

Ocho pasos para el desarrollo de una investigación

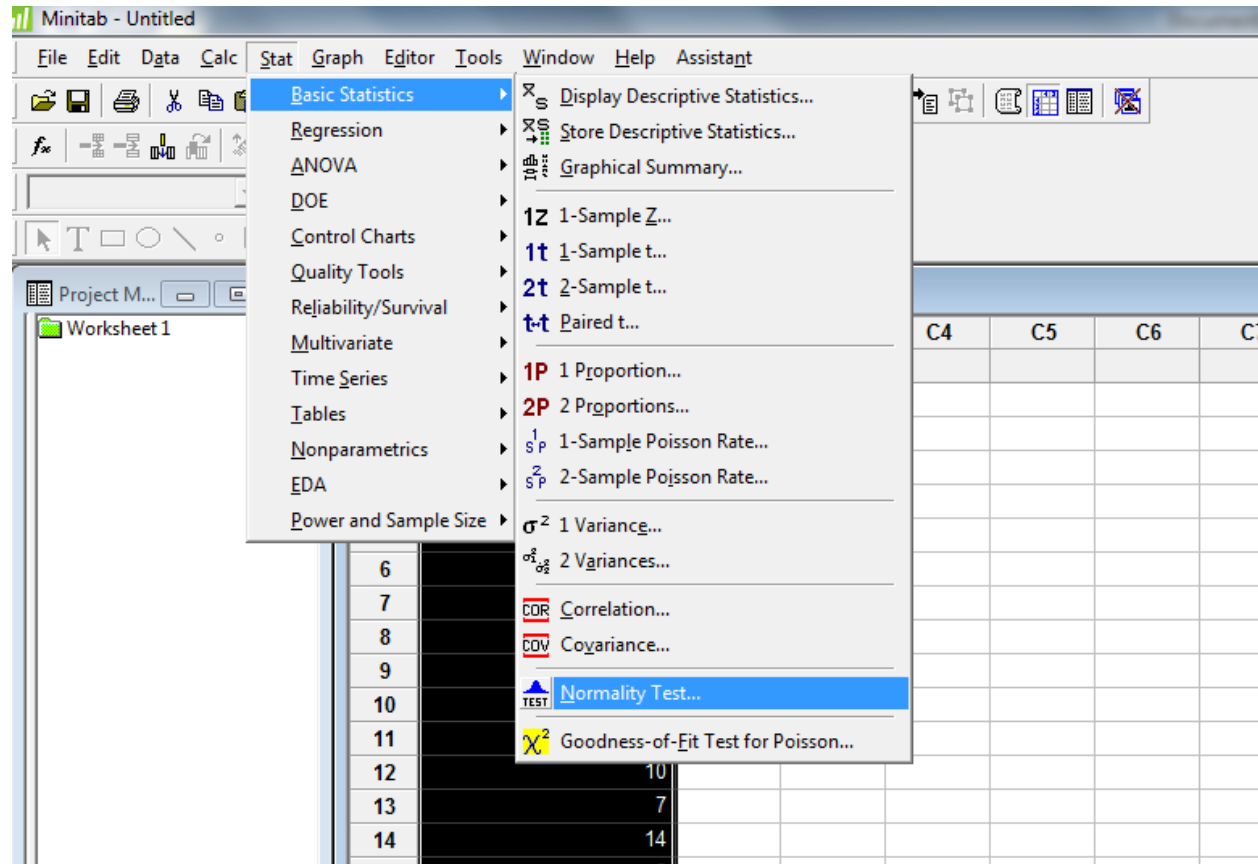
**Ejemplos del Suplemento 1 resueltos con  
Minitab**

## Tabla de Contenido

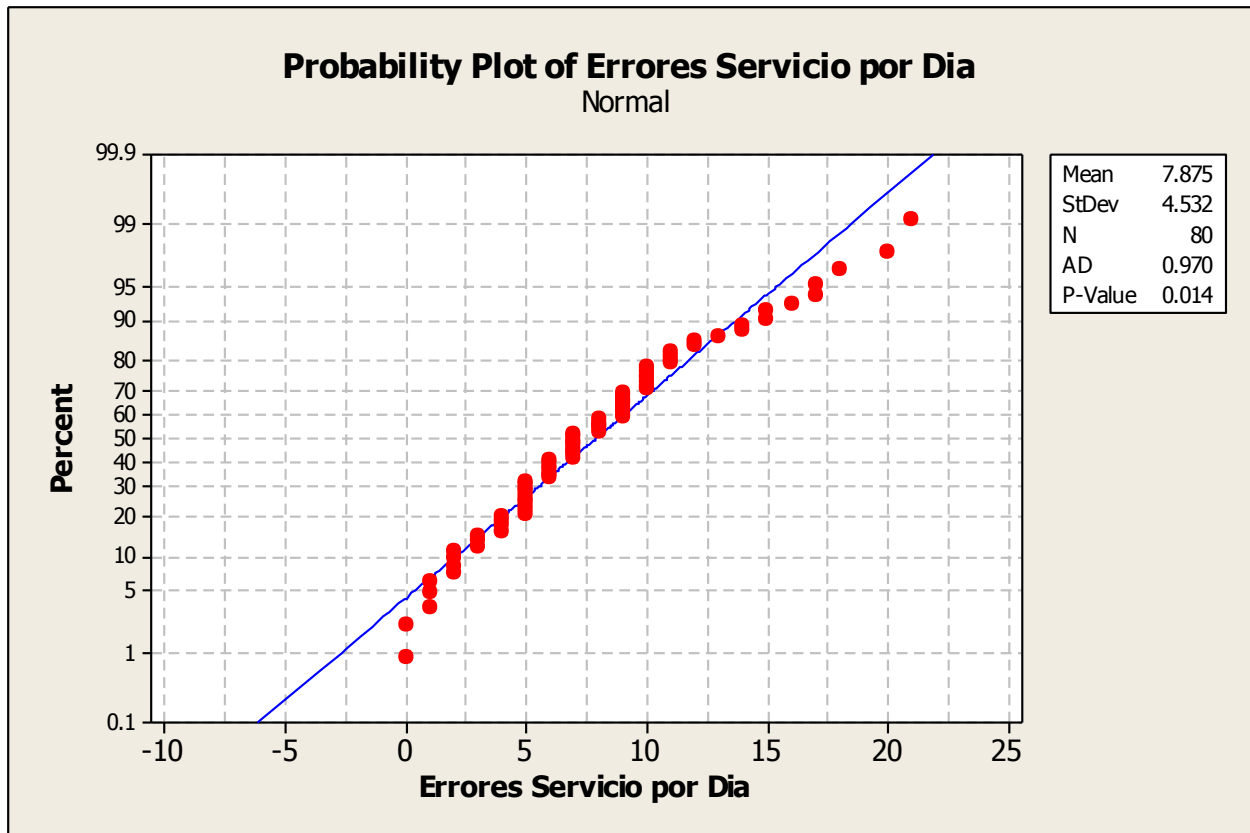
<b>Prueba de Normalidad</b> .....	<b>3</b>
<b>Regresión Lineal</b> .....	<b>7</b>
<b>Prueba t</b> .....	<b>10</b>
<b>Prueba Z</b> .....	<b>13</b>
<b>Análisis de Varianza</b> .....	<b>16</b>
<b>Prueba del Signo</b> .....	<b>19</b>
<b>Ji Cuadrado o <math>X^2</math></b> .....	<b>21</b>
<b>Coefficiente de Correlación de Spearman (rs)</b> .....	<b>23</b>
<b>Prueba de Kendall Tau (<math>\tau</math>)</b> .....	<b>26</b>
<b>Prueba de Kruskal Wallis</b> .....	<b>28</b>
<b>Prueba de Mann Whitney</b> .....	<b>31</b>
<b>Wilcoxon Sign-Rank</b> .....	<b>33</b>
<b>Coefficiente Alfa de Cronbach</b> .....	<b>35</b>

## Prueba de Normalidad

En Minitab, luego de entrar los datos en una sola columna, se selecciona la secuencia **Stat>Basic Statistics>Normality Test:**



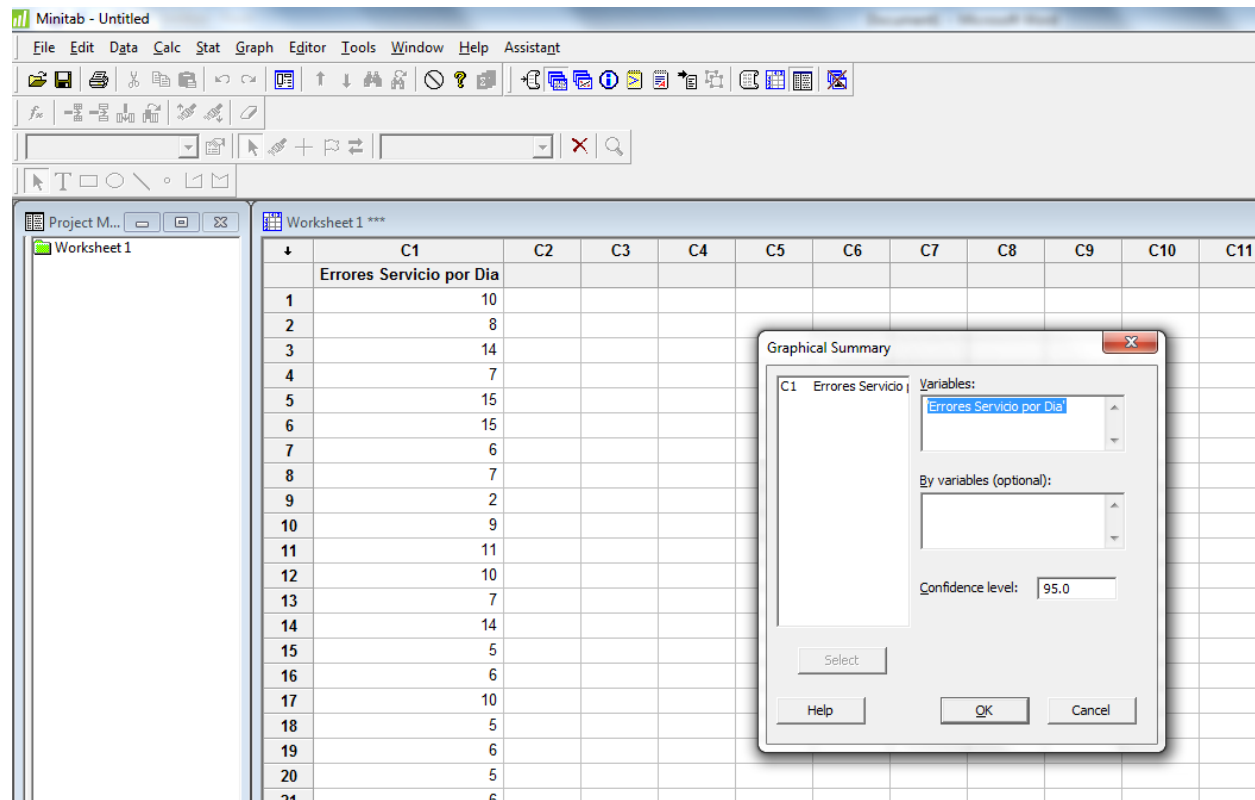
Esto dará un resultado gráfico con la siguiente forma:



El valor P (P-Value) menor de 0.05 nos dice que hay evidencia en contra de la distribución normal de los datos. La gráfica muestra que los datos no siguen la línea de distribución normal.

