

Curso
Métodos Estadísticos
Guía MINITAB Módulo 2

Nombre del profesor: MSc. Sofía Arbuola Briceño
Correo electrónico: s.arbuola@gmail.com



ÍNDICE

TIPOS DE DATOS VRS TIPOS DE GRÁFICOS	3
GRÁFICO DE DISPERSIÓN, CORRELACIÓN O TENDENCIA	4
DIAGRAMA DE PARETO	8
GRÁFICO DE HISTOGRAMA	12
GRÁFICO DE CAJAS Y BIGOTES	16

Tipos de datos

Variables cuantitativas

Continuas

- Toma cualquier valor.
- Puede ser decimal.
- Ejemplo: temperatura, presión, altura, edad.

Discretas

- Toma valores enteros.
- Ejemplo: número de alumnos, cantidad de estudiantes.

Variables cualitativas

Discreta Nominal

- Categorías no ordenadas.
- Ejemplo: género, grupo sanguíneo, lugar de nacimiento.

Discreta Ordinal

- Categoría ordenada.
- Ejemplo: escolaridad, grado de satisfacción.

Tipos de gráficos

Continua-Continua	Discreta-Discreta	Continua-Discreta
<ul style="list-style-type: none">• Histograma.• Gráfico de dispersión (gráfico XY).• Caja y bigotes.	<ul style="list-style-type: none">• Gráfico de barras.• Diagrama de Pareto.• Gráfico de pastel.	<ul style="list-style-type: none">• Caja y bigotes.• Diagrama de Pareto.• Gráfico multivariable.

Gráfico de dispersión, correlación o tendencia

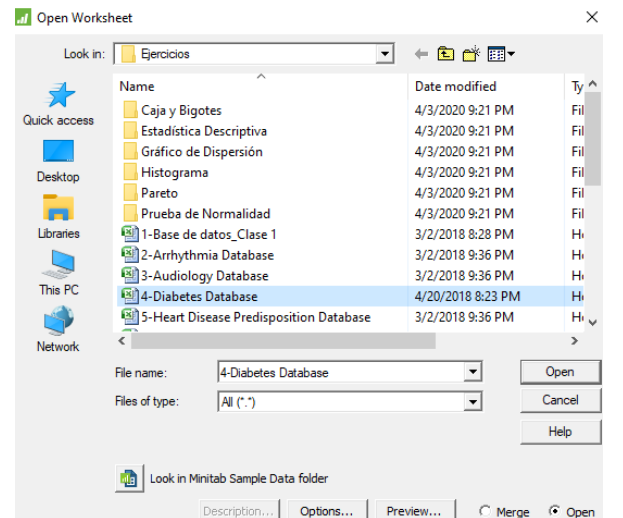
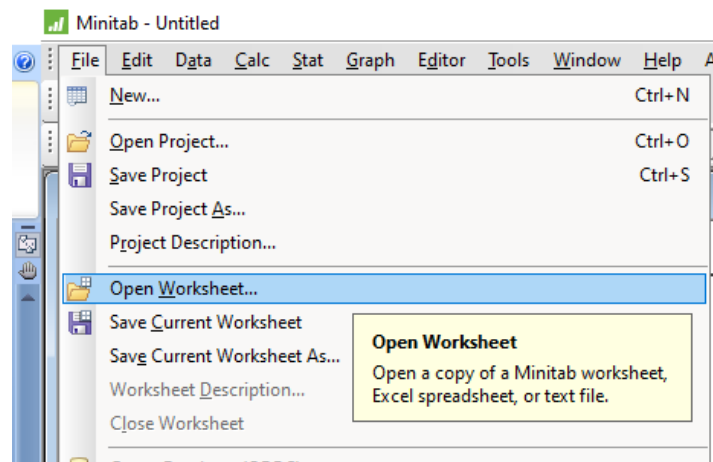
Este gráfico permite comprobar si existe relación entre dos variables de interés. En Minitab lo que se va a calcular es el Coeficiente de correlación (R), en este caso podemos calcular el valor usando el R-simple (coeficiente de correlación de Pearson) o R-cuad (R-cuadrado).

- **R simple:**

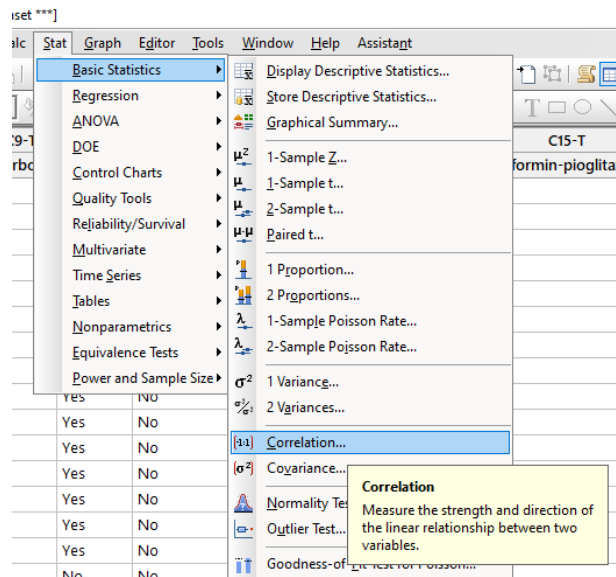
De Minitab se calcula el coeficiente de correlación de Pearson. Existe una correlación alta entre las variables cuantitativas si el valor es lo más cercano a 1 ó -1.

