



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE INGENIERÍA

**Sistema de consulta basado en inteligencia de negocios para el
análisis de la información derivada de la evaluación al desempeño
de los profesores de la Facultad de Ingeniería, UNAM**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Ingeniero en Computación

P R E S E N T A :

**Alan Jesús Pérez Shibayama
Iván Arellano García**



**DIRECTORA DE TESIS:
Ing. Gabriela Betzabé Lizárraga Ramírez**

Prólogo	4
Situación Actual	5
Planteamiento del problema	5
Objetivo General	5
1. CAPÍTULO I. Marco Teórico	7
1.1. Datos, información y conocimiento	7
1.1.1. Datos	7
1.1.2. Información	7
1.1.3. Conocimiento	8
1.2. Base de Datos	8
1.3. Diccionario de datos	8
1.4. Sistema Gestor de Bases de Datos	9
1.5. Extracción, transformación y consolidación de datos (ETL)	9
1.5.1. Extracción de datos	9
1.5.2. Transformación de datos	10
1.5.3. Consolidación de Datos	10
1.6. Data Warehouse (Almacén de datos)	11
1.7. Modelo estrella	11
1.8. Modelo copo de nieve	12
1.9. Arquitectura Cliente-Servidor	12
2. CAPÍTULO II. Generalidades del concepto de Inteligencia de Negocios	13
2.1. Introducción a la Inteligencia de Negocios	13
2.2. Breve historia de la Inteligencia de Negocios	13
3. CAPÍTULO III. IBM Data Stage 7.5.2 e IBM Cognos 10.1.1 como herramientas de Inteligencia de Negocios	15
3.1. IBM Data Stage 7.5.2	15
3.1.1. IBM Data Stage Administrator 7.5.2	16
3.1.2. IBM Data Stage Designer 7.5.2	16
3.1.3. IBM Data Stage Director 7.5.2	17
3.1.4. IBM Data Stage Manager 7.5.2	18
3.2. IBM Cognos 10.1.1	18

3.2.1.	IBM Cognos Connection 10.1.1.....	19
3.2.2.	IBM Cognos Query Studio 10.1.1.....	20
3.2.3.	IBM Cognos Report Studio 10.1.1.....	21
3.2.4.	IBM Cognos Analysis Studio 10.1.1.....	22
3.2.5.	IBM Cognos Transformer 10.1.1.....	23
3.2.6.	IBM Cognos Framework Manager 10.1.1.....	24
4.	CAPÍTULO IV. Desarrollo del proyecto.....	25
4.1.	Requerimiento.....	25
4.2.	Análisis.....	25
4.3.	Diseño del Data Warehouse.....	26
4.4.	Implementación del proceso ETL.....	32
4.4.1.	Tarea para la extracción de los datos proporcionados por la Facultad de Ingeniería de la UNAM.....	33
4.4.2.	Tarea para la actualización de la tabla DIVISIONES.....	35
4.4.3.	Tarea para la actualización de la tabla DEPARTAMENTOS.....	38
4.4.4.	Tarea para la actualización de la tabla PROFESORES.....	40
4.4.5.	Tarea para la actualización de la tabla ASIGNATURAS.....	42
4.4.6.	Tarea para la actualización de la tabla SEMESTRES.....	44
4.4.7.	Tarea para la actualización de la tabla HECHOS.....	46
4.5.	Implementación de un modelo con IBM Cognos Framework Manager 10.1.1.....	48
4.5.1.	Elaboración de un modelo de metadatos.....	49
4.5.2.	Generación y publicación de un paquete de metadatos.....	52
4.6.	Consultando la información de las encuestas a profesores de la Facultad de Ingeniería de la UNAM.....	54
4.6.1.	Reporte analítico para Directivos.....	54
4.6.1.1.	Main Report – División.....	54
4.6.1.2.	Detail Report – Departamento.....	55
4.6.1.3.	Detail Report – Asignatura.....	56
4.6.1.4.	Detail Report – Profesores.....	57
4.6.1.5.	Detail Report - Detalle Profesores.....	58
4.6.2.	Reporte analítico para Profesores.....	59
4.6.2.1.	Main Report - Reporte Profesores.....	59

4.6.3.	Generación de un cubo de información por medio de IBM Cognos Transformer 10.1.1 ..	61
4.6.3.1.	Creación del cubo de información	62
4.6.3.2.	Consultando la información contenida en el cubo de información por medio de IBM Cognos Analysis Studio 10.1.1	64
Conclusiones	65
Áreas de oportunidad del proyecto	66
Índice de Figuras	67
Bibliografía	67

Prólogo

Las universidades presentan estructuras rígidas y son presionadas por el entorno cambiante, lo que afecta la calidad en ellas. Las Tecnologías de la Información (TI) son indiscutiblemente herramientas indispensables para generar ventaja competitiva. Es posible aprovechar las características de las TI y utilizar modelos de autoevaluación para mejorar la gestión de la calidad en las universidades.

La Inteligencia de Negocios es un modelo que puede ser aplicado a instituciones de educación superior para mejorar la gestión de la calidad en ellas.

En este trabajo se propone la implementación de un modelo de inteligencia de negocios que, haciendo uso de las tecnologías de información y basado en un conjunto de encuestas aplicadas a los alumnos para la evaluación del personal docente de la Facultad de Ingeniería (FI) de la UNAM, permita tanto a los profesores, como a los directivos, identificar fortalezas y debilidades tales como su capacitación, puntualidad, motivación, término del temario y desempeño al momento de impartir cátedra al alumnado, facilitando así la gestión de la calidad y la toma de decisiones en la FI de la UNAM.

Situación Actual

Durante las últimas semanas de clases de cada semestre, en la Facultad de Ingeniería de la UNAM, son aplicadas a los alumnos una serie de preguntas relacionadas con el desempeño de cada profesor a lo largo del semestre, estas encuestas son capturadas y consolidadas dentro de archivos Excel, de estos archivos Excel es posible generar reportes que muestran los resultados obtenidos en el semestre inmediato anterior y son entregados de forma personal y en físico a cada uno de los profesores

El objetivo de hacer un análisis a los datos proporcionados por los alumnos con relación al desempeño de cada uno de los profesores, es ofrecer un reporte que le permita a los docentes identificar fortalezas y debilidades al impartir cada una de sus asignaturas correspondientes.

Planteamiento del problema

En la Facultad de Ingeniería de la UNAM son aplicadas a los alumnos una serie de encuestas, con el objetivo de evaluar el desempeño de los profesores a lo largo del curso, dichas encuestas son capturadas y analizadas. Posteriormente, los resultados arrojados son entregados de forma personal y en físico a cada uno de los profesores y directivos.

Para los profesores los resultados cumplen con su cometido, pues sirven como parámetros para observar fortalezas y debilidades al impartir clases, sin embargo, no existe una aplicación en línea y en tiempo real que les permita evaluar e identificar visualmente el desempeño a lo largo de todos y cada uno de los cursos en los diferentes semestres donde se impartió cátedra, con ello sería fácil identificar si las acciones correctivas que están tomando son las indicadas.

Para los directivos, al igual que los profesores, no existe una aplicación en línea y en tiempo real que permita evaluar el desempeño de cada uno de los profesores a lo largo de los semestres impartidos, además de no contar con reportes personalizados que permitan hacer comparativos y establezcan indicadores, con ello sería fácil identificar si los profesores están tomando medidas para mejorar la calidad de sus clases o bien si han tenido algún retroceso.

Objetivo General

Desarrollar una aplicación que permita:

- A los directivos, tener acceso a un reporte analítico y comparativo de todos y cada uno de los profesores a lo largo de los diferentes cursos en los diferentes semestres donde se impartió cátedra, además de permitirles elaborar consultas personalizadas en base a un cubo de

información, el cual permitirá hacer las consultas con mayor velocidad y utilizar las funcionalidades que nos provee.

- A los profesores, tener acceso a un reporte analítico de su desempeño particular a lo largo de los diferentes cursos impartidos en los diferentes semestres donde impartió cátedra.

1. CAPÍTULO I. Marco Teórico

A continuación se detallan algunos de los conceptos que a lo largo de este proyecto son mencionados y de gran importancia para el entendimiento del mismo.

1.1. Datos, información y conocimiento

Primero es importante mencionar la diferencia que existe entre los datos, la información y el conocimiento, pues aunque de primera instancia aparentan tener un significado parecido, existe una diferencia al referirnos a ellos.

1.1.1. Datos

Los datos [1] son la mínima unidad semántica, y se corresponden con elementos primarios de información que por sí solos son irrelevantes como apoyo a la toma de decisiones. También se pueden ver como un conjunto discreto de valores, que no dicen nada sobre el porqué de las cosas y no son orientativos para la acción.

Un número telefónico o un nombre de una persona, por ejemplo, son datos que, sin un propósito, una utilidad o un contexto no sirven como base para apoyar la toma de una decisión. Los datos pueden ser una colección de hechos almacenados en algún lugar físico como un papel, un dispositivo electrónico, o la mente de una persona. En este sentido las tecnologías de la información han aportado mucho a la recopilación de datos.

1.1.2. Información

La información [1] se puede definir como un conjunto de datos procesados y que tienen un significado (relevancia, propósito y contexto), y que por lo tanto son de utilidad para quién debe tomar decisiones, al disminuir su incertidumbre. Los datos se pueden transformar en información añadiéndoles valor:

- Contextualizando: Se sabe en qué contexto y para qué propósito se generaron.
- Categorizando: Se conocen las unidades de medida que ayudan a interpretarlos.
- Calculando: Los datos pueden haber sido procesados matemática o estadísticamente.
- Corrigiendo: Se han eliminado errores e inconsistencias de los datos.
- Condensando: Los datos se han podido resumir de forma más concisa (agregación).

Por tanto, la información es capaz de cambiar la forma en que el receptor percibe algo, impactando sobre sus juicios de valor y sus comportamientos.

1.1.3. Conocimiento

El conocimiento [1] es una mezcla de experiencia, valores, información, que sirve como marco para la incorporación de nuevas experiencias e información, y es útil para la acción. Se origina y aplica en la mente de los conocedores. En las organizaciones con frecuencia no sólo se encuentra dentro de documentos o almacenes de datos, sino que también está en rutinas organizativas, procesos, prácticas y normas.

El conocimiento se deriva de la información, así como la información se deriva de los datos. Para que la información se convierta en conocimiento es necesario realizar acciones como:

- Comparación con otros elementos.
- Predicción de consecuencias.
- Búsqueda de conexiones.
- Conversación con otros portadores de conocimiento.

1.2. Base de Datos

Para todos los proyectos de Inteligencia de Negocios es importante contar con un repositorio que nos permita almacenar todos y cada uno de los datos de interés, a este almacén se le conoce como base de datos y podemos definirlo de la siguiente manera:

Una base de datos es un “almacén” que nos permite guardar grandes cantidades de datos de forma organizada para que luego sean encontrados y utilizados fácilmente.

1.3. Diccionario de datos

Un diccionario de datos es un conjunto de metadatos que contiene las características lógicas y puntuales de los datos considerados dentro de la Base de Datos, ayuda a identificar los procesos donde se emplean los datos y los sitios donde se necesita el acceso inmediato a la información, se desarrolla durante el análisis de flujo de datos y auxilia a los analistas que participan en la determinación de los requerimientos del sistema. [2]

1.4. Sistema Gestor de Bases de Datos

Otro concepto importante a considerar dentro del ámbito de las bases de datos es el Sistema Gestor de Bases de Datos (SGBD), el cual se define como el conjunto de programas que acceden y gestionan una colección de datos relacionados entre sí. [3] Para nuestro proyecto utilizaremos “ORACLE data base express edition 10g” como nuestro SGBD.

1.5. Extracción, transformación y consolidación de datos (ETL)

En ocasiones los datos se encuentran en diferentes fuentes, o bien no se encuentran del todo bien definidos y además es necesario ubicarlos en nuevos repositorios de datos, para ello es necesario definir los conceptos de Extracción, Transformación y Carga de datos, frecuentemente abreviado ETL (por sus siglas en inglés Extract, Transform and Load).

Un proceso ETL permite a una organización mover datos desde múltiples fuentes, reformatearlos, limpiarlos y cargarlos en otra base de datos, data warehouse, o en otro sistema operacional para apoyar un proceso de organización [4]

1.5.1.Extracción de datos

La primera fase del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o archivos planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Una parte intrínseca del proceso de extracción es la de analizar los datos extraídos, de lo que resulta una revisión que verifica si los datos cumplen la pauta o estructura que se esperaba. De no ser así los datos son rechazados.

Un requerimiento importante que se debe exigir a la tarea de extracción es que esta cause un impacto mínimo en el sistema origen. Si el proceso de extracción es largo y complicado, el sistema origen se podría colapsar, provocando que éste no pueda utilizarse con normalidad para su uso cotidiano. Por esta razón, en sistemas grandes las operaciones de extracción suelen programarse en horarios o días donde este impacto sea nulo o mínimo. [4]

1.5.2. Transformación de datos

La segunda fase de un proceso ETL se conoce como transformación de datos, es en esta fase donde se aplican una serie de reglas de organización o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. [4] Algunas fuentes de datos requerirán alguna pequeña manipulación de los datos. No obstante en otros casos puede ser necesario aplicar diversas transformaciones, a continuación se mencionan solo algunas:

- Seleccionar solo ciertas columnas para su carga (ej. Columnas con valores nulos no deben ser cargadas)
- Traducir códigos (ej. Si la fuente almacena un "1" para hombres y una "2" para mujeres el destino deberá almacenar un "H" para hombres y un "M" para mujeres)
- Obtener nuevos valores calculados (ej. Total venta = cantidad * precio)
- Unir datos de múltiples fuentes
- Generación de campos clave en el destino
- La aplicación de cualquier forma, simple o compleja de validación de datos, y la consiguiente aplicación de la acción que en cada caso se requiera:
 - o Datos correctos: Entregar datos a la siguiente etapa
 - o Datos erróneos: Ejecutar políticas de tratamiento de excepciones (ej. Rechazar un registro completo, dar valor nulo a campo erróneo)

1.5.3. Consolidación de Datos

La tercera y última fase de un proceso ETL se conoce como consolidación o carga de datos, es en ésta etapa en donde todos los datos seleccionados en la fase de transformación son cargados en la fuente destino. Es importante mencionar que en esta etapa ya se debe tener bien definido un modelo de datos, es decir, un conjunto de herramientas conceptuales para describir datos, sus relaciones, su significado y sus restricciones de consistencia.

Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. [4]

La siguiente figura describe de manera general el flujo que sigue un proceso de ETL.

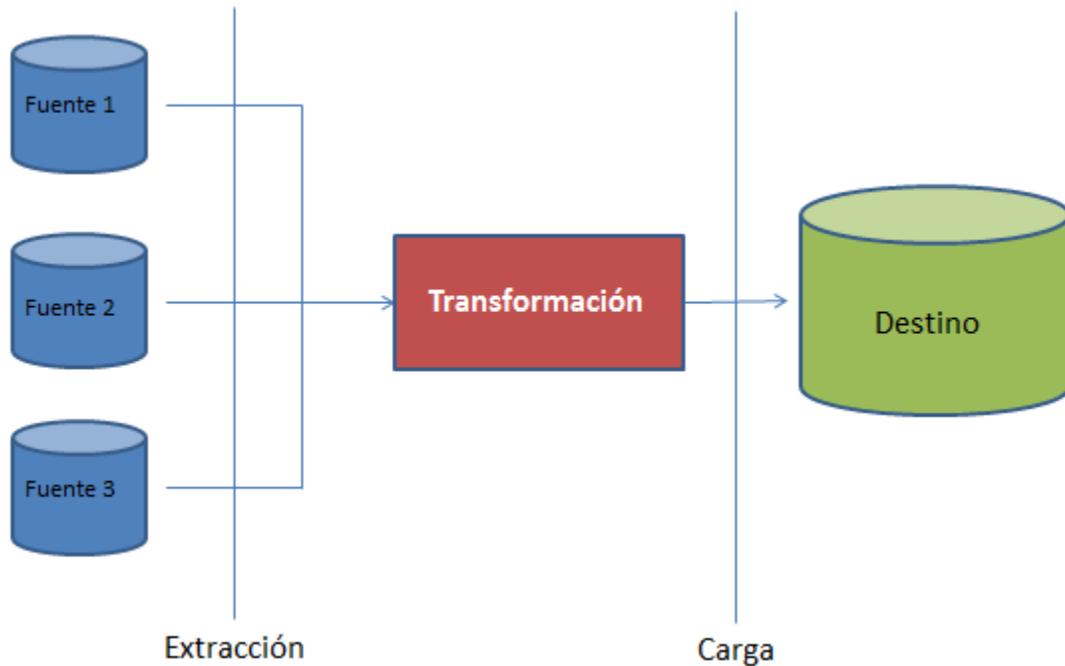


Figura 1 Diagrama con las fases de un proceso ETL

1.6. Data Warehouse (Almacén de datos)

Una vez definidos los conceptos básicos de Bases de Datos podemos hablar de un nuevo concepto, data warehouse (DW), el cual, si bien es una Base de Datos, es una colección de datos orientada a un determinado ámbito, integrado y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza.

Los data warehouse mantienen un historial de los registros de manera que se puede hacer una auditoría de los mismos y disponer de un rastreo de toda la historia de un valor a lo largo del tiempo.

1.7. Modelo estrella

El modelo estrella es una de las estructuras bajo la cual se guía la creación de un data warehouse y es el más sencillo en estructura. Consiste en un modelo con una tabla centralizada, denominada tabla de hechos, que contiene los datos para el análisis, rodeada de varias dimensiones, las cuales contienen información descriptiva. Lo característico del modelo de estrella es que sólo existe una dimensión para cada dimensión.

Este modelo es ideal por su simplicidad y velocidad para ser usado en análisis multidimensionales, además de permitir acceder tanto a datos agregados como de detalle.

1.8. Modelo copo de nieve

El modelo copo de nieve es una variedad más compleja del modelo estrella. Lo que distingue a la estructura en copo de nieve del modelo estrella, es que las dimensiones en este modelo representan relaciones normalizadas y forman parte de un modelo relacional de base de datos. El uso más común de este tipo de arquitectura es cuando las dimensiones están muy grandes o complejas y es muy difícil representar los datos en esquema estrella.

1.9. Arquitectura Cliente-Servidor

Una de las incógnitas que existe por lo regular al desarrollar un proyecto de Inteligencia de Negocios es la manera en que los datos van a ser consultados, es aquí donde el concepto de arquitectura toma gran relevancia y existen diferentes tipos de arquitecturas que permiten realizar dicha tarea, sin embargo, una de las más comunes es la arquitectura cliente-servidor.

La arquitectura cliente-servidor es un modelo de aplicación distribuida en el que las tareas se reparten entre los proveedores de recursos o servicios, llamados servidores, y los demandantes, llamados clientes. Un cliente o usuario realiza peticiones, el servidor, quien le da respuesta.

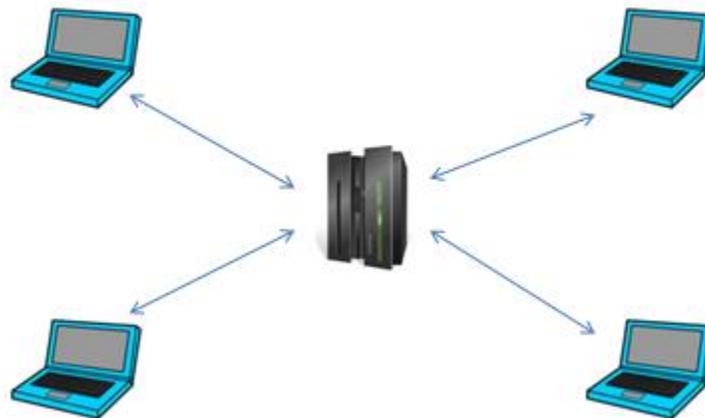


Figura 2 Diagrama de arquitectura Cliente-Servidor

2. CAPÍTULO II. Generalidades del concepto de Inteligencia de Negocios

2.1. Introducción a la Inteligencia de Negocios

Howard Dresner propuso la Inteligencia de Negocios como un término general para describir los conceptos y métodos para mejorar la toma de decisiones organizacionales mediante el uso de sistemas basados en hechos de apoyo. [6]

Desde un punto de vista asociado directamente con las tecnologías de la información, podemos definir Inteligencia de Negocios como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada en información estructurada, para su explotación o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre de la organización. [7]

2.2. Breve historia de la Inteligencia de Negocios

En un tiempo, las organizaciones dependían de sus departamentos de sistemas de información para proporcionarles reportes estándar y personalizados. Esto ocurrió en los días de los mainframes y minicomputadoras, cuando la mayoría de los usuarios no tenía acceso directo a las micro computadoras. Sin embargo, esto comenzó a cambiar en los años 70's cuando los sistemas basados en servidores se convirtieron en una realidad.

Aun así estos sistemas eran usados principalmente para transacciones de negocios y sus capacidades de realizar reportes se limitaba a un número predefinido de ellos. Los sistemas de información se sobrecargaban y los usuarios tenían que esperar por días o semanas para obtener sus reportes en caso que requirieran reportes distintos a los estándares disponibles.

