

## Cubo de Información para el Análisis de Indicadores de Ingreso en Programas Acreditados Data Cube used to Browse Incoming Students Indicators on Accredited Programs

Villegas Téllez, R.<sup>1</sup>, Torres Frausto, D.A.<sup>2</sup>, Soto Patiño, J.C.<sup>3</sup>

<sup>1</sup> Ingeniería Informática, Instituto Tecnológico Superior de Irapuato  
Carretera Irapuato - Silao km 12.5, El Copal, 36821 Irapuato, Guanajuato. México.

<sup>2</sup> Ingeniería Informática, Instituto Tecnológico Superior de Irapuato  
Carretera Irapuato - Silao km 12.5, El Copal, 36821 Irapuato, Guanajuato. México.

<sup>3</sup> Departamento de Tutorías, Instituto Tecnológico Superior de Irapuato  
Carretera Irapuato - Silao km 12.5, El Copal, 36821 Irapuato, Guanajuato. México.  
<sup>1</sup>rovillegas@itesi.edu.mx, <sup>2</sup>datorres@itesi.edu.mx, <sup>3</sup>jusoto@itesi.edu.mx

Fecha de recepción: 6 de junio 2017

Fecha de aceptación: 18 de agosto 2017

**Resumen.** En el Instituto Tecnológico Superior de Irapuato el Departamento de Tutorías procesa una gran cantidad de información relacionada con los alumnos de nuevo ingreso, tal como escuela de procedencia, nivel socio-económico, datos del examen de ingreso CENEVAL, exámenes psicométricos, entre otros rubros. Toda esta información es muy valiosa para la institución, ya que le permite conocer mejor a sus estudiantes y guiarlos adecuadamente durante su estancia en la institución.

El presente trabajo expone el desarrollo de una base de datos Data Warehouse usada para almacenar todos los datos que provienen de las diversas fuentes de información. Esta base de datos es la base del *cubo de información* que le permite al Departamento de Tutorías un análisis rápido que agiliza y sustenta la toma de decisiones, tal como el otorgamiento de becas, estrategias para disminuir el índice de deserción, desarrollo de plataformas de estudio, entre otras más.

**Palabras Clave:** Alumnos de Nuevo Ingreso, Cubo de Información, Data Warehouse, Departamento de Tutorías.

**Summary.** The Instituto Tecnológico Superior de Irapuato Tutorial Center processes a large amount of information related to incoming students, such as high school data, socioeconomic level, CENEVAL and psychometric tests among other items. All this information is very valuable for the institution, since it allows a better knowledge of the students and guide them properly throughout their stay in the institution.

This paper presents a Data Warehouse database development used as a storage of multiple data sources. This database is required by a data cube that allows the Tutorial Center a quick analysis to streamline and support decision options, such as grants and scholarships, strategies to decrease the dropout rate, development of study platforms, and more.

**Keywords:** Data Cube, Data Warehouse, Incoming Students, Tutorial Center.

## 1 Introducción

En su libro de Los Numerati (Baker, 2009) piensa que somos prolíficos en lo que se refiere a la producción de datos. Engrosamos nuestros expedientes digitales todos los días por el simple hecho de usar teléfonos celulares, laptops y tarjetas de crédito. Todo comenzó con los chips de las computadoras. Hasta la década de los ochenta, estas pequeñas piezas de silicio que contenían millones de transistores microscópicos aun eran una novedad. Pero con el paso de los años se abarataron y se desarrollaron aún más, y ahora los fabricantes los incluyen en prácticamente todo lo que pueda beneficiarse de su tecnología.

Las únicas personas capaces de dar sentido a los datos que generamos son los mejores matemáticos, científicos e ingenieros de la computación. Ellos saben cómo convertir en símbolos los fragmentos de nuestra vida. ¿Por qué es necesario esto? Imagina que deseas llevar un registro de todo lo que comes en un año. Si eres como yo era en cuarto de primaria, vas a la papelería y compras un paquete de tarjetas, en cada una de las cuales anotas una comida. Pastel de carne. Espinacas. Tapioca. Cereal. Al cabo de unos días tendrás una pila de tarjetas cada vez más grande. El problema es que no hay manera de contarlas o analizarlas. Son solo un montón de palabras. También son símbolos, por supuesto, y cada uno representa una cosa o un concepto. Pero es casi imposible sumarlos o restarlos, o vaciarlos en una gráfica que muestre una tendencia. Si pones esas palabras en una columna, formarán los que los expertos en computación llaman “datos no estructurados”. Así es como llaman a ese gran caos. Un mejor método consistiría en marcar todas las carnes con una C, todas las verduras con una V, todos los dulces con una D y así sucesivamente. Cuando hayas reducido las palabras a símbolos, podrías vaciarlos en una hoja de cálculo y estimar, digamos, cuantas veces comiste carne o dulces en una semana dada. Luego podrías elaborar una gráfica para relacionar tu alimentación con cambios en tu peso o con la cantidad de granos que tienes en la cara.

Cuando tratamos de extraer información de una pila de datos, necesitamos herramientas que nos ayudan a encontrar lo que es relevante, lo que es importante para explorar diferentes escenarios. Un informe, ya sea

impreso en papel o visualizado en la pantalla, es en el mejor de los casos una representación bidimensional de los datos, una tabla con columnas y filas. Eso es suficiente cuando sólo tenemos dos factores a considerar, pero en el mundo real necesitan herramientas más potentes.

Según (Kay, 2004), los cubos de datos multidimensionales son extensiones de las tablas 2-D, al igual que en la geometría un cubo es una extensión tridimensional de un cuadrado. La palabra cubo nos trae a la mente un objeto 3-D, y podemos pensar en un cubo de datos 3-D como un conjunto de tablas de 2-D de estructura similar apiladas una encima de la otra.

Pero los cubos de datos no están restringidos a sólo tres dimensiones. La mayoría de los sistemas de procesamiento analítico en línea (OLAP) pueden construir cubos de datos con muchas más dimensiones, Microsoft SQL Server Analysis Services 2000, por ejemplo, permite un máximo de 64 dimensiones. Podemos pensar en un cubo de datos en 4-D como consistente en una serie de cubos 3-D, aunque la visualización de entidades tales dimensiones superiores en términos espaciales o geométricos puede ser un problema.

Dado lo anterior podemos decir que un cubo de información es una estructura de datos que utiliza ciertas dimensiones para representar el análisis de la información que proviene de una base de datos del tipo Data Warehouse. Una definición correcta de Data Warehouse la da Inmon (2005) citado por (Nedim & Clare, 2016), quien define el concepto de Data Warehouse como una colección de bases de datos diseñada como base para las funciones principales de los sistemas de soporte de decisiones DSS.

Este tipo de base de datos tiene ciertas características especiales que le permiten a las organizaciones almacenar grandes cantidades de datos, con la finalidad de facilitar la toma de decisiones basadas en datos históricos que la organización ha generado a través del tiempo.

## 1.1 Contexto

En el (Instituto Tecnológico Superior de Irapuato, 2017) el Departamento de Tutorías es el departamento responsable de administrar toda la información concerniente a los alumnos. Su objetivo consiste en desarrollar la formación integral del estudiante mediante el acompañamiento académico y personal para con ello formar hábitos de estudio y trabajo, incentivar la superación personal, asentar los conocimientos temáticos, mejorar los indicadores de calidad e integrar las competencias profesionales establecidas en los Programas Educativos.

Sus objetivos específicos son:

- Favorecer la integración del estudiante a la Educación Superior, al Instituto Tecnológico en cuanto a contexto y particularmente a sus grupos escolares.
- Desarrollar hábitos de estudio y para el trabajo, así como estrategias para aprender requeridos para la labor académica y su inserción posterior en el campo laboral.
- Configurar las condiciones motivacionales para que los estudiantes afronten con éxito su formación académica.
- Desarrollar en el estudiante el sentido de autonomía y las habilidades necesarias para la afrontar y satisfacer sus necesidades académicas y personales.
- Brindar alternativas de apoyo en contenidos de alta dificultad en los programas educativos ofertados en el Instituto.
- Definir e integrar metas para los proyectos personales y profesionales.
- Incorporar al profesorado de tiempo completo en las actividades de Tutoría.
- Dinamizar la actividad docente bajo la detección de necesidades de aprendizaje en el estudiantado.
- Identificar en los estudiantes las necesidades en materia tutorial para generar estrategias de mejora y/o de potencialización.
- Sensibilizar a la Comunidad Tecnológica de la importancia del desarrollo integral del estudiantado.
- Sistematizar las acciones de apoyo requeridas para afrontar las necesidades estudiantiles.
- Favorecer las condiciones de acompañamiento que permitan una mayor permanencia, mejor rendimiento académico y la continuidad de los apoyos académicos en los estudiantes de alta vulnerabilidad.
- Mejorar los indicadores de deserción, reprobación y eficiencia Terminal del Instituto.
- Valorar las actividades y el impacto generado por el Programa de Tutorías.

La información que procesa y analiza el Departamento de Tutorías para cumplir con sus objetivos proviene de diferentes sistemas ajenos al tecnológico y cada sistema genera sus propios reportes en diferentes formatos, por lo que unificar, procesar y analizar estos datos implica un tiempo considerable siendo necesario el apoyo de alumnos de servicio social, y es aquí donde el presente trabajo adquiere relevancia.

Para eficientar el procesamiento de la información y disminuir el personal involucrado en esta tarea, se desarrolló una base de datos Data Warehouse que permite el almacenamiento de todos los datos que provienen de las diversas fuentes de información. Esta base de datos emplea un paquete DTSX para el registro de la información y es la base para la integración de un *cubo de información* que le permite al Departamento de Tutorías un análisis más rápido que agiliza y sustenta la toma de decisiones, tal como el otorgamiento de becas, estrategias para disminuir el índice de deserción, desarrollo de plataformas de estudio, entre otras más.

Aunado a lo anterior fue necesario el desarrollo de un sistema de procesamiento cuya interfaz le facilita al responsable del Departamento de Tutorías la ejecución del proceso ETL (almacenamiento de todos los datos que provienen de las diversas fuentes de información) y la visualización del cubo de información en una hoja de cálculo de Excel para así generar tablas dinámicas y graficas de la información que desea analizar.

Por último se dispuso de un equipo de cómputo perteneciente al Departamento de Tutorías para la implementación del Data Warehouse (SQL Server), paquete DTSX (SQL Server Integration Services), cubo de información (Business Intelligence – Analysis Services) y sistema de procesamiento (Visual Studio), además de la elaboración de un manual para la correcta ejecución del sistema.

