



PROCESAMIENTO DE DATOS EN LAS INVESTIGACIONES EDUCTIVAS UNIDAD II

FASE DE PROCESAMIENTO DE LOS CONJUNTOS DE DATOS ESTRUSTURADOS

Octubre 2025

Dr.C. Omar Pérez Jacinto

Cel: +53 53750444

OBJETIVO GENERAL:

Establecer la finalidad y alcance de la «fase de procesamiento» en el proceso de transformación «datos / información» en la lógica interna de las investigaciones educativas.

«INFORMACIÓN»

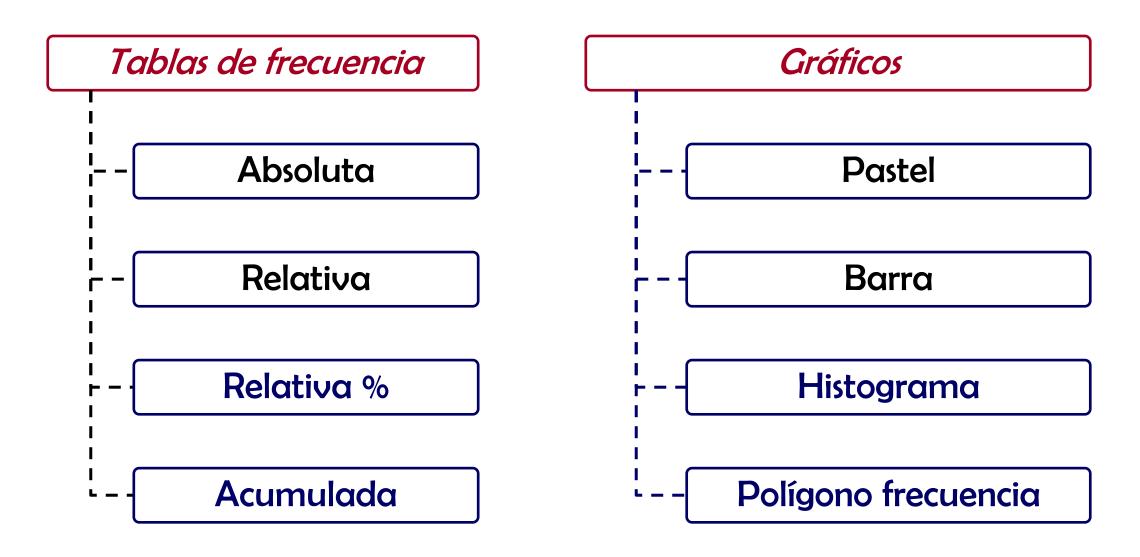
Conjunto organizado de datos procesados que explican los hechos desde un sistema teórico

Proceso de transformación



Representación simbólica de hechos tangibles surgidos de una fuente del medio circundante

ORGANIZACIÓN Y PRESENTACIÓN DE LOS «CONJUNTOS DE DATOS ESTRUCTURADOS»



TABLAS Y GRÁFICOS EN FUNCIÓN DE LAS ESCALAS

Escala Frecuencia Gráfico

Nominal f_{ab} - % Pastel

Ordinal f_{ab} - f_{ac} Barras simples

Razón (Discreta) $f_{ab} - f_{ac}$ Barras simples

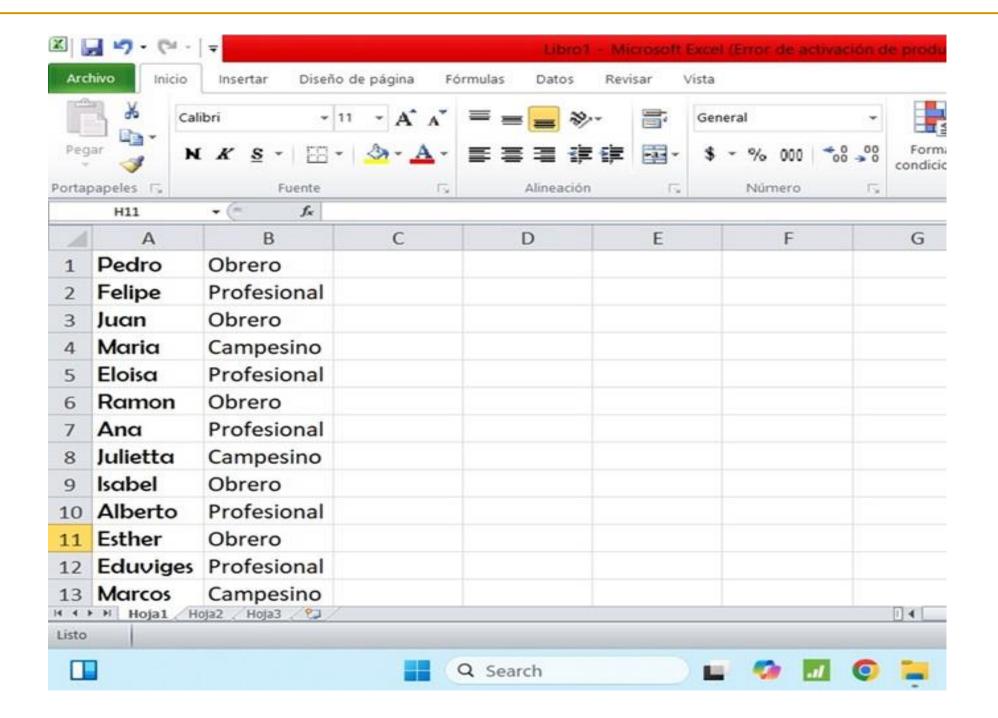
Razón (Continua) f_{ab} - mc Histograma Polígono de frecuencia

«CARACTERIZACIÓN DEL COMPORTAMIENTO DE UNA VARIABLE CLASIFICATORIA EN UNA MUESTRA»

En una Universidad, se indaga acerca de la procedencia social de los estudiantes que matriculan la Carrera de Ingeniería...

... para lo cual, se entrevistan a los 120 estudiantes que matricularon en 1º Año para conocer su procedencia social (hijos de profesionales, obreros y campesinos)...

... para procesar el conjunto de datos, se organizan y presentan en una tabla de frecuencia y un grafico ilustrativo de acuerdo con su procedencia social.

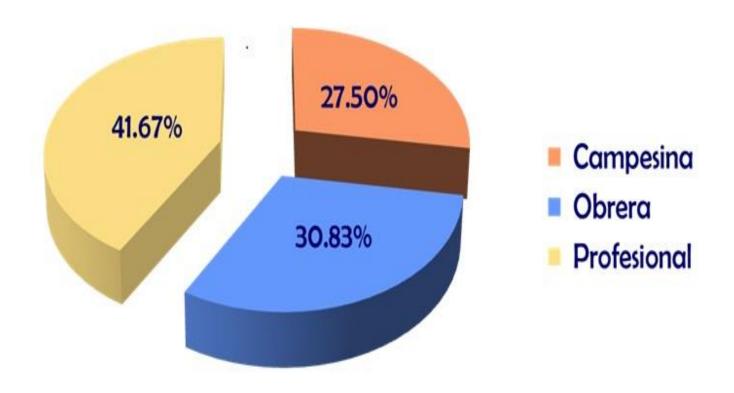


«Organización y presentación en la tabla de frecuencia» Ejemplo de tabla para variable nominal

Procedencia Social de Estudiantes 1º Año

r roccachela social ac Estadiantes in Anio							
Procedencia Social	f _{ab}	f _{rel}	%	f _{ac}			
Campesina	33	0.28	27.50	No			
Obrera	37	0.31	30.83	No			
Profesional	50	0.42	41.67	No			
Total	120	1,00	100 %	No			

Gráfico circular



TÉCNICAS Y PROGRAMAS PARA EL PROCESAMIENTO DE DATOS PROGRAMA DE PROCESAMIENTO MINITAB

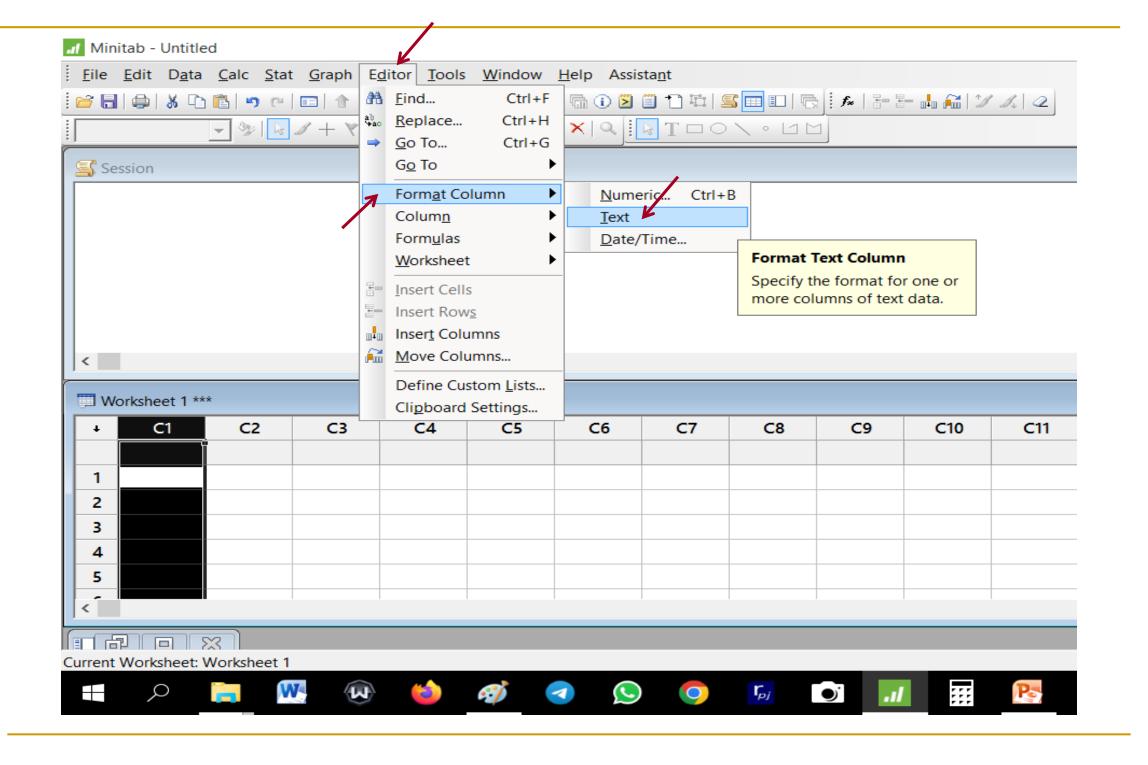


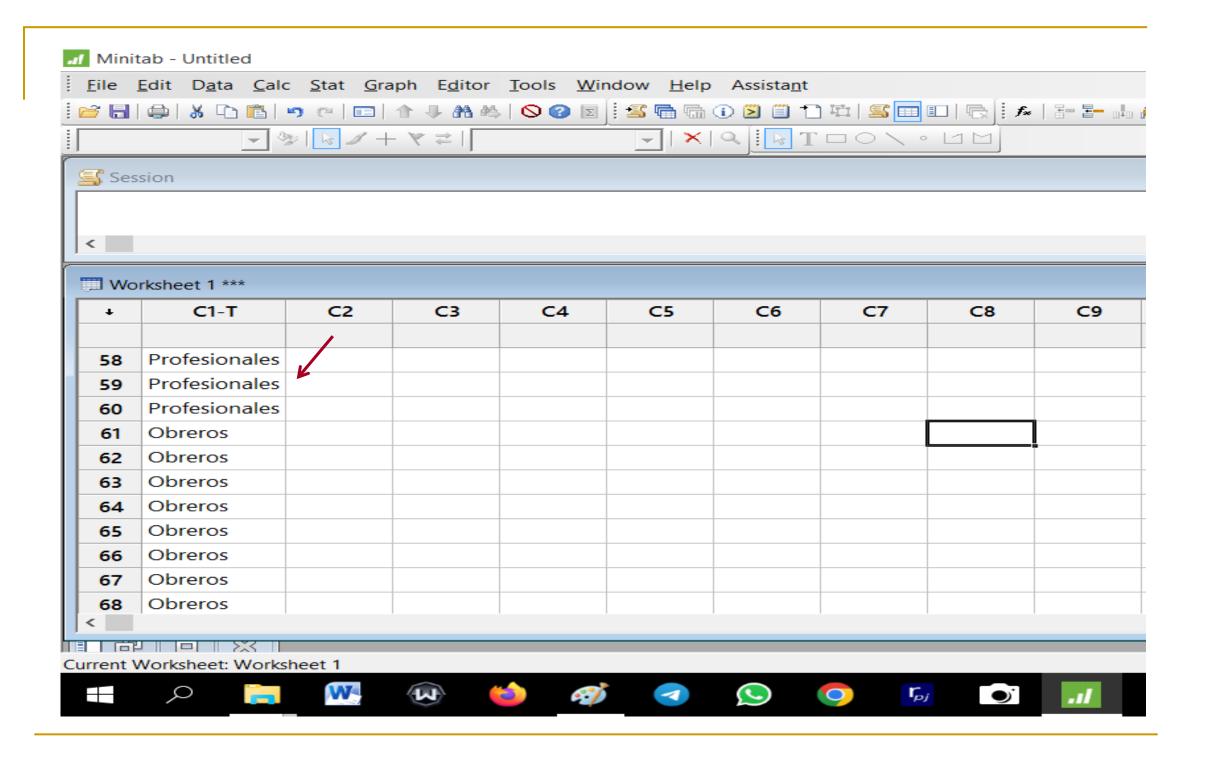
Minitab 17

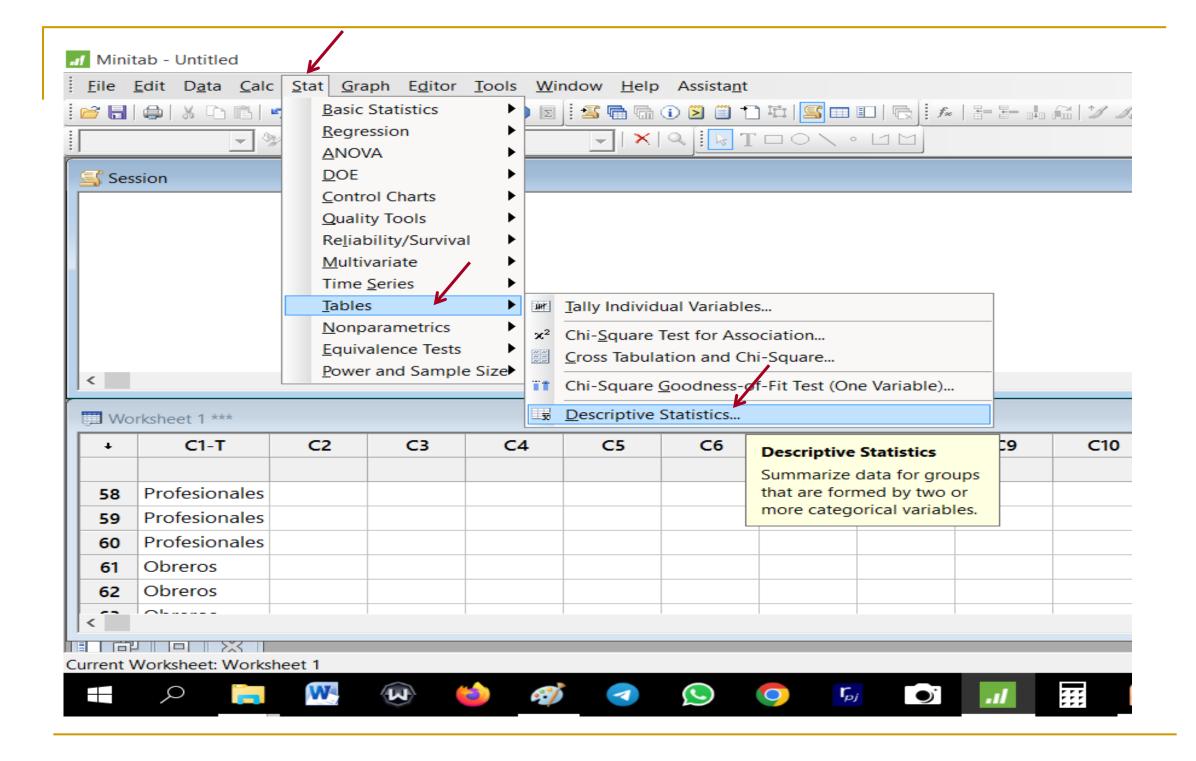
El software de herramientas estadísticas más utilizado en todo el mundo para mejorar la calidad.

