

Programas de Cómputo (Base de Datos) para el Manejo de Registros

Mario García Podestá, MSc, PhD

*Facultad de Veterinaria y Zootecnia,
Universidad Peruana Cayetano Heredia. Lima-Perú
garciam@terra.com.pe, m_garcia_p@hotmail.com*

EL DATO Y LA INFORMACIÓN

Dato, según el Diccionario de la Lengua Española significa “Antecedente necesario para llegar al conocimiento exacto de algo”. En base a esta definición se puede deducir que un dato es sólo un medio (un camino) para conocer u obtener el significado de algo (el objetivo), y ese significado es lo que se denomina información. No obstante, esta diferencia entre “dato” e “información”, estos términos son generalmente utilizados indistintamente, de allí que sea importante hacer una distinción formal en el contexto de bases de datos y tecnología de almacenamiento de datos. La información tiende a ser subjetiva, debido a que el proceso de percepción y decisión es realizado por personas, en tanto que el dato es (o debería ser) objetivo. En otras palabras, el dato procesado viene a ser la información.

La Base de Datos. La base de datos es el conjunto de datos organizados en una forma tal que permite extraer la información. Si bien, esta definición no implica el medio de almacenamiento de los datos, los medios más adecuados son los computarizados (magnéticos y ópticos). El origen de las bases de datos se remonta a los registros gubernamentales, médicos, comerciales y de bibliotecas de épocas pasadas. En la era actual, las computadoras han permitido realizar un importante avance en el manejo de la información, aunque el software inicial enfatizaba el manejo de los registros sin darle importancia a la estructura del sistema. Recién a inicios de los '70 se desconectó la organización lógica del sistema del método físico del almacenamiento, dando nacimiento a las bases de datos relacionados (Relational Database Management System - RDBMS), sistema que perdura hasta la fecha.

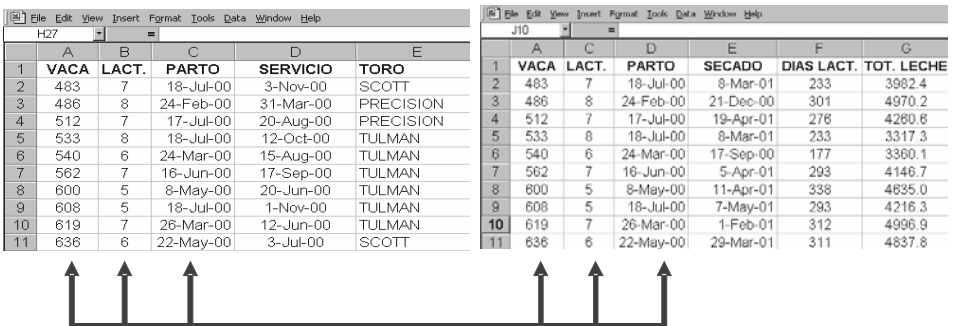
No fue sino hasta mediados de los '80 que los programas comerciales de base de datos comenzaron a hacerse populares (RIM, RBASE 5000, PARADOX, OS/2 Database Manager, Dbase III), para hacerse más complejos y costosos en la década del '90 [PowerBuilder (Sybase), Oracle Developer, VB (Microsoft), etc]. También fue el inicio del desarrollo de herramientas de productividad personal como Excel/Access (Microsoft) y ODBC, y de los prototipos de Object Database Management Systems (ODBMS). El fin del siglo pasado coincidió con el boom del desarrollo de las bases de datos en el Internet (Active Server Pages, Front Page, Java Servlets, JDBC, Enterprise Java Beans, ColdFusion, Dream Weaver, Oracle Developer 2000).

¿SE REQUIERE UNA BASE DE DATOS?

La respuesta está dada en el tipo y cantidad de datos por colectarse, la magnitud de registros por considerarse y el periodo de tiempo en que se recolectará. La forma más fácil (y esto expresado en una forma relativa) y común de almacenar información en pequeña escala, utilizada por estudiantes y profesionales del campo agrario es a través de hojas de cálculo. La facilidad radica en que el usuario sólo tiene que utilizar un programa comercial como Excel (Microsoft), abrir un archivo nuevo, colocar el nombre de las variables en las columnas y comenzar a digitar los datos de cada registro. Este tipo de formato es ideal y oportuno para el manejo de información con pocas variables (ejm. pesos de carcasas en centros de beneficio), o cuando esta información es colectada en un periodo breve y cada observación o situación se introduce en forma total (ejm. encuestas). Sin embargo, son las que ofrecen el mayor nivel de error en el ingreso de la información al carecer de filtros y otros medios verificadoros; su uso se dificulta cuando se tienen que ubicar observaciones previas para añadir información adicional (ejm. registros de producción de leche), fuera de que es el peor sistema para almacenamiento de texto. Por otro lado, la interrelación entre registros es compleja, difícil y por lo tanto, se tiende a repetir el ingreso de información básica para cada registro.

