

Análisis estadístico con Microsoft Excel

Microsoft Excel ofrece un conjunto de herramientas para el análisis de los datos (denominado Herramientas para análisis) con el que podrá ahorrar pasos en el desarrollo de análisis estadístico o de ingeniería. Cuando utilice una de estas herramientas, deberá proporcionar los datos y parámetros para cada análisis, la herramienta utilizará las funciones macro estadísticas o de ingeniería apropiadas y mostrará los resultados en una tabla de resultados. Algunas herramientas generan gráficos además de tablas de resultados.

Para ver una lista de las herramientas de análisis disponibles, elija Análisis de datos en el menú Herramientas. Si este comando no está en el menú, ejecute el programa de instalación para instalar las Herramientas para análisis de la forma siguiente :

Activar las Herramientas para análisis

- 1 En el menú Herramientas, elija Macros automáticas. Si en la lista del cuadro de diálogo Macros automáticas no aparece Herramientas para análisis, haga clic en el botón "Examinar" y busque la unidad, directorio y archivo de la macro automática Herramientas para análisis, o bien ejecute el programa de instalación.
- 2 Seleccione la casilla de verificación "Herramientas para análisis".

Nota Las macros automáticas que instale en Microsoft Excel permanecerán activas hasta que las quite.

Utilización de los cuadros de diálogo de Analisis de datos

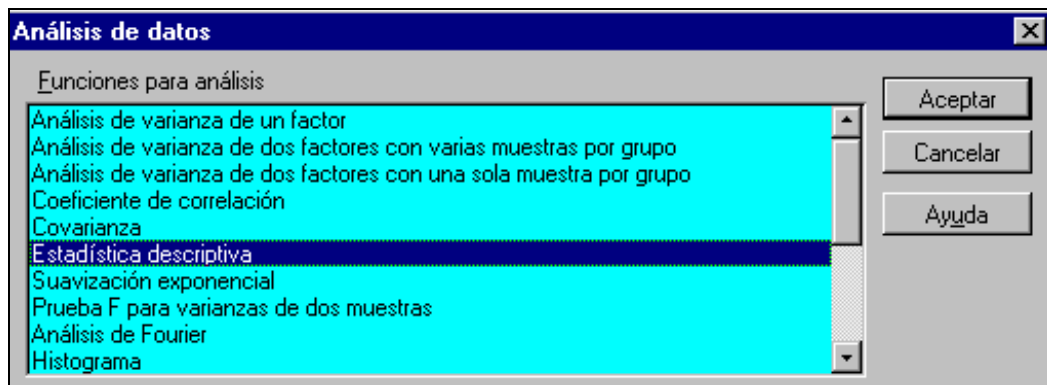
El funcionamiento de Analisis de datos es interactivo. Siempre debe rellenar el campo (o campos) indicado como **Rango de Entrada** con (el rango de) los datos que se van a analizar. Los resultados se escriben por defecto en una hoja nueva. Si se desea obtenerlos en la misma hoja debe seleccionar la opción y el del **Rango de salida** con la coordenada superior izquierda de la zona en donde desea que se muestren los resultados. El resto de los campos puede o no seleccionarse según el tipo de análisis que desee efectuar.

Veamos un ejemplo de como debe procederse usando como guia el análisis "Estadística descriptiva".

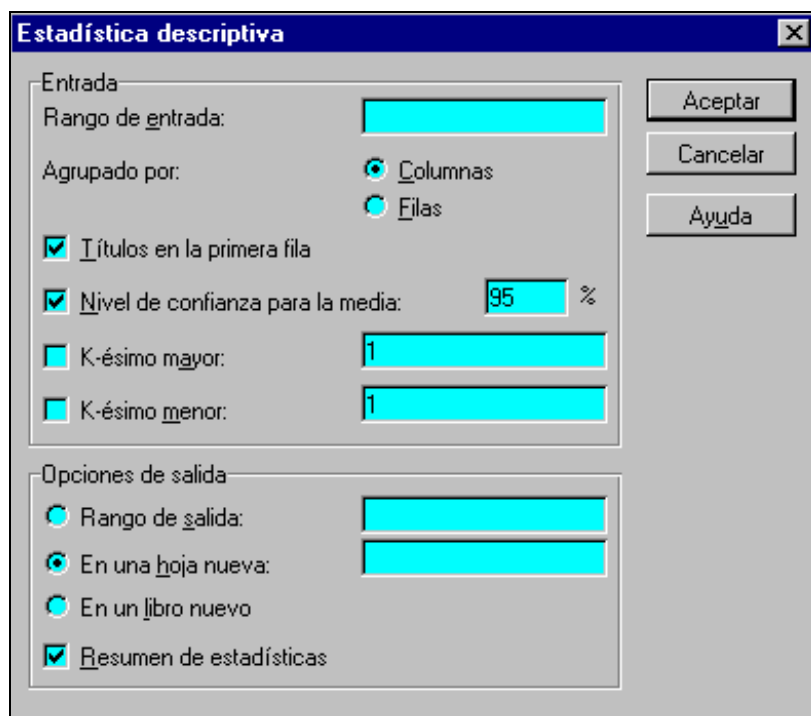
Empezamos por seleccionar el analisis del menú Herramientas :

