

4.1 Uso de herramientas para la representación de datos

“Las PC se han convertido en la herramienta más poderosa que hemos creado. Son herramientas de comunicación, de creatividad y pueden ser modeladas por sus usuarios.” Bill Gates

4.1.1 Introducción

La aparición de herramientas informáticas accesibles para el procesamiento de la información jugó un rol muy importante en el vertiginoso crecimiento del desarrollo de la humanidad durante el siglo pasado. Esto ha sido gracias a las facilidades que estas tecnologías brindan para analizar muchos datos en relativamente poco tiempo. Dentro de este tipo de herramientas entran las conocidas hojas de cálculo, inventadas por Daniel Bricklin, que es precisamente conocido como el *padre de las hojas de cálculo electrónicas*. Bricklin cursaba apenas su primer año del programa de *Master of Business Administration* de la Universidad de Harvard cuando, a raíz de un error cometido en un cálculo inicial mientras procesaba números, tuvo un pensamiento:

Si tan solo tuviese un pedazo de papel mágico en el cual cambiara un número al inicio de un conjunto de cálculos, y esto provocara que todos los números se recalcularan automáticamente. No me molestaría repetir los cálculos la primera vez, pero sí la segunda, tercera, o centésima vez (...) si tan solo tuviera una hoja de cálculo electrónica. (Bricklin, 1978).

Su idea se convirtió en VisiCalc, la primera hoja de cálculo y la aplicación fundamental que hizo que las computadoras dejaran de ser solo un *hobby* para convertirse también una herramienta en los negocios y en las empresas.

Las hojas de cálculo no son más que un tipo de archivo electrónico en el cual es posible manipular datos cualitativos o cuantitativos que se disponen en forma de tablas. Es posible, entre otras cosas, utilizar una hoja de cálculo para representar gráficamente diferentes tipos de variable o para calcular distintos parámetros estadísticos. Por otra parte, también es posible, con gran facilidad, convertir un grupo de datos iniciales en parámetros de mayor interés para el análisis, según el problema en cuestión. En adición, hay que tener en cuenta que en estadística resulta de extrema importancia mantener un orden lógico en la tabulación de las observaciones para poder comenzar cualquier análisis y representación gráfica.

Microsoft Excel es una herramienta para la creación de hojas de cálculo que es desarrollada por la empresa Microsoft y resulta ampliamente utilizada en nuestros días para el desempeño de múltiples tareas. Excel ofrece una plataforma amigable para la construcción de tablas de datos y el desarrollo

La hoja de cálculo de Microsoft Excel permanece como una herramienta idónea para la realización de análisis estadístico, tanto para estudiantes, usuarios principiantes como para maestros y usuarios avanzados.

de gráficos para la representación de estos datos, así como para el uso de diferentes fórmulas matemáticas y lógicas. Adicionalmente, esta herramienta incluye un comando para el análisis de datos que se puede utilizar para análisis estadísticos más complejos y avanzados. Es por ello que esta lectura tiene como objetivo mostrar y ejemplificar el uso de la herramienta Microsoft Excel para la representación y el análisis estadístico de datos. Para ello se expondrán diversas funcionalidades que presenta dicho programa.

4.1.2 Representación de datos mediante tablas

Características generales de Microsoft Excel

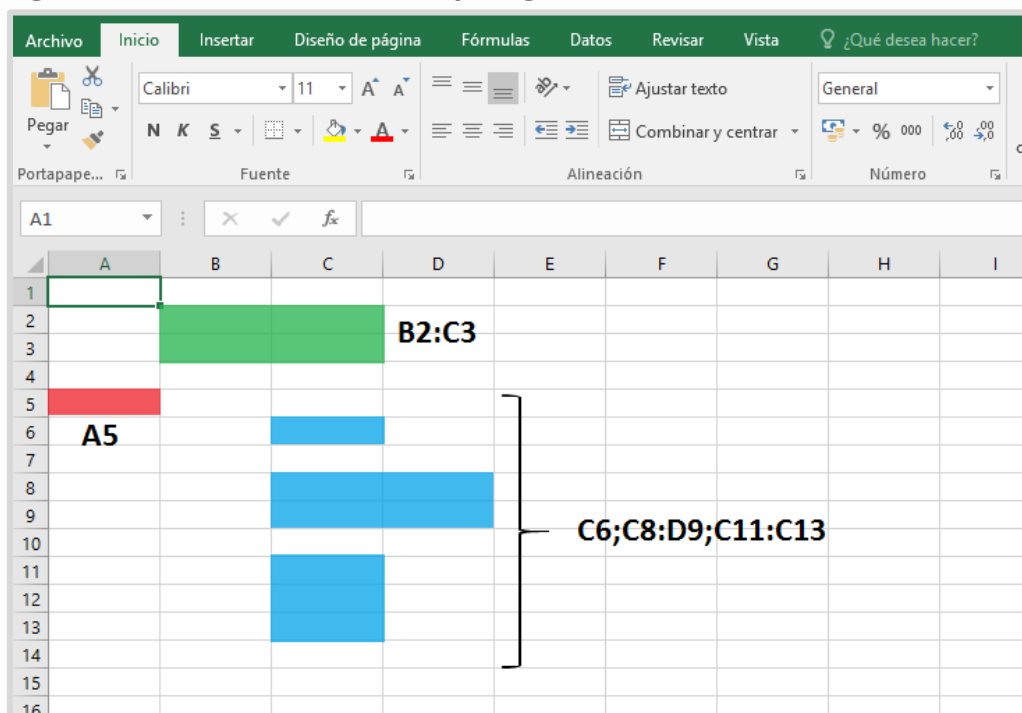
Como se mencionó anteriormente, la estructura básica de una hoja de cálculo son las tablas de datos, que están compuestas por celdas. La *celda* es la unidad elemental de almacenamiento de información y es donde se insertan tanto los datos como las fórmulas que determinan cálculos matemáticos. Por otra parte, a veces es necesario realizar diferentes operaciones, no sobre una única celda, sino sobre un grupo de celdas específico, conocido como **rango de celdas**, que puede ser continuo o discontinuo. Estos conceptos en Microsoft Excel se identifican de la siguiente forma:

Celda: se identifica por la letra de la columna a la cual pertenece, seguida por el número de la fila donde está ubicada. Por ejemplo: **A5** (ver Figura 1).