Figura 1

Registros electrónicos en hojas de cálculo para el almacenamiento y análisis de la producción de leche y de los servicios de inseminación artificial en un establo lechero. Nótese la duplicación de información



En la Figura 1 se puede observar como ejemplo los archivos de una empresa pecuaria en formato de hoja de cálculo. Se tiene el registro para acumulados individuales de producción de leche y el registro de servicios de inseminación artificial, donde la identificación de la vaca, número de lactancia y fecha de parto se encuentran repetidos. La integridad de este tipo de información duplicada es endeble pues un cambio en una hoja no necesariamente se ejecuta en la otra, de allí que ocurra con frecuencia que los diferentes usuarios (productor, administrador, veterinario, zootecnista, capataz) no tengan las mismas vacas, el mismo número de vacas, o las mismas fechas de parto en los diferentes listados (hojas de cálculo).

Una regla sencilla que puede ayudar en la toma de decisiones para utilizar la hoja de cálculo o las bases de datos como medio para el almacenamiento y análisis de datos queda resumida como sigue: Si se tiene un reducido número de variables, un limitado número de registros y los datos pertenecientes a un registro son recolectados en un momento y no son repetitivos con el tiempo se recomienda el uso de hojas de cálculo; caso contrario nos inclinamos por las aplicaciones de base de datos.

CARACTERÍSTICAS Y VENTAJAS DE UNA APLICACIÓN CON BASE DE DATOS

Interfaz amigable con el usuario. El software utilizado por las bases de datos permite utilizar una amplia gama de herramientas gráficas, imágenes, fondos, tipos de fuentes, etc. que pueden satisfacer las necesidades del cliente y facilitar, a través del diseño de las pantallas, el ingreso de los datos. La creación de menús, de botones de acción, de mecanismos de búsqueda y de presentación de los campos (espacios dedicados para el ingreso de la información de cada variable) permite desarrollar un medio ambiente de satisfacción entre el sistema y el usuario.

Validación de los datos por el sistema. El ingreso de datos puede ser verificado por una serie de parámetros especificados por el usuario. Por ejemplo, se puede impedir el ingreso de información adicional si el animal tiene fecha de muerte o de venta. La creación de estas restricciones, usualmente tomando como criterios a los rangos biológicos (ejm. El peso del ternero al nacimiento debe estar entre 25 y 55 kg), intervalos de tiempo (ejm. La fecha del peso al año de edad debe estar entre los 11 y 13 meses posteriores a la fecha de nacimiento del animal), y a criterios de sentido común (ejm. la fecha del servicio de inseminación artificial debe ser posterior a la fecha del parto) se deben hacer mediante un trabajo conjunto entre el programador y el profesional pecuario responsable. Estas restricciones, llamadas reglas de validación pueden ser colocadas a diferentes niveles (antes de pasar al siguiente campo, antes de salvar el registro, antes de cerrar la aplicación, etc.) y generar mensajes de error genéricos o específicos (Figura 2).

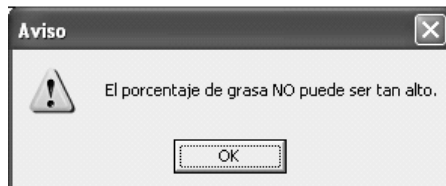
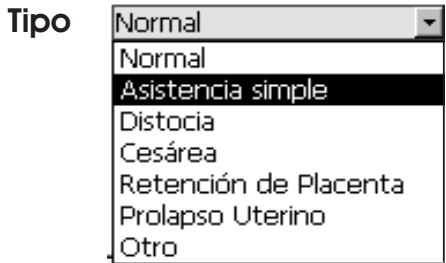


Figura 2
Típico mensaje de error

Utilización de información pre-determinada

Permite el uso de listados de opciones pre-determinadas por el programador (en variables donde se conozca las posibles respuestas como puede ser el N° de parto, sexo, etc.) o fijadas por el propio usuario para el llenado de información en determinadas variables (Figura 3). En general, se puede considerar que entre el 40 y el 70% de las variables utilizadas pueden ser cubiertas por este tipo de herramienta, lo cual reduce enormemente los errores de tipeo y de valores no esperados. Además, facilita el ingreso de datos ya que es más fácil y rápido la selección de una opción que el ingreso manual de la misma.

