

## **Tema 5**

### **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 es una herramienta que se puede utilizar para trabajar con conjuntos de datos más grandes y complejos gracias a su capacidad para manejar datos a gran escala, generalmente con una estructura tabular. Algunas de las cosas que se pueden hacer con Excel son: filtrar, ordenar y agrupar datos, optimizar fórmulas y funciones, utilizar funciones de tablas y tablas dinámicas, crear gráficos, analizar hipótesis, usar complementos, macros y herramientas.

Es una herramienta transversal y poderosa para trabajar con datos, pero es importante tener en cuenta que en casos donde se requiera un análisis más avanzado o el manejo de volúmenes de datos masivos, puede ser necesario recurrir a herramientas específicas de Big Data.

### **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.**

Algunos consejos para trabajar de manera eficiente con grandes volúmenes de datos en Excel pueden ser:

- Optimización de fórmulas y funciones: es preferible evitar realizar fórmulas complejas y pesadas que puedan ralentizar el rendimiento de Excel y utilizar otras funciones más eficientes que permitan reducir el tiempo de cálculo. Además el uso de algunas funciones como por ejemplo las de búsqueda (como BUSCARV o BUSCARH) pueden ser útiles para recuperar datos específicos de grandes conjuntos de datos.
- Uso de funciones de tabla y tablas dinámicas: una buena opción es convertir los datos en tablas para facilitar la gestión y análisis de grandes conjuntos de datos. Estas tablas ofrecen ventajas como que se expanden automáticamente al meter nuevos datos, filtros automáticos, entre otros, que permiten agilizar el trabajo. Una buena práctica es el uso de funciones condicionales como SUMAR.SI o CONTAR.SI. También, con las tablas dinámicas se pueden visualizar los datos de diferentes formas mediante el uso de los campos de filas, columnas, valores y filtros. De esta forma, permiten resumir y analizar grandes volúmenes de datos de forma dinámica, obteniendo la información relevante más rápidamente. Igualmente, una buena práctica es el uso de gráficos para visualizar los datos e identificar tendencias, patrones y/o relaciones en los datos.

- Utilizar macros, herramientas y otros complementos: los complementos de Excel son herramientas adicionales que se pueden agregar a Excel para realizar ciertas tareas.
  - Macros: su uso permite automatizar tareas repetitivas mediante la grabación de unas acciones concretas, de esta forma se pueden realizar esos procesos de forma sencilla y ahorrar tiempo. Una vez creada la macro, se puede ejecutar tantas veces como queramos.
  - Herramientas: como por ejemplo Power Query y Power Pivot sirven para importar, transformar y analizar grandes volúmenes de datos de manera más efectiva y eficiente. En general, las herramientas de Excel son fundamentales para realizar todo tipo de tareas relacionadas con el manejo, análisis y presentación de datos.
- Filtrar, ordenar y agrupar datos: para que se muestren solo los que sean relevantes para cada análisis de forma rápida y sencilla, resumir la información y facilitar su visualización. En relación con esto, es recomendable evitar celdas que no aporten valor, como puede ser el caso de celdas en blanco o filas/ columnas vacías, ya que pueden ralentizar el programa. Una buena opción también es utilizar etiquetas y comentarios para organizar los datos y registrar los cambios que se realizan.
- Dividir los datos en distintas hojas de cálculo: de esta forma se facilita la manipulación de los datos, trabajando cada hoja por separado, incluyendo enlaces entre ellas si es necesario para acceder a información relacionada y combinando los resultados de distintas hojas de cálculo si se requiere.

### **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: es una herramienta de Excel que permite importar datos de distintas fuentes (como archivos csv, xlsx, xml, txt, json, bases de datos, páginas web, entre otros), transformarlos (combinar y dividir columnas, eliminar duplicados, unir tablas, etc, en definitiva, limpiar datos) y cargarlos en Excel. De esta manera, se facilita el proceso de importar y limpiar datos antes de comenzar con el análisis de los mismos en Excel.

Power Pivot: es una herramienta de Excel para modelado de datos que permite crear modelos de datos más complejos y eficientes que las tablas de Excel y realizar análisis. Permite importar grandes conjuntos de datos desde múltiples fuentes, de manera que puedes tener distintas tablas de datos, combinarlas, crear relaciones entre ellas y calcular medidas personalizadas.

Power BI: es una herramienta de Excel que sirve principalmente para crear visualizaciones. Es una herramienta de análisis de datos de empresas que permite importar, transformar y visualizar grandes conjuntos de datos, creando informes interactivos y paneles de control. También ofrece la posibilidad de trabajar en equipo y realizar predicciones.

Solver: es una herramienta integrada en Excel que permite encontrar soluciones óptimas para problemas complejos. Se define un modelo matemático que consta de una función objetivo que queremos maximizar o minimizar, variables que se pueden ajustar y restricciones a esas variables. Funciona mediante técnicas de programación (lineales y no lineales).