Rango de celdas continuas: se identifica por la primera celda, el caracter dos puntos (:) y la última celda. Por ejemplo: **B2:C3** (ver Figura 1).

Rango de celdas discontinuas: se identifica por las celdas y rangos de celdas continuas, separados por el carácter punto y coma (;). Por ejemplo: **C6;C8:D9;C11:C13** (ver Figura 1).

Figura 1: Identificación de celdas y rangos de celda en Microsoft Excel



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

Los ficheros que se crean con Microsoft Excel se denominan *libros*, y pueden contener tantas hojas como la disponibilidad de memoria permita, las cuales pueden ser hojas de cálculo u hojas de gráficos. Cada hoja puede tener 16,384 columnas, identificadas por letras (hasta XFD), y 1,048,576 filas, identificadas por números ([más información sobre los límites de Microsoft Excel](#)). Cada dato debe colocarse en una celda específica, y estas pueden almacenar muchos tipos de datos, que van desde números y fechas hasta textos y datos personalizados, en dependencia de la información que se desee guardar.

Funciones en Microsoft Excel y su uso en la Estadística

Muchas veces resulta de interés crear nuevas variables a partir de relacionar valores almacenados en las hojas de cálculo. Para ello es posible hacer uso de las fórmulas y funciones.

Las **fórmulas** son ecuaciones que efectúan cálculos numéricos o lógicos, se escriben en las celdas o en la Barra de Fórmulas (ubicada en la parte superior) y siempre comienzan por el símbolo igual a (=). Las fórmulas pueden contener funciones, referencias a celdas o rango de celdas (mediante sus respectivos identificadores); valores constantes, sean numéricos o de texto; y operadores matemáticos o de comparación (suma, resta, multiplicación, mayor que, etc.).

Por otra parte, las **funciones** son un tipo de fórmulas ya definidas y que ejecutan cálculos a partir de valores de entrada específicos denominados argumentos. En Excel existe una amplia variedad de funciones que pueden utilizarse para ejecutar operaciones, ya sean simples o complejas. Los argumentos pueden ser números, texto, valores lógicos (como *verdadero* o *falso*), matrices o referencias de celda. También podrían ser constantes, fórmulas u otras funciones, o bien no tener argumentos, como por ejemplo la función HOY (), que muestra la fecha. Las funciones podrían ser creadas de forma completamente personalizada mediante el uso de Macros, contenido que escapa de los objetivos de este curso.


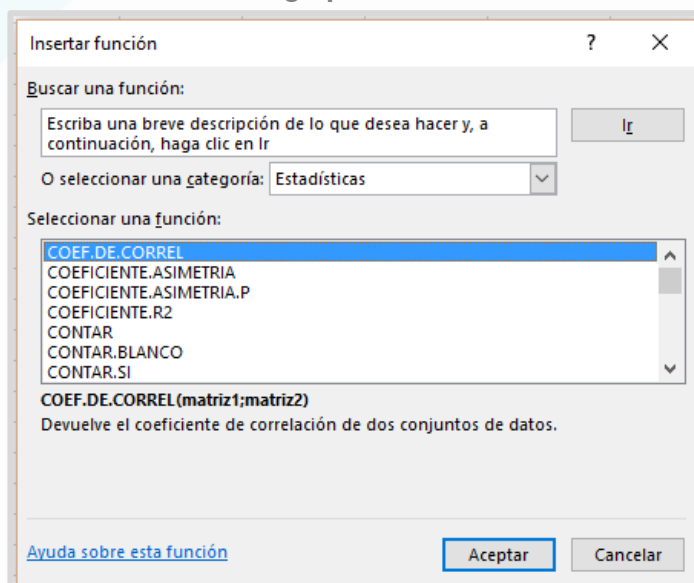
Las funciones son adicionadas mediante el botón ubicado justo a la izquierda de la Barra de Fórmulas (), el cual abre una ventana de diálogo (ver Figura 2) que permite buscar la función deseada, o bien seleccionarla de una lista a partir de elegir una categoría específica. A continuación se muestra un resumen de las funciones de mayor utilidad en el campo de la estadística descriptiva, las cuales pueden encontrarse en la categoría “Estadísticas” (más información sobre las funciones de Excel).

Figura 2: Ventana de diálogo para insertar una función en Microsoft



Excel

Fuente: captura de pantalla del software Microsoft Excel (Microsoft, 2016).

