

Excel para Big Data

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.

Excel ofrece un amplio repertorio de herramientas que permiten realizar análisis de datos grandes y complejos. Excel permite organizar y estructurar los datos de forma muy efectiva, lo que facilitará el trabajo con conjuntos de datos además que es muy flexible a la hora de poder adaptarse a las necesidades de cada usuario y se puede personalizar.

Las fórmulas que tiene Excel nos sirven para calcular, manipular, analizar datos, automatizar tareas repetitivas, entre otras muchas funciones y facilitan el trabajo con los datos. En Excel también disponemos de las tablas dinámicas, una de las funciones más destacables, que permiten resumir y analizar datos de forma rápida y eficiente, y facilitan la visualización de patrones y tendencias. Para hacer análisis estadístico también se puede utilizar y puedes crear gráficos y objetos visuales para facilitar la comprensión de patrones y tendencias. Otras funciones avanzadas son Power Query y Power Pivot, con las que puedes importar y analizar datos de forma eficiente.

Además Excel también permite trabajar con datos externos, y se pueden importar archivos como hemos visto en ejercicios anteriores, para manipular datos de diferentes fuentes.

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.

- Utilizar fórmulas eficientes (que calculen rápida y precisamente optimizando el rendimiento de excel) y evitar fórmulas volátiles (que se recalculan automáticamente y relantiza el rendimiento de Excel). Usar las funciones condicionales como IF, IF ERROR, y por ejemplo, la que para mi es la fórmula más útil que hemos aprendido estos días: VLOOKUP (BUSCARV). Será de gran ayuda si quieres buscar información de una tabla de datos. Si buscas un valor específico dentro de una tabla, a través de un valor en común, esta formula te devuelve el valor correspondiente. Las fórmulas complejas desagilizan el rendimiento de Excel.

- También usar tablas y referencias estructuradas facilitará la manipulación de datos, dentro de una estructura organizada. Usar rangos nombrados: Poner un título a rangos de datos importantes y utilizar la referencia de este rango en la formula.

- Utiliza los gráficos para visualizar la información que tienes cuando sea necesario. Para ver tendencias de tus datos y poder ver en forma de imagen una representación de tus datos, crea gráficos para poder facilitar la interpretación. Pero úsalos de forma puntual para no sobrecargar el documento de excel si estas trabajando ya que con grandes volúmenes de datos, demasiados gráficos pueden relantizar el funcionamiento de Excel.

- Dividir los datos en diferentes hojas de cálculo, de forma organizada. Si un archivo contiene muchos datos puede ocasionar una sobrecarga y empeorará el rendimiento de Excel. Al igual que maximizar la eficiencia de los datos que dispones: eliminar duplicados, combinar las columnas que puedas, etc. Con Power Query puedes hacerlo de forma sencilla.

- Como no podía ser otro: utiliza tablas dinámicas. Serán muy útiles para filtrar, clasificar, analizar y resumir datos de forma sencilla, clara y concisa.

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.

- Power Query: Con esta herramienta, puedes importar a Excel datos de diferentes fuentes externas como CSV, bases de datos, entre otras. A la hora de importar grandes conjuntos de datos, Power Query establece conexiones con distintas formas de datos y facilita la optimización y la transformación de datos gracias a herramientas como el poder filtrar datos, la posibilidad de cambiar el tipo de datos, eliminar información duplicada. Es decir, facilita extraer, transformar y cargar datos desde diversas fuentes de información, lo que permite así crear cuadros de mando e informes avanzados.

-Power Pivot: Es un complemento de excel con el podemos analizar los datos de forma eficaz y crear modelos de datos más complejos; relacionar y combinar diferentes conjuntos de datos de orígenes distintos y hacer análisis multidimensionales. Es decir, permite un manejo más avanzado de las tablas y los gráficos dinámicos.

