

<https://doi.org/10.47300/actasidi-unicyt-2022-28>

EXCEL COMO HERRAMIENTA DE ANALÍTICA EMPRESARIAL

Rodriguez-Rivas, José Gabriel

Tecnológico Nacional de México / Instituto Tecnológico de Durango

Durango, México

gabriel.rodriguez@itdurango.edu.mx

ORCID: 0000-0002-7031-5097

RESUMEN

La inteligencia de negocios utiliza herramientas de software especializadas en el tratamiento de los datos de las empresas, con capacidad para presentar reportes, cuadros de mando, tablas y gráficos de forma amigable. Las pequeñas y medianas empresas en ocasiones no tienen acceso a estas herramientas por su alto costo, o bien, por la necesidad de contar con gente especializada para su uso. Excel es una de las herramientas más utilizadas a nivel global y de fácil acceso para todas las empresas sin importar su tamaño o capacidad económica. El objetivo de este estudio fue desarrollar en Excel un cuadro de mando para el análisis y visualización de datos con segmentación de datos, y con capacidad de trabajar con grandes volúmenes de información y así realizar una mejor toma de decisiones basada en datos. El estudio se realizó bajo un diseño cuantitativo, de tipo no experimental y transversal. El alcance del estudio es descriptivo. Para la recolección de los datos se utilizó la técnica de datos secundarios. Como resultado, se realizó un cuadro de mandos con capacidad de segmentación de datos por diferentes variables, que permiten identificar las tendencias o hallazgos que están ocultos dentro de los datos y, además, con la posibilidad de trabajar con datos más allá de la capacidad que tiene Excel de mostrarlos en las celdas.

Palabras clave: Análisis de datos, Dashboard, Excel, Inteligencia de negocios

ABSTRACT

Business intelligence uses specialized software tools for processing company data, with the ability to present reports, dashboards, tables, and graphs in a user-friendly way. Small and medium-sized companies sometimes do not have access to these tools because of their high cost, or because of the need to have specialized people for their use. Excel is one of the most used tools globally and easily accessible to all companies regardless of their size or economic capacity. The objective of this study was to develop a dashboard in Excel for the analysis and visualization of data with data segmentation, with the ability to work with large volumes of information and thus make better decisions based on data. The study is conducted under a quantitative, non-experimental, and cross-sectional design. The scope of the study is descriptive. For data collection, the secondary data technique was used. As a result, a dashboard was obtained with data segmentation capacity by different variables that allow to identify the trends or findings that are hidden within the data and, with the possibility of working with data beyond the capacity that Excel must show them in the cells.

Keywords: Data analytics, Dashboard, Excel, Business intelligence

1. INTRODUCCIÓN

Los datos son un recurso valioso para las pequeñas y medianas empresas. Datos que se pueden extraer de diferentes fuentes como son los sistemas de cobro o puntos de venta, los sistemas de contabilidad, redes sociales, entre otros. Los datos por si solos no dicen nada, necesitan ser procesados para obtener valor de ellos y generar conocimiento. La pirámide DIKW (Data Information Knowledge Wisdom) es un modelo jerárquico usado en la gestión del conocimiento para ilustrar las transiciones que los datos experimentan hasta obtener suficiente valor para la toma de decisiones informada (Toro, 2022).

La creciente necesidad de las pequeñas y medianas empresas (PYMES), de contar con herramientas de análisis de datos que les permita visualizar en todo momento los indicadores y tendencias de las operaciones de la empresa, hace inevitable que se realicen los esfuerzos necesarios para adquirir las herramientas de análisis de datos que estén a su alcance. Estas herramientas de software les permitirá identificar áreas de oportunidad para realizar una toma de decisiones basada en datos de una manera más eficiente, de acuerdo con la información obtenida en el análisis de los datos.

En el mercado existen muchas herramientas para el análisis de datos, algunas muy costosas, lo que las hace inalcanzables para las PYMES. Algunas otras herramientas tienen versiones libres y gratuitas, pero hay poca información de cómo usarlas e incluso poca disponibilidad de tutoriales o foros de ayuda. En este sentido, Excel es una de las herramientas de ofimática más utilizada a nivel mundial. Es utilizada indistintamente por contadores, administradores, ingenieros, estudiantes, entre otros. Adicionalmente, Excel es una herramienta que en la mayoría de las veces viene instalado en las computadoras que se venden comúnmente. Ante este escenario, surgen las siguientes preguntas de investigación:

- ¿Qué ventajas ofrece Excel para el procesamiento, análisis y visualización de datos?
- ¿Existen limitantes en Excel para el análisis de grandes volúmenes de datos?

Objetivo general

Desarrollar en Excel un cuadro de mando para el análisis y visualización de datos con segmentación de información, con capacidad de trabajar con grandes volúmenes de datos para una mejor toma de decisiones basada en datos.