1-Cargamos la base de datos desde Excel

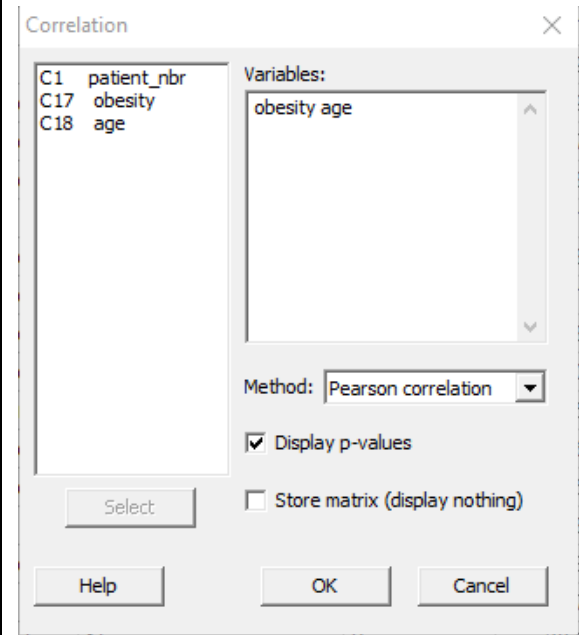
File->Open Worksheet->Seleccionamos "All"->
Cargamos la base de datos en Excel



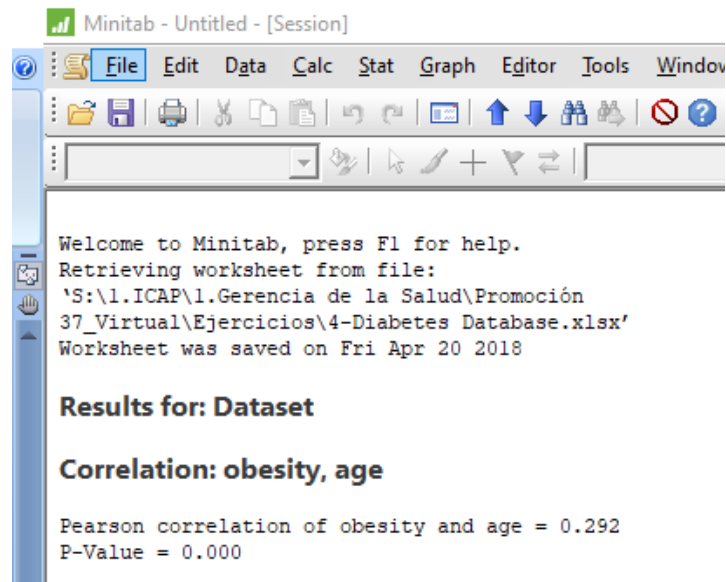
2- Estadísticas -> Estadísticas básicas -> Correlación



3-->Ingresar las variables cuantitativas al formulario



4- En la sesión se genera el valor del coeficiente de correlación de Pearson



▪ **R-cuad:**

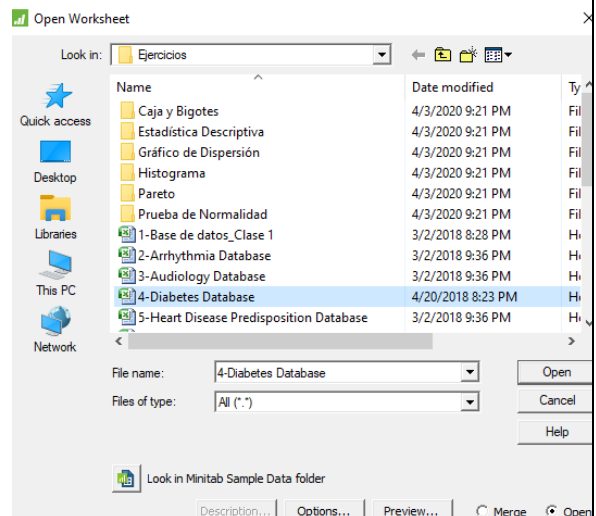
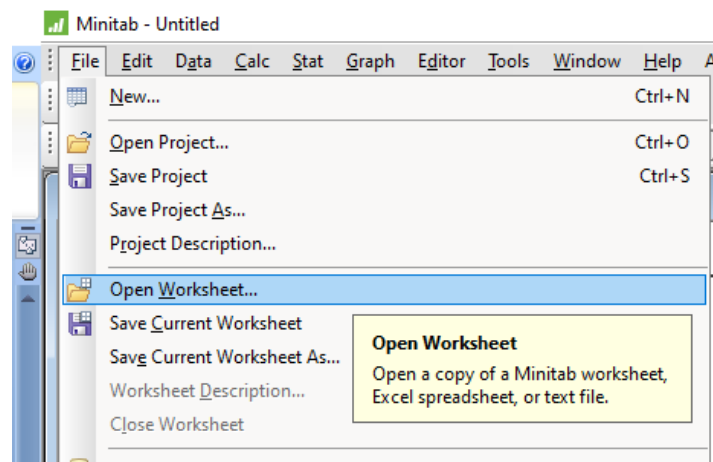
De Minitab podemos calcular el cuadrado del coeficiente de correlación de Pearson y obtener el gráfico de dispersión con la ecuación de la recta de mejor ajuste.