Ambas permiten que se automaticen tareas repetitivas mediante la grabación de pasos. Es decir, por ejemplo, si más de una vez realizas los mismos cambios a tus datos, estos se pueden grabar y usarlos de forma automatizada las próximas veces. Así que sirven como atajo para agilizar el trabajo.

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.

- Límites de capacidad: Excel es más limitado que otras herramientas cuando hay volúmenes de datos muy grandes. El límite de capacidad en filas y columnas en Excel, es inferior al de otras herramientas.
- Velocidad de procesamiento: Tarda más que otras herramientas a la hora realizar análisis complejos o consultas. Y a medida que aumentas el volumen de datos, Excel resulta más problemático y lento.
- Procesamiento distribuido: Excel no ofrece capacidades avanzadas de procesamiento distribuido, a diferencia de otras herramientas de Big Data que distribuyen el procesamiento en servidores para tener un mejor rendimiento y una mejor escalabilidad.
- Colaboración multiusuario: La colaboración multiusuario en excel presenta menos compatibilidad. Excel no es una herramienta óptima para que un equipo trabaje a la vez sobre un documento. Y la versión en línea (que es la que mejora la experiencia multiusuario) tiene las funcionalidades de Excel más limitadas.
- Seguridad de datos: En comparación con otras herramientas, la protección de datos que ofrece no es tan segura.

- Acceso a funcionalidades: Otra limitación que tiene respecto a otras herramientas, es que el acceso a todas las funcionalidades, requiere pagar una licencia, cuando en otras herramientas hay acceso gratuito (aunque no en todas).

En conclusión, excel es una herramienta ideal para usuarios **no** profesionales en Big Data o que trabajen con bases de datos limitadas, ya que para tareas pequeñas, rápidas, es muy útil, intuitivo y sencillo de utilizar, además de ser muy conocido. Pero cuando se precisa analizar tablas de datos muy grandes o hacer análisis de gran complejidad, Excel queda limitado.

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.

Al ser una herramienta tan conocida, hay diferentes recursos donde se puede aprender sobre excel.

En plataformas de cursos como Udemy, Coursera, linkedin learning y Edx, se pueden encontrar distintas formaciones sobre excel y big data, por ejemplo este curso: <https://www.edx.org/es/learn/excel/universitat-politecnica-de-valencia-excel-avanzado-importacion-y-analisis-de-datos>

En youtube hay distintos creadores de contenido que publican videos para formar en excel y big data, como por ejemplo: <https://www.youtube.com/watch?v=hDD7jdWSO3U>

También hay libros disponibles: https://www.amazon.es/Analyzing-Power-Pivot-Business-Skills/dp/150930276X?source=ps-sl-shoppingads-lpcontext&ref_=fplfs&psc=1&smid=A1AT7YVPFBWXBL

Y webs y blogs que publican información sobre excel: <https://dominaexcel.org/>

Mis conclusiones sobre el curso son muy positivas. La mayoría de personas conocemos Excel pero con este curso he aprendido a ir más allá, sobre técnicas y funciones que ayudan a optimizar el uso.

La forma de organizar las clases también la considero muy positiva ya que cada vez que aprendíamos algo nuevo, lo aplicábamos al momento, combinando muy bien la teoría y la práctica para asimilar mejor todos los conceptos.

En cuanto al contenido, cada formula y función ha sido muy interesante y me gustaría destacar las tablas dinámicas. Considero que es de las funciones más interesantes que presenta Excel y facilitan mucho la forma de trabajar analizando datos.

Cómo hemos visto previamente en las limitaciones, quizás Excel no es la herramienta ideal para profesionales en el análisis de datos con grandes volúmenes de información pero por su practicidad y popularidad sé que seguiré utilizandola mucho.

Desde las funciones, las fórmulas, hasta las tablas dinámicas, Power Query, Power Pivot todo son recursos que permiten un análisis de datos complejos muy rápido, ideal para analizar y comprender los datos de manera eficiente y efectiva consiguiendo información significativa.