

Clase 2 – Excel

Procesamiento de datos.

Andrés Felipe Ochoa Muñoz

andres.ochoa@correounivalle.edu.co

Tabla de contenido.

- Ingreso y validación de datos.
- Filtros.
- Formulas.
- Gráficos.
- Módulo Análisis de Datos.

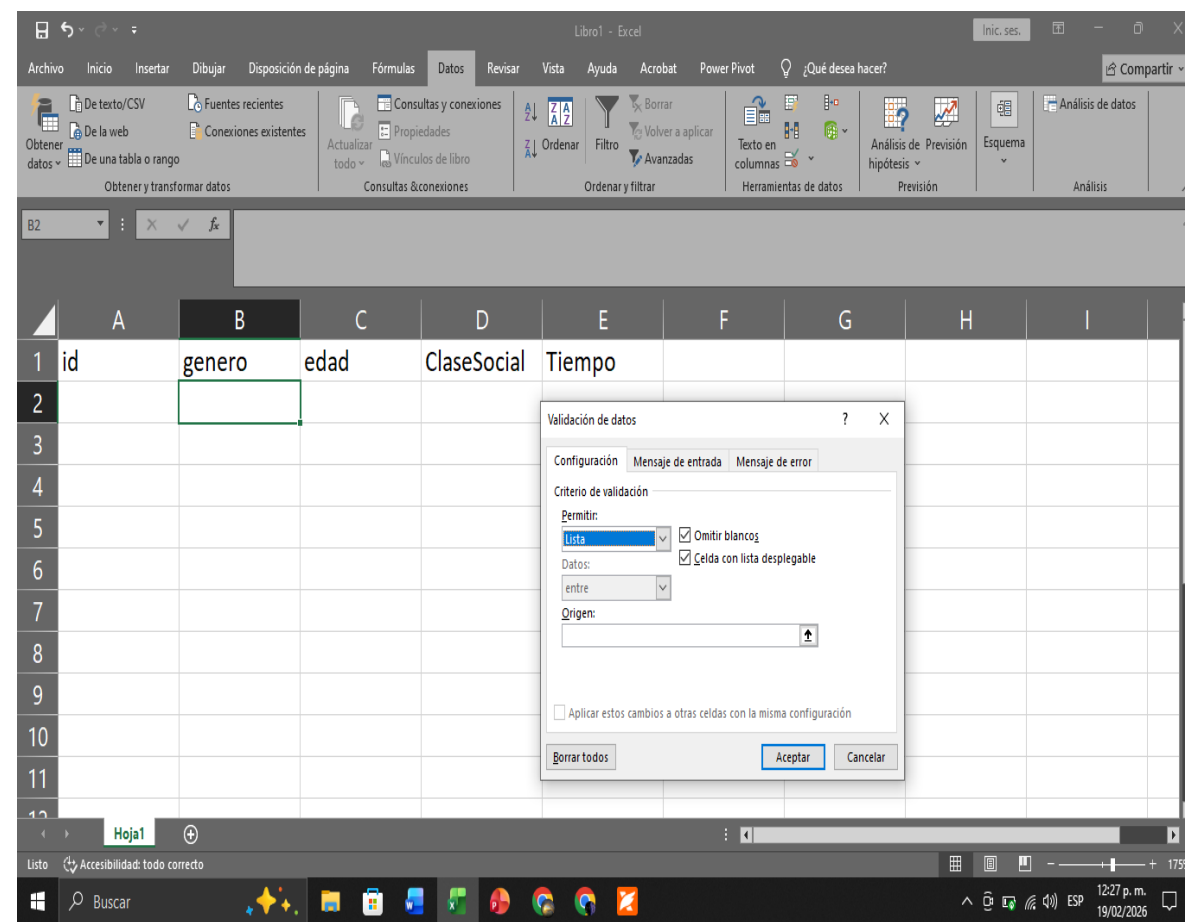
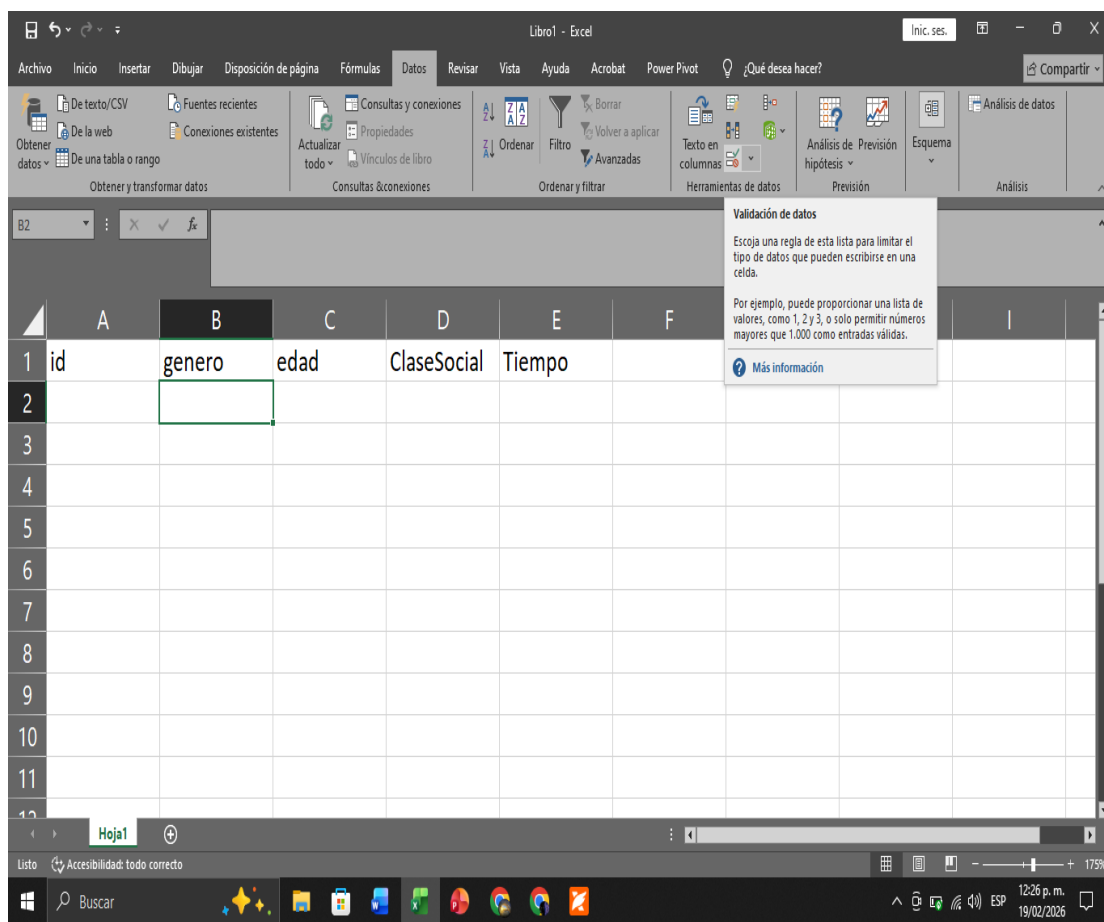
Ingreso de datos.