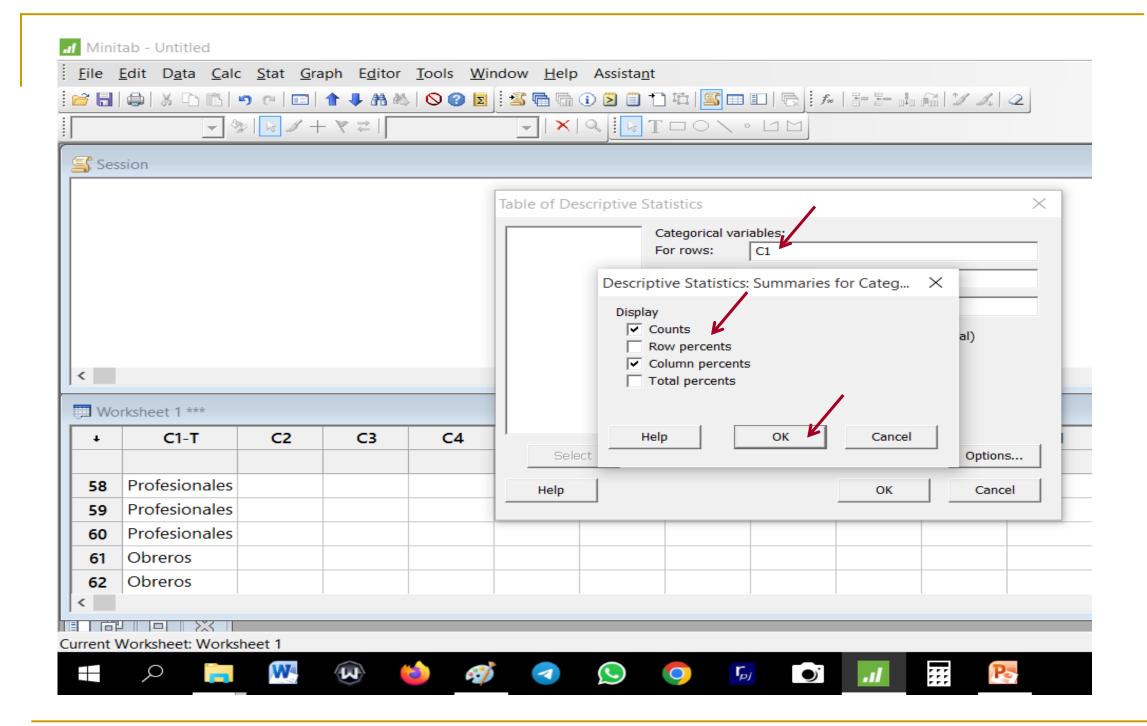
Minitab Inc. es el principal software del mundo para la enseñanza de la Estadística; también, es el software utilizado con mayor frecuencia en Seis Sigma, la principal metodología del mundo para el mejoramiento de la calidad.

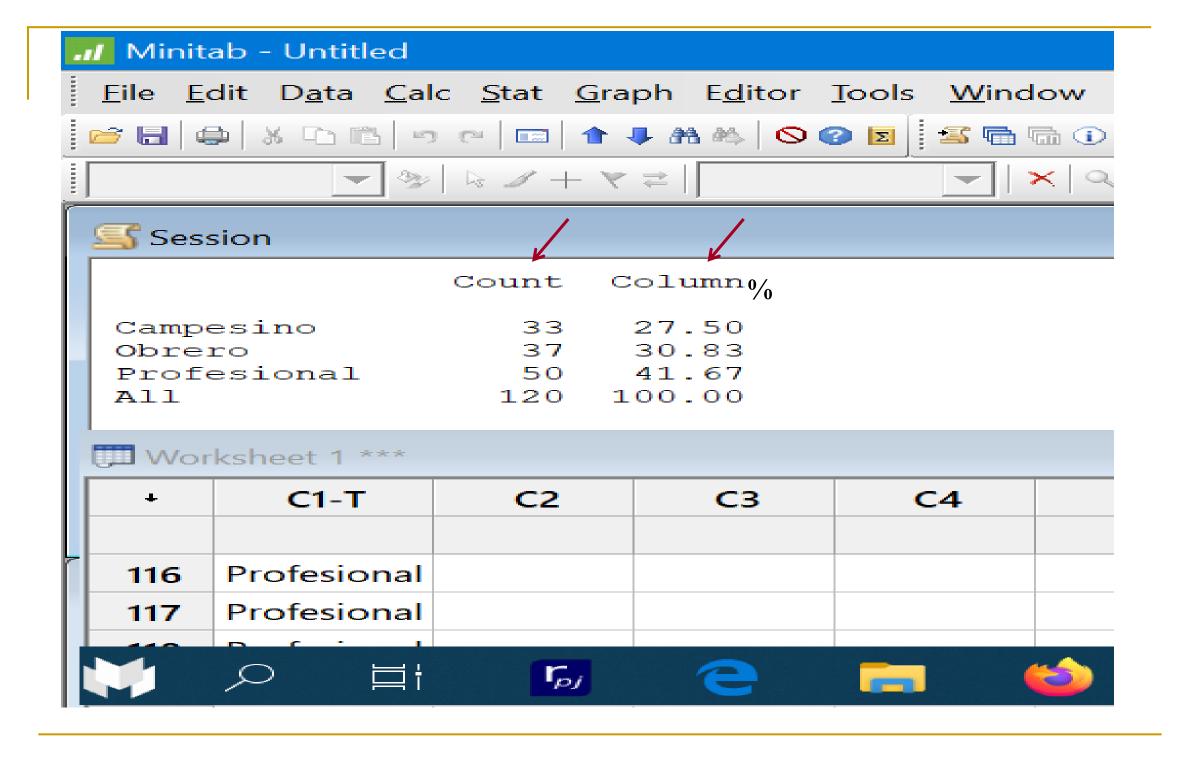
https://drive.google.com/file/d/1T6Oc5FD1BE7r-e_aubmhsNj2B0WY6fJn/view?usp=sharing

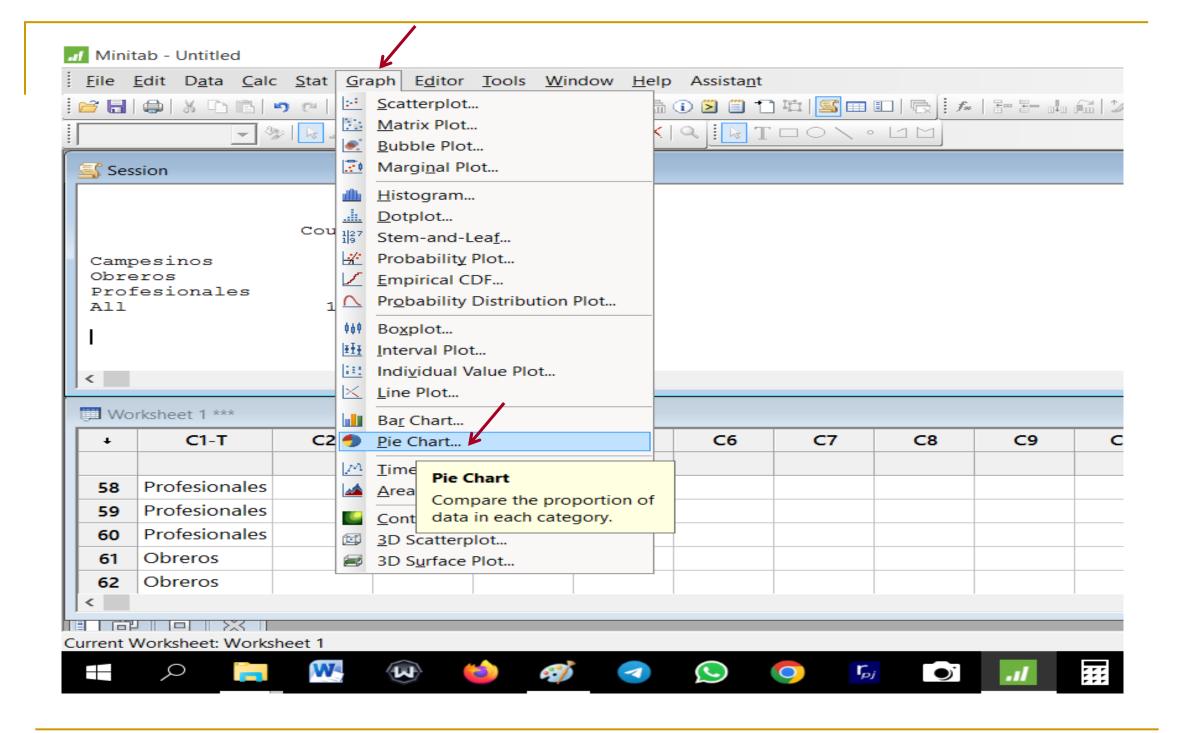


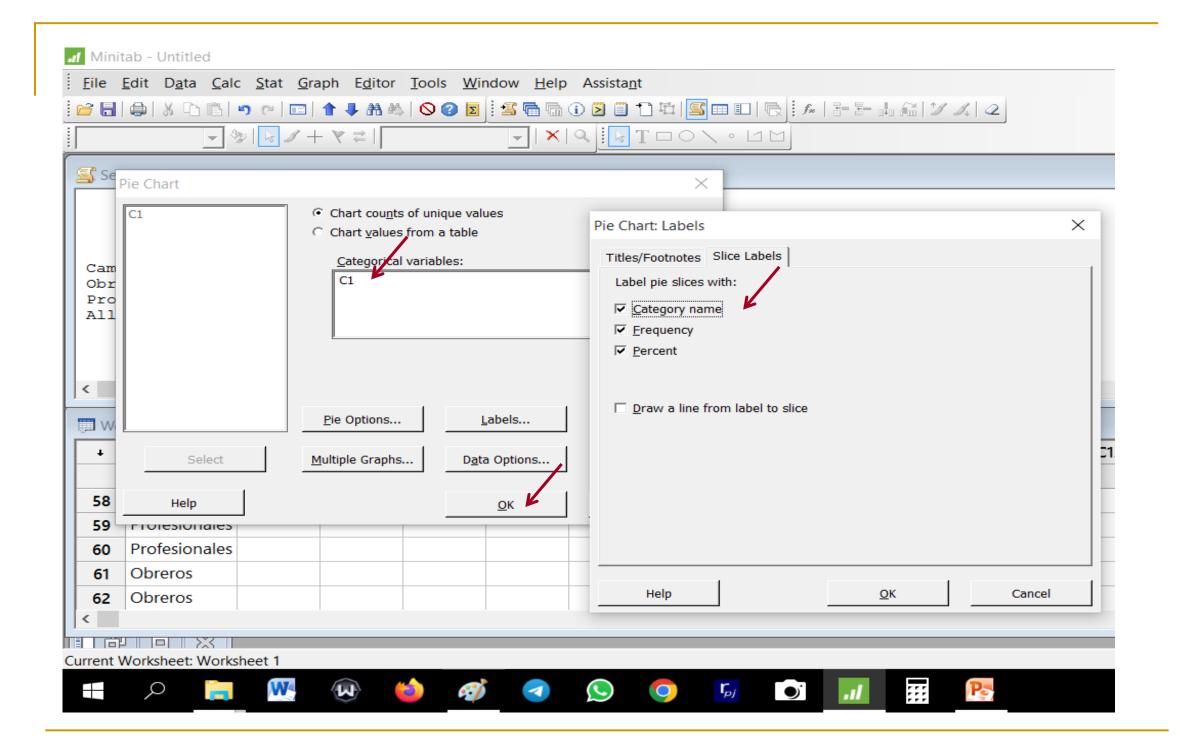


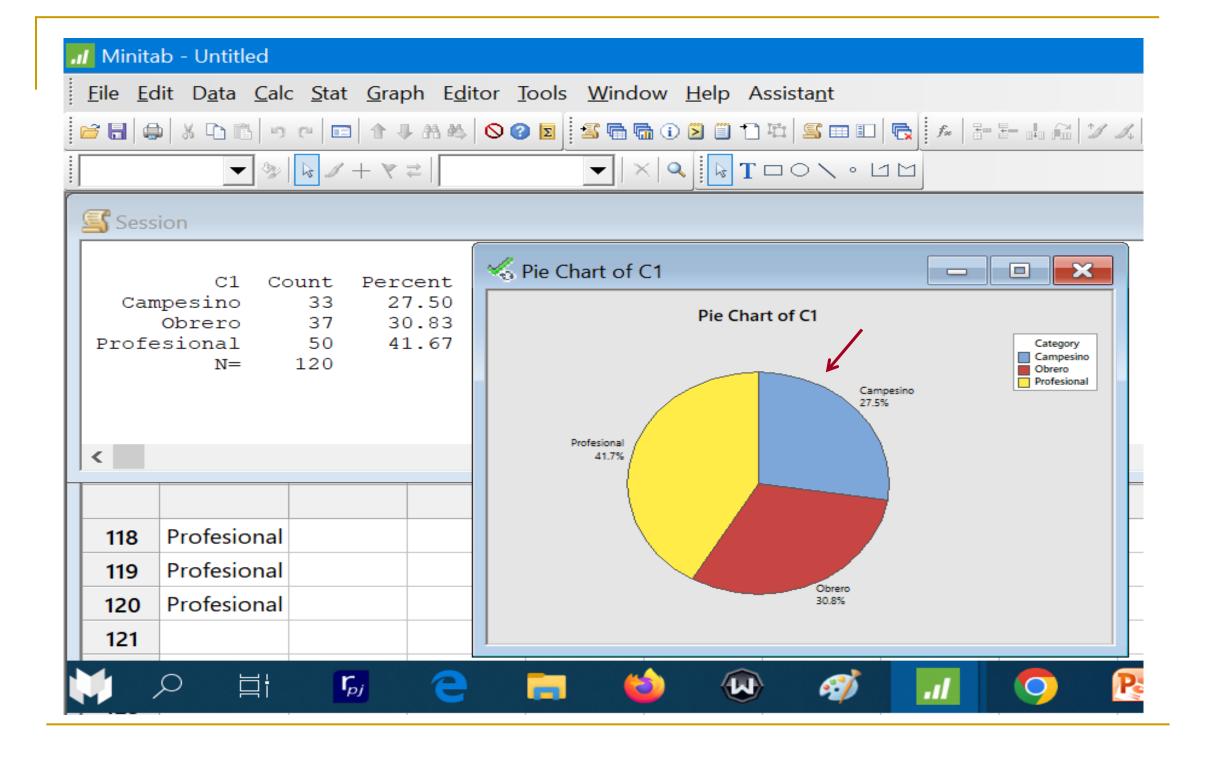






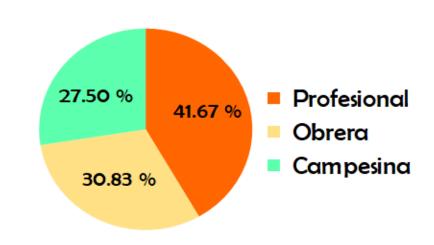






ANÁLISIS

Procedencia Social	f_{ab}	%
Campesina	33	27.50
Obrera	37	30.83
Profesional	50	41.67
Total	120	100 5



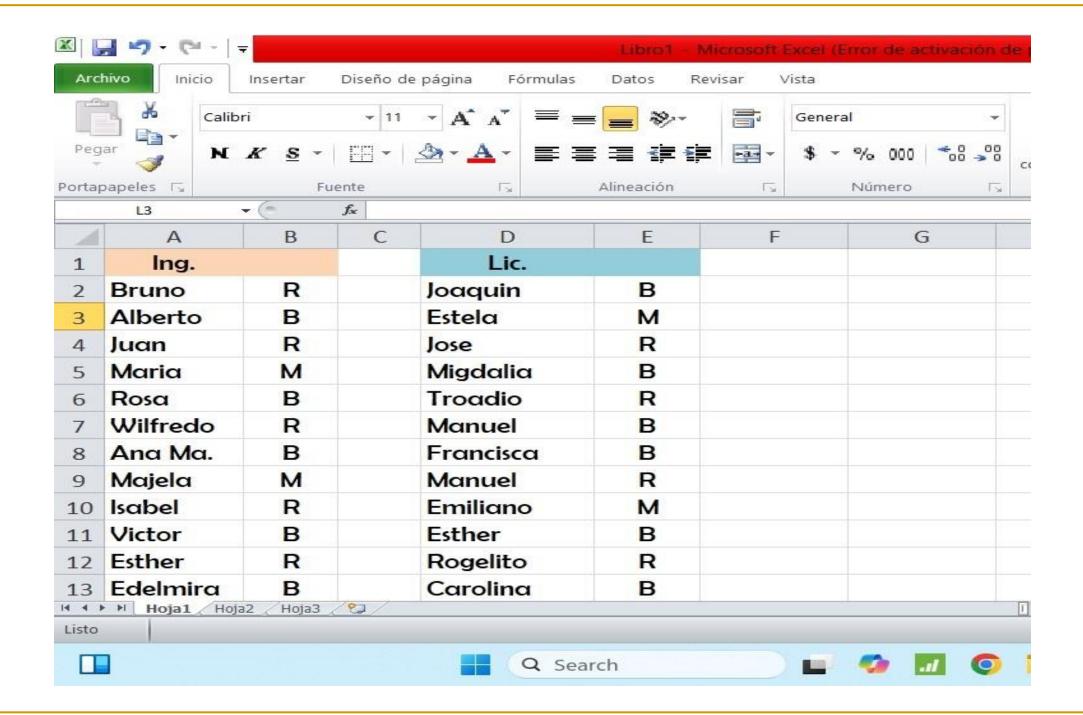
Análisis valorativo

Es probable que, en las matriculas de la carrera de ingeniería, generalmente predominen los estudiantes hijos de *«profesionales»,* seguidos de los hijos de *«obreros»* y sean minoría los hijos de *«campesino»*

«Organización y presentación en la tabla de frecuencia» Ejemplo de tabla comparativas para variable ordinal

En una investigación en una Universidad, se indaga acerca del rendimiento académico de los estudiantes de las Carreras de Ingeniería y Licenciatura...