El R-cuad es el porcentaje de variación de la variable de respuesta que explica su relación con una o más variables predictoras. Por lo general, mientras mayor sea el R-cuad, mejor será el ajuste del modelo a sus datos.

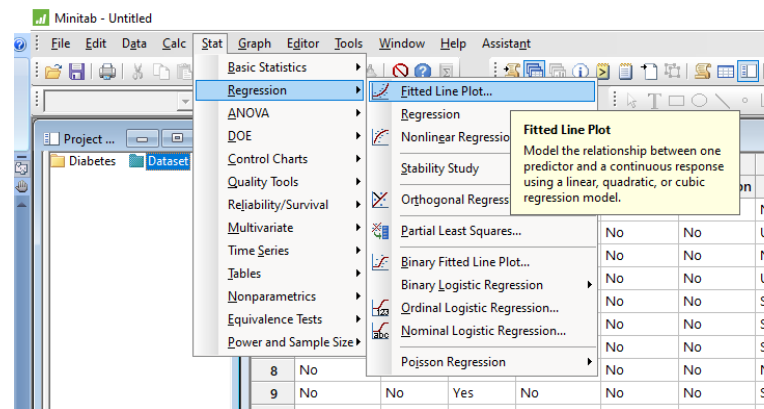
El R-cuad siempre se encuentra entre 0 y 100%. El R-cuadrado también se conoce como el coeficiente de determinación o determinación múltiple.

1-Cargamos la base de datos desde Excel

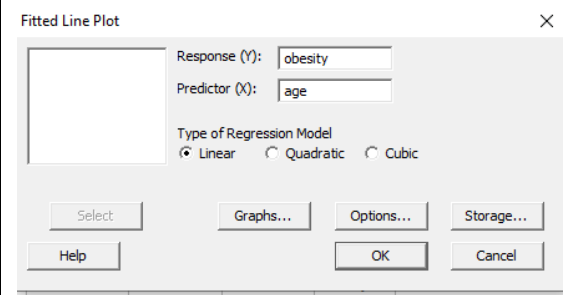
File->Open Worksheet->Seleccionamos "All" ->
Cargamos la base de datos en Excel



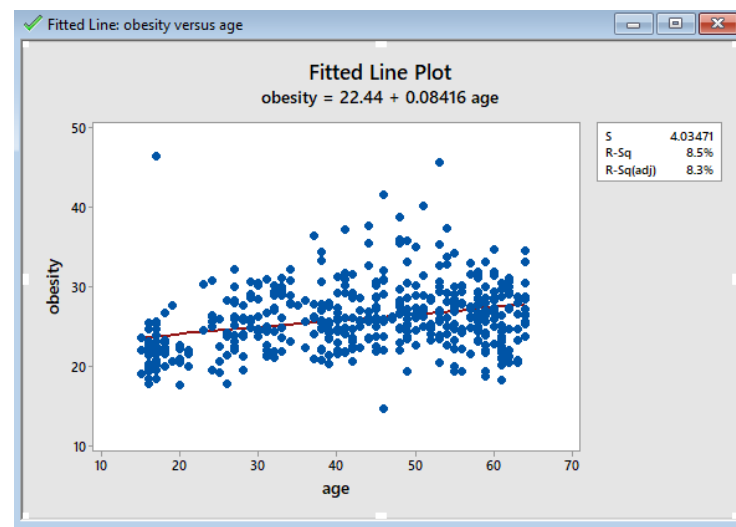
2- Estadísticas -> Regresión -> Gráfico con línea de mejor ajuste



3-->Ingresa las variables al formulario -> Modelo Lineal



4-Se obtiene el gráfico de dispersión con la ecuación de la recta de mejor ajuste, se debe revisar el valor obtenido de R-cuad o R-Sq.



Existen otros elementos a considerar como **error estándar de la regresión (S)**, el cual mide en las unidades de la variable de respuesta y representa la distancia estándar que separa a los valores de los datos de la línea de regresión ajustada. Mientras más bajo sea el valor de S, mejor predecirá la respuesta el modelo. Si compara diferentes modelos, el modelo con el menor valor de S indica el mejor ajuste. Por otra parte el **R-cuad (ajustado)**, es el porcentaje de variación en la variable de respuesta que es explicado por su relación con una o más variables predictoras, ajustado para el número de predictores en el modelo. El valor de R2 ajustado incorpora el número de predictores del modelo para ayudarle a elegir el modelo correcto.

Diagrama de Pareto

Está basado en el principio que dice “unas pocas causas son las que crean los mayores efectos”. El gráfico o diagrama de Pareto indica claramente qué causas crean los mayores problemas en la organización.