## 1.2 Antecedentes

**Análisis de rendimiento académico estudiantil usando Data Warehouse y Redes Neuronales.** Integrantes del Departamento de Ingeniería Informática y Ciencias de la Computación de la Universidad de Atacama, Chile. Desarrollaron el proyecto Análisis de rendimiento académico estudiantil utilizando Data Warehouse y Redes Neuronales, en el cual (Zambrano Matamala, Rojas Díaz, Carvajal Cuello, & Acuña Leiva, 2011) implementan un DW en base a información obtenida de un sistema de base de datos no relacional.

El DW se diseñó para el análisis del comportamiento de aprobación y avance en una malla curricular con datos reales de los estudiantes de la carrera de Ingeniería Civil en Computación e Informática de la Universidad de Atacama, así como la generación de reportes para dar valor a la información y brindar apoyo a la toma de decisiones.

Se utilizó la arquitectura de Redes Neuronales (RNA) como algoritmo de aprendizaje, la cual se diseñó una vez obtenido el DW, esto para la predicción de rendimiento de los estudiantes usando los algoritmos de Matlab, para estimar el comportamiento de un estudiante en el siguiente semestre. La RNA se alimenta de algunos de los datos obtenidos por medio del DW como tasa de aprendizaje.

Se obtuvieron resultados positivos con la realización del Data Warehouse y la generación de reportes, los cuales fueron generados en poco tiempo, demostrando que la implementación del DW facilitó y redujo el tiempo para consultar y analizar datos para predecir su compartimento futuro, para la toma de decisiones dentro de dicha institución, logrando obtener los siguientes resultados:

- Porcentaje de Aprobación de Asignaturas Acumuladas (PAAA) por semestre de permanencia para las distintas cohortes a partir del año 2001 al 2007.
- Porcentaje de Aprobación de Asignaturas por Semestre (PAS) de los estudiantes regulares de la carrera y de los estudiantes en situación no regular.
- Porcentaje de Aprobación por Semestre (PAS) en asignaturas de la carrera por semestre de permanencia para estudiantes regulares y no regulares.
- Porcentaje de Asignaturas Aprobadas Acumulada (PAAA) de la carrera por semestre de permanencia para cada zona geográfica.
- Porcentaje de Aprobación por Semestre (PAS) en asignaturas de la carrera (por semestre de permanencia para las cohortes de la 2001 a la 2004).
- Cantidad Acumulada de asignaturas que se cursan por nivel de asignaturas (solo pregrado).
- Puntajes promedios de pruebas por año (todos).

Los resultados obtenidos mediante la predicción RNA fueron positivos. Esta tecnología permitió obtener la siguiente información futura de la cantidad de asignaturas inscritas por un alumno, mostrando que el complemento entre DW y RNA es una herramienta potente para poder predecir el comportamiento futuro de un indicador de gestión al ajustarse bien la predicción obtenida por la arquitectura RNA a lo que la tendencia histórica del DW arroja.

Lo anterior permitió obtener la siguiente información:

- Cantidad de asignaturas aprobadas por un alumno en un semestre determinado.
- Resultados de la estimación de la cantidad de asignaturas inscritas por un alumno.
- Salida estimada de la cantidad de asignaturas aprobadas en el siguiente semestre.

**Data Warehouse para el análisis académico de la escuela Politécnica Nacional.** Los alumnos de la Escuela Politécnica Nacional de Quito, desarrollaron el proyecto Data Warehouse para el análisis académico de la escuela, donde (Guevara Lenis & Valencia Arcos, 2007) hacen uso del mismo como herramienta de apoyo para llevarse a cabo la visión de la misma (donde se describe como una de las universidades de mayor prestigio referente del país en el desarrollo científico-técnico y con actividad académica de excelencia).

Con base a resultados estadísticos y consolidados en base a análisis, se podrá determinar el cumplimiento o no de las metas; y mediante la toma de decisiones estratégicas encontrar su respectiva solución, esto facilitara el análisis de toda la información de una manera más sencilla y concreta, al estar concentrada en una sola fuente o base de datos, consumiendo menos tiempo y recursos.

Para la realización del Data Warehouse se utilizó Microsoft SQL Server 2000, para la transformación de datos DTS se utilizó una herramienta integrada en SQL y para la realización del cubo Analisis Services integrada de igual manera en SQL, para la creación de reportes la herramienta DWExplorer que permite interactuar con los cubos y exportar reportes a varios formatos.

Problemas presentados: En la recolección y la consolidación de la información en un único directorio. Se obtuvieron resultados positivos con la realización del Data Warehouse al realizar una recopilación de múltiples bases de datos, unificándolas en una sola, logrando obtener los resultados de las consultas de una forma más rápida y productiva para el directivo de la Escuela Politécnica Nacional; con lo cual gracias a los reportes y gráficos generados, facilitaron la visualización global de un conjunto de preguntas seleccionadas tales como:

- a. Número de estudiantes matriculados por periodos.
- b. Número de estudiantes por Carrera.
- c. Número de estudiantes por Sexo.
- d. Número de estudiantes retirados por periodo.
- e. Estudiantes exonerados por periodo.
- f. Estudiantes aprobados por periodo.
- g. Promedio de créditos que toman los estudiantes que toman por periodo.
- h. Promedio de materias por estudiante.
- i. Promedio de los parciales por periodo.
- j. Promedio de notas finales por periodo.
- k. Estudiantes con x matrícula por periodo.
- l. Anulaciones de matrícula por periodo.

**Uso de la Tecnología Data Warehouse en Unidades Educativas de Nivel Medio: Consideraciones Teóricas.** La utilización de un Data Warehouse en instituciones de nivel medio, podría ser según (Cabrera Torres, Emma, & Iván, 2014) muy necesaria por los beneficios que éste representa para dichas instituciones. Dentro de estos beneficios, se citan los siguientes:

- Un Data Warehouse, es una herramienta dinámica que permite el manejo de información generada desde las diferentes áreas de la Institución Educativa.
- Mediante la utilización de un Data Warehouse, los docentes podrían realizar reportes para conocer las tendencias de sus estudiantes respecto a notas por cada período de tiempo, cantidad de estudiantes aprobados y reprobados por año, asistencias, inasistencias y tardanzas de docentes y estudiantes, entre otros.
- El uso y aprovechamiento apropiado de la información académica, va a permitir a la institución mejorar sus procesos de toma de decisiones, debido a que van a contar con información confiable, completa y disponible en el momento que lo deseen.
- Permite recolectar información de los distintos periodos de tiempo escolares, con el fin de analizar la misma y realizar proyecciones que en el futuro permitan determinar cómo se desenvuelve la Institución Educativa.

En general el concepto se originó a mediados de los años ochenta, en nuestro entorno, dicha tecnología aspira proporcionar un modelo de arquitectura para la salida de datos de los sistemas operacionales y los ambientes de decisión. Para iniciar la ruta del Data Warehouse, se parte de los conceptos y para el efecto se ha tomado de referencia pensamientos de varios autores; tales como:

- “Un Data Warehouse, es un repositorio de información coleccionada desde múltiples fuentes, bajo un esquema uniforme y que usualmente reside en un solo sitio.” (Han & Kamber, 2001)

- “Es una colección de datos orientados al sujeto, integrados, de tiempos variantes y no volátiles, que sirven de soporte para el proceso de toma de decisiones.” (Mohania, 2009)
- “Almacenamiento de información homogénea y fiable, en una estructura basada en la consulta y el tratamiento jerarquizado de la misma, en un entorno separado de los sistemas operacionales.” (Humphries, Hawkins, & Dy, 2009)
- “Colección de datos orientado a temas, integrado, no volátil, de tiempo variante, que se usa para el soporte del proceso de toma de decisiones gerenciales.” (Giner de la Fuente & Gil Estallo, 2004)

**Data Warehouse - Universidades Andaluzas y Consejería de Educación y Ciencia de la Junta de Andalucía.** La presente organización en España (Universidades Andaluzas y Consejería de Educación y Ciencia de la Junta de Andalucía, 2005), presenta una guía de usuario en la que define la *arquitectura de un sistema Data Warehouse* de la siguiente manera.

Las organizaciones habitualmente disponen de un conjunto de sistemas transaccionales u operacionales (OLTP - On-Line Transaction Processing). Los datos de estos sistemas son parte de la infraestructura de la organización y se componen de información de detalle, no redundante y de constante actualización. Este componente del sistema informático, de la organización, es lo que se denomina como Operational Data.

A partir de esta información, los datos requeridos por los analistas de gestión normalmente hacen referencia, en primera instancia, a información agregada, basándose en períodos de tiempo más amplios y con un nivel de acotación o filtrado, de la información, adaptado a necesidades puntuales. Este tipo de análisis de la información requiere de sistemas diseñados a tal efecto, dotados de una infraestructura que permita la libre navegación por la información, e independientes de los sistemas operacionales a efectos de no penalizar al sistema corporativo y al mismo tiempo de obtener la máxima dinámica en los procesos de análisis. Estos sistemas de información (OLAP-On-Line Analytical Processing) se conocen como almacén de datos o Data Warehouse.

Para poder acceder a los datos almacenados en un Data Warehouse de manera eficiente, por parte de un usuario no técnico, es necesario saber cómo están estos datos almacenados, es decir, se necesita saber su estructura huyendo de terminología técnica. Este catálogo o inventario de datos, contenidos en el Data Warehouse, se conoce con el nombre de Metadata.

Por último necesitamos el engine o la tecnología que transforma las peticiones de los usuarios en información tabulada y que en el caso de las Universidades de Andalucía está basada en el software, líder en tecnología Data Warehouse, y más concretamente, desde el punto de vista de usuario final, en la herramienta DSSAgent de la compañía MicroStrategy.

### 1.3 Problemática

Actualmente el Departamento de Tutorías del Instituto Tecnológico Superior de Irapuato administra parte importante de la información relacionada con el ingreso de los alumnos. Esta información que consiste, por ejemplo, en datos tales como la escuela de procedencia, resultados de exámenes psicométricos y CENEVAL; e incluso condición familiar y económica, resulta de mucho interés para la dirección académica de la institución.

La información que gestiona el Departamento de Tutorías le permite a la Dirección Académica conocer la condición de sus alumnos e identificar riesgos y áreas de oportunidad con la finalidad de tomar decisiones de manera oportuna y bajo una base sólida de información. Así mismo, los *organismos acreditadores* al realizar la visita *in situ* para revisar el plan de estudios de las carreras, solicitan información relacionada al ingreso de los alumnos; de tal forma que es importante para el Departamento de Tutorías tener esta información procesada y lista para generar reportes resumidos que les permitan a los evaluadores de los organismos acreditadores analizar el comportamiento de la matrícula de ingreso de cada una de las carreras en cuestión.