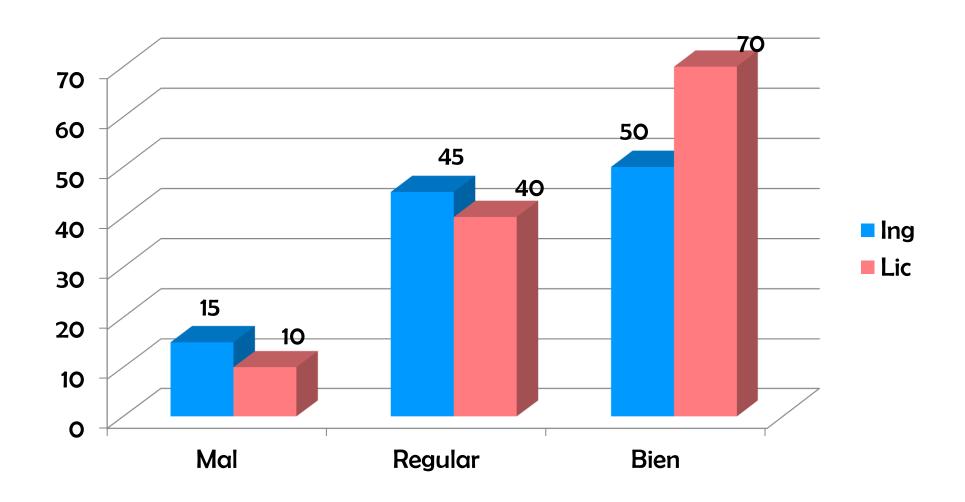
... para lo cual, se revisan los registros de notas de 110 estudiantes de Ingeniería y 120 estudiantes de Licenciatura del 1º Año de las respectivas carreras y se construye una tabla de frecuencia de acuerdo con la el registro de notas (Bien, Regular y Mal) y se obtiene también la frecuencia relativa porcentual y la acumulada.



Ejemplo de tablas comparativa para variable ordinal

Carrera	INGENIERÍA		LICENCIATURA	
	f_{ab}	f_{ac}	f _{ab}	f _{ac}
Mal	15	15	10	10
Regular	45	60	40	50
Bien	50	110	70	120
Total	110	110	120	120

Gráfico de barras simples



PROGRAMA DE PROCESAMIENTO MINITAB

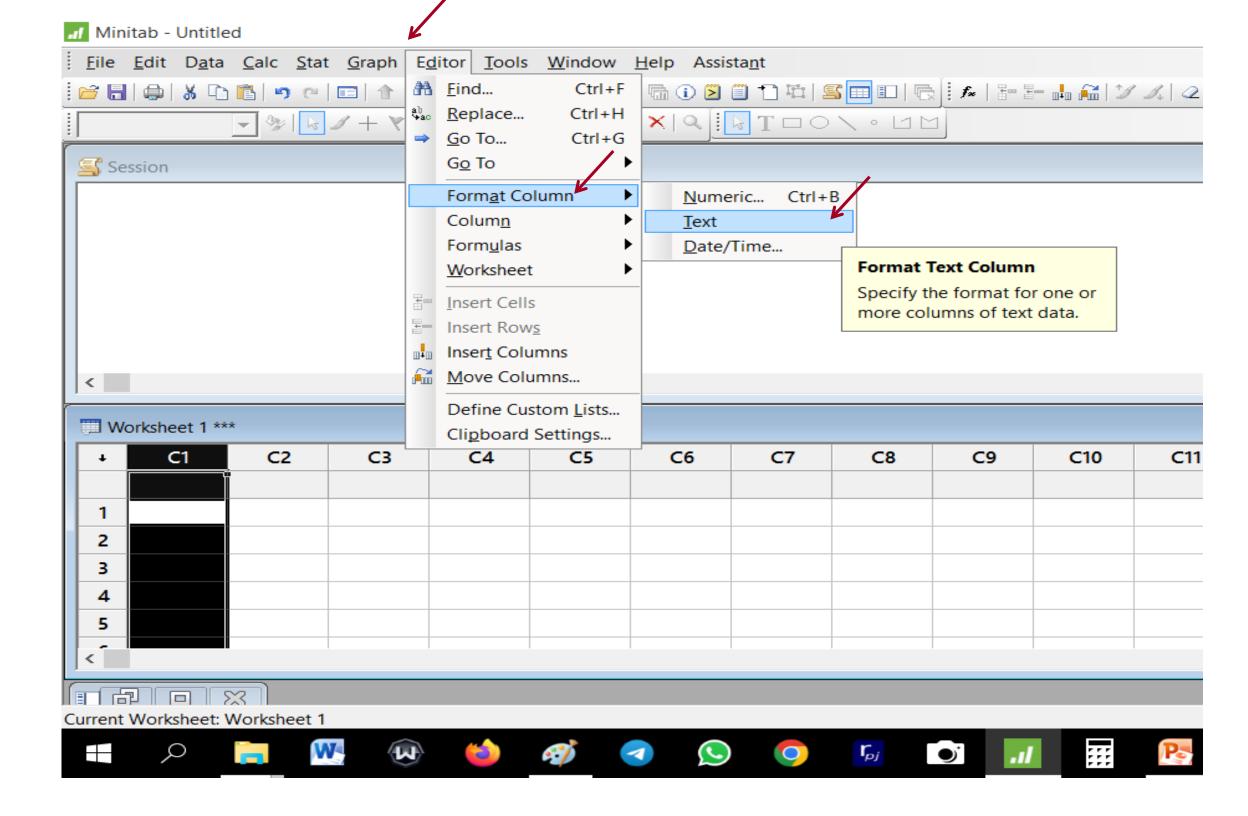


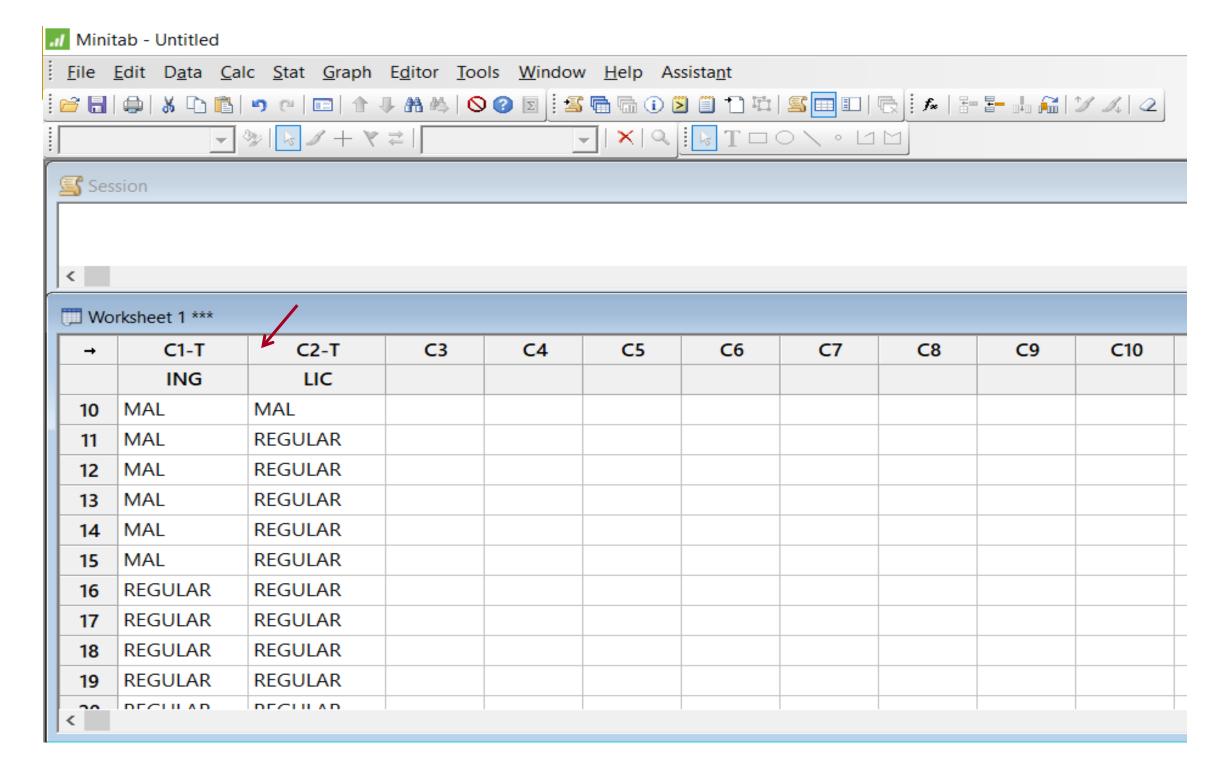
Minitab 17

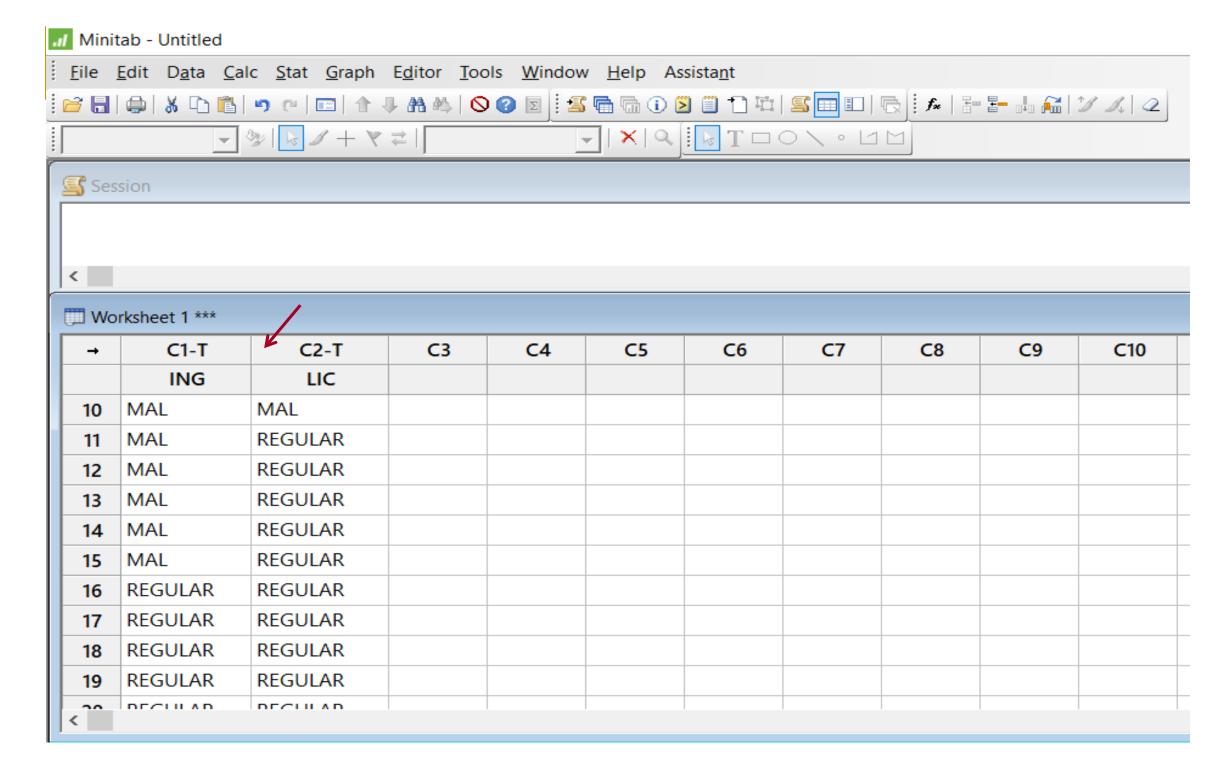
El software de herramientas estadísticas más utilizado en todo el mundo para mejorar la calidad.

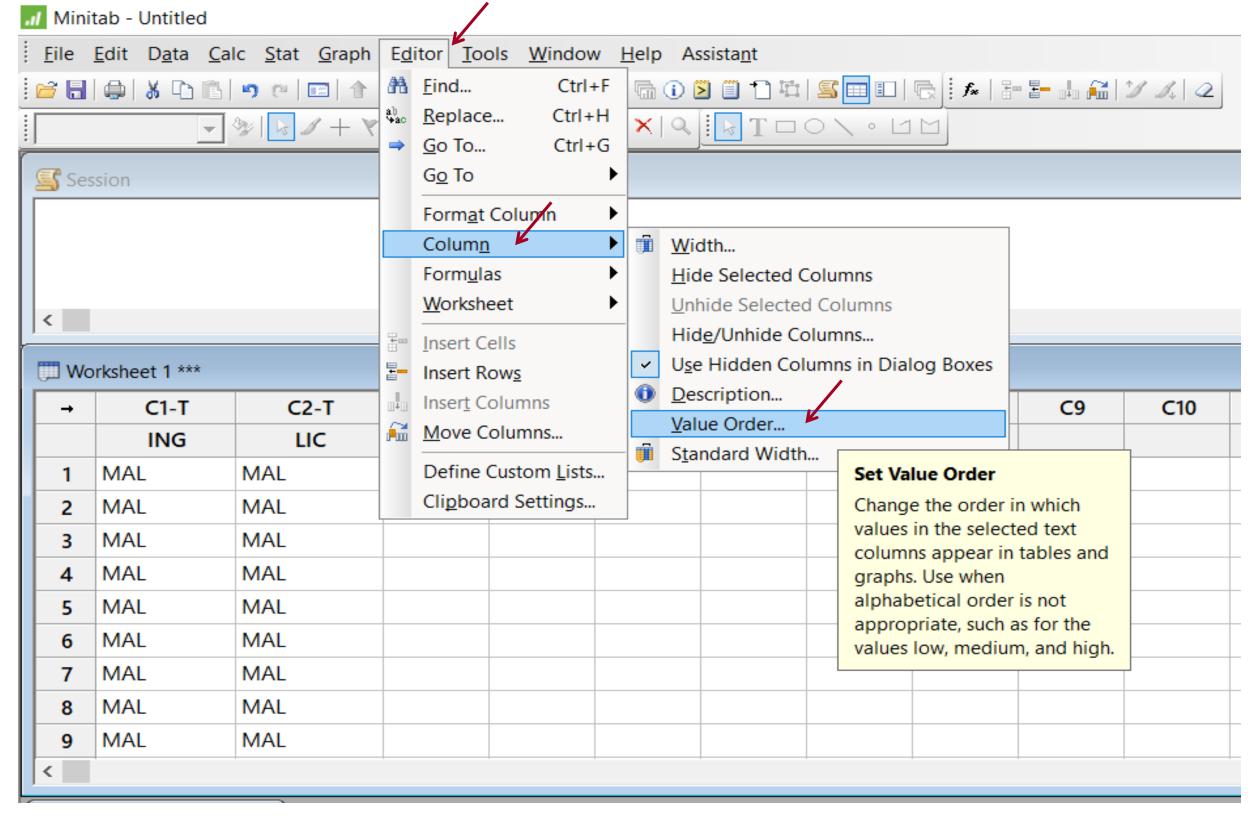
Minitab Inc. es el principal software del mundo para la enseñanza de la Estadística; también, es el software utilizado con mayor frecuencia en Seis Sigma, la principal metodología del mundo para el mejoramiento de la calidad.