Figura 3
Formato de opciones de tipo de parto



No permite el ingreso de registros idénticos

Se pueden definir variables en determinadas tablas cuyos valores no pueden duplicarse. Por ejemplo, en una tabla de identificación de toros, el número o nombre del animal no puede figurar en dos registros. Sin embargo, esto no implica que en una tabla relacionada, el nombre del toro pueda repetirse muchas veces (por ejemplo, en la tabla de pesos mensuales).

No hay restricciones para la cantidad de entradas (N° de registros). A diferencia de otro tipo de aplicaciones (por ejemplo, las hojas de cálculo), los sistemas de base de datos no están seriamente limitados por el número de variables o por la cantidad de registros. Aplicaciones de base de datos desarrollados en programas comerciales como MS Access permiten crecer hasta 2 giga bites o tener cerca de 1,000 formularios y reportes, lo cual en términos generales es mucho más de lo que necesita la gran mayoría de usuarios. Bases de datos más potentes como Oracle escapan de estas limitaciones y permiten crecer en forma ilimitada. En realidad, los límites están fijados por la capacidad física disponible en la computadora o del servidor.

Permite el almacenamiento de imágenes, sonido y video

Los tipos de datos más generalizados son los de texto, número, fecha, dinero y verdadero / falso en sus diferentes modalidades; sin embargo, es factible ingresar imágenes (fotos), sonidos y videos como parte de los registros, y levantar este tipo de información a través de procesos de selección o búsqueda, utilizando los mismos mecanismos existentes para texto o números (por ejemplo, ubicar el registro de una muestra de semen y la imagen escaneada del PCR respectivo).

ANIMAL	
Productor: La Estanzuela	
Bovine	
Identificación	
Nº / Nombre	1026
Productor	La Estanzuela
Sexo	Hembra
Raza	Holstein
Tipo Racial	Puro por cruce
Nacimiento	
Fecha	05-Mar-01
Peso	
¿Parto Múltiple?	<input type="checkbox"/>
¿Nacido aquí?	<input checked="" type="checkbox"/>
Muerte / Saca	
Descarte	<input type="checkbox"/>
Fecha	
Grupo	
Grupo	*****
Experimento	No
Observ.	

Actualización y eliminación de registros relacionados. Estos sistemas permiten realizar cambios globales que afectan información almacenada en diferentes tablas relacionadas. Por ejemplo, el nombre de una vaca cuya información se encuentra en las tablas de partos, de servicios de inseminación artificial, de producción de leche, de ocurrencias clínicas, de consumo de concentrado y de control de pesos podría ser actualizado con sólo cambiar su identificación en la tabla de animales. También es factible y de acuerdo a la programación realizada, eliminar todos los registros de un determinado animal con sólo eliminar su registro de identificación en la tabla de animales.

Utilización de valores por defecto. Se puede prefijar valores por defecto a algunas variables de importancia con valores de escasa variabilidad. Por ejemplo, en el caso de un estudio de centros de engorde, la variable de región de procedencia de los animales podría estar prefijada para “Sierra”, de modo que al abrir un registro nuevo, esta información aparezca de inmediato, lo cual facilita el ingreso de la información de campo. No obstante, y en caso de ser necesario, este valor puede ser cambiado por el usuario y colocar el valor apropiado.

Codificación de los datos. La representación visual en letras es más entendible para el usuario, tanto para el que ingresa la información como para el que realiza la consulta o búsqueda; sin embargo, el sistema de base de datos es mucho más eficiente (en términos de velocidad, accesibilidad y espacio de disco duro) si trabaja con números (por ejemplo, al revisar la relación de razas de bovinos, es preferible encontrar valores como Holstein, Brown Swiss, Brahman, Guzerat, etc., pero el sistema se optimiza con valores de 1, 2, 3, 4, 5, etc. en lugar del nombre de las razas). Este problema se puede resolver fácilmente pues estos programas permiten codificar la información haciendo que el sistema muestra texto (nombres de razas) pero en realidad se está almacenando números (códigos de razas).

Fácil manejo de los datos. El manejo de la información para la búsqueda de registros específicos, selección de grupos de registros en base a determinado valor o rango de valores, o para la creación de informes, sean estos de datos crudos o informes

con acabados profesionales, se pueden desarrollar fácilmente. Es más, las aplicaciones creadas con el debido profesionalismo deben incluir estas herramientas con acceso directo del usuario a través de botones de selección.

Indexación de los datos. El sistema de base de datos permite indexar variables que son o pueden ser requeridas en forma constante para agrupar u ordenar registros. Esta herramienta mejora la eficiencia de la base de datos, ahorrando tiempo en el análisis; factor que fácilmente se aprecia cuando se trabaja con miles de registros.

