

EXCEL COMO AUXILIAR EN LA GENERACIÓN DE REPORTES EN SISTEMAS DE INFORMACIÓN

EXCEL AS AN AUXILIARY IN THE GENERATION OF REPORTS IN INFORMATION SYSTEMS

Eleazar Rosas Centeno

Universidad de Guanajuato
e.rosascenteno@ugto.mx

María Susana Ávila García

Universidad de Guanajuato
susana.avila@ugto.mx

María Isabel Vázquez de la Rosa

Universidad de Guanajuato
mi.vazquez@ugto.mx

Hugo Armando Aguilera García

Universidad de Guanajuato
ha.aguileragarcia@ugto.mx

Resumen

Los sistemas de información ofrecen la funcionalidad de analizar información y generar reportes que sean de interés a los varios niveles gerenciales y de administración de una empresa u organización. Sin embargo, nuevas necesidades de información generan nuevos requisitos de software y por lo tanto la necesidad de extender la funcionalidad de estos sistemas, lo que a su vez tiene costos y riesgos de desarrollo de software asociados. En este trabajo se propone usar la herramienta de Excel como un auxiliar en la generación de reportes que puedan asistir la administración de información. Nuestra propuesta incluye un flujo de trabajo en el que existen dos transformaciones de información: (a) de una base de datos no relacional a una relacional, es decir de MongoDB a MySQL, (b) de una base de datos relacional a una hoja de cálculo, es decir de MySQL a Excel. Finalmente, en Excel se propone el uso de tablas pivote para asistir el análisis dinámico de la información. Como caso de estudio se presenta el Sistema de Registro de Asistencias del Centro de Auto-Aprendizaje de Idiomas del

Departamento de Estudios Multidisciplinarios de la Universidad de Guanajuato en la Sede Yuriria.

Palabra(s) Clave: MongoDB, MySQL, Tablas pivote, VBA, Visual Basic para Aplicaciones.

Abstract

Information systems offer the functionality to collect, store, and analyses data and generate reports for a business or organization. However, new needs of information generate new software requirements and therefore the need to extend the functionality of the systems incurring in additional software development costs and risks. This work proposes the use of Microsoft Excel as a tool to assist the analysis of data and the generation of new reports providing meaningful information to assist the decision making in the administration of the center. Our proposal comprises two transformations of the data: (a) from a non-relational database to a relational database, and (b) from a relational database to Microsoft Excel. We take as a case of study the attendance registration systems of the Self Access and Learning Language Centre at the Department of Multidisciplinary Studies of the University of Guanajuato in Yuriria.

Keywords: MongoDB, MySQL, Pivot tables, VBA, Visual Basic for Applications.

1. Introducción

Excel es una herramienta popular que permite analizar datos, realizar cálculos y generar estadísticas de interés para diferentes usuarios. Las tablas y gráficas pivote son una herramienta que permite el análisis de la información de manera dinámica facilitando la generación de estadísticas. Esta flexibilidad es algo de lo que algunos sistemas de información carecen ya que su diseño y desarrollo considera las necesidades iniciales de los *stakeholders*, de tal manera que cuando nuevas necesidades de analizar y presentar la información surgen los cambios en los sistemas de información implican un esfuerzo considerable en su desarrollo. En este trabajo presentamos el caso de estudio del Sistema de Registro de Asistencias al Centro de Auto-Aprendizaje Dirigido de inglés reportado en [Rodríguez, Contreras,

García, & Rosa, 2016]. Este sistema fue desarrollado en Meteor JS, usando una base de datos no relacional en MongoDB para ofrecer la funcionalidad de registrar las asistencias de los estudiantes al centro y generar reportes de asistencia de estudiantes durante el semestre en curso a profesores que incluyen en sus criterios de evaluación la asistencia a dicho centro.

Mientras los reportes generados por el sistema han cumplido con las necesidades iniciales de la administradora del centro, la acumulación de datos históricos sobre la asistencia de estudiantes al centro abre la posibilidad de analizar estos datos y generar estadísticas interesantes no solo para la administradora, sino para la institución. Sin embargo, este tipo de análisis no es posible realizarlo directamente en el sistema por no contar con esta funcionalidad.