Algunos de los hitos históricos que nos han llevado hasta la situación actual del concepto de Inteligencia de Negocios son los siguientes:

- 1969: Introducción del concepto de base de datos relacionales. [8]
- 1970's: Desarrollo de las primeras bases de datos y las primeras aplicaciones empresariales (JD Edwards, Siebel). Estas aplicaciones permitieron realizar "Entrada de datos" en los sistemas, aumentando la información disponible, pero no fueron capaces de ofrecer un acceso rápido y fácil a dicha información. [8]

- 1980s: Creación del concepto Data warehouse, y aparición de los primeros sistemas de reporte. A pesar de todo, seguía siendo complicado y funcionalmente pobre. Existían relativamente potentes sistemas de bases de datos pero no había aplicaciones que facilitaran su explotación. [8]
- 1989: Introducción del término de Inteligencia de Negocios. [8]
- 1990s: Proliferación de múltiples aplicaciones de Inteligencia de Negocios.
- 2000s: Consolidación de las aplicaciones de Inteligencia de Negocios en plataformas de Inteligencia de Negocios (Oracle, SAP, IBM, Microsoft). [8]

3. CAPÍTULO III. IBM Data Stage 7.5.2 e IBM Cognos 10.1.1 como herramientas de Inteligencia de Negocios

En el ámbito de Inteligencia de Negocios existen una gran cantidad de compañías que nos ofrecen soluciones completas para el desarrollo de un proyecto de Inteligencia de Negocios, algunas de dichas compañías son:

- SAP
- Pentaho
- Oracle
- Micro Strategy
- IBM

Decidimos trabajar con las herramientas de IBM por 2 razones, la primera es que ambas herramientas cuentan con interfaces sencillas que permiten la realización de tareas específicas y complejas de una forma simple, ordenada y sobre todo eficaz, la segunda es por ser herramientas dominantes dentro del mercado de Inteligencia de Negocios.

IBM Data Stage 7.5.2 provee una interfaz para la realización del proceso ETL, mientras que IBM Cognos 10.1.1 provee las herramientas para el análisis de la información.

3.1. IBM Data Stage 7.5.2

Como mencionamos anteriormente IBM cuenta con 2 herramientas para realizar tareas específicas, en el caso de tareas relacionadas con la extracción, transformación y carga de datos nos provee la herramienta IBM Data Stage en su versión 7.5.2.

Originalmente bajo el nombre de Ardent DataStage, para luego ser adquirido por Ascential en el año 2001 y comprado por IBM en el año 2005, IBM Data Stage es un set de herramientas que permiten la extracción, integración y transformación de altos volúmenes de datos desde estructuras simples hasta muy complejas.

Basado en una arquitectura cliente-servidor, IBM Data Stage 7.5.2 es administrado y configurado por medio de 4 herramientas cliente:

- IBM Data Stage Administrator 7.5.2
- IBM Data Stage Designer 7.5.2

- IBM Data Stage Director 7.5.2
- IBM Data Stage Manager 7.5.2

3.1.1. IBM Data Stage Administrator 7.5.2

IBM Data Stage Administrator 7.5.2 es una interfaz muy sencilla que permite crear, modificar, eliminar, configurar y administrar los diferentes proyectos contenidos dentro del Servidor de IBM Data Stage 7.5.2. En la siguiente figura podemos apreciar la interfaz de esta herramienta cliente, donde se encuentra nuestro proyecto nombrado “TESIS”.

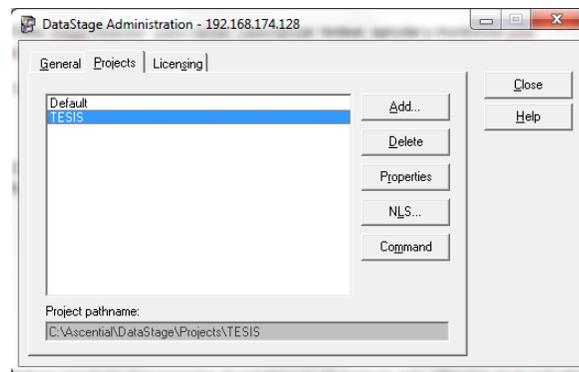


Figura 3 Interfaz IBM Data Stage Administrator 7.5.2

3.1.2. IBM Data Stage Designer 7.5.2

IBM Data Stage Designer 7.5.2 es la interfaz utilizada para crear y diseñar tareas Data Stage, adicionalmente también nos permite compilar y ejecutar dichas tareas Data Stage. Básicamente la interfaz consta de 3 secciones como podemos apreciar en la figura 4, del lado izquierdo se encuentra el repositorio de objetos creados y una paleta con los objetos disponibles para usar, mientras que nuestra área de trabajo se encuentra del lado derecho.

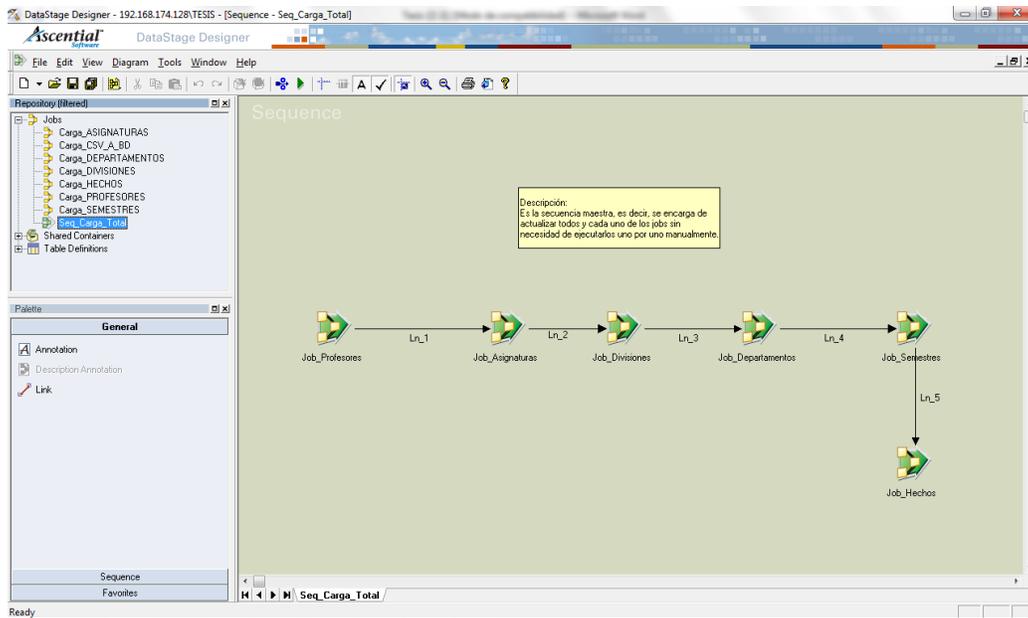


Figura 4 Interfaz IBM Data Stage Designer 7.5.2

3.1.3. IBM Data Stage Director 7.5.2

IBM Data Stage Director 7.5.2 es la interfaz utilizada para validar, calendarizar, ejecutar y monitorizar tareas Datastage. En la figura 5 podemos observar la interfaz utilizada por esta herramienta, donde se localizarán todos las tareas o secuencias creadas con información relevante, como la fecha y el tiempo de ejecución de cada objeto, así como su descripción y su estatus.

Job name	Status	Started	On date	Last ran	On date	Elapsed time	Description
Carga_ASSIGNATURAS	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	
Carga_CSV_A_BD	Finished (see log)	10:54 a.m.	04/01/2007	10:54 a.m.	04/01/2007	00:00:02	
Carga_DEPARTAMEN...	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	Source: C:\Ascential\DataStage\Projects\TESIS\BASES...
Carga_DIVISIONES	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	
Carga_HECHOS	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	
Carga_PROFESORES	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	Source: TB_ENCUESTAS_PROFESORES
Carga_SEMESTRES	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:02	Source: TB_ENCUESTAS_PROFESOREStarget: TB...
Seq_Carga_Total	Finished	09:02 a.m.	13/01/2007	09:02 a.m.	13/01/2007	00:00:22	

Figura 5 Interfaz IBM Data Stage Administrator 7.5.2

3.1.4. IBM Data Stage Manager 7.5.2

IBM Data Stage Manager 7.5.2 es la interfaz usada para visualizar y editar el contenido del repositorio, además de importar o exportar tareas Data Stage. Como se puede apreciar en la figura 6, la interfaz está dividida en 2 secciones, del lado izquierdo se encuentran los tipos de objetos utilizables por IBM Data Stage 7.5.2, mientras que en el lado derecho se localizan los objetos existentes en el repositorio.

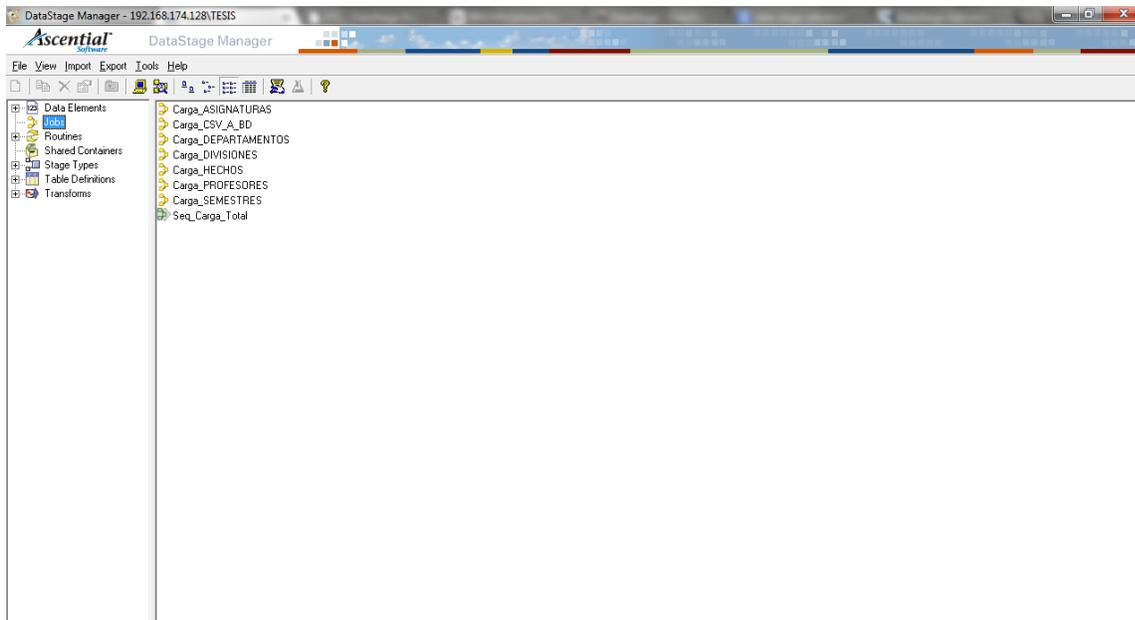


Figura 6 Interfaz IBM Data Stage Manager 7.5.2

3.2. IBM Cognos 10.1.1

Para la explotación de la información IBM nos provee de la herramienta IBM Cognos 10.1.1, la cual es una de las herramientas más utilizadas en la actualidad en cuanto a la realización de proyectos de Inteligencia de Negocios se refiere.

De acuerdo con la siguiente figura (Cuadrante mágico de Gartner), la solución para Inteligencia de Negocios proveída por IBM es bastante completa, manejable y una de las líderes del mercado.

El Cuadrante mágico de Gartner proporciona de manera gráfica un posicionamiento competitivo de cuatro tipos de proveedores de tecnología: [9]

Líderes: Ejecutan bien en contra de su visión actual y están en buena posición para mañana. [9]

Visionarios: Comprenden dónde va el mercado o tienen una visión para el cambio de las reglas del mercado, pero aún no ejecutan bien. [9]

Buena Posición: Centran su éxito en un pequeño segmento, pero están fuera de foco y no son muy innovadores. [9]

Retadores: Ejecutan bien hoy o pueden llegar a dominar una gran parte del mercado, pero no tienen claro hacia dónde se dirige el mercado. [9]

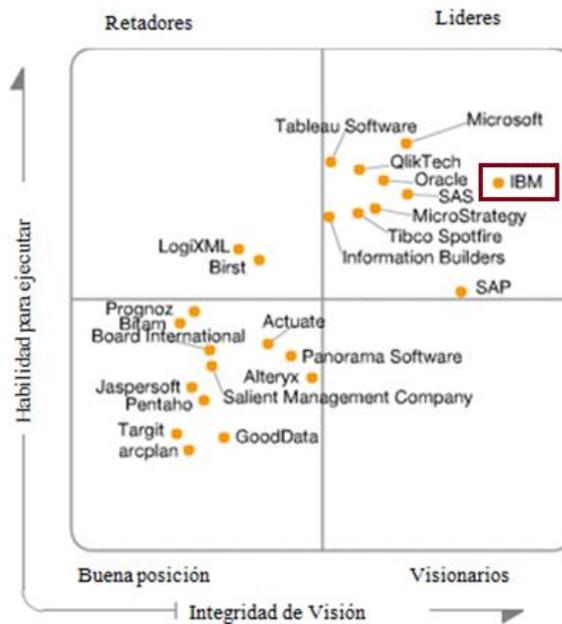


Figura 7 Cuadrante mágico de Gartner Inc [9]

IBM Cognos 10.1.1 a su vez es una suite con diferentes aplicaciones orientadas a cubrir un tipo de necesidad, la mayoría de ellas se maneja 100% desde un explorador Web, tanto para desarrollar o diseñar informes como para consultarlos o realizar tareas de análisis.

Entre las principales herramientas que maneja la suite se encuentran las siguientes:

3.2.1. IBM Cognos Connection 10.1.1

IBM Cognos Connection 10.1.1 es el portal y el punto de acceso a todos los datos disponibles dentro de IBM Cognos 10.1.1, en él es posible crear informes, análisis, consultas, accesos directos y visualizar cada uno de los paquetes existentes, incluso es posible personalizar este portal de acuerdo a las necesidades de la organización, incluyendo URLs. En la figura 8 se presenta la interfaz del portal de IBM Cognos Connection 10.1.1, en donde podemos observar las diferentes herramientas con las que cuenta la solución.

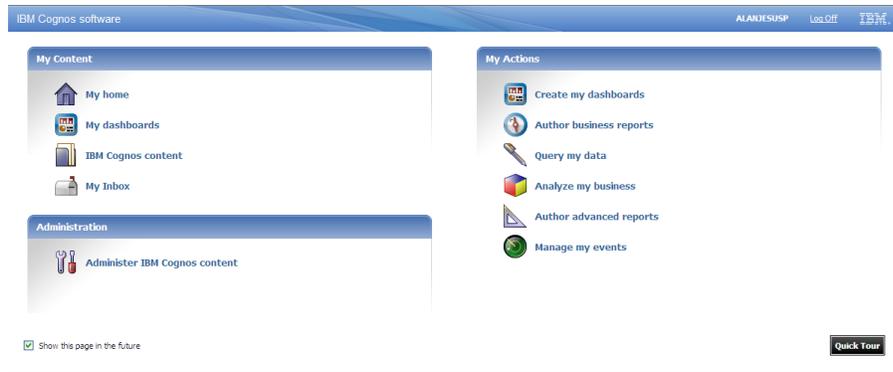


Figura 8 Interfaz IBM Cognos Connection 10.1.1

3.2.2. IBM Cognos Query Studio 10.1.1

Es la herramienta más simple, y se podría considerar más bien como un complemento. Su objetivo es permitir al usuario realizar consultas sencillas para resolver rápidamente cuestiones puntuales que le puedan surgir.

Permite acceder a la misma estructura de datos que utilizan las otras herramientas, tanto si se trata de un modelo relacional como si la estructura es dimensional. Con IBM Cognos Query Studio 10.1.1 se puede crear un informe en segundos arrastrando campos desde el explorador de datos hasta el área de diseño de informes. Permite también aplicar filtros, ordenaciones, operaciones de agrupación de datos e incluso crear gráficas. La figura 9 muestra la interfaz de esta herramienta, dividida en 2 secciones, del lado izquierdo se encuentran los objetos de consulta mientras que del lado derecho se encuentra el área de trabajo. Todas las opciones para realizar cálculos o insertar gráficos y formato se encuentran en la barra de herramientas localizada en la parte superior.

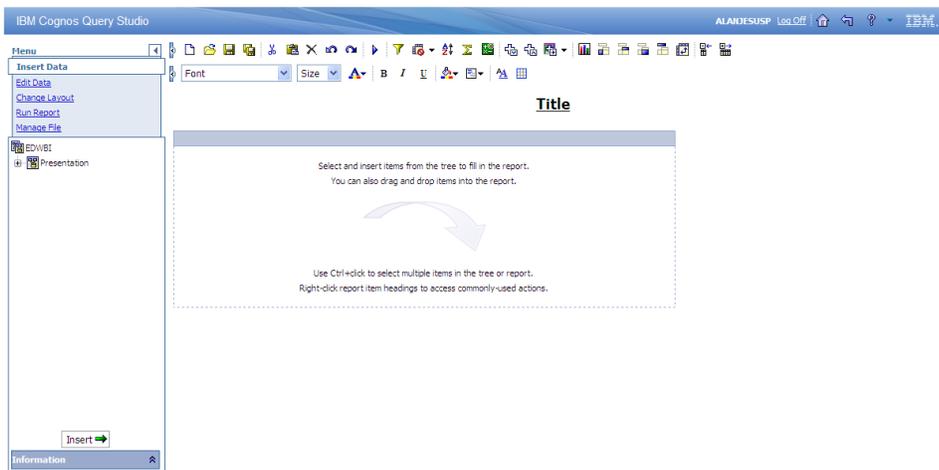


Figura 9 Interfaz IBM Cognos Query Studio

3.2.3. IBM Cognos Report Studio 10.1.1

Es la aplicación principal para la creación de informes. Se asemeja a IBM Cognos Query Studio 10.1.1, pero es mucho más completa.

Existen diferentes tipos de gráficas que se pueden incluir en los informes, mostrar de manera individualizada o guardar para formar parte de un cuadro de mando que se mostraría en el portal. Las opciones de utilización de parámetros son también bastante completas y como en todos estos tipos de herramientas, se pueden definir filtros, ordenar, agrupar y trabajar con agregados, crear subtotales, campos calculados y formateados condicionales.

Estos objetos pueden ser de diferentes tipos: origen de datos, datos específicos del informe y herramientas de diseño. Cada objeto que se incrusta en el informe tiene sus propiedades configurables, y mediante estas se puede llegar a un nivel muy alto de personalización, tanto en los datos que se muestran como en el diseño del formato. Se puede trabajar tanto con estructuras de datos relacionales como con estructuras dimensionales, sólo hay que tener en cuenta que en función del tipo de origen existen diferencias en cuanto a las propiedades aplicables a los datos, e incluso en cuanto al comportamiento en el área de diseño. Aunque no es obligatorio hacerlo así, para mostrar datos de estructura dimensional, lo más apropiado es utilizar informes de tipo tabla de referencias cruzadas.

A continuación, en la figura 10, se muestra la interfaz de esta herramienta, a la izquierda muestra un explorador de objetos desde el que se puede acceder a la estructura de datos, y a otros objetos insertables en los informes. A la derecha se encuentra el área de diseño del informe, donde se pueden arrastrar estos objetos e ir componiendo así la estructura.

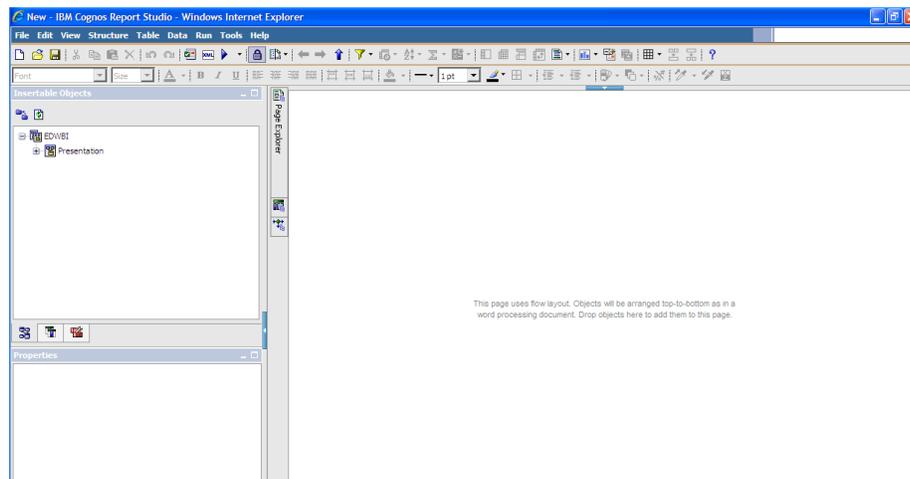


Figura 10 Interfaz IBM Cognos Report Studio 10.1.1

3.2.4. IBM Cognos Analysis Studio 10.1.1

IBM Cognos Analysis Studio 10.1.1 tiene una función similar a IBM Cognos Query Studio 10.1.1, pero para orígenes multidimensionales. Esta herramienta permite la navegación por estructuras multidimensionales como cubos OLAP. También puede trabajar con orígenes de datos relacionales, siempre que estén modelados dimensionalmente desde IBM Cognos Framework Manager 10.1.1.

El objetivo principal de este software es permitir que el analista de la organización pueda navegar por los datos cargados en las estructuras dimensionales sin depender del soporte del área de tecnologías de información. Utilizando IBM Cognos Analysis Studio 10.1.1 un usuario de negocio puede realizar análisis complejos y comparativos de datos para descubrir tendencias, riesgos y oportunidades.

El área de trabajo es similar a la de IBM Cognos Query Studio 10.1.1 e IBM Cognos Report Studio 10.1.1 y las opciones estándar para la creación de informes son muy similares a las de IBM Cognos Query Studio 10.1.1, con aplicación de filtros, ordenaciones, operaciones de agrupación de datos, creación de gráficas, etc.