El problema principal radica en el *procesamiento de la información* de los alumnos de nuevo ingreso. La forma en cómo se gestiona esta información es a través de varios archivos de diversos sistemas de información independientes y ajenos al mismo departamento; por consiguiente este proceso es poco eficaz y muy tardado. Lo anterior también genera inconsistencia en la información derivado del múltiple uso de archivos como fuentes de datos y la intervención del servicio social para la integración completa de la información, por lo que en resumen podemos decir que el procesamiento de la información de los alumnos de nuevo ingreso es lento y poco fiable.

### 1.4 Objetivos

*General:* Facilitar el análisis y la administración de la información académica y personal de los alumnos del Instituto Tecnológico Superior de Irapuato, implementando un cubo de información para el Departamento de Tutorías que le permita reducir el tiempo de consulta en al menos la mitad del tiempo que requiere actualmente.

*Específicos:*

1. Analizar las necesidades y requerimientos del Departamento de Tutorías para generar un modelo de datos para el almacenamiento de la información.
2. Desarrollar una base de datos Data Warehouse para almacenar la información en un repositorio que facilite su manipulación.
3. Implementar un paquete DTSX (proceso ETL - Extract, Transform and Load) empleando Integration Services para migrar la información de los archivos independientes a la base de datos.
4. Desarrollar un cubo de información empleando Business Intelligence para facilitar el análisis de la información almacenada en la base de datos.
5. Realizar las pruebas de análisis de la información procesada en el cubo de información para validar que el proceso ETL migro la información de manera correcta.
6. Desarrollar un sistema de procesamiento que le permita al usuario final ejecutar de forma transparente el paquete DTSX y actualizar el cubo de información para facilitar su manipulación en una hoja de cálculo.
7. Realizar las pruebas del funcionamiento del programa de procesamiento para validar su funcionamiento.

## 1.5 Marco metodológico

El presente trabajo contiene las siguientes etapas:

**Análisis de las necesidades y requerimientos.** El Departamento de Tutorías a través del Dr. Juan Carlos Soto Patiño mostró las necesidades que tiene en dicha área en relación a la información de los alumnos de primer ingreso al instituto. A su vez facilitó los documentos (archivos) generados por diversas áreas ajenas al área tutorial con la información necesaria para poblar la base de datos Data Warehouse.

**Desarrollo de la base de datos Data Warehouse.** Los requerimientos levantados brindaron la base y sustento para el diseño y desarrollo de la base de datos Data Warehouse. La base de datos fue diseñada empleando un modelo de datos Entidad-Relación (E-R) y un modelo Relacional para brindar un diagrama confiable que permitió su desarrollo con el Sistema Manejador de Base de Datos SQL Server 2014.

**Implementación del paquete DTSX.** Esta etapa es importante para el desarrollo de este trabajo debido a que la *extracción, transformación y carga* de datos se lleva a cabo con el desarrollo e implementación del paquete ETL utilizando SQL Server Integration Services. Este paquete a través de controles de datos “Data Flow” permitió: 1) la correcta extracción de la información proveniente de archivos independientes, 2) la transformación necesaria para eliminar la inconsistencia de los datos y 3) la migración a la base de datos Data Warehouse como un *repositorio histórico* de la información.

**Desarrollo y pruebas del cubo de información.** El cubo de información es una tabla dinámica desarrollada en la plataforma Business Intelligence (Analysis Services) para la cual fue necesario configurar: 1) el origen de datos que en este caso consiste de la base de datos Data Warehouse, 2) la selección de tablas y vistas, 3) codificar dos funciones para realizar cálculos automáticos, 4) definir las medidas y dimensiones del cubo, y 5) el procesamiento del cubo. Las pruebas de esta etapa consistieron en la manipulación del cubo realizando algunas consultas analíticas de la información de nuestro Data Warehouse.

**Desarrollo y pruebas del sistema de procesamiento.** Se consideró necesario un sistema de procesamiento que le permita al usuario final ejecutar el paquete DTSX y actualizar el cubo de información asegurando que la información de los archivos independientes fue migrada satisfactoriamente a la base de datos Data Warehouse y que está disponible en el cubo de información conectado a una hoja de cálculo que le permita a la persona una manipulación sencilla en un entorno amigable y fácil de compartir con directivos de la institución.

Adicionalmente es importante mencionar en este marco metodológico el paradigma de desarrollo e implementación del presente trabajo y los resultados preliminares.

**Paradigma de Desarrollo e Implementación.** La metodología ágil Scrum fue el proceso que se utilizó para llevar a cabo el desarrollo y la implementación de todo el sistema en el Departamento de Tutorías, la cual nos permitió ejecutar buenas prácticas para el trabajo colaborativo de un equipo de desarrollo. Un aspecto

relevante consistió en integrar como parte del equipo al Dr. Juan Carlos Soto Patiño (encargado del Departamento de Tutorías) en su rol de Product Owner.

**Resultados preliminares.** De manera general permitieron conocer el decremento en el tiempo que invierte el Departamento de Tutorías para realizar algunas consultas previamente definidas, y así determinar el cumplimiento de nuestro objetivo general. Se muestra una tabla comparativa que implica el tiempo invertido actualmente versus el tiempo invertido empleando el cubo de información.

## 2 Desarrollo del cubo de información

En esta sección se presentarán cada una de las etapas necesarias para el desarrollo del cubo de información permitiendo alcanzar cada uno de los objetivos planteados en la introducción.

### 2.1 Recolección de requerimientos

Para la recolección de los requerimientos fue necesaria una entrevista con el Dr. Juan Carlos Soto Patiño encargado del Departamento de Tutorías, para presentar una propuesta del funcionamiento que tendría el cubo de información dentro del departamento; esta propuesta fue aprobada haciendo principal énfasis en la necesidad de una herramienta tecnológica para sustentar la información ante los organismos acreditadores.

Asimismo fueron entregados los archivos (Excel) con información de la generación 2016, siendo suficientes para alimentar el sistema y realizar algunas pruebas, de los cuales se indicaron los campos a ocupar para poder realizar los análisis pertinentes. Los documentos que se proporcionaron son “TUT\_GEN15\_15NOV16” y “TUT\_GENR13\_15NOV16”, los cuales fueron solicitados al responsable del *Departamento de Control Escolar* para obtener información del alumno tal como: número de control, nombre, carrera y escuela de procedencia.

Otro documento importante es el *Examen Vocacional*, el cual se aplica a los alumnos al ingresar al ITESI para conocer datos tales como: sexo, ciudad, si trabaja, salario, ocupación del padre, percepción mensual promedio del padre. Además, contiene información de hábitos de estudio (concentración, memoria, comunicación oral, comunicación escrita, administración de tiempo, lectura de libros, toma de apuntes, presentación de evaluaciones, relaciones interpersonales y motivación para estudiar).

Un último documento es el *Cuestionario de Salud* que contiene la información de: trastornos padecidos en los últimos seis meses, si han tenido relaciones sexuales, si actualmente han tenido relaciones sexuales, utilización de algún método anticonceptivo, nerviosismo y si tiene hijos.

### 2.2 Data Warehouse

**Modelos de Datos.** Los modelos de datos empleados para el diseño de los diagramas de la base de datos fueron el Modelo Entidad-Relación en el cual se identificaron cuatro entidades junto con sus atributos y relaciones, y el Modelo Relacional, el cual permitió transformar cada entidad en una tabla con campos definidos. Este último diagrama fue la base para la construcción de la base de datos.

**Tabla 1.** Muestra la tabla alumno de la base de datos.

Campo	Descripción	Tipo de dato
NoCtrl	Numero de control	Varchar
Nombre	Nombre	Varchar
Ap_Pat	Apellido paterno	Varchar
Ap_Mat	Apellido materno	Varchar
Genero	Genero	Varchar
Carrera	Carrera	Varchar
Ciudad	Ciudad	Varchar
Esc_Proc	Escuela procedencia	Varchar
Pts_Ceneval	Puntos examen ceneval	Smallint
Año_Ingreso	Año de ingreso	Smallint

**Base de datos.** La base de datos fue desarrollada con SQL Server 2014 Management Studio (véase la Figura 1), el cual es un entorno que permite la configuración y administración de múltiples tablas. La base de datos está conformada por las tablas *Alumnos*, *Edo\_Socioeconomico*, *Hab\_Estudios* y *Salud*. Dichas tablas contienen únicamente la información previamente definida por Departamento de Tutorías y tienen definida el campo “NoCtrl” (campo de los alumnos útil para identificarlos en el contexto de la organización) como llave primaria, llave que para el lenguaje SQL es importante debido a que permitirá la unión entre las demás tablas. A continuación se presentan las tablas creadas en la base de datos con los campos, descripción y tipo de dato.

**Tabla 2.** Muestra la tabla Edo\_Socioeconomico de la base de datos.

Campo	Descripción	Tipo de dato
NoCtrl	Numero de control	Varchar
Salario_Padre	Ingreso familiar	Varchar
Personas_Hogar	Numero de personas que viven en el hogar	Varchar
Ocupacion_Padre	Ocupación del padre	Varchar
Casa	Casa propia, renta o prestada.	Varchar
Hijos	Es padre o madre	Varchar
Trabaja	Trabaja actualmente	Varchar

**Tabla 3.** Muestra la tabla Hab\_Estudios de la base de datos.

Campo	Descripción	Tipo de dato
NoCtrl	Numero de control	Varchar
Concentracion	Concentración	Tinyint
Memoria	Memoria	Tinyint
Comunicacion_Oral	Comunicación oral	Tinyint
Comunicacion_Escrita	Comunicación escrita	Tinyint
Administracion_Tiempo	Administración del tiempo	Tinyint
Lectura_Libros	Lectura de libros	Tinyint
Toma_Apuntos	Toma de apuntes	Tinyint
Presentacion_Evaluaciones	Presentación de evaluaciones	Tinyint
Relaciones_Interpersonales	Relaciones interpersonales	Tinyint
Motivacion_Estudio	Motivación para estudiar	Tinyint

**Tabla 4.** Muestra la tabla Salud de la base de datos.

Campo	Descripción	Tipo de dato
NoCtrl	Numero de control	Varchar
Transtornos	Sufre de algun transtorno	Tinyint
Rel_Sexuales	Ha tenido relaciones sexuales	Tinyint
Act_Rel_Sex	Mantiene relaciones actualmente	Tinyint
Metodo_Anticonceptivo	Métodos anticonceptivos	Tinyint
Nerviosismo	Sufre nerviosismo	Tinyint
Estado_Emocional	Estado en el que se encuentra	Tinyint