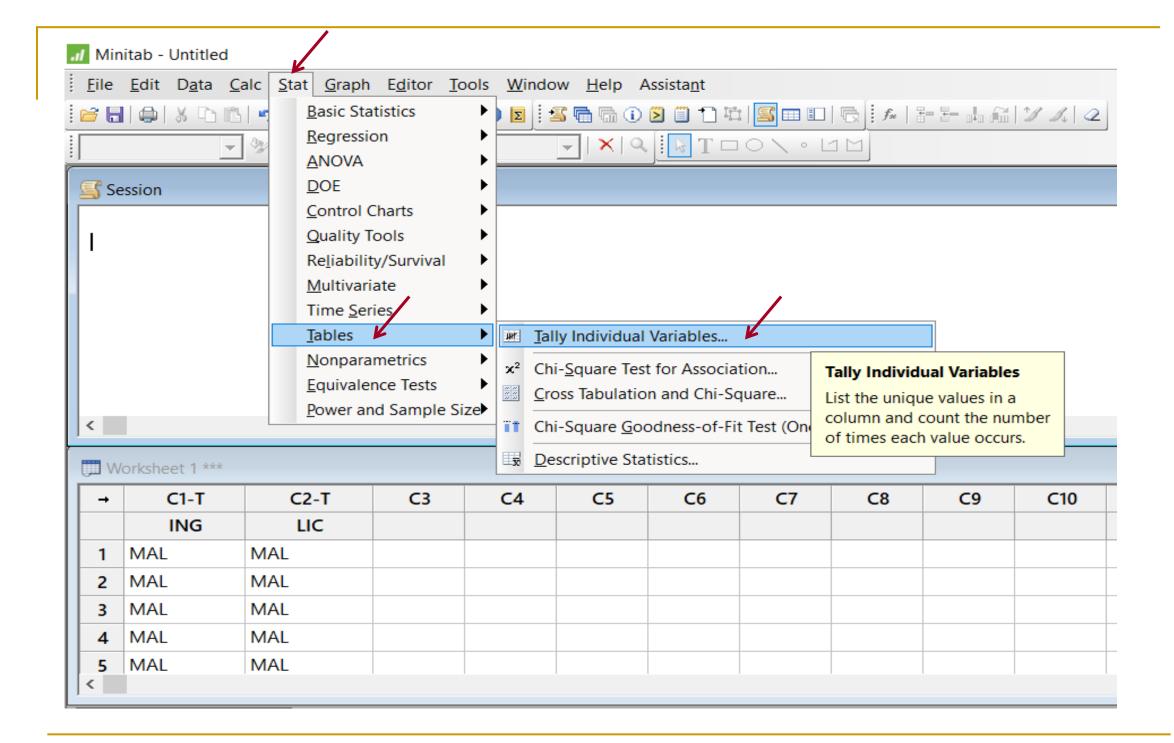
https://drive.google.com/file/d/1T6Oc5FD1BE7r-e_aubmhsNj2B0WY6fJn/view?usp=sharing