También podemos buscar un análisis más completo de los datos, incluyendo el resultado de normalidad. Para esto, use la secuencia **Stat>Basic Statistics>Graphical Summary**. Esto lo llevará a un menú para escoger las variables a analizarse con el intervalo de confianza deseado.

Escoja la columna donde tiene los datos para análisis y oprima **OK**.

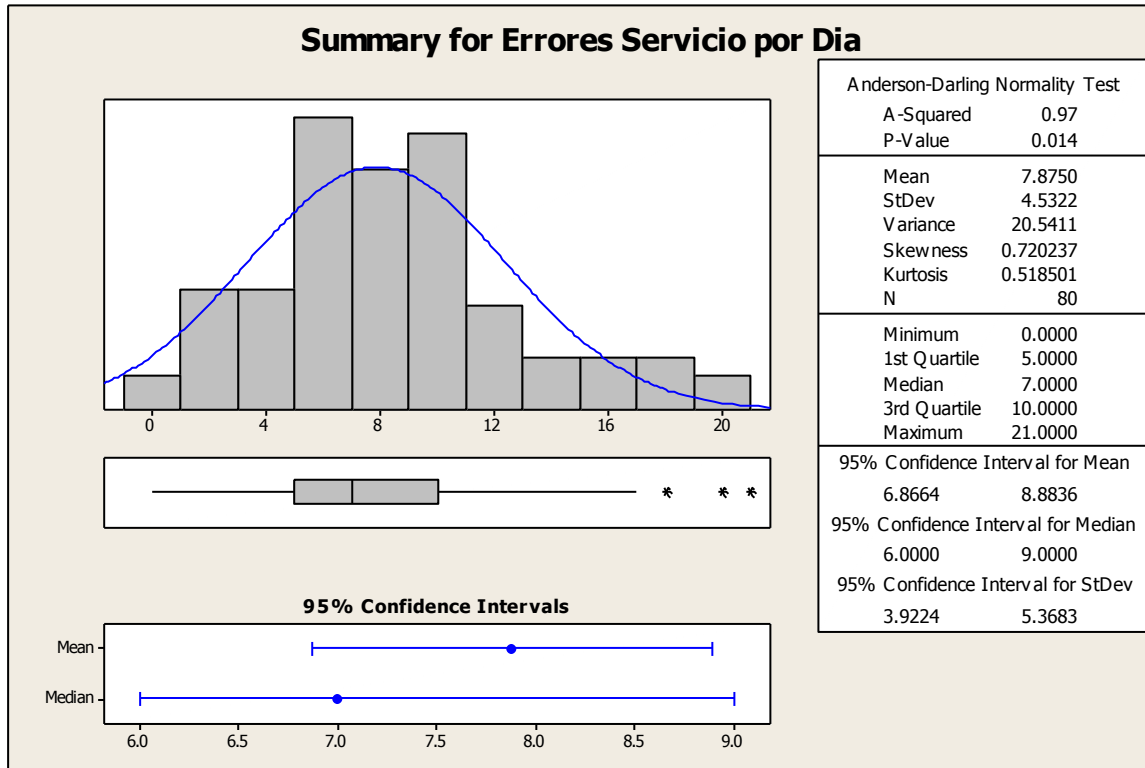


The screenshot shows the Minitab software interface. The main window displays a worksheet with the following data:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	Errores Servicio por Dia										
1	10										
2	8										
3	14										
4	7										
5	15										
6	15										
7	6										
8	7										
9	2										
10	9										
11	11										
12	10										
13	7										
14	14										
15	5										
16	6										
17	10										
18	5										
19	6										
20	5										
21	6										

The 'Graphical Summary' dialog box is open, showing the variable 'C1 Errores Servicio por Dia' selected in the 'Variables' field. The 'Confidence level' is set to 95.0. The 'Select' button is highlighted.

Esto le demostrará un gráfico con la normalidad, promedio, desviación estándar, varianza, cuartiles, y los intervalos de confianza entre otros.





En el menú del Fitted Line Plot, escoja la columna de la respuesta (Y) y la columna del predictor (X). Escoja **linear** y presione **OK**.

The screenshot shows the Minitab interface with a worksheet containing the following data:

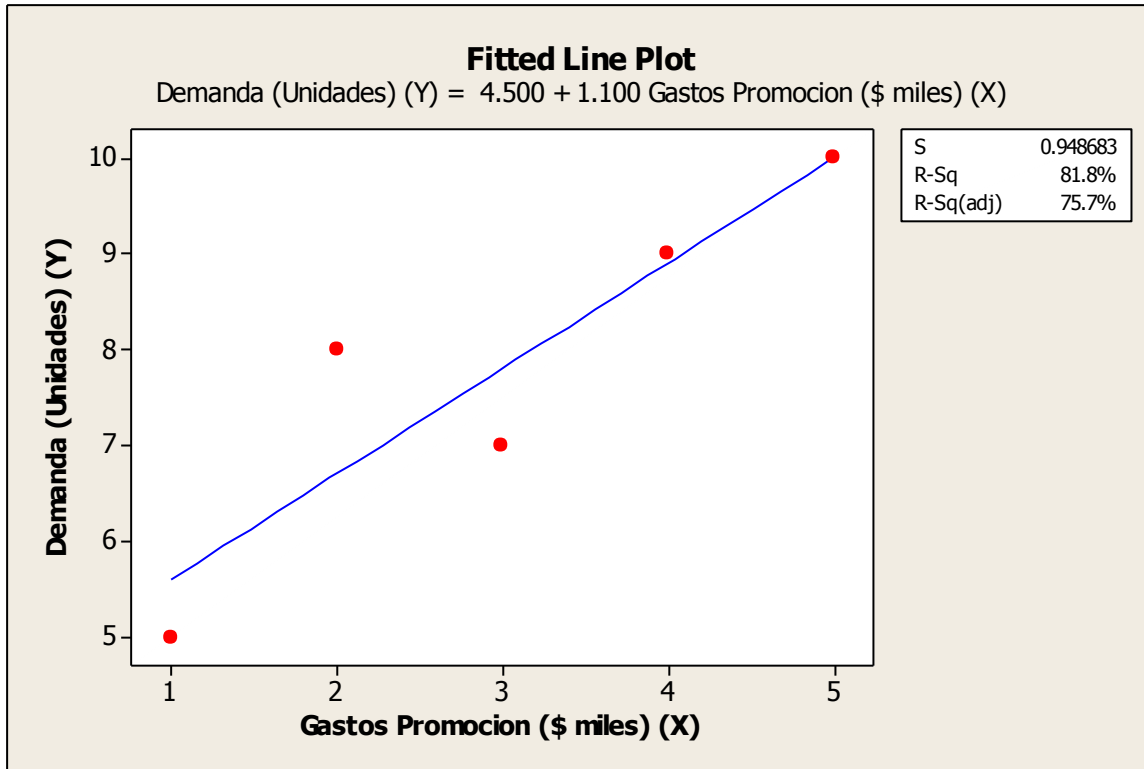
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Gastos Promocion (\$ miles) (X)	Demanda (Unidades) (Y)								
1	1	5								
2	2	8								
3	3	7								
4	4	9								
5	5	10								

The Fitted Line Plot dialog box is configured as follows:

- Response (Y): a (Unidades) (Y)
- Predictor (X): Gastos Promocion
- Type of Regression Model:  Linear,  Quadratic,  Cubic



Eso le debe dar un gráfico parecido al siguiente:



Y debe aparecer una información en el “Session Window” como la siguiente, en donde encuentras, la ecuación de regresión y la desviación estándar. El coeficiente de correlación de Pearson, en Minitab está dado por el valor del **Error** debajo de la columna de **MS**.

También puede calcularse buscando la raíz cuadrada del coeficiente de determinación dado por **R-Sq** que en este caso es de 81.8%

The regression equation is  
 Demanda (Unidades) (Y) = 4.500 + 1.100 Gastos Promocion (\$ miles) (X)

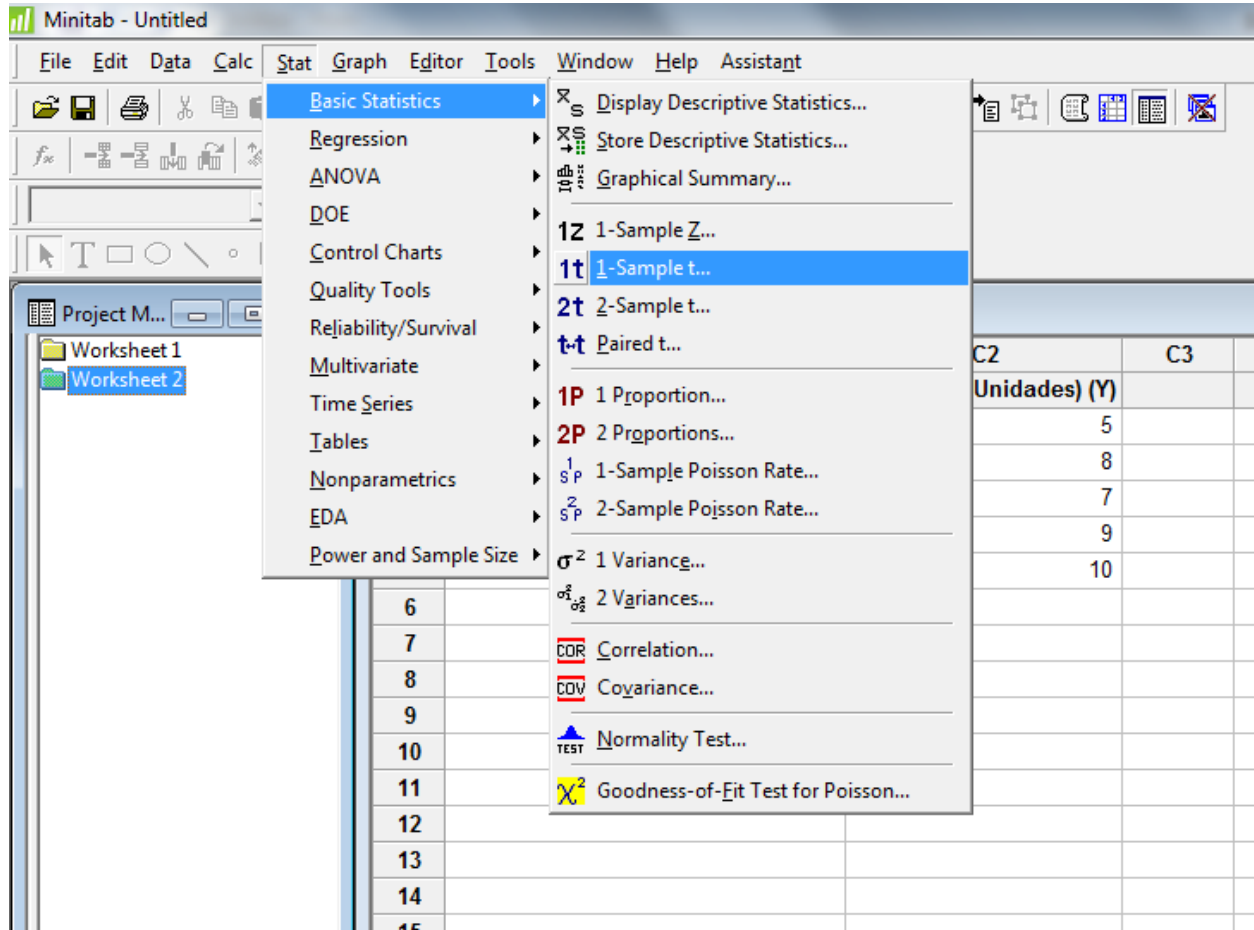
S = 0.948683    R-Sq = 81.8%    R-Sq(adj) = 75.7%