2. MARCO CONCEPTUAL

De acuerdo con Kielstra et al. (2007), los ejecutivos de las PYMES toman decisiones de negocios basándose en los datos que tienen a su disposición. Estos datos muchas veces provienen de la experiencia, intuición e incluso en la tradición. En este sentido, Steiber y Alänge (2016), identifican el término HIPPO (highest paid persons in the organization), o las personas mejor pagadas en la organización. Muchas de las ocasiones estas personas son las que toman las decisiones basándose en su rango o posición, y no precisamente porque ellos tengan a su disposición la mejor información.

Por otro lado, el análisis de datos es el proceso de identificar, limpiar, transformar y modelar los datos para detectar información significativa y útil. Después, los datos se convierten en una historia a través de informes, tableros o cuadros de mando (dashboard), con el fin de realizar el proceso crítico de la toma de decisiones (Microsoft, 2022 a).

Asimismo, las herramientas de inteligencia de negocios o business Intelligence (BI), son herramientas de software especializadas en el tratamiento de los datos de negocios, con capacidad de presentar reportes, cuadros de mando, tablas y gráficos de forma amigable, permitiendo al usuario acceder a datos históricos y actuales (IBM, 2022)

Como lo menciona Gartner (2022), las plataformas de análisis e inteligencia de negocios (ABI) permiten a los usuarios menos técnicos, incluidos los empresarios, modelar, analizar, explorar,

compartir y administrar datos, y colaborar y compartir hallazgos, habilitados por las tecnologías de Información (TI) y aumentados por inteligencia artificial (IA).

Barón et al. (2021), identificaron que el sector empresarial es donde más se usan las herramientas de inteligencia de negocios, seguido del sector salud, educación, tecnología y el sector turismo.

De acuerdo con las especificaciones, Excel tiene como límite 1,048,576 total de filas. Esta limitante para el análisis de datos representa un obstáculo, sobre todo cuando se tiene que trabajar con datos con un número mayor a esa restricción. Sin embargo, en sus versiones 2010 y 2013 Excel ofrece el complemento Power Query que permite trabajar con archivos mayores a esa cantidad. Para versiones 2016 y superiores se incorporó en la pestaña de datos. Power Query es una tecnología de conexión de datos que permite conectar y combinar orígenes de datos para satisfacer las necesidades de análisis de datos (Microsoft, 2022 b).

3. MATERIALES Y MÉTODOS

El estudio se realizó bajo un diseño cuantitativo, de tipo no experimental y transversal. El alcance del estudio es descriptivo. Para la recolección de los datos se utilizó la técnica de datos secundarios, la cual involucra la utilización de registros públicos y archivos electrónicos (Briones, 2002). Los datos utilizados se descargaron de portales de datos abiertos. Se utilizó dos archivos. El primer archivo se llama “*Supermarket sales*” y contiene un registro histórico de ventas de 3 supermercados diferentes. Tiene 1,000 registros con datos ficticios y con información acerca de la sucursal, ciudad, sexo, tipo de cliente, precio unitario, modalidad de pago, entre otras variables. Se descargo del sitio: <https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales>.

El segundo archivo (220830COVID19MEXICO.csv) contiene información referente a las pruebas realizadas por COVID en México, y tiene 5,462,866 registros. Se descargó del portal de datos abiertos del gobierno de México. La información más actual está disponible en el sitio <https://datos.gob.mx/>. También, se puede descargar las bases históricas desde el sitio: <https://www.gob.mx/salud/documentos/datos-abiertos-bases-historicas-direccion-general-de-epidemiologia>. Este archivo sirvió para verificar que Excel puede trabajar con grandes cantidades de datos

El estudio se desarrolló en 2 partes. En la primera, se usó el archivo de ventas de supermercado para realizar el cuadro de mandos. En la segunda, se utilizó Power Query para abrir el archivo COVID19-MX y usar tablas dinámicas para analizar la información.

4. RESULTADOS Y DISCUSIÓN

Antes de realizar el cuadro de mando, primero se deben considerar algunos aspectos importantes. Por ejemplo, el tipo de gráficos a utilizar, es decir, que sean los más adecuados para visualizar la información, y aquí es donde debemos conocer cuantas variables se van a mostrar, cuantos puntos de datos se van a trazar, si se están comparando categorías o si son datos temporales, entre otros aspectos importantes.

En una segunda hoja se usaron tablas dinámicas para generar los resultados intermedios y así alimentar el cuadro de mandos. Por ejemplo, para mostrar el mejor y el peor mes de ventas, se utilizó 2 tablas dinámicas donde se usó la opción de filtros de valor, enseguida los diez superiores y 10 inferiores limitándolos a un único valor.

