

## INTRODUCCIÓN A EXCEL PARA BIG DATA

(Realiza una breve explicación sobre como Excel puede ser utilizado para trabajar con conjuntos de datos más grandes y complejos)

Excel puede utilizarse para trabajar con Big Data en ciertos contextos, aunque tiene limitaciones en comparación con herramientas especializadas en Big Data como Hadoop, BigQuery, Power BI, etc. Algunos casos de uso podrían ser:

- A. **Importación de datos:** Se pueden importar conjuntos de datos grandes desde diversas fuentes, cómo por ejemplo, bases de datos, archivos CSV, archivos de texto, entre otros.
- B. **Análisis básico:** Ofrece capacidades básicas de análisis de datos, como filtros, tablas dinámicas y gráficos, que pueden ser útiles para realizar análisis preliminares en conjuntos de datos grandes.
- C. **Procesamiento de datos:** Mediante fórmulas, funciones y herramientas como Power Query y Power Pivot, se puede realizar cierto nivel de procesamiento y transformación de datos para prepararlos para su análisis avanzado.
- D. **Visualización de datos:** Excel proporciona diversas opciones para visualizar datos a través de gráficos y tablas dinámicas, lo que facilita la comprensión y presentación de resultados.
- E. **Automatización de tareas:** A través de Macros y programación en VBA, se pueden automatizar tareas repetitivas o realizar análisis personalizados de grandes conjuntos de datos.

## MANEJO DE GRANDES CONJUNTOS DE DATOS

(Indica entre 3 y 5 consejos para trabajar de manera eficiente con grandes volúmenes de datos en Excel)

Aunque manejar grandes volúmenes de datos en Excel puede resultar complicado, haciendo uso de buenas prácticas, y utilizando determinadas estrategias, como las que se exponen en los siguientes consejos, se pueden manejar.

Algunas acciones que podemos realizar son:

- A. **Usar tablas:** Se pueden convertir los datos en tablas para facilitar la gestión de grandes conjuntos de datos, ya que ofrecen funcionalidades como filtrar, ordenar y realizar cálculos automáticos que facilitan el trabajo.

- B. **Optimizar las fórmulas:** Se debe evitar usar fórmulas que requieran cálculos con todo el conjunto de datos en cada celda, se deben utilizar funciones como SUMIF, COUNTIF, en vez de fórmulas matriz.
- C. **Filtrar datos de forma inteligente:** para visualizar y analizar segmentos específicos sin cargar toda la información al mismo tiempo.
- D. **Dividir los datos:** Si es posible, se ha de dividir el conjunto de datos en partes más pequeñas para trabajar con ellas por separado. Esto puede reducir la carga en la memoria y mejorar el rendimiento de Excel.
- E. **Utilizar gráficos dinámicos:** En lugar de mostrar grandes tablas de datos, es mejor utilizar gráficos dinámicos para visualizar resúmenes y tendencias clave de los datos. Los gráficos dinámicos son más efectivos para comunicar información en grandes conjuntos de datos.

## HERRAMIENTAS DE ANÁLISIS PARA BIG DATA

(Menciona herramientas )

Para el manejo, transformación y análisis de grandes volúmenes de datos existen herramientas diseñadas específicamente para Big Data, como pueden ser:

- A. **Power Query:** es una poderosa herramienta para importar, limpiar y transformar grandes volúmenes de datos antes de cargarlos en Excel, esto, ayuda a reducir el tamaño del archivo y mejorar el rendimiento.
- B. **Power Pivot:** es una herramienta poderosa de análisis de datos que se integra con Microsoft Excel y permite trabajar con grandes volúmenes de datos de manera eficiente.
- C. **Tableau:** es una herramienta de preparación de datos desarrollada por Tableau que permite limpiar, combinar y transformar datos de manera visual e intuitiva. Ofrece capacidades para unificar datos de múltiples fuentes y prepararlos para análisis en Tableau Desktop.
- D. **Google Data Studio:** es una herramienta gratuita de visualización de datos de Google que también ofrece capacidades básicas de preparación de datos. Permite conectar datos de diversas fuentes, transformarlos y crear informes interactivos y paneles de control.
- E. **Google BigQuery:** es un servicio de almacenamiento y análisis de datos en la nube ofrecido por Google Cloud Platform. BigQuery permite consultar y analizar grandes conjuntos de datos de forma rápida y escalable utilizando SQL estándar. También ofrece integración con otras herramientas de Google Cloud para análisis avanzados y machine learning.

## LIMITACIONES Y CONSIDERACIONES

Cuando se trabaja con Big Data, Excel tiene varias limitaciones y consideraciones importantes que hay que tener en cuenta. A continuación, voy a mencionar algunas de las limitaciones más destacadas y consideraciones al utilizar Excel para Big Data:

### A. Limitaciones de Excel para Big Data:

1. Capacidad de manejo de datos: Excel tiene limitaciones en cuanto a la cantidad de datos que puede manejar eficientemente. Con grandes volúmenes de datos, Excel puede volverse lento, consumir mucha memoria y tener problemas de rendimiento.

2. Tamaño de archivo: Los archivos de Excel tienen un tamaño máximo limitado, lo que puede ser insuficiente para grandes conjuntos de datos. Esto puede llevar a la pérdida de datos o a la corrupción del archivo si se fuerza a Excel a manejar más datos de los que puede soportar.

3. Datos no estructurados: Excel está optimizado para trabajar con datos estructurados en tablas y celdas. Para datos no estructurados, como texto sin formato, imágenes, audio, video, etc., Excel no proporciona las herramientas necesarias para su análisis eficiente.