- **DESVPROM:** devuelve el promedio de las desviaciones absolutas de la media de los puntos de datos.
- **PROMEDIO:** devuelve el promedio de sus argumentos.
- **PROMEDIOA:** Devuelve el promedio de sus argumentos, incluidos números, texto y valores lógicos.
- **PROMEDIO.SI:** Devuelve el promedio (media aritmética) de todas las celdas de un rango que cumplen unos criterios determinados.
- **PROMEDIO.SI.CONJUNTO:** Devuelve el promedio (media aritmética) de todas las celdas que cumplen múltiples criterios.
- **CONTAR:** Cuenta cuántos números hay en la lista de argumentos.
- **CONTARA:** Cuenta cuántos valores hay en la lista de argumentos.
- **CONTAR.BLANCO:** Cuenta el número de celdas en blanco de un rango.
- **CONTAR.SI:** Cuenta el número de celdas, dentro del rango, que cumplen el criterio especificado.
- **CONTAR.SI.CONJUNTO:** Cuenta el número de celdas, dentro del rango, que cumplen varios criterios.
- **COVARIANZA.P:** Devuelve la covarianza, que es el promedio de los productos de las desviaciones para cada pareja de puntos de datos.
- **COVARIANZA.M:** Devuelve la covarianza de ejemplo, que es el promedio de las desviaciones de los productos para cada pareja de puntos de datos en dos conjuntos de datos.
- **DESVIA2:** Devuelve la suma de los cuadrados de las desviaciones.
- **MEDIA.GEOM:** Devuelve la media geométrica.

- **MEDIA.ARMO:** Devuelve la media armónica.
- **CURTOSIS:** Devuelve la curtosis de un conjunto de datos.
- **MEDIANA:** Devuelve la mediana de los números dados.

(Sitio Web Soporte de Microsoft Office, 2016)

Organización de la información

Cuadros estadísticos

Para una correcta comprensión y análisis de los datos almacenados resulta sumamente importante la organización de la información dentro de las tablas. Una forma muy útil de hacer esto es mediante el uso de cuadros estadísticos. “Un *cuadro estadístico* es un arreglo matricial que contiene las variables con sus correspondientes categorías y los datos observados de manera concisa, reflejando de forma fácil la relación o comparación entre las variables para un estudio determinado” (Bazán Baca, 2010, p. 50). Según el número de variables que se presentan, los cuadros pueden ser clasificados en unidimensionales (una sola variable), bidimensionales (dos variables) o multidimensionales (más de dos variables).

Las partes principales de un cuadro estadístico son: **número**, **título**, **encabezado**, **columna matriz**, **cuerpo** y **pie**. La ubicación convencional de cada uno de ellos se puede apreciar en la Figura 3.

Figura 3: Partes generales de un cuadro estadístico

Número	Título		
Columna matriz	Encabezado		
		Datos	
Pie			

Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

- **Número:** si en un estudio o investigación en particular hay más de un cuadro, es necesario enumerarlos para diferenciarlos. En este caso se suele anteponer la palabra *cuadro* o *tabla* y, a continuación, colocar el número.
- **Título:** el título indica de manera clara y precisa el contenido del cuadro. Un título completo es aquel que responde a preguntas como ¿qué estamos observando?, ¿cómo se clasifican los datos?, ¿cuándo se realizó la observación? o ¿dónde se realizó la observación?

- **Encabezado:** contiene los nombres de las categorías en que se subdividan y clasifiquen las variables observadas.
- **Columna matriz:** contiene las diferentes clasificaciones según como se midan las variables y sus correspondientes categorías o rangos. Normalmente suelen estar ordenados por algún criterio (geográfico, importancia, alfabético, etc.).
- **Cuerpo:** contiene celdas donde se almacenan los datos recogidos.
- **Pie:** parte inferior del cuadro, donde se colocan las notas, llamadas y la fuente de los datos.

La Tabla 1 muestra un ejemplo de un cuadro estadístico que relaciona diversos parámetros económicos de algunos países latinoamericanos (ver Lectura 1 de este curso).

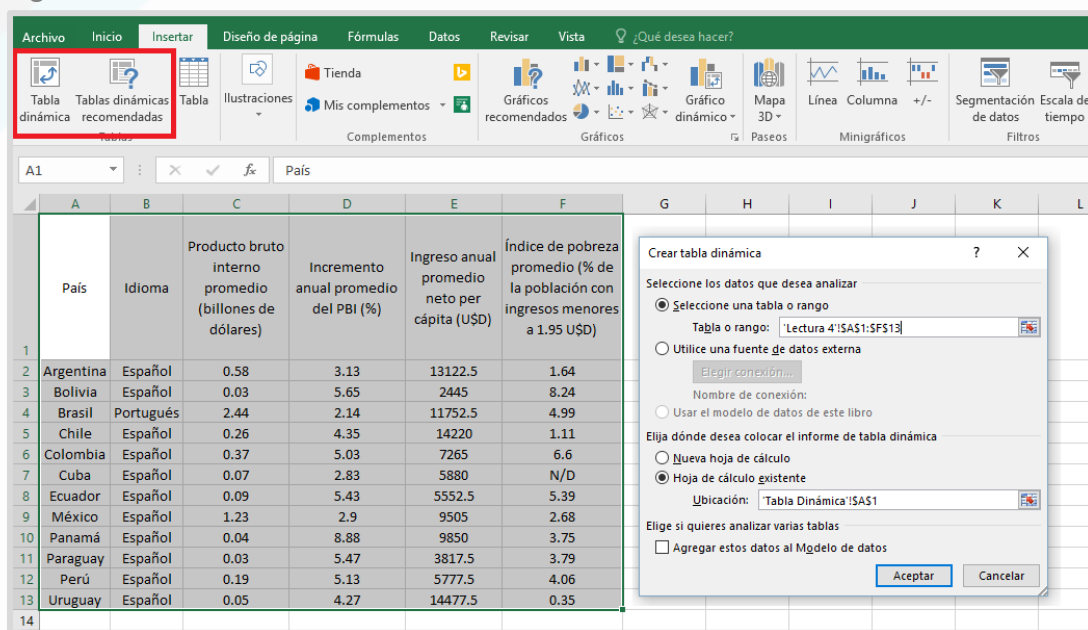
Una herramienta de Excel que resulta de extrema utilidad para la construcción de este tipo de representación y organización de datos son las *tablas dinámicas*, ya que al ser interactivas permiten modificar la presentación de los datos. Para construir una tabla dinámica en Excel utilizando los mismos datos de la Tabla 1, una vez insertados en una hoja de cálculo vamos al menú Insertar y seleccionamos el botón Tabla dinámica. Aparece entonces un cuadro de diálogo como el mostrado en la Figura 4, en el cual seleccionamos los datos que van a estar en la tabla y la ubicación que tendrá esta, que podría ser en otra hoja de cálculo (la selección de los datos podría ser automática si tenemos previamente seleccionada alguna celda que se encuentre en el conjunto de datos). La opción Agregar estos datos al Modelo de datos es necesario que sea señalada si estamos trabajando con varias tablas dinámicas relacionadas de alguna forma.