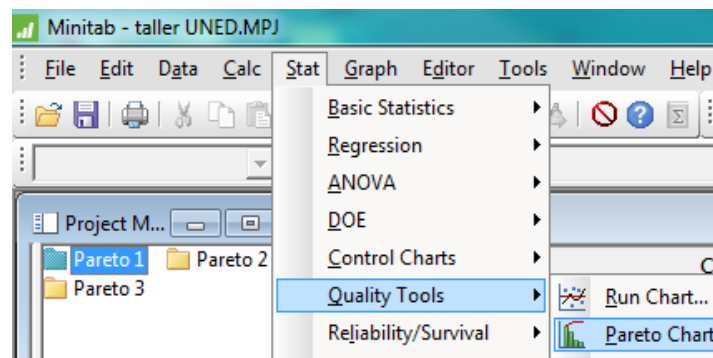
Consiste en que aproximadamente el 80% de los problemas (defectos) se deben a tan sólo el 20% de causas (“Principio de Pareto” o “Ley 80-20”)

Podemos realizar un diagrama de Pareto dependiendo del tipo de variable:

- Por frecuencia, para variables cuantitativas

1-Cargamos la base de datos, que contiene la variable cuantitativa: Quejas.

File->Open Worksheet->Seleccionar la base de datos

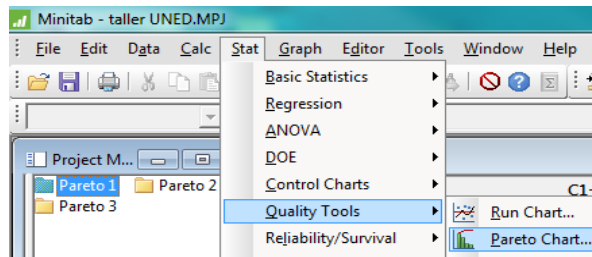


Base de datos:

↓	C1-T	C2
	Causas	Quejas
1	Falta de parqueo	11459
2	Mala atención farmacia	3047
3	Larga espera en control	10896
4	Acceso a emergencia	8632
5	Mala asignación de citas	9571
6	Falta de personal	2896
7	Mal diagnóstico	14782
8	Limpieza	4521

2- Ingresamos a Herramientas de calidad

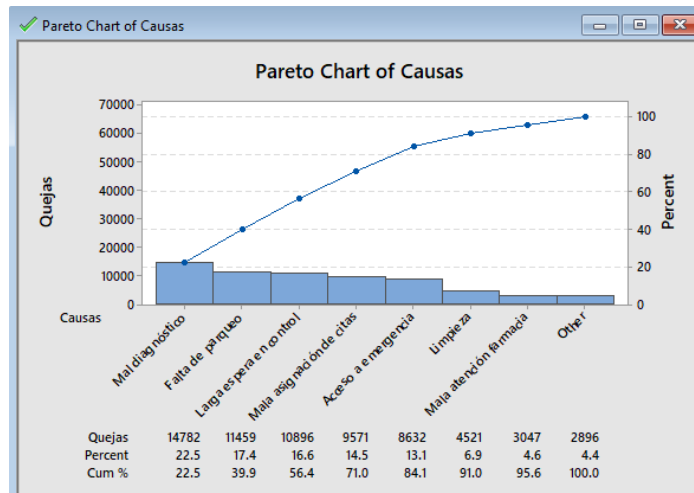
Stat->Quality Tools->Pareto Chart



3- Ingresamos las variables (BY Frequencies).



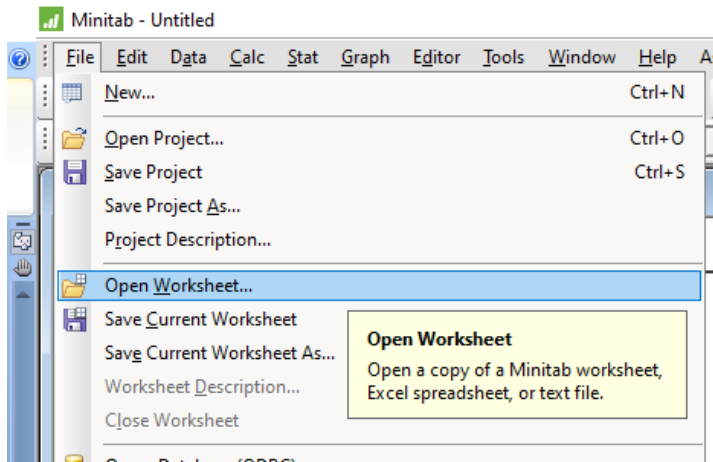
4- El resultado es el diagrama de Pareto, que indica cuáles son las causas que más afectan al Hospital.



- Por variable, para variables cualitativas

1-Cargamos la base de datos, que contiene la variable cualitativa: Turno.

File->Open Worksheet->Seleccionar la base de datos



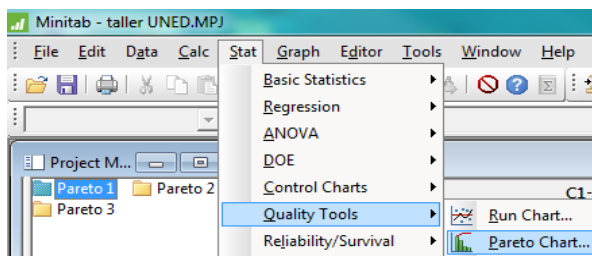
Base de datos:

The screenshot shows a Minitab data table with two columns: 'C1-T Defectos' and 'C2-T Turno'. The data is as follows:

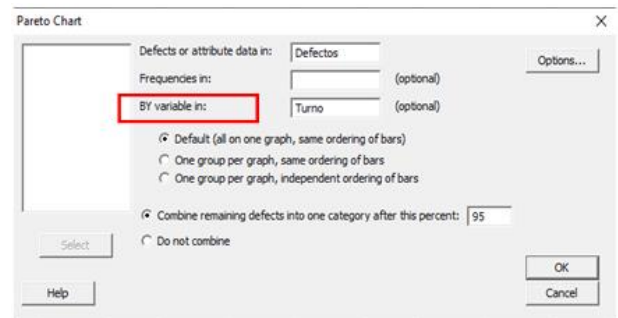
	C1-T Defectos	C2-T Turno
1	Raya	Día
2	Raya	Día
3	Desconchado	Día
4	Desconchado	Día
5	Mancha	Día
6	Raya	Día
7	Otro	Día
8	Otro	Tarde
9	Desconchado	Tarde
10	Desconchado	Tarde

2-Ingresamos a Herramientas de calidad

Stat->Quality Tools->Pareto Chart



3-Ingresamos las variables (BY variable).



4-El resultado es el diagrama de pareto, que indica cuales son los defectos que más afectan por turno.

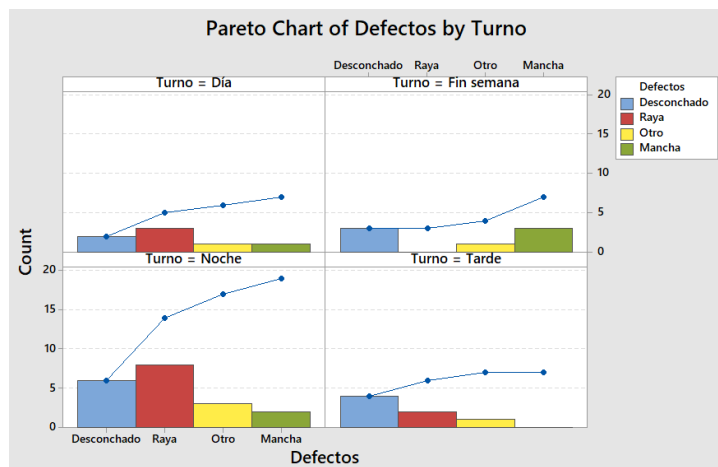


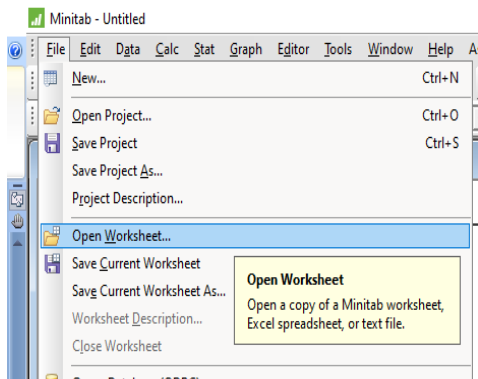
Gráfico de Histograma

Representa gráficamente la dispersión o variabilidad que presenta una serie de datos, por la frecuencia u ocurrencia de los datos de la variable.

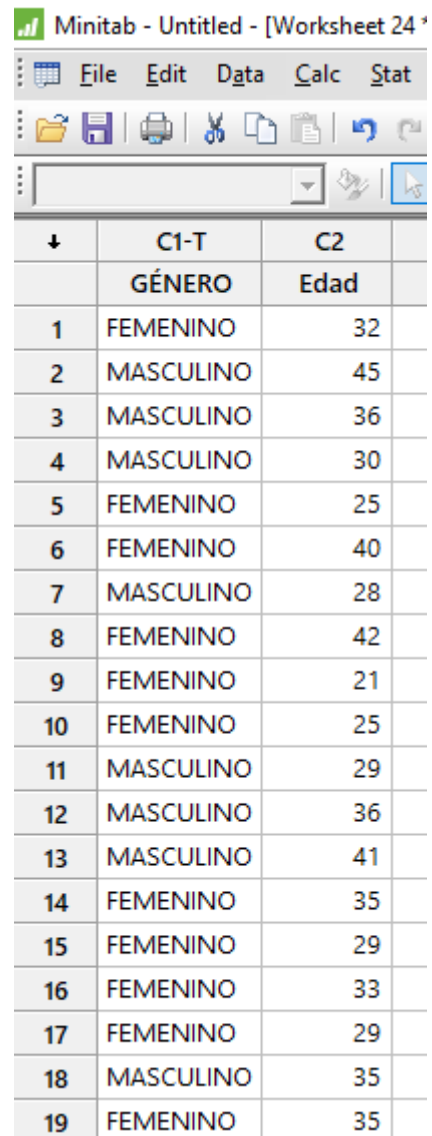
- Variable con 1 serie de datos (simple)

1-Cargamos la base de datos.

File->Open Worksheet->Seleccionar la base de datos



En este ejemplo se requiere graficar la distribución de la variable con 1 serie de datos de "Edad".

