

EXCEL PARA BIG DATA



1. Introducción a Excel para Big Data

Excel, es una herramienta poderosa, es la opinion de sus 750 millones de usuarios, una herramienta que no necesita muchos conocimientos técnicos como la programación o el lenguaje estructurado (structured query language SQL), además ofrece una forma muy visual de poder usarlo.

Las empresas trabajan con archivos de datos cada vez más grandes. ¿Cómo se obtiene información allí de forma rápida y clara?

Pues bien, Excel tiene una gran cantidad de opciones, como el formato condicional, el filtrado, los subtotales y las tablas dinámicas.

Excel puede cargar datos externos y acondicionarlos a tu modelo de trabajo para que puedas extraer la información necesaria de ellos.

2. Manejo de grandes conjuntos de datos

La herramienta de Microsoft a logrado permanecer vigente gracias a sus clásicas y reutilizadas funciones y formulas, entre estas:

CONCATENAR
MAX y MIN
INDICE + COINCIDIR
BUSCAR y HALLAR
BUSCARV
Comparaciones con SI
LARGO
SUMA, CONTAR, PROMEDIO

Pero como podrias emplear estas armas para trabajar con grandes volúmenes de datos, pues bien a continuación te ofrecemos estos importantes consejos a tomar en cuenta.

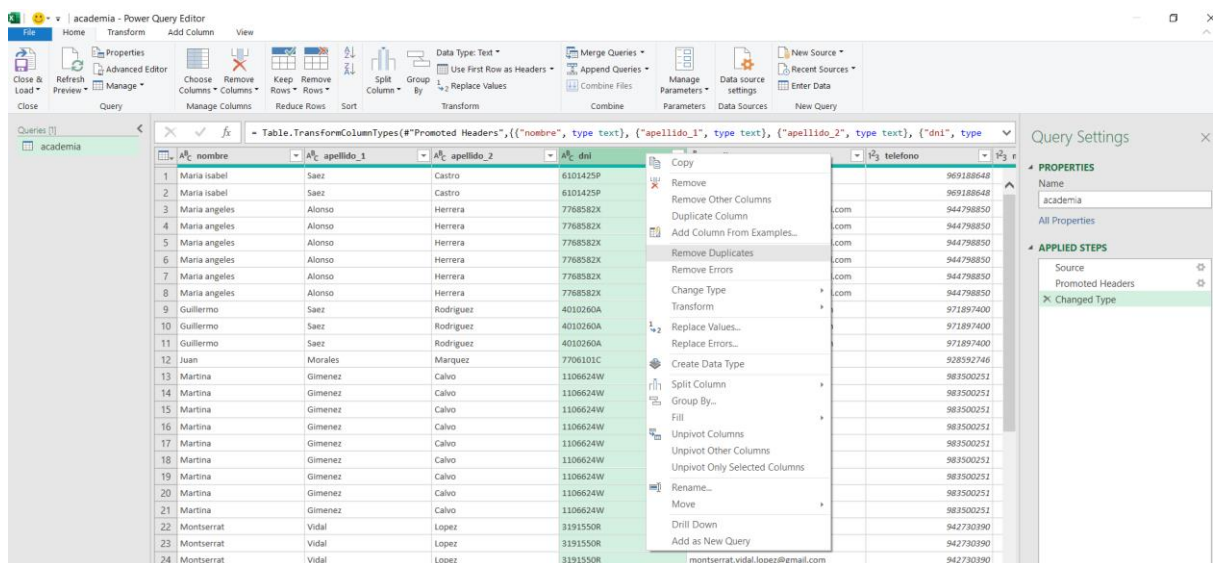
- Primero, haz una limpieza de los datos con los que vas a trabajar, para identificar las impurezas ordena, Excel tiene funciones sencillas para ordenar tus datos y hacer más reconocibles visualmente los errores. Estructura tus datos de manera lógica, usa encabezados significativos para cada columna para que los datos sean más comprensibles y fáciles de navegar.
- Segundo, usa las tablas dinámicas son una poderosa herramienta para resumir y analizar grandes cantidades de datos. Te permiten agrupar y filtrar datos fácilmente, calcular totales y realizar otros cálculos avanzados. Utiliza la función de filtro para mostrar solo la información relevante, esto facilita el trabajo con grandes volúmenes de datos.

- Por último Excel permite una colaboración perfecta en grandes conjuntos de datos. Utiliza plataformas de almacenamiento en la nube como OneDrive o SharePoint para compartir y colaborar en archivos de Excel en tiempo real con colegas o miembros del equipo..

3. Herramientas de análisis para Big Data

Desde siempre Excel ha permitido a sus usuarios involucrarse con las bases de datos, sin ir más lejos sus Macros han resuelto más de un quebradero de cabeza, ahora ofrece **Power Query**, una nueva alternativa para tratar más a fondo tus datos y para que además de solo cargar tus archivos que vengan en otros formatos (.txt,.csv,.xml) , puedas, antes de cargarlos, transformarlos (ETL), puedes unir columnas, dividirlos, crear nuevas columnas basandote en las condiciones de tu tabla previa y hasta hacer un join entre ellas.