Además ofrece funcionalidades más orientadas a orígenes dimensionales, como la navegación con drill up / drill down, o la creación de filtros de contexto.

En la figura 11, de la misma forma que las herramientas anteriores, a la izquierda muestra un explorador de objetos desde el que se puede acceder a la estructura de datos, y a otros objetos insertables en los informes. A la derecha se encuentra el área de diseño del informe, donde se pueden arrastrar estos objetos e ir componiendo así la estructura.

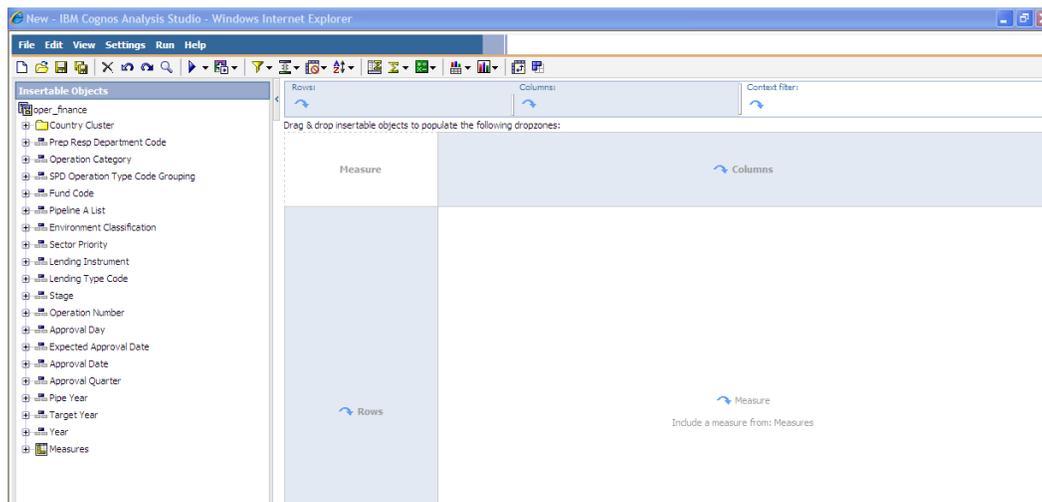


Figura 11 Interfaz IBM Cognos Analysis Studio 10.1.1

3.2.5. IBM Cognos Transformer 10.1.1

Las herramientas mencionadas anteriormente permiten acceder a cubos OLAP como origen de datos. IBM Cognos Transformer 10.1.1 es la herramienta que permite construir cubos OLAP.

Aunque las herramientas de reporteo de IBM Cognos pueden trabajar en ROLAP y atacar bases de datos relacionales, para realizar tareas analíticas lo más eficiente suele ser utilizar un cubo OLAP como origen de datos, es decir, trabajar en MOLAP. Con un volumen de datos controlado los tiempos de respuesta en la utilización de los informes pueden ser mucho mejores.

Con IBM Cognos Transformer 10.1.1 se definen los orígenes de datos, se modela la estructura multidimensional que va a conformar el cubo, se valida, y se procede a la construcción del mismo. Esta herramienta se instala en modo cliente, no se accede por web. Se utiliza de manera independiente para generar los cubos, que es lo que utilizan el resto de las herramientas web de IBM Cognos 10.1.1.

En la figura 12 se muestra la interfaz de esta herramienta, la cual es un poco más compleja, en ella se puede apreciar nuestra fuente de datos, las métricas definidas, el mapa de dimensiones y una vista previa de las categorías, así como el nombre del cubo que posteriormente será publicado dentro de IBM Cognos Connection 10.1.1.

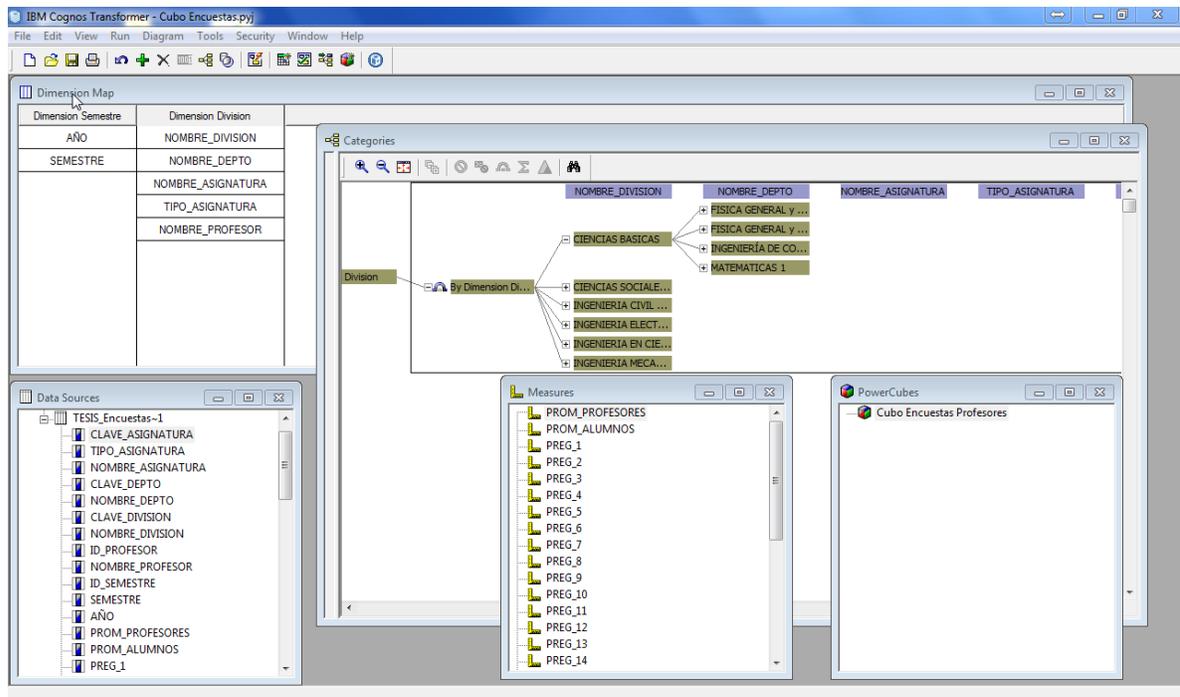


Figura 12 Interfaz IBM Cognos Transformer 10.1.1

4. CAPÍTULO IV. Desarrollo del proyecto

4.1. Requerimiento

Se requiere la implementación de un sistema de Inteligencia de Negocio con los datos de las encuestas aplicadas a los alumnos para la evaluación del desempeño de los profesores de la Facultad de Ingeniería de la UNAM a lo largo de los diferentes semestres donde se impartió cátedra.

4.2. Análisis

En la figura 14 podemos observar los pasos o fases que deben considerarse al desarrollar un proyecto de Inteligencia de Negocio.

Aplicado a nuestra problemática, la fuente de información será un archivo Excel que contiene los datos provenientes de las encuestas aplicados al alumnado para evaluar a los profesores proporcionado por la Facultad de Ingeniería de la UNAM, el cual, por medio de un proceso de ETL diseñado en IBM Data Stage 7.5.2, será almacenado en nuestro Data Warehouse, una vez en nuestro Data Warehouse, los datos serán explotados por las herramientas de IBM Cognos 10.1.1 para la realización tanto de los reportes analíticos, como del cubo de información.

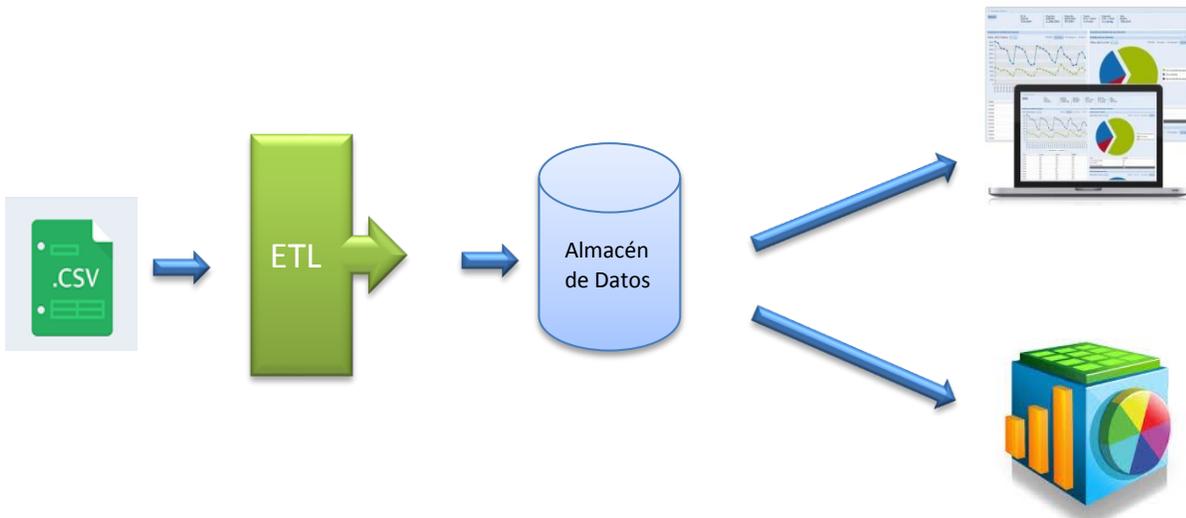


Figura 14 Fases en un proyecto de Inteligencia de Negocios

4.3. Diseño del Data Warehouse

Lo primero a realizar en el modelado de las diferentes tablas consideradas para nuestro proyecto, es el diseño de nuestro Data Warehouse, en la figura 15 se puede observar el diagrama del modelo estrella propuesto para este trabajo. El modelo estrella expuesto está compuesto por 5 dimensiones, cada tabla consta a su vez de una llave primaria y al menos un campo descriptivo, mientras que la tabla de hechos cuenta con todas las métricas que serán utilizadas para el análisis de datos y todas y cada una de las llaves primarias de las tablas.

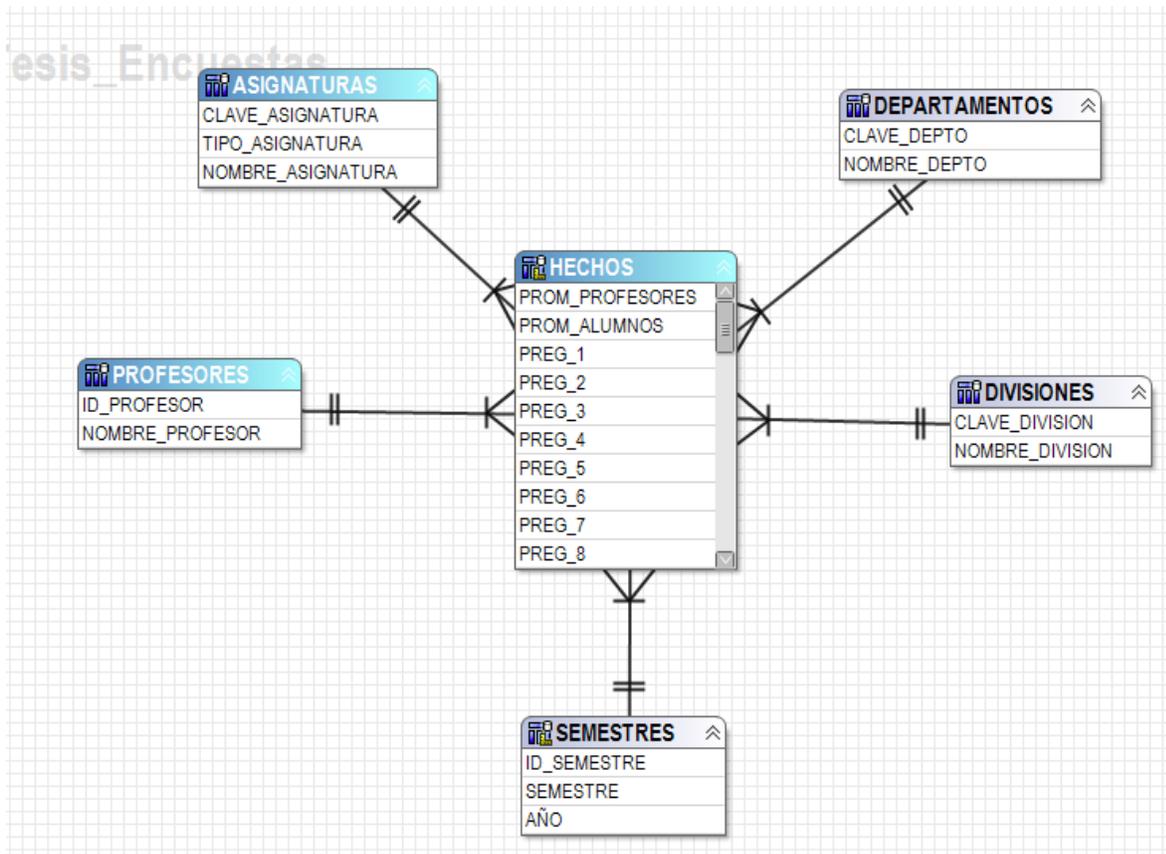


Figura 15 Modelo físico estrella del Data Warehouse propuesto

Es posible también notar que todas las dimensiones tienen establecidas la siguiente relación con la tabla de hechos:

- Cada elemento de la tabla de hechos puede tener únicamente un valor en la dimensión. (1,1)
- Cada elemento de la dimensión puede tener uno o más valores en la tabla de hechos (1,n)

4.3.1. Diccionario de datos

Una vez definido nuestro modelo de datos, es necesario contar con un diccionario de datos donde se describen las entidades y los atributos involucrados en nuestro modelo estrella:

- Tabla DIVISIONES

La tabla DIVISIONES almacena la información correspondiente a las divisiones existentes en la Facultad de Ingeniería de la UNAM. La clave de la división funge como nuestra llave primaria y cuenta con una columna referente al nombre de cada división. . A continuación se presenta el código SQL que muestra las columnas que conforman la tabla DIVISIONES, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como aprecia en la figura 16.

```
select
  b.column_id,
  a.*,
  b.nullable,
  b.data_type
from ALL_COL_COMMENTS a,
  ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
  a.column_name = b.column_name and
  a.owner = 'TESIS' AND
  A.TABLE_NAME = 'DIVISIONES'
order by b.table_name, b.column_id
```

COLUMN_ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
1	TESIS	DIVISIONES	CLAVE_DIVISION	Clave de la División. Primary Key	N	NUMBER
2	TESIS	DIVISIONES	NOMBRE_DIVISION	Nombre de la División	Y	VARCHAR2

Figura 16 Descripción tabla DIVISIONES

- Tabla DEPARTAMENTOS

La tabla DEPARTAMENTOS almacena la información correspondiente a los departamentos existentes en la Facultad de Ingeniería de la UNAM. La clave del departamento funge como nuestra llave primaria y cuenta con una columna referente al nombre de cada departamento. A continuación se presenta el código SQL que muestra las columnas que conforman la tabla DEPARTAMENTOS, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como se aprecia en la figura 17.

```

select
    b.column_id,
    a.*,
    b.nullable,
    b.data_type
from ALL_COL_COMMENTS a,
     ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
      a.column_name = b.column_name and
      a.owner = 'TESIS' AND
      A.TABLE_NAME = 'DEPARTAMENTOS'

order by b.table_name, b.column_id

```

COLUMN_ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
1	TESIS	DEPARTAMENTOS	CLAVE_DEPTO	Clave del Departamento. Primary Key	N	NUMBER
2	TESIS	DEPARTAMENTOS	NOMBRE_DEPTO	Nombre del Departamento	Y	VARCHAR2

Figura 17 Descripción tabla DEPARTAMENTOS

- Tabla PROFESORES

La tabla PROFESORES almacena la información correspondiente a los profesores existentes en la Facultad de Ingeniería de la UNAM. El RFC del profesor funge como nuestra llave primaria y cuenta con una columna referente al nombre de cada profesor. A continuación se presenta el código SQL que muestra las columnas que conforman la tabla PROFESORES, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como se aprecia en la figura 18.

```

select
    b.column_id,
    a.*,
    b.nullable,
    b.data_type
from ALL_COL_COMMENTS a,
     ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
      a.column_name = b.column_name and
      a.owner = 'TESIS' AND
      A.TABLE_NAME = 'PROFESORES'

order by b.table_name, b.column_id

```

COLUMN_ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
1	TESIS	PROFESORES	ID_PROFESOR	Identificador único de cada profesor. Primary key	N	NUMBER
2	TESIS	PROFESORES	NOMBRE_PROFE...	Nombre del profesor	Y	VARCHAR2

Figura 18 Descripción tabla PROFESORES

- Tabla ASIGNATURAS

La tabla ASIGNATURAS almacena la información correspondiente a las asignaturas existentes en la Facultad de Ingeniería de la UNAM. La clave de la asignatura funge como nuestra llave primaria y cuenta con una columna referente al nombre de cada asignatura, así como una columna que especifica si la asignatura en turno es de tipo teoría o práctica. A continuación se presenta el código SQL que muestra las columnas que conforman la tabla ASIGNATURAS, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como se aprecia en la figura 19.

```
select
    b.column_id,
    a.*,
    b.nullable,
    b.data_type
from ALL_COL_COMMENTS a,
     ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
      a.column_name = b.column_name and
      a.owner = 'TESIS' AND
      A.TABLE_NAME = 'ASIGNATURAS'
order by b.table_name, b.column_id
```

Column ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
1	TESIS	ASIGNATURAS	CLAVE_ASIGNATURA	Clave Asignatura. Primary key	N	NUMBER
2	TESIS	ASIGNATURAS	TIPO_ASIGNATURA	Tipo de la asignatura. Teoría o práctica.	Y	VARCHAR2
3	TESIS	ASIGNATURAS	NOMBRE_ASIGNATURA	Nombre de la asignatura	Y	VARCHAR2

Figura 19 Descripción tabla ASIGNATURAS

- Tabla SEMESTRES

La tabla SEMESTRES almacena la información correspondiente a todos los semestres de cada año donde las encuestas a los profesores fueron aplicadas. El id semestre funge como nuestra llave primaria y cuenta con una columna referente al año, así como una columna con información al número de semestre. A continuación se presenta el código SQL que muestra las columnas que conforman la tabla SEMESTRES, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como se aprecia en la figura 20.

```

select
  b.column_id,
  a.*,
  b.nullable,
  b.data_type
from ALL_COL_COMMENTS a,
  ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
  a.column_name = b.column_name and
  a.owner = 'TESIS' AND
  A.TABLE_NAME = 'SEMESTRES'
order by b.table_name, b.column_id

```

COLUMN_ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
1	TESIS	SEMESTRES	ID_SEMESTRE	Id semestre, compuesto por el año y semestre. ej. 20102. Primary Key	N	NUMBER
2	TESIS	SEMESTRES	SEMESTRE	Semestre, puede ser 1 ó 2	Y	NUMBER
3	TESIS	SEMESTRES	AÑO	Año	Y	NUMBER

Figura 20 Descripción tabla SEMESTRES

- Tabla HECHOS

La tabla HECHOS almacena únicamente las métricas (calificaciones, promedios, etc) y como habíamos mencionado antes, las llevas primarias de las dimensiones, las cuales fungen como llave primaria para esta tabla. A continuación se presenta el código SQL que muestra las columnas que conforman la tabla HECHOS, además del esquema al que pertenece dicha tabla, tipo de dato, si permite nulos y algunos comentarios como se aprecia en la figura 21.

```

select
  b.column_id,
  a.*,
  b.nullable,
  b.data_type
from ALL_COL_COMMENTS a,
  ALL_TAB_COLUMNS b
where a.table_name = b.table_name and
  a.column_name = b.column_name and
  a.owner = 'TESIS' AND
  A.TABLE_NAME = 'HECHOS'
order by b.table_name, b.column_id

```

id	COLUMN_ID	OWNER	TABLE_NAME	COLUMN_NAME	COMMENTS	NULLABLE	DATA_TYPE
▶	1	TESIS	HECHOS	PROM_PROFESORES	Promedio de la calificación del profesor	Y	NUMBER
	2	TESIS	HECHOS	PROM_ALUMNOS	Promedio de la calificación del alumno	Y	NUMBER
	3	TESIS	HECHOS	PREG_1	Pregunta 1	Y	NUMBER
	4	TESIS	HECHOS	PREG_2	Pregunta 2	Y	NUMBER
	5	TESIS	HECHOS	PREG_3	Pregunta 3	Y	NUMBER
	6	TESIS	HECHOS	PREG_4	Pregunta 4	Y	NUMBER
	7	TESIS	HECHOS	PREG_5	Pregunta 5	Y	NUMBER
	8	TESIS	HECHOS	PREG_6	Pregunta 6	Y	NUMBER
	9	TESIS	HECHOS	PREG_7	Pregunta 7	Y	NUMBER
	10	TESIS	HECHOS	PREG_8	Pregunta 8	Y	NUMBER
	11	TESIS	HECHOS	PREG_9	Pregunta 9	Y	NUMBER
	12	TESIS	HECHOS	PREG_10	Pregunta 10	Y	NUMBER
	13	TESIS	HECHOS	PREG_11	Pregunta 11	Y	NUMBER
	14	TESIS	HECHOS	PREG_12	Pregunta 12	Y	NUMBER
	15	TESIS	HECHOS	PREG_13	Pregunta 13	Y	NUMBER
	16	TESIS	HECHOS	PREG_14	Pregunta 14	Y	NUMBER
	17	TESIS	HECHOS	PREG_15	Pregunta 15	Y	NUMBER
	18	TESIS	HECHOS	PREG_16	Pregunta 16	Y	NUMBER
	19	TESIS	HECHOS	PREG_17	Pregunta 17	Y	NUMBER
	20	TESIS	HECHOS	PREG_18	Pregunta 18	Y	NUMBER
	21	TESIS	HECHOS	PREG_19	Pregunta 19	Y	NUMBER
	22	TESIS	HECHOS	PREG_20	Pregunta 20	Y	NUMBER
	23	TESIS	HECHOS	PREG_21	Pregunta 21	Y	NUMBER
	24	TESIS	HECHOS	PREG_22	Pregunta 22	Y	NUMBER
	25	TESIS	HECHOS	PREG_23	Pregunta 23	Y	NUMBER
	26	TESIS	HECHOS	PREG_24	Pregunta 24	Y	NUMBER
	27	TESIS	HECHOS	PREG_25	Pregunta 25	Y	NUMBER
	28	TESIS	HECHOS	PREG_26	Pregunta 26	Y	NUMBER
	29	TESIS	HECHOS	PREG_27	Pregunta 27	Y	NUMBER
	30	TESIS	HECHOS	PREG_28	Pregunta 28	Y	NUMBER
	31	TESIS	HECHOS	PREG_29	Pregunta 29	Y	NUMBER
	32	TESIS	HECHOS	CC_1	Calificación concepto 1	Y	NUMBER
	33	TESIS	HECHOS	CC_2	Calificación concepto 2	Y	NUMBER
	34	TESIS	HECHOS	CC_3	Calificación concepto 3	Y	NUMBER
	35	TESIS	HECHOS	CC_4	Calificación concepto 4	Y	NUMBER
	36	TESIS	HECHOS	CC_5	Calificación concepto 5	Y	NUMBER
	37	TESIS	HECHOS	CC_6	Calificación concepto 6	Y	NUMBER
	38	TESIS	HECHOS	CC_7	Calificación concepto 7	Y	NUMBER
	39	TESIS	HECHOS	GPO	Grupo	Y	NUMBER
	40	TESIS	HECHOS	ID_PROFESOR	Identificador profesor. Primary key	N	NUMBER
	41	TESIS	HECHOS	CLAVE_ASIGNATURA	Identificador asignatura. Primary key	N	NUMBER
	42	TESIS	HECHOS	CLAVE_DIVISION	Identificador division. Primary key	N	NUMBER
	43	TESIS	HECHOS	CLAVE_DEPTO	Identificador departamento. Primary key	N	NUMBER
▶	44	TESIS	HECHOS	ID_SEMESTRES	Identificador semestre. Primary key	N	NUMBER

Figura 21 Descripción tabla HECHOS

4.4. Implementación del proceso ETL

Ahora que las tablas de nuestro proyecto están consideradas en el modelo estrella es necesario definir los diferentes procesos que permitan la consolidación de los datos en ellas.