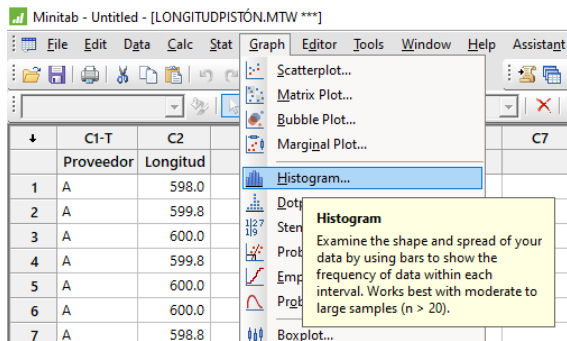


Minitab - Untitled - [Worksheet 24 ']

	C1-T	C2
	GÉNERO	Edad
1	FEMENINO	32
2	MASCULINO	45
3	MASCULINO	36
4	MASCULINO	30
5	FEMENINO	25
6	FEMENINO	40
7	MASCULINO	28
8	FEMENINO	42
9	FEMENINO	21
10	FEMENINO	25
11	MASCULINO	29
12	MASCULINO	36
13	MASCULINO	41
14	FEMENINO	35
15	FEMENINO	29
16	FEMENINO	33
17	FEMENINO	29
18	MASCULINO	35
19	FEMENINO	35

2-Seleccionamos el gráfico de Histograma

Graph->Histograma

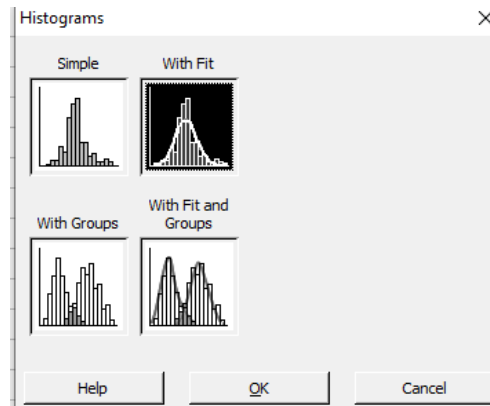


3- Se habilita un cuadro de selección.

Simple: Genera el histograma sin mostrar la curva de mejor ajuste.

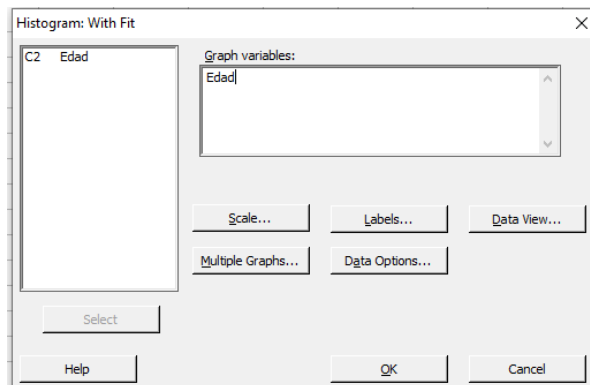
Con ajuste: Muestra la curva de mejor ajuste o la curva de distribución de los datos.

Con grupos: Cuando existe más de una serie de datos.

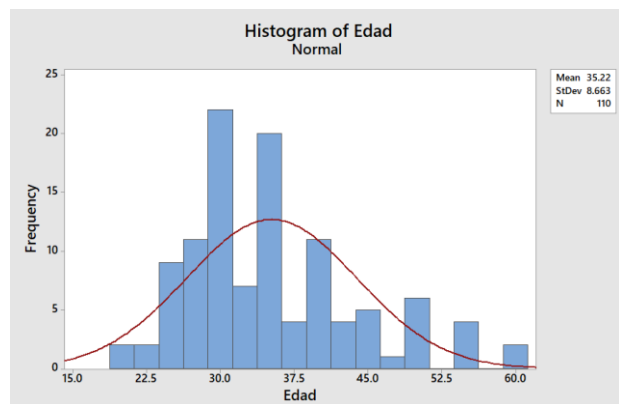


4-Ingresamos los datos al formulario.

En este ejemplo la variable que se graficará será la Edad.



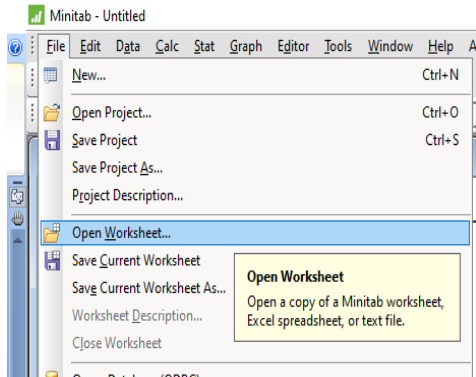
5-Se despliega el gráfico de Histograma, con la visualización de la distribución de frecuencia de la variable Edad.



- Variables con 1 o más series de datos (con grupos)

1-Cargamos la base de datos.

File->Open Worksheet->Seleccionar la base de datos

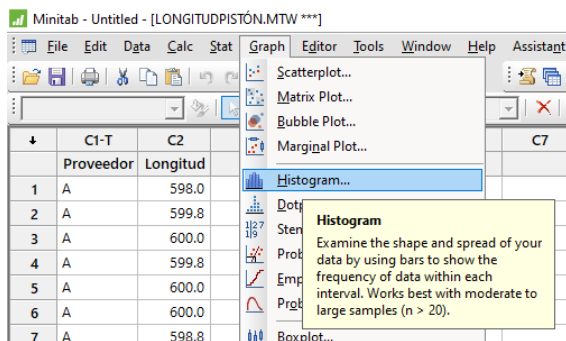


En este ejemplo se requiere graficar la distribución de la variable longitud, que contiene 2 series de datos para el proveedor A y B.