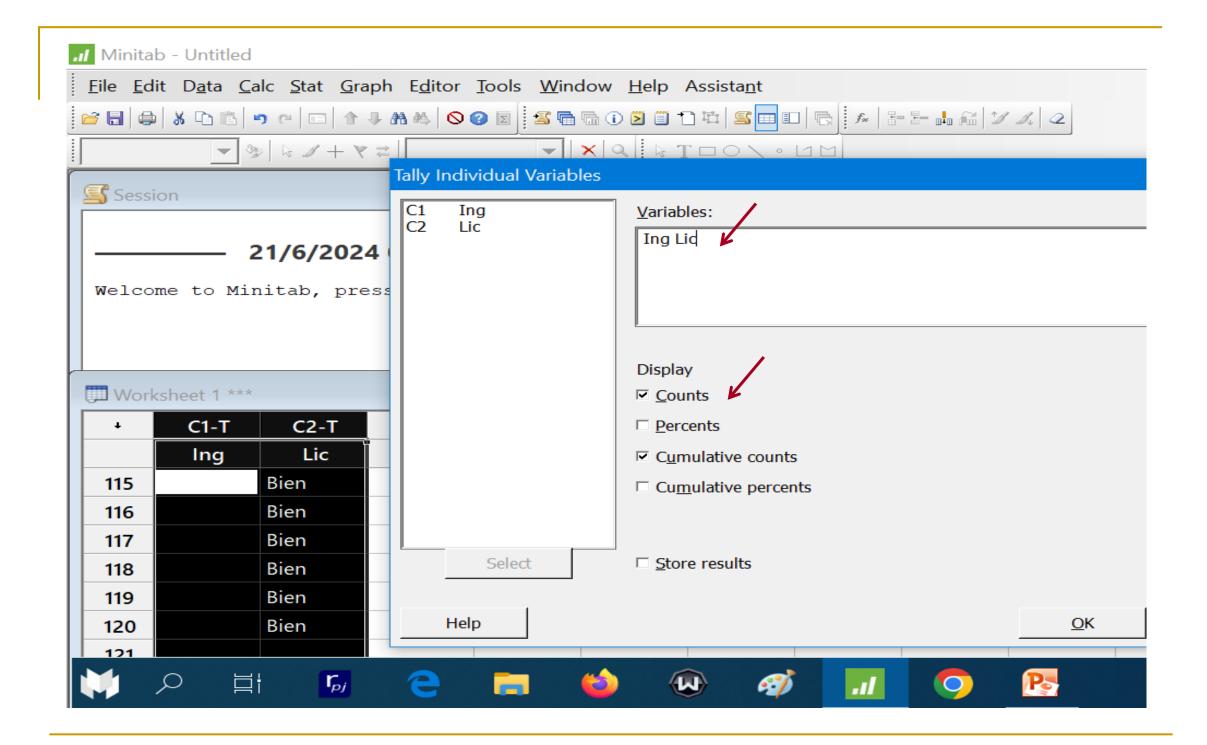


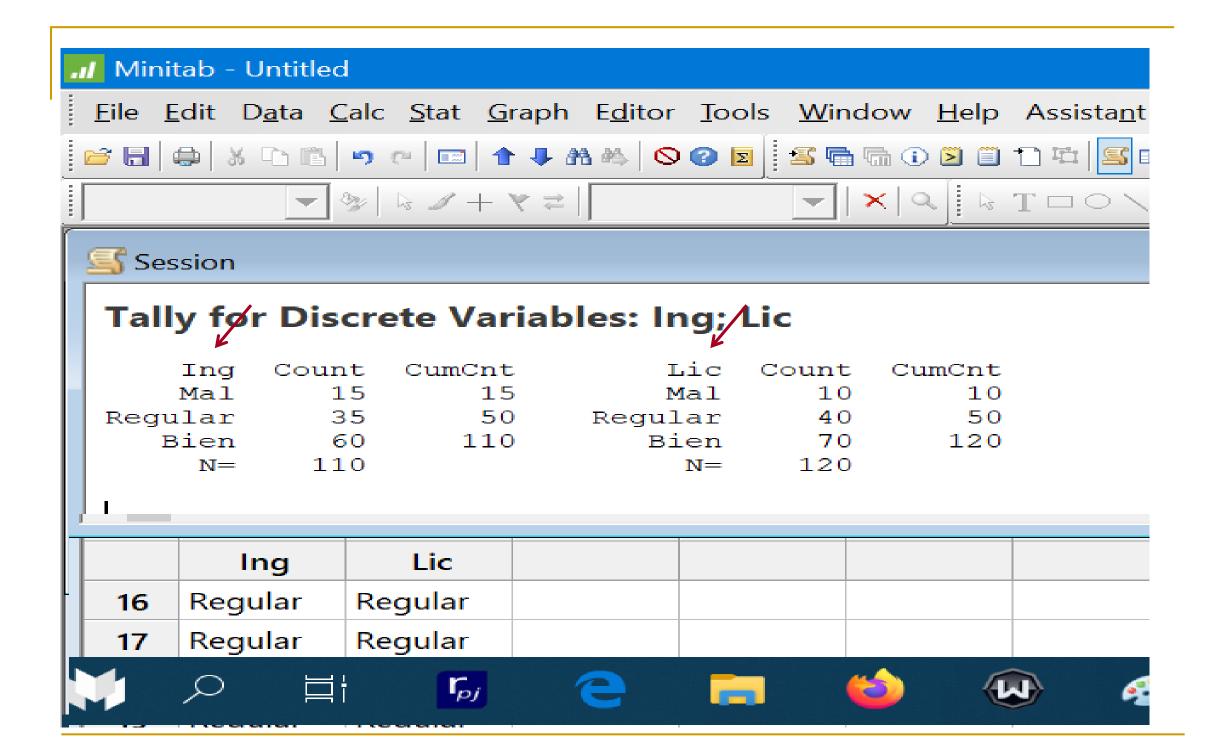


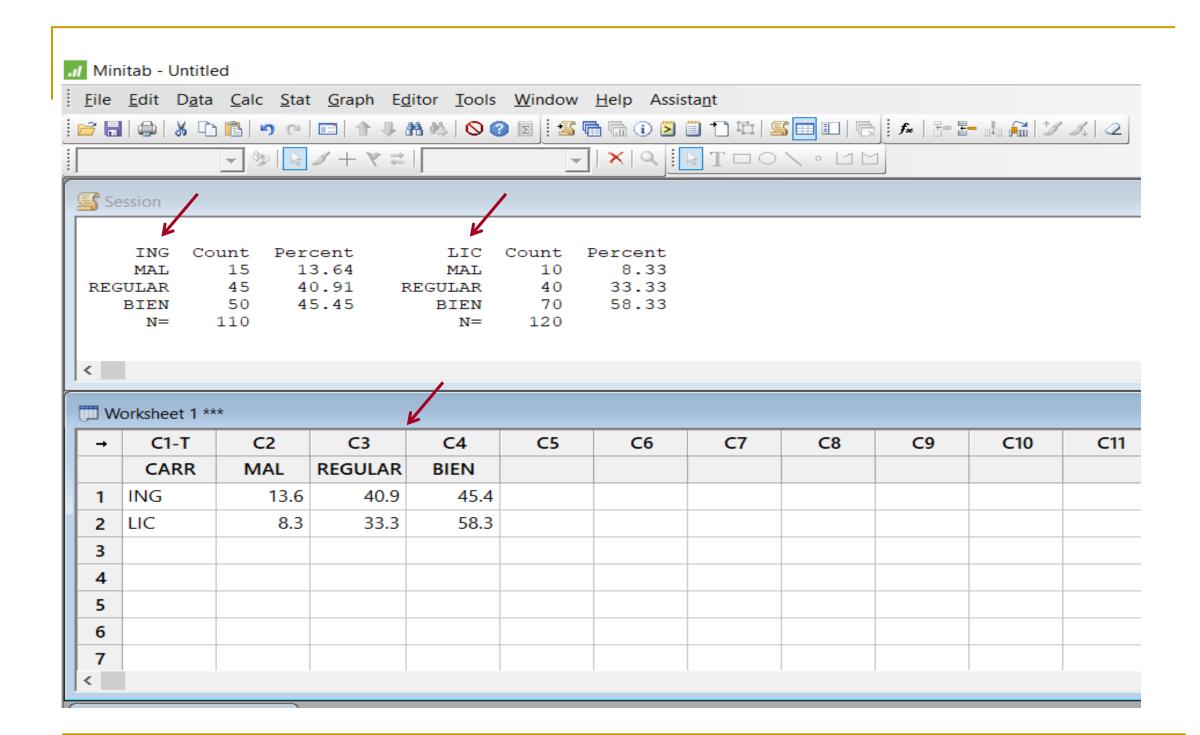


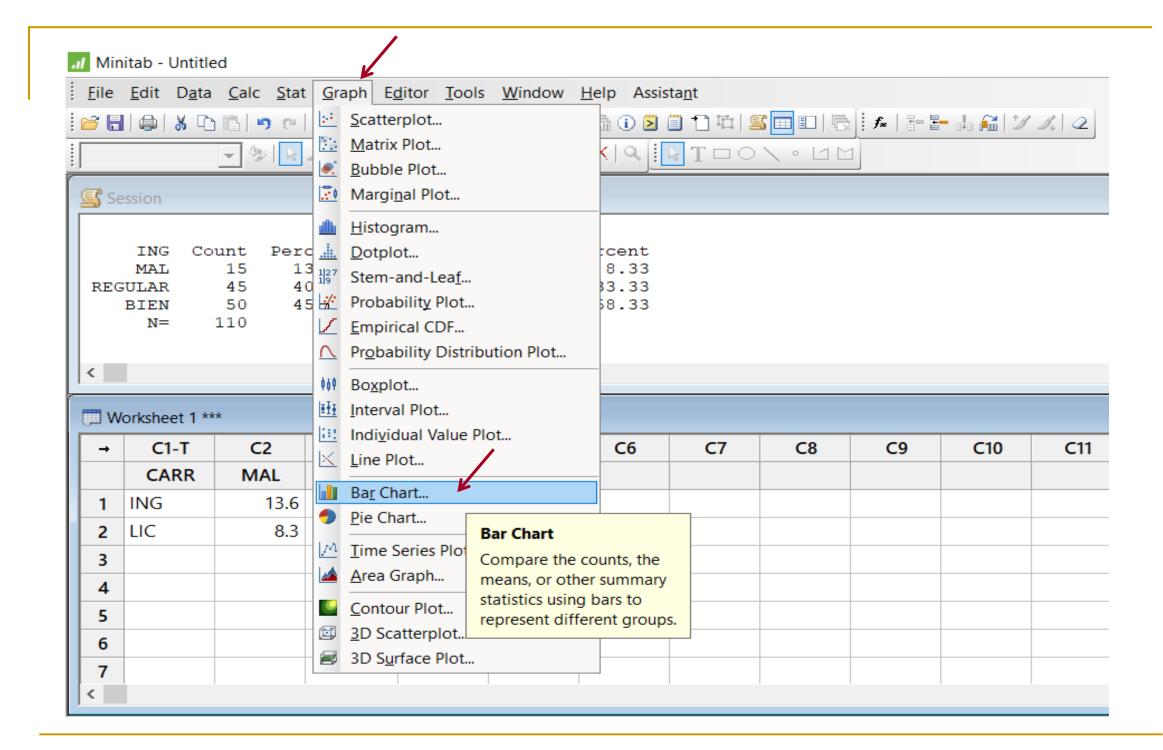


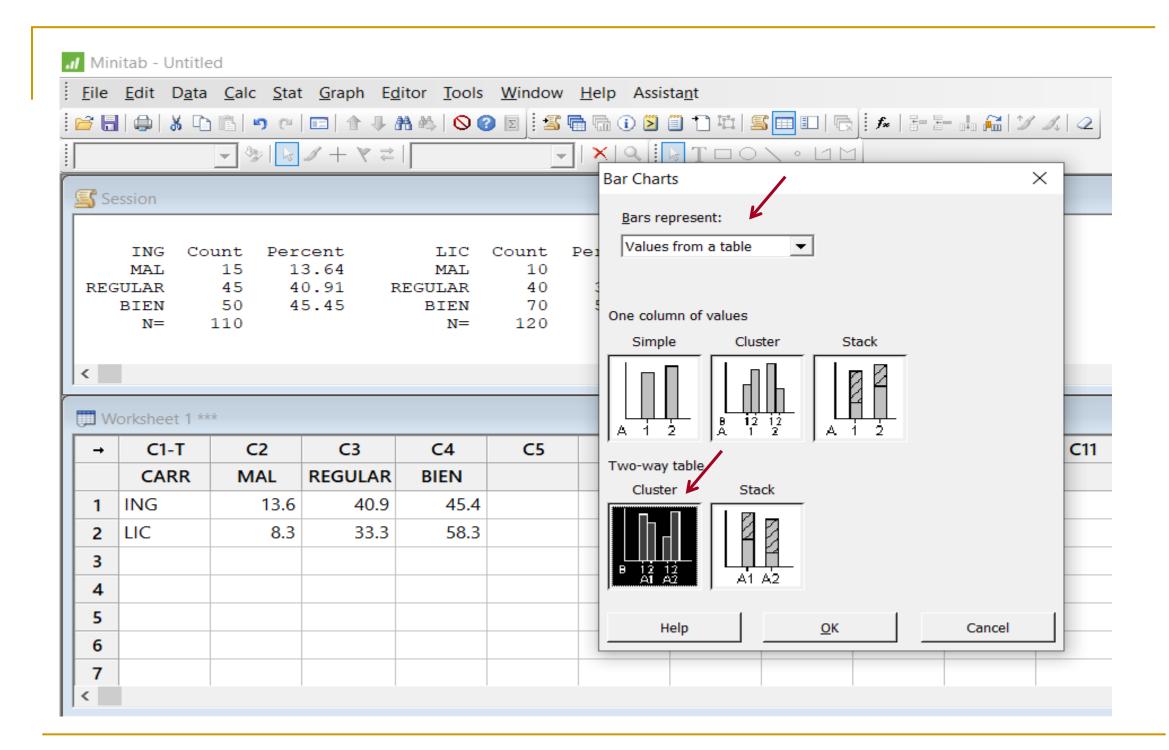


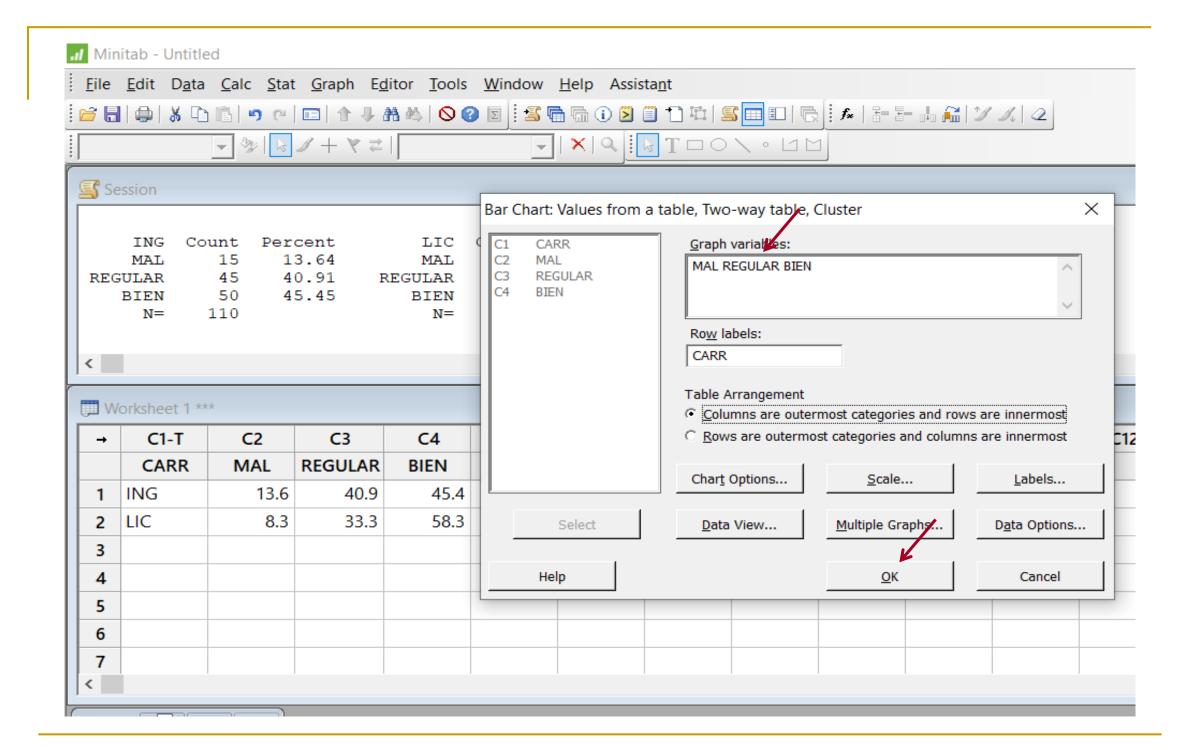


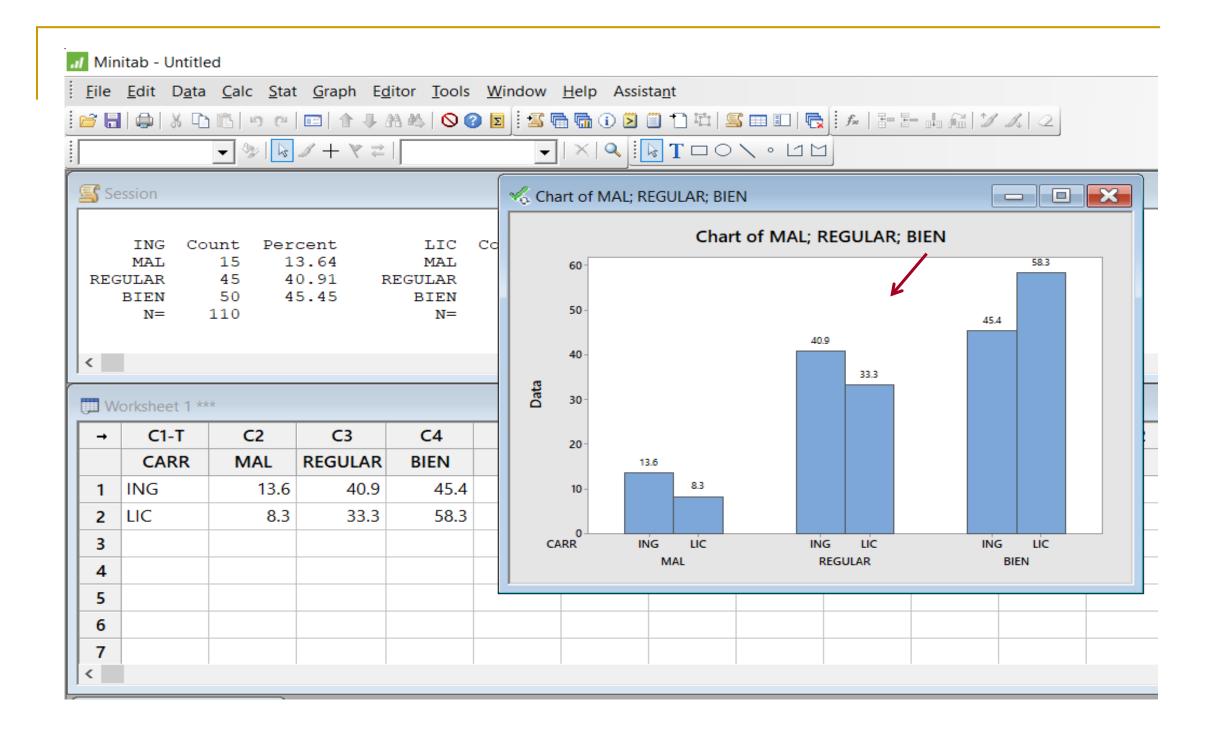






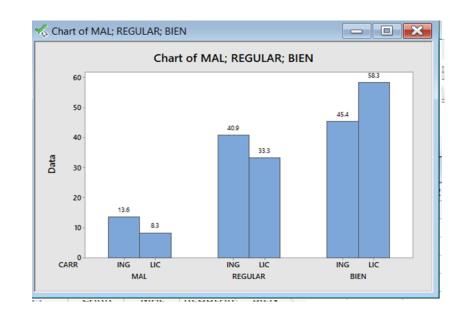






ANÁLISIS

Carrera	INGEN	NIERÍA	LICENCIATURA	
	f_{ab}	fac	f_{ab}	Fac
Mal	15	15	10	10
Regular	45	60	40	50
Blen	50	110	70	120
Total	110	110	120	120



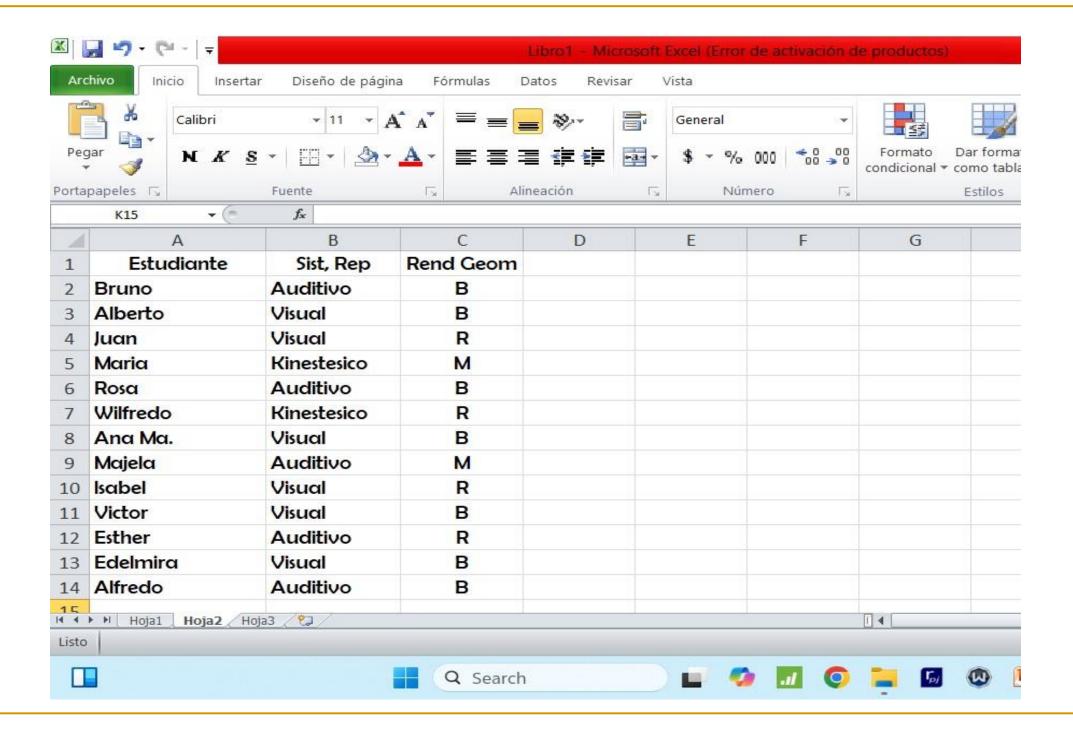
Análisis valorativo

Es probable que de manera general, los estudiantes de la Licenciatura logren mejor rendimiento académico que los estudiantes de Ingeniería. «Organización y presentación en la tabla de frecuencia» Ejemplo de tabla doble entrada relación entre dos variables.

Con el objetivo de indagar si existía alguna *relación* entre los sistemas representativos primarios de los estudiantes y el rendimiento en la asignatura de Geometría...

... se aplican instrumentos estandarizados para determinar el Sistema Representativo Primario a una muestra de 100 estudiantes de preuniversitario...

Posteriormente, se les aplicó a todos los estudiantes de la muestra un test de dicha materia, y se computaron los resultados y se construyó la tabla de contingencia.

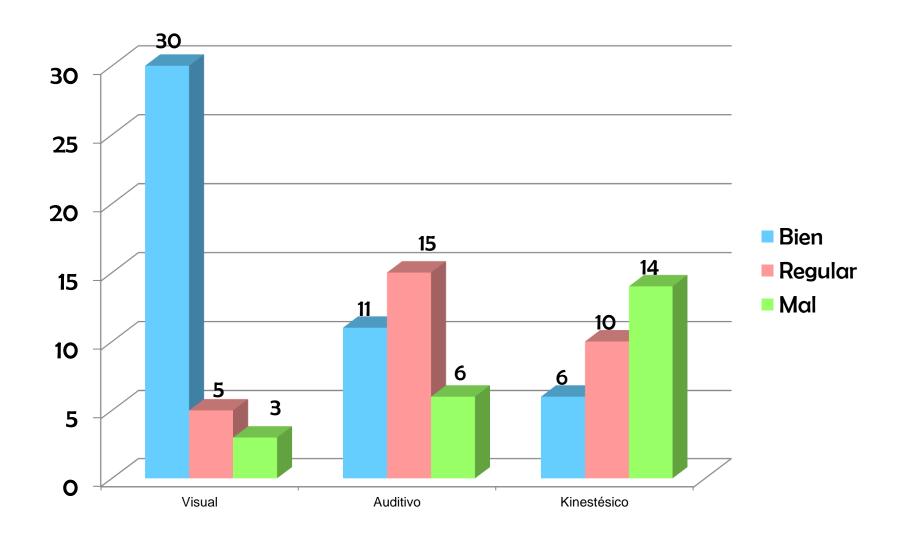


PRESENTACIÓN PARA UNA VARIABLES NOMINAL Y ORDINAL TABLA DE CONTINGENCIA

Sistema predominante

Rendimiento	Visual	Auditivo	Kinestésico	Sub total
Bien	30	11	6	47
Regular	5	15	10	30
Mal	3	6	14	23
Sub total	38	32	30	100