- Usualmente los conjuntos de datos se estructuran de forma que las filas se dejan para unidades de observación y las columnas se dejan para las variables.
- La primera fila del conjunto de datos usualmente se utiliza para el encabezado del nombre de las variables. Además, la primera columna para el id, la cuál es el identificador de cada unidad de observación o código de la encuesta realizada.
- Antes de ingresar los datos, se recomienda utilizar reglas de **validación de datos** con el fin de que se ingrese correctamente la información.

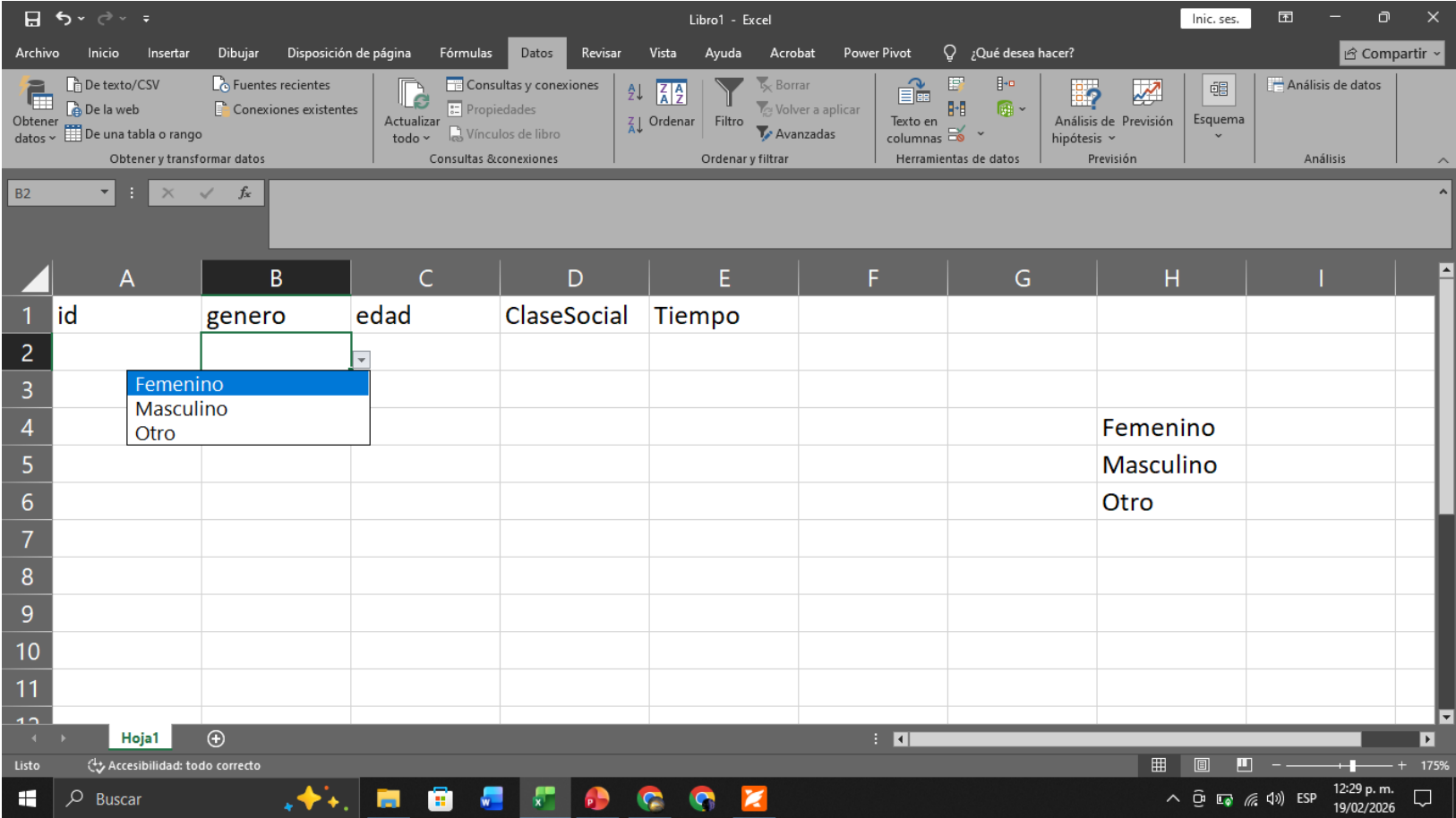
Ejemplo 1: Tiempo de espera en una sucursal bancaria.

1. Crear un archivo Excel, el cuál tenga el nombre “TiemposSucursal.xlsx”.
2. Crear en el encabezado (primer fila) los siguientes nombres: id, genero, edad, ClaseSocial, Tiempo.
3. Para todas las columnas (menos el id) crear reglas de validación de datos, el genero que tenga tres categorías (femenino, masculino, otro). Edad sea una variable entre 18 y 100 años, ClaseSocial tenga tres categorías (baja, media, alta). La variable tiempo definida en minutos se encuentre en el rango de 0 a 300 minutos.
4. Ingresar datos simulados.