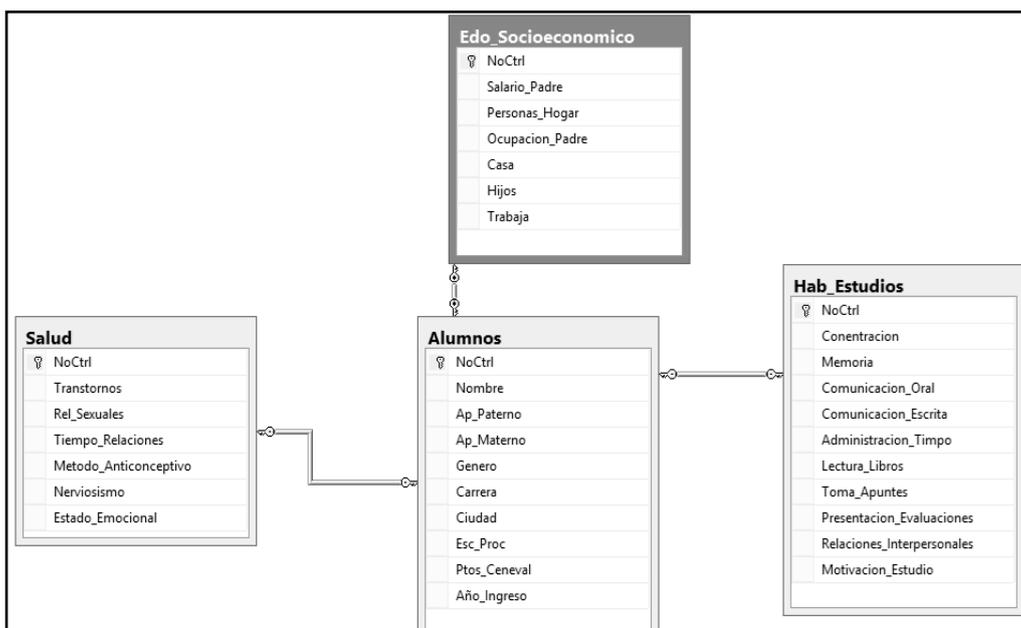


Figura 5. Diagrama de la base de datos en SQL Server.

## 2.3 Plantillas de información

Teniendo en cuenta los datos o campos que le son de utilidad al área tutorial, se creó una plantilla en Excel de tipo Libro de Excel 97-2003 pertinente. Esta plantilla permite organizar los datos considerados como una fuente de datos externos, los cuales se registraron manualmente con información de los documentos de Excel previamente proporcionados. Con ésta información se realizaron distintas hojas de acuerdo a cada tabla que se encuentra en la base de datos en un solo libro llamado “Plantilla”, clasificando la información entregada por el área tutorial.

## 2.4 Paquete DTSX

**Proceso ETL.** Para el desarrollo del paquete DTSX que ejecuta al proceso ETL (véase la Figura 2), se utilizó la plataforma Business Intelligence y la herramienta SQL Server Integration Services. Se configuraron dos diagramas de flujos de datos (Data Flow) compuestos por tareas de flujo de datos, los cuales permitieron la transformación necesaria de los dtos obtenidos de las diferentes plantillas en un solo tipo de dato, a fin de evitar la inconsistencia de la información al procesarla y almacenarla en el Data Warehouse.

*Fuente de origen.* Dentro del flujo de datos (Data Flow) se configuró la fuente de origen de donde se obtuvo la información. Esta fuente de origen recuperó la información de la plantilla de Excel para proveer a la base de datos toda la información de los archivos independientes.

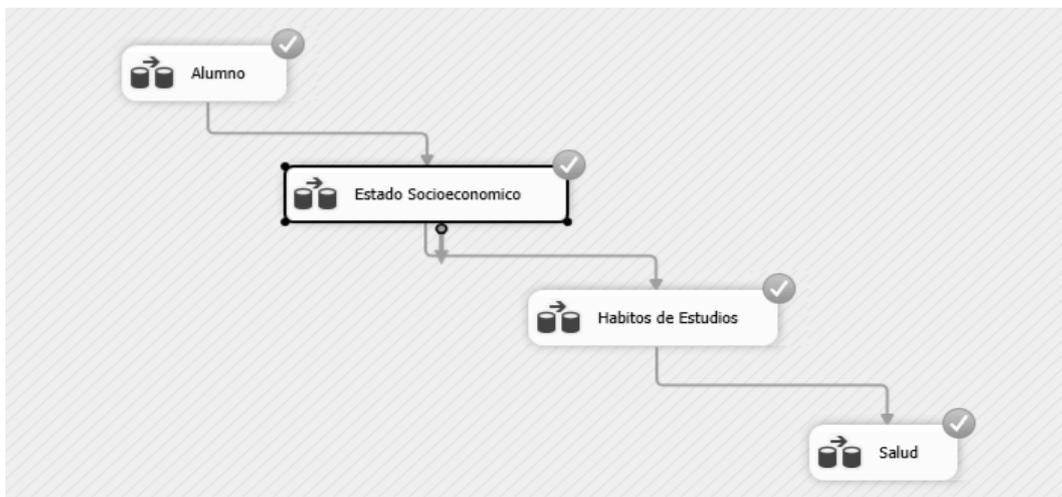
*Transformación.* Una vez seleccionado la fuente de origen, se prosigue con la herramienta Transact-SQL conversión de datos, la cual permite la integración de todos aquellos datos que se deseen transformar; realizando un nuevo campo con su tipo de dato transformado para evitar inconsistencias en la base de datos.

*Fuente de destino.* Por último se seleccionó la fuente de destino para cargar y almacenar los datos ya transformados de la plantilla de Excel a la nueva base de datos previamente desarrollada en SQL Server.



**Figura 6.** Flujo de datos en donde se muestra el proceso ETL; Extrae de Excel, Transforma y Carga a la base de datos Data Warehouse. Esta figura únicamente muestra la información de una tabla (alumno).

**Paquete DTSX.** El desarrollo y ejecución consistió en agrupar los cuatro flujos de datos (uno para cargar cada tabla de la base de datos Data Warehouse) en un flujo de control (véase la Figura 3).



**Figura 7.** Ejecución exitosa del paquete DTSX, cuatro flujos de datos para cuatro tablas de la base de datos Data Warehouse.

Para verificar la ejecución correcta del paquete DTSX se realizó una consulta a la base de datos en SQL Server, con la intención de validar la carga correcta de la información obtenida de la plantilla de Excel. Para lo anterior, se solicita en la consulta información proveniente de las tablas Alumnos y Edo\_Socioeconomico como: nombre, apellido paterno y materno, género, carrera, ocupación del padre y salario del padre, uniéndolas por medio de un join, mostrando información únicamente del sexo masculino (véase la Figura 4).

SQLQuery1.sql - G4...s (G405\Rocha (55))\*

```
select A.Nombre, A.Ap_Paterno, A.Ap_Materno, A.Genero, A.Carrera, S.Ocupacion_Padre, S.Salario_Padre from Alumnos A
join Edo_Socioeconomico S
on A.NoCtrl = S.NoCtrl
where A.Genero = 'Masculino'
```

	Nombre	Ap_Paterno	Ap_Materno	Genero	Carrera	Ocupacion_Padre	Salario_Padre
276	Thomas Eduardo	Delgado	Huerta	Masculino	Ing. Gestión Empresarial	Profesionista Asalariado	32000
277	Juan Antonio	Magdalena	Gutiérrez	Masculino	Ing. Sistemas Computacionales	Empleado	4000
278	Héctor Alexis	Villanueva	Hernández	Masculino	Ing. Electromecánica	Empleado	8000
279	Luis Oswaldo	Huichapa	López	Masculino	Ing. Sistemas Automotrices	Comerciante	NULL
280	ROBERTO CARLOS	BELTRAN	SOLIS	Masculino	Ing. Mecatrónica	Desempleado	12000
281	Luis Enrique	Sanchez	Ramirez	Masculino	Ing. Informática	Obrero	3500
282	Miguel Ángel	Sanchez	Mendoza	Masculino	Ing. Informática	Obrero	3200
283	Emmanuel Nicolás	Barroso	Salazar	Masculino	Ing. Industrial	Profesionista	3000
284	Jorge Antonio	Sanchez	Hernandez	Masculino	Ing. Industrial	Agricultor	4000
285	Jose Francisco	Gutierrez	Gasca	Masculino	Ing. Mecatrónica	Empleado	6400
286	GUILLERMO ALEJA...	LEÓN	CHAIRES	Masculino	Ing. Mecatrónica	Jomalero	6000
287	jose alonso	chagolla	ramirez	Masculino	Ing. Sistemas Computacionales	Jomalero	6000
288	JOSE LUIS	CARDENAS	MARTINEZ	Masculino	Ing. Sistemas Computacionales	Empleado	0
289	JUAN ANTONIO SA...	ZARATE	MEDRANO	Masculino	Ing. Electromecánica	Empleado	5000
290	OMAR ELIM	OSORIO	PEÑA	Masculino	Ing. Electromecánica	Empleado	NULL
291	GARCIA	ERICK AGU...	ALEJANDRE	Masculino	Ing. Mecatrónica	Comerciante	6000
292	Diego de Jesús	Vargas	Jiménez	Masculino	Ing. Mecatrónica	Obrero	5000
293	raymundo	patlan	palacios	Masculino	Ing. Gestión Empresarial	Independiente	4000
294	Ivan	Lara	Garcia	Masculino	Ing. Bioquímica	Comerciante	4000
295	BRANDON ISRAEL	GUZMÁN	VEGA	Masculino	Lic. en Biología	Obrero	3000
296	BRANDON	CASTILLO	VALLEJO	Masculino	Ing. Logística	Empleado	15000
297	Cesar Everardo	Palomino	Castillo	Masculino	Ing. Sistemas Computacionales	Empleado	9500

Consulta ejecutada correctamente. G405 (12.0 RTM) G405\Rocha (55) Prototipo

Figura 8. Consulta a la base de datos Data Warehouse para verificar la carga de los datos desde un paquete DTSX.

## 2.5 Cubo de información

Para la creación del cubo de información se utilizó la plataforma de Visual Studio Business Intelligence (Analysis Services). Una vez cargado el nuevo proyecto en el explorador de soluciones se desarrollaron algunas configuraciones.

**Origen de datos.** Se conectó la base de datos Data Warehouse para proveer la información que alimentará el cubo de información (véase la Figura 5).