Grafica barras simples



PROGRAMA DE PROCESAMIENTO MINITAB

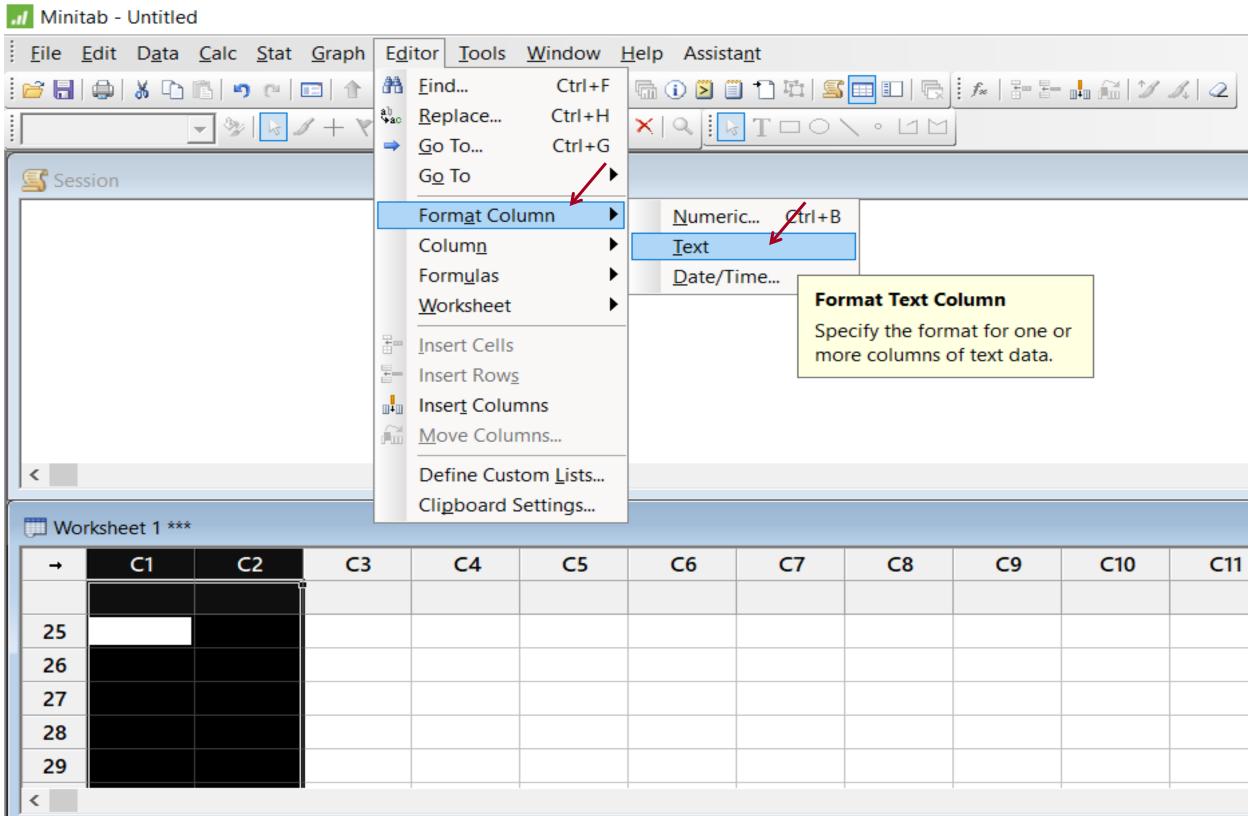


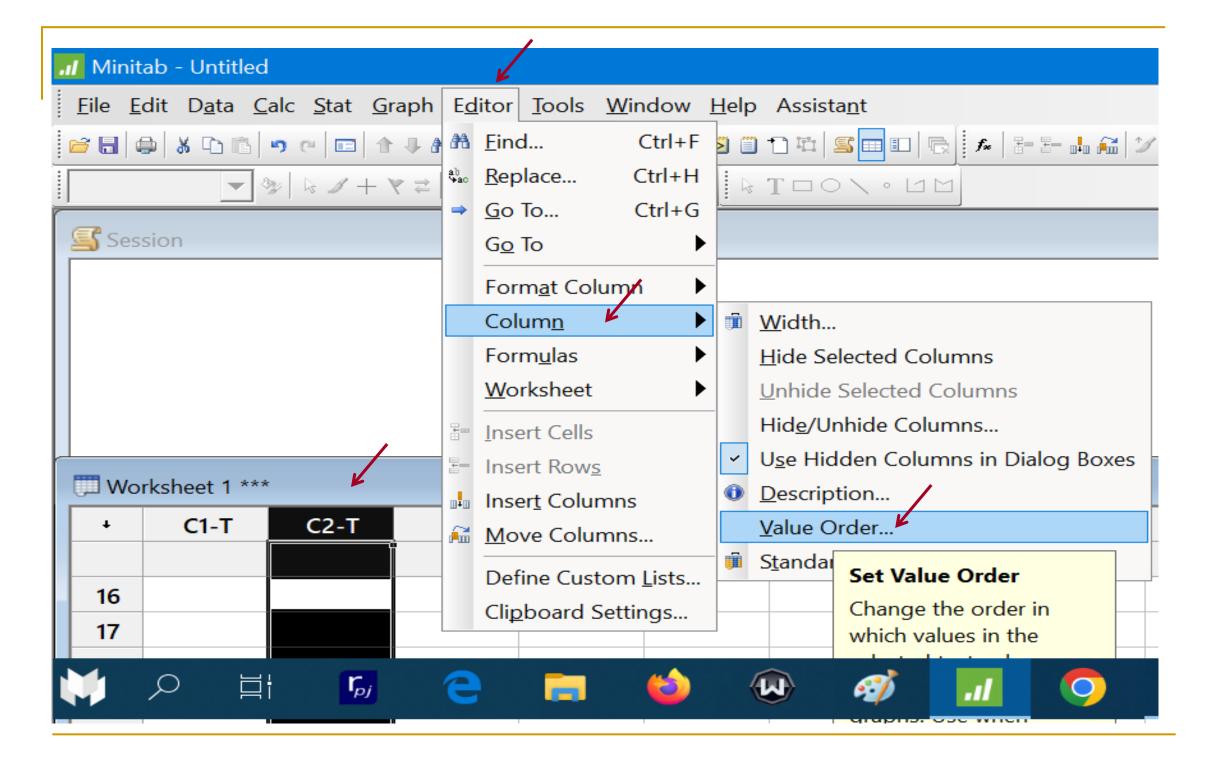
Minitab 17

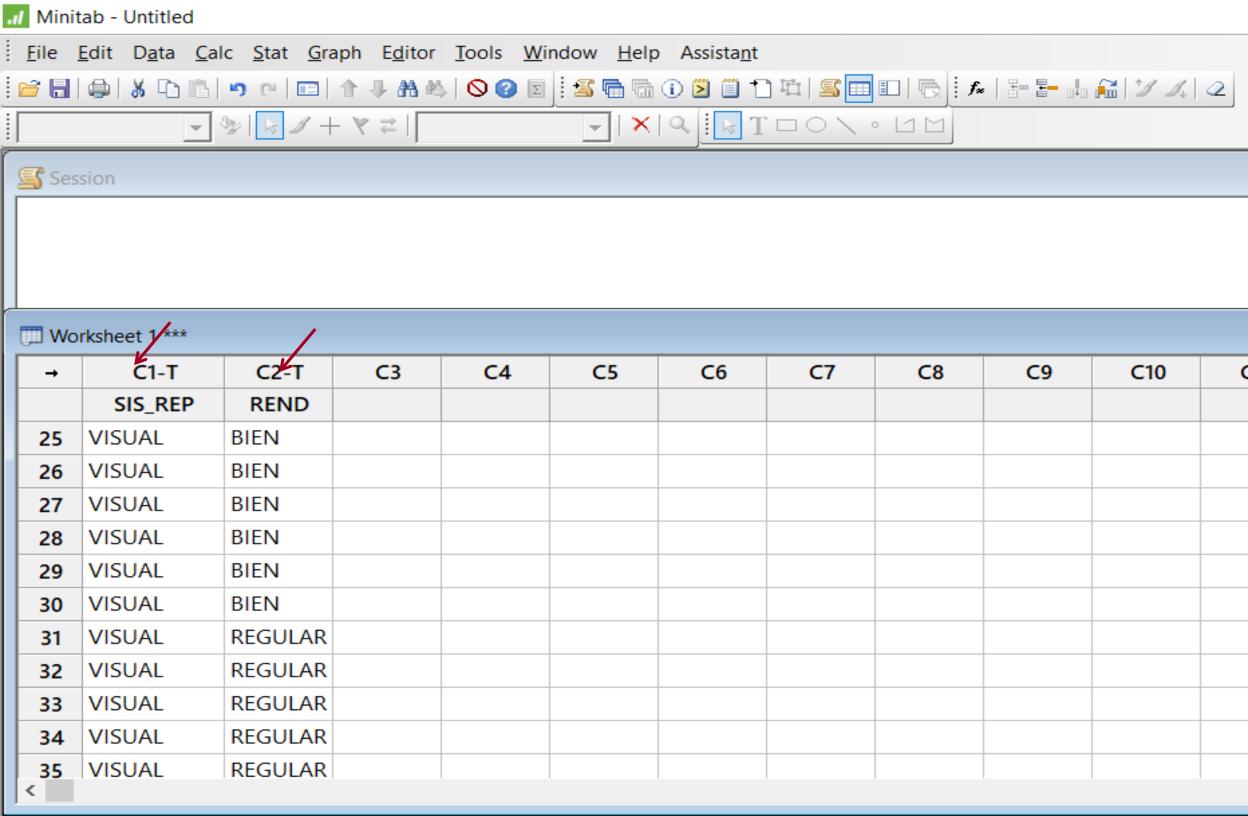
El software de herramientas estadísticas más utilizado en todo el mundo para mejorar la calidad.

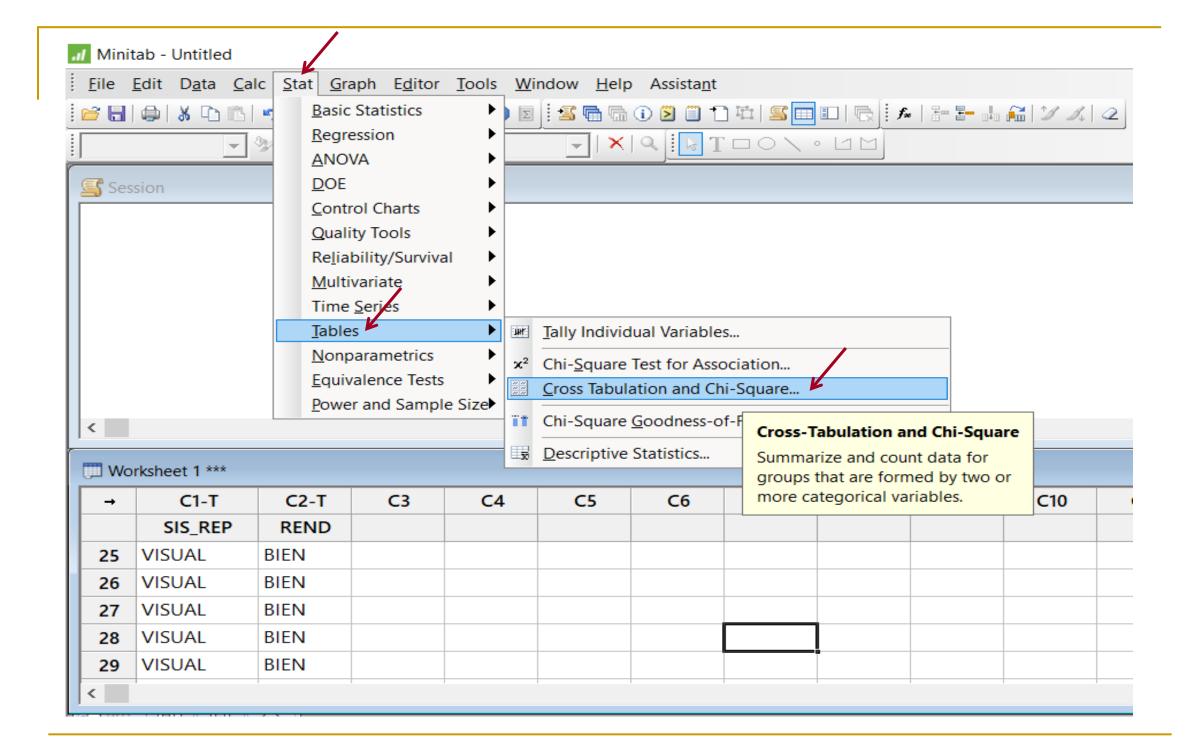
Minitab Inc. es el principal software del mundo para la enseñanza de la Estadística; también, es el software utilizado con mayor frecuencia en Seis Sigma, la principal metodología del mundo para el mejoramiento de la calidad.

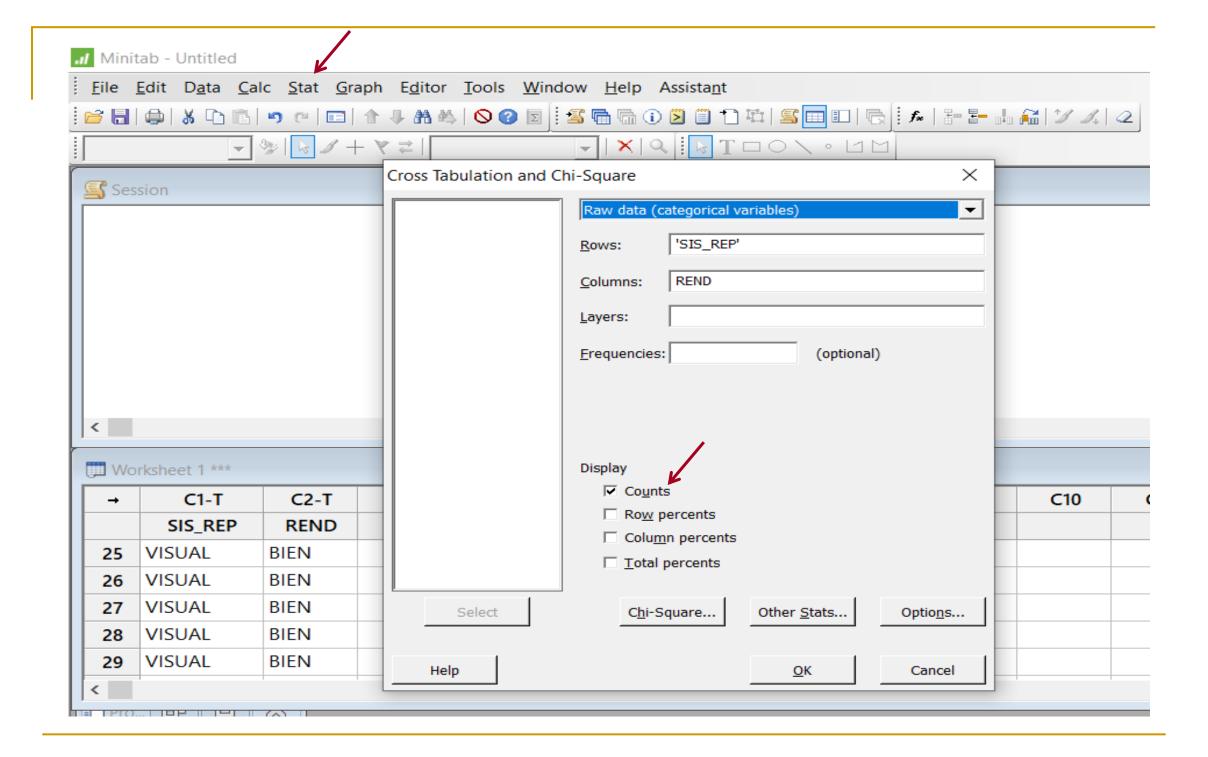
https://drive.google.com/file/d/1T6Oc5FD1BE7r-e_aubmhsNj2B0WY6fJn/view?usp=sharing