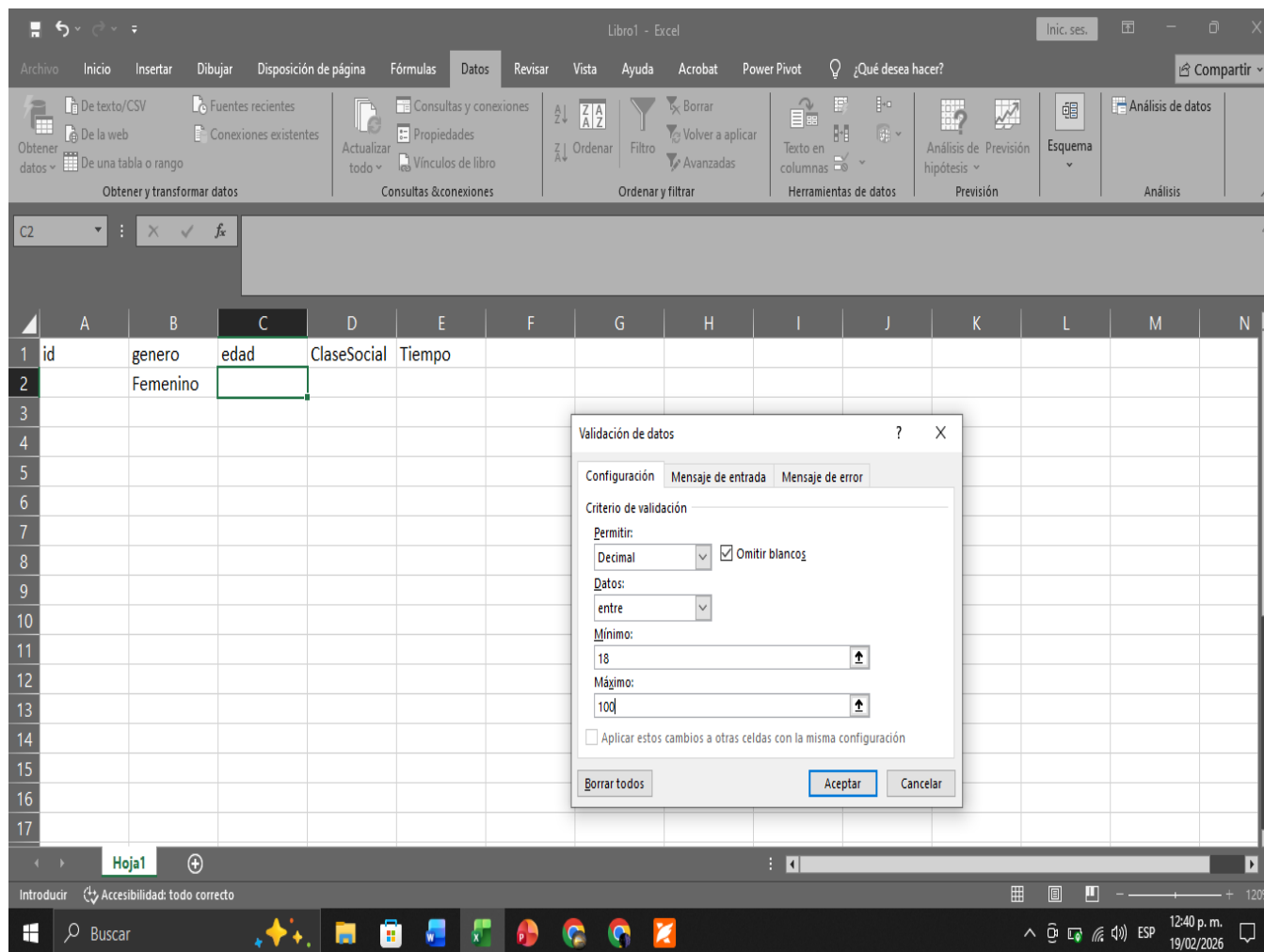
Validación de datos: opción Lista.



Validación de datos: opción Lista.

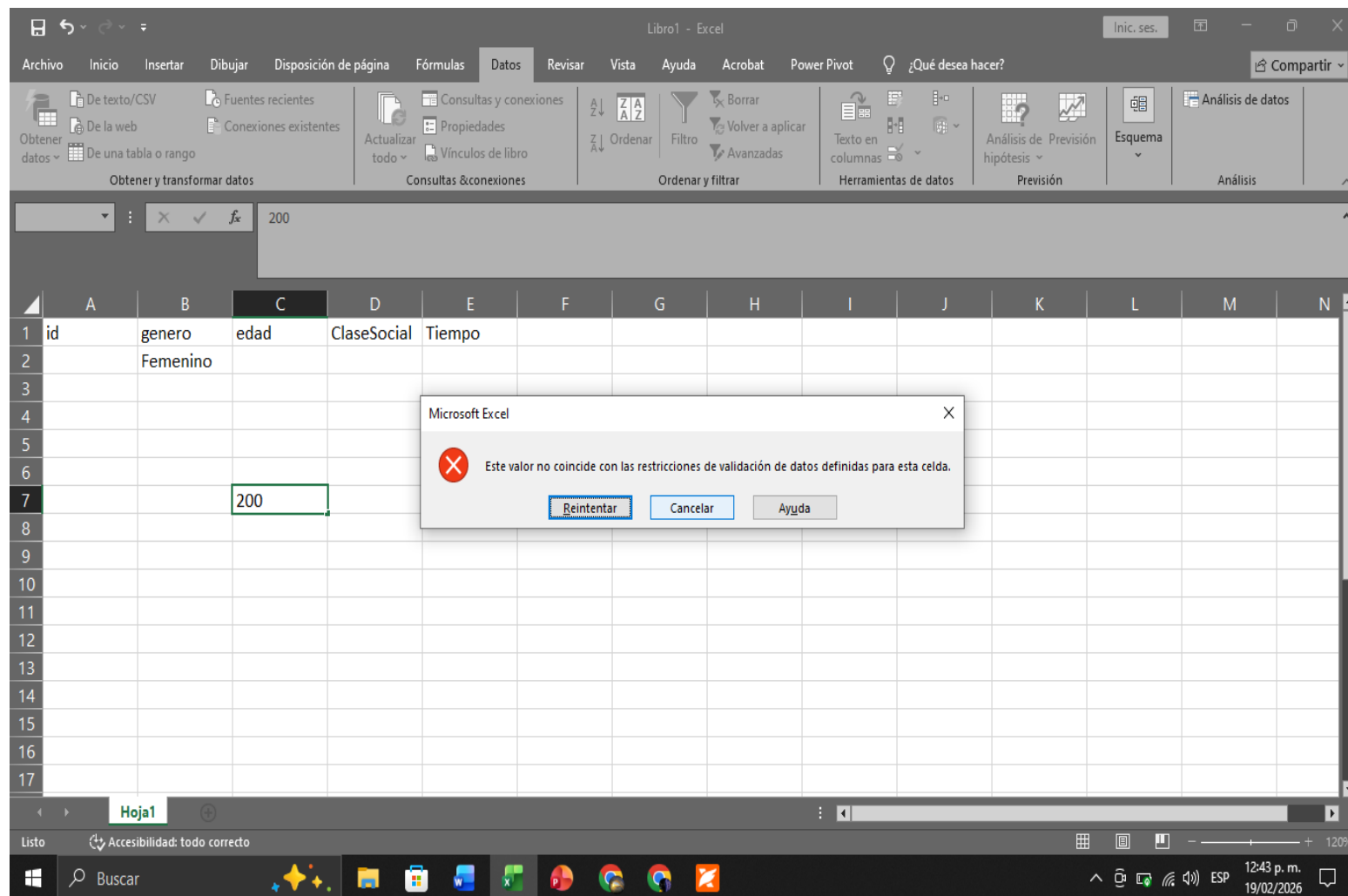


Validación de datos: opción decimal.