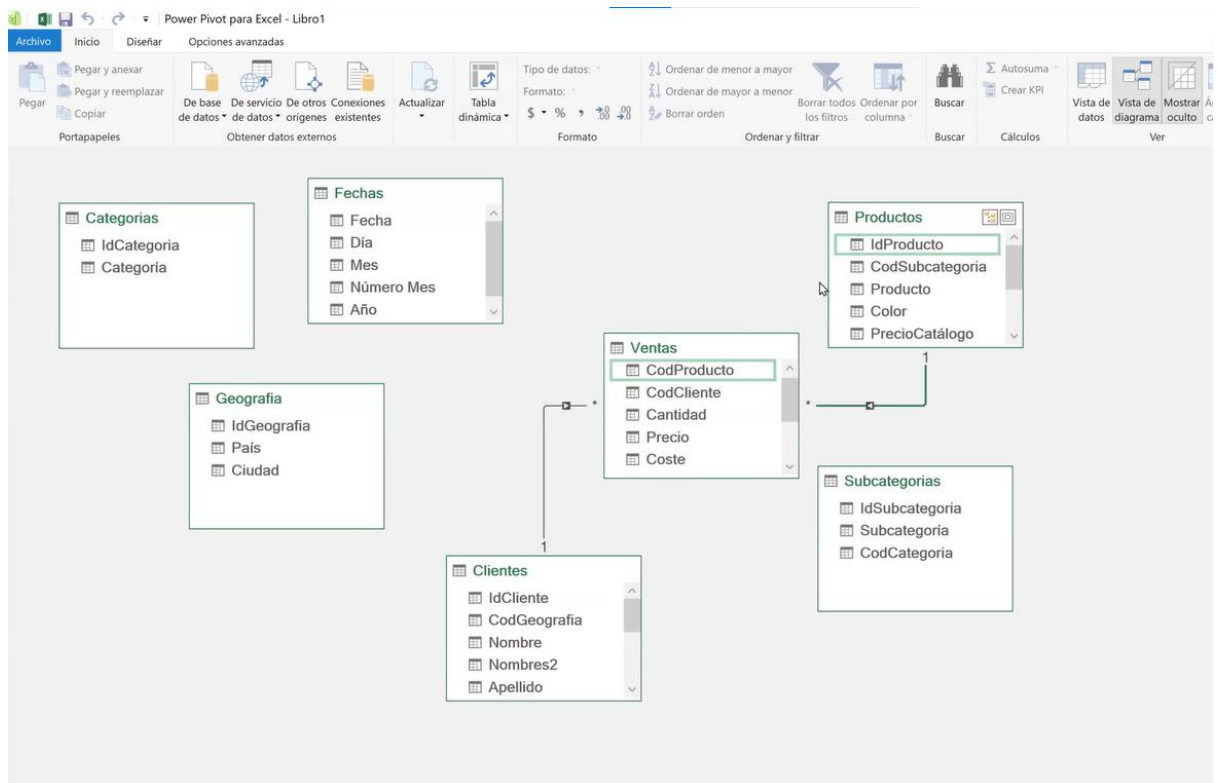
Automatiza el proceso de limpieza y transformación de datos, ahorrándote tiempo y esfuerzo valiosos.



Power Query está estrechamente integrado con otras herramientas de Microsoft, como Excel y Power BI. Esto significa que los usuarios pueden aprovechar los datos y flujos de trabajo existentes, haciéndolos más fáciles de adoptar y utilizar.

Otra de las alternativas que ofrece Excel es **Power Pivot**, un paquete de análisis y modelado de datos integrado en Microsoft Excel. Incorpora algunas funciones clave de inteligencia empresarial al corazón de Excel, lo que permite a los usuarios obtener información más rápidamente. Esta funcionalidad incluye modelado de datos, mayor capacidad de datos en comparación con las hojas de cálculo de Excel y el potente lenguaje de fórmulas DAX.

Power Pivot transformará la capacidad de un analista para trabajar con grandes cantidades de datos relacionados, eliminando fórmulas de búsqueda intensivas y brindándoles una manera de producir análisis rápidos y repetibles.



4. Limitaciones y consideraciones

En el ámbito del Big Data, donde se manejan grandes volúmenes de datos, una de las principales desventajas de Excel es su capacidad limitada para manejar bases de datos extremadamente grandes y complejas.

A medida que el tamaño de los datos aumenta, Excel puede volverse lento, lo que afecta la eficiencia y la rapidez en el procesamiento de la información.