Analysis of Variance

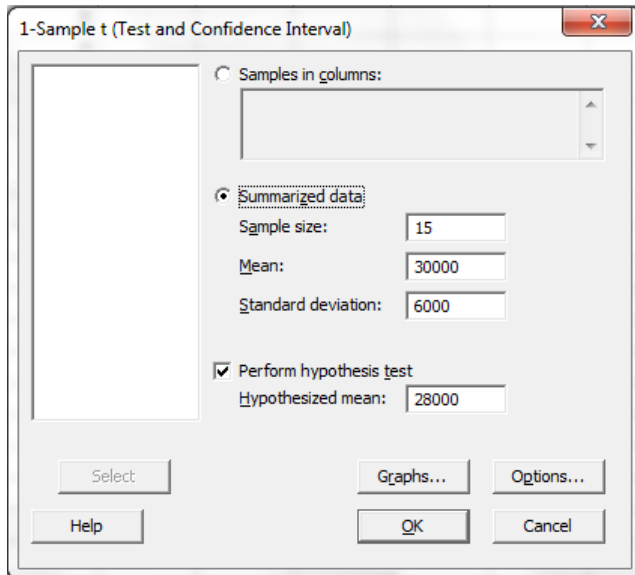
Source	DF	SS	MS	F	P
Regression	1	12.1	12.1	13.44	0.035
Error	3	2.7	0.9		
Total	4	14.8			

## Prueba t

La prueba t se puede realizar utilizando Minitab con la secuencia **Stat>Basic Statistics>1-sample t**

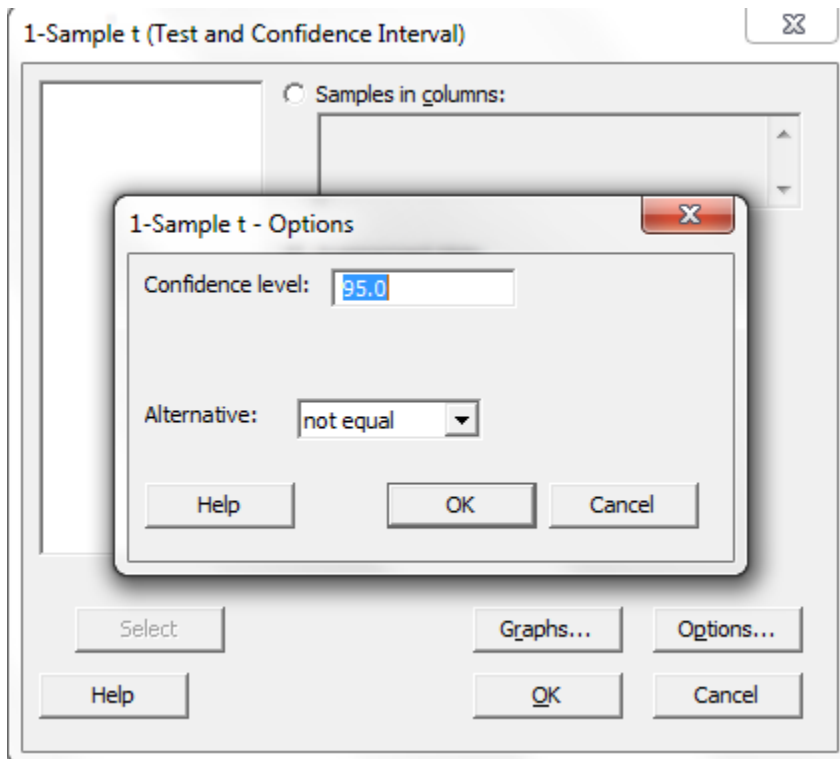


Con el ejemplo presentado, llene la información en la sección de **Summarized Data**



Seleccione la hipótesis que quiere verificar.

En el botón que dice **Options...** aparece la información del intervalo de confianza a seleccionar (el “default” es 95%) y el tipo de hipótesis que quiere realizar (el “default” es **not equal**)



En el "Session Window" va a tener un resultado como el siguiente:

### One-Sample T

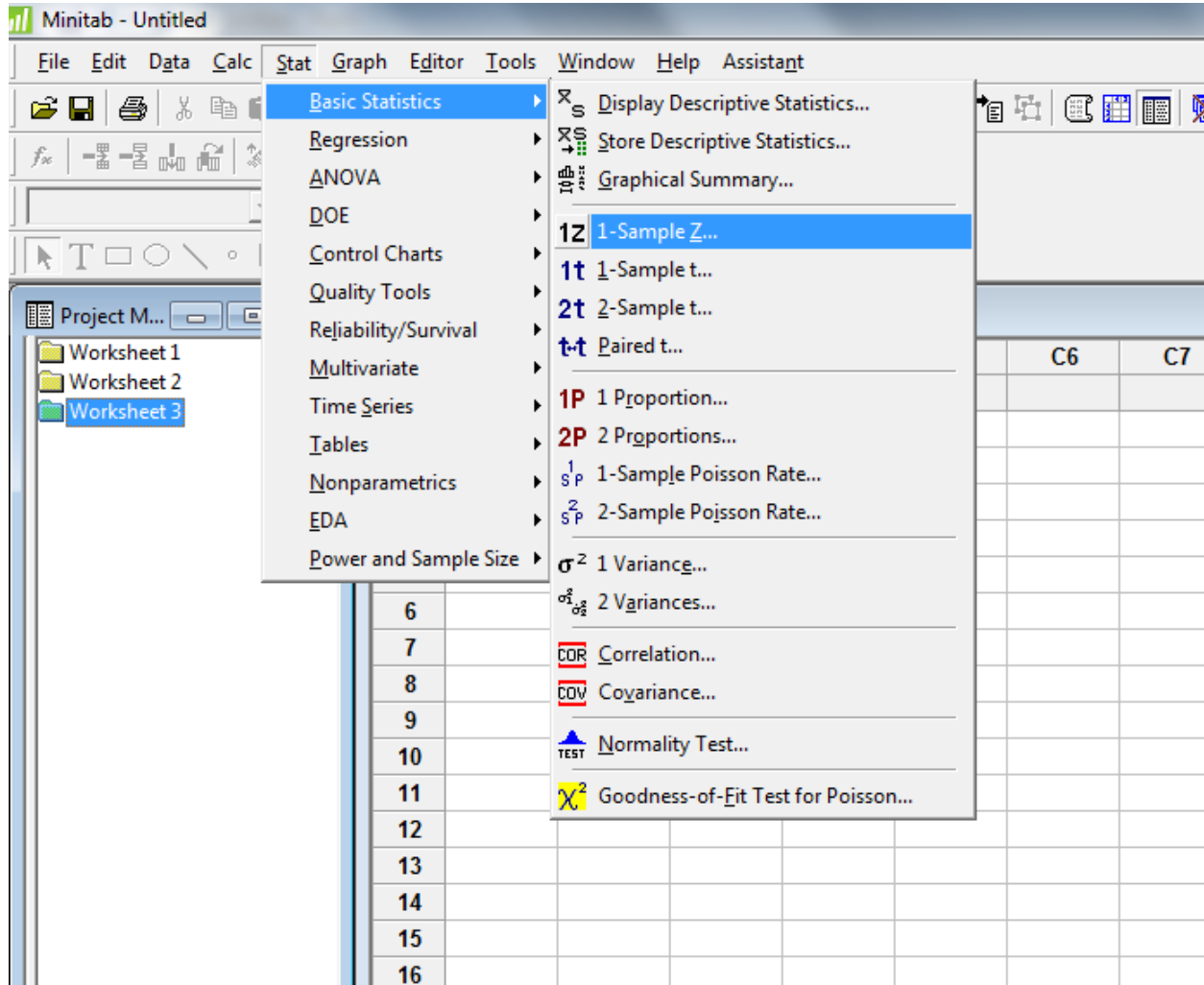
Test of mu = 28000 vs not = 28000

N	Mean	StDev	SE Mean	95% CI	T	P
15	30000	6000	1549	(26677, 33323)	1.29	0.218

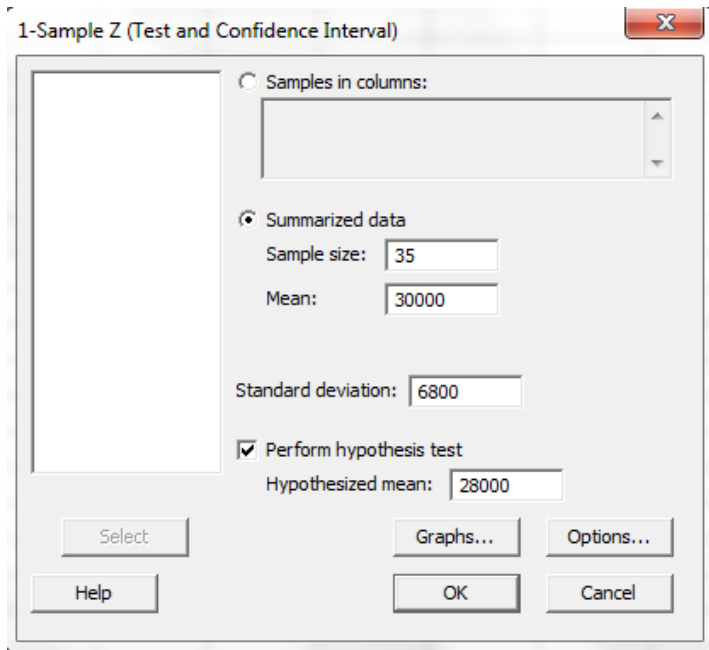
Donde la t estadística es 1.29 y el P-value o la probabilidad es de 0.218

## Prueba Z

La prueba t se puede realizar utilizando Minitab con la secuencia **Stat>Basic Statistics>1-sample Z**