Herramientas / Análisis de datos

Aparecerá la lista de opciones de Análisis de datos en donde seleccionaremos Estadística descriptiva :



Seleccionando Aceptar aparece el cuadro de diálogo de Estadística descriptiva :

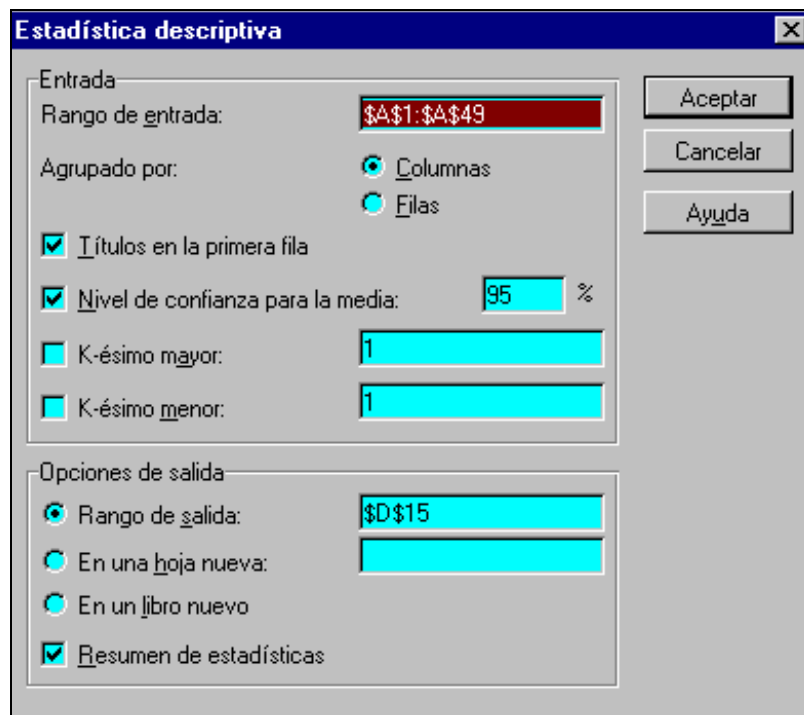


Cómo hemos indicado lo único que es "obligatorio" suministrar son los datos a analizar (Rango de entrada) y el lugar en donde se desea escribir los resultados (Rango de salida). Se pueden especificar ambos por el procedimiento de señalar con el cursor pero debemos asegurarnos, antes de hacerlo, que el campo que vamos a rellenar se halla seleccionado. Veámoslo con un ejemplo.

- Por defecto el cursor se encuentra situado en el campo "Rango de entrada". Para rellenarlo solamente debemos poner el ratón sobre la hoja que contiene los datos y seleccionar este rango. En nuestro ejemplo seleccionaremos la primera fila desde la celda 1 hasta la 49 es decir el rango \$A\$1 :\$A\$49.
- A continuación debemos activar la selección del rango de salida para lo cual :
 1. Marcaremos el boton "Rango de salida" en la zona "Opciones de salida",
 2. haremos click dentro del campo "Rango de salida" de forma que aparezca el cursor en su interior

3. Hecho lo anterior podemos indicar la zona en donde queremos que aparezca el resultado seleccionando con el cursor **una única celda** que se constituye en la celda superior izquierda del rango de salida.
- Si se desean seleccionar otras opciones se procederá a marcar los botones o opciones de selección y una vez finalizado se pulsará "Aceptar" con el ratón, lo que ejecutará el análisis escribiendo los resultados en la posición escogida.