Limitaciones en la Colaboración en Tiempo Real, aunque Microsoft ha mejorado las capacidades de colaboración en tiempo real en Excel, sigue estando por detrás de otras herramientas especializadas en este aspecto. La colaboración en documentos de Excel, especialmente cuando múltiples usuarios trabajan simultáneamente, puede ser menos intuitiva y eficiente en comparación con otras plataformas diseñadas específicamente para el trabajo colaborativo.

Finalmente, las cuestiones de escalabilidad y la integración con otras herramientas y sistemas pueden ser desventajas considerables de Excel.

Excel no siempre es la mejor opción para proyectos que requieren escalabilidad. A medida que las necesidades de los usuarios crecen, especialmente en grandes organizaciones o en aplicaciones de datos complejas, Excel puede no ser suficiente para manejar la expansión sin problemas.

5. Recursos adicionales y conclusiones

Básicamente, Excel está diseñado para funcionar con hojas de cálculo. Los usuarios pueden crear varias hojas dentro de un solo libro, lo que facilita la organización y categorización de datos. Esta estructura basada en cuadrícula permite la entrada y gestión eficiente de datos.

Las características de una hoja de cálculo, como sus fórmulas y funciones versátiles, potentes herramientas de visualización de datos y sólidas opciones de clasificación y filtrado, permiten a los usuarios gestionar y analizar datos de manera eficiente.

Fórmulas y funciones

Una de las características destacadas de Excel es su capacidad para realizar cálculos y operaciones con datos mediante fórmulas y funciones.

Los usuarios pueden crear cálculos complejos combinando operadores matemáticos, referencias de celdas y funciones integradas. Algunas funciones comúnmente utilizadas incluyen SUMA, PROMEDIO, SI y BUSCARV.

Visualización de datos

Excel ofrece una variedad de tipos de gráficos de visualización de datos, como gráficos de barras, gráficos de líneas y gráficos circulares, para ayudar a los usuarios a visualizar tendencias y patrones de datos. Esta es una de las principales características de Excel.

Crear gráficos es tan simple como seleccionar datos y elegir un tipo de gráfico, lo que lo convierte en una excelente herramienta para crear informes y presentaciones informativos.

Tablas dinámicas y gráficos dinámicos

Las tablas dinámicas son herramientas poderosas para resumir y analizar grandes conjuntos de datos. Esta es una de las funciones avanzadas de Excel. Permiten a los usuarios crear resúmenes dinámicos, aplicar filtros y reorganizar datos para obtener información desde diferentes ángulos. PivotCharts trabaja de la mano con PivotTables para proporcionar representaciones visuales de los datos resumidos.

Validación de datos y listas desplegables

Excel permite a los usuarios crear listas desplegables y aplicar reglas de validación de datos para garantizar la precisión y coherencia de los datos. Esta característica es particularmente útil cuando varias personas ingresan datos en la misma hoja de cálculo, ya que ayuda a evitar errores.

Formato condicional

El formato condicional permite a los usuarios formatear celdas automáticamente según criterios específicos. Esta es una de las características más importantes de Excel y ayuda a resaltar puntos de datos, tendencias o valores atípicos importantes, lo que facilita la identificación de información clave en sus datos.

Clasificación y filtrado de datos

Excel proporciona sólidas capacidades de clasificación y filtrado que facilitan la organización y localización de datos específicos dentro de un gran conjunto de datos. Los usuarios pueden ordenar los datos alfabéticamente o numéricamente y aplicar filtros para mostrar solo la información relevante.

Importación y exportación de datos

Excel facilita la importación de datos de diversas fuentes, incluidas bases de datos, páginas web y archivos de texto. También ofrece múltiples opciones de exportación, lo que permite a los usuarios guardar su trabajo en formatos como CSV, PDF y más.

Colaboración y uso compartido

Con la integración de Microsoft 365, Excel admite la colaboración en tiempo real, lo que permite que varios usuarios trabajen en una hoja de cálculo simultáneamente. Los usuarios también pueden compartir archivos a través de OneDrive o SharePoint, mejorando la colaboración y el control de versiones.

En resumen el módulo de excel, me ha dejado gratamente sorprendida, pensaba que a lo mejor yo tendría un poco más de conocimiento sobre esta herramienta ya que la suelo usar, pero me he dado cuenta que nada más lejos de la realidad que Excel es un sin fin de fórmulas y funciones nuevas que aprender y explorar.

También quiero destacar, haber tenido todo el material preparado para cada una de las prácticas, me ha hecho más fácil seguir y realizar los ejercicios, que además eran bastante legibles y claros, fáciles de entender también gracias a los títulos y colores que se usaron.

No tenía conocimiento de Power Query y Power Pivot, y he podido ver que tienen también mucho que aportar en el mundo del Big Data.

Enlaces de interés

https://cincodias.elpais.com/cincodias/2023/07/05/lifestyle/1688550381_876235.html#

<https://www.xataka.com/basics/microsoft-excel-21-funciones-trucos-consejos-para-exprimir-al-maximo-app-hojas-calculo>

<https://learn.microsoft.com/en-us/power-query/best-practices>