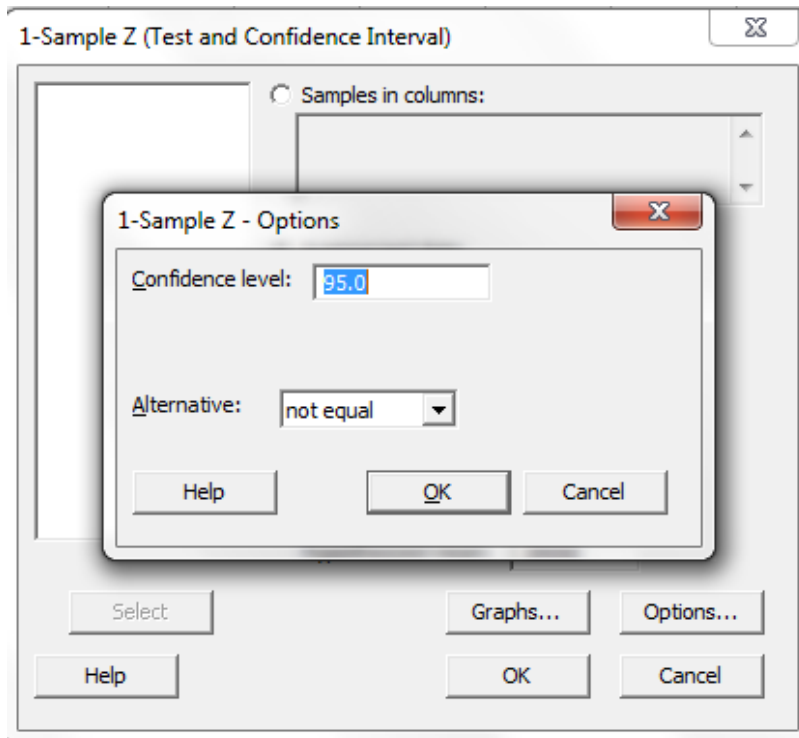


Con el ejemplo presentado, llene la información en la sección de **Summarized Data**



Seleccione la hipótesis que quiere verificar.

En el botón que dice **Options...** aparece la información del intervalo de confianza a seleccionar (el "default" es 95%) y el tipo de hipótesis que quiere realizar (el "default" es **not equal**)



En el "Session Window" va a tener un resultado como este:

### One-Sample Z

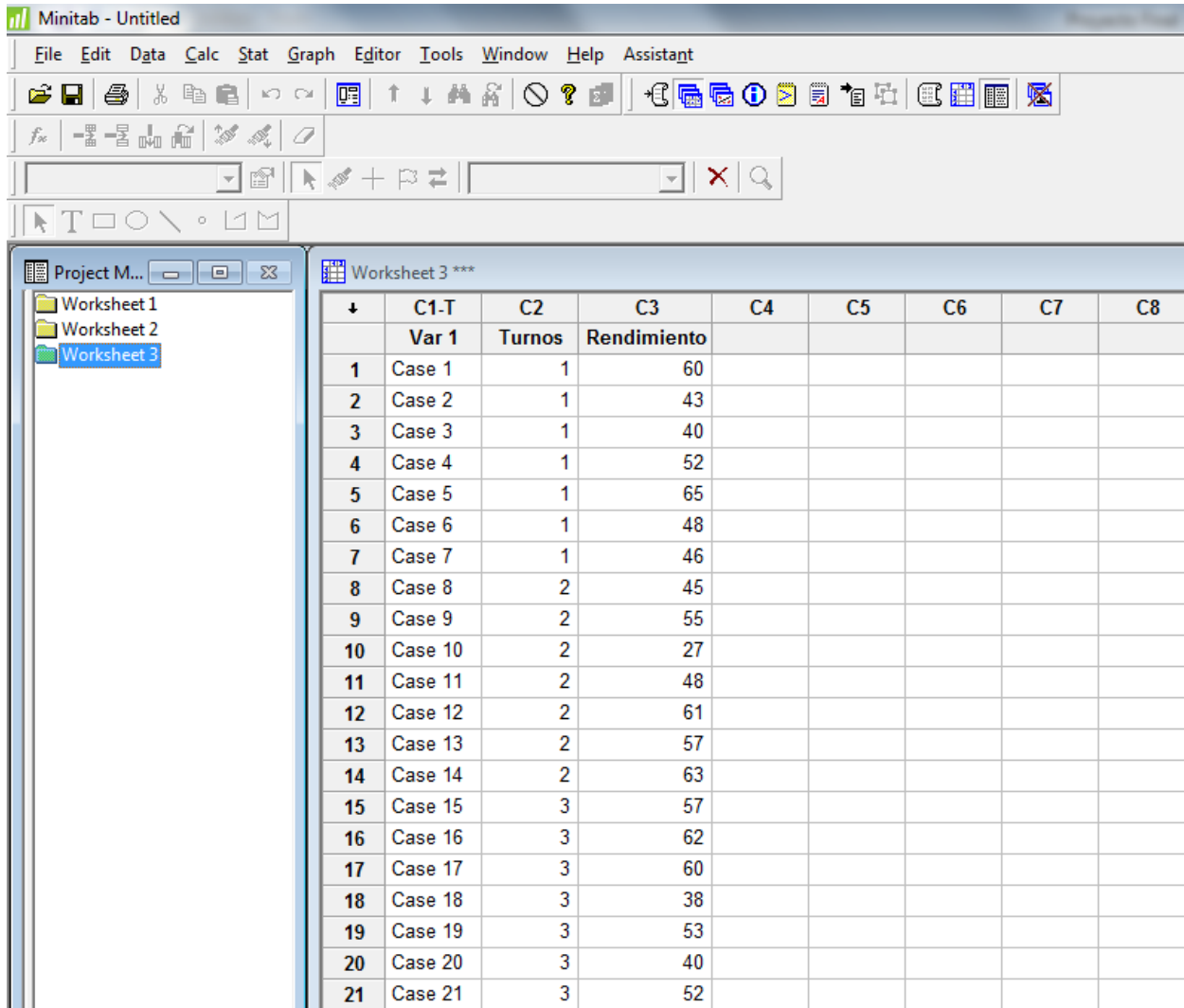
Test of mu = 28000 vs not = 28000  
The assumed standard deviation = 6800

N	Mean	SE Mean	95% CI	Z	P
35	30000	1149	(27747, 32253)	1.74	0.082

Donde la Z estadística es 1.74 y el valor P es de 0.082

## Análisis de Varianza

Utilizando Minitab se definen dos variables en cada columna, una para los turnos como variable independiente y la otra para el rendimiento como variable dependiente. Los turnos se identifican con los números del 1 al 3.



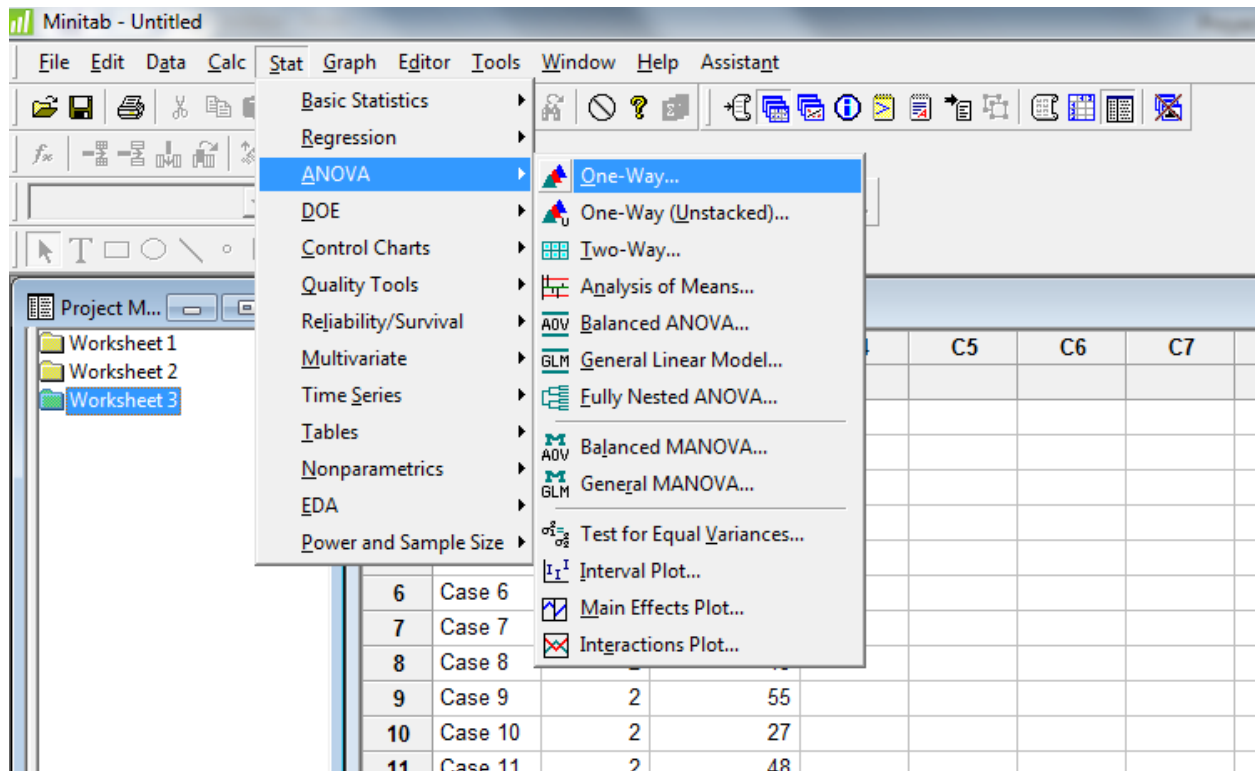
The screenshot shows the Minitab software interface. The main window displays a worksheet titled 'Worksheet 3 \*\*\*' with the following data:

	C1-T	C2	C3	C4	C5	C6	C7	C8
	Var 1	Turnos	Rendimiento					
1	Case 1	1	60					
2	Case 2	1	43					
3	Case 3	1	40					
4	Case 4	1	52					
5	Case 5	1	65					
6	Case 6	1	48					
7	Case 7	1	46					
8	Case 8	2	45					
9	Case 9	2	55					
10	Case 10	2	27					
11	Case 11	2	48					
12	Case 12	2	61					
13	Case 13	2	57					
14	Case 14	2	63					
15	Case 15	3	57					
16	Case 16	3	62					
17	Case 17	3	60					
18	Case 18	3	38					
19	Case 19	3	53					
20	Case 20	3	40					
21	Case 21	3	52					



Luego se debe utilizar la siguiente secuencia:

**Stat>ANOVA>One Way...**

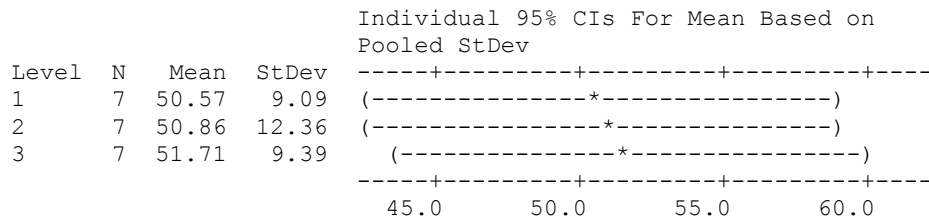


El resultado en Minitab es el siguiente:

### One-way ANOVA: Rendimiento versus Turnos

Source	DF	SS	MS	F	P
Turnos	2	5	2	0.02	0.977
Error	18	1942	108		
Total	20	1947			

S = 10.39    R-Sq = 0.25%    R-Sq(adj) = 0.00%

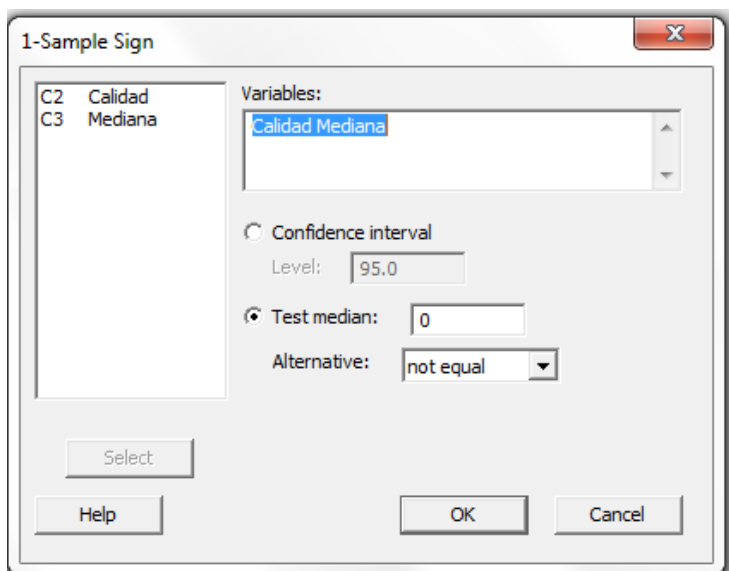
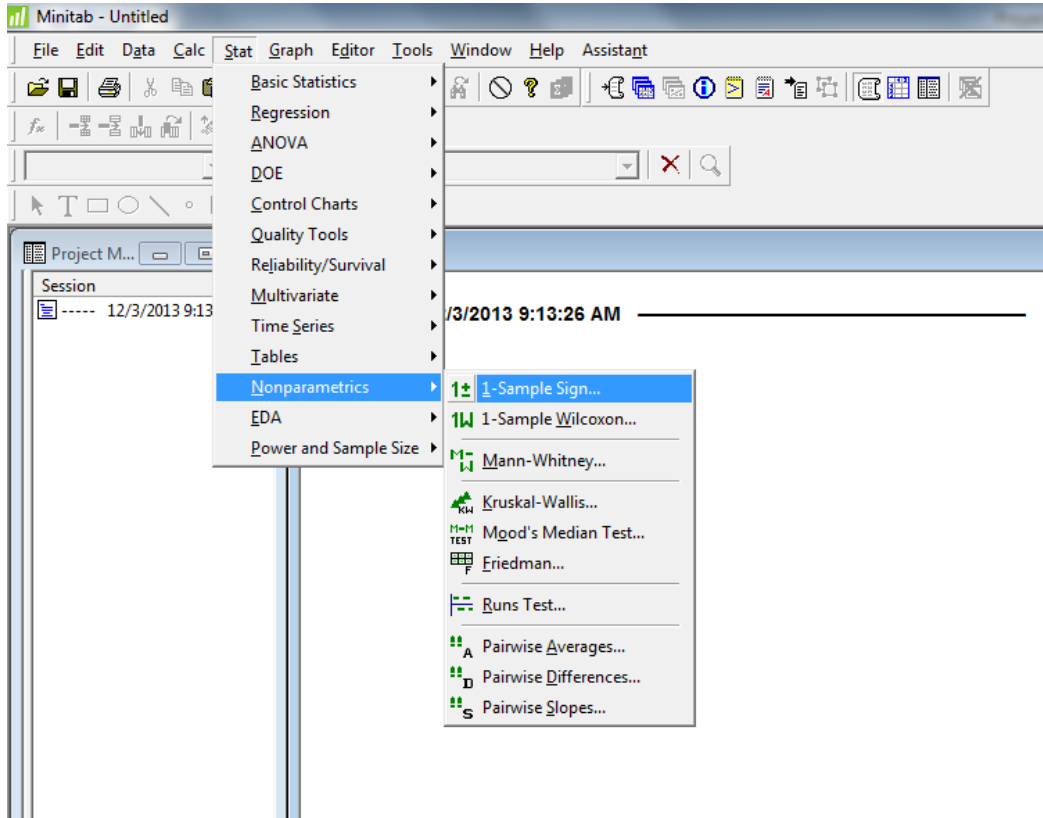


Pooled StDev = 10.39

## Prueba del Signo

Para hacer la prueba del signo, en Minitab se accede a través de

**Stat>Nonparametrics>1-Sample Sign**



El resultado aparece en el “Session Window” con la siguiente información:

**Sign Test for Median: Calidad, Mediana**

Sign test of median = 0.00000 versus not = 0.00000

	N	Below	Equal	Above	P	Median
Calidad	7	0	0	7	0.0156	5.000
Mediana	7	0	0	7	0.0156	4.000

La prueba se repite con “test median” para los valores 1, 2 y 3. Luego los resultados se suman obteniendo los siguiente:

Probabilidad de  $(k \leq 2) = 0.01562 + 0.09375 + 0.23436 = 0.3438$ .

## ***Ji-Cuadrado o $X^2$***

Entrando los datos del ejemplo del libro, se trabaja la tabla en Minitab de la siguiente manera:

Minitab - Untitled

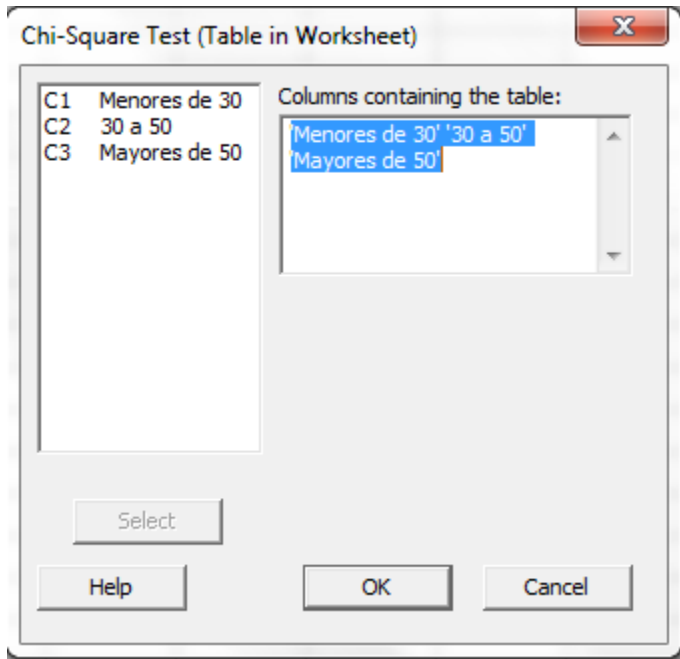
File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Project M... Worksheet 1 \*\*\*

	C1	C2	C3	C4	C5	C6	C7
	Menores de 30	30 a 50	Mayores de 50				
1	200	100	100				
2	600	900	400				
3	400	600	500				
4	700	500	0				
5							
6							
7							
8							
9							
10							
11							

Para el análisis, se utiliza la siguiente secuencia: ***Stat>Tables>Chi-Square Test (Two-Way Table in Worksheet)...***

Y selecciona las columnas con la información como el sub menú a continuación:



El resultado aparece en el "Session Window":

### Chi-Square Test: Menores de 30, 30 a 50, Mayores de 50

Expected counts are printed below observed counts  
Chi-Square contributions are printed below expected counts

	Menores de 30	30 a 50	Mayores de 50	Total
1	200	100	100	400
	152.00	168.00	80.00	
	15.158	27.524	5.000	
2	600	900	400	1900
	722.00	798.00	380.00	
	20.615	13.038	1.053	
3	400	600	500	1500
	570.00	630.00	300.00	
	50.702	1.429	133.333	
4	700	500	0	1200
	456.00	504.00	240.00	
	130.561	0.032	240.000	
Total	1900	2100	1000	5000

Chi-Sq = 638.444, DF = 6, P-Value = 0.000

## Coeficiente de correlación de Spearman ( $r_s$ )

El coeficiente de correlación de Spearman se calcula con datos que estén ordenados (“**Ranked**”) y luego se hace una correlación de Pearson. En el ejemplo del libro, los datos están ya organizados (“ranked”). Para resolverlo con Minitab se debe seguir la secuencia: **Data>Rank** y luego se debe preparar unas columnas adicionales para colocar los datos ordenados (“ranked”).

The screenshot shows the Minitab interface with a worksheet containing data for tennis brands and two surveys. A 'Rank' dialog box is open, showing the selection of columns for ranking.

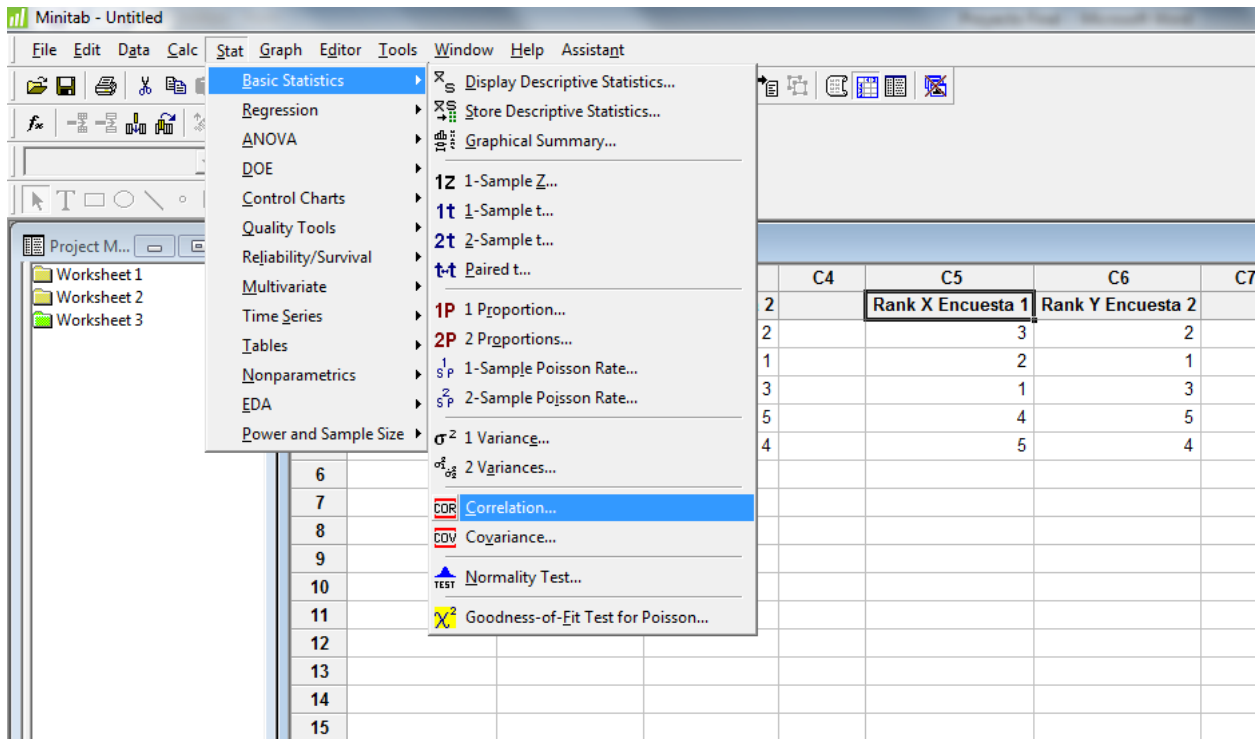
	C1-T	C2	C3	C4	C5	C6	C7	C8
	Marcas de Tenis	X = Encuesta 1	Y = Encuesta 2		Rank X Encuesta 1	Rank Y Encuesta 2		
1	Converse	3	2					
2	Adidas	2	1					
3	Nike	1	3					
4	Puma	4	5					
5	New Balance	5	4					
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								