Tabla 1: Valores de variables económicas (promedio tomado del año 2011 al 2014) de algunos países latinoamericanos

País	Idioma	Producto bruto interno promedio (billones de dólares)	Incremento anual promedio del PBI (%)	Ingreso anual promedio neto per cápita (USD)	Índice de pobreza promedio (% de la población con ingresos menores a 1.95 USD)
Argentina	Español	0.58	3.13	13,122.5	1.64
Bolivia	Español	0.03	5.65	2,445	8.24
Brasil	Portugués	2.44	2.14	11,752.5	4.99
Chile	Español	0.26	4.35	14,220	1.11
Colombia	Español	0.37	5.03	7,265	6.60
Cuba	Español	0.07	2.83	5,880	N/D
Ecuador	Español	0.09	5.43	5,552.5	5.39
México	Español	1.23	2.90	9,505	2.68
Panamá	Español	0.04	8.88	9,850	3.75
Paraguay	Español	0.03	5.47	3,817.5	3.79
Perú	Español	0.19	5.13	5,777.5	4.06
Uruguay	Español	0.05	4.27	14,477.5	0.35

Fuente: Unesco Institute for Statistics, 2016, <https://goo.gl/jvUSLO>

Figura 4: Creación de una tabla dinámica en Microsoft Excel 2016



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

Una vez construida la tabla dinámica, aparecerá en la parte superior el menú Herramientas de tabla dinámica, donde es posible cambiar algunas propiedades, como por ejemplo los datos de entrada o del diseño de la tabla. Adicionalmente, aparecerá en la zona derecha (ver Figura 5) el menú de los Campos de la tabla dinámica, los cuales son necesarios de para poder observar los datos. Para ello es posible arrastrar los campos que se deseen añadir hacia una de las cuatro etiquetas posibles: Filtros, Columnas, Filas o Valores. Las tres últimas se corresponden con el posicionamiento que tendrán los datos dentro de la tabla, la etiqueta *Filtros* permite filtrar los datos para que solo muestre los que sean de un tipo específico. La otra forma de añadir campos consiste en señalarlos. De esta manera se ubican automáticamente en las etiquetas predeterminadas según el tipo de valores que tengan. Los campos no numéricos se agregan al área Filas; las jerarquías de fechas y horas se agregan al área Columnas, y los campos numéricos se agregan al área Valores. Para usar esta última variante se deben tener bien definidos los valores de las celdas que componen los campos. Por ejemplo, si se tiene una lista de valores numéricos, pero al menos uno de los valores es un texto (como en el caso del índice de pobreza en el ejemplo), automáticamente será ubicado en la etiqueta Filas y no en la de Valores. Es por ello que a veces resulta recomendable seleccionar las etiquetas de forma manual.

Figura 5: Tabla dinámica de ejemplo creada a partir de los datos de la Tabla 1

Idioma	País	Producto bruto interno promedio (billones de dolares)
Español	Argentina	0.58
Español	Bolivia	0.03
Español	Chile	0.26
Español	Colombia	0.37
Español	Cuba	0.07
Español	Ecuador	0.09
Español	México	1.23
Español	Panamá	0.04
Español	Paraguay	0.03
Español	Perú	0.19
Español	Uruguay	0.05
Español	Total general	2.94

Campos de tabla di...

Seleccionar campos para agregar al informe:

Buscar

- País
- Idioma
- Producto bruto interno promedi...
- Incremento anual promedio del PB...
- Ingreso anual promedio neto per c...
- Índice de pobreza promedio (% de ...)

Arrastrar campos entre las áreas siguientes:

FILTROS

Idioma

COLUMNAS

FILAS

País

VALORES

Producto brut...

Aplazar actualización... **ACTUALIZAR**

Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

La Figura 5 presenta la tabla dinámica creada en el ejemplo, donde se observa el PBI por países con un filtro según el idioma que muestra solo aquellos países donde se habla español. El uso de las tablas dinámicas dentro de la estadística resulta muy útil, tanto para organizar y presentar los datos, como para poder filtrar, entre mucha información, aquella que sea de nuestro interés. Para más información sobre el uso de tablas dinámicas es posible consultar el sitio de soporte de Microsoft.

Distribución de frecuencias

Como ya se mencionó en la primera lectura de este curso, otra de las formas comunes de resumir y organizar datos estadísticos es mediante la construcción de tablas de frecuencia. Los diferentes tipos y usos de las tablas de frecuencias ya fueron discutidos al inicio del curso. No obstante, resulta importante conocer de qué forma pueden ser construidas estas tablas, lo cual resulta relativamente fácil de hacer utilizando Microsoft Excel. Para ello haremos uso nuevamente de los ejemplos mostrados en la primera lectura.