↓	C1-T	C2
	Proveedor	Longitud
97	A	599.8
98	A	599.0
99	A	599.6
100	A	599.4
101	B	601.6
102	B	600.4
103	B	598.4
104	B	600.0

2-Seleccionamos el gráfico de Histograma

Graph->Histograma

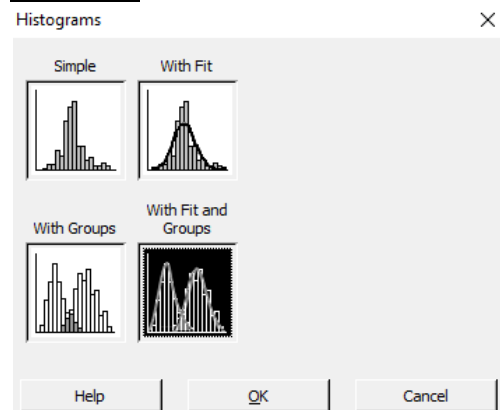


3- Se habilita un cuadro de selección.

Simple: Genera el histograma sin mostrar la curva de mejor ajuste.

Con ajuste: Muestra la curva de mejor ajuste o la curva de distribución de los datos.

Con grupos: Cuando existe más de una serie de datos.



4-Ingresamos los datos al formulario.

En este ejemplo existen datos para un Proveedor A y B, por lo que completamos el espacio de “Categoría por grupo”

Histogram: With Fit and Groups

C1 Proveedor
C2 Longitud

Graph variables:
Longitud

Categorical variables for grouping (0-3):
Proveedor

Graph variables form groups

Scale... Labels... Data View...

Select Multiple Graphs... Data Options...

Help OK Cancel

5-Se despliega el gráfico de Histograma, con la visualización de las dos series de datos.

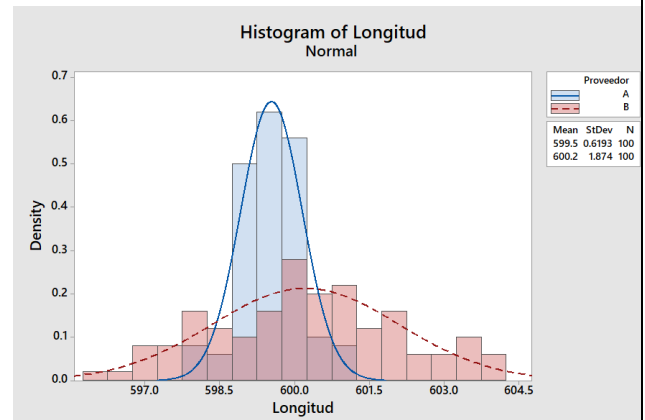
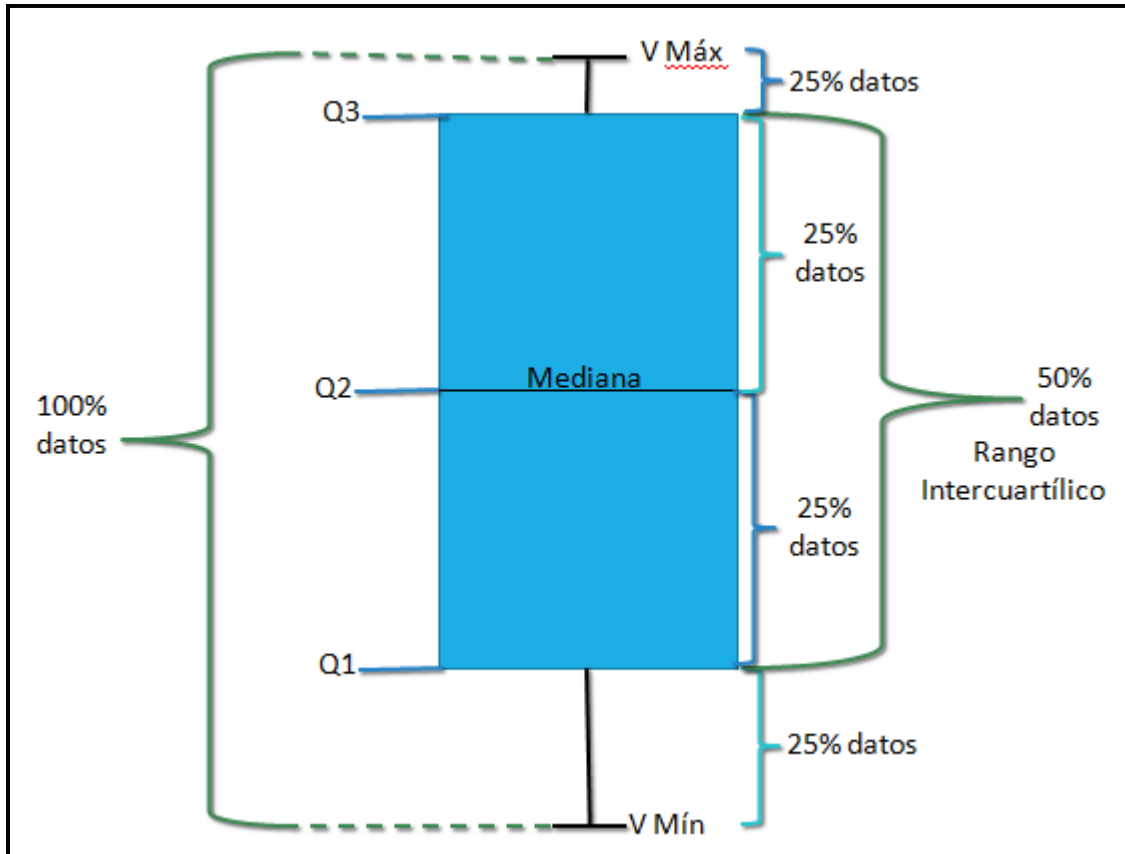