The 'Rank' dialog box shows the following configuration:

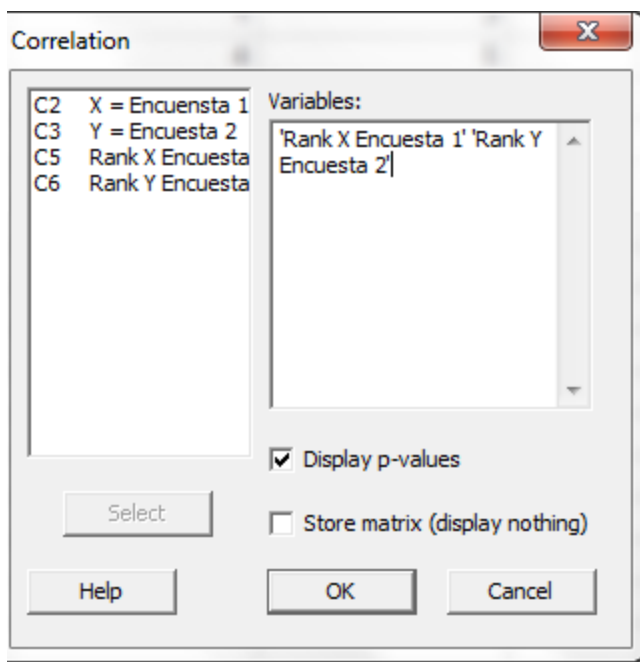
- Rank data in: Y = Encuesta 2
- Store ranks in: Rank Y Encuesta

Luego, con los datos organizados se puede realizar la correlación con la siguiente

secuencia: **Stat>Basic Statistics>Correlation**



Y en el submenú seleccione las columnas de los datos “ranked” y presione **OK**:





Los resultados en el "Session Window":

**Correlations: Rank X Encuesta 1, Rank Y Encuesta 2**

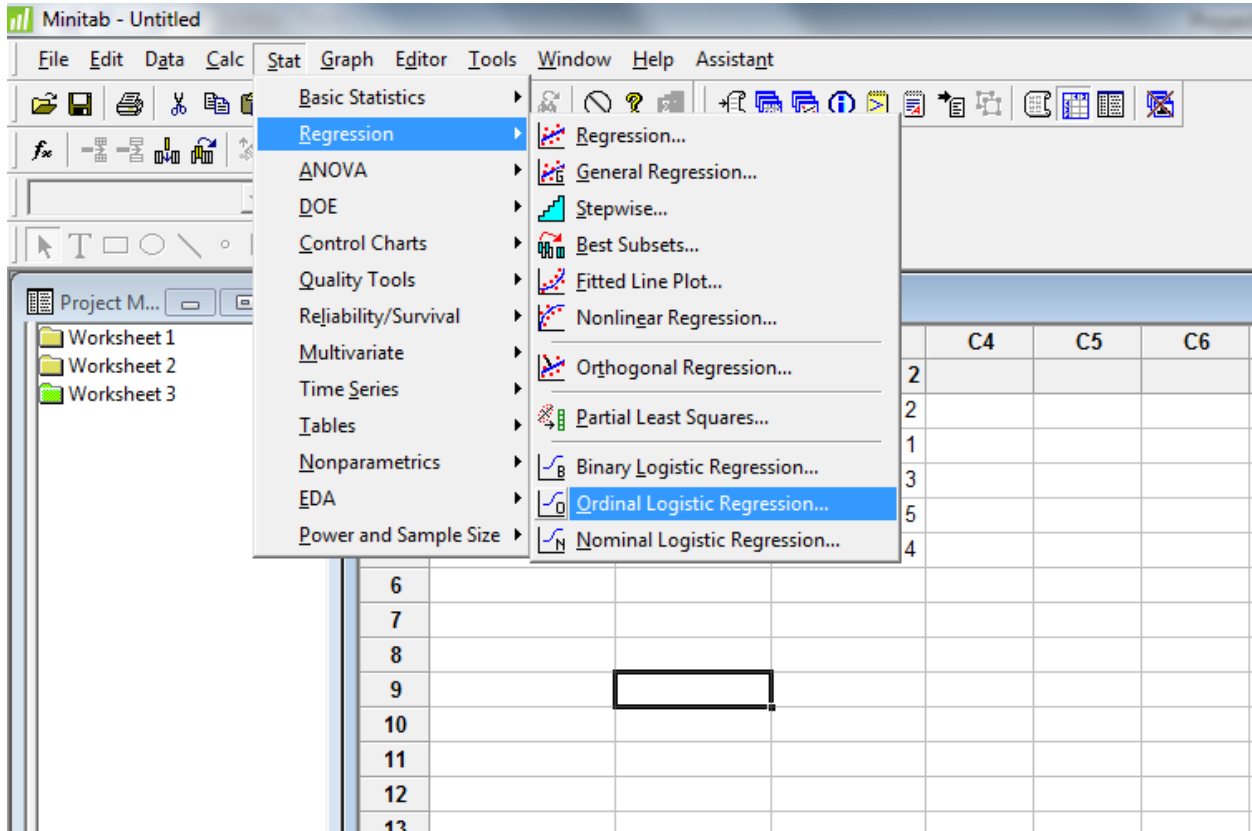
Pearson correlation of Rank X Encuesta 1 and Rank Y Encuesta 2 = 0.600  
P-Value = 0.285

Aunque dice "Pearson", al usar los datos "ranked", en Minitab se interpreta como la correlación de Spearman.

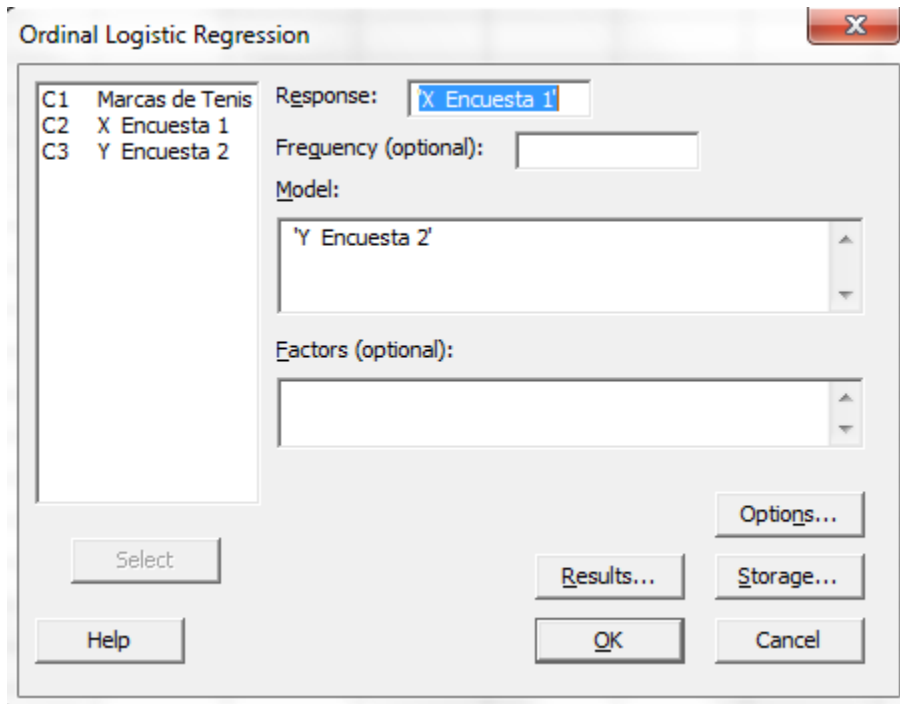
## Prueba de Kendall Tau ( $\tau$ )

Para la prueba Kendall Tau se utiliza la siguiente secuencia:

**Stat>Regression>Ordinal Logistic Regression**



Luego se compara una encuesta contra la otra. En el espacio para "Response" se coloca la "X" y en el "Model" se coloca la "Y" y luego se presiona **OK**:



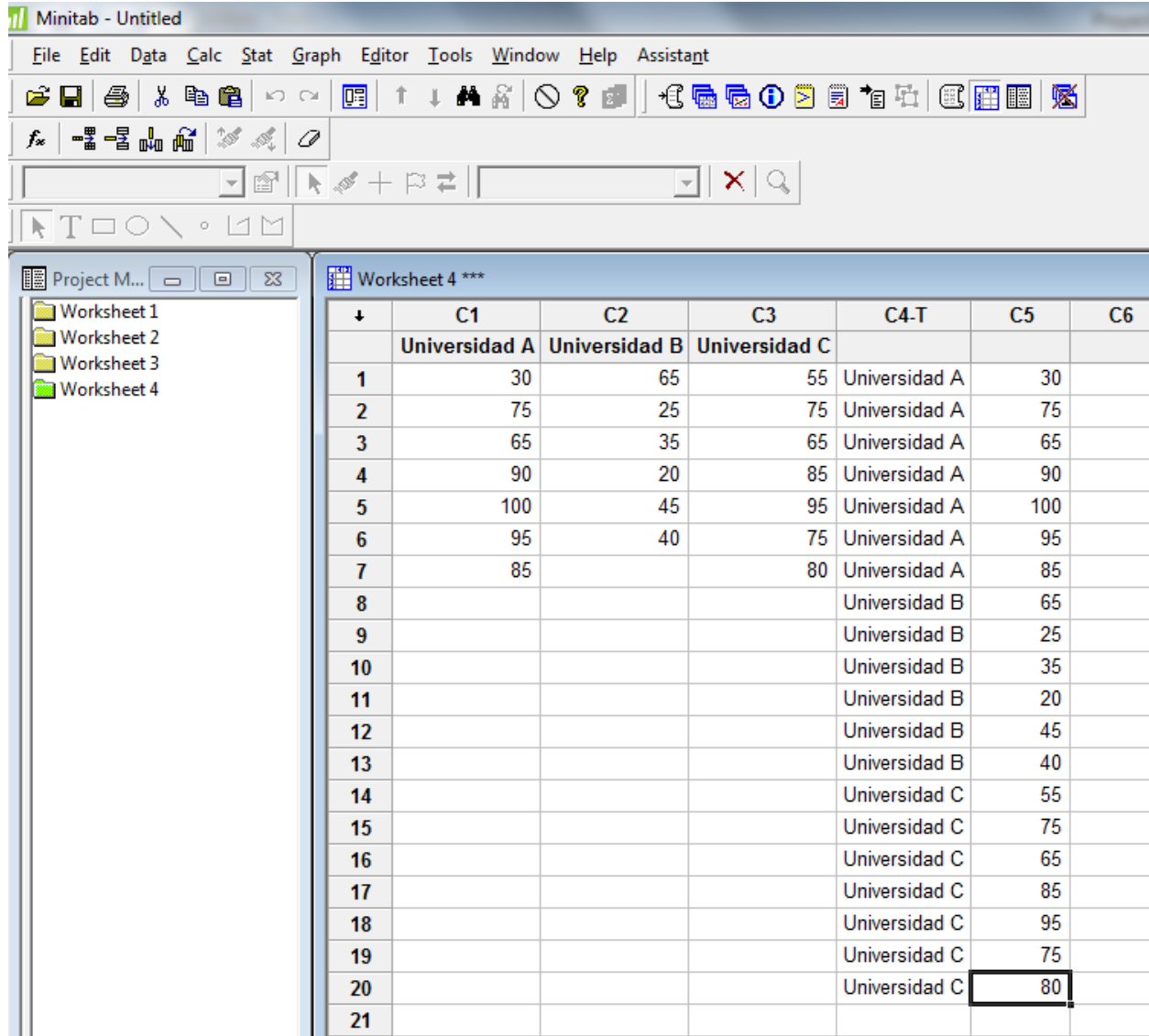
En el "Session Window" tendrá el siguiente resultado:

Measures of Association:  
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	7	70.0	Somers' D 0.40
Discordant	3	30.0	Goodman-Kruskal Gamma 0.40
Ties	0	0.0	Kendall's Tau-a 0.40
Total	10	100.0	

## Prueba Kruskal-Wallis

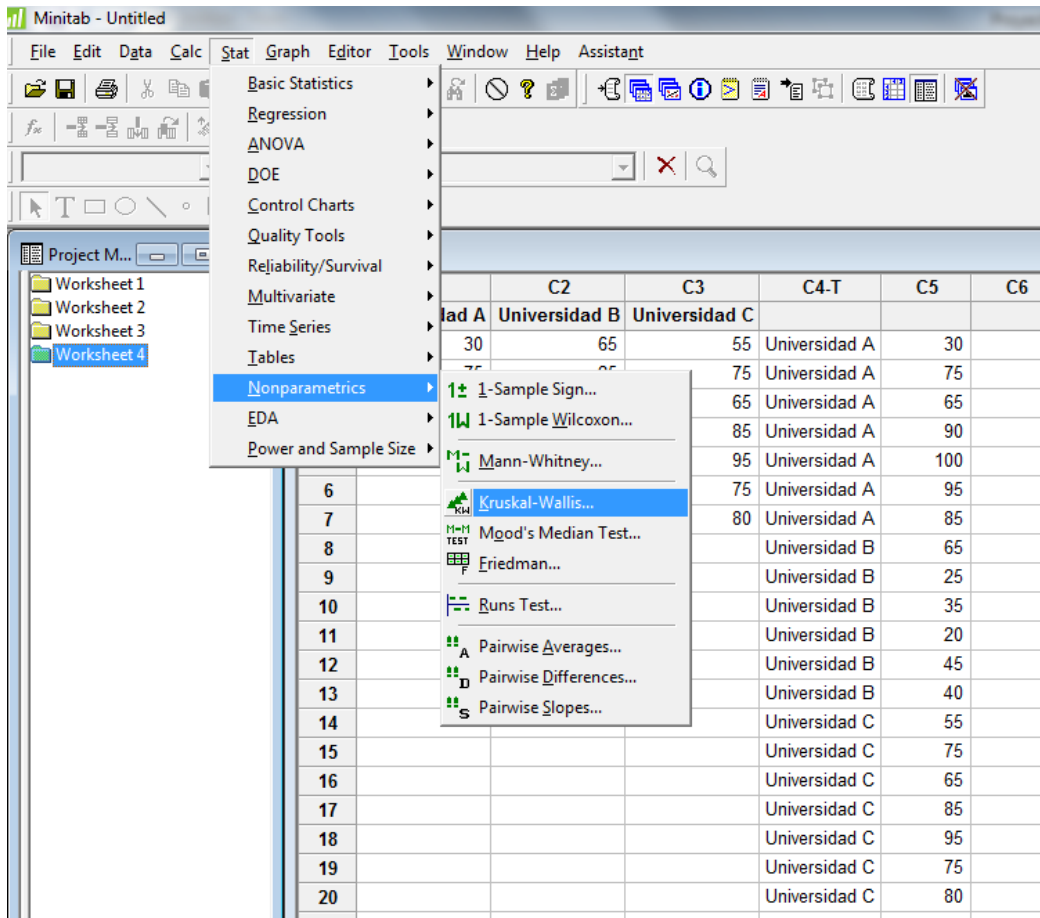
Para esta prueba se deben organizar los datos en dos columnas



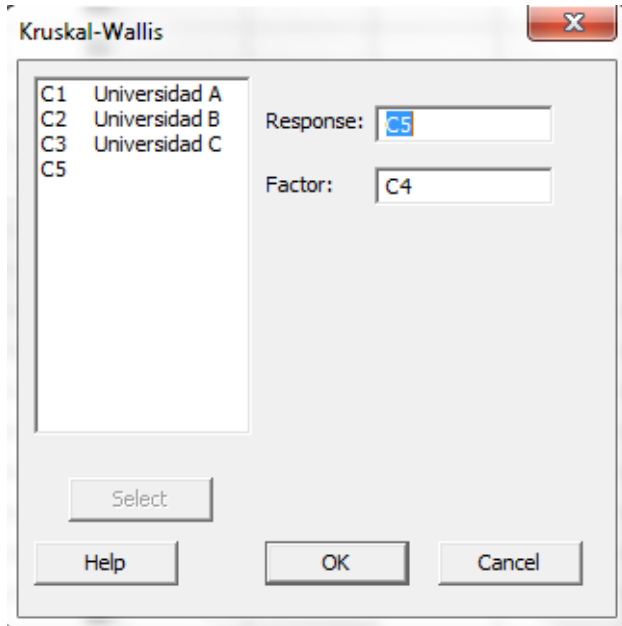
The screenshot shows the Minitab software interface with a worksheet titled 'Worksheet 4 \*\*\*'. The worksheet contains data for a Kruskal-Wallis test. The data is organized into columns: C1 (Universidad A), C2 (Universidad B), C3 (Universidad C), C4-T (Universidad A, B, C), C5, and C6. The data is as follows:

	C1	C2	C3	C4-T	C5	C6
	Universidad A	Universidad B	Universidad C			
1	30	65	55	Universidad A	30	
2	75	25	75	Universidad A	75	
3	65	35	65	Universidad A	65	
4	90	20	85	Universidad A	90	
5	100	45	95	Universidad A	100	
6	95	40	75	Universidad A	95	
7	85		80	Universidad A	85	
8				Universidad B	65	
9				Universidad B	25	
10				Universidad B	35	
11				Universidad B	20	
12				Universidad B	45	
13				Universidad B	40	
14				Universidad C	55	
15				Universidad C	75	
16				Universidad C	65	
17				Universidad C	85	
18				Universidad C	95	
19				Universidad C	75	
20				Universidad C	80	
21						

Luego, se utiliza la siguiente secuencia: **Stat>Nonparametrics>Kruskal Wallis**



En el submenú:



El resultado en el "Session Window" es el siguiente:

#### **Kruskal-Wallis Test: C5 versus C4**

Kruskal-Wallis Test on C5

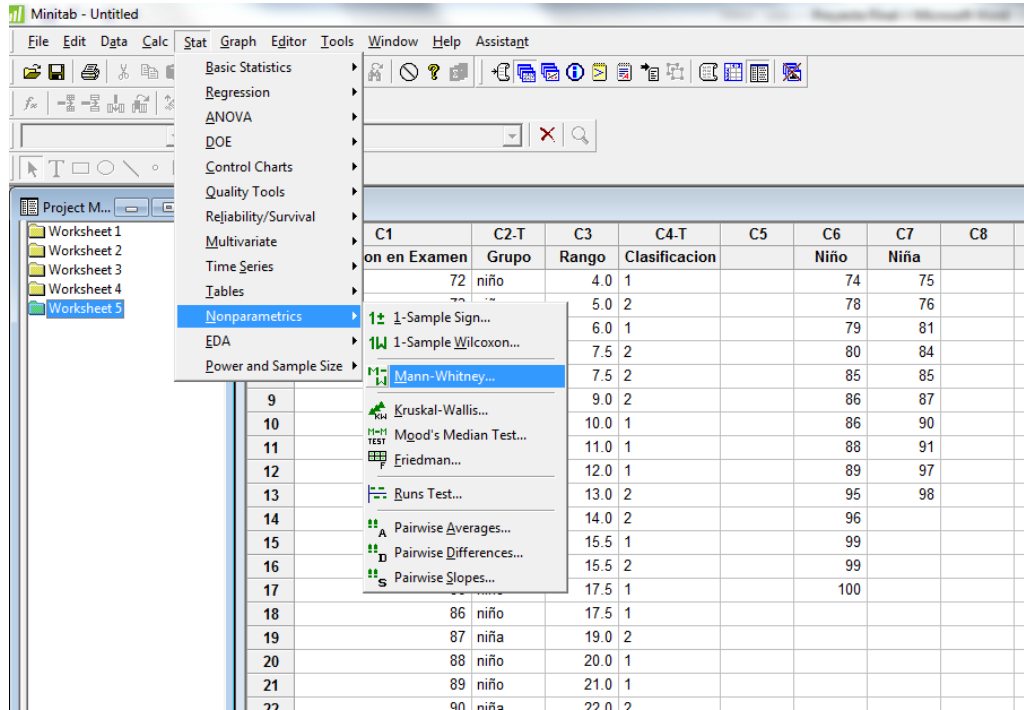
C4	N	Median	Ave Rank	Z
Universidad A	7	85.00	13.6	1.70
Universidad B	6	37.50	4.5	-2.97
Universidad C	7	75.00	12.6	1.15
Overall	20		10.5	

H = 8.92 DF = 2 P = 0.012

H = 8.98 DF = 2 P = 0.011 (adjusted for ties)

