EL MODELO DE DATOS DE MICROSOFT EXCEL

PRESENTACION

Como parte de una serie de iniciativas para fortalecer a la Institución, al personal y a la forma en que realizan sus actividades, el área de informática presenta este taller de capacitación, por medio del cual se pretende renovar la forma en que los trabajadores realizan las rutinas diarias. Se trata de entender el conjunto de objetos con que cuenta MS Excel para el modelado y presentación de datos, dar un primer acercamiento a esta técnica para que el usuario sea capaz de mejorar sustancialmente su productividad.

Este material está orientado especialmente al personal administrativo de zonas, pues son la parte operativa que tiene la responsabilidad de la información generada en campo, mediante la cual se toman decisiones y se plantean estrategias para el cumplimiento de las metas de la institución. Obviamente, también puede ser distribuido a todas las personas que sientan interés por aprender nuevas técnicas y actualizarse en el uso de este programa.

Este taller es un primer acercamiento hacia rutinas específicas, en donde se ha detectado que hay algún tipo de carencia de conocimiento técnico. Se pretende ahondar en estos temas y realizar otros talleres en otras temáticas que presenten las mismas carencias. Como área de informática estamos atentos y dispuestos para el mejoramiento del contenido, esperando la colaboración del personal interesado para elaborar un material de mayor relevancia; así mismo, estamos en plena disposición para cualquier duda o apoyo al respecto.

El formato es escrito y en video, es así para que el personal se sienta en confort y lo estudie en el momento que considere más propicio. Podrán leer con calma y repasar los videos al tiempo en que lo practican en sus equipos. No pretendemos que reemplace sus antiguas bases de datos de forma inmediata, sino mostrarle que hay otras alternativas, y que los nuevos proyectos que realice los haga con una mayor solidez teórica y técnica. Si se siente con capacidad de rehacer sus bases de datos... haga un respaldo de las mismas y manos a la obra.

Esperamos que sea de mucho provecho:

Dr. Enoc Gordillo Argüello, Director de INIFECH.

Lic. Alexis de Jesús González Bonifaz, Jefe de Informática

Equipo de Soporte Técnico del área de Informática.

Junio de 2022.

BREVE HISTORIA DE MICROSOFT EXCEL

Excel es el software especializado en hojas de cálculo más longevo y depurado del mundo, a casi cuarenta años de la primera versión para Mac, Excel experimenta una especie de resurgimiento debido a la masificación del internet y a disciplinas como la inteligencia de negocios B I, (por sus siglas en inglés). Hoy en día, la ciencia de datos es cada vez más útil en un mundo en donde el manejo de la información es crucial.

Visual Basic fue decisivo para posicionar al producto. En 1993 se incluyó como parte de una serie de actualizaciones importantes, agrupadas en la versión 5. Desde entonces, la integración del programa con un lenguaje de programación fácil de aprender, con el que se pueden automatizar tareas rutinarias, lo hicieron indispensable para realizar el trabajo de millones de usuarios, tanto en el trabajo como en el hogar.

Para la versión 2010, Excel muestra la intención de la compañía de desarrollar un producto para modelar y manipular datos de manera profesional, desde entornos heterogéneos, hasta fuentes de datos locales y minería de datos en internet. El concepto de “la nube” y “el internet de las cosas”, está cada vez más presente en los medios privados y públicos, las tecnologías se perfilan para desenvolverse en este tipo de entornos.

En la versión 2013, las innovaciones mostraban un motor de datos muy robusto y capaz de realizar un gran número de tareas en “la nube”. El ecosistema POWER (Power query, Power pivot, Power view y Power map) en el contexto de Business Intelligence, estaba por fin integrado y a punto para trabajar en colaboración con todas las demás herramientas. La conexión a datos almacenados dentro o fuera del programa se logra con Power Query, posteriormente se modela la información con Power Pivot, finalmente se presenta la información con Power Map, o se crean vistosos informes o dashboards con una serie de herramientas como lo son las tablas y gráficos dinámicos.

Con la masificación de Excel el producto se fue utilizando para todo, incluso en tareas donde no es idóneo hacerlo, por ejemplo las bases de datos. Si bien Excel cuenta con herramientas muy potentes para manipular información, no es recomendable implementar una base de datos desde la plataforma de hojas de cálculo, para ello se cuenta con programas como Access, para un entorno de pequeñas bases de datos, o SQL Server, para entornos profesionales. Estos programas especializados sobresalen por el enfoque relacional que manejan de manera nativa, y por el aspecto de la seguridad, para lo cual tienen muchas herramientas muy sofisticadas, especialmente SQL Server.

Lo que sí es Excel, es una excelente herramienta para el acceso y presentación de información. Desde aquellas primeras versiones, hasta el actual y novedoso motor de datos, la tecnología ha evolucionado para poner a nuestra disposición filtros avanzados, un gran número de proveedores, conectividad, poderosas tablas dinámicas, novedosas funciones integradas y diversidad en herramientas de desarrollo; mención especial a los complementos de Java Script y Python para Excel.

BASES DE DATOS – NORMALIZACION

Por desgracia todo este poder informático tiene su costo, y es que no sería posible sin antes haber resuelto algunas reglas de estructuración de la información conocida como normalización de datos. Para personas recién llegadas a estos temas parecerá trabajo de más, pero es un trabajo que recompensa con creces a la hora en que tienes prisa por resolver alguna de esas urgencias de tu organización por algún tipo de reporte.

Muchos hablan de 3 y otros de 5 reglas normales, pero aquí trataremos de resumirlas así:

1. Cada entidad debe formar una tabla, la cual debe contener campos que estén relacionados de manera única con esa entidad, las filas deben ser valores únicos identificados por una clave. Por ejemplo, la entidad EMPRESA debe contener información relacionada únicamente con la empresa, como es el caso del RFC, domicilio, representante legal, etc. Cada empresa es una fila única que se identifica por una clave conocida como ID, la tabla debe ser plana y simple, sin celdas o columnas combinadas y sin espacios en blanco.