Para nuestro proyecto y por medio de la herramienta IBM Data Stage 7.5.2 se generará una tarea para cada una de las tablas involucradas en el modelo estrella previamente planteado, además de una secuencia global que ejecuta de forma automática todos y cada uno de las diferentes tareas involucradas en este proyecto, con el objetivo de ser calendarizado y ejecutado cada semestre y permita la actualización de la información. En la figura 22 se puede apreciar el flujo de la secuencia global, es importante mencionar que el orden de los objetos en dicha secuencia es el orden de ejecución de cada uno de ellos.

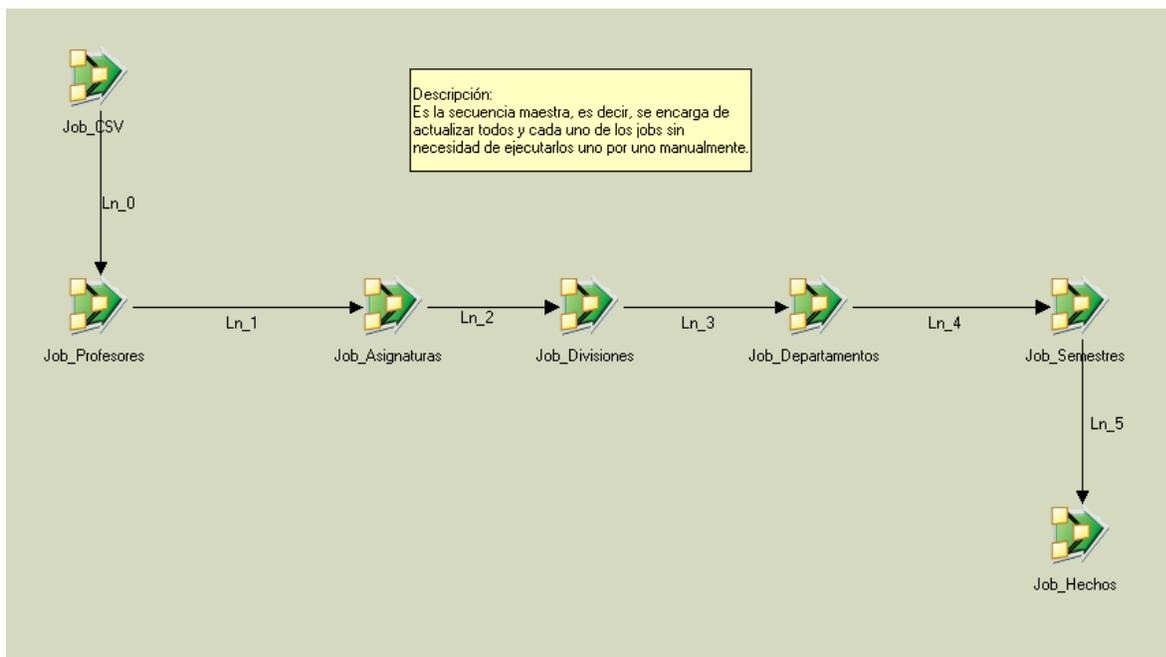


Figura 22 Secuencia_Carga_Total

- Actualización de la fuente de datos proporcionada por la FI de la UNAM
- Actualización de la tabla DIVISIONES
- Actualización de la tabla DEPARTAMENTOS
- Actualización de la tabla PROFESORES
- Actualización de la tabla ASIGNATURAS
- Actualización de la tabla SEMESTRES
- Actualización de la tabla HECHOS

4.4.1. Tarea para la extracción de los datos proporcionados por la Facultad de Ingeniería de la UNAM

El primer paso es definir nuestra fuente de datos y elaborar una tarea que nos permita almacenarlos en nuestra base de datos.

Los datos proporcionados por la Facultad de Ingeniería de la UNAM correspondientes a las encuestas a profesores están contenidos dentro de un archivo Excel, por lo que podemos definir que nuestro origen de datos, en primera instancia, será este documento Excel, sin embargo, es necesario almacenar esta información dentro de nuestra base de datos, para ello se creará una tabla que se encargará de guardar la información exactamente igual a como viene de nuestro archivo Excel, este proceso se realizará por medio de una tarea simple en la interfaz de diseño de IBM Data Stage 7.5.2. La tabla que almacenará esta información tendrá el nombre de TB_ENCUESTAS_PROFESORES.

Como podemos observar en la figura 23 la tarea Carga_CSV_A_BD consiste únicamente en localizar el archivo Excel y cargarlo directamente a la tabla TB_ENCUESTAS_PROFESORES. El archivo Excel está ubicado en la ruta: C:\Ascential\DataStage\Projects\TESIS\BASES DE DATOS.csv, donde se encuentra instalado el servidor de IBM Data Stage 7.5.2, mientras que la tabla TB_ENCUESTAS_PROFESORES se encuentra dentro de nuestra base de datos.

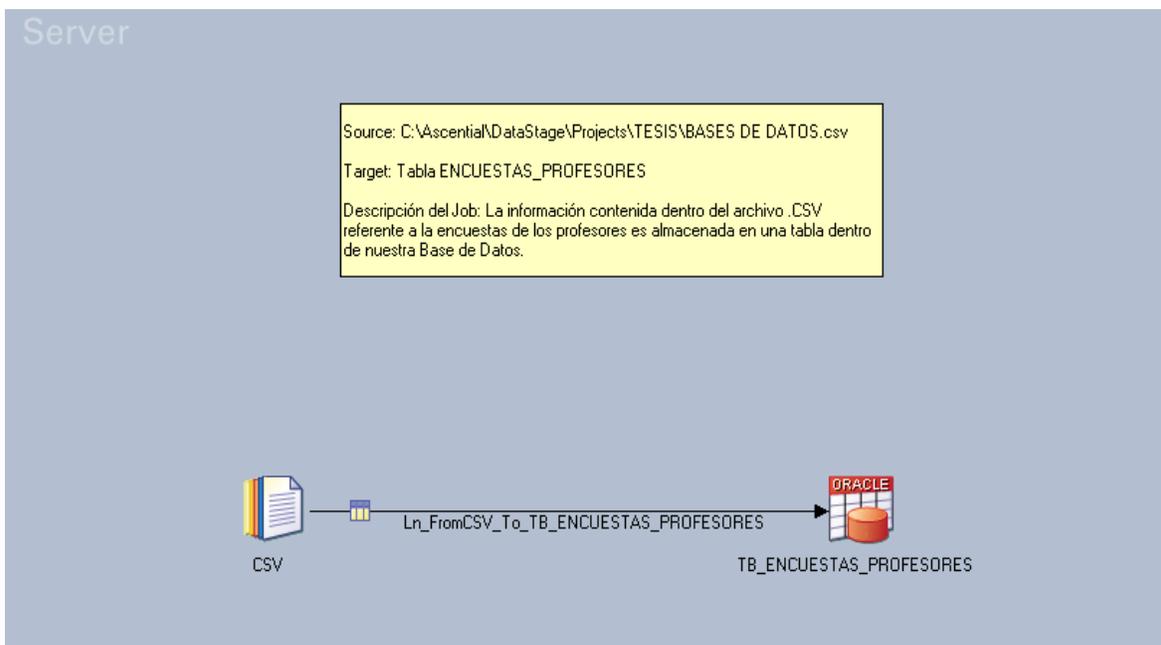


Figura 23 Tarea Carga_CSV_A_BD

Es importante mencionar que el archivo Excel debe seguir la siguiente estructura de columnas:

Numero de columna	Nombre de columna
1	SEC
2	DIV
3	DEP
4	ASIG
5	GPO
6	TIPO
7	NP
8	NOMBRE_ASIGNATURA
9	ID_PROFESOR
10	NOMBRE_PROF
11	INSCRITOS
12	ENCUESTADOS
13	RECEPCION
14	CAL_PROF
15	CAL_ALUM
16	SEMESTRE
17	P01
18	P02
19	P03
20	P04
21	P05
22	P06
23	P07
24	P08
25	P09
26	P10
27	P11
28	P12
29	P13
30	P14
31	P15
32	P16
33	P17
34	P18
35	P19
36	P20
37	P21
38	P22
39	P23

40	P24
41	P25
42	P26
43	P27
44	P28
45	P29
46	CC1
47	CC2
48	CC3
49	CC4
50	CC5
51	CC6
52	CC7
53	CALIFICACION
54	NOM_DIV
55	NOM_DEP

Tabla 1 Estructura archivo fuente

En caso que el archivo Excel no siga con la estructura descrita en la Tabla 1, la tarea al ser ejecutada, abortará y no hará la carga de los datos a la tabla temporal TB_ENCUESTAS_PROFESORES.

4.4.2. Tarea para la actualización de la tabla DIVISIONES

Una vez almacenada la información de las encuestas en nuestra base de datos lo que prosigue es cargar, basados en la tabla ENCUESTAS_PROFESORES, las diferentes dimensiones y la tabla de hechos con la información concerniente a cada una de ellas.

Para hacer la carga de la información a nuestra tabla DIVISIONES se creó la tarea Carga_DIVISIONES, la cual selecciona la clave de cada división con su descripción, en este caso el nombre de la división, y las almacena en la tabla DIVISIONES.

Como podemos apreciar en la figura 24 la tarea Carga_DIVISIONES consta de 3 etapas:

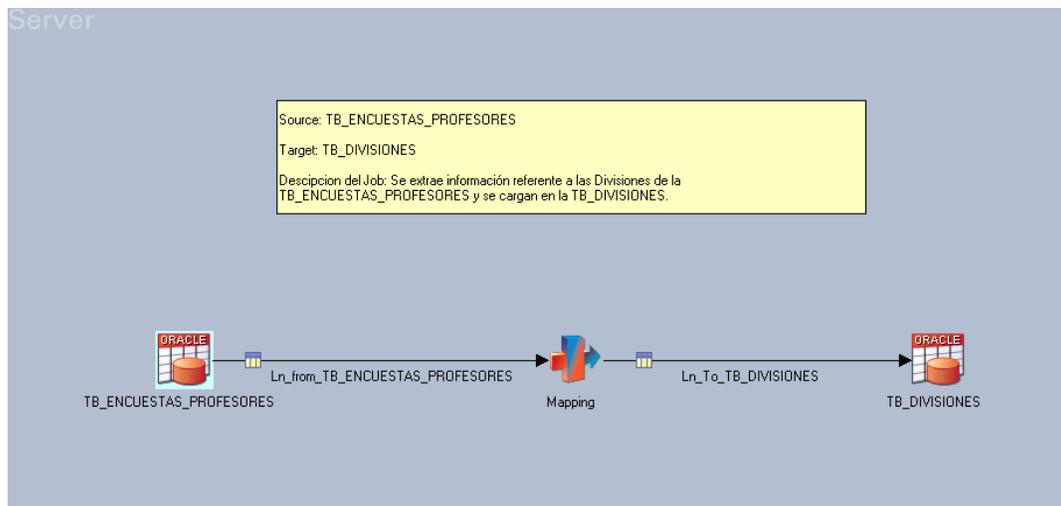


Figura 24 Tarea Carga_DIVISIONES

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta descrita en la figura 25, las columnas de interés para nuestra tabla destino DIVISIONES.

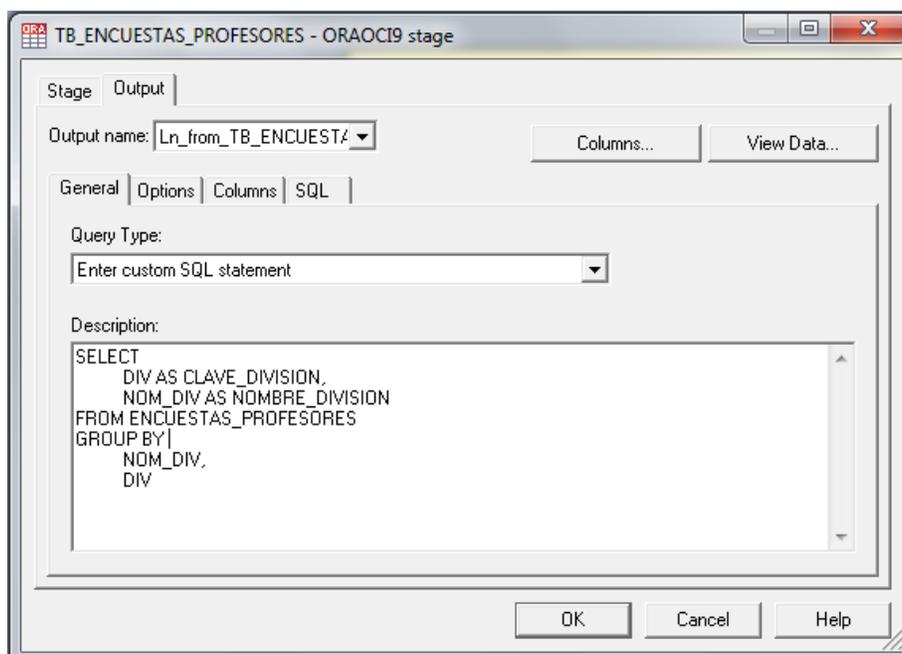


Figura 25 Objeto Base de Datos Origen para la tarea Carga_DIVISIONES

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_DIVISIONES. En la figura 26 se muestra el mapeo establecido.

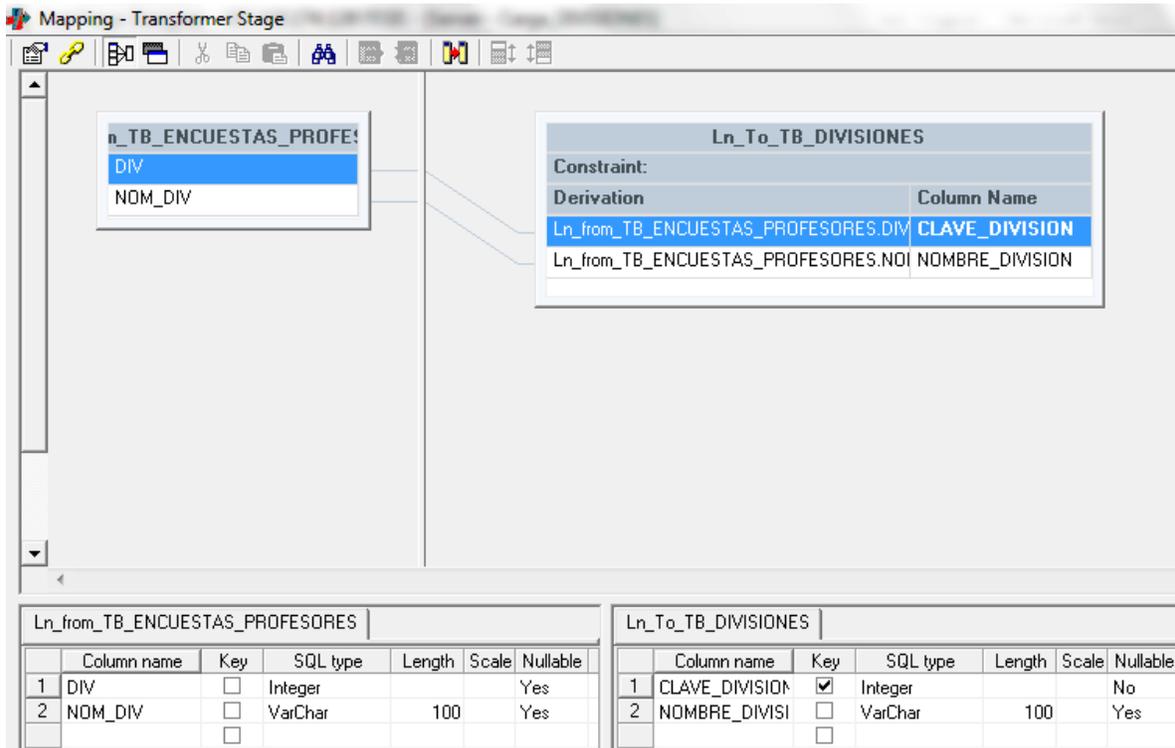


Figura 26 Objeto transformación para la tarea Carga_DIVISIONES

- TB_DIVISIONES

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 27, se definen los parámetros para la conexión a nuestra base de datos para la tabla DIVISIONES, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

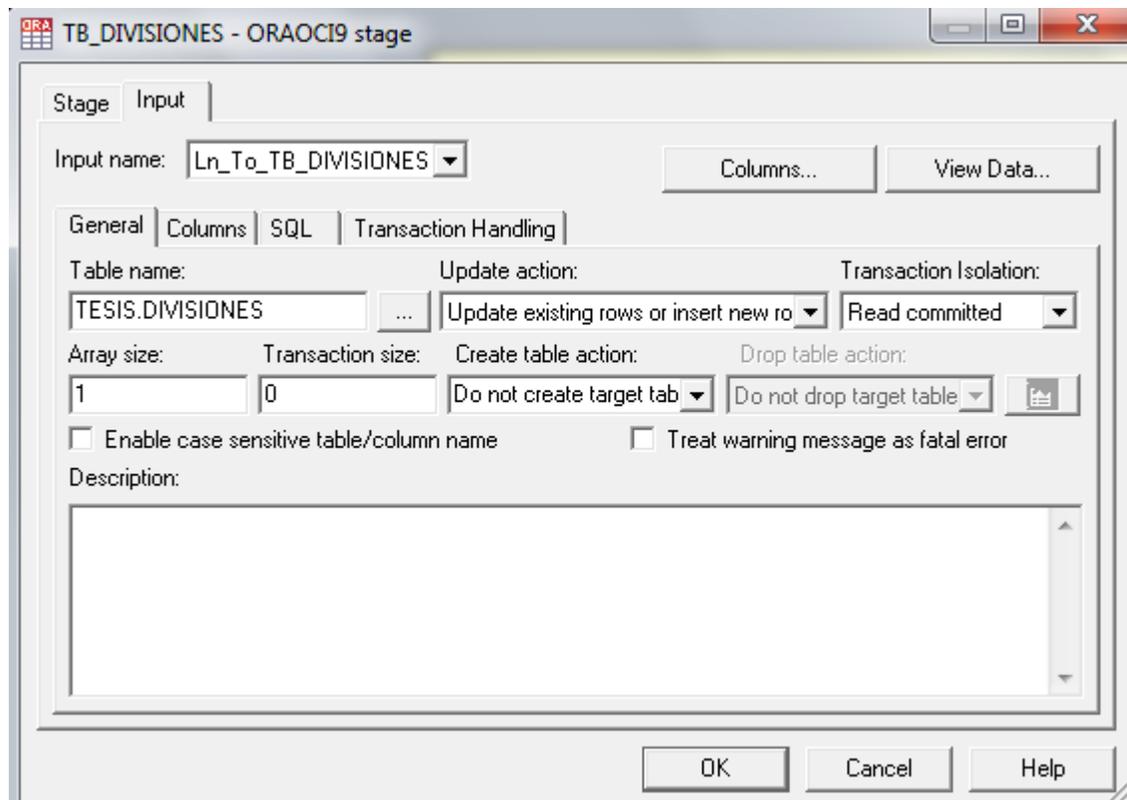


Figura 27 Objeto Base de Datos destino para la tarea Carga_DIVISIONES

4.4.3. Tarea para la actualización de la tabla DEPARTAMENTOS

Para hacer la carga de la información a nuestra tabla DEPARTAMENTOS se creó la tarea Carga_DEPARTAMENTOS, la cual selecciona la clave de cada departamento con su descripción, en este caso el nombre del departamento, y las almacena en la tabla DEPARTAMENTOS.