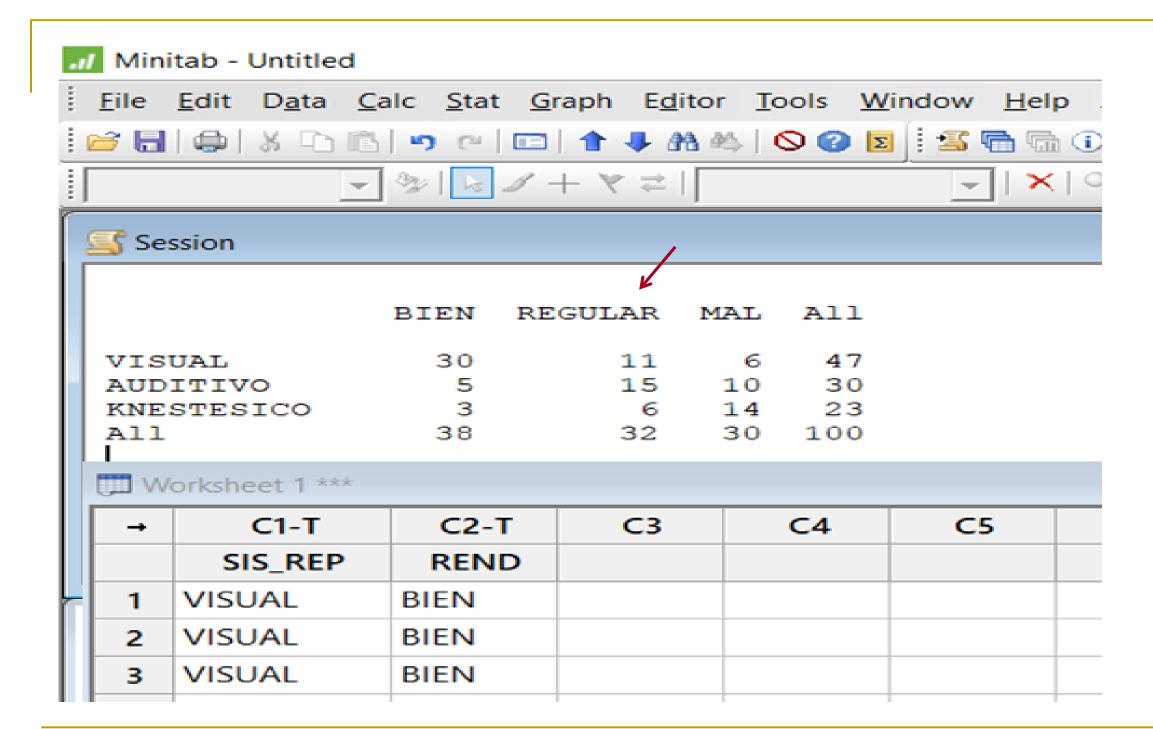


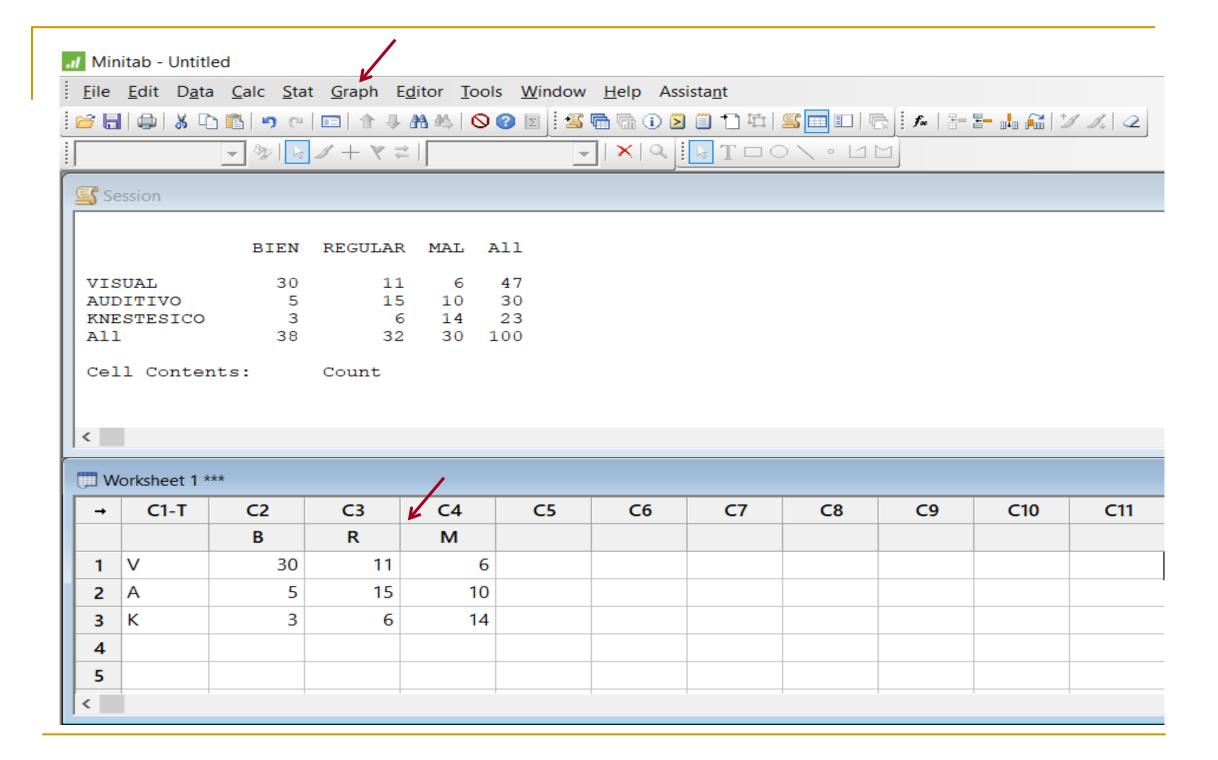


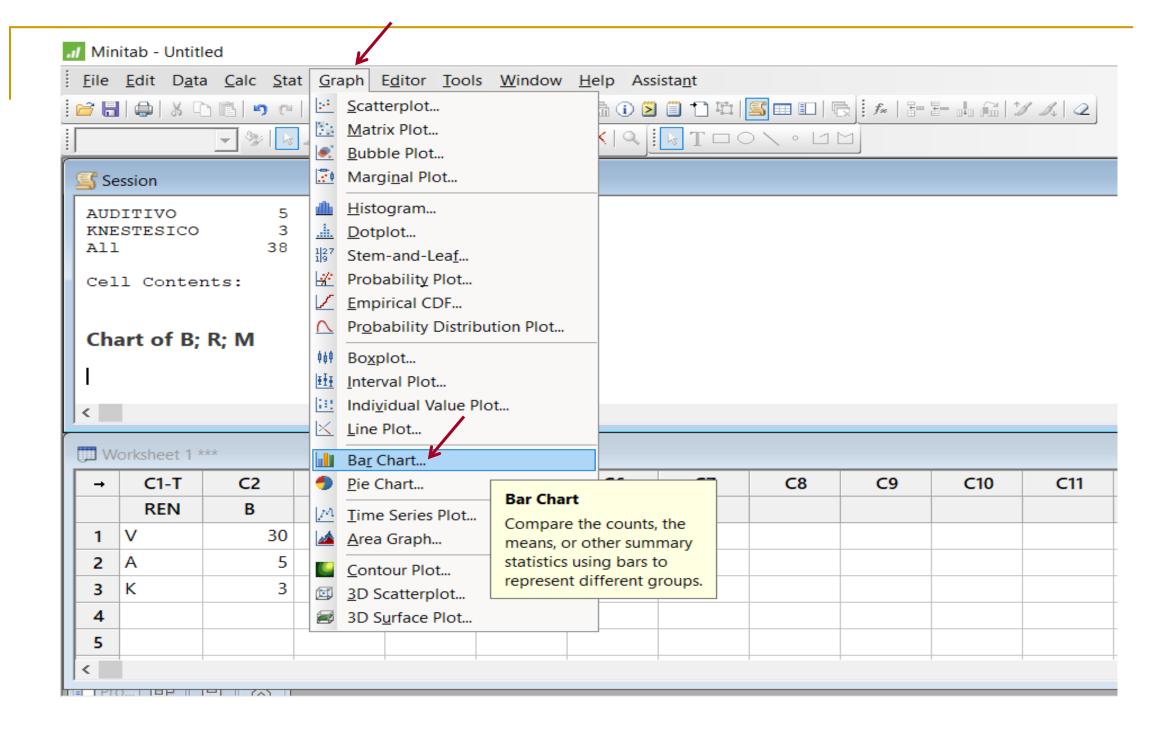


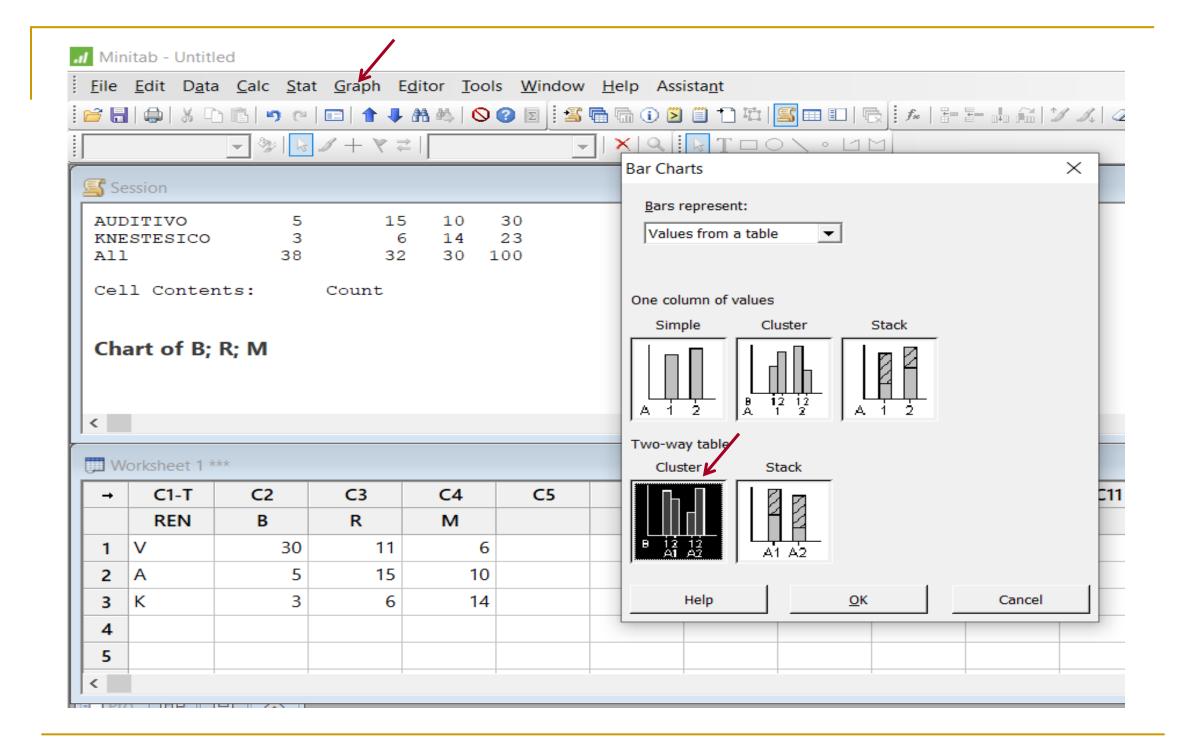


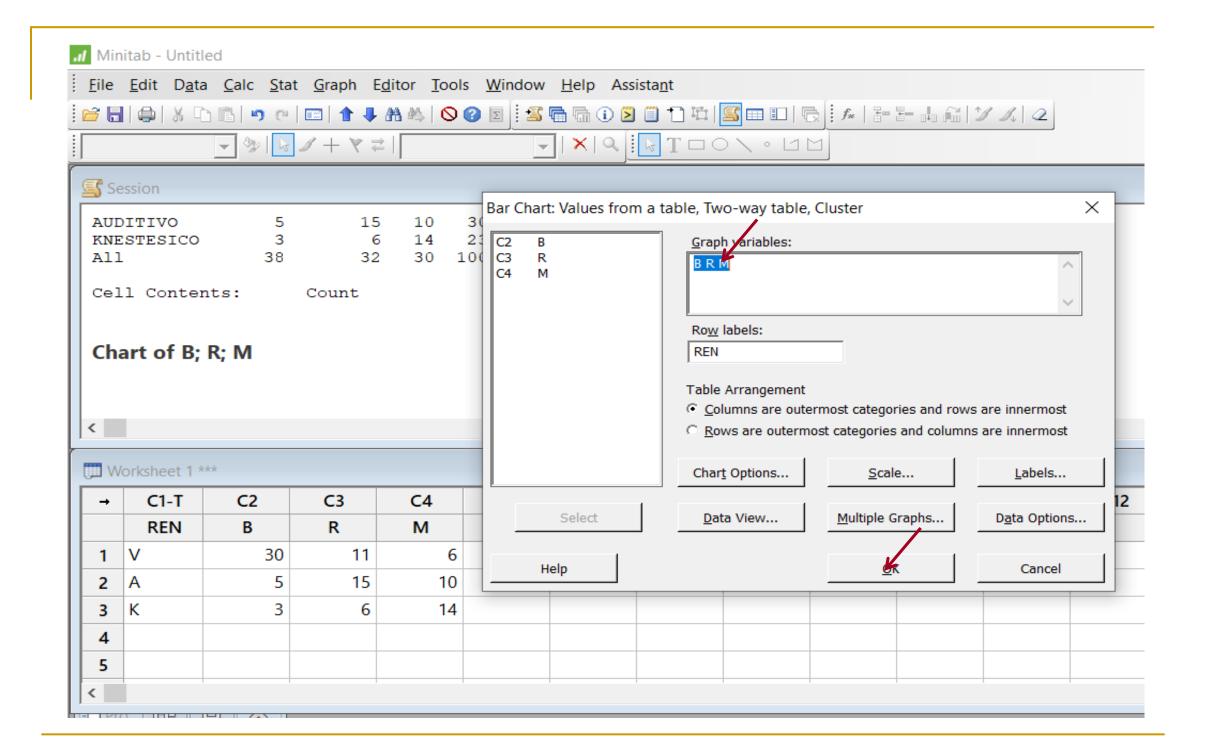


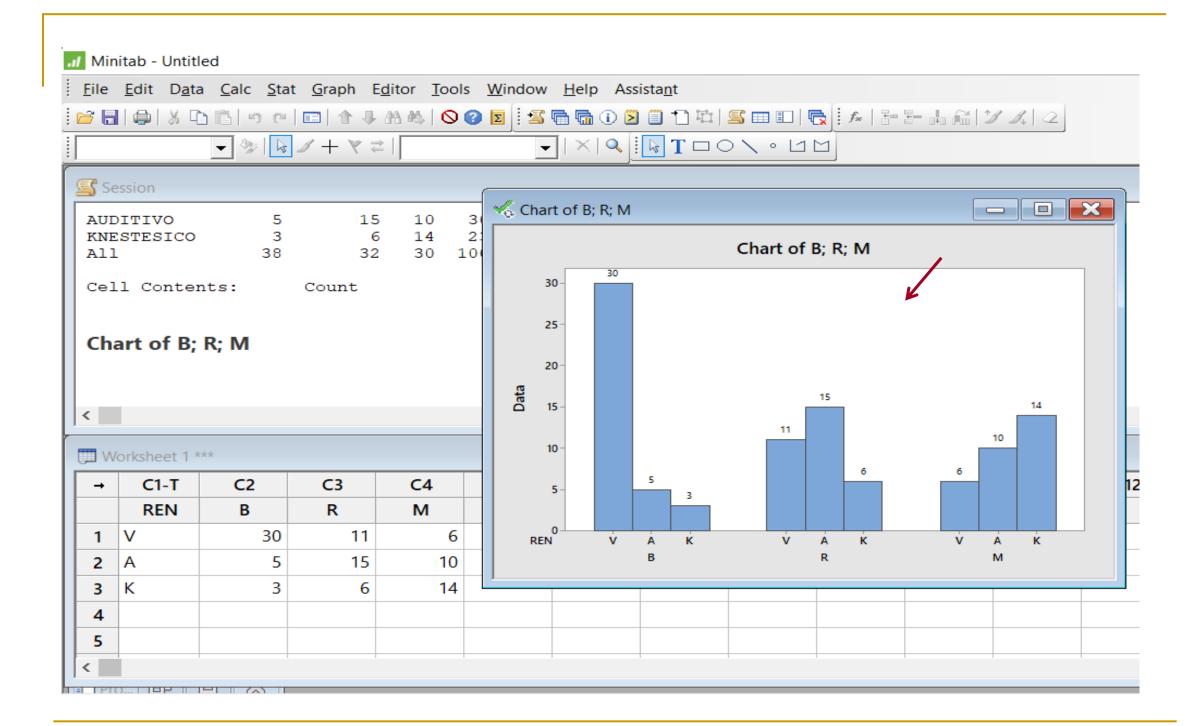








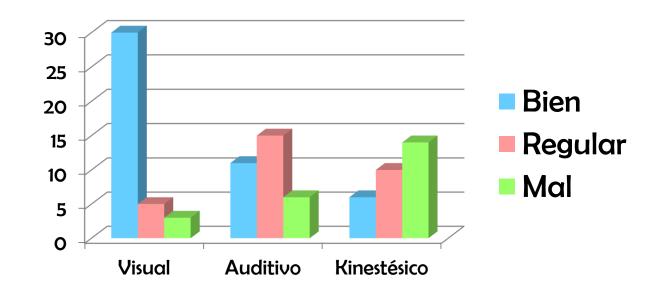




ANÁLISIS

Sistema predominante

	•			
Notas	Visu	Auditi	Kinesté	Sub
	al	VO	sico	total
Bien	30	5	3	38
Regular	11	15	6	32
Mal	6	10	14	30
Sub	47	30	23	100



Análisis valorativo.

En función de la organización y presentación del conjunto datos en la tabla de frecuencia y el grafico, es probable que exista relación o asociación entre el sistema representativo y el rendimiento académico en Geometría

Actividad 3.1

Significando que el rendimiento académico intrínsecamente corresponde a un escalamiento ordinal, en la siguiente tabla se presenta el registro de las calificaciones de la prueba de Matemática de un grupo de 30 estudiantes en los *rangos ordinales del 1º al 10º*.

Registro en escalamiento Ordinal

90	8ō	10 ^{<u>a</u>}	9º	9º	8₀
10º	9º	80	10º	9 <u>a</u>	7 º
90	80	10º	9º	8 º	80
9º	10º	7 º	8ō	90	80
80	9º	80	10º	9⁰	7 º

Elabore la tabla de frecuencia y el gráfico correspondiente.

SÍNTESIS DE LOS CONJUNTOS DE DATOS

MEDIDAS DE TENDENCIA CENTRAL

MEDIDAS DE TENDENCIA DISPERSIÓN

Estadígrafos

- Moda (Mo)
- Mediana (Me)
- Media (x̄)

Estadígrafos:

- Varianza (S²)
- Desviación Estándar (S)
- Coeficiente de Variación (Cv)

MEDIDAS DE TENDENCIA CENTRAL

- Moda (Mo): Categoría o rango de mayor frecuencia en el conjunto de datos.
- Mediana (Md): Valor que registra el elemento que ocupa la posición central del conjunto de datos ordenados.
- Media () Suma de todos los valores registrados, dividido entre el total.

Para datos agrupados

$$M_{d} = L_{i} + \frac{\frac{N}{2} - \sum f_{M_{d}-1}}{f_{M_{d}}} * A \qquad \qquad \overline{x} = \frac{\sum mc \times f_{ab}}{n}$$

MEDIDAS DE DISPERSIÓN

- Varianza (S²): Estadígrafo que determina el valor medio del cuadrado de la desviación de cada elemento respecto a su media.
- Desviación estándar (S): La raíz cuadrada de la varianza.
- Coeficiente de variación (Cv): Estadígrafo que determina la dispersión relativa de un conjunto de datos.

Para datos agrupados

$$S^{2} = \frac{1}{n-1} \sum (mc_{i} - \overline{x})^{2} fab_{i} \qquad S = \sqrt{S^{2}} \qquad Cv = \frac{S}{\overline{x}} \times 100$$

EJEMPLO:

Se indaga acerca del peso corporal entre los estudiantes del centro. Peso en Kg en una muestra de 30 estudiantes.

65.4	64.2	64.8	65.3	64.5
64.0	64.7	65.1	64.9	65.8
64.4	65.3	<u>60.2</u>	62.4	63.2
63.7	63.5	62.1	62.7	63.9
62.8	64.2	65.1	70.5	<u>71.2</u>
68.4	69.3	68.1	68.4	69.3

Construcción de la tabla de frecuencia Algoritmo para la determinación de los intervalos

1º Cálculo del Recorrido

$$R = 71.2 - 60.2 = 11$$

2º Determinación del # de intervalos

Fórmula de Sturges K = 1 + 3.323 log N

$$K = 1 + 3.323 \log (30) = 1 + 3.323 (1.48) = 5.91$$

 $K \approx 6$

3º Cálculo de la amplitud del intervalo

$$A = 11/6 \approx 2$$

La tabla de frecuencia describe el peso en Kg de 30 estudiantes.

Rango	mc	f_{ab}	f acum	Calcular
[60-62[61	1	1	Mediana ?
[62-64[63	7	8	Media ?
[64-66[65	6	14	- Uawiawa-a2
[66-68[67	8	22	■ Varianza?
[68-70[69	5	27	Desviación Estándar?
[70-72[71	3	30	Coeficiente de variación?

CÁLCULO DE LA MEDIANA

Rango	\mathbf{f}_{ab}	f acum	
[60-62[1	1	$M_d = L_i + \frac{\frac{N}{2} - \sum f_{M_d - 1}}{\epsilon} * A$
[62-64[7	8	$f_{M_d} - L_i + f_{M_d}$
[64-66[6	14	
[66-68[8	22	$M_d = 66 + \frac{\frac{30}{2} - 14}{2} * 2 = 66,25$
[68-70[5	27	$M_d = 66 + \frac{8}{8} \times 2 = 66,25$
[70-72[3	30	

CÁLCULO DE LA MEDIA

Rango	mc	f ab	mc * f _{ab}	
[60-62[61	1	61	$\sum mc \times Fab$
[62-64[63	7	441	$\overline{x} = \frac{\sum mc \times Fab}{n}$
[64-66[65	6	390	
[66-68[67	8	536	$\overline{x} = \frac{1986}{30} = 66,2$
[68-70[69	5	345	30
[70-72[71	3	213	
		total	1986	

Construcción auxiliar para el cálculo de medidas de variación

Rango	mc	f_{ab}	$\overline{\mathcal{X}}$	$(mc-\overline{x})^2$	$\left[(mc - \overline{x})^2 * fab \right]$
[60-62[61	1	66,2	27,04	27,04
[62-64[63	7	66,2	10,24	71,68
[64-66[65	6	66,2	1,44	8,64
[66-68[67	8	66,2	0,64	5,12
[68-70[69	5	66,2	7,84	39,2
[70-72[71	3	66,2	23,04	69,12
				Total	220,80