Con la variable fecha de compra se usó 2 gráficos de líneas. Uno para ver las ventas totales por mes y ver las tendencias de las compras con respecto al tiempo. En el segundo, las ventas totales por mes desglosado por género. De esta manera, se puede obtener cual es el sexo que más compra en un determinado mes.

También, se agregó un gráfico de dispersión que mostrará la relación que pudiera existir entre las ventas y la hora. Con este gráfico se obtienen las horas donde hay más ventas y así dedicar

más personal atendiendo este departamento. También, se puede obtener que horas del día son las menos concurridas y de esta manera utilizar al personal en otras actividades.

Se incluyó un gráfico de anillo para comparar los métodos preferidos de pago, mostrando las proporciones de cada uno de ellos. De la misma forma, se usó un gráfico de anillo para mostrar el tipo de cliente. Adicionalmente, se añadió un gráfico de barras horizontales para mostrar las ventas agrupadas por categoría y sexo, mostrando cuáles son las categorías de productos preferidos por sexo, lo que permitirá conocer donde se tienen que realizar esfuerzos adicionales para incrementar las ventas de determinados productos.

Adicionalmente, se agregó una segmentación de datos. La segmentación de datos proporciona botones en los que se puede activar/desactivar para filtrar tablas, tablas dinámicas e incluso gráficos. La segmentación ofrece la posibilidad de mostrar solamente la información deseada, facilitando la visualización y la comparación de diferentes variables. Por ejemplo, en la figura 1, se observan 3 segmentadores que sirven para filtrar los datos por ciudad, tipo de cliente y por sexo.

Figura 1.
Dashboard ventas de supermercado.



Los segmentadores se pueden combinar. Por ejemplo, se puede segmentar la información por la ciudad de Yangón y la ciudad de Mandalay, actualizándose los gráficos y totales de esas 2 ciudades. También se actualizarán los valores de los tarjetones que muestran las ventas totales y el volumen de compras. De igual forma, si se desea conocer las ventas de la ciudad de Yangón, del sexo femenino y que sean miembros, bastara con presionar o activar los botones correspondientes.

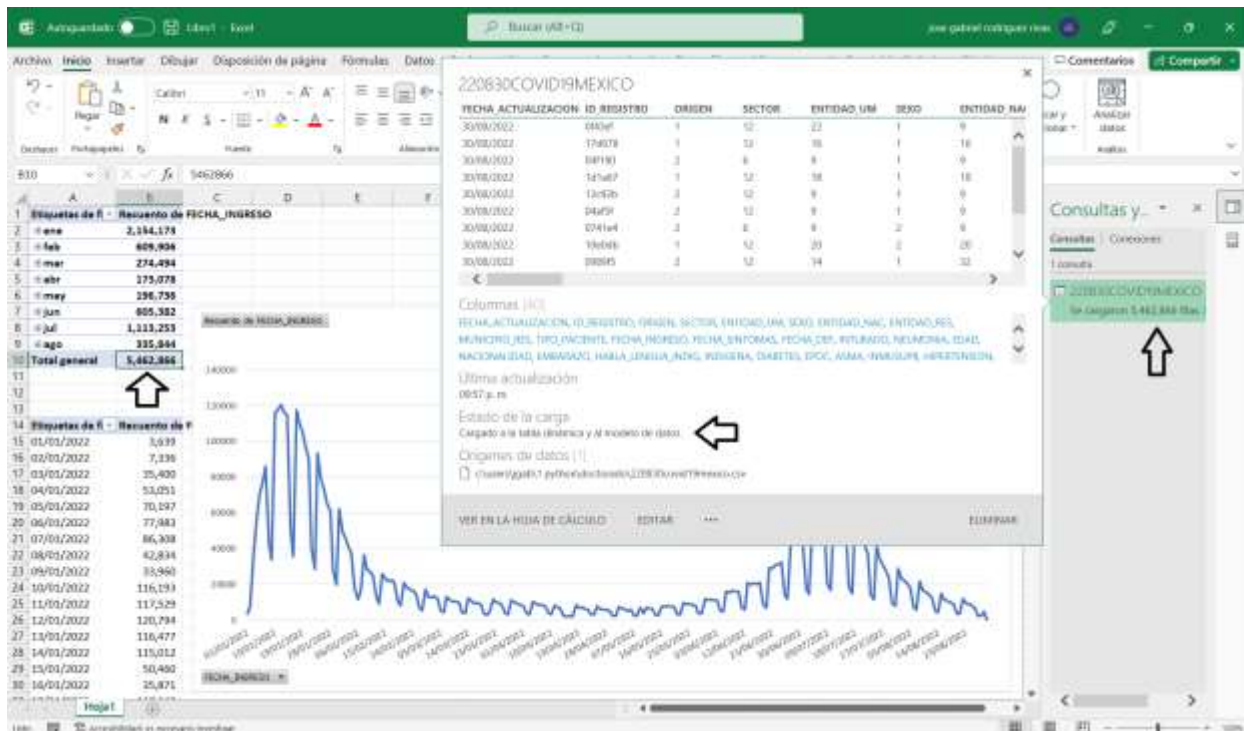
Para la segunda parte se abrió el archivo COVID desde el menú datos, opción obtener datos externos y enseguida en la ventana de importar el archivo, también se especificó que se realizaría únicamente la conexión con el archivo externo y, además, se activó la opción de agregar estos datos al modelo de datos. Enseguida, se insertó una tabla dinámica especificando que se incluyera desde el modelo de datos, y no desde una tabla o rango como comúnmente se hace.