Como podemos apreciar en la figura 28 la tarea Carga_DEPARTAMENTOS consta de 3 etapas:

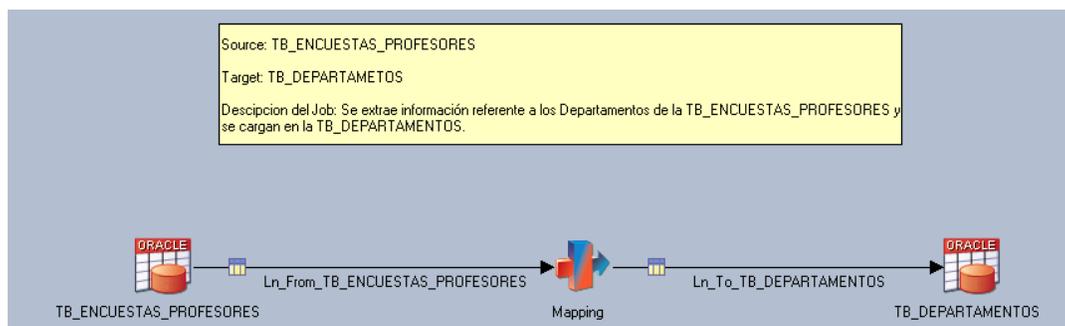


Figura 28 Tarea Carga_DEPARTAMENTOS

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta descrita en la figura 29, las columnas de interés para nuestra tabla destino DEPARTAMENTOS.

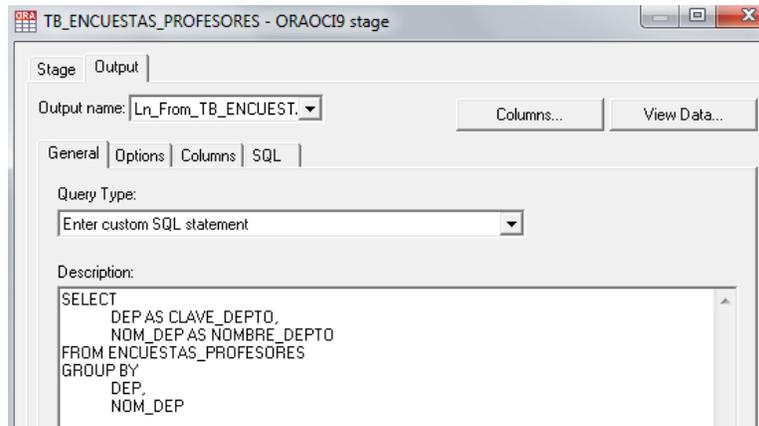


Figura 29 Objeto de Base de Datos Origen para la tarea Carga_DEPARTAMENTOS

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_DEPARTAMENTOS. En la figura 30 se muestra el mapeo establecido.

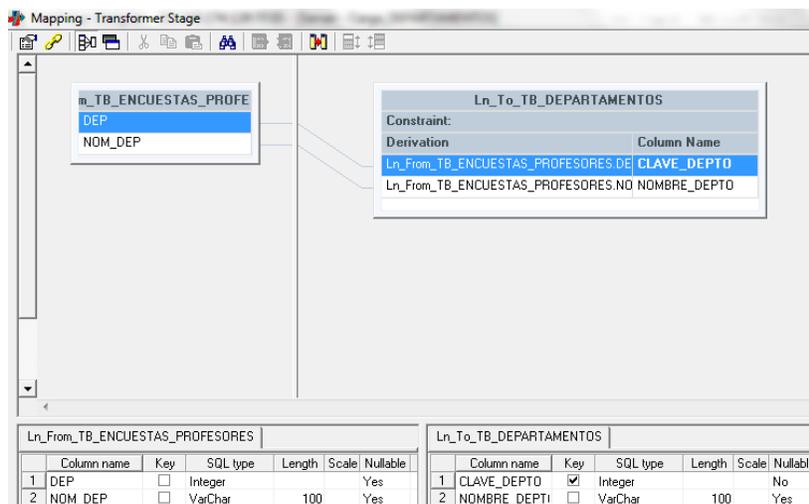


Figura 30 Objeto transformación para la tarea Carga_DEPARTAMENTOS

- TB_DEPARTAMENTOS

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 31, se definen los parámetros para la conexión a nuestra base de datos para la tabla DEPARTAMENTOS, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

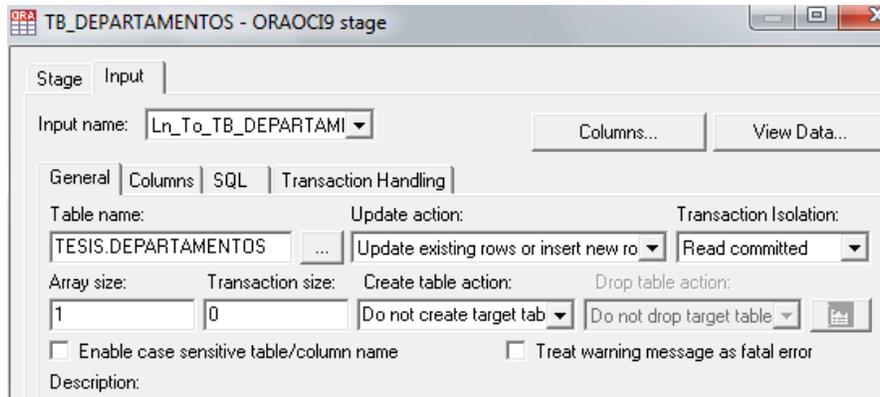


Figura 31 Objeto de Base de Datos destino para la tarea Carga_DEPARTAMENTOS

4.4.4. Tarea para la actualización de la tabla PROFESORES

Para hacer la carga de la información a nuestra tabla PROFESORES se creó la tarea Carga_PROFESORES, la cual selecciona únicamente el identificador de cada profesor con su descripción, en este caso el nombre del profesor, y las almacena en la tabla PROFESORES.

Como podemos apreciar en la figura 32 la tarea Carga_PROFESORES consta de 3 etapas:

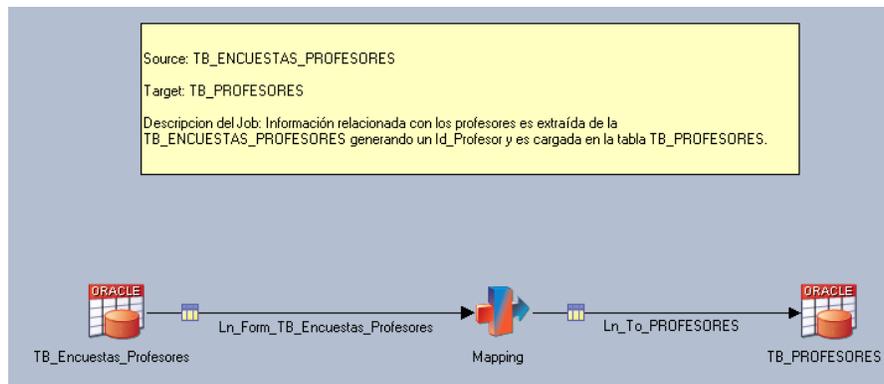


Figura 32 Tarea Carga_PROFESORES

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta mostrada en la figura 33, las columnas de interés para nuestra tabla destino PROFESORES.

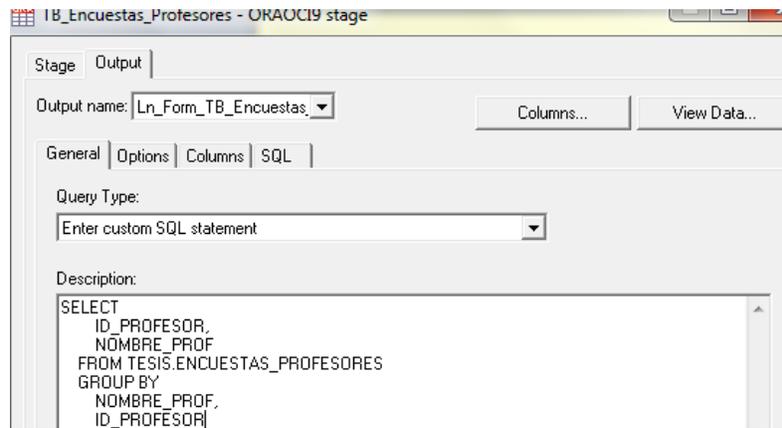


Figura 33 Objeto de Base de Datos origen para la tarea Carga_DEPARTAMENTOS

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_PROFESORES. En la figura 34 se muestra el mapeo establecido.

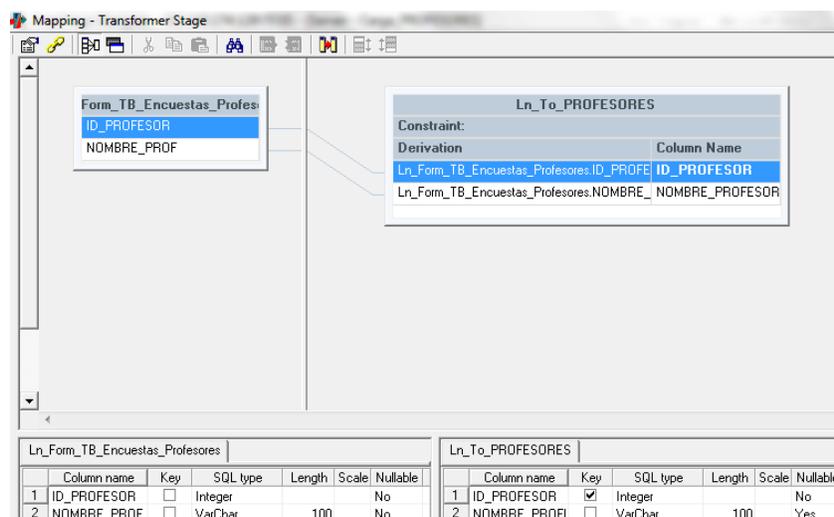


Figura 34 Objeto transformación para la tarea Carga_PROFESORES

- TB_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 35, se definen los parámetros para la conexión a nuestra base de datos para la tabla PROFESORES, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

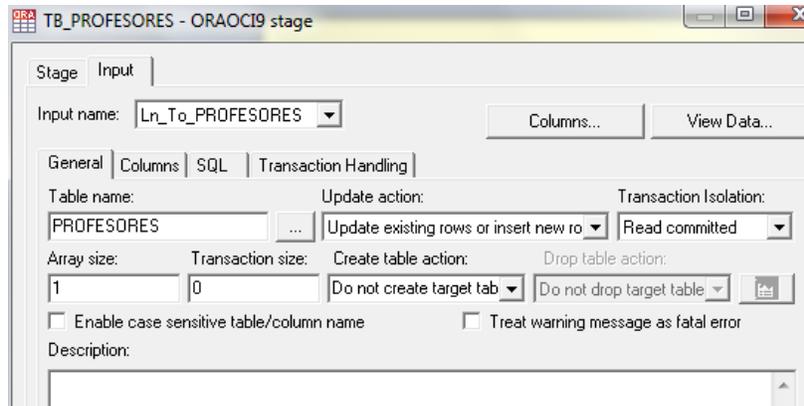


Figura 35 Objeto de Base de Datos destino para la tarea Carga_PROFESORES

4.4.5. Tarea para la actualización de la tabla ASIGNATURAS

Para hacer la carga de la información a nuestra tabla ASIGNATURAS se creó la tarea Carga_ASIGNATURAS, la cual selecciona la clave de la asignatura con su descripción, en este caso el nombre del profesor y el tipo de asignatura, y las almacena en la tabla ASIGNATURAS.

Como podemos apreciar en la figura 36 la tarea Carga_ASIGNATURAS consta de 3 etapas:

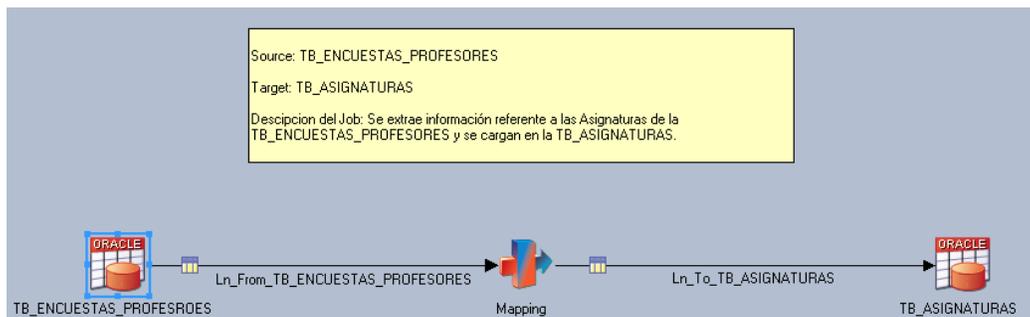


Figura 36 Tarea Carga_ASIGNATURAS

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta mostrada en la figura 37, las columnas de interés para nuestra tabla destino ASIGNATURAS.

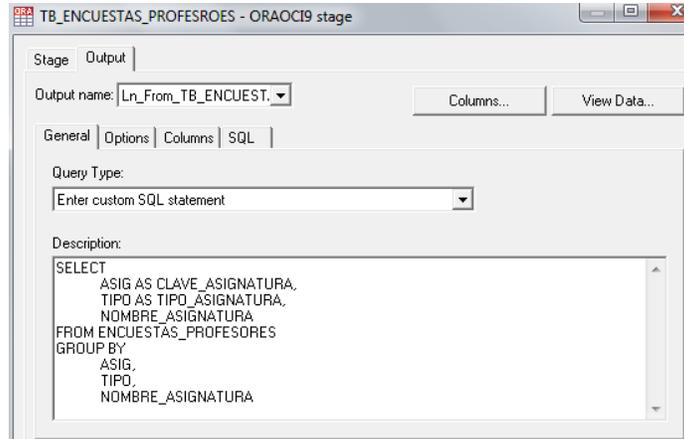


Figura 37 Objeto de Base de Datos origen para la tarea Carga_ASIGNATURAS

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_ASIGNATURAS. La figura 38 muestra el mapeo establecido.

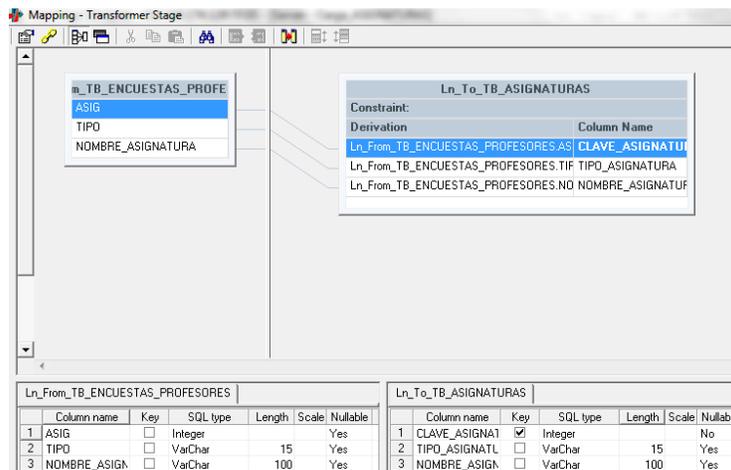


Figura 38 Objeto transformación para la tarea Carga_ASIGNATURAS

- TB_ASIGNATURAS

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 39, se definen los parámetros para la conexión a nuestra base de datos para la tabla PROFESORES, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

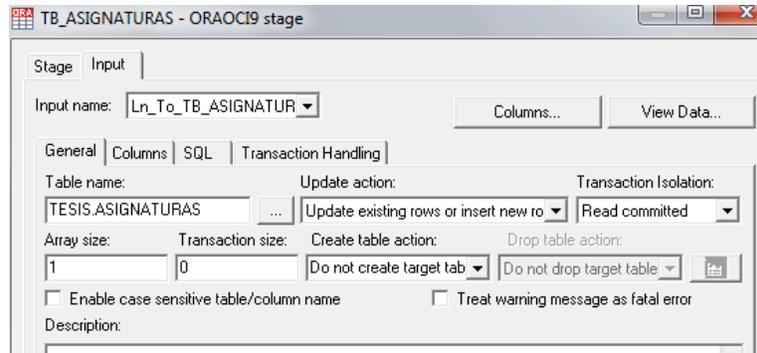


Figura 39 Objeto de Base de Datos destino para la tarea Carga_ASIGNATURAS

4.4.6. Tarea para la actualización de la tabla SEMESTRES

Para hacer la carga de la información a nuestra tabla SEMESTRES se creó la tarea Carga_SEMESTRES, la cual selecciona el identificador para cada semestre y su descripción, en este caso el año y el número de semestre, y las almacena en la tabla SEMESTRES.

Como podemos apreciar en la figura 40 la tarea Carga_SEMESTRES consta de 3 etapas:

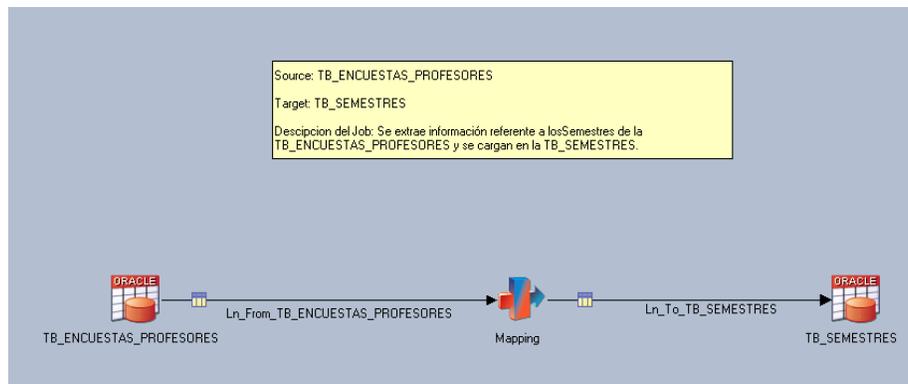


Figura 40 Tarea Carga_SEMESTRES

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta mostrada en la figura 41, las columnas de interés para nuestra tabla destino SEMESTRES.

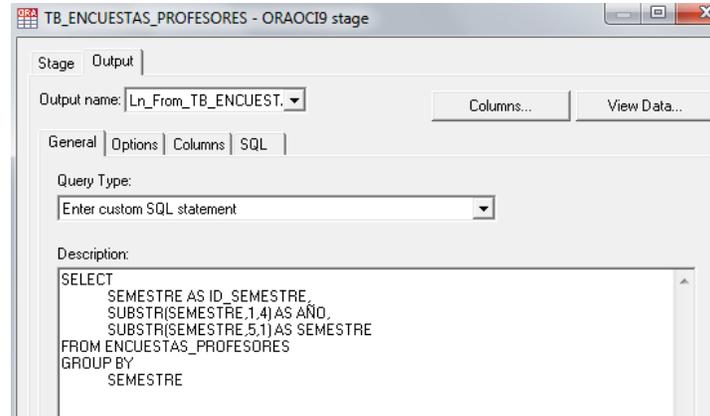


Figura 41 Objeto de Base de Datos origen para la tarea Carga_SEMESTRES

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_SEMESTRES. En la figura 42 se muestra el mapeo establecido.

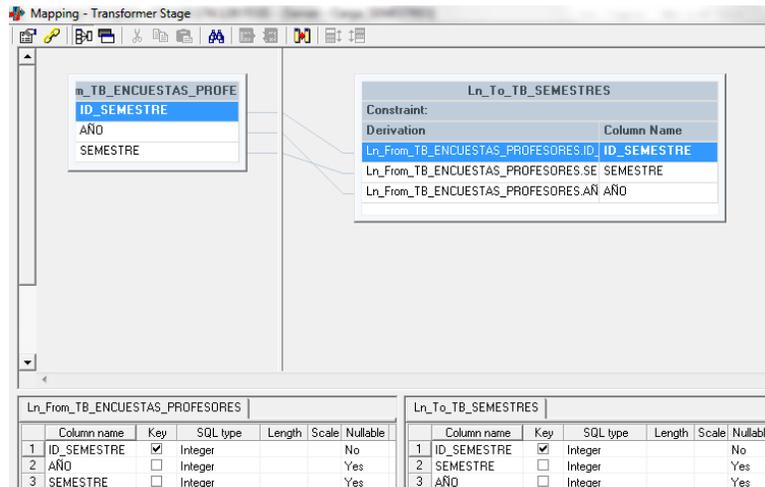


Figura 42 Objeto transformación para la tarea Carga_SEMESTRES

- TB_SEMESTRES

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 43, se definen los parámetros para la conexión a nuestra base de datos para la tabla SEMESTRES, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

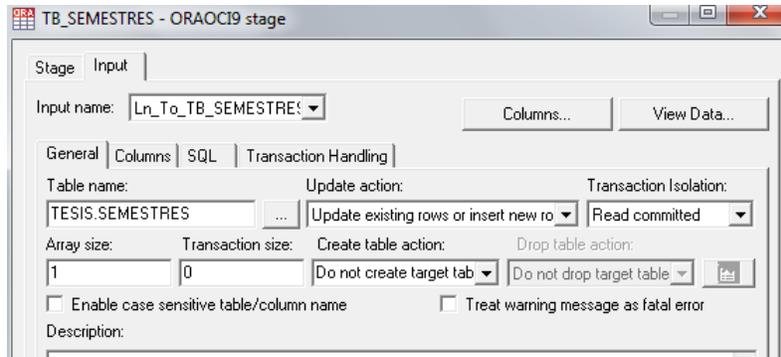


Figura 43 Objeto de Base de Datos destino para la tarea Carga_SEMESTRES

4.4.7. Tarea para la actualización de la tabla HECHOS

Para hacer la carga de la información a nuestra tabla HECHOS se creó la tarea Carga_HECHOS, la cual selecciona todas las métricas de interés necesarias para el análisis de la información, así como de todas y cada una de los identificadores de las tablas dimensiones.