En nuestro caso el aspecto de la caja de diálogo, una vez rellena será el siguiente :



y el aspecto de la salida, que empieza en la celda \$D\$15, será :

| <i>FUELI.</i> | |
|---------------------|--------------|
| Media: | 1.854166667 |
| Error: | 0.220793185 |
| Media: | 2 |
| Moda: | 0 |
| Desv: | 1.529700059 |
| Varia: | 2.33998227 |
| Curtosis: | -0.803962141 |
| Coef. de asimetría: | 0.404804585 |
| Rango: | 5 |
| Mínimo: | 0 |
| Máximo: | 5 |
| Suma: | 89 |
| Cuentas: | 48 |
| Nivel de confianza: | 0.444178177 |

La salida queda seleccionada por defecto, con lo que se puede mejorar su aspecto de forma muy simple indicando que la columna ajuste su anchura al texto, lo que se hará con la opción de menú

Formato / Columna / Ajustar a la selección

que dará a los resultados el aspecto siguiente :

| <i>FUELI.</i> | |
|---------------------------|--------------|
| Media | 1.854166667 |
| Error típico | 0.220793185 |
| Mediana | 2 |
| Moda | 0 |
| Desviación estándar | 1.529700059 |
| Varianza de la muestra | 2.33998227 |
| Curtosis | -0.803962141 |
| Coefficiente de asimetría | 0.404804585 |
| Rango | 5 |
| Mínimo | 0 |
| Máximo | 5 |
| Suma | 89 |
| Cuenta | 48 |
| Nivel de confianza(95.0%) | 0.444178177 |

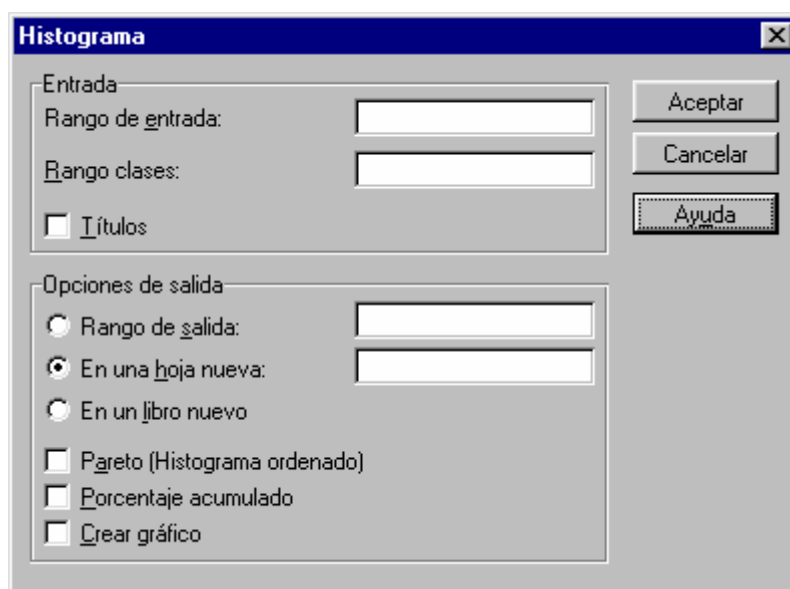
El aspecto de esta salida de resultados se basa en las opciones elegidas. Otra combinación de selecciones puede dar lugar a un aspecto distinto.

Realización de análisis estadísticos con “Análisis de datos”

Histograma

Calcula las frecuencias individuales y acumulativas de rangos de celdas de datos y de clases de datos. Genera datos acerca del número de apariciones de un valor en un conjunto de datos. Por ejemplo, en una clase con 20 alumnos se desea obtener la distribución de calificaciones mediante una categoría de puntuación por letras. Una tabla de histograma presentará los límites de las calificaciones por letras así como el número de calificaciones que hay entre el límite más bajo y el actual. La calificación más frecuente es la moda de los datos.

El aspecto del cuadro de diálogo es el siguiente :



Opciones del cuadro de diálogo Histograma

Rango de entrada

Escriba la referencia de celda correspondiente a los datos que desee analizar.