Por otro lado, Excel es una herramienta de la suite de Microsoft Office de fácil acceso y popular en el análisis de datos. Esta herramienta ofrece varias funciones para generar datos estadísticos y objetos útiles en el análisis dinámico de la información como las tablas y gráficas pivotas. Esta herramienta ha sido exitosamente reportada en [Lin, Cheng, & Gwynedd, s.f.]. En esta investigación se muestra como las tablas pivotas son una poderosa herramienta de Microsoft Excel utilizada para crear tablas interactivas que extraen, organizan y resumen automáticamente grandes cantidades de datos. Además, los informes de tabla dinámica se pueden usar para analizar datos, hacer comparaciones, detectar patrones y relaciones y descubrir tendencias. Se ha descubierto que las tablas son una herramienta muy útil para comunicarse con los clientes dentro del área del análisis estadístico de datos en el área farmacéutica esto debido a que es rápido, fácil e interactivo. Se utiliza principalmente estos informes de tabla dinámica para el análisis exploratorio antes de que se generen las tablas de informes formales.

Las tablas pivotas son flexibles son creadas con la finalidad de crear estadísticas dinámicas sobre los aspectos de interés. Estas tablas han sido empleadas en diferentes áreas de aplicación, como medicina, debido a que la medicina genera grandes volúmenes de datos que pueden llegar a ser opacos. Como es referido en [Kang, 2016] las tablas pivotas se han utilizado para el estudio de datos médicos por diferentes razones, una de ellas es que la suite de office es ampliamente utilizada

en todos los dominios además que las hojas de cálculo pueden ser compartidas de un investigador a otro. Como resultado de esta investigación se han podido registrar los datos de los pacientes de un hospital, así como la bitácora con su diferente sintomatología, lo cual ha permitido a los doctores y enfermeras visualizar y analizar de una manera clara y fácil la información referente a los pacientes.

Otra característica propia de las tablas pivote en Excel es la facilidad de creación lo que permite ser utilizadas por personas con una formación básica en la suite de Microsoft Office. Por su parte Dierenfeld & Merceron [2012] reportan como en el área de docencia se analiza el nivel de aprendizaje de los alumnos a través de una plataforma llamada Moodle. El objetivo es conocer si los estudiantes utilizaban los recursos de la plataforma, cuando la utilizaban y cuáles eran los recursos más utilizados. Para ello se realizó una exportación de los datos de la plataforma Moodle hacia hojas de cálculo en Excel para hacer el análisis. Como resultado se lograron crear diferentes estadísticas, una de ella con las actividades que sus alumnos obtenían en los test de la plataforma, obteniendo el número exacto de objetivos alcanzados o no alcanzados y a partir de ahí tomar decisiones para mejorar su curso.

En este trabajo presentamos una propuesta para definir una solución a la rigidez que pueden presentar los sistemas de información para el análisis dinámico de datos por medio de transformaciones de datos y el uso de Excel.

2. Métodos

El desarrollo de este trabajo siguió una metodología de desarrollo clásica: definición de requisitos, diseño, codificación, y pruebas.

Análisis de Requisitos

El análisis de requisitos se realizó con la *stakeholder* principal del proyecto, la administradora del Centro de Auto Aprendizaje de Idiomas (CAADI). En una entrevista semi-estructurada se recolectó información sobre sus necesidades en cuanto a la generación de reportes. Como necesidades principales se marcaron el asistir los reportes que tiene que generar en auditorías sobre el funcionamiento del

centro. Además de identificar las necesidades de generación reportes sobre las visitas de los alumnos que el CAADI que incluyen:

- Horarios en los que más se visita el CAADI.
- Actividad que más se realiza en el CAADI.
- Edad de las personas que más asisten al CAADI.
- Número de alumnos que asisten por carrera.
- Porcentaje de visitas por género.

Diseño

El flujo de trabajo ilustrado en la figura 1 muestra los pasos principales de la solución propuesta. El primer paso consiste en transformar los datos de la base de datos en Mongo DB a MySQL. Esto se realizó utilizando el software Pentaho Data Integration. Se realizó un mapeo de las colecciones de MongoDB con tablas en MySQL.