Como podemos apreciar en la figura 44 la tarea Carga_HECHOS consta de 3 etapas:

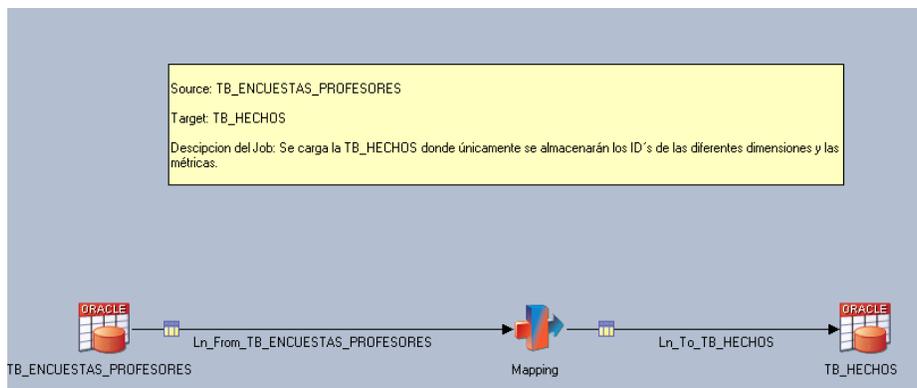


Figura 44 Tarea Carga_HECHOS

- TB_ENCUESTAS_PROFESORES

Objeto tipo Base de Datos ORACLE que funge como nuestro origen de datos para esta tarea, en este objeto se definen los parámetros para la conexión a nuestra base de datos para la tabla ENCUESTAS_PROFESORES, además se especifican, por medio de una consulta mostrada en la figura 45, las columnas de interés para nuestra tabla destino HECHOS.

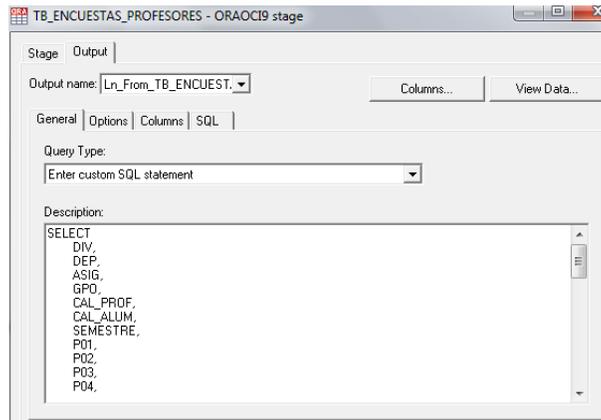


Figura 45 Objeto de Base de Datos origen para la tarea Carga_HECHOS

- Mapeo

Objeto tipo transformación que nos permite hacer el mapeo y la homologación de los tipos de datos entre nuestro origen de datos TB_ENCUESTAS_PROFESORES y nuestro destino TB_HECHOS. La figura 46 muestra el mapeo establecido.

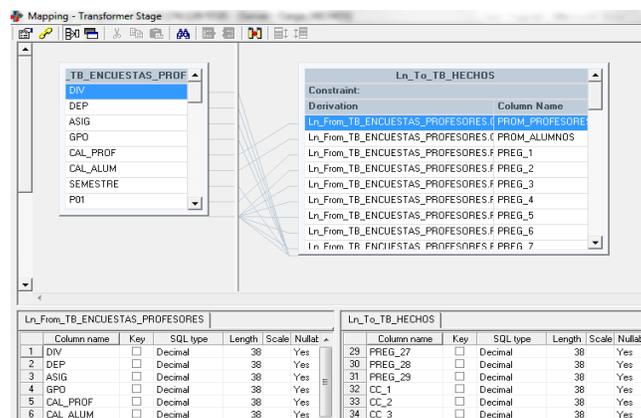


Figura 46 Objeto transformación para la tarea Carga_HECHOS

- TB_HECHOS

Objeto tipo Base de Datos ORACLE que funge como nuestro destino de datos para esta tarea, en este objeto, como se muestra en la figura 47, se definen los parámetros para la conexión a nuestra base de datos para la tabla HECHOS, así como el tipo de carga de datos a realizar, en nuestro caso y basados en la llave primaria de dicha tabla, el tipo de carga de datos consiste en primero actualizar datos existentes y después insertar nuevos datos.

Es importante mencionar que para la creación de la llave primaria para la tabla HECHOS, además de considerar los identificadores de las dimensiones, se añadió la columna del grupo como parte de dicha llave primaria, esto debido a que un profesor puede dar uno o más grupos para una asignatura en un determinado semestre, de esta forma se garantiza que el registro es único dentro de la tabla HECHOS.

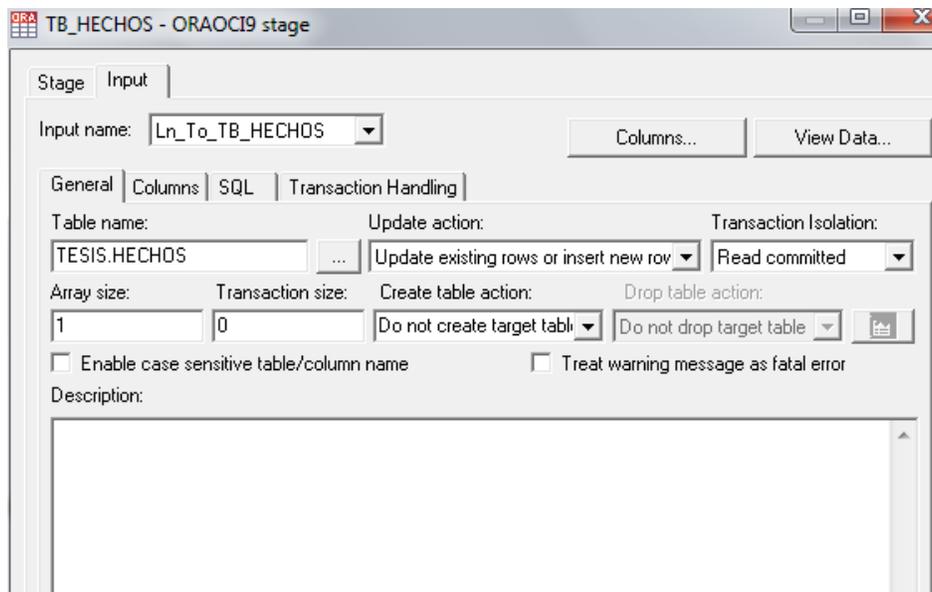


Figura 47 Objeto de Base de Datos destino para la tarea Carga_HECHOS

4.5. Implementación de un modelo con IBM Cognos Framework Manager 10.1.1

El proceso de modelar la información dentro de IBM Cognos 10.1.1 también conlleva una serie de pasos que tienen como objetivo principal evitar que la información mostrada sea errónea e inconsistente, además de facilitar la administración de la misma. A continuación los pasos sugeridos para la realización de este proyecto:

4.5.1. Elaboración de un modelo de metadatos

Una de las mejores prácticas en el modelado de metadatos dentro de IBM Cognos Framework Manager 10.1.1 es la generación de 3 capas que permitan mantener un orden y un control sobre todos los objetos de un proyecto. [10]

En la figura 48 se pueden apreciar las 3 capas utilizadas para nuestro:

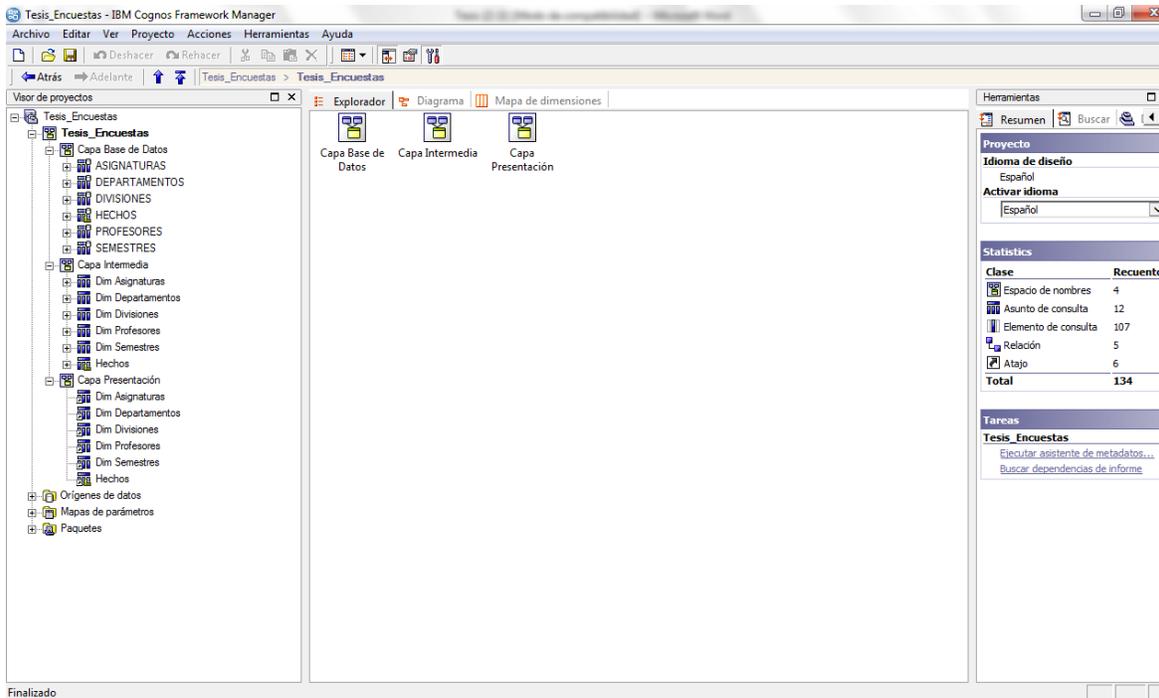


Figura 48 Organización de metadatos en IBM Cognos Framework Manager 10.1.1

- **Capa de BD:**

En esta capa están alojados los metadatos exactamente igual a como se encuentran en la Base de Datos. Aquí se generan las relaciones entre los objetos involucrados en el modelo.

- **Capa intermedia:**

En esta capa es donde son realizados todos los cálculos y filtros, además de la organización de todos los objetos involucrados en el modelo de metadatos.

- **Capa de Presentación:**

En esta capa son definidos los objetos que serán presentados o visualizados dentro del paquete de metadatos.

Los metadatos en IBM Cognos Framework Manager 10.1.1 son organizados como jerarquías, como se puede ver en la Figura 48.

El nivel más bajo de la jerarquía se denomina “Elemento de Consulta”, este objeto representa una columna en una tabla. Ej. Nombre Asignatura

El siguiente nivel de la jerarquía se denomina “Objeto de Consulta”, este objeto es el encargado de agrupar 1 o más “Elementos de consulta”. Ej. Asignaturas contiene Nombre Asignatura, Tipo asignatura y Clave Asignatura.

Por último, el nivel más alto de la jerarquía se denomina “Espacio de nombre”, este objeto es el encargado de agrupar 1 o más “Objetos de Consulta”. Ej. Capa Intermedia contiene Dim Asignaturas, Dim Profesores, Dim Semestres, Dim Divisiones, Dim Departamentos y Hechos.

A continuación se detallan las relaciones establecidas dentro de nuestro modelo de metadatos:

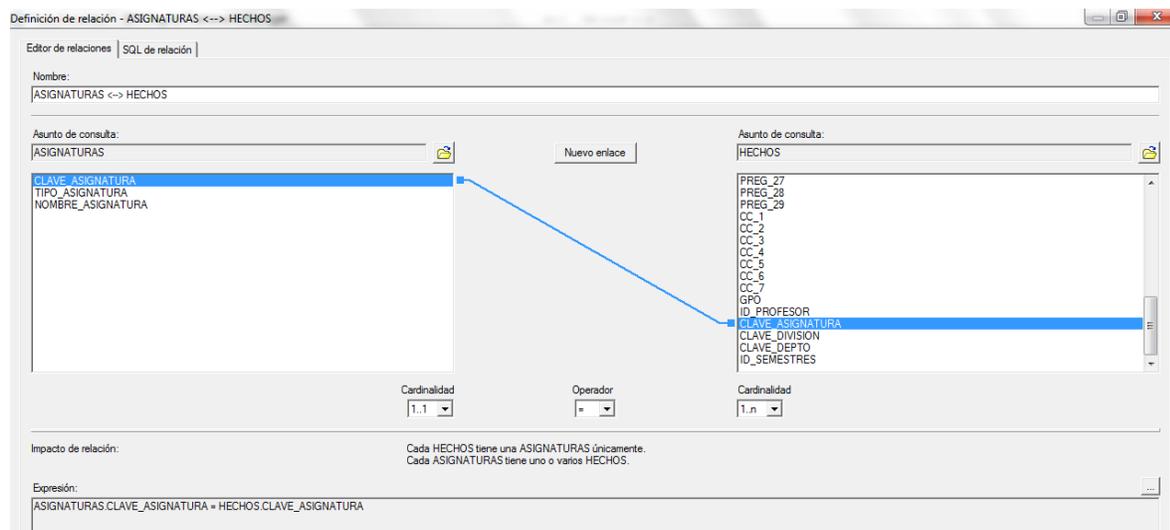


Figura 49 Relación HECHOS-ASIGNATURAS

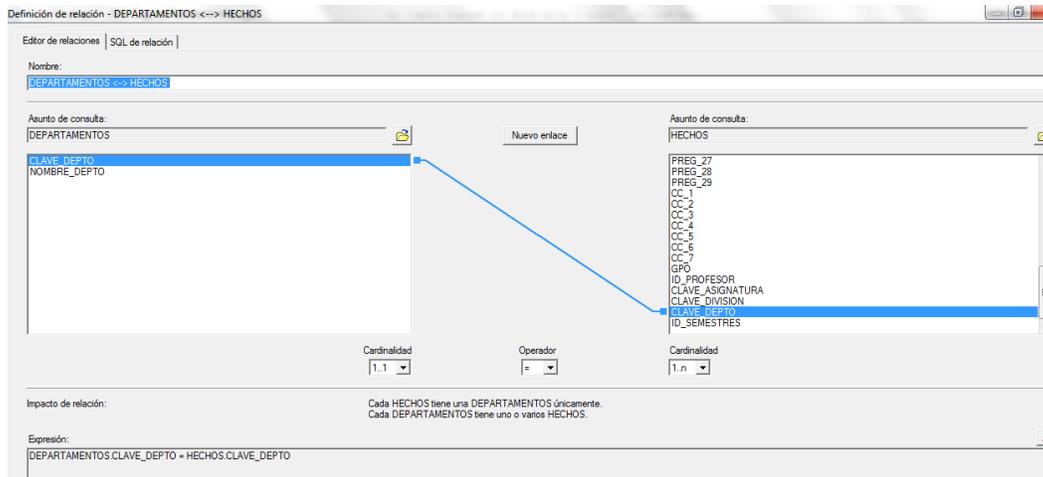


Figura 50 Relación HECHOS-DEPARTAMENTOS

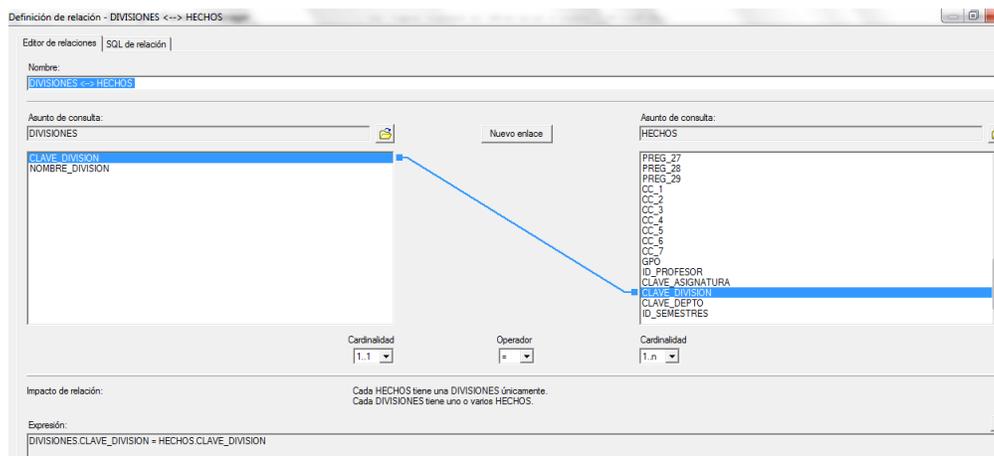


Figura 51 Relación HECHOS-DIVISIONES

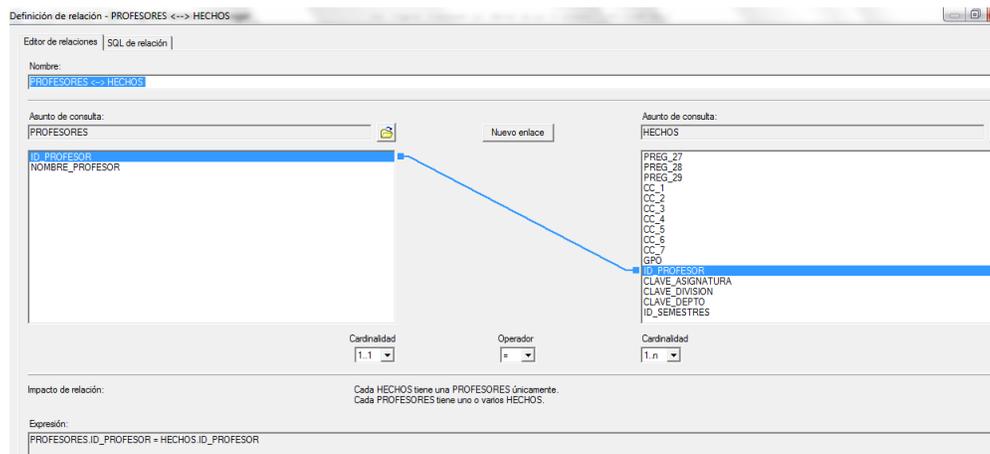


Figura 52 Relación HECHOS-PROFESORES

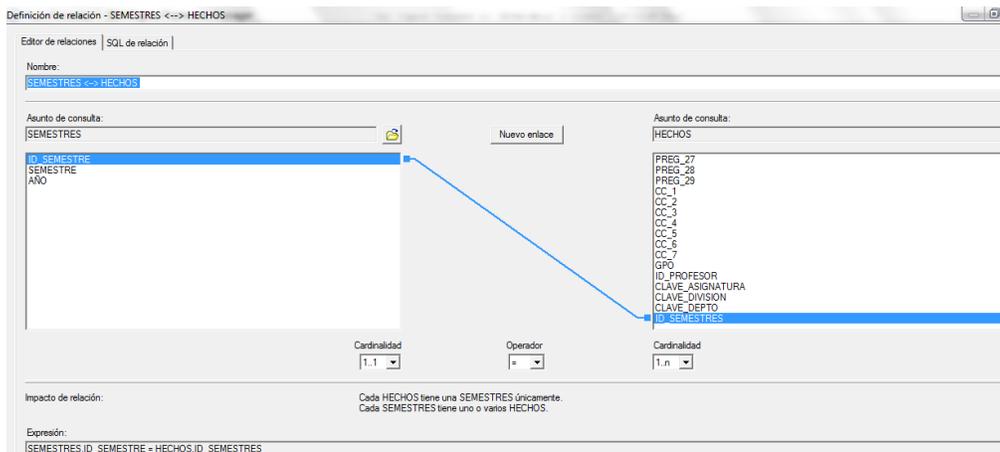


Figura 53 Relación HECHOS-SEMESTRES

4.5.2. Generación y publicación de un paquete de metadatos

Como mencionamos anteriormente la mayor parte de las herramientas de la suite de IBM Cognos 10.1.1 son ejecutadas y administradas por el portal de IBM Cognos Connection 10.1.1, por lo tanto, nuestro modelo de metadatos creado dentro de IBM Cognos Framework Manager 10.1.1 tendrá que ser generado y posteriormente publicado dentro del portal, de esta forma las herramientas para el análisis de la información podrán hacer uso de dicho modelo de metadatos. Es importante mencionar la diferencia entre un modelo de metadatos y un nuevo concepto, un paquete de metadatos. Un modelo de metadatos es un diseño lógico de todos y cada uno de los objetos involucrados en el mismo, mientras que un paquete de metadatos representa todos los elementos que serán visualizados dentro de nuestro portal de IBM Cognos Connection 10.1.1. Los paquetes de metadatos son generados a partir de un modelo de metadatos.