Para que la regla de validación quede en las demás celdas se recomienda copiar y pegar (pegado especial) y seleccionar la casilla validación.

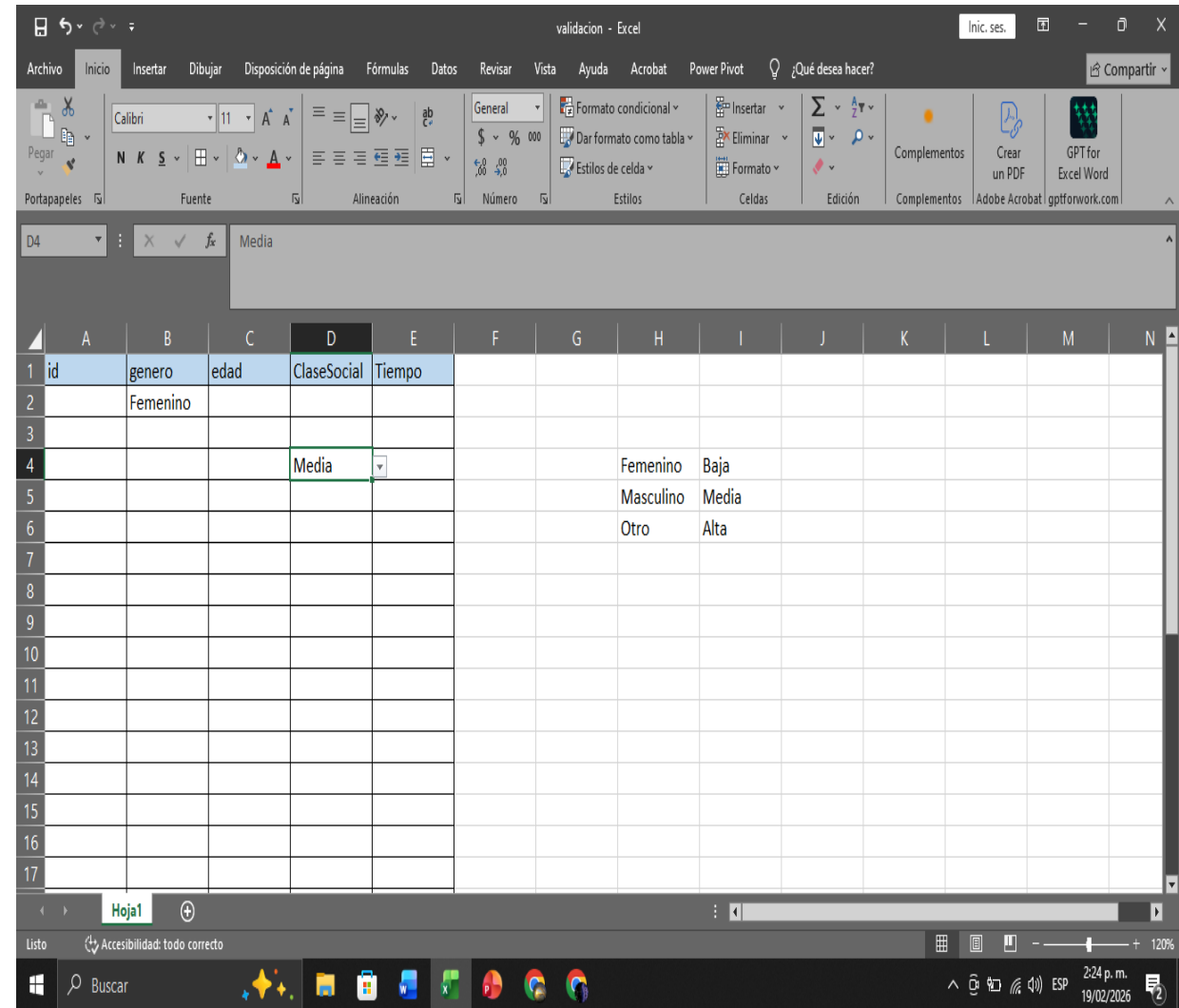
Validación de datos: opción decimal.



Si se digita mal un dato se pueda generar una alerta.

Validación de datos.

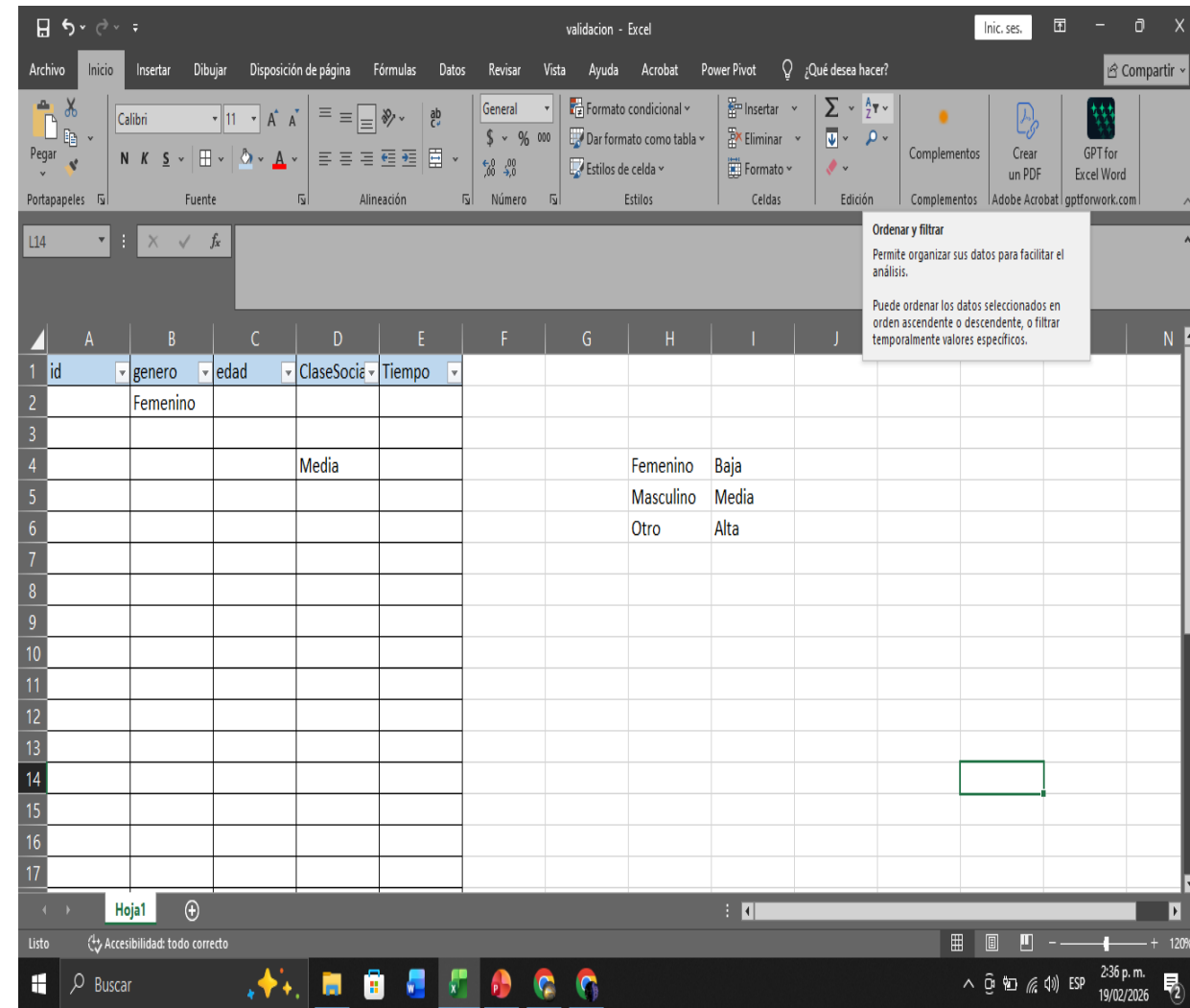
- Utilice validación de datos con la opción lista para crear las categorías de clase social (baja, media, alta).
- Utilice validación de datos con la opción decimal para ingresar los tiempos de espera en minutos, rango de 0 a 300 minutos.



