

TEMA 5: EXCEL PARA BIG DATA

Laura Valladares Roldán

1. Introducción a Excel para Big Data.

- Realiza una breve explicación sobre cómo Excel puede ser utilizado para trabajar con conjuntos de datos más grandes y complejos.

Cuando un archivo contiene un conjunto de datos muy grande Excel puede avisarnos de que el conjunto de datos es demasiado grande para la cuadrícula de Excel y que perderemos datos si decidimos guardar en ese libro. Esto significa que se excede el máximo de filas o columnas que Excel puede manejar. Para evitar esta situación podemos guardar una copia truncada del archivo original. También usar Power Query para cargar una base de datos grande sin fallos.

Las empresas pueden utilizar Excel para la gestión de presupuestos, análisis financiero, seguimiento de inventario, gestión de proyectos, análisis de ventas, seguimiento de clientes, recopilación de datos, gestión de recursos humanos, planificación estratégica, análisis de costes, seguimientos de KPIs, creación de informes ejecutivos, análisis de escenarios y datos de marketing. Para realizar estas tareas Excel es muy útil ya que permite organizar los datos para procesarlos y analizarlos, así como transformar los datos y presentarlos visualmente.

2. Manejo de grandes conjuntos de datos.

- Indica entre 3 y 5 consejos para trabajar de manera eficiente con grandes volúmenes de datos en Excel, como la optimización de fórmulas y el uso de funciones de tabla.

Cuando es necesario trabajar con grandes volúmenes de datos en Excel es recomendable: Evitar repetir datos, ordenar los datos en diferentes hojas para aprovechar el espacio y minimizar el número de celdas en blanco para evitar la ralentización de la hoja de calculo y consumir menos espacio.

Con Power Query algunos consejos son: filtrar los datos desde el principio, utilizar y definir el tipo de datos adecuado, verificar la validez de los datos, organizar correctamente nuestro trabajo y crear funciones que podamos reutilizar.

3. Herramientas de análisis para Big Data.

- **Menciona herramientas como Power Query y Power Pivot que permiten importar, transformar y analizar grandes conjuntos de datos de manera más eficiente en Excel.**

Un complemento de Excel que se puede utilizar para hacer análisis de datos eficientes es Power pivot, este permite combinar grandes volúmenes de datos y realizar varios análisis con la información de manera rápida. Permite filtrar los datos y cambiar las columnas y tablas mientras se importan, las tablas se organizan por páginas con pestañas individuales, no puede modificar celdas individuales y permite definir jerarquías para usar en cualquier parte de un libro.

Otro complemento es Power Query que permite importar datos externos o conectarse a ellos, permite integrar datos de distintos orígenes para obtener una única vista. Con ello puede combinarse en una hoja de calculo o modelo de datos que puede ir actualizándose. Tiene la capacidad de generar un código personalizado con el que acceder a tu propio código M para grabar cada transformación aplicada a los datos, también un lenguaje fácil de aprender.

4. Limitaciones y consideraciones.

- **Discute las limitaciones de Excel en comparación con herramientas específicas de Big Data, como la capacidad de procesamiento y la escalabilidad.**

Con Power pivot una limitación es que no se admite Visual Basic para Aplicaciones en la ventana de Power Pivot y no puede modificar celdas individuales de una tabla. Con Power Query la combinación de consultas o tablas y/o anexar es necesaria para que los datos consigan el resultado final esperado.

Es importante tener en cuenta que el formato .xls tiene un límite de 65.536 filas en cada hoja y el formato del archivo .xlsx tiene un límite de 1.048.576 filas por hoja. Power Pivot admite archivos de hasta 2GB de tamaño y permite trabajar con hasta 4GB de datos en memoria. También se pueden compartir los libros con otras personas publicándolo en SharePoint que Excel Services ofrece.

5. Recursos adicionales y conclusiones.

- Proporciona enlaces a recursos adicionales donde los usuarios puedan aprender más sobre cómo utilizar Excel para manejar y analizar Big Data de manera más avanzada. También, realiza una breve conclusión de los principales aprendizajes del módulo.

Como conclusión, manejar Big Data está a nuestro alcance y permite realizar unos análisis muy potentes con las herramientas adecuadas. Una de las grandes ventajas de Power Query es que cumple la función de un ETL, extrae, transforma y carga datos. Este editor te permite extraer información de distintas fuentes de datos y transformarlas y cargarlas en algún sitio para su uso.

Algunas diferencias entre Power Query y Power Pivot son, que el primero se encarga de importar y transformar los datos y el segundo de modelar y calcular con los datos, principalmente. El lenguaje en el primero es M y en Power Pivot es DAX. Ambas herramientas poseen grandes beneficios como una interfaz fácil de usar, o la capacidad de manejar grandes datos.

Algunos enlaces de interés son:

<https://blog.camaracomercioexterior.org/excel/5-consejos-para-trabajar-grandes-conjuntos-de-datos-en-excel/>

<https://support.microsoft.com/es-es/office/power-pivot-an%C3%A1lisis-de-datos-eficaz-y-modelado-de-datos-en-excel-a9c2c6e2-cc49-4976-a7d7-40896795d045>

<https://support.microsoft.com/es-es/office/acerca-de-power-query-en-excel-7104fbee-9e62-4cb9-a02e-5bfb1a6c536a>

<https://www.crehana.com/blog/negocios/que-es-power-query/>