Rango clases (opcional)

Escriba un rango de celdas con un conjunto opcional de valores límite que defina el rango clases. Estos valores deberán estar en orden ascendente. Microsoft Excel cuenta los puntos de datos que hay entre el número de clases actual y el número de clases mayor, si los hubiera. Se contará un número de una clase determinada si es igual o menor que el número de clase situado por debajo de la última clase. Se contarán todos los valores inferiores al primer valor de clase y superiores al último valor de clase. Si omite el rango de clase, Microsoft Excel creará un conjunto de clases distribuidas desigualmente entre los valores mínimo y máximo de los datos.

Rango de salida

Escriba la referencia correspondiente a la celda superior izquierda de la tabla de resultados. Microsoft Excel determinará el tamaño del área de resultados y mostrará un mensaje si la tabla de resultados reemplaza datos ya existentes o si sobrepasa los límites de la hoja de cálculo.

En una hoja nueva

Haga clic aquí para insertar otra hoja en el libro actual y pegar los resultados en la celda A1. Para darle un nombre a la nueva hoja de cálculo, escríbalo en el cuadro.

En un libro nuevo

Haga clic aquí para crear otro libro y pegar los resultados en la celda A1 del nuevo libro.

Pareto (Histograma ordenado)

Active esta casilla para presentar los datos en orden de frecuencia descendente en la tabla de resultados. Si está desactivada, Microsoft Excel presentará los datos en orden ascendente y omitirá las tres columnas situadas más a la derecha que contengan datos ordenados.

Porcentaje acumulado

Active esta casilla para generar una columna de tabla de resultados con porcentajes acumulados y para incluir una línea de porcentaje acumulado en el gráfico de histograma. Desactívela para omitir los porcentajes acumulados.

Crear gráfico

Active esta casilla para generar un gráfico de histograma incrustado con la tabla de resultados.

Títulos

Active esta casilla si la primera fila y la primera columna del rango de entrada contienen títulos. Desactívela si el rango de entrada no tiene títulos. Microsoft Excel generará los títulos de datos apropiados para la tabla de resultados.

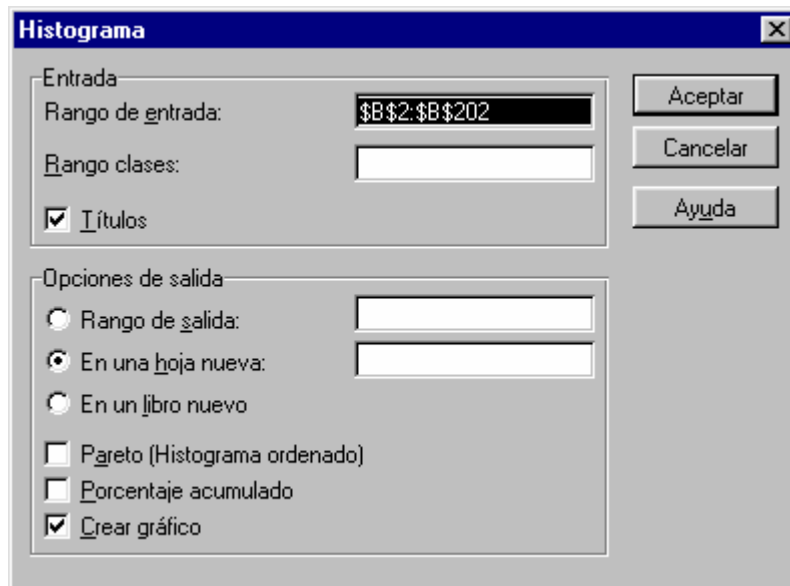
Ejemplo

Supóngase que se desea realizar un gráfico de 200 valores correspondientes a valores de pólizas de una compañía de seguros. Los datos se hallan disponibles en una hoja de cálculo, "SEGUROS.XLS" en la columna 1, filas 2 a 201.