Multiusuario. Estas aplicaciones, si bien pueden ser utilizadas por un único usuario, están desarrolladas para ambientes de multiusuarios. Decenas y centenas de usuarios pueden estar ingresando, re trayendo y analizando la información en forma simultánea desde diferentes máquinas enlazadas por un LAN -local area network- o en una Intranet, así como a través del Internet. Así mismo, otra gran ventaja es que se pueden crear niveles de acceso que pueden ir desde la posibilidad de modificar el diseño y estructura del sistema, pasando por la autorización de ingreso y modificación de los registros, hasta la capacidad de 'sólo-lectura'. Igualmente, se pueden desarrollar áreas o componentes de acceso a determinados grupos de usuarios, con uso de contraseñas para limitar el ingreso. Por ejemplo, los estudiantes pueden tener acceso a la lectura de las historias clínicas y los resultados de laboratorio, pero no tener acceso a la información personal de los dueños de las mascotas.

SITUACIÓN ACTUAL DE LAS APLICACIONES CON BASE DE DATOS EN EL CAMPO DE LA INVESTIGACIÓN Y PRODUCCIÓN PECUARIA

La mayor parte de los proyectos, muchos experimentos y el manejo de la información pecuaria a nivel de finca requieren de una aplicación con base de datos, que sea especialmente desarrollado para cubrir sus necesidades inmediatas y futuras. El registro de datos en tarjetas, fichas y cuadernos, así como el uso de las mal llamadas 'bases de datos' en hojas de cálculo y hasta en procesadores de texto debería ser cosa del pasado, pero lamentablemente son las formas más comunes en nuestro medio. Esta metodología de trabajo, que de ser bien llevada, puede en algunos casos resolver nuestras necesidades, debería ser dejada de lado y emplearse sistemas con base de datos *ad hoc*.

Laboratorios de diagnóstico, centros de producción, bancos de semen, estaciones experimentales, unidades de producción incluidas las fincas productoras de ganado lechero, de carne o de doble propósito, bancos de sueros, y muchos otros que almacenan sus resultados en docenas de diskettes o CD's, pierden grandes oportunidades de realizar evaluaciones y análisis históricos, toda vez que los registros se encuentran distribuidos en cientos de archivos y muchas veces en diferentes formatos.

La experiencia demuestra que instituciones o grupos de trabajo, que organizaron su metodología de trabajo y desarrollaron un eficiente sistema informático con base de datos, lograron automatizar el almacenamiento de la información, minimizando el tiempo y esfuerzo que demanda el ingreso de los datos a las computadoras y redujeron el error proveniente del tipeo, la duplicidad de registros y datos fuera de los rangos biológicos permisibles. Asimismo, el análisis de la información se hizo más fácil y versátil pues puede re-analizarse bajo diferentes ópticas y en cualquier momento.

LIMA (Livestock Information Management Database) es un claro ejemplo de una aplicación, disponible en español e inglés, y de distribución gratuita, que sirve para el manejo integral de la información de establecimientos ganaderos (ganado vacuno, ovino, caprino, bubalino, y de camélidos sudamericanos) y está permitiendo cuantificar una serie de deficiencias tanto en el manejo reproductivo como en la producción de leche en diferentes países en desarrollo, tanto a nivel de finca como de cuenca.

LECTURAS RECOMENDADAS

Balter A. Mastering Microsoft Access 2000 Development. Sams Publishing. USA. 1342p. 1999.

García M. User manual for AIDA (Artificial Insemination Database Application). 36p. 1996.

García M, Goodger WJ, Bennett T, Perera BMAO. Establishing current status and constraints of AI Services in Asia and Latin America using progesterone radioimmunoassay. Proc. Final Research Coordination Meeting of the FAO/IAEA Coordinated Research Project on "Use of RIA and Related Techniques to Identify Ways of Improving Artificial Insemination Programmes for Cattle under Tropical and Sub-Tropical Conditions". Uppsala, Suecia. 1999.

García M. LACALINIA. Manual de Usuario. 37p. 2001.

García M. SPeRM (Semen Processing Records Management). User Manual for SPeRM. Universidad Peruana Cayetano Heredia. Lima-Perú. 60p. 2002.

García M. LIMA (Livestock Information Management Application). User Manual for Lima. Universidad Peruana Cayetano Heredia. Lima-Perú. 2ª ed. 65p. 2003.

Saleme C, Saavedra E. Base de datos. Pontificia Universidad Católica. Lima-Perú. 2002.

Scott Barker F. Microsoft Access. Power Programming. Sams Publishing. USA. 1334p. 1999.