En el caso de nuestro proyecto únicamente la “Capa de Presentación” será incluida dentro de nuestro paquete de metadatos, pues esta capa contiene los elementos necesarios para la generación de los reportes analíticos. La figura 54 muestra los elementos incluidos dentro del paquete a publicar en IBM Cognos Connection 10.1.1.

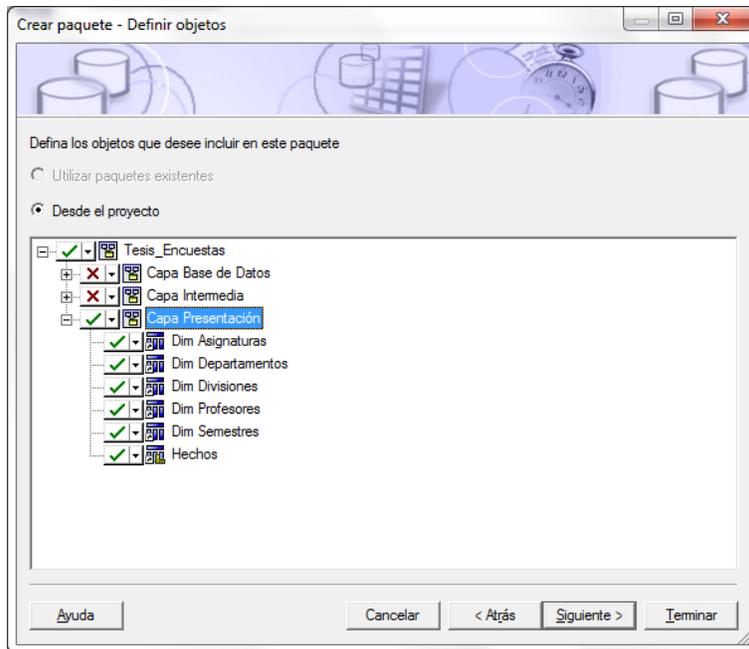


Figura 54 Generación del paquete de metadatos dentro de IBM Cognos Framework Manager 10.1.1

Una vez generado el paquete de metadatos lo que necesitamos es publicarlo dentro de nuestro portal de IBM Cognos Connection 10.1.1, con esto nuestro modelo estará disponible para las herramientas de la suite de IBM Cognos 10.1.1. La figura 55 nos muestra el paquete publicado dentro del portal de IBM Cognos Connection 10.1.1.

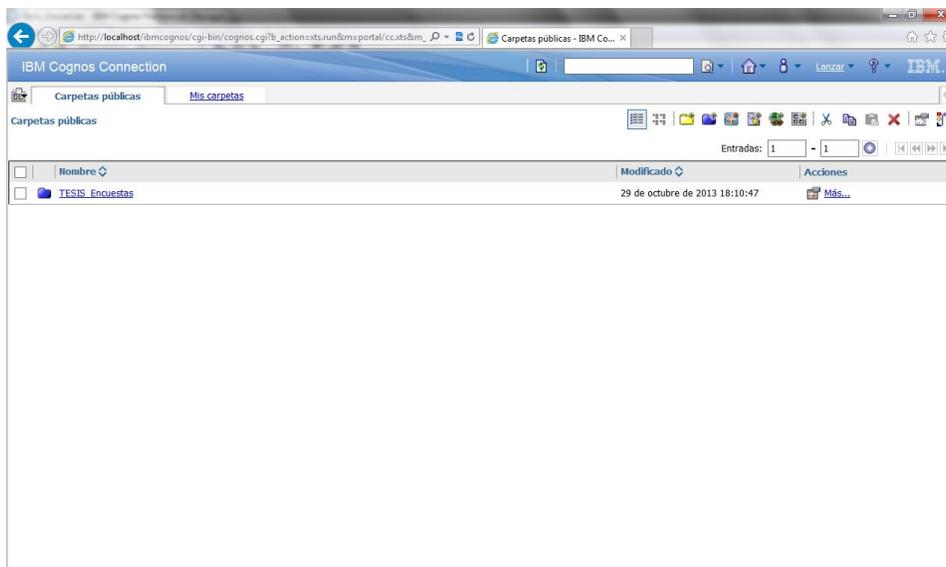


Figura 55 Paquete de metadatos TESIS_Encuestas dentro de IBM Cognos Connection 10.1.1

4.6. Consultando la información de las encuestas a profesores de la Facultad de Ingeniería de la UNAM

Uno de los objetivos de este proyecto de Inteligencia de Negocios es la creación de 2 reportes analíticos, un reporte analítico para directivos y un reporte analítico para profesores, donde se puedan hacer consultas a cerca de los resultados obtenidos de las encuestas a profesores de la Facultad de Ingeniería de la UNAM. Ambos reportes serán desarrollados por medio de IBM Cognos Report Studio 10.1.1 y almacenados dentro de nuestro portal de IBM Cognos Connection 10.1.1 para su consulta.

4.6.1. Reporte analítico para Directivos

El reporte analítico para directivos tendrá como objetivo mostrar a los usuarios un comparativo que partirá desde lo más general (Información a nivel División) hasta lo más específico (Información a nivel profesor), heredando todos y cada uno de los diferentes parámetros que en el transcurso de las diferentes consultas se vayan eligiendo, para ello se elaboraron los siguientes 5 reportes:

- 1 - Main Report – División
- 2 - Detail Report – Departamento
- 3 - Detail Report – Asignatura
- 4 - Detail Report – Profesores
- 5 - Detail Report - Detalle Profesores

Como podemos observar, cada uno de los reportes están anteceditos por un número secuencial, esta secuencia es la que se seguirá una vez que el reporte principal sea ejecutado.

4.6.1.1. Main Report – División

Este reporte (Figura 56) deberá ser el primero en ser ejecutado y en él se mostrará el primer comparativo, el cual consiste en una relación entre el promedio tanto de los profesores y alumnos de las diferentes divisiones con las que cuenta la facultad de ingeniería de la UNAM.

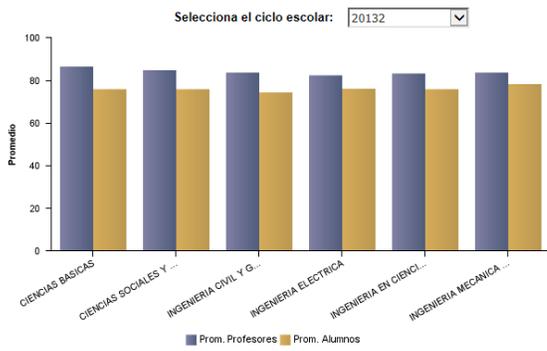


Figura 56 Main Report – División

El reporte cuenta con un parámetro para la selección del semestre (?P_Smestre?) y con un parámetro para la división (?P_Division?), el primero puede ser elegido por medio de la lista desplegable, mientras que el segundo se podrá elegir cuando el usuario presione sobre una de las barras dentro de la gráfica. Ambos parámetros serán heredados al reporte “2 - Detail Report – Departamento” filtrando la información de acuerdo a ambos criterios.

Un aspecto importante a mencionar es que de forma predeterminada el reporte, de primera instancia, filtra el último semestre cargado dentro de nuestra base de datos, sin embargo, este valor puede ser cambiado las veces que sea necesario.

4.6.1.2. Detail Report – Departamento

Este reporte (Figura 57) tiene como objetivo hacer un comparativo entre el promedio tanto de los profesores y alumnos de los diferentes departamento con los que cuenta la facultad de ingeniería de la UNAM, filtrando los parámetros heredados del “Main Report – División”, es decir, va a mostrar el semestre y únicamente los departamentos correspondientes a la división elegidos.



Universidad Nacional Autónoma de México
 Facultad de Ingeniería
 Análisis de Encuesta a Profesores (Departamento)

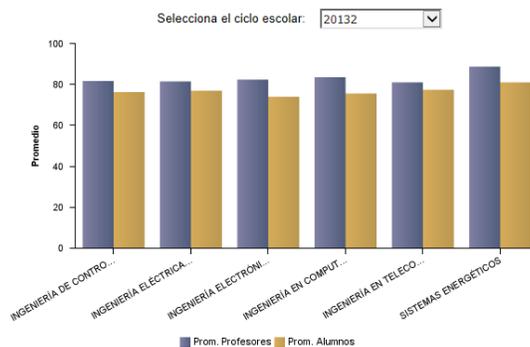


Figura 57 Detail Report – Departamento

Adicionalmente el reporte “Detail Report - Departamento” considera un nuevo parámetro (?P_Departamento?), el cual hace referencia a los departamentos que serán heredados al siguiente reporte junto con el semestre y la división. En este reporte el semestre nuevamente podrá ser seleccionado sin necesidad de regresar al reporte principal “Main Report – División” y de igual forma será heredado, asimismo el reporte finaliza la consulta una vez que el usuario elige el departamento de interés.

4.6.1.3. Detail Report – Asignatura

Este reporte (Figura 58) entra en un poco más de detalle y hace un comparativo entre el promedio tanto de los profesores como de los alumnos con las diferentes asignaturas dentro del semestre, división y departamentos seleccionados anteriormente.



Universidad Nacional Autónoma de México
 Facultad de Ingeniería
 Análisis de Encuesta a Profesores (Asignatura)

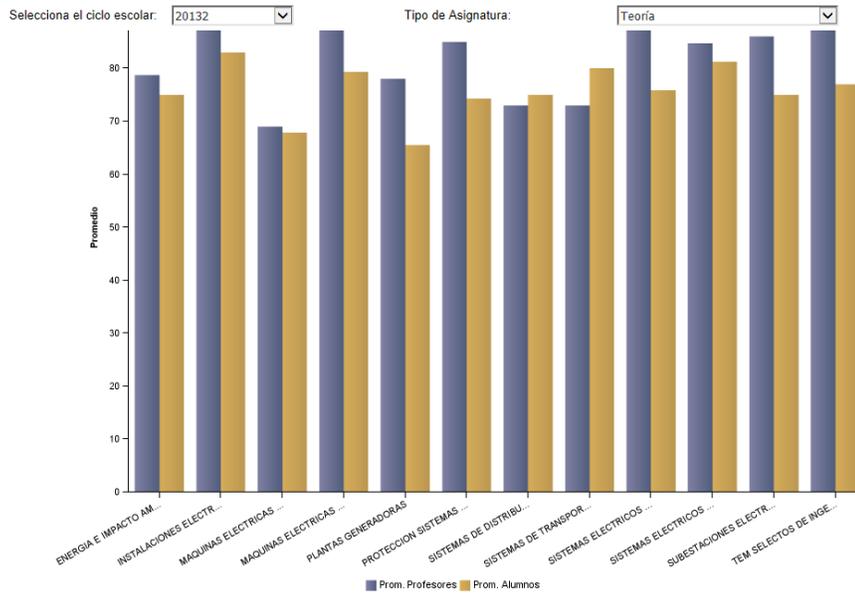


Figura 58 Detail Report – Asignatura

Adicionalmente es necesario considerar dos nuevos parámetros, uno corresponde a el nombre de la asignatura (?P_Asignatura?) y el otro corresponde a el tipo de la asignatura (?P_Tipo_Asignatura?), dichos parámetros de igual forma serán heredados al siguiente reporte, el primero eligiendo una de las asignaturas de interés en la gráfica de barras y el segundo por medio de la lista desplegable para el tipo de asignatura.

4.6.1.4. Detail Report – Profesores

En el reporte “Detail Report – Profesores” (Figura 59) se hace un comparativo específico para cada profesor de la asignatura seleccionada en el reporte “Detail Report – Asignatura”, filtrando y heredando todos los parámetros elegidos.

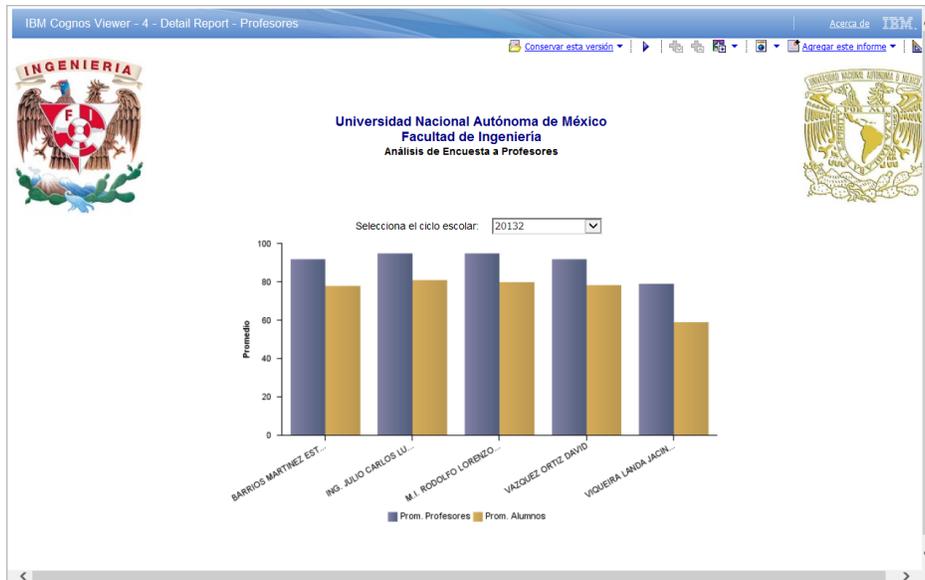


Figura 59 Detail Report – Profesores

Adicionalmente es necesario considerar un nuevo parámetro (?P_Profesor?), que corresponde al nombre del profesor y el cual será filtrado al último reporte de detalle.

4.6.1.5. Detail Report - Detalle Profesores

El reporte “Detail Report - Detalle Profesores” (Figura 60) es el reporte más detallado y muestra, para el profesor seleccionado en el reporte “Detail Report – Profesores”, la calificación que obtuvo en cada una de las preguntas contenidas dentro de la encuesta.

Pregunta 2	Pregunta 3	Pregunta 4	Pregunta 5	Pregunta 6	Pregunta 7	Pregunta 8	Pregunta 9	Pregunta 10	Pregunta 11	Pregunta 12	Pregunta 13	Pregunta 14	Pregunta 15	Pregunta 16	Pregunta 17	Pregunta 18	Pregunta 19	Pregunta 20	Pregunta 21	Pregunta 22	Pregunta 23
62	90	95	100	100	100	100	100	100	100	100	67	100	95	100	95	81	76	100	90	100	43

Figura 60 Detail Report - Detalle Profesores

4.6.2. Reporte analítico para Profesores

El reporte analítico para Profesores tiene como objetivo mostrar de manera específica el resultado obtenido para cada profesor durante los diferentes semestres impartidos, para ello será necesaria la elaboración de los siguientes reportes:

- 1 - Main Report - Reporte Profesores
- 2 - Detail Report - Reporte Profesores Detalle

Nuevamente podemos observar que cada uno de los reportes está antecedido por un número secuencial, esta secuencia es la que se seguirá una vez que el reporte principal sea ejecutado.

4.6.2.1. Main Report - Reporte Profesores

Este reporte deberá ser el primero en ser ejecutado y como se muestra en la figura 61, pedirá al profesor ingresar su RFC a 10 dígitos y el semestre que desea consultar, de esta forma será posible consultar únicamente la información que corresponda a dicho RFC.

The screenshot shows a web browser window with the URL `http://localhost/ibmcognos/cgi-bin/cognos.cgi?b_action=cognosViewer&ui.action=run`. The page title is "1 - Main Report - Reporte P...". The main content area displays the following information:

- Logos for "INGENIERIA" and "UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO".
- Text: "Universidad Nacional Autónoma de México", "Facultad de Ingeniería", "Análisis de Encuesta a Profesores".
- Form fields:
 - "Ingresa tu RFC:" with the value "PIFC770616".
 - "Selecciona el ciclo escolar:" with a dropdown menu showing a list of semesters: 20132, 20131, 20122, 20121, 20112, 20111, 20102.
- A "Finalizar" button.

Figura 61 Página de consulta Main Report - Reporte Profesores

Una vez seleccionados los 2 parámetros el reporte "Main Report – Reporte Profesores" (Figura 62) mostrará todas las asignaturas que el profesor haya impartido en el semestre seleccionado, incluyendo la división, departamento y el grupo al que pertenecen, además del promedio obtenido en las encuestas

aplicadas a los alumnos. Adicionalmente se mostrará un gráfico que muestra un comparativo de los promedios obtenidos a lo largo de los últimos 6 semestres para cada una de las asignaturas.



Universidad Nacional Autónoma de México
Facultad de Ingeniería
Análisis de Encuesta a Profesores



Semestre: 20132

Semestre	División	Departamento	Asignatura	Tipo Asignatura	Grupo	Nombre Profesor	Promedio
20132	INGENIERIA ELECTRICA	INGENIERIA EN COMPUTACION	BASES DE DATOS	T	2	HERVAEZ HERVAEZ LUCRALA, N	83
20132	INGENIERIA ELECTRICA	INGENIERIA EN COMPUTACION	COMPUTACION PARA INGENIEROS	T	2	HERVAEZ HERVAEZ LUCRALA, N	83
20132	INGENIERIA ELECTRICA	INGENIERIA EN COMPUTACION	ESTRUCTURA Y PROGRAMAC.CO.ME	T	2	HERVAEZ HERVAEZ LUCRALA, N	76

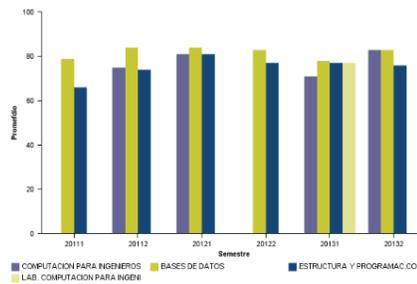


Figura 62 Main Report - Reporte Profesores

El reporte “Main Report – Reporte Profesores” utiliza 3 parámetros, el primero es el identificador del profesor (?P_Profesor?), el cual puede ser introducido únicamente al ejecutar el reporte y corresponde al RFC a 10 dígitos, el segundo parámetro es el semestre (?P_Semestre?), el cual puede ser seleccionado al ejecutar el reporte o bien utilizando el menú desplegable durante la consulta del reporte, y por último, el parámetro correspondiente a la asignatura (?P_Asignatura?), el cual permite realizar la consulta a un nivel de detalle más específico, pues permitirá desplegar el resultado obtenido para cada una de las preguntas aplicadas dentro de la encuesta a los alumnos. Los 3 parámetros son heredados al reporte “Detail Report - Reporte Profesores Detalle” (Figura 63) y son utilizados para filtrar la información, una vez en el reporte de detalle sólo el parámetro correspondiente al semestre puede ser nuevamente seleccionado.

	Pregunta 2	Pregunta 3	Pregunta 4	Pregunta 5	Pregunta 6	Pregunta 7	Pregunta 8	Pregunta 9	Pregunta 10	Pregunta 11	Pregunta 12	Pregunta 13	Pregunta 14	Pregunta 15	Pregunta 16	Pregunta 17	Pregunta 18	Pregunta 19	Pregunta 20
89	91	86	93	93	85	89	87	95	95	100	93	97	93	94	100	97	95	88	10

Figura 63 Detail Report - Reporte Profesores Detalle

4.6.3. Generación de un cubo de información por medio de IBM Cognos Transformer 10.1.1

Antes de generar el cubo de información con la información de las encuestas a profesores es importante mencionar la diferencia de hacer las consultas a los datos basados en un modelo de IBM Cognos Framework Manager 10.1.1 y un cubo generado en IBM Cognos Transformer 10.1.1. El primero hace las consultas en tiempo real, es decir, una vez que el reporte obtiene los parámetros de la consulta, IBM Cognos Framework Manager 10.1.1 funge como un intermediario entre la base de datos y el reporte obteniendo así la información de interés, mientras que un cubo de información consolida dicha información de interés, con el objetivo de hacer las consultas únicamente al cubo, mejorando el tiempo de respuesta en casos donde se maneje demasiada información. Por supuesto el cubo de información tendrá que tener a su vez un diseño propio donde se consideren los datos deseados y tendrá que ser generado cada vez que sea necesario actualizar la información en él, sin embargo este proceso también puede ser automatizado sin la necesidad de hacer alguna tarea manualmente.

Para iniciar el diseño de nuestro cubo de información es necesario primero definir la fuente de datos, IBM Cognos transformer 10.1.1 nos permite utilizar nuestro modelo de IBM Cognos Framework Manager 10.1.1 como fuente de datos, con ello nos aseguramos que los datos mostrados en el cubo de información consideren las relaciones y la cardinalidad que previamente se definieron para nuestro modelo de datos.

4.6.3.1. Creación del cubo de información

Una vez que nuestro modelo de datos es importado en IBM Cognos Transformer 10.1.1 es necesario hacer el diseño de nuestro cubo de información, esto con el fin de establecer la jerarquía de los atributos dentro de nuestras dimensiones y las métricas que podrán relacionarse con dichas dimensiones. Las dimensiones son elementos fundamentales en un cubo de información y su principal función es describir las métricas existentes en el modelo. Dentro de las dimensiones vamos a distinguir los atributos y las jerarquías definidas, partiendo del atributo más general hasta el atributo más específico.

Para nuestro modelo de datos consideramos 2 dimensiones:

- Dimensión Semestre (Figura 64)

Esta dimensión cuenta con un nivel para el año y un nivel para el semestre, jerarquizados en ese orden.