Empezamos por seleccionar el análisis del menú Herramientas :

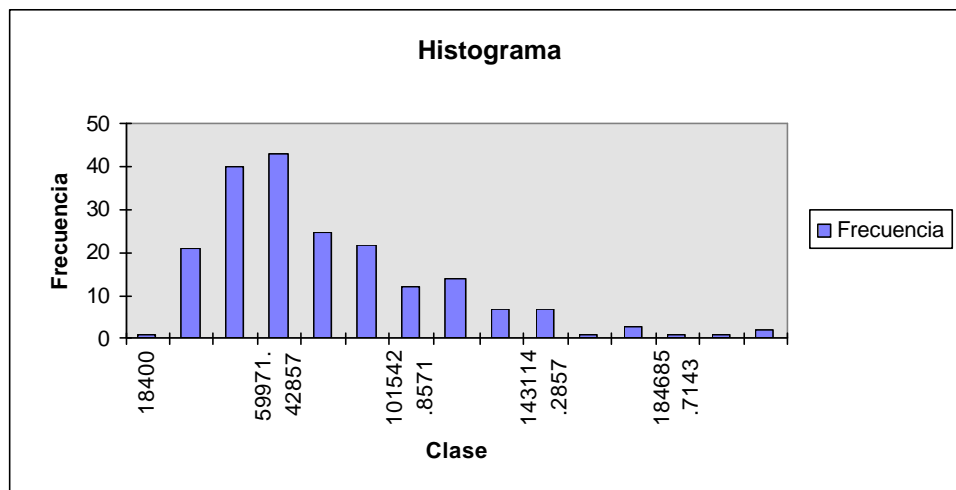
Herramientas / Análisis de datos

Aparecerá la lista de opciones de Análisis de datos en donde seleccionaremos **Histograma**. Pulsando **Aceptar** aparece el cuadro de diálogo de **Histograma** que llenaremos de la forma siguiente ::

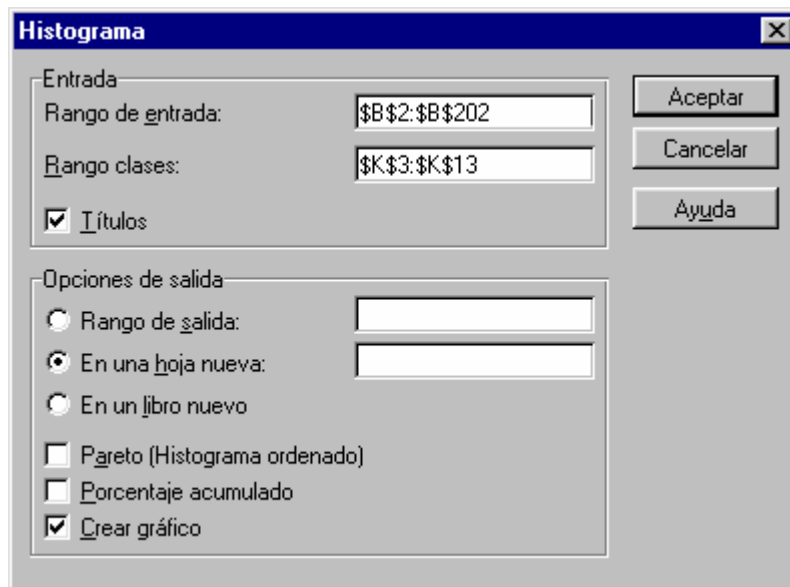


Pulsando Aceptar obtendremos la tabla y el gráfico siguientes, en una nueva hoja :

| Clase | Frecuencia |
|------------|------------|
| 18400 | 1 |
| 32257.1429 | 21 |
| 46114.2857 | 40 |
| 59971.4286 | 43 |
| 73828.5714 | 25 |
| 87685.7143 | 22 |
| 101542.857 | 12 |
| 115400 | 14 |
| 129257.143 | 7 |
| 143114.286 | 7 |
| 156971.429 | 1 |
| 170828.571 | 3 |
| 184685.714 | 1 |
| 198542.857 | 1 |
| y mayor... | 2 |