4. Escalabilidad: A medida que la cantidad de datos y la complejidad del análisis aumentan, Excel puede no escalar de manera efectiva para manejar la carga de trabajo, lo que resulta en tiempos de respuesta lentos y problemas de rendimiento.

### B. Consideraciones al usar Excel para Big Data:

1. Uso prudente de la memoria: Al trabajar con grandes conjuntos de datos en Excel, es crucial administrar cuidadosamente la memoria del sistema para evitar bloqueos o ralentizaciones. Limitar el número de fórmulas complejas y optimizar el uso de recursos puede ayudar a mitigar este problema.

2. Filtrado y resumen inicial: Utiliza funciones de filtrado y resumen en Excel para reducir la cantidad de datos que hay que manipular en un momento dado. Esto puede ayudar a mejorar el rendimiento y la eficiencia del análisis.

3. Utilizar herramientas complementarias: Combina Excel con herramientas complementarias como Power Pivot, Power Query y Power BI para ampliar sus capacidades de manejo y análisis de datos. Estas

herramientas pueden ayudar a trabajar con grandes conjuntos de datos de manera más eficiente.

4. Considerar herramientas de Big Data: Para análisis avanzados de Big Data, considera utilizar herramientas y plataformas diseñadas específicamente para Big Data, como pueden ser Hadoop, Spark, bases de datos NoSQL, etc. Estas herramientas ofrecen capacidades de procesamiento distribuido y escalabilidad que Excel no proporciona.

En resumen, Excel puede ser útil para análisis inicial y manipulación de datos en entornos de Big Data, pero es importante tener en cuenta sus limitaciones y consideraciones para asegurar un análisis eficiente y preciso en grandes volúmenes de datos. Combinar Excel con herramientas especializadas en Big Data puede proporcionar un enfoque más completo y escalable para el análisis de datos a gran escala.

## **RECURSOS ADICIONALES Y CONCLUSIONES**

Para aprender a utilizar Excel de manera avanzada para Big Data, hay recursos en línea, cursos y tutoriales que proporcionen instrucciones detalladas y ejemplos prácticos. Algunos enlaces y recursos adicionales que pueden ser útiles:

1. Microsoft Excel Official Website: sitio web oficial de Microsoft Excel que ofrece documentación, tutoriales y recursos de aprendizaje: [Microsoft Excel](<https://www.microsoft.com/es-es/microsoft-365/excel>)
2. Microsoft Learn - Excel for Data Analysts: Recursos de aprendizaje de Microsoft que cubren aspectos avanzados de Excel para análisis de datos: [Microsoft Learn](<https://learn.microsoft.com/en-us/training/browse/?products=excel>)
3. LinkedIn Learning: Excel para análisis de datos: Curso en LinkedIn Learning que cubre técnicas avanzadas de Excel para análisis de datos: [LinkedIn Learning](<https://www.linkedin.com/learning/excel-para-analisis-de-datos>)
4. Keepcoding :=)
5. ExcellsFun YouTube Channel: Canal de YouTube con una amplia gama de tutoriales sobre Excel, incluyendo funciones avanzadas para análisis de datos: [ExcellsFun](<https://www.youtube.com/user/ExcellsFun>)
6. Exceljet: Sitio web con tutoriales detallados y ejemplos prácticos para dominar Excel, incluyendo técnicas avanzadas para análisis de datos: [Exceljet](<https://exceljet.net/>)

En este módulo hemos aprendido a modo introductorio la diferencia entre dato e información, los diferentes roles dentro de la analítica de datos y diferentes herramientas para el análisis de datos.

En cuanto a Excel, como introducción, hemos aprendido qué es, para qué sirve y por qué es útil aprender Excel, cómo funciona, los beneficios de aprender Excel, así como su entorno y componentes.

Tras esta introducción, hemos aprendido la diferencia entre una fórmula y una función, a manejar funciones básicas y avanzadas de Excel, Data Viz y gráficos, las tablas dinámicas realizando ejercicios prácticos sobre cada uno como:

- Funciones para suma, resta, división, conteo...
- Funciones condicionales
- Creación de tablas
- Creación de gráficos a partir de las tablas
- Creación de tablas dinámicas, aprendiendo cómo añadir filtros y cómo cambiar las relaciones entre los datos, seleccionando que datos de nuestra tabla queríamos mostrar y dónde.

Profundizando, hemos aprendido la parte de Data Acquisition con Excel y las herramientas y complementos para la importación y transformación de datos como:

- Power Query : con el que podemos realizar el procesamiento de datos de extracción, transformación y carga (ETL). Hemos aprendido las transformaciones que podemos realizar, como añadir o quitar columnas, renombrar columnas, corregir el tipo de datos de una columna, añadir columnas condicionales, reordenar y filtrar los datos, etc.
- Power Pivot: que es una extensión de Excel, que ayuda a manejar grandes volúmenes de datos y realizar análisis más avanzados de manera más eficiente utilizando fórmulas DAX.
- Power BI

Por último hemos visto herramientas externas y complementos, como:

- VBA, que es un lenguaje de programación desarrollado por Microsoft y que permite automatizar tareas repetitivas, crear funciones personalizadas, manipular datos y crear aplicaciones dentro de Excel, Word, Access y Power Point.
- Asociado al VBA, hemos aprendido a grabar macros, que son secuencias de comandos para automatizar tareas.

A modo de conclusión, Excel puede usarse como un primer paso para el análisis y creación del conjunto de datos, pues permite hacer un primer análisis de los datos sin necesidad de tener conocimientos de programación, aunque es cierto, que para análisis más profundos y en los que se requiera mayor capacidad de

procesamiento y escalabilidad es mejor utilizar las herramientas diseñadas para Big Data.