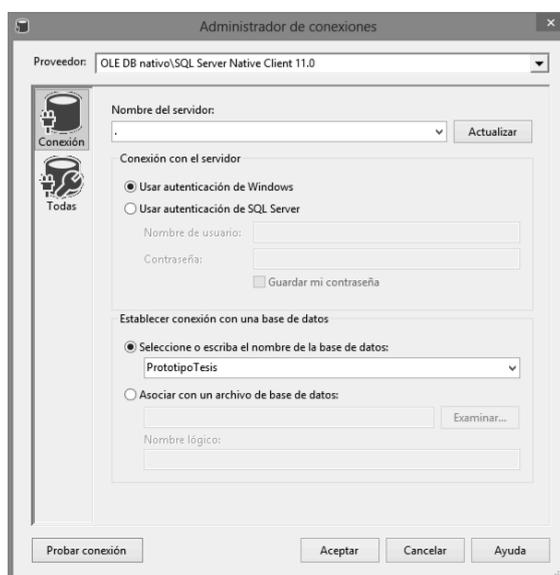
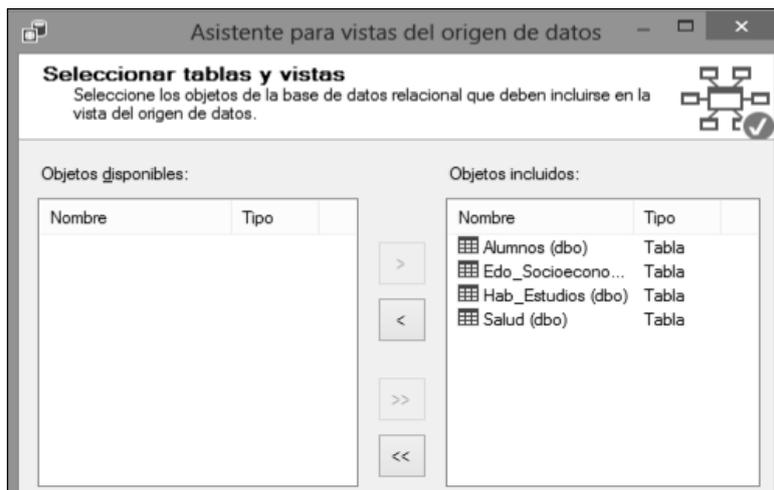


Figura 9. Configuración de la base de datos Data Warehouse como origen de datos.

**Tablas y vistas.** Una vez configurado el origen de datos, se seleccionaron las tablas y vistas del cubo de información (véase la Figura 6), en nuestro caso seleccionamos todas las tablas creadas en la base de datos: Alumnos, Hab\_Socioeconomico, Hab\_Estudios, Salud.

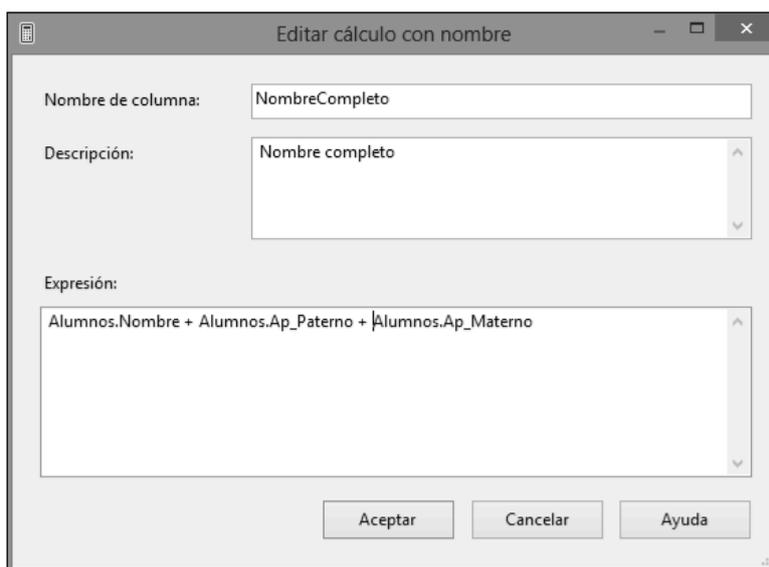


**Figura 10.** Selección de las tablas y vistas del origen de datos.

**Funciones.** Para calcular el *ingreso per capita* de cada uno de los alumnos en base al sueldo del tutor y número de personas que dependen de él fue necesario desarrollar una función que hiciera el cálculo automáticamente (véase Figura 7). Así mismo se desarrolló una segunda función (véase Figura 8) para definir el nombre completo de cada alumno, este compuesto por tres campos de la base de datos (nombre, apellido paterno y apellido materno).



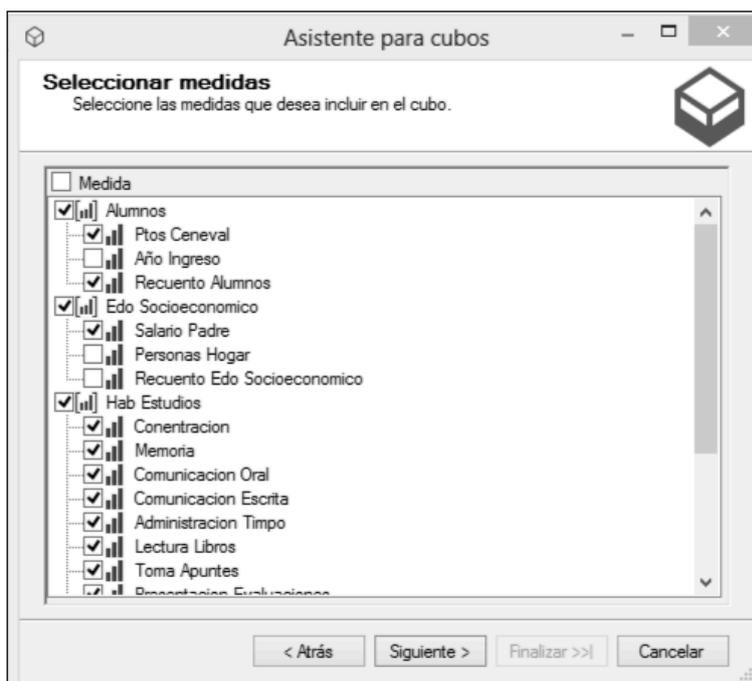
**Figura 11.** Función del ingreso per capita de los alumnos.



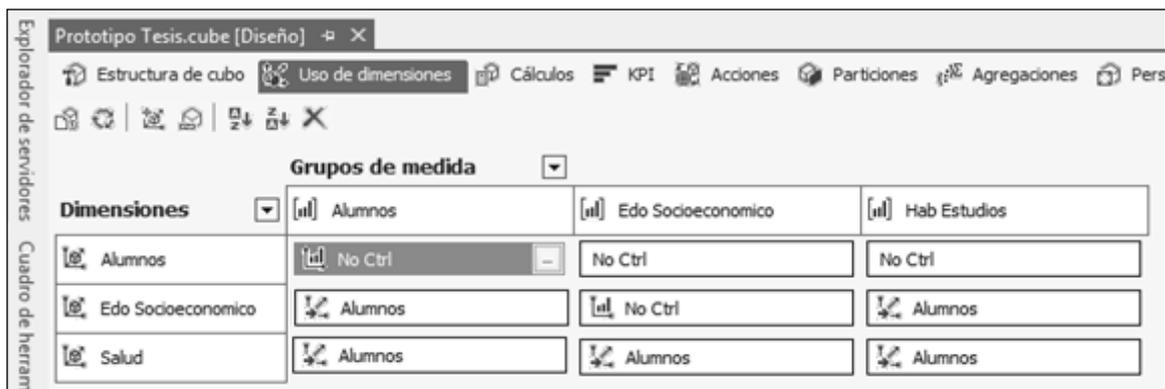
**Figura 12.** Función de nombre completo de los alumnos.

**Medidas y dimensiones.** Para obtener las medidas del cubo de información (véase Figura 9) las cuales se consideran indispensables para la toma de decisiones del área tutorial, como por ejemplo el ingreso per capita podría sustentar el otorgamiento de becas, se definieron las medidas y dimensiones del cubo. Para las medidas (véase Figura 10) se consideraron:

- Pts\_Ceneval (Alumnos).
- Salario\_Padre (Edo\_Socioeconomico).
- Hábitos de estudio (Todos los valores).



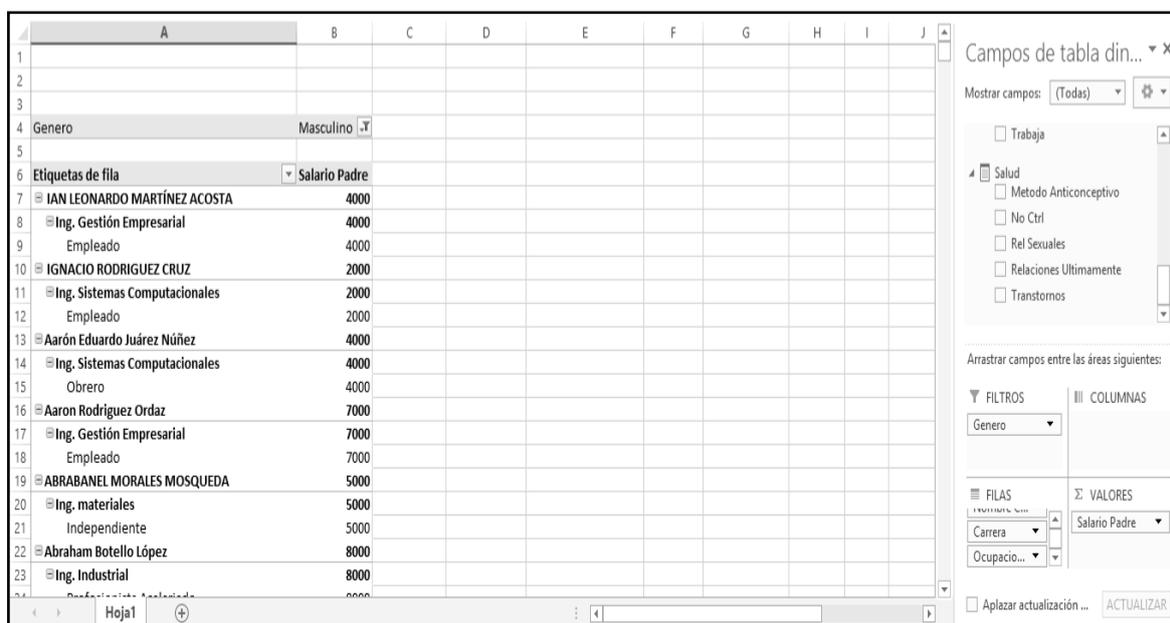
**Figura 13.** Selección de las medidas que tendrá el cubo de información.



