## EXCEL PARA BIG DATA

- 1.Excel puede ser utilizado para trabajar con conjuntos de datos grandes y complejos, ya que incluso en la era del BIG DATA sigue siendo una herramienta esencial gracias a la forma en que ha sabido adaptarse y la cantidad de profesionales que están familiarizados con esta herramienta.
- 2. Para poder trabajar de forma eficiente con grandes volúmenes de datos en Excel, hay algunos consejos como los siguientes:
  - Ordenar los datos: Según los valores de una o varias columnas podemos segmentar los datos que queremos analizar de una forma estructurada. También podemos identificar valores únicos y analizarlos por separado.
  - <u>Evitar el uso de formulas volátiles</u>: Funciones como NOE, HOY, RAND, OFFSET... recalculan cada vez que hay un cambio en el libro de trabajo, por lo que ralentizan bastante el proceso y habría que intentar minimizar su uso.
  - <u>Utilizar técnicas de fórmulas más rápidas</u>: Por ejemplo, utilizar IFERROR mejor que IF+ISERROR o utilizar INDEX/MATCH en lugar de VLOOKUP.
  - <u>Usar tablas dinámicas</u>: Son una función de Excel extremadamente útil que proporciona resúmenes de grandes colecciones de datos, ya que te permite escoger las columnas y filas con las que deseas trabajar.
  - Realizar búsquedas avanzadas de caracteres: Para realizar una búsqueda avanzada se usan los caracteres comodín. Esta búsqueda es útil cuando entre tantos datos no estas seguro del dato exacto que buscas.

Hay tres caracteres de este tipo:

- \* (asterisco): representa caracteres que deben estar presentes en la frase que buscas. Por ejemplo, Ex \* podría significar Excel, Excels, Example, Expert, etc.
- ? (signo de interrogación): representa un solo carácter. ¿Por ejemplo, Tr? Mp podría significar Trump o Tramp.
- ~ (tilde): se utiliza para identificar un carácter comodín (~, \*,?) en el texto.
- 3. Además de estos consejos hay varias herramientas que nos permiten hacer análisis BIG DATA en Excel.

Una de estas herramientas es **Power Query**, conocida como **obtener y transformar** ya que es utilizada para la importación de datos externos y permite transformarlos según las necesidades requeridas. Desde Power Query podemos dar forma a los datos, como quitar columnas y filtrar, y también podemos combinar datos desde diferentes orígenes.

Una vez se ha realizado la consulta que queramos podemos cargar la tabla en Excel para trabajar con ella.

Otra de las herramientas utilizadas en Excel para BIG DATA es **Power Pivot.** 

Esta herramienta nos permite importar millones de filas de datos desde varios orígenes en un único libro de Excel. También nos permite procesar millones de filas al mismo tiempo y realizar cálculos más rápidamente. Se pueden importar y combinar datos desde cualquier ubicación para realizar análisis masivos.

Además, amplia la capacidad de manipulación de datos de Excel utilizando el lenguaje de formulas DAX, con el cual se puede habilitar agrupaciones, cálculos y análisis más complejos y sofisticados.

En el siguiente cuadro podemos comparar algunas de las tareas:

Tarea	En Excel	En Power Pivot		
Importar datos de orígenes diferentes, como bases de datos corporativas grandes, fuentes de distribución de datos públicas, hojas de cálculo y archivos de texto en el equipo.	Importar todos los datos de un origen de datos.	Filtrar los datos y cambiar las columnas y las tablas mientras se importan.		
Crear tablas	Las tablas pueden estar en cualquier hoja de cálculo del libro. Las hojas de cálculo pueden tener más de una tabla.	Las tablas están organizadas por páginas con pestañas individuales en la ventana de Power Pivot.		
Editar datos en una tabla	Puede modificar los valores de celdas individuales de una tabla.	No puede modificar celdas individuales.		
Crear relaciones entre tablas	En el cuadro de diálogo Relaciones.	En la vista de diagrama o el cuadro de diálogo Crear relaciones.		
Crear cálculos	Use las fórmulas de Excel.	Escriba fórmulas avanzadas con el lenguaje de expresiones Expresiones de análisis de datos (DAX).		
Crear jerarquías	No disponible	Defina jerarquias para usar en cualquier parte de un libro, incluido Power View.		
Crear indicadores clave de rendimiento (KPI)	No disponible	Cree KPI para usarlos en informes de Power View y tablas dinámicas.		
Crear perspectiva	No disponible	Crear perspectivas para limitar el número de columnas y tablas que los usuarios de un libro ven.		
Crear tablas dinámicas y gráficos dinámicos	Crear informes de tablas dinámicas en Excel.	Haga clic en el botón Tabla dinámica en la ventana de Power Pivot.		
	Crear un gráfico dinámico.			
Mejorar un modelo para Power View	Crear un modelo de datos básico.	Realizar mejoras como identificar los campos predeterminados, las imágenes y los valores únicos.		

Tarea	En Excel	En Power Pivot
Usar Visual Basic para Aplicaciones (VBA)	Usar VBA en Excel.	VBA no se admite en la ventana de Power Pivot.
Agrupar datos	Agrupar en una tabla dinámica de Excel.	Use DAX en columnas calculadas y campos calculados.

4. Aunque Excel es una herramienta poderosa en el manejo de datos, tiene ciertas limitaciones a la hora de trabajar con grandes volúmenes de datos.

Una de las grandes desventajas es que tiene una capacidad limitada para manejar bases de datos extremadamente grandes y complejas. El formato .xlsx tiene un límite de 1.048.576 filas por hoja, lo que hace que se ralenticen tareas y procesamiento de la información.

Esto afecta también a la escalabilidad, ya que a medida que la cantidad de datos aumenta, puede volverse además de lento, inmanejable, y la estructura de una hoja de calculo puede volverse confusa y desordenada según se van agregando mas datos, lo que dificulta la organización y búsqueda de la información.

Otro de los inconvenientes de Excel con respecto a otras herramientas es la seguridad de los datos, que pueden ser fácilmente vulnerados. Aunque Excel ofrece protección por contraseña y restricciones, no son tan robustas como otras soluciones más especializadas, por lo que no es complejo hacer copias y que estas puedan ser manipuladas o distribuidas sin autorización y haya una fuga de información confidencial.

También la limitación en la colaboración a tiempo real hace que no sea una de las herramientas más eficientes en cuanto a trabajo colaborativo. Aunque Microsoft ha mejorado los servicios en la nube, sigue estando por detrás de otras herramientas más específicas para BIG DATA.

5. Finalmente podemos apreciar que, aunque Excel sigue siendo una herramienta poderosa y muy útil, es importante aprender a manejar otras herramientas que estén mas enfocadas al BIG DATA si vamos a trabajar con grandes volúmenes de datos y en proyectos en los que van a trabajar varias personas.

Lo ideal seria combinar Excel con estas herramientas y utilizarlo para hacer tareas en local combinado, como hemos visto, con Power Query o Power Pivot.

A continuación, tenemos algunos enlaces en los que podemos apoyarnos y ampliar nuestros conocimientos sobre Excel para abordar estas tareas:

Funciones básicas de Excel:

https://blog.hubspot.es/marketing/tips-teclado-formulas-excel

También hay muchos recursos en youtube, como cursos completos:

https://www.youtube.com/playlist?list=PL42LRSNpwreCW2WsKs1L8nC 3tMlZkEFL