Tableau: es una herramienta de visualización y análisis de datos que, aunque es independiente, se puede integrar con Excel. Por ejemplo, puedes conectarte con Excel para importar datos y, en Tableau, realizar visualizaciones más avanzadas y análisis más profundos.

R y RExcel: R es un lenguaje de programación diseñado para realizar análisis estadísticos y crear gráficos. RExcel es un complemento de Excel que facilita la integración de Excel y R. De esta forma, permite ejecutar scripts y comandos de R desde Excel.

Python: es un lenguaje de programación utilizado en aplicaciones web, desarrollo de software, ciencia de datos y machine learning, popular en el ámbito del análisis de datos. Excel tiene integrada una opción para insertar código Python dentro de las celdas de Excel, de forma que permite analizar datos en Excel utilizando bibliotecas de Python (como Pandas o Matplotlib).

Estas son algunas de las herramientas que pueden ampliar las capacidades de Excel para manejar grandes volúmenes de datos y realizar análisis avanzados. En función de las necesidades y el tipo de análisis que queramos realizar, existen herramientas que se adaptan mejor.

#### **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.**

Excel es una herramienta versátil, transversal y poderosa para el análisis de datos a nivel individual o en entornos de trabajo más pequeños, pero es importante tener en cuenta que presenta limitaciones cuando se compara con herramientas específicas de Big Data. Cuando necesitemos realizar análisis más avanzados o manejar volúmenes de datos masivos, puede que necesitemos recurrir a este tipo de herramientas.

Algunas de las limitaciones de Excel en este sentido son:

- Capacidad de procesamiento: cantidad de trabajo (análisis de datos) que puede realizar la herramienta en un tiempo determinado. Excel tiene una capacidad de procesamiento limitada, lo que hace que los tiempos de cálculo puedan ser más largos, especialmente con grandes conjuntos de datos, mientras que otras herramientas específicas de Big Data están optimizadas para procesar grandes volúmenes de datos al dividir las tareas.
- Escalabilidad: capacidad de la herramienta para adaptarse y aumentar la capacidad de procesamiento cuando crece el tamaño de los datos. Excel tiene limitaciones para manejar datos cuando el tamaño de los mismos va aumentando, mientras que las herramientas específicas de Big Data escalan horizontalmente para poder manejarlos más eficientemente.
- Memoria: mientras que Excel dispone de la memoria disponible en el dispositivo local, las herramientas específicas de Big Data pueden almacenar y procesar datos en sistemas distribuidos, por lo que la limitación es menor.
- Datos no estructurados: Excel está diseñado para analizar datos en forma de hojas de cálculo y tablas, pero hay otro tipo de datos no estructurados (como imágenes, texto o audio) que se pueden analizar mejor mediante herramientas específicas de Big Data.
- Complejidad de análisis: las herramientas específicas de Big Data, proporcionan capacidades de análisis más avanzadas que las de Excel, entre las que se encuentran opciones como el procesamiento de datos en tiempo real, la generación de insights, el análisis predictivo, el aprendizaje automático o la inteligencia artificial.

**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.**

Para aprender más sobre cómo utilizar Excel para manejar y analizar Big Data de manera más avanzada existen, entre otros, cursos en línea en algunas plataformas como LinkedIn o Udemy, sitios web como Exceljet, comunidades online como grupos de LinkedIn o MrExcel, así como libros dedicados específicamente a Excel. Algunos ejemplos son:

- Sitio web oficial de Excel: <https://support.microsoft.com/es-es/excel>
- Canal de Youtube: ExcellsFun <https://www.youtube.com/user/ExcellsFun>
- Exceljet: <https://exceljet.net/>
- InfluExcel: <https://influexcel.com/tutoriales-excel/>

Breve conclusión de los principales aprendizajes del módulo:

Los datos son la base de cualquier análisis, pero para que ese dato nos sirva es necesario procesarlos para convertirlos en información que nos ayude a tomar decisiones. Existen múltiples herramientas de análisis de datos, de entre las que destaca Excel por ser la primera que hay que conocer, es la herramienta líder. Excel nos permite manipular datos en formato tabular (filas y columnas), los cuales podemos proteger a 3 niveles (celda, hoja y libro).

El programa permite tener un historial de versiones, hacer comentarios, compartir el archivo, etc. Podemos utilizar fórmulas, funciones, tablas, gráficos, macros, etc, que nos ayuden en el análisis y la visualización de los datos. También podemos importar archivos en distinto formato directamente a Excel y transformarlos con el uso de herramientas como Power Query o Power Pivot.

Hay multitud de complementos que se pueden utilizar en Excel y distintas herramientas que nos facilitan el trabajo. Excel es una herramienta con múltiples funcionalidades en el análisis de datos, de la cual siempre queda algo nuevo por aprender, por lo que este módulo ha sido muy interesante y me ha permitido aprender cosas que me pueden ser de gran utilidad en el trabajo.