Filtros.

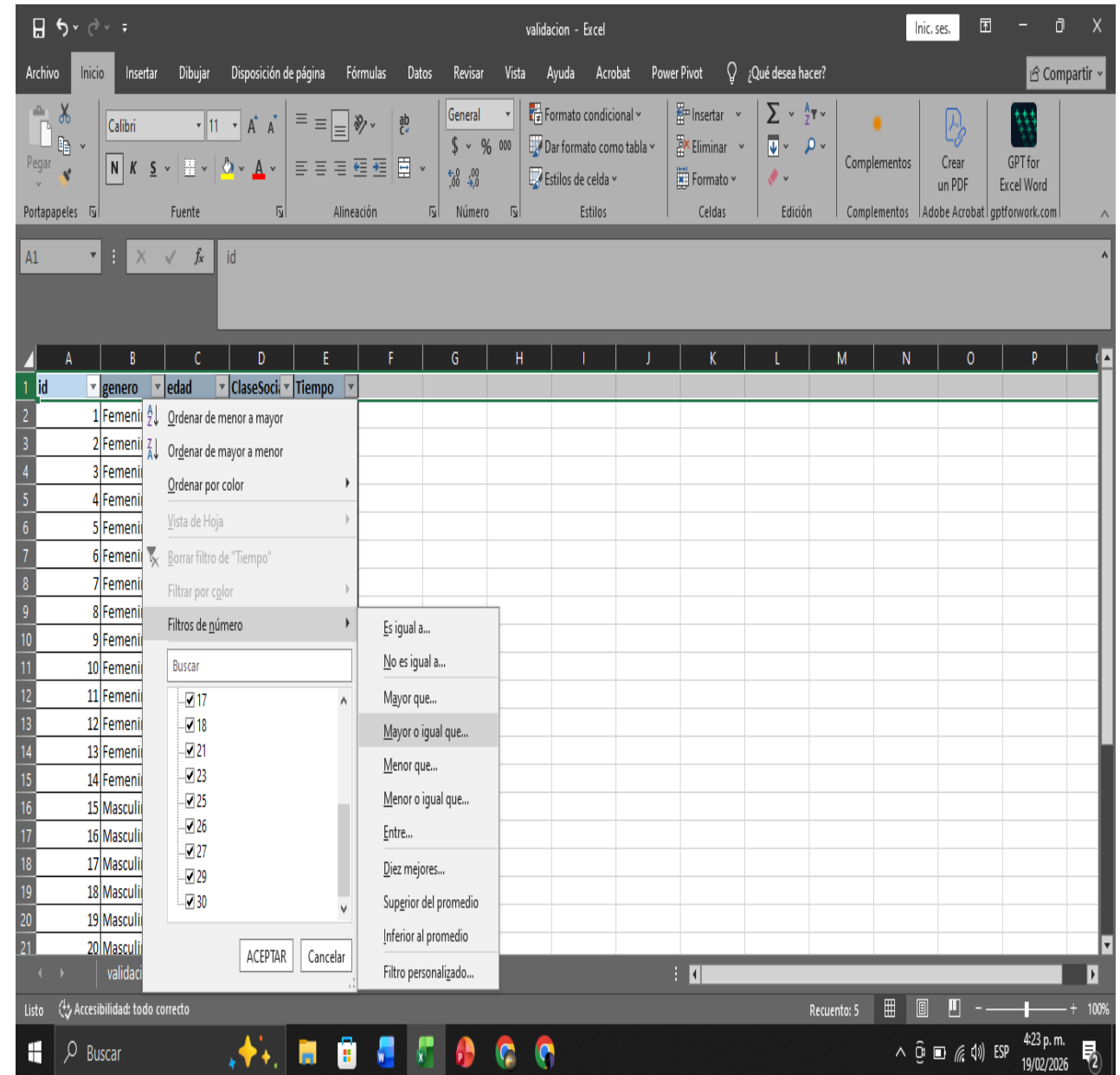
Filtros.

- Los filtros en Excel permiten visualizar rápidamente filas específicas en conjuntos de datos grandes, ocultando temporalmente la información no deseada. Se activan en la pestaña **Datos** > **Filtro** o con Ctrl + Shift + L.



Filtros.