Para obtener una tabla de frecuencias a partir de datos cualitativos, como la preferencia del color de zapatos de una determinada muestra de personas (ver Tablas 3 y 5 de la Lectura 1), el primer paso consiste en determinar la frecuencia de repetición de cada valor cualitativo. Para ello primero es necesario listar todos los casos posibles, seleccionamos las celdas donde se ubicarán los valores de frecuencia y hacemos uso de la función CONTAR.SI a través del botón de Agregar función. Seleccionamos entonces el rango donde

están todos los valores a contar y, como criterio, ponemos los valores de la columna que tiene la lista de los diferentes casos posibles. Una vez hecho esto no damos en Aceptar, sino que presionamos la combinación de teclas CONTROL+MAYUS+ENTER, para que la función se aplique a todas las celdas de resultado. La Figura 6 resume este procedimiento para el ejemplo empleado.

Figura 6: Uso de la función CONTAR.SI para determinar frecuencia de datos cualitativos

The screenshot shows an Excel spreadsheet with a data table in columns A-E and rows 1-12. The data consists of color names. A red box highlights the range A1:E12. A blue box highlights the range G5:G11, which contains the criteria: Blanco, Negro, Marrón, Gris, Rojo, Verde, and Beige. A dialog box titled 'Argumentos de función' is open, showing the formula bar with '=CONTAR.SI(A1:E12;G5:G11)'. The 'Rango' is A1:E12 and the 'Criterio' is G5:G11. The result of the formula is 11. The dialog box also shows the formula bar with '=CONTAR.SI(A1:E12;G5:G11)' and the result '11'.

Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

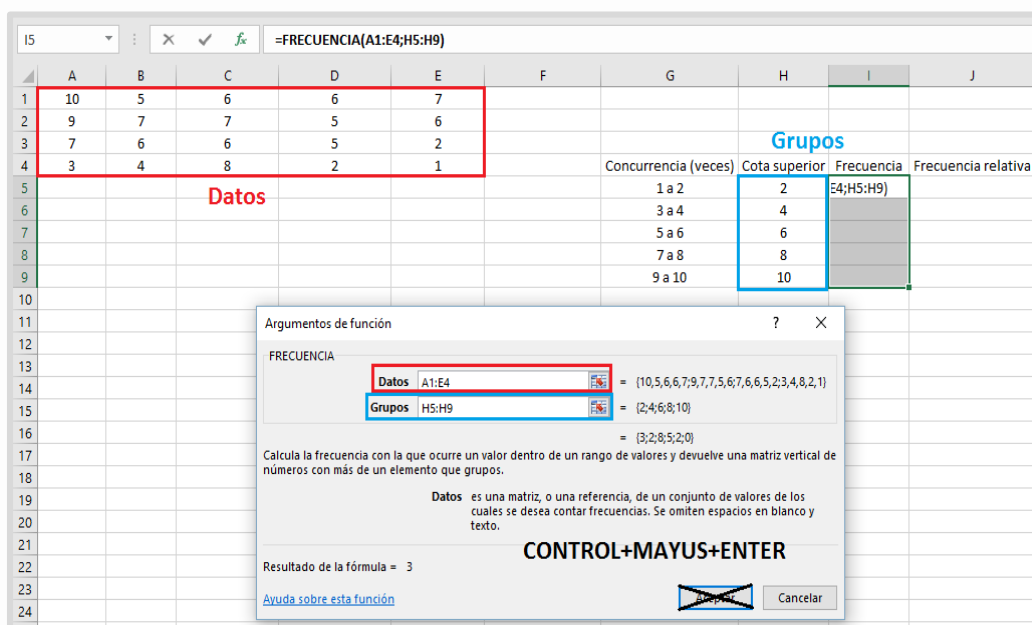
Para obtener las otras dos columnas de la tabla, que corresponden con la *frecuencia relativa* y la *frecuencia porcentual*, basta con calcular el total (60, en el ejemplo) y luego insertar las fórmulas correspondientes en cada caso, que están explicadas en la Lectura 1.

Como se mencionó en la primera Lectura, en el caso de datos cuantitativos se requiere comúnmente definir de forma previa un número de categorías y la amplitud de cada categoría para luego construir la correspondiente tabla de frecuencias. En Microsoft Excel existe una función muy útil para realizar esta separación en categorías, la función FRECUENCIA. Para mostrar su uso utilizaremos el ejemplo de la frecuencia de compra en un centro comercial de una región dada (ver Tabla 5 y 7 de la Lectura 1).

Para poder determinar los valores de frecuencias utilizando esta función, y una vez definidas las categorías con su respectivo rango numérico, es necesario construir una columna que comprenda las cotas superiores de cada rango. Posteriormente se seleccionan las celdas donde se ubicará el resultado y se inserta la función FRECUENCIA. En el campo Datos se añaden las celdas con los datos a contar, y en el campo Grupos se seleccionan las celdas de la columna correspondiente con la cota superior de cada rango. De forma similar al ejemplo anterior, no damos en Aceptar, sino que presionamos la combinación de teclas CONTROL+MAYUS+ENTER para que la función se aplique a todas las celdas de resultado. La Figura 7 ilustra este proceso. Los valores de las columnas correspondientes a la frecuencia relativa y la

frecuencia porcentual se determinan de forma similar a lo explicado para el caso de datos cualitativos.