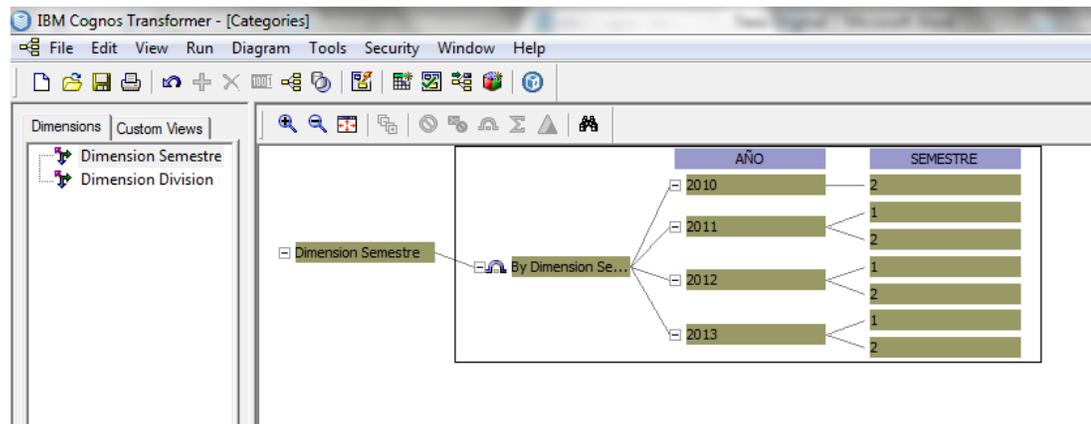


Figura 64 Dimensión Semestre

- Dimensión División (Figura 65)

Esta dimensión considera la estructura con la cual son organizadas todas y cada una de las materias impartidas en la Facultad de Ingeniería de la UNAM, partiendo de lo más general (Divisiones) hasta lo más específico (Profesores).

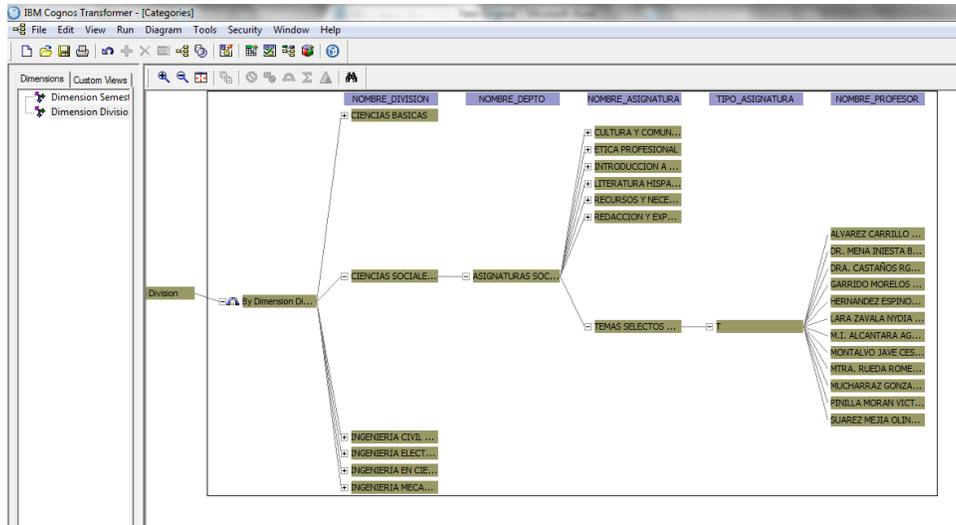


Figura 65 Dimensión División

En el caso de las métricas, son consideradas las 29 preguntas contenidas en la encuesta, además del promedio obtenido tanto por los profesores como los alumnos.

Una vez que nuestro cubo de información fue diseñado es necesario nuevamente publicar y generar dicho cubo dentro de nuestro portal de IBM Cognos Connection 10.1.1, de esta forma los usuarios podrán acceder a él. En la figura 66 se puede observar que el Cubo Encuestas Profesores ha sido generado.

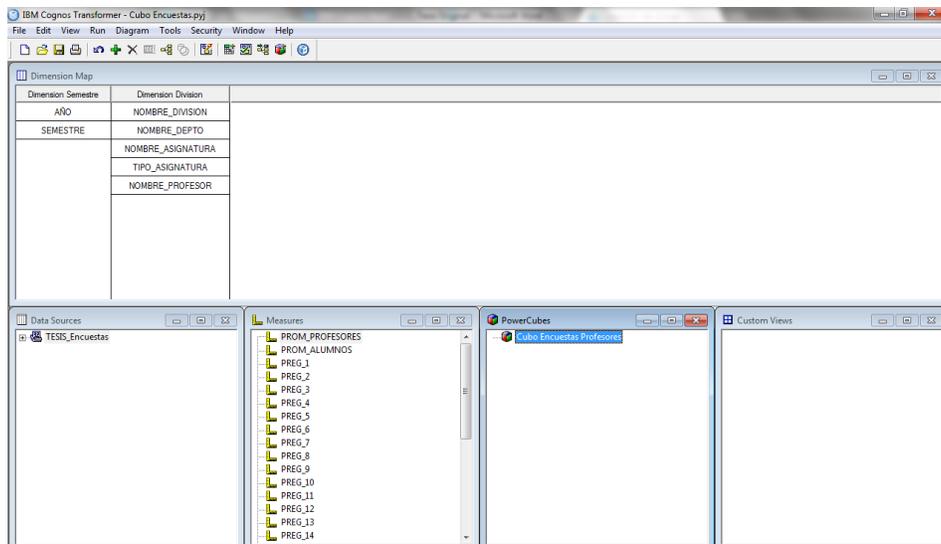


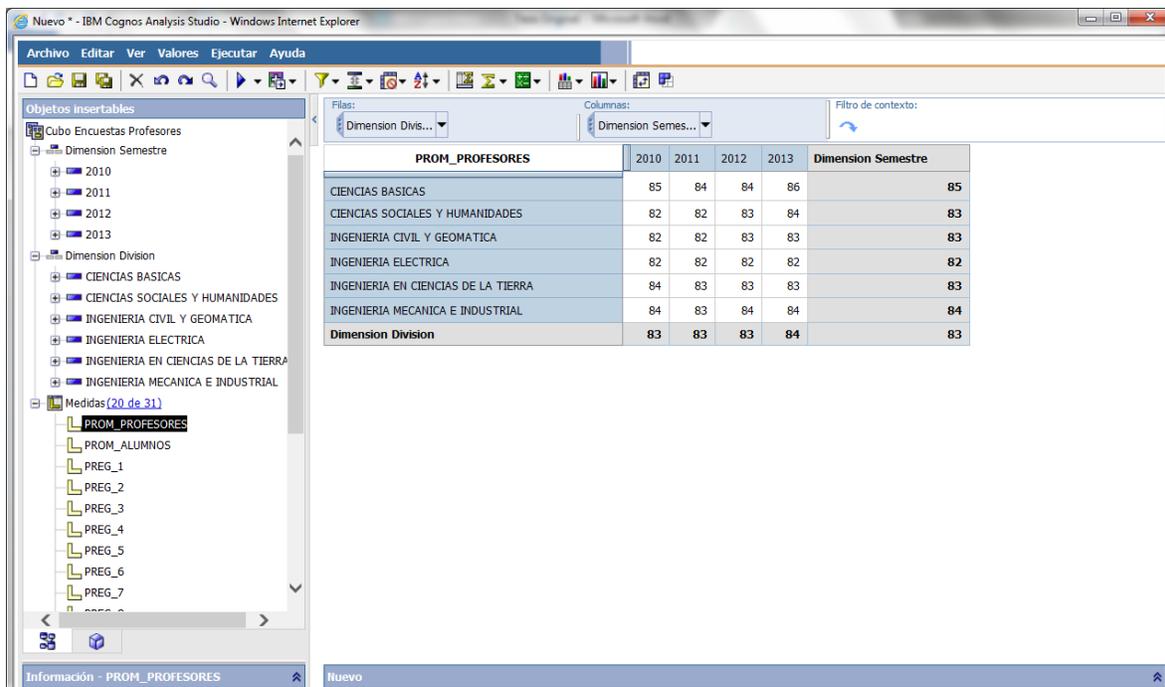
Figura 66 Generación cubo Encuestas Profesores

4.6.3.2. Consultando la información contenida en el cubo de información por medio de IBM Cognos Analysis Studio 10.1.1

El objetivo principal de la creación de un cubo de información es el permitir a los directivos de la Facultad de Ingeniería de la UNAM crear consultas personalizadas en tiempo real, que les permitan visualizar y comparar los resultados obtenidos por cada uno de los profesores a lo largo de los diferentes semestres, para ello se hará uso de la herramienta IBM Cognos Analysis Studio 10.1.1, localizada dentro de nuestro portal de IBM Cognos Connection 10.1.1.

Al ejecutar la herramienta IBM Cognos Analysis Studio 10.1.1 únicamente será necesario elegir como origen de datos el cubo que previamente generamos y publicamos dentro de IBM Cognos Connection 10.1.1 por medio de IBM Cognos Transformer 10.1.1.

Dentro de IBM Cognos Analysis Studio 10.1.1 los usuarios podrán crear consultas únicamente arrastrando y soltando los objetos de interés dentro del área de trabajo de la interfaz, crear filtros, cálculos e incluso crear gráficos. En la figura 67 se presenta una consulta simple a nuestro cubo Encuestas Profesores.



The screenshot displays the IBM Cognos Analysis Studio 10.1.1 interface. On the left, the 'Objetos insertables' (Insertable Objects) pane shows a hierarchy: 'Cubo: Encuestas Profesores' containing 'Dimension Semestre' (2010, 2011, 2012, 2013), 'Dimension Division' (Ciencias Basicas, Ciencias Sociales y Humanidades, Ingenieria Civil y Geomatica, Ingenieria Electrica, Ingenieria en Ciencias de la Tierra, Ingenieria Mecanica e Industrial), and 'Medidas (20 de 31)' (PROM_PROFESORES, PROM_ALUMNOS, PREG_1 through PREG_7). The main workspace shows a query grid with 'Filtros' set to 'Dimension Divis...' and 'Columnas' set to 'Dimension Semes...'. The grid data is as follows:

PROM_PROFESORES	2010	2011	2012	2013	Dimension Semestre
CIENCIAS BASICAS	85	84	84	86	85
CIENCIAS SOCIALES Y HUMANIDADES	82	82	83	84	83
INGENIERIA CIVIL Y GEOMATICA	82	82	83	83	83
INGENIERIA ELECTRICA	82	82	82	82	82
INGENIERIA EN CIENCIAS DE LA TIERRA	84	83	83	83	83
INGENIERIA MECANICA E INDUSTRIAL	84	83	84	84	84
Dimension Division	83	83	83	84	83

Figura 67 Interfaz IBM Cognos Analysis Studio 10.1.1

Conclusiones

1. El proyecto realizado cumple con el objetivo planteado al inicio del mismo, se desarrolló una aplicación que permite:
 - A los directivos, tener acceso a un reporte analítico y comparativo de todas y cada una de las divisiones, departamentos, asignaturas y profesores a lo largo de los diferentes cursos en los diferentes semestres donde se impartió cátedra, adicionalmente y por medio de un cubo de información, les permitirá realizar consultas personalizadas con las bondades que ofrece el trabajar con cubos de información, tales como drill down y drill up y visualizar la información partiendo de las dimensiones establecidas en él.
 - A los profesores, tener acceso a un reporte analítico de su desempeño particular a lo largo de los diferentes semestres donde se impartió cátedra, en dicho reporte podrán también comparar los resultados obtenidos durante los 5 semestres anteriores. También podrán visualizar el resultado obtenido en cada una de las preguntas aplicadas en la encuesta.
2. Tanto los reportes analíticos como el cubo de información permitirán tanto a los profesores como a los directivos identificar fortalezas y debilidades, tales como su capacitación, puntualidad, motivación, término del temario y desempeño al momento de impartir cátedra al alumnado, facilitando así la gestión de la calidad y la toma de decisiones en la Facultad de Ingeniería de la UNAM.
3. Se diseñó un modelo estrella que permite tener la información obtenida de las encuestas aplicadas a los alumnos de forma consistente y organizada.
4. Se crearon las tablas consideradas en el modelo estrella para almacenar toda la información existente y la información futura.
5. Se automatizó un proceso ETL, el cual tiene la capacidad para actualizar las tablas consideradas en el modelo estrella, de esta forma el Data Warehouse será incremental y permitirá almacenar información de futuros semestres.
6. Se presentó una visión general de las fases principales con las que un proyecto de Inteligencia de Negocios debe contar, además se demostró que un proyecto de Inteligencia de Negocios puede ser aplicado a distintos sectores, en este caso al educativo.

Áreas de oportunidad del proyecto

Es importante mencionar que existen áreas de oportunidad en nuestro proyecto, las cuales van relacionadas con la seguridad.

En nuestro reporte analítico para profesores se consideró el RFC para restringir el acceso a los diferentes profesores, sin embargo esta información en la actualidad es fácil de conocer, pues únicamente se necesitan datos generales de una persona para averiguarlo. Una de las soluciones a este problema sería utilizar el CURP en lugar del RFC, si bien también puede ser obtenido se necesita de otra serie de datos más específicos para averiguarlo. Otra alternativa sería la implementación de un registro único para cada profesor, que únicamente dicho profesor pueda saber y que podría constar del CURP con la adición de un número secuencial, de esta forma averiguar el registro de identificación sería una tarea muy compleja. Esta tarea podría ser desarrollada dentro del proceso ETL.

Otro aspecto que mejoraría nuestro proyecto es el uso de una herramienta para la gestión de usuarios, IBM Cognos 10.1.1 nos permite trabajar en conjunto con herramientas de esta índole, con ello la administración tanto de los mismos usuarios, como de los privilegios con los que deben contar sería más sencilla. Una herramienta con la que IBM Cognos 10.1.1 puede trabajar para realizar dicha tarea es Active Directory.

Otra área de oportunidad que puede tener nuestro proyecto es el uso de un servidor de correo, que trabajando en conjunto con IBM Cognos 10.1.1 y una herramienta de gestión de usuarios, permitirá enviar al término de cada semestre un correo electrónico con el reporte analítico para profesores, elaborado en nuestro proyecto, a cada uno de los diferentes profesores. Es importante mencionar que IBM Cognos 10.1.1 nos permitirá enviar a cada profesor únicamente la información concerniente a él, basado en el registro único del profesor, con ello la confidencialidad estaría asegurada.

Índice de Figuras

Figura 1 Diagrama con las fases de un proceso ETL.....	11
Figura 2 Diagrama de arquitectura Cliente-Servidor	12
Figura 3 Interfaz IBM Data Stage Administrator 7.5.2.....	16
Figura 4 Interfaz IBM Data Stage Designer 7.5.2	17
Figura 5 Interfaz IBM Data Stage Administrator 7.5.2.....	17
Figura 6 Interfaz IBM Data Stage Manager 7.5.2.....	18
Figura 7 Cuadrante mágico de Gartner Inc [9].....	19
Figura 8 Interfaz IBM Cognos Connection 10.1.1	20
Figura 9 Interfaz IBM Cognos Query Studio.....	20
Figura 10 Interfaz IBM Cognos Report Studio 10.1.1.....	21
Figura 11 Interfaz IBM Cognos Analysis Studio 10.1.1.....	22
Figura 12 Interfaz IBM Cognos Transformer 10.1.1.....	23
Figura 13 Interfaz IBM Framework Manager 10.1.1.....	24
Figura 14 Fases en un proyecto de Inteligencia de Negocios	25
Figura 15 Modelo estrella del Data Warehouse propuesto	26
Figura 16 Descripción tabla DIVISIONES	27
Figura 17 Descripción tabla DEPARTAMENTOS	28
Figura 18 Descripción tabla PROFESORES.....	28
Figura 19 Descripción tabla ASIGNATURAS	29
Figura 20 Descripción tabla SEMESTRES.....	30
Figura 21 Descripción tabla HECHOS	31
Figura 22 Secuencia_Carga_Total.....	32
Figura 23 Tarea Carga_CSV_A_BD	33
Figura 24 Tarea Carga_DIVISIONES.....	36
Figura 25 Objeto Base de Datos Origen para la tarea Carga_DIVISIONES.....	36
Figura 26 Objeto transformación para la tarea Carga_DIVISIONES.....	37
Figura 27 Objeto Base de Datos destino para la tarea Carga_DIVISIONES.....	38
Figura 28 Tarea Carga_DEPARTAMENTOS.....	38
Figura 29 Objeto de Base de Datos Origen para la tarea Carga_DEPARTAMENTOS.....	39
Figura 30 Objeto transformación para la tarea Carga_DEPARTAMENTOS.....	39
Figura 31 Objeto de Base de Datos destino para la tarea Carga_DEPARTAMENTOS.....	40
Figura 32 Tarea Carga_PROFESORES	40
Figura 33 Objeto de Base de Datos origen para la tarea Carga_DEPARTAMENTOS	41
Figura 34 Objeto transformación para la tarea Carga_PROFESORES	41
Figura 35 Objeto de Base de Datos destino para la tarea Carga_PROFESORES	42
Figura 36 Tarea Carga_ASIGNATURAS.....	42
Figura 37 Objeto de Base de Datos origen para la tarea Carga_ASIGNATURAS.....	43
Figura 38 Objeto transformación para la tarea Carga_ASIGNATURAS.....	43
Figura 39 Objeto de Base de Datos destino para la tarea Carga_ASIGNATURAS.....	44
Figura 40 Tarea Carga_SEMESTRES.....	44

Figura 41 Objeto de Base de Datos origen para la tarea Carga_SEMESTRES	45
Figura 42 Objeto transformación para la tarea Carga_SEMESTRES	45
Figura 43 Objeto de Base de Datos destino para la tarea Carga_SEMESTRES	46
Figura 44 Tarea Carga_HECHOS.....	46
Figura 45 Objeto de Base de Datos origen para la tarea Carga_HECHOS	47
Figura 46 Objeto transformación para la tarea Carga_HECHOS.....	47
Figura 47 Objeto de Base de Datos destino para la tarea Carga_HECHOS.....	48
Figura 48 Organización de metadatos en IBM Cognos Framework Manager 10.1.1	49
Figura 49 Relación HECHOS-ASIGNATURAS.....	50
Figura 50 Relación HECHOS-DEPARTAMENTOS.....	51
Figura 51 Relación HECHOS-DIVISIONES.....	51
Figura 52 Relación HECHOS-PROFESORES	51
Figura 53 Relación HECHOS-SEMESTRES	52
Figura 54 Generación del paquete de metadatos dentro de IBM Cognos Framework Manager 10.1.1 ...	53
Figura 55 Paquete de metadatos TESIS_Encuestas dentro de IBM Cognos Connection 10.1.1.....	53
Figura 56 Main Report – División.....	55
Figura 57 Detail Report – Departamento.....	56
Figura 58 Detail Report – Asignatura	57
Figura 59 Detail Report – Profesores	58
Figura 60 Detail Report - Detalle Profesores	58
Figura 61 Página de consulta Main Report - Reporte Profesores.....	59
Figura 62 Main Report - Reporte Profesores	60
Figura 63 Detail Report - Reporte Profesores Detalle	61
Figura 64 Dimensión Semestre	62
Figura 65 Dimensión División.....	63
Figura 66 Generación cubo Encuestas Profesores.....	63
Figura 67 Interfaz IBM Cognos Analysis Studio 10.1.1.....	64

Bibliografía

- [1] Datos, información y conocimiento [online]. Sinnexus. Fecha de consulta, 2014. Disponible en: http://www.ull.es/view/institucional/bbtk/Referencias_Normas_IEEE/es
- [2] Diccionario de datos [online]. IPN, Polilibros. Fecha de consulta, 2013. Disponible en: http://148.204.211.134/polilibros/portal/Polilibros/P_proceso/Bases_de_datos_vers22_Victor_Alvarez/POLILIBRO/UNIDAD3/BASE%20DE%20DATOS-81.htm#inicio
- [3] Definición de SGBD [online]. Leandro Alegsa. Fecha de consulta, 2013. Disponible en: <http://www.alegsa.com.ar/Dic/sghd.php>
- [4] Extracción, transformación y carga [Online]. Wikipedia. Fecha de consulta, 2013. Disponible en: http://es.wikipedia.org/wiki/Extract,_transform_and_load
- [5] Esquema estrella [online]. ETL-tools.info. Fecha de consulta, 2013. Disponible en: http://etl-tools.info/es/bi/almacenedatos_esquema-estrella.htm
- [6] Business Intelligence [Online]. Scientific Paper org. Fecha de consulta, 2013. Disponible en: http://www.scientificpapers.org/wp-content/files/1102_Business_intelligence.pdf
- [7] ¿Qué es Business Intelligence?. Sinnexus. Fecha de consulta, 2013. Disponible en: http://www.sinnexus.com/business_intelligence/
- [8] Breve historia de Business Intelligence. Paul Orquizo. Fecha de consulta, 2013. Disponible en: <http://www.businessintelligence.info/definiciones/historia-business-intelligence.html>
- [9] Cuadrante mágico de Gartner 2013 [Online]. Antonio Rivas. Fecha de consulta: 2013, Disponible en: www.bi.dev42.es/2013/03/10/cuadrante-magico-de-gartner-2013/
- [10] IBM Cognos Business Intelligence V10.1 Handbook. Dean Browne, Brecht Desmeijter, Rodrigo Frealdo Dumont, Armin Kamal, John Leahy, Scott Masson, Ksenija Rusak [Mayo 2013]