Gráfico de Cajas y Bigotes

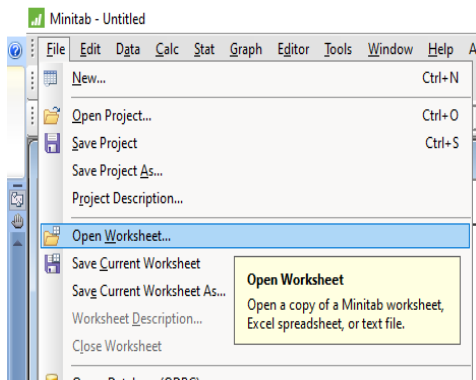
Un diagrama de caja y bigotes es un gráfico basado en cuartiles, mediante el cual se visualiza un conjunto de datos.

Es un gráfico que suministra información sobre los valores mínimo y máximo, los cuartiles Q1, Q2 o mediana y Q3, y sobre la existencia de valores atípicos y la dispersión de los datos.



1-Cargamos la base de datos.

File->Open Worksheet->Seleccionar la base de datos



Base de datos:

A screenshot of the Minitab software interface showing a worksheet titled 'Minitab - Untitled - [CRECIPLANTAS.MTW]'. The worksheet contains a data table with the following structure:

↓	C1-T	C2	C3
	Fertilizante	Altura	
1	Ninguno	23	
2	Ninguno	18	
3	Ninguno	18	
4	Ninguno	17	
5	Ninguno	25	
6	Ninguno	8	
7	Ninguno	23	
8	Ninguno	20	
9	Ninguno	24	
10	Ninguno	22	
11	Ninguno	26	
12	Ninguno	16	
13	Ninguno	27	
14	Ninguno	16	
15	Ninguno	17	
16	Ninguno	14	
17	Ninguno	17	
18	Ninguno	25	
19	Ninguno	13	

2- Seleccionamos el gráfico de caja y bigotes

Graph->Boxplot

	C1-T	C2
	Fertilizante	Altura
1	Ninguno	23
2	Ninguno	18
3	Ninguno	18
4	Ninguno	17
5	Ninguno	25
6	Ninguno	8
7	Ninguno	23
8	Ninguno	20
9	Ninguno	24
10	Ninguno	22
11	Ninguno	26
12	Ninguno	16
13	Ninguno	27

3-Se habilita un cuadro de selección

Una Y: corresponde a una variable de una misma muestra. Puede ser de dos tipos:

- Simple: se utiliza cuando se requiere graficar la dispersión de una variable. Ejemplo: Una muestra A presenta diferentes datos de altura.
- Con grupos: se utiliza cuando en una misma muestra se presentan en diferentes clases o categorías para la variable. Ejemplo: Una muestra A presenta tres tipos de aditivos (x1, x2, x3), con diferentes datos de altura para cada uno.

Múltiples Y: corresponde cuando la variable a graficar se presenta en más de una muestra.

- Simple: se utiliza cuando la variable a graficar se presenta en dos o más muestras. Ejemplo: Se requiere graficar para la muestra A y la muestra B la variable altura.
- Con grupos: se utiliza cuando en una misma muestra se presentan en diferentes clases o categorías para la variable. Ejemplo: Una muestra A y otra muestra B presentan tres tipos de aditivos (x1, x2, x3), con diferentes datos de altura para cada uno.

Boxplots

One Y

Simple With Groups

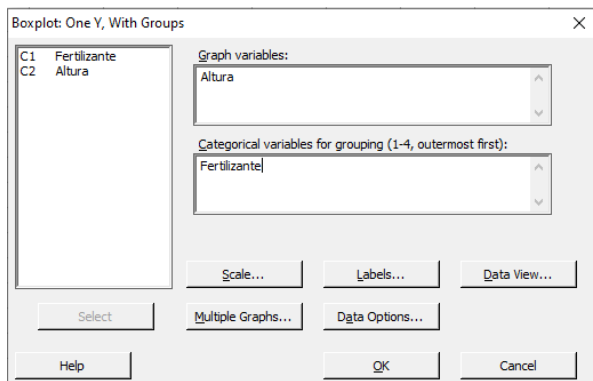
Multiple Y's

Simple With Groups

Help OK Cancel

4-Se ingresan los valores a graficar.

En este ejemplo se tiene una muestra de fertilizante (Una Y) con varios grupos de aditivos para obtener diferentes datos de altura.



5-Se obtienen los gráficos de cajas y bigotes