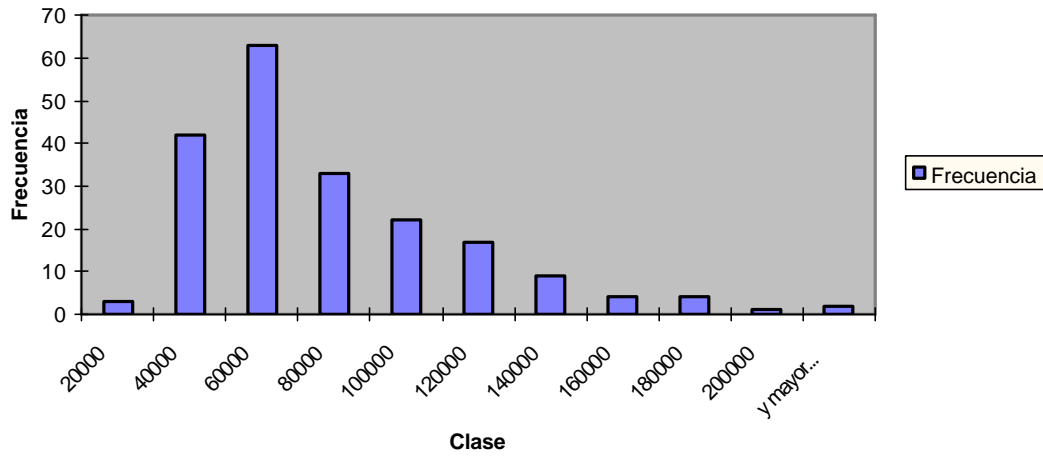
El aspecto de la tabla y el gráfico se puede mejorar indicando al programa los rangos de clases sobre los que se desea agrupar los valores. Si, p.ejemplo se escriben estos rangos en la columna K, el aspecto del diálogo Histograma será :



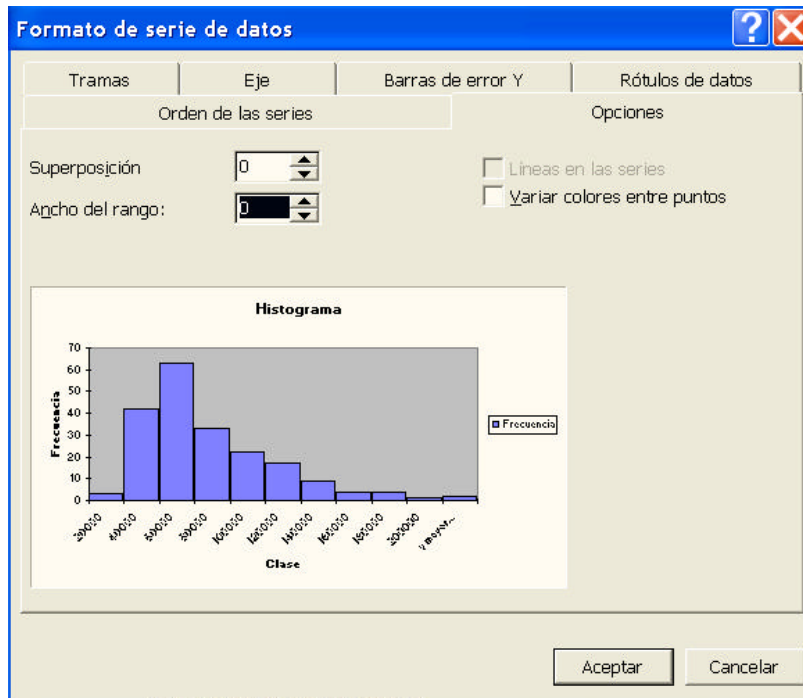
y el gráfico y la tabla que se obtienen al pulsar Aceptar:

| <i>Clase</i> | <i>Frecuencia</i> |
|--------------|-------------------|
| 20000 | 3 |
| 40000 | 42 |
| 60000 | 63 |
| 80000 | 33 |
| 100000 | 22 |
| 120000 | 17 |
| 140000 | 9 |
| 160000 | 4 |
| 180000 | 4 |
| 200000 | 1 |
| y mayor... | 2 |