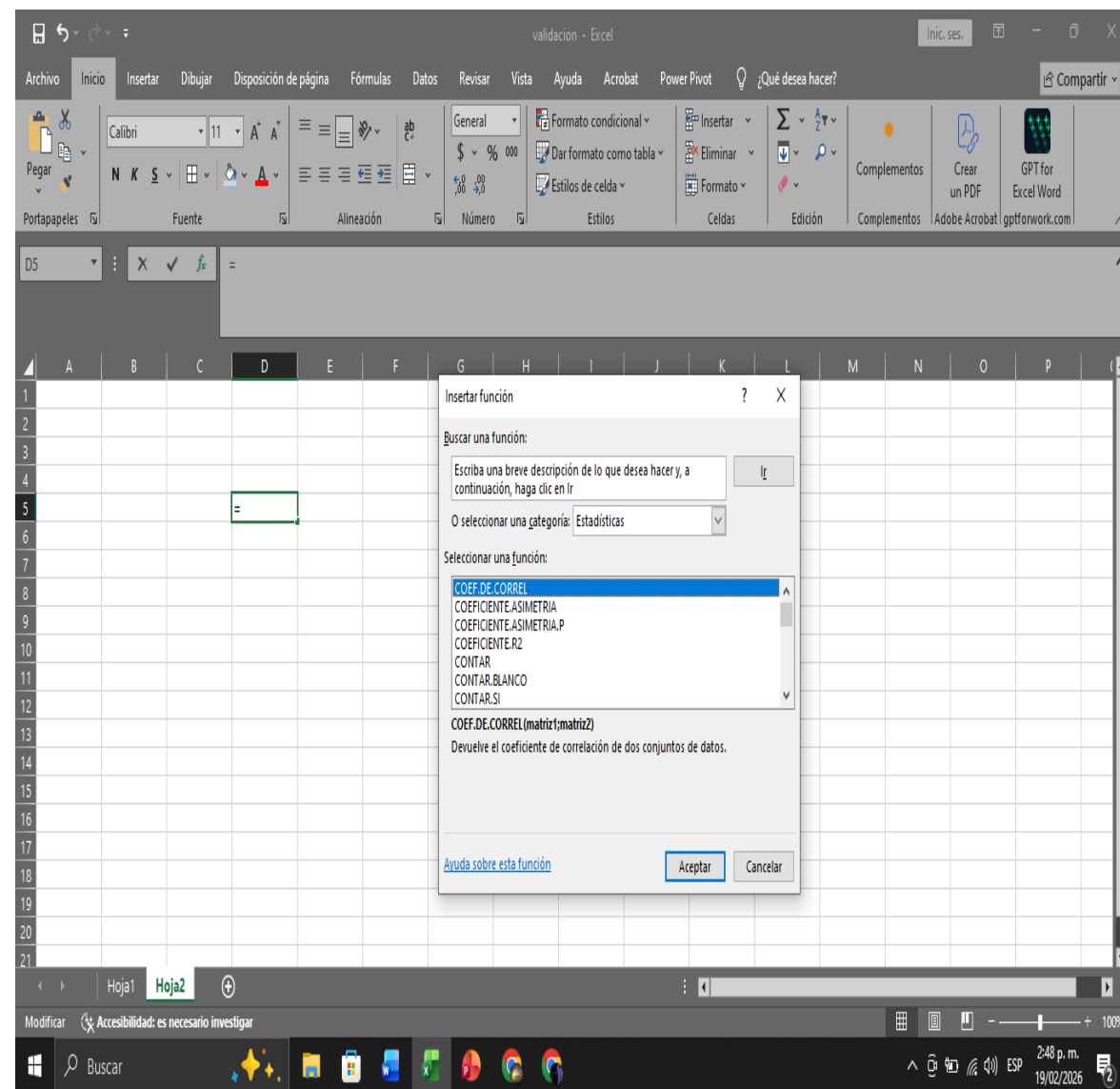
- Es posible filtrar variables **cuantitativas** con base alguna condición, por ejemplo: mayor que, menor que, mayor o igual a, etc.
- Para las variables **cualitativas** podemos filtrar con base a alguna categoría, por ejemplo: sólo género femenino.



Fórmulas.

Fórmulas.

- Las fórmulas en Excel, que siempre inician con un signo igual (=), permiten realizar cálculos matemáticos, lógicos y de búsqueda utilizando operadores básicos (+, -, *, /) o funciones integradas (como **SUMA**, **PROMEDIO**, **BUSCAR**).
- Son fundamentales para automatizar tareas, trabajando con referencias de celdas para actualizar resultados automáticamente.



Fórmulas - - - SI().

- Una de las fórmulas bastante útiles es: SI(), la cuál permite trabajar estructuras condicionales.
- Ejemplo: Construya una columna llamada “TiempoCod”, utilice la fórmula SI para crear una condición, donde si es mayor o igual 20 minutos indique “demorado” y menor de 20 minutos indique “normal”.

=SI(E2>=20;"demorado";"normal")

Fórmulas - - - BUSCAR()

La función BUSCAR() nos puede ayudar a codificar una variable numérica a una variable categórica. Por ejemplo, recodificar la variable tiempo en tres categorías:

- Eficiente --- (0;10]
- Normal --- (10;20]
- Demorado --- (20;30]

=BUSCAR(E2;\$J\$5:\$K\$7;\$L\$5:\$L\$7)

The screenshot shows an Excel spreadsheet with a table of data. The formula bar at the top displays the formula: `=BUSCAR(E2;J5:K7;L5:L7)`. The spreadsheet has columns labeled 'id', 'genero', 'edad', 'ClaseSoci', 'Tiempo', 'TiempoCod', and 'TiempoRecod'. The 'Tiempo' column contains numerical values, and the 'TiempoRecod' column contains categorical values. A small table in the background shows the mapping of 'Tiempo' values to 'TiempoRecod' categories:

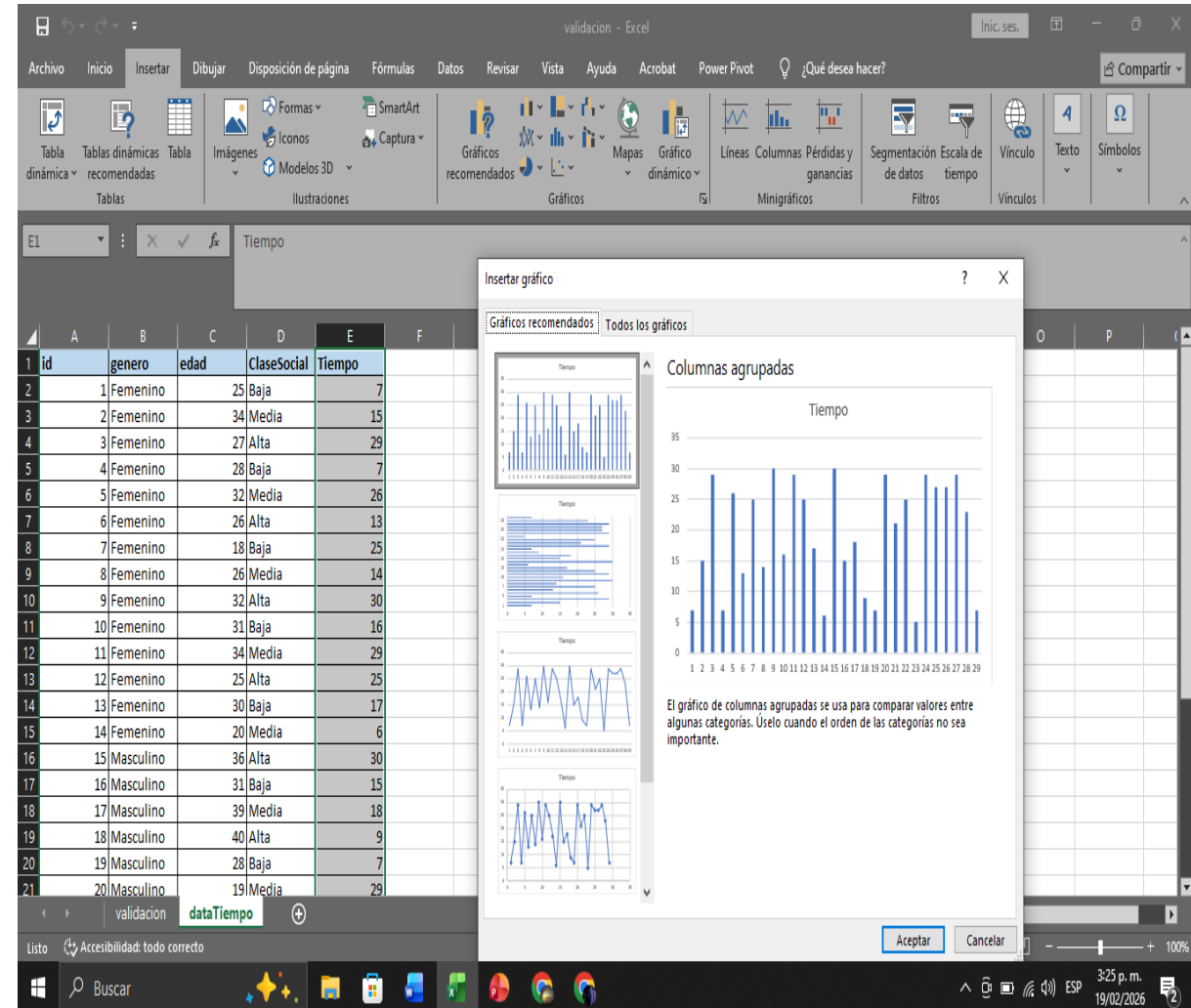
Tiempo	TiempoRecod
0	eficiente
10,0001	normal
20,00001	demorado

Gráficos.

Gráficos.

Una manera sencilla de realizar un gráfico es con la siguiente ruta:

1. Seleccionar una variable.
2. Insertar.
3. Gráficos recomendados
4. Todos los gráficos.



The screenshot shows the Microsoft Excel interface with the 'Insertar' tab selected. The 'Gráficos recomendados' section in the 'Insertar gráfico' dialog box is active, displaying several chart options. The 'Columnas agrupadas' chart is highlighted, showing a bar chart with the title 'Tiempo'. The dialog box also includes a description: 'El gráfico de columnas agrupadas se usa para comparar valores entre algunas categorías. Úsalo cuando el orden de las categorías no sea importante.'

id	genero	edad	ClaseSocial	Tiempo
1	Femenino	25	Baja	7
2	Femenino	34	Media	15
3	Femenino	27	Alta	29
4	Femenino	28	Baja	7
5	Femenino	32	Media	26
6	Femenino	26	Alta	13
7	Femenino	18	Baja	25
8	Femenino	26	Media	14
9	Femenino	32	Alta	30
10	Femenino	31	Baja	16
11	Femenino	34	Media	29
12	Femenino	25	Alta	25
13	Femenino	30	Baja	17
14	Femenino	20	Media	6
15	Masculino	36	Alta	30
16	Masculino	31	Baja	15
17	Masculino	39	Media	18
18	Masculino	40	Alta	9
19	Masculino	28	Baja	7
20	Masculino	19	Media	29

Tipos de Gráficos.

Algunos de los gráficos más utilizados en estadística son:

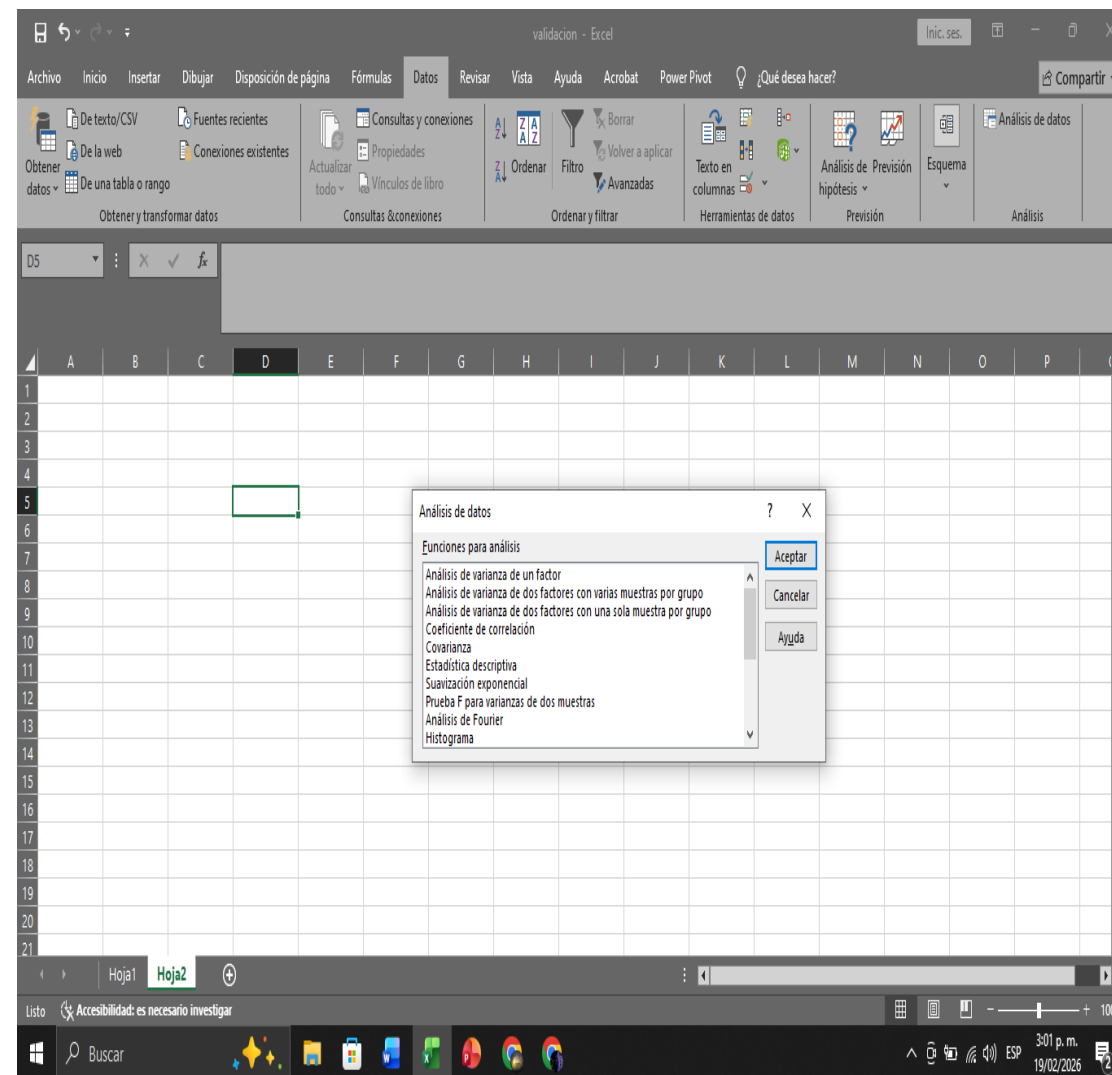
Gráfico	Equivalencia	Tipo de variable
Columnas	Diagrama de Barras	Variable cualitativa y cuantitativa discreta.
Circular	Diagrama circular o de sectores.	Variable cualitativa.
Líneas	Series de Tiempo	Variable cuantitativa eje y, en el eje x se coloca una variable temporal.
XY (Dispersión)	Diagrama de Dispersión	Variable cuantitativa eje y y otra cuantitativa en el eje x.
Histograma	Histograma	Variable cuantitativa continua.
Diagrama de cajas y bigotes.	Diagrama de cajas y bigotes (boxplot).	Variable cuantitativa eje y, variable cualitativa eje x.