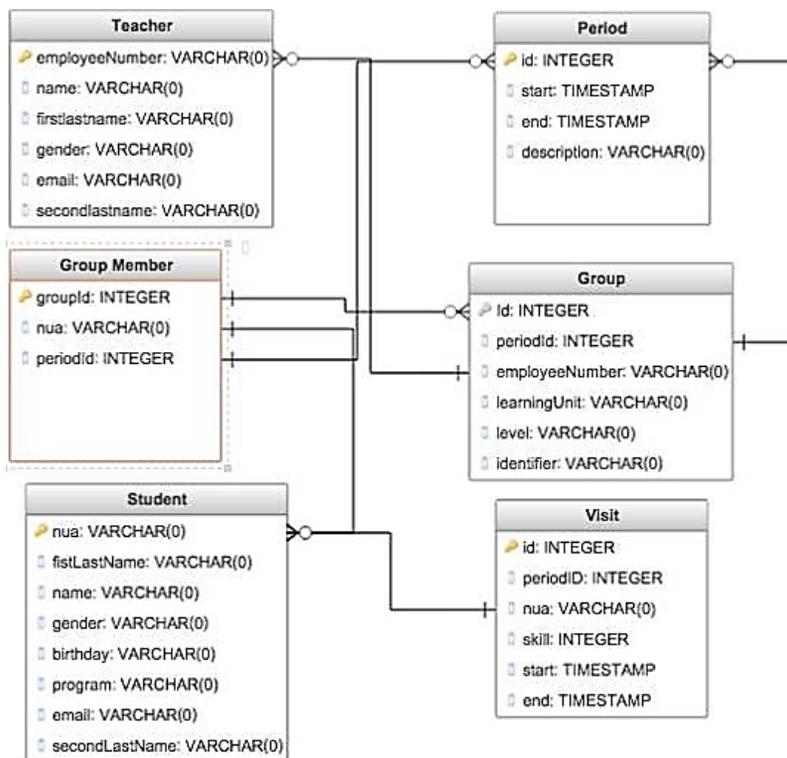


Fuente: Elaboración propia.

Figura 1 Flujo de Trabajo Propuesto.

El diseño de las colecciones de Mongo DB se pueden encontrar en [Rodríguez *et al.*, 2016] y el de la base de datos en MySQL se muestra en la figura 2. El segundo paso consiste en la transformación de los datos de la base de datos en MySQL a

archivos de valores separados por coma (CSV) que puedan ser leídos en Excel. Esta transformación se basó en la propuesta reportada en [Kamalakant L Bawankule, 2014] donde se explica la utilización de la librería Poi para la creación de documentos en Excel a partir de la selección de datos en MySQL. Finalmente, el tercer paso consiste en leer los datos en Excel y crear las tablas pivote para que el usuario pueda crear las estadísticas y reportes necesarios de manera dinámica, de acuerdo a lo indicado en las figuras 1 y 2.



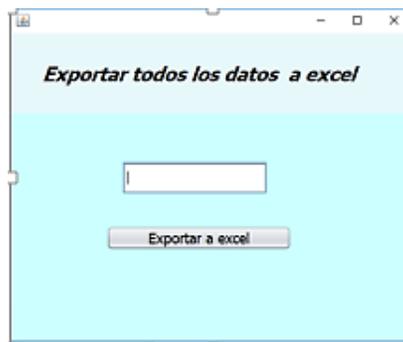
Fuente: Elaboración propia.

Figura 2 Esquema que integra los datos.

Desarrollo

La solución propuesta en este trabajo incluye el desarrollo de una aplicación en Java usando la librería Poi haciendo una conexión al manejador de base de datos de MySQL con una interfaz que permita visualizar los atributos de las tablas que se pasarán al archivo CSV, así como de pasar las tablas de manera más rápida sin visualizar el contenido. El desarrollo se realizó en Java usando el IDE de Netbeans. El sistema ofrece dos opciones: migración de los datos de las consultas de unión

de las tres tablas, y la migración de las tablas por separado. En la primera opción, el programa exporta todos los datos a Excel realizando una consulta SQL la cual selecciona las principales tablas: Student, Visit y Period. Estas tablas se utilizarán para exportar datos en un archivo de Excel y poder generar reportes, lo cual se ha simplificado con un solo botón que hace toda la exportación de las tres tablas necesarias para realizar las estadísticas para el CAADI, como se presenta en la figura 3.



Fuente: Elaboración propia.

Figura 3 Exportación de todos los datos.

El sistema también permite la opción de exportar las tablas en diferentes hojas de cálculo en Excel, una para cada tabla, indicando la ruta donde se guardará los archivos de Excel de manera estática, esto para que el usuario pueda llevar un mejor control del repertorio dentro de su computadora donde se estará generando los reportes, con la interfaz de la figura 4. La creación de la tabla pivote y de la generación de estadísticas necesitadas por la administradora se realizó usando Visual Basic para Aplicaciones.



Fuente: Elaboración propia.

Figura 4 Exportación por separado de todos los datos.

3. Resultados

La solución propuesta ha permitido dar solución a las necesidades de información de la administradora del CAADI. El resultado de las transformaciones permitió obtener hojas de cálculo con la estructura mostrada en las figuras 5 y 6.