[Video: Entidades](file:///C:\Users\adan\Desktop\VIDEOS%20CREXCEL\Proyecto%20Curso%20Excel\entidades.mp4)

1. Cada campo o columna representa una propiedad o atributo de la entidad y se debe procurar la mayor segmentación de la información para que posteriormente sea más fácil catalogar, ordenar o filtrar. Por ejemplo: en el campo dirección con frecuencia se combina calle+numero+colonia, lo más conveniente sería la segmentación en un campo separado para cada concepto, esto favorecerá un filtrado más eficiente con criterios más complejos, pero fáciles de implementar.

[Video: Reglas 1 y 2](file:///C:\Users\adan\Desktop\VIDEOS%20CREXCEL\Proyecto%20Curso%20Excel\regla%201y2%20ok.mp4)

1. La relación entre tablas se realiza a través de las claves, se agrega un campo donde va la clave de la tabla relacionada, y esta columna (s) es el único dato “extraño” que se permite dentro de los campos de cada entidad. Por ejemplo, en la relación empresa-teléfono planteamos que es de tipo UNO a VARIOS, pues una EMPRESA puede tener VARIOS números telefónicos. Visto así, la tabla teléfonos contiene un campo denominado EMPRESA\_ID con el cual se relacionan.

[Video: Relación uno a varios](file:///C:\Users\adan\Desktop\VIDEOS%20CREXCEL\Proyecto%20Curso%20Excel\muestra%20de%20relaciones%20ok.mp4)

1. Por último, tenemos la relación muchos a muchos. Pongamos como ejemplo que una OBRA puede ser auditada por varios supervisores, y un SUPERVISOR puede auditar varias OBRAS. Este caso la relación se resuelve creando una tabla “puente” la cual únicamente contendrá la clave de las tablas a relacionar.

Finalmente aclaremos: es muy probable que pienses que sí, está muy bien todo esto, pero quieres un reporte semanal en donde el jefe te pide colores, campos de varias tablas, subtotales, gráficos, etc. Entonces cómo lo vas a hacer si lo único que tienes son tablas simples por separado. Bien, ahí está precisamente el potencial de todo esto, pues se trata de que NO manipules directamente la base de datos, los reportes son únicamente el reflejo de la base, y puedes hacer cientos de ellos sin que los datos originales sufran ningún cambio por motivo de formato.

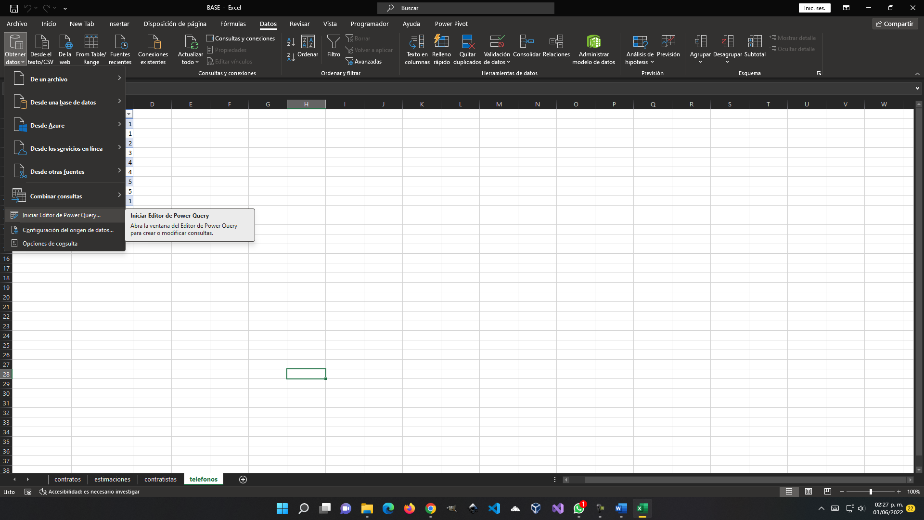
En los próximos videos veremos cómo crear esos reportes junto con gráficos y otros vistosos objetos, sólo te pedimos que sigas estudiando y practicando.

[Bonus video: tablas integradas](file:///C:\Users\adan\Desktop\VIDEOS%20CREXCEL\Proyecto%20Curso%20Excel\tablas%20integradas.mp4)

ENTENDIENDO EL MODELO DE DATOS

Para no confundirnos con información técnica sobre los objetos de Excel orientados a datos, diremos que el modelo de datos es un conjunto de objetos que el programa ha optimizado de manera que los trabaje de forma eficiente. Recordemos que muchas veces los datos de origen no están normalizados, así que Excel los interpreta lo mejor que puede para luego trabajar con ellos.

Cuando iniciamos algún proyecto de acceso a datos lo primero que hacemos es abrir o vincular de alguna manera la fuente de los mismos. Esta acción lo realizamos con Power Query, esta herramienta se encuentra ahora integrada al programa, en realidad a la hora de utilizar la barra de datos estamos utilizando el Set de Power Query.



La interface de usuario de Power Query tiene tres objetivos principales: primero, realizar acciones generales sobre los datos a modo de preparación, por ejemplo agregar la primera fila como encabezados de columna. Segundo, transformar los datos actuales, por ejemplo pone formato de fecha a una columna que subió como texto. Y tercero, agregar datos nuevos, como por ejemplo una columna con alguna función integrada.

[Video: Conexión a la base de datos](file:///C:\Users\adan\Desktop\VIDEOS%20CREXCEL\Proyecto%20Curso%20Excel\powque1\powque1.mp4)