## Prueba de Mann-Whitney

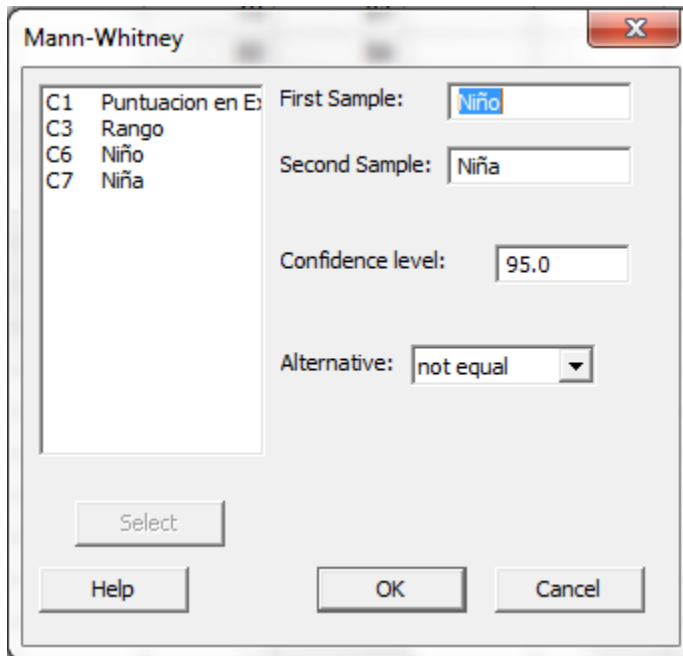
Primero se asignan los datos en dos columnas, una para niño y otra para niña. Luego se debe seguir la secuencia:



The screenshot shows the Minitab software interface with the 'Stat' menu open and 'Mann-Whitney...' selected. The data table is as follows:

C1	C2-T	C3	C4-T	C5	C6	C7	C8
Calificación	Grupo	Rango	Clasificación		Niño	Niña	
72	niño	4.0	1		74	75	
73	niño	5.0	2		78	76	
74	niño	6.0	1		79	81	
75	niño	7.5	2		80	84	
76	niño	7.5	2		85	85	
77	niño	9.0	2		86	87	
78	niño	10.0	1		86	90	
79	niño	11.0	1		88	91	
80	niño	12.0	1		89	97	
81	niño	13.0	2		95	98	
82	niño	14.0	2		96		
83	niño	15.5	1		99		
84	niño	15.5	2		99		
85	niño	17.5	1		100		
86	niño	17.5	1				
87	niña	19.0	2				
88	niño	20.0	1				
89	niño	21.0	1				
90	niña	22.0	2				

Y en el submenú, se coloca la información que se demuestra y luego se presiona **OK**:



El resultado en el “Session Window”:

### **Mann-Whitney Test and CI: Niño, Niña**

	N	Median
Niño	17	86.00
Niña	13	84.00

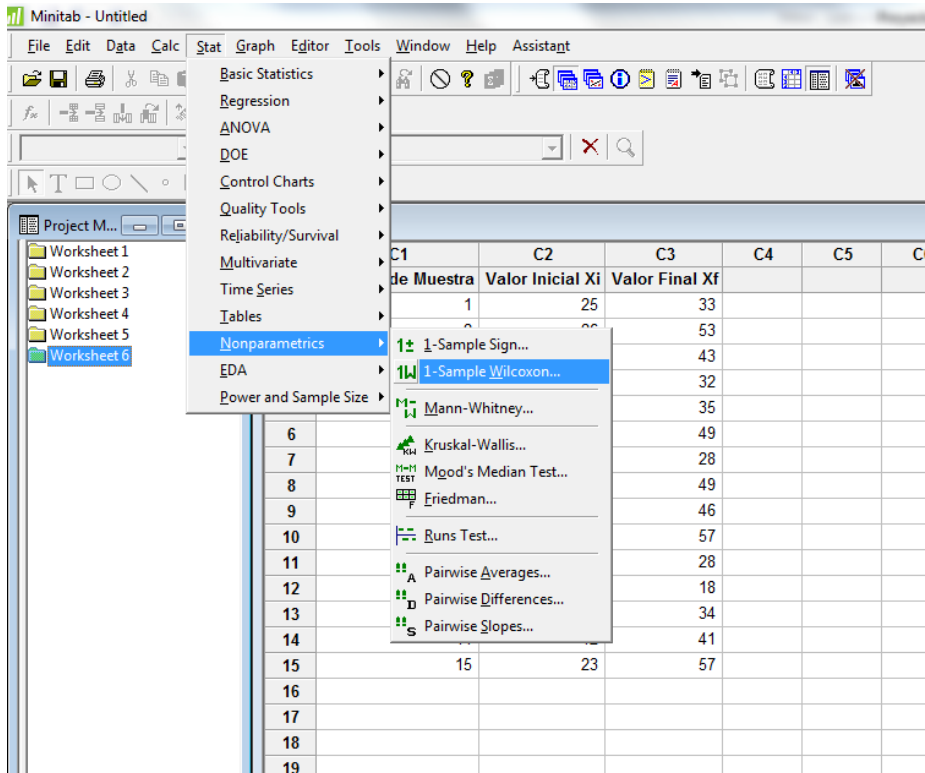
Point estimate for ETA1-ETA2 is 2.00  
95.1 Percent CI for ETA1-ETA2 is (-6.00,10.00)  
W = 273.5  
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.6909  
The test is significant at 0.6908 (adjusted for ties)



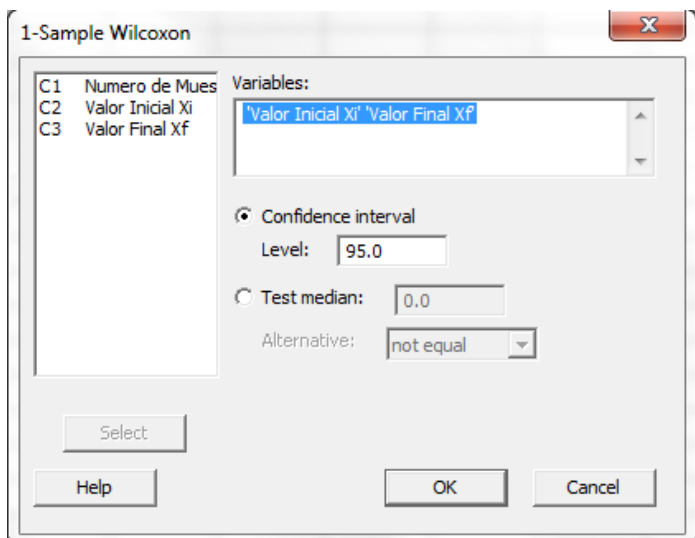
## Wilcoxon signed-rank

Utilizando los datos del ejemplo en el libro, se utiliza la secuencia:

**Stat>Nonparametrics>1-Sample Wilcoxon**



Y en el submenú:



El resultado en el "Session Window":

**Wilcoxon Signed Rank CI: Valor Inicial Xi, Valor Final Xf**

		Estimated	Achieved	Confidence	
	N	Median	Confidence	Lower	Upper
Valor Inicial Xi	15	34.8	95.0	25.0	42.5
Valor Final Xf	15	40.5	95.0	33.5	47.0

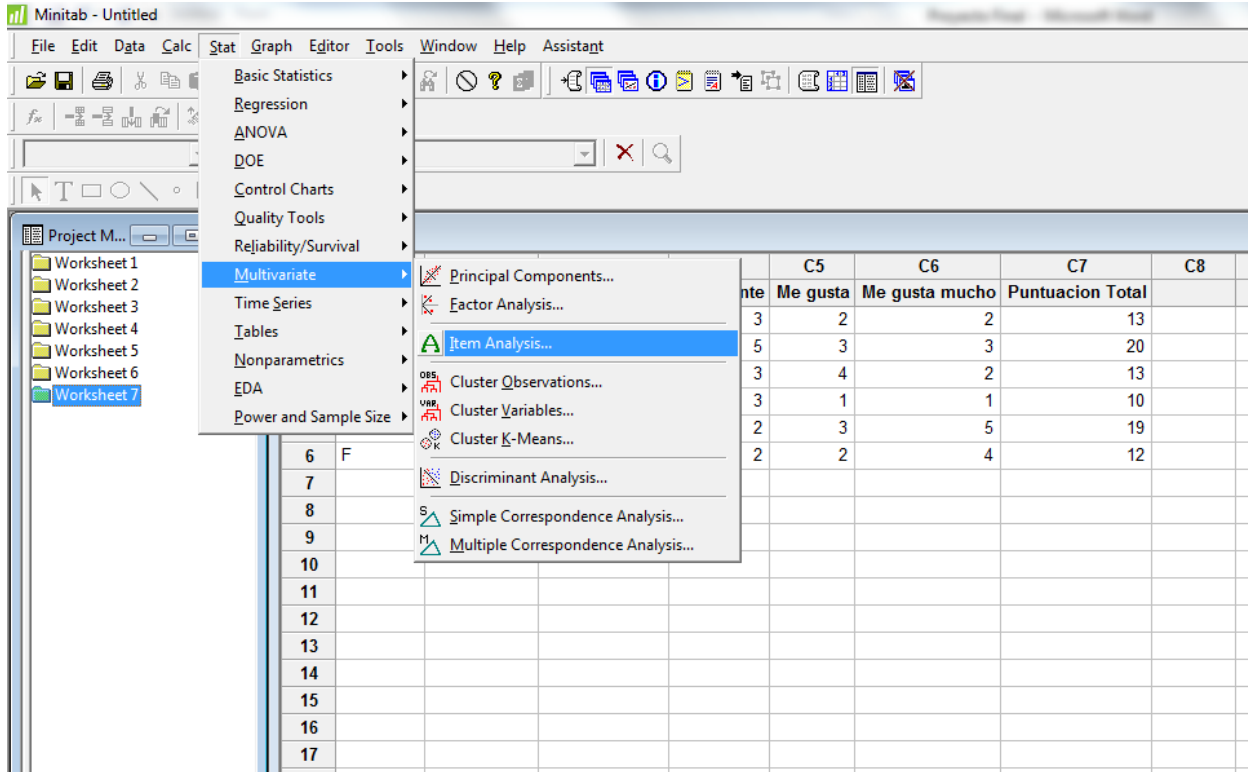
**Wilcoxon Signed Rank Test: Valor Inicial Xi, Valor Final Xf**

Test of median = 0.000000 versus median not = 0.000000

	N for	Wilcoxon		Estimated	
	N	Test	Statistic	P	Median
Valor Inicial Xi	15	15	120.0	0.001	34.75
Valor Final Xf	15	15	120.0	0.001	40.50

## Coeficiente Alfa de Cronbach

Se debe utilizar la siguiente secuencia: **Stat>Multivariate>Item Analysis**



En el submenú, se seleccionan todas las columnas y luego se presiona **OK**. En el “Session Window” el resultado es:

Omitted Variable	Adj. Total	Adj. Total	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
No me gusta	26.000	6.899	0.711057	*	0.668768
Me gusta poco	25.833	7.250	0.664628	*	0.699112
Indiferente	26.000	7.694	0.284747	*	0.757320
Me gusta	26.500	7.556	0.441625	*	0.737011
Me gusta mucho	26.167	7.250	0.490374	*	0.718136
Puntuacion Total	14.500	4.037	1.000000	*	0.618609

Colaboración de:

Ing. Juan Lugo Rivera                      diciembre 2013

Revisión de:

Dr. Walter López Moreno                      diciembre 2013