Nua	Name	Primer Apellido	Segundo Apellido	Programa	Genero	Cumpleaños	Email
-----	------	-----------------	------------------	----------	--------	------------	-------

Fuente: Elaboración propia.

Figura 5 Estructura de datos de alumnos en Excel.

Nua	Periodo	Habilidad	Inicio	Fin	Descripción
-----	---------	-----------	--------	-----	-------------

Fuente: Elaboración propia.

Figura 6 Estructura de periodo y visitas de alumnos al CAADI en Excel.

Las tablas pivote permiten responder a varias preguntas propuestas por la administradora, las cuales se definieron al principio de la metodología como:

- Horas que más se visita el CAADI. Mediante la visualización de una tabla pivote es necesario seleccionar todos los datos que se exportaron anteriormente, enseguida seleccionar el campo "Cuenta de id" en el área de valores" y el campo "inicio" en el área inicio, de esta manera se obtiene el conteo de las horas y minutos en la que más se visita el centro.
- Actividad que más se realiza en el CAADI. Para poder visualizar esto se siguen los pasos del primer inciso, utilizando el id en el área "VALORES" de la tabla pivote y utilizando el campo "habilidad" en el área "FILAS" esto nos dará un conteo de cuál es la actividad más realizada en el centro.
- Edad de las personas que más asisten al CAADI. Usando tablas pivote es posible obtener una estadística de cuál es la edad que más se repite dentro de los alumnos del centro esto usando con el mismo método el campo "id" y "cumpleaños".
- Número de alumnos que asisten por carrera. Se puede visualizar la estadística mediante los campos "id" y "Programa educativo". Porcentaje de visitas por género.

Las estadísticas que se realizaron mediante el uso de VBA y las PivotTables se obtienen mediante la interfaz de la figura 7.



Fuente: Elaboración propia.

Figura 7 Estructura de periodo y visitas de alumnos al CAADI en Excel.

4. Discusión

En este proyecto se buscó utilizar todas las herramientas accesibles al departamento del CAADI de una manera que se pudieran crear reportes que pudieran asistir a la administradora del centro en sus reportes.

Una de las razones por las que se decidió usar MySQL para los reportes es porque se tiene contemplada una migración del sistema del de registro de asistencias de CAADI y la migración y transformación de los datos de MongoDB a MySQL es una prioridad, por lo que la transformación y migración de datos desde el gestor de MySQL a Excel es algo que permanecerá en el flujo de tareas en el futuro.

Se ha integrado una de las características más importantes de Excel, las tablas pivotes que son de fácil creación y fáciles de manejar nos han permitido obtener las estadísticas necesarias para la *stakeholder* del proyecto, lo cual significará una optimización al centro de aprendizaje y mejor calificación en futuras auditorias.

Mientras la aplicación en VBA para la creación de las tablas pivote y las estadísticas la migración de los datos a Excel permitirá que el usuario pueda realizar sus propias configuraciones de las tablas pivote para generar estadísticas y reportes dinámicos

sin estar limitados a la rigidez de un sistema de información. Esta herramienta permitirá hacer un análisis que ayudará a tomar decisiones sobre la distribución del material didáctico y las actividades dentro del área que haga falta optimizar.

5. Bibliografía y Referencias

- [1] Dierenfeld, H., & Merceron, A. (2012). Learning Analytics with Excel Pivot Tables. 1st Moodle Research Conference, 115-121: <http://research.moodle.net/mod/data/view.php?d=3&rid=15>.
- [2] Kamalakant L Bawankule, N. B. R. (2014). Design and implementation of massive MYSQL data intelligent export system to excel by using Apache's POI libraries. IOSR Journal of Computer Engineering (IOSR-JCE), 16(5), 58-65: <http://www.iosrjournals.org/iosr-jce/papers/J016545865.pdf>
- [3] Kang, S. (2016). Research in brief. The Lancet Diabetes & Endocrinology, 4(1), 22. [https://doi.org/10.1016/S2213-8587\(15\)00486-6](https://doi.org/10.1016/S2213-8587(15)00486-6)
- [4] Lin, W. L., Cheng, J., & Gwynedd, U. (s.f.). Pivot Table's A Powerful Tool for Data Exploration in the Pharmaceutical Industry, (c), 1-8.
- [5] Rodríguez, J. B. R., Contreras, I. C., García, M. S. Á., & Rosa, M. I. V. de la. (2016). Desarrollo de un Sistema de Registro de Asistencias para Mejorar las Prácticas de Trabajo del CAADI Yuriria. *Pistas Educativas*, 38(122).