En la tabla dinámica se arrastró la fecha de ingreso a la sección de filas y a la sección de valores. Automáticamente se agrupan las fechas por mes. Se copió la tabla dinámica, y en la copia se desagrupó para obtener los totales por día, y a partir de la segunda tabla dinámica se generó un gráfico de líneas para ver la evolución de las pruebas realizadas por COVID en el año 2022 al 30 de agosto.

En la figura 2, en la ventana lateral derecha muestra la conexión al archivo indicando que se cargaron 5,462,866 filas. Esta cantidad coincide con el total general que se muestra en la primera tabla dinámica. Al pasar el apuntador del ratón sobre el archivo de la conexión, muestra una ventana emergente con la descripción de los datos de conexión indicando que se cargó en el modelo de datos.

Los archivos de este trabajo se pueden descargar desde la dirección:
<https://drive.google.com/drive/folders/1QeyxoMLOnsmrFrpvnyooMGz6YPUKVG1v?usp=sharing>

Figura 2.
Tablas dinámicas con conexión a archivo de 5,462,866 registros



5. CONCLUSIONES

Se realizó un cuadro de mandos con segmentación de datos que permitirá identificar tendencias que pudieran estar ocultas en los datos. Además, se realizó la conexión a un segundo archivo con más registros de los que se pueden cargar en una hoja de cálculo de Excel, analizándolo con tablas dinámicas. Este cuadro de mandos servirá para tomar decisiones basadas en datos, lo cual, es prioritario para las empresas de cualquier tamaño. No importa qué herramientas se utilicen, lo relevante es una correcta presentación a través de tablas, gráficos e indicadores, pero, aún más importante es su correcta interpretación. Siempre debemos tener en mente la máxima que dice: “una imagen vale más que mil palabras”. En consecuencia, el análisis visual hace que

las tendencias y los patrones clave en los datos sean significativamente más fáciles de ver y entender.

REFERENCIAS

- Barón, R. E., García, E. C. W. y Sánchez, G. S. K. (2021). La inteligencia de negocios y la analítica de datos en los procesos empresariales. *Revista Científica de Sistemas e Informática*, 1(2), 37-53. <https://doi.org/10.51252/rcsi.v1i2.167>
- Briones, G. (2002). Metodología de la investigación cuantitativa en las ciencias sociales. Programa de especialización en teoría, métodos y técnicas de investigación social. *Instituto colombiano para el fomento de la educación superior*. Bogotá, Colombia. ARFO Editores e impresores Ltda.
- Gartner (2022). Analytics and Business Intelligence Platform Reviews and Ratings. <https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms>
- IBM (2022). ¿Qué es Business Intelligence? <https://www.ibm.com/mx-es/topics/business-intelligence>
- Kielstra, P., McCauley, D., y Kenny, M. (2007). "In search of clarity-Unravelling the complexities of executive decision making." *Economist Intelligence Unit*, 9-14.
- Microsoft (2022 a) Introducción al análisis de datos de Microsoft. <https://docs.microsoft.com/es-es/learn/paths/data-analytics-microsoft/>
- Microsoft (2022 b). Acerca de Power Query en Excel. <https://support.microsoft.com/es-es/office/acerca-de-power-query-en-excel-7104fbee-9e62-4cb9-a02e-5bfb1a6c536a>
- Steiber, A., y Alänge, S. (2016). The Silicon Valley Model. In *The Silicon Valley Model* (pp. 143-155). Springer, Cham. <http://dx.doi.org/10.1007/978-3-319-24921-6>
- Toro, L. F. J. (2022). Ciencia de los datos con Python. Ecoe Ediciones.

i

ⁱ Los autores del trabajo autorizan a la Universidad Internacional de Ciencia y Tecnología (UNICYT) a publicar este resumen en extenso en las Actas del Congreso IDI-UNICYT 2022 en Acceso Abierto (Open Access) en formato digital (PDF) e integrarlos en diversas plataformas online bajo la licencia CC: Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

La Universidad Internacional de Ciencia y Tecnología y los miembros del Comité Organizador del Congreso IDI-UNICYT 2022 no son responsables del contenido ni de las implicaciones de lo expresado en este artículo.