Figura 7: Uso de la función FRECUENCIA para determinar frecuencia de datos cuantitativos



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

4.1.3 Representación de datos mediante gráficos

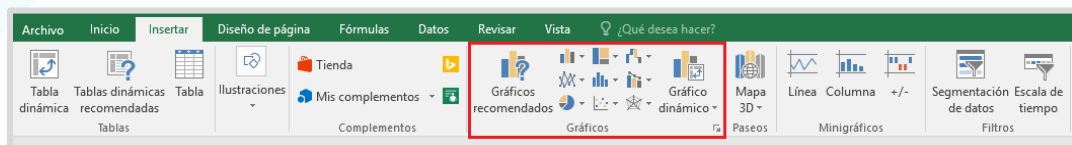
En la primera lectura se introdujo la noción del uso e importancia de los gráficos estadísticos para la visualización de los datos. El objetivo del uso de gráficos en la estadística consiste en poder mostrar a primera vista aspectos que resulten importantes de las variables en estudio, pues, como se dice comúnmente, una imagen vale más que mil palabras. Resulta válido aclarar que esta forma de presentación de la información requiere necesariamente de una tabla o cuadro estadístico previamente diseñado.

A modo general, y de forma similar a los cuadros estadísticos, las partes esenciales de un gráfico son: número, título, cuerpo, leyenda y fuente. Las partes del número, título y fuente son similares al caso de los cuadros estadísticos. El cuerpo es la parte central del gráfico, y allí se muestran representaciones geométricas que buscan establecer contrastes entre las variables y sus categorías. Por otra parte, la leyenda consiste en aclaraciones respecto de los distintos aspectos que pueden componer el cuerpo de un gráfico, expresados a través de colores, tramas, etcétera. Al igual que los cuadros estadísticos, los gráficos pueden ser unidimensionales, bidimensionales o pluridimensionales, según el número de variables que se presenten.

Microsoft Excel presenta una amplísima variedad de tipos de gráficos para analizar datos, ya sean cualitativos o cuantitativos. La construcción de cualquiera de estos gráficos es similar y de extrema sencillez en todos los casos. Se realiza seleccionando las celdas que contienen los datos que se desean graficar y luego se selecciona, en el menú Insertar, en la sección Gráficos, el tipo de gráfico que se desea realizar (ver Figura 8). Es válido



notar que la sección Gráficos se encuentra organizada por distintas categorías de gráficos, lo cual facilita muchas veces la elección correcta.

Figura 8: Insertar gráficos en Microsoft Excel 2016





Fuente: captura de pantalla del software Microsoft Excel (Microsoft, 2016), con modificación propia.

Representación de datos cualitativos

Entre los gráficos de variables cualitativas más usados tenemos los gráficos de pastel () y los gráficos de barras (), que fueron mostrados en la Lectura 1. El gráfico de pastel, también llamado gráfico de sectores, es útil para mostrar todas las categorías de una variable en un mismo círculo que esté particionado de forma proporcional al número de casos de cada categoría. De esta forma es posible clasificar las particiones con colores diferentes, por ejemplo, y así obtener información visual inmediata acerca de las variables. La Tabla 6 de la Lectura 1 muestra un ejemplo del uso de este tipo de gráficos mediante la utilización de los datos de la Tabla 3 de esa misma lectura. Por otra parte, el gráfico de barras es una representación rectangular en el plano cartesiano. Indica en el eje de las abscisas (X) la variable con sus categorías, y, en el eje de las ordenadas (Y), el número y/o el porcentaje de unidades de análisis que se presenta.

Representación de frecuencias

Entre los principales gráficos de variables cuantitativas se encuentran los ya mencionados gráficos de barras, los histogramas de frecuencias () y los polígonos de frecuencia (). El histograma de frecuencias es un tipo especial de gráfico de barras en el cual el ancho de cada barra estará relacionado con la amplitud de su categoría correspondiente. Así, tanto la altura de la barra como su ancho es información visual de rápido análisis. El polígono de frecuencias es un gráfico de líneas que une los puntos medios superiores de las barras del histograma con el fin de tener una visión global de la distribución de frecuencias cuando son suavizadas.

4.2 Uso de herramientas informáticas para el análisis de datos

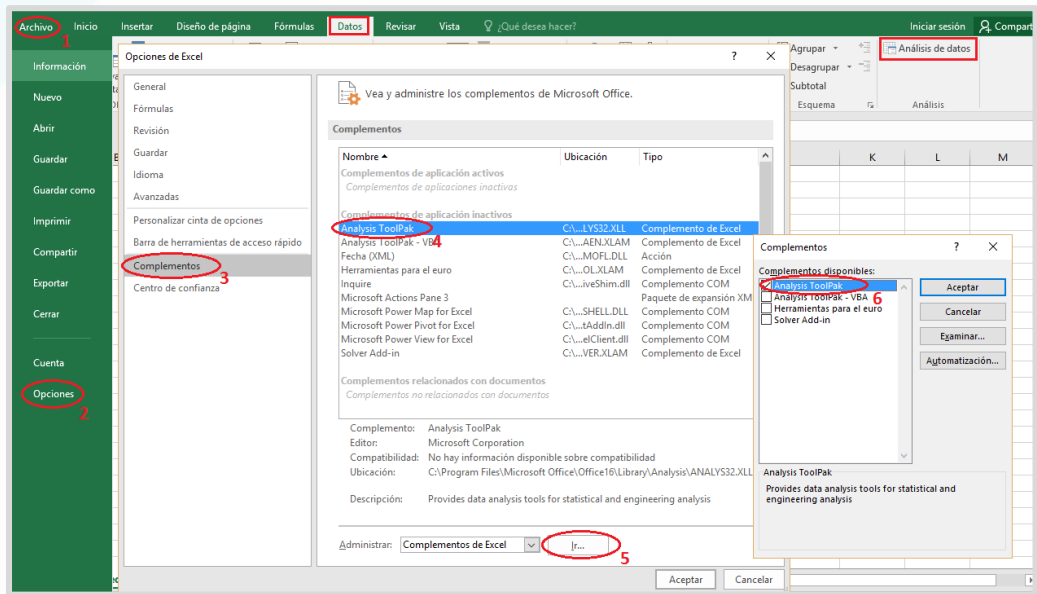
4.2.1 Herramienta de Microsoft Excel para análisis de datos