Módulo análisis de datos.

Módulo análisis de datos.

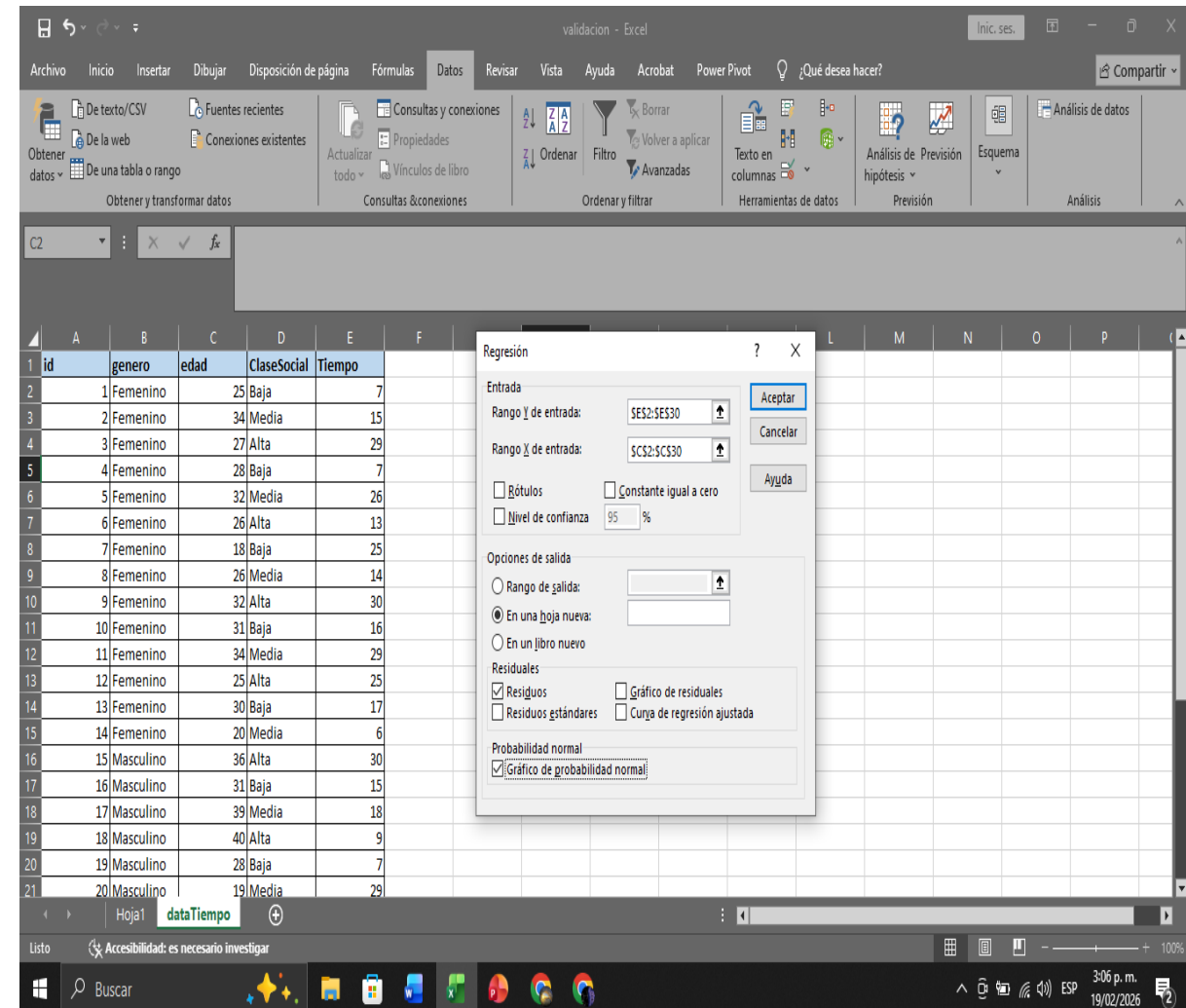
Para utilizar el módulo de análisis de datos es necesario activar desde la siguiente ruta:

1. Haz clic en archivo.
2. Selecciona opciones.
3. En el menú lateral, haz clic en complementos.
4. En la parte inferior, donde dice administrar, selecciona: Complementos de Excel, luego haz clic donde ir...
5. Marca la casilla herramientas para análisis.
6. Haz clic en aceptar.



Módulo de análisis de datos.

- Una herramienta útil dentro del módulo, es la herramienta de Regresión.
- Por ejemplo, podemos realizar una regresión lineal del Tiempo en función de la Edad, con el fin de analizar si existe una correlación entre la Edad y el Tiempo de espera en la sucursal del banco.



Módulo de análisis de datos.

Explora las siguientes herramientas:

- Generación de números aleatorios > distribución normal.
- Muestra > 5 muestras de una variable.
- Estadística descriptiva > selecciona alguna variable.

¡¡GRACIAS!!