**Figura 14.** Selección de las dimensiones que tendrá el cubo de información. En la figura se muestran tres dimensiones (Alumnos, Edo\_Socioeconomico y Salud) y dos tablas de medidas (Edo\_Socioeconomico y Hab\_Estudios)

*Aclaración 1. La tabla de Hab\_Estudios no aparece como dimensión ya que todos los campos son utilizados como medidas para el cubo de información.*

**Procesamiento del cubo.** Este procesamiento consiste en verificar que el cubo de información no tiene ningún error en alguna de las configuraciones previamente descritas. Una forma de validar el cubo es procesarlo y manipularlo por ejemplo en una hoja de cálculo, permitiéndonos analizar la información que tiene como origen la base de datos Data Warehouse, tal y como se muestra en la Figura 11.



**Figura 15.** Procesamiento y manipulación del cubo de información en una hoja de cálculo de Excel. En esta hoja se muestran cada uno de los alumnos de nuevo ingreso, junto con el salario del padre o tutor, además de indicar la ocupación del padre y la carrera a la que pertenece el alumno.

## 2.6 Sistema de procesamiento

Para poder ejecutar Paquetes DTSX desde un sistema de procesamiento se requiere de una función job o tarea de SQL Server, el cual permite ejecutar de manera automática algunos procesos o en nuestro caso el paquete DTSX mediante el Agente SQL Server.

Para crear un job, se debe entrar al explorador de objetos del Administrador de SQL Server y desplegar el agente de SQL Server, donde se encuentra la carpeta Trabajos, clic derecho nuevo trabajo. Si se expande esta carpeta se pueden ver los trabajos creados.

Copiamos de la opción de propiedades del paquete DTSX la ruta de acceso, para pegarla en la configuración de cada job. Una vez configurado el job que ejecutará nuestro paquete DTSX, se inician haciendo clic sobre el en la opción iniciar trabajo, esto para verificar que está ejecutándose correctamente (véase Figura 12).

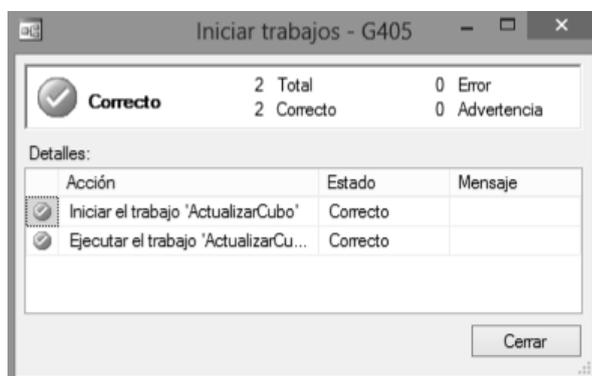


Figura 16. Ejecución exitosa del job que ejecuta el paquete DTSX.

Así mismo se codificó un procedimiento almacenado (Stored Procedure) en lenguaje Transact-SQL dentro de SQL Server Management para ejecutar el job previamente configurado (véase la Figura 13). De tal forma que el procedimiento almacenado iniciará el job para ejecutar automáticamente el paquete DTSX sin necesidad de abrir el proyecto de Business Intelligence, esto permitirá al sistema de procesamiento ejecutar el paquete DTSX de forma transparente al usuario final.

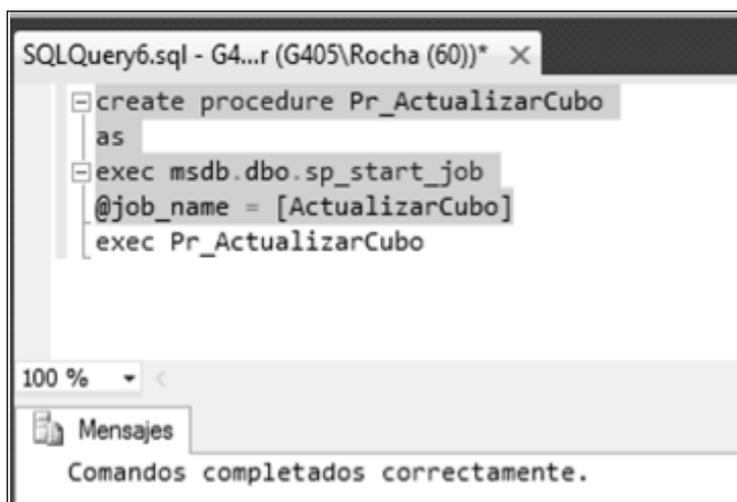


Figura 17. Codificación del procedimiento almacenado que inicia automáticamente el job. Un procedimiento almacenado puede ser ejecutado desde el código de programación de un sistema desarrollado en Visual Studio.

Por último se desarrolló un sistema de procesamiento (véase Figura 14) que permite cargar los datos (desde una hoja de cálculo como archivo independiente), procesar el cubo de información y abrir el cubo de información en una hoja de cálculo para su manipulación como una tabla dinámica. Las ventajas de visualizar el cubo de información desde una hoja de cálculo de Excel es facilitar al usuario final el guardar la información y transportarla en un archivo amigable y con posibilidades de obtener gráficas automáticas.



Figura 18. Sistema de procesamiento que muestra una interfaz amigable al usuario final.

La hoja de cálculo es visible desde el programa de Excel como se muestra en la Figura 15, y es a través de este archivo como el Departamento de Tutorías podrá manipular el cubo de información, teniendo acceso a las dimensiones del cubo y diversas opciones tales como la definición de filtros, columnas, filas y valores.

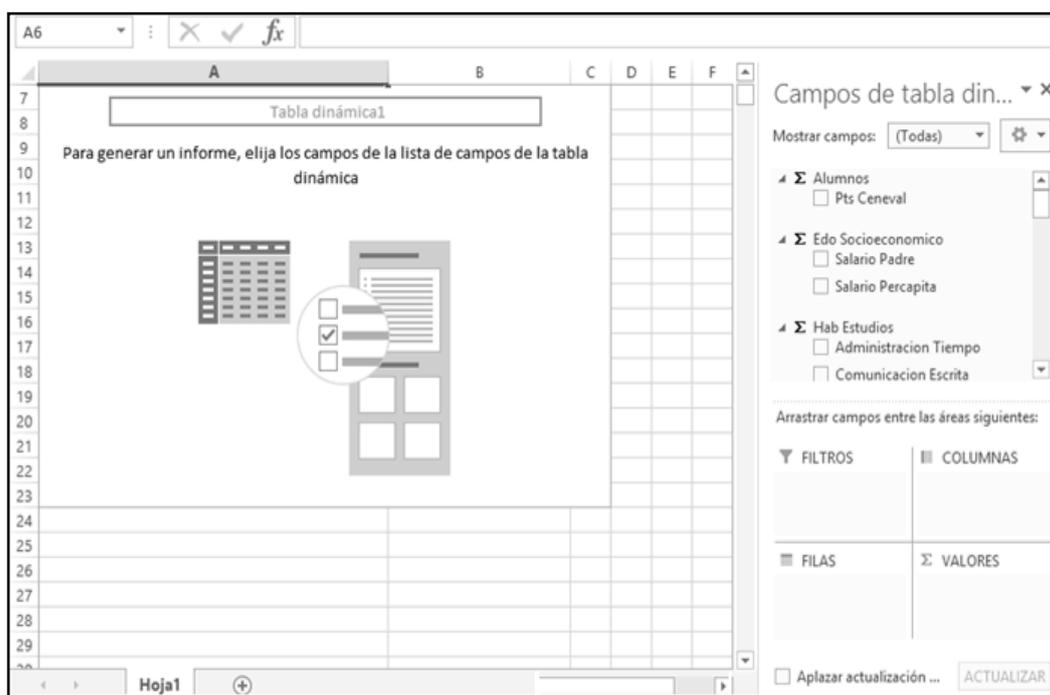


Figura 19. Interfaz del cubo de información desde una hoja de cálculo en Excel.

## 2.7 Paradigma del Desarrollo e Implementación

La forma en la cual se llevó a cabo el desarrollo del sistema completo fue siguiendo la metodología SCRUM, la cual plantea (Rivera Hernández, 2010) como una metodología ágil o proceso en el que se aplican de manera regular un conjunto de mejores prácticas para trabajar colaborativamente, en equipo, y obtener el mejor resultado posible de un proyecto. Estas prácticas se apoyan unas a otras y su selección tiene origen en un estudio de la manera de trabajar de equipos altamente productivos.

La metodología Scrum según (Softeng Software Engineers, 2010) es una metodología ágil y flexible para gestionar el desarrollo de software, cuyo principal objetivo es maximizar el retorno de la inversión para su empresa (ROI). Se basa en construir primero la funcionalidad de mayor valor para el cliente y en los principios de inspección continua, adaptación, auto-gestión e innovación. Esta metódica de trabajo promueve la innovación, motivación y compromiso del equipo que forma parte del proyecto, por lo que los profesionales encuentran un ámbito propicio para desarrollar sus capacidades.

Esta metódica de trabajo promueve la innovación, motivación y compromiso del equipo que forma parte del proyecto, por lo que los profesionales encuentran un ámbito propicio para desarrollar sus capacidades. Los beneficios que conlleva se listan en:

- *Cumplimiento de expectativas*: El cliente establece sus expectativas indicando el valor que le aporta cada requisito / historia del proyecto, el equipo los estima y con esta información el Product Owner establece su prioridad. De manera regular, en las demos de Sprint el Product Owner comprueba que efectivamente los requisitos se han cumplido y transmite se feedback al equipo.
- *Flexibilidad a cambios*: Alta capacidad de reacción ante los cambios de requerimientos generados por necesidades del cliente o evoluciones del mercado. La metodología está diseñada para adaptarse a los cambios de requerimientos que conllevan los proyectos complejos.
- *Reducción del Time to Market*: El cliente puede empezar a utilizar las funcionalidades más importantes del proyecto antes de que esté finalizado por completo.
- *Mayor calidad del software*: La metódica de trabajo y la necesidad de obtener una versión funcional después de cada iteración, ayuda a la obtención de un software de calidad superior.
- *Mayor productividad*: Se consigue entre otras razones, gracias a la eliminación de la burocracia y a la motivación del equipo que proporciona el hecho de que sean autónomos para organizarse.
- *Maximiza el retorno de la inversión (ROI)*: Producción de software únicamente con las prestaciones que aportan mayor valor de negocio gracias a la priorización por retorno de inversión.
- *Predicciones de tiempos*: Mediante esta metodología se conoce la velocidad media del equipo por sprint (los llamados puntos historia), con lo que consecuentemente, es posible estimar fácilmente para cuando se dispondrá de una determinada funcionalidad que todavía está en el Backlog.
- *Reducción de riesgos*: El hecho de llevar a cabo las funcionalidades de más valor en primer lugar y de conocer la velocidad con que el equipo avanza en el proyecto, permite despejar riesgos eficazmente de manera anticipada.