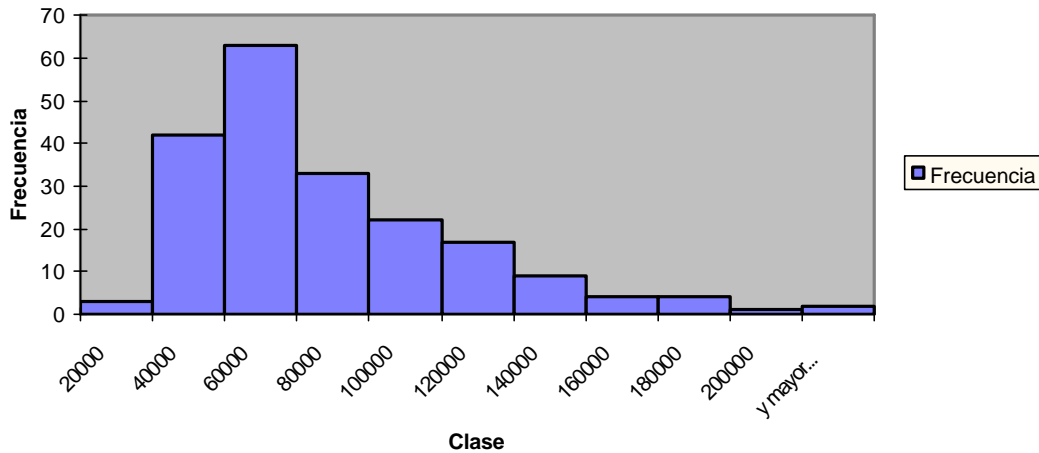
Histograma



Podemos mejorar el aspecto del gráfico si abrimos el gráfico y seleccionando la serie de datos apretamos el botón derecho y escogemos la opción "Formato de la serie de datos". En esta opción debemos escoger: "Opciones" y poner el "Ancho de rango" a 0



Histograma



Estadística descriptiva

Genera un informe de estadísticas de una sola variable para datos del rango de entrada, y proporciona información acerca de la tendencia central y dispersión de los datos.

El aspecto de los cuadros de diálogo entrada y la salida de resultados puede verse en el apartado anterior en el epígrafe "Utilización de los cuadros de diálogo de Análisis de datos".

Opciones del cuadro de diálogo Estadística descriptiva

Rango de entrada

Escriba la referencia de celda correspondiente a los datos que desee analizar. La referencia deberá contener dos o más rangos adyacentes organizados en columnas o filas.

Rango de salida

Escriba la referencia correspondiente a la celda superior izquierda de la tabla de resultados. Esta herramienta genera dos columnas de información por cada conjunto de datos. La columna de la izquierda contiene los títulos de las estadísticas y la derecha, las estadísticas. Microsoft Excel escribirá una tabla de estadística de dos columnas por cada columna o fila del rango de entrada, dependiendo de la opción que se haya seleccionado en el cuadro "Agrupado por".

En una hoja nueva

Haga clic aquí para insertar otra hoja en el libro actual y pegar los resultados en la celda A1. Para darle un nombre a la nueva hoja de cálculo, escríbalo en el cuadro.

En un libro nuevo

Haga clic aquí para crear otro libro y pegar los resultados en la celda A1 del nuevo libro.

Agrupado por

Haga clic en el botón "Filas" o "Columnas" para indicar si los datos del rango de entrada están organizados en filas o en columnas.

Títulos en la primera fila/Títulos en la primera columna

Si la primera fila del rango de entrada contiene títulos, active la casilla de verificación "Títulos en la primera fila" o, por el contrario, si hay títulos en la primera columna, active la casilla "Títulos en la primera columna". Desactívela si el rango de entrada carece de títulos; Microsoft Excel generará los títulos de datos apropiados para la tabla de resultados.

Resumen de estadísticas

Active esta casilla si desea que Microsoft Excel genere un campo por cada una de las siguientes estadísticas de la tabla de resultados: Media, Error típico (de la media), Mediana, Moda, Desviación estándar, Varianza de la muestra, Curtosis, Coeficiente de asimetría, Rango, Mínimo, Máximo, Suma, Cuenta, Mayor (#), Menor (#) y Nivel de confianza.

K-ésimo mayor

Active esta casilla si desea incluir una fila correspondiente al valor k-ésimo mayor de cada rango de datos en la tabla de resultados. Escriba en el cuadro el número que se utilizará para k. Si k es 1, dicha fila contendrá el máximo del conjunto de datos.

K-ésimo menor

Active esta casilla si desea incluir una fila correspondiente al valor k-ésimo menor de cada rango de datos en la tabla de resultados. Escriba en el cuadro el número que se utilizará para k. Si k es 1, dicha fila contendrá el mínimo del conjunto de datos.

Nivel de confianza para la media

Active esta casilla si desea incluir una fila correspondiente al nivel de confianza de la media en la tabla de resultados. Escriba en el cuadro el nivel de confianza que desee usar, por ejemplo, con un valor 95%, se calculará el nivel de confianza de la media con un nivel de significación del 5%.