luis Amigó.
- Martínez, C. (2012). *Estadística y muestreo*. Ecoe Ediciones.
- Telesup (2017). Excel, una poderosa herramienta para el manejo de información. Instituto Tecnológico Telesup. Consultado el 13 de febrero de 2020. <https://telesup.edu.pe/excel-una-poderosa-herramienta-para-el-manejo-de-informacion/>