La metodología empleada en el desarrollo del sistema es Scrum. La única modificación que se tiene que realizar en cuanto a la ejecución de la metodología es el Daily Scrum Meeting, puesto que el equipo (team) de desarrollo solo coincide en reuniones semanales para la ejecución de algún sprint. La asignación de roles y las herramientas utilizadas se definen en la tabla 5.

**Tabla 5.** Roles y herramientas empleadas en Scrum.

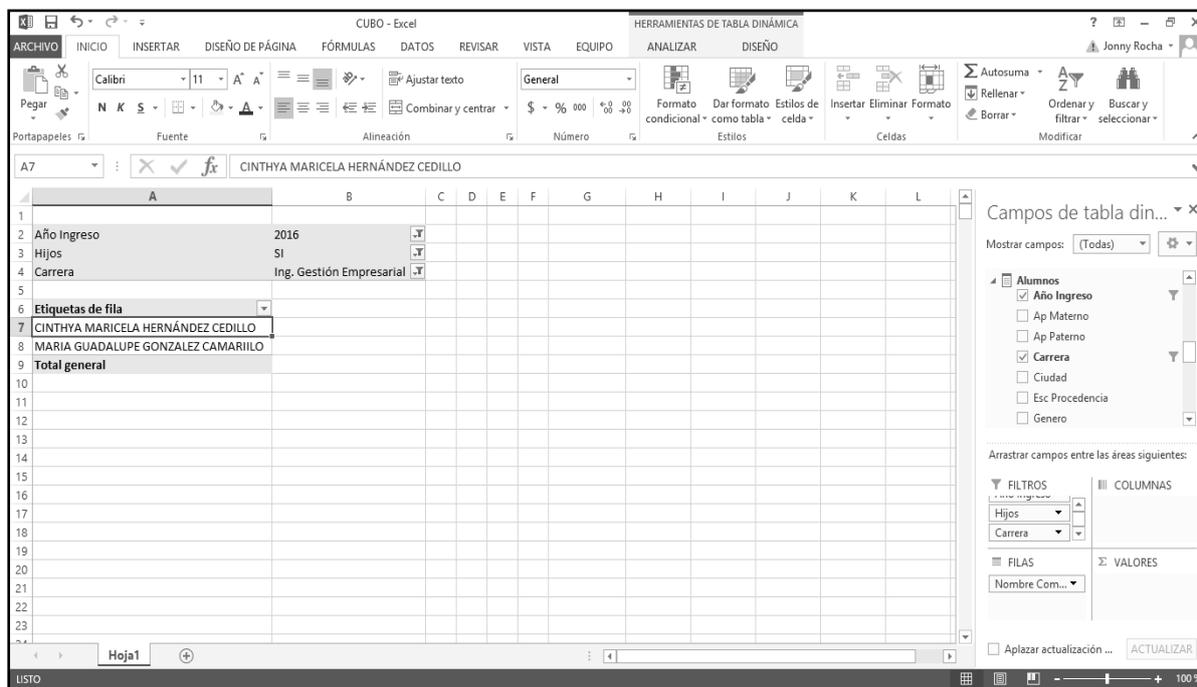
Recurso	Componente
Roles Scrum	<ul style="list-style-type: none"> <li>o Scrum Master (MTI Rodrigo Villegas Téllez)</li> <li>o Product Owner (Dr. Juan Carlos Soto Patiño)</li> <li>o Team (MTI David Antonio Torres Frausto, Rubicelia Jasso Hernández y Juan Rocha Jaramillo)</li> </ul>
Herramientas de Desarrollo	<ul style="list-style-type: none"> <li>o Visual Studio 2010</li> <li>o SQL Server 2014</li> <li>o Componentes Scrum:                             <ul style="list-style-type: none"> <li>▪ Product Backlog</li> <li>▪ Sprint Backlog</li> <li>▪ Sprint Planning</li> </ul> </li> </ul>

- Burndown Chart
- Weekly Meeting (variante del Daily Meeting)

## 2.8 Resultados Preliminares

Las consultas desarrolladas a través de este sistema fueron determinadas por el Departamento de Tutorías y los resultados obtenidos fueron los siguientes:

- ¿Qué alumnos (nombre completo) de Ing. Gestión Empresarial de la generación 2016 tiene hijos? (véase Figura 16)



**Figura 20.** Resultado de una consulta en el cubo de información empleando una hoja de cálculo en Excel.

- ¿Qué alumnos (nombre completo) de la generación 2016 de género femenino provienen de la escuela CBTis 65? (véase Figura 17)

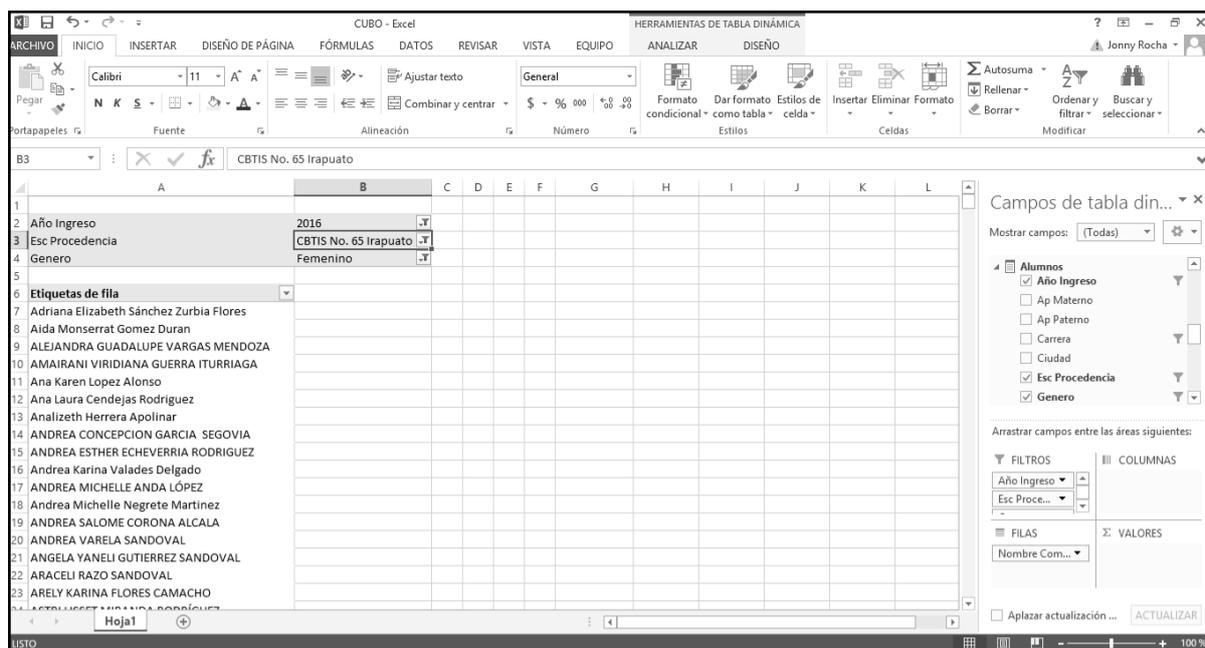


Figura 21. Resultado de una consulta en el cubo de información empleando una hoja de cálculo en Excel.

- ¿Qué alumnos (nombre) de la generación 2016 de género masculino tienen por apellido paterno López? (véase Figura 18)

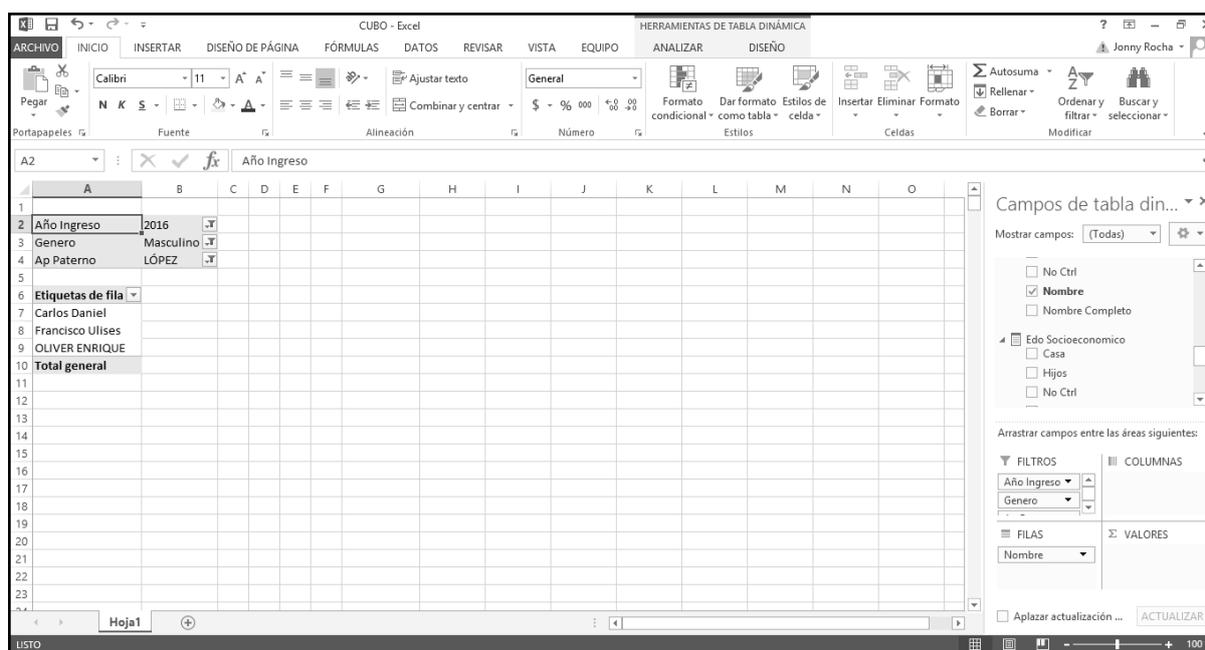


Figura 22. Resultado de una consulta en el cubo de información empleando una hoja de cálculo en Excel.