Para el análisis estadístico en específico, Microsoft Excel presenta una característica sumamente poderosa: las herramientas de Análisis de datos. Estas herramientas permiten ahorrar pasos en el desarrollo de análisis estadístico, pues, partiendo de los datos iniciales y de los parámetros de cada análisis, es posible directamente obtener tablas de resultados e incluso gráficos en algunos casos. Estas herramientas no suelen estar instaladas de forma predeterminada en Microsoft Excel, así que a veces es necesario activarlas previamente, proceso que en algunos casos podría solicitar el disco de instalación.

Para activar estas herramientas es necesario acudir al menú de Complementos (Archivo > Opciones > Complementos) y acceder (mediante el botón Ir...) al complemento Analysis ToolPak o Herramientas para análisis —dependiendo de la versión de Excel—, y activar la pestaña correspondiente. Una vez activada es posible acceder a esta herramienta en el menú Datos. La Figura 9 resume este proceso de activación.

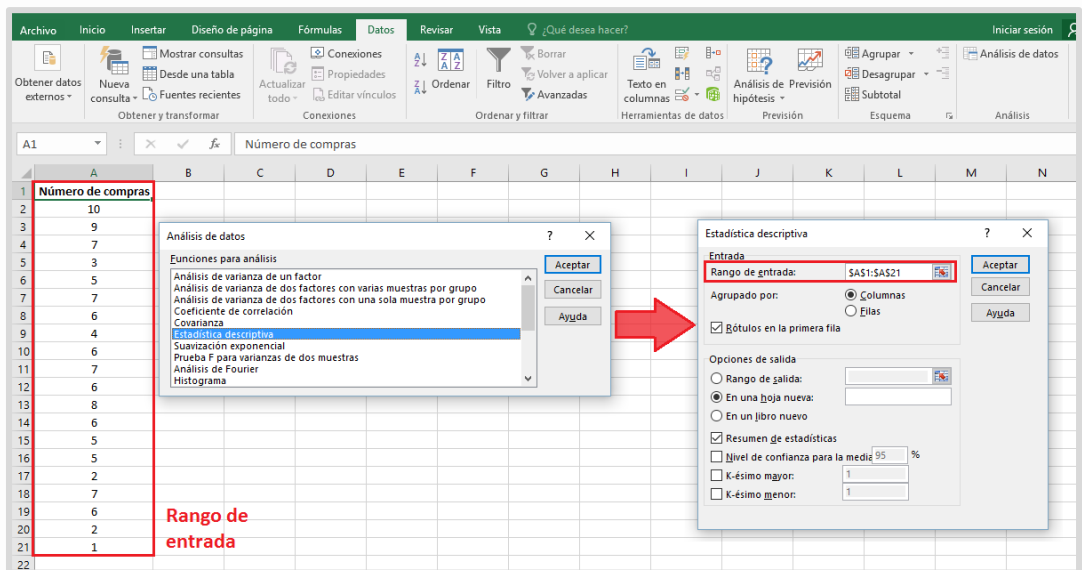
Esta herramienta presenta una elevada cantidad de facilidades para el trabajo en estadística, todas listadas en el cuadro de diálogo que se observa cuando accedemos a ella (ver ejemplo en las Figuras 10 y 12). Su uso es relativamente fácil (más información), pues solo consiste en completar los datos que se solicitan una vez elegido el tipo de análisis que se desea, lo cual normalmente no requiere de realizar cálculos o análisis previos. Mostraremos aquí un ejemplo de su uso dentro de la estadística descriptiva y para la obtención de gráficos. Para ello emplearemos los datos de la frecuencia de compra en un centro comercial, mostrados en la Lectura 1 y utilizados anteriormente. Las Figuras 10 y 11 muestran el proceso y el resultado, respectivamente, del análisis de estadística descriptiva para estos datos.

Figura 9: Activación de las herramientas para análisis en Microsoft Excel 2016



Fuente: captura de pantalla del software Microsoft Excel (Microsoft, 2016), con modificaciones propias.

Figura 10: Uso de la herramienta de Análisis de Datos de Microsoft Excel para Estadística descriptiva



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

Figura 11: Resultado del uso de la herramienta de Análisis de Datos de Microsoft Excel para Estadística descriptiva

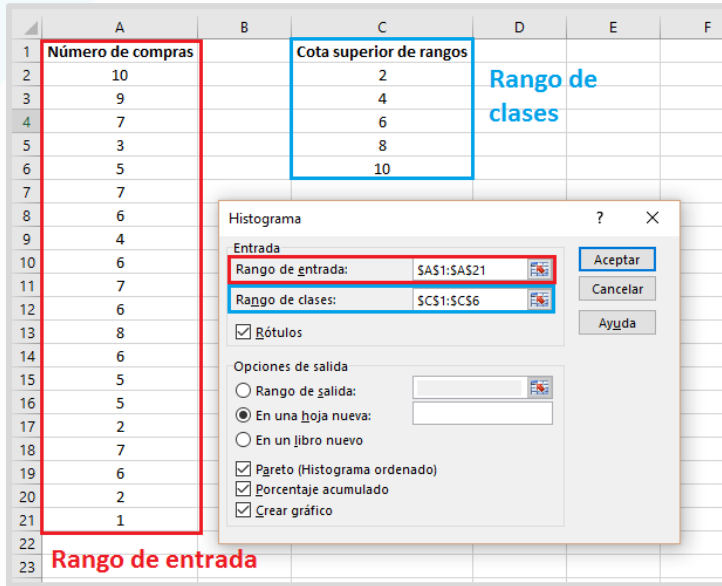
	A	B
1	<i>Número de compras</i>	
2		
3	Media	5.6
4	Error típico	0.52012144
5	Mediana	6
6	Moda	6
7	Desviación estándar	2.32605381
8	Varianza de la muestra	5.41052632
9	Curtosis	-0.10064052
10	Coefficiente de asimetría	-0.2944149
11	Rango	9
12	Mínimo	1
13	Máximo	10
14	Suma	112
15	Cuenta	20
16		

Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

En el cuadro de diálogo que aparece una vez elegida la opción de Estadística descriptiva, es necesario especificar el Rango de entrada, que no son más que las celdas que contienen los valores a analizar, así como la dirección en que están ordenados, sea en Filas o Columnas. Si dentro del rango de entrada se encuentra un título en la primera fila o columna (como en el caso del ejemplo) es necesario marcar la casilla Rótulos en la primera fila/columna. En las opciones de salida es posible especificar dónde mostrar los resultados, ya sea en un rango dentro de la misma hoja de la entrada, o en una hoja nueva, o incluso en un libro nuevo. La casilla Resumen de estadísticas es necesaria para que los resultados muestren una tabla con los distintos parámetros de estadística descriptiva resumidos. Adicionalmente, es posible especificar algunos otros parámetros de salida, como por ejemplo el nivel de confianza deseado para la media.

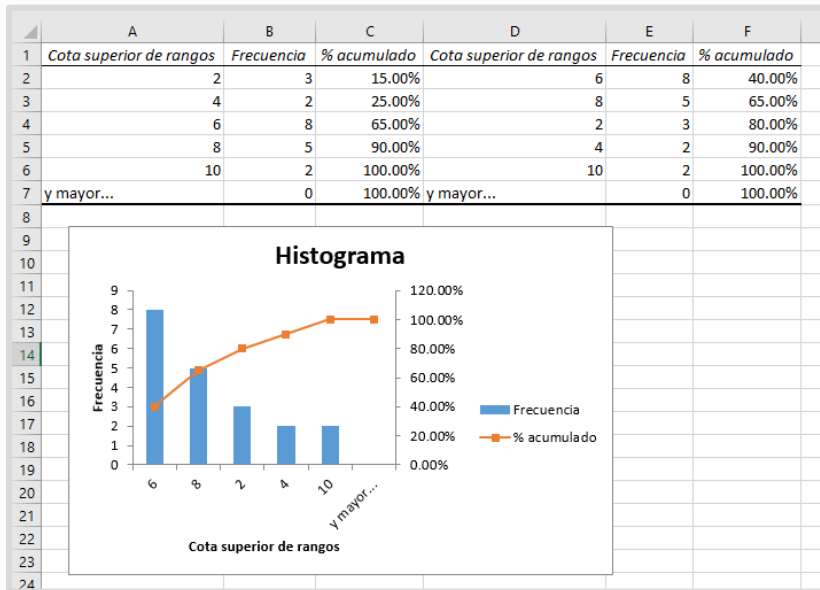
Otro ejemplo del potente uso de esta herramienta consiste en la obtención de tablas de frecuencias y gráficos de histogramas mediante la opción Histograma. En este caso es necesario también incluir los datos de los Rangos de clases, los cuales corresponderían con la cota superior de las categorías deseadas. Esta opción permite, no solo crear las tablas de frecuencia, sino también determinar los porcentajes acumulativos y mostrar los gráficos de histograma o de Pareto. El *gráfico de Pareto* es un gráfico muy utilizado cuando se quiere mostrar las categorías más representativas de una variable cualitativa en forma de frecuencia descendente. Es muy útil para tomar algunas decisiones, como en el caso de los controles de calidad en los que, por ejemplo, se busca controlar los defectos más importantes. Las Figuras 12 y 13 muestran respectivamente el proceso y el resultado de la construcción de histogramas para los datos del ejemplo anterior.

Figura 12: Uso de la herramienta de Análisis de Datos de Microsoft Excel para construcción de histogramas



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

Figura 13: Resultado del uso de la herramienta de Análisis de Datos de Microsoft Excel para construcción de histogramas



Fuente: elaboración propia a base de Microsoft Excel (Microsoft, 2016).

Referencias

Bazán Baca, J. F. (2010). *Texto de estadística computacional con R, Excel, Minitab y SPSS*. Perú: Universidad Nacional del Callao.

Bricklin, D. (1978). *Special short paper for the HBS advertising course*. USA: Harvard Business School.

Dretzke, B. (2012). *Statistics with Microsoft Excel* (5ª ed.). EE. UU.: Pearson.

Hernández González, S., y Cuevas Acosta, J. H. (2013). *Programas informáticos de uso libre y su aplicación en la enseñanza de la estadística*. Revista Investigación Operacional, 34(2), 166-174.

Microsoft Office. (2016). *Ayuda de Excel*. Recuperado de <https://support.office.com/es-ar/excel>

Quirk, T. J. (2016). *Excel 2016 for Business Statistics. A Guide to Solving Practical Problems*. USA: Springer.

Unesco Institute for Statistics. (2016). Unesco Institute for Statistics. Recuperado en septiembre 2016 de Unesco Institute for Statistics: <https://goo.gl/jvUSLO>