Cálculo de las medidas de dispersión

$$S^{2} = \frac{1}{n-1} \sum (mc_{i} - \bar{x})^{2} f_{ab_{i}} \qquad S^{2} = \frac{1}{30-1} 220,80 = 7,61$$

$$S^2 = \frac{1}{30-1} 220,80 = 7,61$$

$$S = \sqrt{\frac{1}{n-1}} \sum (mc_i - x)^2 f_{ab_i}$$

$$S = \sqrt{S^2} = \sqrt{7,61} = 2,76$$

$$Cv = \frac{S}{\overline{x}} \times 100$$

$$Cv = \frac{2,76}{66,2} \times 100 = 4,17$$

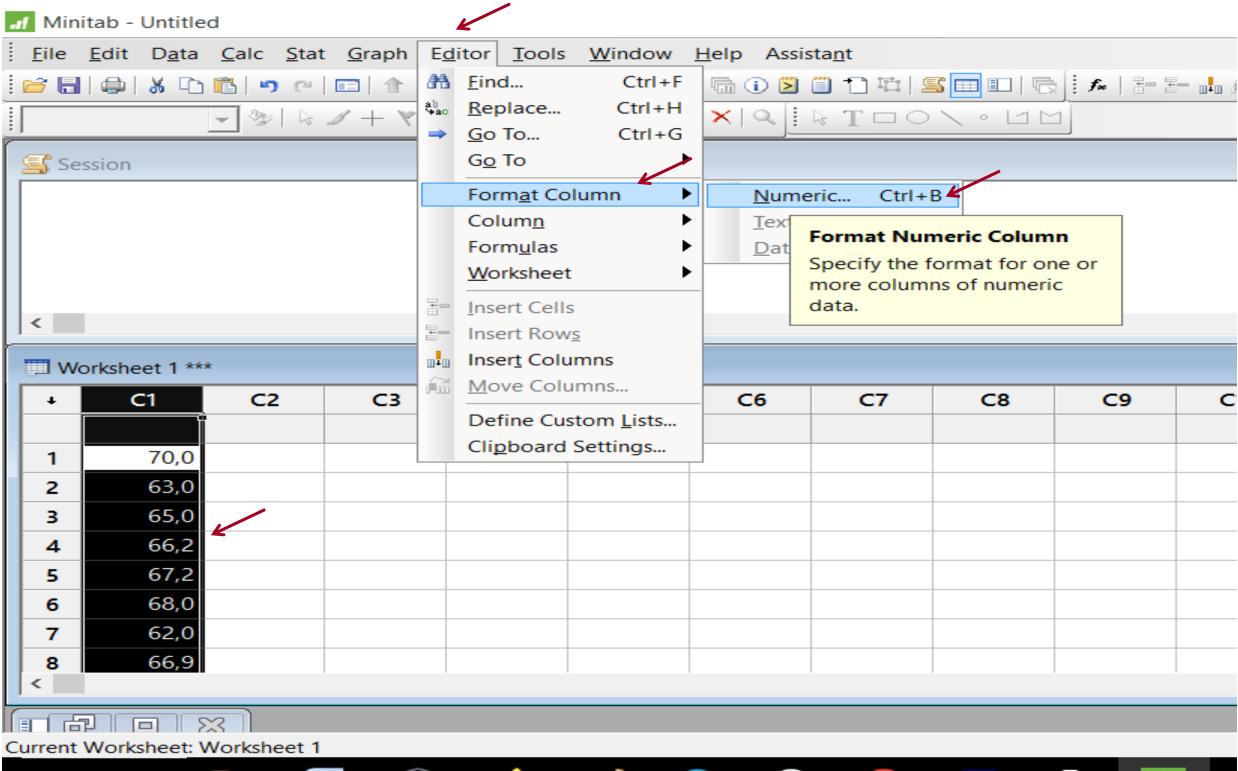
PROGRAMA DE PROCESAMIENTO MINITAB



Minitab 17

El software de herramientas estadísticas más utilizado en todo el mundo para mejorar la calidad.

Minitab Inc. es una compañía privada cuya sede principal se encuentra en State College, Pensilvania, y tiene subsidiarias en el Reino Unido, Francia y Australia. con representantes y distribuidores en muchos países alrededor del mundo.

















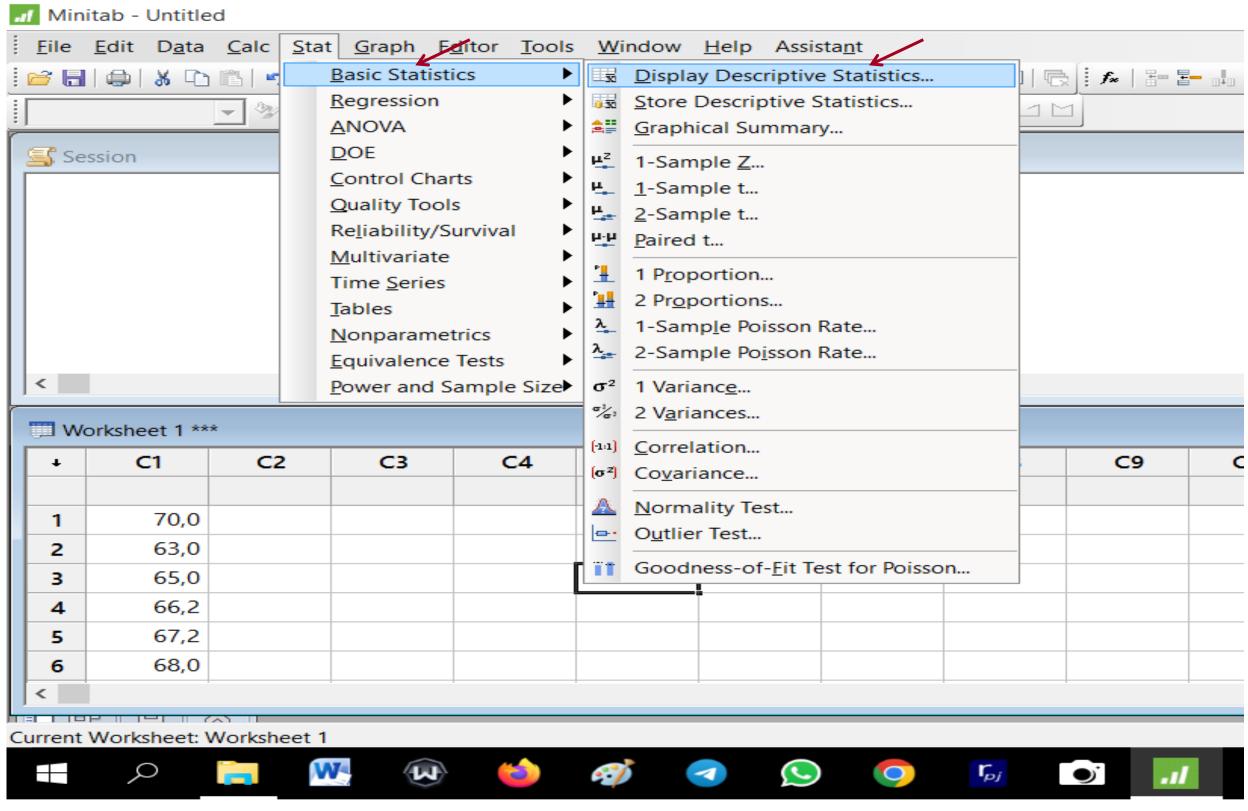


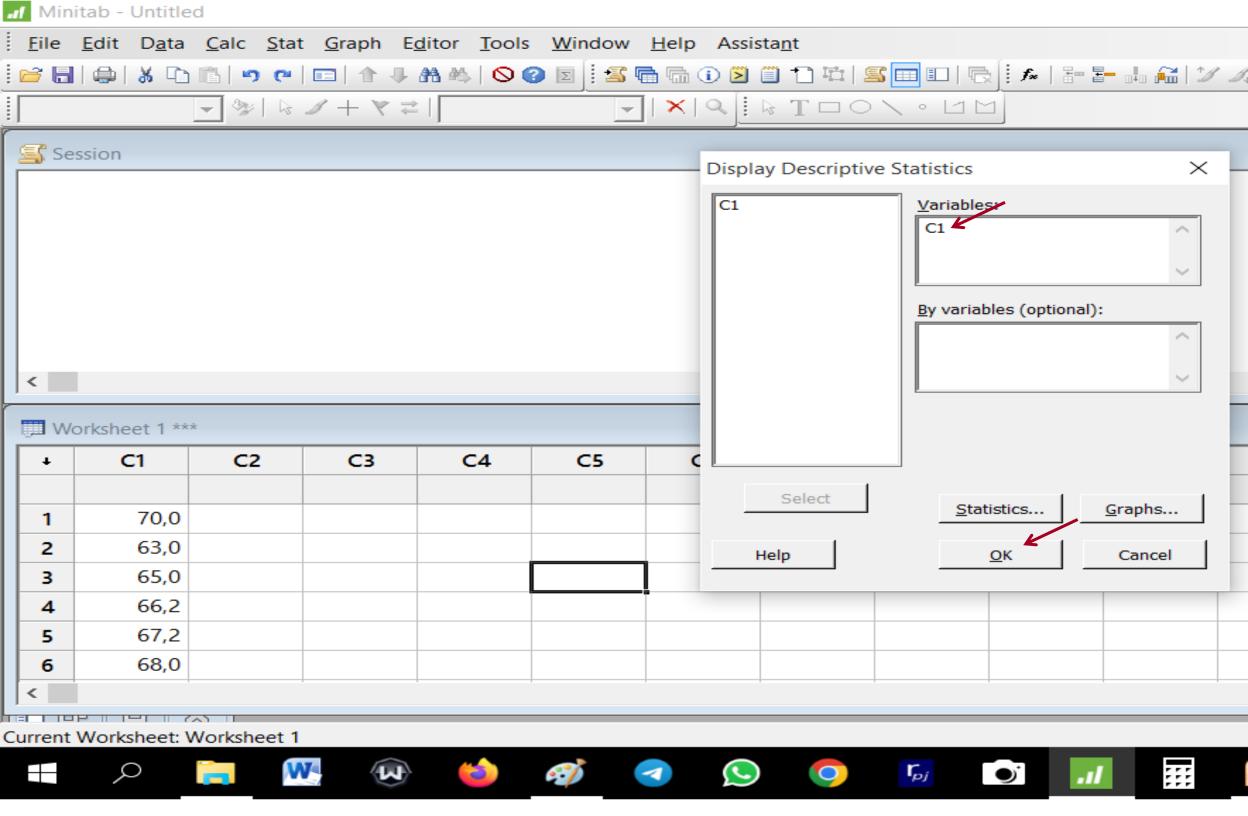


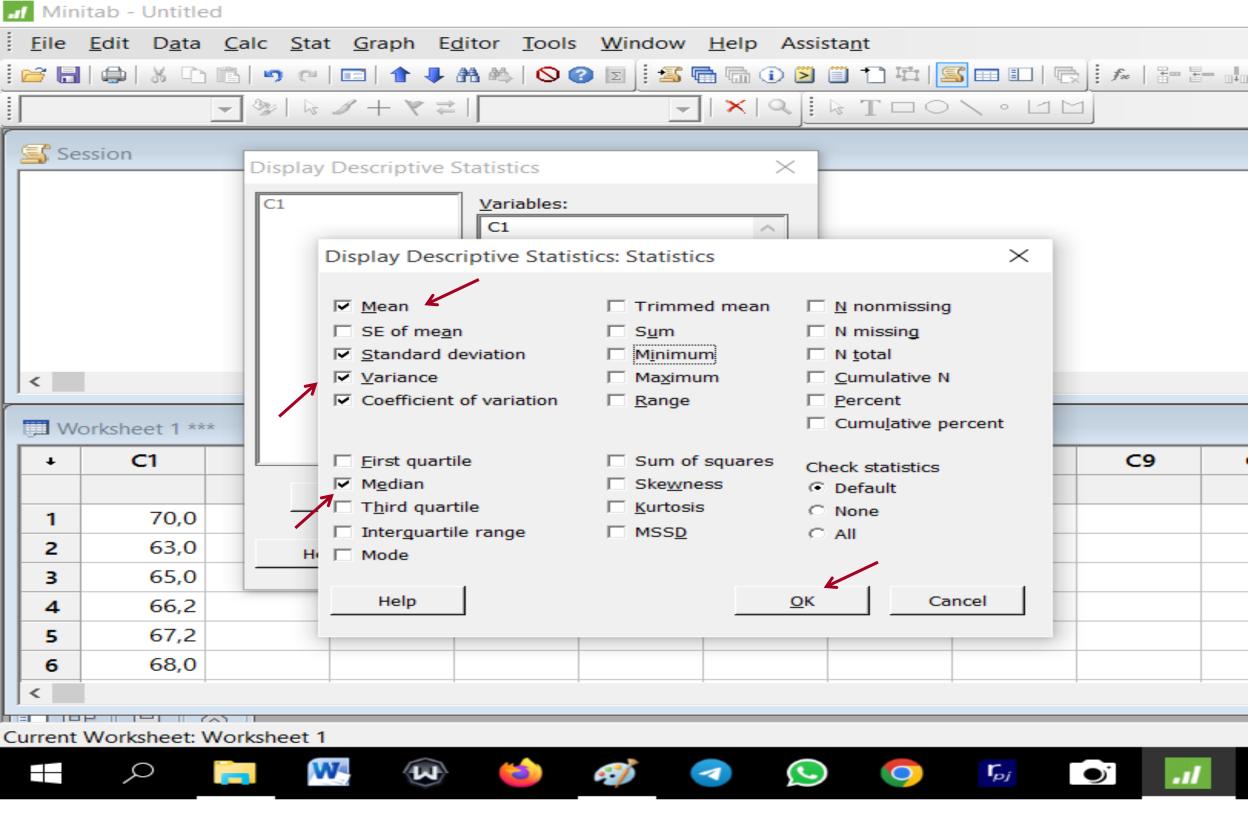


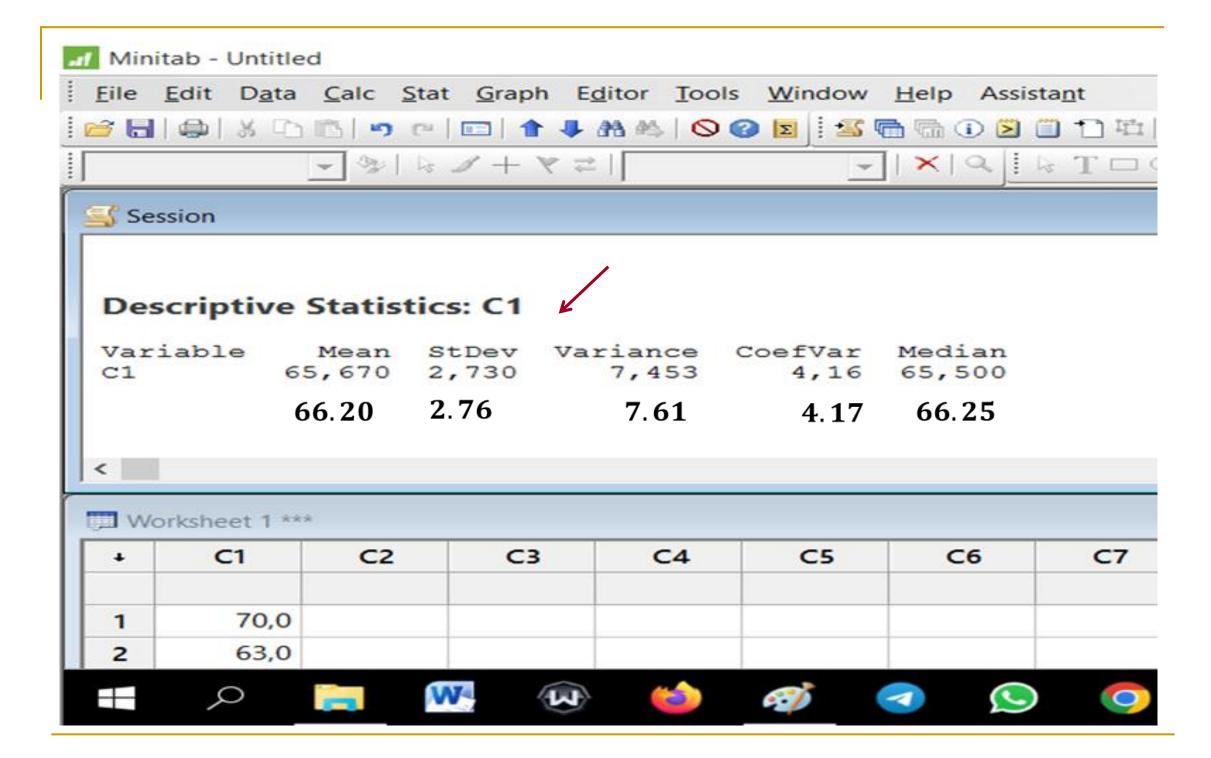












CONCLUSIONES:

La «presentación en tablas, gráficos y la síntesis en estadígrafos de tendencia central y dispersión», en el proceso de transformación de «datos-información», facilita el análisis para comparar, descubrir regularidades, formular hipótesis, llegar a conclusiones y socializar sus resultados científicos.

La *naturaleza intrínseca del atributo* objeto de estudio determina la *escala de medición* para su registro y ésta a su vez, condiciona la selección de uno de los diferentes tipos de *tablas, gráficas y estadígrafos* para la organización y presentación y síntesis del conjunto de datos.

Actividad 4.1

Significando que el rendimiento académico intrínsecamente corresponde a un escalamiento ordinal, en la siguiente tabla se presenta el registro de las calificaciones de la prueba de Matemática de un grupo de 30 estudiantes en los rangos ordinales del 1º al 10º.

Registro en escalamiento Ordinal

9º	8 º	10 ^{<u>a</u>}	9⁰	9º	8ō
10º	90	80	10º	9 <u>a</u>	7 º
9 º	8ō	10º	9º	8 <u>o</u>	8ō
9⁰	10º	7 º	8ō	90	8ō
8 <u>o</u>	9º	8 <u>o</u>	10º	90	7 º

Calcular el estadígrafo que corresponda al escalamiento empleado e interprete el resultado.