- ¿Qué alumnos (matricula, nombre y apellido paterno) de Ing. Industrial de la generación 2016 tienen hijos y han tenido relaciones sexuales recientemente? (véase Figura 19)

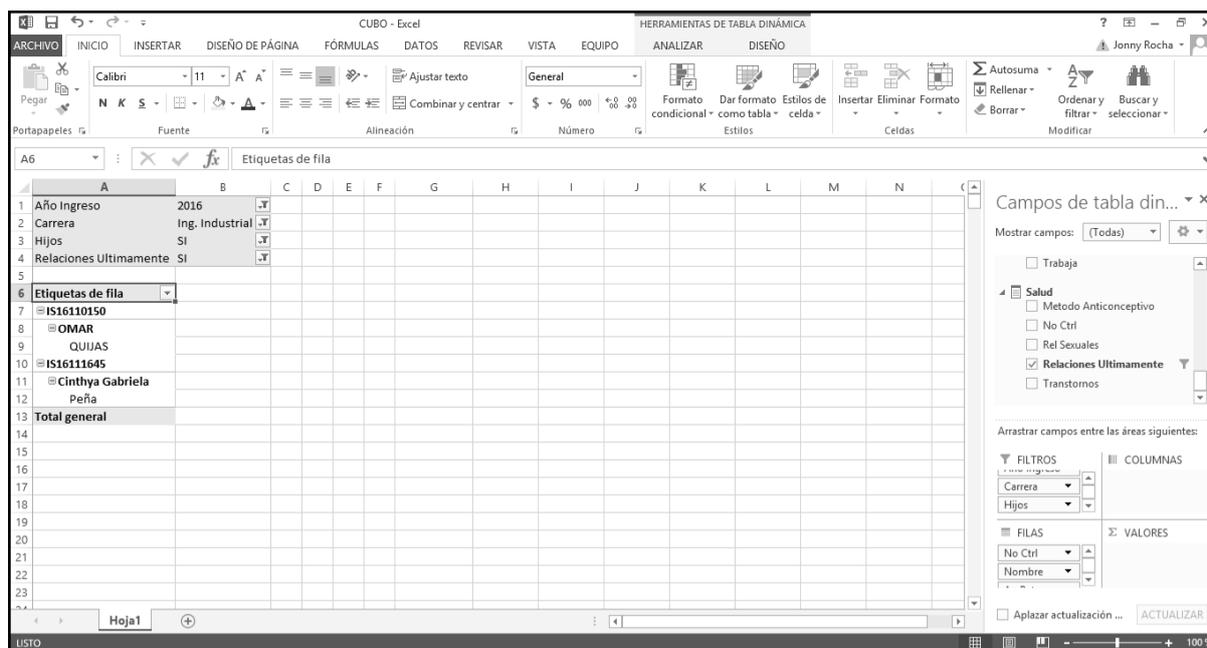


Figura 23. Resultado de una consulta en el cubo de información empleando una hoja de cálculo en Excel.

- ¿Qué alumnos (nombre completo) de la generación 2016 han tenido relaciones sexuales recientemente utilizado algún método anticonceptivo? (véase Figura 20)

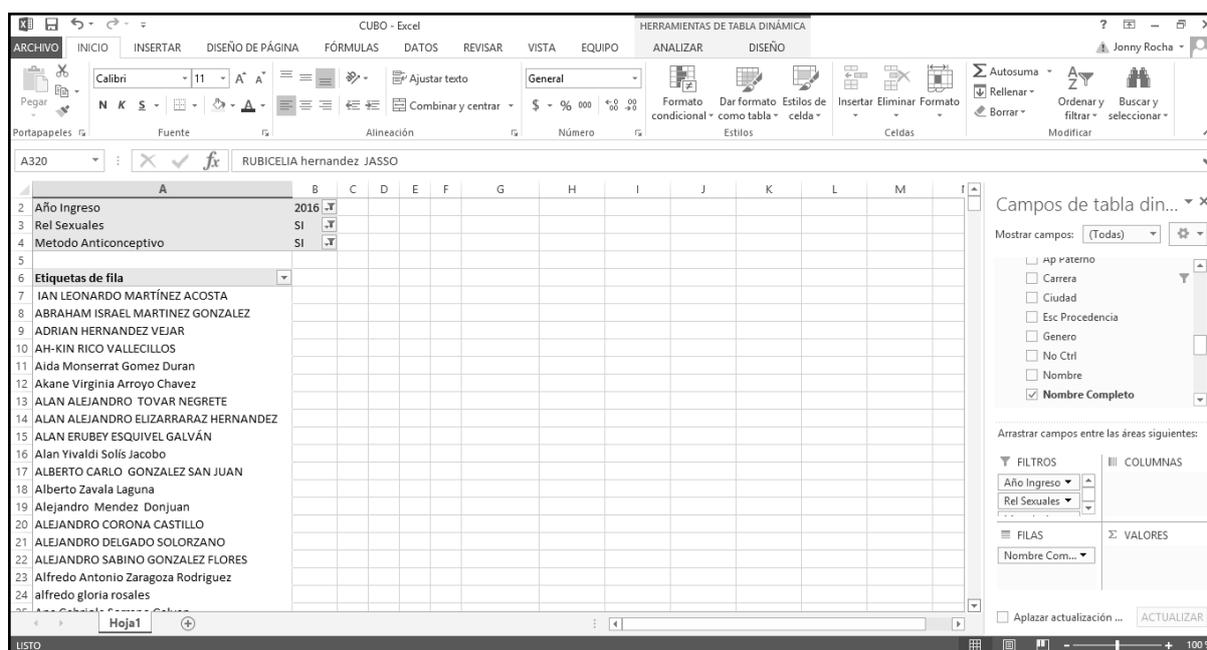


Figura 24. Resultado de una consulta en el cubo de información empleando una hoja de cálculo en Excel.

### 3 Conclusiones y trabajos futuros

Una de las ventajas más palpables que nos ofrece el cubo de información consiste en la reducción de tiempo que conlleva desarrollar las consultas para analizar y conocer mejor la situación de los alumnos del tecnológico.

La medición del tiempo para cada una de las consultas definidas en los *resultados preliminares* se muestra en la tabla 6, siendo esta tabla el resultado final que determina el cumplimiento de nuestro objetivo general.

**Tabla 6.** Tabla comparativa del tiempo de respuesta de consultas.

Consulta	Tiempo Actual	Tiempo con cubo de información
¿Qué alumnos (nombre completo) de Ing. Gestión Empresarial de la generación 2016 tiene hijos?	220 minutos	2 minutos
¿Qué alumnos (nombre completo) de la generación 2016 de género femenino provienen de la escuela CBTis 65?	220 minutos	1 minuto
¿Qué alumnos (nombre) de la generación 2016 de género masculino tienen por apellido paterno López?	220 minutos	3 minutos
¿Qué alumnos (matrícula, nombre y apellido paterno) de Ing. Industrial de la generación 2016 tienen hijos y han tenido relaciones sexuales recientemente?	60 minutos	2 minutos
¿Qué alumnos (nombre completo) de la generación 2016 han tenido relaciones sexuales recientemente utilizando algún método anticonceptivo?	240 minutos	2 minutos

Otra ventaja es la disposición organizada y de fácil acceso a la información por medio de un sistema que será útil como medio de consulta ante organismos acreditadores tales como el *Consejo Nacional de Acreditación en Informática y Computación*.

El desarrollo del cubo de información para el análisis de indicadores de ingreso en programas acreditados como los registrados en el Instituto Tecnológico Superior de Irapuato es una herramienta tecnológica que será una base sólida para la toma de decisiones, además de agilizar el procesamiento de este tipo de información permitiendo un análisis rápido y confiable al interior de la institución y su exterior, como lo son para aquellos organismos acreditadores cuya función primordial es la de evaluar y verificar que los procesos académicos de cada institución se lleven adecuadamente ofreciéndole en todo momento a la sociedad programas de estudio con calidad.

Como trabajo futuro se planea implementar un módulo que permita generar reportes impresos de manera automática, así como predefinir algunas consultas analíticas que son necesarias en cada ciclo escolar, lo que daría como resultado un sistema ad hoc al área tutorial con menos tiempo de respuesta a las necesidades de este departamento.

**Agradecimientos.** A la Dirección Académica del Instituto Tecnológico Superior de Irapuato por la apertura del Departamento de Tutorías para lograr este proyecto que será una herramienta tecnológica que apoyará la toma de decisiones bajo una base que la sustente.

### Referencias

1. Baker, S. (2009). *Los Numerati*. Editorial Planeta Mexicana S.A de C.V.
2. Cabrera Torres, A. A., Emma, M. C., & Iván, R. A. (2014). Uso de la Tecnología Data Warehouse en Unidades Educativas de Nivel Medio: Consideraciones Teóricas. *Revista Ciencia UNEMI*(11), 51 - 57. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/5210252.pdf>
3. Giner de la Fuente, F., & Gil Estallo, M. d. (2004). *Los sistemas de información en la sociedad del conocimiento*. ESIC.

4. Guevara Lenis, J., & Valencia Arcos, J. (2007). *Data Warehouse para el análisis académico de la escuela politécnica nacional*. Escuela Politécnica Nacional del Ecuador. EPN.
5. Han, J., & Kamber, M. (2001). *Data Mining Concepts and Techniques*. Morgan.
6. Humphries, M., Hawkins, M. W., & Dy, M. C. (2009). *Data Warehousing: Architecture and Implementation*. Prentice Hall.
7. Instituto Tecnológico Superior de Irapuato. (2017). *Página Oficial del Instituto Tecnológico Superior de Irapuato*. Obtenido de [http://www.itesi.edu.mx/vida\\_en\\_el\\_campus/estudiantes/tutorias.html](http://www.itesi.edu.mx/vida_en_el_campus/estudiantes/tutorias.html)
8. Kay, R. (2004). Data Cubes. *ComputerWorld The Voice of Business Technology*. Obtenido de <http://www.computerworld.com/article/2564238/business-intelligence/data-cubes.html>
9. Mohania, M. (2009). *Data Warehousing and Knowledge Discovery*. Springer.
10. Nedim, D., & Clare, S. (2016). An Evaluation of the Challenges of Multilingualism in Data. *In Proceedings of the 18th International Conference on Enterprise Information Systems, 1*, 196-206.
11. Rivera Hernández, G. (2010). *Metodologías utilizadas en la Oficina de Ingeniería de Software para el desarrollo de software*. Irapuato.
12. Softeng Software Engineers. (2010). *Metodología Scrum para desarrollo de software - aplicaciones complejas*. Obtenido de <https://www.softeng.es/es-es/empresa/metodologias-de-trabajo/metodologia-scrum.html>
13. Universidades Andaluzas y Consejería de Educación y Ciencia de la Junta de Andalucía. (2005). *DATA WAREHOUSE*. Obtenido de [https://www.gestion.uco.es/gestion/datawarehouse/doc/m\\_usuario/01000\\_usuario.pdf](https://www.gestion.uco.es/gestion/datawarehouse/doc/m_usuario/01000_usuario.pdf)
14. Zambrano Matamala, C., Rojas Díaz, D., Carvajal Cuello, K., & Acuña Leiva, G. (2011). Análisis de rendimiento académico estudiantil usando data warehouse y redes neuronales. *Ingeniare. Revista chilena de ingeniería*, 19(3), 369-381. Obtenido de [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-33052011000300007&lng=es&nrm=iso&tlng=es](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-33052011000300007&lng=es&nrm=iso&tlng=es)